THE STUDY OF INTERRELATIONSHIPS IN THE ANALYSIS OF FAILURE TIME DATA WITH APPLICATION TO CLINICAL TRIALS

SATE SONGHORABADI

Doctor of Philosophy,

University of Edinburgh

1984

UNIVERSITY OF EDINBURGH

ABSTRACT OF THESIS (Regulation 7.9)

Name of Candidate ....SATE...SONGHORABADI.....

Address .....

Degree .....DOCTOR OF PHILOSOPHY..... *Date* .....19TH APRIL, 1984.....

Title of Thesis .....THE STUDY OF INTERRELATIONSHIPS IN THE ANALYSIS OF FAILURE

.....TIME DATA WITH APPLICATIONS TO CLINICAL TRIALS.....

Generally it is recognised that scientifically designed clinical trials play an important part in the development and the evaluation of medical treatments. Such trials fundamentally contain natural administrative and ethical conflicts.

In the course of this thesis we will look at the developments in the analysis of failure time data and deal with study of interrelationships within clinical trial data. The general utilisation of such analytical methods have been made possible by the distribution of fast computerised processing power.

In the area of survival distributions we will consider various empirical distributions and perform a comparative study of the non-paramtric and parametric methods and deal with the recent developments in the area of semi non-parametric methods, using the Cox's proportional hazard model. We will perform an assessment of power efficiencies of tests for computer simulated clinical trial data, under varying, sample sizes, censoring levels, significance limits, asymptotic normality and likelihood tests, time dependency assumptions, and a range of treatment and prognostic effect values.

We consider interrelationships of relevance in the context of trials to be those of prognostic effects as well as the event time variabilities under a multivariate failure time context.

We will deal with two data sets, both of which relate to breast cancer. Initially we consider a data set from a clinical trial organised in Edinburgh, and study prognostic and treatment effects for a set of risk factors such as local recurrence, metastatic recurrence and death. Finally we use a data set on breast cancer patients purely for the assessment of prognostic effects. In the latter study we consider a set of accepted prognostic effects as well as a set of measurements dealing with tumour change and extent. In the discussions of the above we present various models in order to test time periods to and from intervening events in a multivariate study. We will also consider time dependency of various effects in order to check on the persistance of an effect on the time scale.

ACKNOWLEDGEMENTS

I wish to express particular appreciation and thanks to Dr. Robin Prescott, my supervisor, for his consistent guidance, help and kindness during the different stages of the course of this thesis.

I am grateful of Professor A.P.M. Forrest for making the various grants available and so the opportunity to write this thesis.

I am indebted to Dr. Helen Stewart, Director of the Scottish Cancer Trials Office, Mrs. Sue Robinson, the Co-ordinator of the Scottish Cancer Trials, Mr. Walter Lutz, Director of the Medical Computing and Statistics Unit, for the many useful and helpful conversations in the last few years.

The final completion of this thesis was made possible by the care and accuracy of Mrs. Wendy Taylor in typing the thesis and showing much patience with the mathematical parts, and to whom I am most grateful.

Throughout the development of this thesis various members of the Edinburgh University, notably the Staff of the Medical Computing and Statistics Unit, the Scottish Cancer Trials Office, the Edinburgh Region Computing Centre and the Edinburgh University Library have been the source of valuable help and to all of whom I owe special thanks.

C O N T E N T S

C H A P T E R 1

INTRODUCTION

Statistical inference has been increasingly regarded as a necessary tool for the assessment of risks in its various forms. This necessity to examine and compare risks is becoming an essential part of the methodology of a large number of subjects that deal with risk in its varied and distinct forms such as occupational hazards, industrial developments, environmental risks and patient management in hospitals. As an abstract formulation we can regard the general problem as that of choosing between two or more courses of action knowing that the courses of action have risk values attached to them. Part of these risks are in terms of costs and benefits to the individual and partly to the collective society. We can thus identify a set of general questions by which finding relevant answers for each particular context is the essential part of the methodology. How much information is sufficient for discriminating between the courses of action? What are the acceptable levels of benefit for introducing a new course of action? What are the appropriate/

appropriate measures of risk? What are the conflicting rights of the individuals and institutions, and finally, how do we collect the relevant information?

The principal part of the notion of risk and its appraisal is introduced as soon as one considers social and human dimensions of a decision. In contrast, within the framework of most natural experiments the concept of risk does not usually arise and is substituted with that of deriving optimal rules for obtaining appropriate measures at minimum cost and time in collecting the relevant information.

The methodology we are dealing with in this thesis relates to that of a clinical trial and analysis of failure time data for a clinical trial. The principal aims are to show that for this particular application, within the limits of controlled experiments how concepts such as control of concomitant information, exploratory approach in analysis and that of study of association between various risks may be employed to provide a better understanding of the data.

1.1 HISTORY.

In 1693 E. Halley the well-known discoverer of the Halley's Comet produced a life table of the population of Breslau in Germany. This data was based on the city records and was published in the philosophical transactions of the Royal Society of London, with the title of "An estimate of the degree of mortality of mankind, drawn from the curious tables of the birth and funerals in the city of Breslau." The data was composed of the age and time of death and more importantly the cause of Death was specified to be small pox or other causes.

This final small detail on the cause of death in Halley's data, later on led Daniel Barnoulli in 1760 to reformulate the problem. In his paper which was read at the Royal Academy of Science, Paris, Barnoulli adopts an ingenious and simple argument to derive for each individual, who died of small pox, his determined length of life had the risks of death from small pox been eliminated.

However, the method is based on the assumption that the disease affects the total population in a uniform manner, and thus the method is not sensitive to the possibility of structural variability for smaller subgroups such as, a small subgroup of patients being strong and thus more immune from the disease. One rather obvious source of structural variability was pointed out by D'Alembert (1761), the eminent French mathematician of the time. He noted that the probability of contracting small pox as well as dying from it may well be dependent on age.

At the time when d'Alembert and Bernoulli were constructing the early life tables, mathematical tools had not been developed for a more refined analysis. The method is based on a deterministic analysis of the numbers in a time period while it does not provide a probabilistic interpretation. Further it seems that although Halley may have been interested in a functional form of parameters to investigate the total population and possibly a population distribution (being and astronomer himself.) Bernoulli adopted a non-parametric approach at each interval, based on a number of cases, to determine the expected values. (The distinction between parametric and non-parametric methods will be discussed more extensively later).

In actuarial studies a similar problem arises where a population is measured for the risk of death. At the time of analysis some members of the population may not have completed their time to the response of interest (death) and therefore no information is available on their time of death. By 1825, Probability Theory had been well developed and Gompertz (1825) had produced a function to approximate such a population survival distribution with the above property of some cases not contributing to death times. This distribution known as Gompertz-Makeham has been the central theme of many models in actuarial theory. The model proposed by Gompertz and further by W.M. Makeham (1875) is very realistic in that, the basis of its philosophy is to allow separate risks of withdrawals from the population with a response, such as death due to a particular cause (e.g. cancer), or due to other causes. In fact by ignoring the possibility of different

rates/

rates of death due to different causes would at times invalidate the conclusions of the study. However, the above flexible approach allows a check on the assumptions regarding the relevant causes of death. The importance of this approach in allowing different risks was not introduced into medical studies until the mid-1950's with the contribution of J. Cornfield in application to clinical trials. Studies carried out as late as 1939 by Bernstein, Binham and Ach came to an invalid conclusion through overlooking the problems of choice of relevant response rates as the final events of interest. Prior to the works of Cornfield, similar developments were taking place in another branch of applied mathematics. Emergence of complex mechanical devices and early electronic networks required mathematical models for a representation of the logical flow of the chance of failure, and a final assessment of the probability of failure of the system. These areas were named reliability and life-testing. At present a major application of these techniques is related to development of defence systems in U.S.A. and U.S.S.R. There are many similarities between reliability studies and survival studies of a population. The conceptual simplicity of the electronic systems were partly responsible for the emergence of recent trends in multivariate failure time analysis. Any device may be composed of a number of components each with its own risk of development of a failure. These components may be in series and thus failure of a single component can result in a system failure or may be in parallel. A medical example of the latter would be a study of kidney failure where damage to one kidney would not be fatal. The basis of this approach in medical studies and population studies has been laid by Fix and Neyman in 1951 and Chiang in 1960.

As/

As was pointed out a sound methodology had been developed by the mid-1950's to apply statistical methods to clinical trials. The Epochal Streptomycin trial conducted under the auspices of the Medical Research Council first reported in 1948 by MRC and later by Bradford-Hill (1962) may also be considered as one of the contributors to present trends. What is important about this study is its impact on medicine by the introduction of scientific attitudes to the study of treatments. Further development of methodology in clinical trials was diversified from those of analytical methods derived from reliability and life testing to a shift of emphasis towards the proper scientific practice of considering a good design as a primary aim.

Within the medical literature Peto et al (1977) proposed a major set of guide lines for the conduct of trials. Most of the emphasis in their report is on the construction of a well-designed trial. For the analysis of data however they adopt a standard statistical method for use in a clinical trial.

Some of the works of J. Cornfield were responsible for early application of statistical analysis methods in clinical trials. He also pointed out some of the problems of statistical interpretation, for example, in the area of multiple risks. Although the framework of risks associated with components of a system failure is simple enough for mechanical applications, in medicine there are more major difficulties. Some of the developments of the thesis will be related to these difficulties.

Finally, it seems that a change of emphasis has taken place. In early studies of the development of risk of a disease, most applications were on communicable diseases. An epidemic develops

and/

and initially there is a high risk of failure (death) from contracting the disease. With the passage of time chance of progression of the disease decreases and falls to zero. That is for survivors within a relatively short period of time there is often a possibility of return to normality. The present context for chronic diseases must assume that from the start of the process, failure begins and so with any secondary event the chances of death increase.

1.2 Some Methodological Concepts in Clinical Trials.

In this section we present some of the special features of clinical trials. Basically the aim of a clinical trial is the management of the unknown in a clinical setting, so that some knowledge or dogma that has been obtained due to historical reasons may be refuted or substantiated. The information gained is then useful in practic in the administration of treatments. In this respect a trial does not differ from an experiment in the natural sciences. However any form of a scientific enquiry which involves the collection of data within the human environment is open to various constraints. Some are related to the impact of the study on the subject under study and some are related to the actual validity of conclusions drawn from the study. Although none of the above problems undermine the fact that the final scientific answer is important, they do make a contribution to the quality of the data which is gathered and the role data gathering plays in the administrative and ethical areas. From a medical point of view the question is not only of legitimacy of the approach in terms of how scientific the trial is, but also whether the trial can be administratively and ethically accepted. Difficulties in the management of the unknown is present in many areas. In other forms of trials that

may/

may take place outside of medical fields the experimental unit may be subject to far greater risks. In fact the introduction of any new policy can be thought of as posing initial high risks. Within the framework of medicine, the problem of risk is due to the rights of the individuals and how the uncertain effects and its conclusion may benefit the society through the works of institutions.

In here a distinction may be made between two types of risks involved. One form of risk is due to possible progression of disease or expected status of disease over time if there is no intervention. The other risk is related to the new method of control of progression of the disease with expected side effects over time. Depending on the phase of testing of a new drug clearly a different level of risk may be present in treatment.

Three stages have been recognised in the development of a new drug. We will in here mention these three phases but the area of particular interest for our study is mainly related to one phase only and deals with controlled trials.

Initial study of a new drug is often referred to as a phase one trial. There is little emphasis on actual statistical testing but more on obtaining insight into acceptable dosage and practical limits in administration.

Next stage is a screening study to assess efectiveness of drug under study and its value in performing further controlled studies. Finally a phase three trial is the stage where a comparison of two or more treatment regimes is needed.

The phase two trials have been at times the subject of controversy as to their place between phase one and phase three. Often a balance is made between the level of advance of the disease and the risk it subjects the patient to with that of accepted value of the treatment.

The first and foremost motivitation in proceeding with a trial is to find scientifically valid answers with the minimum number of patients in the shortest period of time. A well designed trial has been encouraged from various approaches by many authors. Peto et al (1977) consider the roles of factorial designs in trials. Simons (1979) considers the role of stratification in design stages of a trial and Brown (1980) discusses the role of cross over trials although it is not relevant to survival studies. We have mentioned these methods for completeness and consider some of them during analysis of trials in later chapters. At the centre of these approaches lies the principle of randomisation of the patients to the various arms of a trial. Randomisation is seen from a scientific view to hold a central role. Also it has been received increasingly by the medical profession as having an important place in all assessments of comparative patient management.

An alternative to controlled clinical trials is the use of historical controls which has found favour in certain clinical circles. The latter approach does not resolve the important problems of personal bias of the investigator, and the passed on institutional dogma. At the present time the value of randomised controlled clinical/

clinical trials is recognised by most medical investigators, although their proper practice in data collection and interpretation have been the subject of discussion, in different situations. Ethics and value of scientific refutability form the framework of discussion in this circumstance.

In the past, two general types of historical controls have been reported. One group is related to comparison of patient groups treated by different methods at different times within the same institution, and a second type which allows the comparisons to take place across various institutions. Neither of these two methods provide a satisfactory basis for allowing a like with like comparison of two groups that have been treated by different methods without making unjustifiable assumptions. Clearly the problem of final interpretation is that, it becomes difficult to distinguish effects due to treatment with those of institutional and/or time variability. For example, Pocock (1974) has reported the unreliability of historical control results from three cancer chemotherapy co-operative groups. In this study a comparison is made between similar treatments which are used consecutively. 19 such instances were identified with the changes in the death rate ranging from - 46% to +24% and with 4 instances giving a significant difference at the 2% level. The phase two trials that we mentioned may at times be defined to belong to this class of historical controls.

If a treatment is found to perform a major significant improvement on cure, the weight of such evidence may be so overwhelming that a controlled trial is not necessary and thus confounding of treatment/

treatment effect and time effect is judged unimportant. Although it must be emphasized what may seem very overwhelming evidence to ignore time confounding for some is not necessarily overwhelming evidence to others at all times.

Problems of historical controls are not only confined to their philosophical position. In practical terms there are some further difficulties. Missing information is usually a probelm in statistical analysis and the time gap between treatment methods does not provide a uniform setting for the recording of relevant information. Prognostic indicators are often subject to various forms of interpretation and again across institutional variability combined with time variability can introduce additional bias. In terms of analysis the historical control data analysis require relatively more control of various factors. These effects will make the analysis firstly more complex and secondly more dependent on model assumptions and open to differing interpretations. The above were some of the problems of historical controls given that patient environment does not introduce its own bias. Eligibility criterion, wrong patient mix, adjuvant patient care, observors perception of patients final status are all various factors that open the ways for introducing bias from medical participants in a trial.

Although we have put randomisation as the central argument of a scientific approach to trials, there are a few other issues involved in a good statistical design. For reasons of efficiency and representativeness one can use multi-centre trials with reasonable levels of stratifications. Further, depending on the/

the form of questions, one may proceed with a cross over or factorial design trial.

For a good scientific conclusion, there is a need to organise a trial with a sufficiently large number of patients. In order for a trial to be able to detect differences of clinical importance between the treatments and be likely to judge this difference as statistically significant, either the period of accrual of patients has to be long enough to allow a large number of homogenous patients to be allocated to various treatments, or alternatively a multi-centre approach could be adopted by which a number of institutions such as hospitals and medical centres refer the decision making to a central trials office. The last approach at times can lead to an introduction of more heterogenity in the total population, due to environmental, varied practice or institutions or population structure differences of the different areas. In here a distinction must be made between institutional variability that is controlled by the randomisation and those of historical controls. In controlled trials although extra variability is introduced by the institution, the within institutional strata randomisation ensures that no bias is involved in the final assessment. The long accrual period also has a slight similarity with historical controls in that it spans through time. However, the distributional variability of the patients prior to treatment allocation can be thought of as being more consecutive in controlled trials.

Once a large number of patients are allocated and randomised to different treatments, then the patients are followed up for a long period. Continuously the patients are monitored for development/

ment of patterns of progression of the disease, with respect to survival, side effects, disease spread, together with treatment, stratifying and concomitant variables. Further it is necessary to perform the analysis of the data at various times with up-dated follow-up information mainly for ethical and administrative purposes. It is likely that at the time of analysis some patients may not have responded for each particular time measurement. This effect is known as censoring of the survival time for the patients, in that no response is known and survival time is cut off by other events before the patient has had a sufficiently long period of follow-up for responding. Censoring is a special effect present in study of failure time data. A few special problems arise in presence of censoring. The major one is related to "lost to follow-up" cases. It is possible that in certain trials a group of patients produce a different distribution as regards to the number of patients that are lost at time of analysis. Such effects are mainly due to administration of the trials and are undesirable. In the next section we will deal with censoring in more detail.

The randomisation can provide a good setting for control of administrative bias. However it provides no guarantee that differences between the groups towards the end of study are only due to treatment effects. It is important that together with the formulation of an a priori hypothesis, a framework is set up so that the patients in the two groups are in some sense comparable in terms of their known prognostic indicators and follow-up procedures. This framework in practice is extended to a protocol that all participants agree to conform to. In this way the data collection and interpretation of effects and some of the clinical practices are standardised.

From a scientific point of view the emphasis on the better design of a trial will clearly enhance the reliability of a conclusion that is drawn from a trial. Much of the respectability of hard data sciences such as physics and chemistry is attributed to the development of good calibration and development of instruments for proper measurement. The development of better recording facilities and computer storage and analysis may go in some way to provide more uniform standards in clinical assessment.

Some of the prognostic indicators later form the basis of further analysis of survival times. At times such analysis can suggest a path for formulation of a new hypothesis. In here there is a need to distinguish between two forms of questions that may arise. All the above discussions have dealt with the value of a treatment in terms of the individual survival times. However, other failure time indicators related to progress of disease, side effects and changing prognostic indicators at times can be used to provide information on the biological nature of the treatment and disease.

This latter distinction between the two types of question is made due to the recognition of the fact that trials are not experiments in the pure hard data sense of the word. What may be termed in the 'hard' sciences as data dragging and problems of multiplicity may justifiably be recognised as locally valued exploratory data analysis in the clinical trial data context. The problem is that what is often considered as valuable research is related to the unknown and it is in the area of the unknown where clinical judgement may be thought to be at its strongest value. This type of explor-

atory/

atory analysis therefore can provide a framework for reduction of the data and secondary analysis. Part of the benefits of local exploratory analysis will be in the formulation of new hypothesis and part of the benefit may be in terms of an improvement in the quality of the data that is collected. However it must be emphasised that a proper placing of secondary (exploratory) analysis is achievable only by a utilisation of diverse and relevant methods of analysis.

1.3 Trends, Philosophy and Ethics.

In the previous section an overview of the main topics of clinical trials was given. In this section of this chapter some trends and developments in the light of the present definitions will be given. Clinical trials play an important role both in terms of the value of the information they produce and in their impact on the general public. They introduce problems of ethics in a situation where there are conflicting interests and risks involved. Further, to resolve the real problems that exist and to arbitrate between conflicting risks and advantages we use scientific methodology. This is at a time when the distinction between science in its pure sense of the work and its applications are diminishing.

In the previous pages, we discussed the setting within which a trial is performed and we touched on a few topics that determine the design stages of a trial. We will now continue with the quality and form of the data that arises and the type of information that is considered to be essential for providing an answer to the questions on trials.

The minimum data required for the analysis of a trial is the information on treatment allocation to the individual patients and the survival distributions at the end of the study. A slightly more elaborate analysis may also require auxilliary covariate information on the prognostic indicators. In the course of this thesis we will mention some of the established methods for an extensive analysis and concentrate on the proportional hazard method of Cox (1972), read at the Royal Statistical Society.

The proportional hazards model and some of the recent extensions constitute a major development in the methods of analysis. The model allows a comparison of the history of the disease by use of prognostic indicators that may change through time. For a statistical method of analysis, the approach can allow an expansion of the methodology of analysis of event time variability.

In here we will mention a few recent approaches that have been attempted in various fields. Later in the course of the thesis we will concentrate on cancer trial data only.

1. Di prete 1981, Considers a study of <u>duration</u> of employment in which adult members of a labour force pass from various states of unemployment to employment.
2. Hannan, Tuma and Greenveld 1978, consider effects of income and other effects on the <u>periods</u> of marriage and divorce.
3. Hannan and Carroll 1981, study of effects of various characteristics in society that lead to various forms of government and the <u>times</u> of remaining in one political status.

4./

4. Crowley and Hu 1977, study heart transplant data and various characteristic variables in determining survival times.
5. P.K. Anderson and N.K. Rasmusson (1982) consider times of admission of a group of women attending psychiatric hospitals.

Although the above studies arise in different settings, all deal with the progression or development of a process through time. This parallels the progress of disease in time and possible events that may occur in this process. The emphasis in here is not so much that of desirability of the approach in a clinical setting but more in dealing with practicality in providing a flexible model for the interpretation of the data.

The need for organised experimentation arose in the natural science due to a need to replace occasional fragmentary experience with harder unbiased evidence. In such contexts the experimental unit is an inanimate object with no morals, collective memory or values. The need to perform experiments on human subjects in general arises out of a wish to answer important questions on the nature and treatment of various diseases with some degree of scientific and ethical accountability. The final result is scientific and technical progress for the benefit of society. In the biomedical fields in particular the institutional demands and individual rights play a major part in the final outcome of the study. In general two types of experiments are identified in this context, therapeutic and non-therapeutic. We will now give a brief description of the two.

Non/

Non-therapeutic experiments are primarily performed for the purpose of gaining new knowledge and not so much for reasons of benefit to the subject. An example is the use of healthy human volunteers in early phases of drug testing.

More important are the therapeutic experiments. The primary aim is to benefit the patients by intervening in the progress of a disease. However similtaneously the intervention is organised in a controlled manner so that a valid scientific conclusion may be possible at the end of the study. On the scientific importance of such trials, M. Baum, R. Kay and H. Scheurlen (1982) have written:

> "Over-enthusiastic and uncritical adoption of a conceptual framework by some clincians has led to therapeutic dogma and consequent erection of new ethical constraints. Factors outside the control of the clinicians which are active in hindering progress are an increasing public awareness of the problem, the clamour for informed consent, scrutiny by the legal profession, the involvement of national government agencies and the escalating costs of treatment. Those developments also force us to reconsider the scientific fundamentals of clinical trials as opposed to other approaches to scientific questioning".

The key word in statistics is information and evidence and it has always dealt with 3 practical problems. What are the assumptions of analysis? What are the assumptions of collection? and finally, how relevant is the data? The above problems are particularly relevant in trials in that results may not be know for a long period. As far as the attitudes of the clinicians involved in treatment and measurement are concerned, changes may take place. This may result in premature withdrawal from a trial with the result that the objectives of the trial are not fulfilled. Alternately, their assessment of patients may change over a period of a trial.

This final remark will be emphasized to some extent in the course of the thesis on the effect of varying definitions such as progression of the disease that may arise. These changes of concept may affect the clinicians from many directions, from those of personal motivation to be right to those of individual responsibility. The final effect is that there is potential for conflict between the scientific objectives of the trial and the subjective decisions of the clinicians. In here science is dependent to some extent on the background assumptions. In the physical sciences performing standard uniform methods of measurement is possible, but in a clinical setting even with a willingness to conform systematically with the protocol, the measurement will not necessarily be free from preconceptions. One further difficulty mentioned in the last section is human involvement as an experimental subject and the fact that individual rights are at the forefront of any responsibility. There are different modes of ethics present. First of all cancer is a problem and it is ethical for our institution to find the relevant answers. Also it is ethical to utilize resources efficiently and be aware of their value and obtain relevant inference. Further, there are clinical ethics based on the personal judgement of the physician and finally there are interests of the individual and a choice preference he or she may want to excercise.

In here a difference exists between the observational requirements of the natural sciences and the ethical attitudes of the individual physician, mainly due to the limited form of information available to them at a time. For example during the progress of a trial a physician may gain the impression from incomplete data that one/

one treatment is more successful that another, posing him an ethical dilemma as to whether to continue with the trial or withdraw. It is difficult to consider any of the above mentioned problems in isolation from the role of computer and information networks in the developments of future procedures. Science as a common arbitrator is confronted with many information techniques ranging from multivariate statistical methods to those of data base management systems. A general and undisciplined use of the above methods would lead to an increased likelihood of eithical conflict. On the other hand a utilisation of relevant methods of secondary analysis, in the correct context and specified fully by a protocol in the beginning of the study may contribute towards a better participation. With respect to the role of feedback of information, Prescott (1978) based on patient entry into Edinburgh trials, indicates that with a feedback of information it may be possible to maintain the level of interest in a multicentra trial.

1.4 Definitions and Mathematical Functions.

In this section we will develop and define some of the initial concepts in survival or failure time analysis. Before we commence with various definitions that we need in this thesis it must be emphasised that the titles survival or failure time analysis are a little misleading in that basically we are interested in an analysis of progress of various events in time and this event in time need not be death or regression but can be discharge from hospital, or any other event not necessarily representing a failure.

In the study of survival time three mathematical functions are often used. These are survival function, hazard rate and the density function. These functions are in fact different transformations of one another. For reasons of interpretation however a particular function is usually used and in the course of the thesis we will mention certain practical advantages of each.

For all of our cases we have a time t_i available which in the observed period for that case until a particular event of interest for example death. Clearly t_i is always greater than zero. We define for the density function of T the function f(t) and for the distribution function F(t) as is the usual practice in the statistical literature. We can thus define a more useful function for these applications, namely survival function S(t) giving.

$$S(t) = 1 - F(t) = Pr\ \{T > t\} = Pr\ (\text{survival for a case exceeds } t)$$

Also

$$f(t) = \frac{-d}{dt} S(t)$$

(and as usual $\int_{-\infty}^{\infty} f(t)dt = 1$)

Another useful function is the hazard rate or hazard function. In epidemiology, this is named as a force of mortality. We have here the hazard function given by:

$$\lambda(t) = \frac{f(t)}{S(t)}$$

$$\lambda(t) = Pr\ \{t < T < t + dt \quad T > t\ \}$$

$$= Pr\ \{\text{death in a small interval } dt \text{ given survival up until time } t\}$$

We/

We can now explore the functions in relation to each other.

$$\int_0^t \lambda(u)\,du = \int_0^t \frac{f(u)}{S(u)}\,du = [-\log\{1 - F(u)\}]_0^t$$

$$= -\log[1 - F(t)] = -\log S(t)$$

which implies the important relationship

$$S(t) = e^{\int_0^t \lambda(u)\,du}$$

These concepts have been defined here for a continuous case but can be extended to discrete form of T.

In practice what distinguishes survival analysis from most other branches of statistical analysis is that at the end of the study or at the time of analysis we dc not have a failure time for some of the cases. That is we know that they have survival up until the last follow-up and also know that they will fail in the future. This effect is known as censoring of the failure times and will be discussed for the rest of this section.

For each case we will have a time y_i or c_i, available, indicating resprctively that the observed time was terminated by a failure or that the case has not had enough follow-up time to produce a failure. In industrial applications two types of censoring namely type I and type II are usually used. Both of these types of censoring imply that all cases are put on trial simultaneously at time zero. If/

If a fixed maximum time of failure is considered sufficient before the end of a trial we will have a type I censoring and if the stopping criterion is taken to be the ratio of censored to sample size we will have a type II censoring. An example is the situation of monitoring a set of light bulbs on time. We will not develop these concepts any further but continue with a form of censoring that will be used later in the thesis.

In biomedical applications a different type of censoring is produced by the data and usually named as random censoring. Patients are entered into a trial at different times and then are observed after treatment for a number of years. We therefore have a time t_i for case i and it is,

$$t_i = \text{Min}\,(y_i , c_i) \text{ , that is we observe either censoring or failure whichever is first.}$$

and

$$\delta_i = \begin{cases} 0 \text{ if } y_i > c_i & , \ t_i \text{ is censored} \\ 1 \text{ uf } y_i \leq c_i & , \ t_i \text{ is not censored (This notation will be used in the development of various models)} \end{cases}$$

In practice some further complications arise. What we have defined to be death or censoring can be in fact a subset of a final outcome of a more complex process with more end points.

For example at times a patient decides to leave the geographic area within which a trial is prepared and thus the case is a lost to follow-up. At times terminations other than one

of/

of interest occurs; say a death from a second illness or a car accident and thus the final result can be open to different interpretations.

Generally we assume that censoring times are independent of death times. This assumption is quite valid for most trials. However if dropping out of the treatment is more common for one arm of trial the effect is at least loss in efficiency and more seriously a possible introduction of bias. In the thesis we will also discuss the possibility of analysis of data with more than one type of failure and in these contexts certain types of dependence on death times and censoring can be tested.

1.5 Outline of Thesis.

Ethics and certain scientific stands give trials properties that are slightly different from scientific experimentation in the natural sciences. The role of large, cheap and accessible computer information banks and fast end processing is new to this area and is changing the statistical methodology which can be applied. The recent developments in the field of failure time analysis originating mainly from Cox (1972) and his proportional hazard models are the main subject of our discussion.

We will study the applications of this model to clinical trial data with various forms of interrelations. The variety of interrelations will be defined to be both in terms of covariate effects and actual events where more than one event is present on the time scale. Further we will study the flexibility of the method/

method in dealing with the different forms of interrelationship that arise in situations where the regression parameters are not necessarily fixed and their influence can best be described in terms of a process through time rahter than a cause and effect situation. The major emphasis of our discussion will be on the exploratory use of the analysis and the variety of the models available in the framework of proportional hazards. At times when the limitation of the proportional hazard method makes them inappropriate for example in the study of the actual distribional shape of the hazard rates, we will discuss altervative methods.

In the context of the proportional hazard models with intervening events and time dependent covariates, there is a deviation from the traditional regression approach. Proportional hazards do not provide the same restrictive assumptions in the distinction between the exogenous and endogenous variables (in this framework fixed covariates and final response times). With the use of this extra flexibility a greater number of models are available for analysis with the proportional hazard assumptions. The need for this flexibility aims at a different interpretation of cause and effect, and more on the interpretation of the structure of change.

Although from a scientific point of view it is desirable that all measurements are made under uniform conditions throughout the trial, it may not be possible to isolate totally clinical judgement which may change with experience, from clinical measurement. Further the individual behaviour of the patients and also the long time scale in data collection may also play a role/

role in the different pattern of development occuring for different subgroups of patients. We will not analyse data according to all of the above possibilities. However time dependency does provide a good construct for such an analysis. We will be looking at certain aspects of time dependengy within the thesis.

In outline the structure of the rest of the thesis is as follows:- In chapter two we will consider the non parametric methods and their advantages. Also in this chapter we will study a general group of tests that have been used in the analysis of trial data. Chapter three deals with parametric methods and the various advantages and the disadvantages of the exponential, Weibull and a few less known but more complex distribution. In chapter four we deal with the semi-parametric proportional hazard of Cox (1972). We will consider various regression forms of the proportional hazard models and consider its position in relation to different parametric and non-parametric methods. Chapter five considers a realistic simulation method for the generation of clinical trial data. These simulations are carried out for treatments and one covariate parameter in the presence of proportional hazard assumptions and deviation from it, using the various approaches of analysis described above. Chapters six and seven consider and analysis data from a clinical trial which was organised in Edinburgh. In particular in chapter seven we will consider multivariate approaches to survival analysis and the effect of different intervening events in the patient progress. Chapter eight considers a different data/

data set for the purpose of study of various prognostic indicators and the use of time dependency in prognosis. Finally, in chapter nine we will bring together the findings from the earlier chapters.

C H A P T E R 2

NON- PARAMETRIC METHODS

2.1 Initial Developments of Life Tables.

In this chapter we introduce the non-parametric methods of analysis of survival data. These methods are closely related to the life tabel method originally proposed by Halley (1693) and which was mentioned in the introduction. Such life tables according to their particular applications have been refered to as population life tables, clinical life tables and cohort life tables. We do not intend to discuss the difference between the applications but to concentrate on the clinical life tables, because of their relevance to failure time data. In here however we generalise the area of applications by rephrasing to length of stay in a particular state; for example time from entry to a hospital to time of death or operation.

Some of the developments outside the field of clinical life tables such as those of competing risks are relevant to multiple failure time analysis and we refer to these methods in chapter 6, as multivariate competing risks. There have also been developments in which some of the methodology and techniques initially used in the analysis of failure time data have found use /

use in life tables outside of clinical studies. An example is Breslow (1982) on the use of Cox's method for cohort studies. We will return to this area again later in chapter 4 when we allocate a chapter to Cox's approach. Although at times we refer to similar developments in neighbouring fields we concentrate on applications to clinical trial data methodology. Two other types of life tables that are used extensively in other applications are population life tables and cohort tables. The population life tables require two sources of data. These are (a) census data on number of individuals alive in a particular age group and (b) vital statistics on number of deaths in a given year for each age group. Cohort studies on the other hand concentrate on describing the actual survival experience of a group born at about the same time.

In clinical life table data we use data from a group of patients and the data refers to entry to a particular state until removal from that state. The nature of removal from the state however often has to be conditional either on removal due to response, e.g. death, progression of disease or it can be a removal due to withdrawal, censoring, death from other causes etc. Further we are interested in a comparison of two or more treatments and thus the analogy with population studies is not carried further. In population studies one is oten comparing a survival rate of a group with a rate from census data or vital statistics, much of which is historically based information. In trials however, we refer to two arms of a trial. The comparison of interest is performed based on a measure/

measure of difference between rates of failure between the two treatment arms. There are some trials based on more than two treatment options but the principle of the analysis is the same. In terms of structure however the method of clinical life tables as proposed originally are similar to population life tables in that the data is grouped into intervals and that the probability of survival is estimated for each time interval. Chiang (1966) produced variance estimates of the probability of survival at any fixed point in time for describing the two treatment groups. Later it was discovered by Kuzma (1967) that these estimates can underestimate the variance of survival probabilities quite significantly if the censoring percentage is high.

Better estimates of survival rates can be obtained by use of parametric methods, given that the relatively restictive assumptions of the parametric distribution of interest are not violated. This conflict of interest between robustness of an estimating procedure versus its efficiency is part of the disussions in chapter 3 and 4.

2.2 Product Limit of Survival Times.

In this section we will describe the product limit estimate or the Kaplan and Meier estimate and later show that by the method of Johanson (1978) that the product limit estimator can be derived as maximum likelihood estimators. The product limit estimate of survival for n observed response times and censoring/

censoring times was initially proposed as a descriptive method rather than a method of inference. However recently it has become the most commonly used method of estimation of life tables in the context of clinical forms of survival data.

The product limit method is different from the methods of the previous section in that rather than using a fixed time interval it is based on forming a rank set of survival times in such a way that, for equivalent death and censoring times it is defined that the censoring times should have a rank greater than its equivalent death time. The product limit estimators are of special interest in that they form the basis of a large number of non parametric tests and are closely linked to the proportional hazard model. We will now proceed with a derivation of the product limit estimators using the maximum likelihood estimation. Throughout we will assume a continuous time scale so that there are no tied events for each rank. All results however may be generalised to tied data with slight extensions.

First we order the survival data into a rank order.

$$t_{(1)} < t_{(2)} < \ldots < t_{(n)}$$

Further for each $t_{(i)}$ there exists an indicator variable $\delta_{(i)}$ such that,

$$\delta_{(i)} = \begin{cases} 1 \text{ if } t_{(i)} \text{ is a response} \\ 0 \text{ if } t_{(i)} \text{ is a censored observation.} \end{cases}$$

we then define

$$P_i = P \left[\text{Surviving at } t_{(i)} \text{ given survival until } t_{(i)} \right]$$

$$= P \left[T > t_{(i)} \setminus T \geq t_{(i-1)} \right]$$

giving

$$\hat{P}_i = \begin{cases} 1 - \frac{1}{n_i} & \text{if } \delta_{(i)} = 1 \\ 1 & \text{if } \delta_{(i)} = 0 \end{cases} \quad \text{where } n_i = n-i+1,\ i = 1 \ldots n \qquad (2.2.1)$$

From definition of chapter 1 we obtain

$$\hat{S}(t) = \prod_{t_{(i)} \leq t} \hat{P}_i = \prod_{t_{(i)} \leq t} \left(1 - \frac{1}{n_i}\right)^{\delta_{(i)}}$$

$$= \prod_{t_{(i)} \leq t} \left(\frac{n - i}{n - i + 1}\right)^{\delta_{(i)}}$$

Then the corresponding estimate of standard error is

$$\text{s.e. } [\hat{S}(t)] = \hat{S}(t) \left\{ \sum_{t_{(i)} \leq t} \frac{\delta_{(i)}}{(n - i)(n - i + 1)} \right\}^{\frac{1}{2}}$$

We will now express the survival distribution in terms of a likelihood function based on product limit estimates and show that(2.2.1) can in fact be considered as the maximum likelihood estimates of the likelihood function.

Likelihood L = (terms due to cases dead) (terms due to censored times)

$$L = \prod_{i=1}^{n} \Pr [T = t_{(i)}]^{\delta_{(i)}} \Pr [T > t_{(i)}]^{1-\delta_{(i)}}$$

$$= \prod_{i=1}^{n} [\Pr (T = t_{(i)})^{\delta_{(i)}} \sum_{j \geq i}^{n} \Pr (T = t_{(i)})^{1-\delta_{(i)}}]$$

We let R_i be probability of an event at T giving

$$L = \prod_{i=1}^{n} R_i^{\delta_{(i)}} \left[\sum_{j \geq i} R_j \right]^{1-\delta_{(j)}}$$

we can thus define the hazard rates conveniently as

$$\lambda_i = R_i \left[\sum_{j \geqslant i}^{n} R_i \right]^{-1} \quad \text{with} \quad \sum_{j=1}^{n} R_j = 1$$

we thus have

$$1 - \lambda_i = 1 - \frac{R_i}{\sum_{j \geqslant i} R_j} = \frac{\sum_{j \geqslant i+1} R_j}{\sum_{j \geqslant i} R_j} \qquad (2.2.2)$$

and

$$\sum_{j \geqslant i+1} R_j = (1 - \lambda_i) \sum_{j=i}^{n} R_j = (1 - \lambda_i)(1 - \lambda_{i-1}) \sum_{j=i-1}^{n} R_j$$

$$= (1 - \lambda_i)(1 - \lambda_{i-1}) \ldots (1 - \lambda_2) \sum_{j=2}^{n} R_j$$

$$= (1 - \lambda_i)(1 - \lambda_{i-1}) \ldots (1 - \lambda_2(1 - \lambda_1) \sum_{j=1}^{n} R_j$$

$$= \prod_{j=1}^{i-1} (1 - \lambda_j) \qquad \text{since} \quad \sum_{j=1}^{n} R_j = 1 \qquad (2.2.3.)$$

Also

$$\lambda_i = \frac{R_i}{\sum_{j \geqslant i} R_j} = \frac{R_i}{\prod_{j=1}^{i-1} (1 - \lambda_j)}$$

giving $$R_i = \lambda_i \prod_{j=1}^{i-1} (1 - \lambda_j) \qquad (2.2.4)$$

Thus expressing the likelihood in terms of the hazard rates

we get

$$L = \prod_{i-1}^{n} \left(\lambda_i \prod_{j=1}^{i-1} \{1 - \lambda_j\} \right)^{\delta(i)} \left(\prod_{j=1}^{i-1} \{1 - \lambda_j\} \right)^{1 - \delta(i)}$$

$$= \prod_{i=1}^{n} \left[\lambda_i^{\delta(i)} \left(\prod_{j=1}^{i-1} \{1 - \lambda_j\} \right) \right]$$

$$= \lambda_1^{\delta(1)} \circ \lambda_2^{\delta(2)} (1-\lambda_1) \circ \lambda_3^{\delta(3)} (1-\lambda_1)(1-\lambda_2) \ldots \lambda_n^{\delta(n)} (1-\lambda_1) \ldots \ldots (1-\lambda_{n-1})$$

But $\lambda_n^{\delta(n)} = 1$ by definition

Therefore

$$L = \left(\lambda_1^{\delta(1)} \ldots\ldots \lambda_n^{\delta(n)} \right) (1-\lambda_1)^{n-1} \circ (1-\lambda_2)^{n-2} \circ (1-\lambda_{n-1})^{n-(n-1)} \circ (1-\lambda_n)^{n-n}$$

$$= \prod_{i=1}^{n-1} \lambda_i^{\delta(i)} (1-\lambda_i)^{n-i}$$

Now we take logarithm of the likelihood in order to obtain a maxima with respect to λ_i

$$\frac{\delta \text{ Log } L}{\delta \lambda_i} = \frac{\delta_{(i)}}{\lambda_i} + \frac{(n-i) \circ (-1)}{1-\lambda_i} = 0$$

$$= (1-\lambda_i)\delta_{(i)} + (n-i)(-1)\lambda_i = 0$$

$$\Rightarrow \quad \hat{\lambda}_i = \delta_{(i)} / (\delta_{(i)} + n - 1)$$

But/

But $\delta_{(i)}$ is either 0 or 1, giving

$$\hat{\lambda}_i = \frac{\delta_{(i)}}{1 + n - i} \qquad (2.2.5)$$

For the product limit estimator by definition & (2.2.5) we get

$$\hat{R}_i = \hat{\lambda}_i \prod_{j=1}^{i-1} (1 - \hat{\lambda}_j)$$

$$= \frac{\delta_{(i)}}{1 + n - i} \prod_{j=1}^{i-1} \left(1 - \frac{\delta_{(i)}}{n - j + 1}\right)$$

Thus by definition of $\delta(t)$ we have

$$\hat{R}_i = \frac{\delta(i)}{1 + n - i} \prod_{j=1}^{i-1} \left(\frac{n - j}{n - j + 1}\right)^{\delta_{(i)}} \qquad (2.2.6)$$

Therefore we conclude that the $\hat{R}_i$ are the required maximum likelihood estimators of survival times.

2.3 Nonparametric methods for two treatments

First we consider the log rank test, Peto (1972 a), which is also named mantel, Mantel-Haenzel, Mantel-Peto-Cox, or Savage-Mantel-Peto-Cox statistics. In this method, which is based on the observed and expected values of numbers of events in a particular time (under the Null hypothesis) we derive a form of chi squared test which is indirectly related to the ranks of survival times. The ranks are then transformed to a comparative ratio of numbers responding and numbers at risk. By this method any probability value that we obtain is used for the main objective of discrimination between treatments./

treatments. This enables us to infer that the difference between observed and expected values of the survival rates is either compatible with the Null hypothesis of no treatment difference, or that it is due to the effects of the alternative hypothesis that there are treatment differences. There are certain assumptions necessary for an analysis based on the log rank test. Later we will compare these assumptions with those of the Wilcoxon test and present a general form of test which incorporates both tests as special cases. At this stage however we only mention that the method can be derived and is related to Cox's proportional hazard model of Chapter 4.

Initially the same procedure as that of Kaplan and Meiers is used to transform the survival times. Similarly a vector or survival times is obtained based on

$$t_{(1)} < t_{(2)} < \quad . \quad . \quad . \quad . \quad . \quad . \; t_{(r)}$$

Thus at the beginning of each of these time points, say $t_{(j)}$ we form a 2 x 2 table to categorise the total number of patients at risk, according to treatment grouping and status at end of $t_{(j)}$ period.

	Number in group at time $t_{(j)}$	Observed events (deaths) in group at $t_{(j)}$	Alives
Group 1	N_{1j}	O_{1j}	$(N_{1j} - O_{1j})$
Group 2	N_{2j}	O_{2j}	$(N_{2j} - O_{2j})$
Total	N_j	O_j	$(N_j - O_j)$

Then we have a contingency table giving $E_{ij} = \frac{N_{ij}\, 0_j}{N_j}$ as the expected number of responses at t_j in group i.

We thus represent the above as a 2 x 2 x r table with the log rank statistics

$$X^2_{LR} = \frac{\left[\sum_{j=1}^{r} (0_{1j} - E_{ij})^2\right]}{\sum_{j=1}^{r} V_j} \equiv \frac{(0 - E)^2}{V} \qquad (2.3.1)$$

Using the hypergeometric distribution and the corresponding moment generating functions we have the first two moments giving

$$E_{ij} = \frac{N_{1j}\, 0_j}{N_j} \quad \text{and} \quad V_j = \frac{N_{1j}\, N_{2j} 0_j\, (N_j - 0_j)}{N_j^2\, (N_j - 1)}$$

which can be used in (2.3.1)

For a single level X^2 test where r = 1 we can then present a X^2 for a single level of j giving

$$X^2 = \frac{N_j\, (0_{1j}\, [N_{2j} - 0_{2j}] - 0_{2j}\, [N_{1j} - 0_{1j}]\,)^2}{N_{1j}\, N_{2j}\, 0_j\, (N_j - 0_j)}$$

Now by referring to the table of the chi square distributions with one degree of freedom we can accept or reject the null hypothesis of equality of survival rates for the two treatment groups against the alternative of different survival rates.

The log rank test is based on the Kaplan and Meier estimates. It acts indescriminantly in combining expected rates of number of failures. Like the Kaplan and Meier estimates the expected values of numbers of events in each category is obtained by a ratio of numbers/

numbers of events by the number at risk. However in some circumstances a more efficient estimation of the survival differences may be possible if a weighting is attached to the expected number of events. We will consider those conditions later in this section when we deal with Gehan's generalisation of the Wilcoxon test.

The special property of the Wilcoxon test is that continuously the contribution to the likelihood is weighted by the total number at risk at $t_{(j)}$. This statement is analogous to a special form of time dependancy in proportional hazards. In terms of interpretation however the null hypothesis is slightly different between the two tests in that for the Wilcoxon test the null hypothesis is based on the equality of the survival rates between the two groups together with equality of the censoring rates. In the log rank test this latter assumption is not required.

Thus in the Wilcoxon test early events are weighted slightly higher than late events. The log rank test may be expressed in vector form by -

$$X^2 = (O - E)' [V]^{-1} (O - E)$$

The notations for E, O and V are expressed in matrix form below for a similar expression of the Wilcoxon test; equivalently the Wilcoxon test can be expressed as -

$$[\ell' (O - E)]' [\ell' V \ell]^{-1} [\ell'(O - E)]$$

where for a comparison of two treatments we let,

$$\ell' = (\ell_1 \ldots \ell_r),\ O' = (O_{11} \ldots O_{1r}),\ E' = (E_{11} \ldots E_{1r})$$

and/

and

$$V = \begin{bmatrix} V_1 & & \underset{\sim}{0} \\ & \ddots & \\ \underset{\sim}{0} & & V_r \end{bmatrix}$$

Further ℓ_j is set to N_j numbers at risk at $t_{(j)}$

The above formulation was first used by Tarone (1975) on a test for departure from trends. Tarone and Ware (1977) show that the difference between the logrank rest and the Wilcoxon test is in fact due to the choice of weights as a function of the number of individuals at risk at the time of each death. Once again with one degree of freedom we have a chi-square ,

$$X^2_{TW} = \frac{\sum_{j=1}^{r} w_j \, (O_{1j} - E_{1j})^2}{\sqrt{\sum_{j=1}^{r} w_j^2 \, V_j}} \qquad (2.3.2)$$

where

$$V_j = N_{1j} \, N_{2j} \, O_j \, (N_j - O_j) \, / \, N_j^2 \, (N_j - 1)$$

and $E_{1j} = N_{1j} O_j \, / \, N_j$

Thus this general result gives the logrank test for $W_j = 1$ and the Wilcoxon test for $w_j = N_j$. Taron and Ware suggest a different function of weights, namely $w_j = \sqrt{N_j}$ and claim that it has better efficiency over a range of alternatives.

The above approach is closely related to time dependency scaling of the proportional hazards. The Wilcoxon test considers the distribution of censoring times as well as death times. There is however no reason why w_j must be defined as a function of N_j alone/

alone. In fact later in chapters on proportional hazards we will consider time dependencies as a function of metastatic or other intervening events.

Now we expand the logrank and Wilcoxon test to the multivariate situation. For this purpose the general Tarone and Ware statistics generalisation is used. We will continue with O_{ij} , E_{ij} and V_{ij} presentations.

In cases where there are a number of subgroups we present for say a set of different levels of an outcome, the following formulations.

Level	d_0 d_1 . . . . . d_k	Total at $t_{(j)}$
Event	O_{0j} O_{1j} O_{kj}	O_{+j}
at risk	N_{0j} N_{1j} N_{kj}	N_{+j}

We then have a longrank type null hypothesis for the equality of survival rates 0 to k and a Wilcoxon type null hypothesis for equality of all survival rates and also all censoring rates.

Once again we have a chi-squared test with (k - 1) degrees of freedom,

$$X^2 = (O - E)' \quad V^{-1} \quad (O - E) \qquad (2.3.3)$$

where

$$(O - E) = \sum_j w_j \quad (O_{ij} - E_{ij})$$

By the first 3 moments of the hypergeometric distributions we get - /

get -

$$E_{ij} = N_{ij} \; 0_{+j} \; / \; N_{+J}$$

$$V = \sum_{j} w_j^2 \sum_{i} V_{ij}$$

and

$$\sum_{i} V_{ij} = \left(\frac{0_{+j}\,[N_{+j} - 0_{+j}]}{N_{+j} - 1}\right) \begin{bmatrix} \frac{N_{0j}}{N_{+j}}\left(1 - \frac{N_{kj}}{N_{+j}}\right) \cdots & \frac{-N_{pj}N_{\ell j}}{N_{+j}^2} \\ \vdots & \vdots \\ -\frac{N_{pj}N_{\ell j}}{N_{+j}^2} \cdots \frac{N_{kj}}{N_{+j}} & \left(1 - \frac{N_{kj}}{N_{+j}}\right) \end{bmatrix}$$

Now referring to (2.2.1) and (2.2.2) it can be seen that (2.2.3) is a generalisation of the previous tests. Once again $w_j = 1$ gives a general form of the logrank tests. $w_j = N_{+j}$ gives the Wilcoxon test and $w_j = \sqrt{N_{+j}}$ gives a Tarone and Ware type statistic.

2.4. Stratification.

In the introduction we mentioned uses of stratification in conjunction with randomisation, and considered it to be a proper method of conduction of a trial at times. We did not consider the necessary analytical techniques in the development of the methods. We will now consider stratification methods in conjunction with the non parametric methods of analysis which can describe the general advantages of stratification.

In many trials apart from the treatment assessment information an array of different types of exploratory data is

also/

also collected on patients, often referred to as prognostic indicators or covariates. A few examples of each data that we will be referring to are , age, node status, size of tumour etc. This type of covariate information is a reflection of the underlying make up of the group of patients to whom the inference is relating. A proper randomisation in a large sample would imply that the patient variability between the two groups are suitable. In some trials however purely leaving the allocation of patients to randomisation may not provide a satisfactory final outcome of the patient mix. In practice the type of adjuvant care or therapy can be dependent on the prognostic conditions of the patient; this condition can provide a framework by which the two arms of the trial are not comparable. An example is a situation where the amount of radio-therapy given may be influenced by size of the tumour, and thus the size of the tumour may mainly influence the survival rates of the two arms of a trial. In other situations where there is a perfectly standard treatment for all patients, it may be known from preset that a group of patients that have less advanced disease, will be generally better in survival regardless of the type of treatment. Such differences can lead possibly to a correlated prognostic and treatment effect and furthermore may bias the inference. The remedy is often a prospective stratification. The utilisation of stratification has been subject of some controversy. Peto et al (1976) considers stratification often as unnecessary and unjustified administrative inconvenience. The basis of this view is that for large trials often the gain in power of tests is nominal where randomisation guarantees comparable/

comparable treatment groups. An alternative view which is in favour of stratification considers, firstly small trials to be common in practice and an important part of research, secondly for large trials an interim analysis can be based on small numbers of patients which consequently may condition the conclusions on the type of patient mix.

It should be pointed out that although stratification adds a form of control on the randomisation procedure it in no way influences the chances of treatment assignment to a treatment arm. Apart from stratification at design the relation between stratification and analysis is also important and at this point we can make a few comments which can also apply to the methods that we will be considering later. In either of the situations where the sample size is large enough to achieve a balance of treatment arms in terms of prospective effects, or the situation of stratified trials and balanced prognostic effects, it is useful to account for any of the possible survival differences that as a priori is considered relevant to the trial. There are two major aims for this type of analysis which may not have been so clear from the discussion of stratification and design. Firstly, we may aim to study the whole patient group and so it may be of some importance to know characteristics of prognosis in different strat\ and account for heterogenity of patient survival rates. However, care is needed in the interpretation of such assessments for it is possible that any such inferences are conditioned on sample size and/or other prognostic effects. Secondly, it is possible to use prognostic factors to define the treatment comparisons/

comparisons, we may thus obtain a test statistic for each stratum of a prognostic variable, in order to compare treatment effects within each stratum. In the next section we will present the results of a trial data by which we obtain tests statistics by considering a fraction of data belonging to a particular prognostic category. It is then necessary to obtain an overall observed and an overall expected (extent of exposure) comparison. The overall test is important in that even if a treatment difference for a prognostic group is not significant, the overall test provides a test by which if the directions of the influence of the treatments are in the same direction they can give an overall influence. This overall test will then remove any consequences of a possible correlation between prognostic variable and treatment effect. An important condition where the above consideration is important is when the tests for the different strata in fact do not point to a treatment difference in the same direction. Once again we will study such effects more carefully in the data analysis section of this chapter and the subsequent chapters.

2.5 Comparative application of non-parametric tests.

Consider a trial in which a set of treatments have been allocated and further stratification has been performed on some of the prognostic variables, either prospectively or retrospectively. The development of the logrank rank test are then useful in expressing group differences in a single statistic. The development of the previous section reinforce the notion that the logrank test is a useful test for trials and further that it can be set within a larger theoretical framework.

In/

In practice we are interested in a comparative study of the general expressions of (2.3.1), (2.3.2) and (2.3.3) as applied to a set of clinical trial data from Edinburgh. At this stage we will consider the relation of the special forms of the w vector of the last section to the various hazard rates of the prognostic group. We will not deal with inferences drawn for the shape of the hazard rates, since chapters 4 and 6 models are more appropriate for this. The actual data will be discussed in greater detail in chapter 7, with some history of the topics and problems related to the treatment of breast cancer.

Basically the data consists of 561 cases treated for breast cancer either by radical mastectomy or simple mastectomy plus radiotherapy to the axilla.

In here we record the survival times only and the primary purpose in the use of this data is to assess the relative merit of the arms of the trial for the total group of patients and then according to various prognostic factors. In chapter 7 we deal with the situation of more than one response variable and consider intervening events, such as development of local and metastatic disease.

The Wilcoxon as was explained attaches more weight to early events and thus gives a slightly different chi-squared value to the logrank test for most of the groups in our data. Earlier we presented these tests as 2 x 2 x r contingency tables. In fact the logrank is the most powerful test given that the

second/

second order interaction is negligible.

In general there may be more than one set of independent variables acting and thus we will perform an analysis based on the various subgroups of patients. Thus our data can be expressed by various probability values, related through a likelihood function by the following formulation.

Likelihood (particular subgroup) = Π terms due to cases dead
= Π terms due to cases censored
= Π terms due to time dependencies.

In the above likelihood formulation we have introduced time dependency. It is difficult to establish a complete meaning of time dependency without resorting to empirical hazard rates. We will do so in chapter 3. For the present section it is important to consider a comparison of the Wilcoxon and the logrank test, using a trial data. Such a comparison is intended to serve as a representation of the effects of the two tests for different hazard rates. As we will indicate in the discussion of the data the two tests produce the same interpretation of the data. On the use of time dependency however we confine their difference to that of having a different &vector of weights in the overall X^2 test. The effect of such weights is of importance only if the variability of the difference of the proportion of rates is of relevance.

In an extreme situation that is rarely detected in practice one may encounter crossing survival rates. However/

However in the comparison of the logrank and the Wilcoxon test any time dependency if it exists will be reflected by the influence of late events versus early events. For the purpose of estimation the logrank and the Wilcoxon test in fact ignore the last term of the likelihood. In fact practically for most situations one can assume that the effects due to the last term of the above likelihood are negligible. The slight difference that we will detect for the logrank and the Wilcoxon test is due to the structural differences between the two tests. This structural difference however is essential for the power of the tests in the presence of the most relevant alternative hypothesis. In the chapters on proportional hazards, time dependency effects can in fact be tested directly by more suitable methods.

In the first steps of the analysis we will obtain the product limit estimates of the two treatments and the corresponding hazard rates, Figs. (2.5.1) and (2.5.2). The mthod used for the plot of the hazard rates is described by Johnson and Johnson (1981). We use a grouping period of 30 months for all the hazard rate plots. By the use of the logrank test and then the use of the Wilcoxon test we obtain the probability values for the difference between the two survival rates, Table (2.5.1)

	No. of cases	No.of Responses	Expected Responses	X^2_{LR}	X^2_W	df	P_{LR}	P_W
Radical Surgery	288	135	162.05	10.04	12.23	1	.0015	.0005
Simple Surgery + XRT	273	161	133.95					

Table 2.5.1

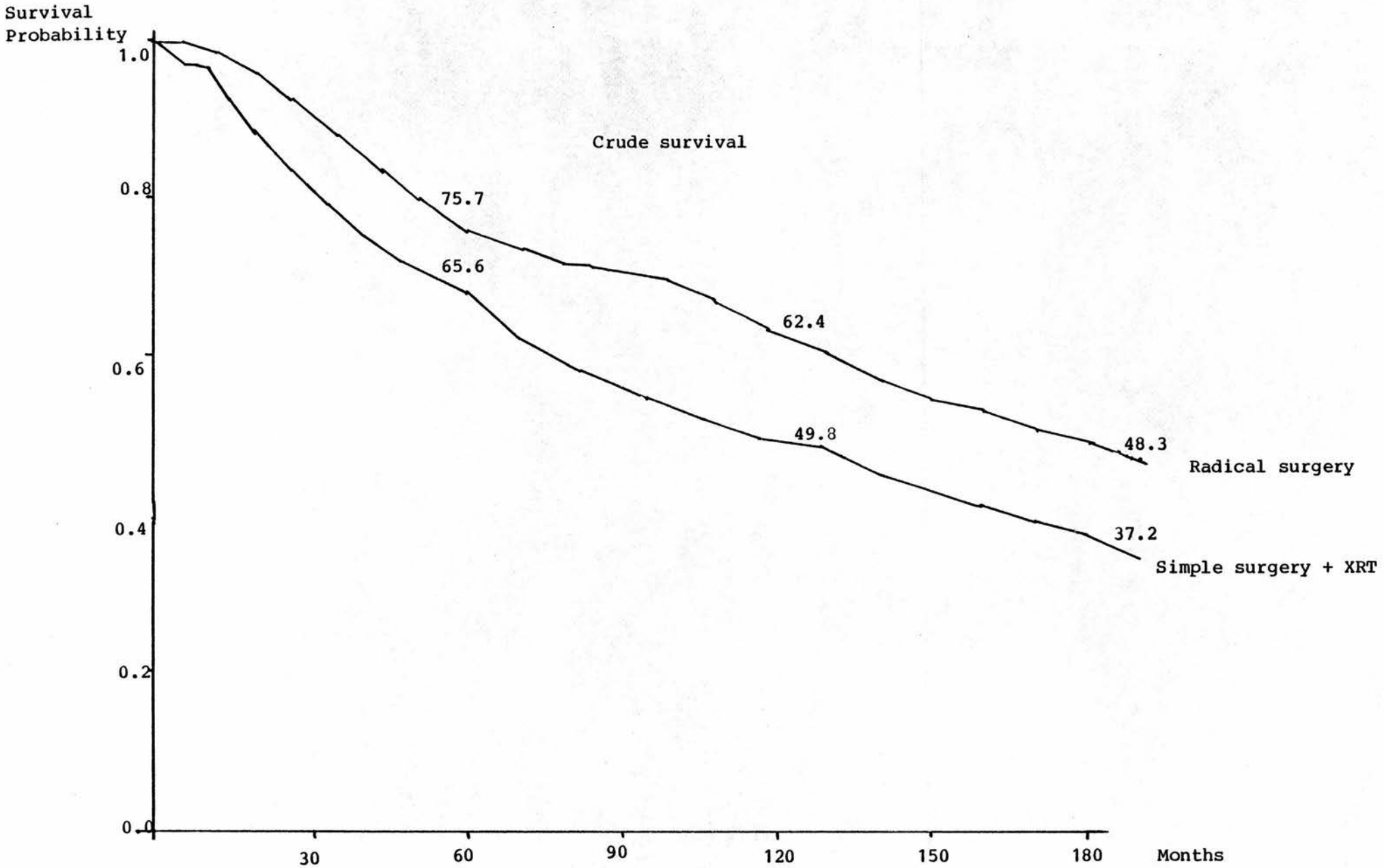

Survival
Probability
1.0
0.8
0.6
0.4
0.2
0.0
Crude survival
75.7
65.6
62.4
49.8
48.3
Radical surgery
37.2
Simple surgery + XRT
30
60
90
120
150
180
Months

Figure (2.5.1)

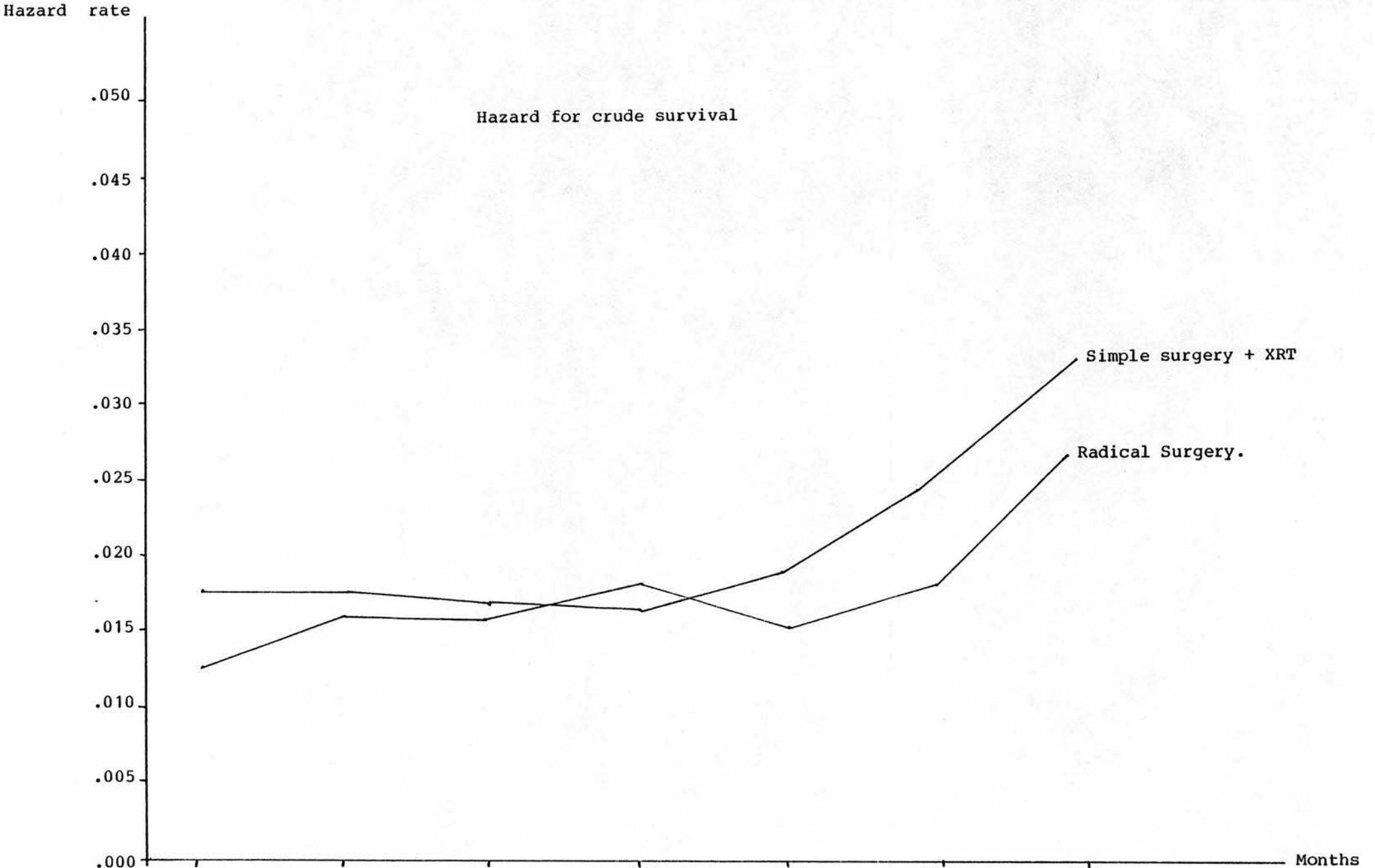

Hazard rate
Hazard for crude survival
.050
.045
.040
.035
.030
.025
.020
.015
.010
.005
.000
Simple surgery + XRT
Radical Surgery.
Months
15
45
75
105
135
165
195

Figure (2.5.2)

Later we use the modified version of the logrank and Wilcoxon test, so that apart from obtaining the difference between actual survival causes of the total treatment groups, we also obtain an overall treatment comparison adjusted for prognostic variability. In the process of obtaining the adjusted comparisons, a comparison for each level of the prognostic indicators is estimated and the final adjusted comparison is based on weighted differences of the observed and expected values of each subgroup.

The primary purpose in the comparison of each logrank statistic and the Wilcoxon test is to study a difference in their corresponding chi-squared and probability values. Further we examine, the shape of the hazard function and the association between the patterns of differences between the rates of failure and the way in which the Wilcoxon test puts more emphasis on early events.

Another manner of looking at the effects of time scale will be done in Chapter 7 by use of the regression like models of the life tables. In these models we relate the shape of the hazard and the time dependency indicators.

The prognostic indicators that we use are namely, Age, Node, Stage, Size and Menopausal status.

	No. of Cases	No. of Deaths	Expected No. of Deaths	X^2_{LR}	X^2_w	df	P_{LR}	P_w
Premenopausal	163	59	98.53					
Menopausal	38	20	20.26	25.12	21.97	2	.000	.000
Postmenopausal	359	216	176.21					
Pre & R	89	27	35.74					
Pre & S	74	32	25.26	3.17		1	.0752	
Meno & R	21	11	11.66					
Meno & S	17	9	8.34	0.09		1	.7618	
Post & R	178	97	115.21					
Post & S	181	119	100.79	6.21		1	.0127	
	ADJUSTED (R V S) for Logrank			9.05		2	.0021	
Node Status								
NØ	375	184	210.02					
N1	181	112	185.98	10.96	10.84	1	.0007	.0009
NØR	199	83	102.54					
NØS	179	101	81.46	8.56		1	.0034	
N1R	88	51	58.19					
N1S	93	61	53.81	1.90		1	.1683	
	ADJUSTED (R V S) for Logrank			9.94		1	.0016	
Tumour size								
T1	56	17	35.76					
T2	397	213	208.11	13.82	10.62	2	.0010	.0047
T3	107	65	51.13					

T1R/

Table Continued.

Tumour size (contd)	No. of Cases	No. of Deaths	Expected No. of Deaths	X^2_{LR}	X^2_w	df	P_{LR}	P_w
T1R	37	13	10.96)					
			)	1.08		1	.2981	
T1S	19	4	6.04)					
T2R	198	40	115.88)					
			)	12.76		1	.0004	
T2S	199	123	97.12)					
T3R	53	32	33.84)					
			)	0.04		1	.8343	
T3S	54	33	32.16)					
	ADJUSTED (R V S) for Logrank			8.4		1	.0038	
S1	307	147	171.76)					
			)					
S2	141	79	69.91)	9.41	10.97	2	.0090	.0067
			)					
S3	112	69	53.33)					
S1R	164	65	85.15)					
			)	11.41		1	.0007	
S1S	143	82	61.84)					
S2R	67	35	39.65)					
			)	1.11		1	.2914	
S2S	74	44	39.33)					
S3R	57	35	35.67)					
			)	0.03		1	.8714	
S3S	55	34	33.33)					
	ADJUSTED (R V S) for Logrank			9.38				
Age								
-40	31	14	17.51)					
			)					
40-50	168	66	96.40)					
			)	18.65	23.70	3	.0003	.0000
50-60	174	96	88.75)					
			)					
60+	188	120	93.34)					
-40 R	20	8	9.23)					
			)	.48		1	.4863	
-40 S	11	6	4.77)					
40-50/								

Table continued

Age (contd)	No. of Cases	No. of Deaths	Expected No. of Deaths	X^2_{LR}	X^2_w	df	P_{LR}	P_w
40-50R	89	35	35.65)					
			)	.03		1	.8721	
40-50S	79	31	30.35)					
50-60R	85	39	50.64)					
			)	5.70		1	.0170	
50-60S	89	57	45.36)					
60+R	94	53	67.07)					
			)	6.78		1	.0092	
60+S	94	67	52.93)					
	ADJUSTED (R V S) for Logrank			10.52		1	.0012	

Table (2.5.2)

A point to note regarding the tests in the table (2.5.2) is that we intend to compare the logrank test with the Wilcoxon test. The most noticeable source of discrepancy if any in terms of magnitude will be detectable in the study of the actual prognostic indicators rather than treatment comparisons. For this reason a comparison of the two tests based on prognostic differences will suffice. A consideration of the results of table (2.5.2) indicates that the radical surgery without radiotherapy is producing longer survival times than the simple surgery with radiotherapy.

The prognostic factors indicate that stage one, two and three are respectively ordered in terms of their progress of the disease and the later risks of development of the disease. The stage one group produce a treatment difference that is much greater than stage two and three tumours. This is an indication that the actual value of the stage may be interacting with treatment. It is not possible now to discuss this point further or substantiate with a formal test. In Chapter 6 we will do so.

The indication of different values of the treatment effects appears for some other prognostic indicators. Menopausal status and age indicate that post menopausal patients and for age over 50's group, the treatment differences are at their heighest. This, as was pointed out earlier, may be due to sample size rather than the treatment effects on prognostic strata. The number of patients in the menopausal group are rather low and thus with the present method we will not study the effect of treatment status further.

The/

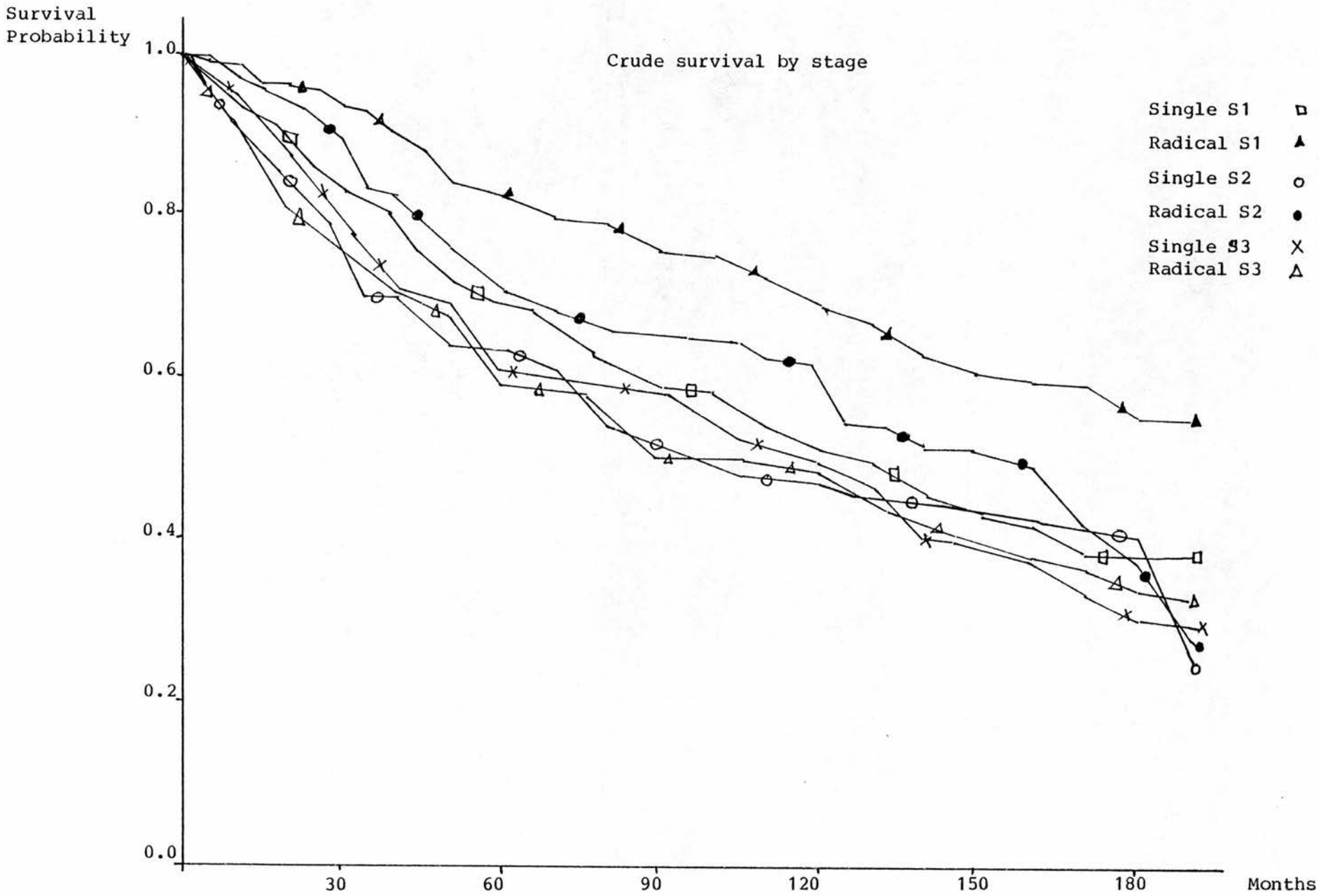

Survival Probability
Crude survival by stage
Single S1
Radical S1
Single S2
Radical S2
Single S3
Radical S3
1.0
0.8
0.6
0.4
0.2
0.0
30
60
90
120
150
180
Months

Figure (2.5.3)

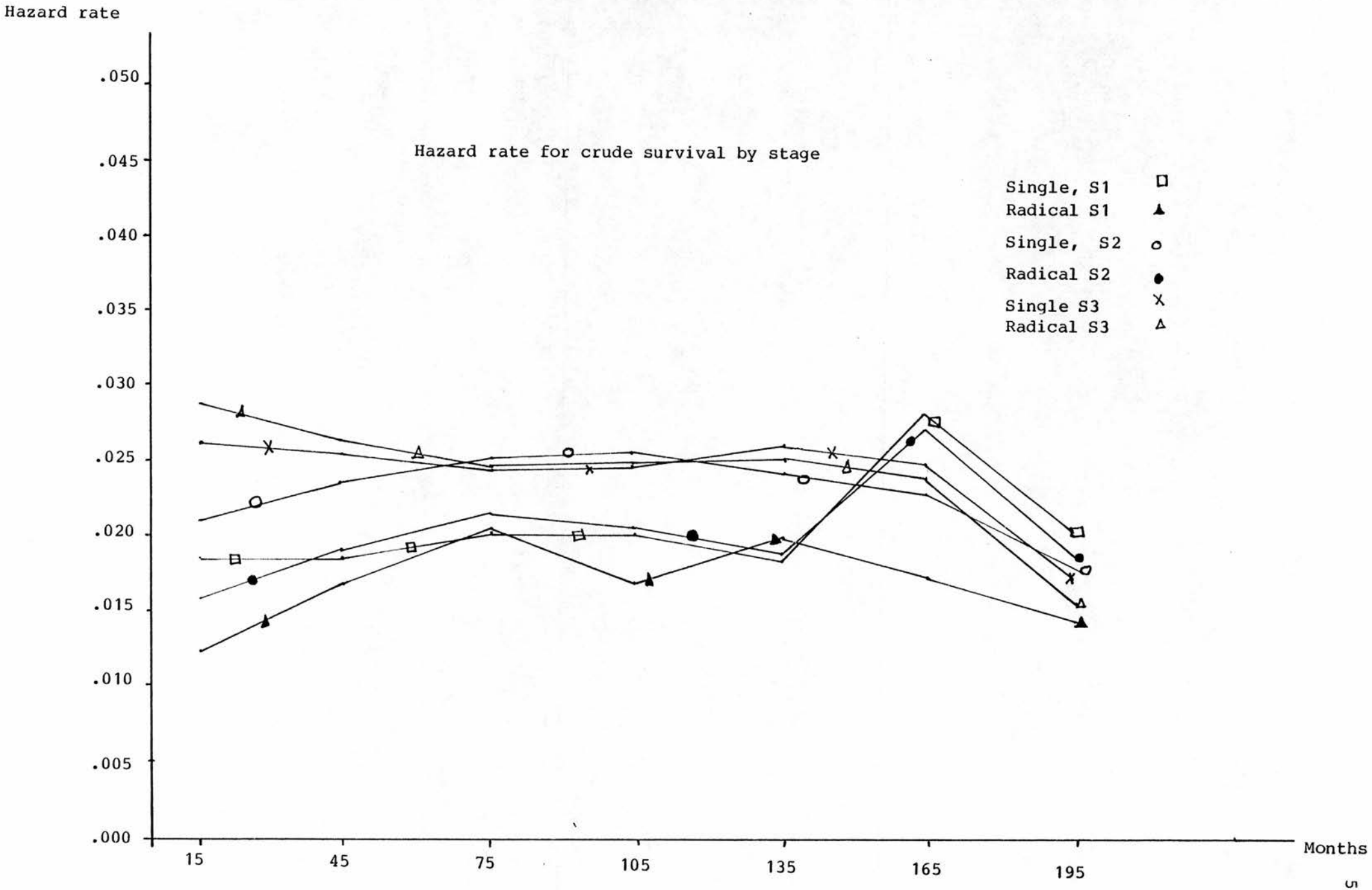

Hazard rate
Hazard rate for crude survival by stage
Single, S1
Radical S1
Single, S2
Radical S2
Single S3
Radical S3
.050
.045
.040
.035
.030
.025
.020
.015
.010
.005
.000
15
45
75
105
135
165
195
Months

Figure (2.5.4)

The adjusted rates on logrank statistics for each of the prognostic indicators carry the same message as the unadjusted rates. That is we detect a better survival rate for the radical surgery group.

In the case of age, the subgroups contain a reasonable number of cases in each category and a statement in a descriptive manner may be made regarding interaction between age and treatment. The survival rates for the 50-60 group give a significance level for the treatment difference of 0.0170 and for the 60+ group a level of 0.0092. The younger patients give probability levels that are not significant in terms of treatment differences. This effect is more notable for the 40-50 group. It must be noted however that this apparent difference is not a statistical indication of a difference in treatment effectiveness for the different age groups. Such formal tests will be performed in Chapters 6 and 7.

On considering the hazard rates and the corresponding logrank tests, there is an indication that the treatment effects are in a similar direction for all prognostic subgroups. However it must be pointed out that in terms of extent of the risks on the time scale they are not always similar. Figures (2.5.5) and (2.5.6) together with the corresponding logrank tests suggest that older patients produce a higher failure rate when they are treated with simple mastectomy and radiotherapy. Further there seems to be an indication that risks are reduced for the 50+ group 7 years after treatment, while risks remain the same for the rest. Figures (2.5.7) and (2.5.8) with the corresponding logrank tests suggest a similar pattern for the menopausal status, which conforms to the age/

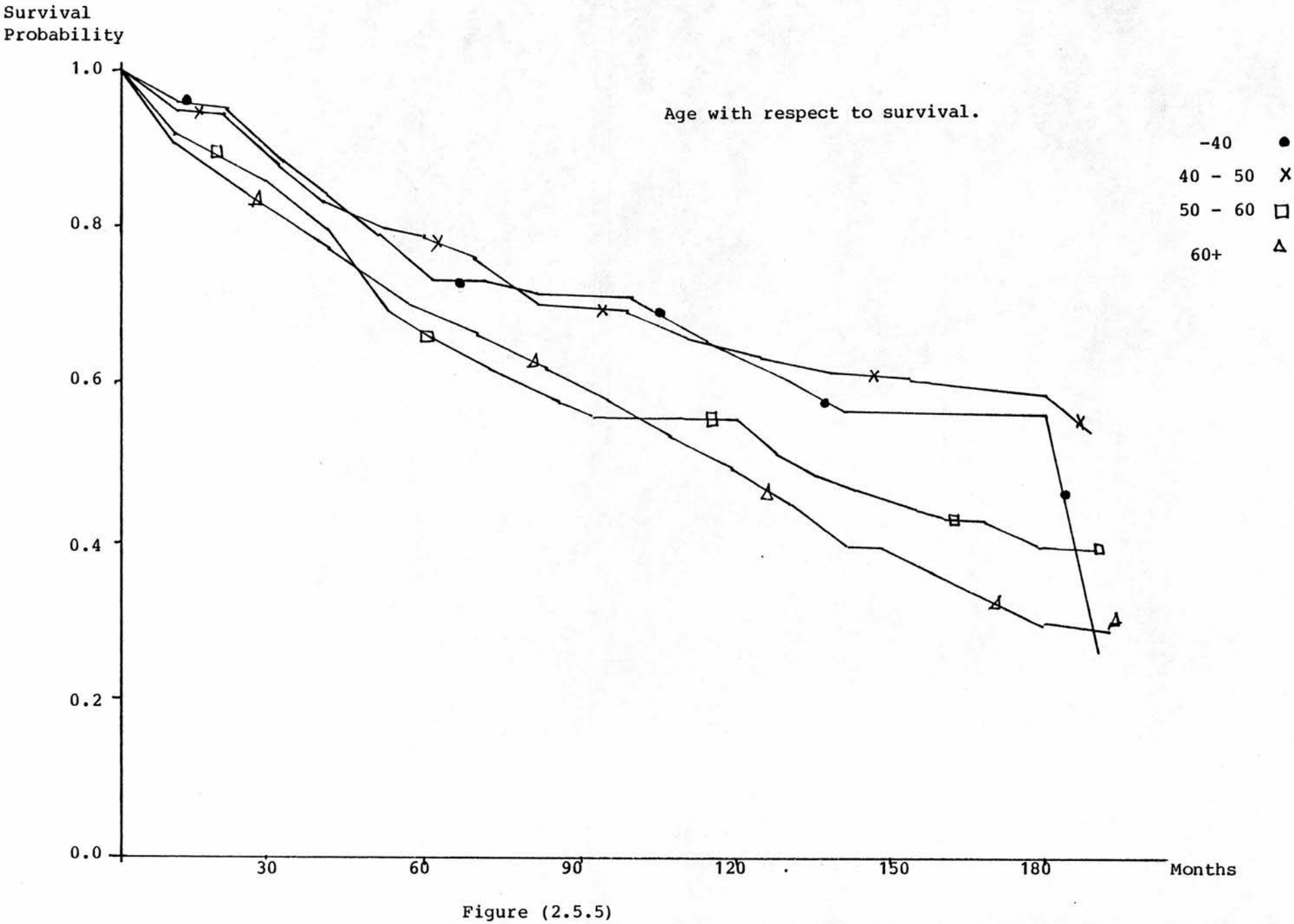

Survival
Probability
1.0
0.8
0.6
0.4
0.2
0.0
Age with respect to survival.
-40
40 - 50
50 - 60
60+
30
60
90
120
150
180
Months

Figure (2.5.5)

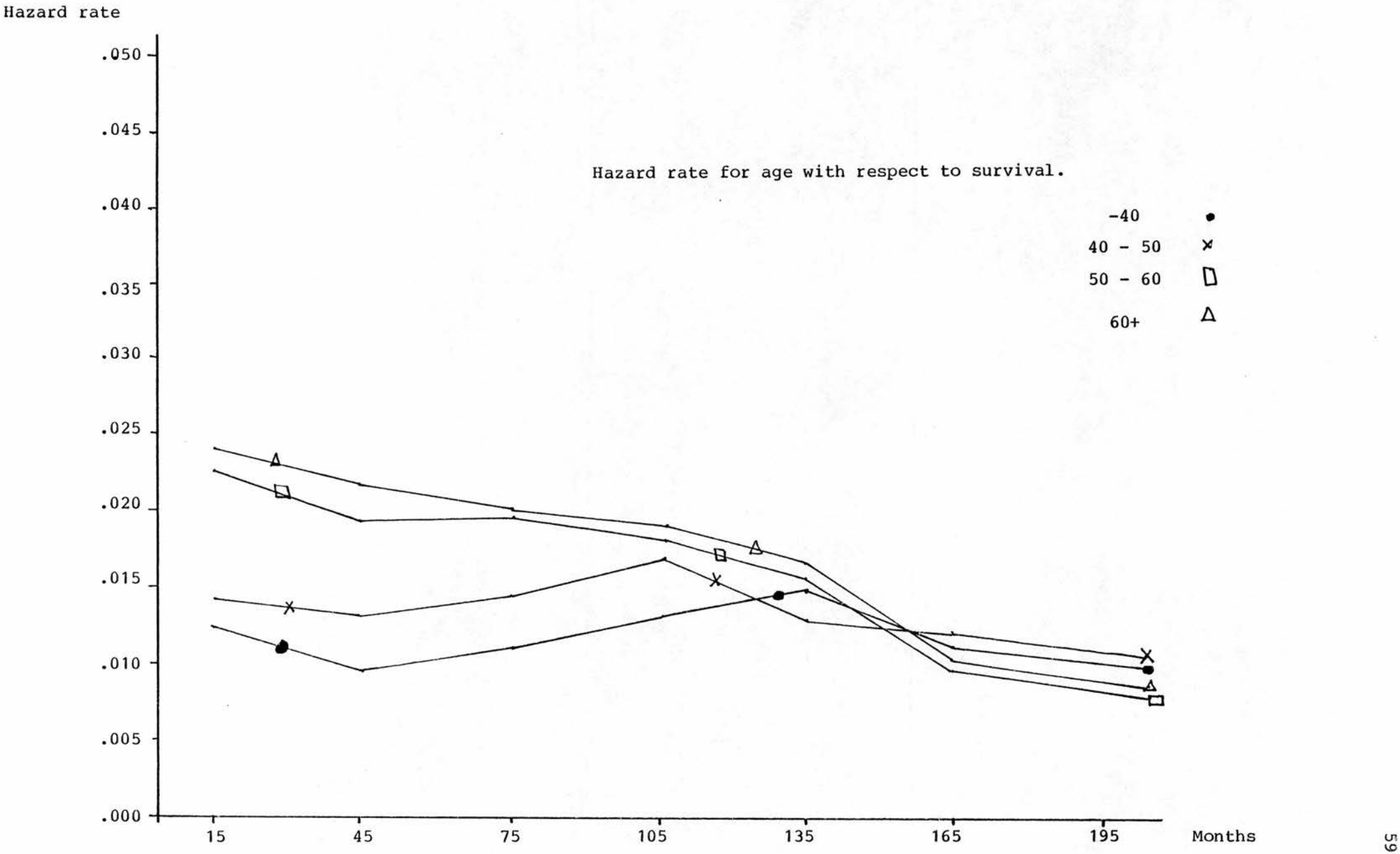

Hazard rate
Hazard rate for age with respect to survival.
-40
40 - 50
50 - 60
60+
.050
.045
.040
.035
.030
.025
.020
.015
.010
.005
.000
15
45
75
105
135
165
195
Months

Figure (2.5.6)

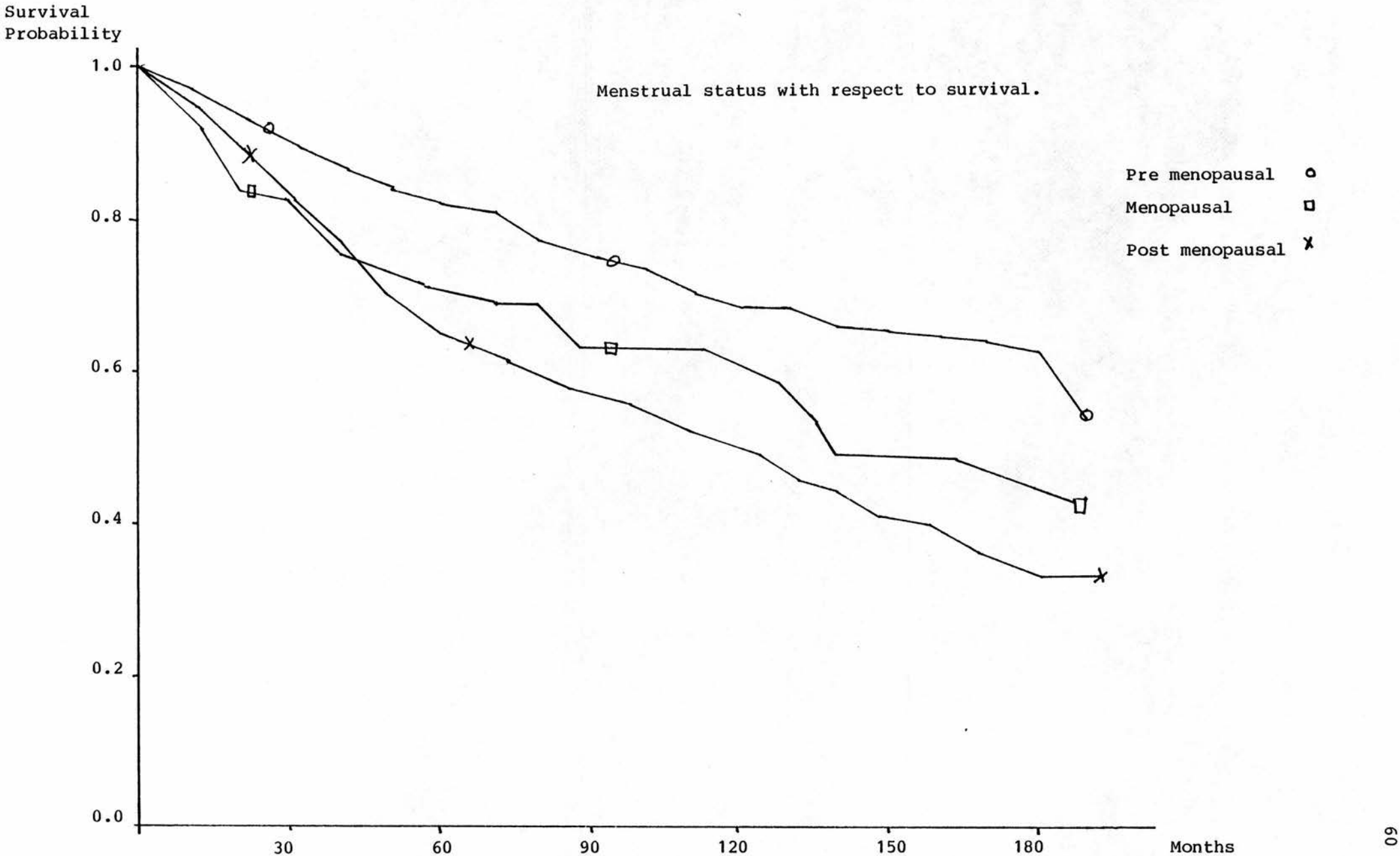

Survival
Probability
1.0
0.8
0.6
0.4
0.2
0.0
Menstrual status with respect to survival.
Pre menopausal
Menopausal
Post menopausal
30
60
90
120
150
180
Months

Figure (2.5.7)

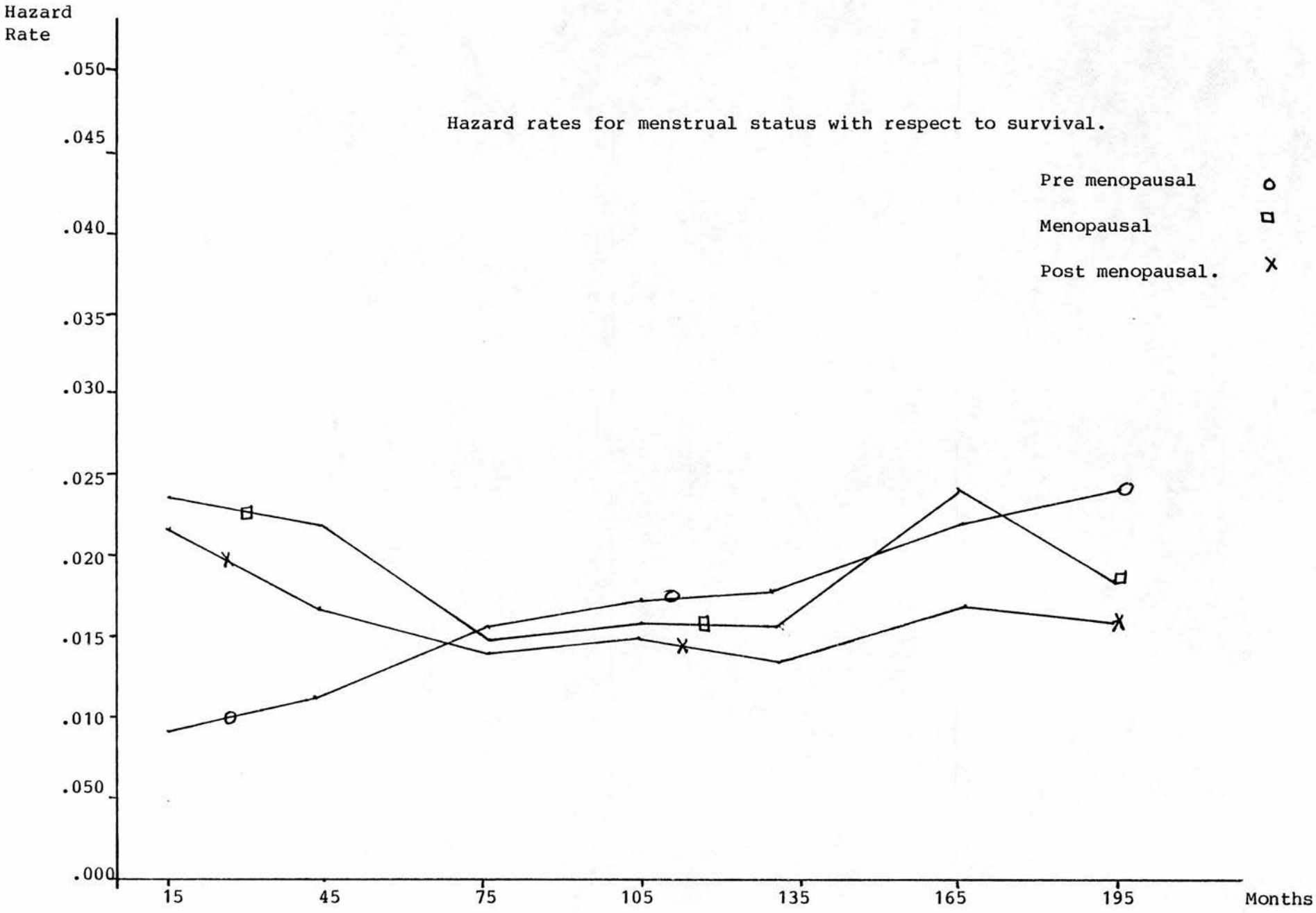

Hazard
Rate
.050
.045
.040
.035
.030
.025
.020
.015
.010
.050
.000
Hazard rates for menstrual status with respect to survival.
Pre menopausal
Menopausal
Post menopausal.
15
45
75
105
135
165
195
Months

Figure (2.5.8)

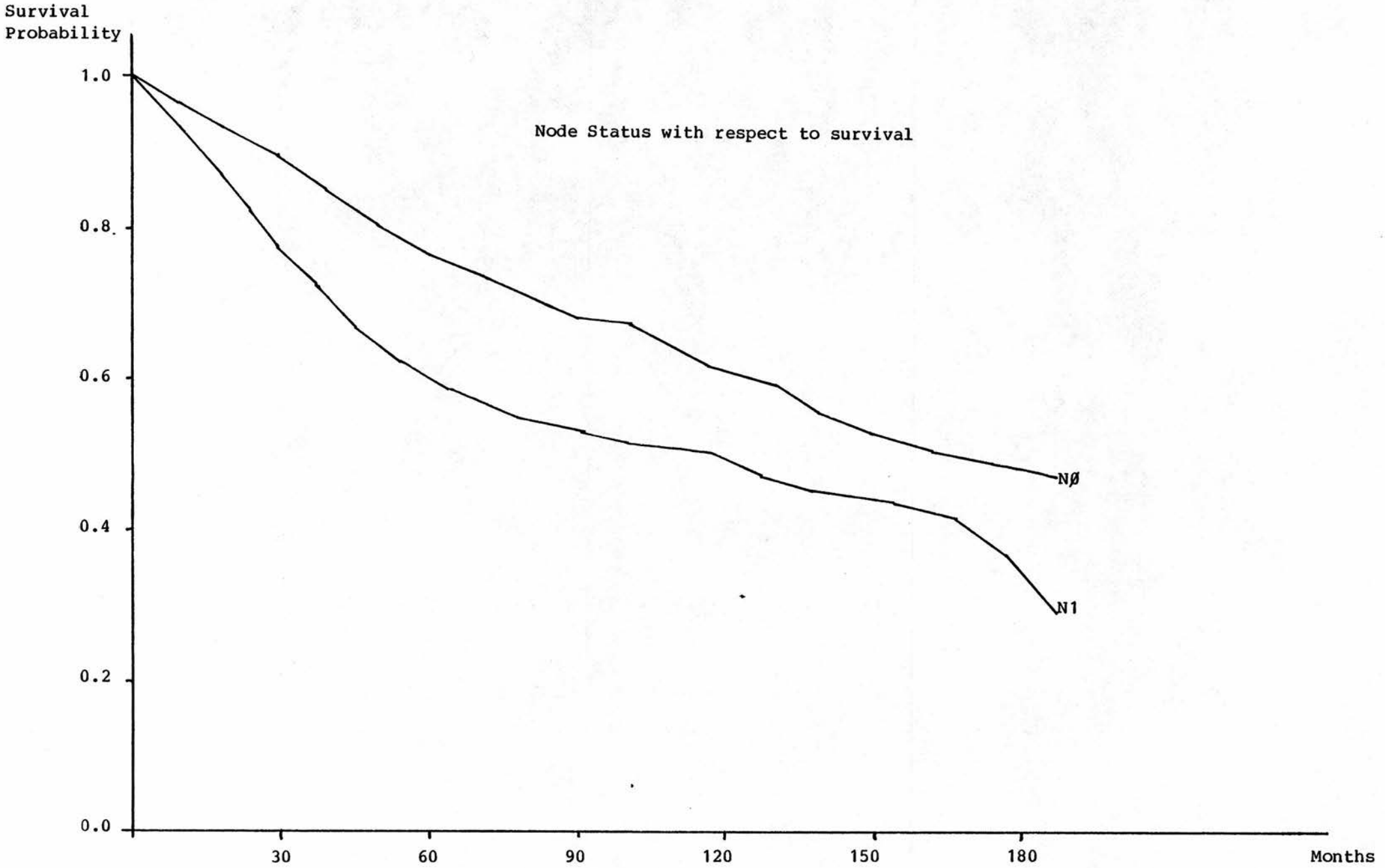

Survival Probability
1.0
0.8
0.6
0.4
0.2
0.0
Node Status with respect to survival
NØ
N1
30
60
90
120
150
180
Months

Figure (2.5.9)

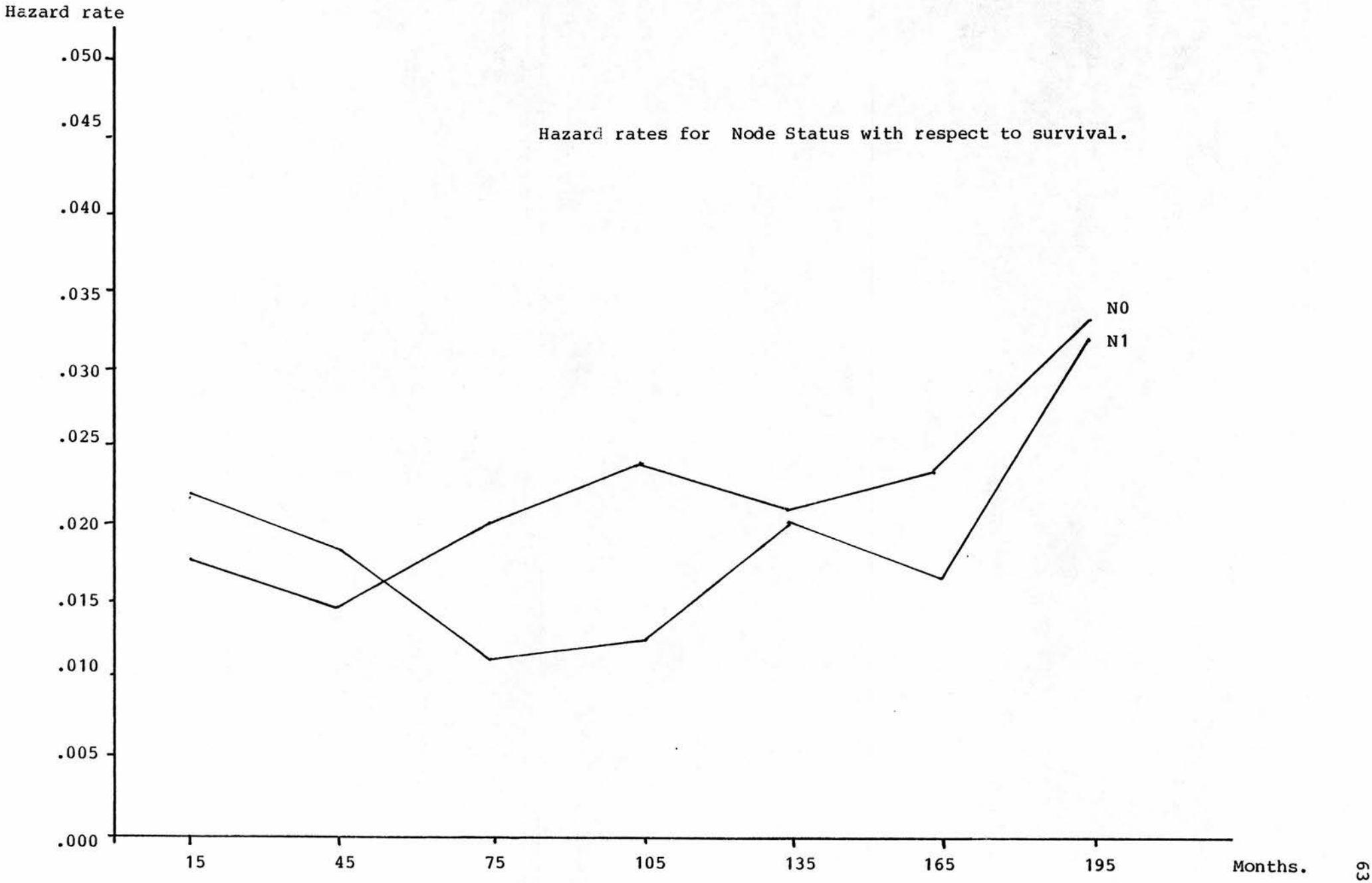

Hazard rate
.050
.045
.040
.035
.030
.025
.020
.015
.010
.005
.000
Hazard rates for Node Status with respect to survival.
N0
N1
15
45
75
105
135
165
195
Months.

Figure (2.5.10)

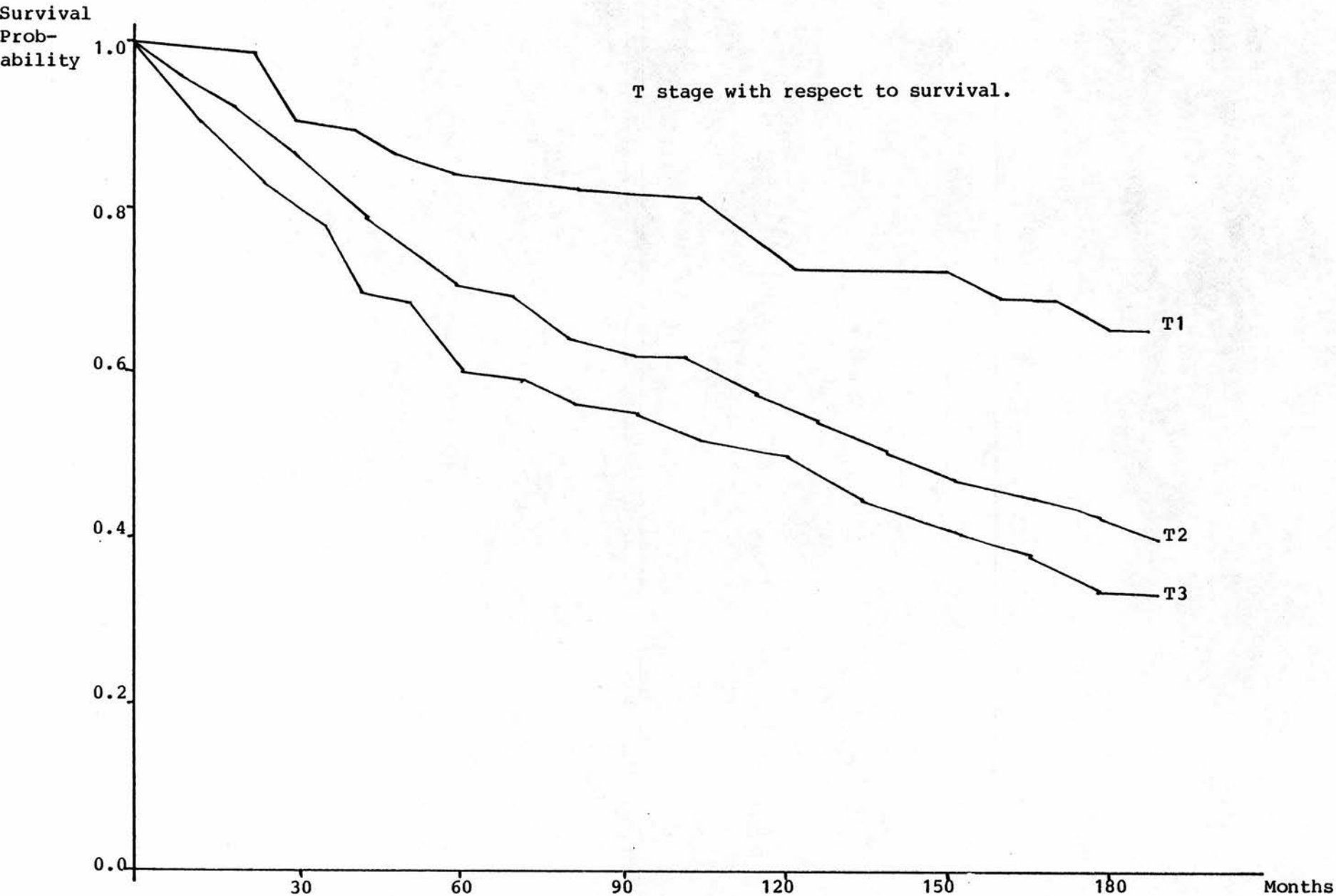

Survival Prob-ability
1.0
0.8
0.6
0.4
0.2
0.0
T stage with respect to survival.
T1
T2
T3
30
60
90
120
150
180
Months

Figure (2.5.11)

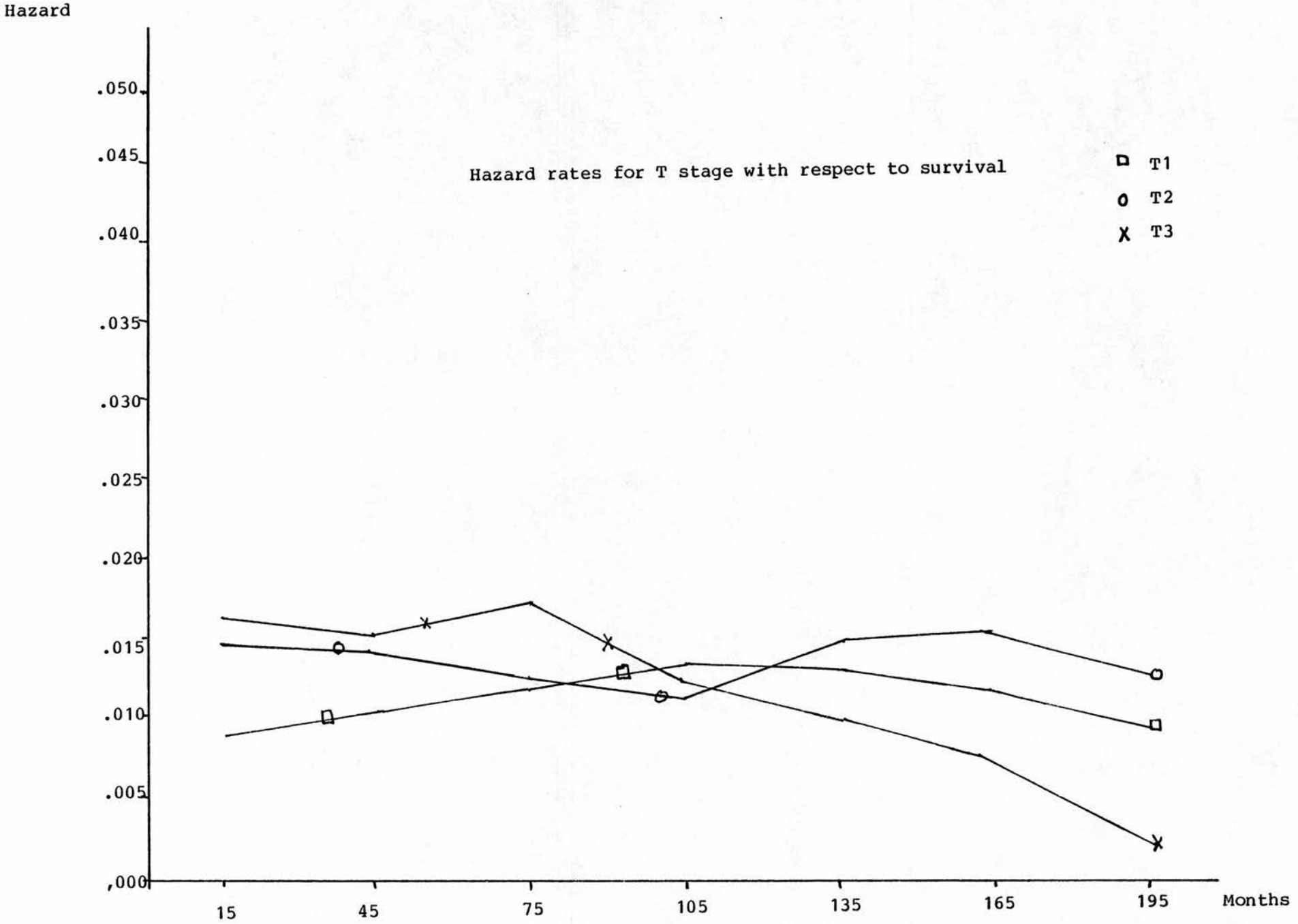

Hazard
.050
.045
.040
.035
.030
.025
.020
.015
.010
.005
,000
Hazard rates for T stage with respect to survival
T1
T2
T3
15
45
75
105
135
165
195
Months

Figure (2.5.12)

age interpretations. Both menopausal and premenopausal groups produce higher initial hazard rates than the postmenopausal groups but the rates later converge. In Chapter 6 we will perform tests on such time dependencies.

