Spring 1984

STATISTICAL METHODS FOR ANALYSIS OF CANCER CHEMOPREVENTION EXPERIMENTS

STEPHEN MICHAEL KOKOSKA
University of New Hampshire, Durham

Follow this and additional works at: https://scholars.unh.edu/dissertation

INFORMATION TO USERS

This reproduction was made from a copy of a document sent to us for microfilming. While the most advanced technology has been used to photograph and reproduce this document, the quality of the reproduction is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help clarify markings or notations which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure complete continuity.

2. When an image on the film is obliterated with a round black mark, it is an indication of either blurred copy because of movement during exposure, duplicate copy, or copyrighted materials that should not have been filmed. For blurred pages, a good image of the page can be found in the adjacent frame. If copyrighted materials were deleted, a target note will appear listing the pages in the adjacent frame.

3. When a map, drawing or chart, etc., is part of the material being photographed, a definite method of "sectioning" the material has been followed. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.

4. For illustrations that cannot be satisfactorily reproduced by xerographic means, photographic prints can be purchased at additional cost and inserted into your xerographic copy. These prints are available upon request from the Dissertations Customer Services Department.

5. Some pages in any document may have indistinct print. In all cases the best available copy has been filmed.

8419556

Kokoska, Stephen Michael

STATISTICAL METHODS FOR ANALYSIS OF CANCER CHEMOPREVENTION
EXPERIMENTS

University of New Hampshire   PH.D.   1984

University
Microfilms
International   300 N. Zeeb Road, Ann Arbor, MI 48106

STATISTICAL METHODS FOR ANALYSIS OF
CANCER CHEMOPREVENTION EXPERIMENTS

BY

Stephen Michael Kokoska
B.A., Boston College, 1978
M.S., University of New Hampshire, 1980

DISSERTATION

Submitted to the University of New Hampshire
in Partial Fulfillment of
the Requirements for the Degree of

Doctor of Philosophy
in
Mathematics

May, 1984

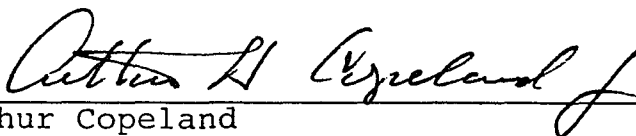This dissertation has been examined and approved.

_____
Dissertation director, L. David Meeker
Professor of Mathematics

_____
Samuel Shore
Associate Professor of Mathematics

_____
Kenneth Constantine
Assistant Professor of Mathematics

_____
Arthur Copeland
Professor of Mathematics

_____
Henry Thompson
Associate Professor of Nutrition

_____5/1/84_____
Date

TABLE OF CONTENTS

LIST OF TABLES

ABSTRACT

STATISTICAL METHODS FOR ANALYSIS OF
CANCER CHEMOPREVENTION EXPERIMENTS

by

Stephen Michael Kokoska

University of New Hampshire, May, 1984

The experimental systems studied in this dissertation
are designed to investigate the effect of diet on incidence
rates of cancer. These investigations involve the chemical
induction of tumors in experimental animals in order to test
the chemopreventative effects of various substances. Both
tumor number and rate of tumor development are important in
evaluating the effects of a chemopreventative agent. This
is made difficult, when multiple tumors occur, by the confounding
of the number of induced tumors and their time of detection.
This confounding occurs because experiments are terminated
before all induced tumors have been detected. Fewer tumors
observed in one treatment group, as compared to another, may
be the result of a decreased number of induced tumors, a
slowing of tumor growth rate, or both. Current statistical
procedures do not consider this factor and therefore, do not
reliably discriminate between these biologically different
possibilities.

This study provides the cancer researcher with statistical

procedures which directly address this problem of confounding
of tumor number and detection time distributions. The method
of maximum likelihood is used to simultaneously estimate the
parameters characterizing these two confounded distributions.
In order to compare treatments the likelihood ratio test is
used to detect overall group differences and a technique is
described to isolate which factor(s) (tumor number and/or
rate of development) is(are) contributing to a group difference.
Numerical results are used to discuss the sensitivity of the
estimation procedure subject to changes in the experimenter
controlled variables in order to design more accurate and
efficient experiments and better utilize resources.

CHAPTER I

INTRODUCTION AND BACKGROUND

Description of Experiments

The experimental systems studied in this dissertation are designed to investigate the effect of diet on incidence rates of cancer. These investigations involve the use of animal tumor models in order to test the chemopreventative effects of vitamin A, selenium, and other substances. Experimental animals are randomly assigned to treatment or control diets before or after exposure to a carcinogen depending on the focus of the experiment. Among the response variables used to compare treatments are the number of induced cancers and the rate at which they develop.

Throughout this study the term tumor is used to mean malignant tumor or cancer. We distinguish two types of animal tumor models that involve the chemical induction of cancers in laboratory animals. An experimental system is of type A if both the number of tumors per animal and their times to detection are available. Examples of such are experimental skin and mammary tumor systems. Type B systems are those which require that an animal be sacrificed in order to estimate its induced tumor burden. In such experiments only the number of induced cancers and an upper bound on their detection times is available for use in comparing treatments. Included among this type of experimental system are those

involving cancers of the liver, kidney, and lung.

An effective chemopreventative agent should decrease the number of tumors per animal and/or retard the tumor growth rate. Current statistical procedures for treatment comparisons do not account for the confounding of tumor number and cancer growth rate parameters. As a result the experimenter cannot reliably discriminate between a chemopreventative agent which decreases the number of induced tumors and one which slows their rate of development. This study is intended to provide the cancer researcher with statistical procedures which directly address this problem of confounding by developing, testing, and implementing methods of simultaneous assessment of tumor number and growth rate parameters in experiments of type A and B.

Statement of the Problem

Both tumor number and rate of tumor development are important in evaluating the effects of a chemopreventative agent on experimentally induced neoplasia. This is made difficult, when multiple tumors occur, by the confounding of the number of induced tumors and their time of detection. This confounding occurs because experiments are terminated before all induced tumors have been detected. Fewer tumors observed in one treatment group, as compared to another, may be the result of a decreased number of induced tumors, a slowing of tumor growth rate, or both. Current statistical test procedures do not consider this factor and therefore,

do not reliably discriminate between these biologically different possibilities.

To illustrate this confounding problem, Figure 1 exhibits the effects of changes in the number of tumors and rate of tumor development on the tumor incidence curve (describing the proportion of tumor bearing animals) and, T_{50}, the time of 50% tumor incidence. We define λ and μ to be the mean number of induced tumors and the mean time to tumor detection (which is related to the rate of tumor development), respectively. A reduction in induced tumors can seem to shift the incidence curve to the right without a change in rate of tumor development. However, if the number of induced tumors is unchanged, and the rate of tumor development slows, a similar shift can occur. The statistical tests described below are intended to discriminate between these two actions and provide a more accurate analysis of experiments of this nature.

The objectives of this study are;

1) to develop procedures for the joint estimation of tumor number and parameters related to tumor growth rate for each experimental treatment group,

2) to develop procedures for comparison among experimental treatment groups in order to evaluate the significance of treatment effects,

3) to investigate the influence of experimental parameters including

    a) frequency or duration of examination intervals,

    b) duration of experiment,

    c) number of animals per treatment group

Figure 1

Graph (a)
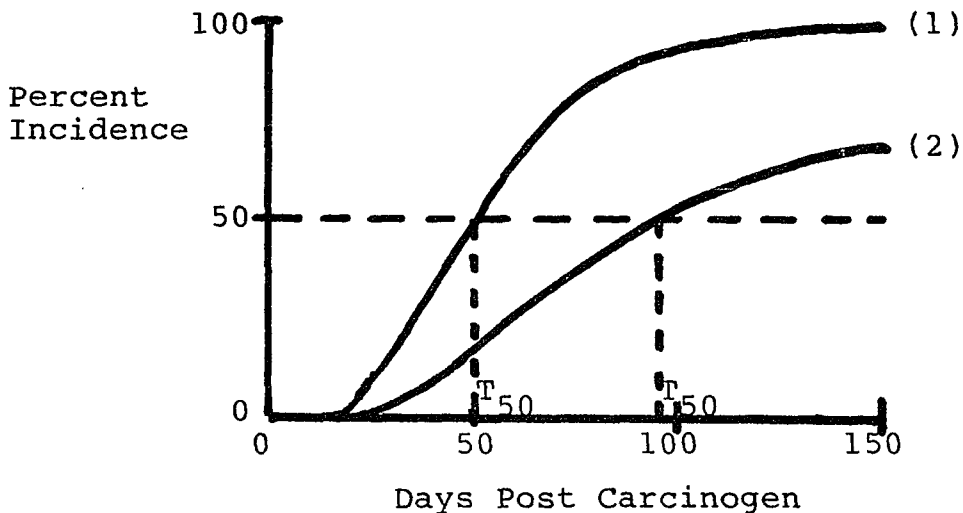
Graph (b)
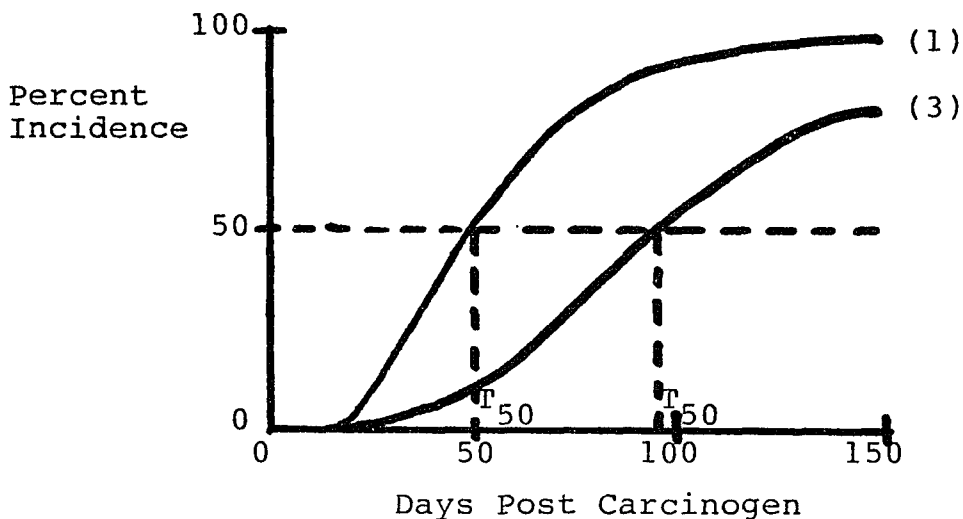
Figure 1: In graph (a) curve (1) has parameters $\lambda=5$, $\mu=100$ and curve (2) has parameters $\lambda=1.32$, $\mu=100$. The growth rate parameters remain constant but a decrease in the number of induced tumors shifts the incidence curve to the right. Note the increase in T_{50} from 50 to 95.
In graph (b) curve (1) is unchanged but curve (3) has parameters $\lambda=5$, $\mu=210$. Here the number of induced tumors remains fixed, and a change in the growth rate parameters causes a similar shift in the incidence curve.

in order to design more accurate and efficient experiments and better utilize resources.

The Model

The mathematical model characterizing type A and type B experimental systems is based on a two-stage theory of tumor development which represents the carcinogenic process by a combined initiation and promotion stage and a progression, or development, stage. It is assumed the carcinogen initiates a susceptible cell with the potential for neoplastic growth. If the cell is promoted, it expresses that potential and loses its ability to control growth and cell division. Thus a single, previously normal, cell is promoted to the neoplastic state. In the progression stage, the clone arising from a promoted cell develops into a clinically detectable cancer.

Let the random variables M and T denote the number of promoted tumors per animal in a given treatment group and the time to detection of a randomly selected tumor on that animal, respectively. Let μ_M and σ_M^2 be the mean and the variance of the random variable M, respectively, and let F(t) be the cummulative distribution function (cdf) of the random variable T. Let J(t) be the number of observed tumors per animal at time t. The following theorem demonstrates the dependence of the number of detectable tumors at time t upon the mean number of promoted tumors and the detection time distribution.

Theorem 1.1   J(t) has mean $\mu_M F(t)$ and variance

$$(\sigma_M^2 - \mu_M)F^2(t) + \mu_M F(t) \ .$$

<u>Proof</u> Let $P_M(s)$ be the probability generating function for the random variable M, the number of promoted tumors per animal. Since $J(t)|M=m \sim B(m,F(t))$ then

$$P_{J(t)}(s) = E(s^{J(t)}) = \sum_{m=0}^{\infty} E(s^{J(t)}|M=m) \cdot P(M=m)$$

$$= \sum_{m=0}^{\infty} (sF(t) + 1 - F(t))^m \cdot P(M=m)$$

$$= P_M(sF(t) + 1 - F(t)) \qquad (1)$$

Therefore

$$E(J(t)) = P'_{J(t)}(1) = P'_M(1)F(t)$$

$$= E(M)F(t) = \mu_M F(t)$$

and

$$Var(J(t)) = P''_{J(t)}(1) + E(J(t)) - E^2(J(t))$$

$$= P''_M(1)F^2(t) + \mu_M F(t) - \mu_M^2 F^2(t)$$

$$= (\sigma_M^2 - \mu_M + \mu_M^2)F^2(t) + \mu_M F(t) - \mu_M^2 F^2(t)$$

$$= (\sigma_M^2 - \mu_M)F^2(t) + \mu_M F(t)$$

Using equation (1) we have the following theorem concerning the distribution of J(t).

<u>Theorem 1.2</u> If M is distributed as a Poisson random variable with parameter λ then J(t) is distributed as a Poisson random variable with parameter $\lambda F(t)$.

<u>Proof</u>

$$P_{J(t)}(s) = e^{\lambda(sF(t) + 1 - F(t) - 1)} = e^{\lambda F(t)(s-1)}$$

which is the probability generating function for a Poisson random variable with parameter $\lambda F(t)$.

Suppose the M promoted tumors of a given animal are detected at times $T_1, \ldots, T_M$. The animal's tumor-free time, $T_{(1)}$, is defined to be the minimum of the set $\{T_1, \ldots, T_M\}$. The expected proportion of tumor bearing animals in a given treatment group at time t, or incidence I(t), can now be calculated.

$$I(t) = P(T_{(1)} \leq t) = P(J(t) > 0)$$

$$= 1 - P(J(t) = 0)$$

$$= 1 - P_{J(t)}(0)$$

$$= 1 - P_M(1-F(t)) \tag{2}$$

Equation (2) further illustrates the confounding of tumor number and detection time parameters. A change in the incidence rate may be caused by a change in number of promoted tumors, the detection time parameters, or both.

Current Statistical Procedures

In this section we use the previous results concerning J(t) and the incidence rate to demonstrate that existing statistical tests for treatment comparisons in experiments of type A and B do not distinguish between a change in the mean number of induced tumors and/or a change in the rate of tumor development. We consider an experiment of type A and assume that each of the animals of a treatment group survives until the end of the experiment.

Let t* denote the length of the experiment and $J_i(t^*)$ denote the number of tumors detected on animal i during the

experiment. Then, from Theorem 1.1, $J_i(t^*)$ has mean $\mu_M F(t)$. The comparison of mean number of tumors across treatments is usually based on likelihood ratio tests [1,2] or on t-tests or ANOVA using the transformed values $\sqrt{J_i(t^*)}$ [3]. In comparing two groups corresponding to treatments 1 and 2, the parameters being compared are $\mu_1 F_1(t^*)$ and $\mu_2 F_2(t^*)$, not μ_1 and μ_2.

Tumor incidence rates are frequently used in a comparison of treatments. The 2x2 chi-square test is the most common test used to compare tumor incidence. However, since I(t) has the form of equation (2), in such comparisons both the mean number of promoted tumors and the detection time distribution influence the results of the comparisons.

Tumor latency generally refers to the rate of tumor development. The methods of survival analysis, based on tumor-free-times, are also frequently used to compare treatments [4,5]. One measure of tumor latency is the median tumor-free-time, t_{50}, defined by the equation

$$I(t_{50}) = P(J(t_{50}) \geq 1) = 1 - P_M(1 - F(t_{50})) = 1/2 .$$

Clearly, t_{50} depends strongly on the distribution of M as well as the cdf F(t).

Suppose an animal has M promoted tumors with detection times $T_1,\ldots,T_M$ and ordered times $T_{(1)},\ldots,T_{(M)}$. Both the ordered times and the "gap times" (duration between ordered detection times) can be used to test for differences among treatments [6,7,8]. These tests, however, have two major drawbacks: (a) while they test for differences among treatments, the two basic mechanisms underlying such differences (tumor

promotion and tumor development) are still confounded. For example, as μ_M increases the "gaps" will tend to decrease as if the rate of development was increasing. In addition (b), the number of tumors and or "gaps" considered is determined by the group with the fewest tumors. This can result in the exclusion (from such tests) of a large proportion of the actual tumor detection time data.

As outlined above, the chemical induction of multiple tumors on experimental animals complicates the analysis of the experiments. In contrast to survival experiments, where an organism can die only once, the number of censored tumor observations is unknown. Analyses based on multiple-failure times also suffers from the confounding of the number of tumors and their detection times during an experiment of finite duration and from the need to exclude a (possibly) large proportion of the observed detection times. These difficulties have led to the statistical procedures developed in this dissertation which explicitly acknowledge and represent the confounding inherent in experiments of type A and B.

Dissertation Overview

In type A and B experimental systems the number of induced tumors is, generally, well-described by a Poisson distribution. The Gamma family supplies a wide range of possible distributions and these are used to characterize the tumor detection times. Given these underlying distributions, the joint likelihood of observed samples can be calculated.

The method of maximum likelihood is used to estimate the

parameters characterizing the two confounded distributions -
the number of tumors and the detection time distribution
parameters. The likelihood equations for both type A and B
experiments are derived and the estimators for tumor number
and parameters related to tumor growth rate are studied. For
type A experiments the existence of a unique solution to the
likelihood equations is proven. For type B experiments the
existence of an essentially unique solution to the likelihood
equations is shown to be true as the number of observations
increases. By applying the invariance principle for maximum
likelihood estimators the mean time to tumor detection is
also estimated. In order to construct confidence regions for
the true value of the mean number of promoted tumors per
animal and the mean time to tumor detection, the variance-
covariance matrix for the corresponding estimators is derived.
This also enables the experimentor to conduct simple hypothe-
sis tests concerning these parameters.

The statistical techniques developed provide insight
into designing more efficient experiments. For experiments
of type A computer simulations are used to test the robustness
over distributional assumptions of the estimators for the
mean number of promoted tumors and the mean time to tumor
detection. Analytical methods detail the sensitivity of the
estimators to the number of animals per group, length of
experiment, and frequency of examination intervals.

The likelihood ratio test is used to assess the statisti-
cal evidence in support of H_o (all animals in two treatment
groups come from the same population) versus H_a (these two

groups arise from different populations). This analysis first examines the difference between treatment groups in terms of the parameters considered simultaneously. Once a significant difference is established between treatment groups, further statistical tests are outlined which enable the experimenter to detail along what dimensions the differences occur.

CHAPTER II

ONE SAMPLE PROBLEM – ESTIMATION OF PARAMETERS

Introduction

In this chapter the objective is to estimate, using the sample from a type A or B experiment, the value of the parameters characterizing the distributions of M, the number of promoted tumors per animal, and T, the time to tumor detection. M is assumed to have a Poisson distribution with expected value λ, and T is assumed to be distributed as a Gamma random variable, $\Gamma(\alpha,\beta)$, with expected value $\alpha\beta$ and variance equal to $\alpha\beta^2$. The parameters of interest are λ, the mean number of promoted tumors per animal, α and β, parameters related to the tumor growth rate, and their product $\mu=\alpha\beta$, the mean time to tumor detection.