For the size of the tumour, there is a slight indication that hazard rates are of similar pattern for all groups during most of the time scale. However, the larger tumours after an initial period of constant risk produce lower levels in later stages of the disease. The main purpose for putting this emphasis on time dependency of size and age is to relate the findings to logrank and Wilcoxon tests. The above points are similarly noted for the differences between the two tests. Wilcoxon test for the categories of age survival gives a chi-squared value of 23.7 against logrank value of 18.05. The reverse is true for the size of the tumour. That is the Wilcoxon test gives a chi-squared value of 10.62 and the logrank test gives a higher value of 13.82, indicating that the differences may be due to later events. The two tests do not differ to an important degree and other main effect prognostic categories show even lower differences. In fact the difference between the two tests of size may be coincidental. Inferences about variables of prognostic importance from this study will be considered again in Chapters 6 and 7. In these chapters a more detailed model will be used and indirectly we will explain some of the differences between the logrank and the Wilcoxon tests.

C H A P T E R 3

PARAMETRIC METHODS AND HAZARD FUNCTIONS

In the previous chapter we discussed a set of non-parametric statistical methods for the analysis of survival data. In this chapter we will be dealing with the parametric methods. Within the descriptions of this chapter we will discuss a few possible hazard functions from empirical data.

3.1 Commonly used parametric methods in survival analysis.

These methods follow the general philosophy of parametric statistics by which we assume that time to a critical event is a random variable and based on this postulate we may assign a frequency distribution to the survival times. Basically the distribution functions must be able to approximate to the empirical life-tables which present the cumulative proportion of cases surviving against the time scale of events. It is often difficult to visualise differences between classes of survival functions or identify them purely based on an inspection of the distribution function, in that the survival distribution is always a decreasing function. However for purposes of defining and classifying between distribution functions, their transformation to hazard functions plays an important role, so that by a visual display of such functions a pattern of events can be observed. The hazard rates in fact present the rate of change of the survival curves and thus the/

the pattern of the hazard can be useful for the purpose of identification between the empirical and frequency distributions. In the early stages of the chapter we are not interested in the effects of treatments or prognostic variates, but rather, in the possible families of distribution functions that may be useful in the applications of fitting parametric distribution functions to life-tables.

First we describe 3 rather general methods and the plots of their hazard functions.

Name of Distribution	Hazard Rate	Death Density Function	Survivorship Function
Exponential	$\lambda(t) = \lambda$	$f(t) = \lambda\exp(-\lambda t), \lambda>0;\ t>0$	$S(t)=\exp(-\lambda t)$
Weiball	$\lambda(t) = \mu\upsilon t^{\upsilon-1}$	$f(t) = \mu\upsilon t^{\upsilon-1}\exp(-\mu t^{\upsilon}) \mu>0\ \upsilon>0;\ t>0$	$S(t)=\exp(-\mu t^{\upsilon})$
Rayleigh	$\lambda(t)=\lambda_0+2\lambda_1 t$	$f(t)=(\lambda_0+\lambda_1 t)\exp(-\lambda_0 t-\lambda_1 t^2)$ $\lambda_0>0,\lambda_1>0,t>0$	$S(t)=\exp(-\lambda_0 t-\lambda_1 t^2)$

The first two of the above distribution are in fact members of the same set and the final distribution will be referred to as a special case of Taulbee's approach later in this chapter. Figures (3.1.1) to (3.1.9) illustrate the various functions for the three distributions at variable parameter values.

In the previous chapter we presented some of the empirical hazard rates for the old Edinburgh trial data, in specific subgroups of patients. It is important to note that at this stage we mention hazards in general terms for the total population. In practice hazard rates can show different quantitative failure rates for different subgroups. Under a parametric context for a comparative study/

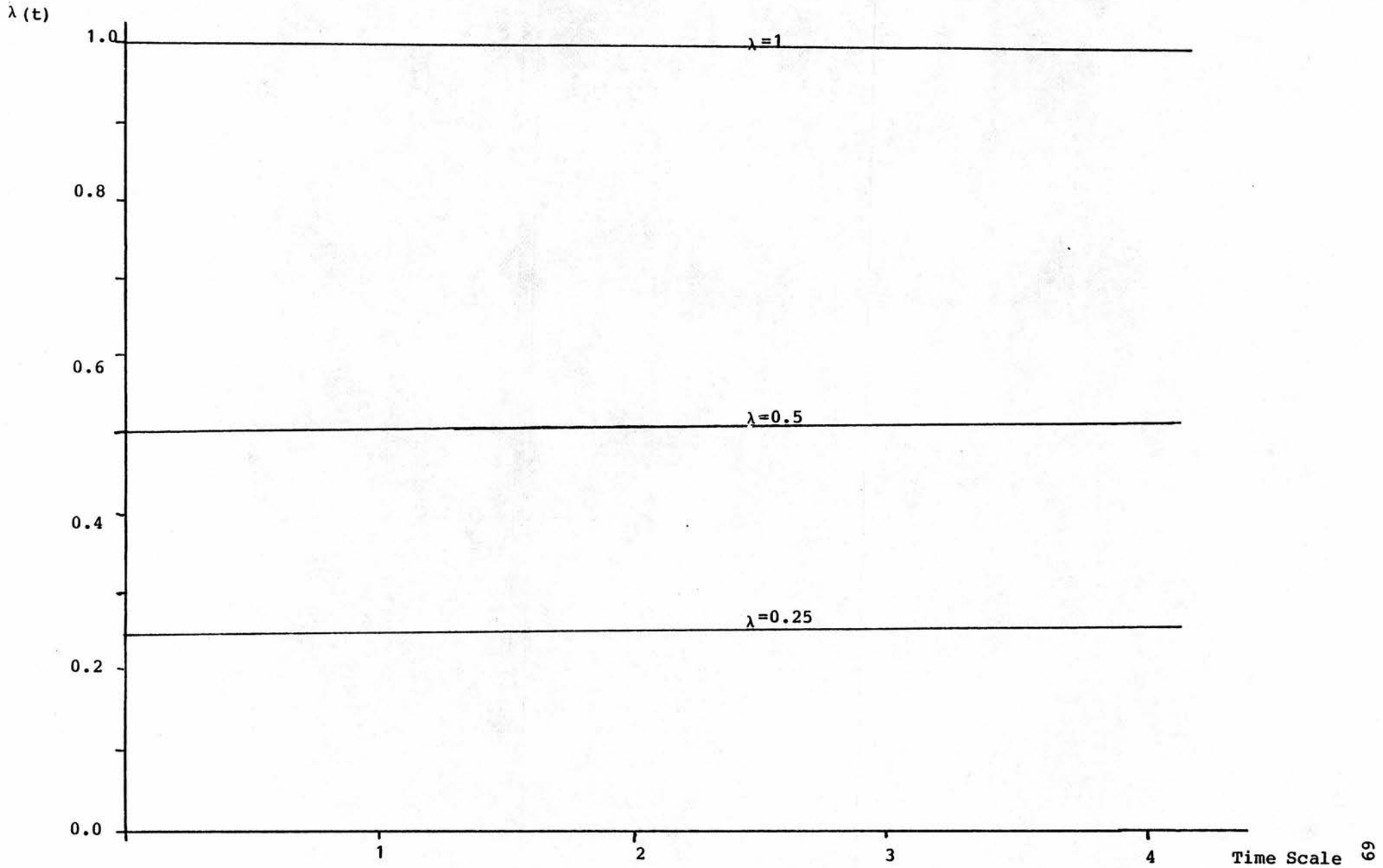


Figure (3.1.1) Hazard rate for the exponential distribution, $\lambda(t)$

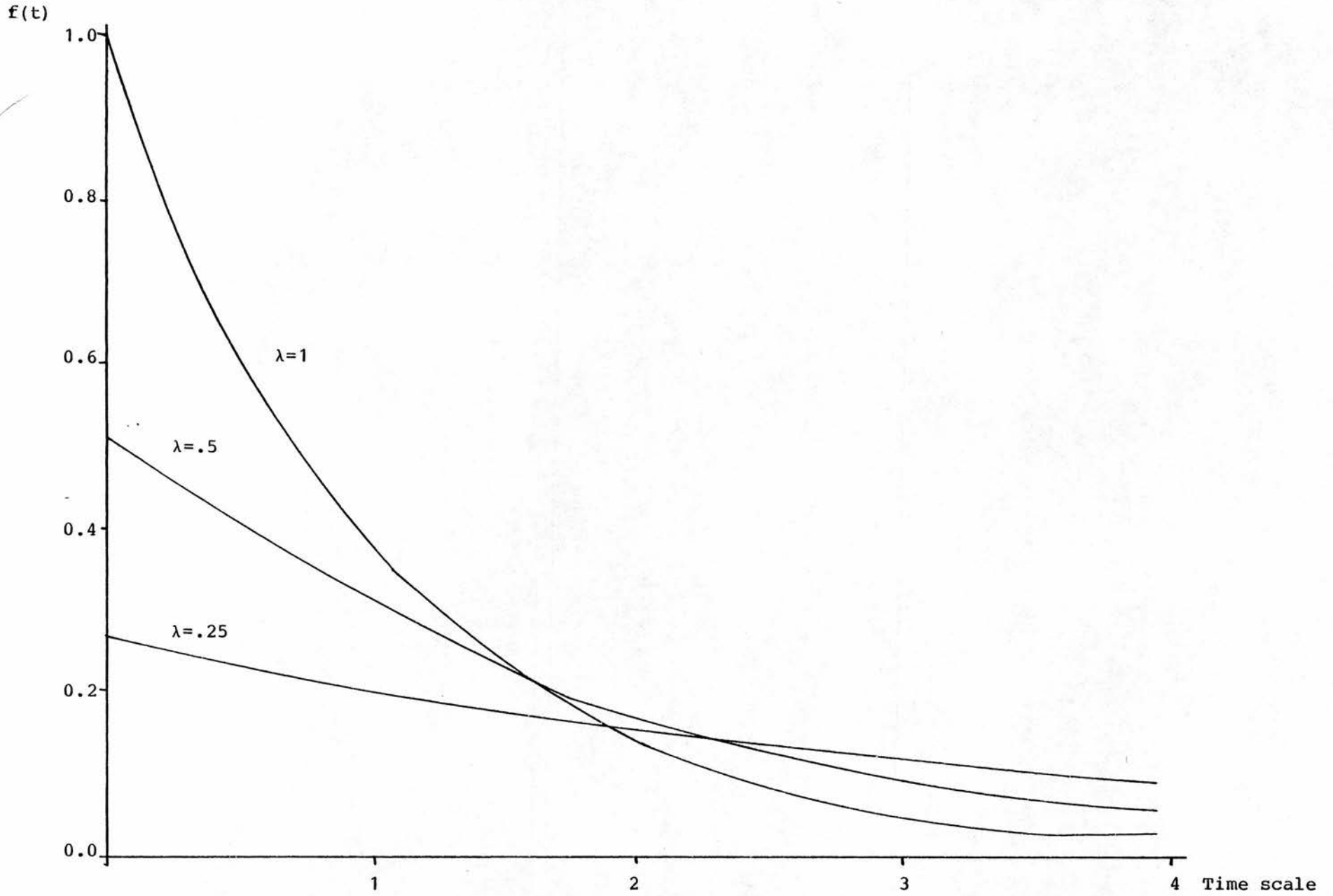


Figure (3.1.2) Death density function for the exponential distribution, f(t)

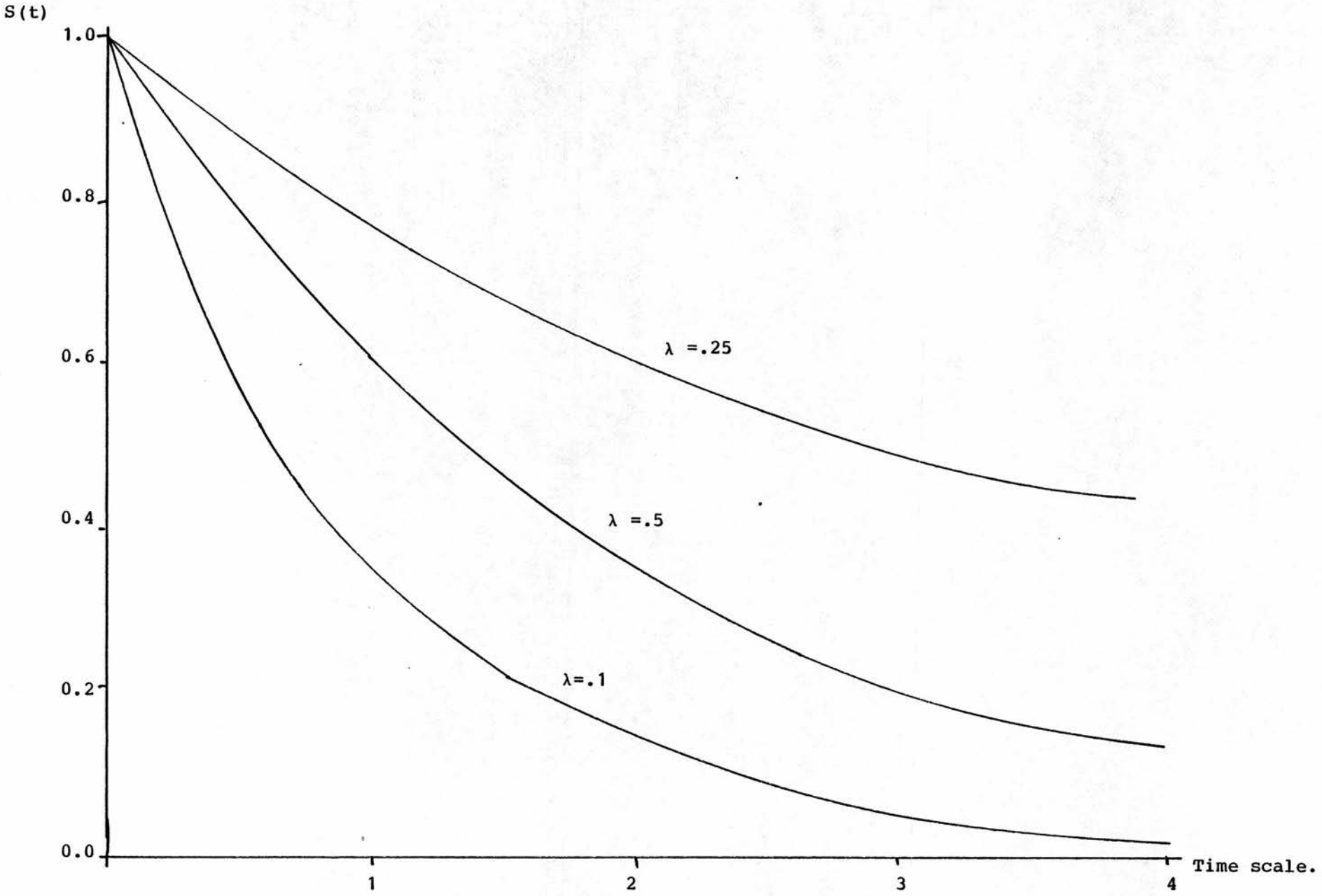


Figure (3.1.3) Survivorship function for the exponential distribution, S(t)

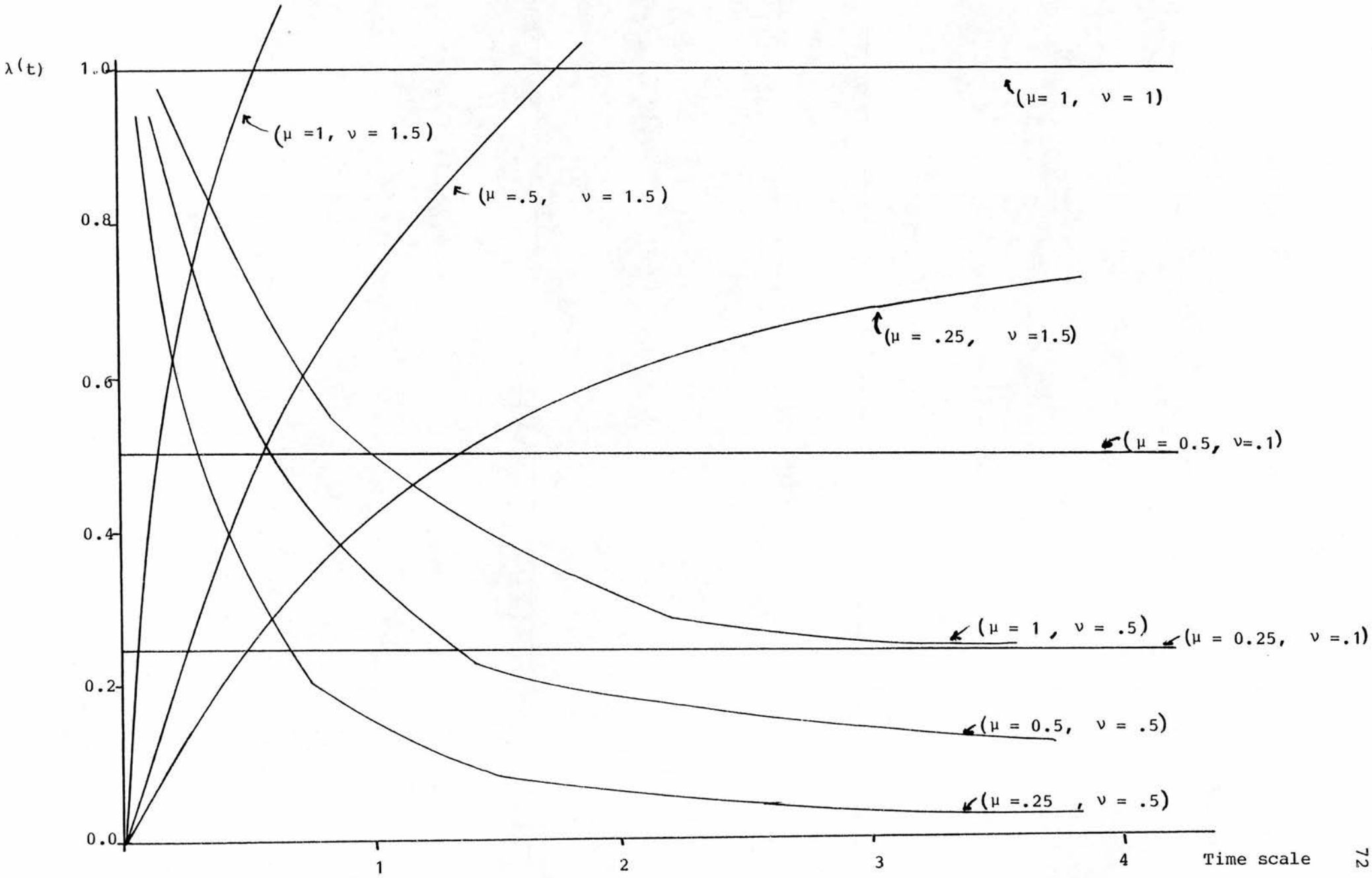


Figure (3.1.4) Hazard rate for the Weiball distribution λ(t)

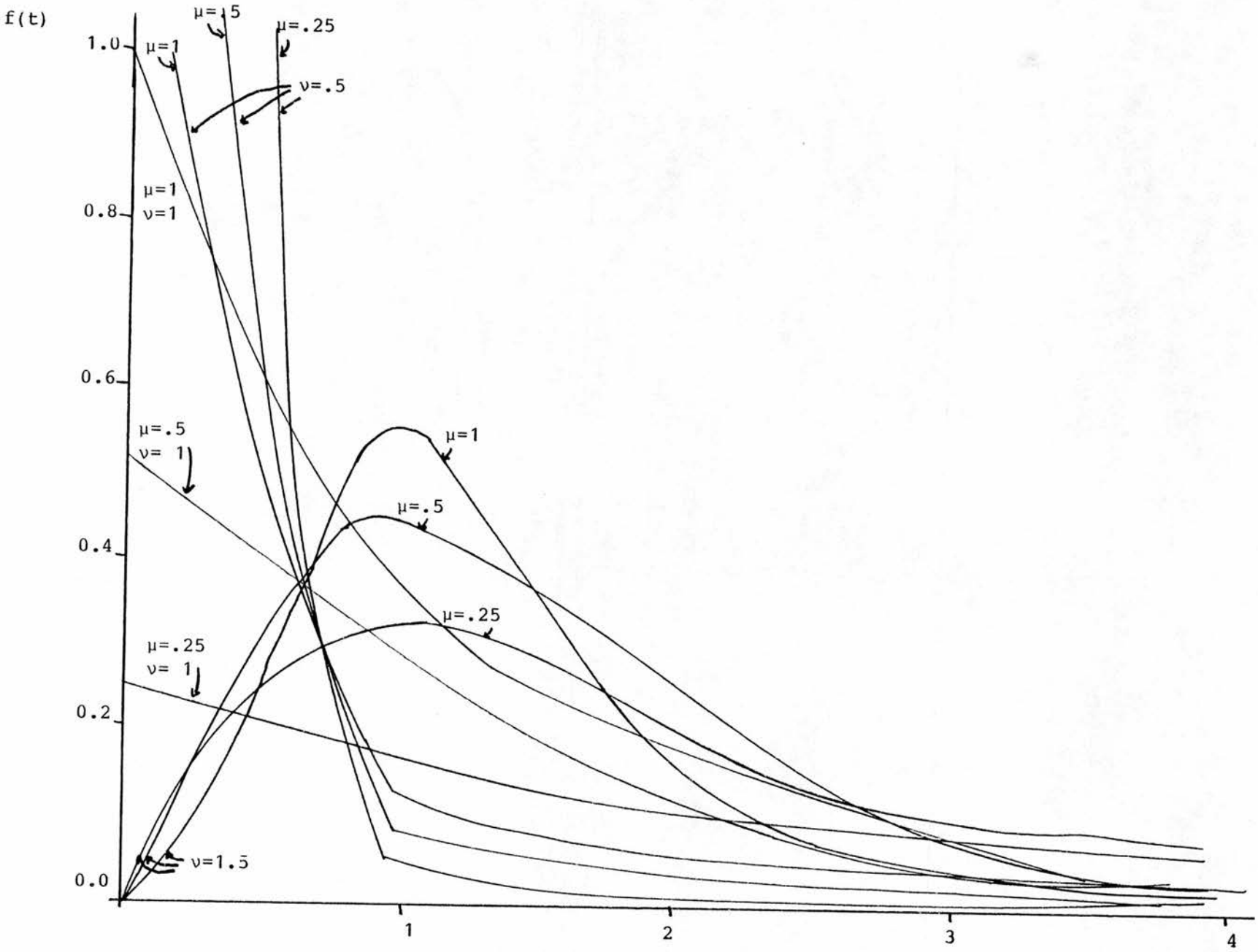


Figure (3.1.5) Death density function for the Weiball distribution, f(t)

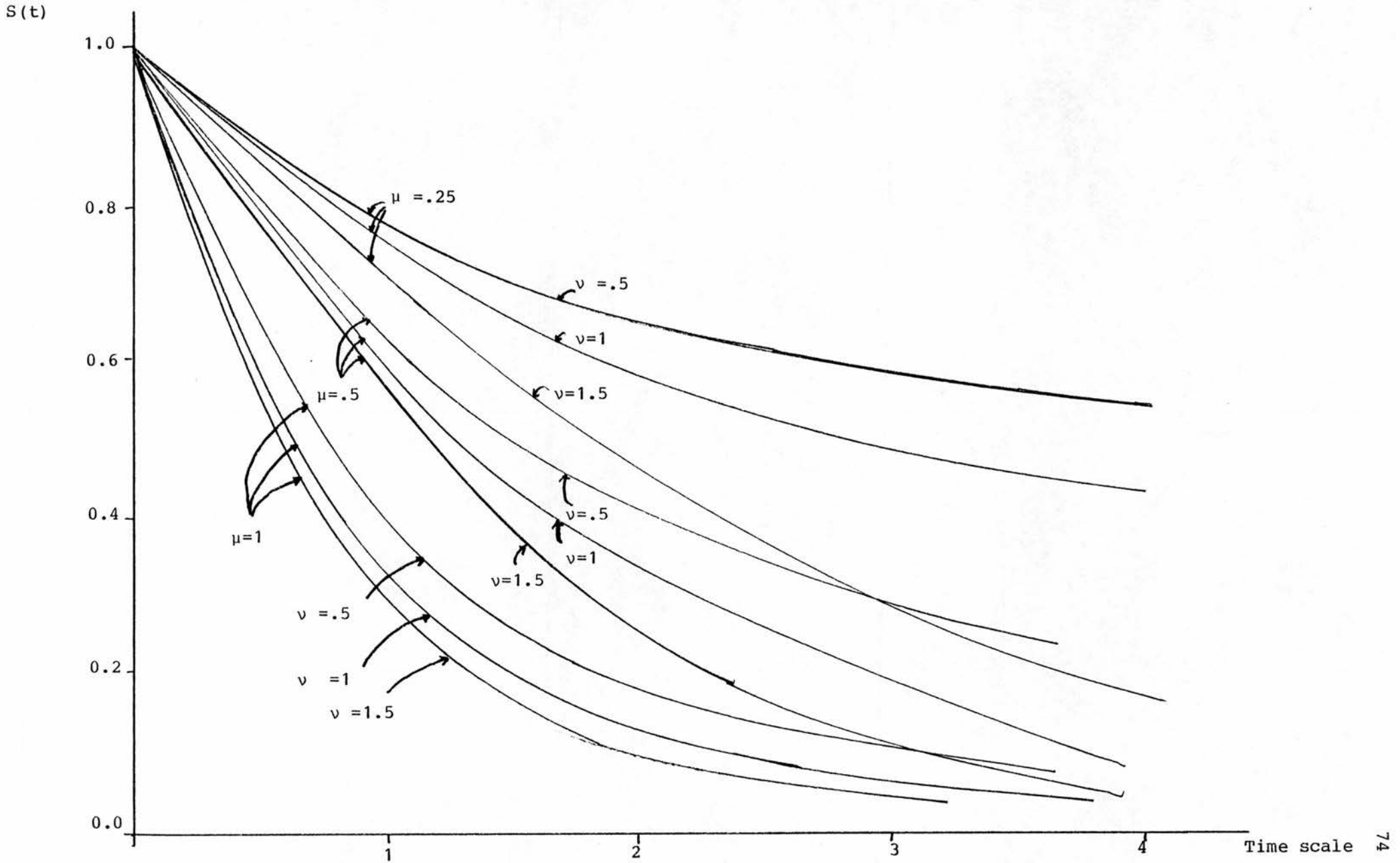


Figure (3.1.6) Survivorship function for the Weiball distribution S(t)

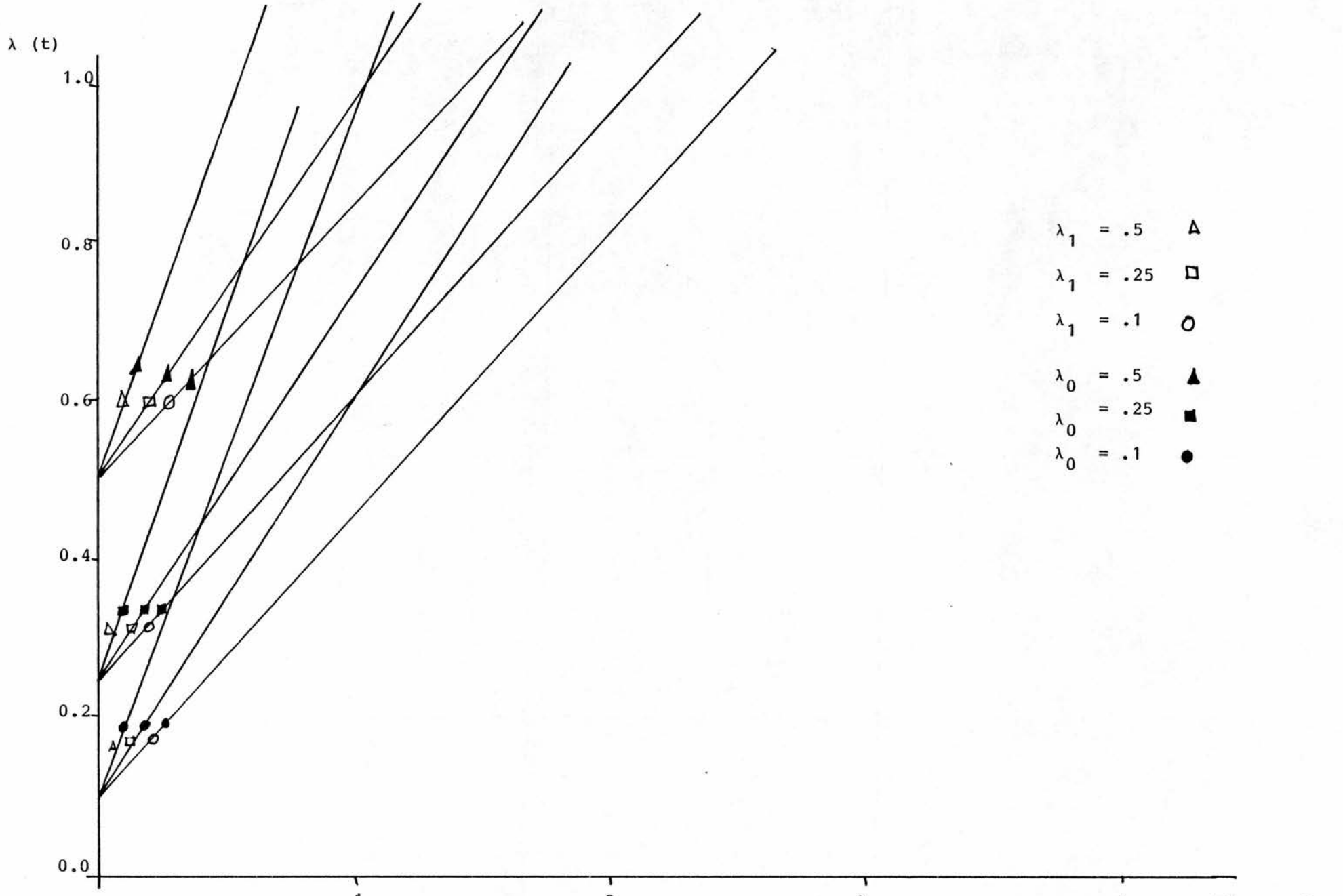


Figure (3.1.7) Hazard rate for the Rayleigh distribution

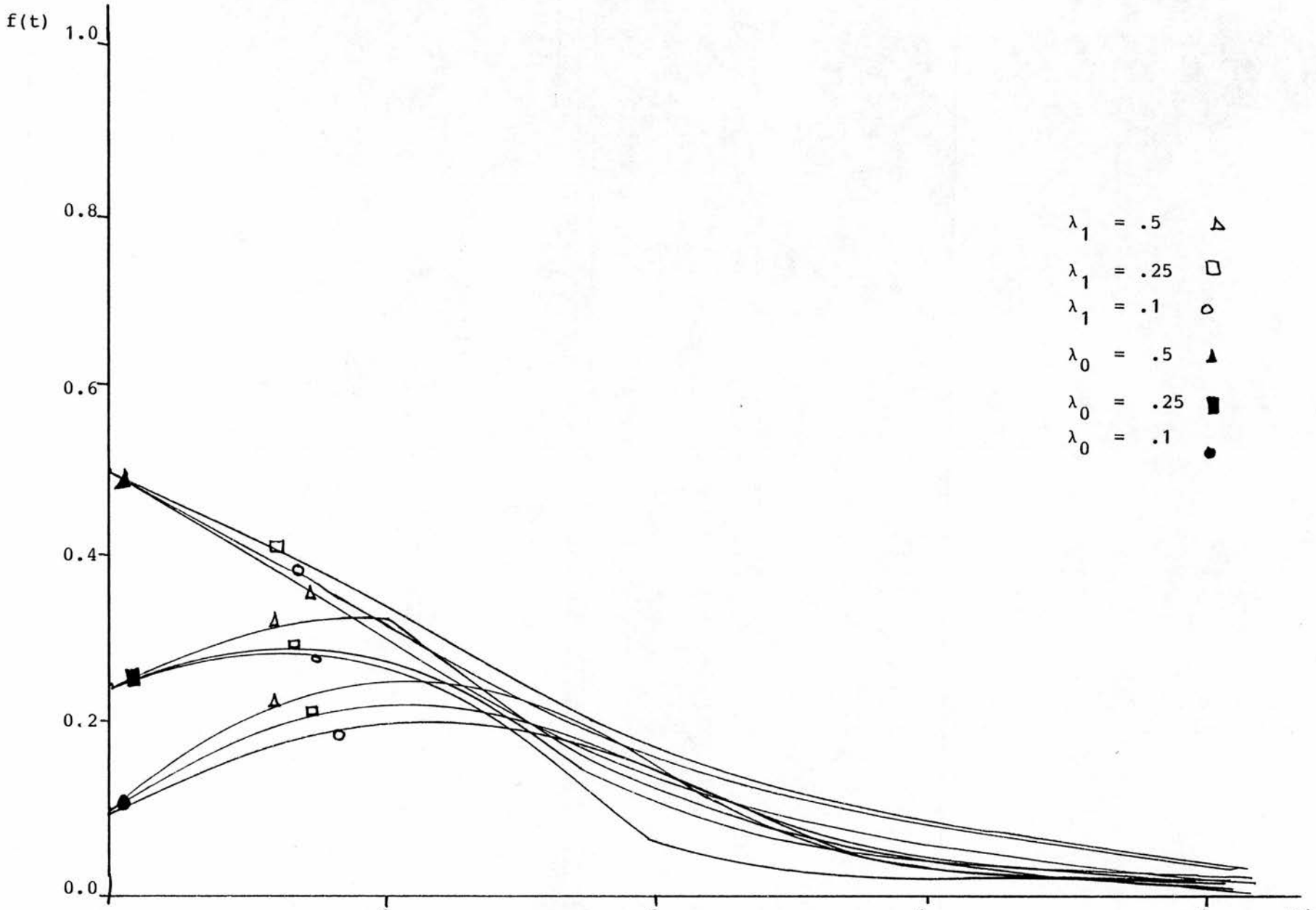


Figure (3.1.8) Death density function for the Rayleigh distribution, f(t)

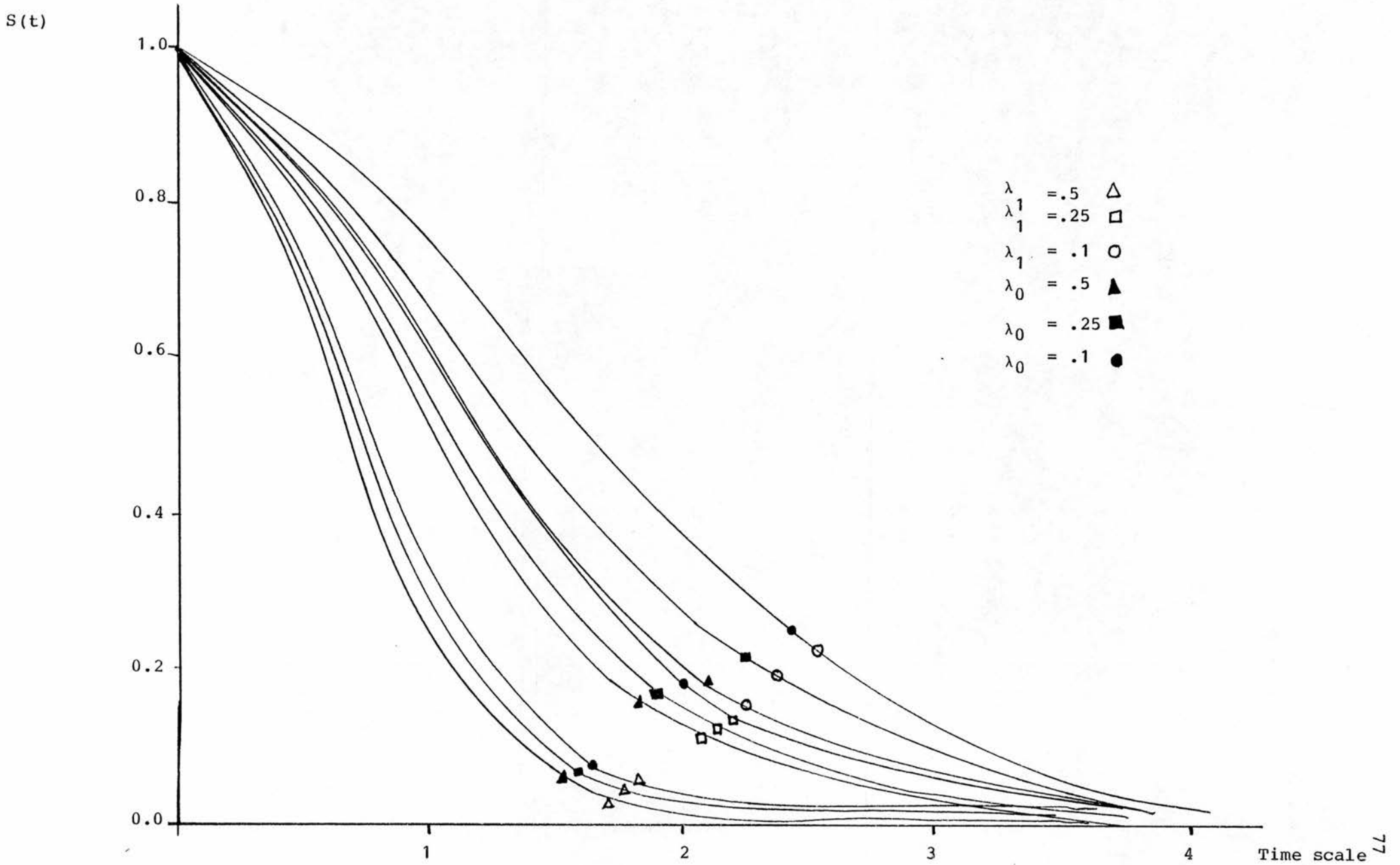


Figure (3.1.9) Survivorship function for the Rayleish distribution, S(t)

study of the survival rates we are interested in the magnitude of a parameter that can best describe the differences between the various subgroups of patients. Later in this chapter we will derive the necessary estimates of the above distributions with covariate effects present. In applying such methods we make certain assumptions on the actual form of the hazard rates in choosing a particular model for describing the relevant differences. By a visual inspection of the hazard rates one can then judge how well the data conforms to the assumptions of the statistical method.

3.2 Examples of hazard functions and families of distribution for survival analysis.

The most common parametric distributions used in clinical trials for the survival of patients are the exponential and the Weiball distribution. The Weiball offers a wide range of increasing and decreasing hazard rates, with the exponential function being a special case for a constant hazard model. As will be shown, these two distributions belong to a family of proportional hazard models with covariates. The assumption of proportional hazards requires that the hazard rates for all subgroups must be a multiple of a base time hazard rate for the entire set of subgroups. S. Gore (1981) Figure (3.2.1) shows the example of breast cancer trial data in which the assumptions of proportional hazards are violated. Without a visual inspection of the hazard functions, there is a strong temptation to use one of the robust proportional hazard models, such as that described in Cox's (1972) paper. An inspection of the hazard function can also lead to choosing a more efficient analysis based/

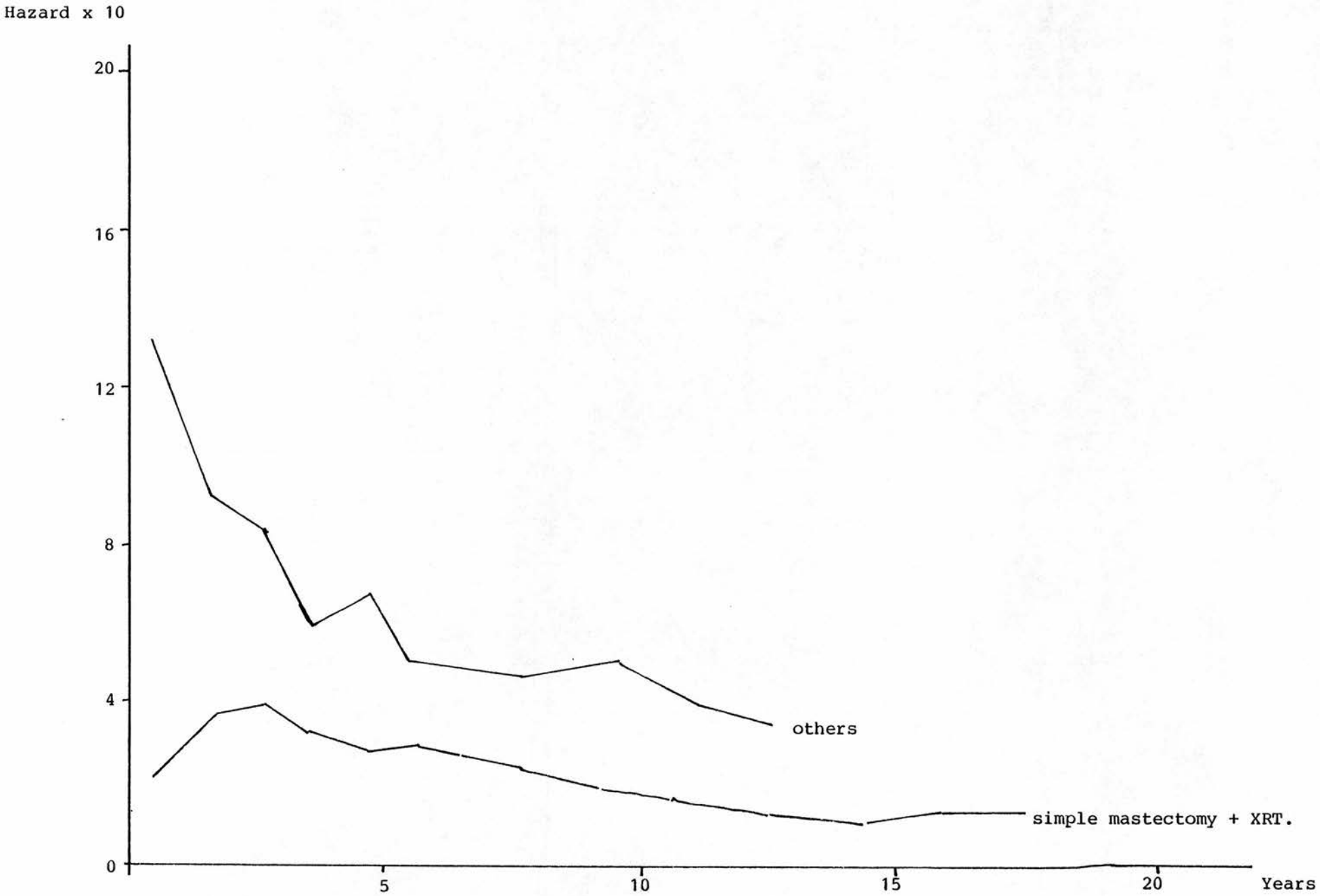


Figure (3.2.1) Estimated hazard function for treatment.

based on parametric methods, while an approach purely based on tests of significance using completely non-parametric methods may over-generalise the pattern of failure.

In the plots of exponential, Weiball and Ragleigh distribution some forms of constant increasing and decreasing failure rates were presented. Later we will discuss some u-shaped and cone-shaped hazard rates that can arise from a trial data.

Turner et al (1976) consider a general 3 parameter family of survival distributions. This family is able to generate an extensive number of distributions that can be used in a survival analysis. The general survival function is given by

$$S(t) = \{1 + \frac{n}{|n|}\,[\tfrac{1}{2}\,(1 - \frac{n}{|n|} + \beta\,|n|pt)^{-1/p}\,]\}^{-1/n} \qquad (3.2.1)$$

For $t \geq 0$, $\beta > 0$, $p > 0$

and $-\infty < n < \infty$

The probability density function is

$$f(t) = \beta\{\frac{n}{|n|}\,[\,S(t)^{(1-np)/(1+p)} - S(t)^{(1+n)/(1+p)}\,]\,\}^{1+p} \qquad (3.2.2)$$

These functions provide a set of highly flexible distributions with many differing shapes for the hazard functions, such as increasing, decreasing, constant and cone shaped hazards. This variability of the hazard rates can be mimicked by the range of the distributions of form, Gamma, Weiball, Lognormal, Ragleigh, Single hit and Arheus distributions. The most important advantage in the use of Turner's family of distributions is that all these distributions can be defined/

defined by only 3 parameters. This offers a formal test for comparison of shapes of hazards. In a related paper by Bertanou et al (1978) the Turner's family of distributions with concomitant variables is used. They use a maximum likelihood estimator for the parameters using a method of Hazelrig et al (1978), The most important problem in the general use of this approach in survival analysis so far has been that of adopting an estimation procedure capable of dealing with the complications of the censored survival data.

Bertanou et al (1978) compare life expectancy in two groups of children treated without surgery in Tetralogy of Fallot. The data is not based on a randomised trial, but is formed of clinical information and autopsy data. In this approach the analysis begins with the study of the possibility of detecting changing risk patterns among subgroups. Further, a comparison of estimates within each group can easily be made without making the restrictive assumptions that the two subgroups have similarly shaped survival distributions. It may well be expected that this approach by being a parametric one provides a better approximation to the hazard functions. The result is an effective procedure for estimating parameters of the distribution. In their conclusion Bertanou et al (1978) support the generally accepted view that "the natural history of person born with tetralogy of Fallot is determined primarily by the severity of the pulmonary sterosis, as demonstrated by the tendency of the person with pulmonary uteria to die at a young age than those without pulmonary uteria or the group as a whole".

Using/

Using this parametric approach, the conslusions regarding the results, both in terms of survival times of treatment main groups and the subgroups are the same as the alternative non-parametric or single parametric approaches. However, within the present setting they provide the corresponding hazard functions for the different groups, Figure (3.2.2).

It is clear that the highest risk period for pulmonary steriosis is the first two years and unlike pulmonary uterisia, the risk of death does not decline in the later years, probably due to relatively high risks in the second decade.

With the parametric estimation of the hazard rates clear cut functions are produced that are intelligible in reducing the data on the timescale. An empirical plot would yield the same patterns and the same information. However the Turner's generic family of survival curves has the distinct advantage that by inclusion of extra parameters, there is a possibility of testing the hazard functions and obtaining a distribution from its hierarchy that yields the best parsimonious fit. The difficulty with this approach can be the interpretation of the results if the estimating parameters of the hazards differ greatly. Since there are 3 parameters it will be difficult to decide on the meaning of such patterns. An extreme example is a situation where one group has a higher initial hazard rate followed by a constant hazard rate and another group having initially a low hazard rate followed by a high rate. The outcome can be two survival curves that cross each other somewhere in the time/

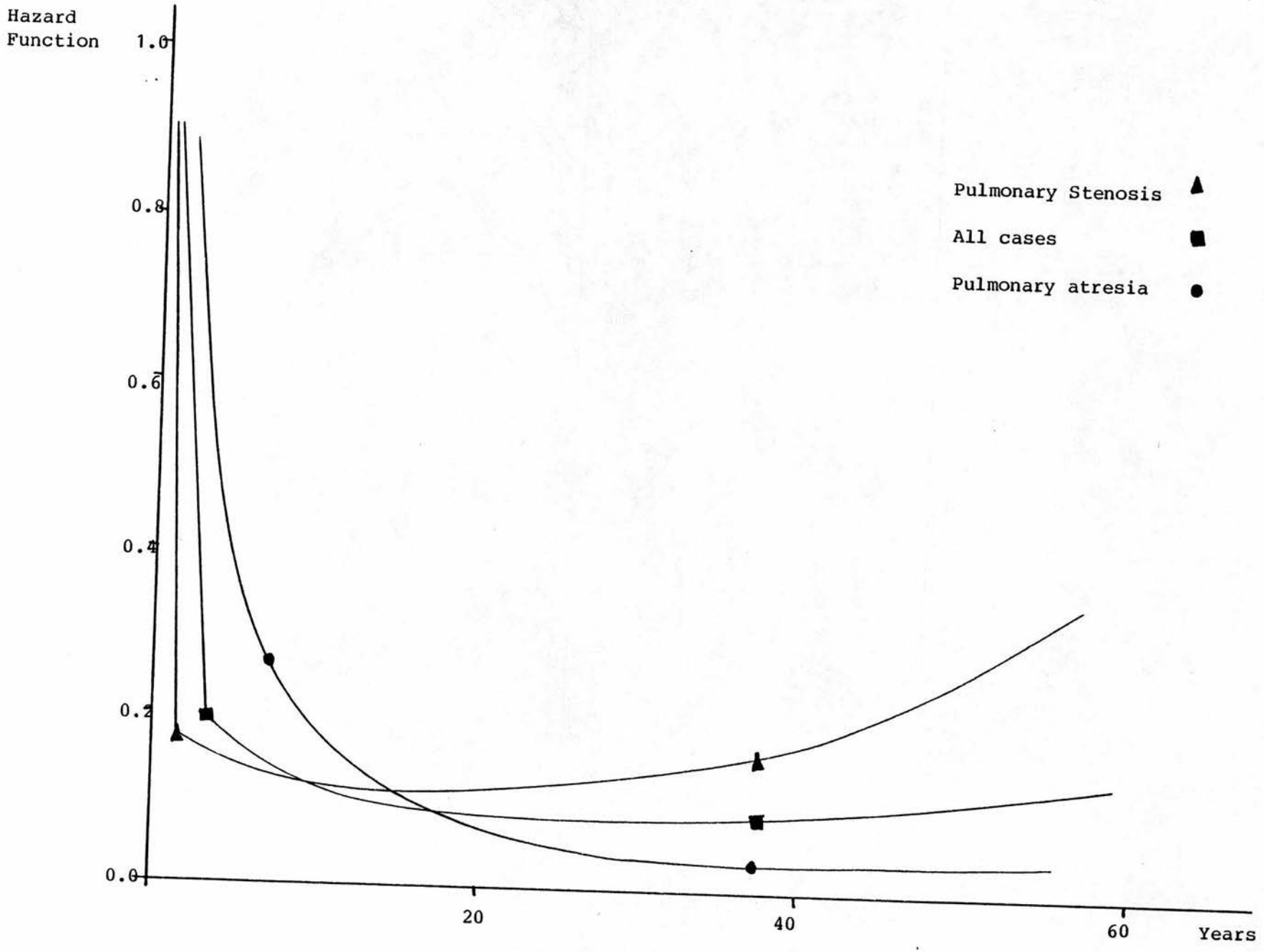

Hazard
Function
1.0
0.8
0.6
0.4
0.2
0.0
20
40
60
Years
Pulmonary Stenosis
All cases
Pulmonary atresia

Figure (3.2.2)

time scale. In the extension of the Weiball model with covariates to a proportional hazards model of the next section we will discuss these points in less extreme situations.

With a non-parametric approach tests can also be constructed to assess hazards. However if a parametric approach is justified, there will be a loss in efficiency in adopting a non-parametric method. It must also be emphasized that in practice the Turner's generic family may be too generous in providing a range of distributions, where the main aim is to assess effects of treatments and covariates.

Barlow et al (1978) adopt a more confined approach in classification of survival distributions. In their terminology, they adopt a failure rate rather than hazard rates. Three classes of distributions are defined by their terminology, (a) increasing failure rates, (b) decreasing failure rates, (c) u-shaped failure rates. The Turner's family also includes a cone shaped hazard which belongs to the Arhenous distribution, as was mentioned earlier in this section.

An example of an increasing failure rate would be a healthy population of over 50 years of age. In such a group one would expect that the effects of old age will become increasingly dominant and hence with increasing age the number of deaths will increase. A possible hazard curve for such a population is the Weiball distribution with shape parameter p = 1.5 Figures (3.2.3) and (3.2.4) represent the /

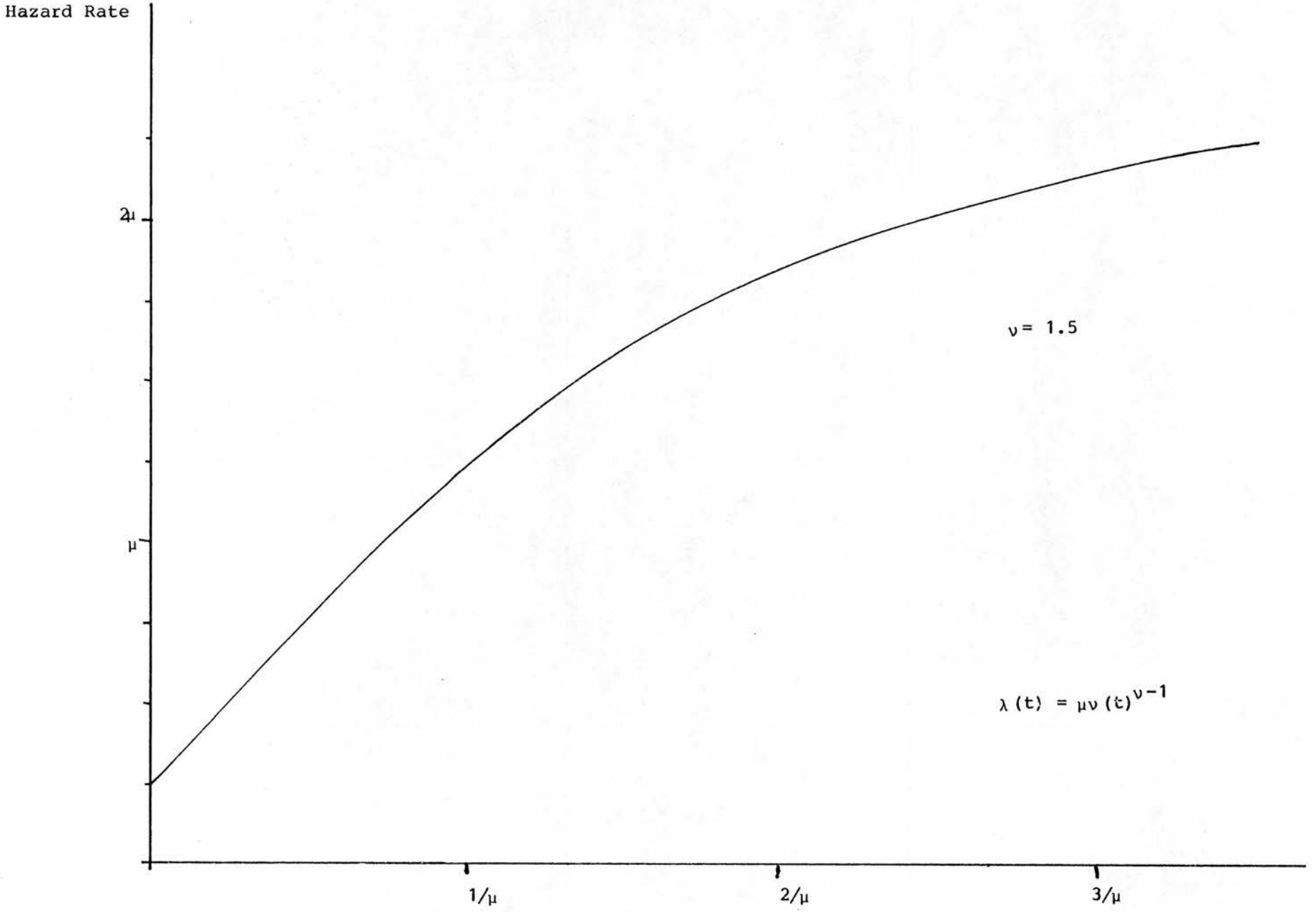


Figure (3.2.3)

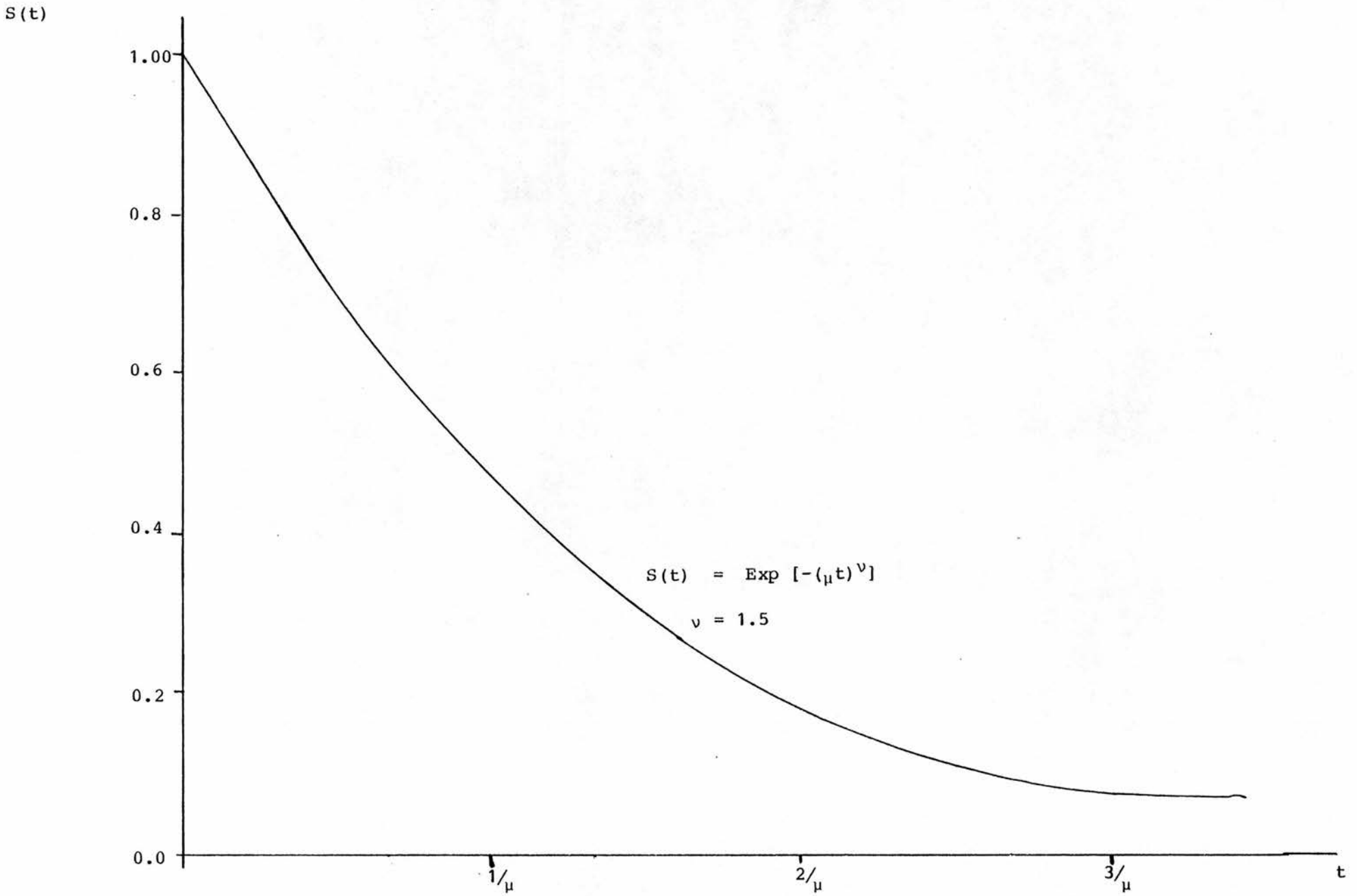


Figure (3.2.4)

the hazard rate with the corresponding survival rates.

The time immediately after a major operation is a critical period for the patients. Often the patient is recovering from anaesthetics, which add extra risk to the survival of the patients. However if there is no progression of disease and the population is young enough not to be affected by old age a possible survival distribution would have a relatively high rate of death in the beginning of the time scale. With the passage of time the normal functions of body can take over and the survival rates could decrease and conform to a healthy population. A Weiball distribution with the parameter $p = 0.5$ can be a possible distribution to approximate such a population. Figures (3.2.5) and (3.2.6) represent the corresponding hazard and survivor function.

Had we not taken the above assumptions, regarding the age of patients, then the effects of old age become increasingly dominant in our population. Further assuming that after the operation there is still some possibility of the progress of disease, as the case may be in a population of post-menopausal Stage I and II brease cancer patients treated by mastectomy and operative radiotherapy, then the hazard rate will be composed of a declining hazard rate followed by an increasing hazard. In fact the life table of all ages of population of a country shows such a hazard rate. At birth the newly born experiences the highest risk of illness and death; with development and growth of the child the risks decrease until later in life new risks of death develop due to old age. Figure (3.2.3)

The/

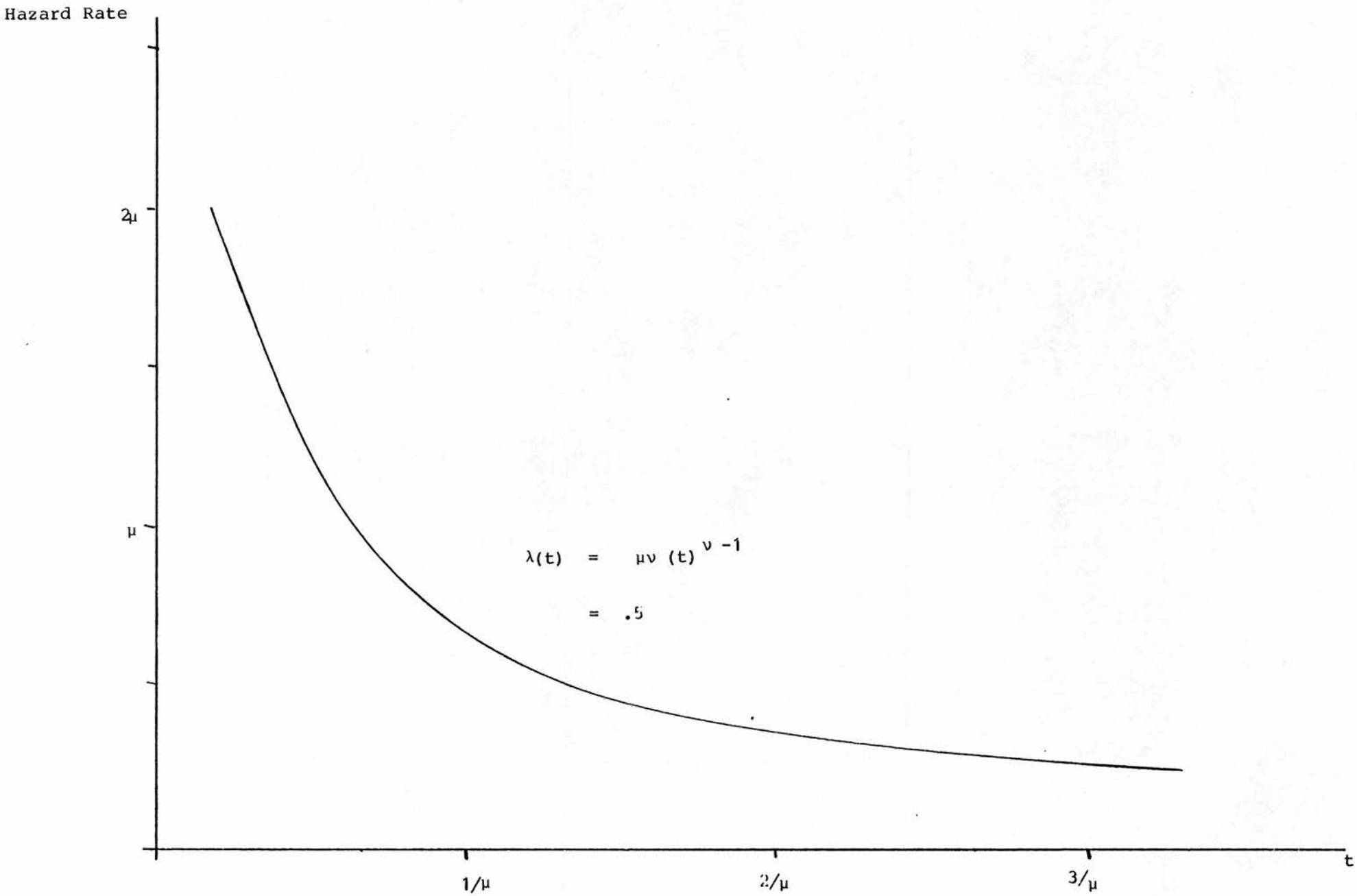


Figure (3.2.5)

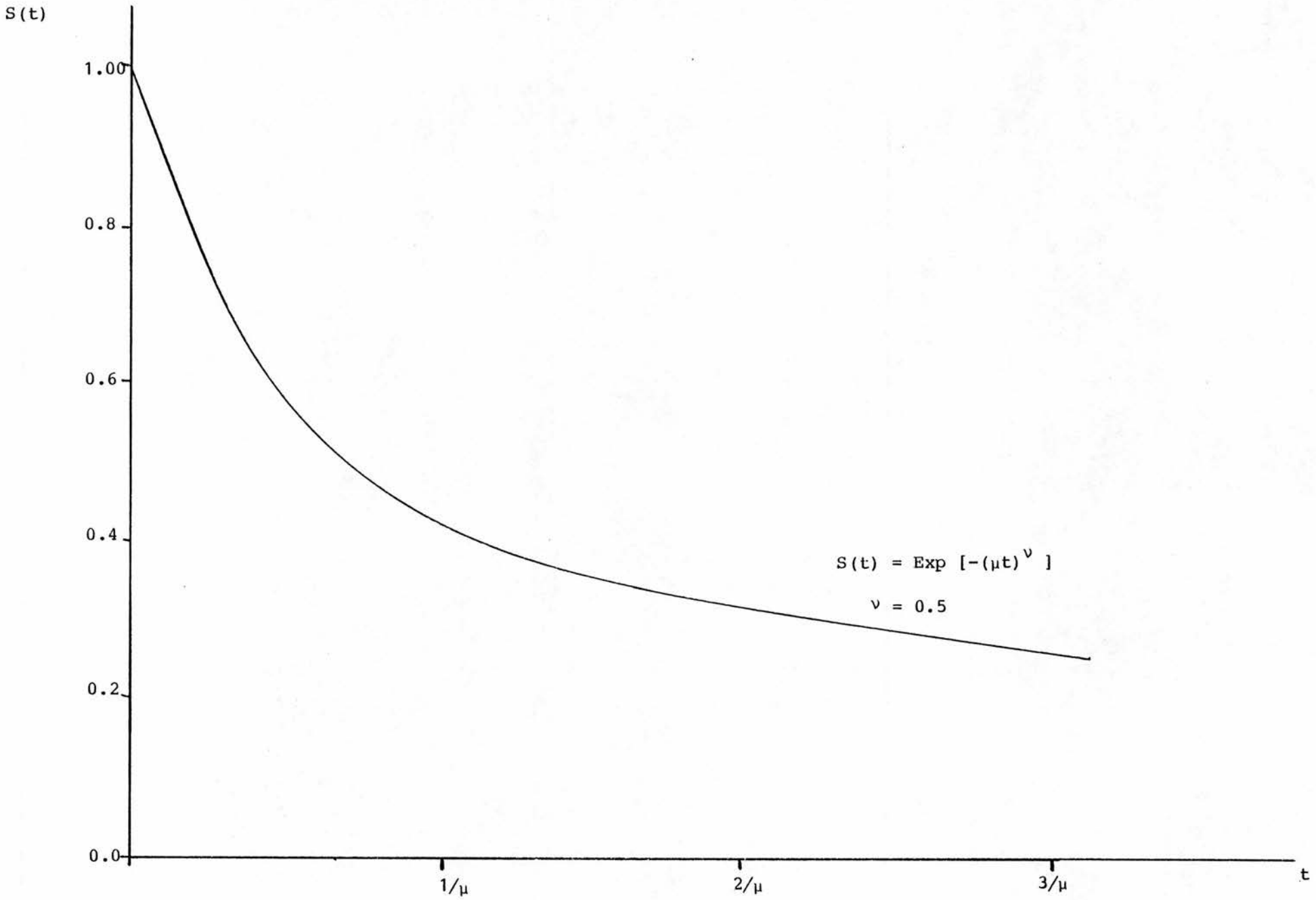


Figure (3.2.6)

The u-shaped pattern of failure, Figure (3.2.7) can be interpretated as suggesting two forms of failure. One due to risks of early life and birth and the other due to old age. In a clinical trial situation if such a pattern is apparent, the constant hazard period in the middle tends to be much shorter. This gives rise to one of the important methodological problems in clinical trials; that is defining the relevant causes of death.

In the previous example on breast cancer treatment, three processes were taking place, each of which can contribute to death. The first factor is the side effects of the treatment, that is mastectomy and radiotherapy in the initial period. Secondly there are risks due to the general progress of the disease either locally or due to metastatic disease and finally there is death due to old age.

The next example of a hazard function we consider is a cone shaped hazard. The u-shaped hazard was a combination of a decreasing hazard rate followed by an increasing hazard rate. The cone shaped hazard is the reverse of this. It begins with an increasing failure rate, reaches a peak and then falls. Time to development of metastatic disease in cancer patients can have a cone shaped hazard.

In the early stages, the disease is confirned to local areas and hence there is a low probability of development of metastatic disease, depending on the form of cancer. With the passage of time the chances of developing metastatic disease increases. For operable breast/

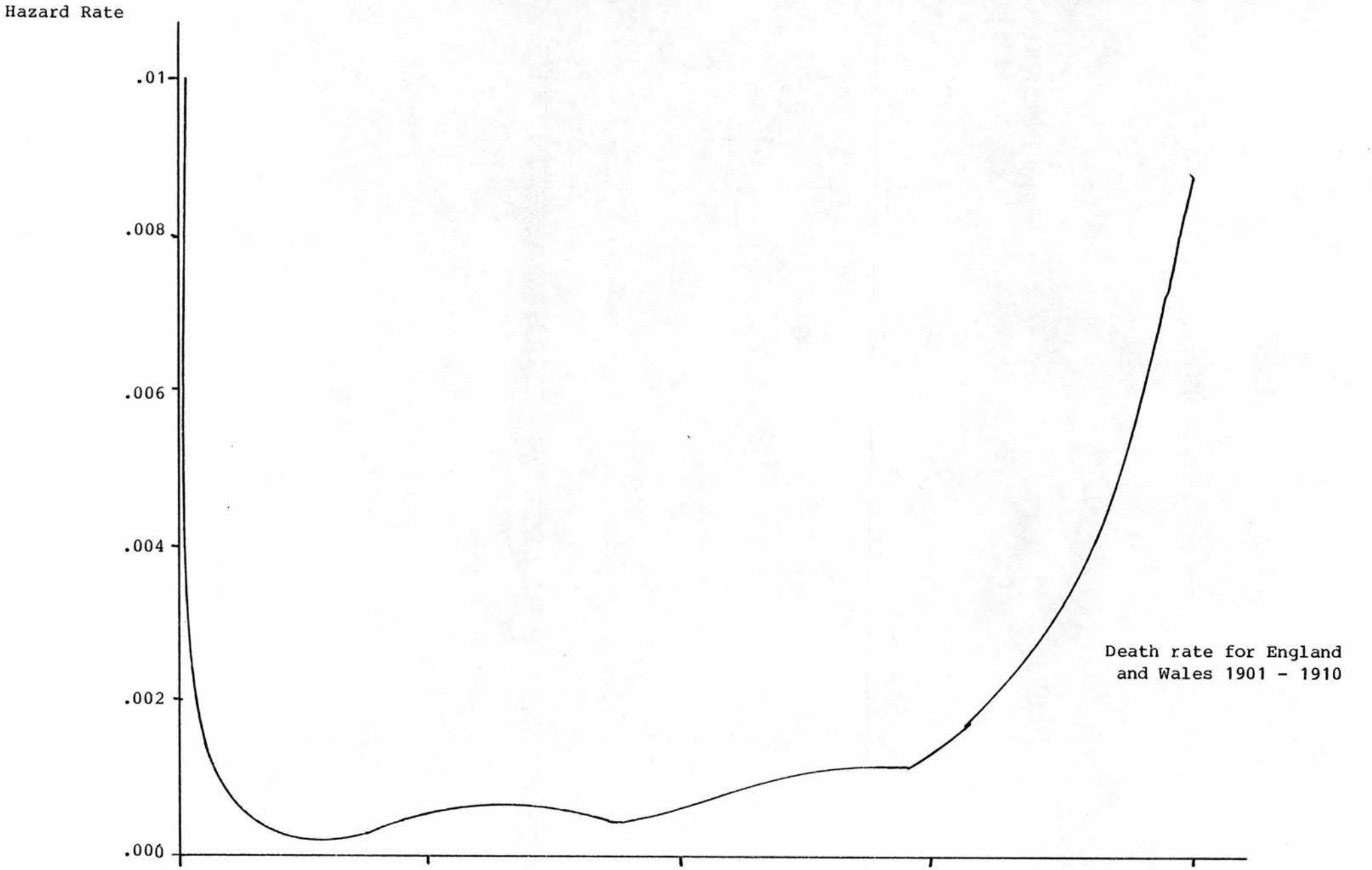

Hazard Rate
.01
.008
.006
.004
.002
.000
20
40
60
80
Years of age.
Death rate for England
and Wales 1901 - 1910

Figure (3.2.7)

breast cancer patients this peak may be reached within 5 years of the detection of the disease. In such a group of patients, there will be some patients with a better prognosis who will not develop metastatic disease. These patients can be increasingly distinguished from the rest who have developed metastatic disease by passage of time. If a patient has not developed metastatic disease in the first five years, the chances of devloping metastatic disease diminishes in the subsequent years. For this reason the peak in the 5th year hazard rate should begin to fall. Figures (3.2.8) and (3.2.9) relate to the hazard rates and survival functions of the above discussion.

Prout, Slack and Bross (1973) discuss a rather interesting population of invasive bladder cancer patients. Criterion for entry into the trial is that patients must have non-invasive bladder cancer, but also must pass a test indicating that there is no metastatic disease present. After 10 years of follow-up the hazard rate is observed. The hazard curves show two separate peaks for the population. Proust et al consider the reason for the appearance of two cones to be due to the population being composed of two very different prognostic groups. This effect is also later indicated by biological evidence. One major entry criterion is a negative result on metastatic disease test. Patients who enter the trial must have shown a negative result with the test. However a group of patients who do not show any evidence of metastatic disease and are test negative are in fact metastatic patients who have not been detected by the test. The first peak of the hazard is due to these patients. The second peak is due to the rest of the population, who are non-metastatic at the time of entry, but develop metastatic disease later/

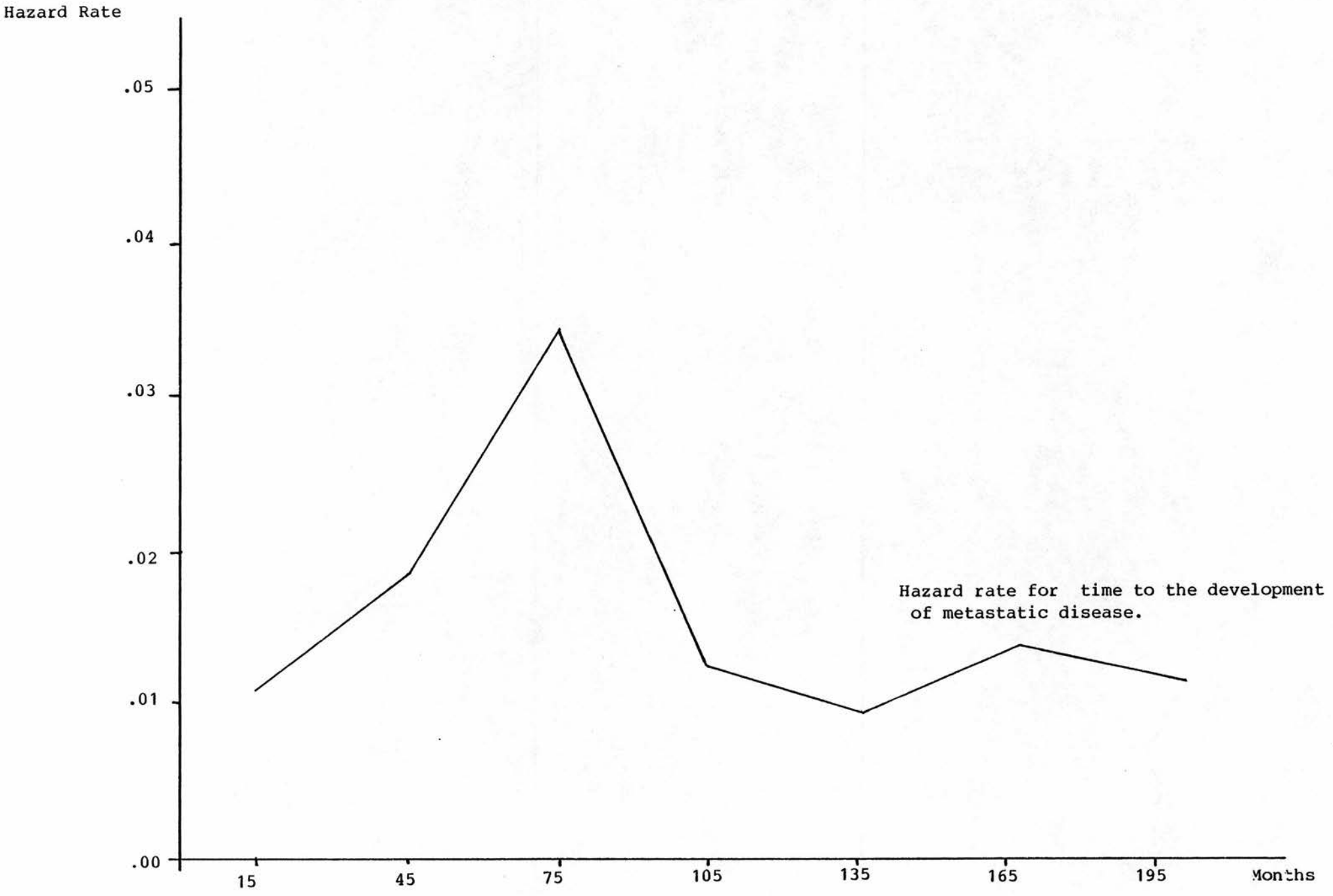

Hazard Rate
.05
.04
.03
.02
.01
.00
15
45
75
105
135
165
195
Months
Hazard rate for time to the development of metastatic disease.

Figure (3.2.8)

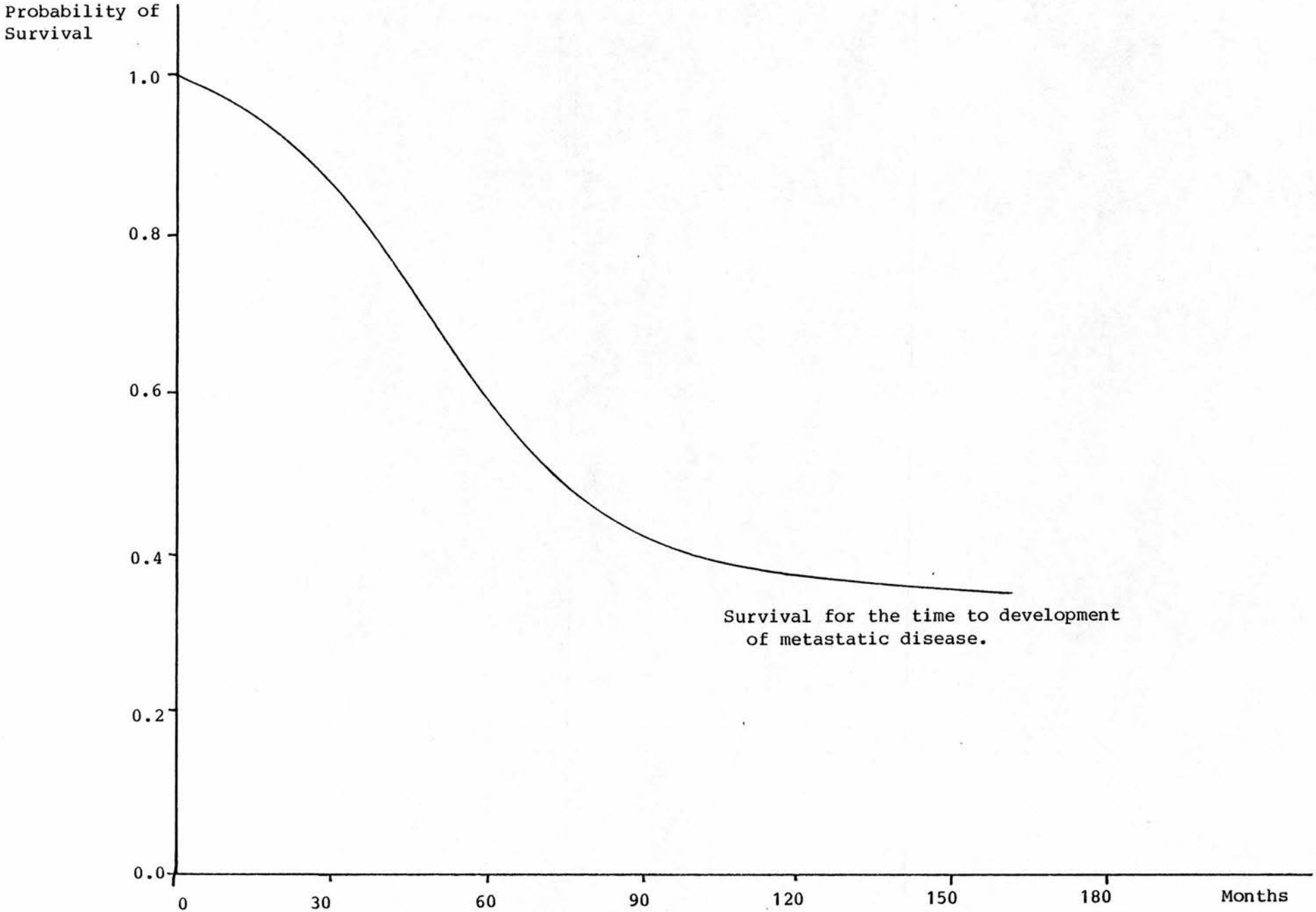

Probability of Survival
1.0
0.8
0.6
0.4
0.2
0.0
0
30
60
90
120
150
180
Months
Survival for the time to development of metastatic disease.

Figure (3.2.9)

later in the course of the progress of disease. Figures (3.2.10) and (3.2.11) present the hazard and survivorship functions for this population.

Finally in a paper, L.E. Rutquist et al (1982) set out to answer the question "is breast cancer a curable disease?" They consider cure to be synonymous with a pattern of survival rates conforming to survival rates of a normal healthy population. For reasons of comparison they note that there exists two different mortality rates, one due to the uncured cases assumed to be constant over time and the other for the cured patients subject to risks of a normal healthy population. Therefore they assume a two parameter model representing sums of two exponential models as appropriate. Further they consider a log normal distribution giving a low initial mortality which rapidly increases to a maximum and with a slow decrease in mortality after the maximum has occured. In their conclusion it is noted that excess mortality from breast cancer is noted at least 18 years after treatment.

One point to note in the above study as well as in some of the previous methods is that for purposes of inference they adopt a X^2 test of goodness of fit for the comparison of the expected and observed values of the survival distributions. Another commonly used method for the estimation of relevant parameters is the maximum likelihood method. We will discuss this approach in more detail within the discussions of the covariates.

In the above discussions much importance was attached to the shape of the hazard rates. Examples of empirical data were discussed/

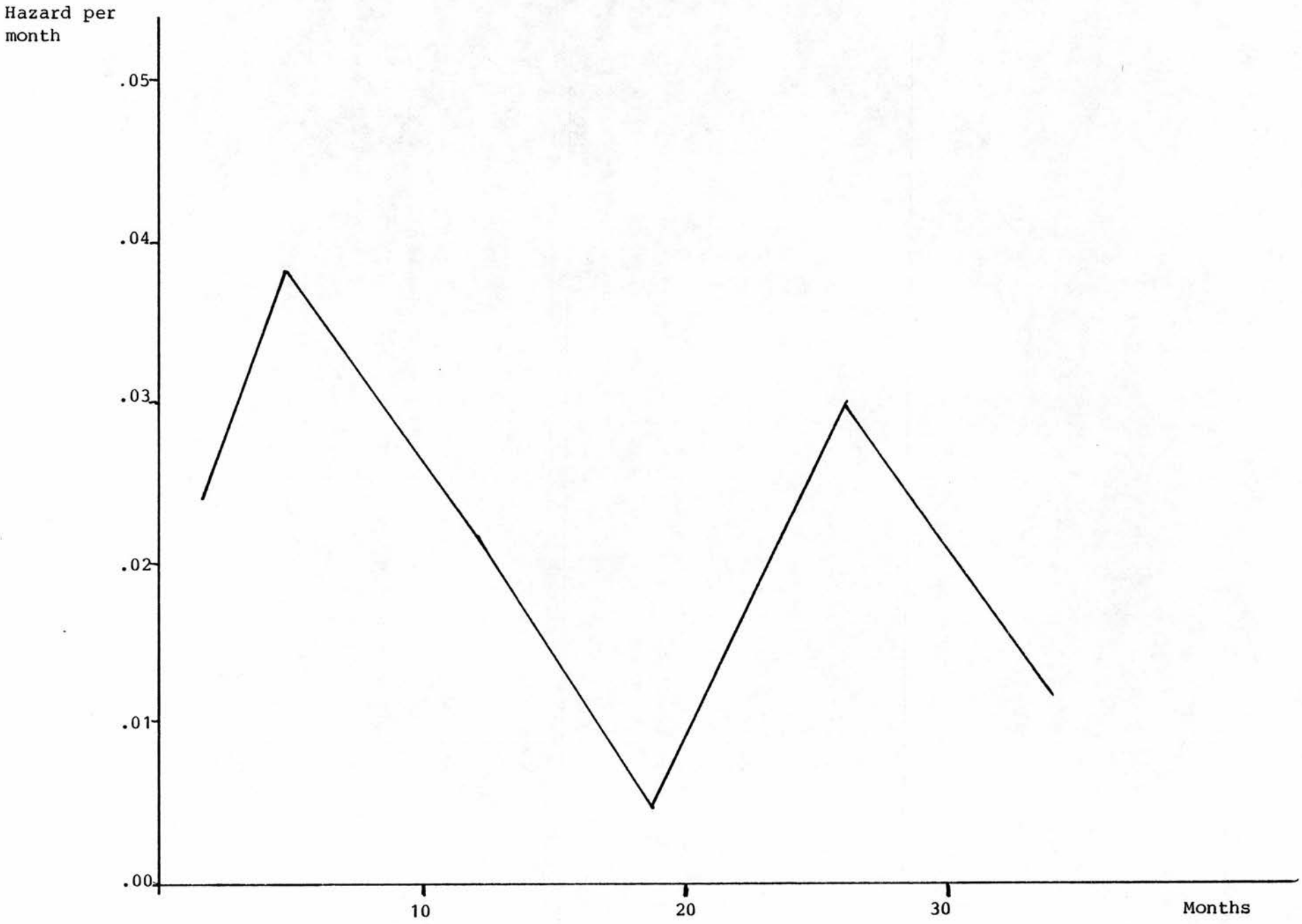


Figure (3.2.10)

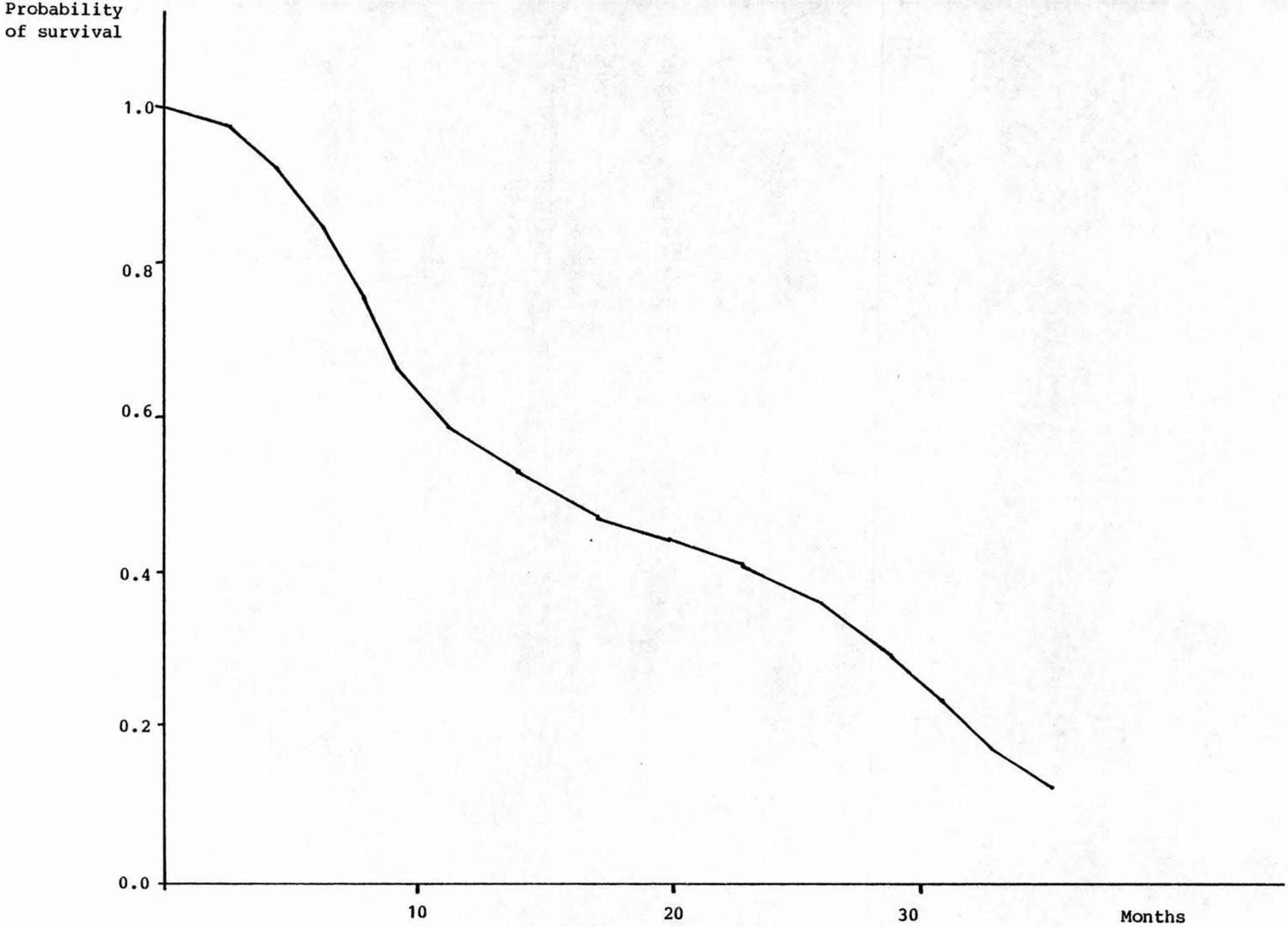


Figure (3.2.11)

discussed and some parametric distributions were mentioned that can approximate the distributions. The shape of the hazard may give useful information as far as the biological nature of the progress of disease is concerned. It can also introduce tests of significance. For example we could test the null hypothesis of a constant hazard (exponentially distributed density function) against the alternative with increasing or decreasing hazards (Weiball density functions). In the next section we will discuss parametric distributions purely for the purpose of testing the effects of treatments and other concomitant variables, by the use of covariates.

3.3 Inclusion of Covariates.

Once a decision is made on the shape of the hazards that may be fitted to the population, an additional function may be combined with the hazard function to form a hazard function for a specific sub-population. This additional information is related to the extra function and is refered to as the concomitant information. Examples of concomitant information are indicators for the treatment effects, age of patient at entry, stage of disease, size of tumour and other prognostic indicators. These additional sets can be used either singly or in combination to estimate parameters so that a distinct survival distribution may be fitted to each subgroup. The estimated value of such parameters will be used to assess the significance of the survival differences between two or more subgroups.

In the above discussion we have made a necessary distinction between the hazard functions and the concomitant variables. The former provides information on the rate of failure of the patients while the /

the latter defines subgroups of patients. The distinction may be more complex and difficult at times in deciding which parameter is appropriate for assessment and comparison of the subgroup. Using the extreme example of crossing survival curves, an interpretation of the parameter estimates can depend to some extent on the weighting attached to the various points in time. However it is not a problem that one often encounters in practice. We proceed now with the development and representation of parametric statistical methods which are useful in clinical trials.

For each case entered into the trial, in addition to failure time or censoring time t_i and the indicator variable δ_i, (o for censored, 1 for uncensored response), there exists a vector $Z_i = (Z_{1i} \ldots . Z_{ri})$ of covariate indicators or explanatory variable indicators. Then according to the previous definitions of the hazard rates for each subgroup we can represent the hazard rate as .

(Hazard at time t, for subgroup k) = (General hazard at time t)
(Function of variable indicator for subgroup k)

In the simple case of the exponential distribution with the general hazard rate λ_0 we can write the above as

$$\lambda(t_1 Z_i) = \lambda_0 \circ EXP (\beta' Z_i)$$

where EXP $(\beta' Z_i)$ is a mathematically convenient function for representing multiplicative effects of indicator variables. β is a vector to be estimated and represents a set of coefficients associated with the covariates and is used for the testing of prognostic/

prognostic effects indicators. One point to note at this stage is that our formulations need not be as restrictive as the above formulation. Later we will discuss a group of models that are based on the following formulation.

(Hazard at time t, for subgroup k) = (General hazard at time t for subgroup k) ° (Function of variable indicators for subgroup k)

The former models are in general named as proportional hazard model and an example of the latter model is the accelerated failure time model.

The proportional hazard models are expressed as -

$$\lambda(t, Z_i) = \lambda_o(t) \exp(\beta Z_i)$$

where $\lambda_o(t)$ is a function of time referring to a base line hazard rate. In the case of the exponential it is not dependent on time and in the case of the Weiball it is expressed as $\lambda_o(t) = \mu\upsilon\, t^{\upsilon-1}$, where μ and υ are scale and shape parameters.

One point in introducing the concept of covariates is that it enables us to compare different treatments for a single disease. Further it is possible to identify auxillary factors that influence survival times. The use of concomitant information is an approach for identifying the factors that are associate with the survival times in relative terms. This latter emphasis is different to the discussion of earlier parts of this chapter on parametric methods, which dealt with parametric estimation of survival times and a possible interpretation based on the functional form of the

parametric/

parametric models. The procedure commonly used for the estimation and testing of the effects of the covariates is termed as the maximum likelihood estimation and is dealt with by S.D. Silvey (1975). Basically, to assess the effects of factors influencing the survival times we require a function that can express the survival experience of all cases.

Thus

$$\text{likelihood function} = \prod_{\substack{\text{all} \\ \text{patients}}} (\text{likelihood of survival experience of a case})$$

Further we can distinguish censored cases and responding cases, and thus we write;

$$\text{likelihood function} = \prod_{\text{deaths}} (\text{death density function}) \prod_{\text{alives}} (\text{survival function})$$

By the definitions of the hazard functions, survival functions and density functions we can write the above equivalently as,

$$\text{likelihood function} = \prod_{\text{deaths}} (\text{hazard function}) \prod_{\text{all}} (\text{survival function})$$

Each of the above function in brackets can have a mathematical formulation, based on insight into the distributional form of the data. Further each of these formulations may be defined by a set of parameters. Our intention is to use a procedure to estimate the best values of parameters that can explain the survival experience of the population with the least number of parameters and with an acceptably low difference between estimated, expected and actual survival times. Later in this chapter we will develop the above formulation for the exponential, Weiball and Taulbee approach. Also based on the distributions we will define general families of functions.

When/

When parametric methods are used in conjunction with covariate effects more care is needed for identifying the correct functions. What is crucial in survival analysis as in any branch of applied statistics is obtaining a reasonable fit to the data. A commonly used indicator of a good model for the data is the pattern of the residuals, where the residuals are defined to be a function of the difference between predicted and observed values. Further a plot of the data can at times indicate whether the theoretical model's range can fall within the variability of the data.

An example is the situation where it is assumed that there exists a constant hazard rate for a population. Therefore the best model to fit is conjectured to be the exponential distribution, which has a constant hazard rate. Once the data is fitted and the values of the residuals are compared, possible short comings of the model may become apparent. If the conjecture is substantiated by the data, then the outcome would be a set of residuals that follow a constant pattern through time. In situations that the hazard rate is increasing or decreasing a similar pattern will be reflected by the residuals.

In the formulation of the likelihood with the proportional hazard assumtion the hazard rate is assumed to be dependent on the covariates only through the EXP (βZ) function. It is however possible that in some situations with the passage of time the effects of covariates may change. One manner in which we can test the time dependency assumtions of covariates in a proportional hazard model is by formulation of a likelihood such as;

$$\text{likelihood function} = \prod_{\substack{\text{death}\\\text{cases}}} (\text{hazard function}) \prod_{\substack{\text{all}\\\text{cases}}} (\text{survival function}) \prod_{\substack{\text{all}\\\text{cases}}} (\text{time depencency function}) \qquad (3.3.1)$$

In the following sections of this chapter we will study in detail different methods of the estimation of covariate effects for different shapes of hazard rates. Before doing so we will remark on the various advantages of the approaches that have been discussed so far.

Earlier we mentioned Turner's family of distributions as a flexible multi-parametric method by which it is possible to obtain a close fit to the subgroups of the data as a method for data reduction. For reasons of comparison between subgroups, there may be situations where it is sufficient to use a multi-parametric method for a base line hazard rate together with a simple single parameter relative risk. Clearly estimation of a large number of nuisance parameters is an important consideration in such a study. Alternatively in other situations a multi-parametric method may be used with a multi-parameter relative risk for each subgroup. This approach has the disadvantage that the interpretation of the inference of the subgroup may not be easy. The likelihood function (3.3.1) has a major advantage in that the interpretation of the covariate effects is much simpler than when separate distributions are fitted to different subgroups.

In terms of the survival curves the above formulation of the proportional hazards may be interpreted as follows. Between the range of all/

of all possible survival curves for the set of prognostic and treatment groups, there exists a base hazard rate. All other survival curves can further be generated after multiplying the base hazard rate by the corresponding function of EXP ($Z\ \beta$), which is a scaler for each subgroup. If the proportional hazard assumption holds the final term in (3.3.1) contributes nothing to the covariate effects. If alternatively the hazards behave as non proportional rates then a time dependent functional form of Z(t) must be used instead of Z.

3.4 Polynomial Hazard Rates.

Taulbee (1979) discusses a generalised form of the Ragleith distribution, in which the hazard has a polynomial pattern.

$$\lambda_m(t) = \lambda_o + \lambda_1 t + \lambda_2 t^2 + \ldots \lambda_m t^m = \sum_{k=0}^{m} \lambda_k t^k \qquad (3.4.1)$$

where m refers to the degree of the polynomial.

In the presence of covariate effects Z_j for j = 1,...S we may then adopt a substitution for λ_k such as $\lambda_k \exp(B_k Z_i)$ giving

$$\lambda_k(t,Z_i) = \lambda_k t^k \exp(B_k Z_i)$$

$$\text{for } k = 0, 1 \ldots m$$

Where in the above definitions we have considered B_k to be a parameter set to be estimated for each of k = 0 to m. Further for each B_k there exists a representation $B_k = (B_{k1}, \ldots B_{ks})$ and for the vector Z_i ther is a representation $Z_i = (Z_i, \ldots .Z_{is})$.

Where/

Where s is number of covariates and i refers to a particular case.

EXP ($B_k Z_i$) is numerically the most convenient function although in general we can express the above as

$$\lambda_k (t,Z_i) = \lambda_k t^k \quad h(Z_i, B_k) \qquad (3.4.2)$$

and let h (B_k, Z_i) be for example EXP ($B_k Z_i$), $(1+B_k Z_i)$ or $(1+ B_k Z_i)^{-1}$. In here we adopt a general definition of B_k. However later we will adopt a restricted f orm of B_k where $B_0 = B_1 = B_2$, giving a proportional hazard type of the model.

In analogous manner to that of other parametric models such as the Weiball, it seems necessary that a good prior knowledge is required for use of any particular hazard shape. Further the functional form of $h(B_k, Z_i)$ is related to the derivation of the functional form of the subgroup hazard rates from that of the base line hazard, $\lambda_k (t,0)$. In particular this relation is important for the proportional hazard restriction form of the model. We will discuss these points later with the use of a particular form of Rayleigh distribution with increasing hazards, that is $\lambda_0 > 0$ and $\lambda_1 > 0$. By substituting (3.4.2) and expanding (3.4.1) we write the general hazard function as

$$\lambda_m (t,Z_i) = \lambda_0 h(Z_i, B_0) + \lambda_1 h(Z_i, B_1) t + \ldots + \lambda_m h(Z_i, B_m) t^m$$

$$= \sum_{k=0}^{m} \lambda_k h(Z_i, B_k) t^k$$

Using/

Using the definitions from the introduction we have

$$S(t) = \text{EXP}\left[-\int_0^t \lambda(u)du \right]$$

giving

$$S_m(t,Z_i) = \text{EXP}\left[-\int_0^t \sum_{k=0}^{m} \lambda_k h(Z_i, B_k) u^k du \right]$$

$$= \text{EXP}\left[-\sum_{k=0}^{m} \lambda_k h(Z_i B_k) \left\{ \int_0^t u^k du \right\} \right]$$

$$= \text{EXP}\left[-\left\{ \sum_{k=0}^{m} \frac{\lambda_k}{k+1} h(Z_i, B_k)\, t^{k+1} \right\} \right]$$

Further using usual approaches in the construction of the likelihoods we have

$$\text{Likelihood} = \prod_{\substack{\text{All} \\ \text{deaths}}} (\text{Hazard function}) \prod_{\text{All cases}} (\text{Survivor function})$$

$$= \prod_{\substack{\text{deaths} \\ (i)}} \lambda_m(t, Z_i) \prod_{\text{all}\ (i)} S_m(t,Z_i)$$

$$= \prod_{\substack{D \\ (i)}} \sum_{k=0}^{m} \lambda_k h(Z_i, B_k) t^k \cdot \prod_{\substack{A \\ (i)}} \text{EXP}\left\{ -\left[\sum_{k=0}^{m} \frac{\lambda_k}{k+1} h(Z_i, B_k) t^{k+1} \right] \right\}$$

Then the likelihood function L, for a 2 degree hazard is given by (we eliminate subscript i, for the moment for brevity).

$$L = \prod_{D} \left[\lambda_0 h(Z, B_0) + \lambda_1 h(Z, B_1) t + \lambda_2 h(Z, B_2) t^2 \right]$$

$$\cdot \prod_{A} \text{EXP}\left\{ -\left[\lambda_0 h(Z, B_0) t + \frac{\lambda_1 h(Z, B_1) t^2}{2} + \frac{\lambda_2 h(Z, B_2) t^3}{3} \right] \right.$$

We now reinsert subscript i

giving

$$\ell = \ln L = \sum_{\substack{\text{All} \\ (i)}} \delta_i \ln(F1_i) + F2_i$$

where,

$$F1_i = [\lambda_0 h(Z_i, B_0) + \lambda_1 h(Z_i, B_1)t + \lambda_2 h(Z_i, B_2)t^2]$$

and

$$F2_i = [-\{\lambda_0 h(Z_i, B_0)t + \frac{\lambda_1 h(Z_i, B_1)t^2}{2} + \frac{\lambda_2 h(Z_i, B_2)t^3}{3}\}]$$

Now we do a differentiation of the necessary parameters for maximum likelihood estimates

$$\frac{\partial \ell}{\partial \lambda_0} = \sum_{\substack{\text{All} \\ (i)}} \delta_i \frac{h(Z_i, B_0)}{F1_i} + (-h(Z_i, B_0)\, t)$$

$$\frac{\partial \ell}{\partial \lambda_1} = \sum_{\substack{\text{All} \\ (i)}} \delta_i \frac{h(Z_i, B_1)t}{F1_i} + \left(-\frac{h(Z_i, B_1)t^2}{2}\right)$$

$$\frac{\partial \ell}{\partial \lambda_2} = \sum_{\substack{\text{All} \\ (i)}} \delta_i \frac{h(Z_i, B_2)t^2}{F1_i} + \left(-\frac{h(Z_i, B_2)t^3}{3}\right)$$

$$\frac{\partial \ell}{\partial \beta_{0j}} = \sum_{\substack{\text{All} \\ (i)}} \delta_i \left(\frac{1}{F1_i}\right) \frac{\partial h(Z_{ij}, B_{oj})}{\partial B_{oj}} + (-\lambda_o t) \frac{\partial h(Z_{ij}, B_{oj})}{\partial B_{oj}}$$

$$\frac{\partial \ell}{\partial \beta_{1j}} = \sum_{\substack{\text{All} \\ (i)}} \delta_i \left(\frac{1}{F1_i}\right) \frac{\partial h(Z_{ij}, B_{1j})}{\partial B_{1j}} (\lambda_1 t) + \left(\frac{-\lambda_1 t^2}{2}\right) \frac{\partial h(Z_{ij}, B_{1j})}{\partial B_{1j}}$$

$$\frac{\partial \ell}{\partial B_{2j}} = \sum_{\substack{\text{All} \\ (i)}} \delta_i \left(\frac{1}{F1_i}\right) \frac{\partial h(Z_{ij}, B_{2j})}{\partial B_{2j}} (\lambda_2 t^2) + \left(\frac{-\lambda_2 t^3}{3}\right) \frac{\partial h(Z_{ij}, B_{2j})}{\partial B_{2j}}$$

Where $\delta_i = 1$ for deaths and $\delta_i = 0$ for censored cases.

B_{oj}, B_{1j}, B_{2j} are covariates for each degree of the polynomial. (In this case a 2 degree polynomial.)

$\lambda_o, \lambda_1, \lambda_2$ refer to the hazards polynomial,

and i is a subscript for each case and j is the number of covariate effect under test.

In the above formulations we have allowed B_k to vary depending on the degree of the polynomial that approximates the hazard rate. This generality is violating the proportional hazards assumption. A restriction such as $B_k = B$ for all k, converts the approach to a proportional hazards version. In such a situation the hazard is,

$$\lambda_m (t, Z_i) = \left(\sum_{k=0}^{m} \lambda_k t^k \right) \circ h(Z_i, B)$$

giving for a 2 degree hazard

$$\lambda_2(t, Z_i) = (\lambda_0 + \lambda_1 t + \lambda_2 t^2) \circ \text{Exp}(B Z_i)$$

Further $\lambda_1(t, Z_i)$ using the above generality is an example of the Rayleigh distribution hazard rate, with covariates

$$\lambda_1 (t, Z_i) = (\lambda_0 + \lambda_1 t) \circ \text{Exp}(B Z_i)$$

for $\lambda_0 > 0$ and $\lambda_1 > 0$

We proceed with this approach in an analysis of the Edinburgh trial data/

data, using maximum likelihood estimation. For our particular use we adopt a method of maximising the likelihood function using the P3R programme of the BMDP for estimation of non-linear regression models by the Newton-Raphson procedures. This programme is a flexible enough procedure for the estimation of the relevant parameters of the above named functions.

The programme requests the actual likelihood function, the derivaatives with respect to the estimating parameters and a loss function. In the last section we produced the necessary functions and derivatives for a general Taulbee approach. Initially we analyse the data for a linear hazard model with the proportional hazard assumption. In this analysis we use the treatment option given by Radiotherapy and simple surgery against radical surgery as the main effect of the study. Later with use of the other covariates we approach the analysis with the Weiball and the exponential models. Throughout we use a survival time scale in months.