The Method of Maximum Likelihood

The method of maximum likelihood is a constructive procedure for obtaining point estimates and can be used to estimate the population parameters λ, α, β. In addition, when the sample size is large the resulting estimators have certain desirable properties. We begin by recalling the definition of the likelihood function of a sample.

<u>Definition</u> Let the random variables $X_1,\ldots,X_n$ have a joint probability mass function (pmf) or probability density function (pdf) $f(\underline{x}|\underline{\theta})$ where $\underline{x}=(x_1,\ldots,x_n)$ is the vector of

12

observations and the parameter $\underline{\Theta}=(\Theta_1,\ldots,\Theta_m)\varepsilon\Omega\underline{\subseteq}\mathbb{R}^m$. The likelihood function $L(\underline{x}|\underline{\Theta})$ of the sample is the joint probability(density) of the observations as a function of the parameter $\underline{\Theta}$.

The method of maximum likelihood selects as estimates $\hat{\Theta}_1,\ldots,\hat{\Theta}_m$ those values of the Θ_i's that maximize the likelihood function. That is we choose $\hat{\underline{\Theta}}$ (if it/they exist(s)) such that

$$L(\underline{x}|\hat{\underline{\Theta}}) \geq L(\underline{x}|\underline{\Theta}) \qquad\qquad \underline{\Theta}\varepsilon\Omega$$

First we consider the case of n independent observations taken from the same distribution and one unknown parameter Θ. To find the maximum likelihood estimate for Θ, we form the likelihood function $L(\underline{x}|\underline{\Theta})=f(x_1|\Theta)\cdots f(x_n|\Theta)$ where $f(x|\Theta)$ may be a pmf or pdf. We assume $\underline{x}=(x_1,\ldots,x_n)$ is the vector of observations, $\Theta\varepsilon\Omega=[a,b]\subseteq R$, the support of $f(x|\Theta)$ is independent of Θ, and $L>0$. We select the maximum likelihood estimate $\hat{\Theta}$, the value of Θ which maximizes the function $L(\underline{x}|\Theta)$.

The likelihood function often contains complicated products and is therefore difficult to work with. In this case it is easier to work with the natural logarithm of L, ln L, rather than L itself. Since ln L is a monotonic function of L, finding $\hat{\Theta}$ which maximized ln L is equivalent to maximizing L itself.

In the more general case, n independent observations are taken from the same distribution with m unknown parameters. These m parameters must be estimated simultaneously. The notion of maximum likelihood is the same, we must select a set of admissible values of the parameters $\Theta_1,\ldots,\Theta_m$ which

make the likelihood function an absolute maximum. The following theorem, stated in Kendall and Stuart [9], formalizes this concept.

Theorem 2.1 Let $f(x|\underline{\Theta})$ be a pmf or pdf such that $\underline{\Theta}=(\Theta_1,\ldots,\Theta_m)\varepsilon\Omega\subseteq R^m$ where Ω is a connected open set, and the following regularity conditions are true;

a) the support of $f(x|\underline{\Theta})$ is independent of $\underline{\Theta}$,

b) $\underline{\Theta}$ may take on any value in Ω,

c) the likelihood function has continuous second partial derivatives.

Then the necessary conditions for a local extrema of the likelihood function are

$$\frac{\partial\ln L(\underline{x}|\underline{\Theta})}{\partial\Theta_i} = 0 \qquad\qquad i=1,\ldots,m \qquad\qquad (3)$$

and for a solution to (3) to be a local maximum, the matrix

$$N(\underline{\hat{\Theta}}) = \left(\frac{\partial^2\ln L(\underline{x}|\underline{\hat{\Theta}})}{\partial\Theta_i\partial\Theta_j}\right) \quad i=1,\ldots,m \ , \ j=1,\ldots,m \qquad (4)$$

must be negative definite.

Theorem 2.1 does not imply the existence of a solution to the system of equations in (3). It is also possible that two or more solutions to (3) yield the same maximum value of the likelihood function. In order to state a uniqueness theorem we will need the following definitions.

Definition A statistic is a function of observable random variables with no unknown parameters.

We will use the following as a definition of jointly

sufficient statistics. The alternate definition reduces to a condition (equation (5)) on the likelihood function.

Definition Let $S_1, \ldots, S_k$ be statistics based on the random sample $X_1, \ldots, X_n$. Then $S_1, \ldots, S_k$ are jointly sufficient statistics for the estimation of the parameters $\Theta_1, \ldots, \Theta_m$ if the likelihood function can be factored into two nonnegative functions

$$L(\underline{x}|\underline{\theta}) = g(\underline{s}|\underline{\theta})h(\underline{x}) \tag{5}$$

where $g(\underline{s}|\underline{\theta})$ is a function of $\underline{s}=(s_1, \ldots, s_k)$ and $\underline{\theta}$ and $h(\underline{x})$ is independent of $\underline{\theta}$.

Assume the likelihood function can be factored as in (5) then

$$\ln L(\underline{x}|\underline{\theta}) = \ln g(\underline{s}|\underline{\theta}) + \ln h(\underline{x}) \tag{6}$$

Taking partial derivatives of (6) with respect to Θ_i and calculating the system of equations defined in (3) we note that the partial derivative of $\ln h(\underline{x})$ with respect to Θ_i is 0. So the system in (3) contains only $\underline{s}$ and $\underline{\theta}$. Therefore all the information relevant to $\underline{\theta}$ is contained in the statistics $s_1, \ldots, s_k$. The number of sufficient statistics is a determining factor in establishing uniqueness of a solution to (3).

Theorem 2.2 Suppose the log likelihood function has the form $\ln L(\underline{x}|\underline{\theta}) = B(\underline{\theta}) + h(\underline{x}) + \sum_{j=1}^{k} Q_j(\underline{\theta})R_j(\underline{x})$ where $\underline{\theta}=(\Theta_1, \ldots, \Theta_m)\varepsilon\Omega\subseteq\mathbb{R}^m$. Let Q° denote the interior of the range of the 1-1 transformation $c:\underline{\theta} \to Q_1, \ldots, Q_k$. If there exists a solution, $\hat{\underline{\theta}}$, to (3) such that $c(\hat{\underline{\theta}})\varepsilon Q^\circ$ and the regularity conditions hold, and the number of sufficient statistics k equals m, then $\hat{\underline{\theta}}$ is the unique maximum likelihood estimate.

The proof, found in Kendall and Stuart [9], consists of showing that the matrix $N(\hat{\theta})$ of (4) is negative definite if $\hat{\theta}$ is a solution to (3). However, under the regularity conditions, there must be a minimum (or saddle point) between any two maxima. Since $N(\hat{\theta})$ is negative definite, there can be only one maximum. Therefore, in this case joint sufficiency ensures that the likelihood equations have a unique solution and that this solution is a maximum of the likelihood function.

If there are more sufficient statistics than unknown parameters Theorem 2.2 does not guarantee the existence of a unique solution to the likelihood equations. However, under very general conditions the joint maximum likelihood estimators converge in probability, as a set, to the true value of the parameter $\underline{\theta}_o$, (see Kendall and Stuart [9]). Therefore, given two maximum likelihood estimators $\hat{\underline{\theta}}_1$ and $\hat{\underline{\theta}}_2$, then they are essentially the same in the sense that the difference $(\hat{\underline{\theta}}_1 - \hat{\underline{\theta}}_2)$ converges in probability to 0.

Properties of Maximum Likelihood Estimators

Assuming the regularity conditions on the joint distribution of the sample, the maximum likelihood estimators have several desirable statistical properties [9].

Property 1 As a function of the sample size, n, the maximum likelihood estimators for the parameters $\theta_1, \ldots, \theta_m$ are asymptotically unbiased. The maximum likelihood estimators converge in probability to the population parameters $\theta_{10}, \ldots, \theta_{mo}$.

Unbiased estimators are frequently compared using the

magnitude of their variances.

Definition Among estimators that are asymptotically
unbiased, an estimator with minimum variance in large samples
is called an efficient estimator, or simply efficient.

Property 2 The maximum likelihood estimators for
$\Theta_1,\ldots,\Theta_m$ are efficient estimators.

Property 3 The maximum likelihood estimators are
asymptotically multivariate normally distributed with
variance-covariance matrix Σ equal to the inverse of the
information matrix where

$$\Sigma(\underline{\Theta}) = I(\underline{\Theta})^{-1} = [-E \{ \frac{\partial^2 \ln L(\underline{x}|\underline{\Theta})}{\partial\Theta_i \partial\Theta_j} \}]^{-1} \qquad i,j=1,\ldots,m$$

Often we are interested in estimating functions of the
parameters $\Theta_1,\ldots,\Theta_m$, not just the parameters themselves.
We have the following theorem [20].

Theorem 2.3 If on the basis of a given sample $\hat{\Theta}_1,\ldots,\hat{\Theta}_m$
are maximum likelihood estimates of the parameters $\Theta_1,\ldots,\Theta_m$,
then $g_1(\hat{\Theta}_1,\ldots,\hat{\Theta}_m),\ldots,g_m(\hat{\Theta}_1,\ldots,\hat{\Theta}_m)$ are maximum likelihood
estimates of $g_1(\Theta_1,\ldots,\Theta_m),\ldots,g_m(\Theta_1,\ldots,\Theta_m)$ if the
transformation from $\Theta_1,\ldots,\Theta_m$ to $g_1,\ldots,g_m$ is one-to-one.
If the estimates of $\Theta_1,\ldots,\Theta_m$ are unique, then the
estimates of $g_1,\ldots,g_m$ are unique.

Given the more general transformation $\Theta_1,\ldots,\Theta_m \rightarrow g_1,\ldots,g_k$
where the g_i's are continuous we use the notion of the previous
theorem to find estimates of $g_1,\ldots,g_k$. We define the maximum
likelihood estimates of $g_1,\ldots,g_k$ to be $g_1(\hat{\Theta}_1,\ldots,\hat{\Theta}_m),\ldots,$
$g_k(\hat{\Theta}_1,\ldots,\hat{\Theta}_m)$. The variance-covariance matrix of the maximum
likelihood estimator $\underline{\hat{Y}}=(\hat{g}_1,\ldots,\hat{g}_k)$ can now be approximated.

<u>Theorem 2.4</u> Let $\underline{\hat{X}}^t=(\hat{\Theta}_1,\ldots,\hat{\Theta}_m)$ be a vector of maximum likelihood estimators for the parameters $\Theta_1,\ldots,\Theta_m$ with variance-covariance matrix Σ. Let $\underline{Y}^t=(g_1,\ldots,g_k)$ be a vector of functions of $\Theta_1,\ldots,\Theta_m$ such that the partial derivatives of g_i , $i=1,\ldots,k$, exist. Let $\underline{\hat{Y}}^t=(\hat{g}_1,\ldots,\hat{g}_k)$ be the vector of maximum likelihood estimators for the functions $g_1,\ldots,g_k$. Then the variance-covariance matrix for $\underline{\hat{Y}}$ can be approximated by

$$\Sigma_{\underline{\hat{Y}}} \approx \frac{dg}{d\Theta} \Sigma \frac{dg}{d\Theta}^t$$

where the matrix

$$\frac{dg}{d\Theta} = \begin{pmatrix} \dfrac{\partial g_1}{\partial \Theta_1} & \cdots & \dfrac{\partial g_1}{\partial \Theta_m} \\ \vdots & & \\ \dfrac{\partial g_k}{\partial \Theta_1} & \cdots & \dfrac{\partial g_k}{\partial \Theta_m} \end{pmatrix}$$

<u>Proof</u> Bury [10].

The Maximum Likelihood Estimators

In this section we apply the general concepts about maximum likelihood estimators to the specific mathematical model characterizing type A and type B experiments. We will assume the following;

1) the number of promoted tumors per animal, M, has a Poisson distribution with parameter λ, $(\lambda \geq 0)$. Therefore the pmf for the random variable M is

$$P(M=m) = p(m;\lambda) = \frac{e^{-\lambda}\lambda^m}{m!} \qquad m=0,1,2,\ldots$$

2) the time to detection, T, of a randomly selected tumor has a Gamma distribution with parameters α and β ($\alpha>0$, $\beta>0$). Therefore the pdf for T is

$$f(t;\alpha,\beta) = \begin{cases} \dfrac{1}{\beta^\alpha\Gamma(\alpha)} \, t^{\alpha-1} \, e^{-t/\beta} & t\geq 0 \\[2ex] 0 & \text{otherwise} \end{cases}$$

where $\Gamma(\alpha)$ is the Gamma function evaluated at α,

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} \, e^{-x} \, dx$$

and the cdf for T is given by

$$F(t;\alpha,\beta) = \int_0^t \frac{1}{\beta^\alpha\Gamma(\alpha)} \, y^{\alpha-1} \, e^{-y/\beta} \, dy$$

$$= \int_0^{t/\beta} \frac{w^{\alpha-1}e^{-w}}{\Gamma(\alpha)} \, dw$$

which is the incomplete Gamma function evaluated at t/β,

3) in type A experiments all animals survive until the end of the experiment t*,

4) tumor promotion is independent,

5) tumors grow independently.

Type A Experiments

We first consider the estimation of the mean number of promoted tumors and the detection time parameters which are related to the rate of tumor development in experiments of type A. By Theorem 1.2, J(t*) has a Poisson distribution

with parameter $\lambda F(t^*;\alpha,\beta)$. Therefore, the probability of observing m tumors in a randomly selected animal, given the length of the experiment is t^*, is given by

$$p(m|t^*) = \frac{e^{-\lambda F(t^*;\alpha,\beta)}(\lambda F(t^*;\alpha,\beta))^m}{m!} \quad .$$

Now we consider the likelihood function for an entire treatment group. Let m_i, $i=1,\ldots,n$ be the number of observed tumors in animal i. Let t_{ij}, $j=1,\ldots,m_i$ be the time to detection of tumor j in animal i. The likelihood of observing m_i tumors in animal i with tumor times $t_{i1},\ldots,t_{im_i}$ can be derived as follows.

$$L_i(\lambda,\alpha,\beta) = P(M=m_i)\cdot f(t_{i1}|t^*)\cdots f(t_{im_i}|t^*)$$

$$= \frac{e^{-\lambda F(t^*;\alpha,\beta)}(\lambda F(t^*;\alpha,\beta))^{m_i}}{m_i!} \cdot \frac{\prod\limits_{j=1}^{m_i} \frac{1}{\beta^\alpha \Gamma(\alpha)} t_{ij}^{\alpha-1} e^{-t_{ij}/\beta}}{F(t^*;\alpha,\beta)^{m_i}}$$

The likelihood function for the entire treatment group is calculated by multiplying the individual animal likelihood functions together. Therefore

$$L(\lambda,\alpha,\beta) = \prod_{i=1}^{n} L_i(\lambda,\alpha,\beta)$$

$$= \prod_{i=1}^{n} \frac{e^{-\lambda F(t^*;\alpha,\beta)}(\lambda F(t^*;\alpha,\beta))^{m_i}}{m_i!} \cdot \frac{\prod\limits_{j=1}^{m_i} \frac{1}{\beta^\alpha \Gamma(\alpha)} t_{ij}^{\alpha-1} e^{-t_{ij}/\beta}}{F(t^*;\alpha,\beta)^{m_i}}$$

Let $K=m_1!\cdots m_n!$, then $L(\lambda,\alpha,\beta)$ simplifies to

$$L(\lambda,\alpha,\beta) = \frac{e^{-n\lambda F(t^*;\alpha,\beta)} \lambda^{(\sum\limits_{i=1}^{n} m_i)}}{K(\beta^{\alpha}\Gamma(\alpha))^{(\sum\limits_{i=1}^{n} m_i)}} \prod\limits_{i=1}^{n} \prod\limits_{j=1}^{m_i} t_{ij}^{\alpha-1} e^{-t_{ij}/\beta} \qquad (7)$$

We seek estimates $\hat{\lambda},\hat{\alpha},\hat{\beta}$ that maximize the likelihood

function $L(\lambda,\alpha,\beta)$. In this case, it is simplier to work with

$\ln L(\lambda,\alpha,\beta)$ rather than $L(\lambda,\alpha,\beta)$ itself. Therefore, taking

the natural logarithm of both sides of (7) we have

$$\ln L(\lambda,\alpha,\beta) = -n\lambda F(t^*;\alpha,\beta) + (\sum\limits_{i=1}^{n} m_i)\ln \lambda - \ln K$$

$$-(\sum\limits_{i=1}^{n} m_i)\alpha\ln \beta - (\sum\limits_{i=1}^{n} m_i) \ln \Gamma(\alpha)$$

$$+(\alpha-1)\sum\limits_{i=1}^{n} \sum\limits_{j=1}^{m_i} \ln t_{ij} - \frac{1}{\beta}\sum\limits_{i=1}^{n} \sum\limits_{j=1}^{m_i} t_{ij} \qquad (8)$$

We will refer to equation (8) as the likelihood function for

type A experiments. Before solving for $\hat{\lambda},\hat{\alpha},\hat{\beta}$ the following

theorem guarantees there is a unique maximum for (8).

Theorem 2.5 Given a likelihood function of the form in

(8) for a group of animals in a type A experiment, there

exists a unique solution to the likelihood equations and

this solution yields a maximum for the likelihood function.

Proof We seek a maximum likelihood estimate $(\hat{\lambda},\hat{\alpha},\hat{\beta})\epsilon\Omega=$

$\{(\lambda,\alpha,\beta)|\lambda\geq0, \alpha>0, \beta>0\}$. The regularity conditions hold, so

left to show is

1) the presence of at least one maximum at some point in Ω,

2) the sufficiency of three statistics.

To make certain a maximum exists on the set Ω, we note the following. Holding all other parameters constant $L(\lambda,\alpha,\beta) \to 0$ as $\lambda \to \infty$, or as $\alpha \to \infty$, or as $\beta \to \infty$, and as $\lambda \to 0$ ($\sum_{i=1}^{n} m_i \neq 0$), or as $\alpha \to 0$, or as $\beta \to 0$. Thus $L(\lambda,\alpha,\beta)$ is small outside some compact subset of Ω and therefore a maximum exists. Now let

$$s_1 = \sum_{i=1}^{n} m_i \ , \quad s_2 = \sum_{i=1}^{n} \sum_{j=1}^{m_i} \ln t_{ij} \ , \ \text{and} \ s_3 = \sum_{i=1}^{n} \sum_{j=1}^{m_i} t_{ij} \ .$$

Then $\ln L(\lambda,\alpha,\beta)$ in (8) can be written as

$$\ln L(\lambda,\alpha,\beta) = -n\lambda F(t^*;\alpha,\beta) + s_1(\ln \lambda - \alpha \ln \beta - \ln \Gamma(\alpha))$$

$$+ s_2(\alpha-1) - \frac{s_3}{\beta} - \ln K \tag{9}$$

Equation (9) is in the required form and the number of sufficient statistics is equal to the number of unknown parameters. Thus by Theorem 2.2 a unique maximum exists for the likelihood function.