First we fit a model with a zero rate hazard. This is equivalent to an exponential model with the proportional hazard assumption. With m = 0 we have

$$\lambda(t, Z) = \lambda_0 \circ \mathrm{Exp}(\beta Z)$$

$$L = \prod_{D} \lambda_0 \mathrm{Exp}(\beta Z) \circ \prod_{A} \mathrm{Exp}\left[-\lambda_0 \mathrm{Exp}(\beta Z)\, t\right]$$

giving the estimated parameters

λ_0 = .0041 S.E. = 0.00309)
) df = 559 lnL = -1741.82
β = 1.4821 S.E. = 0.4021)

Now we can expand the model by allowing the hazard to have a straight line passing through the origin. If we set $\lambda_0 = 0$, then hazard is

$$\lambda(t, Z) = (\lambda_1 t) . \text{Exp}(\beta Z)$$

The present model is not suitable for the purpose of analysis in that we have introduced two types of restriction, one indicating $\lambda_0 = 0$ and the other assessing the proportionality of hazards. As a more suitable model we use the next memeber of this calss of distribution.

We now fit a model of the hazard form that allows the straight line hazard not to pass through the origin. We therefore have to estimate both parameter λ_0 and λ_1 simultaneously as well as the βparameter for the covariates.

We thus obtain the following estimated parameter for the model given by

$$\lambda(t, Z) = (\lambda_0 + \lambda_1 t) . \text{Exp}(\beta Z)$$

$$L = \prod_D \lambda_0 \text{Exp}(\beta Z) + \lambda_1 \text{Exp}(\beta Z) t . \prod_A \text{Exp}\left[-\lambda_0 \text{Exp}(\beta Z) t + \frac{\lambda_1 \text{Exp}(\beta Z) t^2}{2}\right]$$

λ_0	= 0.008762	S.E.	= 0.00186	)	
λ_1	= 0.000438	S.E.	= 0.00382	)	d.f. = 558,
β	= 1.4951	S.E.	= 0.4037	)	ln L = 1739.85

The value of λ_1 is close to zero. In fact there is little improvement/

ment over the original straight line model with λ_0 parameter used as the hazards function only. Now, although the base line is approximated better and is less restricted, the estimator of the treatment effect is virtually unchanged. This indicates that the covariate effect part of the hazard namely $\text{Exp}(\beta Z)$ is consistant if we can assume the proportionality of the hazards.

The next model we consider, relaxes the proportional hazards assumption. This is a useful model for checking the proportionalities of linear type. Returning to the original derivations of the model we can express the hazard rates of the next model as,

$$\lambda(t, Z) = \lambda_0 \text{Exp}(\beta_0 Z) + \lambda_1 \text{Exp}(\beta_1 Z)t$$

$$L = \prod_D \lambda_0 (\text{Exp}(\beta_0 Z) + \lambda_1 \text{Exp}(\beta_1 Z)t \cdot \prod_A \text{Exp}[-(\lambda_0 \text{Exp}(\beta_0 Z)t + \frac{\lambda_1 \text{Exp}(\beta_1 Z)t^2}{2})]$$

giving the estimator

λ_0	= 0.009674	S.E. = 0.00257	)
λ_1	= 0.000511	S.E. = 0.00376	) d.f. = 557
β_0	= 1.12	S.E. = 1.311	) ln L = -1737.97
β_1	= 1.4848	S.E. = 0.4121	)

The value of β_0 is not significant. A comparison of the log likelihood of this model with 557 degrees of freedom and the previous model with 558 degrees of freedom gives the difference of $-2\ln L = 3.76$, which according to the chisquared distribution is not significant. We therefore do not reject the proportionality of hazards assumtion.

By/

By the above models the linear structure of the hazard shapes may not allow an efficient estimation of the effects. This point is expressed more vividly when we deal with the Weiball models of the next section. In the discussions of the exponential model estimator it will be made clear that the actual value of λ_0 is arbitrary in so far as the comparison of β's for different subgroups are concerned.

So far in the study of the application of polynomial hazard rates, the above linear hazard rate has been the most appropriate. We will now estimate some of the subgroup covariate effects by this model and observe their contribution to a proper explanation of patient survival variation. The prognostic categories that are of particular interest now and which will be discussed in more detail later are, menopausal status, initial size of the tumour and node histology status. Later in chapter 6 we will define the indicators in greater detail. In here we will only use them for the purpose of illustration.

We fit covariate effect models to the data for each of the above main effects in presence of the treatment effects. Consistantly we note that there is a reduction in treatment effect of $\hat{\beta}$ estimator and a comparison of the covariate functions does not show any important difference in the treatment effect estimators. This indicates that the treatment effect is stable for the different prognostic groups.

Model with treatment and node,

$\beta_{\text{treatment}}$ = 1.4521 S.E. = 0.4072

β_{node} = 1.2182 S.E. = 0.6410

Model/

Model with treatment and size,

$\beta_{treatment}$ = 1.4486 S.E. = 0.4513

β_{size} = 1.1087 S.E. = 0.6834

Model with treatment and menopausal status, where we consider post menopausal and menopausal group as one category and pre-menopausal as another.

$\beta_{treatment}$ = 1.2730 S.E. = 0.4481

$\beta_{menopausal}$ = 1.3051 S.E. = 0.5913

3.5 Exponential distribution for censored survival data with covariates.

The exponential distribution has been used extensively as a basis for study of survival distributions. The simplicity of this model has been the main reason for its common usage. However, at times it has been used in situations where the assumption of constant hazards has been violated. The model is simple to estimate and has only one parameter for defining the failure rate which is not dependent on time. Thus in this model the risk of death is independent of time.

In an early demonstration of the exponential survival distribution Boag (1949) applied the distribution to the survival of cancer patients. David (1952) examined the distribution in relation to the field of reliability and applied the method to 26 mechanical survival situations. Several authors, Halperin (1952) and Epstein and Sobel (1953, 1954) investigated the maximum likelihood estimation of the one parameter case, λ, for censored data. Fiegle and Zelen (1965) investigated the problem

of /

of estimation of constant linear hazards λ with covariates. The exponential distribution at times has been termed as the "memory less distribution" since the hazard rate is not a function of time.

The median of the distribution is $\ln(2) / \lambda$ mean is $1/\lambda$ and the variance is $1/ \lambda 2$. Where λ is interpreted as the force of mortality. The larger the value of λ , the shorter is the mean life. The estimation for the uncensored case of the distribution is relatively simple. Using the maximum likelihood estimation there is a closed bound solution for the mean and variance.

For a situation of random censorship, defined in the introduction to be the most common censoring in trials, we have the likelihood

$$L = \prod_{\text{Deaths}(i)} f(t_i) \prod_{\text{Censored}(i)} S(t_i)$$

$$= \prod_{\text{Deaths}(i)} \lambda e^{-\lambda t_i} \prod_{\text{Censored}(i)} e^{-\lambda t_i}$$

$$= \prod_{i=1}^{n} (\lambda e^{-\lambda t_i})^{\delta_i} (e^{-\lambda t_i})^{1-\delta_i}$$

$$= \prod_{i=1}^{n} \lambda^{\delta_i} (e^{-\lambda t_i})$$

$$= \lambda^{n^*} \operatorname{Exp}\left(-\lambda \sum_{i=1}^{n} t_i\right)$$

where n^* = number of uncensored or deaths

giving

$$\ln L = n^* \ln(\lambda) - \lambda \sum_{i=1}^{n} t_i$$

$$\frac{\partial \ln L}{\partial \lambda} = \frac{n^*}{\lambda} - \sum_{i=1}^{n} t_i = 0$$

=> the maximum likelihood estimator of λ has value

$$\hat{\lambda} = [\sum_{i=1}^{n} t_i / n^*]^{-1}$$

The second derivative of the log likelihood with respect to λ yields

$$\frac{\partial^2 \ln L}{\partial \lambda^2} = -n^* / \lambda^2$$

so that $\frac{\hat{\lambda} - \lambda}{\sqrt{\lambda^2/n^*}}$ is approximately normally distributed as N(0,1)

Using the asymptotic normality results on likelihoods. Further a transformation by the delta method gives

$$\hat{\lambda} \sim N(\lambda, \lambda^2/n^*)$$

The exponential distribution with only one parameter λ is rather simple to obtain. The next stage of the development of the exponential distribution is to use covariates. The use of covariates with an exponential hazard rate may be developed with the maximum likelihood estimation and the Newton-Raphson procedure. The reason for the use of the Newton Raphson procedure is that we often do not have a closed bound solution of the estimator. The interpretation of the results are also straight forward if the assumption of time independent hazards holds. In the above we obtain a method for the estimation of the hazard rate λ . In the situation of analysis with covariates and use of maximum likelihood estimation, λ in fact is not needed and it is possible to derive an inference for the covariates by setting λ, the base hazard rate value to 1. In the section on the Weiball we will derive/

derive functions for a maximum likelihood estimation of covariates and the Weiball shape parameter.

The exponential with covariates is a special case of that procedure and will be discussed in more detail there.

3.6 the Weiball distribution.

The Weiball distribution was originally used by a Swedish physicist, Waladdi Weiball, who was interested in measuring the breaking strength of materials. The main reason for the initial interest on the Weiball distribution was that unlike the exponential it was able to fit the data, even when the breaking rate was not constant. Later A.C. Cohen produced maximum likelihood estimators for estimation of uncensored and censored cases.

The Weiball distribution is an extension of the exponential distribution. In the graphical presentation of the first section of this chapter, it was shown how, with a shape parameter set to one the Weiball distribution reduces to the exponential. Apart from the situation with the shape parameter set to one, the hazard function in the Weiball is time dependent, and thus the rate of failure changes with the passage of time.

For the distributional definitions of Weiball in section 3.1 we have; the median of the Weiball is given by $(\ln (2) / \mu)^{1/\nu}$, the mean is given by

$$\frac{\Gamma(1 + 1/\nu)}{\mu^{1/\nu}}$$

and the variance is

$$\frac{\Gamma(1 + 2/\nu) - (\Gamma[1 + {}^{1}/_{\nu}])^2}{\mu^{2/\nu}}$$

In here ν is related to the hazard rate. The actual interpretation of the ν for cases greater that one and less than one is the same as those for applications in which rate of death changes because of the underlying biological process. The estimation procedure is more complex than the exponential case. Neither in uncensored nor in censored models does there exist a closed maximum likelihood estimator. We now proceed with the derivation of the maximum likelihood estimator for a Weiball model with covariates.

In the last section concomitant information was introduced into the likelihood function for the Taubles general model. According to the hazard functions, the subgroups affect the rate of events in terms of intensity with a relationship of exp (β Z). However, the covariate part of the model does not effect the shape of the base line hazard rates.

The effect of the covariate on the Weiball hazard is represented by

$$\lambda(t , Z) = \nu\mu(t)^{\nu-1} \cdot e^{Z\beta}$$

The latter part Exp (Z β) refers to covariates and is independent of time. This is basically a similar assumption as the one used in the last section to derive some of the results, for concomitant information.

According to the relations in section 3.1 and the above hazard rates, the density function and survival function for the Weiball with covariates are

$$f(t, Z) = \nu\mu(t)^{\nu-1} e^{Z\beta} .Exp[- \mu t^{\nu} e^{Z\beta}]$$

$$S(t, Z) = \mu \, Exp [- \mu t^{\nu} e^{Z\beta}]$$

All the above expressions of the hazard rate, density function and the survival function can be considered as a generalisation of the exponential distribution with concomitant variables , simply by allowing ν to be set to one. These results can serve as a general purpose model for the different forms of the exponential and Weiball models, when the emphasis is on the estimation of the covariate effects. Using the formulations from previous sections, the likelihood function is -

$$\text{likelihood} = \prod_{i=1}^{n} [\mu\nu t^{\nu-1} Exp(\beta Z_i)]^{\delta i} Exp[(-t^{\nu} \mu c^{\beta Z_i})]$$

where as before δ_i = 1 for death and
= 0 for censorings

and n is the total sample
and n* is the number of deaths.

$$= \prod_{i=1}^{n} \mu^{\delta_i} [\nu t^{\nu-1} Exp(\beta Z_i)]^{\delta_i} Exp [(-t^{\nu} e^{\beta Z_i})] e^{\mu}$$

The value of μ is independent of the time t and covariate effect Exp (βZ_i). Thus it is a scaling measure and μ does not have an effect on the comparative values of $\hat{\beta}$ and $\hat{\nu}$. In the following expression for the log likelihood, the terms involving μ are omitted.

$$\ln L = \sum_{i=1}^{n} \delta_i \ln[\nu t_i^{\nu-1} Exp(\beta Z_i)] + (-t_i^{\nu} e^{\beta Z_i})$$

$$= \sum_{i=1}^{n} \delta_i(\ln \nu + (\nu-1)\ln t_i + \beta Z_i) + (-t_i^{\nu} e^{\beta Z_i})$$

$$= n^* \log \nu + \sum_{i=1}^{n} \delta_i[(\nu-1)\ln t_i + \beta Z_i] + (-t_i^{\nu} e^{\beta Z_i})$$

$$\ln L = n^* \ln \nu + \sum_{i=1}^{n} \delta_i (\nu-1 \, t_i + \beta Z_i) + (-t_i^{\nu} e^{\beta Z_i}) + \delta_i \ln t_i$$

The last part $\delta_i \ln t_i$ is not dependent on parameters $\hat{\nu}$ and $\hat{\beta}$. Thus in terms of the proportionality of the likelihood we let

$$R_i = t_i^{\nu} e^{\beta Z_i} \quad \text{and} \quad \ln R_i = \nu \log t_i + \beta Z_i$$

giving

$$\ln L = n^* \ln \nu + \sum_{i=1}^{n} \delta_i (\ln R_i) - R_i$$

Now we obtain the derivatives of the logarithms of the likelihood. These derivatives at values equal to zero give the best estimators of the maximum likelihood function.

$$\frac{\partial \ln L}{\partial \beta_i} = \sum_{i=1}^{n} \delta_i Z_{ij} - t_i^{\nu} c^{\beta Z_i} \cdot Z_{ij} = \sum_{i=1}^{n} (\delta_i - R_i) Z_{ij} = 0$$

$$\frac{\partial \ln L}{\partial \nu} = \frac{n^*}{\nu} + \sum_{i=1}^{n} \delta_i \ln t_i + (-t_i^{\nu} c^{\beta Z_i}) \cdot \ln t_i = \frac{n^*}{\nu} + \sum_{i=1}^{n} (\delta_i - R_i) \ln t_i = 0$$

In the procedure for estimator of $\hat{\beta}_j$ and $\hat{\nu}$ we also need to know the second derivatives of the logarithms of the likelihood. These values are used in the maximising procedures of the Newton-Raphson as well as deriving the information matrix to obtain the variance covariance esimators.

$$\frac{\partial^2 \ln L}{\partial \beta_j \, \partial \, \beta_R} = \sum_{i=1}^{n} - t_i^{\nu} e^{\beta Z_i} \cdot Z_{ij} \cdot Z_{ik} = - \sum_{i=1}^{n} R_i Z_{ij} Z_{ik}$$

$$\frac{\partial^2 \ln L}{\partial \beta_j \partial \nu} = \sum_{i=1}^{n} - t_i^{\nu} e^{\beta Z_i} \cdot Z_{ij} \cdot \ln t_i = - \sum_{i=1}^{n} R_i Z_{ij} \ln t_i$$

$$\frac{\partial^2 \ln L}{\partial \nu^2} = \frac{-n^*}{\nu^2} + \sum_{i=1}^{n} - \left(t_i^{\nu} \; e^{\beta Z_i} \right) . \ln t_i . \ln t_i =$$

$$\frac{-n^*}{\nu^2} + \sum_{i=1}^{n} - R_i (\ln t_i)^2$$

The above functions are thus the necessary functions that may be used in conjunction with a standard Newton-Raphson maximisation. S.D. Silvey (1975) describes such a procedure.

3.7. Interpretation of the models with use of the old Edinburgh Trial Data.

In this section we perform an analysis of the old Edinburgh trial data, with the parametric methods. The general purpose is to give a comparative illustration of the parametric and non-parametric methods as discussed in the last section. First we perform an exponential and then a Weiball model analysis with only one regression coefficient.

The first covariate we test is the treatment effect, that is a comparison of survival times for simple surgery and radio-therapy against radical surgery. A t shape parameter value fixed to one we are in fact using an exponential model. Further the variation of the shape parameter from one indicates a Weiball model.

Shape parameter set to one $\beta_{option} = 1.4821$ S.E. = .4021 d.f.= 559

Shape parameter estimated = 1.36 $\beta_{option} = 1.6321$ S.E.= .4043 d.f. = 558.

(S.E. of the shape parameter = 1.18)

We continue the estimation procedure with inclusion of another covariate/

covariate, the menopausal status. Once again we consider post-menopausal and menopausal as one category and the pre-menopausal as a separate category.

Shape parameter	β option	S.E.	β meno	S.E.	d.f.
set to one	1.3819	.4921	1.3722	.3821	558
estimated = 1.53	1.3521	.4185	1.3986	.2581	557

In both of the above models we note that option and menopausal status play an important role in describing the survival rate of the patients. Then we consider the addition of tumour size.

Shape parameter	β option	S.E.	β meno	S.E.	β size	S.E.	d.f.
set to one	1.4181	.4931	1.3843	.3840	1.1927	.3192	557
estimated=1.54	1.4961	.4166	1.3506	.2931	1.1159	.2901	556

Now we add a term for node status to the above models -

Shape para.	β option	S.E.	β meno	S.E.	β size	S.E.	β node	S.E.	d.f.
set to one	1.423	.4930	1.4134	.3872	1.3741	.3793	1.4721	.5128	556
est.= 1.54	1.478	.4381	1.3902	.2881	1.1462	.3121	1.531	.6321	555

The above models show that the survival differences of the patients can be attributed to the above covariate indicators. So far we have not considered significance levels of the different estimators for the parametric methods. In Chapter 6 we put more emphasis on the analysis and interpretation of the data rather than a comparison of the analytical methods. In summary the above models indicate that for the above covariates there is very little to choose from the exponential and the Weiball. The results of the Taublees family for the 2nd term, also show very similar results which, because of their close similarity are not detailed here.

Now/

Now we introduce the concept of interaction and its use in the framework of a parametric model. In the second stage of the last analysis with the exponential and the Weiball models, the information from option and menopausal status played the most important role. One advantage in use of a regression model is that we are able to do a formal test of interaction effects. These tests assess if the effect of covariates acting simultaneously is any different from an addition of the two effects acting independently. Once again we represent menopausal status in two categories of pre-menopausal and menopausal + postmenopausal. The effect of the latter two categories of the menopausal status can be seen from the shape of the hazards, which are in fact very similar. Further, for the present purpose such a transformation of the menopausal status suffices.

We begin with a model which was presented at above and included menopausal status and the treatment option as the only two effects. Now we continue with a test of an interaction effect for treatment and menopausal effects.

Shape parameter	β option	S.E.	β meno	S.E.	β size	S.E.	d.f.
set to one	1.3839	.4938	1.2121	.3109	.2127	.3782	557
estimated= 1.52	1.3210	.4179	1.2382	.2052	.6171	.6312	556

This result indicates that all the necessary information may be contained within the two main effects. We, therefore conclude that the radical treatment group perform better in terms of survival time. The effect of treatment is consistently the same for the various categories of the prognostic indicators, size, node and menopausal/

menopausal status. The behaviour of the various categories of the indicators is as may be expected. That is, the smaller tumours younger patients and the node negative tumours are the good prognosis groups and the older patients, larger tumours and node positive are the group providing the worst survival times.

None of the main effects of the prognostic values show an important interaction with treatment effects. That is all subgroup variability of the survival times can be described in an additive manner.

The final model of the Weiball and the exponential distribution with all three covariates and treatment effect included shows very similar estimators of the prognostic main effects in comparison to models with treatment and one covariate effect included, thus once again suggesting that prognostic values are consistently the same given the present framework of the Weiball model.

3.8. Families of distribution with covariate effects.

The Taulbee or Turner family of distribution can provide a flexible set of distribution for use in failure time analysis. When we deal with covariates there is another approach to classifying distribution according to a combination of hazard rates and covariate effects. The most commonly used method is to assume that the population has a single underlying failure rate according to the inherent nature of disease. Further any difference in failure rates for the subgroup originates from a separately identified covariate effect. This class are termed on the proportional/

hazards model and the exponential , Weiball and the polynomial models of the previous section were based on its assumptions. An alternative useful approach is to consider the failure rate to have a function dependent on time and the covariate structure. This group is known as the accelerated failure time model and we will consider them later in this section.

The group of regression models with the assumptions of the proportional hazards are generalised as models of the form

$$\lambda(t, Z) = \lambda_0(t) \text{ Exp}(\beta Z) \qquad (3.8.1)$$

Now if we let $\lambda_0(t)$ to be independent of time and set $\lambda_0(t) = \lambda$ we have an exponential distribution with covariates. Alternatively, if we let $\lambda_0(t)$ to be time dependent with a shape parameter μ and set

$$\lambda_0(t) = \mu\nu\, t^{\nu-1} \qquad (3.8.2)$$

We have a Weiball distribution. In case of the Reyleigh family of distribution or a restricted Taulbee approach we deal with hazards of the form,

$$\lambda_0(t) = \lambda_0 + \lambda_1 t \qquad (3.8.3)$$

In terms of the reduction of the data into a useful statistic, it is clear that estimation of the β gives the relevant information on effects due to membership of a particular subgroup. The assumption that must hold is that, the membership into a particular subgroup does not effect the sape of the $\lambda_0(t)$. That is, there exists a base line hazard rate for the total population and any effects due to covariates take the form of a proportional effect introduced as Exp (β Z).

In the estimation of the maximum likelihood for the exponential/

exponential and the Weiball, the actual scale of the hazard curve plays an arbitary role in the relative effects of the covariates. What matters is in fact only the shape parameter of the Weiball. The important assumption that must hold again due to the proportional hazards, is the fact that regardless of the subgroup, the total population must have the same shape parameter. In terms of interpretation we require that the base line hazard rate can be projected on to the subgroup rates for the entire population.

Up until now we have considered a general proportional hazard model in which, the entire population has had the same base line hazard rate. There is an extension with which we can allow more than one base line hazard rate. However the information contributing to the covariate effects is inherently the same. These models are useful in situations that a population is composed of different strata. The information regarding membership of a particular strata is not testable, but information regarding some other covariate must be estimated by allowing for the strata effects. The model has the form.

$$\lambda_j (t, Z) = \lambda_{0j}(t) \text{ Exp } (\beta Z)$$

where j refers to a particular strata. This functional form of $\lambda_{0j}(t)$, can in fact be used for any of the parametric models. The semi parametric model of Cox (1972) can also be used by the above definitions and interpretations. As an example we can allow different hazard rates for the different strata, say young patients and older patients. This topic in general is also related to the time dependency of the covariates and it will be studied later in Chapter 7.

Apart from the proportional hazards model there is another group of regression models with a multiplicative effect on the regression parameters, namely, accelerated failure time models. The general formulation of the model is

$$\lambda(t, Z) = \lambda_0(t e^{-Z\beta}) e^{-Z\beta}$$

For this model unlike the model discussed previously, the effect of the covariates under test can have a direct effect on the base line hazard rate that is estimated. It is important to note that both the proportional hazards models and the accelerated failure time models are log-linear models with additive effects of the hazard function, the covariates and the logarithm of the time.

These models are most useful in terms of a generalised model for the estimation of the regression parameters. The method mentioned based as Turner's family of distribution is also useful in that it provides a useful way of classifying hazard rates. However, the main advantages of the proportional hazards compared to that of Turner's family or accelerated failure time is that the interpretation of events is much simpler.

3.9 Parametric, non-parametrics and Cox's approach.

In the above approaches and derivations an assumption has consistently been used in order to try to distinguish between the different survival rates. We kept the postulate that the time to a critical event is a random variable and that it can be explained by a continuous function. In the last chapter, however, the methods initially/

initially began with the reduction of the data into some form of a rank order. This reduction ultimately implies a loss of precision, in distinguishing the survival rates for the subgroup of the data. The advantage however in use of a non-parametric method based on ranks is that non-parametric tests are more robust. Extensive comparative studies of non-parametric and parametric methods have been done by various authors and we will deal with those in Chapter 5. The Cox's method which offers a practical comprcmise between parametric and non-parametric methods is also considered in Chapter 5. We will perform simulations to assess small sample properties of the Cox's method for trial data.

The analytical results of Chapter 2 and the present chapter have been based on the analysis of the old Edinburgh trial. As may be expected there are no qualitative differences in terms of the conclusions of the results. However there are slight variations by which we can reiterate the theoretical results of the earlier part of this section on hazard rates. The importance of parametric methods in here is not only that of precision alone, but rather due to an ease by which parametric methods are able to provide a conceptual frame for classifying the distributions of survival data into families of mathematical models. This flexibility to classify distribution is however compensated by a greater loss in robustness. Although all the families of distributions mentioned in this section provide flexible frameworks within which a large number of distributions for survival analysis are placed, it is difficult to imagine what may be done with an estimating procedure more complex than the Taulbee approach. In fact so far as a description of the progress of the disease matters, a plot of the/

the empirical hazard rates may suffice. If one is prepared to take the position, that hazard rates are mainly useful indescribing the biological progress of disease, then any robust general approach must lie somewhere between the non-parametric and parametric methods. The basic assumption then is that the actual rates of events are not important and need not be parameterised, but the difference between the failure rates in various groups must be estimated as precisely as possible. The final result will add to the robustness of the general method.

Cox (1972) presented a proportional hazard model by which the data is reduced to ranks and thus adopts an estimating procedure for which the rates of events are not important. The method offers a robust and flexible approach for the analysis of survival data and it is discussed in the next chapter.

CHAPTER 4

Cox's Proportional Hazards Model

For all of this chapter we will be dealing with the study of the method proposed by Cox (1972) for survival data. In this chapter we cover topics related to the usefulness of the method as applied to clinical trials. Some of the derivations from the original appraoch and the derivation of some of the central results are covered so that we may deal with the advantages and the disadvantages of the approach.

There are a few major factors that distinguish the method of the previous 2 chapters form the proportional hazards approach. The latter method is more efficient than the non-parametric methods that were discussed in the last chapter. The method in fact allows comparison of the covariate effect to be made without making unduly restrictive assumptions. In relation to the completely non-parametric methods however it is more suitable for providing a useful conceptual model for considering and testing the relationships of the effects efficiently in particular when several covariates are tested. Further it considers the relative effects of covariates as the relevant information for analysis and thus is more robust than the parametric methods where the requirement is closer approximation to the survival rates for the various groups.

There are certain requirements that must be satisfied in use

of/

of the method. One is related to the proportional hazards assumption in a non-parametric setting. The other requirement is on the type of information that is available on each case i. We must have a set of covariates $Z_i(t)$, so that predictions on the survival times of the population may be made. The time we consider, from definitions of the previous chapter can be either time to the terminating event, e.g. death, or to the follow-up event e.g. censoring. However, the functional form of $Z_i(t)$ refers to development of the covariate process in the survival time scale.

In the above discussion we mentioned the core of the topics of this section later we will consider these topics in greater detail.

4.1 Development of survival functions.

We use a similar methodology to that used for the parametric methods. $S(t)$ is the survival function; $f(t)$ is the density function and the hazard function is given by $\lambda(t)$. Such that if T is a random variable representing failure time, then for sufficiently short periods of time h, the hazard rate at time t is give by

$$\lambda(t) = \lim_{h \to o+h} \frac{1}{h} \Pr (t \leqslant T \leqslant t + h \setminus T \geqslant t) \qquad (4.1.1)$$

What the above basically implies is that the rate of failure is the conditional probability of an event at time t, given that, the individual has survival until a time immediately previous to it. For a continuous distribution we mentioned a similar definition in the introduction. However we now have a situation in which time T has a discrete distribution/

distribution and observed times have values $t_1 < t_2 < \ldots t_n$

It follows that,

$$f(t) = Pr\ (T = t)$$

$$S(t) = \sum_{j \setminus t_j < t} f(t_j)$$

$$\& \quad \lambda(t) = Pr\ (T = t \setminus T \geq t) \qquad (4.1.2)$$

The above formulation of (4.1.2) is an extension of previous definition of (4.1.1) with the difference that T is now discrete. The theoretical distinction between discrete and continuous forms of the hazard rate does not prohibit extension of the proportional hazards to a discrete analogue. For the above distribution with a covariate set Z a corresponding survival function is

$$S(t,Z) = \left[S_0\ (t)\right]^{Exp\ (Z\ \beta)}$$

Where $S_0(t)$ represents a base line survival rate at Z=0 and has a corresponding base line hazard rate given by $\lambda_0(t)$. We will return to the above formulation in section 4.4 for the construction of the likelihood.

In the context of the general proportional hazards model we can express the hazard rates as,

$$\lambda(t,\ Z) = \lambda_0\ (t)\ \ r\ (Z,\beta\ ) \qquad (4.1.3)$$

Where all the relevant information regarding the difference for survival rates is decomposed by the relative rates of failure in the $r(Z, \beta)$ function. Cox uses an exponential decomposition of the $r(Z, \beta)$ giving

$$\lambda(t,\ \ Z) \quad = \lambda_0\ (t)\ \ Exp\ (\beta\ Z) \qquad (4.1.4)$$

For the base line hazard rate of the proportional hazard model Cox uses a discrete form of $\lambda_0(t)$, based on ranks of times. The aim is that by use of this form of base line hazard rate robustness may be introduced into the model, for the estimation of what remains relevant, i.e. the relative risks. The base line hazard rate is a form of a nuisance parameter and we will deal with nuisance parameters later. As to the interpretation of the $\lambda_0(t)$ within the statistical theory, in chapter 2 we showed a maximum likelihood estimation of the Kaplan and Meier estimation and they are essentially the same. The values of the β estimators are similar in interpretation to the case of the parametric models.

The derivation of the Kaplan and Meier estimates as maximum likelihood estimators justifies the use of a discrete distribution and a parametric decomposition. In this form of the proportional hazards model, the discrete form of the base line hazard variability is removed and the data is transformed to a base line of Kaplan and Meier estimates. Cox (1972) discusses both discrete and continuous failure time data and shows a unification in the approach by which both the discrete and the continuous cases can be accommodated in essentially the same way. The term $\lambda_0(t)$ is a transformation of survival times in to the rank based product limit estimates for the hazard rates. Thus we have ranks $t_{(1)} < \ldots < t_{(k)}$

$$\hat{\lambda}(t) = \sum_{i=1}^{k} \frac{m_{(i)}}{r_{(i)}} \delta(t - t_{(i)}) = \frac{\text{no. of deaths at } t_{(i)}}{\text{no.at risk at } t_{(i)}} \quad (4.1.5)$$

&

$$\hat{S}(t) = \prod_{t_{(i)} < t} \left[1 - \hat{\lambda}(t_{(i)})\right] = \prod_{t_{(i)} < t} \left[1 - \frac{m_{(i)}}{r_{(i)}}\right]$$

The function $\delta(t - t_{(i)})$ represents a dirac delta of values 0 or 1. It is 1 in case of a failure at $t_{(i)}$ and 0 elsewhere. $m_{(i)}$ refers to number of events at rank (i) and $r_{(i)}$ referes to number at risk. This is a generalisation of the Kaplan and Meier estimates. In order to avoid problems associated with censoring times tied with failure times we adopt the convention of letting censoring occur just after the failure.

Now, regarding the relative risk part of the equation (4.1.3), we intend to categorise our population accoring to a set of measurements available on the patients. The measurements in this context are referred to as covariates. The β's are values that must be estimated and they provide information on the effects of covariates. Once again similar to the definitions of chapter 3 we refer to β's as the regression parameters.

In equation (4.1.3) we separated the effects into $r(\beta,Z)$ and a time dependent function $\lambda_0(t)$. Depending on the form of the covariate effects it is possible that the explanatory variable Z_i, be also a function of time. That is the contribution of the covariate is allowed to be a random variable, that changes with time, so a formulation such as $Z_1(t), \ldots Z_s(t)$ may be more appropriate. If our population consists of n patients and s covariate measurements, then a s x n matrix set as follows can define all auxilliary information. (Information apart from death not censorings).

$$\begin{bmatrix} Z_{11}(t) & Z_{21}(t) \ldots\ldots & Z_{n1}(t) \\ Z_{12}(t) & Z_{22}(t) & Z_{n2}(t) \\ \vdots & \vdots & \vdots \\ \cdot & \vdots & \vdots \\ \cdot & & \\ Z_{1s}(t) & \ldots\ldots\ldots\ldots & Z_{ns}(t) \end{bmatrix} \qquad (4.1.6)$$

Thus a general form of the Cox (1972) proportional hazards model with regression parameters is given by

$$\lambda(t, Z(t)) = \lambda_0(t) \text{ Exp } (\beta \hat{Z}(t)) \qquad (4.1.7)$$

Before proceeding with the discussions of the various assumptions necessary for the estimation of the regression parameters, β, we discuss the role of $\lambda_0(t)$ in the framework of the model.

4.2 Role of the Nuisance functions and the relative risks.

Meaningful isolation of relevant information is the major intention in much of statistical work. This intention can be achieved at times only by estimation of parameters that specify a distribution. In some complex processes we require a reduction of the data in a more elaborate manner.

The figures (4.2.1) and (4.2.2) present the survival rates and the disease free interval for a group of (337) patients who were entered into a randomised adjuvent chemotherapy trial, in the South East of Scotland for four years from 1.4.74. Our purpose in presenting these results is to consider the relevance of the proportional hazards to such studies. A comparison of the rates of failure by survival A and B is sufficient in giving relative rates of failure. However a more robust and thus a less restrictive estimating procedure may be achieved by realising that the relvant information in terms of the difference between survival rates of A and B in either figure is in fact contained within the shaded region C. Thus the information relating to shape of A or B at times need not play a significant role in the interpretation of the data.

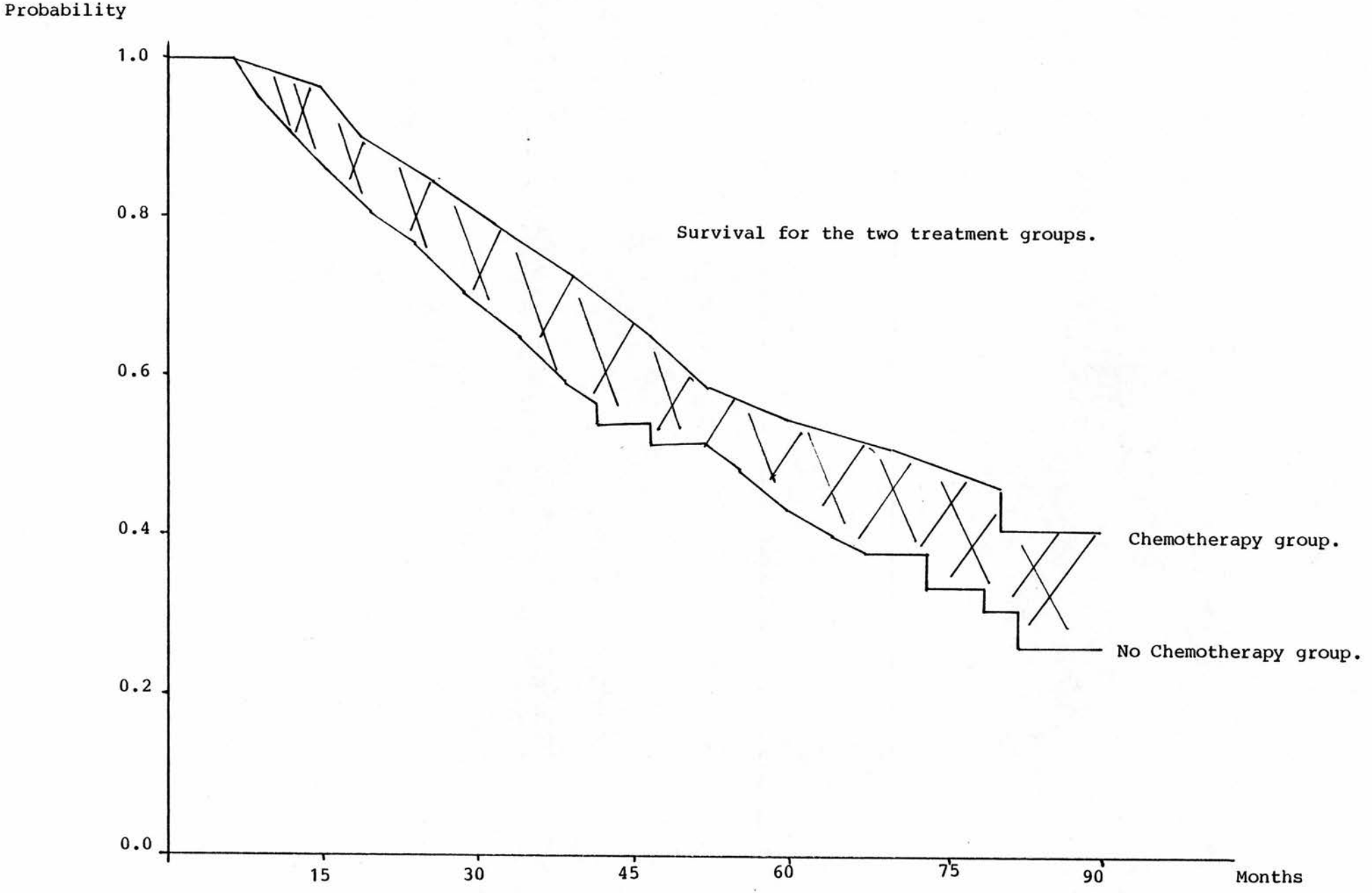

Probability
1.0
0.8
0.6
0.4
0.2
0.0
Survival for the two treatment groups.
Chemotherapy group.
No Chemotherapy group.
15
30
45
60
75
90
Months

Figure (4.2.1)

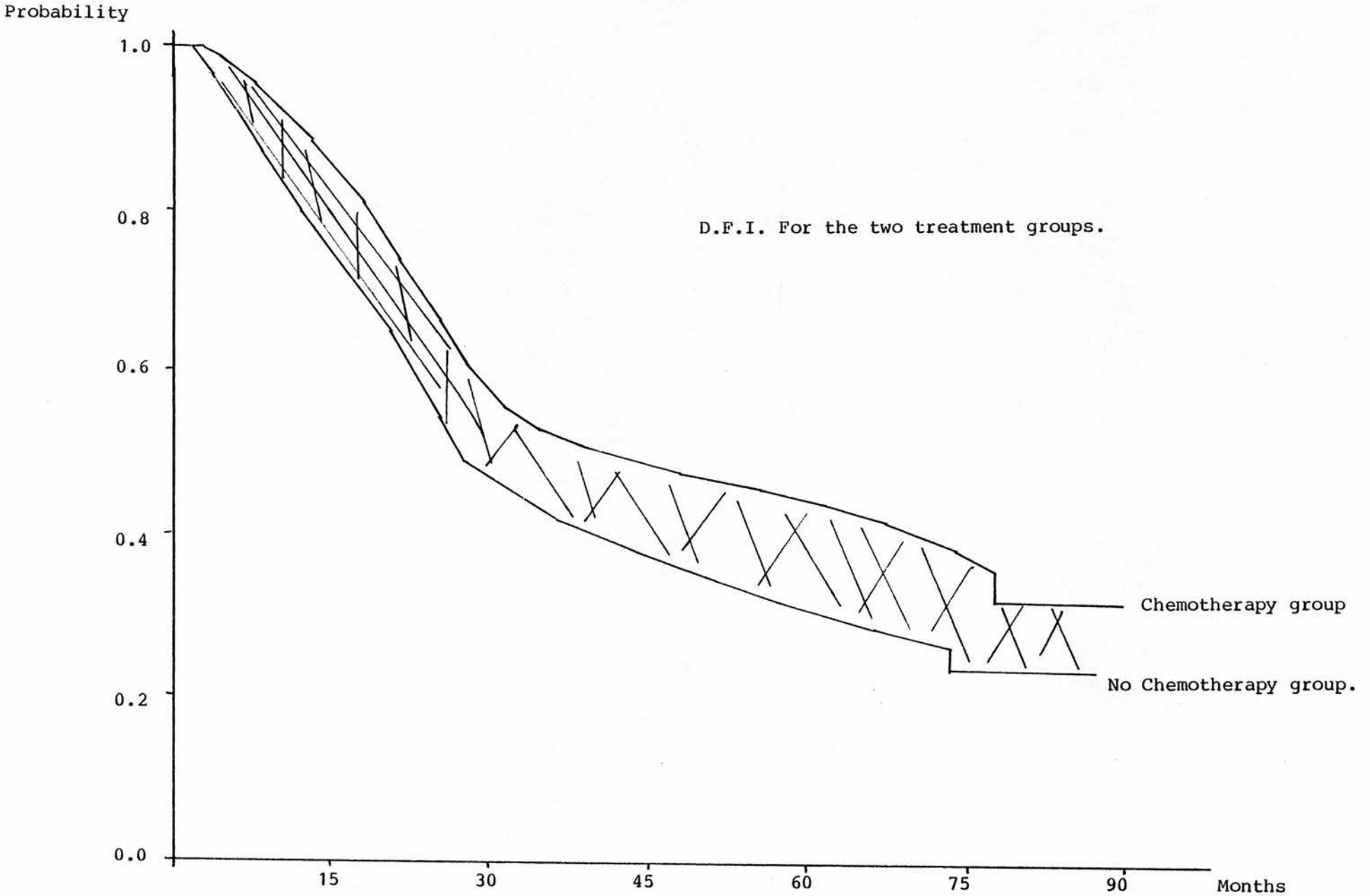

Probability
1.0
0.8
0.6
0.4
0.2
0.0
15
30
45
60
75
90
Months
D.F.I. For the two treatment groups.
Chemotherapy group
No Chemotherapy group.

Figure (4.2.2)

Parameters relating to actual shapes of curve A & B are referred to as nuisance parameters. In model (4.1.3) , $\lambda_0(t)$ is a function that must be estimated, in relation to effects of covariates Z_i and its estimator is a nuisance parameter. In fact a set of sufficient statistics for the estimation of parameters generating the shaded region is composed of the β parameters of the Z covariates. Thus in the case of a clinical trial we perform a trial L, and obtain a data set (L, Z) , where for each element of Z a measurement has been made to assess its value for aparticular patient i.

In the most elementary form of applying the probability theory, we have 3 general abstractions. A sample space X which is the set that conclusions refer to, a subset of X which is the total data set, a reduction of X, given by the model M, and a further abstraction P, which is a probability measure on model M and represents a form of variability of the data from the model.

In the context of models of survival time P is in fact composed of a subset P_Z for each particular covariate and a P_H for the hazard rate. In an ideal situation we would like to have a one to one mapping of $Z \rightarrow P_Z$ for each covariate. In the case of a trial with the model (4.1.3) form, the above restriction would rquire specifying the distribution of the survival function and the covariates for each form of risk that depends on the covariate set. However for reasons of generality and robustness a reduction is made. If in a trial the actual form of the hazard rate for each competing risk is not of interest, then the two subsets of P, namely P_Z and P_H can be defined as follows. P_Z

relates/

relates to a probability measure in terms of relative risks of covariates and P_H to probability measures for the hazard rates. The nuisance function in proportional hazards models is related to P_H.

There are certain points that must be considered in relation to the nuisance parameters. The actual trial and the way it is planned plays arole for maximising the support we obtain from the data. Since the relevant information is related to the covariates rather than time, we can maximise this form of information by the usual procedure of randomisation and possible stratification of prognostic indicators and treatments. The maximisation of support in no way needs to be related to a time factor. The model achieves its robustness by transforming the time scale into a rank order, and thus the new scale is sufficient to measure the amount of support the data gives to various values of β. The maximum likelihood approach provides a setting for optimising these values and hence obtain the various required estimates. This data reduction in Cox's approach as described so far requires the proportional hazards assumption. That is we expect the hazards for the subgroups to be multiples of a base line hazard. In the above paragraphs we discussed the issues related to nuisance parameters and some of the necessary assumptions that are related to it. The expansion of the above can include time dependency of hazards, multiple competing risks, censoring and stratification of the data, when this type of model is used for analysis. Later in this chapter and in Chapter 7 we will return to these points.

Continuing for the moment with the proportional hazard situation,

situation, the relative risk part of the model provides the necessary framework for extraction of relevant information. Here β's must be estimated and they give a representation of the dependence of the distribution of survival time T on the subgroups. The covariate set Z provides the necessary information on treatments, categories of prognostic indicators or some other measurements that are considered to be relevant at the beginning of the study. β is a 1 x S vector and it follows that one element of β has to be estimated for each Z ($j = 1, \ldots S$). The actual functional form of $r(Z, \beta)$ is oftentaken to be of the type $\text{Exp}(\beta Z)$. Cox adapts the above exponential decomposition but also allows Z's to be time dependent of the form $\text{Exp}(\beta, Z(t))$. By allowing the time dependent form of Z(t) to operate we are in fact allowing the data to generate a model with non-proportional hazard assumptions. This inclusion of time dependent covariates allows us to assess and test if the effect of certain prognostic indicators diminishes over time. It is difficult to separate topics such as time dependency, censoring with dependent effects, and competing risks when some of the withdrawals are in fact events due to other causes.

Thomas (1980) concentrates on the functional form of $r(Z, \beta)$ and produces a set of relative risk functions e.g.

$1 + (Z \times \beta)$, $1 + (\text{Exp}(\beta) \times Z)$, etc.

Gore (1981) considers an exponential decomposition form of

$$\text{Exp}(\beta_1 e^{-p_1 t} Z_1 + \beta_2 e^{-p_2 t} Z_2)$$

Kalbfleisch and Prentice (1972) suggest time dependent covariates such as

$$\text{Exp}(\beta_1 Z_1 + \beta_2 t Z_1)$$

Effron (1977) shows that $r(Z, \beta)$ can in fact be any positive function and uses a logistic dependence form ratio $\log(1 + \exp(\beta Z))$

4.3 Limitations and assumptions of the model.

In the formulation of likelihood functions we have used 2 functions and considered these to contain relevant information. They are namely $\lambda(t)$ and $r(Z, \beta)$. A fuller likelihood may contain an extra function given by

$$\text{likelihood} = \lambda(t) \cdot r(Z, \beta) \cdot \gamma(\beta, Z, t) \qquad (4.3.1)$$

The form of the hypotheseis of Gore (1981) can also be tested by an expression of the form

$$\lambda(t) \cdot \text{Exp}(Z_1 \beta_1 + Z_2 \beta_2) \cdot \text{Exp}(-Z_1 \beta_1 P_1 t - Z_2 \beta_2 P_2 t)$$

given that $e^{-P_i t} \doteq 1 - P_i t + (\frac{P_i t}{2})^2 + \ldots + (\frac{-P_i t}{n!})^n$

and that powers of greater than or equal to two are of negligible effects.

A slightly different analysis may use a model of form

$$\text{Exp}(Z \beta_1 + Z \beta_2 t^* + z \beta_3 t^{*2})$$

where $t^* = (t - \bar{t}) / \sigma_t$

It is interesting to note what kind of effect is produced by weighting the relation between Z and t differently in a family of transformations such as $y^{(\alpha)} = y^{1/\alpha}$ and $\ln(y)$ where y is a function of t.

For example a substitution for $\gamma(\beta, Z, t)$ in (4.3.1) may give

$\beta Z \ln(t)$ or $\beta Z (t)^{1/\alpha}$

We/

We will now examine the above time dependency concepts for the family of the Weiball distribution as described in Chapter 3. The survival rates for a two group sample can be expressed as

$$S_i = \text{Exp}\left[-\int_0^t \alpha_i t^{\gamma_i} e^{\beta_i} dt\right] = \text{Exp}\left[\frac{-\alpha_i t^{\gamma_i + 1}}{\gamma_i + 1} e^{\beta_i}\right]$$

The above is a proportional hazard expansion of Weiball distribution however for a time dependency effect we will allow α_i, γ_i as well as β_i to depend on group membership.

$$\ln\left(S_i\right) = \frac{-\alpha_i t^{\gamma_i + 1}}{\gamma_i + 1} e^{\beta_i}$$

For an expression of the Lehman alternatives we have

$$\frac{\ln\left(S_1\right)}{\ln\left(S_2\right)} = C\, t^{\gamma_1 - \gamma_2} e^{(\beta_1 - \beta_2)},$$

$$\text{for } C = \frac{\alpha_1(\gamma_2 + 1)}{\alpha_2(\gamma_1 + 1)}$$

$$= C\, e^{(\beta_1 - \beta_2 + (\gamma_1 - \gamma_2)\ln(t))}$$

The relative risk expression may thus be expressed as

$$\text{Exp.}\left(\gamma^*_t \ln(t) Z_1 + \beta^*_1 Z_1\right)$$

where γ^*_t and β^*_1 are parameters that must be estimated and Z_1 is the indicator of the subgroups. The ln(t) transformation will thus prvide a natural scaling for the range of the Weiball distribution.

The $\gamma(\beta, Z, t)$ part of (4.3.1) in fact contains no relevant information if we deal with a proportional hazard situation in which cases are/

are monitored continuously and censoring is non-informative. More important however is a situation where there is a need to test the effect of a covariate according to the time scale. An example is a test for assessing the persistancce of a prognostic indicator as an indicator of short survival. In practice measurements on patients are done in the beginning of the study and thus time dependency of the covariates may be assessed by the above function. We list some of the theoretical problems that may give us unwanted assumptions. The list is not composed of a set of mutually exclusive topics and the severity of the assumptions is not often significant in trials.

(1) There exists a minimum observation time and depending on this minimum observation time some information regarding censorings may be lost. Also in some studies the minimum practical observation time may not correspond to the minimum observation time at the analysis. For example, recording of death may be correct to day of death, but analysis is performed in weeks of survival.

(2) One further implication of the existence of a minimum observation time in (1) is that the data is discrete and some ties may be present in the data.

(3) Time between ranks are assumed to be non informative. (Kalbfleisch (1980) considers a Baysian approach with Gamma prior distributions between ranks.

(4) Censoring times may be informative with respect to certain covariates. A related situation is where a second cause of the event is recorded. The problem is that an auxillary cause of death, if included among censorings may cause censoring patterns to be informative, in the sense that by excluding or including a particular cause of death from, or into the event set, inconsistant conclusions may be possible.

(5)/

(5) Effect of treatment or covariates may not be consistent in time. The $r(Z, \beta)$ relative risk assumes that a single function independent of time is sufficient. The topics of (4) & (5) refer to trials in which the time variable interacts with covariate effects. In such cases although we can assess the effect by the $r(Z, \beta(t))$ part of the model, the conclusions are limited unless we move to a multivariate competing risk model.

We illustrate the above points by the following example. The figure (4.3.1) represents possible outcomes that may be recorded for a case. C refers to censoring, D to an event of interest say death and O to some other event, or auxillary event.

For a situation that all assumptions hold we require the concentration of events to have the patterns of (4.3.2a), (4.3.2b), (4.3.2c) figures for the censoring times, death times and the other auxillary event respectively. The shaded regions refer to areas with higher concentration of events. The censoring and auxillary event in (4.3.2a) and (4.3.2c) are uniformly and equally distributed and so do not provide useful information. If on the other hand the auxillary event was somehow related to lost to follow-up because of the effects of treatments, or the event of interest is metastatic disease and auxillary case is death with no previous metastatic sign, then instead of (4.3.2c) we may obtain distributions such as (4.3.2d), (4.3.2e) and (4.3.2f) for the auxillary events, where the events for treatments A and B are not uniformly distributed. We will return to this topic after more development of the mathematics. Further in Chapter 5 we will study the implication of some of the above assumptions in small samples/

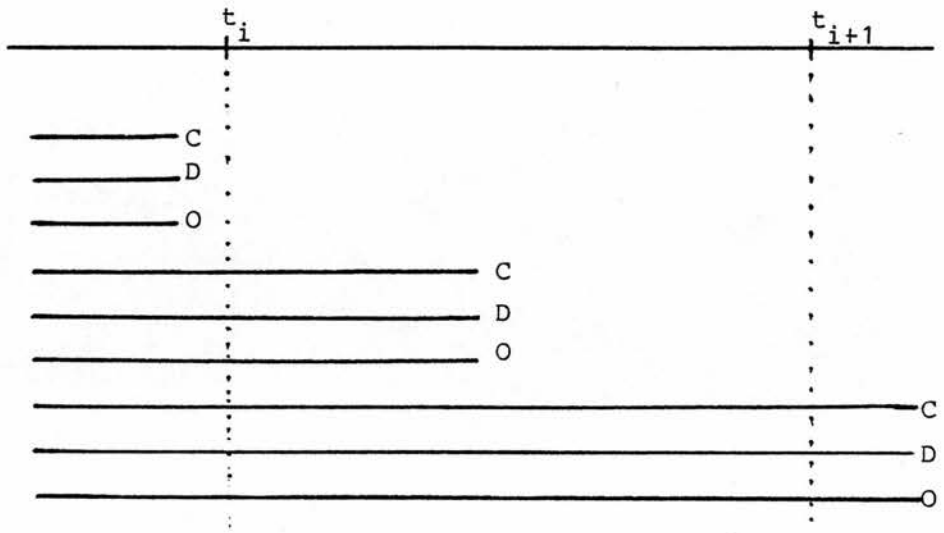


Figure (4.3.1)

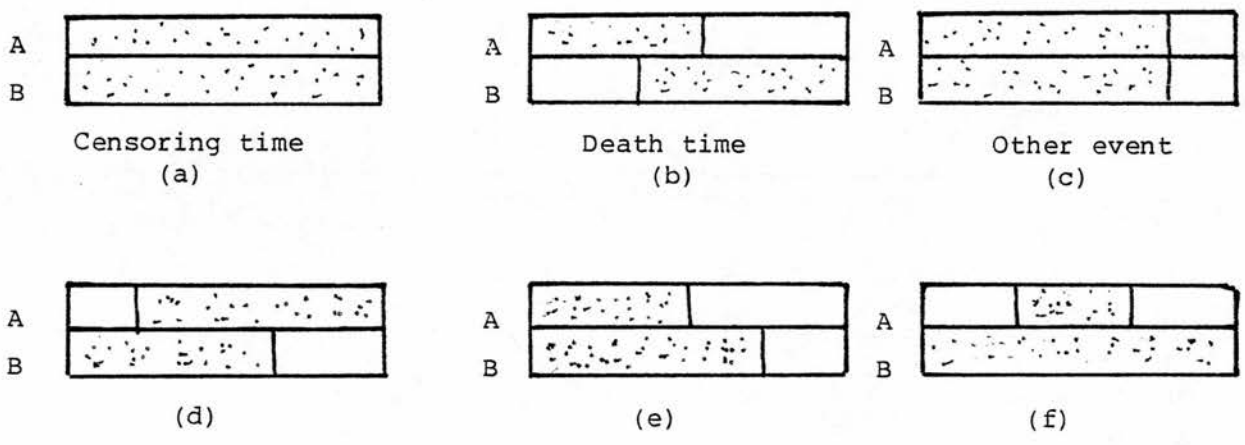


Figure (4.3.2)

samples for a realistic simulation of clinical trial data.

4.4 The construction of the likelihood and its properties.

Now we consider the methods for the estimation of the regression parameter and the construction of the relevant likelihood equations. Censoring is dealt with in the manner of Chapters 2 & 3. We observe minimum of either T_i, the failure time, or C_i the censoring time. The above statements can be expressed as

$(T_i \leqslant C_i) \Longrightarrow$ Failure

$(T_i > C_i) \Longrightarrow$ Censoring time precedes failure time.

According to definition of proportional hazards we have

$$S_o(t) = \text{Exp}\{ -\int_o^t \lambda_o(u)\, du \} \qquad (4.4.1)$$

and $$S_Z(t) = \text{Exp}\{ -\int_o^t \text{Exp}(\beta Z)\, \lambda_o(u)\, du \}$$

giving $$S_Z(t) = [S_o(t)]^{\text{Exp}(\beta Z)} \qquad (4.4.2)$$

which is an example of a Lehman Alternative, by which a reduction of relevant information may be made by the ratios of the two survival distributions. A simple example is to take a single covariate case with treatment covariate Z, set to 1 for new treatment and Z set to 0 for controls. Hence

$$S_1(t) = [S_o(t)]^{\text{Exp}\beta} \qquad (4.4.3)$$

Thus the two survival distributions are related in a multiplicative manner. The relation between a function of the ratios of $S_1(t)$

&/

& $S_o(t)$ is equivalent to a constant transformation of β and does not involve the time factor. In other words $S_o(t)$ can be projected onto $S_1(t)$ and by a function of β . The general aim of the derivations of this section is to estimate the values of β's with nonparametric hazard of the form $\lambda_o(t)$. Although the method considers $\lambda_o(t)$ and functions of β the method can also be used to generate a transformation of $\lambda_o(t)$ to estimate the survival functions.

In the original approach Cox, adopts a conditional argument to construct the likelihood. At any moment in time, there exists a particular risk set $R(t_{(i)})$. Any failure at the unique i'th in time namely $t_{(i)}$ must have arisen from this set. Therefore, probability of failure at $t_{(i)}$ given the risk set $R(t_{(i)})$ (or given survival up until $t_{(i)}$) is

$$L_i = \text{Exp}(\beta Z_i) \Big/ \sum_{l \in R(t_{(i)})} \text{Exp}(\beta Z_l) \qquad (4.4.4)$$

For the population of size n we have the likelihood function to be composed of the following function and allow censorings to occur without contributing to the likelihood.

$$L = \prod_{i=1}^{n} [L_i]^{di} = \prod_{i=1}^{n} \left[\text{Exp}(\beta Z_i) \Big/ \sum_{l \in R(t_{(i)})} \text{Exp}(\beta Z_l) \right]^{d_i} \qquad (4.4.5)$$

for $d_i = 0$ censored

$d_i = 1$ Death

(4.4.5) refers to a situation where no ties are present in the data.

Later/

Later we will consider a more general risk set for dealing with more than one event at a given time.

Now for a general likelihood (4.4.4) & (4.4.5) with the proportional hazard assumption we generate a population of size 5 and allocate the likelihood factor contributions, as in table (4.4.1)

Rank	Survival time	Censoring = 0 Death = 1	Z_i	Contribution to Likelihood $[L_i]^{d_i}$
1	1	1	0	$1/3 + 2\exp(\beta)$
2	5	1	1	$\exp(\beta)/2 + 2\exp(\beta)$
3	10	0	1	$[\exp(\beta)/2+\exp(\beta)]^0$
4	20	1	0	1/2
5	30	1	0	1/1

Table (4.4.1)

On taking logarithms of the Cox's conditional log likelihood function of survival time we have

$$\ln L = \sum_{\substack{\text{Deaths} \\ i}} \beta Z_i - \sum_{\substack{\text{Death} \\ i}} \log \{ \sum_{\substack{\text{Risk set} \\ \text{at time } t_{(i)}}} \exp(\beta Z_l) \} \quad (4.4.6)$$

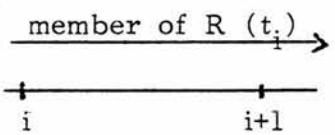
The argument is straight forward with conditionality and a single factor as in equation (4.4.4) for the case i.

However Peto (1972b) raised some related points regarding treatment of ties and censorings. Later Kalbfleish and Prentice (1973) chalenged this terminology since the conditionality for a single case does not carry over to the full set population. As is clear from Table (4.4.1) , the 3rd survival time is equal to 10 and is a censored time. This implies that there is no contribution for this case to the likelihood function. The Cox's conditionality argument makes an extrapolation on the state of failure of the remaining set. This extrapolation is related to the/

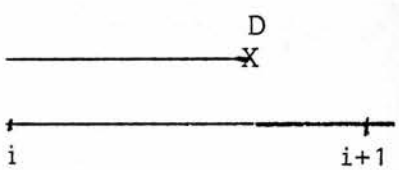
the pattern of ties in the presence of possible censoring and assumes that the distance between the events does not contain relevant information. The extent of the assumptions can be judged by considering the types of events that can occur.

In reducing the time of events into ranks we produce 3 types of observable events. Within any ranking point say ith to ith + 1, for the risk set at $t_{(i)}$ the information on time may contain any of the following 3 groups.

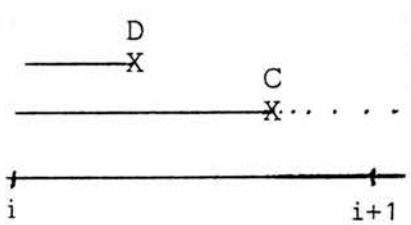
First class are those present at the beginning and the end of the time period. The consequence to the likelihood is that full information is contained within the risk set



Next class are those that die within the period. The consequence to the likelihood is that maximum information is contained if deaths occur just prior to i+1. Clearly a death can occur anywhere within the minimum observable time. In here we have considered death to be the event of interest.



Finally the group that are not present during all or part of the period. That is cases with death at $t_{(i)}$ and cases with censoring at $t_{(i)}$ to $t_{(i+1)}$. In this situation the consequence to the likelihood is made most realistic by ranking censorings after deaths.



Next we will show how the general likelihood is a deviation from the full likelihood. Later we will mention situations where the two likelihoods give close estimators. We begin with an explanatory definition for (4.4.4)

$$L_i = \frac{\exp(\beta Z_{(i)})}{\sum_{l \in R_{(t_i)}} \exp(\beta Z_l)} \cdot = \Pr(\text{Individual (i) fails} \setminus \text{risk set at } t_{(i)} \,\&, t_{(i)} \text{ has at least one death})$$

$$= \Pr(\text{Death at } t_{(i)} \setminus \text{all previous censoring and present failure information at } t_{(i)})$$

The last expression is in fact the set of sufficient information necessary to obtain column 5 of table (4.4.1).

Any L_i for a failure at $t_{(i)}$ is more generally conditional on "past history", which was expressed as risk set at $t_{(i)}$, and the fact that the event is a failure. Thus sufficient information for "past history" is, 1 to (i-1), censoring and failure information + the ith censoring information. In other words the probability regarding death at $t_{(i)}$ is made conditional upon the information regrading the occurrence of all previous deaths and censorings, and also including the information regarding censoring at current time. The probability regarding a censoring at $t_{(i)}$ is conditional upon the information regarding the occurrence of the previous deaths and censorings. This is expressed as

L_i = Pr (Death at $t_i \setminus$ 1to (i-1) censoring & death & ith censoring)

& L_i = Pr (Censored at $t_i \setminus$ 1 to (i-1) censoring & deaths)

The/

The full likelihood is a combination of the above and using the previous definitions, by the general arguments of the survival analysis we take the full likelihood to be

$$L = \prod_{\text{Deaths}} (\text{Hazard}) \cdot \prod_{\text{All}} (\text{Survivals}) \qquad (4.4.7)$$

Let $D = [\, i : t_i \leqslant c_i \,]$ be no. of cases prior to the last death

$$L = \prod_{i=1}^{D} \lambda_0(t_i)\, e^{\beta Z_i(t_i)} \prod_{i=1}^{n} \exp\left[-\int_0^{t_i} \lambda_0(u) e^{\beta Z_i(u)} du \right] \qquad (4.4.8)$$

The above is from (4.4.7)and (4.4.1)

In a population of size 3 with the first two cases failing at t_1 & t_2 and the third case censored at c_3 we obtain, on expansion of (4.4.8) such that for i=1 we have

$$L = \lambda_0(t_1)\, e^{\beta Z_1(t)} \exp\left[-\int_0^{t_1} \lambda_0(u)\, e^{\beta Z_1(t)} du \right]$$

for i =1, 2 we have

$$L = \lambda_0(t_1)\, e^{\beta Z_1(t)}\, \lambda_0(t_2)\, e^{\beta Z_2(t)} \exp\left[-\int_0^{t_1} \lambda_0(u) e^{\beta Z_1(t)} du \right]$$

$$\cdot \exp\left[-\int_0^{t_1} \lambda_0(u) e^{\beta Z_2(t)} du \right] \exp\left[-\int_{t_1}^{t_2} \lambda_0(u) e^{\beta Z_2(t)} du \right]$$

for i=1,2,3 we have

$$L = \lambda_0(t_1)\, e^{\beta Z_1(t)}\, \lambda_0(t_2) e^{\beta Z_2(t)} \exp\left[-\int_0^{t_1} \lambda_0(u)\, e^{\beta Z_1(t)} du \right]$$

$$\cdot \exp\left[-\int_0^{t_1} \lambda_0(u) e^{\beta Z_2(t)} du \right]$$

$$\cdot \exp\left[-\int_{t_1}^{t_2} \lambda_0(u) e^{\beta Z_2(t)} du\right] \exp\left[-\int_{0}^{t_1} \lambda_0(u) e^{\beta Z_3(t)} du\right]$$

$$\cdot \exp\left[-\int_{t_1}^{t_2} \lambda_0(u) e^{\beta Z_3(t)} du\right] \exp\left[-\int_{t_2}^{c_3} \lambda_0(u) e^{\beta Z_3(t)} du\right]$$

By rearanging the exponential term for each integral period $\int_{t_{i-1}}^{t_i}$ we get

$$\prod_{i=1}^{D} \exp\left[-\int_{t_{i-1}}^{t_i} \lambda_0(u) \sum_{j \in R} e^{\beta Z_j(u)} du\right] \quad (4.4.9)$$

The rest of (4.4.8) are the contributions of the deaths, and is the first part of the equation, given by

$$\prod_{i=1}^{D} \lambda_0(t_i) \exp\left[\beta Z_{ji}(T_i)\right] \quad (4.4.10)$$

Note (4.4.8) is (4.4.9) & (4.4.10)

By combining (4.4.9) & (4.4.10) we obtain

$$L = \prod_{i=1}^{D} \left(\left[\exp\left[-\int_{t_{i-1}}^{t_i} \lambda_0(u) \sum_{j \in R} e^{\beta Z_j(u)} du\right] \lambda_0(t_i) \sum_{j \in R} e^{\beta Z_j(t_i)}\right] \cdot \frac{\exp(\beta Z_{ji}(T_i))}{\sum_{j \in R} e^{\beta Z_j(t_i)}} \right) \quad (4.4.11)$$

Note that the above is equivalent to (4.4.8) with the term $\sum_{j \in R} e^{\beta Z_j(t_i)}$ introduced to the equation. The part (2) of the equation is clearly the/

the usual partial likelihood and part (1) gives the extra contributions for the full likelihood. If $\lambda_0(t)$ is unknown then part (1) provides little information about β's. Thus (4.4.11) reduces to the likelihood (4.4.5).

Now we will return to the generalisation of the closing parts of section 4 .3, using the above mathematical notations. The parts (1) and (2) of the full likelihood (4.4.11) have two time quantities, $\lambda_0(u)$ and the other $Z(u)$. We can often assume that part (2) of the likelihood does not contribute to the information on covariates. This is a true assumption by an independency of $Z(u)$ from $\lambda_0(u)$ in the integral of t_{i-1} to t_i.

The problem of tied observation was treated by Cox (1972) in the following manner.

Say two observations are tied a & b. Due to the fact that we do not know the order of these events, the actual probability contributed to the likelihood is

$$\left[\frac{\exp(\beta Z_a)}{\sum_{j \in R} \exp(\beta Z_j)} \cdot \frac{\exp(\beta Z_b)}{\sum_{\substack{j \in R \\ a \notin R}} \exp(\beta Z_j)} \right] + \left[\frac{\exp(\beta Z_b)}{\sum_{j \in R} \exp(\beta Z_j)} \cdot \frac{\exp(\beta Z_a)}{\sum_{\substack{j \in R \\ b \notin R}} \exp(\beta Z_j)} \right]$$

Cox's approximation is

$$\prod_{i=a,b} 2\exp(\beta Z_i) \Big/ \prod_{i=a,b} \left(\sum \exp(\beta Z_j) \right) - \exp(\beta Z_i)$$

Peto (1972b) suggested an approximation to the tied ranks distribution and/

and later Kalbfleish (1972) referred to this likelihood as marginal likelihood. He pointed out that this assumption prohibits use of time dependent covariates. The expression for the above is

$$\prod_{i=a,b} \exp(\beta Z_i) \Big/ \begin{bmatrix} r \\ m \end{bmatrix} \Big(\sum_{j \in R} \exp(\beta Z_j) / r \Big)^m$$

where r is no. at risk at time ties have occurred and m is no. of events tied.

By use of the approach proposed by Cox in dealing with ties the calculations become exceedingly cumbersome. The ratio of calculations in fact multiply as the number of ties increase in the sample. However in our study this method is used mainly because the use of the alternative approach implies the prohibition of the use of time dependent covariates. The partial likelihood of (4.4.11) can be expressed as a function of the log likelihood of k distinct deaths as

$$\ln L = L(\beta) = \sum_{i=1}^{k} [\beta' Z_i - \ln (\sum_{j \in R_i} \exp(\beta' Z_j))] \quad (4.4.12)$$

For use of the maximisation method of Newton-Raphson we first require two derivatives of the likelihood with respect to β. Different derivations of the likelihood may introduce different restrictions on the form of $Z_i(t)$. However the following can be obtained without loss of generality.

$$\frac{\delta L(\beta)}{\delta \beta_p} = \sum_{i=1}^{k} Z_{ip} - \frac{\sum_{j \in R_i} \exp(\beta' Z_j) \cdot Z_{jp}}{\sum_{j \in R_j} \exp(\beta' Z_j)} \quad (4.4.13)$$

Thus/

Thus we want solutions to

$$\sum_{i=1}^{k} \left(Z_{ip} - \frac{\sum_{j \in R_i} \left[Z_{jp} \exp(\beta' Z_j) \right]}{\sum_{j \in R_i} \exp(\beta' Z_j)} \right) = 0$$

The equation can be solved by Newton Raphson procedure and use of the following 2nd partial derivatives.

$$\frac{\delta^2 L(\beta)}{\delta\beta_p \delta\beta_q} = - \sum_{i=1}^{k} \left(\frac{\left(\sum_{j \in R_i} \left[\exp(\beta' Z_j) \right] \right) \left(\sum_{j \in R_i} \left[Z_{jp} Z_{jq} \exp(\beta' Z_j) \right] \right)}{\left(\sum_{j \in R_i} \exp(\beta' Z_j) \right)^2} - \frac{\left(\sum_{j \in R_i} Z_{jp} \exp(\beta' Z_j) \right) \left(\sum_{j \in R_i} Z_{jq} \exp(\beta' Z_j) \right)}{\left[\sum_{j \in R_i} \exp(\beta' Z_j) \right]^2} \right) \quad (4.4.14)$$

We can thus estimate the β values in (4.4.12) and thereby assess the effects of the Z concomitant variables. Using the above derivations we will now proceed with a few commonly used testing procedures.

For testing the global null hypothesis that all coefficients are identically zero Cox gives the the efficient score statistics of Rao based on

$$Q = U'(0) \; I^{-1}(0) \; U(0)$$

where U is a vector of all first derivatives given by (4.4.13) . I is the information matrix and is composed of elments given by second derivatives as in(4.4.14). Q has a chi-squared distribution with r concomitant variables and v degrees of freedom. In studies where

r/

r is large like the situation of most clinical trials, Cox suggests the use of significance tests for subset of parameter estimates. The two commonly used tests are the asymptotic likelihood and the asymptotic normality tests.

For the likelihood ratio test with one degree of freedom we have,

$$-2 \left(L \left(\hat{\beta}_a \right) - L \left(\hat{\beta}_b \right) \right)$$

where $\hat{\beta}_a$ and $\hat{\beta}_b$ are vectors of parameter estimates that are included in the likelihood model. $\hat{\beta}_a$ in fact spans a space which is a subset of $\hat{\beta}_b$ and the two often have a dimensional difference of one. The test then has a chi-squared distribution with one degree of freedom under the null hypothesis that the concomitant information missing in the likelihood of $\hat{\beta}_a$ has an estimator zero in $\hat{\beta}_b$.

With the assumption of asymptotic normality for a one-sided $\alpha\%$ significance level we have

$$\Pr \left(\hat{\beta}_p / \left[I^{-1} \left(\hat{\beta}_p \right) \right]^{\frac{1}{2}} > t_\alpha \right) = \alpha$$

where t_α refers to the percentage point of the t distribution with α significance level. The above tests are used extensively in the simulation studies of the Chapter 5.

In the final part of this section we will show that the first and the second derivatives, namely (4.4.13) and (4.4.14) are in fact translation invarient. This result is of interest when we consider transformation/

transformation of the covariates for the fast convergence of the iterative methods.

The values of β are translation invariants under a translation of Z_j to $(Z_j + a)$ where a is any vector of constants. At this stage we substitute values of $(Z_j + a)$ with Z_j in the equation (4.4.13) & (4.4.14) and show that the ratios remain the same. We have

$$\sum_{i=1}^{k} \left([Z_{ip} + a_p] - \frac{\sum_{j \in R_i} \left[(Z_{jp} + a_p) \exp (\beta' [Z_j + a]) \right]}{\sum_{j \in R_i} \exp (\beta' [Z_j + a])} \right)$$

$$= ka_p + \sum_{i=1}^{k} [Z_{ip} - \frac{\exp(\beta' a) \left(\sum_{j \in R_i} Z_{jP} \exp(\beta' Z) + a_p \sum_{j \in R_i} \exp(\beta' Z) \right)}{\exp(\beta' a) \sum_{j \in R_i} \exp(\beta' Z_j)}]$$

$$= ka_p + \sum_{i=1}^{k} [Z_{ip} - \frac{\sum_{j \in R_i} Z_{jp} \exp(\beta' Z_j)}{\sum_{j \in R_i} \exp (\beta' Z_j)} - \frac{a_p \left[\sum_{j \in R_i} (\exp(\beta' Z_j)) \right]}{\sum_{j \in R_i} \exp (\beta' Z_j)}]$$

$$= ka_p + \sum_{i=1}^{k} \left(Z_{ip} - \frac{\sum_{j \in R_i} Z_{jp} \exp(\beta' Z_j)}{\sum_{j \in R_i} \text{Exp} (\beta' Z_j)} \right) - ka_p$$

$$= \sum_{i=1}^{k} \left(Z_{ip} - \frac{\sum_{j \in R_i} Z_{jp} \exp (\beta' Z_j)}{\sum_{j \in R_i} \exp (\beta' Z_j)} \right)$$

and hence by letting the equations once again to be set to zero, we will have the same values for the β estimators.

For the 2nd derivatives we have -

$$\frac{\delta^2 L(\beta)}{\delta\beta_p \delta\beta_q} = -\sum_{i=1}^{k}\left[\frac{\left(\sum_{j\in R_i}\exp(\beta' Z_j)\right)\left(\sum_{j\in R_i} Z_{jp} Z_{jq}\exp(\beta' Z_j)\right)}{\left(\sum_{j\in R_i}\exp(\beta' Z_j)\right)^2} - \frac{\left(\sum_{j\in R_i} Z_{jp}\exp(\beta' Z_j)\right)\left(\sum_{j\in R_i} Z_{jq}\exp(\beta' Z_j)\right)}{\left(\sum_{j\in R_i}\exp(\beta' Z_j)\right)^2}\right] = (*)$$

again using the same substitution we get

$$= -\sum_{i=1}^{k}\left[\frac{e^{2\beta' a}\left(\sum_{j\in R_i}\exp(\beta' Z_j)\right)\left(\sum_{j\in R_i}(Z_{jp}+a_p)(Z_{jq}+a_q)\exp(\beta' Z_j)\right)}{e^{2\beta' a}\left(\sum_{j\in R_i}\exp(\beta' Z_j)\right)^2} - \frac{e^{2\beta' a}\left(\sum_{j\in R_i}(Z_{jp}+a_p)\exp(\beta' Z_j)\right)\left(\sum_{j\in R_i}(Z_{jq}+a_q)\exp(\beta' Z_j)\right)}{e^{2\beta' a}\left(\sum_{j\in R_i}\exp(\beta' Z_j)\right)^2}\right]$$

$$= \sum_{i=1}^{k} -\left[\left(\frac{a_p \sum_{j\in R_i} Z_{jq}\exp(\beta' Z_j) + a_q \sum_{j\in R_i} Z_{jq}\exp(\beta' Z) + a_p a_q \sum_{j\in R_i}\exp(\beta' Z_j)}{\sum_{j\in R_i}\exp(\beta' Z_j)}\right)\right.$$

$$-\left(a_p\left(\sum_{j\in R_i}\exp(\beta' Z_j)\right)\left(\sum_{j\in R_i} Z_{jq}\exp(\beta' Z_j)\right) + a_q\left(\sum_{j\in R_i}\exp(\beta' Z_j)\right)\right.$$

$$\left.\left.\left(\sum_{j\in R_i} Z_{jp}\exp(\beta' Z_j)\right) + a_p a_q\left(\sum_{j\in R_i}\exp(\beta' Z_j)\right)^2\right) \Big/ \left(\sum_{j\in R_i}\exp(\beta' Z_j)\right)^2\right]$$

$$+(*)$$

$$= (*)$$

The above results indicate that the a translation of Z_j to $Z_j + a$ leaves the function of the second derivatives of the log likelihood with respect to β values the same.

As was mentioned the functions of second derivatives are used in the estimation of the variance of the β estimators. It is defined to be $\text{Var}(\hat{\beta}) = \left(\frac{-\delta^2 L(\beta)}{\delta\beta_p \delta\beta_q}\right)^{-1}$ The value of the second partial derivatives are also used in the estimator procedure of Newton Raphson where a function is formed to obtain a convergence of the equation (4.4.12).

The method is iterative and it spans the likelihood surface until it finds the required maxima. The rate of convergence to the maxima depends on the slope and shape of the likelihood surface. Primarily the rate of convergence is slow if a number of covariates have a large scale range and these show a degree of correlation. The consequence is that the variance covariance matrix at inversion will have a determinant which is almost zero. The problem of scale range can be remedied by subtracting the mean value of the covariate effects from the covariate if they have a large scale range.

In Chapter 6 we will use the above and alternative methods based on the categorisation of the continuous variables.

4.5 Covariate interaction and time dependency..

In/

4.5 Covariate Interaction & Time-Dependency.

In this section we will consider the proportional hazards model in terms of possible functional forms of the relative risk part of the model. We will describe the possible functions that may be useful and efficient in an analysis of clinical trial data. We will relate these functions to appropriate hazard rate patterns and in the future chapters some of the topics and models of this section will be used in analysis and interpretation of the results. In here we will keep the definition of $\lambda_0(t)$ to be as that of previous sections of this chapter. The $r(Z, \beta, t)$ function measures the relative risk differences in relation to the base line hazard and the projected subgroup hazard rates. We reiterate the point that this difference may be due to various forms of time-dependency or purely due to fixed covariate effects. This distinction is not in most circumstances very clear especially in an exploratory analysis or a situation of measurement over time. This effect may be referred to as time confounding and is related to the influence of the various covariates on each other within the time scale. An example is a situation where treatment effect comparisons may show a different relative risk pattern for younger patients and the older patients. Such an effect is testable by a complete model of age and treatment. A different approach may attempt to test the adequacy of the functional form of $r(Z, \beta)$ by inclusion of a treatment and time dependent covariate based on the time scale itself. Much of this section is related to various developments of $r(Z, \beta, t)$ and the way it can influence, time dependency, interaction and confounding. In a situation of stratified analysis of the data with say an exponential decomposition/

decomposition of the relative risks we have

$$\lambda_k(t,Z) = \lambda_{ok}(t)\, Z_{ok}\, \mathrm{Exp}(\, Z_1 \beta_1 + Z_2 \beta_2 + \ldots\,)$$

Where Z_{ok} is set to be dummy variable and conditions the analysis on the strata of interest k.

That is

$$Z_{ok} = \begin{cases} 1 \text{ case belongs to strata of interest, } k \\ 0 \text{ case does not belong to the relevant strata } k \end{cases}$$

In effect by repeating the analysis for the various strata it will result in a different base line hazard being produced for each strata set. The significance of this point is merely attributed to the method by which stratified analysis may be incorporated into the general procedures. The resultant effect on the partial likelihood argument is the introduction of a conditionality parameter such that,

Pr (subject i failing at t_i\ presence until c_i or t_i and also membership of strata) $= Z_{ok}\, r(Z_1, \beta_1) / \sum_k Z_{ok} r(Z_1\ \beta_1)$

$$= \frac{Z_{ok} r(Z_1\ \beta_1)}{\Sigma_{k^*}\, r(Z_1 \beta_1)}$$

where k* is the new risk set and excludes all cases not belonging to the particular strata. A point that must be noted is that Z_{ok} function in the above may be interpreted as a function adjusting $\lambda_{ok}(t)$ rather than one acting on $r(Z,\beta)$.

Then/

Then

$$\lambda_{ok}(t) = \left[1 - \frac{m_{(i)}}{r_{(i)}} \right] Z_{ok}$$

An example is a situation of separate analysis for the older patients, such that

if age $>$ d then $Z_{ok} = 0$

and age $\leqslant$ d then $Z_{ok} = 1$, where d is a constant on the age scale.

Now we consider a situation where Z_1 is related to a categorised separation of a continuous variable say time or size. Suppose we set $Z_1 = (-1,0,1)$. We then test the effect of Z_1, with above categorisation assumption, that the relative risk at the lower level of $Z_1 = -1$ is related to the middle level of $Z_1 = 0$ by the same scale which relates the middle level of Z_1 to the higher level of $Z_1 = 1$. A more elaborate analysis will allow the 3 levels of Z_1 to act independently by introducing $(Z_1)^2$ such that $\text{Exp}\ (Z_1 \beta_1 + (Z_1)^2 \beta_2)$ In the case of interaction effects being present in the data between the two covariates we may have expressions of the form

$$\text{Exp}(Z_1 \beta_1 + Z_2 \beta_2 + Z_1 Z_2 \beta_{12}).$$

Under this assumption we are testing the multiplicative effect of Z_1 & Z_2 on each other and on $\lambda_0(t)$. i.e. the relative risks.

$$\text{Exp}\ (Z_1 \beta_1) \circ \text{Exp}(Z_2 \beta_2) \cdot \text{Exp}(Z_1 Z_2 \beta_{12})$$

For an actual trial we can represent the various subgroups for say/

say, treatment A and treatment B and node positive and negative groups such that hazard rates for the two treatment groups are

$$\lambda(t,Z_t) = \begin{cases} \lambda_0(t) & \text{For patients of group A treatment} \\ \lambda_0(t)\, e^{\beta t} & \text{For patients of group B treatment} \end{cases}$$

The common $\lambda_0(t)$ base line hazard is clearly a nuisance parameter. The relative riks $e^{\beta t}$ represents the effect of treatment. The greater its deviation from 1, the greater is the importance of new treatment.

For the two prognostic subgroups a similar pattern may be represented

$$\lambda(t,Z_t,Z_n) = \begin{cases} \lambda_0(t) & \text{group with treatment A, node negative} \\ \lambda_0(t)\, e^{\beta t} & \text{group with treatment B, node negative} \\ \lambda_0(t)\, e^{\beta n} & \text{group with treatment A, node positive} \\ \lambda_0(t)\, e^{\beta t\, \beta n} & \text{group with treatment B, node negative.} \end{cases}$$

The above structure however is considering β_t and β_n to be of similar effect if they are present singly or both simultaneously. There is an extension to the model by which one can test the effect of both treatment and node, while one is testing their effective simultaneous presence.

$$\lambda(t,\ Z_t,\ Z_n) = \begin{cases} \lambda_0(t) & \text{group with treatment A, node negative} \\ \lambda_0(t)\, e^{\beta t} & \text{group with treatment B, node negative} \\ \lambda_0(t)\, e^{\beta n} & \text{group with treatment A, node positive} \\ \lambda_0(t)\, e^{\beta t}\, e^{\beta n}\, e^{\beta I} & \text{group with treatment B, node negative} \end{cases}$$

Then if β_I is significantly different from 0 then there is a suggestion/

suggestion that the new treatment may be more effective for one prognostic group than the other. In the above we have dealt with binary treatment and prognostic categories. In a case of 3 categories of a prognostic indicator say size divided into 3 separate classes small, medium and large tumours, an expansion of the concept of interaction is possible. Like the example of the node we may have alinear interaction of the size with treatment. However due to the fact that there are 3 levels of size present we may have various quadratic effects acting. That is the larger tumours may be behaving in a way completely different to those of small and medium. We may then introduce two different sets of covariates $Z_s = (-1,0,1)$ and $Z_s^2 = (-1,0,1)^2$ so that a test of size effect may be done in such a way that various main effects and interaction effects of size are independent.

In a situation where time may effect the influence of certain covariates, we may represent the time interaction by

$\text{Exp}(\beta_1 Z_1 + \beta_t . Z_1 t^*)$, for t^* a function of time.

A Taylor series expansion of the time dependent effect gives.

$$(1 + \beta_t Z_1 t^*) \cdot \text{Exp}(\beta_1 Z_1) \quad \text{for } (\beta_t . Z_1 . t^*)^j \to 0$$

$$\text{as } \beta_t \to 0 \text{ for } j \geq 2$$

In the previous model the factors of order $j \geq 2$ have been considered insignificant by the Taylor series expansion. Clearly other possible situations for detecting departures of specific type from/

from the proportional hazard assumption can require a model of form

$$\lambda(t, Z) = \lambda_0(t) \left[1 + (Z_1 \beta_t t^*) + \frac{(Z_1 \beta_t t^*)^2}{2!}\right] . \text{Exp} (\beta Z)$$

When considering time dependencies the functional form of t^* is also of importance for an efficient analysis. It may be necessary to transofrm t to $\frac{t - t}{\sigma_t}$. Alternatively if the effect on covariate is influenced exponentially with time we use $(\ln(t) - \ln(\bar{t}))$. An analgous approach may use a transformation of the time scale t^* to 0 or 1 scale, so that effects of intervening events such as metastatic recurrence may be studied.

In here we must make a n important distinction between the various forms of time dependency which have been considered. It can be that a measurement over time like age is considered an independent value which affects the survival time. It may be that age is considered to have a time scale which is inappropriate under the proportional hazard assumption and therefore study of departures of particular types based on the functional form of t^* may be of interest. Finally we may be interested in the study of intervening events like the metastatic recurrences.

In the analysis of the data presented in the next chapter we will use a functional form of the $r(Z, \beta)$ referred to in the Cox's paper on the exponential decomposition of the relative risk Exp (β Z). We will return to this topic of time dependency in Chapter 7 and 8 where a more detailed study and analysis of trial data will be performed. As we mentioned in section (4.3) the/

the ln(t) is a natural transformation for testing time dependency of Weiball form, in a proportional hazards setting. In Chapter 8 we will relate these topics to concepts of change and random covariate effects. In Chapter 7 we will consider various functional forms of the time dependency such as logarithmic or linear time dependency. Further we will study effects of intervening effects by transformations of the time scale in to binary 0 or 1 scale.

CHAPTER 5

SIMULATION OF PATIENT ACCRUAL TIME TO RESPONSE. In a clincial trial with proportional and non proportional hazard rates.

In this chapter we will describe a method for generating random samples of survival times with a given distributional assumption. The distributional form of the generated sample will clearly play a major role in value of an analysis method. Further we will develop a method of producing different levels of censoring times as an analogous situtation to that of random arrival of patients into the trial, and early analysis when some patients are still alive.

It is intended that by such an approach a comparative study of the generated small samples of survival data may be made with varying values of covariates, censoring percentages, sample sizes and the hazard rate of cases.

For reasons of comparison we explain type I and type II errors in the context of the present study and finally the results are presented and discussed.

5.1 Generation of survival times.

In Chapter 3 we described some of the possible distributional forms of the survival times. We also presented some of/

of the empirical results to show that different patterns of failure rates do occur in practice. One major aim for any system of generating random samples is that the method should be flexible, so that we may produce a range of survival times with a good level of control over the many factors under study.

We will present a manageable method of simulating random survival distributions with proportional and non-proportional hazards, relevant to failure time analysis. Further for a realistic simulation of trials we will develop an approach for accrual and censoring times.

We confine the study to the most commonly used distributions, exponential and the Weiball under covariate constraints. The method of generation provides a good methodology for producing distributions both in the framework of covariates and also in terms of time-dependency. However it does not extend to censoring. Later on in this section we will describe an algorithm for censorings.

The conditional survival distribution function of Weiball survival in presence of covariates may be presented by

$$S(t, Z) = \text{Exp}\ [- (\mu t)^{\nu} e^{Z \beta}] \qquad (5.1.1)$$

Where ν is the shape parameter in the Weiball distribution. Clearly at $\nu = 1$ we have a special case of the exponential distribution. We thus have the survival time T is always greater than or equal to zero, Z is a vector of explanatory indicators, β is the vector of parameter that eventually has to be estimated and μ is a parameter for/

for "adjusting" the rate of the hazard functions. The conditional probability density functions and the conditional hazard function for T then follow from (5.1.1).

$$-\frac{\partial S(t, Z)}{\partial t} = f(t, Z) = \mu\nu(\mu t)^{\nu-1} e^{Z\beta} \text{Exp}[-(\mu t)^{\nu} e^{Z\beta}] \qquad (5.1.2)$$

and

$$\frac{f(t, Z)}{S(t,Z)} = \lambda(t, Z) = \mu\nu(\mu t)^{\nu-1} e^{Z\beta} \qquad (5.1.3)$$

In making the functions more manageable we use a two stage transformation. In its present form it is not easy to recognise a probability distribution function of the above. However after the transformations we will relate the distribution to the extreme value distribution. We let Y have a probability density function $f_Y(y)$. If h(y) is either increasing or decreasing in y, then U = h(Y) has the density function given by

$$F_U(u) = f_Y[h^{-1}(u)] \left|\frac{dy}{du}\right| \qquad (5.1.4)$$

A useful method is finding the density function of Y = log T. Therefore we use function h(t) = log t, giving $h^{-1}(y) = \text{Exp}(y)$.

$$\frac{\partial(h^{-1}(y))}{\partial y} = \text{Exp}(y) \qquad -\infty < y < \infty$$

Now substituting for t = exp (y) in f(t,Z) and multiplying by $|\text{Exp}(y)|$ we get

$$f(y,Z) = \mu\nu(\mu e^{y})^{\nu-1} e^{Z\beta} \text{Exp}[-(\mu e^{y})^{\nu} e^{Z\beta}] e^{y}$$

This gives

$$S(y,Z) = \text{Exp}[-(\mu e^{y})^{\nu} e^{Z\beta}]$$

which/

which can be derived from using $f(y,Z) = - \partial S(y,Z) / \partial y$

For manageability we use a further transformation of Y,

$$W = Y\nu - \alpha\nu + Z\beta$$

where $\alpha = - \log(\mu)$

The density function of interest is

$$h^{-1}(w) = \frac{w}{\nu} + \alpha - \frac{Z\beta}{\nu}$$

$$\frac{\partial(h^{-1}(w)}{\partial(w)} = \frac{1}{\nu}$$

Then we obtain $f(w,Z)$ by substituting $y = \frac{w}{\nu} + \alpha - \frac{Z\beta}{\nu}$

and multiplying by $\left|\frac{1}{\nu}\right|$ we have the probability density function of w and Z as ,

$$f(w, Z) = e^{-\alpha}\nu(e^{-\alpha}e^{\frac{w}{\nu} + \alpha - \frac{Z\beta}{\nu}})^{\nu-1} e^{Z\beta} . \text{Exp.}[-(e^{-\alpha}e^{\frac{w}{\nu} + \alpha - \frac{Z\beta}{\nu}})^{\nu}e^{Z\beta}]e^{\frac{w}{\nu} + \alpha - \frac{Z\beta}{\nu}} . \frac{1}{\nu}$$

$$= \nu(e^{\frac{w}{\nu} - \frac{Z\beta}{\nu}})^{\nu} e^{Z\beta} \text{Exp}[-(e^{\frac{w}{\nu} - \frac{Z\beta}{\nu}})^{\nu} e^{Z\beta}] \frac{1}{\nu}$$

$$= (e^{w - Z\beta}) e^{Z\beta} \text{Exp}[-(e^{w - Z\beta}) e^{Z\beta}]$$

$$= (e^{w}) \text{Exp}[-(e^{w})] = \text{Exp}(w - e^{w}) \qquad (5.1.5)$$

The above is an example of the extreme value distributed random variable for w, with the distribution function.

$$F(w) = \text{Exp}(-e^{-w}) \qquad -\infty < w < \infty \quad (5.1.6)$$

Now (5.1.6) is in a convenient form so that the required distributions may be generated.

In/

In cases where there are no covariates present the survival distribution reduces to

$$S(t) = \mu \, Exp[- (\mu t^{\nu})]$$

and it follows that the generator is

$$y = \frac{w}{\nu} + \alpha \qquad \text{where} \quad \mu = \exp(-\alpha)$$

When ν is equal to 1, the Weiball distribution reduces to an exponential distribution with the survival distribution given by

$$S(t) = \mu \, Exp (- \mu \, t)$$

and the generating function is $y = w + \alpha$

In here we will not discuss the actual values for ν , α, β and μ. However later we will mention the actual values that are used in the study.

The extreme value distributed random variables can be easily generated using the operation of two logarithms on a set of unifromly distributed random variables, between 0 and 1, so that

$W = \log(-\log U)$, for U, uniformly distributed between 0 and1.

One result that may be of practical importance although we do not use it further in this thesis is the pattern by which the lognormal distribution can mimic the standard extreme value distribution. This will allow a similar expansion of the methods, so that other distribution may be approximated to produce other shapes of the hazard functions, using the same procedures.

Standardised Cumulative Distribution Functions

X	Extreme Value	Lognormal
-2.0	.00063	.0002
-1.5	.02140	.0196

Standardised Cumulative Distribution Functions (cont'd)

X	Extreme Value	Lognormal
-1.0	.1321	.1324
-0.5	.3443	.3471
0.0	.5704	.5700
0.5	.7440	.7423
1.0	.8558	.8546
1.5	.9224	.9207
2.0	.9577	.9579
2.5	.9775	.9773
3.0	.9881	.9884
3.5	.9937	.9939
4.0	.9967	.9968

The above extreme value distribution is the standardised extreme value distribution with

$$S(X) = \mathrm{Exp}[- \exp(-1.28254\, x - 0.57722)]$$

The Lognormal then has the distribution

$$S(X) = (\sqrt{2\pi})^{-1} \int_{-\infty}^{u(x)} \exp(-\tfrac{1}{2} u^2) du$$

with

$$U(x) = 6.52771 \ \mathrm{Log}_{10} (X + 2.74721) - 2.68853.$$

We will now describe the present extreme value distribution more specifically. We will use the standard probability density function of X as a unimodal $\exp(w - e^w)$, with skewness -1.14, Kurtosis 2.4, variance $[\pi 2/6] = (1.282)^2$ and the mean -0.5722, which is the negative of Eulers constant.

Before we proceed with a discussion of the distributional properties of the extreme value distribution generations for various sample

sample sizes, we mention a few words on the uniform random number generator.

The uniform random number generator used in the computer procedures is based on the standard random number generator of the Unix operating system version V Library of Programs. The system is available on the PDP II Computer at the Medical Computing and Statistics Unit of the Edinburgh University. For all of the simulations we have used the seventh edition of the Unix operating system and its various software on the PDP II computer. The uniform random generator asks for 2 initial seeds to begin the simulations. We have used values 1 and 2 as the initiators of our generations. In order to allow the random numbers to reach stability we proceed with the generation of 500 random numbers and then use the 501 st generator as the first effective random number for the survival samples. The returning value of the generator is within the range of values 0 and 1. We repeat the generations for the various values of sample size and note that the generations conform to a mean value of $\frac{n+1}{2}$ and the variance of $\frac{n^2-1}{12}$ for various sample sizes that are greater than 20 over the range of values that we examined.

Prior to proceeding with the generation of random survival samples, the extreme value distributed random variables were generated, for different sample sizes. The purpose is to assess the capability of the generator in conforming to the above specifications.

10/

10 Standard extreme value distributed random variables with the sample size of 50

Sample No.	Mean	Variance	Skewness	Kurtosis
1	-.732	1.504	-0.271	9.581
2	-.716	1.249	-0.543	-0.135
3	-.397	1.166	-0.133	-0.146
4	-.727	1.241	-0.151	-0.681
5	-.428	1.262	-2.351	-0.683
6	-.495	1.058	0.223	-1.711
7	-.658	1.417	-0.446	5.597
8	-.280	1.005	-3.107	-2.084
9	-.391	1.150	-0.730	-0.331
10	-.456	1.361	-0.606	0.984.

5 Standard extreme value distributed random variables with the sample size of 100.

Sample No	Mean	Variance	Skewness	Kurtosis
1	-0.581	1.517	-3.745	9.480
2	-0.496	1.075	-0.177	-1.505
3	-0.454	1.412	-2.540	9.669
4	-0.637	1.295	-0.775	0.545
5	-0.568	1.354	-0.965	1.821

5 Standard extreme value distributed random variables with the sample size of 200.

Sample No.	Mean	Variance	Skewness	Kurtosis
1	-0.597	1.362	-2.187	9.462

5 Standard extreme value distributed random variables with the sample size of 200 (cont'd)

Sample No.	Mean	Variance	Skewness	Kurtosis
2	-0.494	1.364	-1.476	3.721
3	-0.458	1.209	-0.362	-0.724
4	-0.462	1.315	-1.005	2.408
5	-0.663	1.270	-1.164	1.421

5 Standard Extreme value distributed random variables with the sample size of 1000

Sample No.	Mean	Variance	Skewness	Kurtosis
1	-0.592	1.260	-1.151	2.656
2	-0.584	1.288	-1.207	2.384
3	-0.575	1.272	-1.118	2.343
4	-0.557	1.271	-1.118	2.670
5	-0.459	1.188	-0.678	2.521

2 Standard extreme value distributed random variables with the sample size of 30,000

Sample No.	Mean	Variance	Skewness	Kurtosis
1	-0.591	1.293	-1.184	2.712
2	-0.588	1.287	-1.125	2.354
3	-0.569	1.285	-1.159	2.674

One important point to note is that the above are random samples and that there has been no selection.

The sample with n = 30,000 clearly shows that we are generating the/

the appropriate distribution and therefore asymptomatically all moments are stabilised and conform to the theoretical values. In so far as the small sample properties are concerned the first moment mean is stabilised at $n = 100$, variance at 200, Kurtosis at $n = 1000$ and Skewness is the last to stabilise at before $n = 30,000$ over the range of values examined.

For practical purposes we only study small sample properties of the statistical methods, therefore is is of interest to know how well mean and variance conform to the theoretical values. However for the sake of consistency it is important to know that given a large enough sample size the distribution conforms fully, as far as skewness and Kurtosis are concerned with the generated population.

5.2 Illustration of the method of generation.

Now for the purpose of illustrating the generation of survival times in the simulation procedure we define a population with the following characteristics.

Let the hazard be constant and fixed at, exp(-5.2983) giving $\lambda = 0.005$. Allow two parameters β_1 and β_2 to have values 0.99 and 0.49 respectively. Let the covariate indicator vectors Z_1 and Z_2 be equally and uniformly allocated to values of 0 to 1. Finally it is assumed that none of the survival times have been censored.

Since we are assuming a constant hazard we use the generating/

generating expression as

$$Y = W + \alpha - Z_1\beta_1 - Z_2\beta_2$$

since there are 2 covariates each at 0 or 1 level,, 4 distributions are generated by the subroutine:

Namely, S(t,0,0) , S(t,1,0), S(t,0,1) and S(t, 1, 1). These survival distributions may be obtained by substituting in the general distribution for

$$S(t, Z_1, Z_2) = \text{Exp}[- e^{-\alpha} t. e^{(Z_1 \beta_1 + Z_2 \beta_2)}]$$

where $\beta_1 = 0.99$

and $\beta_2 = 0.49$

giving

$S(t, 0, 0) = \text{Exp}[-0.005\ t]$

$S(t, 0, 1) = \text{Exp}[-0.005\ t.\ 1.6323]$

$S(t, 1, 0) = \text{Exp}[-0.005\ .t.2.6912]$

$S(t, 1, 1) = \text{Exp}[-0.005\ .\ t.\ 4.3928]$

With a perfect approximation to the above theoretical distributions the estimated values of β_1 and β_2 must correspond to values 0.99 and 0.49 respectively. The 4 distributions are illustrated in the figure (5.2.1), & table (5.2.1) give the details of the sample size of 50 which was generated.

Survival Time	Hazard Function	$\sum_{i=1}^{2} z_i\beta_i$	W	Y	z_1	z_2
44.0	0.00500	0.49000	-1.00926	3.79906	0.00000	1.00000
30.0	0.00500	1.48000	-0.39407	3.42425	1.00000	1.00000
64.0	0.00500	0.49000	-0.63759	4.17073	0.00000	1.00000
62.0	0.00500	0.49000	-0.67770	4.13062	0.00000	1.00000
270.0	0.00500	0.49000	0.79077	5.59908	0.00000	1.00000
264.0	0.00500	0.00000	0.27990	5.57822	0.00000	0.00000
19.0	0.00500	0.49000	-1.84074	2.96758	0.00000	1.00000
16.0	0.00500	0.99000	-1.53203	2.77629	1.00000	0.00000
220.0	0.00500	0.00000	0.09956	5.39788	0.00000	0.00000
20.0	0.00500	0.99000	-1.28939	3.01893	1.00000	0.00000
65.0	0.00500	1.48000	0.35743	4.17575	1.00000	1.00000
64.0	0.00500	0.99000	-0.13930	4.16902	1.00000	0.00000
68.0	0.00500	1.48000	0.41424	4.23256	1.00000	1.00000
7.0	0.00500	1.48000	-1.77981	2.03851	1.00000	1.00000
151.0	0.00500	0.99000	0.71370	5.02202	1.00000	0.00000
71.0	0.00500	0.99000	-0.03391	4.27441	1.00000	0.00000
16,0	0.00500	0.99000	-1.51903	2.78929	1.00000	0.00000
80.0	0.00500	0.00000	-0.90993	4.38838	0.00000	0.00000
84.0	0.00500	0.99000	0.12893	4.43724	1.00000	0.00000
357.0	0.00500	0.00000	0.58049	5.87881	0.00000	0.00000
11.0	0.00500	1.48000	-1.34890	2.46941	1.00000	1.00000
81.0	0.00500	1.48000	0.58743	4.40575	1.00000	1.00000
79.0	0.00500	0.49000	-0.42989	4.37843	0.00000	1.00000
242.0	0.00500	0.00000	0.19377	5.49208	0.00000	0.00000
4.0	0.00500	0.00000	-3.84875	1.44957	0.00000	0.00000
71.0	0.00500	0.49000	-0.54172	4.26660	0.00000	1.00000
30.0	0.00500	0.00000	-1.87210	3.42621	0.00000	0.00000
23.0	0.00500	0.49000	-1.64768	3.16063	0.00000	1.00000
2.0	0.00500	1.48000	-2.86412	0.95420	1.00000	1.00000
18.0	0.00500	0.49000	-1.90699	2.90133	0.00000	1.00000
74.0	0.00500	0.99000	0.00538	4.31370	1.00000	0.00000
12.0	0.00500	0.99000	-1.78716	2.52116	1.00000	0.00000
77.0	0.00500	0.99000	0.04566	4.35398	1.00000	0.00000
119.0	0.00500	0.49000	-0.02331	4.78501	0.00000	1.00000
85.0	0.00500	0.49000	-0.36473	4.44359	0.00000	1.00000
0.0	0.00500	0.99000	-5.07574	-0.76742	1.00000	0.00000
121.0	0.00500	0.00000	-0.50017	4.79814	0.00000	0.00000
201.0	0.00500	0.49000	0.49686	5.30518	0.00000	1.00000
35.0	0.00500	0.99000	-0.75237	3.55595	1.00000	0.00000
2.0	0.00500	1.48000	-2.81221	1.00611	1.00000	1.00000
242.0	0.00500	0.99000	1.18442	5.49274	1.00000	0.00000
16.0	0.00500	1.48000	-1.00116	2.81715	1.00000	1.00000
48.0	0.00500	0.99000	-0.42716	3.88116	1.00000	0.00000
12.0	0.00500	0.49000	-2.26778	2.54053	0.00000	1.00000
405.0	0.00500	0.99000	1.69616	6.00448	1.00000	0.00000
385.0	0.00500	0.49000	1.14530	5.95361	0.00000	1.00000
15.0	0.00500	0.49000	-2.08965	2.71867	0.00000	1.00000
154.0	0.00500	0.99000	0.73326	5.04158	1.00000	0.00000
11.0	0.00500	1.48000	-1.40793	2.41039	1.00000	1.00000
43.0	0.00500	0.49000	-1.04437	3.76394	0.00000	1.00000

Table (5.2.1) Actual survival times generated from distributions of figure (5.2.1)

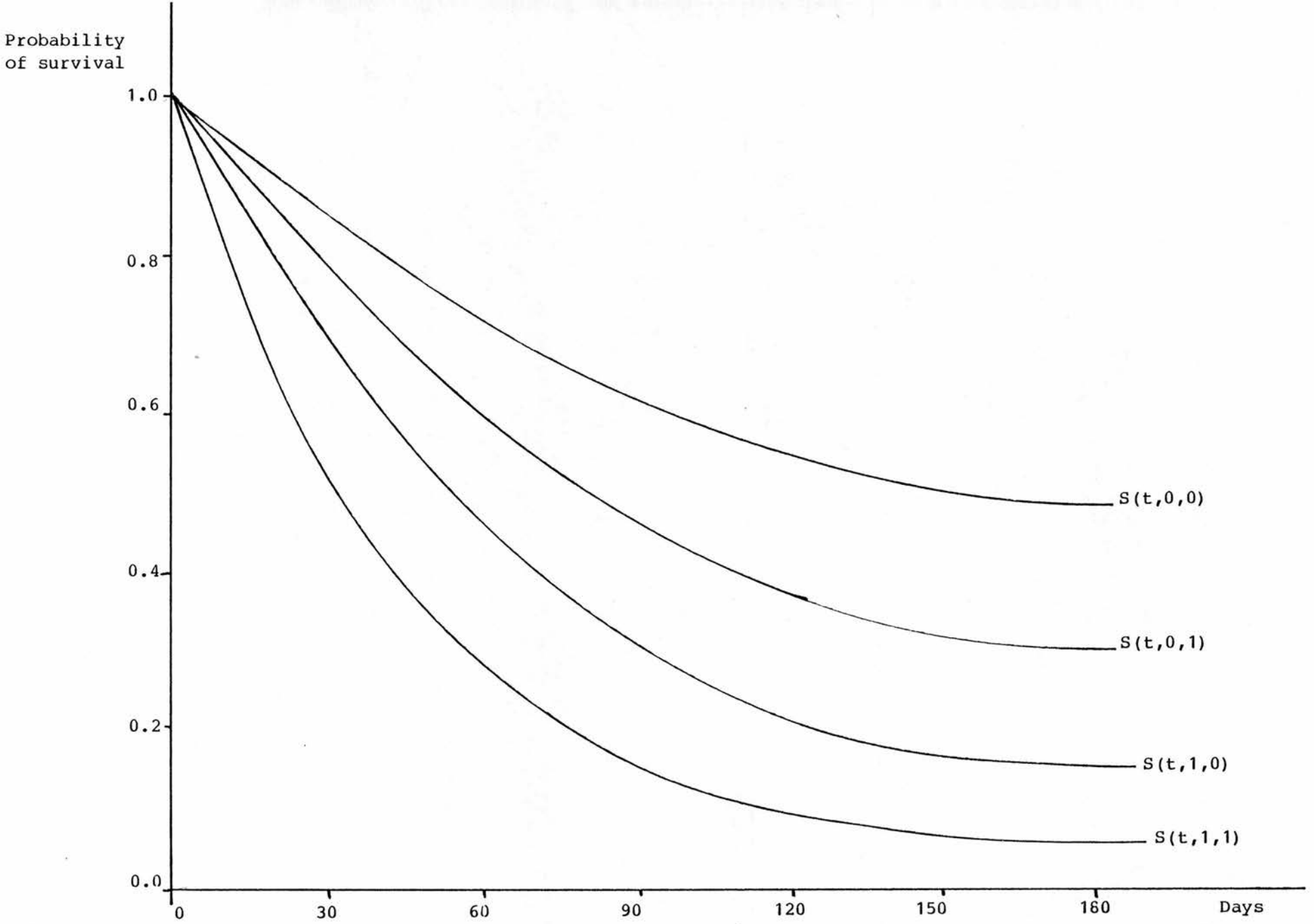


Figure (5.2.1) Theoretical distributions for a generated sample.

The values for the columns of the table (5.2.1) correspond to

$$Y = W + \alpha - Z_1 \beta_1 - Z_2 \beta_2$$

As an example for the first row we have

$$\alpha = -\ln (0.005) = 5.2983$$

$$Z_1\beta_1 + Z_2 \beta_2 = 0.49$$

$$W = -1.00926$$

$$Y = -1.00926 + 5.2983 - 0.49 = 3.79906$$

$$\text{Time} = \text{Int} [\text{Exp} (3.79906)] = 44$$

We can plot the above data using life table analysis methods to derive cumulative survivals (as mentioned before). The purpose at this stage is not to do a detailed comparison of the estimation methods bur rather a general overview of the survival generator.

The following 4 tables give comparisons of the cumulative survival estimates using the product limit estimation, as in Figure (5.2.2) and the theoretical value of the exponential distributions in Figure (5.2.1)

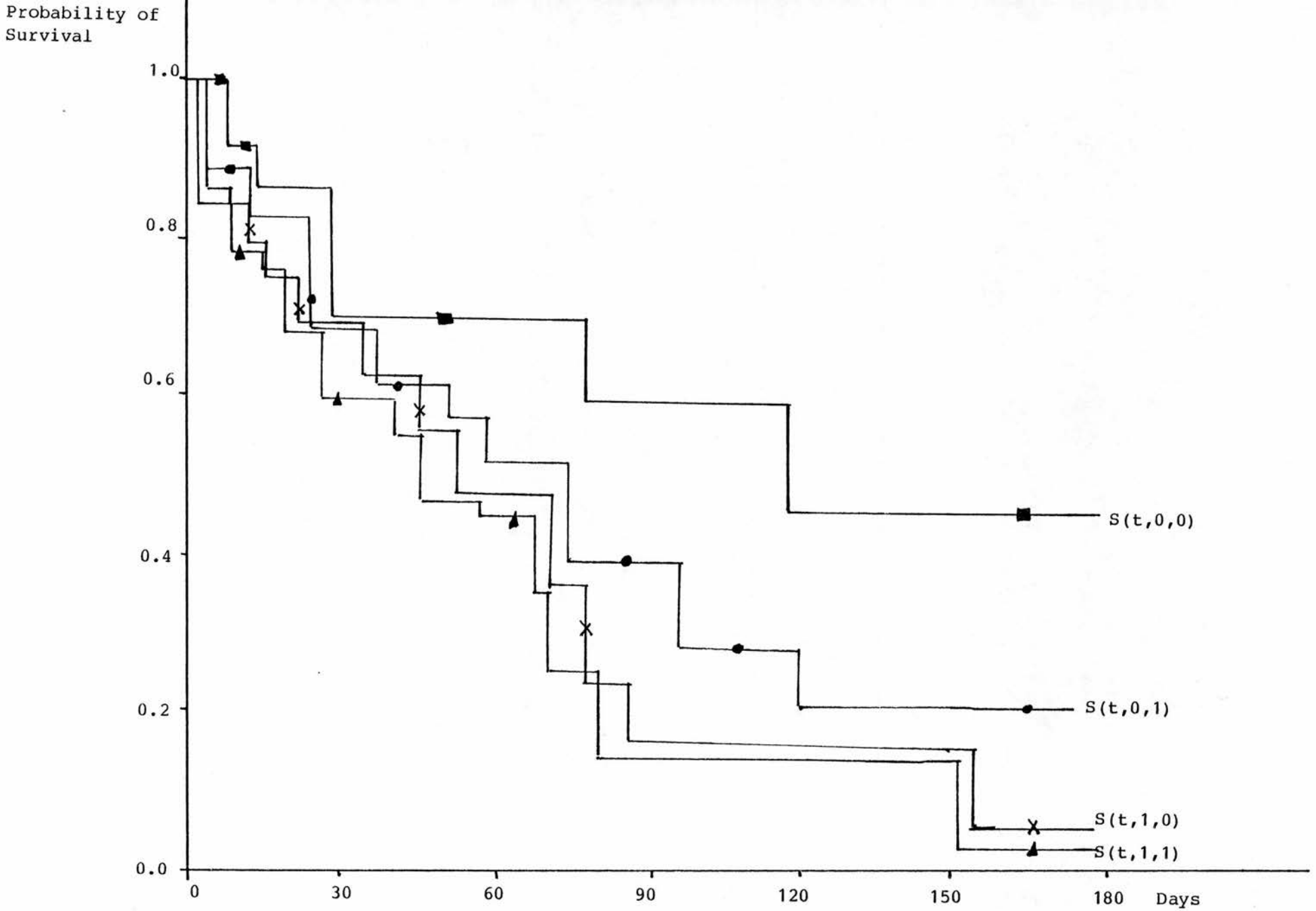


Figure (5.2.2) Actual distributions for a generated sample.

Time	Cumulative Survival	S.E.	Theoretical S(to,o)
4	.8750	.1169	.9801
30	.7500	.1531	.8601
80	.6250	.1712	.6703
121	.5000	.1768	.5460
220	.3750	.1712	.3328
242	.2500	.1531	.2981
264	.1250	.1169	.2671
375	.0000	.0000	.1533

Table (5.2.2)

Time	Cumulative Survival	S.E.	Theoretical S(t,o,1)
12	.9375	.0605	.9067
15	.8750	.0827	.8847
18	.8125	.0978	.8633
19	.7506	.1083	.8563
23	.6875	.1159	.8288
43	.6250	.1210	.7040
44	.5625	.1240	.6983
62	.5000	.1250	.6028
64	.4375	.1240	.5431
71	.3750	.1210	.5601
79	.3125	.1159	.5247
85	.2500	.1083	.4997
119	.1875	.0976	.3780
201	.1250	.0827	.1938
270	.0625	.0605	.1104
385	.0000	.0000	.0431

Table (5.2.3)

Time	Cumulative Survival	S.E.	Theoretical S(t,1,0)
0	.9333	.0644	1
12	.8667	.0878	.8508
16	-	-	-
16	.7333	.1142	.8063
20	.6667	.1217	.7640
35	.6000	.1265	.6244
48	.5333	.1288	.5241
64	.4667	.1288	.4226
71	.4000	.1265	.3846
74	.3333	.1217	.3694
77	.2667	.1142	.3548
84	.2000	.1033	.3229
131	.1333	.0878	.1310
134	.0667	.0644	.1259
405	.0000	.0000	.0043

Table (5.2.4)

Time	Cumulative Survival	S.E.	Theoretical S(t,1,1)
2	-	-	-
2	.8182	.1163	.9570
7	.7273	.1343	.8574
11	-	-	-
11	.5454	.1501	.7853
16	.4545	.1501	.7036
30	.3636	.1450	.5174
65	.2727	.1343	.2398
68	.1818	.1163	.2245
81	.0909	.0867	.1687
242	.0000	.0000	.0049

Table (5.2.5)

5.3 Generation of Censored Survival Times.

As we described earlier, in the first step we generate a survival distribution S(t) according to a set of covariates and the shape of the hazard function. A sorted plot of this sample will show a survival pattern as in figure (5.3.1). It is assumed that all patients enter the study at time zero. In this sample there are no censored cases. In terms of clinical trials it is assumed that sufficient time has passed since the start of the trial to allow an observation of full length of survival.

In real data from a clinical trial the situation is slightly different. Patients do not arrive simultaneously into the study. Patients are not observed for a full length of survival,either because they drop out of the study, or analysis is performed at a time that not all patients have had the chance of producing a complete survival time.

First we consider the problem of arrival or accrual period. This kind of follow-up study is composed of two periods, accrual and follow-up period. The accrual is a period to allow a sufficient number of patients entre the study so that a reasonable statistical comparison may be made of the patients. Thus the accrual period in such a study becomes a function of the required sample size and the rate of arrival of patients. In effect the accrual period is prejudged by the value of treatment difference which the study is designed to detect. In general it may be assumed that all patients entre the study uniformly, and that there are no trends or seasonal patterns present in the covariates according to the accrual period/.

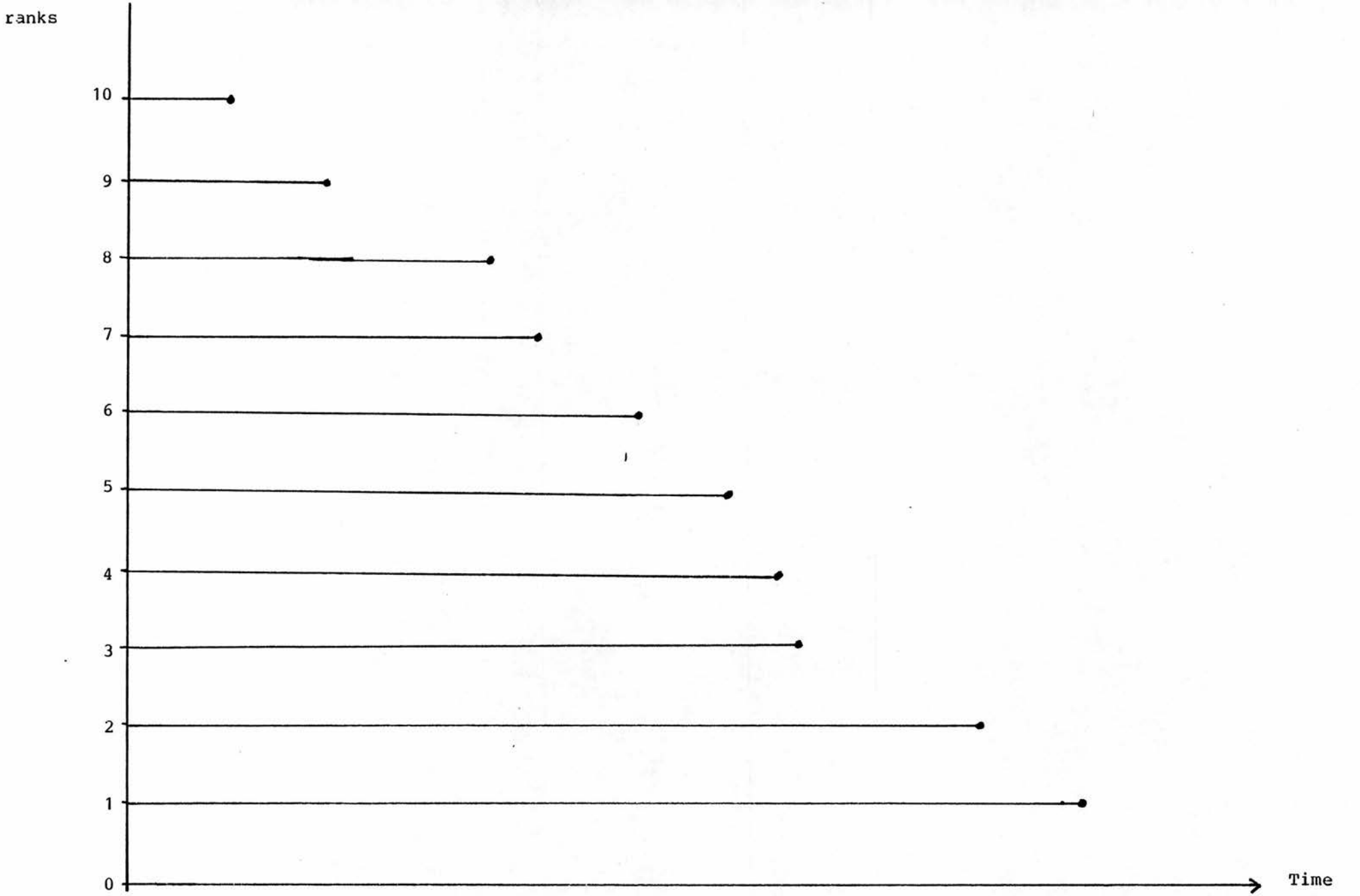


Figure (5.3.1) Tables of sorted relative ranks of survival time against time

period, In practice randomisation of the patients provides this condition for the main treatments. However the conditions may not hold for the prognostic factors. An example can be a situation where by chance younger patients are entered in the first year of a study and older patients in the second year.

Another period we consider is the follow-up time. In so far as the clinical trial procedure is concerned a good clinical trial provides conditions and procedures so that all patients may be followed-up and that at the end of study or time of the interim analysis, survival status is recorded for all of the patients. Such a condition guarantees that censoring, if it occurs, is only due to reasons of treatment, disease and patients and not due to the procedures of follow-up and withdrawals. In the simulations of this chapter we assume that the above condition holds. Follow-up period is in practice often dependent on many external factors, due to constraints from management of patient care. Further it is customary to do a number of interim analysis of the data. For an efficient unambiguous analysis there should be no crossing patterns present in the survival rates. It is often the case in practice and any variability in the number of survivors may be attributed to the survival rates of the subgroups of patients.

There is one complicating factor in practice which arises in multiple failure studies. If there is more than one form of failure responsible for the reduction of the cases, the usual approach of analysis is then by classifying one set of end points/

points as deaths and another set as censored. Such a methodology requires a different method of generation of censoring allocations. This method can also be used in a way to allow a different censoring rate to be present for each covariate set, so that the problem of lost to follow-up for different groups can be assessed. Within the present simulations we assume that there is only one single cause of death and that all censorings are by definition due to the fact that not enough time has passed since the entry of a case for a death to take place.

In the introduction we mentioned the various forms of censoring and stated that in trials often we are interested in random censoring, by which arrival of patients is not fixed but occurs within an accrual period and thus any censoring at time of the analysis is a random censoring. This latter procedure is the method we use for the generation of the random samples. However as will be described later rather than fixing the time of the analysis we allow time of the analysis to vary slightly so that we can have fixed censoring percentages at time of each analysis. Thus we summarise.

Total length of study = Accrual period + Follow-up period.

For each individual patient we have a time t_i which is generated from S(t), and a time from start of the trial to the entry of patient i. We will let a_i denote this time and is a uniformly distributed random variable between 0 and A.

Thus the total length of study for the patient i is -

$$T_i = a_i + t_i$$

We can now transform the figure (5.3.1) into the figure (5.3.2) in/

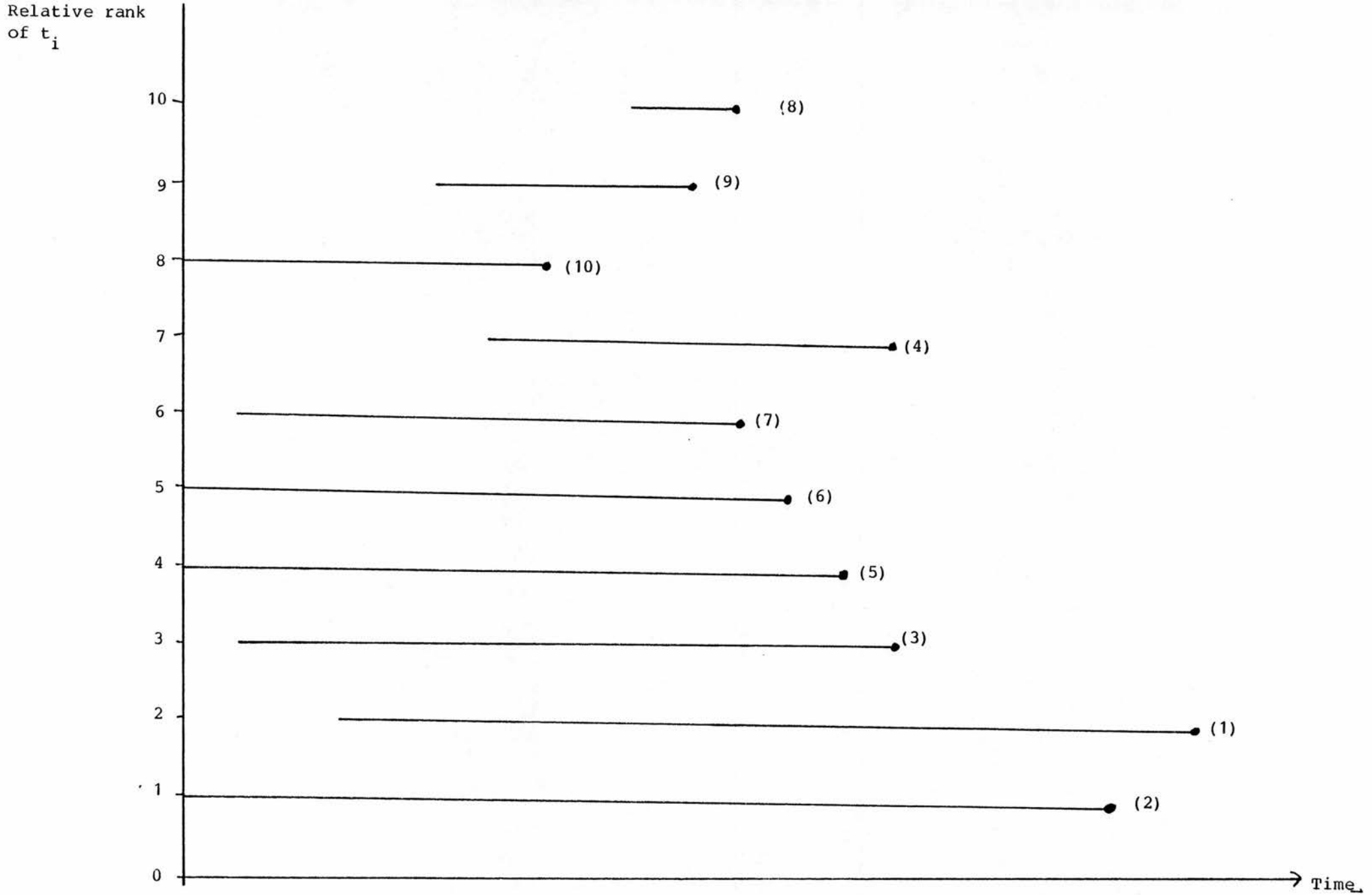


Figure (5.3.2) Relative ranks of t_i and a_i, sorted according to t_i

in which the accrual period is also represented. Now we can produce the figure (5.3.3) in which the data is sorted accrding to the values of $a_i + t_i$.

As we mentioned before most trials are analysed in one of the following two situations. Either a final analysis is performed prior to the minimum sufficient time for producing a complete survival time, or that the clinical trial results have been formed and discussed at an interim stage. In both cases we can generalise to the following: every trial analysis has a fixed value I which is a point in the time of the study when survival information prior to it are complete and all possible events after this time are taken to be censored.

A crucial factor before the start of a trial is a decision on the likely number of events. Two factors that are in practice of considerable interest in design of a trial, are the accrual period and the number of patients in the study. Based on these assessments a decision is finally made on the appropriate times at which interim analysis and the main analysis may be performed. Now it seems proper to summarise some of the generalisations that have been made in the course of our discussions.

(1) It has been assumed that censoring is synonymous with non-informative censoring. Therefore there are no situations that patients leave the study due to side effects or other forms of failure. Thus patient censoring times are distributed in the same manner between subgroups. An example of the violation of the above assumption would be a higher dropout rate from an arm of a trial due to/

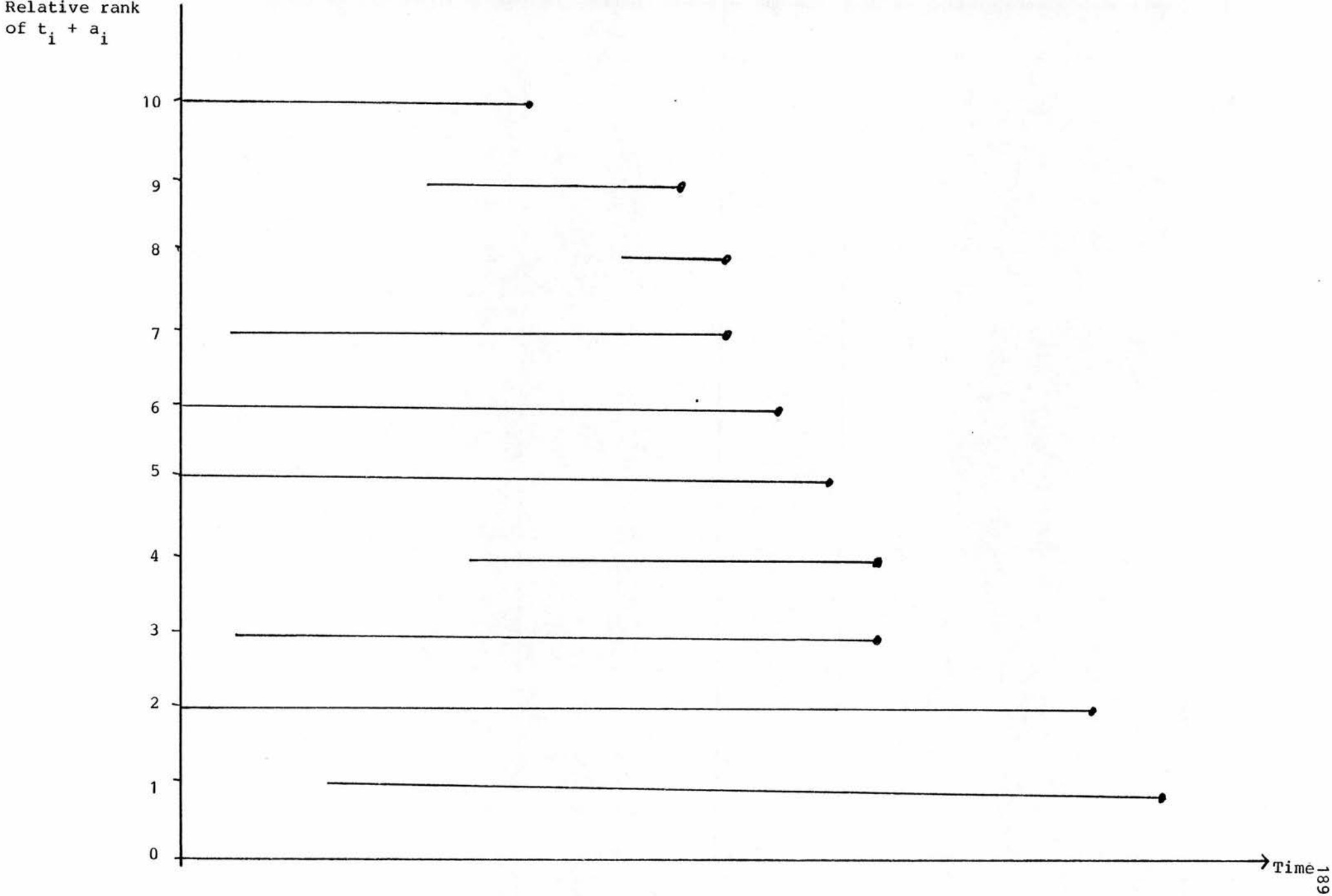


Figure (5.3.3) Relative ranks of $t_i + a_i$, sorted according to $t_i + a^i$

to a particular prognostic indicator or treatment.

(2) Further if the relative rate of failure is not constant between subgroups, that is there exists a time dependent relative risk between their corresponding failure rates, then the level of censoring must follow a trend in time. That is the presence of time dependency has a direct effect on number of cases that are dead or are censored at any given time. As an example if one subgroup has a higher failure rate at the end of study, a measure of censoring percentage will give a different relative difference at early and late stages of time scale. Aternatively no time trends will imply a constant relative rate of censoring.

The relation between (1) and (2) is an interesting one and represents the relationship of problems of competing risks and the time dependency of the hazard rates. Due to the methods of generation of censoring times we do not proceed with the study of method (1). However by use of differing failure rates we will study the effects of time dependency. In a descriptive manner we consider (1) to be a causal situation within which we have a fixed set for a cause of death. Using the example given in (1) a good analysis may indicate a link if it exists between a prognostic variable and a particular cause of death. In the example (2) however we are describing the failure rate as a form of a function of time. This function of time however need not be of a continuous form as described in the example given in (2). In fact the description of (1) and (2) and their examples can be exchanged at times in the language of the other, with (2) being slightly more flexible. We can describe cause of death in (1) in terms of time dependency of (2)/

(2) by letting time dependency be a parametric function of the type failure. An example is the effect of old age or survival distributions. For reasons of consistency of the conclusions one may use an approach by which deaths suspected of old age are dealth with as censored, or alternatively define a functional form of the old age and incorporate this function into the relative risk as a diagnostic check of the relative risk. As we pointed out in this section we will concentrate on the simulation of time dependent type, and later in Chapters 7 and 8 we will consider introduction of types (1) and (2) in appropriate application with various functional forms of time dependency.

Finally in this section we discuss some of the points regarding Censoring times. A purpose of this study is to evaluate the power of different tests according to their level of censoring. An analytical assessment is impossible, thus we require some criterion. Such a criterion must be general enough to be relevant to real life practice and thus easy to draw relevant conclusions from. In the next section we will discuss such a criterion. However on the point of censoring the accrual period and follow-up period both can be thought of as some form of fixed variables and thus we can generate different levels of censoring according to the relationship between them.

As an alternative to the above we can fix the level of censorings. Thus a 10% censoring in a sample of 50 implies that the interim time is somewhere between $(a_{45} + t_{45})$ and $(a_{46} + t_{46})$ see figure (5.3.3). Again since we are only considering a fixed percentage/

percentage value for the level of censoring at interim analysis, in case of ties being present, the exact value of I may be adjusted by a uniform distribution between 45th and 46th survival times. All accrual time plus survival times after this point in time are censored.

The procedure is thus fixed by a set of covariates, the percentage of censoring, a fixed value of hazard rate and a fixed value of accrual period for the sample . It is important to note that the value of hazards must be fixed so that reasonable survival times are produced. Similarly the value of accrual period also has to be adjusted so that a realistic sample is generated. For the following simulations we fix the base line hazard at a constant value of Exp(-5) = 0.00674, with an accrual period of 50 units (which may be considered as months). For each single simulation level of the above values we repeat the simulations 300 times. this value of 300 is set to be fixed for all simulations presented in this study.

5.4 Criteria for the Comparison of the methods of analysis.

A widely held view among statisticians involved with the design of clinical trials is that, the sample size and power assessment are the most crucial factors in proceeding with a scientifically sound trial. This scientific stand is in practice often confronted with management constraints that eventually lead to a form of compromise in the design of trial.

In/

In the introduction we mentioned some of the drawbacks and difficulties in the analysis of data with small samples in the presence of a number of covariate effects and a process of time dependency. We termed such effects in general to be interrelations. Later we will analyse real clinical trial data with an exploratory emphasis and use of such interrelations. In the later chapters we we will discuss functional forms of time dependency. We will now deal with small sample properties of the Cox's proportional hazard models. The main aim is to establish a criterion for a comparative study of small sample sizes, under trial design constraints. The factors that we have taken into account in the generation of the survival samples have been chosen with a particular emphasis on crucial design factors in a realistic clinical trial. These factors are accrual period, censoring, sample size and interim analysis time. In the process of the generations we have constantly adopted a generating procedure that we have considered useful for a range of applications in the later chapters. However a more detailed theoretical study based on asymptomatic properties can take a route different from the one we have taken. Areas which may pose interesting directions are situations of competing risk generations and stratified analysis with varying accrual periods. It seems that study of small sample properties of methods as opposed to a more theoretical study of the asymptomatic properties of the method of analysis is a useful approach in a better understanding of the scientific design constraints. S.D. Silvey (1975) discusses in detail small sample properties of statistical tests with simple hypothesis (only one estimator) and composite hypothesis (with more than one estimator or nuisance parameters). Such properties are related to a function of different

different errors that take part in any statistical hypothesis, namely, Type I error and Type II error, represented by $E\alpha$ (false positive) and $E\beta$ (false negative) respectively. (Although α and β are often denoted in this context from now we refer to them as $E\alpha$ and $E\beta$ to avoid confusion with β covariates).
In this work we are dealing with survival data, however these definitions can be generalised to all statistical tests. In fact in a clinical trial setting another branch of tests are used that are based on a simple proportion and at times this simple proportion can provide all the necessary information regarding a new treatment.

The tests based on time to failure are playing an increasing role. Such tests are based usually on producing some form of approximation to the probability distribution of failure times or life tables. In the chapters 2 and 3 we made the necessary distinction between the distribution free tests and parametric approaches that are related to life tables and discussed their properties. In the evaluation of the sample size however at start of treatment, some form of parametric assumption must be made. The most common is to assume an exponentially distributed survival time, with the proportion of survivors approximated by

$$S(t) = e^{-\lambda t}$$

For N patients with the mean survival time M, and the hazard estimated by M^{-1}, which is asymptotically $L \sim N(\lambda, \lambda^2 / N)$. These results are analogous to the results of the Chapter 2 on the discussion of the exponential distribution. Much of the early works in efficiency studies were based on the study of the asymptotic properties of the proportional hazards model and the exponential approximation. Such comparisons/

comparisons are useful in practice due to conflicting benefits of the methods. For example, although the exponential analysis is more efficient given that the sample is generated from an exponential distribution, the proportional hazard model has a far better robustness property, when the data is not exponentially distributed. Various authors have discussed asymptotic properties of the Cox's approach. Kalbfleish (1974) discusses the asymptotic efficiency for a single covariate model. Efron (1977) discusses conditions for full asymptotic efficiency, Kay (1979) provides a comparison of two covariate models with the exponential and Kalbfleish and Macintosh (1977) expand the results to time dependent situtation. The results indicated that for covariates not dependent on time, the estimations based on ranks is fully efficient for $\beta = 0$ and has good properties for $\beta \neq 0$. For the case of time-dependent covariates the asymptotic properties are related to the ratio of the hazard rates. From a different position properties of the proportional hazard models have been studied in relation to the logrank test. (Crowley (1974) and Tarone and Ware (1979).) It is shown that asymptotically the Cox's method can lead to the logrank test. Lustbar (1980) derives the Wilcoxon test as a special case of Cox's model with a time dependent covariate. It is shown that the two fully distribution-free tests in fact differ only in their choice of weights as functions of the number of cases at risk. A more detailed discussion of those tests was given in Chapters 2 & 3.

Although the above studies are useful in allowing some form of comparison between the methods, they do not allow comparisons for differing sample sizes, censoring rates and censoring methods. The point/

point is made by Oaks[a] (1981), that in small sample studies the expectations may be different and preculiarities may be present.

In the study of the proportional hazard models and the parametric models, two tests have been used in general. One is the maximum likelihood estimation with asymptotic normality assumption and the other is the likelihood ratio test. These results are given in more detail for parametric methods in Chapter 3 and for proportional hazards in Chapter 4.

It seems proper now to follow the study of small sample properties in the following directions:

(1) Study more than 1 covariate with constant hazards.

(2) Study non-constant hazards.

(3) Study non-proportional hazards.

In the first instance we study the different methods on exponentially generated samples with two covariates, firstly, as a matter of comparing relative power of the small samples and secondly as an expansion of the 1 covariate study. As we mentioned earlier two types of error are of interest. Now we develop these definitions so that they may be used as a criterion for the comparisons. Type I error represented by E_α is under the control of statistician at the end of study and type II error, E_β is dependent on sample size and the value of the covariate. The hypothesis of special interest for all practical reasons is that of $\beta_2 = 0$. The power of this test is then noted for the varing levels of β_1 and β_2 over the different simulations. Another hypothesis we consider is for a composite test of (β_1, β_2). The main purpose for this test is a theoretical one and may/

may be useful in exploratory stepwise regression techniques. The theoretical basis of this point will be made clear later. The rate of acceptance of the null hypothesis as the actual β_1 & β_2 values differ from zero will give a measure of type II error. Namely it is a function of the proportion of times that we may wrongly accept the null hypothesis when it is false. At values of β_1& β_2 fixed at zero however, the same proportion is a function of the type I error which has the proportion of times we wrongly reject the null hypothesis when it is true. The last statement in computational terms may be represented by an equivalent rephrasing in which value of β_1 is fixed at β_{01} and value of β_2 is set to β_{02} and we test the null hypothesis as $(\beta_1\ \beta_2) = (\beta_{01}\ \beta_{02})$. The power of a test is a function of the alternative hypothesis and is related to E_β , in the following way. The sample space of an observation in any test may be divided into two regions. One region is called the acceptance region and if the estimates fall into this space we accept the null hypothesis. The rest of the space is called the critical region.

Thus -

Prob(estimator falls within the critical region$\backslash H_1$) = $1-E_\beta$ = Power.

and Prob(estimator falls within the acceptance region$\backslash H_1$) = E_β

similarly for the null hypothesis.

Prob(estimator falls within the critical region$\backslash H_0$) = E_α

So far we have defined the power and acceptance region in terms of the null hypothesis and the alternative hypothesis. Going back to the opening section of this chapter we rephrase by saying that in general we seek a critical region such that the power is as large as possible. Then in addition to the control of probability of

Type/

Type I error at E_α we shall have minimized the probability of Type II error at E_β.

These definitions in terms of the survival analysis are always further complicated by the fact that we are only interested in testing a subset of the estimators that define our parameter space of the critical region and thus we always deal with a composite hypothesis, as opposed to a simple hypothesis where all the distribution is fully defined. This point in general is related to the effects of the hazard functions in the estimation of the relevant covariate estimators.

The above points regarding the composite hypothesis for small samples is mainly a problem of illustration in here rather than a theoretical one. Neyman and Pearson (1933) justify a method in testing a single hypothesis against a simple alternative. That is if we are choosing between two completely specified distributions then problems of finding a best critical region is simple and they provide a solution. Further results of Lehman & Scheffer (1950) permits us to reduce the above problem of finding a most powerful region for a composite hypothesis to a familiar problem of finding a best critical region for a simple hypothesis.

Now we illustrate the problem for a simple case as that of obtaining an area of overlap of the error distribution of the estimator and the sample parameter distribution. Figure (5.4.1) represents the error regions for a one sided test. The null hypothesis relates to the estimator with $N(\mu_0, \Sigma_0^2)$ and the alternative is $N(\mu_1, \Sigma_1^2)$. If we adopt a significance legel of E_α for H_0 then, chances of type I error is E_α , that we obtain the wrong conclusion when/

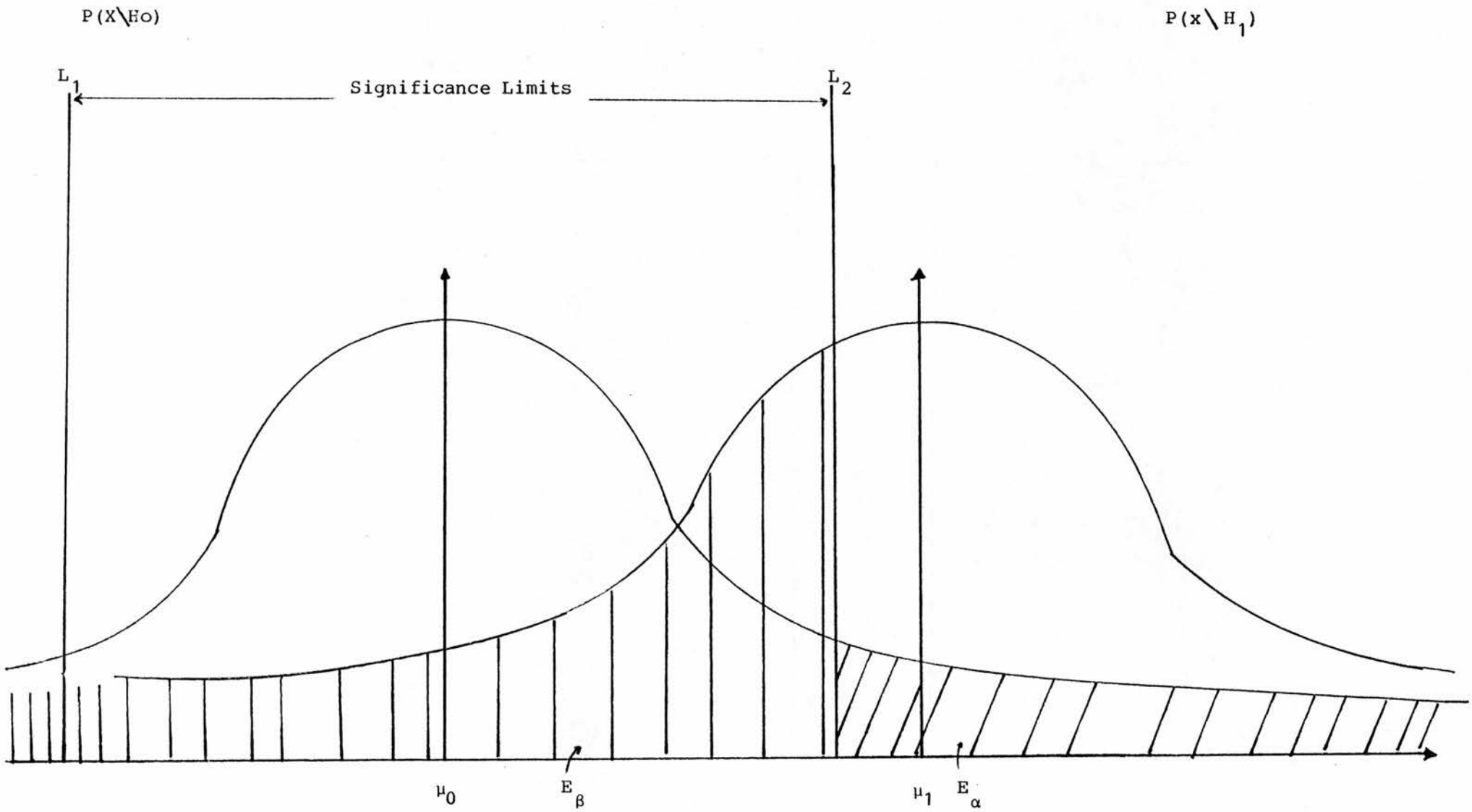


Figure (5.4.1) Errors of Type I and Type II. Line L1 is not relevant to the one sided tests of hypothesis.

when H_0 is true.

The possibility of obtaining a type II error is E_β , that is opting for the wrong conclusion when H_0 is false. The actual value of the percentage of the rejections is clearly dependent on the form of the hypothesis that we adopt, and the actual value of the coefficients associated with the covariates. These will produce an indication of type I and type II errors. For a single covariate situation we have.

H_0: = 0

Reality	Conclusion: $\beta = 0$	Conclusion: $\beta \neq 0$
$\beta = 0 \Rightarrow H_0$ true	$1 - E_\alpha$	Type I = E_α
$\beta \neq 0 \Rightarrow H_0$ false	Type II = E_β	$1 - E_\beta$ = Power

In the above the percentage of rejections are represented by the number of cases that fall into the second column of the table. Given that the null hypothesis $H_0 : \beta = 0$ is true, we obtain a measure of E_α , the observed significance level. Given that the null hypothesis is false we obtain a measure of the power of the test, which is a fuction of sample size, censoring percentage and the magnitude of the coefficients of the covariates.

The final remark on the method of simulations relates to the allocation of the covariates in each sample generation. As was remarked earlier the variance covariance matrix of the covariates plays an important role on the type II error. It is well known that if different subgroups are numerically divided in an equal manner then the efficiency is at maximum. Further the covariances between these values play an important role, in that, depending on the value of each/

each covariate a high correlation can reduce the efficiency.

5.5 Description of parameter range for the trial simulations.

In the generation of the (Z_1, Z_2) matrix we use the same random number generator as before. We randomly allocate -1, and 1 values as a dichstomos function to Z_1 and Z_2 for each patient. We obtain these values by dividing the 0 to 1 range of the uniform random numbers into appropriate segments. Thus there are 4 sets of patients in the data with the Z_1 and Z_2 values set to (-1, -1), (-1, 1), (1,-1) and (1,1).

A well designed trial would allocate equal numbers of patients to each arm of the trial. Any other covariate set is not usually controlled. All other uncontrolled effects due to a large sample size at times do level out in terms of treatment effects and within each type approximately equal numbers are usually allocated to each arm of a trial. The importance of equal subgroups is mainly noticed in the power of the tests. Tests usually are at their maximum efficiency if they are composed of equal subgroups.

In the generation of the covariate sets we use a uniform random number generator. The covariate sample generations however are fixed so that for a given sample size every simulation is composed of exactly the same covariate sets. Within the sample set however we intend to balance the treatment effect so that there is a 50:50 likelihood of allocations to a particular treatment. For the covariate set a different ratio of the two covariate values are used, so that we do not have a symmetrical relationship between β_1 and β_2. Thus a particular set/

set of values of $(\beta_1, \beta_2) = (a,b)$ does not necessarily correspond to a power efficiency value for $(\beta_1, \beta_2) = (b,a)$. The consequence of this effect is similar to the use of a lower sample size for β_1 in comparison with β_2

At this point a few remarks are needed regarding the different possibilities of the generation of the covariate effects. One method of generation of the covariate effects would be to allocate a different set of Z_1, Z_2 variables at each simulation. Such a method implies that any power assessment is complicated by the sample covariate variability. An alternative is to fix the Z_1, Z_2 variable set for any required proportion within the covariate categories and treatment groups. The resultant consequence is that the final results are conditional on the generated proportions within the corresponding treatment and covariate groups. Bryson and Johnson (1981) discuss a method for the generation of the earlier approach and point out some of the theoretical problems with generation of monotone likelihoods in such generated samples. However, this problem could be avoided since it is realistic to add a restriction within generated simulations so that no subgroup should be generated which contains less than say 10% of the total number of sample size. The rest of the procedure would then be confined to dividing the uniform random distribution range from 0 to 1 into the relevant segments as required. However, we are restricting ourselves to a fixed covariate set for all samples and so the above problem does not arise. The method for the generation of the 40:60 ratio prognostic covariate set, Z_1, and the 50:50 ratio treatment set , Z_2, is to divide the uniform random number scale to the following categories. We let generations corresponding/

ponding to uniform random numbers between 0 and 0.4 to be the low level of Z_1 and values of random numbers greater than 0.4 and less than 1.0 to correspond to the high level of Z_1. Further we subdivide the two parts of the range of the uniform random number corresponding to high and low levels of Z_1 into two equal parts. So that within 0 to 0.4 range there is a 50:50 chance of allocation to high and low levels of Z_1 and within 0.4 to 1.0 there is a 50:50 chance of allocation to low and high levels of treatment effect. As an asymptotic property of the sample the ratios of the margninals of the treatment and covariate indicators will then approach the required ratios. As we pointed out earlier the major emphasis is on small sample properties and although the above approach may be justified under certain theoretical conditions, in a simulation of a clinical trial it is sufficient to condition our results on a generated sample that conforms fully with the required ratios.

Now we summarise the properties of the generated sample and the value of each parameter.

The random variable W has the standard extreme value distribution.

In terms of the survival times it is related to it by the function.

$$Y = \frac{W}{p} + \alpha - \frac{Z_1 \beta_1}{p} - \frac{Z_2 \beta_2}{p}$$

Where Y is the log of survival times. (In the process of derivation of the above Weiball generating function we have used ν in place of p)

$\alpha = 5$ giving $\lambda = 0.00674$ in $\lambda = \text{Exp}(-\alpha)$

β_1 values range contains (-1, -.5, -.2, -.1, 0,.1,.2,.5, 1)

β_2 values are (0, .1, .2, .5, 1)

p is/

p is varied from Exp (-.3) to Exp(+.3)

For the censoring patterns we distinguish between accrual times and the survival times. The survival times are generated by the above function. The accrual times are uniformly distributed between 0 and 50. The levels of censoring are fixed at 0, 5%, 10% and 30%. The significance levels are set at 0.05, 0.01, 0.005. The sample sizes vary at 25, 50, 100. We set the null hypothesis to be

Ho : $\beta_2 = 0$ and later

Ho : $(\beta_1, \beta_2) = (0, 0)$ and vary the values of β_1 & β_2 in the region that we mentioned.

The low end of the magnitude of $|\beta_1|$ and $|\beta_2|$ where $\beta_1 = \beta_2 = 0$ The power represents the type I errors. We repeat the simulations for the one sided alternative hypothesis which is the necessary condition of some clinical trials. The above range of sample size and censoring levels form the complete simulated sets. However, if the results are at times very close to each other we will only comment on their overlap.

5.6 Discussion of the simulation results.

In the first instance we refer to figures (5.6.1) to (5.6.6) in which a representation of the power efficiency is given for the null hypothesis of Ho : $\beta_2 = 0$. Clearly by the figures we obtain almost parallel lines for the range of the various covariate values. In fact the main determinant of the efficiency is the value of the β_2 magnitudes. At the β_2 values equal to zero we obtain a representation of the type I error in all cases, and this value is consistent over a range of factors. The most striking feature of the results/

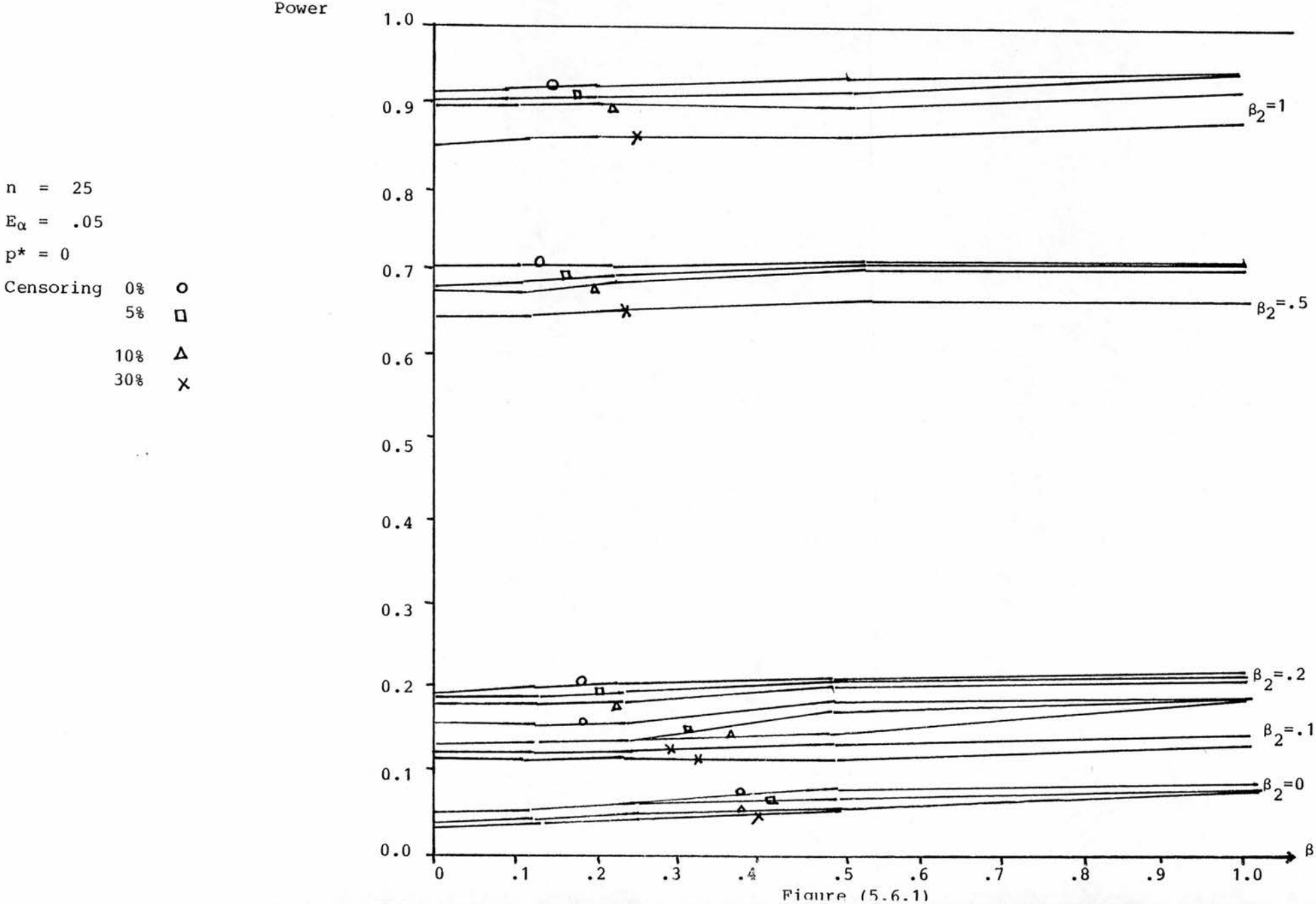
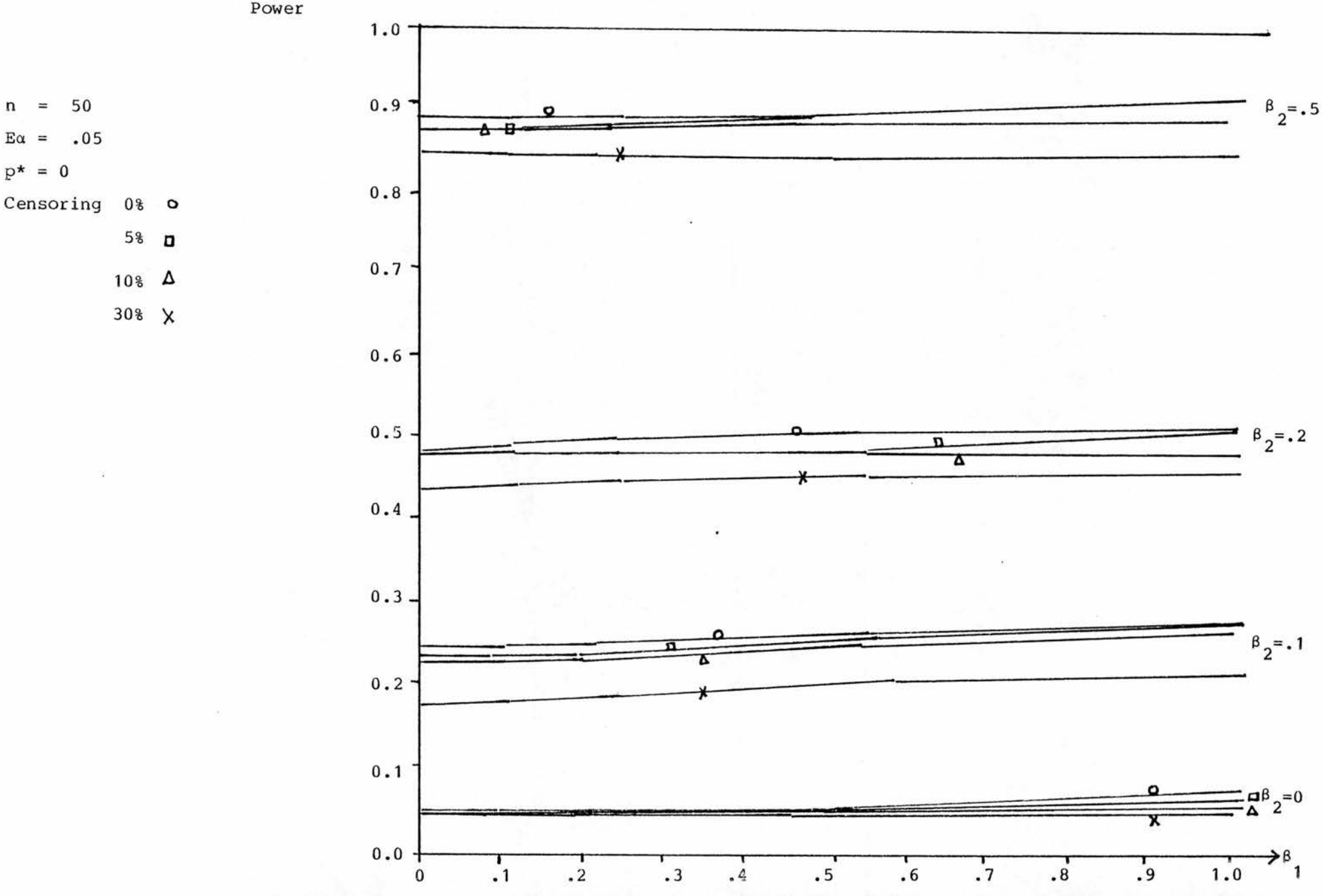

Power
n = 25
E_α = .05
p* = 0
Censoring 0% o
5% □
10% Δ
30% X
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
β_2=1
β_2=.5
β_2=.2
β_2=.1
β_2=0
β_1
0
.1
.2
.3
.4
.5
.6
.7
.8
.9
1.0

Figure (5.6.1)


Power
n = 50
Eα = .05
p* = 0
Censoring 0%
5%
10%
30%
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
β2=.5
β2=.2
β2=.1
β2=0
0
.1
.2
.3
.4
.5
.6
.7
.8
.9
1.0
β1

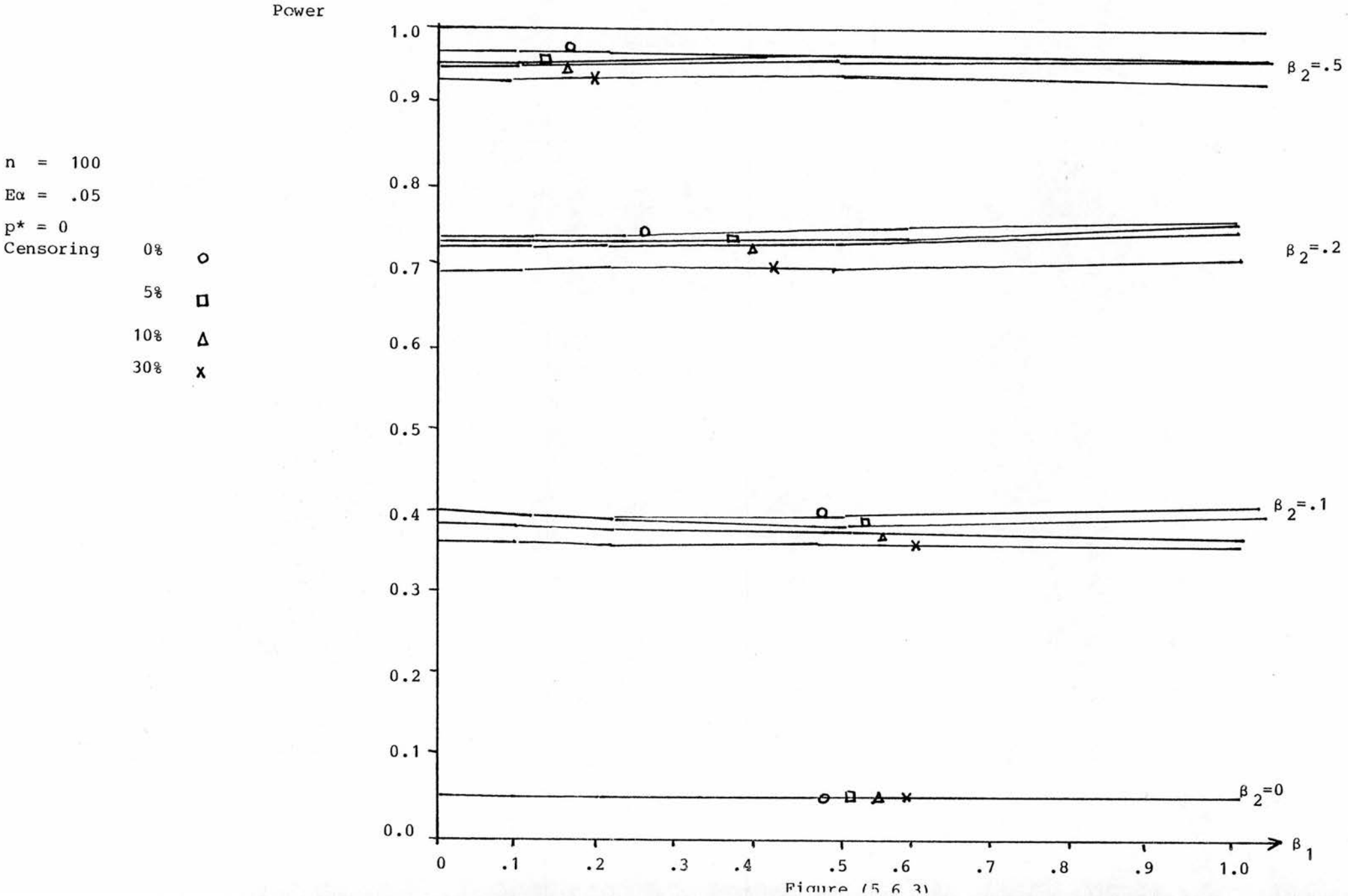

Power
n = 100
Eα = .05
p* = 0
Censoring
0%
5%
10%
30%
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
β 2=.5
β 2=.2
β 2=.1
β 2=0
0
.1
.2
.3
.4
.5
.6
.7
.8
.9
1.0
β 1

Figure (5.6.3)

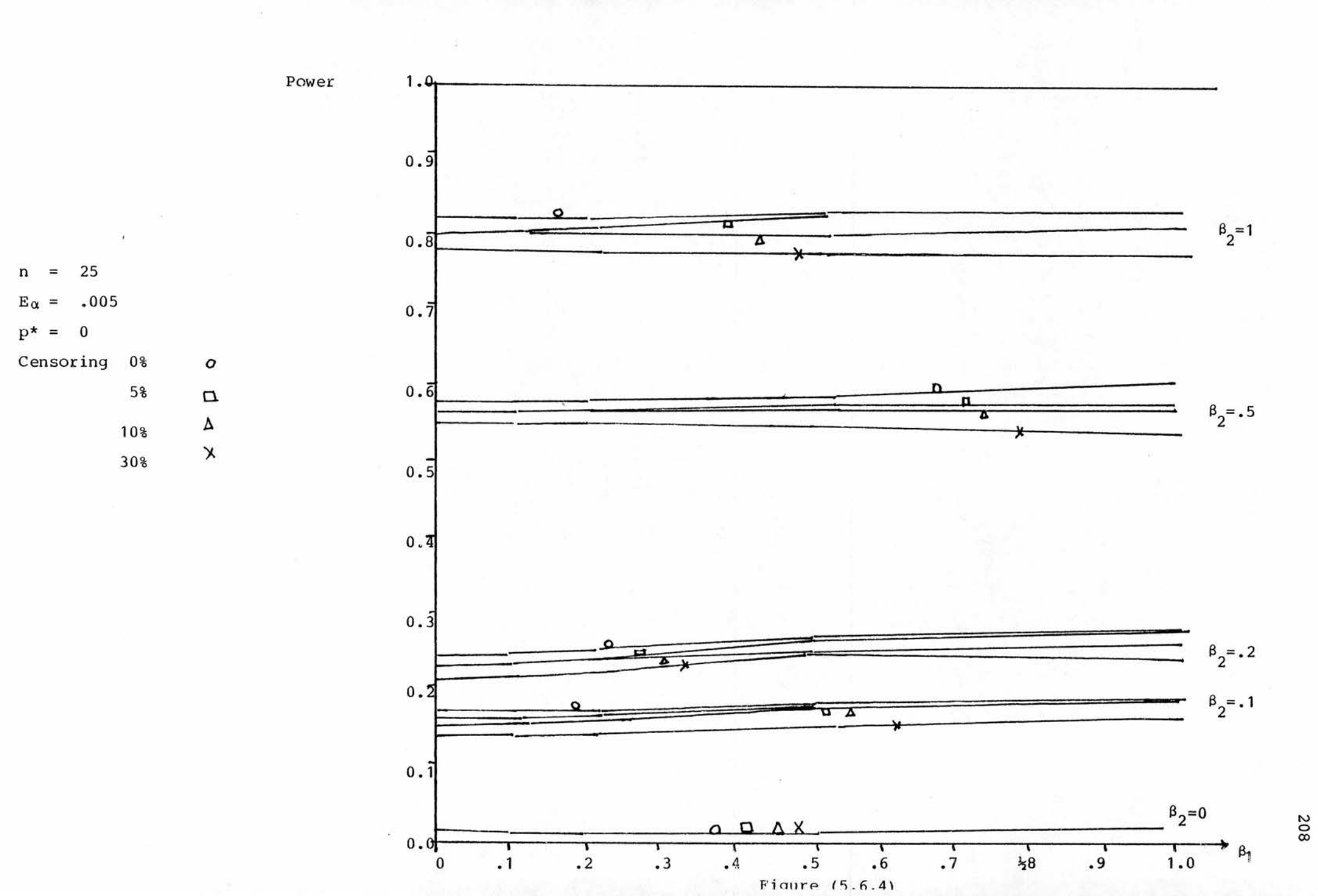
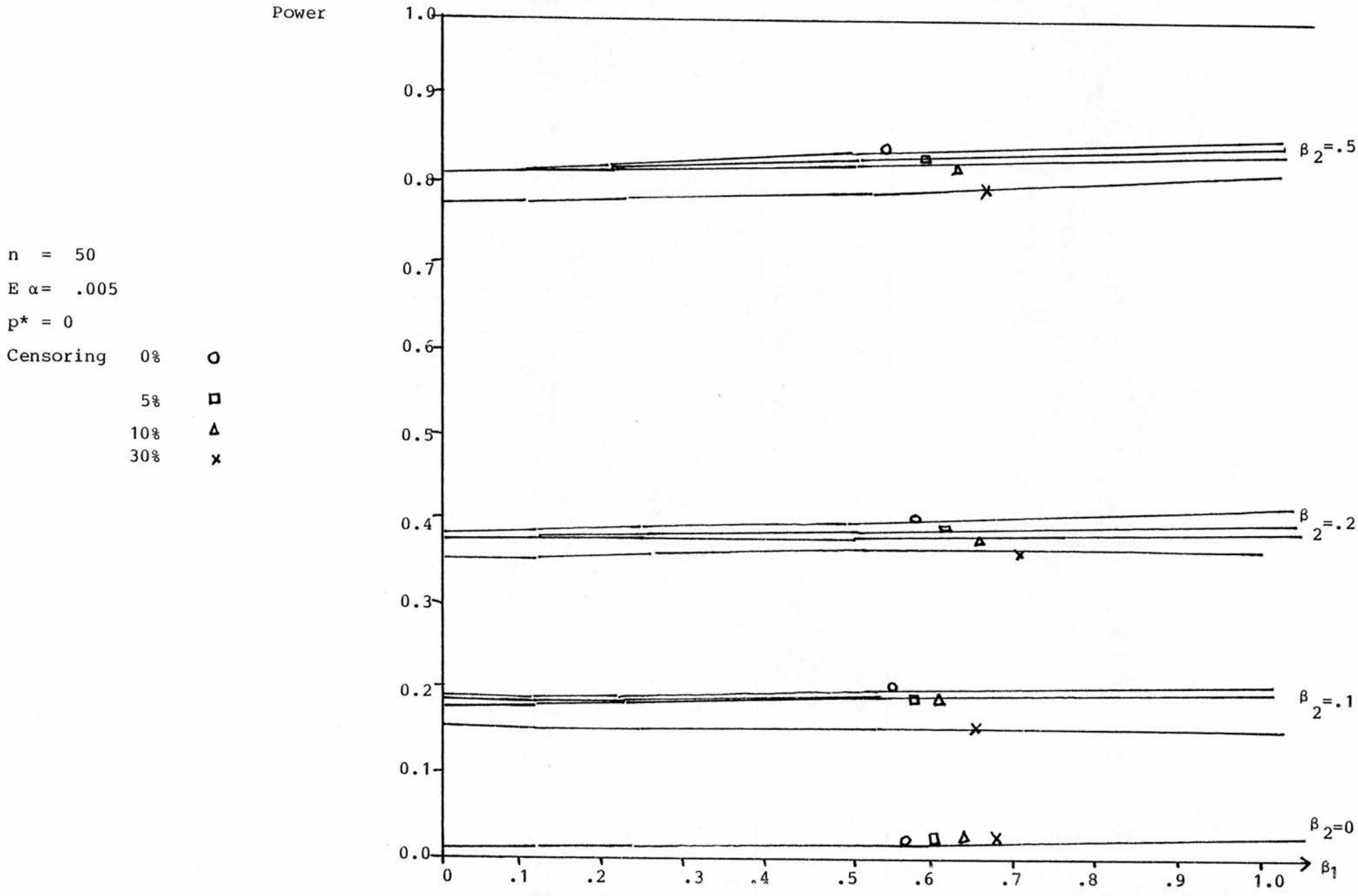

Power
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
n = 25
Eα = .005
p* = 0
Censoring 0%
5%
10%
30%
β2=1
β2=.5
β2=.2
β2=.1
β2=0
0
.1
.2
.3
.4
.5
.6
.7
⅛8
.9
1.0
β1

Figure (5.6.4)


Power
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0
.1
.2
.3
.4
.5
.6
.7
.8
.9
1.0
β1
β2=.5
β2=.2
β2=.1
β2=0
n = 50
E α= .005
p* = 0
Censoring 0%
5%
10%
30%

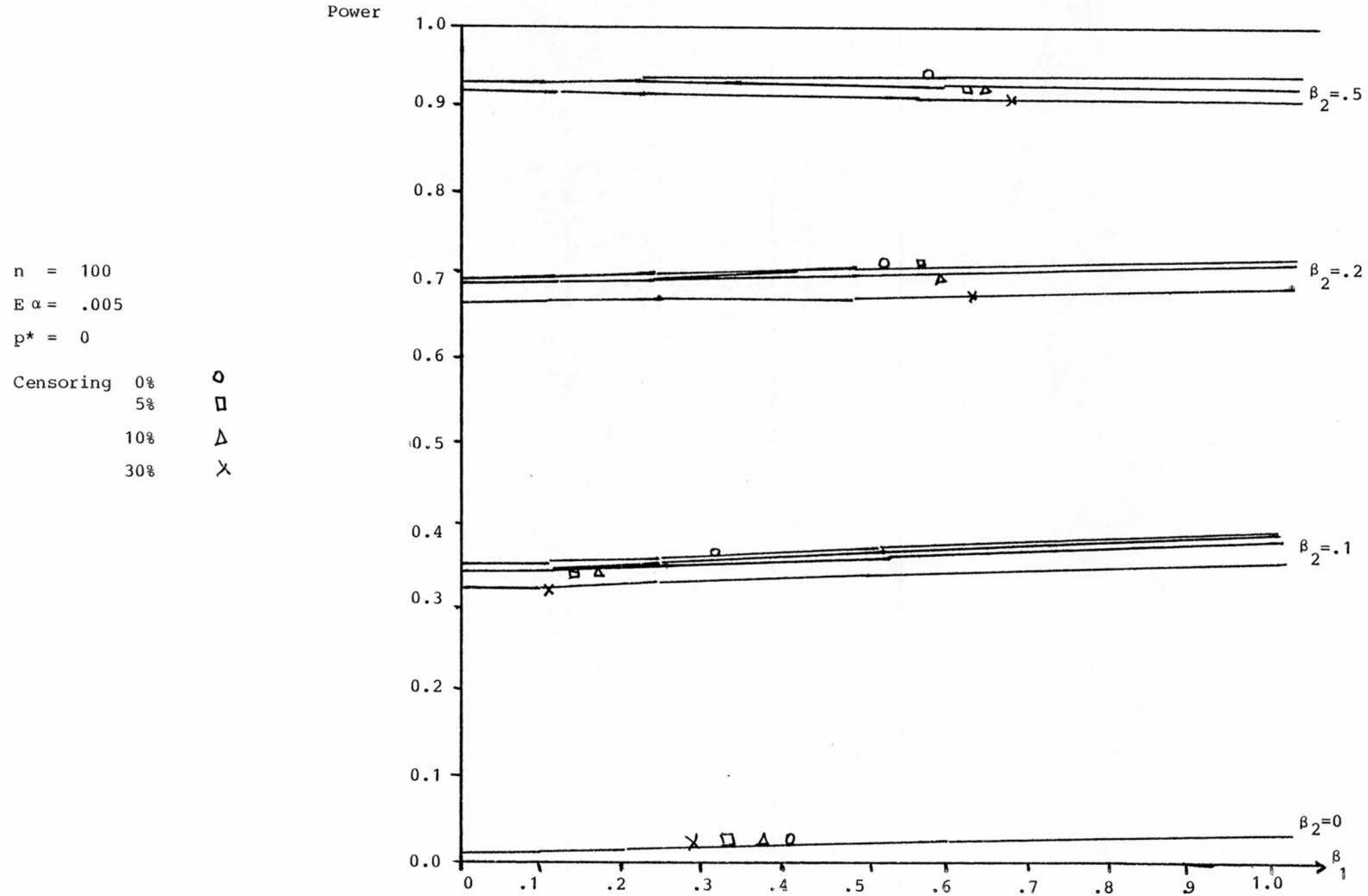

Power
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0
.1
.2
.3
.4
.5
.6
.7
.8
.9
1.0
β1
β2=.5
β2=.2
β2=.1
β2=0
n = 100
E α = .005
p* = 0
Censoring 0%
5%
10%
30%

Figure (5.6.6)

results of the above null hypothesis is the consistency of power values of Cox's test regardless of the value of the β_1 covariates. Clearly as we may expect the efficiency of the tests do deviate to some extent according to the value of sample size,censoring and the significance level, however, none of these factors seem to effect the lack of influence of β_1 covariates in the power of treatment effect tests. This finding is clearly in contrast with a view expressed by C.L. Chastang (1983) where it is reported that the efficiency of treatment effect is dependent on the value of prognostic effects, even when it is not included in the model. We will return to this hypothesis of treatment effect later in this chapter when we consider alternative parametric models in the study of proportional and non-proportional hazard distributed samples. Now we will consider the results of the tests of the simple treatment effect hypothesis in more detail.

At the value of $\beta_2 = 1$ and $n = 25$, $\alpha = 0.05$, we have a separation in censoring levels of almost 7% in power over the range of β_1 values, (Figure 5.6.1). An increase in the sample size to 50 diminishes the separation of the 0% censoring and 30% censoring (Figure 5.6.2). At $\beta_2 = 0$ we have a difference of 3% for the range of β_1 values. At $\beta_2 = .5$ we have a separation of almost consist- and 5% over the range of β_1 values.

A point to note is that the decline in the value of power of tests due to censoring seems to be affected by sample size to some extent. At the higher sample size of 100 the separation between the 0% censoring and 30% censoring at value of $\beta_2 = .5$ are almost 4% (Figure 5.6.3), while the same separation in $n = 25$ is 7%, Figure (5.6.1)

When/

When we consider a lower significance level of $\alpha = 0.005$, the separation between values of censoring levels declines so that at $n = 100$ and $\beta_2 = .5$ the 0% censoring and 30% censoring have a maximum separation of 3%, Figure (5.6.6). The same separation in power for n = 25, Figure (5.6.4) is 6%. Up until now we have dealt with simple tests of hypothesis, now we discuss a set of power curves for the composite test. The following simulations have a slight change of emphasis. The previous simulations asserted the power of tests for a practical assessment of the β_1 & β_2 values in a trial. What follows ia presented for theoretical interest and completeness. Tolley (1978) discusses a group of non parametric tests in survival analysis where a composite test of hypothesis may be of interest. Such tests have certain computational advantages when dealing with a large data set and a stepwise variable selection is adopted. The results of Tolley imply that that large sample distribution of a composite test has a chi-squared distribution with q degrees of freedom. The value of the test statistics is then given by:

$$Q_q = Q_{(r)} - Q_{(r-q)}$$

Where there are (r) concomitant variables in the fuller model and (r-q) in the simpler model. In a more complex hypothesis with Ho ; $C\beta = 0$ where C is a (q x r) contrast matrix the value of Q_q is then:

$$Q_q = U'(0)\, I^{-1}(0)\, C' \left[C\, I^{-1}(0)\, C' \right]^{-1} C\, I^{-1}(0)\, U(0)$$

where as in the notations of section 4.4 on Cox's method we consider U(0) as asymptotically normal with zero mean vector, covariance matrix I(0) and the test statistics;

$$U'(0) \left[I^{-1}(0) \right] U(0)$$

We will next present the results and show that the Cox's method has /

has good, predictable small sample properties where β_1 and β_2 are independent. In the simulation again we use a 50:50 allocation of Z_2 values and a 40:60 allocation of Z_1 values. Where Z_1 is taken to be a prognostic effect. One form of trial that has been used to some extent recently is based on 2 x 2 factorial trials. Such designs by randomisation will allocate 50:50 ratio to both Z_1 and Z_2 indicators. Although the ratio of the simulations in our results are different from the requirements of a 2 x 2 trial the good properties of the composite test efficiency may be attributed to the suitability of the Cox' s method when used for simple tests of 2 x 2 trials. Thus for more conclusive results in this respect, simple tests based on a symetric 50:50 proportion of binary variables of Z_1 and Z_2 are needed.

In the comparisons of the small sample properties of what follows we will use a few terms that need an explanation. The term maximum deviation is used when two power curves that are compared have a similar pattern and thus we only report the maximum deviation between the two graphs, since it is the most suitable descriptor . Relative efficiency is used when two single simulations are compared and is the difference between them. Finally we use the term balanced in situations where changes of censoring or sample size does not effect relative efficiency by a major degree.

Initially we concentrate on the situation of $\beta_1 > 0$ and $\beta_2 > 0$ with varying levels of significance of type I error, α, censoring levels and sample size. Later we will look at the situation of $\beta_1 < 0$ and finally at non-proportional hazards. In the final part we consider both/

both increasing and decreasing non proportionaity with positive and negative values of β_1 . When we discuss efficiency or power of tests it must be noted that in the interesting situtations we are dealing with tests that have power less than the full efficiency of 1. In the figures of power representation often a pattern of converging power curves appear. Up to a particular value of $(\beta_1\ \beta_2)$ set the tests deviate from the entity they are estimating. However at higher values of $(\beta_1\ \beta_2)$ or high sample sizes, the variability due to factors of interest like censoring and significance levels are not apparent due to the dominance of the covariate effects. Thus the efficiency values converge towards the maximum full efficiency of 1.00.

We first consider type I error which relates to the number of times the null hypothesis was rejected when it is true. Figure (5.6.7) to (5.6.12) contain such information for the proportional hazard cases. The generated samples conform to the α level probability limit of the type I error. Asympototic properties of the type I error are best summarised in the $\beta_1 = 0$ & $\beta_2 = 0$. At value of $\beta_1 = 0$ and $\beta_2 = 0$ we have a balanced configuration of the power curves in that we note by differing the value of censoring and significance level the power variability is small or nill for sample sizes at 25, 50 and 100.

We now consider in this paragraph simulations where $\beta_1 = 0$. From the no censoring to a 30% censoring with the sample size of 25 and the 5% significance level in the range of β_2 values the maximum deviation is a 5% loss. These results in fact complement the earlier results on the simple tests. Once again at a higher sample size or the

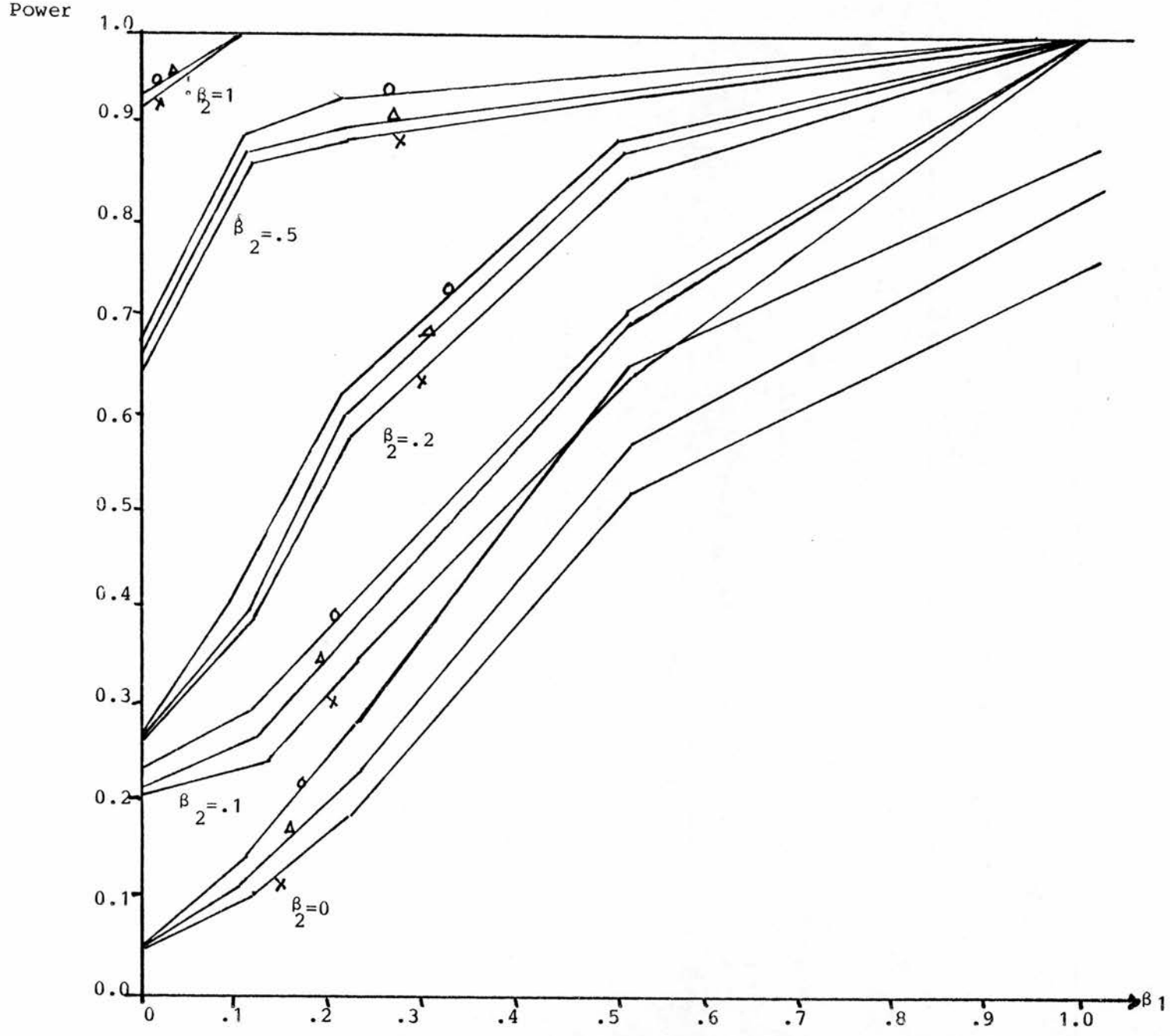

Power
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0
.1
.2
.3
.4
.5
.6
.7
.8
.9
1.0
β1
β2=1
β2=.5
β2=.2
β2=.1
β2=0
n = 25
Eα = .05
p* = 0
Censoring 0%
10%
30%

Figure (5.6.7)

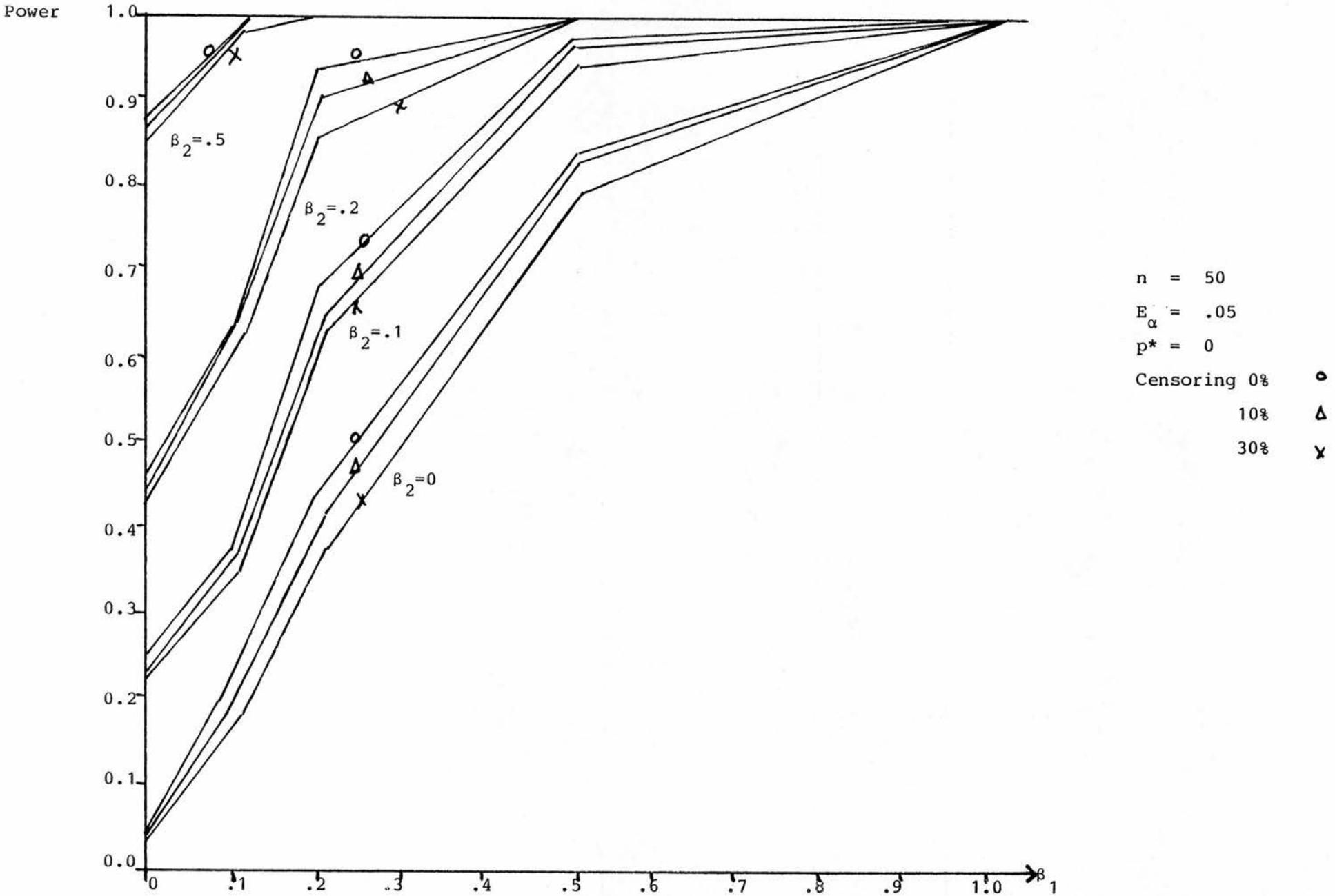

Power
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
β2=.5
β2=.2
β2=.1
β2=0
0
.1
.2
.3
.4
.5
.6
.7
.8
.9
1.0
β1
n = 50
Eα = .05
p* = 0
Censoring 0%
10%
30%

Figure (5.6.8)

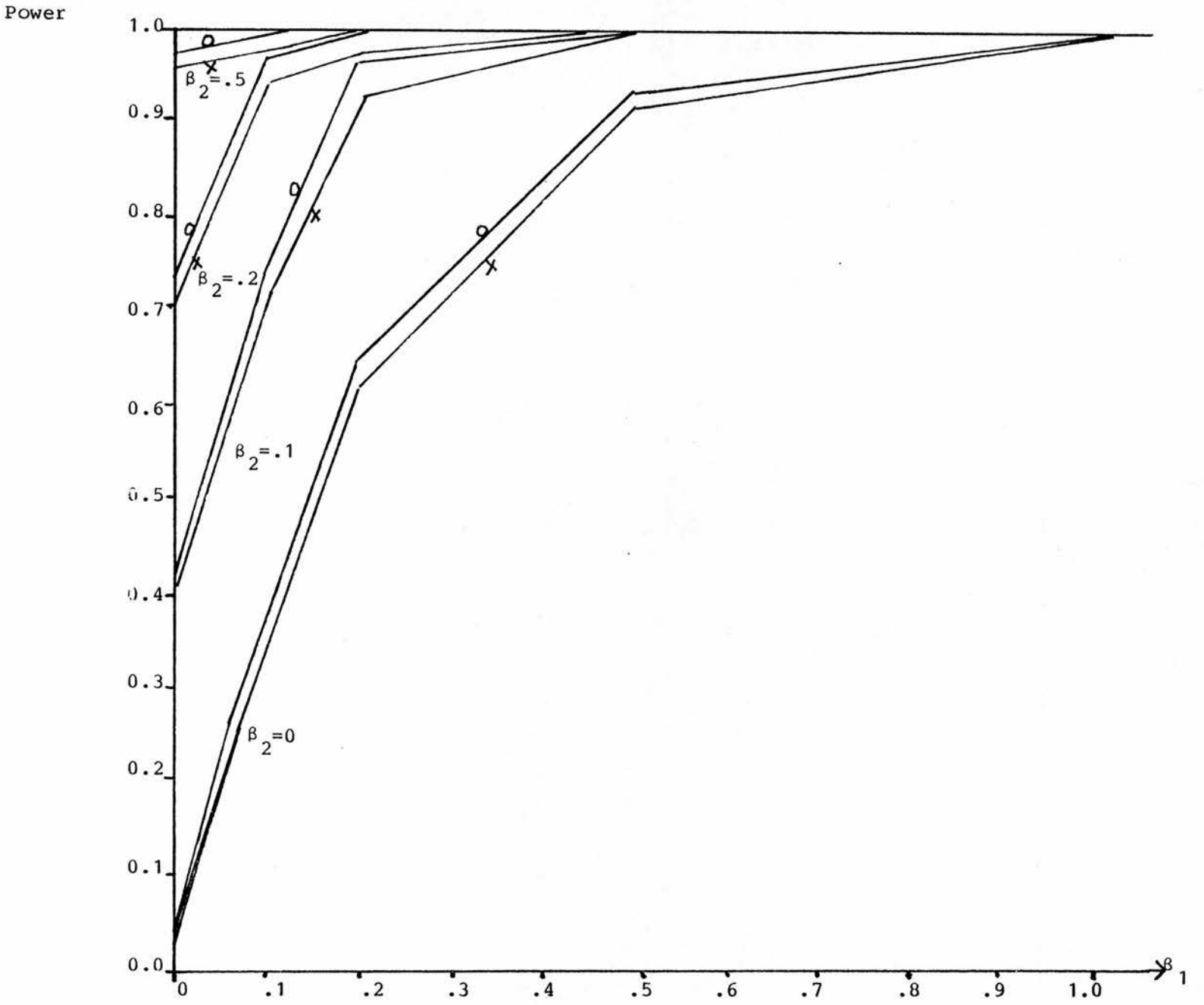

Power
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
β2=.5
β2=.2
β2=.1
β2=0
0
.1
.2
.3
.4
.5
.6
.7
.8
.9
1.0
β1
n = 100
Eα = .05
p* = 0
Censoring 0% o
30% x

Figure (5.6.9)

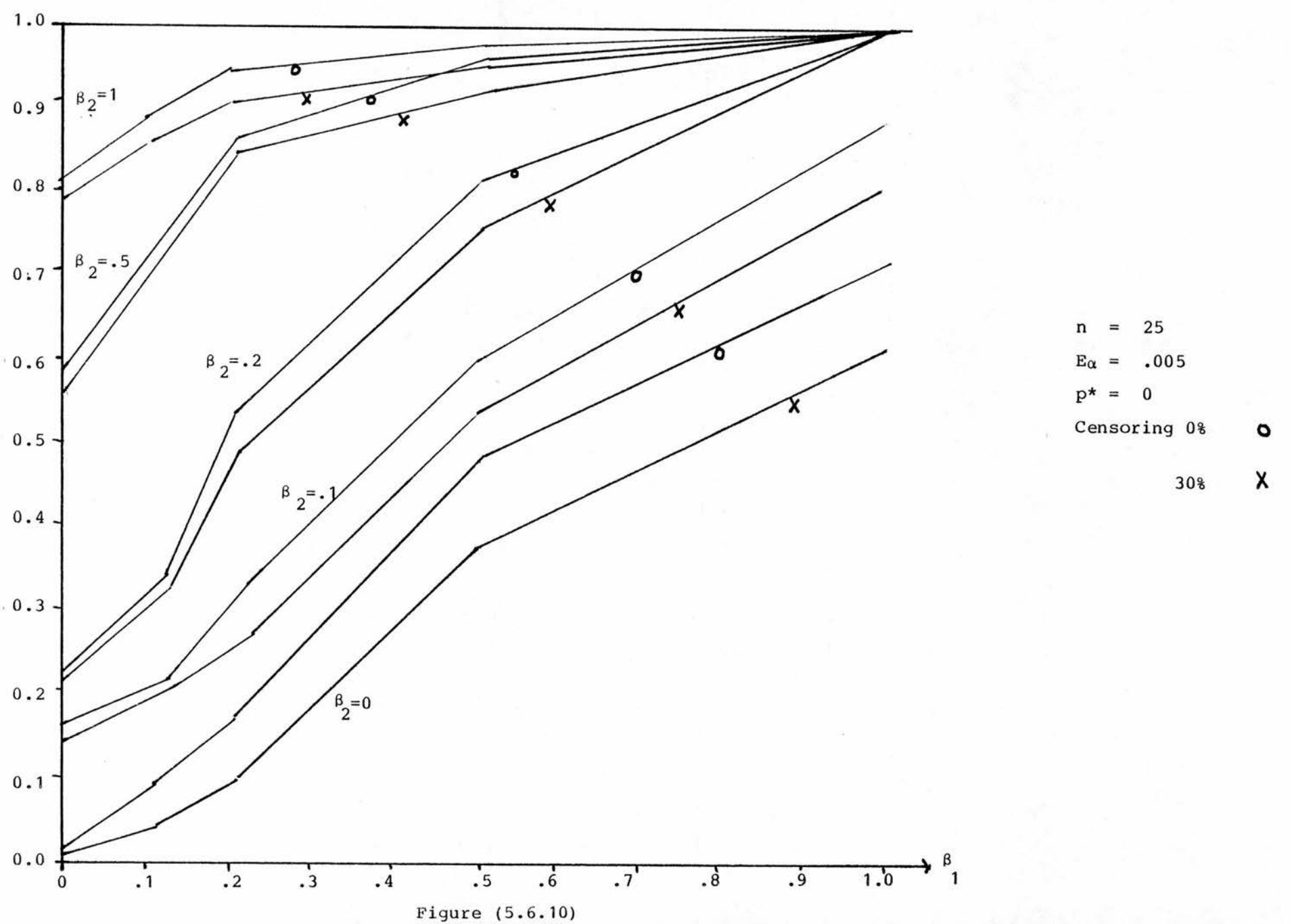
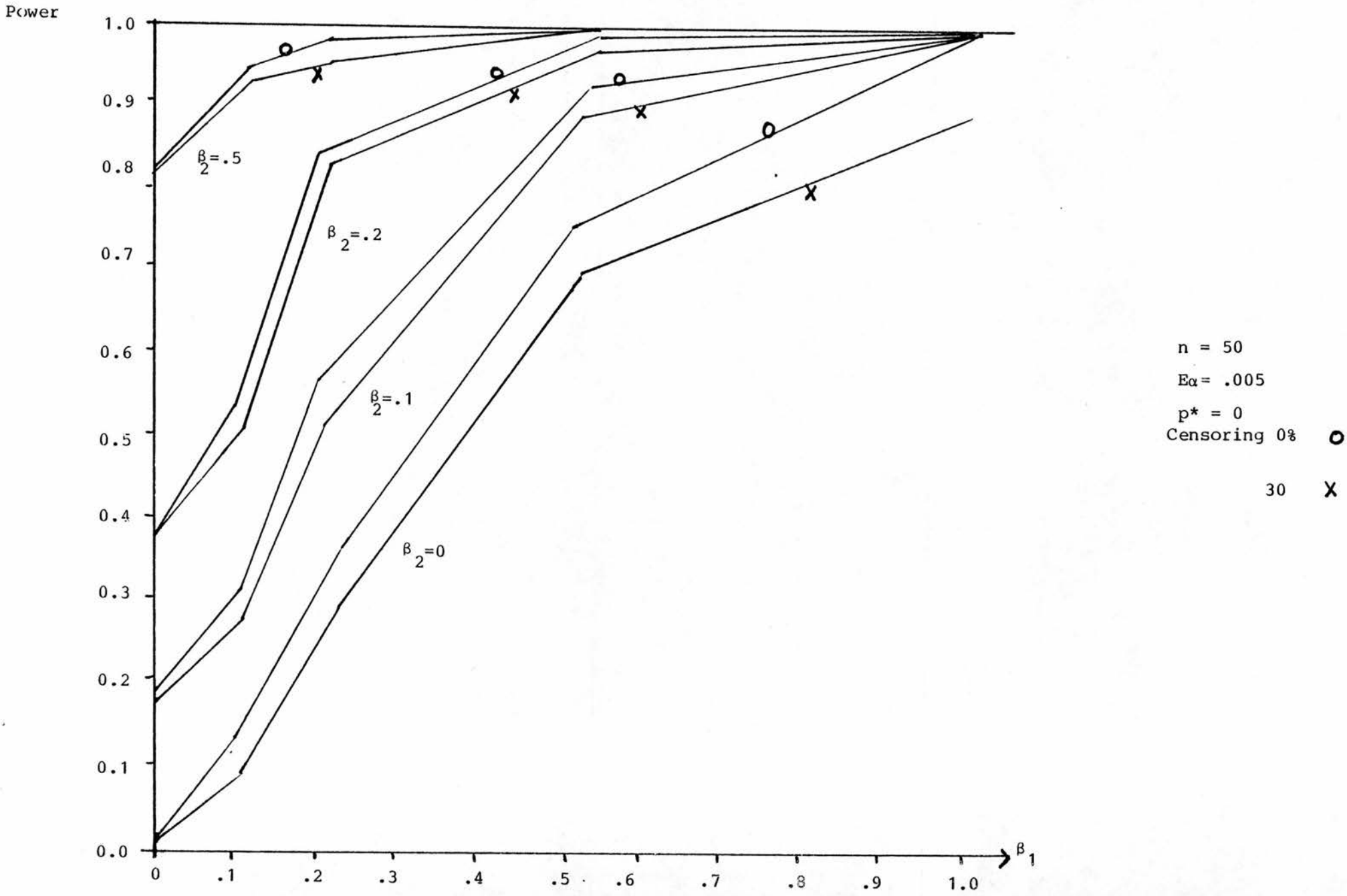

Power
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
β_2=1
β_2=.5
β_2=.2
β_2=.1
β_2=0
0
.1
.2
.3
.4
.5
.6
.7
.8
.9
1.0
β_1
n = 25
$E\alpha$ = .005
p* = 0
Censoring 0% o
30% X

Figure (5.6.10)


Power
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
β_2=.5
β_2=.2
β_2=.1
β_2=0
0
.1
.2
.3
.4
.5
.6
.7
.8
.9
1.0
β_1
n = 50
Eα = .005
p* = 0
Censoring 0% O
30 X

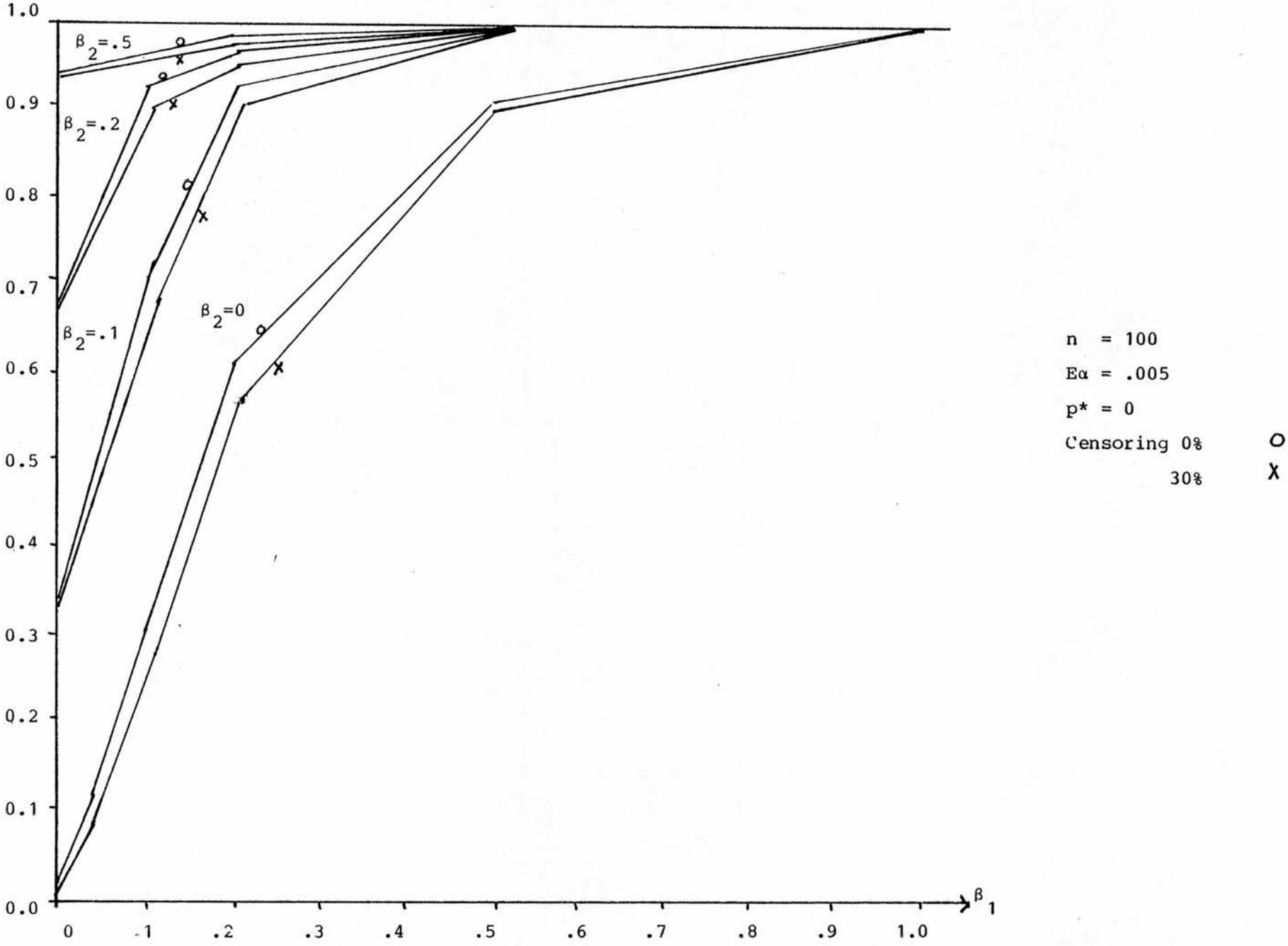

Power
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
β2=.5
β2=.2
β2=.1
β2=0
β1
0
.1
.2
.3
.4
.5
.6
.7
.8
.9
1.0
n = 100
Eα = .005
p* = 0
Censoring 0%
30%

Figure (5.6.12)

the significance level of $\alpha = 0.005$ the differences diminish.

For the above configuration the likelihood ratio test and the asymptotic normality give reasonably close efficiency values. The figure (5.6.13) presents these ratios for $\alpha = 0.005$ at sample size of 25 at β_1 and range of β_2 values. The maximum deviation between the two tests is 1% and it is at 30% censoring. (This is the only figure presented since other sample sizes and α- values do not produce figures different from the general pattern.)

Up until now we have been considering $\beta_1 = 0$ values. We now consider the changing of values of β_1 to .1, .2, .5 & 1.0 and repeat for each corresponding value of β_2. Clearly for $\beta_2^* > \beta_1^*$, there is a slightly higher small sample power distributed at (β_2^*, β_1^*) against (β_1^*, β_2^*) figure (5.6.7) to (5.6.13). This represents a slight lack of symmetry for the Z_1 and Z_2 ratios thus resulting in a higher relative power for β_2. This imbalance is most noticeable at the lower sample sizes and decreases with increasing sample sizes at 50 and 100. In the same figures we have also represented the different censoring values. As may be expected the value of small sample power decreases with increasing censoring levels. Although again by an increase in the sample size the effect of censoring is minimised. In general a decrease in the significance level also produces a reduction in power. Now with reference to the above figures we consider the magnitude of relative efficiency for β_1 & β_2 values. In general we note that by an increase in the β_2 values the efficiency increases for fixed values of β_1. On grounds of relative efficiency for a 25 sample size/

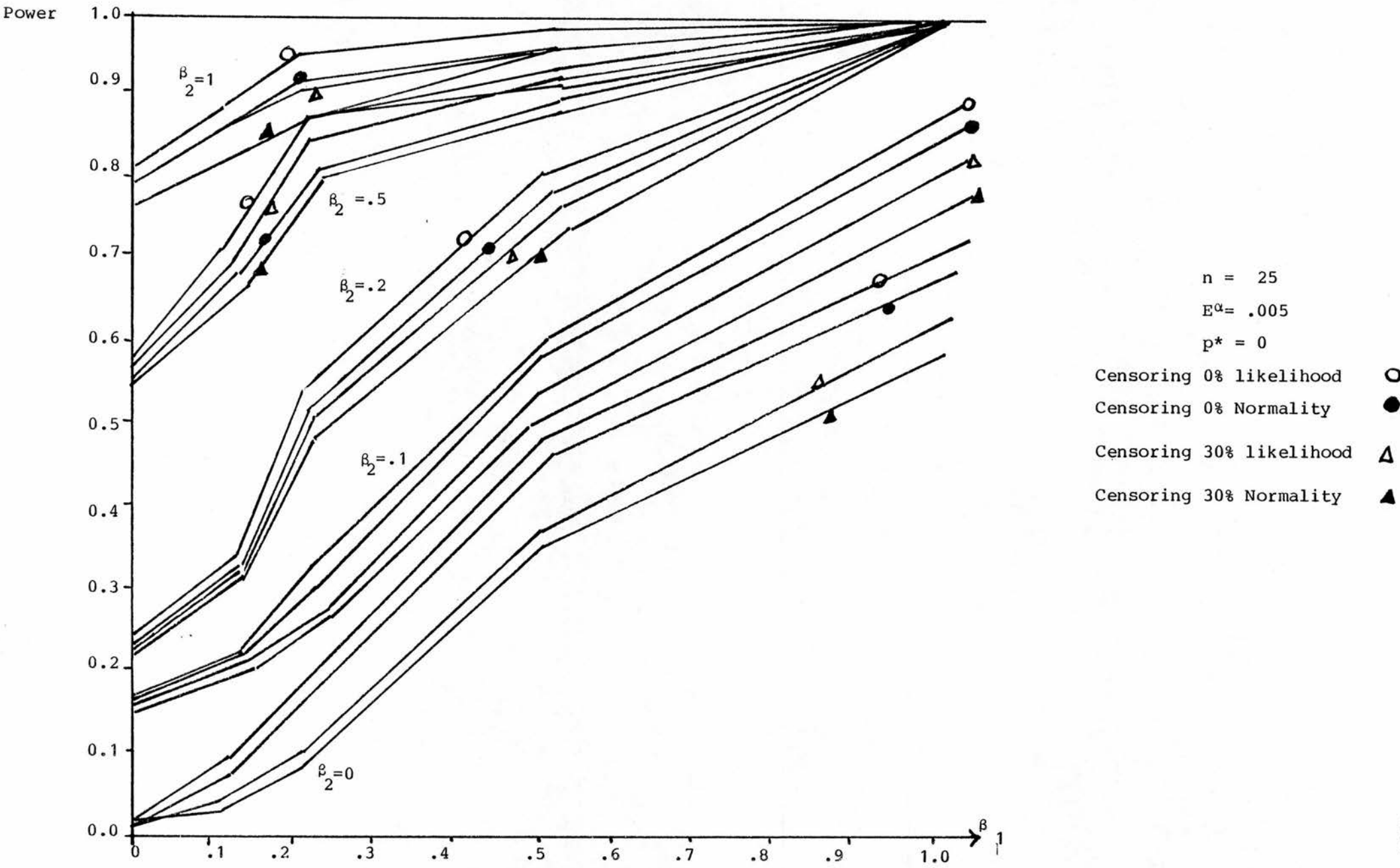

Power
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
β_2=1
β_2 =.5
β_2=.2
β_2=.1
β_2=0
0
.1
.2
.3
.4
.5
.6
.7
.8
.9
1.0
β
1
n = 25
E^{α}= .005
p* = 0
Censoring 0% likelihood
Censoring 0% Normality
Censoring 30% likelihood
Censoring 30% Normality

Figure (5.6.13)

size with censoring present we note that by an increase of $\beta_2 = 0$ to$\beta_2 = .1$, the relative efficiency between no-censoring and 30% censoring deviates from a 11% loss to a stable 4%. This 4% relative loss of power for censoring is in fact consistantly the same for higher values of β_2, figure (5.6.7).

For the sample size of 50 a value of relative loss of 5% occurss regularly for most values of β_1 and β_2 from 0 to 1, figure (5.6.8). The sample size of 100 gives a stable 2% loss of efficiency with 30% censoring, figure (5.6.9). Thus apart from the 30% censoring for the sample size of 25 with no covariate effects present, the loss of efficiency is very reasonable and at worst cases of the sample size of 25 a 30% censoring produces a relative loss below 10%. This 10% loss however will be discussed later and is far less for a balanced effect.

A point that may be made here is that if we consider 10% and 5% censoring we obtain stable losses of efficiency throughout simulation even at lower sample sizes. At the significance level of 0.005 there is a maximum efficiency loss of 12% at the sample size of 25, figure(5.6.10). This value does not stabilise and constantly diminishes reaching a value of 7% for the higher value of β_2. However at the sample size of 50 and the same significance level the loss in efficiency due to censoring stabilises at 5% for value of $\beta_2 > .1$ and at 8% for $\beta_2 = 0$, figure (5.6.11). For the sample size of 100 the loss in efficiency is a regular 4%. Thus once again reasonable efficiency losses are produced at 30% censoring, figure (5.6.12). The 5% and 10% censorings even with a significant level of 0.005 produces constantly very low efficiency losses. Once again we note that although the/

the 10% loss at 30% censoring is not problematic it is pointed out that some of this value may be attributed to lack of balance for the covariate effect.

We can thus summarise that in all sample sizes a reduction in the values of α reduces the power. In relative terms, the increase in power of the test due to increase in the sample size however is greater at low sample sizes, and at low type I error levels. Again in relative terms the higher censoring effects appear with lower sample sizes and low value of α. The relative difference due to significance level and censoring effects are thus minimised for sample size 100, within the range of our simulations. We also note that the improvement in the power of the tests from sample size of 25 to 50 is greater than the improvement from the sample size of 50 to 100.

As we pointed out earlier an imbalance has been introduced into the covariate effects. Thus the power of the tests are slightly different for (β_1^*, β_2^*) and (β_2^*, β_1^*) values. In other words a $(\beta_1\ \beta_2)$ value referring to a particular covariate and treatment effect does not refer to a set within which covariate and treatment effects have been exchanged. The same condition applies to the varying censoring levels and sample sizes. In the higher censorings and low sample sizes the effect of the difference in the comparable magnitudes of β_1 and β_2 appear more substantial. At the sample size of 50 and 100 the relative effects of censoring diminish substantially and the resultant loss of power from 0% to 30% censoring remains the same for (β_1^*, β_2^*) and (β_2^*, β_1^*). The lack of symmetry between the covariates has a resultant power difference of 20% for the/

the worst case of the sample size 25 at the significance level of 5%. This value diminishes for the sample sizes of 50 to 10% at the worst case and at 100 to bout 7% for the significance level of 5%. The relative difference in loss of efficiency for the 30% censoring is 10% for $(\beta_1, \beta_2) = (.5, 0)$ and 6% for $(\beta_1 \beta_2) = (0., .5)$ as the most extreme case for the loss in efficiency. At the sample size of 50 for the same significance level, censoring and (β_1, β_2) values we note 5% and 3% relative losses in efficiency.

At the significance level of 0.005 for the worst case at the 25 sample size and 30% censoring, we obtain a 12% loss in efficiency for the lack of balance in the worst case. Although at the extreme worst case the relative loss is the same for the two significance levels of 0.05 and 0.005, at the latter value the results are more regular and towards a higher range of magnitude. At the 50 sample size the results of 0.005 significance level is similar to the 0.05 level both in terms of the magnitude of the worst case and the regularity of the losses. At the 100 sample size we note a similar pattern as above in relation to loss in efficiency due to the 5% and 0.5% significance levels.

For the relative loss in efficiency of the above discussion, the worst case is the relative loss due to the significance level at 0.005 and 30% censoring giving a 11% loss at 25 sample size. This value is reduced to 5% and 4% losses for the samples of 50 and 100 respectively.

Referring back to the 20% imbalance of the covariate effect

z_1

Z_1 and fully balanced treatment effect Z_2, in relative terms at the 30% censoring level the loss in efficiency is more serious for the unbalanced variable than the balanced variable. In fact the 11% difference at the sample size of 25 diminishes to a reasonable 4% loss which is a stable loss for values of significance limit at 0.05 and 0.005. Although we have considered the only reason for the power differences of the above type to be those of the lack of balance in Z_1 under small sample properties, other investigations Kay(1979) have presented asymptotic results which are compatible with our findings.

In this study we also consider the effect of type of test used. That is the efficiency of the asymptotic normality against the asymptotic likelihood test. The difference between the tests again diminishes for the higher sample sizes and is most pronounced for the sample size of 25. Almost consistently the asymptotic normality turns out to be the more conservative test and this is true in particular as β_1 and β_2 deviate from zero. Censorings do not produce a major difference on the relative difference of the asymptotic likelihood and normality. In magnitude the maximum difference is at 6% for 30% censoring and sample size 25 and significance level of 0.005. All other relative variability of power between the two tests are less than this value. We will return to a comparison of the two tests under conditions that the proportionality of the hazard assumption is not valid, later in this section.

Up until now we have dealt with the generation of exponential samples in our simulations. Most of the results so far are in fact very closely in line with what may be expected of such simulations.