Next, we calculate the system of likelihood equations which must be solved simultaneously to yield the maximum likelihood estimates $\hat{\lambda}, \hat{\alpha}, \hat{\beta}$.

$$\frac{\partial \ln L(\lambda,\alpha,\beta)}{\partial \lambda} = -nF(t^*;\alpha,\beta) + s_1/\lambda = 0 \tag{10}$$

Therefore $\hat{\lambda} = \dfrac{s_1}{nF(t^*;\alpha,\beta)}$ \hfill (11)

$$\frac{\partial \ln L(\lambda,\alpha,\beta)}{\partial \alpha} = -n\lambda \frac{\partial F(t^*;\alpha,\beta)}{\partial \alpha} - s_1 \ln \beta - s_1 \psi(\alpha) + s_2 = 0 \tag{12}$$

(where $\psi(\alpha) = \dfrac{d}{d\alpha} \ln \Gamma(\alpha)$ is the digamma function)

$$\frac{\partial \ln L(\lambda, \alpha, \beta)}{\partial \beta} = -n\lambda \frac{\partial F(t^*; \alpha, \beta)}{\partial \beta} - \frac{\alpha s_1}{\beta} + \frac{s_3}{\beta^2} = 0 \qquad (13)$$

Substituting equation (11) into equations (12) and (13) and simplifying, we have

$$\frac{-s_1}{F(t^*; \alpha, \beta)} \frac{\partial F(t^*; \alpha, \beta)}{\partial \alpha} - s_1 (\ln \beta + \psi(\alpha)) + s_2 = 0 \qquad (14)$$

and

$$\frac{-\beta^2 s_1}{F(t^*; \alpha, \beta)} \frac{\partial F(t^*; \alpha, \beta)}{\partial \beta} - \alpha \beta s_1 + s_3 = 0 \qquad (15)$$

Equations (14) and (15) contain only the parameters α and β, but cannot be solved explicitly. Due to the complexity of this system, a computer algorithm was developed for their solution. The IMSL subroutine ZSPOW [11], which solves a system of non-linear equations, was used. Once the estimates $\hat{\alpha}$ and $\hat{\beta}$ have been calculated, the maximum likelihood estimate $\hat{\lambda}$ is found by substituting $\hat{\alpha}$ and $\hat{\beta}$ into equation (11).

Type B Experiments

In experiments of type B, only the number of observed tumors per animal and an upper bound on their detection times is available for analysis. At any time during the experiment an animal must be sacrificed in order to determine its tumor count. Therefore, the likelihood function for type B experiments is slightly different.

Let $J(t_i)$ be the number of observed tumors per animal at time t_i. We again assume that $J(t_i)$ has a Poisson distribution with parameter $\lambda F(t_i; \alpha, \beta)$. Thus the probability

of observing m tumors in a randomly selected animal at time t_i is given by

$$p(m|t_i) = \frac{e^{-\lambda F(t_i;\alpha,\beta)}(\lambda F(t_i;\alpha,\beta))^m}{m!}$$

We consider the likelihood function for an entire treatment group. Let r_i be the (non-random) number of animals sacrificed ar time t_i , i=1,...,k. Let m_{ij} , j=1,...,r_i be the number of observed (or estimated) tumors in rat j sacrificed at time t_i. The likelihood of observing m_{ij} tumors in animals sacrificed at time t_i is given by

$$L_i(\lambda,\alpha,\beta) = P(M_{i1}=m_{i1},\ldots,M_{ir_i}=m_{ir_i}|T=t_i)$$

$$= \prod_{j=1}^{r_i} \frac{e^{-\lambda F(t_i;\alpha,\beta)}(\lambda F(t_i;\alpha,\beta))^{m_{ij}}}{m_{ij}!}$$

The likelihood function for the entire treatment group is the product of each likelihood function corresponding to a different sacrifice time. Thus,

$$L(\lambda,\alpha,\beta) = \prod_{i=1}^{k} L_i(\lambda,\alpha,\beta)$$

$$= \prod_{i=1}^{k} \prod_{j=1}^{r_i} \frac{e^{-\lambda F(t_i;\alpha,\beta)}(\lambda F(t_i;\alpha,\beta))^{m_{ij}}}{m_{ij}!} \tag{16}$$

Once again we seek estimates $\hat{\lambda},\hat{\alpha},\hat{\beta}$ that maximize the likelihood function $L(\lambda,\alpha,\beta)$ and find it is simplier to work with ln $L(\lambda,\alpha,\beta)$ rather than $L(\lambda,\alpha,\beta)$. Taking the natural logarithm of both sides of equation (16) and simplifying

we have

$$\ln L(\lambda, \alpha, \beta) = \ln \lambda \sum_{i=1}^{k} \sum_{j=1}^{r_i} m_{ij} + \sum_{i=1}^{k} \sum_{j=1}^{r_i} m_{ij} \ln F(t_i; \alpha, \beta)$$

$$- \lambda \sum_{i=1}^{k} r_i F(t_i; \alpha, \beta) - \sum_{i=1}^{k} \sum_{j=1}^{r_i} \ln(m_{ij}!) \qquad (17)$$

We will refer to equation (17) as the likelihood function for a type B experiment.

We consider whether the likelihood function for a type B experiment has a unique maximum. Let the statistics $s, s_1, \ldots, s_k$ be defined by

$$s_i = \sum_{j=1}^{r_i} m_{ij} \qquad \text{and} \qquad s = \sum_{i=1}^{k} s_i$$

Then the likelihood function can be written as

$$\ln L(\lambda, \alpha, \beta) = s \ln \lambda + \sum_{i=1}^{k} s_i \ln F(t_i; \alpha, \beta) - \lambda \sum_{i=1}^{k} r_i F(t_i; \alpha, \beta)$$

$$- \sum_{i=1}^{k} \sum_{j=1}^{r_i} \ln(m_{ij}!)$$

The likelihood function satisfies the regularity conditions and is decreasing (assume $s \neq 0$) outside some compact subset of $\Omega = \{(\lambda, \alpha, \beta) \mid \lambda \geq 0, \ \alpha > 0, \ \beta > 0\}$. Thus, a maximum likelihood estimate is a solution to the system of equations defined in (3) and more than one maximum likelihood estimate may exist. However, in experiments of type B, any maximum likelihood estimator converges in probability to the true value of the parameter.

Thus, any two maximum likelihood estimators are essentially the same in the sense that their difference converges in probability to zero. We assume n, the number of animals in the group, is large and we seek the maximum for the likelihood function in experiments of type B.

We derive the system of likelihood equations as defined in (3) that must be solved simultaneously to yield the maximum likelihood estimates $\hat{\lambda}, \hat{\alpha}, \hat{\beta}$.

$$\frac{\partial \ln L(\lambda, \alpha, \beta)}{\partial \lambda} = \frac{s}{\lambda} - \sum_{i=1}^{k} r_i F(t_i; \alpha, \beta) = 0 \qquad (18)$$

so that

$$\hat{\lambda} = \frac{s}{\sum_{i=1}^{k} r_i F(t_i; \alpha, \beta)} \qquad (19)$$

$$\frac{\partial \ln L(\lambda, \alpha, \beta)}{\partial \alpha} = \sum_{i=1}^{k} s_i \frac{\partial F(t_i; \alpha, \beta)}{\partial \alpha} \frac{1}{F(t_i; \alpha, \beta)}$$

$$- \lambda \sum_{i=1}^{k} r_i \frac{\partial F(t_i; \alpha, \beta)}{\partial \alpha} = 0 \qquad (20)$$

$$\frac{\partial \ln L(\lambda, \alpha, \beta)}{\partial \beta} = \sum_{i=1}^{k} s_i \frac{\partial F(t_i; \alpha, \beta)}{\partial \beta} \frac{1}{F(t_i; \alpha, \beta)}$$

$$- \lambda \sum_{i=1}^{k} r_i \frac{\partial F(t_i; \alpha, \beta)}{\partial \beta} = 0 \qquad (21)$$

Substituting equation (19) into equations (20) and (21), and simplfying we have

$$\sum_{i=1}^{k} \left(\frac{s_i}{F(t_i;\alpha,\beta)} - \frac{r_i s}{\sum_{j=1}^{k} r_j F(t_j;\alpha,\beta)} \right) \frac{\partial F(t_i;\alpha,\beta)}{\partial \alpha} = 0 \qquad (22)$$

$$\sum_{i=1}^{k} \left(\frac{s_i}{F(t_i;\alpha,\beta)} - \frac{r_i s}{\sum_{j=1}^{k} r_j F(t_j;\alpha,\beta)} \right) \frac{\partial F(t_i;\alpha,\beta)}{\partial \beta} = 0 \qquad (23)$$

Equations (22) and (23) contain only the parameters α and β, however they cannot be solved explicitly. By using a computer supplied algorithm, $\hat{\alpha}$ and $\hat{\beta}$ can be calculated, then $\hat{\lambda}$ is found by substituting $\hat{\alpha}$ and $\hat{\beta}$ into equation (19).

Properties of the Estimators

In both type A and type B systems, the experimenter is interested in how a chemopreventative agent affects the number of promoted tumors per animal and the tumor growth rate. Therefore, he would like to estimate not only the number of promoted tumors per animal but also $\mu=\alpha\beta$, the mean of the time to detection distribution. A larger mean time to detection is believed to indicate a slower tumor growth rate. By applying Theorem 2.3, using the transformation $\lambda,\alpha,\beta \rightarrow \lambda,\alpha\beta,\alpha\beta^2$, the maximum likelihood estimate for μ is $\hat{\mu}=\hat{\alpha}\hat{\beta}$ and $\hat{\sigma}^2=\hat{\alpha}\hat{\beta}^2$ estimates σ^2, the variance of T. The maximum likelihood estimators $\hat{\lambda}$, $\hat{\mu}$, and $\hat{\sigma}^2$ are asymptotically multivariate normal, asymptotically unbiased, and efficient.

The estimation procedure, for experiments of type A and B, described above simultaneously considers the effects of

tumor number and detection time. This procedure addresses the confounding problem as illustrated in Figure 1, as shown in the derivation of the mean and variance of J(t), and as shown in the calculation of the incidence rate. Instead of considering only tumor number and estimating λ, then tumor times and estimating μ, this estimation procedure provides the experimenter with a more complete description of a given treatment group.