Next/

Next we will consider the generation of Weiball type of distributions. Initially we generate the samples in such a way so that the assumption of the proportionality of hazards is not violated. Later we will generate samples in which there is non-proportionality of the hazard present. As shown before we will use P to control the shape of hazard rates and for producing non-constant Weiball type hazard rates. Using

$$Y = \alpha + \partial w - Z_1\beta_1\partial - Z_2\beta_2\partial$$

The value of $\partial = P^{-1}$ has been fixed at $P = 1$ so far. Therefore all hazard rates have had constant rates for all subgroups. Now we will vary the value of P and proceed with the generation of samples of varying sizes that produce proportional hazards of Weiball type with increasing or decreasing hazard rates. By the definition of the proportional hazards, such effects should play a nominal role in the estimation of the β's. This is true mainly due to definition, that β's are estimated in terms of relative effects on subgroups. It is however known that in the estimation of partial likelihood in the treatment of ties and also the effect of censorings certain assumptions have been introduced. The results for ($P = 1.5$) increasing hazard and ($P = 0.5$) decreasing hazards are identical when there is no censoring. However with 30% censoring there was a slight deviation of 1 to 2 samples in 300 generations which is nominal. At $P = 0.5$, that is decreasing hazards, with rather high initial failure rate we may notice a larger number of failure times at zero, therefore the chances of producing tied observations at the beginning of survival times is higher and again there is a lower efficiency for these groups. Altogether all βsets that were tried, produced very close efficiency value of order of 3 in 300 generations in the extreme worst cases. Due to the close similarity of these results we will not produce

any/

any graphical presentations. However, we study two sample generators one at P = 0.5 (decreasing hazard) and one at P = 1.5 (increasing hazard).

Fix α at 0.0001, $\beta_1 = .2$, $\beta_2 = .1$ and no censoring n = 25 we obtain the following estimators

P = 1.5	p = 0.5
$\beta_1 = .233$	$\beta_1 = .232$
$\beta_2 = .139$	$\beta_2 = .138$
Var $(\beta_1) = .043$	Var $(\beta_1) = .043$
Var $(\beta_2) = .032$	Var $(\beta_2) = .032$
Lik $(\beta_1, \beta_2) = -16.85$	Lik $(\beta_1, \beta_2) = -16.92$
Lik $(\beta_1, 0) = -18.53$	Lik $(\beta_1, 0) = -18.59$
Lik $(0, \beta_1) = -17.47$	Lik $(0, \beta_2) = 17.51$

The above results clearly indicate very similar estimates for values of β. This close resemblance is mainly due to the non-parametric nature of the method. An interesting question however is related to the study of the effects of the covariates when the actual regression coefficients are time dependent. This effect can best be generated by allowing different subgroups of the patients to have different hazard rates.

In the study of the effects of non-proportionality of the hazards we continue with 2 covariate generating models. The effect of non proportionality can thus be more complicated in that it effects both β_1 & β_2 at similar times, simultaneously. We use the same model as before however the value of P is dependent on value of Z.

Hence

$$P \rightarrow \begin{cases} \neq 1 & \text{if } Z_1 \text{ or } Z_2 = +1 \\ = 1 & \text{if } Z_1 \ \& \ Z_2 = -1 \end{cases}$$

The value of P = 1 for all samples reduces to exponential decomposition of the hazard rates. The value of $P \neq 1$ is however important in that it indicates a measure of deviation from proportionality.

We repeat the simulation for similar ranges of β_1 & β_2 using the same hypothesis with the same sample sizes. This time the value of asymptotic normality and the asymptotic likelihoods are of special interest.

In the above we have assumed that time dependency is acting equally on the high levels of Z_1 & Z_2. This need not be the case in a more restrictive simulation study. One may allow time dependency to be an effect of one of the covariates. A usual manner of analysis is to stratify the data into early and late effects, and thus one produces two base line hazards for the population. In terms of a population with one time dependent effect the two strata should produce contours of the type in figure (5.6.14)in presence of the normality of their $\hat{\beta}$ estimates.

St1 and St2 refer to the two strata for early and late events when both β_1 andβ_2 are greater than zero. Although we have presented one figure with two different contour sets, it may be considered as two different figures for each strata when they are superimposed. For the cases of two time dependent covariates or a situation where time dependency is latent within the population the contour generated by our model may be represented as in figure (5.6.15). We have a continuous time dependent effect influencing covariates in both

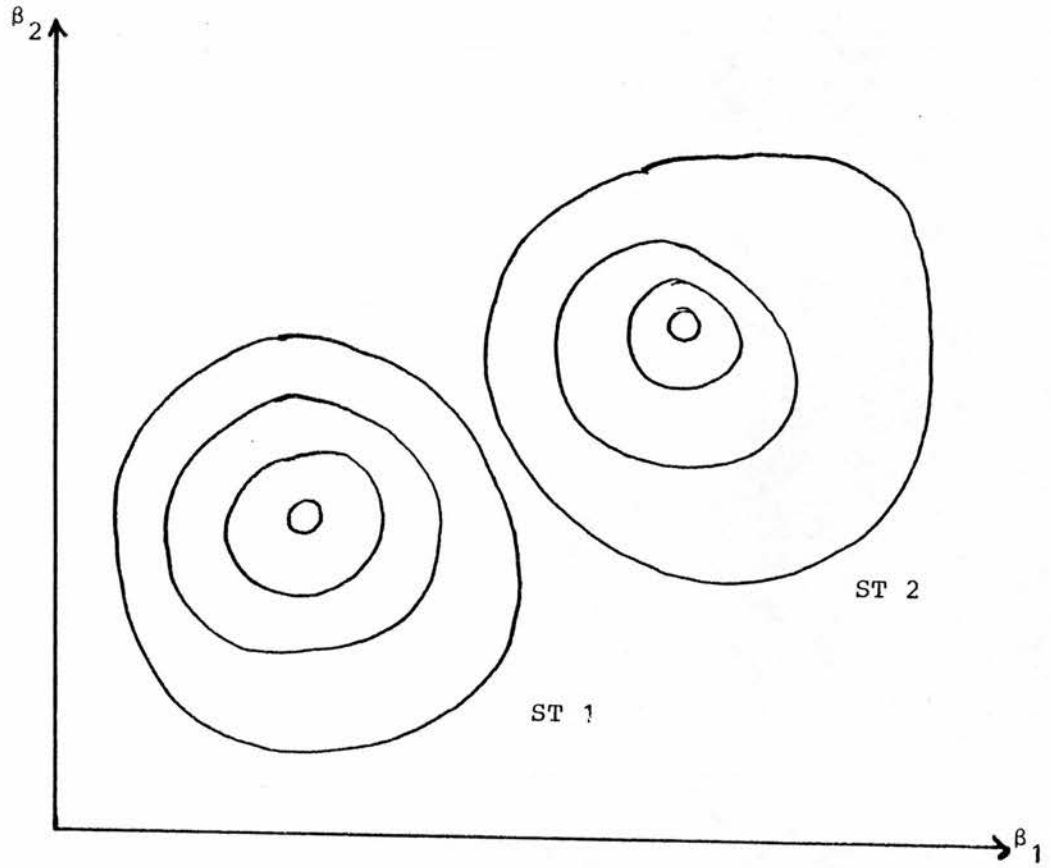


Figure (5.6.14) Example of β_1,β_2 contours with two separate strata in time.

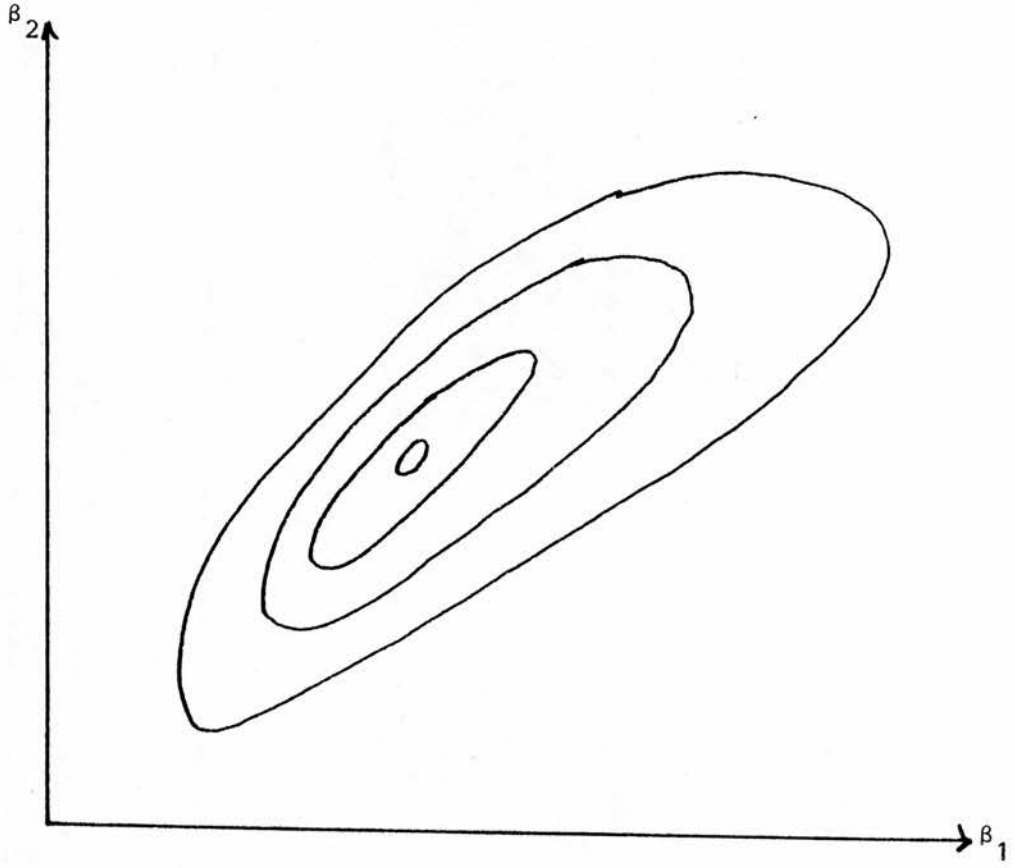


Figure (5.6.15) Example of β_1,β_2 contours with time dependency.

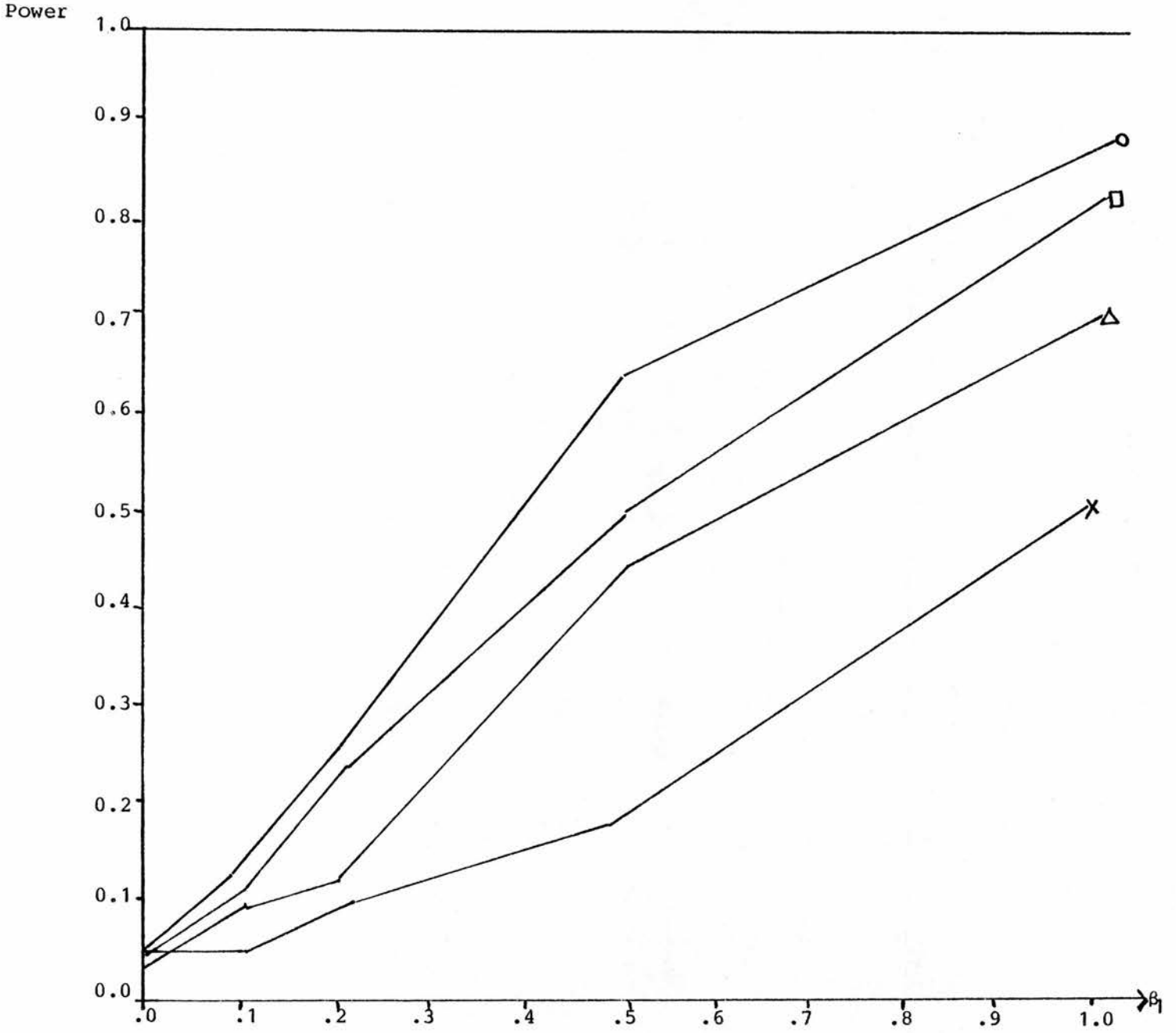
both β_1 and β_2 directions.

For reasons of dimensional symmetry we use a transformation of P to P* = log (P) so that the negative values of P* refer to decreasing hazards. We therefore use values of P* set to -.3, -.2, -.1, 0, .1, .2, .3. At first we consider the effect of

P* on β_1 & $\beta_2 \gtrless 0$ Figure (5.6.16) to (5.6.21).

Consistently we see a reduction in the type I error less than the actual level of significance level α . In relative terms the power decreases with increasing deviation from the proportionality of hazards. Using figure (5.6.16) and (5.6.17) we note that the reduction in type I error is to the reasonable numerical low level of 4% compared to the nominal 5% significance level, for the maximum deviation from proportionality at P* = .3 and P* = -.3. Further it is noted that at $\beta_1 = 0$ and $\beta_2 = 0$, there is not a reduction in the relative loss of power (in terms of the proportional to the non-proportional hazards) with an increase of the sample size, Figure (5.6.18) and (5.6.20). Thus indicating that the loss is due to the systematic effect of time dependency. For the lower values of increasing and decreasing non-proportionality at P* = -.2, -.1, .1 and .2, the reduction in efficiency is also within a range of 4%. In comparison of the relative loss of power for the corresponding magnitudes of the increasing and decreasing hazards once again we note a reduction in type I error of the increasing non-proportionality compared to those of decreasing non-proportionality rates. This effect is at 2% for the maximum difference of P* = .3 and P* = -.3.

Once/



n = 25

Eα = .05

No censoring

β_2 = 0

Non proportionality p* = 0 ○

p* =-.1 □

p* =-.2 △

p* = -.3X

Figure (5.6.16)

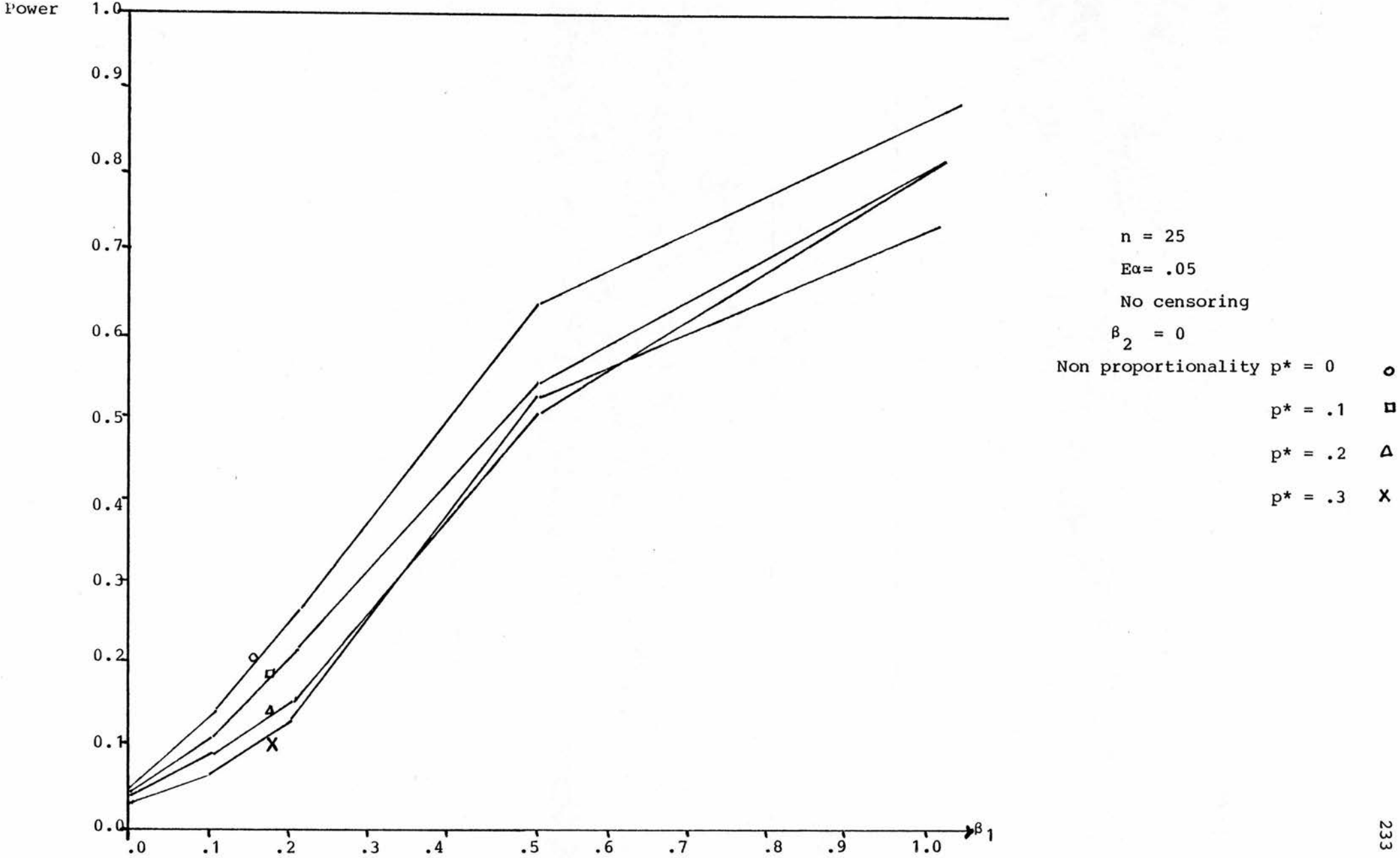

Power
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
.0
.1
.2
.3
.4
.5
.6
.7
.8
.9
1.0
β_1
n = 25
Eα= .05
No censoring
β_2 = 0
Non proportionality p* = 0
p* = .1
p* = .2
p* = .3

Figure (5.6.17)

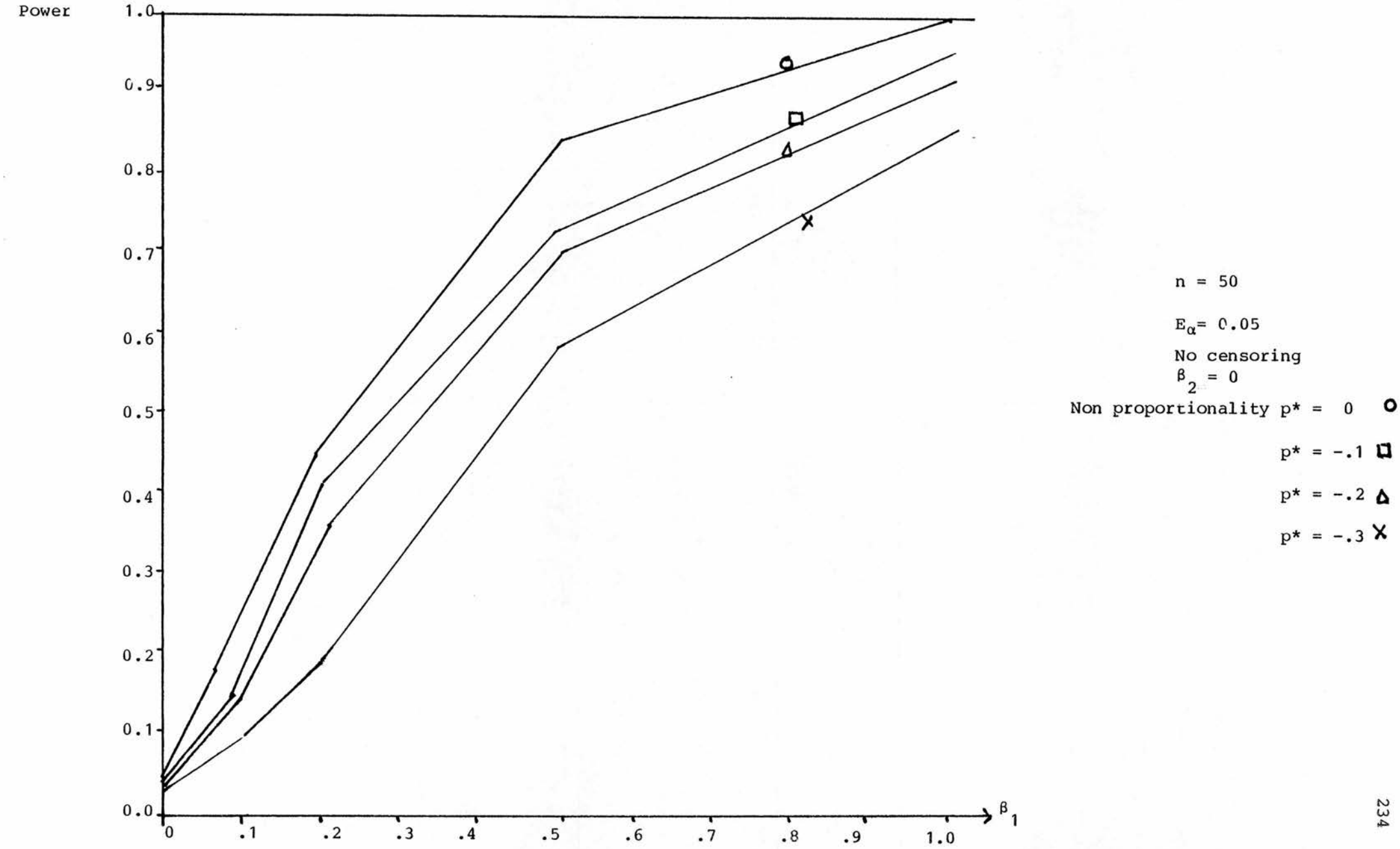

Power
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0
.1
.2
.3
.4
.5
.6
.7
.8
.9
1.0
β_1
n = 50
E_α= 0.05
No censoring
β_2 = 0
Non proportionality p* = 0
p* = -.1
p* = -.2
p* = -.3

Figure (5.6.18)

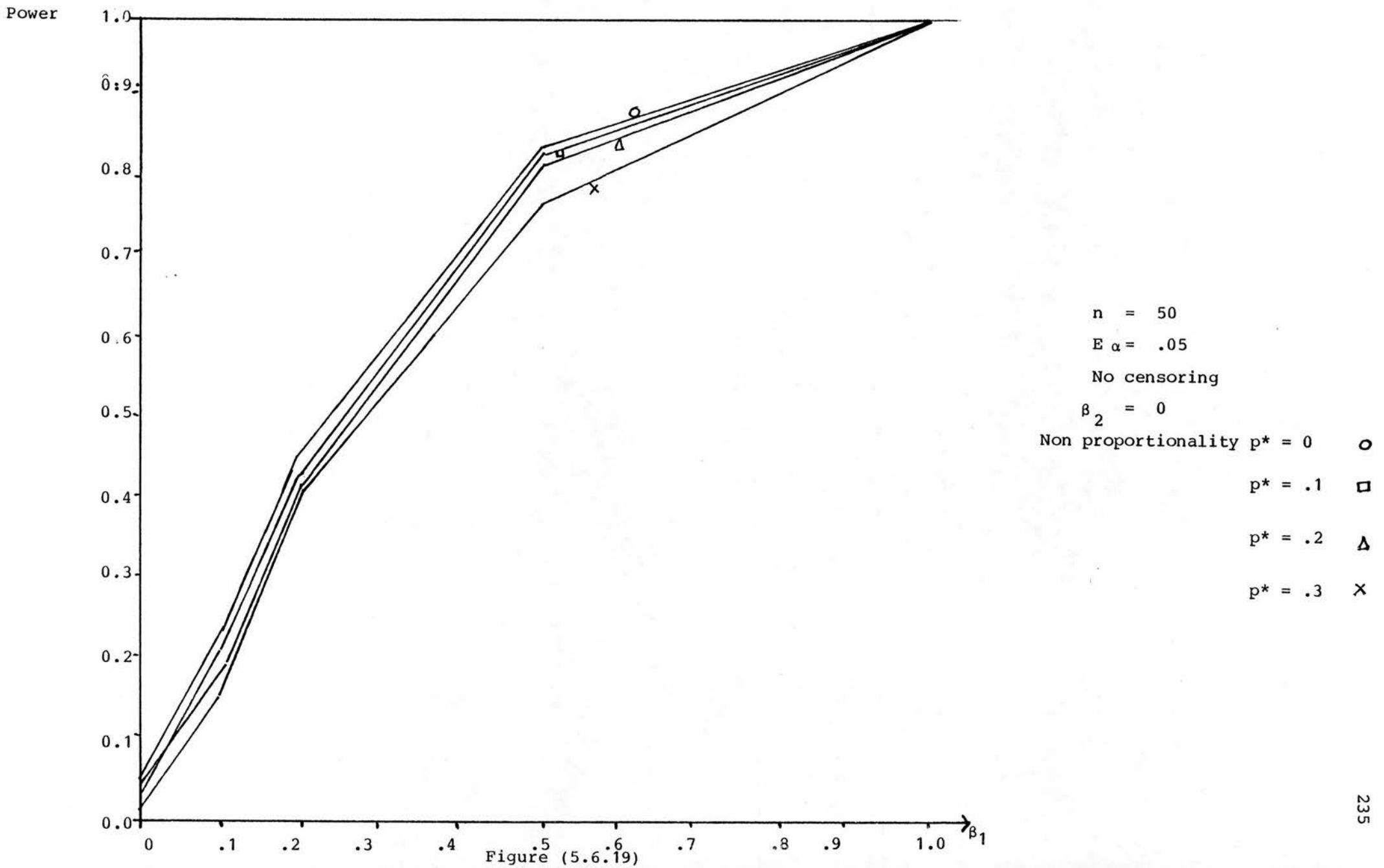

Power
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0
.1
.2
.3
.4
.5
.6
.7
.8
.9
1.0
β_1
n = 50
E α = .05
No censoring
β_2 = 0
Non proportionality p* = 0
p* = .1
p* = .2
p* = .3

Figure (5.6.19)

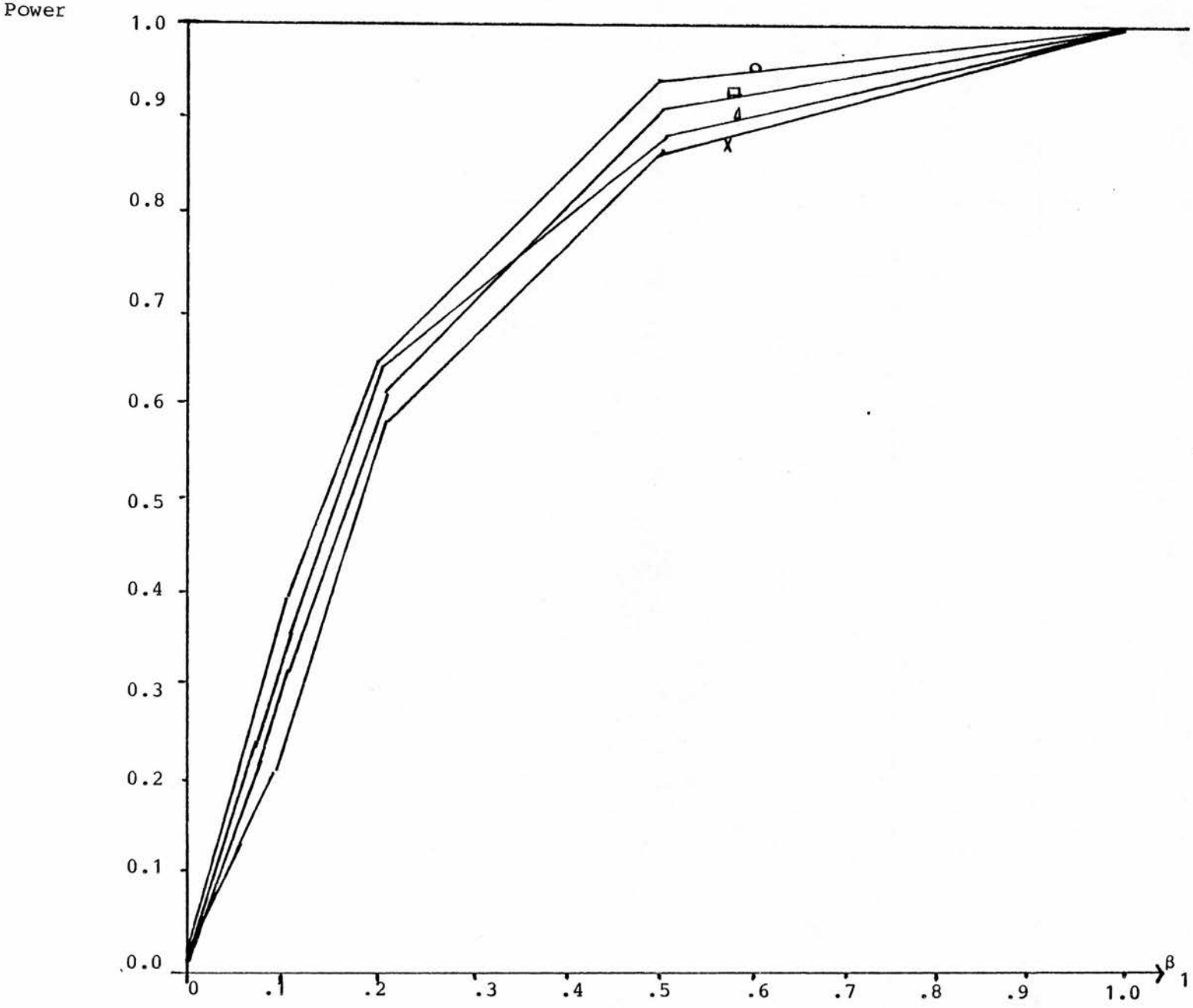

Power
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0
.1
.2
.3
.4
.5
.6
.7
.8
.9
1.0
β1
n = 100
Eα= .05
no censoring
β2 = 0
Non proportionality p* = 0 o
p* = -.1 □
p* = -.2 Δ
p* = -.3 X

Figure (5.6.20)

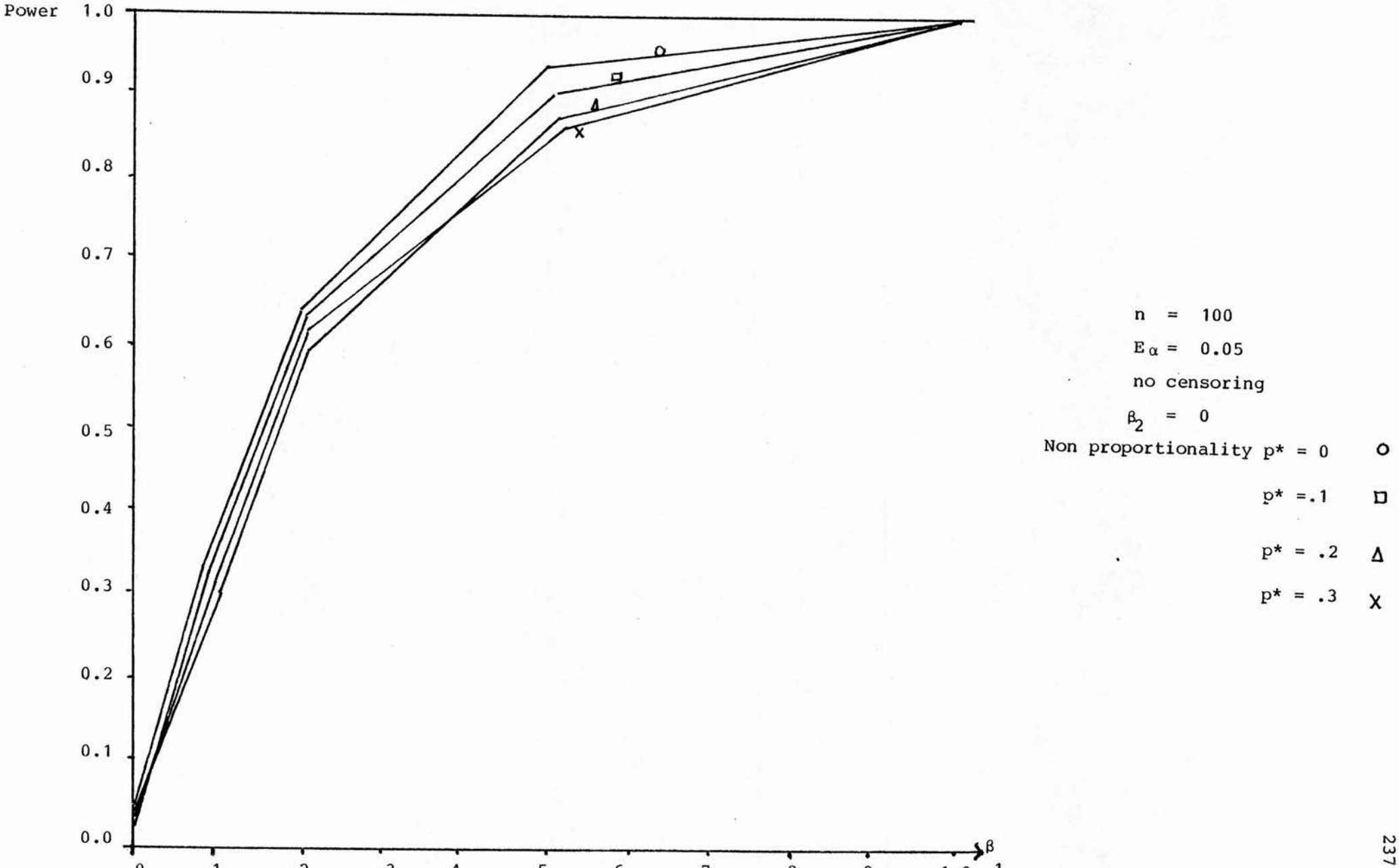

Power
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0
.1
.2
.3
.4
.5
.6
.7
.8
.9
1.0
1
β
n = 100
Eα = 0.05
no censoring
β2 = 0
Non proportionality p* = 0
p* =.1
p* = .2
p* = .3

Figure (5.6.21)

Once again we note the asymptotic likelihood and normality at $(\beta_1, \beta_2) = (0\ 0)$, which indicates a measure of type I errors. The relative loss of asymptotic normality to asymptotic likelihood is 2% at the sample size of 25 with the maximum reduction due to non-proportionality at $P^* = -.3$, Figure (5.6.23). For the sample size of 50, Figure (5.6.24) this relative deviation of type I error reduces to just under 1%, which is at a similar level to the relative difference of type I error for the proportional hazard rates. We thus conclude that the relative difference in type I error in situations of non-proportionality of hazards at the sample size of 25 is at a low value of approximately 2% and the relative loss reduces to those of the proportional hazard situation as the sample size increases to 50.

We continue with the simulations by letting the β_2 value for treatment effect be at zero and the β_1 covariate effects vary over a range of values .1, .2, .5 and 1.0. We note that the non-proportionality of hazards at the sample size of 25 with significance level at 5% produces a maximum loss of power of 25% with decreasing non-proportionality rate of $P^* = -.1$. This loss is at a reasonable 5% level for decreasing rates of $P^* = -.3$. An increase of the sample size to 50 reduces this relative loss for the maximum decreasing rate to 15%. At the sample size of 100 the relative loss of power at $P^* = -.3$ reduces to a reasonable value which is less than 6% and this value is stable for a range of β_1 values at .1,.2 and .5. We thus state that decreasing hazards do lead to a loss of power and the magnitude of this loss at $\beta_2 = 0$ is dependent to some extent on the values of β_1. This effect of dependence of β_1 values on the non-proportionality rates reduces to nominal levels at 50 and 100 sample sizes. For the sample size/

of 25 however, with non-proportionality at a decreasing rate of $P^* = -.3$ and the covariate β_1 effect at a high value of 1 the loss in efficiency is unacceptable.

In terms of increasing hazards again we obtain the same results of loss of power. However there is less relative loss of power compared with decreasing deviations from proportionality (negative value of P*). In considering various sample sizes once again the same conclusions may be drawn. In fact the magnitude of increasing hazard rate at $P^* = .3$ produces very stable values of relative loss of power at less than 6% for the sample size of 50 and lower values for the sample size of 100. At the sample size of 25, the maximum variability due to increasing non-proportionality rate is due to $P^* = .3$ and the value of β_1 at its maximum 1.0. Such an effect in terms of decreasing rates was noted to be 25% and judged unacceptable but now it is reduced to 12%. Although the value of sample size at 50 and 100 gives reasonable values of power loss due to non-proportionality of decreasing and increasing types, extreme caution is needed for sample size of 25 in presence of decreasing non-proportionality rates and higher value of covariate effect β_1.

Up until now we have considered the effect of β_1 values on power for the different sample sizes, now we continue with a few words on the relative gain in power by increasing sample sizes. An increase in sample size of 25 to 50 gives a maximum gain of 14% in relative efficiency of the proportional hazard rates. Change in sample size from 50 to 100 gives an increase of less than 8% in relative power of proportional hazard rates. Clearly the increase in sample size plays/

plays a more significant role for the non-proportional hazard rates. At the maximum non-proportionality of P* = -.3 an increase of 27% is obtained at covariate value β_1 set to 1. This improvement is important in the sense that at higher values of β_1 the loss of power due to non-proportionality reduces to an acceptable level. For the sample size of 50 t0 100 for a similar generation of β_1 and β_2 values we obtain a gain of 13% in power. However this is achieved at a point in which the sample size of 100 with P* = -.3 and β_1 = 1 has the full maximum efficiency at 1.0.

In the study of non-proportionality we observe that Exp(-.3) = P, produces the best representation of the results for β_1 and β_2 values and thus we continue with this sample for values of $\beta_2 > 0$ and $\beta_1 < 0$, using an analgous one sided alternative hypothesis. Figures (5.6.22) to (5.6.25).

First we deal with maximum likelihood estimator with constant hazard rates. At the constant hazard rate the relative efficiency is clearly symmetric about $\beta_1 = 0$, Figure (5.6.22). The figure (5.6.23) presents a decreasing hazard rate and there is clearly a lack of symmetry about $\beta_1 = 0$. As we showed earlier values of $P < 1$ produced a larger level of variability than $P > 1$ hazard rates and now we note that there is a lack of symmetry about $\beta_1 = 0$. The following figure can describe what is happening in terms of converging or diverging forms of the proportionality of the hazards.

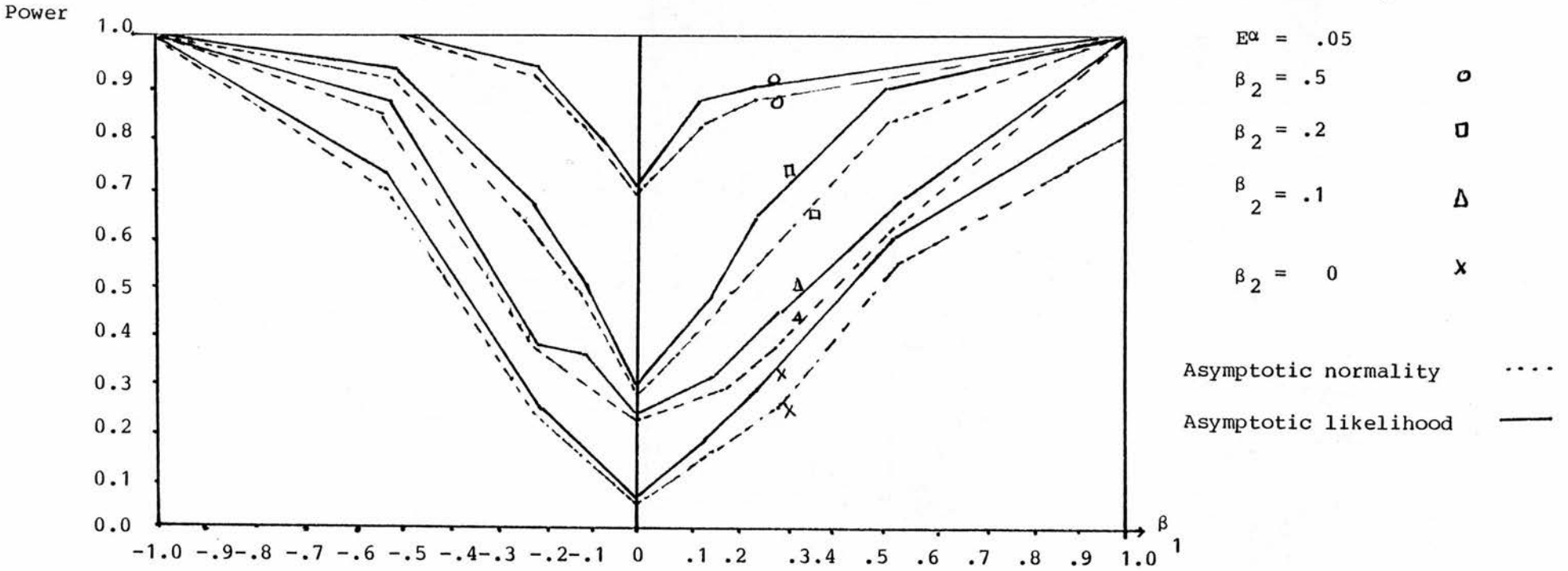

Power
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
-1.0 -.9 -.8 -.7 -.6 -.5 -.4 -.3 -.2 -.1 0 .1 .2 .3 .4 .5 .6 .7 .8 .9 1.0
β1
n = 25
p* = 0
no censoring
Eα = .05
β2 = .5
β2 = .2
β2 = .1
β2 = 0
Asymptotic normality
Asymptotic likelihood

Figure (5.6.22)

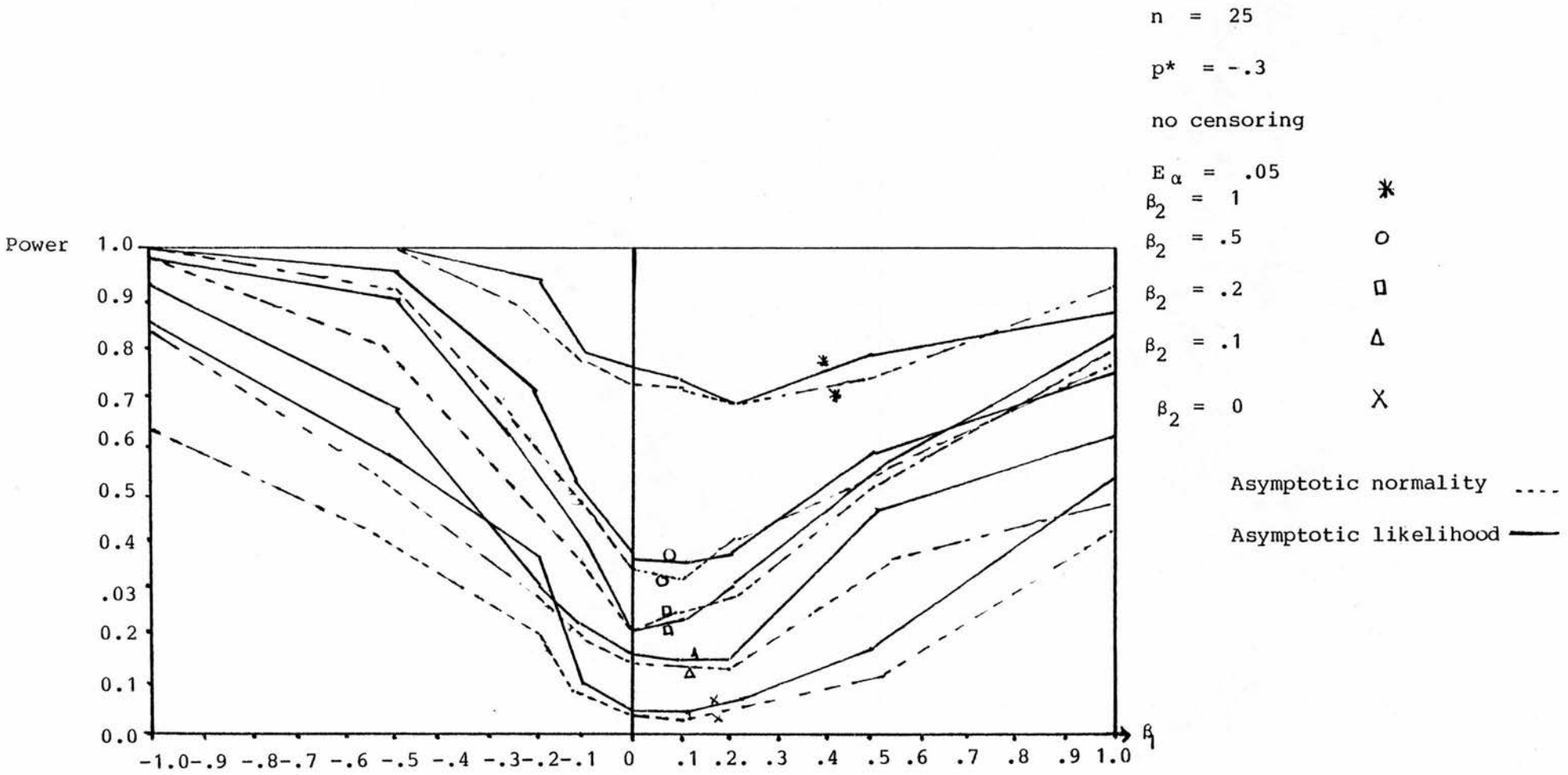
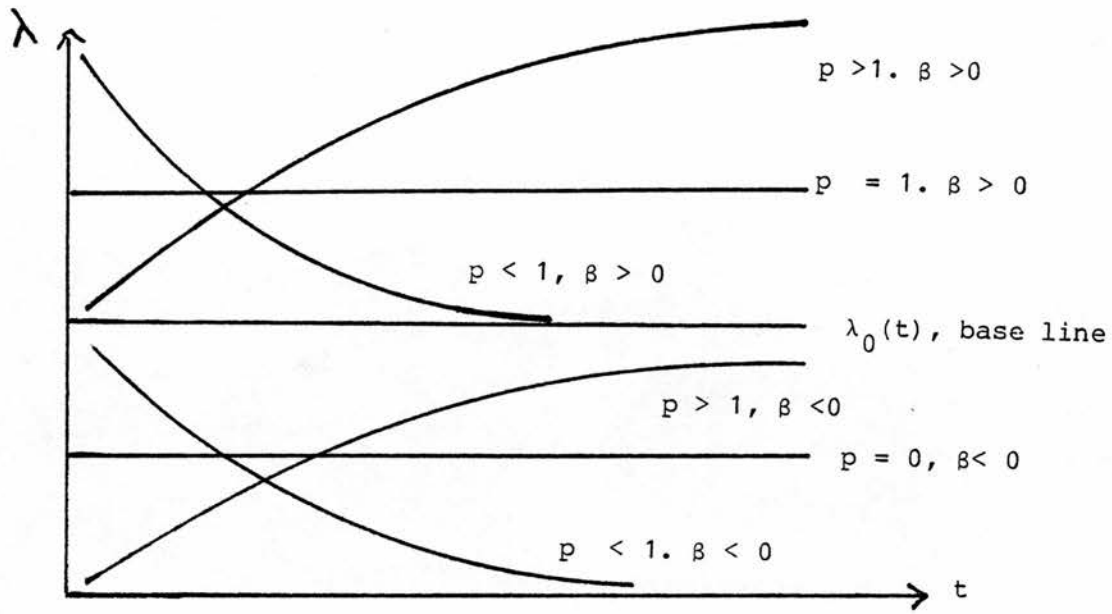

n = 25
p* = -.3
no censoring
E_α = .05
β_2 = 1
β_2 = .5
β_2 = .2
β_2 = .1
β_2 = 0
Asymptotic normality
Asymptotic likelihood
Power
1.0
0.9
0.8
0.7
0.6
0.5
0.4
.03
0.2
0.1
0.0
-1.0 -.9 -.8 -.7 -.6 -.5 -.4 -.3 -.2 -.1 0 .1 .2 .3 .4 .5 .6 .7 .8 .9 1.0
β_1

Figure (5.6.23)



The above figure presents increasing, decreasing and constant Weiball hazard rates together with positive and negative β values. In the discussions of figures (5.6.16) to (5.6.21) we presented results in which for positive values of β, the $P > 1$ simulations were more stable and efficient than $P < 1$ values. Using the above figure we note that for $\beta > 0$, $P > 1$ implies diverging hazards, while $P < 1$ implies converging hazards. In the describtion of figures (5.6.22) and (5.6.23) again the above figure can help, in that for $P < 1$ we note that $\beta < 0$ (diverging hazard) compared to $\beta > 0$ (converging hazard) produces high efficiency. That is in either case as may be expected divergence from the base line hazard produces higher efficiencies.

We once again observe that due to the imbalance of 20% for the covariate effect, corresponding values of power efficiency of (β_1, β_2) and (β_2, β_1) deviate slightly from each other. For the various values of treatment effect β_2 and covariate $\beta_1 > 0$ we note that there are important losses of power with the magnitude as high as 28% for n = 25.

In/

In the previous discussions of non-proportionality with $\beta_2 = 0$ a high efficiency loss in fact presented stable levels of efficiency, and often stable levels of efficiency for the ranges of β_1 & β_2 corresponded with low efficiency levels. In here as we pointed out in the discussion of non-proportionality at sample size of 25 and $\beta_2 = 0$ extreme caution is needed if ever used in practice. However there is a pattern emerging from the sample siz e of 25 simulations which represent the relative efficiency in terms of variability of β_1 and β_2 and P*, Figure (5.6.23)

The worst region in terms of relative loss of power between figures (5.6.22) and (5.6.23), is due to the covariate effect values of β_1 between 0 and .2 for β_2 values greater than 0.2, and β_1 values between 0 and 1 for β_2 values less than .2. The losses for the earlier group ranges with 15% efficiency loss and the latter group produces efficiency loss of 28%

For the negative values of β_1 we observe a pattern is emerging indicating that for values of $\beta_1 < -.2$, the relative loss in efficiency in comparison to the proportional hazard samples is decreasing steadily. In fact between β_1 values of -.1 and 0 for each particular level of β_2, the difference in efficiency is within 24% , while outside of this range of generations of β_1 values the efficiency is closer to the proportional hazard situation.

Thus we summarise that with the sample size of 25, there is a loss of symmetry Figure (5.6.23). This relative difference for $\beta_1 < 0$/

$\beta_1 < 0$ and $\beta_1 > 0$ is restored with an increase in the sample size. It must be noted that $\beta_1 > 0$, $\beta_2 > 0$ with $P^* > 0$ is not comparable with $\beta_1 < 0$ and $\beta_2 > 0$ and $P^* < 0$.

We now proceed with a similar simulation as those of sample size 25 with the non-proportionality set to $P^* = -.3$. However this time we increase the sample size to 50 and 100 to study the effect of deviation from non-proportionality at higher sample sizes, Figure (5.6.24) and (5.6.25). The first point we note is that there is less asymmetry about the $\beta_1 = 0$ axis when the sample size is increased. In the range of β_1 and β_2 values we also note that doubling the sample size from 50 to 100. At values of sample size of 100 and β_2 values greater than zero we obtain a maximum loss in efficiency of 7% for this particular level of non-proportionality, if we confine to positive values of β_1. For the negative values of β_1 the loss in efficiency is even less when proportionality does not hold. At the sample size of 50 the relative loss due to non-proportionality is at 10% compared to 7% for the sample size of 100.

So far in the discussions of non-proportionality we have only mentioned the asymptotic likelihoods. As we pointed out the results from asymptotic normality follow a very close pattern when we deal with a proportional hazard situation. This deviation increases with deviation from proportionality, however remains at minimal levels for all sample sizes of 50 and 100 generations.

Now we continue with the simulations for the non-proportional case/

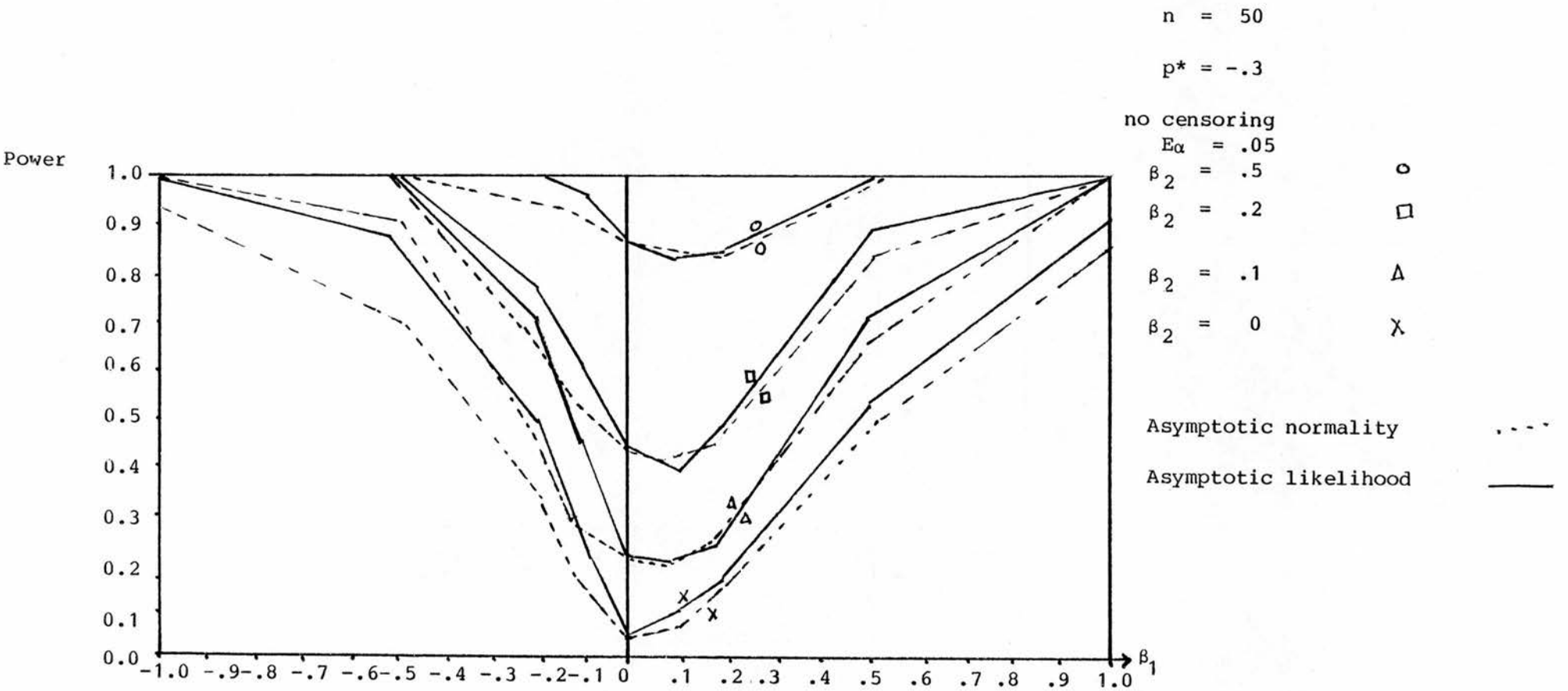

n = 50
p* = -.3
no censoring
Eα = .05
β2 = .5
β2 = .2
β2 = .1
β2 = 0
Asymptotic normality
Asymptotic likelihood
Power
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
-1.0 -.9 -.8 -.7 -.6 -.5 -.4 -.3 -.2 -.1 0 .1 .2 .3 .4 .5 .6 .7 .8 .9 1.0
β1

Figure (5.6.24)

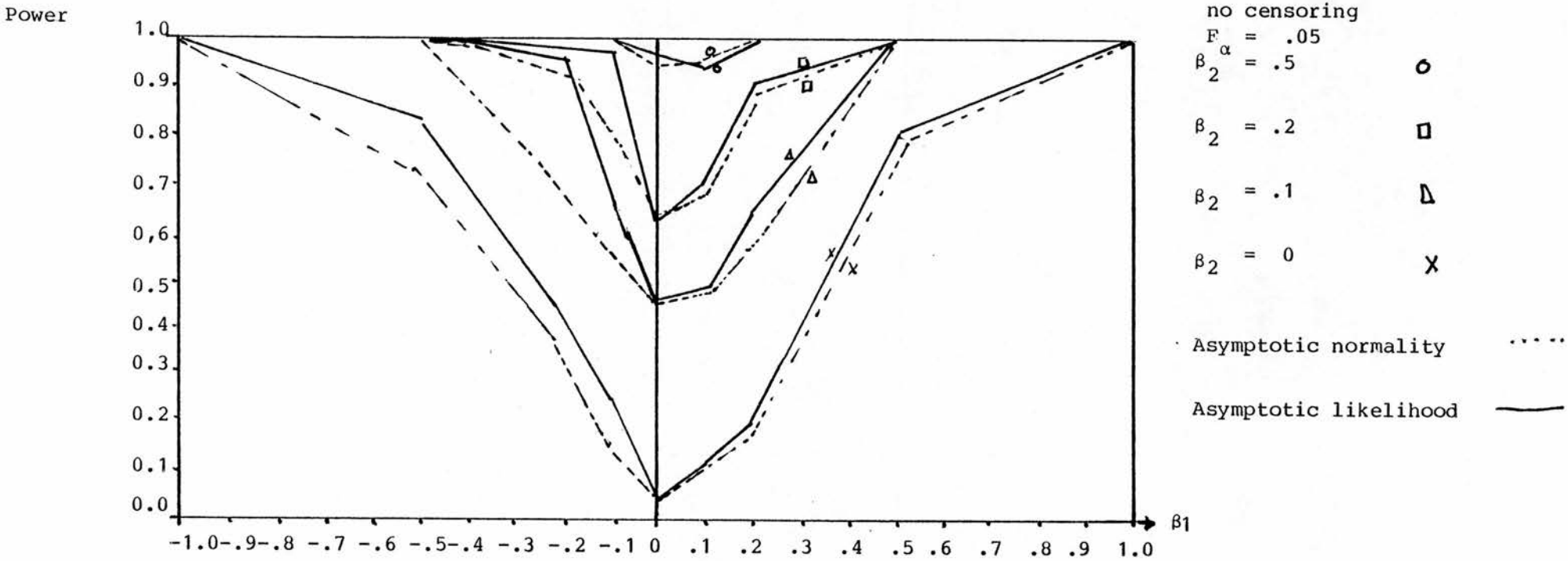

Power
1.0
0.9
0.8
0.7
0,6
0.5
0.4
0.3
0.2
0.1
0.0
-1.0 -.9 -.8 -.7 -.6 -.5 -.4 -.3 -.2 -.1 0 .1 .2 .3 .4 .5 .6 .7 .8 .9 1.0
β1
n = 100
p* = -.3
no censoring
F_α = .05
β_2 = .5
β_2 = .2
β_2 = .1
β_2 = 0
Asymptotic normality
Asymptotic likelihood

Figure (5.6.25)

case and use graphical representation on $\beta_2 > 0$, P* = -.3 and repeat for $\beta_1 > 0$, $\beta_1 < 0$ and sample sizes 25, 50 and 100 and α set to 5% and no censoring present in the data figures (5.6.23) to (5.6.25). There is a slight indication that asymptotic normality behaves in a more symmetric manner on the two sides of the β_1 axis. The asymptotic likelihood however is consistently less conservative than the asymptotic normality test. The relative power difference of the two tests diminishes with increasing sample size. The actual magnitude of parameter $|\beta_i|$ are clearly playing a role in the power of the tests. Generally the increase in value of $|\beta_i|$ reduces the relative power difference of the asymptotic likelihood and normality. This is partly due to the fact that non-proportionality variability is reduced by the increase in sample size and partly by the actual covariate effect becoming more dominant and thus producing a reduction in its variability.

Finally we present the tables (5.6.1), (5.6.2) and (5.6.3) which give the various values of range of β_1 and β_2 values used in the simulations and the corresponding β_1 and β_2 estimates with their variance, under the proportional hazard assumptions. Clearly the β estimators are very close to the actual β values. There is a negligible bias present over the range of the simulations for the given covariate set, which declines with increase in the sample size.

β_1	β_2	$\hat{\beta}_1$	VAR ($\hat{\beta}_1$)	$\hat{\beta}_2$	VAR ($\hat{\beta}_2$)
-1	0	-.9946	.077	.0005	.026
-.5	0	-.4961	.057	.0006	.027
-.2	0	-.1987	.043	.006	.027
-.1	0	-.0993	.041	.0007	.027
0	0	.0004	.031	.0007	.027
.1	0	.1012	.043	.0007	.027
.2	0	.2019	.047	.0007	.027
.5	0	.5051	.059	.0006	.027
1	0	1.0073	.079	.0005	.027
-1	.1	-.9948	.075	.1008	.032
-.5	.1	-.4962	.056	.1008	.033
-.2	.1	-.1978	.043	.1009	.036
-.1	.1	-.0993	.041	.1009	.038
0	.1	-.0004	.030	.1009	.039
.1	.1	.1011	.042	.1009	.037
.2	.1	.2017	.046	.1008	.037
.5	.1	.5049	.057	.1008	.036
1	.1	1.0070	.078	.1008	.033
-1	.2	-.9949	.071	.2010	.042
-.5	.2	-.4965	.054	.2011	.043
-.2	.2	-.1991	.041	.2011	.043
-.1	.2	-.0994	.039	.2013	.044
0	.2	-.0004	.030	.2013	.044
.1	.2	.1011	.041	.2012	.044
.2	.2	.2014	.046	.2012	.042
.5	.2	.5049	.055	.2009	.041
1	.2	1.0069	.074	.2006	.041

-1	.5	-.9956	.064	.5032	.050
-.5	.5	-.4971	.051	.5035	.051
-.2	.5	-.1992	.040	.5038	.051
-.1	.5	-.0996	.037	.5038	.052
0	.5	.0003	.029	.5039	.052
.1	.5	.1010	.037	.5037	.051
.2	.5	.2011	.043	.5034	.050
.5	.5	.5045	.052	.5031	.049
1	.5	1.0063	.069	.5027	.047
-1	1	-.9962	.054	1.0060	.050
-.5	1	-.4978	.047	1.0066	.058
-.2	1	-.1994	.038	1.0068	.064
-.1	1	-.0997	.033	1.0069	.067
0	1	-.0002	.029	1.0070	.069
.1	1	.1008	.035	1.0069	.066
.2	1	.2001	.041	1.0067	.062
.5	1	.5043	.050	1.0065	.059
1	1	1.0059	.057	1.0059	.045

Table (5.6.1) n = 25, no censoring, P* = 0

β_1	β_2	$\hat{\beta}_1$	VAR ($\hat{\beta}_1$)	$\hat{\beta}_2$	VAR ($\hat{\beta}_2$)
-1	0	-.9975	.033	.0003	.015
-.5	0	-.4981	.025	.0003	.016
-.2	0	-.1989	.021	.0002	.016
-.1	0	-.0998	.019	.0003	.016
0	0	-.0003	.015	.0003	.016
.1	0	.1007	.019	.0002	.016
.2	0	.2012	.021	.0003	.016
.5	0	.526	.027	.0002	.016
1	0	1.0039	.034	.0002	.016
-1	.1	-.9975	.033	.1006	.017
-.5	.1	-.4982	.025	.1004	.018
-.2	.1	-.1990	.020	.1006	.021
-.1	.1	-.0997	.019	.1003	.021
0	.1	-.0003	.014	.1002	.021
.1	.1	.1006	.019	.1003	.020
.2	.1	.2011	.021	.1004	.020
.5	.1	.5025	.026	.1004	.000
1	.1	1.0038	.034	.003	.048
-1	.2	-.9979	.031	.2008	.022
-.5	.2	-.4983	.024	.2009	.023
-.2	.2	-.1993	.019	.2009	.023
-.1	.2	-.0999	.018	.2009	.023
0	.2	-.0002	.015	.2008	.023
.1	.2	.1005	.019	.2008	.023
.2	.2	.2010	.021	.2007	.022
.5	.2	.5024	.024	.2007	.022
1	.2	1.0037	.033	.2006	.021.

-1	.5	-.9978	.027	.5019	.027
-.5	.5	-.4982	.023	.5017	.026
-.2	.5	-.1992	.019	.5017	.026
-½1	½5	-½0998	.018	.5015	.026
0	.5	-.0002	.014	.5016	.025
.1	.5	.1005	.018	.5017	.025
.2	.5	.2011	.020	.5017	.026
.5	.5	.5025	.024	.5017	.027
1	.5	1.0037	.029	.5018	.027
-1	1	-.9980	.023	1.0029	.024
-.5	1	-.4983	.021	1.0029	.027
-.2	1	-.1993	.018	1.0030	.030
-.1	1	-.0999	.016	1.0032	.032
0	1	-.0001	.013	1.0032	.033
.1	1	.1005	.017	1.0031	.031
.2	1	.2010	.019	1.0029	.030
.5	1	.5024	.023	1.0028	.028
1	1	1.0037	.025	1.0027	.025

Table (5.6.2) n = 50, no censoring, P*=0

β_1	β_2	$\hat{\beta}_1$	VAR $(\hat{\beta}_1)$	$\hat{\beta}_2$	VAR $(\hat{\beta}_2)$
-1	0	-.9989	.018	.0001	.011
-.5	0	-.4990	.015	.0001	.011
-.2	0	-.1996	.013	.0001	.012
-.1	0	-.1001	.012	.0001	.012
0	0	.0000	.011	.0000	.012
.1	0	.1005	.011	.0001	.012
.2	0	.2008	.012	.0001	.012
.5	0	.5017	.016	.0001	.012
1	0	1.0019	.019	.0002	.011
-1	.1	-.9989	.018	.1004	.012
-.5	.1	-.4991	.015	.1004	.013
-.2	.1	-.1996	.012	.1004	.015
-.1	.1	-.1001	.012	.1005	.015
0	.1	.0001	.011	.1005	.014
.1	.1	.1005	.012	.1004	.014
.2	.1	.2008	.012	.1004	.014
.5	.1	.5017	.015	.1004	.013
1	.1	1.0020	.019	.1004	.013
-1	.2	-.9990	.017	.2008	.015
-.5	.2	-.4991	.014	.2008	.015
-.2	.2	-.1997	.012	.2007	.015
-.1	.2	-.1001	.012	.2007	.015
0	.2	.0001	.011	.2007	.015
.1	.2	.1005	.013	.2007	.015
.2	.2	.2007	.013	.2008	.015
.5	.2	.5016	,014	.2008	.015
1	.2	1.0021	.08	.2009	.015

-1	.5	-.9990	.015	.5015	.018
-.5	.5	-.4992	.013	.5014	.017
-.2	.5	-.1998	.011	.5014	.017
-.1	.5	-.1001	.011	.5013	.016
0	.5	.0001	.010	.5012	.016
.1	.5	.1004	.011	.5013	.016
.2	.5	.2007	.012	.5014	.017
.5	.5	.5017	.015	.5014	.018
1	.5	1.0023	.016	.5015	.018
-1	1	-.9991	.013	1.0021	.014
-.5	1	-.4995	.012	1.0020	.016
-.2	1	-.2000	.011	1.0019	.019
-.1	1	-.1001	.010	1.0017	.019
0	1	.0000	.010	1.0016	.019
.1	1	.1004	.011	1.0017	.018
.2	1	.2007	.012	1.0019	.017
.5	1	.5015	.013	1.0020	.017
1	1	1.0023	.015	1.0022	.016

Table (5.6.3) n = 100, no censoring, P*=0

Earlier we showed that in the study of the Cox's method the exponential and Weiball generated samples are very similar in terms of testing significance of covariates and in fact the interesting situation is that of effect of non-proportionality of the hazards. Now we will consider the simple tests of the treatment effect for the various values of the non-proportional hazard generations. The tests once again correspond to a similar set of β_1 and β_2 values, both greater than zero. In the following generations however we will repeat the simulations and the analysis of the generated sample according to different generalised linear models. As we described in chapter 3 the most commonly used models in this respect are the Weiball and the exponential models. We will report the simulations initially for the proportional hazard generations. In the analysis we will consider (a) the fixed covariate Cox's method, (b) time dependent Cox's method which is more suitable for the non-proportional situation, (c) stratified Cox's method (d) Weiball model with the generalised linear model assumption and finally (e) the exponential model.

The non-proportional generations are all of the Weiball type. This is an arbitrary choice in so far as deviation of the exponential decomposition of the relative risk is concerned. For the purpose of the analysis we deal only with exponential and Weiball parametric models. These two models are in practice the most relevant for the decomposition of the relative risks in survival studies. Due to an introduction of non-proportionality into the generated samples an alternative approach based on the non linear models of Weiball type may be possible. However in this respect the interpretation of the β estimates/

β estimates are not relevant to the aims of the study. In here a point which must be noticed is the important distinction between the non-proportionality mentioned in the parametric Weiball models and the non-proportionality under the semi-parametric constraint of the Cox's method. In the Weiball model the problem is essentially specification of the wrong model in the presence of time dependency, which is similar to that of specifying the Cox's proportional hazard model without time dependency.

A further point that requires some attention is related to the results of the previous study of β_2 efficiency in the presence of β_1. We assumed there is no correlation between β_1 and β_2 and thus there was no confounding effects present. It was generally clear that the value of β_1 does not effect the power of tests for β_2. We will now repeat the same values of β_1 and β_2 and analyse the samples with the above mentioned models. Once again there is a distinction in that absence of confounding between β_1 and β_2 entails a constant relative power for the estimation of treatment effects. For the time dependent situation however any loss of power is essentially attributed to the use of wrong models.

In general we will ignore the exponential model in our discussions, it is presented for illustration in the figures and later we will make a few comments on the inferiority of the exponential model. In all of what follows we take a significance level of 0.05 and no censoring situations. Initially we consider the sample size of 25, figure(5.6.26). At β_2 set to zero there is clear agreement between all the models at power levels around 0.05. This value is constant and does not/

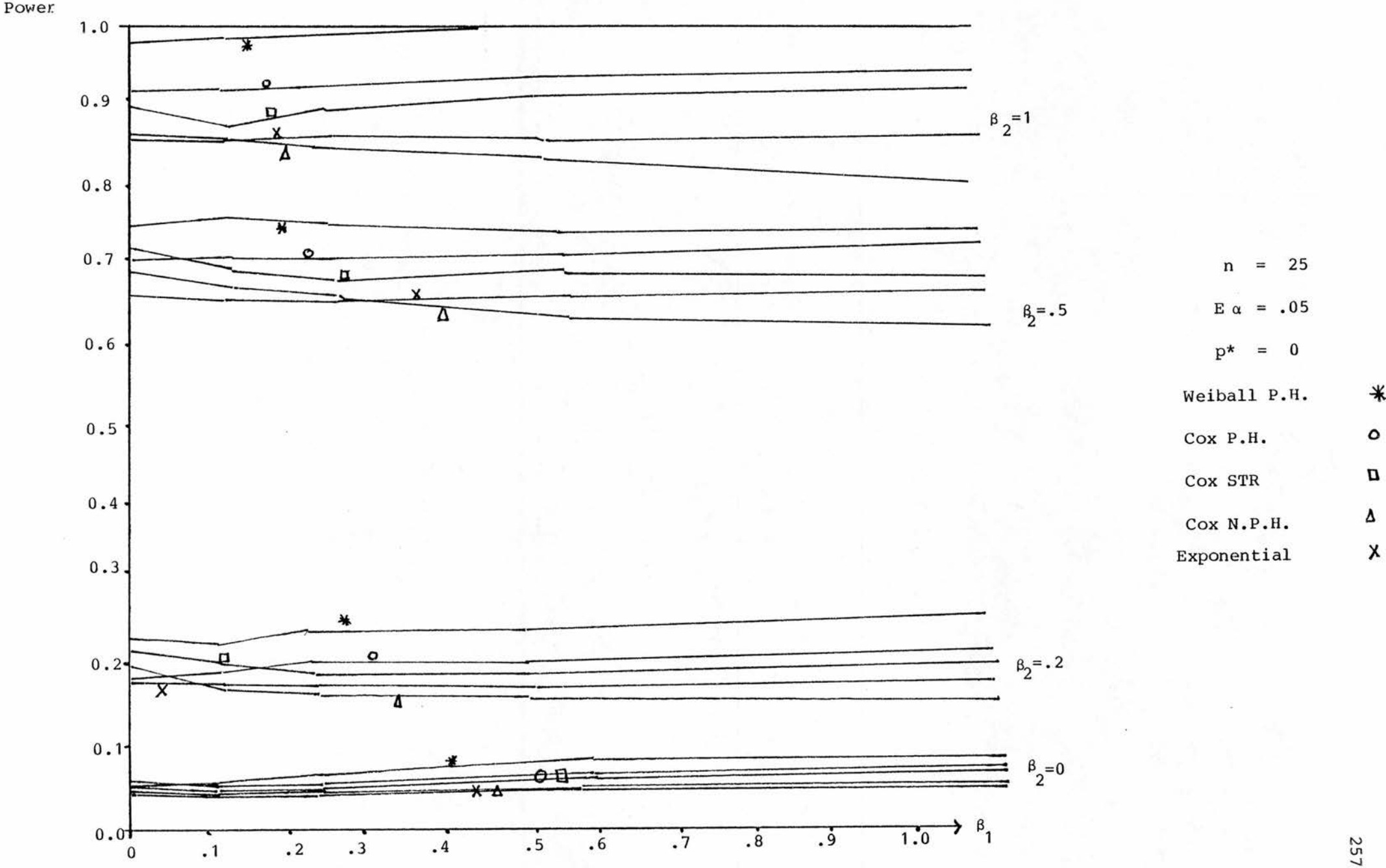

Power
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0
.1
.2
.3
.4
.5
.6
.7
.8
.9
1.0
β_1
β_2=1
β_2=.5
β_2=.2
β_2=0
n = 25
E α = .05
p* = 0
Weiball P.H.
Cox P.H.
Cox STR
Cox N.P.H.
Exponential

Figure (5.6.26)

not change with increasing values of β_1 from 0 to 1. Next we consider values of β_2 greater than zero. We note that with the proportional hazard generations consistantly, the parametric model, Weiball, has more power than the other models. The worst model (ignoring exponential) when the samples have proportional hazards is the Cox's non-proportional hazard model. The other models Cox' proportional hazards and the stratified Cox's model both have power properties between the Weiball and the non-proportional Cox and this is true for all values of β_1 and β_2 simulations.

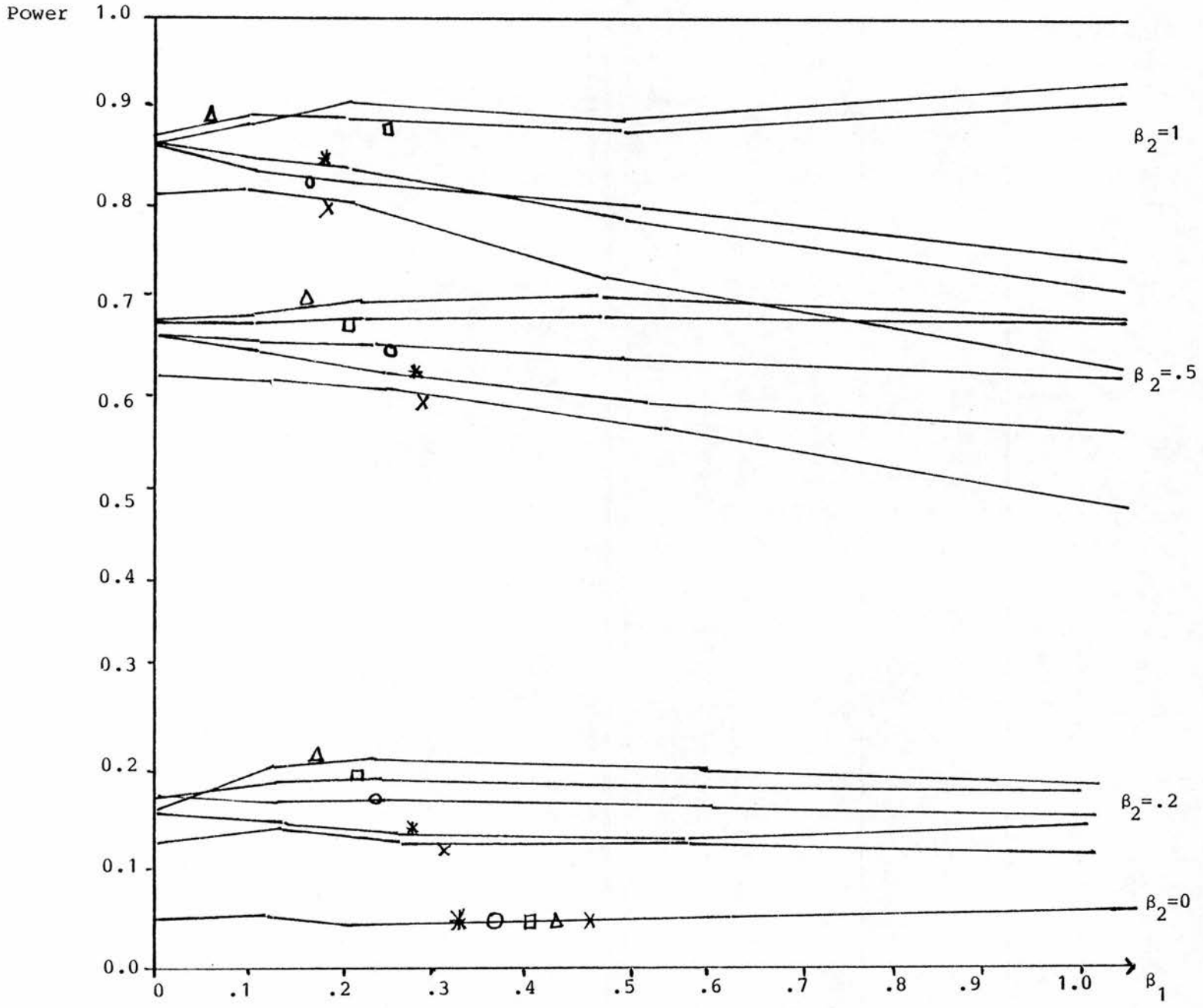
The difference between the power of Cox's non-proportional hazard and the Weiball proportional hazard is about 5%, near $\beta_1 = 0$ and about 9% near $\beta_1 = 1$, when β_2 is equal to 0.2 for both. The difference increases with increasing values of β_2 so that at $\beta_2 = 1$ with values near $\beta_1 = 0$, the difference is 12% and at $\beta_2 = 1$ it is 13%. There is clearly a lack of consistency in the power efficiency of some of the models, in so far as values of β_1 are concerned. The more superior models namely Cox's proportional hazards and the Weiball in fact consistantly produce the same power value for a given β_2 regardless of values of β_1. The difference in power for any given value of β_2 in fact for either of Cox's proportional hazard or Weiball does not vary by more than 3% over the range of β_1 values. For the less appropriate models, the Cox's non-proportional hazard, we note a slight declining trend at n = 25,with increasing β_1 values. As an example at $\beta_2 = 5$ and $\beta_1 = 0$ the power is 69% and the value declines to 63% when $\beta_1 = 1$.

The stratified Cox's model also produced consistantly similar values/

values of power for the test of β_2 regardless of β_1, although there is a slight loss of power compared to the Cox's unstratified model. The difference between their power value is almost consistantly 4%

Next we consider the sample size of 25, significance level 0.05, the increasing non-proportionality at P* =0.3 and the decreasing non-proportionality at P* = -.3. Generally the power values of the P*>0 are slightly superior to similar values of P*< 0. This is mainly due to the convergence or the divergence of the hazard rates for the given range of β_1 and β_2 values.

First we deal with P* = +.3, figure (5.6.27). At β_2 = 0 the power of all tests and all models is once again very close to the value of the significance level 0.05. In fact there is very little to separate the power of tests according to type of the model or the range β_1 values. For values of $\beta_2 > 0$ once again there is a slight difference between the power of tests according to the type of the model. Consistantly the exponential is the worst model in the analysis of non-proportionality. The best model for such samples is the Cox's non-proportional hazard model. The stratified Cox's model also produces relatively superior power values compared to the Weiball or the Cox's proportional hazard models. The non-proportional hazard model produces very consistant power values for β_2 regardless of β_1 values. This point is in fact also true of the stratified models of Cox. The two less powerful tests in the analysis of non-proportional samples are the proportional hazard Cox and the proportional hazard Weiball. At the value of β_2 = .2 we note that the value of β_1 does not effect the power of any of the tests and the maximum difference for the ranges of β_1 value/



n = 25

E_α = .05

p* = +.3

Weiball P.H. ✱

Cox P.H. ○

Cox STR □

Cox N.P.H. Δ

Exponential ×

Figure (5.6.27)

value is 3%.

There is a lack of consistancy in the power of β_1 tests as the magnitude of β_2 increases. This pattern is not present for the correctly specified models namely Cox's non-proportional hazard and the stratified Cox's model. With the proportional hazard model Cox and Weiball however we note a decline in the efficiency of the tests. The decline in efficiency for the β_2 =.1 over the range of β_1 from 0 to 1 is about 9% if a proportional hazard model is used. Before we finish this point however, we must remark that this pattern is present at this magnitude only at the relatively low sample size of 25. The value of P* = -.3 figure(5.6.28), produces non-proportionality which implies generally a higher loss of power compared to P* = .3. At the value of β_2 = 0, the power is at about 0.05 for all tests and all values of β_1. However there is a slightly higher variability over the range of β_1 values for the different models compared to the situation of P* = +.3.

On increasing values of β_2 there is a general increase in the overall power of tests, which indicates that for all models, values of β_2 is the major factor influencing power. The pattern is similar to P* > 0, indicating that non-proportional hazard Cox's model and the stratified Cox's method are the superior models. This once again indicates that the correct specification of model implies a general constant power for the β_2 values regardless of β_1 values. The worst situation occurs for the Weiball and the Cox's proportional hazard model in the P* = -.3. There is once again a slight indication that increasing values of β_1 may influence the power, which is to some/

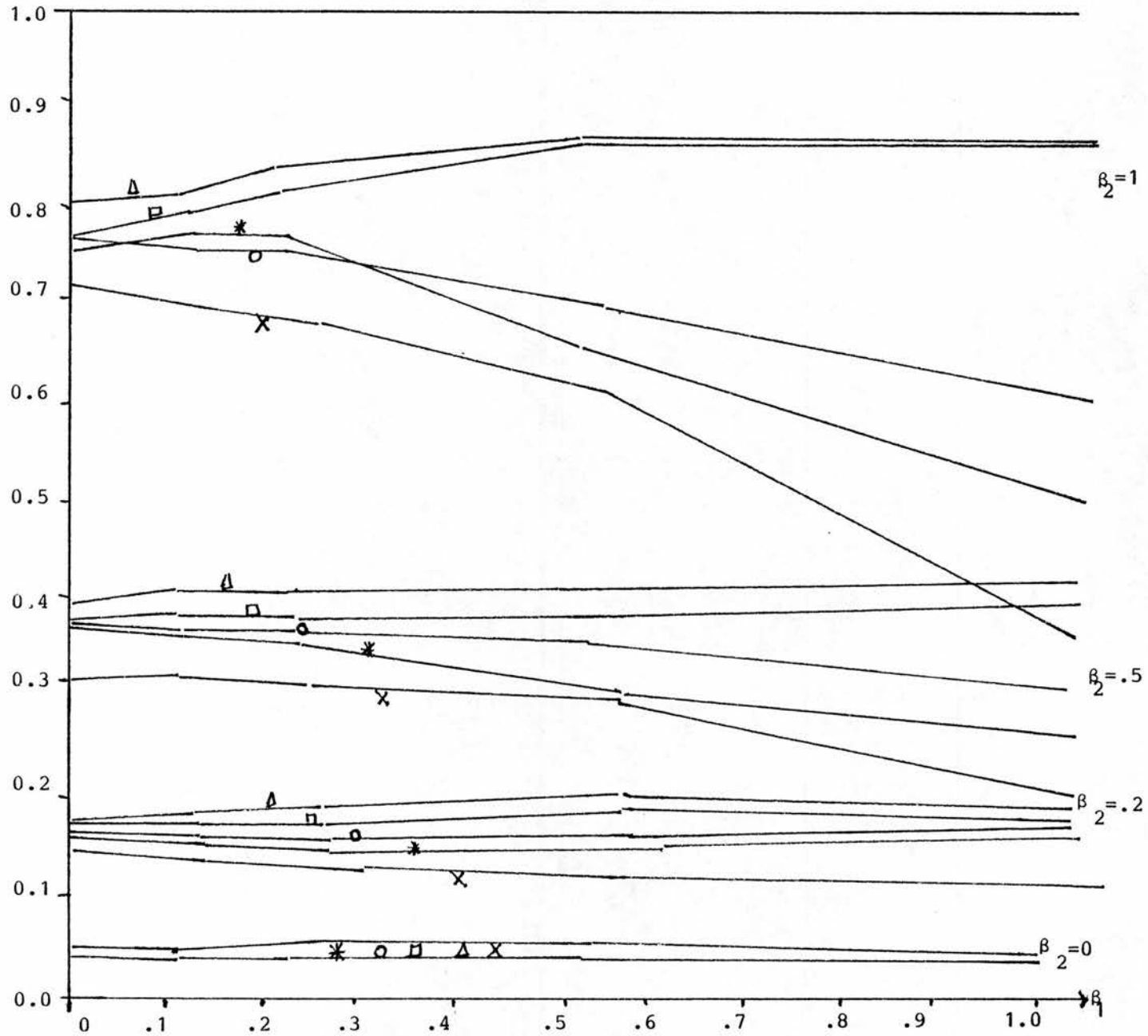

Power
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0
.1
.2
.3
.4
.5
.6
.7
.8
.9
1.0
β1
β2=1
β2=.5
β2=.2
β2=0
n = 25
E α = .05
p* = -.3
Weiball P.H.
Cox P.H.
Cox STR
Cox N.P.H.
Exponential

Figure (5.6.28)

some extent due to a low sample size and partly due to lack of balance due to non-proportionality.

By an increase of sample size to 50, figure (5.6.29) to (5.6.31) and to 100, figures (5.6.32) to (5.6.34) a similar pattern as before is repeated. However the differences between the appropriately fitted models and the unsuitable models in either of proportional and non-proportional hazard situations decline. Under the proportionality of hazards, Weiball, is in fact the most suitable model and produces the highest power of the tests. (It must be noted however that the generation are also of proportional hazard of Weiball type). Cox's proportional hazard is also suitable in that it is not influenced by varying values of β_1. The two relatively unsuitable models are stratified Cox and the time dependent Cox, although their loss of efficiency is relatively small. In the analysis of the generation of non-proportional hazard type both Weiball and proportional hazard Cox decline in power.

As we pointed out earlier, clearly the exponential model is the least suitable model and we have included the model purely for reference in the graphs.

In conclusion, the stratified Cox's model and the time dependent Cox's model are both suitable for the analysis of non-proportional generations. The value of covariate effect β_1 in this respect does not vary the power of the test. As may be expected the power of the test is purely dependent on the magnitude of the treatment effect. In the situation of proportional hazards both Weiball and Cox's proportional hazard/

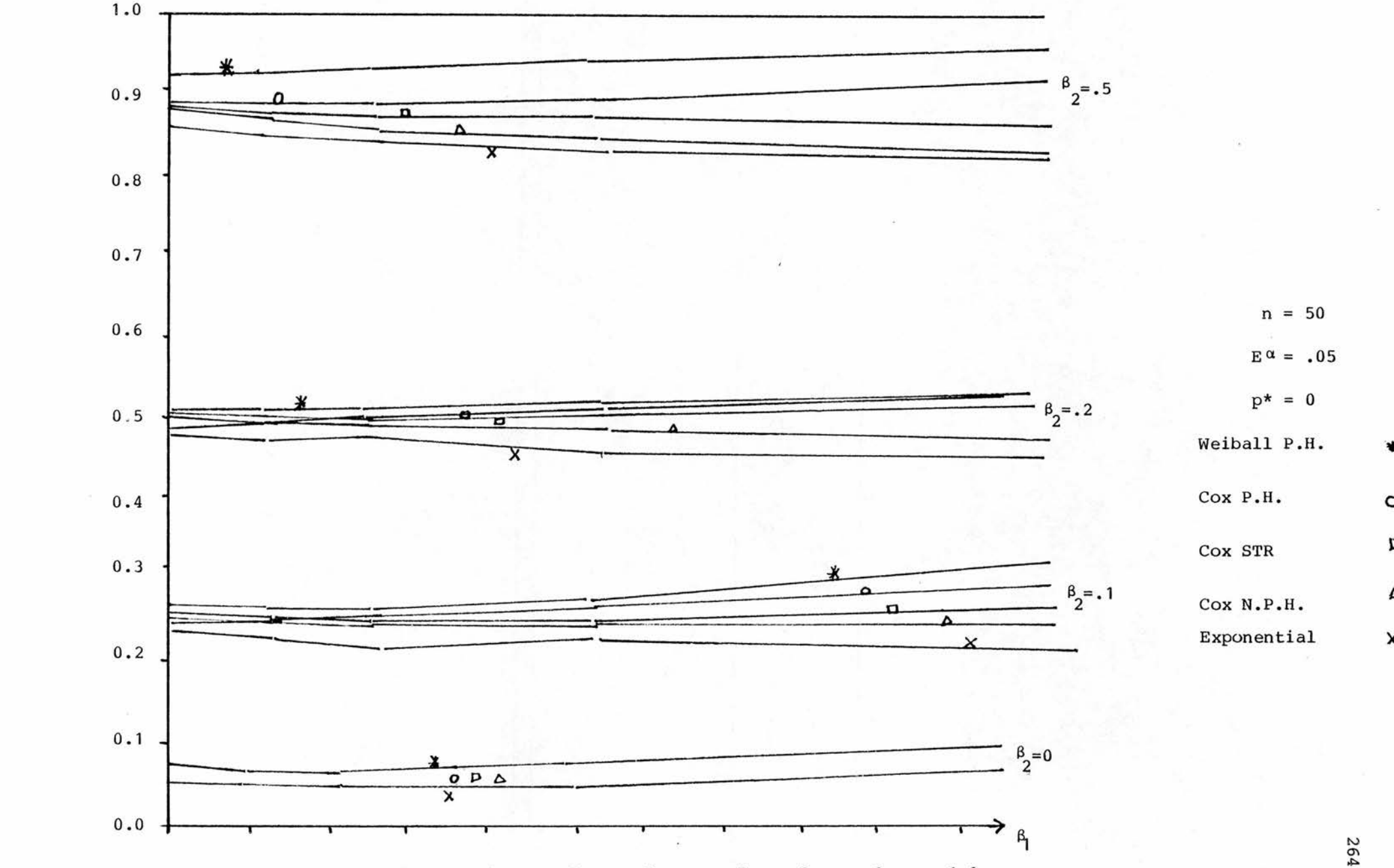

Power
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
β2=.5
β2=.2
β2=.1
β2=0
β1
.1
.2
.3
.4
.5
.6
.7
.8
.9
1.0
n = 50
Eα = .05
p* = 0
Weiball P.H.
Cox P.H.
Cox STR
Cox N.P.H.
Exponential

Figure (5.6.29)

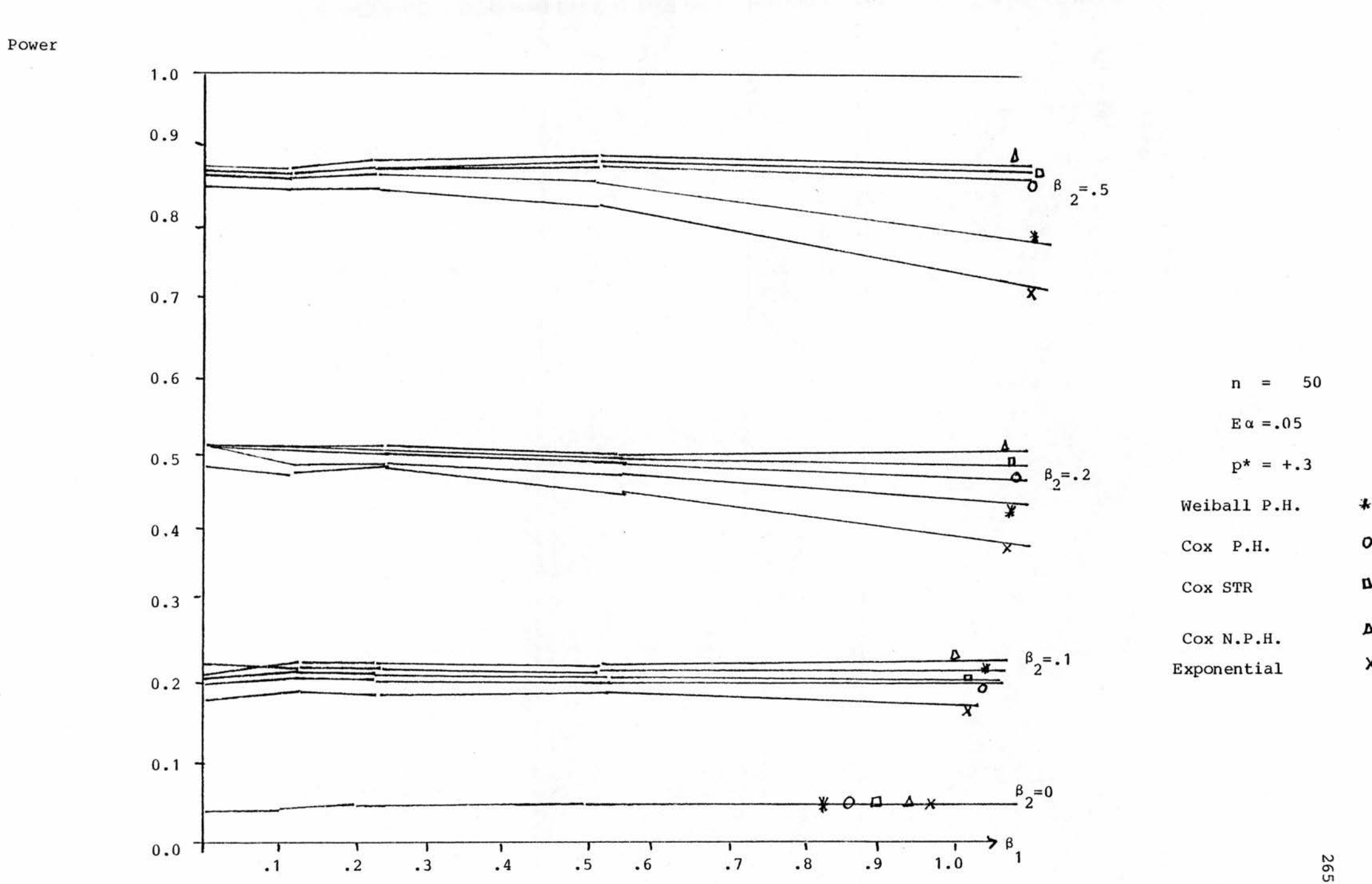

Power
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
.1
.2
.3
.4
.5
.6
.7
.8
.9
1.0
β1
β2=.5
β2=.2
β2=.1
β2=0
n = 50
Eα =.05
p* = +.3
Weiball P.H.
Cox P.H.
Cox STR
Cox N.P.H.
Exponential

Figure (5.6.30)

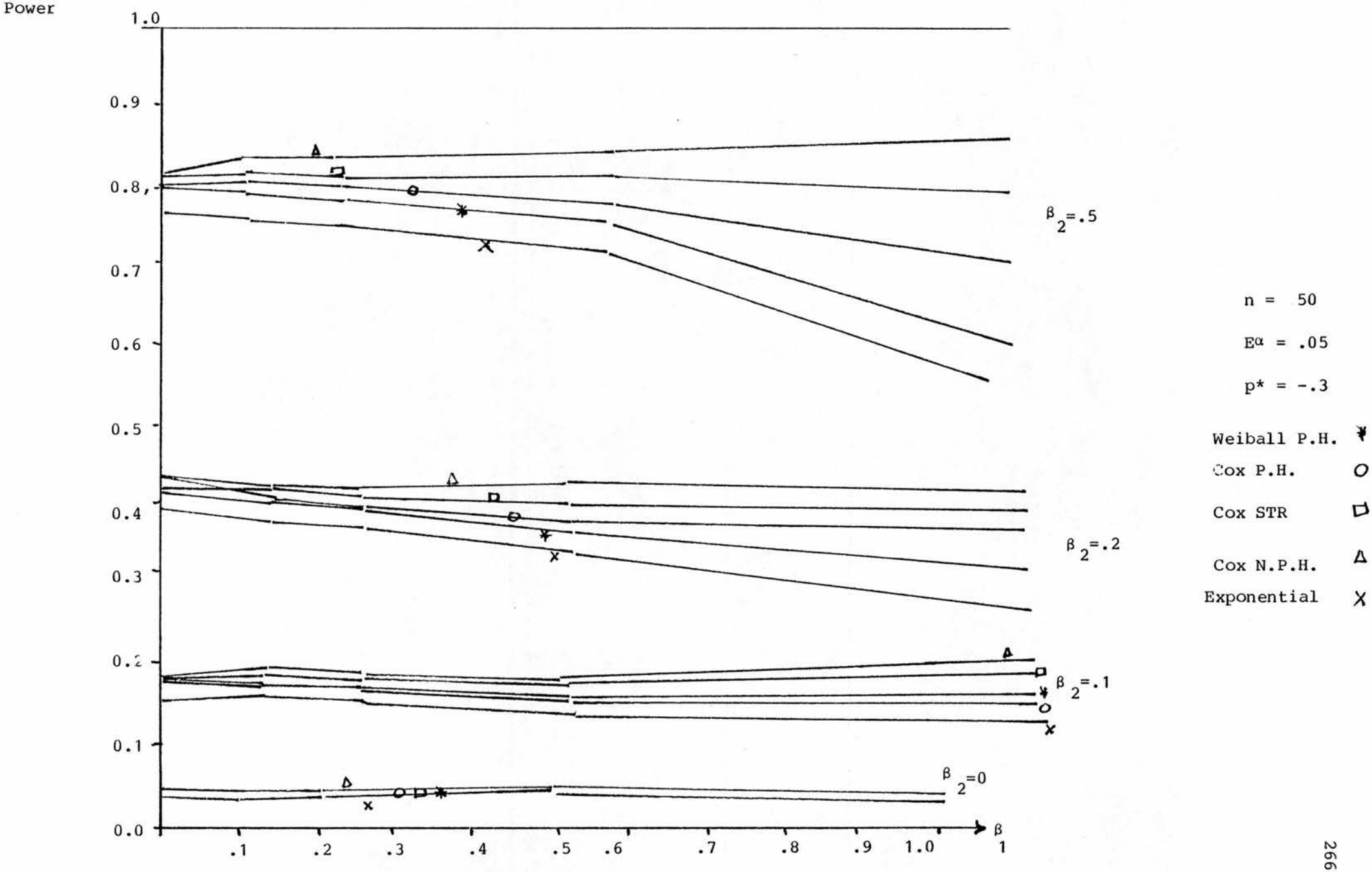

Power
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
.1
.2
.3
.4
.5
.6
.7
.8
.9
1.0
1
β
β 2=.5
β 2=.2
β 2=.1
β 2=0
n = 50
Eα = .05
p* = -.3
Weiball P.H.
Cox P.H.
Cox STR
Cox N.P.H.
Exponential

Figure (5.6.31)

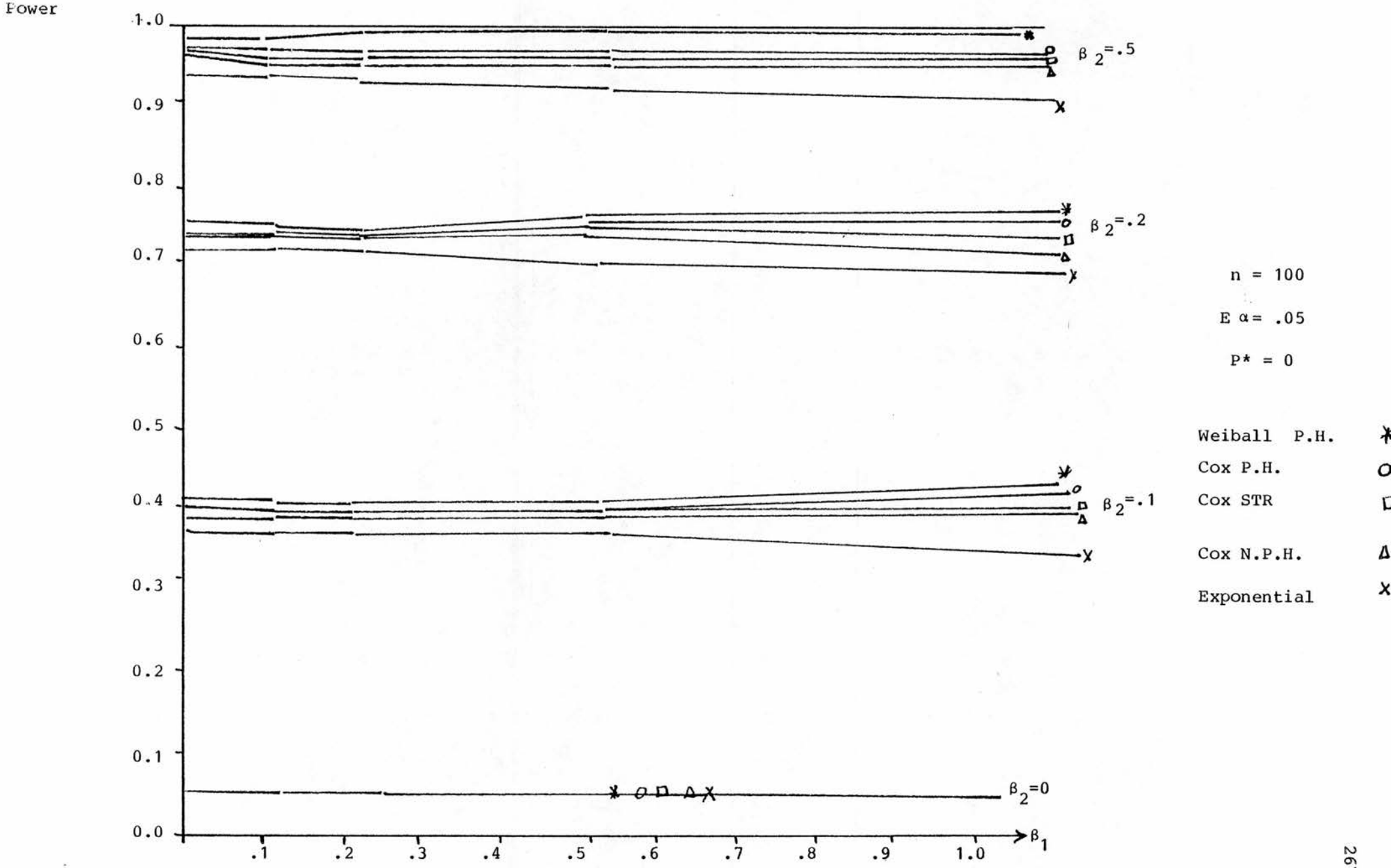

Power
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
.1
.2
.3
.4
.5
.6
.7
.8
.9
1.0
β_1
β_2=.5
β_2=.2
β_2=.1
β_2=0
n = 100
E α= .05
P* = 0
Weiball P.H.
Cox P.H.
Cox STR
Cox N.P.H.
Exponential

Figure (5.6.32)

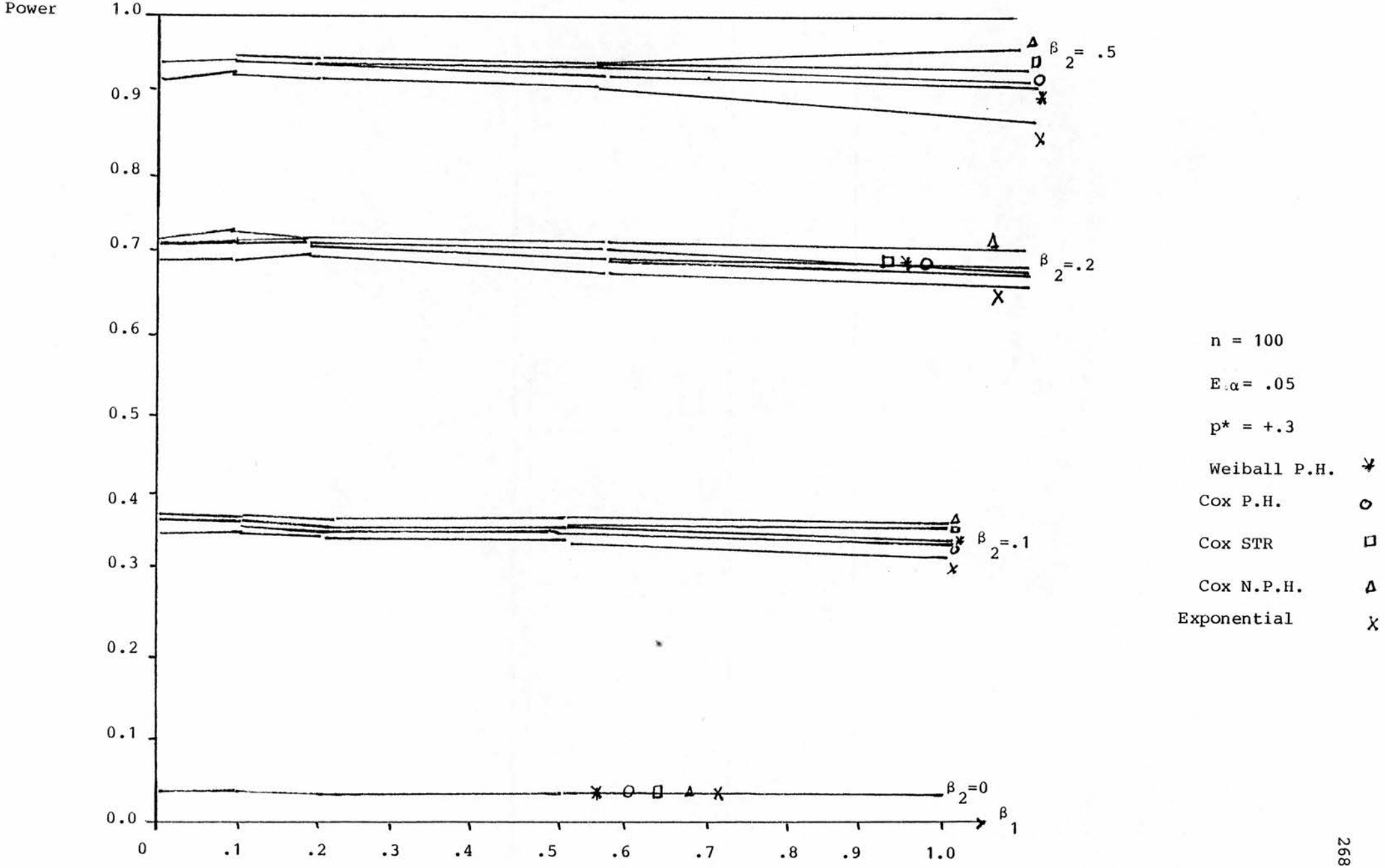

Power
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0
.1
.2
.3
.4
.5
.6
.7
.8
.9
1.0
β 1
β 2= .5
β 2=.2
β 2=.1
β 2=0
n = 100
E.α= .05
p* = +.3
Weiball P.H.
Cox P.H.
Cox STR
Cox N.P.H.
Exponential

Figure (5.6.33)

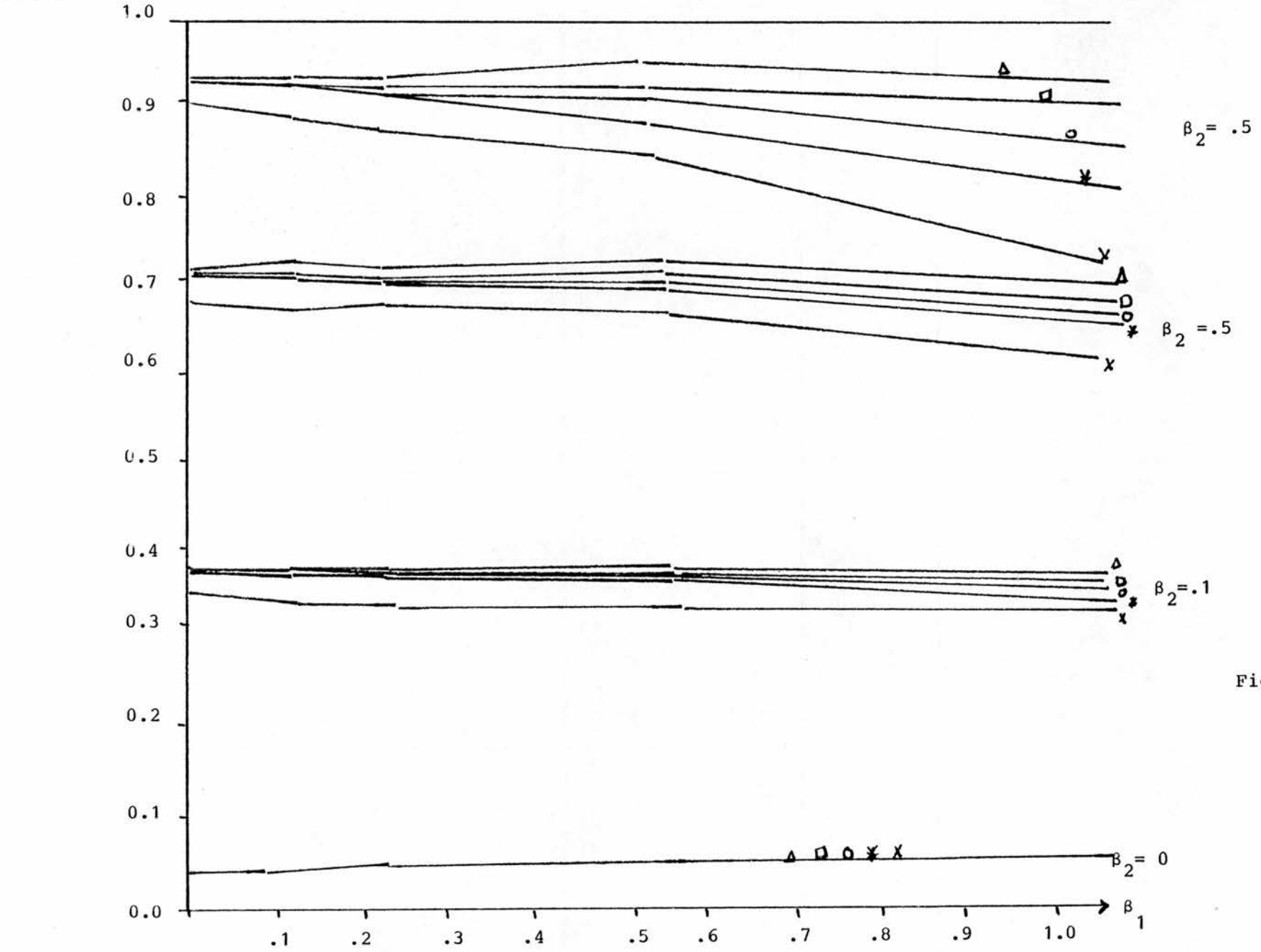

Power
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
.1
.2
.3
.4
.5
.6
.7
.8
.9
1.0
β1
β2= .5
β2 =.5
β2=.1
β2= 0
n =100
Eα =.05
p* =-.3
Weiball P.H.
Cox P.H.
Cox STR.
Cox N.P.H.
Exponential

Figure (5.6.34)

hazard have good power properties. Once again the power is dependent on the magnitude of the treatment effect and is not influenced by the covariate effects. This is in opposition with the findings of a similar study in situations of proportional hazards where unreasonable loss of power is detected due to the magnitude of the covariate effect (C.L. chastong 1983). In our study specification of wrong models does imply a loss of power which, with the small sample size of 25 can become unreasonably dependent on the magnitude of the covariate effect.

For all practical reasons the semi non parametric methods provide a robust construct for the analysis of the data.

CHAPTER 6

ANALYSIS OF THE OLD EDINBURGH DATA

In this chapter we proceed with the analysis of data from a clinical trial. The purpose of this chapter is to illustrate some of the results of the previous sections, using the proportional hazards model. Since the analysis of the breast cancer trial data is the main part of the discussion, we will begin this chapter with a history of the treatment of the disease. In the later sections a general overview of the subject will be presented. Then our data is described and the procedures which were adopted to collect it will be presented. Finally we analyse the data using the general methods with a single event of interest and multiple coefficient models with tests of interactions.

In the present chapter we only consider time independent covariates, however in chapter 7 we will deal with time dependency and multivariate risks.

6.1 Randomised Trials in early breast cancer.

Breast cancer is the most common form of cancer among women in the Western Hemisphere. Despite this, there is no general agreement as to the best treatment of an early case. This disagreement is related/

related to both types of surgery and the value of radiotherapy. Recently various forms of post operative drug treatments have also added a new dimension to the decision making efforts.

Breast cancer is one of the few malignant diseases in which there are well documented data on long term survivals in untreated patients in existence. The earliest efforts for the purpose of the treatment of the disease took place some 80 years ago. However later, during the 1950's, epidemiologists gathered the first empressive arguments against the use of the established treatment of the time which was radical surgery. It seemed that treatment did not cure the patients in terms of their long run survival or proportion of the development of metastatic disease.

Following the above developments many studies were carried out to assess a range of different treatments which consisted mainly of loco-regional treatment by various forms of surgery and radiotherapy. Subsequently some ovarian ablation by oopherectomy or by irradiation has also been used. None of these treatments, however, produce a major improvement in terms of over all survival. Much of this lack of success in treatment had been ascribed to the fact that patients with the possibility of developing metastatic disease are not influenced by the loco-regional treatments and the development of the metastatic disease has not been attacked by the treatment.

At present there are new trials taking place in which surgery is followed up by chemotherapy. The value of such drug treatments occurs/

occurs by not only considering the benefits in terms of local progression but in terms of systemic general progression of the disease. Some of the methods we study in the next chapter are in fact appropriate for the proper assessment of the effects for this form of trial.

Often the evaluation of the treatment of breast cancer is made difficult by the fact that patients differ considerably in their individual form of the development of the disease. Various prognostic factors in the past have been assessed in terms of effects of various treatments. Some of the indicators that have been given an importance in the past are, the size of the initial tumour, axillary node involvement and the menstrual status. In chapter 8 we will deal in more detail with the important prognostic effects. Generally size of the tumour is invariably related to survival and this is a result that has been shown to be true consistantly. Another strong prognostic indicator is the extent of axillary node involvement. This can be measured as a form of index with involved or not involved categories; or by an index representing extent of the involvement by the number of nodes examined and the number that were found to be involved. Age and menstrual status are two factors that are closely related to each other and are of less strength in assessing the chance of survival of a patients due to the disease in comparison with node or size.

It has often been shown that with increasing age the chance of survival increases until the menopause. After the menopause the survival/

survival rates decline at a slower rate. The effect of the other prognostic factors are also present if we do a separate stratified analysis of the various age categories.

6.2 Description of the data.

The objective of this trial has been to assess the pattern of survival rates for a group of patients with the invasive carcinoma of the breast, who were treated in a clinical trial in the South East of Scotland from 1964 to 1971.

In the protocol, the general criterion for selection was taken to be, all female patients between the ages 35 - 69 inclusive. Further it was considered essential that all patients must be suitable for treatment by either arms of the trial, so that a reasonable level of homogeneity of patients is established in terms of prior treatment status.

The two trial options were:

(1) Radical mastectomy: The breast, the pectoral muscles and the axillary cortices were removed.

(2) Simple mastectomy plus post-operative radiotherapy: The breast was removed from the fascia overlying pectoralia major via an elliptical oblique incision. This included the nipple and the areola. Post operative radiotherapy was given by a 2 Mer vander Graff generator. The axilla and the supraclaricular fossa were irradicated using parallel semi-opposed fields to 4250 rad. maximum dose in 10 factions in 4 weeks. The chest wall and the internal mammary nodes were irradiated/

by parallel tangential fields to 4500 rad. in fractions in 4 weeks.

All patients were categorised into stage 1,2 and 3 patients according to the then currentInternational Staging Systems based on TNM, (codes for Tumour size, Nove involvement and Malignancy status respectively). In chapter 8 we consider the development of the TNM staging in greater detail.

Thus the stage I patients were composed of patients with tumours of size 5 cm or less in the maximum diameter, Skin fixation absent or incomplete, nipple may be retracted or pagets disease present, pectoral muscle fixation absent, chest wall fixation absent, no homo-ateral axillary nodes palpable and no distant metastases present.

Stage II patients had primary tumours as in Stage I but also include homolateral axillary nodes palpable, movable and not fixed to one another.

Further certain members of Stage III were also defined as elligible to take part in the study. Such cases may have tumour of any size, skin fixation complete or ulceration not exceeding 3 cm in diameter, peau d'orange present in tumour area only, pectoral muscle fixation complete or incomplete. Stage III patients which were excluded were cases with skin involvement wide of tumour or ulceration greater than 3 cm, peau d'orange wide of tumour, chest wall fixation present, homolateral axillary nodes fixed to each other or to adjacent structures, oedema of the arm, homolateral supra-clavicular/

clavicular or infraclavicular nodes movable or fixed. All Stage IV patients were excluded. These are in fact patients with distant metastasis.

Apart from the above, certain other patients were excluded from the trial:-

(a) Previous treatment for carcinoma of the breast.

(b) Bilateral breast carcinoma

(c) Any other malignancy.

(d) Breast carcinoma having arisen during or presenting in association with pregnancy or lactation.

(e) Previous bilateral ovariectomy or pelvic irradiation

(f) Peripheral vascular disease of the upper limb.

(g) Certain tumours in axillary tail unsuitable for treatment by radiotherapy because of position.

Patients with advanced disease are usually subjected to high risks under operation. On ethical grounds this entails the exclusion of all such patients from the arms of the trial. For scientific reasons a few conditions in this respect are of importance. Advanced patients have often short survival times due to external factors and thus their distribution can mask the treatment survival patterns. The number of such patients is often small and an unbalanced distribution of these patients can make the treatment comparisons controversial. Therefore in order to obtain more uniform groups of patients for the final comparisons it is reasonable to exclude the advanced patients.

All/

All elligible patients were further stratified according to age and stage of the disease. Such criteria were taken to be the minimum data necessary for a random allocation of patients into the trial. Clinical stages form 3 strata (a) stage I, (b) stage II and (c) stage III. Age is also categorised into three strata, (a) 35-44, (b) 45-59 and (c) 60 - 69.

A randomisation office was set up and on receipt of name, age and the stage of the disease at the initial examination from a peripheral hospital, the units concerned were informed of the treatment by telephone and by writing.

The benefits of a stratified allocation of treatments can be maximised by an accurate assessment of the categories. Age seems not to be a major problem since it is a single measurement in a continuous scale. However stage is a collection of various informations based on T, size and N, node staging.

The M staging, presence of metastatic disease in this trial reflects only a group of inelligible patients and it is important that assessment of presence or absence of metastatic disease is very accurate. In order to reduce the chances of including cases with skeletal metastasis, it was stated in the protocol that X-rays of chest and pelvis should be taken in all the cases included in the trial. It stated further that if possible this should be done by the surgical unit and films must be sent with the patients.

The/

The total group of patients who were randomised by this procedure included approximately 50% of patients found at the time of operation to have benign breast disease. These patients were excluded from the trial for all purposes. As we will point out later, this procedure resulted in the allocation of unequal numbers of patients for each comparable strata of the treatment arm. However the final imbalance in terms of the number of the malignant patients is not of practical importance.

6.3 Recording of Information.

A general procedure was adopted so that information on the patients could be standardised and so processed by a computer. However it was noted in the last review of the data performed in 1981 that certain concepts and categories defined by the protocol were not in accord with more recent practice. Most of these changes did not create a major problem of interpretation. The major source of inconsistency among the changing definitions seemed to be concepts related to the recurrence of the disease. In fact in the original protocol there was no mention of the definition of recurrence of the disease, although in the data forms space was allocated to recording of such information.

Basically there were 4 standard forms available for processing

Form 1 - the Initial Examination form, Figure (6.3.1)

Form 2 - The Primary Treatment form , Figure (6.3.2)

Form 3 - The Anniversary Record form, Figure (6.3.3)

Form 4 - The Pathology Report form, Figure (6.3.4)

FORM 1

INITIAL EXAMINATION

A

Serial Number ____

SURNAME _____

GIVEN NAMES _____

ADDRESS _____

_____

COUNTY _____

UNIT _____ SURGEON _____

MARITAL STATE – Enter M or S in box

DATE OF BIRTH – Enter day 01 to 31, month 01 to 12, and last two digits of year.

Day	Month	Year

(AGE:–)

MENSTRUAL STATE

Premenopausal – enter 1
Menopausal – enter 2
Post-menopausal – enter 3

AGE AT MENOPAUSE (years last birthday)

HISTORY AND CLINICAL FINDINGS

Day	Month	Year

DATE FIRST SYMPTOM OR SIGN NOTICED _____

PRIMARY TUMOUR

SIDE – Enter R or L

SITE

Medial half only – enter 1
Lateral half only – enter 2
Central – enter 3
Both halves – enter 4
Whole breast – enter 5
Unknown or other – enter 6

SIZE – greatest diameter in cm.

TMN CATEGORIES

T – enter appropriate number

N – enter appropriate number

CLINICAL STAGE – Enter appropriate number _____

If tumour is STAGE III:

(a) State SKIN INVOLVEMENT (T1, 2 or 3)

(b) State PECTORAL MUSCLE INVOLVEMENT (T1 or 3)

SELECTED TREATMENT OPTION

Enter appropriate code – R1, R2, S1 or S2

FORM 2

PRIMARY TREATMENT

B

Serial Number ____

SURNAME _____

GIVEN NAMES _____

ADDRESS _____

_____

PRIMARY TREATMENT

Date of first treatment _____ Day | Month | Year

SURGERY (enter 0 if NO and 1 if YES for each item below)

Simple mastectomy

Node or nodes removed

Part of pectoral fascia removed

Part of pectoral muscle removed

Radical mastectomy

Closure without skin graft

Closure with skin graft

DAYS IN HOSPITAL _____

RADIOTHERAPY – enter 0 if NO, 1 if YES _____

Min. Dose (enter rads)

Max. Dose (enter rads)

Time (weeks)

SUPPLEMENTARY TREATMENT

If none – enter 0
If ovariectomy – enter 1
If ovarian radiation – enter 2

If ovarian radiation

Completed – enter 1
Incomplete – enter 2

FORM 3

ANNIVERSARY RECORD

C 281

Serial Number

SURNAME _____

GIVEN NAMES _____

ADDRESS _____

_____

Anniversary year

LATE COMPLICATIONS

Enter 0 if NO, 1 if YES in each case

(Oedema of arm
(Limitation of shoulder movement
(Other late complications
Specify _____

LOCAL RECURRENCE

Date of first evidence of local recurrence (if in this anniversary year) — Month | Year

Enter 0 if NO, 1 if YES for each item

(SITE 01 Chest wall
(02 Axilla
(03 Supraclavicular fossa
(04 Internal mammary node

Enter number (above) of first site of recurrence if it occurred in this anniversary year. (If more than one site observed simultaneously enter 05.) Otherwise enter 00.

DISTANT METASTASIS

Date of first evidence of distant metastasis (if in this anniversary year) — Month | Year

Enter 0 if NO, 1 if YES for each item

(SITE 06 Skeleton
(07 Lung
(08 Pleural effusion
(09 Other
Specify _____

Enter number (above) of first site of recurrence if it occurred in this anniversary year. (If more than one site observed simultaneously enter 10.) Otherwise enter 00.

SECONDARY TREATMENT (commenced in this anniversary year)

Enter 0 if NO, 1 if YES for each item

(Surgical excision of metastases
(Radiotherapy
(Hormone therapy (oestrogens, androgens or steroids)
(Endocrine surgery (oophorectomy, adrenalectomy, hypophysectomy)
(Cancer chemotherapy
(Other
Specify _____

DEATH during this anniversary year (Enter 0 if NO, 1 if YES)

If YES: Date of death — Month | Year

CAUSE OF DEATH

If carcinoma of breast — enter 1
If other cause but recurrence of breast carcinoma present — enter 2
If other cause but no evidence of recurrence of breast carcinoma — enter 3

If other cause was:

Complication of primary treatment — enter 1
Complication of secondary treatment — enter 2
Other primary neoplasm — enter 3
Other intercurrent condition — enter 4

Specify _____

(If due primarily to carcinoma of breast enter 0.)

FORM 4

PATHOLOGY

D

Serial Number

SURNAME _____

GIVEN NAMES _____

ADDRESS _____

_____

LABORATORY _____

PRIMARY TUMOUR

Carcinoma – 0
Sarcoma; specify _____ – 1
Non Malignant – 2

Size – greatest diameter in cm.
* N.S. – XX

Descriptive

Scirrhous	– 0	Paget's	– 6
Comedo	– 1	Other; specify _____	– 7
Papillary	– 2	More than one	– 8
Medullary	– 3	* N.S.	– X
Mucoid	– 4	≠ N.A.	– Y
Squamous	– 5		

Differentiation

Well Differentiated	– 0
Moderately Differentiated	– 1
Poorly Differentiated	– 2
Anaplastic	– 3
N.S.	– X
N.A.	– Y

Cell Type

Pleomorphic	– 0
Large Cell	– 1
Small Cell	– 2
Spheroidal	– 3
Duct Cell	– 4
N.S.	– X
N.A.	– Y

Ducts

No intraduct tumour noted	– 0
Intraduct tumour present	– 1
Introduct tumour alone	– 2
N.S.	– X
N.A.	– Y

Form 1 provides the necessary information used by the surgeons to make the initial decisions regarding eligibility and the stratification of the patients. On this form all the information regarding staging is placed. Also site of the disease is recorded. At the end of this form the treatment that is allocated is recorded. The treatment categories R1, R2, S1, S2 include a random division of both the radical mastectomy group and the simple mastectomy plus radiotherapy group into subgroups. At the time of trial design, many of the currently available statistical techniques were unknown, and the initial intention for the design of these subgroups was to allow some crude estimation of the effect of randomisation. These subgroups are ignored in this thesis.

Form 2, primary treatment form: This form records the basic data necessary to categorise patients on the treatments administered and also allows possibility of checking any violations from the allocated treatments of the protocol. It is important to stress that although this trial was initiated at an early time with respect ro randomosed trials, the concept of standardised treatment is clear and the information that was collected for assessing the diversity in terms of surgery and radiotherapy indicates good conformity with the protocol.

Form 3, Anniversary Record; all follow-up information was envisaged to be recorded on this form. Initially the major concern on the follow-up information in the protocol was that of devising a procedure by which all patients may be seen at a specific follow-up clinic, However it was requested that on radiotherapy case records information/

information on oedema of the arm, limitations of the shoulder movement chest pain and dyspnea due to post-irradiation pulmonary fibrosis and post-irradication skin atrophy should be recorded, for the purpose of a retrospective assessment of their occurrence. On the actual diagnosis of recurrence of disease it was requested that information should be provided for site, date of appearance of metastasis and the subsequent treatment. Finally on the anniversary form the cause and time of death is also recorded.

Form 4, pathology report form; this form keeps the information on size of tumour and the number of nodes found to be involved.

Initially 1099 patients were randomised according to one of the two trial options. Of these number 512 were found to have benign disease and so were withdrawn and so thereafter no data was collected on them for the purpose of the trial. The remaining 587 who had histological proof of carcinoma, were formed of 273 patients treated by simple mastectomy and X-ray therapy and 288 patients treated by radical surgery and 26 ineligibles.

In an initial analysis of the data, 87 cases who had breast cancer were withdrawn for reasons of violations of the protocol. Such violations included - case not belonging to proposed protocol population, case having ineligible form of malignancy and protocol violation due to inappropriate treatment. Decisions in regard to trial violation were made by a trial committee and it was decided to exclude all such cases from the final analysis. However in a review analysis of this data/

data some of the follow-up concepts were altered and the number of eligible patients for analysis was increased to 273 for simple mastectomy and radiotherapy and 288 for radical mastectomy. The main reason for this increase in the numbers of eligible patients for analysis was the introduction of a policy of comparison of patients according to the treatment allocated rather than the treatment performed. Table (6.3.1) gives details of the relevant reasons for the exclusion or inclusion of the original deviants and other cases.

Another area that at the time of the review of the follow-up data implied slight changes in the form of concepts adopted was in dealing with the assessment of response due to the treatment following the recurrence of the disease and general concepts such as local and metastatic disease. For this particular trial it is important to consider response to the treatment in terms of the delay in the development of the local disease. Similarly it is of interest to consider disease free survival and the time to metastatic recurrence. In terms of times after the recurrence of the disease it is generally expected that the treatment will not effect the survival of the patients a great deal after the detection of metastatic recurrence.

For the general recurrence categories the position of contralateral disease classification had been reviewed. In the past all secondary tumours were considered to be a metastatic recurrence, however with the new review some had been classed as new malignancies. Therefore it seems important that for future collection of any trial data, allowance must be made for the possibility of changing definitions. It seems that in general all categorisation relating to time such as response/

response or a duration of an interval, will eventually be indexed in more detail in terms of length of duration, extent and form. It may well be the case that metastatic disease will eventually be looked at in terms of extent and duration. This is going to become more common as multiple failure analysis becomes more common. In the past success or failure of surgical treatment has been assessed mainly by survival. With the new drug treatments assessment of response to disease and the detection of recurrence is playing an increasingly important role.

In the last section the data was described. In so far as prognostic information is concerned, the data is held on the initial examination form. Later in this chapter cross tabulations of different prognostic factors will be presented. It must be noted for some of the factors with a continuous scale,it may be desirable to categorise such variables. Age of the patient is such an example by which it is possible to split the population into different groups and then study survival distribution for each category.

For the events after treatment that may contribute to the understanding of the disease treatment process, there are 4 major events that we consider as important. These are local recurrence, metastatic recurrence, death and the last follow-up date. Clearly these events can produce in combination a large number of measurable periods. Using These periods it may be useful to study time to a particular stage of the development of the disease or it may be possible to stratify sub-groups of patients according to some prior event. For example, one can stratify the population according to time to local disease and observe/

observe the distribution of time from local disease to death.

Treatments according to option drawn.	RMX	SMX+XRT	Total
Treatment according to protocol	257	243	500
Correct treatment minor option modification*	6	14	20
Randomised therapy deviations.*			
Immediate XRT given though not indicated	4	1	5
Surgery only - died before XRT	0	4	4
Incorrect oophorectomy	7	9	16
Wrong surgery	14	2	16
Entered for this analysis.	288	273	561
Legitimate withdrawals after randomisation			
Benign disease	277	235	512
Ineligible but malgnant	13	13	26
Total patients randomised	578	521	1099

* A detailed survival study according to malignant withdrawals and exclusions will deviate from the general course of study. - I n terms of conclusions they do not effect the overall results.

Table (6.3.1)

A general discussion on the methods of construction of the likelihood functions for a stratified analysis of the data is given in chapter 7. Figure (6.3.5) represents a possible path for a progression of disease described in the above paragraph

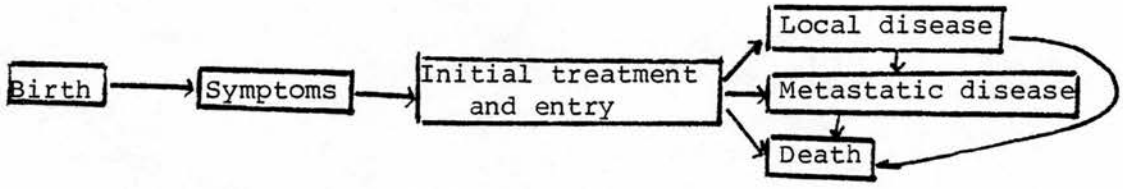


Figure (6.3.5)

In the first instance for the analysis of the data we consider cross tabulations of various categories. Then we study the failure distribution of the population in terms of survival times. For the estimation of the important prognostic factors we use Cox's proportional hazard model later in this chapter. In chapter 7 we consider time dependency of various prognostic indicators with different functional forms of time dependency. Using the same data later in chapter 7 we consider the effect of multiple events present in the time scale with the use of sem-markov hazard models.

6.4 Initial analysis with cross tabulation tables.

A good preliminary study of the data can be performed by a set of cross tabulations. The value of the Pearson chi-square can indicate a possible association between the distributions of the two factors. At this stage we are only trying to assess whether the data is distributed according to expectations of the previous studies. Appendix A presents important cross tabulations for the prognostic factors/

ors. The data is described by the following factors.

1. Menopausal status; premenopausal, menopausal, post menopausal.
2. Side of the lexion; right, left.
3. Site of the lesion: Medical half only, lateral half only, central both halves, whole breast.
4. Size stage: T1, T2, T3.
5. Node stage: N0, N1.
6. State stage: S1, S2, S3. (stratifying factor)
7. Skin involvement: not relevant. T1, T2, T3.
8. Pectoral muscle involvement. Not relevant, T1, T3.
9. Treatment option: Radical mastectomy, Simple mastectomy and Radiotherapy.
10. Disease status: Local & metastatic recurrence, Metastatic recurrence, Local recurrence, None.
11. State: alive, dead.

Most of the above factors (1 to 6) are related to prognostic state of the patient. Skin and pectoral muscle involvement refer to extent and site of early developments of the disease. Disease status and state finally refer to indicators of the progression of the disease at time of last follow-up. The option which defines the treatment allocated is also looked at for assessing distribution of the prognostic factors.

The first tables we will consider are the set (A.1) within Appendix A. As may be expected the largest number of patients are post menopausal (395). There are 38 menopausal and 163 pre-menopausal patients. When we consider the distribution of the 3 categories of menopausal states against other prognostic factors there are no statistically significant associations (except for age). The most significant value is for node status, with $X^2 = 3.0$, 2 d.f. giving the probability/

probability value of 0.22, which is not significant but indicates more node palpability with pre menopausal patients. X^2 = 4.9, 4 d.f.and p = 0.29 is obtained for T stage, indicating smaller tumours for pre menopausal patients and larger tumours for post menopausal patients.

There are 284 right side main lesions and 276 left side lesions. Side of the lesion is not an important factor in defining a patient even when we consider site of the main lesion categories. The most significant association with side is for T stage with X^2 = 3.3, 2 d.f. and p = 0.18, which is not significant. Site of the lesion has been categorised in a way that basically indicates the size of the tumour. There are 286 patients with their lateral half involved, 183 with medial half involved and 67, 22 and 2 with central, both or whole breast involved by the tumour respectively.

The T stage in fact give X^2 = 17.7, 8 d.f. and p = 0.02 and implies T3 (larger tumours) with central and both halves involved. Smaller tumours correspond with the medial half or lateral half alone involved.

Node status gives X^2 = 7.6, 4 d.f., p = 0.10 giving more node positive patients with central or both halves involved. (or perhaps basically with larger tumours)

Stage of the disease was defined to be a combination of the T stage and the Node status and the following tables clearly indicate this:-

T. Stage

		T1	T2	T3	Total
Node State	N0	35	273	67	375
	N1	21	124	40	185
	Total	56	392	107	560

S Stage

		S1	S2	S3	Total
T stage	T1	5	17	4	56
	T2	272	124	1	397
	T3	0	0	107	107
	Total	307	141	112	560

S Stage

Node state		S1	S2	S3	Total
	N0	307	0	68	375
	N1	0	141	44	185
	Total	307	141	112	560

T stage with node state cross tabulation gives a $X^2 = 2.0$ with 2 d.f. and $p = .37$

As presented in tables (a.1), stage is significantly associated with site. Stage 1 patients (good prognosis) more commonly have medial tumours or laterals half alone involved, which reflects the previously mentioned association with size.

Now/

Now we will consider the distribution of the prognostic indicators with the two arms of the trial. It must be reiterated that treatments were allocated before malignancy was diagnosed and therefore some of the patients were later removed for trial purposes since they had benign tumours. The total number of radical surgery patients is 288 and the simple mastectomy and radiotherapy patients are 272. The treatment options were also stratified according to age and the stage of the disease, but again the benign disease exclusions could affect this balance. Table (6.5.1) in fact shows that in most respects a good balance between the treatment groups resulted from the randomisation, with only slightly more T1 patients being allocated to the radical mastectomy group.

The next set of categories which were studied by the cross tabulations were the disease progress indicators and the spread of the initial tumour. In here we must emphasise that the disease indicator such as progression of the disease and final state of the patients will be studied more extensively in the next chapter. The present method of considering the cross tabulations does not allow an independent survival and censoring analysis and the X^2 values reported should not be interpreted as representing value of a treatment at this stage. These tables are presented within section (A.2) of the Appendix A.

Menopausal status shows a degree of association with the spread of the initial tumour indicators, in terms of skin and pectoral muscle involvement. The pre-menopausal patients have a lower level of skin and pectoral muscle involvement, $X^2 = 17.5$, 8 d.f.p=0.02 at/

at X^2 = 11.1, 6 d.f., p = 0.08 respectively. The status of the patients at the end of the study indicates X^2 = 25.8 with 2 d.f. and p< 0.0001 giving much better survival for the younger patients. Surprisingly the disease recurrence does not reach a significance with the present method, giving = = 0.11 with less local or metastatic disease among younger patients.

Side of the lesion does not play an important role for the final or initial disease progress. The highest value for the side is by skin involvement, X^2 = 3.6, 4 d.f., p =.46. Categorisation by site however plays an important role for the pectoral and skin involvement. X^2 = 27.8, 16 d.f., p = 0.03 and X^2 = 24.4, 12 d.f., p = 0.02 respectively for skin and pectoral muscle involvement. As we mentioned previously site is a reflection of the size of the tumour and in general medial or lateral halves involved produce less skin and pectoral involvement than other sites. The same pattern appears with the disease progression. More local or metastatic recurrence is noticed with both halves or central area tumours, X^2 = 21.0, 12d.f . p = 0.06. Following the above, lateral and medial half only, produce best number of survivors X^2 = 8.9, 4 d.f., p = 0.06.

T1 patients described earlier are a better prognostic group. By definition they have less initial skin and pectoral muscle involvement and finally less local or metastatic recurrence and therefore are better survivors.

T stage of the Tumour

	X^2	d.f.	p
Final disease condition	19.74	6	.0031
Final survival status	14.14	2	.0009

Node status does not produce any association with skin and pectoral muscle involvement. However with the node negative patients there is less recurrence of the disease at the end of the study, X^2=22.7 3 d.f., $p < 0.0001$. Further node negative patients are better survivors than the node positive patients

S1 cases are taken to be a good prognosis group and Stage 2 and 3 respectively worse. This is true both for the survival number and for the number of recurrences. Stage 1 groups give the highest proportion of disease free survivors $X^2 = 27.6$, 6 d.f., $p = 0.0001$, and a better number of final survivors, $X^2 = 7.05$, 2 d.f., $p = 0.03$. For the pectoral and skin involvement there is a defined relation between S stage and involvement

In the above discussion prognostic factors that indicate a significant skin involvement also indicate a pectoral muscle involvement. The two are very closely related and often coincide. However pectoral skin involvement does not have a significant association with disease recurrence, $X^2 = 13.8$, 12, d.f., $p = .31$ and $X^2 = 12.1$, 9 d.f., $p = 0.21$ respectively. Nor do skin and pectoral involvement show a significant association with the status at the end of the study, $X^2 = 4.29$, 4 d.f., $p = 0.36$ and $X^2 = 5.51$, 3 d.f., and $p = 0.13$ respectively

Disease recurrence and final status of the patients are very closely related as expected with metastatic recurrence producing a larger portion of dead cases. Treatment option and disease progress will be studied in later sections. In so far as the numberical distributions are concerned, treatment option is not associated with the skin pectoral/

pectoral involvement or disease progress. Disease recurrence and option give X^2 = 1.83 with 3 d.f. and p = .60. However final status of the patients seems to be related to option, X^2 = 8.2, 1 d.f., p.= 0.004.

So far the description of the data has been concerned with sets of categorical variables. The picture emerging is that T stage, S stage, menopausal status, node involvement and treatment options are factors producing the major associations with the categories of final disease status and survival status. Menopausal, T stage and S stage are related in effect to two important continuous variables namely age for menopausal status and size for T stage and therefore S stage. It seems proper to look at the distribution of these variables. Table (6.4.1) gives the means and the standard deviations of age and size, for all the 561 cases; menopausal status is one factor that is of course related to age. The distributions according to the table clearly indicate this. The size of the tumour is similarly related to T stage, and this is clearly shown by the table.

	AGE		SIZE		
	Mean	s.d	mean	s.d.	n
Pre menopausal	44.18	4.28	3.45	1.48	163
Menopausal	48.62	3.65	3.87	1.49	38
Post menopausal	59.91	5.85	3.62	1.40	359
Right side	54.27	9.19	3.71	1.44	284
Left side	54.87	8.73	3.49	1.41	276
Site medial	55.05	8.62	3.46	1.25	183
Site lateral	54.08	9.23	3.51	1.39	286
Site central	55.30	8.28	3.11	1.57	67
Site both	55.41	10.26	4.68	1.86	22
T1	51.44	9.17	1.64	.70	56
T2	54.77	8.93	3.51	1.01	397
T3	55.45	8.73	4.95	1.68	107
NØ	55.42	8.93	3.52	1.38	375
N1	52.84	8.78	3.75	1.54	186
S1	55.22	9.03	3.26	1.14	307
S2	52.51	8.76	3.39	1.080	141
S3	55.38	8.70	4.79	1.85	112
Skin involvement T0	51.72	9.88	3.33	.58	2
T1	50.47	8.30	4.83	2.48	12
T2	56.12	8.78	4.91	1.89	70
T3	56.21	8.50	4.028	1.50	36

	AGE		SIZE		
	Mean	s.d.	Mean	s.d.	n
Not involved	54.31	9.00	3.32	1.15	440
Pectoral muscle involvement TØ	57.73	7.70	3.75	1.98	14
T1	54.94	8.01	5.21	1.84	63
T3	55.97	8.91	4.05	1.51	43
Not involved	54.32	9.03	3.32	1.17	440
Radical MX	54.18	0.12	3.59	1.49	288
Simple MX + XRT	54.97	8.77	3.60	1.38	273
L + M Recurrence	54.47	8.74	3.98	1.44	114
M Recurrence	55.40	8.21	3.89	1.45	141
L Recurrence	55.62	7.62	3.81	1.28	16
No Recurrence	54.14	9.43	3.29	1.37	290
Alive	52.59	8.88	3.31	1.39	265
Dead	56.33	8.66	3.85	1.43	296

TABLE (6.4.1)

6.5 Survival time analysis of the old Edinburgh Trial data.

At randomisation the patients were stratified according to age and clinical stage of the disease. The table (6.5.1) indicates a balanced distribution of patients to the treatment options within each stratum. A comparison of the number of patients allocated to each treatment by the year of entry, also gives an almost uniform pattern of the accrual of the patients. There is a slight deviation for some years. However the reason is that the treatments were allocated prior to histology and so some patients were allocacted a treatment while they were non malignant and so had to be excluded from the trial.

An unstratified comparison of the survival of radical mastectomy patients and simple mastectomy patients gives a log rank, X^2 value of 10.04 with 1 d.f. which is highly signficant ($p = 0.0015$). Figure (2.5.1) gives a plot of the Kaplan and Meier survival probabilities of the two crude survival times.

Further for each strata a separate analysis of the survival times is performed. Certain of the subgroups indicate a highly significant difference between the survival probabilities. Table (2.5.2) refers to a summary of the analysis of the various strata using the logrank test. Generally speaking the treatment effect is consistent for the various subgroups with the most significant differences being indicated by the subgroups with larger number of patients. One interesting pattern that emerges, however, from the survival distributions is indicated in the survival plots of node status,

Age/

age, T stage and menopausal status with respect to survival time, Figures (2.5.1) to (2.5.12). The crude hazard rates of treatments showed proportional rates of failure for the two groups, Figure (2.5.2). This pattern is not so clear once we look at the subgroups of T stage age and menstrual status. The survival patterns can be explored further by a plot of the hazard rates. Clearly the plots indicate that depending on the time of observations of each subgroup the rate of failure is slightly different. At this stage it is not possible to explore this point further and assess the sifnificance of such a hypothesis, but only to observe it. In later chapters more relevant questions with more advanced statistical methods may be asked. These methods will be based on the validation of the proportional hazard assumption. Meanwhile the methods of the present chapter are based on the assumption of proportional hazards. One important point to note is that we have so far only stated slight differences in the significance levels of the different parametric survival families and the various non-parametric tests, as fitted to our data. We have not considered tests of the model assumptions in order to attach a significance level to the model differences.

The analysis so far, presented in table (2.5.2) indicates a poorer survival for all patients treated with simple mastectomy and radiotherapy. A categorisation according to stage indicates a significant difference in the same direction for the stage 1 patients and not a major difference between radical mastectomy and simple mastectomy and radiotherapy, for the stage two and stage three patients. By the age categorisation indicators, patients less than 50 year old do not show a significant difference between the two treatments, while

older/

ENTRY	R x Mx	S Mx + XRT	Total
1964	43	41	84
1965	61	48	109
1966	41	60	101
1967	36	36	72
1968	38	30	68
1969	33	35	68
1970	28	21	49
1971	8	2	10
	288	273	561
Age	54.2 ± 9.2	55 ± 8.8	54.6 ± 9.0
Size	3.6 ± 1.5	3.6 ± 1.3	3.6 ± 1.4
T1	37	19	56
T2	198	199	397
T3	53	54	107
NØ	199	179	378
N1	88	93	181
S1	164	143	307
S2	67	74	141
S3	57	55	112
Pre	89	74	163
Meno	21	17	38
Post	178	181	359

TABLE (6.5.1)

older patients (greater than 50 years old) seem to benefit from a radical surgery treatment. In so far as menopausal status is concerned , post menopausal patients benefit from a radical mastectomy treatment. Node positive patients with radical mastectomy show an improved survival while node negative patients treated by radical mastectomy or simple mastectomy and radiotherapy show similar survival patterns.

6.6 Analysis of the data using the Cox's proportional hazards model.

From the previous section there are certain points that we notice. One is that for certain covariates the relative hazard rate is dependent not only on the covariate understudy but also the time at which the covariate is looked at. That is there seems to be a suggestion that the effects of some covariates are not uniformly the same for the subgroups but are time dependent. There is also a slight form of inconsistency in the manner in which treatments effect patients with different prognostic status.

The effects described above are basically different forms of interaction that may be present in our data. The first set describes a possible interaction between time and a covariate while the latter describes an interaction between the two covariates. Although we have intoduced the idea of interaction in here we are not implying that the interaction is statistically significant and the difference between the significance levels in various strata may be attributed purely to the sample sizes of each subgroup. It is a point we will examine in later parts of this section in more detail.

The/

The model we are concerned with is of the form

$$\lambda(t_1 Z) = \lambda_0(t) \text{ Exp } (\beta_1 Z_1 + \beta_2 Z_2 + \beta_{12} f(Z_1 . Z_2)) \quad (6.6.1)$$

The above is a 3 parameter model representing a Z_1, Z_2 covariate interaction. Alternatively a time dependent covariate model may be represented by

$$\lambda(t_1 Z) = \lambda_0(t) \text{ Exp } (\beta_1 Z_1 + \beta_2 f(Z_1, t) + . \ . \ .) \quad (6.6.2)$$

Depending on the form of the variable under study, the number of covariates will vary and we may end up with more than 3 and 2 covariates in the above models respectively. For example T stage represents size of the tumour and is composed of 3 categories. For a representation of such a variable we require two variables say Z_1 and Z_2 giving

$Z_1 = 0, \ Z_2 = 0$ For T1

$Z_1 = 1, \ Z_2 = 0$ For T2

and $Z_1 = 0, \ Z_2 = 1$ For T3

Using the above parameterisations we can test the significance of T stage values as a prognostic indicator without making assumptions on the order level of the categories. An alternative approach would be to allow a variable Z with values -1, 0 and 1, to indicate a linear categorisation of the T stage values.

In general the numerical values attached to the quantitative covariates should not be a major problem. For the example of T staging there may be a slight loss of efficiency with the latter approach if there is a difference in the pattern of the influence of the size. On the other hand using two covariates for removing the effects of size in the former description with Z_1 and Z_2 is less convenient and time consuming if the effect of size is uniformly the same/

same on survival. For a broad purpose of exploratory analysis of the data thoughtful parameterisation allows the possibility of study of a large number of independent variables without introducing a large number of covariates.

Initially we are only interested in the main effects of the described parameters in exploring the variability of the failure time from randomisation to time of death. However later we will study other failure times and intervening events. The method we adopt to explore the data is in some ways similar to what is usually termed as a stepwise regression method, by which certain levels of introducing variables is adopted for inserting covariate estimates into the model. We set the limits to be probability value of 0.100 for entry and probability level of 0.150 for removal. At each step we estimate all relevant parameters and consider the parameters that are significant and introduce only the most significant into the model. At each stage, if a parameter estimator that is already in the model becomes non significant (because of its association with variables added to the model), we will remove the newly non-significant effects. We will deviate from the above approach in our exploratory approach by considering certain strata variabilities separately. Further unlike the initial stages where we will study the parameters in relation to main effects of covariate only, in the next stage we will consider models of the form with main effects and a corresponding prognostic and treatment interaction term. One point to note is that at any time we mention size and age covariates in this chapter, we will use a normalised transformation of the effect by letting,

$$Z_{ij} \longrightarrow (Z_{ij} - \text{mean}\ (Z_{j})) \ /\ \text{S.D.}\ (Z_{j})$$

where/

where Z_{ij} refers to a covariate for a patient, Mean (Z_j) and S.D.(Z_j) refer to mean and standard deviation of a particular covariate for all patients.

In the first stage of the stepwise procedure both age and menopausal status are highly significant ($p < 0.0001$), with the menopausal status being a slightly stronger prognostic factor than the actual age parameterisation.

We thus continue with a categorised analysis of the relative risks, due to age. In stead of considering age as a continuous variable we categorise the scale into younger than 50 years of age and older than 50 years. The value of the β relative risks are then noted. As we pointed out the relative risk value is signfiicant at a probability value $p < 0.0001$,. Now in a comparison of the age effect tested by the two methods, the age effect as a continuous variable gives a β value of 0.3198 and a standard error of 0.0798, while as a categorised variable almost coinciding with the sectors of menopausal status it presents a β value of 0.3321 and the standard error of 0.0745. We perform an analysis by stratifying the data into premenopausal and post-menopausal groups. The two groups are then analysed by assessing the age effect on them separately. The analysis indicates that the age effect reduces to insignificant levels. Later we will consider the age effect with time dependent parameters so that rather than categorise the age variability we may obtain a similarly flexible indication by a parametric function of the age effect.

Following/

Following the age and menopausal status of the patients the most direct prognostic factors are the actual size ($p < 0.0001$), T stage ($p = .0004$) Node status ($p = .0007$), treatment option ($p = .0016$) and S stage ($p = .0024$). Clearly treatment is a significant effect and is of special importance to our study. At this stage we continue with the stepwise regression as described. Later we will consider forcing the treatment effect in the first step so that we may check consistancy of the model in a situation where treatment effect has a priori precedence. Site of the disease seems to play a marginal role only due to the lateral half involvement ($p = .11$)

Before dispensing totally with the various sites of disease indicators, we consider a stratified analysis for each of the different sites of the initial tumour, we perform a stratified analysis based on each single site as defined in the section 6.4, and the set of covariate effects that have been considered significant up until now. Without presenting too much detail once again age and menopausal status play the most important role in defining the survival rates. The relative risk rates are closely related for each of the strata and there is an indication that the age and menstrual status effects are consistently in a similar direction within the various sites.

The rest of the covariates we have been interested in at this stage for this particular failure time do not reach a significant level. The covariates that we will ignore for the rest of the analysis of this particular event are, side of the initial lesion, other/

other sites of lesion and patient conception of the time from first noticing the tumour to the time of the operation.

The most direct prognostic indicator is the menopausal status of the patient. We introduce this variable into the general model of the Cox's approach.

$$\beta_{men} = 0.03558 \qquad \text{S.E.} = 0.0720$$

with this model the pattern of the significance of the remaining prognostic factor changes to some extent. Actual size remains the most important factor in describing the remaining variability in survival ($P < 0.0001$). Node status becomes more significant ($P = .0002$) than the T categorisation of the size of tumour ($P = .0008$). The S stage is still significant at ($P = .0018$) and finally for this stage of the stepwise regression, the treatment option produces a significant contribution with $P = 0.0029$.

The menopausal variable clearly removes the contributions of actual age in explaining survival totally, with the significance value reduced to, $P = .42$. The premenopausal patients are generally better survivors than the menopausal patients or post-menopausal patients.

We can also state that larger tumours are an indication of a worst survival. A stratified analysis based on node status indicates that this statement is true for both node negative and positive patients. In terms of menopausal status a stratified analysis of menopausal status with covariate analysis of size indicates that the/

the larger size of the initial tumour is consistently an indicator of bad prognosis, for both pre and post menopausal patients, with the effect of size being more significant among younger patients with $P = 0.004$ and $P = 0.021$ as the significance level of size in premenopausal and post-menopausal patients respectively. Similarly a covariate analysis of node indicates that the node positive patients are worst survivors. The effect is once again more significant among the premenopausal group, with the significance levels $P = 0.0007$ and $P = 0.0011$ for the pre and post menopausal groups. Thus we may summarise that the size and node are important contributors to the various survival patterns. However there is a suggestion that their effect is more significant among the younger patients. It seems that by introducing more factors into the model apart from the menopausal status we can still define the survival time more precisely. The most effective prognostic factor after introducing menopausal status was the actual size. Therefore it is the next term to be introduced into the model. The previous indications by the stratified analysis also are suggesting that size is having a consistently similar effect in sense of direction for both pre and post menopausal groups.

$\beta_{men} = 0.3643$ S.E. 0.0723

$\beta_{size} = 0.2079$ S.E. = 0.0402

The above coefficients are positive and therefore indicate that smaller tumours as may be expected are better survivors. It is rather interesting to note that among age and menopausal status, the more important contribution was the menopausal status which is a categorised variable. However, in the case of the size of the tumour some efficiency is lost by categorisation of size into different T stage values. After introducing/

introducing size of the initial tumour, significant value of T stage reduces to P = 0.60. Also the stage significance reduces to P = 0.25. The only covariates which still show significant levels are node status and treatment options with P = 0.0004 and P = 0.0017 respectively.

In the next step we introduce the node status covariate

β_{men} = 0.3864 S.E. 0.0729

β_{size} = 0.1981 S.E. = 0.0399

β_{N} = 0.4315 S.E. = 0.1205

Node negative patients are a better prognostic group than the node positive patients. After introducing the 3 major prognostic factors namely, menopausal status, actual size of the tumour and node histology, the only factor remaining that still shows a significant contribution to survival is the trial option, P = 0.0017. With stratified analysis we study the two effects that do not show any significance, namely T stage and the S stage of cases, to make sure that the reason for this loss of significance is not due to the assumptions of the proportional hazards. The T stage is only a function of the size and this is reflected in the stratified analysis of T stage in the way in which the variability in survival due to the size effect reduces to insignificant levels. For menopausal status we obtain similar β estimators of .34, .38 and .39 on the relative risk factor within T1, T2 and T3 strata respectively.

By the definition of the S stage, node status has a direct role in defining the staging systems. We note that the 3 strata of stage/

stage give similar directions of association for the menopausal status and size of the tumour in terms of survival time.

It is also important to note that the significance of different prognostic factors generally vary with the introduction of different terms into the model. This is due to their inter-relationship. However, the significance of the treatment option becore introducing any term in our model is P = 0.0016 and after introducing 4 terms the significance is P = 0.0017 which is very close to the original value reflecting the similarity of the treatment groups with respect to the distribution of other covariates.

The coefficients of the model after the introduction of the option indicators are:-

β_{men} = 0.3847 S.E. = 0.0730

β_{size} = 0.2040 S.E. = 0.0401

β_{N} = 0.4319 S.E. = 0.1206

β_{Option} = 0.3656 S.E. = 0.1169

The two treatment options are radical mastectomy and simple mastectomy with radiotherapy, and the model indicates that patients may benefit from radical mastectomy in terms of their survival.

We finally perform additional analysis in order to be certain that the general final model is representative for the subgroups and categories that it represents. The main aim would be to check on the possibility of the extistence of smaller subgroups showing a pattern of survival/

survival rates which is different in direction to the general model. We perform an analysis based on a stratified analysis of each category of a covariate with respect to other covariates. At this stage we ignore treatment option but later do formal tests on them. Altogether there is a slight deviation based on the sample size of each covariate set, however it is noted that the stratified analysis does not suggest that there exists a subgroup with a significantly different suggestion of prognostic value in the opposite direction to the general model.

Up until now in the study of the relative effects of the prognostic factors we have introduced the covariates according to their level of significance. Study of the treatment option however is the major objective of a trial. Now in the initial step of the categorising of the patient population we introduce the treatment option. Further for each main effect prognostic factor we introduce a set of first order interaction covariates that act multiplicatively between option and the other prognostic factors in terms of the model (6.6.1). It is important to note that if our intention at this stage was purely a study of the interaction effects, then we would have continued with inserting interaction effects into the above 4 covariate models. However as an alternative to the previous stepwise procedure, we introduce the option effect in the first step to study behaviour of the different models and also check the consistency of the final model. We use a generalisation of the model (6.6.1) as,

$$\lambda(t,Z) = \lambda_0(t) \; \text{Exp}\left(\beta_1 Z_1 + \beta_2 Z_2 + \beta_{12} Z_1 Z_2 + \ldots \right)$$

where β_{12} is introduced for assessing an interaction between treatment and/

and a prognostic effect. The two variables size and age are continuous and may be time dependent. We will deal with these models in Chapter 7.

In what follows in this section we will consider functional forms of $f(Z_1, Z_2)$ from (6.6.1). In the case of binary categorisation of a variable a functional form of $Z_{12} = Z_1 \cdot Z_2$ is sufficient (as we have used 0, 1 to indicate the two levels.) so long as enough consideration has been given to ease of interpretation. In fact most of the variables we will be considering are of the above binary form. The continuous variables like age and size can also be transformed to binary categorisations by considering the high and low levels of their scale and independent dummy variables. Later in this chapter we will consider continuous form of size and age variables. In these conditions a continuous parametric representation may be useful. We will later consider a possible functional form of age and size in the presence of binary treatment effect. Namely models of the form

$$\lambda(t, Z) = \lambda_0(t) \operatorname{Exp}\left(\beta_1 Z_1 + \beta_2 Z_2 + \beta_3 Z_3 + \beta_{22} Z_2^2 + \beta_{33} Z_3^2\right)$$

and

$$\lambda(t,Z) = \lambda_0(t) \operatorname{Exp}\left(\beta_1 Z_1 + \beta_2 Z_2 + \beta_3 Z_3 + \beta_4 Z_2 Z_3 \sqrt{Z_2 + Z_3}\right)$$

where Z_1 refers to treatment, Z_2 to age and Z_3 to size.

Initially we consider categorised variables with binary interaction effects, $Z_{12} = Z_1 \cdot Z_2$ for the different subgroups. We introduce the effect of treatment option into the total sample. The main effect of option is therefore represented in the relative risk function by

$$\beta_{\text{option}} = 0.3677 \qquad \text{S.E.} = 0.1168$$

Now the overall variance due to other prognostic main effects increases and so their significance value is reduced. However their relative significance does not change to the one prior to introducing treatment option.

Menopausal status ($P < 0.0001$) and age ($P < 0.0001$) are the most important factors followed by size ($P < 0.0001$), T stage ($P = 0.0008$), Node status ($P = 0.0009$), and S stage ($P = 0.0044$). Sites of the tumour that have lateral half involvement are again only marginally significant ($P = 0.06$). At this stage prior to introducing other main effects is not possible to try to interpret the value of the interaction effect parameters. It must be noted that interaction effects prior to introduction of the main effects do often show a significance, with a probability value slightly lower than that of the corresponding main effects.

An explanantion is in order in regard to the value of such interaction effects. The main reason being due to the fact that the variability due to the main effect is not removed yet. We further note that the significance of probability level of the interaction paramters are lower than their main effects. The major variability is due to the main effect of age at ($P < 0.0001$) and its interaction effect $P = 0.0021$. In terms of β estimators of the relative risk function we also note less significant values for the interaction effect while the actual magnitude for the direction of the effect is always positive and at a lower level. This pattern implies that the only type of interaction effect that we may expect to find will have a positive multiplicative/

multiplicative effect.

Once again the major variability is due to age and its interaction effect with $\beta = 0.0042$ and $\beta = 0.0042$ respectively, closely followed by menstrual status main effect at $\beta = 0.3492$ and its interaction effect $\beta = 0.3022$. It seems reasonable to add further prognostic factors that show a significant probability value. We thus introduce the main effects into the model one by one depending on their relative significance at each stage. First we introduce menopausal status with option.

$$\beta_{option} = 0.3475 \qquad S.E. = 0.1168$$

$$\beta_{men.} = 0.3492 \qquad S.E. = 0.0720$$

By this covariate all variability due to age is again also explained. The interaction effect for age option and menopausal option also become insignificant. Size ($P < 0.0001$), N ($P = 0.0002$), T ($P = 0.0014$) S($P = 0.0026$) are all significatn. Site with lateral half tumours also increase in significance ($P = 0.028$). Now we introduce the size of the tumour.

$$\beta_{option} = 0.3650 \qquad S.E. = 0.1169$$

$$\beta_{men} = 0.3598 \qquad S.E. = 0.0723$$

$$\beta_{size} = 0.2045 \qquad S.E. = 0.0406$$

The only remaining factor that makes a significant contribution is node histology. We also note that by introducing main effects of the prognostic factors the interaction effects are also explained. Hence finally we obtain the same models with the same prognostic factors as/

as the previous approach, since none of the interaction effects have contributed substantially to the explanation of the survival rates.

$\beta_{men} = 0.3847$ S.E. = 0.0730

$\beta_{size} = 0.2040$ S.E. = 0.0401

$\beta_{N} = 0.4319$ S.E. = 0.1206

$\beta_{option} = 0.3656$ S.E. = 0.1169

In the above discussions the general conclusion is that menopausal status, Size of the initial tumour and node histology are the main prognostic factors that define a survival time for a group of patients. However, size with the T stage and menstrual status with age also show a high level of dependence and introducing one factor generally compensates for the information due to the other factors. The same may be said for the stage of the disease. Stage is a combination of the node and size categories. However it seems that a better assessment may be made by introducing node and size separately. In fact on considering a model of the form with menstrual status, treatment option and the effect of S stage represented by two covariate indicators we obtain

$\beta_{men} = 0.3921$ S.E. = 0.0782

$\beta_{S1} = 0.3952$ S.E. = 0.0281

$\beta_{S2} = 0.3161$ S.E. = 0.9791

$\beta_{Option} = 0.3721$ S.E. = 0.1291

There is not a major difference noted for the new option and menopausal status estimator. We further introduce 2 interaction effects of option with β_{S1} and β_{S2} parameters and they do not reach

a/

a significant level. In terms of interpretation the previous model was probably more straight forward than the present approach since by the latter, one must always refer back to the interpretation of S1 and S2, while the model of size and node give a clearer interpretation.

Now we use a method which is alternative to the stepup procedure and is usually termed as a step down procedure. It is again a study of the relative significance of eaeh factor when other factors are present. We begin with fitting a model to the data in which all the prognostic factors have been introduced. In the consequent steps we remove the effects one by one depending on the level of significance, that the particular estimator contributes in regard to defining the variability of the data. As before we will deviate from the standard procedure however by looking at different strata at each stage. In the stepdown procedure we will only consider the main effects, since up until now there has not been a major inconsistency in the direction of the effects. The probability levels we adopt are again 0.100 for re-entry of a term previously removed and 0.150 for removal of an effect from the model. The main purpose in using this approach is to check on the consistency of the final model of the last section in being able to describe the variability in the survival rates. Further by an extensive comparison of different covariate effects we describe the improvements in the estimators of the treatment parameters and also the ease of interpretation for each prognostic effect.

Finally the most relevant significant levels in the stepdown procedure/

procedure are the significance levels for the removal of effects from the model (unlike the step-up method). As before the value of the significance levels for each effect changes at each step.

The first model we consider is the full model containing all covariate effects that are suspected to play a part in defining the survival times. The following model is hence obtained.

β_{men} = 0.3231 S.E. = .1138

β_{side} = 0.0014 S.E. = .1183

β_{site} = 0.9933 S.E. = .0731

β_{size} = 0.1580 S.E. = .0508

β_{T} = 0.3847 S.E. = .2311

β_{N} = 0.6381 S.E. = .1772

β_{S} = 0.2510 S.E. = .1538

β_{option} = 0.3787 S.E. = .1191

β_{size} = 0.0077 S.E. = .0110

β_{year} = 0.0008 S.E. = .0028

Now by reviewing each of the above terms from the model once again we can assess the relative importance of each factor with the above restrictions. Basically there is no inconsistency with the previous approach and the model considers the same factors as important prognostic indicators. However there are slight differences in the order of their significance. Node histology is the most significant indicator (P = 0.0002) followed by option (P = 0.0014), size (P = 0.0016), menopausal/

menopausal status (P = 0.0045), T(P = 0.0867) and S(P = 0.0936). The rest of the covariates are treated as factors contributing insignificant levels. These factors are respectively, lateral half involvement, age, symptom and side of the tumour. The value of the important prognostic factor seems not to change if the insignificant factors are removed one by one. However, value of the covariates for each significant prognostic factor is slightly different and generally the probability values are higher. The model we obtain after the removal of the non-significant factor is -

β_{men} = 0.3788 S.E. = 0.0731

β_{size} = 0.1653 S.E. = 0.0505

β_{T} = 0.4054 S.E. = 0.2295

β_{N} = 0.6521 S.E. = 0.1739

β_{S} = 0.2617 S.E. = 0.1529

β_{option} = 0.3585 S.E. = 0.1171

Giving menopausal status (P<0.0001) as the most significant factor, Node (P = 0.0001) as the second factor, size (P = 0.0009) and option (P = 0.0022) as significant, and T(P = 0.0687) and S(P=0.0912) as marginally significant. At a more conventional 5% significant level, we can also remove T and S staging. This reduces the model to the initial 4 covariate model. The slight indication for the T significance level shows that size may not have a linear effect in the time scale. This will be looked at more closely later in this section.

Now we continue with the analysis for the assessment of the important/

important prognostic factors for the period of randomisation time to the metastatic spread of the disease. The study of the time to the development of metastatic disease is also important in that it defines the spread of the disease more directly and unlike the time to death is not affected by factors such as death from old age, or other causes. Considered singly initially the major prognostic factor is the size of the initial tumour ($P < 0.0001$) followed by the closely related S stage ($P < 0.0001$) N node status ($p < 0.0001$) and T stage ($P=0.0001$). However the menopausal status is only important after the size effect with ($P = 0.0001$) and age is now even less significant at ($P = 0.0493$). Option is again significant although with a loss in significance. The site of the tumour being lateral however seems to play a more important role with ($P = 0.014$). If we introduce the size effect into the model then the effects due to T stage ($P = 0.4$) and S stage ($P = 0.03$) are reduced. The effect of node status ($P = 0.0001$) remains highly important. The relative importance of age ($P = 0.11$) and option are both reduced. Site of the tumour being in the lateral half is significant only marginally ($P = 0.05$).

$$\beta_{size} = 0.2421 \qquad \text{S.E.} = 0.0426$$

In the study of time to death menopausal status of the patient played the most important role. In the study of the variability due to time to metastatic disease, the most important contributions are made by the size of the tumour followed by the node histology.

$$\beta_{size} = 0.2255 \qquad \text{S.E.} = 0.0421$$

$$\beta_{N} = 0.5025 \qquad \text{S.E.} = 0.1280$$

The menopausal status of the patients is then highly significant ($P=0.0001$)/

(P = 0.0001) only after the above two variables. The menopausal status is then followed in significance level by age (P = 0.02), however it is reasonable to assume that the age effect will be explained by the menopausal status. Finally treatment option is then followed with (P = 0.04). After the introduction of size and node histology, the previously major contributions of T and S reduce to insignificant levels. This represents once again a slight deviation from the analysis of time to death, since the stage of the disease with node and size effects present was showing a marginal significance. Therefore perhaps it is indicative of a node status or size interaction within the time scale to death.

Finally we introduce the treatment option.

β_{size} = 0.2234 S.E. = 0.0431

β_{N} = 0.5043 S.E. = 0.1321

β_{men} = 0.2967 S.E. = 0.0751

$\beta_{options}$ = 0.2641 S.E. = 0.1432

Now, we approach the study of the response variable (time to metastatic disease), with the actual treatment forced into the model. At the same time we are interested in the study of the effects of any possible prognostic factor with the option interactions that may be present. With only the treatment option present we obtain.

β_{option} = 0.2532 S.E. = 0.1254

Once again in consistency with previous study of time to metastatic disease actual size of the tumour plays the most important role, followed by node histology, T stage and S stage.

$\beta_{size/}$

β_{size} = 0.2440 S.E. = 0.0426

β_{option} = 0.2638 S.E. = 0.1254

Following the introduction of the size and option effects we note that the interaction effects for the two variables gives P = 0.9.
Node histology also has a significant effect while size and option effects have both been introduced into the model. However once again unlike the time to death response variable menopausal status is relatively less significant

β_{size} = 0.2283 S.E. = 0.0422

β_{N} = 0.4979 S.E. = 0.1271

β_{option} = 0.2560 S.E. = 0.1255

Again no interaction effect is noted for the node status and option. The most significant factor remaining is the menopausal status. With the entry of this later factor the major factor that influences survival are once again the major factors that influence time to the metastatic disease.

β_{men} = 0.2967 S.E. = 0.0751

β_{size} = 0.2234 S.E. = 0.0431

β_{N} = 0.5043 S.E. = 0.1321

β_{Option} = 0.2641 S.E. = 0.1432

Once again we observe the influence of possible interaction effects with the treatment option and again there seems to be none acting. The site of the tumour being lateral was a factor that for time to death was initially marginally significant and with the removal of other major factors becomes less and less significant. In the study/

study of the time to metastatic disease there seems to be a similar trend present, and at the end lateral half involvement is not significant with a probability value of P = 0.09. The next response variable that we study in relation to the exploratory value of the prognostic indicators is the time from randomisation to the local progression of the disease. Once again we use a stepwise procedure approach similar to the last section. One by one we introduce important prognostic factors and observe their effects. Finally we allow test of interaction between the treatment main effect and the prognostic indicators. The final model of the relative risks parameters that we obtain are similar in terms of order of importance of the prognostic covariates to the parameters obtained for the models of time to metastatic disease. However the magnitude of the various estimators are different. The final model is thus composed of parameters

$$\beta_{men} = 0.2481 \qquad S.E. = 0.0621$$

$$\beta_{size} = 0.2013 \qquad S.E. = 0.0510$$

$$\beta_{N} = 0.5518 \qquad S.E. = 0.1511$$

$$\beta_{Option} = 0.2421 \qquad S.E. = 0.1080$$

Before we finish with this chapter which has been based on categories of the various variables we consider functional forms of the two major prognostic indicators, namely size of the initial tumour and age of the patient, when they are both considered to be continuous variables. We will initially consider a model of the form with the age and size and no interaction in the model with the treatment effect present. A main effect model of size and age with treatment gives a relative risk function with the following parameters

β_{Option} = 0.3511 S.E. = 0.1172

β_{Age} = 0.0068 S.E. = 0.0171

β_{Size} = 0.1990 S.E. = 0.0432

In the earlier discussion in dealing with categorised size variability and age represented by menstrual status we concluded that there is no suggestion of an interaction between size of the initial tumour and age of patients at entry in describing the survival times. However we noticed a slight improvement in the treatment effect estimators of the β parameters. We will now introduce a model of the relative risks by which we assess the age and size effect in a continuous manner. The first model we consider is a relative risk function given by linear effects of age and size as well as their independent quadratic effects, giving the hazard rate

$$\lambda(t_1 Z) = \lambda_0(t) \text{ Exp}(\beta_1 Z_1 + \beta_2 Z_2 + \beta_3 Z_3 + \beta_{22} Z_2^2 + \beta_{33} Z_3^2)$$

when 1 refers to treatment option, 2 to age and 3 to size, and giving the following parameter estimates

β_{Option} = 0.3578 S.E. = 0.1151

β_{Age} = 0.0068 S.E. = 0.0171

β_{Size} = 0.1987 S.E. = 0.0451

β_{22} = 0.0003 S.E. = 0.0241

β_{33} = 0.0170 S.E. = 0.0642

Clearly there is no suggestion of size and age playing a quadratic role in the explanation of the survival times.

Finally/

Finally we consider an interaction of the continuous age and size. In the following model of the relative risk function we will consider an interaction of the form in which not only a multiplicative effect of the prognostic indicator is present but also simultaneously there is an additive effect present. We then have a model of the form

$$\lambda(t,Z) = \lambda_0(t) \; \mathrm{Exp}(\beta_1 Z_1 + \beta_2 Z_2 + \beta_3 Z_3 + \beta_4 Z_2 Z_3 \sqrt{Z_2 + Z_3 \times 10})$$

In the above model, 1 refers again to treatment, 2 to age, 3 to size and 4 to the interaction parameter. Further for the convergence of the maximum likelihood estimator we will transfer the actual covariate indicators so that they are almost normalised. That is we let

$$Z_2 = (\text{Age} - \text{Mean Age}) \; / \; \text{Standard deviation of Age}$$

$$Z_3 = (\text{Size} - \text{Mean size}) \; / \; \text{Standard deviation of size.}$$

and

$$\sqrt{Z_2 + Z_3 \times 10} = [(\text{Age} + 10 \times \text{size} - \text{mean (age} + \text{size} \times 10)) \; / \; \text{Standard deviation of (Age} + \text{size} \times 10)]^{\frac{1}{2}}$$

We thus have

β_{Option}	= 0.3912	S.E. = 0.1018
β_{Age}	= 0.0067	S.E. = 0.0168
β_{Size}	= 0.2012	S.E. = 0.0419
$\beta 4$	= 0.0003	S.E. = 0.0041

Once again there is no suggestion of an additive with multiplicative relative risks of size and age interaction. It seems useful to consider various interactions of the prognostic effects when they are of /

of continuous form. A better approach and to some extent related approach is adopting time dependencies of the continuous effects. In Chapters 7 and 8 we will consider such time dependencies. In the next chapter we will also consider multiple risk approaches. At this stage we will only mention the relevance of the present approach to the multiple risk and continue with the analysis in the next chapter after some further methodological developments. Apart from the response variables looked at so far there are some other response variables that are of interest, like the time from local disease to death or the metastatic disease to death. For such variables we will require an adjustment of the initial time segment for the proper assessment of the treatment and covariate effects. With the approach we have followed up until now one can do such an adjusting by stratifying the response variable according to time from randomisation to the present event. We may then have models of the form

$$\lambda_j (t,Z) = \lambda_j(t) \text{ Exp } (\beta,Z)$$

Where in the above example λ_j refers to hazard rate for time from metastatic disease to death, then j signified categories of the time to metastatic disease. Clearly this is an example of a situation in which we have multiple events. An alternative is to use time dependency as described above rather than stratification of the basic line hazard $\lambda_j(t)$. These two latter considerations make the proportional hazards model a flexible method for such studies.

CHAPTER 7

MULTIVARIATE RISKS

Once we confine the method of analysis to the approaches discussed in the previous section we may not be able to assess the effects of various treatments in the presence of progression, metastatic disease, or other forms of intervening events most efficiently. The main parts of this chapter deal with situations of a trial with multiple events within the time scale of study. Initially we will deal with models in which cases move from one state to another. In particular we refer to those of semimarkov models and the analysis of data in groups. Such models allow a quick analysis to be performed and are useful for an exploratory analysis. More importantly for this thesis they make a good conceptual shift from models of the previous chapter to a situation of multiple risks. Later we will deal with the development of the proportional hazards approach with a functional form of a time dependent factor for the intervening event.

7.1 Initial developments of the methodology.

In 1959, Bartlett in a paper on the impact of the theory of stochastic processes on statistics, stated, "correct specification of statistical problems has only become possible in terms of stochastic process"/

process". Earlier in 1950 Neyman had written a chapter on "competing risks" in his text book on statistics and probability theory. These methods were inferred from a relatively simple illness and death model. His original ideas on this work had arisen from works similar to those of Daniel Bernoulli which were mentioned in the introduction. In particular Neyman was interested in the problem of assessing risks of dying from breast cancer by a comparison of risk of dying from cancer after treatment with that of dying from other causes or being lost to follow-up. This method of Neyman often referred to as Fix-Neyman clearly differs from ordinary survival analysis in that, in the latter there is only one transient state (entry) and one obsorbing state (death), while in the present context one is concerned with different causes of death, progression, regression and possibly other stages.

When there are several end points present, there is a general and almost traditional way of analysing the data based on 3 assessments:- crude probability, partial crude and net probabilities of survival. These concepts have been used by people who have been studying failure time in occupational health studies or the epidemiological studies of chronic diseases. However comment made by Stormer et al (1980) expresses fully the associated problems. "There is now mounting evidence in the biomedical literature to suggest that experimental methodologies are deficient when applied to the investigation of chronic diseases. Chronic disease appears to be substantially more complex than acute disease in several respects: chronic disease is dynamic. It represents the long term cumulative effects of interactions between a host biological system and the surrounding/

surrounding environment. The environmental influences are not static, so chronic disease acquires a time varying characteristic... it is possible that any combination of the above factors be influencing a trial to a significant extent". Although our position does not go along with some of the comments made in the above regarding the generality of the environmental effects, one aspect of the statement holds even within randomised trials; the fact that the complexity of chronic disease requires complex processes by which time varying characteristics may be incorporated.

Such problems initially were related to an approach that was named competing risk. J. Cornfield (1957) on competing risks and clinical trials puts the approach in the following perspective of the language of cause and effect. He defines a formal effect as if individuals died from some extraneous cause and had no chance of dying from cause under study. Further empirical effects relate to those who died from extraneous causes and might have a probability of developing the disease of interest, which differs from probability of those who died from disease of interest. The latter effects are then suggested to be analgous with withdrawal at time of the analysis.

C.L. Chiang (1964) develops the concept of probability for competing risks in a formal manner by defining 3 separate functions:

(1) The crude probability of survival.

$Q_i(t)$ = (probability that individual alive at t will fail in t + h from cause i in the presence of all other risks).

(2) The Net probability of survival

$H_i(t)$ = (probability that individual alive at t will fail in t + h if risk i was the only risk acting).

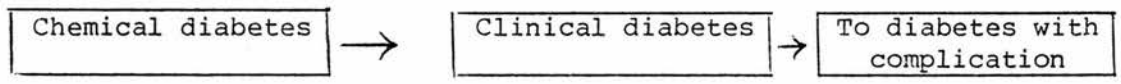
(3)/

(3) The partial crude probability of survival.

$Q_{i,j}(t)$ = (probability that individual alive at t will fail in t + h, from cause i, if cause j was eliminated as a cause of death).

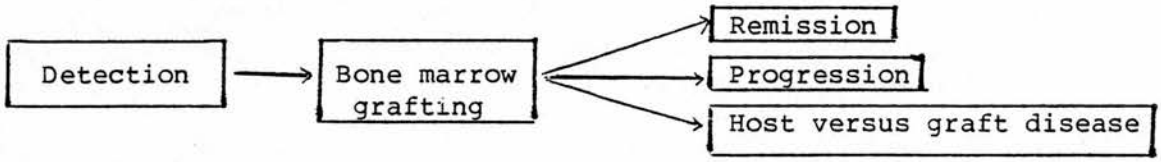
In a discrete situation we may divide the time scale into segments and apply life table approaches. For a continuous case based on distributional assumptions, parametric methods may be used. C.L. Chang (1976) has extended the method and as an alternative approach has used Fix-Neyman model for the two transient states and more than two absorbing states. Such a model of two transient states is a realistic model in which different patients with separate prognostic values can be placed on different transient states. Individuals may thus move from one transient state to another until in a finite time they enter one of the absorbing states. An adequate form of explaining such a phenomenon would then be based on the recorded number of transitions and the times of the transitions between any two states. C.L. Chang (1979) developes this method further for the particular case of chronic conditions. He makes the observation that, the disease advances with time from mild through intermediate stage to severe to death. The cases may die in any one of these states. A few practical situations where the above assumptions can aid in the analysis are given below. Later in this chapter we will describe a more natural method for analysis.

(1) Definition of stages in diabetes.

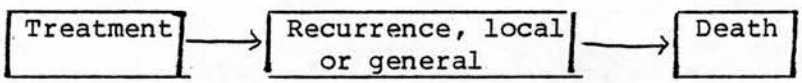


2/

(2) Progression and treatment of leukemia



(3) Breast cancer



What the above examples have in common is that the processes are always irreversible; this is an observation which is useful in the further development of the methodology. One further restriction that has often hindered the general use of such approaches has been that of robustness. It is often possible to develop a general maximum likelihood function for the paths of progression, however if one considers distributions more complicated than the exponential distribution, the method of maximum likelihood estimations becomes unrealistic in terms of the quantity of calculations.

It is clear that what may be required for our form of problem is a model that takes care of the problem of censoring and uses the assumptions of irreversibility. Such an approach is suggested by Lagakos, Summer and Zelen (1978) by which a non-parametric method based on ranks of the sojourn times between the states is used. The main purpose for the use of this approach is that of analysis of the data with an exploratory approach and a better description of the semimarkov models. However later we will extend the methodology to the proportional hazards models in which similar tests can be incorporated into the functional forms of time dependency with less restrictive/

restrictive assumptions, than the present semimarkov models.

The semimarkov models for the partially censored data provide a good construct for situations when patients move from one state to another. We assume that there exists h state denoted by $S_1, S_2, S_3 \ldots S_h$. S ome of these states may be transient states, that is one may assume that the stay in that state is finite. All other states are restricted to absorbing states that is patients after entry into this type of state will remain there until the end of study. Without loss of generality one assumes that the first states are transient and the rest are absorbing. For any case history we have.

$$H = \left(S_0, T_1, S_1, T_2, S_2, \ldots T_m, S_m \right) \qquad (7.1.1)$$

Where T_i refers to the time of transition or sojourn between states S_{i-1} to S_i. For the assumption of a semi-markov process to be true we must have two conditions present. One is that the next state of a patient will only depend on the current state and not on the previous state, and secondly that the sojourn timesbetween states are independent from each other. Therefore the length of a sojourn time will depend only on the adjoining states.

We can thus define the following properties for the semimarkov processes in a more mathematical setting. The case history such as (7.1.1) in fact can be represented by the following terms a(i), a(i,j) and F(t,i,j) where,

a(i) = Pr $(S_0 = i)$, probability that the initial state is i.

a/

$a(i,j) = \Pr(S_{n+1} = j \setminus S_n = i)$, probability that the next state is j given the present state i.

$F(t,i,j) = \Pr(T_n > t \setminus S_{n-1}=i, S_n = j)$, probability that the sojourn time between state n-1 and n exceeds t.

Further we let $F'(t,i,j) = \frac{-\partial F(t,i,j)}{\partial t}$ to be derivative of F with respect to time.

We can thus represent the probability element associated with a single history as

$$a(S_0) \prod_{n=1}^{m} [\, a(S_{n-1}, S_n)\; F'(t_n, S_{n-1}, S_n)]$$

In biomedical studies we require to have an absorbing state related to the censoring times. We can allow such a state to exist and whithout loss of generality let the last state to be a censoring time represented by (h + 1). A case history is then represented by.

$$a(S_0) \prod_{n=1}^{m-1} [\, a(S_{n-1}, S_n)\; F'(t_n, S_{n-1}, S_n)]$$

$$X\, [\, a(S_{n-1}, S_n)\; F'(t_n, S_{n-1}, S_n)\,]^{u(h-S_m)}$$

$$X\, [\, \sum_{j=1}^{h} a(S_{m-1}, j)\; F(t_m, S_{m-1}, j)]^{u(S_m-h-1)} \qquad (7.1.2)$$

where h is the last disease state, h+1 is censoring, and u(i) is set to zero for $i < 0$ and $u(i) = 1$ for $i \geq 0$.

The distribution of F(t,i,j) can take various forms for the different states. A simple method would be to consider an exponential distribution based on $F(t,i,j) = \exp(-\lambda_{ij} t)$. This distribution however may be too restrictive. This choice of the distributional form of the/

the sojourn time is the major drawback in the proper use of this type of method.

A more robust procedure however may be adopted by use of the ranks of the sojourn times. We will present the method by deriving the relevant likelihood functions. Later we will expand the results of Lagekos, Sommer and Zelen by deriving survival estimators based on a predictor -corrector method.

In order to express the (7.1.2) likelihood in terms of non-parametric maximum likelihood estimators a parameterisation is used by which survivorship function $G(t, i, j)$ is given by

$$\sum_{j=1}^{h} a(i,j)\, F(t,i,j) = \prod_{j=1}^{h} G(t,i,j)$$

The full likelihood is then expressed by the above authors as

$$L = \prod_{\substack{\text{transient states} \\ (i)}} a(i)^{l_i} \prod_{i,j} L_{ij} \qquad (7.1.3)$$

where

$$\log L_{ij} = \sum_{k=1}^{M} \Big\{ \sum_{l=j+1}^{h+1} m_{ilk} \log G(r_k,i,j) + \sum_{l=1}^{j-1} m_{ilk} \log G(r_{k-1},i,j)$$

$$+ m_{ijk} \log \left(G(r_{k-1},i,j) - G(r_k, i,j)\right) \qquad (7.1.4)$$

Where $r_1 < \ldots < r_M$ are the distinct sojourn times from the state i into j. m_{ijk} is the number of sojourn times from state i to j of length r_k. l_i is number of subjects starting at state i.

By defining $P_{ijk} = G(r_k, i,j) \;/\; G(r_{k-1}, i,j)$

so that

G/

$$G(r_k, i,j) = \prod_{l=1}^{k} p_{ijl} \qquad (7.1.5)$$

The (7.1.4.) may be rewritten as

$$\sum_{k=1}^{M} \left(\left(N_{ijk} - m_{ijk} \right) \log P_{ijl} + m_{ijk} \log (1-P_{ijk}) \right.$$

where $N_{ijk} = \sum_{l=j}^{h+1} m_{ilk} + \sum_{l=1}^{h+1} \sum_{r=k+1}^{M} m_{ilr}$

It follows that P_{ijk} has the maximum likelihood estimator.

$$\hat{P}_{ijk} = 1 - m_{ijk} / N_{ijk}$$

where $\hat{P}_{ijk} = 1$ if $N_{ijk} = 0$

And also giving

$$\hat{a}(i,j,r_k) = (1-\hat{P}_{ijk}) \prod_{l=1}^{j-1} \hat{P}_{ilk} \prod_{l=1}^{h} \prod_{r=1}^{k-1} \hat{P}_{ilr}$$

The result as presented has an intuitive appeal in that, when there are only two states the P_{ijk} estimator reduces to the analogous product limit estimator. In the situation of K states the results yield a competing risk model given by Hoel (1972).

With the situation of multiple risks the above estimators can in fact be dependent on the assumption inherent in the reparameterisation of the survival rates as in (7.1.5). This point regarding the arbitrariness of the conventions in situations of more than two states is in fact accepted by the authors.

The/

The (7.1.4) may be made more general by reparametering (7.1.5) differently. We will do so with the aim of correcting the estimators closer to the product limit estimator in the situation of single risks. The major problem with adopting a maximum likelihood approach then would be the problems associated with the estimation. It is quite likely that the P_{ijk} will not have a closed estimator. An alternative method would be to use a prediction corrector method by using $a(i,j,r_k)$ (probability of transition from state i into j of duration r_k). The p_{ijk} may then be reparameterised according to

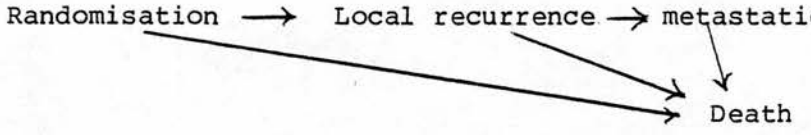
$$G^{(1)}(r_k, i,j) = \prod_{l=1}^{k} P_{ijl}^{(0)} \frac{1-a^{(0)}(i,j,r_k)}{1-\sum_{j=1}^{h} a^{(0)}(i,j,r_1)} = \prod_{l=1}^{k} P_{ijl}^{(1)}$$

where in the above $a^{(0)}$ in one step of estimators are used to form new survival function $G^{(1)}$. Clearly in the first step $a^{(0)}$ values are set to zero. The corrector part in the above model is then the ratio of the probability of a case not making sojourn time less than a particular duration (r_1) from state i to j, over the probability of not making sojourn time less than the same duration from state i to any state. Such a weighing of the transition probabilities will then correct the original probabilities by the ratio of the units of time available for transition at each state for a given time.

We will now continue with the analysis of the data based on the described method. We will later plot the semi markov probability plots based on a single step and a three step procedure. In all of what follows we will be presenting transition rate schemes for the relevant disease states. We will not estimate transition/

transition times to censorings since they do not have the same interpretation in terms of the disease.

The 561 breast cancer patients are observed under a semimarkov setting. We assume there are three transient states.



We assume no local recurrence after a general recurrence which is a justifiable assumption based on clinical definitions. We also assume that there is a further state for censored cases although there is no reason for presenting the probability distributions for these classes of patient.

The data consists of 921 epochs of the 561 patient. All patients begin from state 1 (not a necessary assumption), then all patients transfer from one state to the other until the history of observation for a patient ends in an absorbing state (Dead, censored).

Case Number	Time of Sojourn	Arriving state.
1	52	2
	0	3
	0	4
2	193	5
	192	5
	⋮	⋮

In the above sub sample of the data the first patient has "local and metastatic recurrence in the 52 month, and zero transition time to death.

Case/

Case 2 has a survival for 193 months with no recurrence. Of the 561 patients, 105 have local recurrence, 166 have general recurrence and 63 are dead after the first transition from randomisation, giving

Pr (Transition 1,1) = 0.29 $\pm$ 0.025

Pr (transition 1,3) = 0.50 $\pm$ 0.037

Pr (transition 1,4) = 0.21 $\pm$ 0.029

For the 105 patients with local recurrence, 89 die with metastatic recurrence and 7 die with no metastatic recurrence.

Pr (transition 2,3) = 0.93 $\pm$ 0.038

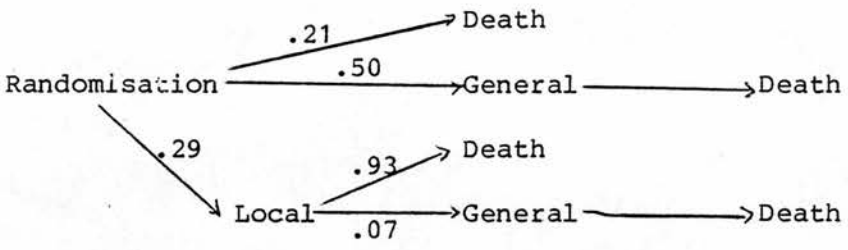
Pr (transition 2,4) = 0.07 $\pm$ 0.027

255 patients have metastatic recurrence and 226 of them die. The rest are censored.

Pr (transition 3,4) = 1.00

The semimarkov appraoch introduces certain assumtpion which are too restrictive and to some extent unnatural for survival studies. We will mention these assumption here for the present analysis but later we will introduce non-proportionality of hazards as a good basis for study of the scale of survival times. The set of metastatic patients are in fact composed of two groups. The group with previous local recurrence and the group with no record of previous recurrence. A property of semimarkovs is that the transitions at any stage do not depend on the previous states and hence in this case we assume t hat the assessment of the progression of the disease from metastatic disease to death is not affected by presence or absence of local disease. Further in the semimarkov models time of transition from previous states do not play any role in the pattern of development of the disease at present state. Hence regardless of the time that/

that a patient becomes metastatic , the analysis of a sojourn time is performed from the moment the patient enters that sojourn time onwards.



We will not present the cumulative probability of transitions, that is conditionalprobability of transition from one state to the other exceeding a time t.

$$P\,r(\text{transition at } t, i, j) = P[\, T_n > t \setminus \text{present state is } i, \text{next is } j]$$

The figure (7.1.1) to (7.1.3) show a plot of the probabilities against months of transition from randomisation, local and general recurrence. Once there is a local or general recurrence there is a fast progression to death, figures (7.1.2) and (7.1.3). There is some similarity between transition from local or general disease to death although the local to death set is very small. In figure (7.1.2) the plot of general recurrence probability appears to start at 0.7, the reason is due to the subgroup showing simultaneous local and general disease. In such cases time of transition from randomisation to local recurrence was recorded as a time from state $1 \longrightarrow 2$ and a zero transition from $2 \longrightarrow 3$.

Although we do not emphasise a statistical test of the various subgroups of the data we will report the calculations of the transition probabilities for the two transient options of the trial. A formal test based on the proportionality of the hazards will be developed later. Figure (7.1.4) to (7.1.7) summarise such relationships/

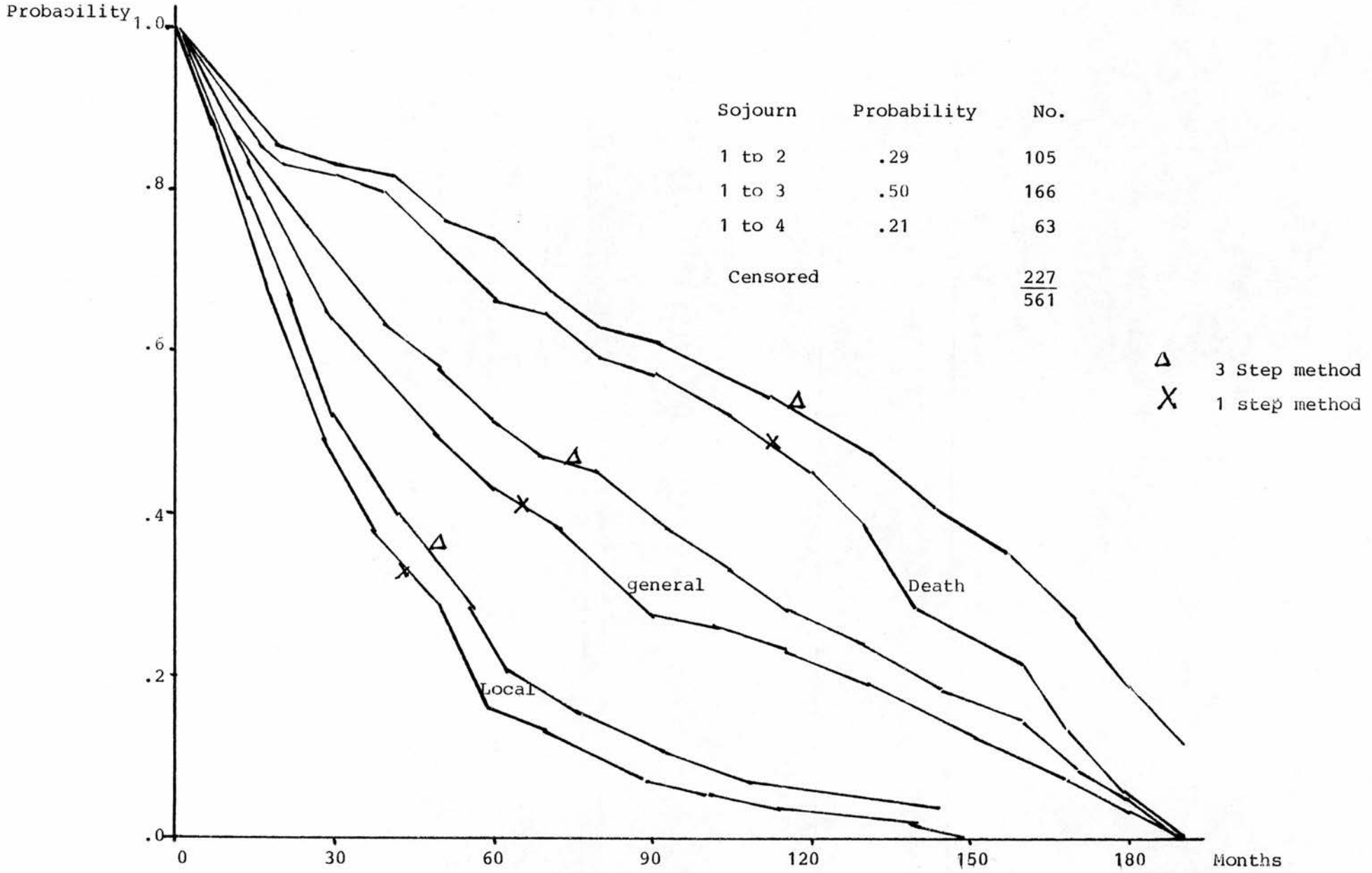
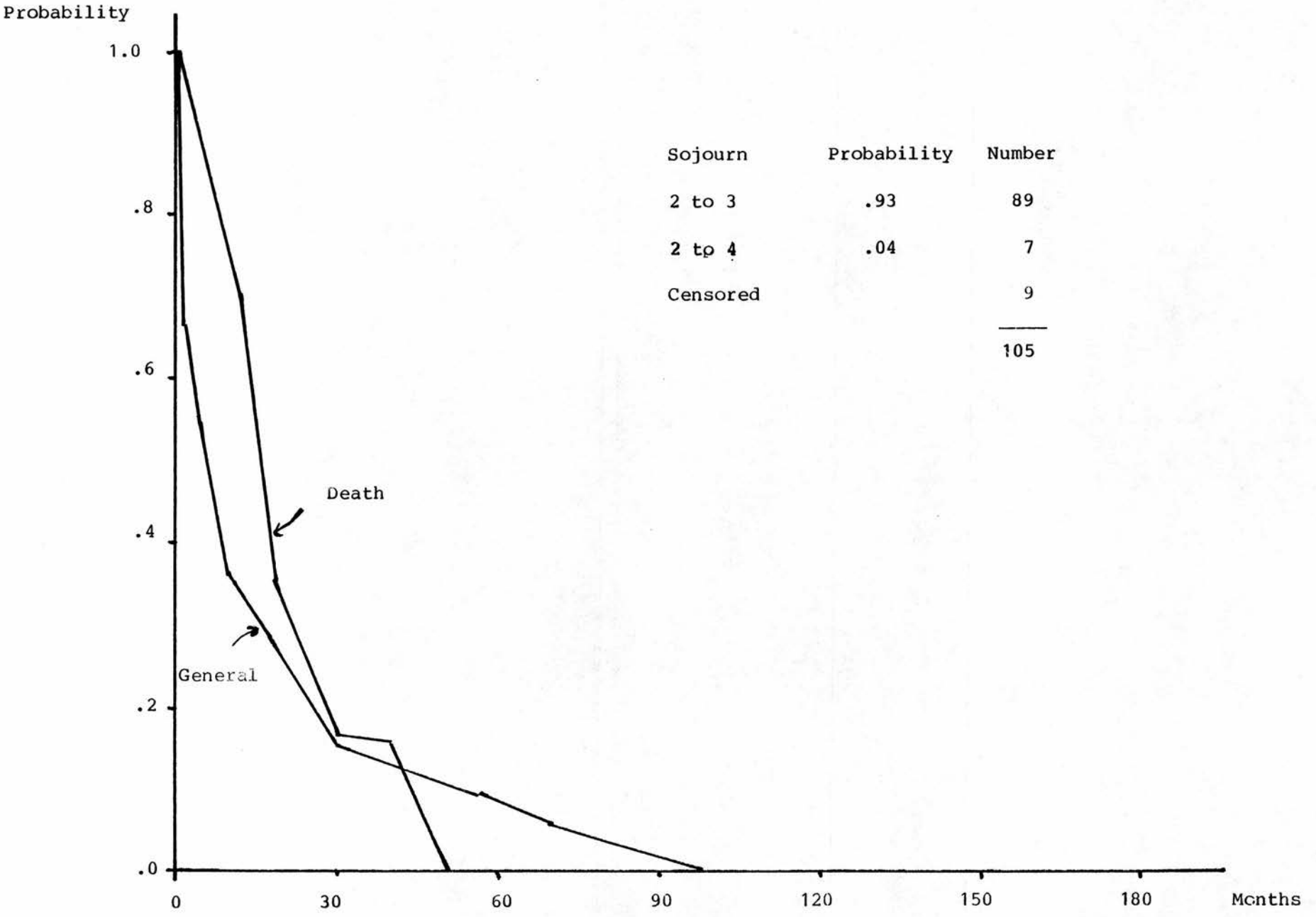
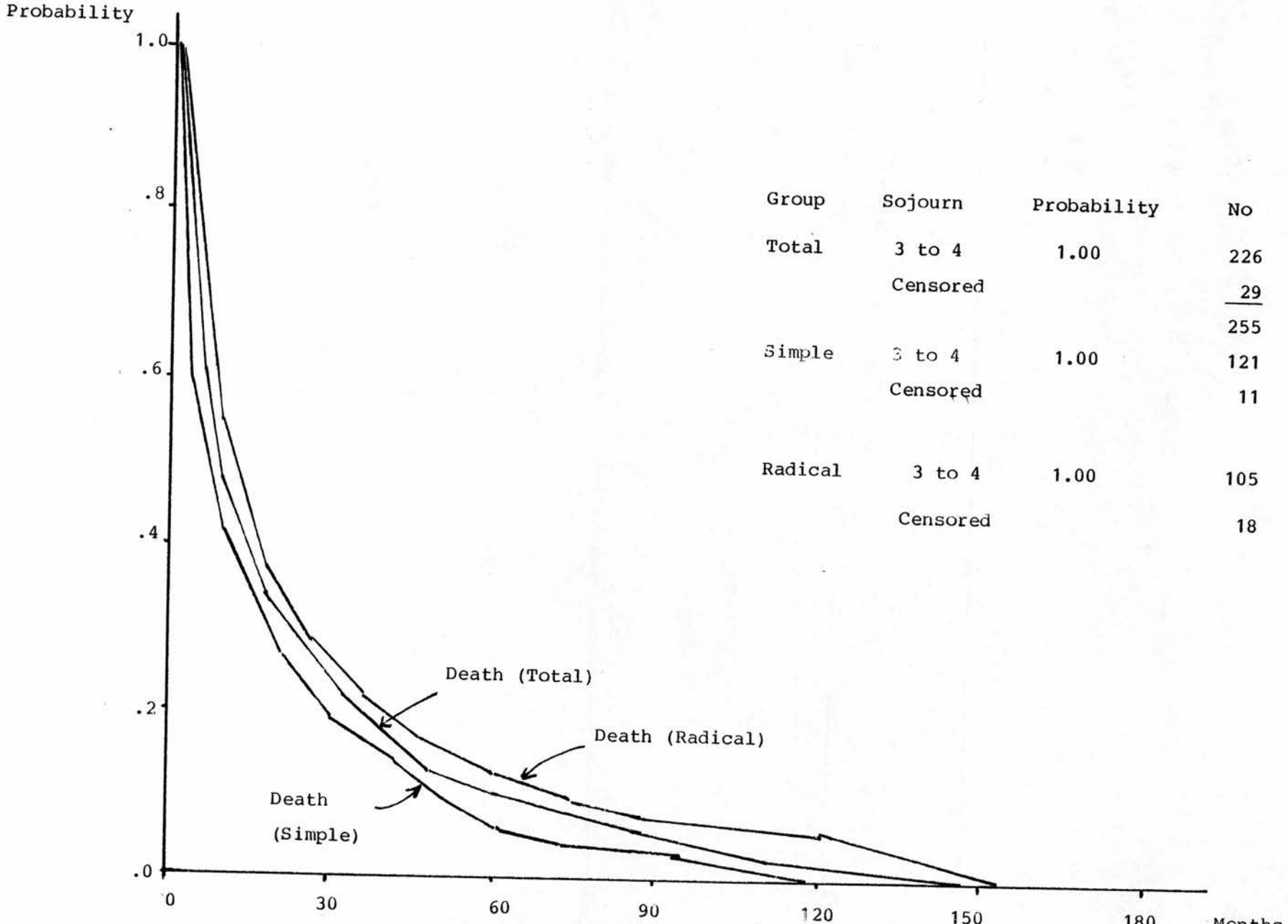


Figure (7.1.1) Total Group



Sojourn	Probability	Number
2 to 3	.93	89
2 to 4	.04	7
Censored		9
		105

Figure (7.1.2) Total group



Group	Sojourn	Probability	No
Total	3 to 4	1.00	226
	Censored		29
			255
Simple	3 to 4	1.00	121
	Censored		11
Radical	3 to 4	1.00	105
	Censored		18

Figure (7.1.3)

relationships. In terms of the interpretations of the figures (7.1.1) to (7.1.7) a point must be emphasised that distinguishes such plots from the earlier Kaplan and Meier estimators. The present plots are transition probabilities as they occur and therefore at the end of a time scale there is always a descent of the probability values to zero.

We will now use the extensions of the method as described earlier. We will use probabilities of transition from one state to another with a given duration for obtaining a corrected value of the analogous survival times (which is the G function). We repeat the method for the different transition times and in fact after three stages of the method the estimated values reach a value such that the fourth step does not contribute. The three step rates are presented in the same figures as the one step method, for the transition times from randomisation. In general we will obtain rates which are closer to those from the product limit estimator For the times other than the initial state at randomisation we will obtain values close to those of the one step methods and therefore are not presented.

In general the method has a drawback in that it assumes that censorings are unimportant. This effect is most important in terms of interpretations of figures if we are considering the lowest levels of survival probability levels when only a few patients remain. The three step method on the other hand presents an improvement on the rates of transitions. In the next section we will use the proportional hazard assumption to study some of the events mentioned in this section.

7.2/

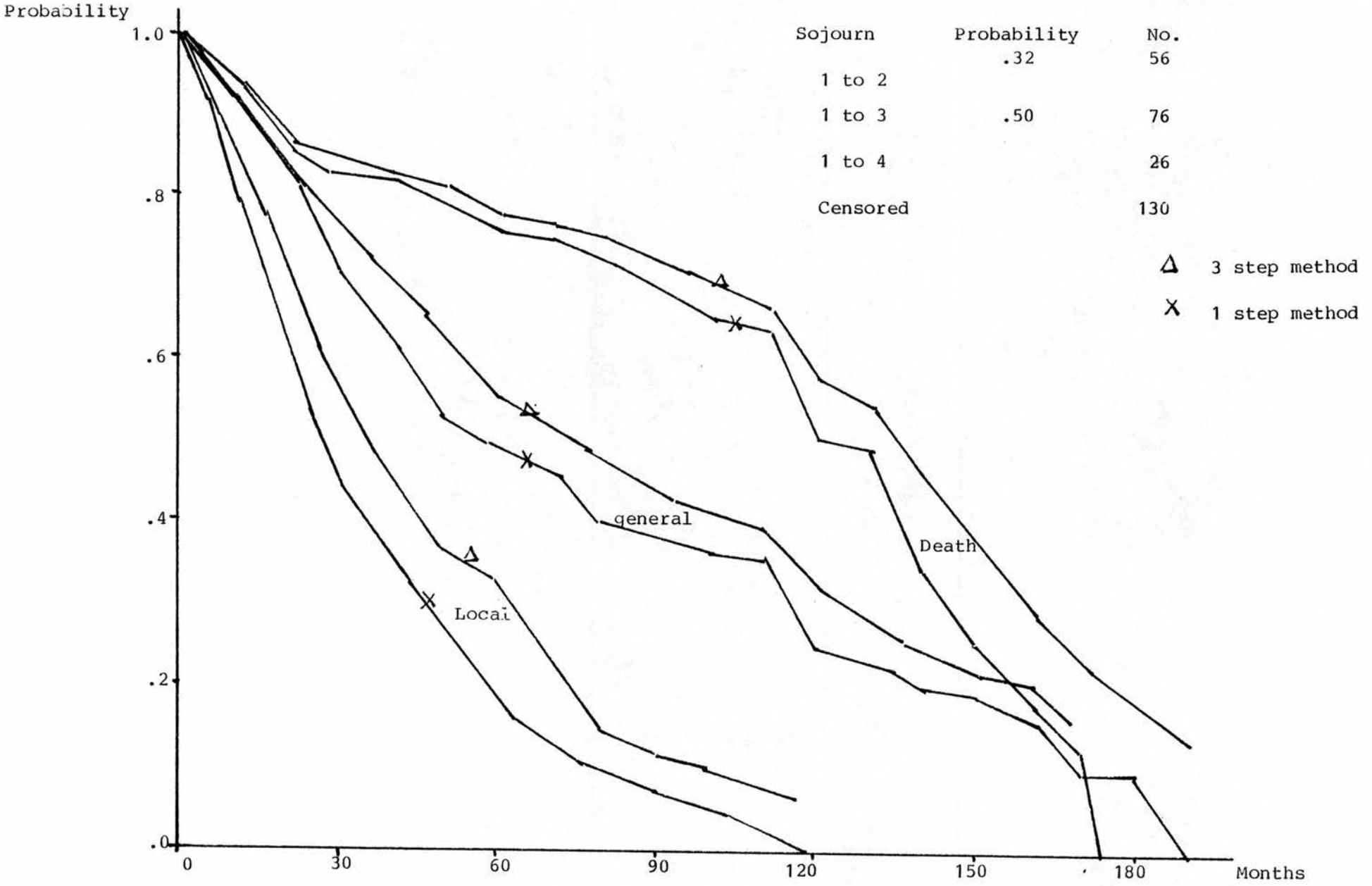


Figure (7.1.4) Radical Surgery.

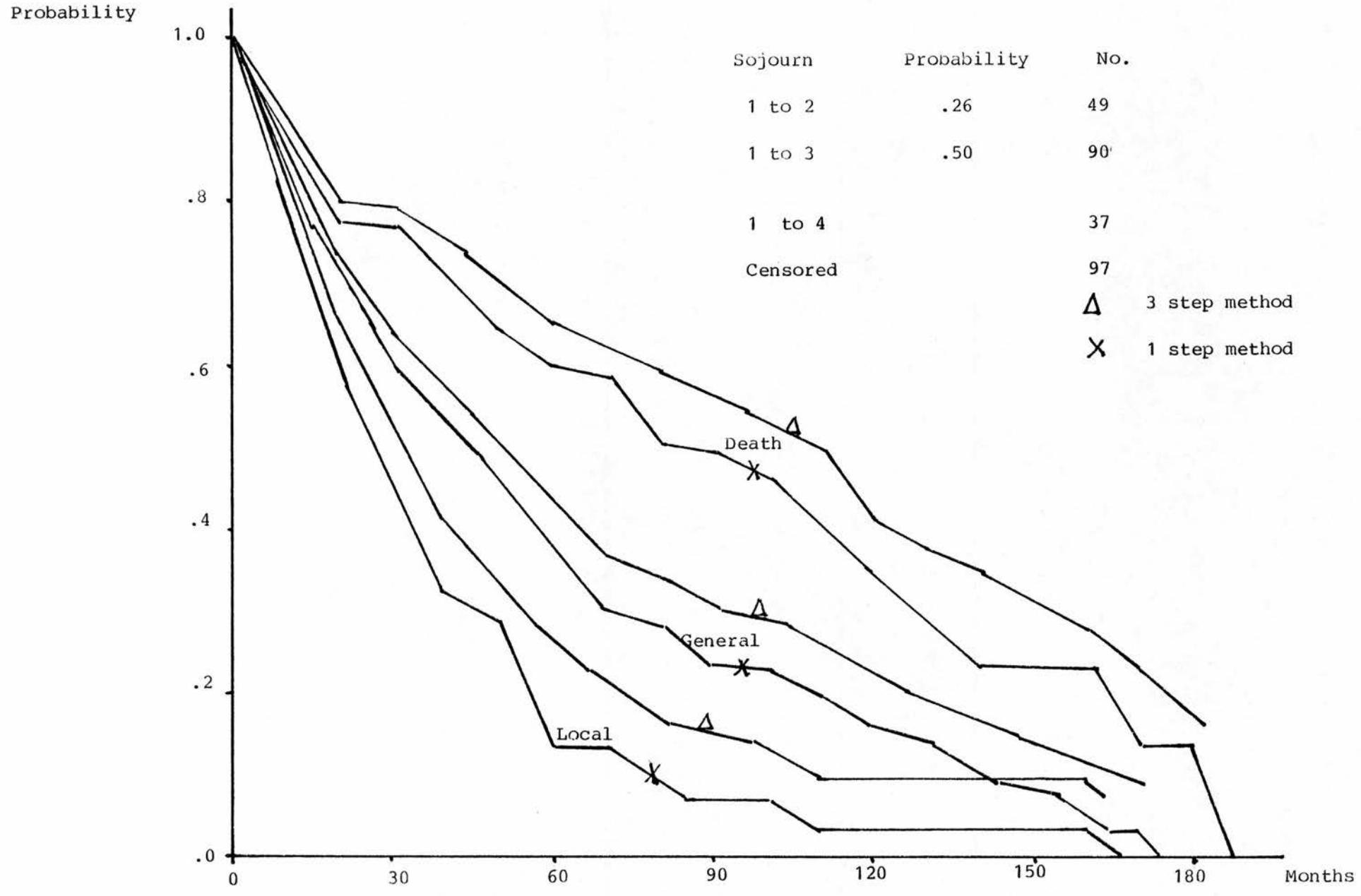
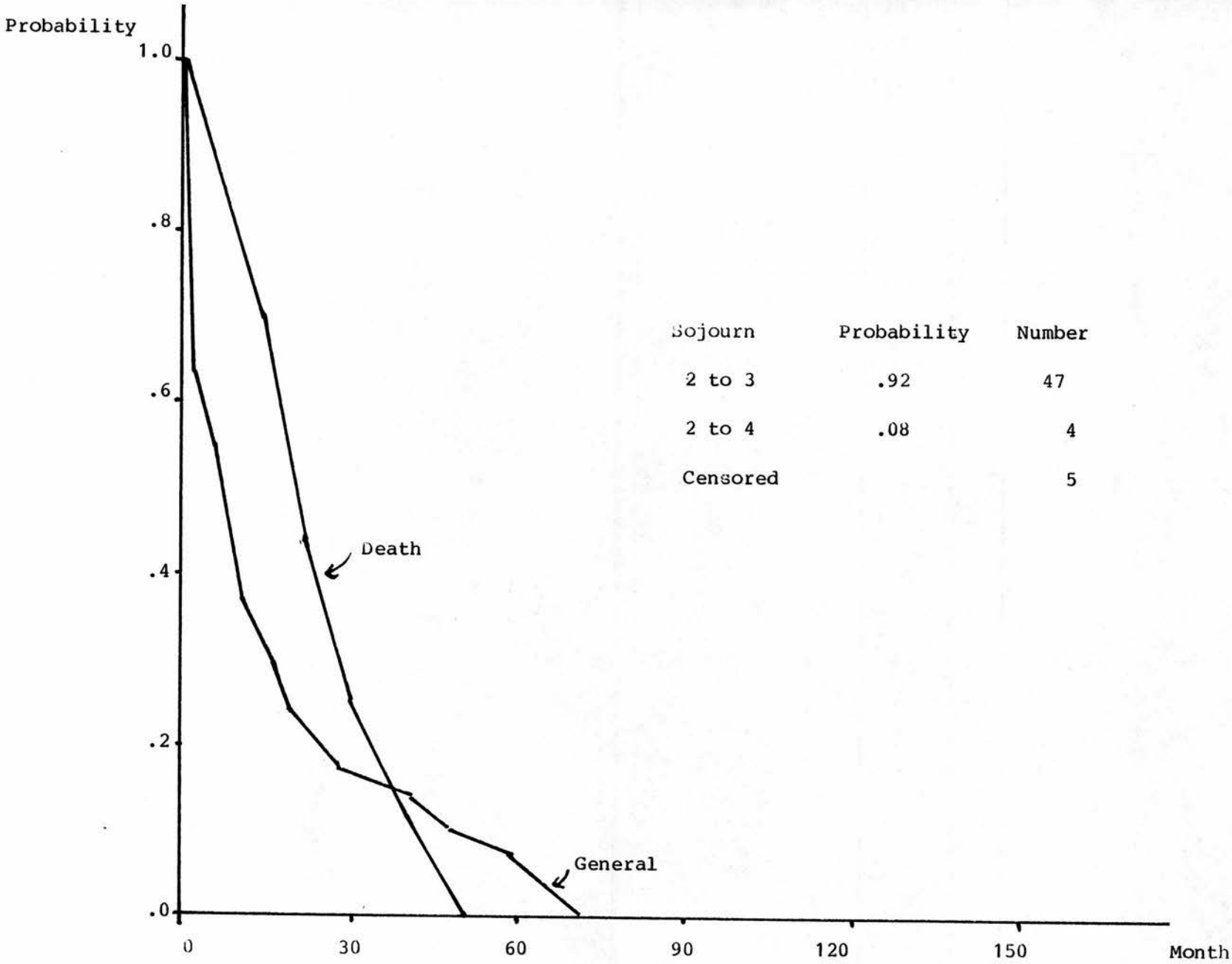
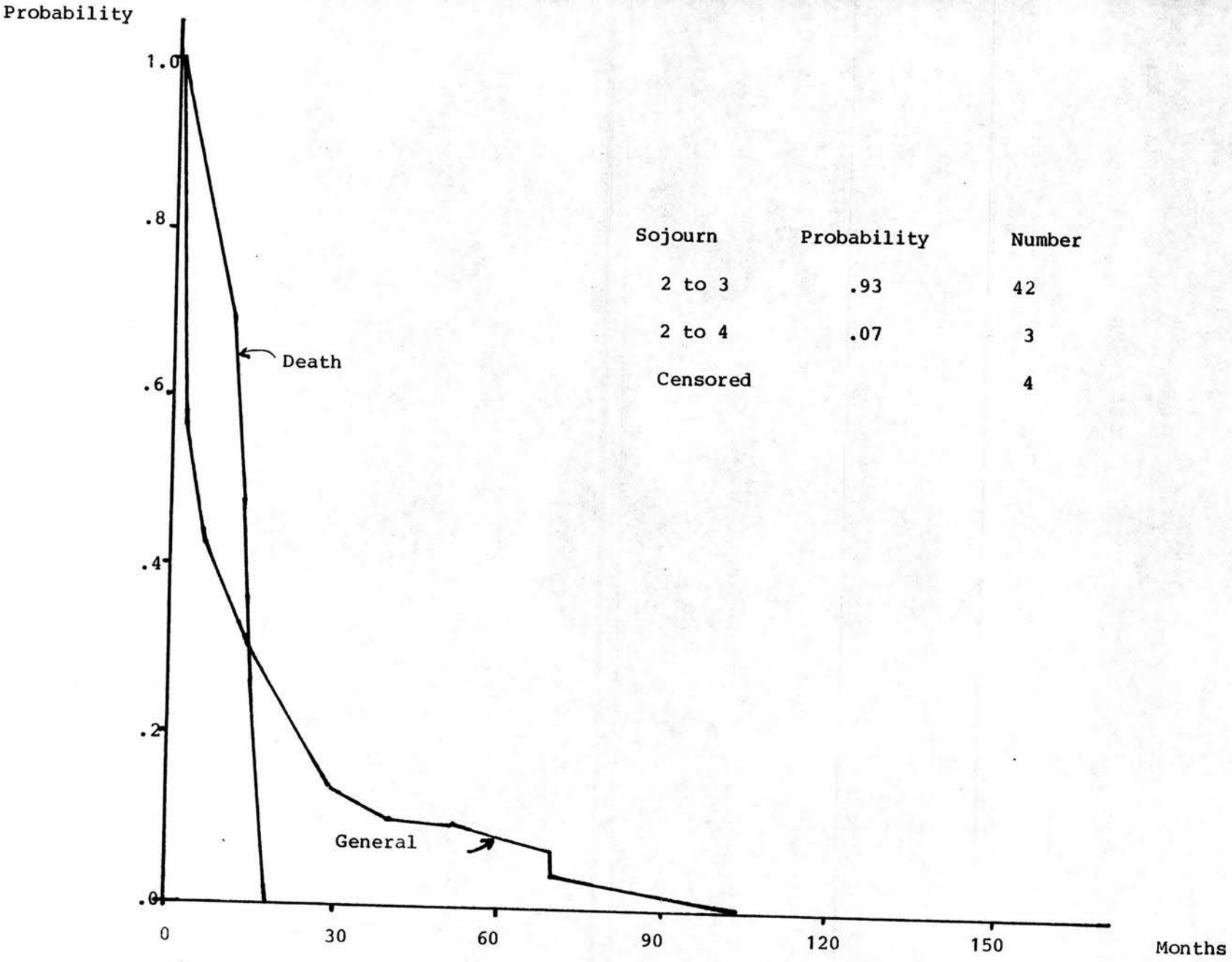


Figure (7.1.5) Simple Surgery.