$\hat{\lambda}$ is an estimate of the mean number of promoted tumors per animal; the total number of tumors we would expect to see if the experiment were allowed to continue to infinity. Equation (11), the maximum likelihood estimator for λ in experiments of type A, further illustrates the confounding of tumor number and detection time parameters and the effect each has on $\hat{\lambda}$. From equation (11) we see that $\hat{\lambda}$ is a function of s_1, the total number of observed tumors at time t^*, and $F(t^*;\alpha,\beta)$, the cdf for the time to detection distribution. As $t^* \to \infty$, then $F(t^*;\alpha,\beta) \to 1$ so that $\hat{\lambda} \to s_1/n = \bar{m}$, the sample mean of the observed tumors per animal. Since most experiments are terminated before all promoted tumors are observed, the point estimate for $F(t^*;\alpha,\beta)$, $F(t^*;\hat{\alpha},\hat{\beta})$, is generally less than 1 which causes the estimate for λ to be larger than $\bar{m}$. Thus, the likelihood equations for experiments of type A illustrate the need to consider tumor number and detection time parameters simultaneously.

In experiments of type B, the maximum likelihood estimate for λ also illustrates the confounding problem. $\hat{\lambda}$ is a function of s, the total number of observed tumors during the entire experiment, and $F(t_i;\alpha,\beta)$, the cdf for the time to detection distribution at times $t_1, t_2, \ldots, t_k$. To demonstrate the effect

each has on $\hat{\lambda}$, we note the following;

1) if more(less) tumors are observed per animal at each time
   t_i, each s_i increases(decreases), thus s increases(decreases)
   and $\hat{\lambda}$ increases(decreases),

2) if the rate of tumor development decreases(increases),
   $F(t_i;\hat{\alpha},\hat{\beta})$ is less(greater) for each t_i, therefore $\hat{\lambda}$
   increases(decreases).

In the remainder of this study we restrict our analysis to
the two parameters λ, and $\mu=\alpha\beta$, the mean of the time to detection
distribution. We focus on μ because, in contrast to α and β,
it has a direct biological interpretation and with λ, is one of
the two prime criteria for treatment comparisons in those
systems with which this study is most concerned. The variance,
$\sigma^2=\alpha\beta^2$, provides a natural second parameter for the description
of $\Gamma(\alpha,\beta)$. However, it is not commonly used for treatment
comparisons in these experiments. The biological significance
of changes in σ^2 across treatments is not obvious.

It would be desirable to conduct the analysis in terms of
the two common parameters μ and σ^2. Direct reparametrization,
however, creates complications and inefficiencies in the
estimation routines which the natural parametrization avoids.
Therefore, we elect to base numerical estimation on λ, α, and
β, to estimate $\mu=\alpha\beta$, and to define treatment differences in
terms of λ and μ. We omit further consideration of σ^2 because
1) it is not of common interest, 2) it has a large sampling
variance, and 3) treatments which differ significantly in
terms of σ^2 would be inconclusive.

CHAPTER III

ONE SAMPLE PROBLEM - CONFIDENCE REGIONS AND HYPOTHESIS TESTS

Introduction

In type A and type B experimental tumor systems, we are most interested in estimating the parameter λ, the mean number of promoted tumors, and the rate of tumor progression as manifest in μ, the mean time to tumor detection. In this chapter we seek a $100(1-\alpha')\%$ confidence region for the actual value $\underline{\mu}^t=(\lambda,\mu)$. We continue to assume that detection times are described by the $\Gamma(\alpha,\beta)$ distribution and so $\mu=\alpha\beta$. A hypothesis test is also described for testing the simple hypothesis H_o: $\underline{\mu}=\underline{\mu}_o$ versus the alternate hypothesis H_a: $\underline{\mu}\neq\underline{\mu}_o$. We assume n (the number of animals in the group) is large and we use the fact that under regularity conditions, the estimators $\underline{\hat{Y}}^t=(\hat{\lambda},\hat{\alpha},\hat{\beta})$ and $\underline{\hat{X}}^t=(\hat{\lambda},\hat{\mu})$ converge in probability to (λ,α,β) and (λ,μ) respectively.

Type A Experiments - Variance-covariance Matrix

In order to construct a confidence region for (λ,μ) and/or conduct the hypothesis described above, we need the variance-covariance matrix for the estimator $\underline{\hat{X}}$. We must first calculate the second partial derivatives of the log likelihood function $\ln L(\lambda,\alpha,\beta)$, the expected value of each second partial derivative with respect to $\underline{\theta}$, and the information matrix for $\underline{\hat{Y}}$. We will need the following lemma.

30

<u>Lemma 3.1</u> Let X be an integer valued random variable with mean μ_1. Let $Y_1, Y_2, Y_3, \ldots$ be independent identically distributed (iid) random variables from a population with mean μ_2. Assume X and $Y_1, Y_2, Y_3, \ldots$ are independent. Then

$$E\left[\sum_{j=1}^{X} Y_j\right] = \mu_1\mu_2 \quad .$$

<u>Proof</u>

$$E\left[\sum_{j=1}^{X} Y_j\right] = P(X=0)\cdot 0 + P(X=1)E\left[\sum_{j=1}^{1} Y_j\right] + P(X=2)E\left[\sum_{j=1}^{2} Y_j\right] + \ldots$$

$$= P(X=0)\cdot 0 + P(X=1)E(Y) + P(X=2)2E(Y) + \ldots$$

$$= E(Y)[P(X=0)\cdot 0 + P(X=1)\cdot 1 + P(X=2)\cdot 2 + \ldots]$$

$$= E(Y)E(X)$$

$$= \mu_1\mu_2$$

The following corollary deals with the statistic S_3 that appears in the likelihood function for type A experiments.

<u>Corollary 3.2</u> Assume a type A experiment. Let $M_1, \ldots, M_n$ be a random sample from a Poisson distribution with parameter $\lambda F(t^*; \alpha, \beta)$. Let T_{ij}, $i=1, \ldots, n$, $j=1, 2, 3, \ldots$ be iid random variables from a Gamma distribution with parameters α and β. Assume M_i, $i=1, \ldots, n$, and T_{ij} are independent and α and β are known. Define the statistic

$$S_3 = \sum_{i=1}^{n} \sum_{j=1}^{M_i} T_{ij}$$

then $E(S_3) = n\lambda F(t^*; \alpha, \beta)\alpha\beta \quad .$

Proof

$$E(S_3) = \sum_{i=1}^{n} E[\sum_{j=1}^{M_i} T_{ij}] = \sum_{i=1}^{n} E(M_i)E(T_{ij})$$

$$= n\lambda F(t^*;\alpha,\beta)\alpha\beta$$

__Theorem 3.3__ Assume a type A experiment. If $\lambda,\alpha,$ and β are known, then the asymptotic variance-covariance matrix, $\Sigma_{\hat{\underline{Y}}}$, for the maximum likelihood estimator $\hat{\underline{Y}}^t=(\hat{\lambda},\hat{\alpha},\hat{\beta})$ is

$$\Sigma_{\hat{\underline{Y}}} = \frac{1}{n|I|} V \qquad \text{where}$$

$$|I|=\lambda[F(t^*;\alpha,\beta)((\frac{\partial^2 F(t^*;\alpha,\beta)}{\partial\alpha^2}+\psi'(\alpha)F(t^*;\alpha,\beta))(\frac{\partial^2 F(t^*;\alpha,\beta)}{\partial\beta^2}$$

$$+\frac{\alpha F(t^*;\alpha,\beta)}{\beta^2})+2\frac{\partial F(t^*;\alpha,\beta)}{\partial\alpha}\frac{\partial F(t^*;\alpha,\beta)}{\partial\beta}.$$

$$(\frac{\partial^2 F(t^*;\alpha,\beta)}{\partial\alpha\partial\beta}+\frac{F(t^*;\alpha,\beta)}{\beta}))$$

$$-(\frac{\partial F(t^*;\alpha,\beta)}{\partial\beta})^2(\frac{\partial^2 F(t^*;\alpha,\beta)}{\partial\alpha^2}+\psi'(\alpha)F(t^*;\alpha,\beta))$$

$$+\frac{\partial F(t^*;\alpha,\beta)}{\partial\alpha}^2(\frac{\partial^2 F(t^*;\alpha,\beta)}{\partial\beta^2}+\frac{\alpha F(t^*;\alpha,\beta)}{\beta^2})$$

$$+F(t^*;\alpha,\beta)(\frac{\partial^2 F(t^*;\alpha,\beta)}{\partial\alpha\partial\beta}+\frac{F(t^*;\alpha,\beta)}{\beta})^2)]$$

and the entries v_{ij} $i,j=1,2,3$ of the matrix V are

$$v_{11}=\lambda^2[(\frac{\partial^2 F(t^*;\alpha,\beta)}{\partial\alpha^2}+\psi'(\alpha)F(t^*;\alpha,\beta))(\frac{\partial^2 F(t^*;\alpha,\beta)}{\partial\beta^2}+\frac{\alpha F(t^*;\alpha,\beta)}{\beta^2})$$

$$-(\frac{\partial^2 F(t^*;\alpha,\beta)}{\partial\alpha\partial\beta}+\frac{F(t^*;\alpha,\beta)}{\beta})^2]$$

$$v_{12}=v_{21}=-\lambda\left[\frac{\partial F(t^*;\alpha,\beta)}{\partial\alpha}\left(\frac{\partial^2 F(t^*;\alpha,\beta)}{\partial\beta^2}+\frac{\alpha F(t^*;\alpha,\beta)}{\beta^2}\right)\right.$$

$$\left.-\frac{\partial F(t^*;\alpha,\beta)}{\partial\beta}\left(\frac{\partial^2 F(t^*;\alpha,\beta)}{\partial\alpha\partial\beta}+\frac{F(t^*;\alpha,\beta)}{\beta}\right)\right]$$

$$v_{13}=v_{31}=\lambda\left[\frac{\partial F(t^*;\alpha,\beta)}{\partial\alpha}\left(\frac{\partial^2 F(t^*;\alpha,\beta)}{\partial\alpha\partial\beta}+\frac{F(t^*;\alpha,\beta)}{\beta}\right)\right.$$

$$\left.-\frac{\partial F(t^*;\alpha,\beta)}{\partial\beta}\left(\frac{\partial^2 F(t^*;\alpha,\beta)}{\partial\alpha^2}+\psi'(\alpha)F(t^*;\alpha,\beta)\right)\right]$$

$$v_{22}=F(t^*;\alpha,\beta)\left(\frac{\partial^2 F(t^*;\alpha,\beta)}{\partial\beta^2}+\frac{\partial F(t^*;\alpha,\beta)}{\beta^2}\right)-\frac{\partial F(t^*;\alpha,\beta)}{\partial\beta}^2$$

$$v_{23}=v_{32}=-(F(t^*;\alpha,\beta)\left(\frac{\partial^2 F(t^*;\alpha,\beta)}{\partial\alpha\partial\beta}+\frac{F(t^*;\alpha,\beta)}{\beta}\right)$$

$$-\frac{\partial F(t^*;\alpha,\beta)}{\partial\alpha}\ \frac{\partial F(t^*;\alpha,\beta)}{\partial\beta})$$

$$v_{33}=F(t^*;\alpha,\beta)\left(\frac{\partial^2 F(t^*;\alpha,\beta)}{\partial\alpha^2}+\psi'(\alpha)F(t^*;\alpha,\beta)-\frac{\partial F(t^*;\alpha,\beta)}{\partial\alpha}^2\right)$$

<u>Proof</u> We proceed by calculating the second partial derivatives of $\ln L(\lambda,\alpha,\beta)$. Let I_{ij} , $i,j=1,2,3$ be the entries in the information matrix.