Sojourn	Probability	Number
2 to 3	.92	47
2 to 4	.08	4
Censored		5

Figure (7.1.6) Radical Surgery



Sojourn	Probability	Number
2 to 3	.93	42
2 to 4	.07	3
Censored		4

Figure (7.1.7) Simple Surgery.

7.2 Time scale variability of the covariate process.

In this section we will study the effect of treatments on patients in terms of secondary response variables. Relevant questions are, given that an event has taken place and it has been a progression of disease. Firstly how is each treatment group behaving and secondly how is each prognostic indicator affecting the disease process within each group.

The secondary events that we have considered so far under the framework of the old Edinburgh trial are related to the various forms of the recurrence of the disease. These results together with the results of the exploratory approach of the non-parametric likelihoods indicated a high degree of compatability between time from randomisation to any secondary event such as local recurrence or metastatic disease. Now we will analyse the effect of covariate and treatment from secondary event to a later event. This form of analysis fits the framework of semimarkov processes in which rates of transition from any state may depend on the state the subject is occupying.

In the section on the construction of the overall likelihood we showed how the probability of a response may be represented given the previous event and censoring numbers prior to a time $t_{(i)}$. Now we will expand and define similar formulation in terms of more than one event of interest. In any given time period we defined two types of events of interest. One event was named to be the responding/

responding event of interest and the other was named a censoring. Now the argument may be expanded to allow various forms of recurrence of the disease to contribute to the partial likelihood. This can be achieved by allowing censoring to contain other events after the event of interest. Therefore for any time interval $t_{(i)} - t_{(i+1)}$ we may have s possible strata in which transitions of various forms are taking place. With the single risk case a full likelihood was represented by

$$\prod_{i=1}^{k} \Pr\left(\text{individual (i) dies} \,\middle\backslash\, \text{Immediately last censoring and present death information}\right)$$

$$\prod_{i=1}^{k} \Pr\left(\text{individual(i) censored is} \,\middle\backslash\, \text{number at risk after censorings and death}\right)$$

The former part of the likelihood is by definition the partial likelihood of Cox (75). Now by grouping the cases into s strata within which a particular response set is available we write

$$\prod_{s \geq 1} \prod_{i=1}^{k_s} \Pr\left(\text{individual (i) responds} \,\middle\backslash\, \text{immediately last censoring present death and transition information}\right)$$

$$\prod_{s \geq 1} \prod_{i=1}^{k_s+1} \Pr\left(\text{individual i is censored} \,\middle\backslash\, \text{number at risk after deaths censorings and transitions}\right)$$

The present development by Gail et al (1980) indicates that the Cox's method may be used in an analgous manner with a stratified analysis of the data. The strata are further defined to be a function of $s = S\{N(t), Z(t), t\}$. The $Z(t)$ and t have the usual interpretations under a Cox's model. However $N(t)$ represents a counting process by which one can define the base line hazard function/

function to vary for the various forms of censoring or events depending on the time of an event. The initial recording event of interest is the appearance of the local disease prior to a metastatic development. Under the present study another failure time of interest is the time to appearance of either local or metastatic disease, usually termed as the disease free interval. As a general rule we define a three parameter function to represent a response variable

R.V. (Entry, Termination, Censoring).

For the disease free interval the function is,

DFI (Randomisation,Local or metastatic disease, Last follow-up)

For the progression of the local disease we may be interested in,

(Evidence of Local disease, Metastatic recurrence or death, last follow-up)

or alternatively (Evidence of local disease, Metastatic recurrence, last follow-up or death).

As we presented the hazard rates in the Chapter 2, the initial period after treatment show converging hazard rates for the two treatments. We suspect that a similar pattern may be present for the time to local and metastatic spread of the disease. We therefore conjecture that there may be a time dependency present and the proportional hazards with a time dependent covariate may be more suitable. That is we may obtain relative risk factors of the form of figure (7.2.1) The period following the above critical events are also of interest. That is we may be interested in time after local disease to death or the development of the metastatic recurrence. We define stages of the progression in the present trial to be R,L, M & D for randomisation, Local recurrence, Metastatic spread/

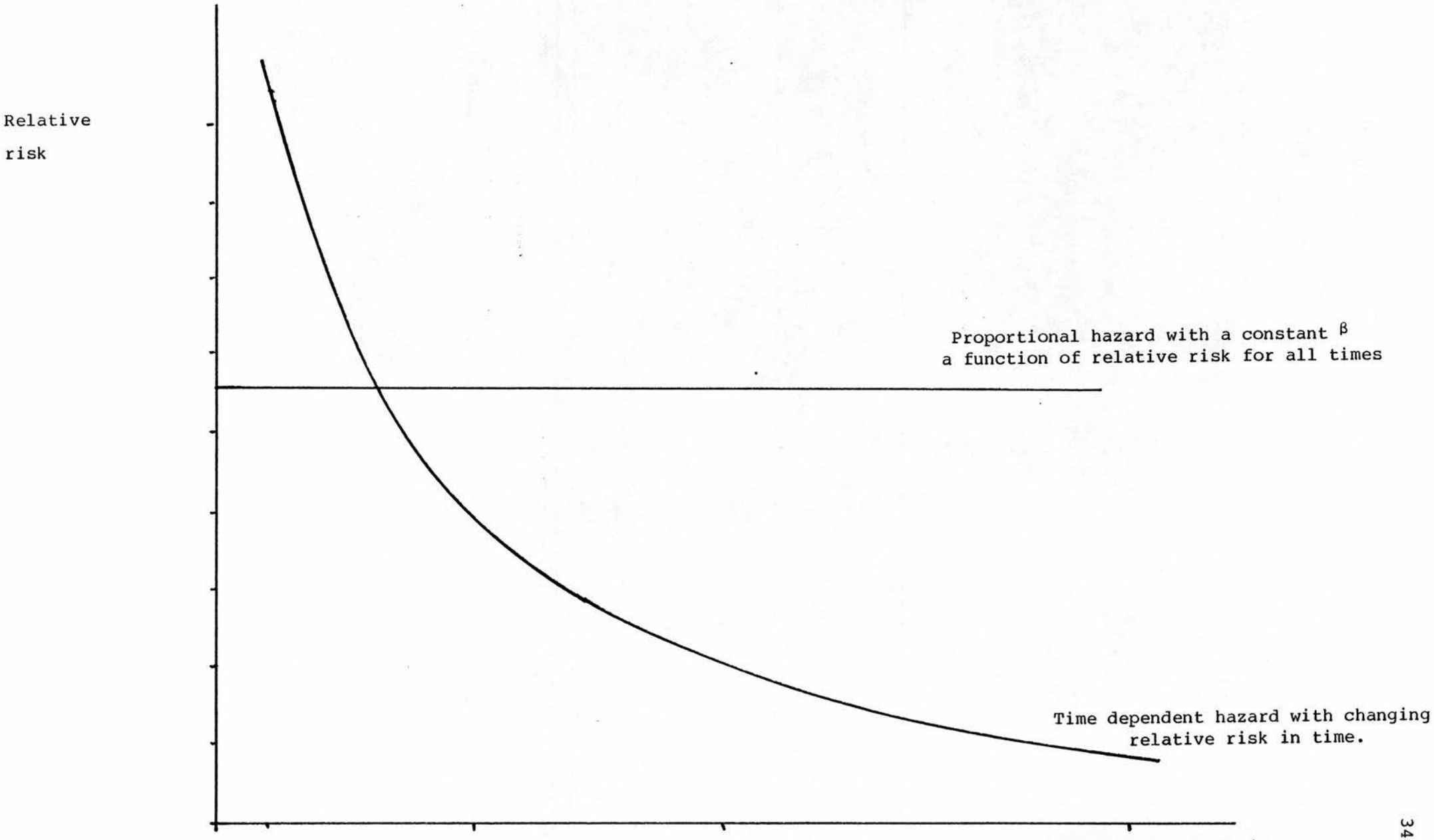


Figure (7.2.1) An example of a time dependent relative risk for time from randomisation to local recurrence.

spread and death. We also define a hazard rate for each of the intervals to be $\lambda^{<i>}_{<j>}$, where <i> refers to the entry set and <j> to the departure set. Thus λ^{R}_{L} is randomisation to local recurrence with other events as censored and λ^{L}_{M} is the hazard for local to metastatic rates. This notation will produce a general enough terminology by which various forms of entry time and termination time may be defined.

Time from local recurrence to death λ^{L}_{D} is then give by (λ^{L}_{M}, λ^{L}_{D}, λ^{M}_{D}). Time from local recurrence to time of death or metastatic recurrence which ever happens first $\lambda^{L}_{M,D}$ is then a function of hazards of the strata (λ^{L}_{M} , λ^{L}_{D}). Any case not indicated as a member of set <i> in $\lambda^{<i>}_{<j>}$ is then excluded from the strata and cases present in <i > set and absent from indicator set <j> are the censored set of study. In the above notation we define each $\lambda^{<i>}_{<j>}$ in terms of a time variable (t). Each covariate set Z(t) would then be associated only to the set of <i> present at time of study. A time dependent function of Z(t) can include information in past history by referring to information in terms of events prior to <i>. The main emphasis of study with this appraoch is to determine separately for each strata the significance of a particular covariate set for a given response variable. This is different to a stratified analysis of the type described in earlier chpaters where the emphasis was on obtaining efficient estimators for a common β obtained from pooling information from all strata. The former approach requires likelihoods of the form

$$\lambda_s(t, Z(t)) = \lambda_{0s}(t) \text{ Exp } (\beta_s Z(t))$$

Where depending on the particular form of the response variable β_s estimator is different. The latter approach requires -

$$\lambda_s(t, Z(t)) = \lambda_{0s}(t) \text{ Exp } (\beta \ Z(t))$$

The last two models clearly differ in their functional form of β and β_s.

An example of the time dependent model in study of the randomisation to death would then be introducing a time dependent covariate $Z(t) = 0$ if time for a single patient is prior to metastatic disease and $Z(t) = 1$ if time is after the metastatic disease. Basically in this approach we are affecting the proportional hazard rate by introducing different weights to the time prior to say a critical event and post critical event, for each fixed covariate set.

Up until this point we have been mainly concerned with the type of covariate that is either fixed at the time of entry of patients or it has been part of an external process from time and the response variable. Strictly the time effect is assumed to be completely related to the covariate set which is fixed from the beginning. This form of analysis is often the most important and often sufficient for analysis. However, the covariate set may have a changing pattern in time. In this situation two different conditions may be of interest. First is the situation where time trends are present and they are due to the processes within the covariate of interest. An example is the situation of age of patients in a low mortality study. We will observe an aging effect and if the duration of survival is short it may be of interest to/

to know a possible trend in ageing. Secondly, two processes may be intertwined. An example is study of long term survival in the presence of ageing, in this situation time trend may be related to the time scale itself and we may be interested in detecting departures of particular type from the model, like non-proportionality of particular type. The latter type of time dependency forms the basis of an analysis in which we will test non-proportionality due to an intervening event. In these analysis we will study the randomisation to death time for the Edinburgh trial and consider the metastatic spread to be the intervening event. In the context of the present study of the old Edinburgh trial, we identify three forms of covariates. One known generally is a fixed prognostic attribute of the case at diagnosis. Clearly these effects are external to the time scale and are inherently related to each individual patient. An example is the effect of Node status or site of main lesion. These effects were generally dealt with in the previous section. Now we introduce the time dependency concept and look at some of the fixed covariates. An example although not part of discussion under the present framework is age of the patients being related to time scale. This time dependency affects the duration and or magnitude of the age effect. Another similar covariate is the size of the initial tumour and the duration of its effect on the survival time. Further we consider a third type, the stochastic type of internal covariate, in which we introduce time dependent covariates of a type related to duration of prior events. This is taken to be the effect of time to local or general recurrence in determining survival past these events.

Another/

Another example of the third type is to consider time dependency, related to the status of patients, where the covariate is inherently related to the survival process. An example of this effect would be allocating $Z_m(t) = 0$ and $Z_m(t) = 1$ for times prior and past metastatic disease.

In Chapter 6 we concluded that age, menopausal status, T, N, S and size are the important factors affecting the survival of patients. This pattern is consistent for both arms of the trial. An interesting form of analysis is then related to the effect of prognostic indicators in time prior to metastatic disease and the significance of the effects after this event.

At this stage we attempt to utilise the size and age information by a time dependent covariate. Further we deal with an internal stochastic time dependency by considering the level of progress of disease due to the appearance of local and metastatic recurrence, (i.e. disease free interval). In terms of the secondary failure times however we consider the local disease also to be an event of interest. This is different to the exploratory approach of the last chapter in which secondary failure times were defined in combination as end points only.

Initially we introduce a model of the form containing size effect only since size can have relevant time dependent properties. Size effect is initially defined to have an external effect on the time scale. This definition will allow a relative risk function to be estimated that projects the base line hazard $\lambda_0(t)/$

$\lambda_0(t)$ on to the corresponding hazard funtion $\lambda_0(t,$ size) only by a linear and constant relation of the relative risk namely $Exp(\beta_{size},$ size). We further introduce the treatment effect by the same procedure and definition and produce a relative risk function.

$$Exp\ (\beta_{size} \cdot \text{ size} + \beta_{treatment} \cdot \text{ treatment})$$

These two models together with the treatment only model of last section will yield the following values for the estimators.

$$RR = Exp\ (\beta_{treatment} \cdot \text{ treatment})$$

$\beta_{treatment} = 0.3677$ S.E. = 0.1168 $X^2 = 9.97$ $p = 0.001$

$$RR = Exp\ (\beta_{size} \cdot \text{ size})$$

$\beta_{size} = 0.2132$ S.E. = 0.0562 $X^2 = 22.60$ $p < 0.0001$

$$RR = Exp\ (\beta_{size} \cdot \text{ size} + \beta_{treatment} \cdot \text{ treatment})$$

$\beta_{size} = 0.2271$ S.E. = 0.0581 $X^2 = 23.60$ $p < 0.0001$

$\beta_{treatment} = 0.3289$ S.E. = 0.1253 $X^2 = 7.92$ $p = 0.0012$

Clearly the models indicate a better relative survival time for the radical surgery treatment versus simple surgery with X-ray therapy. The relative order implies a worst survival of order 1.44 for the simple surgery group.

The size effect is also playing a consistently increasing role. For each two centimeters increase in the size of the initial tumour the relative risk increases by an order of 1.53. The effect of size and treatment given the present time scale is additive in the relative risk sense and there is no significance attached to the slight negative estimate of β_{st} for the size and treatment interaction.

RR/

$RR = Exp\ (\beta_{size} \cdot size + \beta_{treatment} \cdot treatment + \beta_{s.t}\ treatment.size)$

$\beta_{size} = 0.2189$ S.E. = 0.0597 $X^2 = 22.61$ $p < 0.0001$

$\beta_{treatment} = 0.3310$ S.E. = 0.1319 $X^2 = 8.43$ $p = 0.0010$

$\beta_{s.t.} = -0.0521$ S.E. = 0.0732 $X^2 = 0.85$ N.S.

Now we define the relative effect of each covariate to be dependent on a transformation of the time scale. That is firstly introduce a time dependent factor to assess the influence of size over time and secondly to test for the proportionality of the hazard rates of the option effects. As we showed in Chapter 4 the most natural form of a transformation of the time scale is achieved by a log transformation and thus for the time dependent covariates we introduce a log transformation followed by a subtraction of near mean for normalising the variable. Therefore initially all time dependencies are scaled to [log(time in months) - 2].

First we introduce a model with the time dependency of the option effect. By the definitions of the proportional hazards the effect of time must be consistently related to the relative risk function regardless of the time of death.

$RR = Exp\ (\beta_{treatment} \cdot treatment + \beta_t \cdot treatment[log(time)-2])$

$\beta_{treatment} = 0.3782$ S.E. 0.1174 $X^2 = 10.43$ $p = 0.001$

$\beta_t = -0.0921$ S.E. = 0.1131 $X^2 = .7551$ N.S.

Thus there is no indication that the proportional hazard assumption is violated with respect to treatment. There is a slight negative value attached to β_t which indicates that with increasing time the value treatment effect diminishes and that the largest differences due to treatment are in the earlier part of the study.

Figure/

Figure (7.2.2) presents a plot of relative risk functions over time for treatment. The Figure indicates that the relative risks with the inclusion of the time dependent variable is consistenatly positive and that there is no indication of crossing survival curves.

The size of the initial tumour was taken to have a consistent effect within all time periods. Now we proceed with a model to test this assumption. Basically we use a similar method as the method for testing the proportional hazards assumption of treatment effect. The time transformation is again [log(time in months)-2]. Thus the relative risk for the size and time effect is :

$$RR = \text{Exp}\ (\beta_{size}\ \text{size} + \beta_{t}\ \text{size.}\ [\log(\text{time})-2])$$

$\beta_{size} = 0.2991$ S.E. = 0.0619 $X^2 = 24.80$ $p < 0.0001$

$\beta_{t} = -0.0178$ S.E. = 0.0104 $X^2 = 2.94$ N.S.

Once again there is not a significant improvement in the size effect interaction if we allow the relative risk to be time dependent.

In the above model using the time dependency terms we note a negative β which can indicate a diminishing size influence over the time scale. The probability value of X^2 does not reach a significant level at 5% for the time dependency. For the long term survivor there is no suggestion that the size of tumour is playing a minimal role. Nevertheless we present the pattern in figure (7.2.3) in which the values of the relative risks are plotted for a few values of time and/

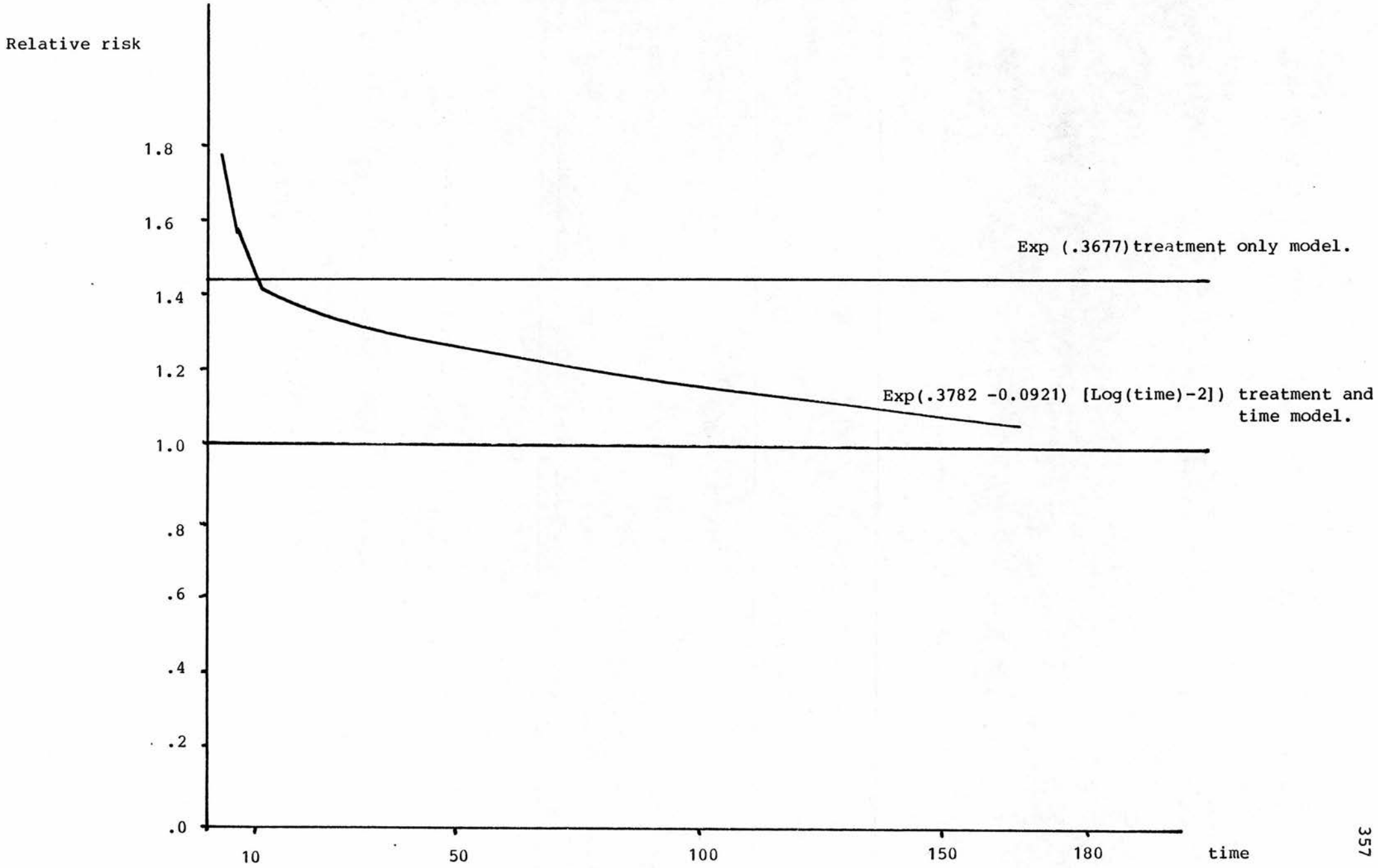


Figure (7.2.2) Relative risks of constant and time dependent model

and size separately. Tables (7.2.1) and (7.2.2) present intermediate values for obtaining the relative risks.

size*	size	[log(time)-c] -2	2.5	3
0	-2	-.6694	-0.5509	-0.4914
2	0	0	0	0
4	2	.6694	0.5509	0.4914

Table (7.2.1) (β_{size} . size + β_{t} . size.Time)

size*	size	[log (time)-c] -2	2.5	3
0	-2	.572	.601	.6117
2	0	1	1	1
4	2	1.953	1.7348	1.63

Table (7.2.2) Relative risk.

At time zero, size plays its maximum role in determining risk of death. The relative risk is intially twice as great for the larger values of size compared with the cases at mean size of 2 centimeters. Again initially the size effect for tumours of less than 2 centrimeters is 60% of the effect of the size effect for the cases with mean size of 2. The values of relative risk converge with time. The relative risks reach 1.6 for larger tumours and 0.61 for smaller tumours at the 180th month.

Now we will consider the same form of time dependency with a different functional form. It seems that the previous log transformation is natural in the sense of non-proportionality of the Weiball/

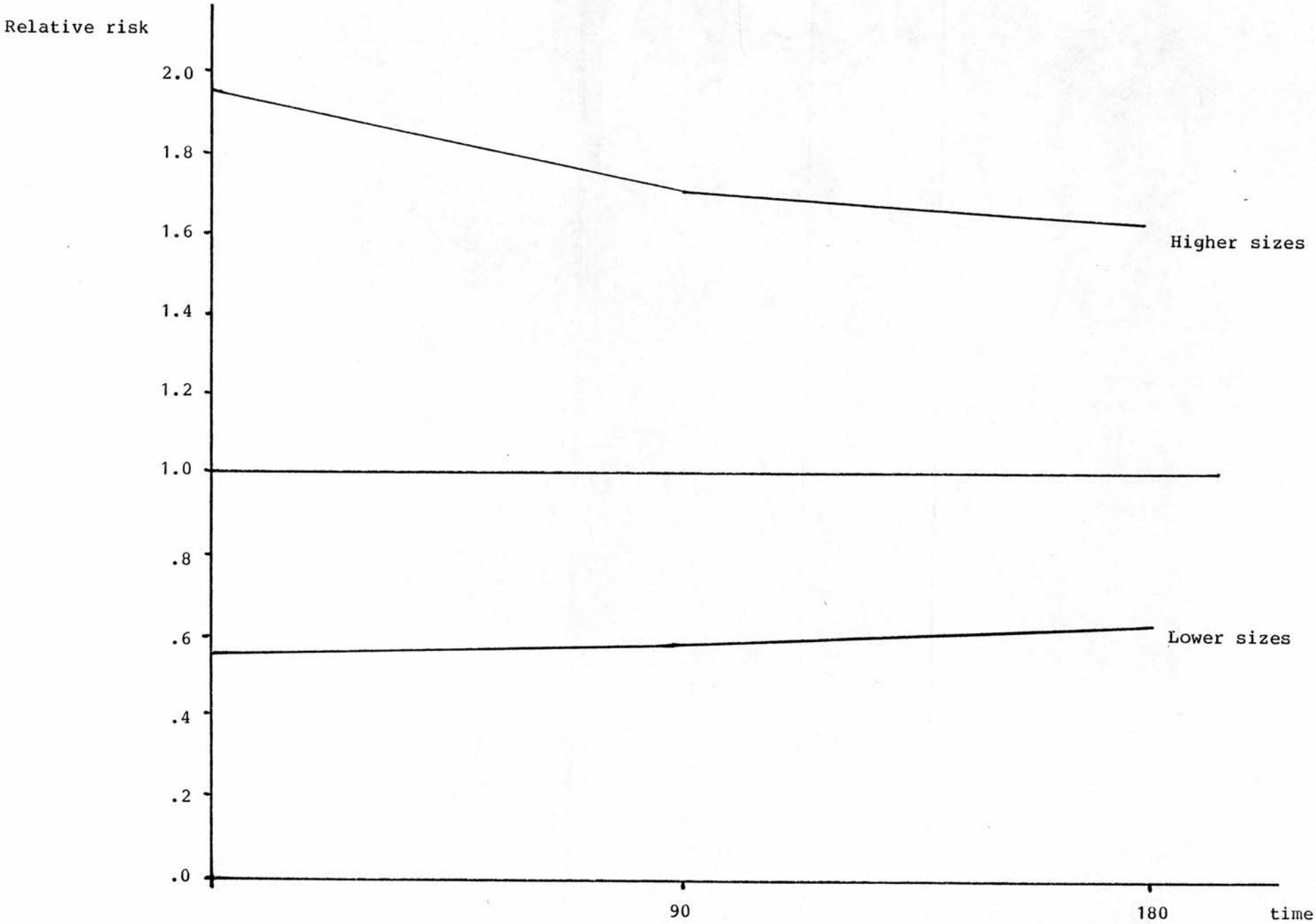


Figure (7.2.3) Relative risk of time dependent model for size. logarithmic function of time dependency

Weiball family. In the present study we suspect that the effect of time non-proportionality for each size effect is linear. That is although we can hold the view that initially the size effect is most significant, the distinction in the present function will be in thenature of the rate of the decline of the hazard rates. Table (7.2.4) refers to the new sets of risk function that are estimated using the new functional form of time dependency. The actual relative risk is then presented as

$RR = Exp\ (\ \beta_{size}\ .\ size\ +\ \beta_t\ size\ [(time/20)-2]$

$\beta_{size} = 0.3081$ S.E. = 0.0528 $X^2 = 34.05$ $p < 0.0001$

$\beta_t = -0.0154$ S.E. = 0.0085 $X^2 = 3.37$ $p = 0.064$

Compared to the previous logarithmic function of time dependency the actual magnitude of β_{size} remains close to the present estimator. The values of the estimator of the standard error of β_t also changes slightly. We refer to tables (7.2.3) and (7.2.4) and figure (7.2.4) for a graphical representation of the relative risks

		[(time in months/20)-2]		
size*	size	-2	2.5	7
0	-2	-.6778	-0.5392	-0.4006
2	0	0	0	0
4	2	0.6778	0.5392	0.4006

Table (7.2.1) (β_{size}. size + β_t size.time)

[(time in months/20)-2] /

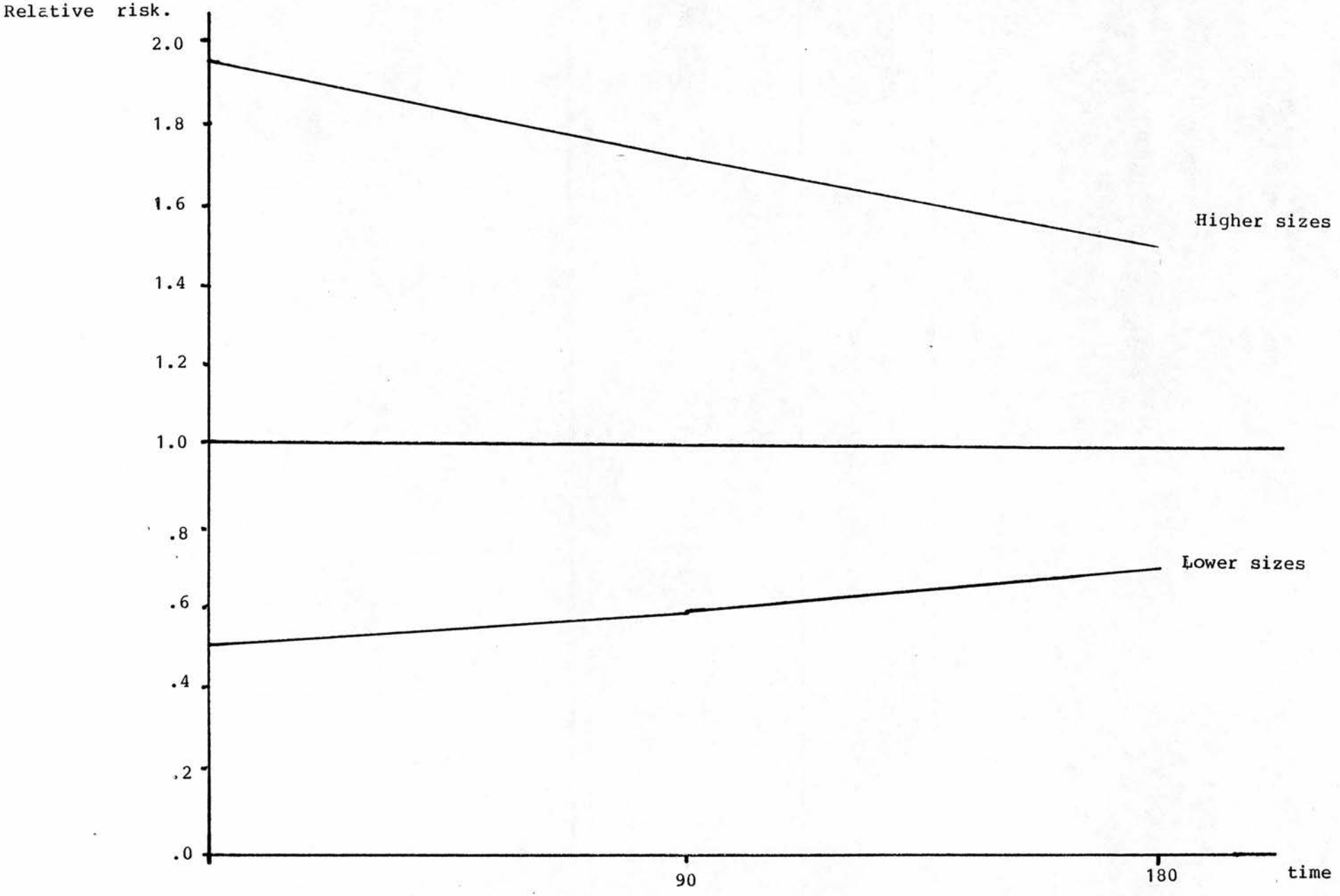


Figure (7.1.4) Relative risk of time dependent model for size. Linear function of time dependency.

		[(time in months/20)-2]		
size*	size	-2	2.5	7
0	-2	.5077	.5832	.6699
2	0	1	1	1
4	2	1.9695	1.7146	1.4927

Table (7.2.4) Relative risk.

We thus conclude that there is a slight suggestion that size of tumour for the long term survivors may play a less important role. At the early part of the time scale size has the maximum effect in determining risks of death. The relative risk is highest for the larger tumours and has a ratio of 2 : 1 for larger tumours versus medium sized tumours. This ratio reduces to 1.5 : 1 for the same sizes after the passage of time in 180th month. One final remark is that the above conclusions are compatible with models with no change over time and models with treatment effect included.

7.3 Event-time variability of the time scale.

Up until now all time effects have been dealt with on the basis of a time scale and the covariate process. That is we have assumed the change in time scale to be due to an external process. Now we deal with covariates in a time scale within which an internal variability is assumed. That is we may estimate the time effect difference for time prior to and after a critical event. Generally it is regarded in breast cancer that the initial treatmeant effect does/

does not play a major role after the development of metastatic disease.

We will initially develop a maximum likelihood function for a general approach based on a parametric method. The likelihood function will be used later to show how all the relevant information may be extracted by a particular test of the hazards. For the present methodology and the development of the likelihood we consider three separate time events.

(1) Death without the recurrence of disease.

(2) Time to the recurrence of the disease.

(3) Time from recurrence of the disease to death.

The situation is presented in Figure (7.3.1)

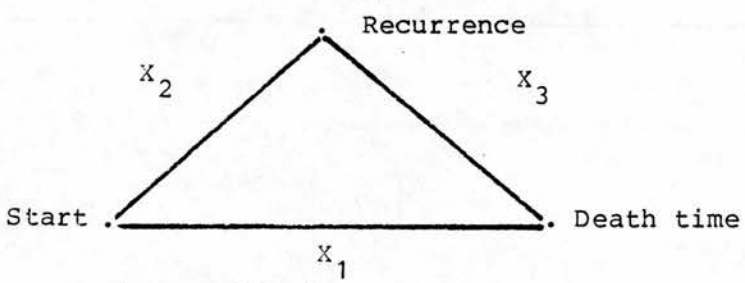


Figure (7.3.1)

We may thus expect that initially all patient groups are subject to risks of both recurrence and of death. We can represent the time of death T as

$$T = \begin{cases} X_1 & \text{if } X_1 \leqslant X_2 \\ X_2 + X_3 & \text{if } X_1 > X_2 \end{cases}$$

We can expand the above definitions so that censorings may also be included in the formulation. In here any recorded censoring may refer/

refer to two censoring paths. One is censoring before the intervening event and death and the other before death but after the intervening event. In here we refer to censoring events as c. Figure (7.3.2) to all the possible outcomes.

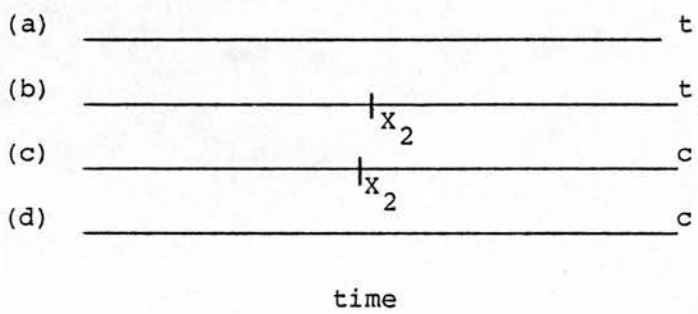


Figure (7.3.2) Types of possible observable events.

(a) refers to an outcome for a case that has a death with no recurrence of the disease being recorded. The only observable time is therefore $x_1 = t$ with the distribution $X_1 \leqslant X_2$.

(b) refers to a case with a recurrence at time x_2 and a death at time t, giving $x_3 = t-x_2$. In this case we have $X_2 < X_1$.

(c) refers to a case recurrent at time x_2 and censoring at time c giving $x_3 > c-x_2$. In this case once again we have a distribution $X_2 < X_1$.

(d) finally, in this part we refer to individuals who are observed but do not show a recurrence and are alive at the end of the study. The distributional restriction of the case is then $X_1 > c$ and $X_2 > c$

In general for most situations it is justifiable to make assumptions on the distribution of X_1, X_2 and X_3, so that random variables have an independence. Such an approach is useful in the estimating/

estimating part of the likelihood function with parametric restrictions for each of the three events. In here we consider a general likelihood function. Later in the discussion of the covariate we will reconsider the assumptions and show the use of a convenient test for the independence of the distributions

In the construction of the likelihood we consider a model in which the two paths to death are independent. That is X_1 is independent of both X_2 and X_3, but the sections of the failure time path with recurrence namely X_2 and X_3 are dependent. The general likelihood function may then later be completed with the usual distribution functions like the Weiball or exponential.

We now introduce the following notations for the distribution of X_1, X_2 and X_3 given X_2 has occured. X_1 has the density function $f(x_1)$ distribution function $F(x_1)$ and the survival function $\overline{F}(x_1)$. X_2 has the density function $g(x_2)$, distribution function $G(x_2)$ and the survival function $\overline{G}(x_2)$. Finally $X_3 \setminus X_2$ has the density function $h(x_3 \setminus x_2)$, the conditional distribution function $H(x_3 \setminus x_2)$ and the survival function $\overline{H}(x_3 \setminus x_2)$.

The maximum likelihood function is then composed of the contributions of the four types of observable events (a), (b), (c) and (d), as in figure (7.3.2).
(a) presents the conditional distribution of the death times given that there has been no recurrence of the disease.

$$F_1(x_1 \setminus X_1 \leq X_2) = \Pr[X_1 \leq x_1 \setminus X_1 \leq X_2]$$

$$=/$$

$$= \frac{\Pr[(X_1 \leqslant x_1) \bigcap (X_1 \leqslant X_2)]}{\Pr[X_1 \leqslant X_2]}$$

$$= \frac{1}{\Pr[X_1 \leqslant X_2]} \int_0^{x_1} \int_{y_1}^{\infty} f(y_1)\, g(y_2)\, dy_2\, dy_1$$

$$= \frac{1}{\Pr[X_1 \leqslant X_2]} \int_0^{x_1} f(y_1)\, \bar{G}(y_1)\, dy_1$$

The differentiation of the distribution function then gives the density function,

$$f_1(x_1 \setminus X_1 \leqslant X_2) = \frac{1}{\Pr[X_1 \leqslant X_2]} f(x_1)\, \bar{G}(x_1)$$

Thus the contribution to the likelihood from case dying at x_1 is

$$f(x_1 \setminus X_1 \leqslant X_2) \cdot \Pr[X_1 \leqslant X_2] = f(x_1)\, \bar{G}(x_1)$$

(b) Gives the joint conditional distribution of recurrence and death after a recurrence.

$$F_{23}(x_2, x_3 \setminus X_2 < X_1) = \Pr[(X_2 \leqslant x_2) \bigcap (X_3 \leqslant x_3) \setminus X_2 < X_1]$$

$$= \frac{\Pr[(X_2 \leqslant x_2) \bigcap (X_3 \leqslant x_3) \setminus (X_2 < X_1)]}{\Pr[X_2 < X_1]}$$

$$= \frac{1}{\Pr[X_2 < X_1]} \int_0^{x_3} \int_0^{x_2} \int_{y_2}^{\infty} f(y_1) g(y_2) h(y_3 \setminus y_2)\, dy_1\, dy_2\, dy_3$$

$$= \frac{1}{\Pr[X_2 < X_1]} \int_0^{x_3} \int_0^{x_2} \bar{F}(y_2) g(y_2) h(y_3 \setminus y_2)\, dy_2 dy_3$$

After differentiating the above distribution functions we will obtain the joint conditional density function.

$$f_{23}(x_2, x_3 \setminus X_2 < X_1) = \frac{1}{Pr[X_2 < X_1]} \bar{F}(x_2) g(x_2) h(x_3 \setminus x_2) \quad (7.3.1)$$

Thus the contribution to the likelihood for an observation with the intervening event time x_2 and the death time x_3 after the recurrence is

$$f_{23}(x_2, x_3 \setminus X_2 < X_1) \cdot Pr[X_2 < X_1] = \bar{F}(x_2) g(x_2) h(x_3 \setminus x_2)$$

(c) The distribution of type (c) is similar to (b) distribution with a difference that the ending point is c rather than x_3. Using (7.3.1) we have

$$\int_{c-x_2}^{\infty} f_{23}(x_2, x_3 \setminus X_2 < X_1) dx_3 = \frac{1}{Pr[X_2 < X_1]} \bar{F}(x_2) g(x_2) \bar{H}(c-x_2 \setminus x_2)$$

giving a likelihood function represented by

$$Pr[X_2 < X_1] \int_{c-x_2}^{\infty} f_{23}(x_2, x_3 \setminus X_2 < X_1) dx_3 = \bar{F}(x_2) g(x_2) \bar{H}(c-x_2 \setminus x_2)$$

(d) Finally the type (d) represent cases with a censoring time c and no time of recurrence recorded. The contribution to the likelihood is

$$Pr[X_1 > c, X_2 > c] = \int_c^{\infty} \int_c^{\infty} f(x_1) g(x_2) dx_1 dx_2$$

$$= \bar{F}(c) \bar{G}(c)$$

Now/

Now we define a total of n patients with n_1, n_2, n_3 and n_4 patients representing the number of cases with (a), (b), (c) and (d) events. The final likelihood is then given by

$$L = \prod_{i=1}^{n_1} f(x_{1i}) \quad \bar{G}(x_{1i})$$

$$\circ \prod_{i=n_1+1}^{n_1+n_2} \bar{F}(x_{2i})\, g(x_{2i})\, h(x_{3i} \setminus x_{2i})$$

$$\circ \prod_{i=n_1+n_2+1}^{n_1+n_2+n_3} \bar{F}(x_{2i}) g(x_{2i})\, \bar{H}\ (c_i - x_{2i} \setminus x_{2i})$$

$$\circ \prod_{i=n_1+n_2+n_3+1}^{n} \bar{F}\ (c_i)\ \bar{G}(c_i)$$

Now by a substitution of a particular form of a distribution form we will be able to estimate the relevant parameters. In here it may be possible to obtain a reasonable estimation proeedure for a constant hazard case using an exponential distribution. However if we adopt a more robust distribution based on the Weiball distribution the method will become very complex.

If we adopt a distribution with the covariate restrictions of the proportional hazards assumptions we will have

$$\lambda(x_{i1}, Z_i) = \mu_{0i}\ (x_{1i}) \quad \text{Exp}\ (\beta_1\ Z_i)$$

As the hazard rate of the ith observation of the x_1 time. The survival function is then by the definitions of the introduction

$$\bar{F}\ (x_{1i}, Z_i) = \text{Exp}[- \int_0^{x_{1i}} - \lambda_0(t)\ \text{Exp}(\beta_1\ Z_i)\ dt]$$

There/

There is clearly a one to one correspondence between the hazards and the survival functions.

Based on the assumptions of the proportional hazards we will have

$$L = \prod_{i=1}^{n_1} f(x_{1i}, Z_i)\ \bar{G}(x_{1i}, Z_i)$$

$$\circ \prod_{i=n_1+1}^{n_1+n_2} \bar{F}(x_{2i}, Z_i)\ g(x_{2i}, Z_i)\ h(x_{3i} \setminus x_{2i}, Z_i)$$

$$\circ \prod_{i=n_1+n_3+1}^{n_1+n_2+n_3} \bar{F}(x_{2i}, Z_i)\ g(x_{2i}, Z_i)\ \bar{H}(c_i - x_{3i} \setminus x_{2i})$$

$$\circ \prod_{i=n_1+n_2+n_3+1}^{n} \bar{F}(c_i, Z_i)\ \bar{G}(c_i, Z_i) \qquad (7.3.2)$$

The above gives a good representation of the full likelihood for a process involving an intervening event. Now temporarily returning to the discussion of Chapter 4 we have.

$$L = \prod_{i=1}^{D} \lambda_0(t_i) e^{\beta Z_i(t_i)} \prod_{i=1}^{n} \mathrm{Exp}\{ -\int_0^{t_i} \lambda_0(u)\ e^{\beta Z_i(u)}\, du\} \text{ as in (4.4.8)}$$

$$L = \prod_{i=1}^{D} \left([\mathrm{Exp}\{ -\int_{t_{i-1}}^{t_i} \lambda_0(u) \sum_{j \in R} e^{\beta Z_j(u)}\, du\}\ \lambda_0(t_i) \sum_{j \in R} e^{\beta Z_j(t_i)}] \right.$$

$$\left. \frac{\mathrm{Exp}\left(\beta Z_{ji}(\tau_i) \right)}{\sum_{j \in R} e^{\beta Z_j(t_i)}} \right) \text{ as in (4.4.11)}$$

The/

The expression (7.3.2) has been presented for a proportional hazard rates with fixed covariates and the probability density functions, f, g and the joint density function h. The expression (4.4.11) in fact can allow time dependent covariates within the time scale (if we assume non informative censoring and a further generalisation for the model so that $Z_i(u)$ and $\lambda_0(u)$ are assumed to be independent within the integration region.) The consequence is to preserve the proportionality of the hazards while testing the lack of fit by a time dependent covariate t. We in fact can have the following hazard rates.

$$\lambda_1(x_{1i}, Z_i) = \lambda(x_{1i}) \text{ Exp } (\beta_1 Z_i)$$

$$\lambda_2(x_{2i}, Z_i) = \lambda(x_{2i}) \text{ Exp } (\beta_1 Z_i + \beta_t t_1 Z_i)$$

$$\lambda_3(x_{3i}, Z_i) = \lambda(x_{3i}) \text{ Exp } (\beta_1 Z_i + \beta_t t_2 Z_i)$$

t_1 and t_2 are described below and are not related to 4.4.8 and 4.4.11.

In so far as a testing of covariate effects is concerned we may be interested in tests of non-proportionality due to either λ_2 or λ_3 which are assessed by functional forms of t_1 or t respectively. Thus (7.3.2) over generalises the process for a relevant test. We then have the following 2 x 2 table.

Contribution of individual to likelihood for a day of survival.

Test of non-proportionality	Before recurrence	After recurrence
For time to recurrence	$t_1 = 1 \quad t_2 = 0$	$t_1 = 0 \quad t_2 = 0$
For time after recurrence	$t_1 = 0 \quad t_2 = 0$	$t_1 = 0 \quad t_2 = 1$

The/

The survival distributions can now be expresses as

$$\overline{F}\,(x_{1i},Z_i) = \text{Exp}\left[-\int_0^{x_{1i}} -\lambda_0(u)\ \text{Exp}(\beta_1 Z_i)\,du\right]$$

$$\overline{G}\,(x_{2i},Z_i) = \text{Exp}\left[-\int_0^{x_{2i}} -\lambda_0(u)\ \text{Exp}(\beta_1 Z_i + \beta_t t_1 Z_i)\,du\right]$$

$$\overline{H}\,(x_{3i}\backslash x_{2i},Z_i) = \text{Exp}\left[-\int_0^{x_{3i}\backslash x_{2i}} -\lambda_0(u)\ \text{Exp}(\beta_1 Z_i + \beta_t t_2 Z_i)\,du\right]$$

Note that in essence since $\lambda_0(u)$ is a nuisance function an adjustment of the initial value of $\lambda_0(u)$ at each iteration should suffice and therefore t_1 and t_2 are otherwise essentially independent from the integration. t_1 is then a function of u for its initial value and independent of the integration by definitions of the partial likelihood. t_2 is conditional on t_1 and has a similar definition from partial likelihoods. We thus can allow adjustment of the time scale before or after the recurrence by introduction of a time dependent covariate t_1 or t_2, using the proportional hazards with the Kaplon and Meier base line hazards. As in (7.3.2) the group 1 to n_1 and $n_1+n_2+n_3+1$ to n are the usual contributors in the absence of recurrence. The relevant part of the distribution of n_1+1 to n_1+n_2, and n_1+n_2+1 to $n_1+n_2+n_3$ is now represented by either t_1 or t_2 depending on the type of test. From now we will return to the language of the proportional hazards model and express t_1 and t_2 as the Z(t) covariates.

If the assumption of the proportional hazards does not hold for a period of the time scale, between the two covariate subgroups the time dependency will be testable.

In/

In the situation of non-proportional hazards the value of Z is allowed to change within the time scale. In this context, rather than assume a covariate effect is acting consistently in a multiplicative manner on the base line hazard, we can test the value of β in particular periods of time. A different deviation of from the base line hazard within the time scale can then be attributed to a priori important event taking place before the last follow-up. In order to assess the value of such an effect in application, we test the impact of the development of metastatic disease. We consider a time dependency of the above type, with Z(t) = 1 if time is after metastatic disease and Z(t) = 0 if time is prior to metastatic disease. Thus we will have a relative risk, initially composed of,

$$\text{Exp}\ (\beta_Z Z + \beta_t Z \times Z(t))$$

We know that the treatment plays an important role in determingin survival. By a rescaling and use of the assumptions of the proportional hazards we did not have sufficient evidence to reject the proportionality assumption for size or treatment. Now we test the assumption of proportional hazards based on the development of a secondary event using the above constructs and details.

$$RR = \text{Exp}\ (\beta_{\text{treatment}} \cdot \text{treatment} + \beta_t \cdot \text{treatment} \cdot Z(t))$$

$\beta_{\text{treatment}} = 0.3982$ S.E. = 0.1170 $X^2 = 11.61$ p = 0.0007

$\beta_t = -0.1239$ S.E. = 0.0781 $X^2 = 3.10$ p = 0.0748

Although again we do not reject the proportionality assumption based on the development of metastatic disease, there is some indication that treatment effect is more substantial prior to metastatic disease.

The/

The main relative risk under study so far has been the survival relative risk based on the various covariate functions. Earlier in this chapter we explained a method for defining different response variables more clearly. Now we will study time dependency with other response variables. This is analgous to the study of competing risks or multivariate failure time study. Initially we will concentrate on stratified analysis based on the log-rank test.

The hazard function λ_L^R analysis indicates that there is not a significant difference between the two arms of the trial, by either the log rank test or the Wilcoxon test. (Chisquared values 1.21 and 0.82 respectively). On plotting the survival curves, for both treatment groups we note quite similar rates. However on plot of the hazard rates there is an indication that the simple mastectomy group are at a slightly higher risk of developing local recurrence than the radical group. This effect is not significant although produces a relatively larger number of locally recurrent patients within the first three years. Figures (7.3.3) and (7.3.4).

By considering that the local disease may be an important intervening effect we will continue with analysis and consider $\lambda_{M,D}^L$ and later λ_D^L. Time of local recurrence to death and local recurrence to metastatic disease do not show a significant difference between the two treatment options. The tests are performed for a stratified analysis as well as a pooled stratified analysis according to time to the development of local disease. The three strata are defined as in figures (7.3.5) to (7.3.8)

The/

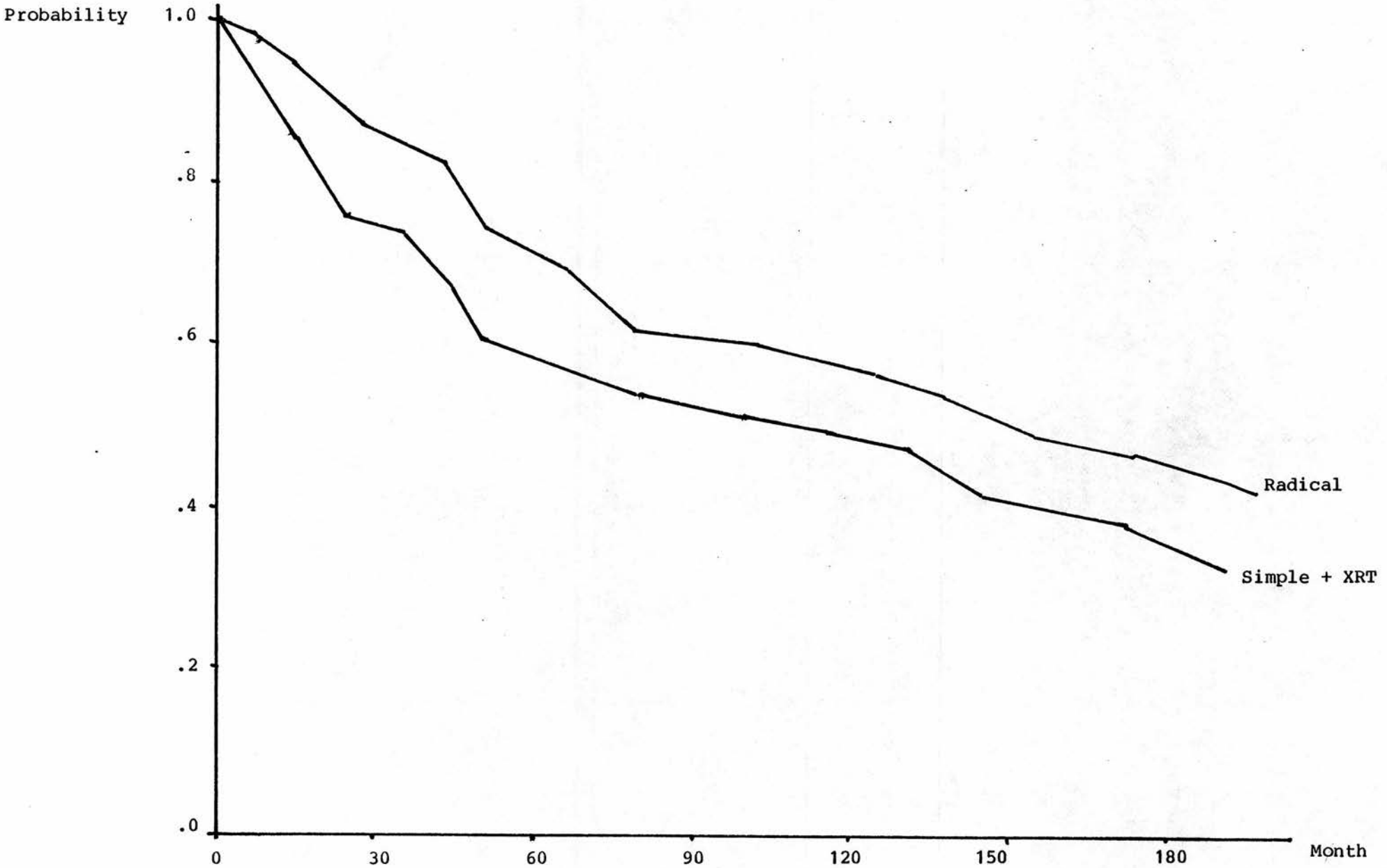


Figure (7.3.3) S_L^R

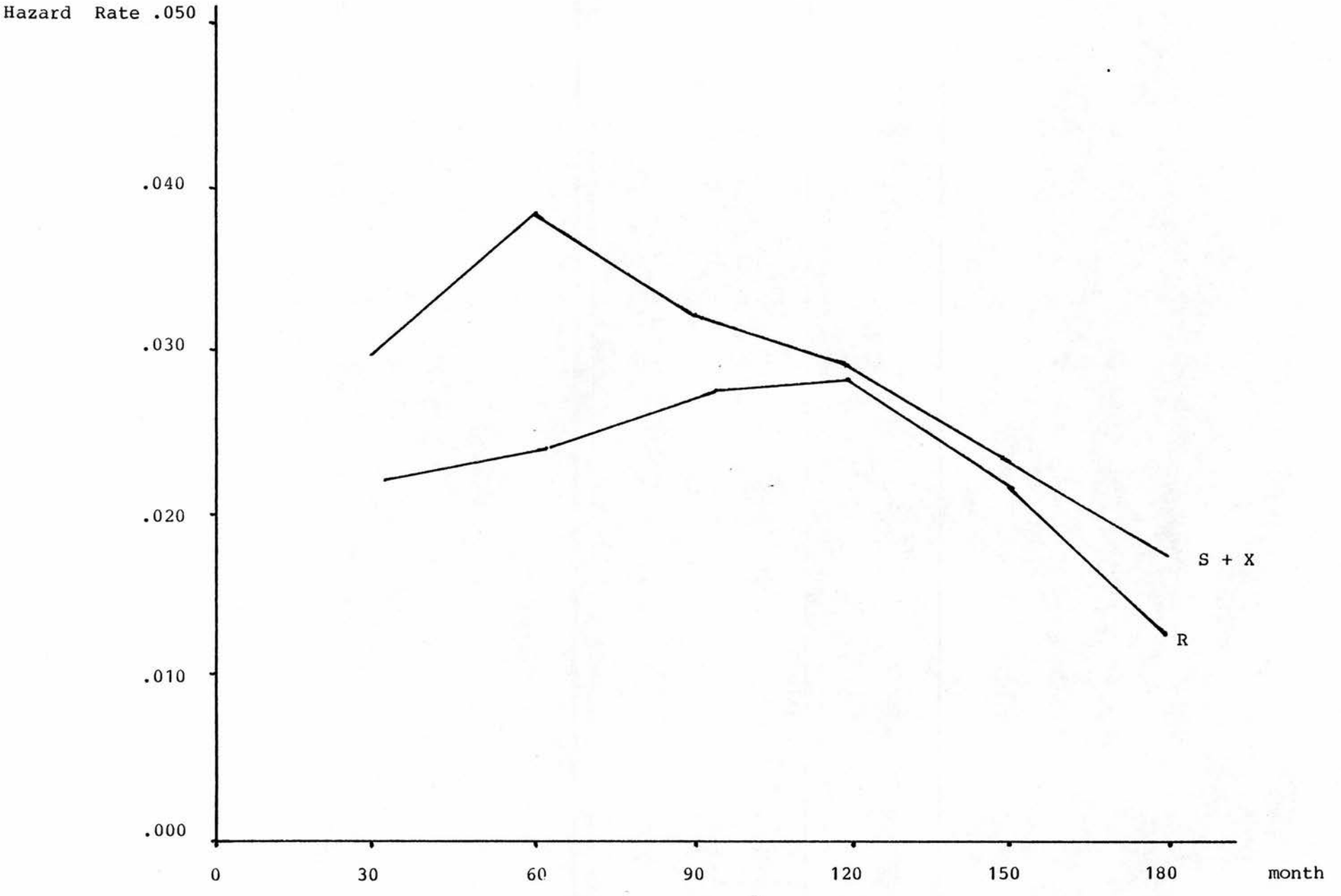


Figure (7.3.4) λ_L^R

The time from randomisation to metastatic disease indicates a similar pattern to that seen for the time from randomisation to death. The radical surgery group show lower risks with the chi-squared value of 4.01 and the probability value of .0457. No difference of statistical value is detected for the actual treatments past the development of metastatic disease namely for the hazards of λ_D^M . The tests for hazards of λ_D^M is performed in a similar manner to those of $\lambda_{M,D}^L$.

One pattern which consistently emerges indicates a higher hazard rate for the simple surgery group in the initial 3 year period after treatment, figure (7.3.9) and (7.3.10). The above figure conforms to the findings for the λ_L^R hazards. We thus consider the DFI $=\lambda_{L,M,D}^R$.

The disease free interval is traditionally an accepted response variable in survival studies of breast cancer and in here the hazards indicate a consistent distributional structure for both local and metastatic periods.

The logrank test indicates a significant difference between the treatments in terms of the DFI. X^2 value based on the logrank test is 4.08 which has the corresponding probability value of p=0.043.

Finally we study the response time of $\lambda_D^{L,M}$ which is the time from the development of the metastatic disease or local recurrence given locals are prior to metastatic disease to the time of death./

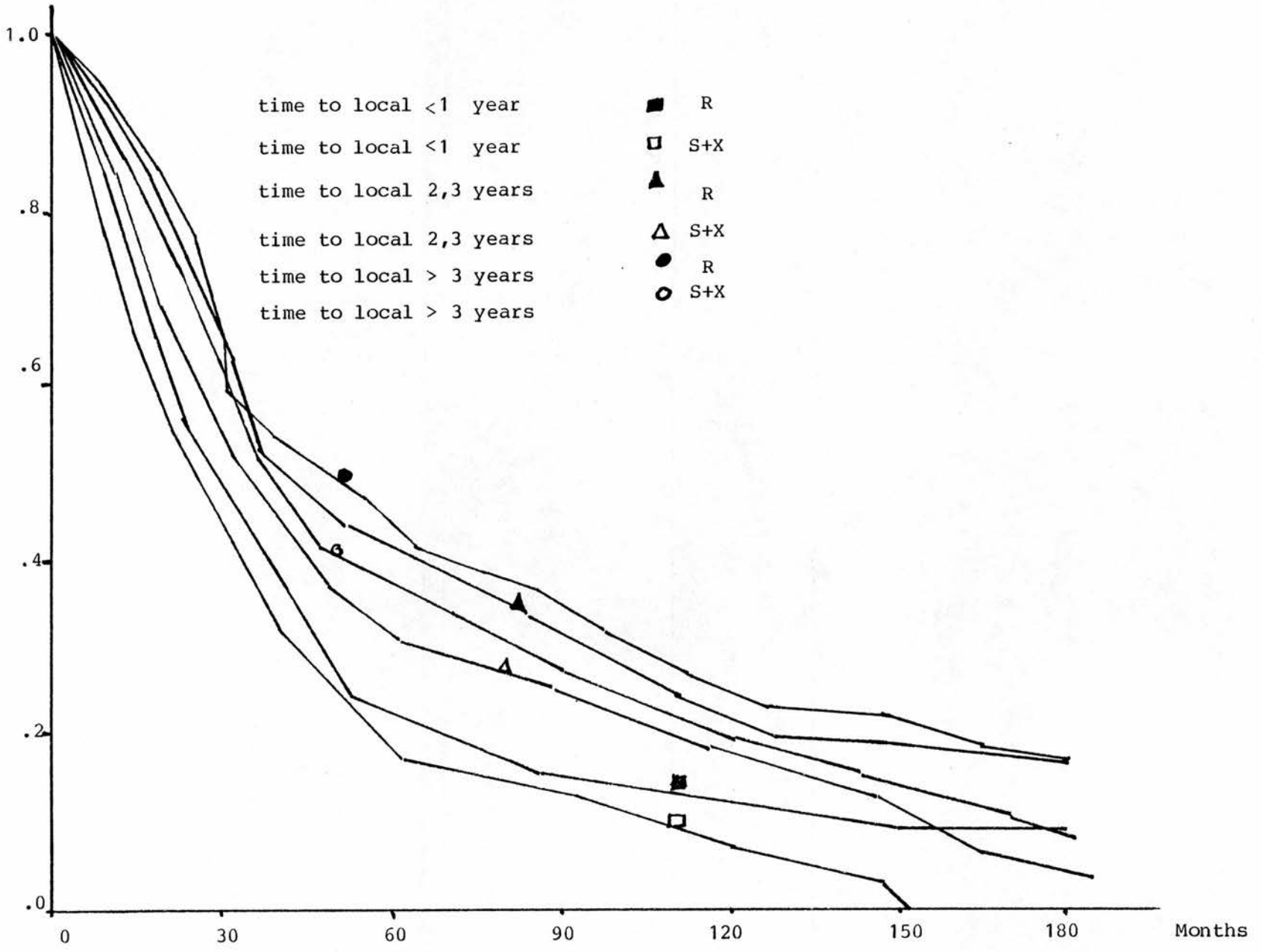


Figure (7.3.5) S_D^L Stratified according to time to local.

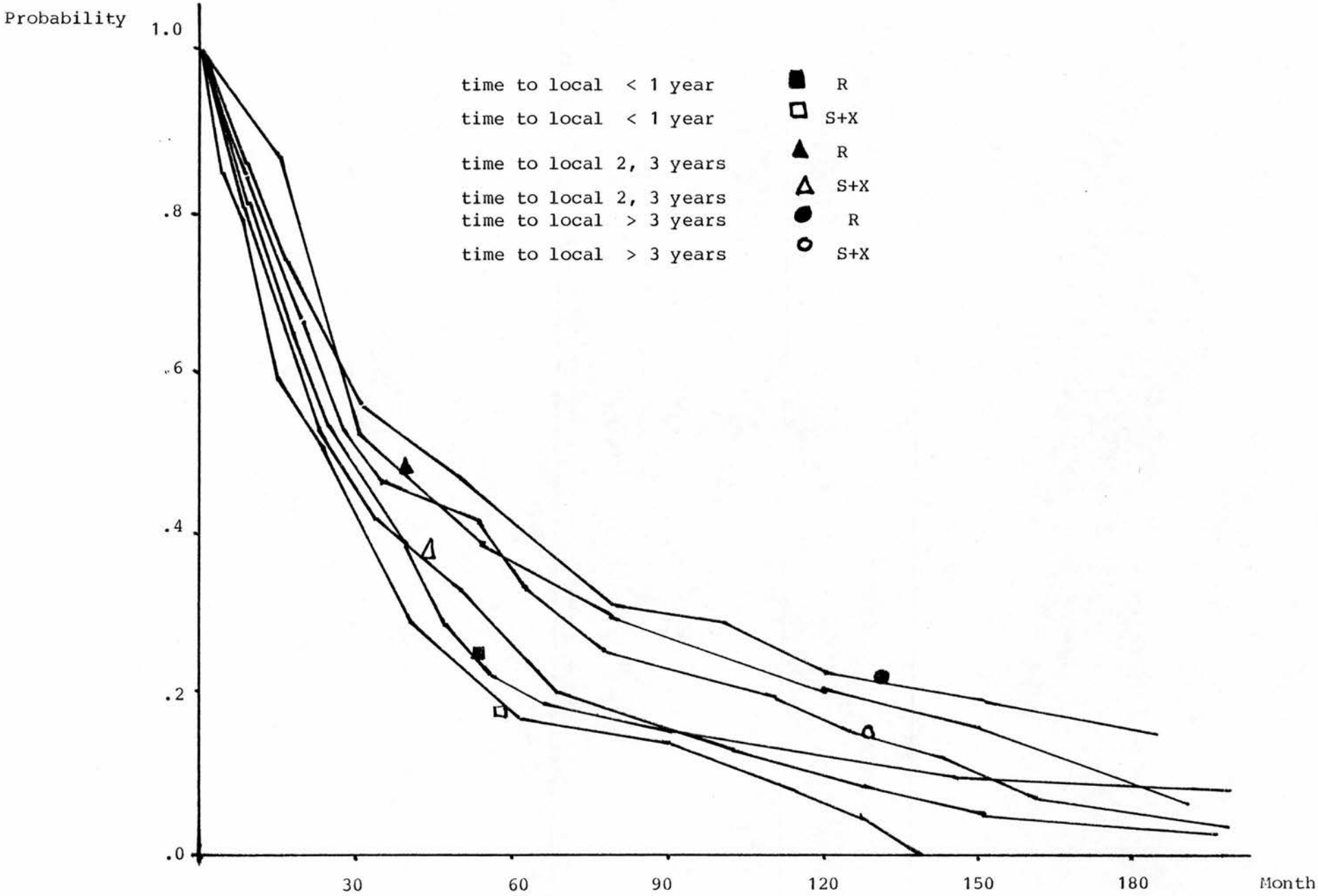


Figure (7.3.6) $S^{L}_{M.D}$ Stratified according to time to local.

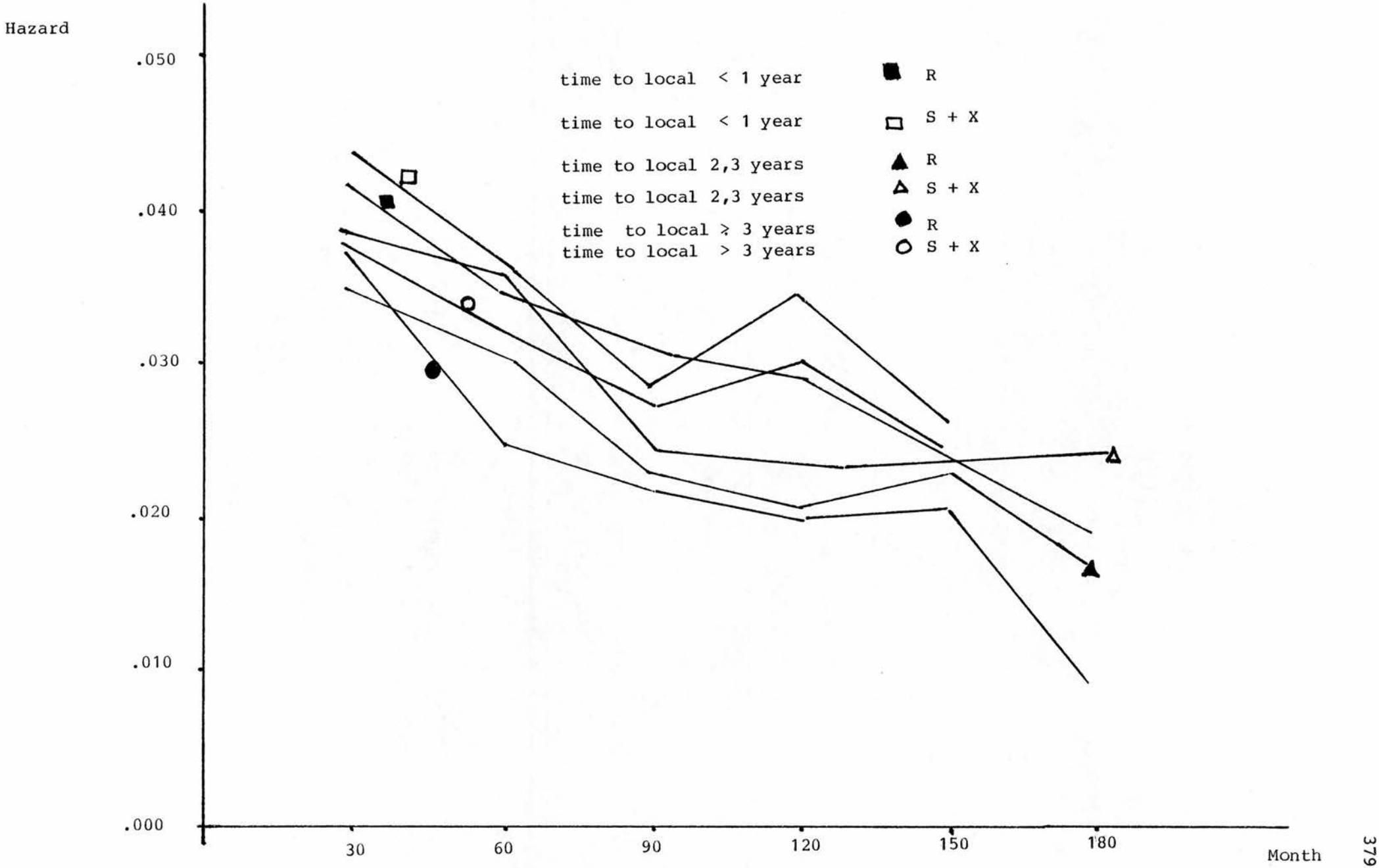


Figure (7.3.7) λ_D^L Stratified according to time to local.

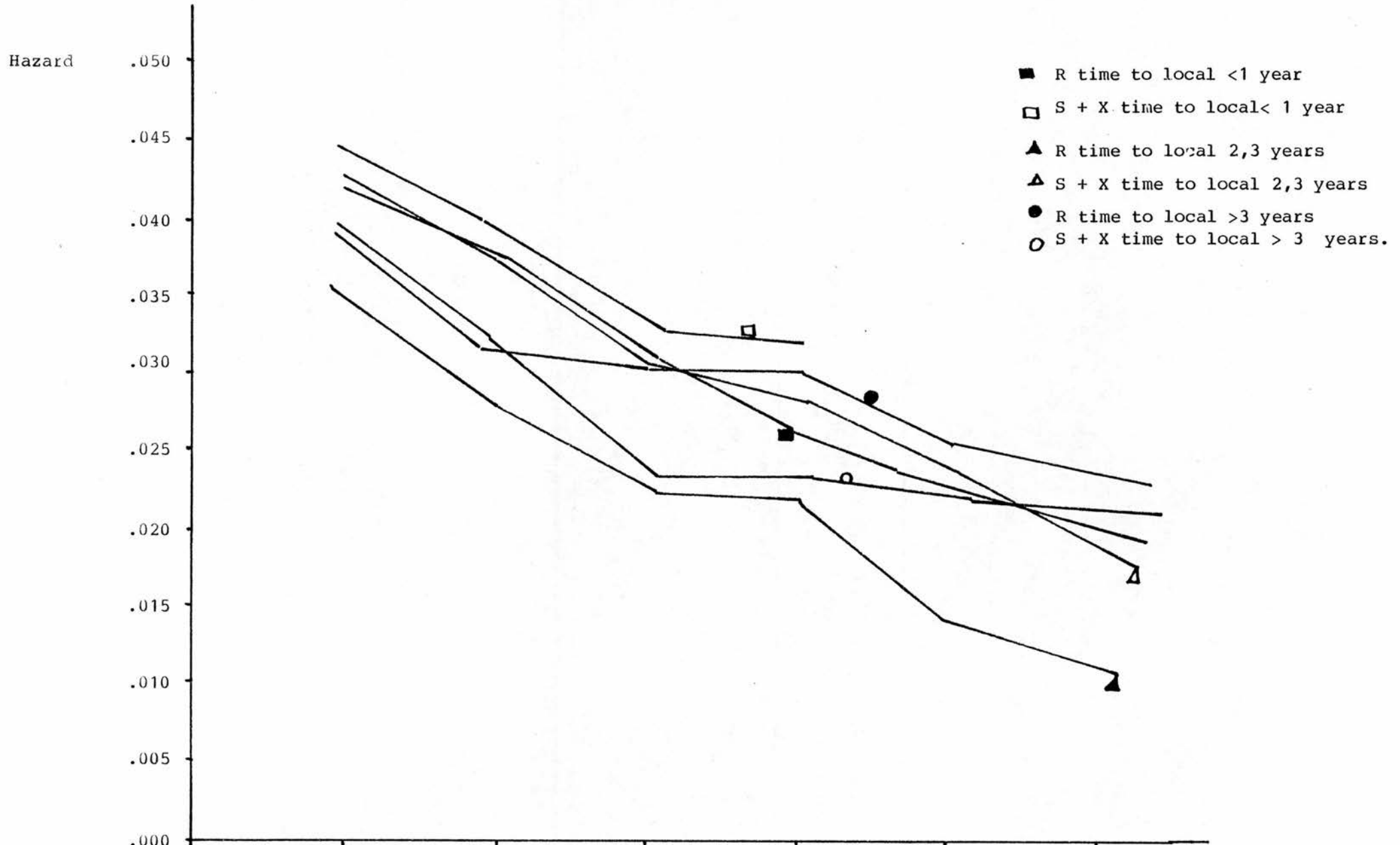


Figure (7.3.8) $\lambda^{L}_{D,M}$ Stratified according to time to local recurrence

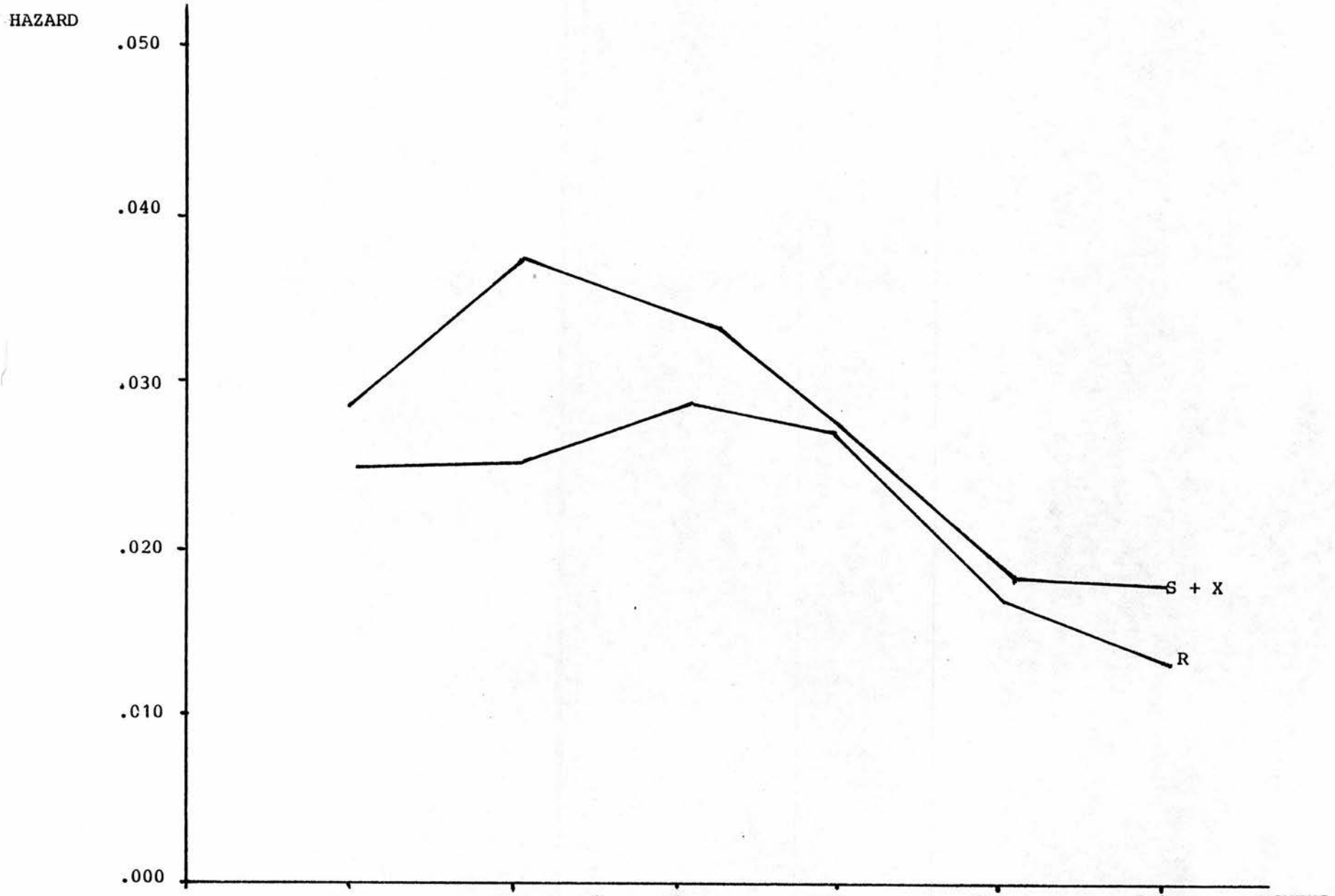


Figure (7.3.9) , λ_M^R

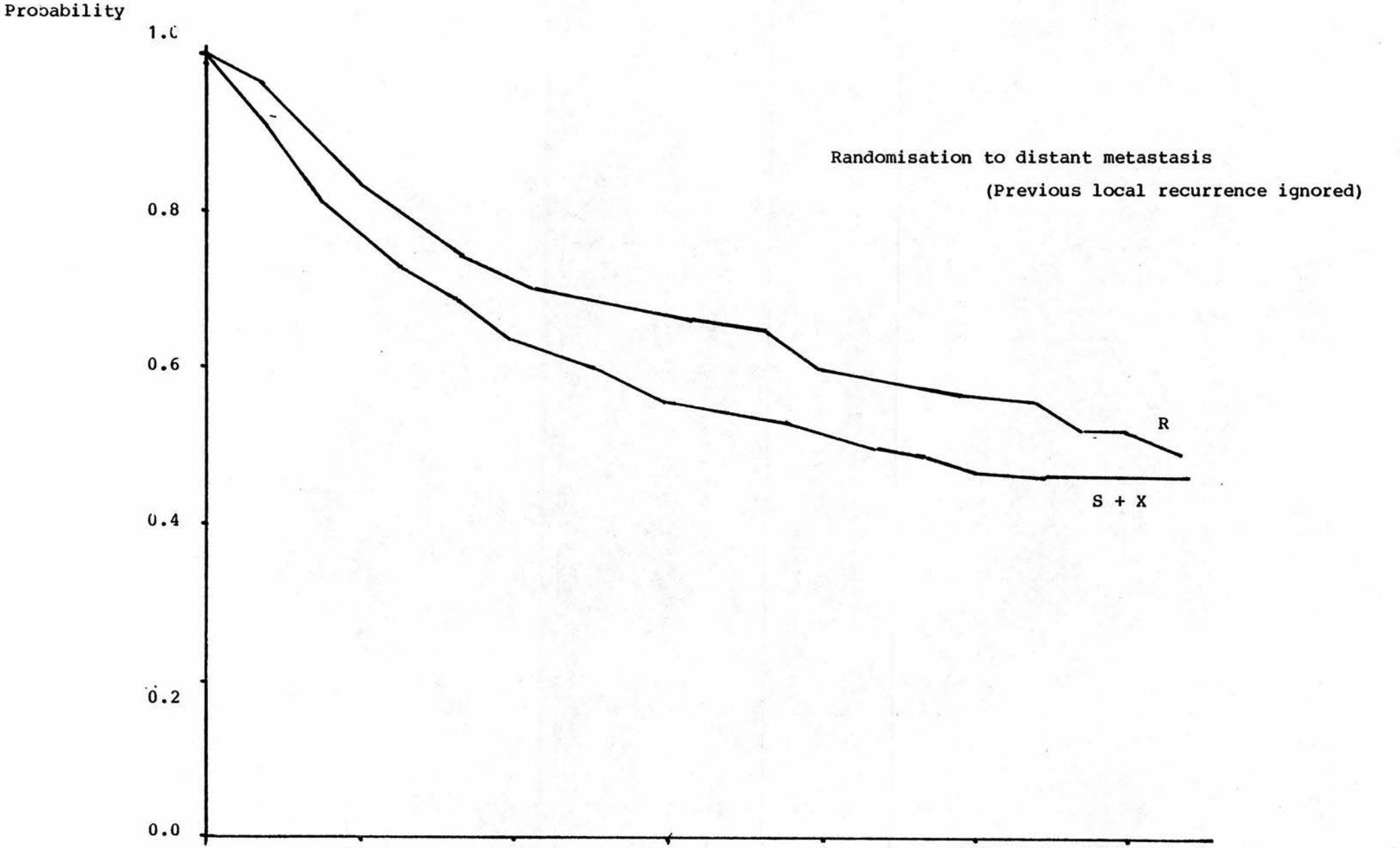


Figure (7.3.10), S_M^R

death. The logrank test for the treatment differences indicates a value of 4.85 with the significance level of 0.0277.

The period to the appearance of the local or metastatic disease is further used as a stratifying variable for a comparative analysis of radical versus simple surgery with XRT in terms of the response variable with the hazard $\lambda_D^{L,M}$. We define three strata, based on DFI.

	Total R	No.R Dead	Total S	No.S.Dead.	X^2	P
DFI $\leq$ 1 (Rvs)	22	21	28	28	9.39	.0022
1 < DFI $\leq$ 3 (Rvs)	41	38	50	49	1.48	.2240
3 < DFI (Rvs)	69	53	61	47	0.66	.4165

Then for patients recurrent after the 1st year their having had radical surgery is less likely to benefit the patients. Figures (7.3.11) and (7.3.12). However for those recurring early there is benefit in terms of survival by a radical surgery. We will show later this is not an indication of interaction.

We will continue the analysis by inclusion of a time dependent covariate related to the disease free interval, using the Cox's proportional hazard model. We will use the formulations which were presented in the early parts of this section on the intervening events. Now we will use such concepts to detect departures of specific type from the proportionality of hazards. In particular we are interested in the group of patients showing an early recurrence of the disease. We will define a function t* which is the time to the detection of recurrence. We first analyse the data according to the relative risk function of time of first recurrence to death, in presence/

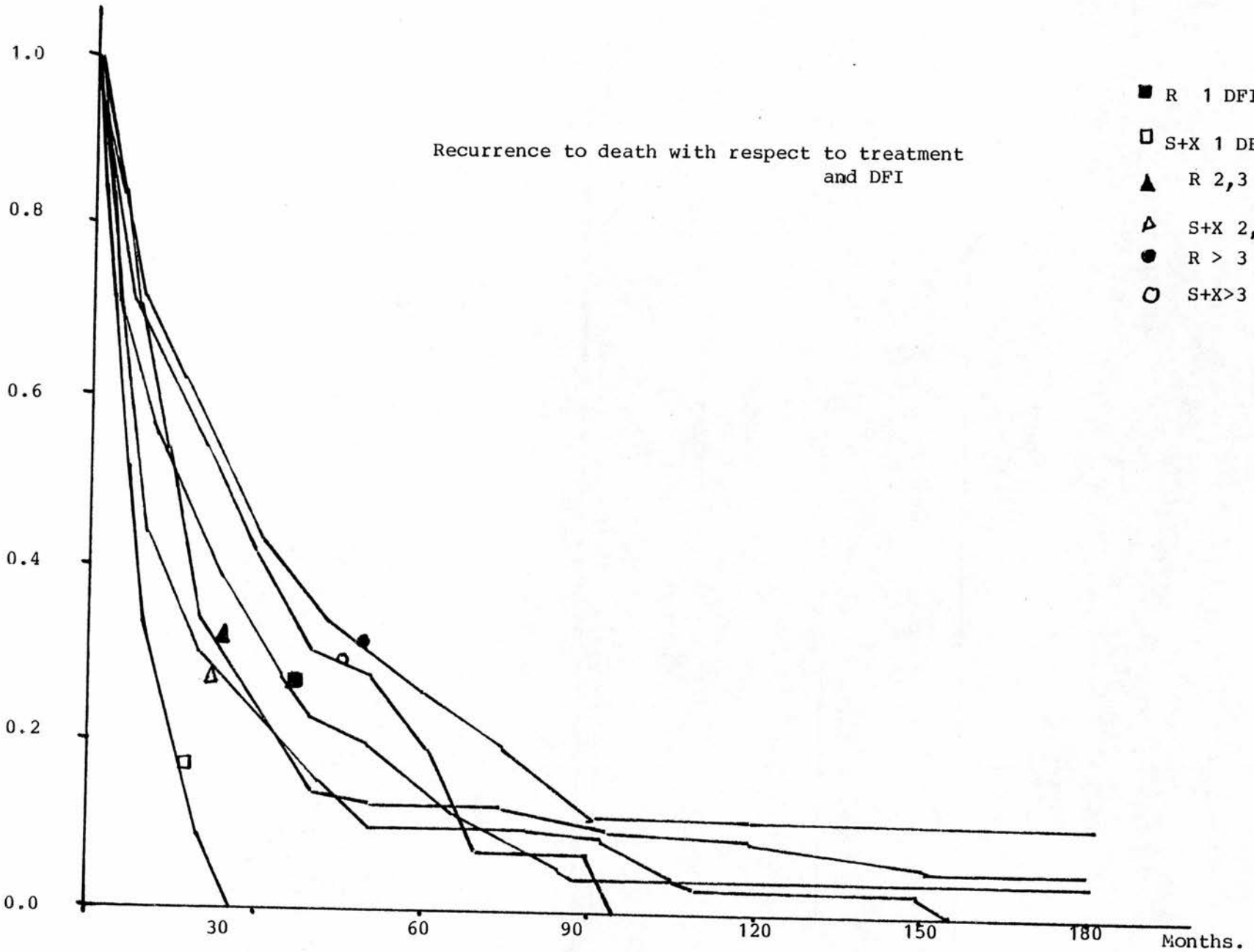


Figure (7.3.11) Survival for DFI

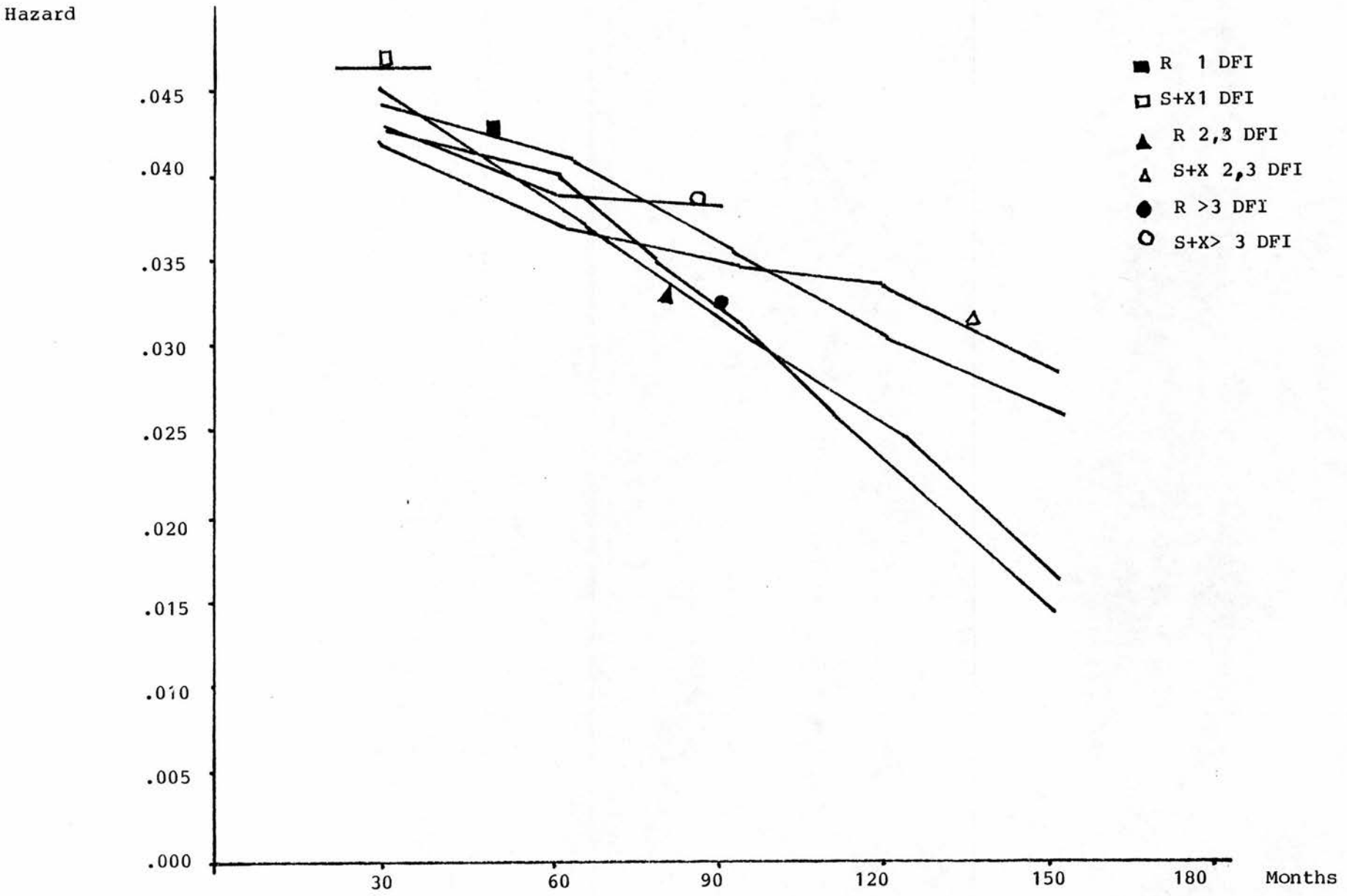


Figure (7.3.12) λ For DFI

presence of treatment effects and time dependency due to the DFI. That is a model of the form,

$$RR_{D}^{L,M} = \text{Exp}\left(\beta_{\text{treatment}}\ \text{treatment} + \beta_{t}\ \text{treatment}(\log(t^*)-2)\right)$$

A model of the treatment effect gives

$\beta_{\text{treatment}}$ = .1475 S.E. 0.0658 X^2 = 5.20 P = 0.0225

which closely approximates the logrank test where no time dependency is included in the model. With the inclusion of a time dependent effect we have

$\beta_{\text{treatment}}$ = .1487 S.E. = 0.0698 X^2 = 5.81 P = 0.016

β_{t} = -.0720 S.E. = 0.0501 X^2 = 2.85 P = 0.0871

There is not sufficient evidence to conclude confounding over the disease free interval. However as a priori one tailed test there is an indication of narrowing of the two treatments.

Finally in this chapter we will develop a methodolgoy for a family of functions for the analysis of an intervening event in a clinical trial using Cox's proportional hazard model. Clearly there are difficulties attached to the analysis of trial data if in the course of progress of disease there are a few routes acting which differ for various patients. By a fixed covariate approach and the proportional hazard assumption we may do a useful analysis as long as there is not a crossover of the hazard rates. The β estimator provides a good basis for the interpretation of data.

One of the problems in such an analysis is that often the present methods of treatment may not affect the total survival time/

time but rather may lead to differing qualities of survival depending on the development of the progression of disease. The example of analysis by the semimarkov procedure gives a representation of the problems involved. The method we will develop in this section in continuation of (7.3.2) allows a formal test to be performed for the intervening event. By testing the rate of change to the event of interest prior to an intervening event and post intervening event for a particular treatment or subgroup it is possible to detect departures from the proportional hazard assumption. Much of the work in this area is concentrated in the actual estimation of the parameters. In line with the developments of the last section we will continue by concentrating on the functional forms of the time dependency. According to the previous definition we considered two forms of logarithmic and linear time dependency.

Now we will develop a functional form by which we may study the pattern of development of risks by adjusting the rate of severity of the intervening event to be a function of the time scale. That is we have a relative risk function of the form

$$\text{Exp}\ (Z_1 \beta_1 + f(Z_1 . t^{\alpha})\ \beta_t)$$

where $-\infty < \alpha < \infty$. The importance of the intervening event may then depend on the component of time prior to and after the event. Figures (7.3.13) and (7.3.14) represent the two possibilities. The figures also present various functions of t^{α}. There is an area of close overlap within t^{α} which covers ln (t) and also Exp(t) for the value of t. In the example if we consider metastatic disease to be an intervening event there exist three time dependence variables, t_m, $t_D - t_m$ and t_D. The t_D represents the survival time within/

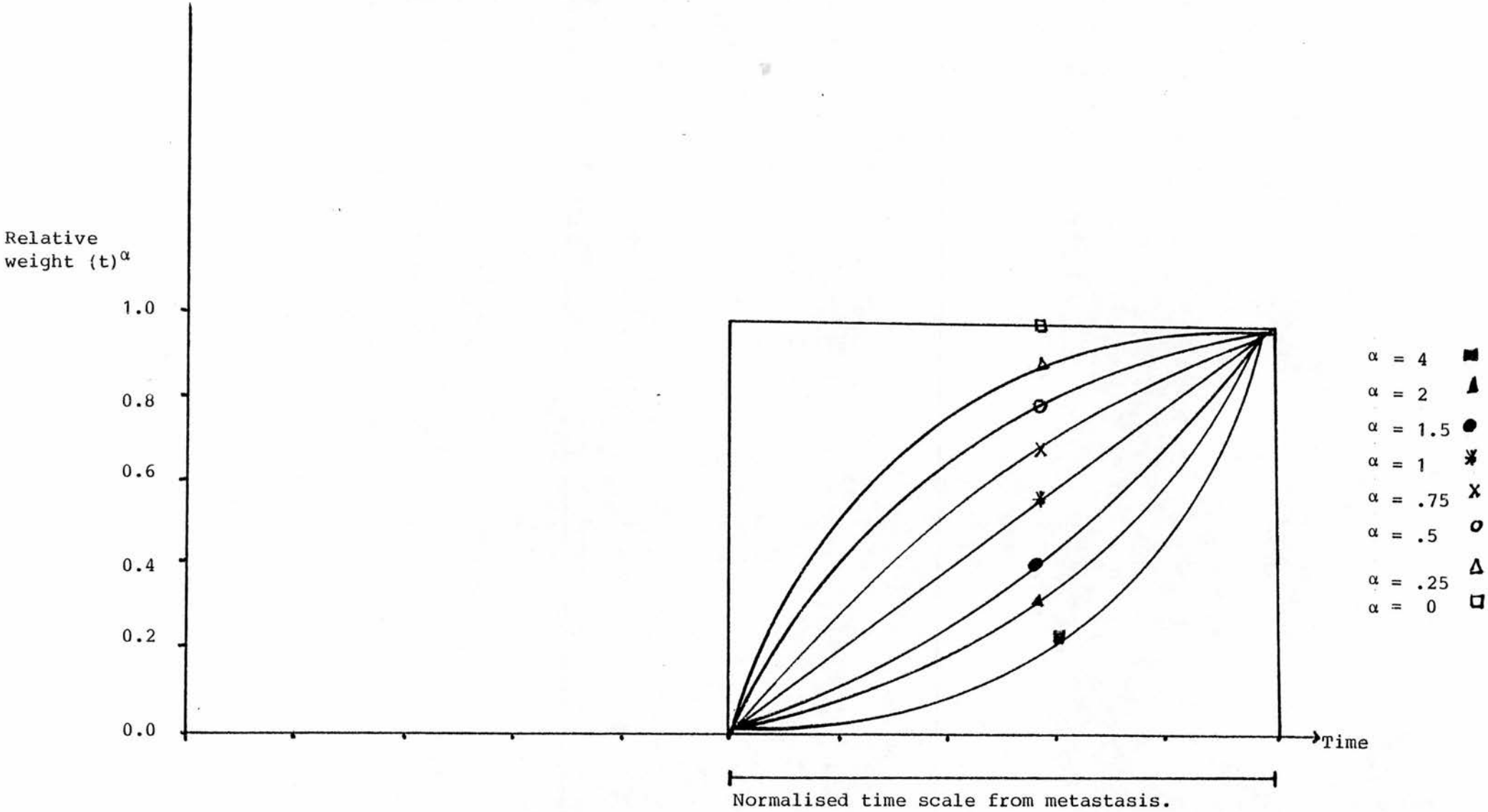


Figure (7.3.13) Functions of various values of α in normalised time after metastasis, to the power of α

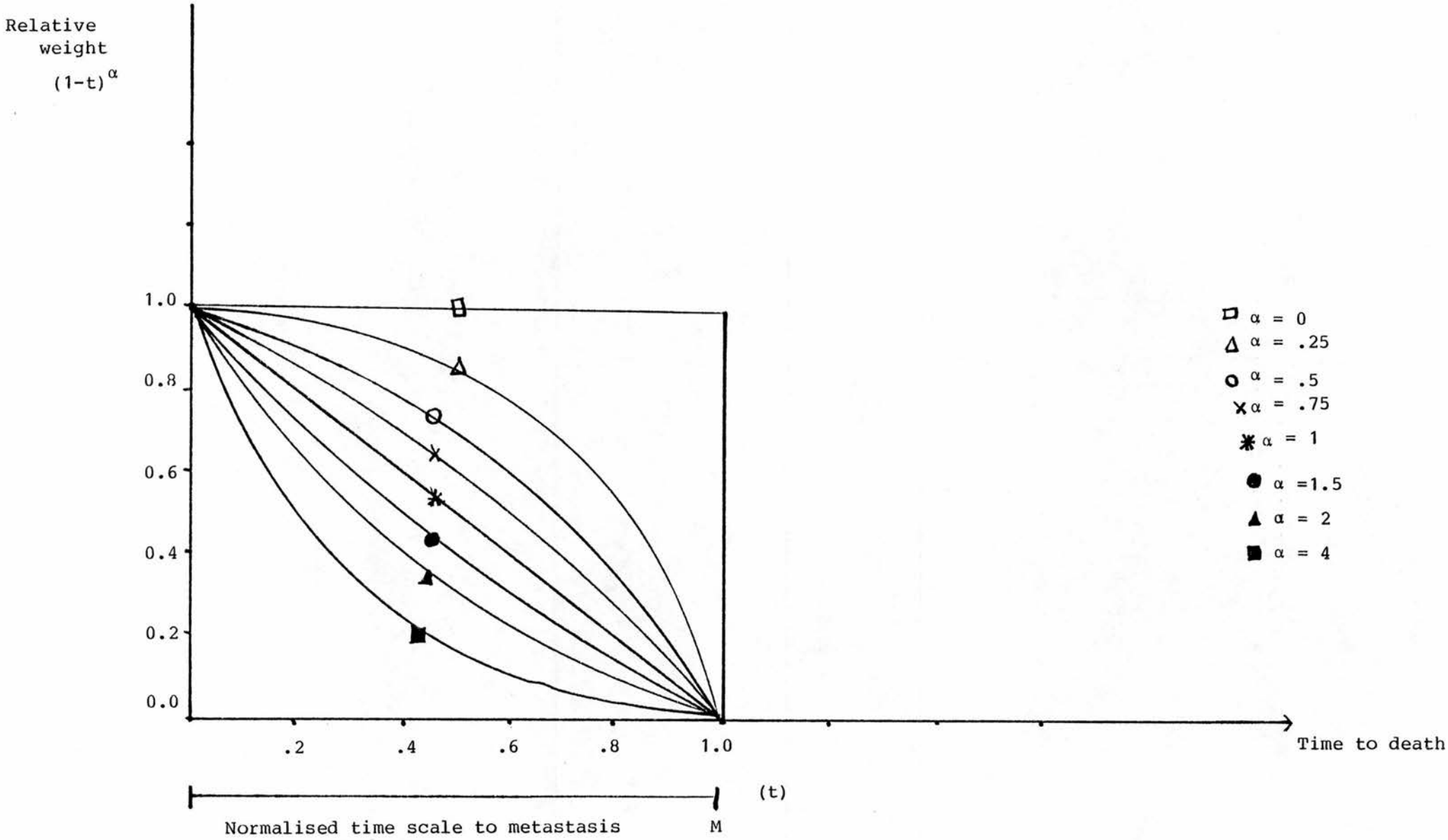


Figure (7.3.14) Functions of various values of α in normalised time to power of α.

within which all remeasurements are made. The t_m and $t_D - t_m$ are periods of time that subdivide t_D. As an extension we let α represent a weight function for the transformation of the periods. Detection of the metastatic disease implies a progression of the disease for both treatment groups. In the comparison of the treatments however we expect the proportionality of the hazards to hold throughout the time scale. By letting $\alpha > 0$ we can test the metastatic disease or other intervening event progress in terms of deviation from proportionality. This transformation is analogous to the Exp $(t_D - t_m)$ type of time dependency, by which the longer the period of survival after intervention, the more risks increase. Figure (7.3.13) with $\alpha > 0$ shows a situation where there is a build up of high risks from intervening events.

Alternatively we consider the time t_m and the test of the period up to the intervening event. A possible transformation as presented in figure (7.3.14) is then by $\alpha = 0$, which implies that non proportionality due to the intervening event may be assumed at a constant risk previous to the detection of the intervening event. Further the transformation of $0 < \alpha < \infty$ is a situation within which cases are initially at high risks of showing a survival pattern more critical than the proportional hazards assumption, but with the passage of time the two treatment groups produce proportional rates. At $\alpha = 1$ we have a replicate transformation of the actual time scale. What is of importance in all these transformations is the magnitude of the relative weights at each period of time, in comparison with the adjoining times. Therefore for reasons of dimensional symmetry and also a faster convergence of/

of the Newton Raphson procedure we use the time scale,

$$t_m \rightarrow \left[\begin{array}{cc} t_m^{\alpha} & / \\ & \delta t_m^{\alpha} \end{array} \right]$$

and $$t_{D-m} \rightarrow \left[\begin{array}{cc} (t_D - t_m)^{\alpha} & / \\ & \delta (t_D - t_m)^{\alpha} \end{array} \right]$$

We thus recall the relative risk function for the treatment only model in the full time scale of randomisation to death.

$\beta_{treatment} = 0.3677$ S.E. = 0.1168 $X^2 = 9.97$ P = 0.001

We will now consider a scaling of the time from metastatic disease to death with α set to a value in the range 1 to 0. Previously we defined α between $-\infty$ to $+\infty$. Clearly in here value of $\alpha = -\infty$ will transform the measure of time dependency to zero, that is t^{-x} as $x \rightarrow \infty$. Given this situation in fact we will return to the model with no time dependency included. The initial value of α we consider is at zero. In the earlier part of this section we derived the value of time dependency according to $Z(t) = 1$ for time after metastatic disease. This is in fact the same as $Z(t) = t^0$. The estimated β values are,

$\beta_{treatment} = 0.3982$ S.E. = 0.1170 $X^2 = 11.61$ p = 0.0008

$\beta_t = -0.1239$ S.E. = 0.0780 $X^2 = 3.096$ P = 0.07488

Indicating there is no suggestion of lack of proportionality of tne type with a constant scale after metastatic disease.

Now we consider a linear effect of the metastatic disease.

That/

That is

Z(t) = linear normalised time after metastasis.

giving, [(t/15) -2], where t is time after metastatic disease. Thus

$$RR = \text{Exp}\,[\beta_{\text{treatment}}\ \text{treatment} + \beta_t\ \text{treatment. time}]$$

$\beta_{\text{treatment}} = 0.3855$ S.E. $= 0.1181$ $X^2 = 10.11$ $P = 0.002$

$\beta_t = -0.1389$ S.E. $= 0.0847$ $X^2 = 2.69$ $P = 0.0956$

Referring to table (7.3.1), we present the transformations of the time scale for nonproportionality. Use of the various power transformations of the time scale is a good check on the consistency of the results that may be obtained. In the present context the non-proportionality does not show a significant deviation from the proportional hazard model; however we note that at $\alpha = 0.4$, the scale of non proportionality is at the most efficient value.

In fact for the present data the different power transformation do not influence the estimator of treatment a great deal. As a general conclusion the appearance of the metastatic disease does not influence the assumptions of the model. The final conclusion of the present chapter in fact conform with the analysis of the Chapter 6. A point of interest howeve is that in the analysis of this chapter we have not considered only one event variable but rather two intertwined processes through time and have concluded that the events through time do not influence the conclusions of our study. The implication in medical terms is that the relative risks between the two treatments according to this data do not provide evidence of a difference for times prior to and post metastatic recurrence. Since the occurrence of metastatic recurrence is an intervening random variable/

variable we use different transformation with α and again there is no suggestion of a deviation from the above finding.

		Estimated value	S.E.	X^2	P
$\alpha = 0$	$\beta_{treatment}$	.3982	.1170	11.22	.0008
	β_t	-.1239	.0780	3.18	.0749
$\alpha = .1$	$\beta_{treatment}$	.3985	.1163	11.25	.0008
	β_t	-.1321	.0775	3.02	.0822
$\alpha = .2$	$\beta_{treatment}$	.3992	.1158	11.43	.0007
	β_t	-.1368	.0775	3.15	.0761
$\alpha = .3$	$\beta_{treatment}$	.4021	.1151	11.75	.0006
	β_t	-.1381	.0749	3.44	.0639
$\alpha = .4$	$\beta_{treatment}$	.4034	.1148	11.76	.0006
	β_t	-.1411	.0721	3.89	.0484
$\alpha = .5$	$\beta_{treatment}$	.4029	.1152	11.78	.0006
	β_t	-.1411	.0734	3.82	.0509
$\alpha = .6$	$\beta_{treatment}$	.4018	.1159	11.45	.0007
	β_t	-.1408	.0751	3.65	.0559
$\alpha = .7$	$\beta_{treatment}$	.4015	.1163	11.26	.0008
	β_t	-.1407	.0775	3.45	.0637
$\alpha = .8$	$\beta_{treatment}$	.3885	.1169	10.83	.0010
	β_t	-.1401	.0809	3.02	.0822
$\alpha = .9$	$\beta_{treatment}$	.3867	.1176	10.53	.0012
	β_t	-.1395	.0825	2.94	.0861
$\alpha = 1.0$	$\beta_{treatment}$	.3855	.1181	9.97	.0016
	β_t	-.1389	.0847	2.79	.095
$\alpha = 1.2$	$\beta_{treatment}$	.3842	.1201	9.83	.0017
	β_t	-.1349	.0897	2.35	.126
$\alpha = 1.5$	$\beta_{treatment}$	.3769	.1211	9.75	.0018
	β_t	-.1211	.0928	2.19	.138

Table (7.3.1)

CHAPTER 8

PROGNOSIS IN BREAST CANCER

The purpose of this chapter is to evaluate the importance of certain prognostic indicators in a group of breast cancer patients. In this section however we make a distinction between indicators that are regularly assessed in the staging of patients and some other indicators that have not been considered a great deal in the past. The present data is related to a group of patients diagnosed as having breast carcinoma and referred to by H.J. Stewart et al (1968). We will deal later with the data and the procedures for its collection and the various measurements made on the patients. Before considering the data however, we will remark on certain important trends in the study of prognostic indicators and the auxiliary indicators that are used in the analysis.

In this study we are not so much concerned with substantiating a major disease indicator but rather to consider if some of the survival time variability of the patients may be attributed to some measurements outside of the usually accepted prognostic indicators. Thus the findings of this study may be of some value in a subsequent sample of patients. In the discussions of what follows we will refer to a number of variables. In here we will describe these variables and later refer to them in their short notation. Through the course of the discussion more necessary details and references to some of the variables will be given.

Tumour/

Tumour Contour, types will be discussed in greater detail further in the next section and figure (8.1.1) refers to the classification of the tumour.

Inoperability, is also referred to in more detail in the discussion. Basically inoperable patients are patients who have had a spread of the disease to the extent that no surgical treatment is performed.

Size, is considered to be the maximal tumour diameter of the initial tumour.

Node refers to the involvement of the axillary nodes according to histological findings.

Extent refers to the depth of the initial tumour.

Grade is the histological grade of the initial tumour and is discussed further.

Presence of complicated change.refers to the type of tumour where there is evidence of abnormal skin distant from the main tumour. These include thickening of skin overlying tumour, blurring of tumour outlines and the dilation of adjacent veins. These effects are observable by X-Ray.

Tumour foci refers to two possible types of tumour, these being either single or multiple foci.

Micro calcification is a method for detecting areas for histological examination. In here we define possible areas where tumour calcification had been shown.

8.1 Methodology and sources of data.

Two studies in the past have mentioned the value of tumour/

tumour contour types; Ingleby et al (1960) and Lane et al (1961). Neither of these studies however were concerned with assessment by use of a probability measure of difference between the patients.

In the paper by H.J. Stewart two methods are discussed in the assessment of the tumour contour types of 157 patients. One is paper section and the other is mamography. Further in the paper they mention a few other measurements on the actual distributions of the contour types, such as presence of complicated disease, extent, tumour change etc. In the present study we will use the same data for assessing survival distributions for different subgroups of patients in a more complete analysis of the data.

The grouping of breast cancer patients by clinical staging is now a good guide to survival assessment. However in 1958 Harmer recognised " at least ten systems all basically the same but each irritatingly different from the next". The present system is attributed to Union International Centra Cancer and is a resultant system from various systems that have been used in the past. A single Manchester system was in use in Britain up until 1958 when the staging was replaced by the TNM effective mainly in Europe. As from 1966, a different general system was adopted in the U.S.A. Finally in 1973 a system was adopted by the UICC and the American Joint Committee on Cancer staging with the (UICC/AJCC) giving the present method. This system distinguishes between pre-treatment and post surgery findings and is based on Node histology, size of the tumour and metastatic status. Further for the size categories distinction/

distinction is made for tumours with fixation to underlying pectorial fascia, and further for Node one cases with moveable homolateral axillary nodes; a distinction is made for a node containing growth and those with no growth.. Given these developments there is still an enormous variation within any single stage. This is partly due to the effectiveness of treatments. If the treatments were more effective for all patients, there would be less emphasis in classifying cases more precisely. However part of the problem in the assessment by classification is that is a crude categorising procedure of a complex biological process of host tumour in time, and is far more complex than an assessment made by a single instantaneous measurement for a single time. A less "subjective" assessment on patient tumour process and survival prediction would ideally require repeated measurements in time. This is however, not practicable in that for clinical reasons it is accepted that any diagnosis of breast cancer requires immediate treatment.

From a different point of view, other studies, J.E. Devitt (1967) have indicated that the clinical stage of breast cancer may not be a measure of degree or extent of growth so much, but a measure of tumour biological potential and host reaction. With the present study we concentrate on the survival time of patients as the only response variable. However we will not only deal with the study of static prognostic indicators and a "frozen" patient resistance but rather we distinguish between static indicators and indicators containing information about changes or progression. The choice of an indicator as static or one containing growth is difficult in that almost any indicator can be considered to contain an indication of growth.

growth. As an example, a variable that we will not consider as time-dependent but may contain such an effect is shape of the actual tumour. The pattern in the growth of the tumour may be related to the form of body resistance to it. (See shape of tumours figure (8.1.1)

At this point we will present example of a type of study where measurement over time has been of some use in the study of breast cancer. It is generally accepted that early treatment improves prognosis of patients. However there is a lack of consistant evidence in regard to the value of early diagnosis in the improvement of survival times. A study was carried out by Bloom (1965) to test whether a prompt diagnosis of breast cancer improves survival as assessed from the date of first symptoms and whether the delay between the appearance of the first symptom and diagnosis has become shorter in the recent years. This study in fact reiterated the commonly held view that cases with a short delay between the appearance of the first symptom and diagnosis have a better long term survival rate than those with long delays.

In this context it may be taken that the delay is in fact a representation of the growth of the tumour. In the studies of time dependencies as in other multivariate studies, the order of incorporating a variable into the model is of some importance. Often studies of the patient classification is measured by the staging of the tumour. If delay is taken to be a prognostic indicator it is measured after staging category effects have been removed. Thus in the above example one problem with measuring the tumour development based on delay time is that it may be confounded with /

with certain other factors inherent for each individual patient. Thus the question may be phrased as that of assessing the value of delay after staging variables have been statistically removed.

With this introduction on the types of models of interest we will return to the description of the data as mentioned earlier in this chapter. Initially we will use cross tabulations to show the numberical association of the indicators between themselves and with the number of cases alive at the end of the study. Later we will use a Cox model assuming a constant relative risk throughout follow-up. Then we will consider the estimation of the Cox's model, allowing for time dependent effects of prognostic variables. Over 2000 mamographs have been studied from 1963 to 1967. Among the cases with mamograms, 306 cases had a diagnosis of first time breast cancer. This group has certain patients for whom the data is inadequate and thus 98 cases have to be removed, so that the remaining patients are a more defined group of patients. The 98 cases that were excluded are largely defined by the information collected at the initial X-ray sessions. 53 of the patients had been previously treated for breast abnormality. 14 of the patients were initially diagnosed in wrong subgroups in terms of their form of malignancy and thus were also excluded. 7 cases had either an unusual malignancy or had post-operative death. Finally 24 cases had inadequate clinical information after diagnosis or mamograph were of inadequate standard.

	Number	% of 306
Previously treated	53	17.3
Uncodable diagnostic error	14	4.6
Atypical malignancy	3	1.0
Post-operative death	4	1.3
Inadequate clinical details	13	4.2
Inadequate films	11	3.1
Total excluded	98	32

The 11 inadequate films were also taken at the beginning of the entry month when the technique was still being perfected. There are 208 remaining cases who had a median follow-up time of 11½ years, with a range from 4 to 18 years. For this group 163 had died at the time of study. No cause of death was recorded for the cases but the actual date of death is available. Of the remaining 45 patients with censored survival data , 17 of the patients had attended on annual review to one year prior to the time of study, 21 were dismissed after 10 years of follow-up and 7 patients were lost to follow-up with less than 10 years of follow-up. Therefore for the 208 patients 78% have a recorded death time. The follow-up information in this study was mainly obtained through extraction of relevant follow-up information from the Cardiff clinical notes in 1981.

Four tumour contours are defined and this definition is related to the type defined by Ingleby et al (1960). They define 3 types of tumour, irregular, smooth or mixed outline. Further they represent better survival for smooth or circumscribed tumours. In the present study an additional subdivision is made. Between the extremes of smooth and spiculated, two categories are defined, namely/

namely mixed tumour with well defined smooth and spiculated parts to their outline and conglomerate tumours which have a mulberry appearance macroscopically and have a blurred and irregular but not definitely spiculated outline on the mamograms. Further 31 or 15% of the 208 maomgrams had evidence of malignancy but no tumours show. Thus there are 5 groups in all

Smooth Conglomorate Mixed Spiculated

Figure (8.1.1) Representation of the contour types.

The mthod of obtaining mamographs was reported in 1968 based on the Egan techniques. The assessment however considered only the first 60 patients and used the Gough and Wessthon technique of paper mounted thin whole breast sections.

Several further radiological features were also recorded during the initial examination. For all cases size was recorded in millimetres. Microcalcification was also noted at the special X-ray review sessions and thus patients were categorised into calcification present within, on the outset or both within and outset. Clinical inoperability is a criterion that is not strictly definable clinically and thus patients with no sign of metastatic disease and operable tumours were recorded as operable cases and if metastatic disease is present or the tumour is inoperable they are classified as inoperable. The point about inoperable and operable patients is that they do represent very different groups of patients and in the final analysis we will distinguish between statements made in this regard.

Axillary/

Axillary node involvement is another well established clinical indicator and thus cases are grouped into node negative and node positive groups. Bloom and Richardson (1957) define three histological grades for lesions, which we use in this analysis. In terms of shape of tumour we distinguish between multiple and single foci.

Picard J.D. (1962) has defined several well recognised features that can occur in the normal breast tissue around the tumour shadow on the mamograms of advanced primary lesion. This is termed as tumour showing complicated change and as thickening and straightening of the travecular shadows, thickening of the skin overlying tumour, blurring of the tumour outline and the dilation of adjacent veins. The above features are present on X-rays when there is oedema present clinically but they were also noted at mamographic review sessions of the data.

Apart from complicated change, extent is also studied, by separation of patients into greater than and less than ¼ inch deep tumours. Clinical size of the tumour and age of the patient complete the data.

Variable No.	Variable name	Description
1	Operability	Inoperable or metastatic,operable
2	Microcalcification	Within, outset,both.
3	Node	Negative, positive
4	Histological grade	I, II, III
5	Foci	Single, multiple
6	Contour	Smooth, spiculated, mixed, conglomorate.
7/		

Variable No.	Variable name	Description
7	Complicated change	Present, absent
8	Extent	Less than ¼ inch deep, greater than ¼ inch deep.
9	Size	
10	Age	

The result of an interim analysis based on 157 mamograms was published by Stewart (1968). In conclusion no significant relationship between contour types and certain prognostic indicators whether considered separately or together was obtained. However a trend was noted contrary to findings of Ingleby and Gershon - Cohen (1960) and Lane et al (1961) suggesting a better prognosis for spiculated tumour and bad prognosis in smooth and also possibly mixed lesions. The 1968 analysis however did not deal with any of the other indicators that we mentioned earlier in terms of survival times.

8.2 Categorical distributions of the prognostic indicators.

Initially we perform a preliminary analysis based on Cross tabulations of the prognostic indicators. Two well known and accepted indicators are node histology and the initial size of the tumour. The extent of the progress of the disease is important in so far as we have to distinguish initially between the inoperable and operable cases. The main group of interest are in fact the operable patients. However, we will discuss the distribution of the inoperables in the early stages of the analysis.

75% of the 208 cases were treated by mastectomy but some were with palliative intent in the presence of clinical inoperability. 5 clinical groups may in fact be defined. 131 cases belong to the accepted operable group of interest. 15 further cases had positive contralateral mamograms and their survival distribution is similar to the operable group. All 15 had a second mastectomy from 1 to 27 years after the first. In contrast to these two groups, there are 3 remaining groups in whom both mean and the median survival times are considerably less. 20 patients had local but clinically inoperably tumours and a further 20 have been termed inoperable solely because of the detection of the involved supraclavicular node at mastectomy. 22 others presented with systemic disease comprise the final group. In terms of the progress of disease the patients were separated into operables and inoperable patients. The main aim is to consider prognostic indicators for the operable group. Clearly the operable patients contain a smaller proportion of node positive patients (43%) to inoperable patients with 55% node positive cases. These results are clearly in line with expectations that inoperable patients are more advanced and thus they contain a higher proportion of patients with axillary node involvement. In fact operable cases as a group of less advanced disease patients have a higher proportion of patients in better prognosis groups. For the grade of the tumour the inoperable cases have 74% of the patients with grade 3 tumours and operable cases 52% grade 3 tumours. This is not surprising and only conforms to what is expected. (Later we will discuss the grade categories in more detail in more detail in relation to other categories).

For/

For the extent of the tumour, operable cases present a 9% proportion with greater than ¼ inch depth as to 32% for the inoperable patients. Although this result conforms to what is expected it is also in line with a hypothisis which assesses extent as a time-dependent indicator of progress. Once again the same conclusions are obtained when we consider complicated change. 22% of the operables present evidence of complicated change in the initial tumour as to 63% of the inoperables.

Multiple foci tumours form a small number of patients altogether . We obtain 16% multiple foci group among the operables and a 23% multiple foci group for the inoperables. One indicator that does show a similar distribution for the operables and the inoperables is the tumour contour shapes. We will study these categories further in terms of survival but at this stage there is no evidence to link tumour contour types with those of the progress of the disease. Calcification present within or on outset also is similarly distributed for the operable groups versus the inoperable group. At the end of follow-up we also note that 36% of operables are still alive, compared to 1% for the inoperable cases.

Node involvement is a further accepted prognostic indicator in so far as this study is concerned. We will at first consider node involvement for the total population and in some important categories mention the distribution of node involvement for the subgroups of operable and inoperable cases.

Two/

Two other variables that may indicate progression process are extent and complicated change. These two do not show statistically significant associations with the node categories. Greater depth tumours are present in 6% of node negative and 17% of node positives. Complicated change amounts to 25% of node negatives and 31% of node positives. It is difficult in here to conclude what extent and positive nodes imply, but it is an indication that in terms of good and bad prognosis value, extent is describing something slightly different from that of node status. Finally for node and survival status at the end of follow-up, there are 30% alive patients with initially no nodes involved and 18% alive with nodes recorded as initially involved.

For the total of grade categories there are 12%, 16% and 17% of tumour with higher extent depth at 0 & 1, 2 and 3 grade levels respectively. By separating operable cases again there is not a major deviation from the above for each subgroup of operable and the inoperable. However by considering node negative patients against the node positive the percentage value of the above categories of the grade change to 18%, 12% and 9% for the node negative and 11%, 12% and 10% for the node positives. Thus in terms of classifying patients into good to bad prognosis there seems to be again an indication that node status and extent may be defining different attributes of tumour progression for each grade category. It must be pointed out in here that the above percentages are presented purely to illustrate distributional patterns of subgroups of patients. In the next section we will present survival distributions and the relevant/

vant statistical tests.

Grade in terms of percentage of cases showing complicated change gives values of 18%, 31% and 38% for grades 0 & 1, 2 and 3 respectively. This is a similar pattern for direction as that of extent. (Although the complicated change values are significant at $p < 0.001$ extent categories are not). By the definitions of extent and complicated change it is possible that they are explaining similar effects of the tumour progress. Once again by subclassifying by the operables and the inoperables we do not obtain a major deviation. However for the node status the same pattern as that of extent emerges. That is for node negative patients and the respective values of grade we obtain 29%, 22% and 20% showing complicated change. While we obtain 31%, 30% and 30% for node positives showing complicated change at grades 0 & 1, 2 and 3 respectively. The conclusion from this pattern is clearly the same as that of extent of tumour. However a point must be emphasised that up until now we have considered relative effects in terms of prognostic distributions and we have not dealt with the survival times.

In terms of contour types we do not detect any interesting distributional patterns for the various values of grade. Grade in relation to multiple foci and calcification distributions give once again a uniform pattern.

Calcification within and at outset together with single or multiple foci tumours also show no significant association with any of /

of the other recorded variables.

The mean size of tumour is 3.25 centimetres. For the operables and the inoperables we do not detect a major difference . Size also gives a similar pattern for the node negative and the node positive patients. However extent and complicated change both show a slightly different mean value at good and bad prognostic levels of extent and change. For the extent of the tumour less than ¼ inch we have the mean size to be 2.51 and for the extent greater than ¼ inch, 4.89 as the mean sizes. (t-test, $p < 0.001$) With the complicated change however this pattern is not represented so significantly. Tumours with no sign of complicated change have a mean size of 3.00 centimetres and tumours with complicated change have a mean size of 3.71 centimetres. (t-test, $p < 0.01$). Once again there is an indication that if complicated change is playing any role in classifying patients it relates to a different group of patients than the size category classification. For various other factors such as contour types and tumour foci calcification a similar value for the mean size distribution is obtained.

Status of patients, at the end of study indicates that contour type, tumour foci and calcification do not play a major role in determining survival of patients. Among the accepted indicator size, node and operability are the major indicators for determining survival. With regard to the present data however two other indicators are also of value; complicated change and the extent of tumour. We pointed out earlier, these two indicators refer to groups of/

of patients that are not identified by their node status, size or operability. Later in this chapter we will discuss those patients in more detail by considering survival probabilities.

Finally we will consider the age distribution of the patient according to various categories. The mean age of patients is 50.2 years. For the operable patients we have a mean age at 50.1 years and for the inoperables a mean age at 50.3 years. therefore the age distributors are very close. Node status categories produce again very close mean ages with the node negative patients being a little older than the node positive patients. Mean age for the grade of the tumour are also very close to the mean value, with higher grade patients, slightly older than lower grade patients. For lower than mean size groups we obtain again that age distribution is the same as the larger tumours. Extent and complicated change also produce the same lack of age diffierences. In the case of foci, calcification and contour pattern again we observe that age distributions are very close to each other in terms of the mean distribution of the various categories.

In the earlier part of this chapter we mentioned that some patients were excluded from the study. Altogether they comprise 32% of the 306 patients. From examination of the features of these excluded patients, we observe that in general the exclusions are uniformly distributed between the various categories of the indicators.

8.3 Prognostic indicators according to survival time. /

8.3 Prognostic indicators according to survival time.

Up until now we have considered groups of patients and the pattern by which they were formed into distinct groups. At this stage we deal with survival status at the end of study and the estimation of the survival functions for each of the distinct groups. The group of operable patients as my be expected have much better survival than the inoperable. For completeness we present the survival rates of the two groups , Figure (8.3.1).

The various categories of the indicators do not suggest a significant difference between any of the inoperable groups. However, we note that some of the indicators do not affect the survival times of the inoperables in the same direction as that of operable groups. This effect is due to chance rather than adequate statistical evidence for a real difference. The most striking effect with respect to inoperables presenting a survival trend in different direction as that of operables is given in figure (8.3.2). By which the two categories of contour types with speculated tumours show a slightly worse survival than smooth contour types for the inoperables, while in comparison of the operable groups the spiculated group do better than the operables with smooth contours. Due to the small numbers of the inoperables we will leave this subgroup and concentrate on the operable group only. Clinically, the operable group are of more interest in terms of prognosis since they are composed of patients with less advanced diseases.

Initially/

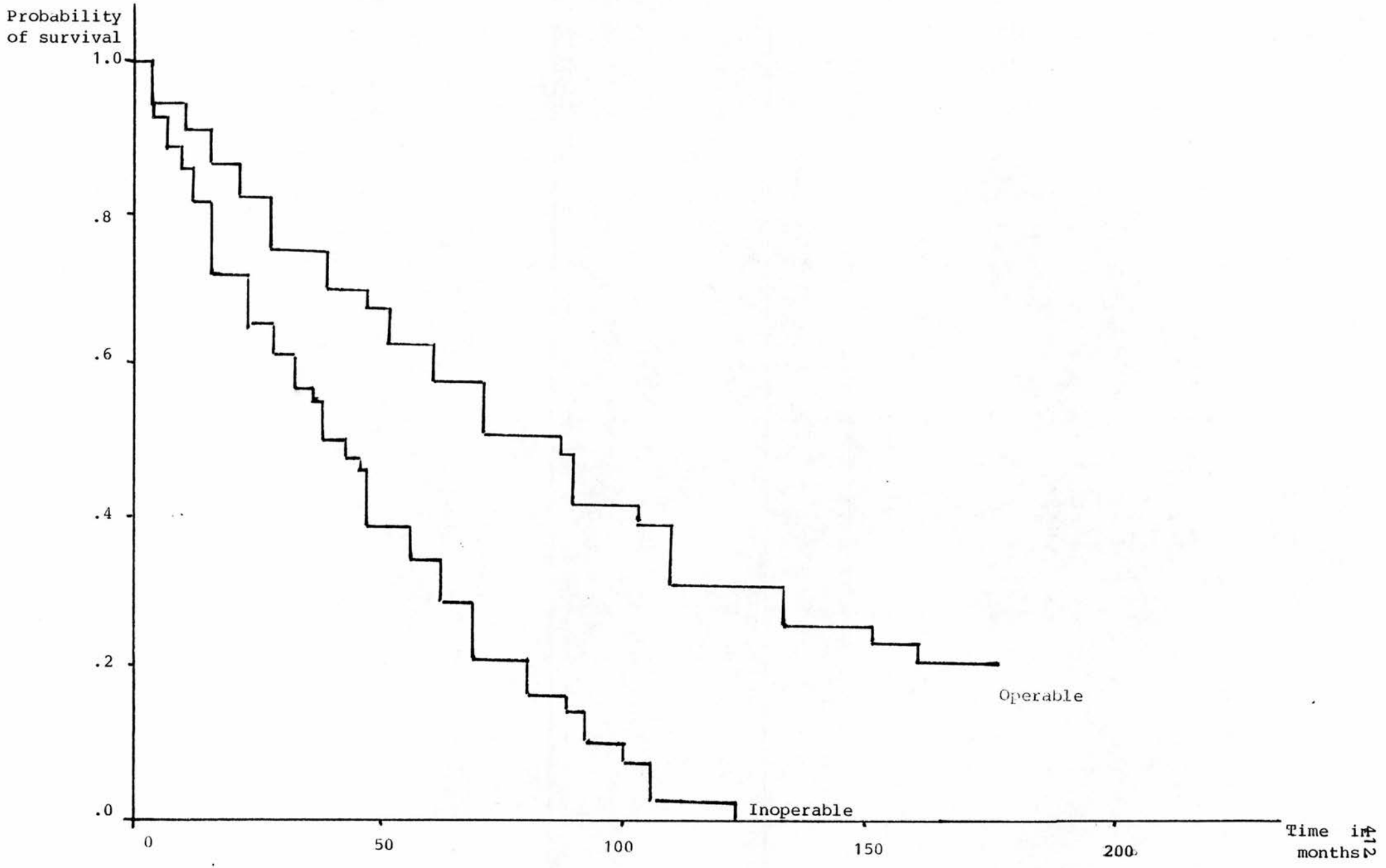


Figure (8.3.1) Survival rates according to operability

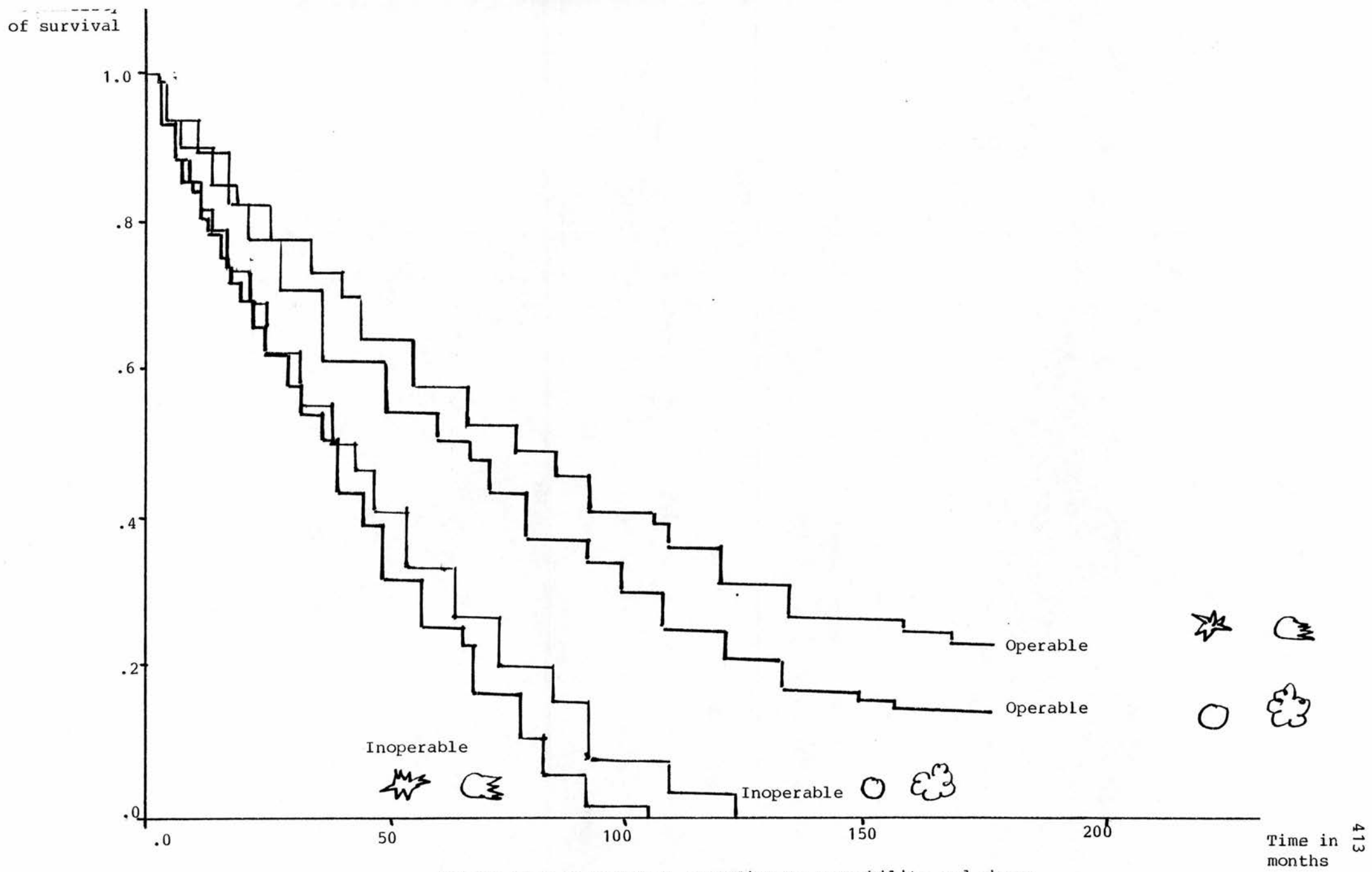


Figure (8.3.2) Survival according to operability and shape

Initially we will deal with the operable patients and the different categories considered independent of time. Later we will consider the time dependency of the various indices with their relevant interpretations.

For the operable group we note that the node negative patients tend to have a much better survival time than the node positive patients. The median survival time of the node negative patients is in fact 8 years and 9 months against 6 years and 4 months for the node positive patients. Node histology is one of the well-accepted prognostic indicators of survival time and we thus introduce it at first step of producing a relative risk function of the survival times for the operable strata.

The total number of patients is 122. There are 68 node negative patients and 54 node positive patients. At the end of the study there are 80 patients with recorded death times and 42 censored times. By use of the Cox's proportional hazard we estimate the corresponding relative risk functions, given by the model .

$$RR = \text{Exp}(\beta_{node} \cdot node)$$

node negative =0, node positive=1

$\beta_{node} = .5319$ S.E. = .2144 $X^2 = 6.30$ p = 0.012

Figure (8.3.3) represents the survival times for the two groups of the node negative and node positive patients in the operable strata.

One/

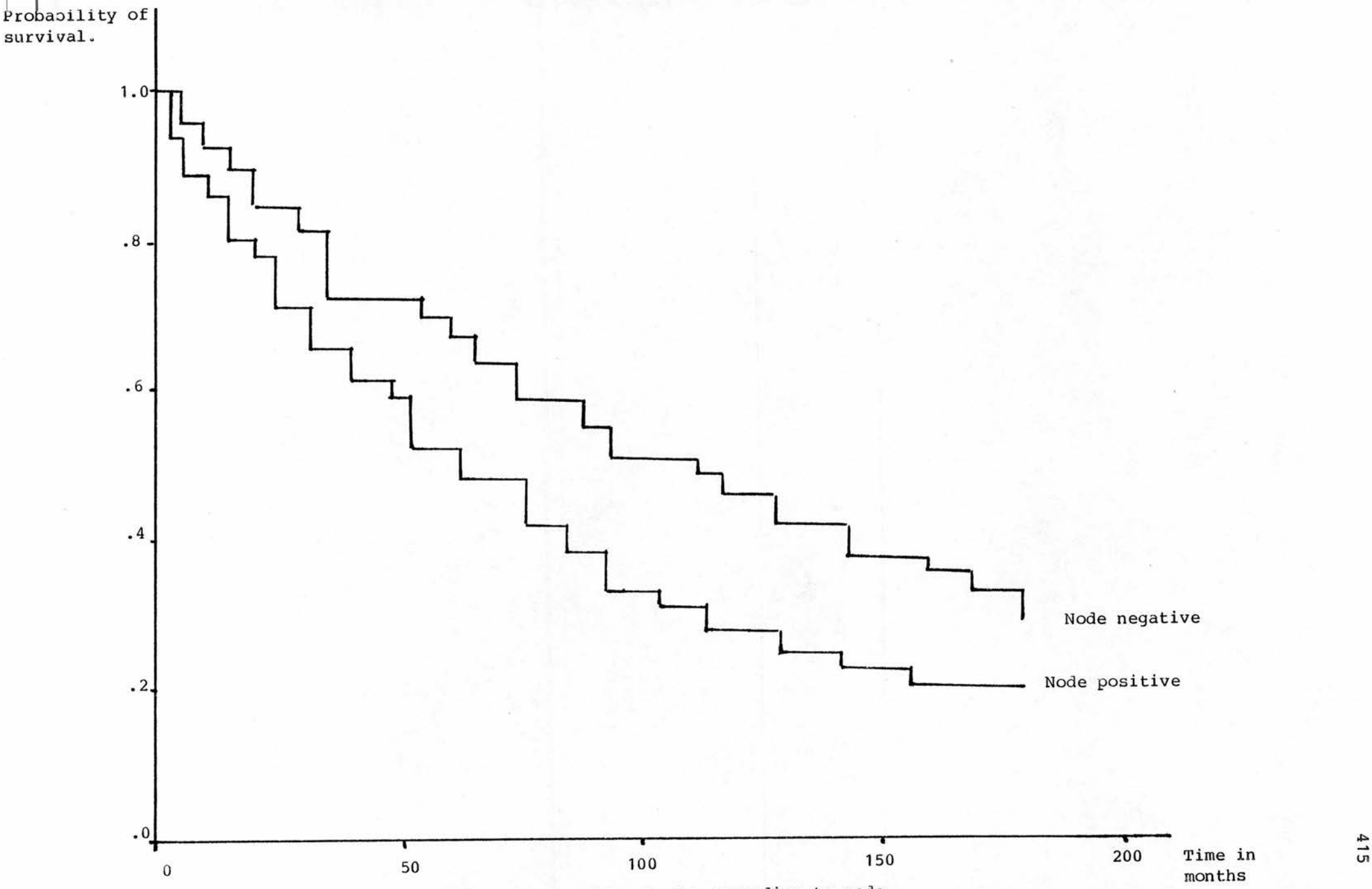


Figure (8.3.3) Survival for the operable strata according to node.

One further accepted indicator is that of size of the initial tumour. Size in the group of operable patients is playing a slightly less important role than the node histology. The prognostic importance of size however reaches a statistically significant level for the operable group. (Among the inoperable group however we do not detect a significant level and observe that the direction of the prognostic value is in the opposite direction to the operables). Model of the relative risk,

$$RR = \text{Exp}\,(\beta_{size}.size)$$

$\beta_{size} = .4581$ S.E. = .2171 $X^2 = 4.78$ $p = 0.029$

Figure (8.3.4) refers to survival rates for size when a split for over and under 3.5 cm. lesions has been made.

The significance of size and node status however remains when either node or size variability is introduced in the presence of the other.

$$RR = \text{Exp}\,(\beta_{size}.\ size + \beta_{node}.node)$$

$\beta_{node} = .4160$ S.E. = .2091 $X^2 = 4.31$ $p = 0.038$

$\beta_{size} = .3891$ S.E. = .1765 $X^2 = 4.26$ $p = 0.037$

In terms of the magnitude of the direction of node and size progression we introduce an interaction term for the relative risk model, giving

$$RR = \text{Exp}\,(\beta_{size}.\ size + \beta_{node}.\ node + \beta_{int.}.\ size.node)$$

$\beta_{node} = .4271$ S.E. = .2087 $X^2 = 4.20$ $p = .040$

$\beta_{size} = .3881$ S.E. = .1854 $X^2 = 4.09$ $p = .043$

$\beta_{int.} = .0092$ S.E. = .0426 $X^2 = .0731$ N.S.

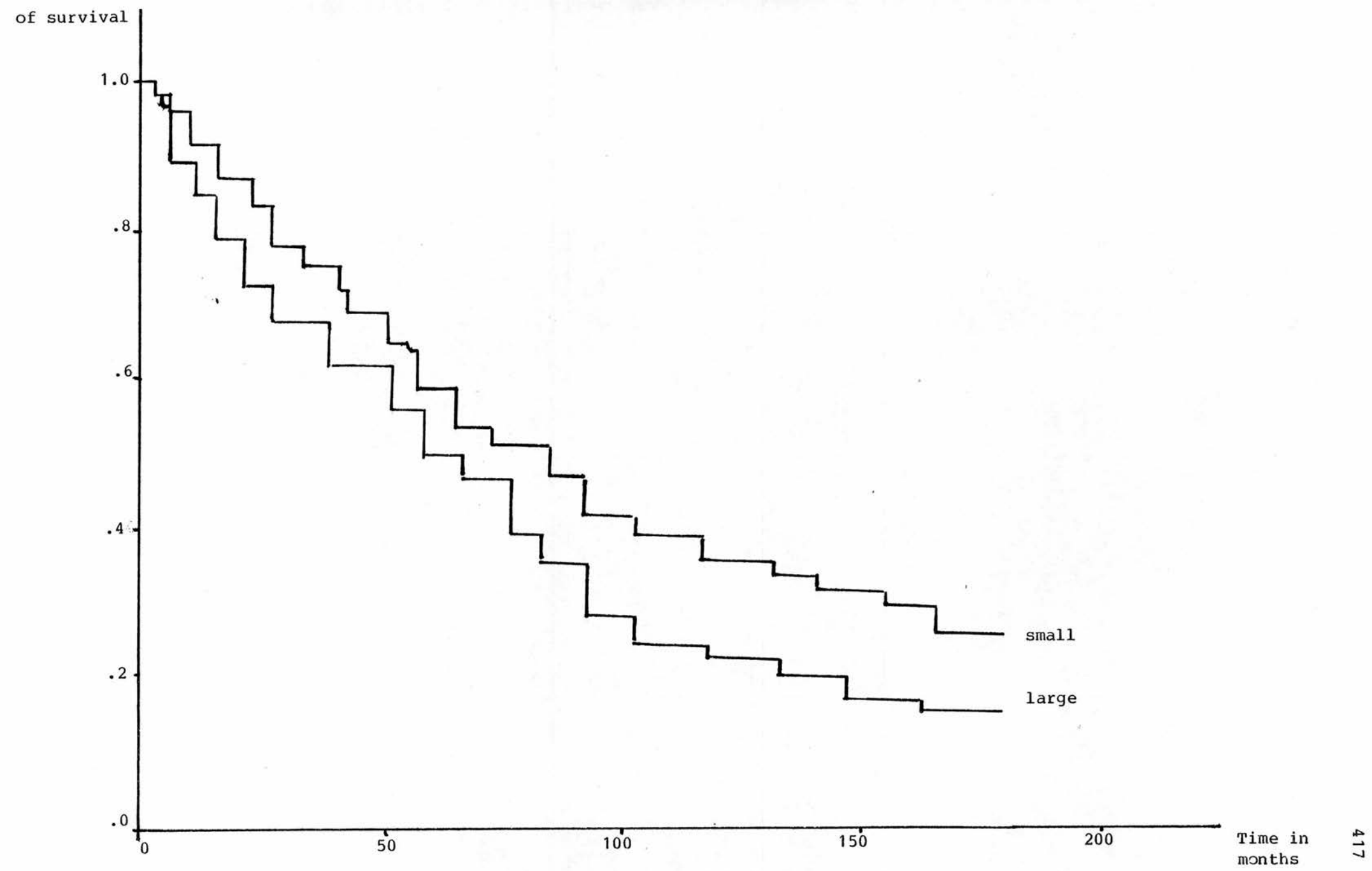


Figure (8.3.4) Probability of survival for the operable strata according to size.

The above model allows the effect of size to be different in node positive and node negative patients. In fact with the introduction of this interaction effect we do not observe any statistically significant improvement to the additive model of size and node.

By a single covariate relative risk model we study the effect of various indicators. Microcalcification is marginally not significant at 6.2% probability level and other indicators namely, tumour foci and contour type are of even less significance. The indicators grade, extent and change , as we may expect, show statistically significant levels in terms of time to death of patients. The most significant contributor is grade given by

$$RR = \text{Exp}\ (\beta_{grade} \cdot \text{grade})$$

β_{grade} = .5902 S.E. = .2091 X^2 = 8.02 p = 0.005

However grade is related to size and node status. Thus after the introduction of node and size in fact we reduce the grade effect.

$$RR = \text{Exp}\ (\beta_{node} \cdot \text{node} + \beta_{grade} \cdot \text{grade})$$

β_{node} = .4311 S.E. = .2109 X^2 = 4.16 p = .041

β_{grade} = .4021 S.E. = .2231 X^2 = 3.95 p = 0.047

$$RR = \text{Exp}\ (\beta_{size} \cdot \text{size} + \beta_{grade} \cdot \text{grade})$$

β_{size} = .4081 S.E. = .2051 X^2 = 4.05 p = 0.044

β_{grade} = .4019 S.E. = .2162 X^2 = 3.84 p =0.050

Next we introduce a model of size, node and grade which indicates very close estimators to the model of node and size, for the estimator/

estimator of node and size and an insignificant estimator for the grade.

$$RR = Exp\ (\beta_{node} \cdot node + \beta_{size} \cdot size + \beta_{grade} \cdot grade)$$

β_{node} = .4408 S.E. = .2010 X^2 = 4.81 p = .028

β_{size} = .4135 S.E. = .2081 X^2 = 3.92 p = .047

β_{grade} = .2850 S.E. = .2567 X^2 = 1.23 N.S.

Both extent and change produce statistically important relative risk patterns in terms of survival. If inserted singly, we obtain:

$RR = Exp\ (\beta_{ext.} \cdot ext)$ (Extent < ¼ inch deep) = 0

(Extent > ¼ inch deep) = 1

$\beta_{ext.}$ = .2425 S.E. = .1162 X^2 = 4.33 p = .037

$RR = Exp\ (\beta_{change} \cdot change)$ (complicated change not indicated) = 0

(complicated change indicated) = 1

β_{change} = .2637 S.E. = .1206 X^2 = 4.81 p = 0.028

Extent is slightly less significant than change. However the two categories represent almost overlapping subgroups of patients in terms of their own good (0 level) and bad (1 level) prognostic indicators.

First we introduce extent & change in the presence of size effect. Their significance level shows little change. Once again extent is less significant than change.

$$RR = Exp\ (\beta_{ext.} \cdot Ext. + \beta_{size} \cdot size)$$

$\beta_{ext/}$

$\beta_{ext.}$ = .2638 S.E. = .1259 X^2 = 4.39 p = 0.036

β_{size} = .3934 S.E. = .1785 X^2 = 4.63 p = 0.031

RR = Exp (β_{change}. change + β_{size}. size)

β_{change} = .2701 S.E. = .1288 X^2 = 4.45 p = .035

β_{size} = .3939 S.E. = .1783 X^2 = 4.61 p = 0.030

Now we will consider the possibility of an interaction effect between extent and size and further change and size. (node interactions later).

RR = Exp($\beta_{ext.}$. Ext + β_{size}. size +$\beta_{Int.}$. ext.size)

$\beta_{ext.}$ = .2641 S.E. = .1248 X^2 = 4.51 p = .034

β_{size} = .3942 S.E. = .1785 X^2 = 4.79 p = .028

$\beta_{Int.}$ = .0176 S.E. = .0712 X^2 = .059 N.S.

RR.. = Exp (β_{change}.change + β_{size}. size + $\beta_{Int.}$.change.size)

β_{change} = .2694 S.E. = .1285 X^2 = 4.39 p = .036

β_{size} = .3941 S.E. = .1780 X^2 = 4.64 p = .031

$\beta_{Int.}$ = .0211 S.E. = .0619 X^2 = .184 N.S.

Thus effect of extent and change is additive in presence of size and the interaction effect is not significant.

We can now consider extent and change in the presence of node histology. In the previous discussions node and size were clearly major contributurs in defining survival time. First we deal with the relative risk for node and change status.

RR = Exp(β_{node}. node + β_{change}. change)

β_{risk}/

β_{node} = .4381 S.E. = .2097 X^2 = 4.71 p = .029

β_{change} = .2517 S.E. = .1183 X^2 = 4.61 p = .032

The extent also presents a similar pattern as that of change

RR. = Exp (β_{node}.node + $\beta_{ext.}$Ext)

β_{node} = .4376 S.E. = .2081 X^2 = 4.57 p = .032

$\beta_{ext.}$ = .2480 S.E. = .1198 X^2 = 4.28 p = .038

The value of change with node is significant and it is interesting to study the effect of size in this respect. That is we assess the survival variability which is unexplained in terms of node and change, by the introduction of size. Before doing so we study the effect of an interaction between change, node, and extent, node.

RR = Exp (β_{node}. node + β_{change}. change + $\beta_{Int.}$. change.node)

β_{node} = .4392 S.E. = .1988 X^2 = 5.15 p = .023

β_{change} = .2524 S.E. = .1182 X^2 = 4.56 p = .032

$\beta_{Int.}$ = .0896 S.E. = .0489 X^2 = 3.78 p = .052

RR = Exp(β_{node}.node + $\beta_{ext.}$.ext. + $\beta_{Int.}$.ext.node)

β_{node} = .4366 S.E. = .2114 X^2 = 4.28 p = .038

$\beta_{ext.}$ = .2593 S.E. = .1274 X^2 = 4.17 p = .041

$\beta_{Int.}$ = .0782 S.E. = .0523 X^2 = 2.47 N.S.

There is a slight indication of an interaction effect with node and change indicating both complicated change and node positivity together add extra risks for survival. However this is not of great importance since it may be a spurious significance. The size variability however has/

has not been included in our model. If we do introduce the size effect by the relative risk function, none of the interactions remain significant.

$$RR = \text{Exp}(\beta_{node}.node + \beta_{change}.change + \beta_{size}.size)$$

β_{node} = .4279 S.E. = .2136 X^2 = 4.21 p = 0.040

β_{change} = .2512 S.E. = .1215 X^2 = 4.41 p = 0.036

β_{size} = .4181 S.E. = .1976 X^2 = 4.51 p = 0.033

In terms of extent no interaction effects are significant and if we introduce a model of size, node and extent,once again a similar pattern emerges as that of change.

$$RR = \text{Exp}(\beta_{node}.node + \beta_{ext.}.Ext. + \beta_{size}.size)$$

β_{node} = .4281 S.E. = .2237 X^2 = 4.40 p = .036

$\beta_{ext.}$ = .2484 S.E. = .1268 X^2 = 3.90 p = 0.048

β_{size} = .3927 S.E. = .2093 X^2 = 4.04 p = 0.045

The main reason for introducing the change concept has been that of considering an effect of tumour initial status by which some of the attributes in terms of external progress of tumour may be explained. This effect is clearly not sufficiently explained by node and size classification alone.

We will now introduce the concept of time dependency in this section for the various prognostic indicators. One reason for this conceptual change of model is to study effect of node or size over a time scale and test how each prognostic effect may eventually diminish/

diminish or increase over time.

First we consider a relative risk function based on node histology and a time dependent function of time, given by t* = ln(time)

$$RR = \text{Exp}(\beta_{node}.\text{node} + \beta_t . \text{node}.t^*)$$

β_{node} = .4284 S.E. = .2071 X^2 = 4.29 p = .038

β_t = -.0785 S.E. = .1211 X^2 = .420 N.S.

There is not a great improvement over the overall likelihood by the introduction of the time dependency factor into the model. Thus we may consider the effect of node histology to be static in terms of prognostic value. Once a patient is node positive the patient is at a higher risk and this risk for the individual patient in relative terms does not decrease or increase over the passage of time.

Size is the second factor we study with respect to time dependency.

$$RR = \text{Exp}\ (\beta_{size}. \text{size} + \beta_t . \text{size}.t^*)$$

β_{size} = .4201 S.E. = .1964 X^2 = 4.32 p = .036

β_t = -.1291 S.E. = .0623 X^2 = 4.19 p = .040

Size effect clearly diminishes with the passage of time. The larger tumour patients are at a higher risk in the early part of the diagnosis. However for the larger tumours that do not correspond with an early death, the prognostic significance of size will eventually diminish in terms of relative risk.

Both/

Both extent and change were two other indicators that produced some significant contributions in explaining the survival variability rates. Now we consider extent and change as two time dependent variables.

$$RR = \text{Exp}\ (\beta_{change} \cdot change + \beta_t \cdot t^* \cdot change)$$

β_{change}	= .2281	S.E.	= .1109	X^2	= 4.22	p	= .040
β_t	= -.0819	S.E.	= .0902	X^2	= .72	N.S.	

$$RR = \text{Exp}(\beta_{ext.} \cdot ext + \beta_t \cdot t^* \cdot ext)$$

β_{ext}	= .2319	S.E.	= .1132	X^2	= 4.24	p	= .039
β_t	= -.0792	S.E.	= .0876	X^2	= .96	N.S.	

Neither change nor extent contributions are affected significantly by the time dependent variability. In terms of interpretation we conclude that change and extent classify patients in the beginning and their effect is consistantly the same in terms of relative risk of death. Node histology therefore has a prognostic effect which may be interpreted in a similar way to that of extent and change.

One interesting question that may be asked is related to the effect of time on the magnitude of the size effect. Given that size is a time dependent factor, that is the sizes of tumour do not all conform to a single fixed relative risk function and some patients are slightly different type of survivors, is there a prognostic factor measureable and static at beginning of the study by which we may separate the time dependency of the size effect. Although the question and the evidence from the data may simply be represented by a few histograms, in terms of

of statistical significance there are a few models all of which according to this data can explain the variability within the data. Primarily we presented a model of relative risks based on node histology, size, complicated change and extent. In the comparison of models of complicated change and extent, there is little to choose between the two, in so far as our study is concerned. For practical reasons however the extent of tumour may be an easier variable for measurement. In the interpretation of the time dependency of the size effect we may conclude that there exists a subgroup of patients in whom largeness of size of the tumour is sign of bad prognosis. By the passage of time in the survival scale however the value of size as a prognostic variable diminishes.

C H A P T E R 9

FINAL SUMMARY

In this the final chapter, we will summarise the findings of the thesis. We will separate the findings into the statistical and medical, and allocate a section to each. In these sections we will present an overview of ideas which may be useful for future research.

9.1 Overview and conclusions of the statistical results.

Initially we identified various hazard shapes which have been reported in the literature. Such methods were useful for presenting in a descriptive manner the patterns of events in time scale for the different subgroups of patients. Further we discussed recent developments in the area of non-parametric methods and the way in which such methods are able to provide a flexible approach for classifying different non-parametric tests, which are often used in survival analysis (such as the logrank and Wilcoxon tests). In the area of parametric methods we considered various analytical methods and in comparing the various assumptions of the methods with empirical data with subgroups, we found the methods theoretically restrictive but practically in terms of conclusions consistent for our data set. Primarily we performed the analysis of the old Edinburgh trial data by the different parametric and non-parametric methods purely for the purpose of comparing the statistical methods. In terms of conclusions we did not find any inconsistencies between any of the parametric and/

and non-parametric methods. However as expected we were able to attribute the slight differences between the two non-parametric log-rank and Wilcoxon tests in the weighting attached to the events within the time scale of study. We defered the discussion of the difference between the various methods (in terms of significance levels) to later chapters where the concept of time dependency is more developed. Parallel with the above discussions we considered multivariate methods and how concepts such as multivariate prognostic factors and multivariate events may be employed in analysis. We considered efficiency and robustness of an approach to be two factors of extreme importance when dealing with the above forms of interrelationships between various events and prognostic factors. A method that we found suitable for this type of analysis, was the Cox's semi non-parametric proportional hazard model.

One important aspect of the Cox's method which can provide a robust framework for the analysis of such data is in the manner in which the actual survival times are transformed into ranks. Before proceeding with the development of models using Cox's method, we presented transition rates between the various states of the old Edinburgh trial, using an explanatory stochastic method which was referred to as the non-parametric semi Markov model. Although the approach was considered to be informative we found the Cox's method more suitable in the manner by which it could provide a check on the model assumptions, using the information on the intervening events. Initially we considered the expansion of the models with fixed relative risks into models that have covariates with an internal variability within the time scale. At times we found checks on the assumption/

assumption of consistency of a prognostic effect useful in a proper interpretation of the data. We refered to such models, models with time dependent prognostic effects. Alternatively an interpretation was possible by employing the concept of intervening event within the time scale.

We found that by the utilisation of the information on an important progression event such as metastatic recurrence, we were able to check on the consistancy of the relative treatment effect for the times prior to and post intervening event. In general such intervening events are random events and we used a family of transformations of the time of the intervening event, in order to check on the consistency of the goodness of fit tests. We found that in practice such a consistency was present and that the proportional hazard of non-parametric type was considered quite suitable, (for the covariate subgroups that we dealt with in the data).

We allocated a full chapter to the simulation methods for a clinical trial study. In this study we presented the small sample properties of the various statistical methods (in particular Cox's method) using simulated data. The method of simulation we adopted had a useful property of being able to generate increasing, decreasing and constant hazard rates with covariates. In fact all the generated samples belonged to the family of proportional hazards of the Weiball type. This property was found useful when we dealt with a simulation study of time dependencies.

An important property of survival studies as discussed before has been that of censoring of the survival times. In developing a simulation/

simulation method we discussed various approaches of generating censored survival times and adopted one which is suitable for a trial data and can give a constant proportion of censored cases. Our initial intention in performing the simulation has been an assessment of the small sample properties of the Cox's method. However later in study of time dependencies with the Cox's method we also considered Weiball and exponential parametric methods. Within these simulations we used a range of sample sizes, significance levels, levels of censoring and a range of treatment and covariate effects. In order to assess the power properties we constantly refered to the asymptotic normality and the likelihood ratio tests.

Initially we discussed the power properties of the simple test of hypothesis for the treatment effect at both treatment effect and covariate effect set to zero. This value is clearly an indicator of type one error. We obtained efficiency values close to the expected values according to the singificance levels. In repeating the simulations for a range of covariate effects we note that the efficiency of simple tests of hypothesis for treatment is not in any way influenced by the value of the covariate effect. Clearly as we expected the efficiency of the tests do deviate to some extent according to the values of sample size, censoring and significance level, however none of these factors affect the lack of influence of the covariate effect value. A point of some interest was that the decline in the power of the simple tests due to censoring which seems to be affected by the sample size to some extent, indicated a lower loss due to censorings for higher sample sizes.

Next/

Next we discussed power curves for the composite tests. These simulations had a change of emphasis in that they were presented for a more theoretical interest and we showed that the Cox's method has good, predictable small sample properties. At first we dealt with the type one error and showed that our results conform with the levels of the significance limit. This finding was true for both the asymptotic likelihood and the asmyptotic normality tests. Later we considered a range of treatment effect and covariate effects. We concluded as expected sample size, significance level and censoring levels do influence the power of tests. However the relative efficiency of the above factor is not influenced significantly by the treatment effect or covariate effect.

In comparing the asymptotic normality and asymptotic likelihoods we note that the asymptotic normality in general is more conservative, as the treatment effect and covariate effects deviate from zero. Up until this point we have summarised simulation results when the generated samples were based on an exponential distribution. Next we deal with the summary of results from the Weiball distribution. We reported the simulations for the samples of Weiball in which the proportionality of hazards had not been violated. We found very close resemblance between the efficiency of simulations on increasing, decreasing and constant hazards (all other factors e.g. sample size, censoring being equal). We attributed this close resemblance to the non-paramtric nature of the Cox's method.

At this stage we deal with results of the Weiball distributed samples/

samples in which the proportionality of the hazards was violated. That is, there existed a degree of time dependency for the covariate and treatment effects. Consistently we noted a reduction in the type one error less than the actual significance level. In fact we noted this power decreases with an increase from the proportionality of the hazards. For the range of treatment and covariate effects (negative and positive) we noted that for the non-proportionality with divergence from base line hazard there was a higher loss of power in comparison with the non-proportionality with a convergence to the base line.

In comparing the asymptotic normality to the likelihood ratio test we observed that the normality test is more conservative. We repeated the simulation for a range of sample sizes and censoring levels and the conclusions were consistently the same. One pattern which emerged was that due to non-proportionality, the power curves for the various composite tests did not have the symmetric pattern of the proportional hazard situation. However, we noticed that with an increase in the sample size there was a reduction in the lack of symmetry about the covariate effect axis.

In the final discussion of the simulation results we studied simple tests of hypothesis in the presence of one covariate effect. We generated non-proportional hazard data of the Weiball type and in order to analyse the data we used exponential mode, Weiball model, Cox's proportional hazard, Cox's stratified and Cox's time dependent models. At the value of treatment effect set to zero we consistently noted that power efficiency is close to the significance level for all models./

models. This pattern was consistently the same for non-proportional hazard samples. In order to study the influence of the covariate effect we increased the value of the covariate effect and again there was little change in the type one errors.

Next we considered non-zero values of the treatment effect for both non-proportional and proportional hazard samples. We concluded that for the analysis of the non-proportional hazard samples, both the stratified Cox's model and the time dependent Cox's model were suitable. This was true for a range of covariate effect values and we noted that the magnitude of the covariate effect did not influence power of the tests. The unsuitable models were the Cox's model with fixed relative risk, the Weiball model and the exponential respectively, with the exponential being much worse than the other two. For these three unsuitable models we noted that the magnitude of the covariate effect does influence the power of the tests. In summary we concluded that specification of the correct model is of some importance, when dealing with proportional hazard models. The Weiball and Cox's models (fixed relative risk) are the most suitable for analysing proportional hazard data and the time dependent Cox's model and stratified Cox's model are less suitable.

In terms of magnitude of the efficiencies we noted that at high sample sizes of 100 there was little to choose in general between the models (except exponential) while at low sample sizes of 25 some of the problems with the specification of the wrong model were more apparent.

9.2/

9.2 Overview and conclusions of the medical results.

On dealing with the applications of the various statistical methods we presented two data sets. Both of which dealt with the primary breast cancer. The first data set was refered to as the South East of Scotland trial data. The major objective of this trial was an assessment of the survival rates for a group of patients treated by radical surgery versus the group treated by simple surgery and radiotherapy. Before commencing the analysis of this data we discussed the important design aspects of this trial such as patient eligibility rules, stratification and data administration. This trial with regard to the magnitude of the data which was collected and the type of events that were expected to take place within the survival time of each patient, is suitable for an exploratory analysis of the various interrelationships such as the multivariate events and multiple prognostic indicators, as discussed earlier.

As we indicated by the study of the cross-tabulations there was in most respects a very uniform balance between the two treatment groups and the various prognostic indicators. Initially we performed an analysis based on the conventiaonal methods. This analysis indicated that the radical surgery patients may have an overall higher survival rate compared with the group treated by simple mastectomy and radiotherapy. We further indicated that the significance levels of the treatment differences as expected was not consistently the sane for the different prognostic subgroups.

At the first stages of the exploratory analysis of the data,

we/

we are interested in the study of the duration from randomisation to any single event of importance, namely any of death, metastatic recurrence, or local recurrence. In this approach we used various stepwise regression methods with the Cox's proportional hazard model. A point to note was that in the approach we considered step-up and step-down procedures, together with a different procedure by which as a priori rule, we forced the treatment effect into the model at first step regardless of its relative significance to the other prognostic factors. The above methods consistently yielded the same model indicating a better survival for the radical surgery group, and with the significant prognostic contributors to the model being, menopausal status, size of the tumour and the node status.

In order to make sure the findings of the models were not dependent upon the model assumptions we used stratified analysis at each step of introducing a new term into the model. Once again we noted that the direction of the effects was consistently the same. In order to assess the multiplicative effect of the various indicators we performed tests of the interaction effects using the Cox's method and we did not find any evidence of interactions for the survival times.

In the next stage of the analysis we considered the time period from randomisation to the development of the metastatic disease. We once again used the above stepwise procedures for model reduction with the same set of prognostic effects. We noted consistently the final model was the model involving, option, size, node and menopausal status. However, we also noted slight deviations in the order of the entry of the various terms. This pattern is not testable at this stage however,/

however, this difference in the order of importance of the prognostic effects for these two response variables is of some importance for the later discussions when we consider time dependency in the time scale of study. The analysis based on the randomisation to local recurrence also presented a consistent model indicating the same set of covariate main effects.

In the summary of the results so far we have confined the conclusions to those of analysis of randomisation time to an important final event. From this point onwards we will summarise the results of the South East of Scotland trial with models of multiple risks in which patients move from one state to another.

In general we attributed the developments within the time scale to be due to intertwined processes. Such processes were combinations of epochal events such as metastatic recurrence, local recurrence or death. Alternatively we may have been interested in the assessment of cumulative risks over time for a single covariate.

Initially we considered an analysis based on the semi-markov models . This approach gave an explanatory stochastic description of the movement of patients from one state to the other. In terms of presentation of the results we derived general expressions for the survival rates in order to obtain close approximations to the transition rates. More important however, are the results by which we represented the above survival rates in terms of exploratory models of the proportional hazard type. Using the proportional hazard models with time dependent covariates, we were able to ask questions such as; given/

given that an event has taken place (such as the progression of disease) how is the relative risk for treatment groups and prognostic effects performing within the time scale. In the study of the empirical hazard rates we consistently refered to hazard rates in which the proportionality of the hazards may have been violated. We were then in the position of testing any possible departures from the assumptions of the model.

At first we considered a test of the time dependency of the treatment effect, by which we concluded that there was no evidence of the proportional hazard assumptions being violated with respect to the treatment effects. We then studied the survival rates according to the time dependency of the size effect by which we concluded there was not sufficient evidence to reject the constant relative risk assumption.

Next we studied the response time of metastatic recurrence or local recurrence to death, given that a recurrence had already taken place. First we stratified the data according to the period of randomisation to the appearance of local or metastatic disease, in order to compare radical versus simple surgery in a model by which time to the appearance of first recurrence was controlled. The pattern emerging was that for patients recurrent after the first year, their having had radical surgery, was less likely to benefit the patient. However for those recurring early there was benefit in terms of survival by a radical surgery.

Finally for this data we considered the analysis of the

data/

data according to the secondary events that had occurred within the full patient survival time. Such secondary events were taken to be based on local or metastatic recurrence of the disease. Initially we dealt with the relative risk function according to the treatment effect and tested the goodness of fit in order to detect departures from the proportional hazard assumptions of type by which the relative risk after the secondary event may be acting differently for the two treatment groups. Once again we noted that the proportional hazard model is quite appropriate and that all conclusions were consistently in line. Although the non-proportionality did not reach significance we noted, however, that the treatment effect was more substantial prior to the development of metastatic disease.

In the previous chapter of the thesis we considered survival rates of a group of cancer patients in order to assess the importance of certain prognostic indicators. Before analysis of the data we made an initial distinction between indicators that are regularly assessed in the staging of patients and some other indicators that have not been considered a great deal in the past. In general the variables we were interested in were; tumour contour type, operability, size, node, extent, grade, change, tumour foci and microcalcification. In the analysis of the data we concentrated on the survival time of the patients as the only response variable. However we did not only deal with a study of static prognostic factors but also studied how changes may occur in the value of the initial prognosis.

At first we noted that the various indicators for the

operables/

operables and the inoperables did not necessarily indicate prognosis in the same direction. The prognostic indicators for the inoperables were not of a significant difference and we consequently confined the analysis to the operable patients. For the operable patients we noted that node negative patients had a better survival rate than the node positives. Further, as expected, size of the tumour conformed with what was expected and patients with smaller tumours showed better survival rates. In terms of the relative value of size and node we noticed that the significance of either factor remained in the presence of the other factor. In terms of direction in the effects of size and node there is no evidence of an interaction. Two further prognostic factors that are found to have an important influence on the final survival pattern of the patients were the extent of the tumour and the presence of complicated change in the tumour. We observed that extent is slightly less significant than change. However, the two categories represented almost overlapping subgroups in terms of good and bad prognosis. The good prognosis being tumour with less than ¼ inch depth and tumour with no evidence of complicated change. We performed test of interaction for the different indicators and found none significant. We further performed tests of time dependency of the indicators and found time dependency of size indicated that size is of more importance during early periods of survival.

Throughout the course of this thesis we have presented the models with an emphasis on interrelationships. A study of such factors will clearly imply a check on the generalisations and the assumptions of the models, and may introduce a diversity of interpretations. According to the data sets that are here examined it has/

has been possible to show a consistency of results in the final analysis. It is however important to consider the impact of such methods in situations where data may be the resultant outcome of constantly evolving treatments and that the analysis is performed in a situation of widely accessible distributed computing procedures.

APPENDIX (A)

Section A.1

Cross tabulation tables for the prognostic factors referred to in section 6.4.

Menopausal Status

		Prem	Meno	Post M.	Total
Node	N0	107	21	247	375
	N1	56	17	112	185
	Total	163	38	359	560

Menopausal Status

T stage		Prem	Meno	Post M.	Total
	T1	22	2	32	56
	T2	114	26	257	397
	T3	27	10	70	107
	Total	163	38	359	560

Side

T stage		Right	Left	Total
	T1	22	34	56
	T2	205	192	397
	T3	57	50	107
	Total	284	276	560

Site

T stage		Medial only	Lateral only	Central	Both halves	Whole Breast	Total
	T1	20	30	6	0	0	56
	T2/						

T Stage	Site: Medial only	Lateral only	Central	Both Halves	Whole Breast	Total
T1	20	30	6	0	0	56
T2	133	203	48	13	0	397
T3	30	53	13	9	2	107
Total	183	286	67	22	2	560

S Stage	Site: Medial only	Lateral only	Central	Both Halves	Whole Breast	Total
S1	108	161	31	7	0	307
S2	44	70	21	6	0	141
S3	31	55	15	9	2	112
Total	183	286	67	22	2	560

Node Status	Site: Medial only	Lateral only	Central	Both Halves	Whole Breast	Total
NØ	133	191	38	12	1	375
N1	50	95	29	10	1	185
Total	183	286	67	22	2	560

Section A.2

	Skin involvement					Pectoral Muscle involvement				Disease Status				Total
	T0	T1	T2	T3	NI	T0	T1	T3	NI	None	L	M	L+M	
Meno-pausal status														
Prem	1	5	16	7	134	8	16	8	131	27	34	3	99	163
Meno	1	1	3	7	26	5	9	2	22	11	7	1	19	38
Post M	0	6	51	22	280	1	38	33	287	76	100	12	171	359
Side														
Right	1	7	41	15	220	9	33	22	220	56	73	9	146	284
Left	1	5	29	21	220	5	30	21	220	58	68	7	143	276
Site														
Med.only	1	6	17	10	149	2	16	15	150	32	48	4	99	183
Lat.only	0	5	37	16	228	7	28	23	228	58	60	10	158	286
Central	1	0	9	7	50	2	11	4	50	18	22	2	25	67
Both	0	1	5	3	13	3	7	0	12	6	9	0	7	22
Whole Breast	0	0	2	0	0	0	1	1	0	0	2	0	0	
T Stage														
T1	1	2	3	0	50	5	1	2	48	3	14	1	38	56
T2	0	1	9	0	381	9	6	1	381	85	90	11	211	397
T3	1	9	58	36	3	0	56	40	11	26	37	4	40	107
Node														

Section A.2 (cont'd)

	Skin involvement					Pectoral muscle involvement				Disease status				Total
	T0	T1	T2	T3	NI	T0	T1	T3	NI	None	L	M	L+M	
Node														
No	2	7	38	27	301	7	42	25	301	63	86	7	219	375
N1	0	5	32	9	139	7	21	18	139	51	55	9	70	185
S Stage														
S1	0	1	5	0	301	7	3	0	297	51	63	5	188	307
S2	1	1	3	0	136	5	4	0	132	36	40	7	58	141
S3	1	10	62	36	3	2	56	43	11	27	38	4	43	112
Total	2	12	70	36	440	14	63	43	440	114	141	16	289	560

A P P E N D I X B

The Fortran maximum likelihood estimation program with use of the Newton-Raphson procedure. The version which is listed performs the necessary calculations for the estimators of proportional hazards model with fixed exponential relative risks.

```
      real  blank,r,rb,idno,beta
      common tdeath(1000),istdvr(40),iselvr(40),beta(40),
     3     rb(40),u2cinv(40),dd(40),z(40),zk(40),
     2          r,idno,ns,ireprt,irej,alnl1,kch,ndatin,distnc,
     4       model,itime,mxhalf,status,stp,ipr,npbase,liktyp,
     1     nvar,maxitr,alnl0,time,ncase,ndsv3,break,curtim,timint,
     5     ivtime,ivstat,iventr,ividno,is,nvtot,istep,niter,ivznp
      dimension bout(40,50)
      data bout/2000*0./
C
CCCC define all file output and input
C
      open(5,file='phr.r5')
      open(6,file='phr.w6')
      open(8,file='phr.r8')
      open(9,file='phr.w9')
      open(11,file='phr.w11')
      rewind 5
      rewind 6
      rewind 8
      rewind 9
      rewind 11
C
CC read all the data and specification for analysis
C
      read (5,903) kch,liktyp,ndatin,nvtot,ivstat,ivtime,iventr,
     1     ividno,is,itime,mxhalf,ipr,npbase,ivznp
      if (mxhalf.eq.0) mxhalf = 5
      read (5,908) stp,timint,distnc,chient
      if (stp.eq.0) stp = .001
      if (timint.eq.0) timint = 1
      if (distnc.eq.0) distnc = 1
      if (chient.eq.0) chient = 1.32
      if (ipr.eq.1) read (5,908) zk
      ncin = 0
      ncsv = 0
      ndin = 0
      ndsv1 = 0
      ndsv2 = 0
      ndsv3 = 0
      sr = 0
  10  read (8,903,end=50) (dd(n),n=1,ndatin)

CC transform any variables if needed

      call transf
      status = dd(ivstat)
      if (status.eq.1) go to 20
      ncin = ncin + 1
```

```
      go to 30
   20 ndin = ndin + 1
   30 if (irej.eq.1.and.(kch.eq.1.or.status.eq.0.or.liktyp.eq.1))
     1  go to 10
      sr = sr + 1
      if (status.eq.1) go to 40
      ncsv = ncsv + 1
      go to 10
   40 if(irej.eq.0) ndsv1 = ndsv1 + 1
      ndsv3 = ndsv3 + 1
      if (ivtime.ne.0) tdeath(ndsv1) = dd(ivtime)
      if (sr.gt.1.and.irej.eq.0) ndsv2 = ndsv2 + 1
      if (liktyp.eq.0.and.kch.eq.0) sr = 0
      go to 10
   50 nin = ncin + ndin
      nsv = ncsv + ndsv2
      ndead = ndsv2
      ncase = nsv
  100 read (5,903,end=800) istep,nvar,ns,ibref,maxitr,ireprt,model,
     1     iswise
      if (iswise.eq.0) go to 110
      nvarsv = nvar
      maxisv = maxitr
      maxitr = -1
  110 read (5,904) iselvr
      do 120 n=1,40
      istdvr(n) = 0
      if (n.le.ns) istdvr(n) = 1
  120 continue
C
CCCC insert MLE or use previous steps
C
      if (ibref.eq.0) go to 150
      do 130 n=1,40
  130 beta(n) = bout(n,ibref)
      go to 300
  150 read (5,908) beta
C
CCC obtain estimate for a constant rate null
C
      if(liktyp.eq.0.or.model.eq.6) go to 300
      do 200 n=1,nvar
      if(beta(n).ne.0) go to 300
  200 continue
      if(model.gt.1) go to 250
      beta(1) = alog(float(ndead)/(ncase - ndead))
      go to 300
  250 beta(1) = - float(ncase - 2 * ndead) / (ncase - ndead)
  300 continue
```

```
      call phr
C
CCC use the MLE either for this step or
CCC  use the MLE for forward selection based on largest
CCC  Chi Squared (or use Forced variables)
C
      if (iswise.eq.0) go to 700
      nvar = ns + 1
      maxitr = maxisv
      ireprt = 0
      chimax = 0
      do 500 n=nvar,nvarsv
      if (u2cinv(n).lt.chimax) go to 500
      nmax = n
      chimax = u2cinv(n)
 500  continue
      if (chimax.ge.chient) go to 600
      go to 700
 600  saviv = iselvr (nvar)
      iselvr(nvar) = iselvr(nmax)
      iselvr(nmax) = saviv
      call phr
      if (nvar.eq.nvarsv) go to 700
      ns = nvar
      istdvr(ns) = 1
      nvar = nvarsv
      maxitr = -1
      go to 300
 700  do 750 n=1,40
 750  bout(n,istep) = beta(n)
      go to 100
 800  write (6,958)
      stop
 903  format (16i5)
 904  format (40i2)
 908  format (8f10.0)
 958  format ('normal ending')
      end
      subroutine phr
C
CCC perform the MLE calculations
C
      real r,rb,rbb,sr,srb,srbb,idno,betnam,unam,beta,r1
      real c,cinv,csave,d1,d2,estims,blank,vmarg,const
      real betasv,u,sarat,sratd,shift
      common tdeath(1000),istdvr(40),iselvr(40),beta(40),
     3     rb(40),u2cinv(40),dd(40),z(40),zk(40),
     2          r,idno,ns,ireprt,irej,alnl1,kch,ndatin,distnc,
     4       model,itime,mxhalf,status,stp,ipr,npbase,liktyp,
```

```
     1      nvar,maxitr,alnl0,time,ncase,ndead,break,curtim,timint,
     5      ivtime,ivstat,iventr,ividno,is,nvtot,istep,niter,ivznp
      dimension u(40),c(820),cinv(820),csave(820),betasv(40),corr(40),
     1      srb(40),srbb(820),sarat(40),sratd(40)
      dimension  msv(1000),nrsksv(1000),srsv(1000)
      dimension ar(40)
      alnlst = -10**10
      nvsq2 = nvar*(nvar+1)/2
      niter = 0
      nsurv = ncase - ndead
      alnl0 = 0
      if (liktyp.eq.1) alnl0 = ndead*alog(float(ndead))
     1                       + nsurv*alog(float(nsurv))
     2                       - ncase*alog(float(ncase))
C
CCCC after initialising clear arrays and iterate
C
100   m = 0
      i = 0
      niter = niter + 1
      alnl = 0
      sr = 0
      do 110 n=1,nvar
      sratd(n) = 0
      sarat(n) = 0
 110  srb(n) = 0
      do 120 nn=1,nvsq2
      c(nn) = 0
 120  srbb(nn) = 0
      break = tdeath(1) - timint/2
      if (itime.eq.0.and.iventr.eq.0) break = 0
      curtim = tdeath(1)
      nphr = 0
      ier = 0
      tlast = 999999.
      if (ireprt.lt.0) rewind 11
C
CCCC get data on each case together with ties if needed
C
 200  rewind 8
 210  call getr
      nphr = nphr + 1
      if (r.gt.0) go to 250
      if (niter.eq.1) go to 895
      alnl = 0
      alr1 = 0
      alr0 = 0
      ier = 2
      go to 510
```

```
 250  if (liktyp.eq.0) go to 300
C
CCCC Ignore the next loop if using partial likelihoods
C
     r1 = r + 1
      alnl = alnl - alog(r1)
      if (status.eq.1) alnl = alnl + alog(r)
      nn = 0
      do 260 n=1,nvar
      if (status.eq.1) sratd(n) = sratd(n) + rb(n)/r
      sarat(n) = sarat(n) + rb(n)/r1
      do 260 n1=1,n
      nn = nn + 1
 260  c(nn) = c(nn) + rb(n)*rb(n1)/(r*r1*r1)
      if (nphr.lt.ncase) go to 210
      go to 500
C
CCCC    Begin partial liklihood estimation for survival times
C
 300  sr = sr + r
      nn = 0
      do 310 n=1,nvar
      srb(n) = srb(n) + rb(n)
      do 310 n1=1,n
      nn = nn + 1
 310  srbb(nn) = srbb(nn) + rb(n)*rb(n1)/r
      if (status.eq.0) go to 210
      if (kch.ne.0.and.time.ne.tdeath(i+1)) go to 210
C
CCCC   ignore censored times and adjust deaths
C
      i = i+1
      m = m+1
      if (ireprt.ge.i.and.niter.eq.1)
     1     write (9,999) i,nphr,irej,time,r,(rb(n),n=1,nvar)
C
CCCC   if death has missing data , skip calculation
C
     if (irej.eq.1) go to 455
      alnl = alnl + alog(r)
      nn = 0
      do 320 n=1,nvar
 320  sratd(n) = sratd(n) + rb(n) / r
C
CCC    loop to
CCC    perform calculations for first and second derivative
C
     if (kch.eq.0.or.i.ge.ndead) go to 350
      if (tdeath(i).eq.tdeath(i+1)) go to 210
```

```
 350  nn = 0
      if (ireprt.ge.0) go to 361
      dt = tlast - tdeath(i)
      tlast = tdeath(i)
      alam0 = m/sr
      vlam0 = m*(nphr-m)/(nphr*sr*sr)
      do 360 n=1,nvar
 360  ar(n) = srb(n)/sr
      write (11,971) tdeath(i),dt,m,nphr,alam0,vlam0,(ar(n),n=1,nvar)
 971  format (2f10.3,2i5,41e15.6)
 361  continue
      msv(i) = m
      nrsksv(i) = nphr
      srsv(i) = sr
      alnl = alnl - m*alog(sr)
      if (niter.eq.1)
     1  alnl0 = alnl0 - m*alog(float(nphr))
      do 400 n=1,nvar
      sarat(n) = sarat(n) + m*srb(n)/sr
      do 400 n1=1,n
      nn = nn + 1
 400  c(nn) = c(nn) + m * (srbb(nn)/sr -  srb(n)*srb(n1)/(sr*sr))
      if (ireprt.ge.i.and.niter.eq.1)
     1   write (9,998) sr,(srb(n),n=1,nvar)
      if (ireprt.ge.i.and.niter.eq.1) write (9,998)
 455  m = 0
C
CCCC   Now for all cases a contribution to liklihood has been made
CCCC    if using ordinary partial likelihood the next few lines
CCCC    are not needed
C
      if (i.ge.ndead) go to 500
      if (kch.eq.0) go to 465
      curtim = tdeath(i+1)
      if (tdeath(i+1).ge.break) go to 210
 460  break = break - timint
      if (tdeath(i+1).lt.break) go to 460
 465  nphr = 0
      sr = 0
      do 470 n=1,nvar
 470  srb(n) = 0
      do 480 nn=1,nvsq2
 480  srbb(nn) = 0
      if (is.eq.1.and.(istep.ne.1.or.niter.ne.1)) go to 210
      if (kch.ne.0) go to 200
      go to 210
C
CCCC  If likelihood is low enough take an estimate for this iteration
C
```

```
 500  if (alnl-alnlst.gt.-stp) go to 540
      ier = 1
      alr0 = 2 * (alnl - alnl0)
      alr1 = 2 * (alnl - alnl1)
 510  do 520 n=1,nvar
 520  beta(n) = (beta(n) + betasv(n)) / 2.
      shift = shift/2
      rr = 0
      rw = 0
      nshift = nshift + 1
      niter = niter - 1
      write (6,906) niter,nshift,ier,shift,alnl,alr0,alr1,rr,rw
      write (6,910) (beta(n),n=1,nvar)
      if (niter.ge.maxitr.or.nshift.ge.mxhalf) go to 800
      go to 100
 540  nshift = 0
      aldiff = alnl - alnlst
      alnlst = alnl
      do 570 n=1,nvar
 570  u(n) = sratd(n) - sarat(n)
C
CCC  Output results of firs iteration,do a test for each variable,
C
      if (niter.gt.1) go to 650
      alnl1 = alnl
      write (6,902)
      do 630 n=1,nvar
      nn = 0
      do 620 n1=1,nvar
      do 620 n2=1,n1
      nn = nn + 1
      csave(nn) = c(nn)
      if (n1.ne.n.and.istdvr(n1).eq.0) go to 610
      if (n2.ne.n.and.istdvr(n2).eq.0) go to 610
      go to 620
 610  csave(nn) = 0.
      if (n1.eq.n2) csave(nn) = 1.
 620  continue
CC      invert matrix
      call linv1p (csave,nvar,cinv,1,d1,d2,ier)
      nn = n*(n+1)/2
      u2cinv(n) = u(n) * u(n) * cinv(nn)
CCC     calculate the chi value
      call mdch (u2cinv(n),1.,psig,ier)
      psig = 1. - psig
      if (iselvr(n).eq.0) go to 624
      if (ipr.eq.1) go to 625
      write (6,904) n,iselvr(n),istdvr(n),
     1   beta(n),sratd(n),sarat(n),u(n),c(nn),cinv(nn),u2cinv(n),psig
```

```
      go to 630
 624     write (6,904) n,iselvr(n),istdvr(n),
     1   beta(n),sratd(n),sarat(n),u(n),c(nn),cinv(nn),u2cinv(n),psig
      go to 630
 625  write (6,929) n,iselvr(n),zk(n),istdvr(n),
     1   beta(n),sratd(n),sarat(n),u(n),c(nn),cinv(nn),u2cinv(n),psig
 630  continue
C
CCCC get correlations and information matrix ready
C
      if (ipr.eq.1) go to 646
      write (6,903)
      if (iselvr(1).eq.0) go to 634
      go to 636
 634  if (nvar.ge.2) go to 635
      go to 636
 635  continue
 636  n11 = 0
      n12 = 0
      do 645 n1=1,nvar
      n11 = n11 + n1
      n22 = 0
      do 640 n2=1,n1
      n22 = n22 + n2
      n12 = n12 + 1
 640  corr(n2) = c(n12) / sqrt(c(n11)*c(n22))
      if (iselvr(n1).eq.0) go to 644
      write (6,901) (corr(n2),n2=1,n1)
       go to 645
 644  write (6,901) (corr(n2),n2=1,n1)
 645  if (nvar.gt.8) write (6,901)
 646  write (6,901)
      write (6,918) alnl0
C
CCC  Invert the information matrix and check within range
C
 650  if (maxitr.lt.0) go to 890
      do 670 nn=1,nvsq2
 670  csave(nn) = c(nn)
      if (ipr.eq.0) go to 690
      nn = 0
      do 685 n=1,nvar
      do 685 n1=1,n
      nn = nn + 1
      if (u(n).ge.0.and.u(n1).ge.0) go to 685
      if (u(n).lt.0.and.beta(n).le.0) go to 675
      if (u(n1).lt.0.and.beta(n1).le.0) go to 675
      go to 685
 675  if (n.eq.n1) go to 680
```

```
      c(nn) = 0
      go to 685
 680  c(nn) = 1
 685  continue
      ns1 = ns
      do 686 n=1,nvar
      if (u(n).ge.0.or.beta(n).gt.0) go to 686
      u(n) = 0
      ns1 = ns1 + 1
 686  continue
 690  do 6905 nn=1,nvsq2
 6905 csave(nn) = c(nn)
CC      invert matrix
      call linvlp (c,nvar,cinv,1,d1,d2,ier)
      if (ier.eq.0) go to 691
      ier = 3
      rw = 0
      rr = 0
      alr0 = 2 * (alnl - alnl0)
      alr1 = 2 * (alnl - alnl1)
      niter = maxitr
      go to 770
C
CCC ignore the following for survival analysis with single loops
C
691   rr = 0
      rw = 0
      shift = 1
      nnsv = 1
      do 692 n=1,nvar
 692  betasv(n) = beta(n)
      do 700 n=1,nvar
      nnsv = nnsv + n - 1
      nn = nnsv - 1
      do 695 n1=1,nvar
      nn = nn + 1
      if (n1.gt.n) nn = nn + n1 - 2
      rr = rr + u(n)*u(n1)*cinv(nn)
      rw = rw + betasv(n)*betasv(n1)*csave(nn)
      beta(n) = beta(n) + u(n1)*cinv(nn)*distnc
 695  continue
      if (ipr.eq.0.or.beta(n).ge.0) go to 700
      if (betasv(n).gt.0)
     1     shift = amin1(shift,betasv(n)/(betasv(n)-beta(n)))
      if (betasv(n).eq.0) beta(n) = 0
 700  continue
      if (niter.eq.1) rr0 = rr
      if (shift.eq.1) go to 750
      do 720 n=1,nvar
```

```
      beta(n) = beta(n)*shift + betasv(n)*(1-shift)
  720 if (beta(n).lt.betasv(n)*stp) beta(n) = 0
  750 shift = shift*distnc
      if (niter.gt.1) go to 770
      write (6,905)
      if (ipr.eq.1) go to 760
      if (iselvr(1).eq.0) go to 758
      go to 765
  758 if (nvar.ge.2) go to 759
      go to 765
  759 continue
      go to 765
  760 write (6,913) (zk(n),n=1,nvar)
  765 write (6,901)
  770 alr0 = 2 * (alnl - alnl0)
      alr1 = 2 * (alnl - alnl1)
      write (6,906) niter,nshift,ier,shift,alnl,alr0,alr1,rr,rw
      write (6,910) (beta(n),n=1,nvar)
      write (6,910) (u(n),n=1,nvar)
      write (6,901)
C
CCCC  stop all loops
C
     if (maxitr.eq.0) go to 870
      if (niter.ge.maxitr) go to 800
      if (rr.ge.stp) go to 100
      if (aldiff.gt.(stp * 10.).and.shift.eq.1) go to 100
C
CCC    final estimatore get SE and MLE for relative risk
C
  800 if (ipr.eq.0) go to 810
      write (6,927)
      r = 0
      if (npbase.ne.0) r = - beta(npbase)
      n = 0
      var = 0
      if (npbase.ne.0) var = cinv(npbase*(npbase+1)/2)
      expr = exp(r)
      sebeta = sqrt(var)
      expse = exp(-sebeta)
      chisq = 0
      if (var.ne.0) chisq = r*r /var
      z0 = 0
      write (6,928) n,z0,r,sebeta,expr,expse,chisq
      var = 0
      do 808 n=1,nvar
      r = r + beta(n)
      if (beta(n).eq.0) go to 808
      if (n.eq.npbase) go to 807
```

```
      var = var + cinv(n*(n+1)/2)
      nn1 = n*(n-1)/2 + 1
      nn2 = n*(n+1)/2 - 1
      n1 = 0
      do 805 nn=nn1,nn2
      n1 = n1 + 1
 805  if (beta(n1).ne.0.and.npbase.ne.n1)    var = var + 2*cinv(nn)
 807  se = sqrt(var)
      chisq = 0
      if (se.ne.0) chisq = (r/se)**2
      expr = exp(r)
      expse= exp(se)
      sebeta = sqrt(cinv(n*(n+1)/2))
      write (6,928) n,zk(n),beta(n),sebeta,expr,expse,chisq
 808  continue
      go to 862
C
CCC   in order to use the following change the above liner
CCC    the following will use estimates for removal of a variable
C
 810  write (6,907)
      if (iselvr(1).eq.0) go to 814
      go to 816
 814  if (nvar.ge.2) go to 815
      go to 816
 815  continue
 816  write (6,910)    (beta(n),n=1,nvar)
      write (6,915)
      do 820 n=1,nvar
 820  csave(n) = 1/csave(n*(n+1)/2)
      write (6,910) (csave(n),n=1,nvar)
      write (6,916)
      do 850 n=1,nvar
      nn1 = n*(n-1)/2 + 1
      nn2 = n*(n+1)/2
      if (iselvr(n).eq.0) go to 825
      write (6,910) (cinv(nn),nn=nn1,nn2)
      go to 850
 825  write (6,910) (cinv(nn),nn=nn1,nn2)
 850  if (nvar.gt.8) write (6,901)
      do 860 n=1,nvar
 860  u2cinv(n) = beta(n) * beta(n) / cinv(n*(n+1)/2)
      write (6,911)
      write (6,913)     (u2cinv(n),n=1,nvar)
      rw0 = 0
      nn = 0
      ns1 = ns + 1
      ndf = nvar - ns
      do 8600 n=ns1,nvar
```

```
      nn0 = n*(n-1)/2 + ns
      do 8600 n1=ns1,n
      nn = nn + 1
      nn0 = nn0 + 1
 8600 cinv(nn) = cinv(nn0)
CC     invert matrix
      call linv1p (cinv,ndf,c,1,d1,d2,ier)
      nnsv = 1
      do 8610 n=ns1,nvar
      nnsv = nnsv + n - ns - 1
      nn = nnsv - 1
      do 8610 n1=ns1,nvar
      nn = nn + 1
      if (n1.gt.n) nn = nn + n1 - ns - 2
 8610 rw0 = rw0 + beta(n)*beta(n1)*c(nn)
  862 if (nshift.ne.0.or.niter.ge.maxitr) write (6,914)
      alrchi = 2.*(alnl-alnl0)
      df = nvar
      if (ipr.eq.1) df = df - ns1
CC     calculate the chi-s value
      call mdch (alrchi,df,psig,ier)
      psig = 1. - psig
      write (6,920)
      write (6,908) alrchi,df,psig
CC     calculate the chi s value
      call mdch (rw,df,psig,ier)
      psig = 1. - psig
      write (6,924) rw,df,psig
      alrchi = 2*(alnl-alnl1)
      df = nvar-ns
CC     calculate the chi s value
      call mdch (alrchi,df,psig,ier)
      psig = 1. - psig
      write (6,921)
      write (6,908) alrchi,df,psig
  870 df = nvar-ns
CC     calculate the chi s value
      call mdch (rr0,df,psig,ier)
      psig = 1. - psig
      write (6,912) rr0,df,psig
CC     calculate the chi s value
      call mdch (rw0,df,psig,ier)
      psig = 1. - psig
      write (6,924) rw0,df,psig
      if (maxitr.ge.0) go to 881
      do 880 n=1,nvar
  880 beta(n) = betasv(n)
      return
  881 if (kch.eq.0) return
```

```
      write (6,930)
      ch0 = 0
      ch1 = 0
      skm0 = 1
      skm1 = 1
      j = ndead + 1
      do 885 i=1,ndead
      j = j-1
      if (j.eq.ndead) go to 882
      if (tdeath(j).eq.tdeath(j+1)) go to 885
882   alam0 = msv(j) / srsv(j)
      alam1 = msv(j) / float(nrsksv(j))
      rbar = srsv(j) / nrsksv(j)
      ch0 = ch0 + alam0
      ch1 = ch1 + alam1
      skm0 = skm0 * (1-alam0)
      skm1 = skm1 * (1-alam1)
      sbres0 = exp(-ch0)
      sbres1 = exp(-ch1)
      write (6,931)   tdeath(j),nrsksv(j),msv(j),alam1,ch1,skm1,sbres1,
     1                                rbar,alam0,ch0,skm0,sbres0
885   continue
890   return
895   write (6,922) nphr,i,key2,(dd(l),l=1,ndatin)
      write (6,923) r,rb(n),(beta(n),n=1,nvar)
      stop
901   format (4x,8f15.5/(13x,8f15.5))
902   format ('   variable           beta',
     1'        observed     expected          u          e(i) ',
     2'       (e(i))**-1      chisq        p'/)
903   format ('correlation matrix')
904   format (i6,i7,6x,i8,6e13.5,f9.3,f11.6)
905   format ('-iteration outputs :'/'     iteration   increment',
     1'    error     difs              log            lr test        ',
     2' lr test           one       two'/
     2'        number     halving     codes    multiplier    likelihood'
     3,'       h0:beta=0     h0:beta=beta0  h0:beta=mle ',
     4'    h0:beta=0'/)
906   format (3i11,6f15.6)
907   format ('MLE of beta'/)
908   format ('LR chi square=',f12.6,f4.1,'df,p=',f9.6)
910   format (4x,8e15.6/(13x,8e15.6))
911   format (' tests to remove')
912   format ('chi square=',f12.6,f4.0,'df,p=',f9.6)
913   format (13x,8f15.3)
914   format ('no convergance')
915   format ('variances,with Asym normality')
916   format ('temps')
918   format ('ln(L) null hypothesis=',f13.4)
```

```
920   format ('significance tests')
921   format ('adds')
922   format ('?initial estimates of beta produce zero or negative',
     1' r for subject ',i5,' of risk set ',i5,' with id = ',a8/
     2' raw data for subject follows'/10f13.6/10f13.6)
923   format ('computed r = ',e16.5/'     n                    rb(n)',
     1'         beta(n)'/(i5,2x,2e15.5))
924   format ('test chi square=',f12.6,f4.0,'df,p=',f9.6)
927   format ('-final  estimate of the relative risk',
     1' function and its asymptotic standard error'/' increment    ',
     2' z      beta     se(beta)      r           x se(r)    cum. rw test')
928   format (i7,f10.3,2f10.6,3f11.4)
929   format (i6,i7,6x,f8.3,i8,6e13.5,f9.3,f11.6)
930   format ('1estimated survival functions at the mle of beta'/
     1'0 survival   num    num          entire  set    - unadjusted for covar
     2iates',10x,'mean',8x,' null functions - i.e. evaluated at z = 0'/
     3'     time      at  dead          hazard    cumulative   survival func
     4tions',10x,'risk',8x,'hazard     cumulative   survival functions'/
     5'              risk                rate       hazard        cox
     6     ',10x,'     ',8x,' rate        hazard        cox                '/)
931   format (f8.1,i7,i5,4x,2f12.7,2f11.5,f13.4,4x,2f12.7,2f11.5)
998   format (38x,f10.4,6x,6e13.5/(3x,10e13.5))
999   format (' ',i4,i5,2x,i3,f8.2,7x,f10.4,6x,6e13.5/
     1    (3x,10e13.5))
      end
      subroutine getr
      real blank,r,rb,idno,beta
      common tdeath(1000),istdvr(40),iselvr(40),beta(40),
     3     rb(40),u2cinv(40),dd(40),z(40),zk(40),
     2          r,idno,ns,ireprt,irej,alnl1,kch,ndatin,distnc,
     4       model,itime,mxhalf,status,stp,ipr,npbase,liktyp,
     1     nvar,maxitr,alnl0,time,ncase,ndead,break,curtim,timint,
     5     ivtime,ivstat,iventr,ividno,is,nvtot,istep,niter,ivznp
      if (is.eq.1.and.(istep.ne.1.or.niter.ne.1)) go to 150
      time = 0
      idno = 0
      tentry = 0
100   read (8,903) (dd(n),n=1,ndatin)
903   format (16I5)
      call transf
      status = dd(ivstat)
      if (irej.eq.1.and.(kch.eq.1.or.status.eq.0.or.liktyp.eq.1))
     1     go to 100
      if (ivtime.ne.0) time = dd(ivtime)
      if (ividno.ne.0) idno = dd(ividno)
      if (iventr.ne.0) tentry = dd(iventr)
      if (tentry.gt.curtim.and.kch.eq.1) go to 100
150   continue
151   do 200 n=1,nvar
```

```
      if (iselvr(n).eq.0) go to 190
      z(n) = dd(iselvr(n))
      go to 200
  190 z(n) = 1
  200 continue
      if (ipr.eq.1) go to 700
      go to (300,400,500,600,600,650),model
      go to 899
  300 betaz = 0
      do 310 n=1,nvar
  310 betaz = betaz + beta(n)*z(n)
      r = exp(betaz)
      do 320 n=1,nvar
  320 rb(n) = r*z(n)
      go to 800
  400 r = 1
      do 410 n=1,nvar
      r = r + beta(n)*z(n)
  410 rb(n) = z(n)
      go to 800
  500 r = 1
      do 510 n=1,nvar
      ebeta = exp(beta(n))
      r = r + ebeta*z(n)
  510 rb(n) = ebeta*z(n)
      go to 800
  600 call comb
      go to 800
  650 call mdlsub
      go to 800
  700 r = 0.
      kz = 0
  710 kz = kz + 1
      if (z(ivznp).lt.zk(kz).or.kz.gt.nvar) go to 720
      r = r + beta(kz)
      go to 710
  720 r = exp(r)
      kzm1 = kz - 1
      do 725 n=1,kzm1
  725 rb(n) = r
      if (kz.gt.nvar) go to 740
      do 730 n=kz,nvar
  730 rb(n) = 0
  740 continue
  800 return
  899 continue
  999 continue
      stop
      end
```

```
      subroutine comb
      real blank,r,rb,idno,beta
      common tdeath(1000),istdvr(40),iselvr(40),beta(40),
     3     rb(40),u2cinv(40),dd(40),z(40),zk(40),
     2          r,idno,ns,ireprt,irej,alnl1,kch,ndatin,distnc,
     4       model,itime,mxhalf,status,stp,ipr,npbase,liktyp,
     1     nvar,maxitr,alnl0,time,ncase,ndead,break,curtim,timint,
     5     ivtime,ivstat,iventr,ividno,is,nvtot,istep,niter,ivznp
      dimension rab(40),rmb(40)
      nm1 = nvar - 1
      b = beta(nvar)
      if(nvar.gt.ns) go to 300
      nm1 = ns
      b = 1
  300 go to (800,800,800,400,500),model
      go to 800
  400 radd = 1
      rmult = 1
      do 450 n=1,nm1
      radd = radd + beta(n)*z(n)
      rab(n) = z(n)
  450 rmult = rmult * (1 + beta(n)*z(n))
      do 460 n=1,nm1
  460 rmb(n) = z(n)*rmult/(1+beta(n)*z(n))
      go to 700
  500 ndum = 0
      do 550 n=1,nm1
      rab(n) = 0
      rmb(n) = 0
      if (z(n).eq.0) go to 550
      go to (530,550),ndum
      i1 = n
      ndum = 1
      go to 550
  530 i2 = n
      ndum = 2
  550 continue
      go to (570,580),ndum
      radd = 1
      rmult = 1
      go to 700
  570 radd = 1 + beta(i1)
      rmult = radd
      rab(i1) = 1
      rmb(i1) = 1
      go to 700
  580 radd = 1 + beta(i1) + beta(i2)
      rmult = radd + beta(i1)*beta(i2)
      rab(i1) = 1
```

```
      rab(i2) = 1
      rmb(i1) = 1 + beta(i2)
      rmb(i2) = 1 + beta(i1)
700   bm1 = 1 - b
      r = radd ** bm1 * rmult ** b
      rb(nvar) = r * (alog(rmult) - alog(radd))
      do 750 n = 1,nm1
750   rb(n) = r * (bm1 * rab(n) / radd + b * rmb(n) / rmult)
 800  return
      end
      subroutine mdlsub
      real r,rb,beta
      common j1(1160),beta(40),rb(40),j2(100),z(40),j3(40),r,j4(17),
     1     nvar,j5(17)
      return
      end
ccc    this the same subroutine as transf with dummy
      subroutine drtans
      common j1(1360),dd(40),j2(106),irej,j3(19),curtim,j4(10)
      irej = 0
      return
      end

      subroutine transf
      common j1(1360),dd(40),j2(106),irej,j3(19),curtim,j4(10)
      do 100 n=10,16
100   dd(n)=0
      irej = 0
      return
      end
```

References:

ANDERSON, P.K. & RASMUSSEN, N.K., (1982) Research Report 8216, statistical research unit, Danish Medical Research Council.

BARLOW, R.E., & PORSCHEN, F. (1976). Asymptotic Theory of Total time on Test processes, with applications to life-testing. Multivariate Analysis IV, P.R. Krishnaiah (Ed.) Amsterdam, N. Holland, 227 - 237.

BARTLETT (1959) An introduction to stochastic processes, with special reference to methods and applications. Cambridge University Press.

BAUM, M., KAY, R., SCHEURLEN, H., (1982) Clinical Trials in Early Breast Cancer, 2nd Heidelberg Sumposium, Heidelberg, Dec. 14-17, 1981.

BECK, (1975) Stochastic survival models with competing risks and covariates - unpublished Ph.D. University of California, Berkely, Canlffornia.

BERNOULLI, D., (1760), Essai d'une nouvelle analyse de la mortalite causee par la petite verole, & des avantages de l'inoculation pour la prevenir. Academie Des Sciences, Paris, Histoire avec les Memoires, pp. 1 - 45.

BERNSTEIN, F., BINHAM, Z.W., & ACHS, S. (1939). Is or is not cancer dependent on age? Am. J. Cancer, 37. 298 - 311. Oct. 1939.

BERTRANON, E.G., BLACKSTONE, E.H., HAZELRIG, J.B., TURNER,M.E., KIRKLIN, J.W., (1978), The American J. of Cardiology, 42, 458-465.

BLOOM, H.J.G., & RICHARDSON, W.W., (1957) Histological grading and prognosis in breast cancer a study of 1409 cases of which 359 have been followed for 15 years. British Journal of Cancer, II (3), Sept. 359-377

BLOOM, H.J.G. (1965) The Influence of delay on the natural history and progression of breast cancer. British Journal of Cancer, 19, 228-262.

BOAG, J.W., (1949), Maximum likelihood estimates of the proportion of patients cured by cancer therapy, J. of Royal Stat. Soc. Vo. B 11, pp. 15-53.

BRANDFORD-HILL,A ., (1962), Statistical methods in clinical and preventive medicine; Epochal streptomycin trial, Edinburgh, 44-45.

BRESLOW, N.E., (1982), Multiplicative models and the analysis of cohort studies. Paper presented at the Royal Statistical Soc. Conference,York.

BROWN, B.M. (1980), The crossover experiments for clinical trials. Biometrics (36) , 69-80.

BRYSON, M.C., & JOHNSON, M.E., (1981). The incidence of monotone likelihood in the Cox model. Technometrics, 23, 381-384.

CHASTANG, C.L., (1983). A simulation study in a two-covariate survival model: 4th International Congress in Clinical Biostatistics incorporating a mini-symposium on diabetics. 11th-15th Sept. 1983, Paris.

CHIANG, C.L. (1960) A stochastic study of life tables and its applications: I Probability distribution of the biometric functions. Biometrics, 16, 618-635.

CHIANG, C.L., (1964). A stochastic model of competing risks of illness and competing risks of death. pp. 323-354 in: Stochastic Models in medicine and biology, J. Gurland (Ed.) U. Wisconsin Press, Madison.

CHIANG, C.L. (1966), On the Formula for the Variance of the Observed Expectation of life. - E.B. Wilson's approach, Human Biology, Vol.38, pp. 318-319.

CHIANG, C.L. (1976) Multiple transition time in a simple illness death process - Fix-Neyman Model, Mathematical Biosciences, 31, 51-71.

CHIANG, C.L. (1979) Survival and stages of disease. Mathematical Biosciences, 43, 159-171.

CORNFIELD, J. (1957). The estimation of the probability of developing a disease in the presence of competing risks. Am.J. Public Health, 47, 601-607,

COX, D.R., (1972). Regression models and life tables (with discussion) J.R. Stat. Soc. B.34, 187-220.

CROWLEY, J. (1974). A note on some recent likelihoods leading to the log rank test. Biometrika, 61, 533-538.

ROWLEY, J. & HUM (1977) Covariance analysis of heart transplant survival data, J. of Amer. Stat. Assoc. 72, 357, 27-36.

D'ALEMBERT (1761) Sur l'application due calcul des probabilites a l'inoculation de la petit verole, Opuscules II 26-95.

DAVIS, D.J. (1952), An analysis of some failure data. J. Amer. Stat.Assoc. 47, 113-150.

DEVITT, J.E. (1967).Clinical stages of breast cancer. Canad.Med.Ass. Journal, 97, 1257-1261.

EFFRON, B. (1977). Efficiency of Cox's likelihood function for censored data. J. Am. Stat. Assoc. 72, 557-565.

EPSTEIN B., AND SOBEL, M.(1953, 1954).
(1953 Life testing. J. Am. Stat. Assoc. 48, 486-502.)
(1954 Some theorems relevant to life testing from an exponential distribution. Ann. Math. Stat. 25, 373-81.

FIEGLE, P. & ZELEN,M. (1965). Estimation of exponential survival probabilities with concomitant information. Biometrics, 21, 826-838.

FIX, E. & NEYMAN, J. (1951). A simple stochastic Model of recovery, Relapse, Death and Loss of Patients. Human Biology, Vol 24, 205-241.

GAIL, M.H., STANTENER, T.J., BROWN, C.C. (1980). An analysis of comparative carcinogenesis experiments based on multiple times to tumour. Biometrics, 36, 255 -266.

GOMPERTZ, B. (1825). On the nature of the function expressive of the Law of human mortality. Philos. Trans. Roy. Soc. (London) 115; 513-583.

GORE, S. (1981) Unpublished Ph.D. thesis, University of Aberdeen. Long Term survival in breast cancer.

HALLEY, E. (1693). An estimate of the degreee of the mortality of mankind, drawn from curious tables of the births and funerals at the city of Breslau. Philos. Trans. Roy. Soc. (London) 17: 569-610.

HALDERIN, M. (1952) Maximum likelihood estimation in truncated samples. Annals of Mathematical Statistics, Vol 23, pp. 226-238.

HANNAN, M.T., TUMA, N.B., and GREENVELD,L.P., (1978). Income and independence effects on marital dissolution. American J. of Sociology, 84, 3, 611-633.

HANNAN, M.T., & CARROLL, G.R., (1981). Dynamics of formal political structure: an event-history analysis. American Sociological Review, 46, 19-35.

HARMER, M.H. (1958). The British clinical staging of breast cancer. British Medical Journal, NO. 5073, 767-769.

HAZELRIG, J.B., TURNER, M.E., & ACKERMAN, E. (1978). A function minimization computer package for non-linear parameter estimation providing readily accessible maximum likelihood estimators. Computers and Biomedical Research, 11, 51-63.

INGLEBY, H.,& GERSHON-COHEN, J: Comparative anatomy, pathology and roentgenology of the breast. Philadelphia, University of Pennsylvania Press (1960) p.359.

JOHANSON, S. (1978). Product limit estimator as maximum likelihood estimator. Second J. Statistics, 5, 195-199.

JOHNSON, R.C. and JOHNSON, N.L., (1980). Survival Models and data analysis. John Wiley & Sons.

KALBFLEISCH, J.D., & PRENTICE, R.L. (1972). Contribution to the discussion paper by D.R. Cox. J.Roy.Stat. Soc. B. 34, 205-207.

KALBFLEISCH, J.D. & PRENTICE, R.L. (1973). Marginal likelihoods based on Cox's regression and life model. Biometrika, 60, 267-278.

KALBFLEISCH, J.E. (1974). Some efficiency calculations for survival distributions. Biometrika, 61, 31-38.

KALBFLEISCH, J.D., & MACINTOSH, A.A. (1977), Efficiency in survival distributions with time dependent covariables. Biometrika, 64, 47-50.

KALBFLEISCH, J.D., (1980). Non-parametric Bayesian analysis of survival time data. J.Roy. Statistical Soc. B. 40, 214-221.

KAY, R. (1977). Proportional hazard regression models and the analysis of censored survival data. J.Roy. Stat. Soc. C. 26, 227-237.

KAY, R. (1979). Some further asymptotic efficiency calculations for survival data regression models. Biometrika, 66, 91-96.

KUZMA, J. (1967). A comparison of two life tables methods. Biometrics. Vol 23, pp.51-64.

LAGAKOS, S.W., SUMMER, C.J. & ZELEN, M. (1978). Semi-markov models for partially censored data. Biometrika, 65, 2, 311-317.

LANE, N., GOKSEL, H., SALERNO, R.A. & HAAGENSON, C.D., Clinico-Pathologic analysis of the surgical curability of breast cancers; a minimum 10 year study of a personal series. Ann. Surg. (1961) 153, pp 483-498.

LEHMAN, E.L. & SCHEFFE, H. (1950). Completeness, similar regions and unbiased estimation. Sankhya, Vol. 10. 305-340.

LUSTBACLER, E.D., (1980). Biometrika, 67, 3, 697-698.

MAKEHAM, W.M. (1874). On the law of mortality and the construction of annuity tables. J. Inst. Actuaries (London), 8.

MAKEHAM, W.M. (1875). On an application of the theory of the composition of decremental forces. J. Inst. Actuaries. 18, 317-322.

NEYMAN, J. & PERASON, K. (1933). On the problem of the most efficient tests of statistical hypothesis. Phil. Trans. Roy. Soc. A. 231, 289-337.

NEYMAN, J. (1950). First course in Probability and Statistics. New York. Holt.

OAKS, D. (1981). Survival times: aspects of partial likelihood. International statistical review, 49, pp.235-264.

PETO, R. (1972a). Rank tests of maximal power against Lehmann-type alternatives. Biometrika, 59, 472-475.

PETO , R. (1972b). Contribution to the discussion of paper by D.R.Cox. J. Roy.Statistical Soc. B. 34, 205-207.

PETO, R., PIKE, M.C., ARMITAGE, P, BRESLOW, N.E., COX, D.R., HOWARD, S.V., MANTEL, N., MCPHERSON, K., PETO, J. & SMITH, P.S. (1977). Design and analysis of randomised clinical trials requiring prolonged observation of each patient. Part 2. Analysis and examples. Br.J. Cancer, 35, 1-39.

PICARD, J.D. (1962). The role of mammography in prognosis of malignant tumour of the breast. Acta. Un. Int. Cancer (1962) 18, pp 823-825.

POCOCK, S.J. (1979). Allocation of patients to treatments in clinical trials. Biometrics, 35, 183-197.

PRESCOTT, R.J. (1978) . Feed back of data to participants during clinical trials,. Proceedings of an EORTC symposium held in Brussels Belgium, April, 26-29th, 1978. Published in Controversies in Cancer: design of trials and treatment. Masson Publishing, U.S.A. (1979). 55-61.

PRETE, D.I. (1981). Unemployment over the life cycle: Racial differences and the effects of changing economic conditions. American Journal of Sociology, 87, 2, 286-307.

PROUT, G.R., SLACK, N.H. & BROSS, I.D.J., (1970). Irradiation and 5-fluorouracil as adjuvants in management of invasisve bladded carcinoma - A co-operative group report after 4 years. Journal of urology, 104 (I): 116-119

RUTQUIST, L.E., WALLGREN, A, & NILSSON, B., (1982). Is breast cancer a curable disease? Unpublished report.

SILVEY, S.D., (1975) Statistical Inference. Chapman and Hall, London.

SIMONS, R. (1979). Restricted randomisation designs in clinical trials, Biometrics, 35, 503-512, June, 1979.

STEWART, H.J., WILLIAMS, W.J., APSIMON, H.T., GRAVELLE, I.H. & FORREST, A.P.M.: Prognostic factors in breast cancer. Ed. Forrest, A.P.M. & Kunkler, P.B.,Edinburgh, Livingstone. (1968). p.301-308.

STORMER, J.L., (1980). Quantitative management in clinical medicine. Symposium on uses if clinical and biopharmaceutical statistics. 21-23 May, 1980. New York.

TARONE, R.E. (1975). Tests for trend in life tables analysis. Biometrika, 62m 3, 679-682.

TARONE, R.E. & WARE, J. (1977). On distribution -free tests for equality of survival distributions. Biometrika, 64, 1,pp. 156-160.

TAULBEE, J.D. (1979) A general model for the hazard rate with covariables . Biometrics, 35, 439-450.

THOMAS, D.C., (1980). Multivariate methods for risk detection; relating multiple exposures to multiple disease and points. Presented at the Society for Epidermiologic research. Minneapolis, Minn. June, 18-20. 1980.

TOLLERY, H.D. (1978). A non-parametric test for composite hypothesis in survival analysis. Annals of the Ins. of Statistical Mathematics, 30, A, 281-295.

TURNER, M.E., BRADLEY, E.L., & KIRK, K.A., (1976). A theory of growth, Mathematical Biosciences, 29, 367-373.