From $\dfrac{\partial\ln L(\lambda,\alpha,\beta)}{\partial\lambda} = -nF(t^*;\alpha,\beta) + S_1/\lambda$

we find;

i) $\dfrac{\partial^2\ln L(\lambda,\alpha,\beta)}{\partial\lambda^2} = -S_1/\lambda^2$

so that

$I_{11} = (1/\lambda^2)E(S_1)$

$\qquad = (1/\lambda^2)(n\lambda F(t^*;\alpha,\beta)) = nF(t^*;\alpha,\beta)/\lambda$

ii) $\dfrac{\partial^2 \ln L(\lambda,\alpha,\beta)}{\partial\alpha\partial\lambda} = -n\dfrac{\partial F(t^*;\alpha,\beta)}{\partial\alpha}$

so that

$I_{12} = I_{21} = n\dfrac{\partial F(t^*;\alpha,\beta)}{\partial\alpha}$

iii) $\dfrac{\partial^2 \ln L(\lambda,\alpha,\beta)}{\partial\beta\partial\lambda} = -n\dfrac{\partial F(t^*;\alpha,\beta)}{\partial\beta}$

so that

$I_{13} = I_{31} = n\dfrac{\partial F(t^*;\alpha,\beta)}{\partial\beta}$

From $\dfrac{\partial \ln L(\lambda,\alpha,\beta)}{\partial\alpha} = -n\lambda\dfrac{\partial F(t^*;\alpha,\beta)}{\partial\alpha} - S_1(\ln\beta + \psi(\alpha)) + S_2$

we find;

i) $\dfrac{\partial^2 \ln L(\lambda,\alpha,\beta)}{\partial\alpha^2} = -n\lambda\dfrac{\partial^2 F(t^*;\alpha,\beta)}{\partial\alpha^2} - S_1\psi'(\alpha)$

so that

$I_{22} = n\lambda\dfrac{\partial^2 F(t^*;\alpha,\beta)}{\partial\alpha^2} - \psi'(\alpha)E(S_1)$

$\quad\quad = n\lambda\left(\dfrac{\partial^2 F(t^*;\alpha,\beta)}{\partial\alpha^2} - \psi'(\alpha)F(t^*;\alpha,\beta)\right)$

ii) $\dfrac{\partial^2 \ln L(\lambda,\alpha,\beta)}{\partial\beta\partial\alpha} = -n\lambda\dfrac{\partial^2 F(t^*;\alpha,\beta)}{\partial\beta\partial\alpha} - S_1/\beta$

so that

$I_{23} = I_{32} = n\lambda\dfrac{\partial^2 F(t^*;\alpha,\beta)}{\partial\beta\partial\alpha} + E(S_1)/\beta$

$\quad\quad = n\lambda\left(\dfrac{\partial^2 F(t^*;\alpha,\beta)}{\partial\beta\partial\alpha} + \dfrac{F(t^*;\alpha,\beta)}{\beta}\right)$

From $\dfrac{\partial \ln L(\lambda,\alpha,\beta)}{\partial\beta} = -n\lambda\dfrac{\partial F(t^*;\alpha,\beta)}{\partial\beta} - \dfrac{\alpha}{\beta}S_1 + \dfrac{1}{\beta^2}S_3$

we find;

i) $\dfrac{\partial^2 \ln L(\lambda,\alpha,\beta)}{\partial \beta^2} = -n\lambda \dfrac{\partial^2 F(t^*;\alpha,\beta)}{\partial \beta^2} + \dfrac{\alpha}{\beta^2}S_1 - \dfrac{2}{\beta^3}S_3$

so that

$$I_{33} = n\lambda \dfrac{\partial^2 F(t^*;\alpha,\beta)}{\partial \beta^2} - \dfrac{\alpha}{\beta^2}E(S_1) + \dfrac{2}{\beta^3}E(S_3)$$

$$= n\lambda(\dfrac{\partial^2 F(t^*;\alpha,\beta)}{\partial \beta^2} + \dfrac{\alpha F(t^*;\alpha,\beta)}{\beta^2})$$

The variance-covariance matrix $\Sigma_{\hat{Y}}$ is the inverse, I^{-1}, of the information matrix. Inverting the matrix I yields the desired result.

Variance-covariance Matrix for $\hat{X}^t = (\hat{\lambda}, \hat{\mu})$

We are interested in a confidence region for the actual value $\underline{X}^t = (\lambda,\mu)$. Since the maximum likelihood estimator $\hat{\underline{X}}^t = (\hat{\lambda}, \hat{\mu}) = (\hat{\lambda}, \hat{\alpha}\hat{\beta})$ is a vector of functions of the parameters in $\hat{\underline{Y}}$, the variance-covariance matrix for $\hat{\underline{X}}$ can be approximated using Theorem 2.4. $\hat{\underline{X}}$ also tends to a multivariate normal distribution and is an asymptotically unbiased estimator for the actual value (λ,μ).

Theorem 3.4 Assume a type A experiment. If λ,α,β are known, then the approximate variance-covariance matrix $\Sigma_{\hat{\underline{X}}}$ for the maximum likelihood estimator $\hat{\underline{X}}^t = (\hat{\lambda}, \hat{\mu})$ is

$$\Sigma_{\hat{\underline{X}}} \approx \dfrac{1}{n|I|} G$$

where $|I|$ is the determinant of the information matrix for $\hat{\underline{Y}}$, v_{ij} are defined in Theorem 3.3, and the matrix G has as entries g_{ij}, $i,j=1,2$ given by

$$g_{11} = v_{11} \qquad g_{22} = \beta^2 v_{22} + 2\alpha\beta v_{23} + \alpha^2 v_{33}$$

$$g_{12} = g_{21} = \beta v_{12} + \alpha v_{13}$$

<u>Proof</u> We are given that the variance-covariance matrix for $\hat{\underline{Y}}$ is $\Sigma_{\hat{\underline{Y}}} = \frac{1}{n|I|} V$. $\hat{\underline{X}}$ is a maximum likelihood estimator for

$$\underline{x}^t = (h_1(\lambda,\alpha,\beta), h_2(\lambda,\alpha,\beta)) = (\lambda, \alpha\beta) = (\lambda, \mu).$$

Define the matrix H by

$$H = \begin{pmatrix} \dfrac{\partial h_1(\lambda,\alpha,\beta)}{\partial \lambda} & \dfrac{\partial h_1(\lambda,\alpha,\beta)}{\partial \alpha} & \dfrac{\partial h_1(\lambda,\alpha,\beta)}{\partial \beta} \\[2mm] \dfrac{\partial h_2(\lambda,\alpha,\beta)}{\partial \lambda} & \dfrac{\partial h_2(\lambda,\alpha,\beta)}{\partial \alpha} & \dfrac{\partial h_2(\lambda,\alpha,\beta)}{\partial \beta} \end{pmatrix}$$

so that

$$H = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \beta & \alpha \end{pmatrix}$$

Then by Theorem 2.4 the variance-covariance matrix $\Sigma_{\hat{\underline{X}}}$ can be approximated by

$$\Sigma_{\hat{\underline{X}}} \approx H \, \Sigma_{\hat{\underline{Y}}} \, H^t$$

$$\approx \begin{pmatrix} 1 & 0 & 0 \\ 0 & \beta & \alpha \end{pmatrix} \frac{1}{n|I|} V \begin{pmatrix} 1 & 0 \\ 0 & \beta \\ 0 & \alpha \end{pmatrix}$$

$$\approx \frac{1}{n|I|} \begin{pmatrix} v_{11} & \beta v_{21} + \alpha v_{13} \\ \beta v_{21} + \alpha v_{31} & \beta^2 v_{22} + \alpha\beta v_{32} + \alpha\beta v_{23} + \alpha^2 v_{33} \end{pmatrix}$$

Noting the symmetry of the matrix V yields the desired result.

In chapter V we show, as a consequence of this theorem, that the variance of $\hat{\lambda}$ is proportional to λ and the variance

of $\hat{\mu}$ is proportional to $1/\lambda$. This is important in understanding why we get less accurate estimates of μ for small λ.

Type B Experiments - Variance-covariance Matrix

In type B experiments, the variance-covariance matrix $\Sigma_{\hat{\underline{X}}}$, is slightly different. We proceed in the same way by first deriving the information matrix for $\hat{\underline{Y}}$. Recall that the statistic S_i is given by

$$S_i = \sum_{j=1}^{r_i} M_{ij} \ .$$

Then we have $E(S_i) = \lambda r_i F(t_i;\alpha,\beta)$. We can now calculate the variance-covariance matrix for $\hat{\underline{Y}}$ in a type B experiment.

<u>Theorem 3.5</u> Assume a type B experiment. If λ,α,β are known, then the variance-covariance matrix, $\Sigma_{\hat{\underline{Y}}}$, for the maximum likelihood estimator $\hat{\underline{Y}}^t=(\hat{\lambda},\hat{\alpha},\hat{\beta})$ is

$$\Sigma_{\hat{\underline{Y}}} = \frac{1}{|I|} V \qquad \text{where}$$

$$|I| = \lambda \left[\left(\sum_{i=1}^{k} r_i F(t_i;\alpha,\beta) \right) \left(\sum_{i=1}^{k} \frac{r_i}{F(t_i;\alpha,\beta)} \left(\frac{\partial F(t_i;\alpha,\beta)}{\partial \alpha} \right)^2 \right) \right.$$

$$\left(\sum_{i=1}^{k} \frac{r_i}{F(t_i;\alpha,\beta)} \left(\frac{\partial F(t_i;\alpha,\beta)}{\partial \beta} \right)^2 \right)$$

$$+ 2 \left(\sum_{i=1}^{k} r_i \frac{\partial F(t_i;\alpha,\beta)}{\partial \alpha} \right) \left(\sum_{i=1}^{k} r_i \frac{\partial F(t_i;\alpha,\beta)}{\partial \beta} \right)$$

$$\left(\sum_{i=1}^{k} \frac{r_i}{F(t_i;\alpha,\beta)} \frac{\partial F(t_i;\alpha,\beta)}{\partial \alpha} \frac{\partial F(t_i;\alpha,\beta)}{\partial \beta} \right)$$

$$-((\sum_{i=1}^{k} r_i \frac{\partial F(t_i;\alpha,\beta)}{\partial \beta})^2 (\sum_{i=1}^{k} \frac{r_i}{F(t_i;\alpha,\beta)} (\frac{\partial F(t_i;\alpha,\beta)}{\partial \alpha})^2)$$

$$+(\sum_{i=1}^{k} r_i \frac{\partial F(t_i;\alpha,\beta)}{\partial \alpha})^2 (\sum_{i=1}^{k} \frac{r_i}{F(t_i;\alpha,\beta)} (\frac{\partial F(t_i;\alpha,\beta)}{\partial \beta})^2)$$

$$+(\sum_{i=1}^{k} r_i F(t_i;\alpha,\beta))(\sum_{i=1}^{k} \frac{r_i}{F(t_i;\alpha,\beta)} \frac{\partial F(t_i;\alpha,\beta)}{\partial \alpha} \frac{\partial F(t_i;\alpha,\beta)}{\partial \beta})^2))]$$

and the entries v_{ij} , $i,j=1,2,3$ of the matrix V are

$$v_{11}=\lambda^2[(\sum_{i=1}^{k} \frac{r_i}{F(t_i;\alpha,\beta)} (\frac{\partial F(t_i;\alpha,\beta)}{\partial \alpha})^2)(\sum_{i=1}^{k} \frac{r_i}{F(t_i;\alpha,\beta)} (\frac{\partial F(t_i;\alpha,\beta)}{\partial \beta})^2)$$

$$-(\sum_{i=1}^{k} \frac{r_i}{F(t_i;\alpha,\beta)} \frac{\partial F(t_i;\alpha,\beta)}{\partial \alpha} \frac{\partial F(t_i;\alpha,\beta)}{\partial \beta})^2]$$

$$v_{12}=v_{21}=-\lambda[(\sum_{i=1}^{k} r_i \frac{\partial F(t_i;\alpha,\beta)}{\partial \alpha})(\sum_{i=1}^{k} \frac{r_i}{F(t_i;\alpha,\beta)} (\frac{\partial F(t_i;\alpha,\beta)}{\partial \beta})^2)$$

$$-(\sum_{i=1}^{k} \frac{r_i}{F(t_i;\alpha,\beta)} \frac{\partial F(t_i;\alpha,\beta)}{\partial \alpha} \frac{\partial F(t_i;\alpha,\beta)}{\partial \beta})(\sum_{i=1}^{k} r_i \frac{\partial F(t_i;\alpha,\beta)}{\partial \beta})]$$

$$v_{13}=v_{31}=\lambda[(\sum_{i=1}^{k} r_i \frac{\partial F(t_i;\alpha,\beta)}{\partial \alpha})(\sum_{i=1}^{k} \frac{r_i}{F(t_i;\alpha,\beta)} \frac{\partial F(t_i;\alpha,\beta)}{\partial \alpha} \frac{\partial F(t_i;\alpha,\beta)}{\partial \beta})$$

$$-(\sum_{i=1}^{k} \frac{r_i}{F(t_i;\alpha,\beta)} (\frac{\partial F(t_i;\alpha,\beta)}{\partial \alpha})^2)(\sum_{i=1}^{k} r_i \frac{\partial F(t_i;\alpha,\beta)}{\partial \beta})]$$

$$v_{22}=(\sum_{i=1}^{k} r_i F(t_i;\alpha,\beta))(\sum_{i=1}^{k} \frac{r_i}{F(t_i;\alpha,\beta)} \frac{\partial F(t_i;\alpha,\beta)}{\partial \beta})^2)$$

$$-(\sum_{i=1}^{k} r_i \frac{\partial F(t_i;\alpha,\beta)}{\partial \beta})^2$$

$$v_{23}=v_{32}=-[(\sum_{i=1}^{k} r_i F(t_i;\alpha,\beta))(\sum_{i=1}^{k} \frac{r_i}{F(t_i;\alpha,\beta)} \frac{\partial F(t_i;\alpha,\beta)}{\partial\alpha} \frac{\partial F(t_i;\alpha,\beta)}{\partial\beta})$$

$$-(\sum_{i=1}^{k} r_i \frac{\partial F(t_i;\alpha,\beta)}{\partial\alpha})(\sum_{i=1}^{k} r_i \frac{\partial F(t_i;\alpha,\beta)}{\partial\beta})]$$

$$v_{33}=(\sum_{i=1}^{k} r_i F(t_i;\alpha,\beta))(\sum_{i=1}^{k} \frac{r_i}{F(t_i;\alpha,\beta)}(\frac{\partial F(t_i;\alpha,\beta)}{\partial\alpha})^2)$$

$$-(\sum_{i=1}^{k} r_i \frac{\partial F(t_i;\alpha,\beta)}{\partial\alpha})^2$$

<u>Proof</u> We proceed by calculating the second partial derivatives of $\ln L(\lambda,\alpha,\beta)$. Let I_{ij} , $i,j=1,2,3$ be the entries in the information matrix.

From $\frac{\partial \ln L(\lambda,\alpha,\beta)}{\partial\lambda} = (S/\lambda) - \sum_{i=1}^{k} r_i F(t_i;\alpha,\beta)$

we find;

i) $\frac{\partial^2 \ln L(\lambda,\alpha,\beta)}{\partial\lambda^2} = -S/\lambda^2$

so that

$$I_{11} = (1/\lambda^2)E(S) = (1/\lambda^2)(\sum_{i=1}^{k} \lambda r_i F(t_i;\alpha,\beta)$$

$$= (1/\lambda)(\sum_{i=1}^{k} r_i F(t_i;\alpha,\beta))$$

ii) $\frac{\partial^2 \ln L(\lambda,\alpha,\beta)}{\partial\alpha\partial\lambda} = -\sum_{i=1}^{k} r_i \frac{\partial F(t_i;\alpha,\beta)}{\partial\alpha}$

so that

$$I_{12} = I_{21} = \sum_{i=1}^{k} r_i \frac{\partial F(t_i;\alpha,\beta)}{\partial\alpha}$$

iii)
$$\frac{\partial^2 \ln L(\lambda,\alpha,\beta)}{\partial\beta\partial\lambda} = -\sum_{i=1}^{k} r_i \frac{\partial F(t_i;\alpha,\beta)}{\partial\beta}$$

so that

$$I_{13} = I_{31} = \sum_{i=1}^{k} r_i \frac{\partial F(t_i;\alpha,\beta)}{\partial\beta}$$

From $\dfrac{\partial \ln L(\lambda,\alpha,\beta)}{\partial\alpha} = \displaystyle\sum_{i=1}^{k} s_i \frac{\partial F(t_i;\alpha,\beta)}{\partial\alpha}\frac{1}{F(t_i;\alpha,\beta)} - \lambda\sum_{i=1}^{k} r_i \frac{\partial F(t_i;\alpha,\beta)}{\partial\alpha}$

we find;

i)
$$\frac{\partial^2 \ln L(\lambda,\alpha,\beta)}{\partial\alpha^2} = \sum_{i=1}^{k} s_i \left(\frac{\partial^2 F(t_i;\alpha,\beta)}{\partial\alpha^2}\frac{1}{F(t_i;\alpha,\beta)} - \frac{\left(\frac{\partial F(t_i;\alpha,\beta)}{\partial\alpha}\right)^2}{F^2(t_i;\alpha,\beta)}\right) - \lambda\sum_{i=1}^{k} r_i \frac{\partial^2 F(t_i;\alpha,\beta)}{\partial\alpha^2}$$

so that

$$I_{22} = \lambda\sum_{i=1}^{k}\frac{r_i}{F(t_i;\alpha,\beta)}\left(\frac{\partial F(t_i;\alpha,\beta)}{\partial\alpha}\right)^2$$

ii)
$$\frac{\partial^2 \ln L(\lambda,\alpha,\beta)}{\partial\beta\partial\alpha} = \sum_{i=1}^{k} s_i \left(\frac{\partial^2 F(t_i;\alpha,\beta)}{\partial\beta\partial\alpha}\frac{1}{F(t_i;\alpha,\beta)} - \frac{\frac{\partial F(t_i;\alpha,\beta)}{\partial\alpha}\frac{\partial F(t_i;\alpha,\beta)}{\partial\beta}}{F^2(t_i;\alpha,\beta)}\right) - \lambda\sum_{i=1}^{k} r_i \frac{\partial^2 F(t_i;\alpha,\beta)}{\partial\beta\partial\alpha}$$

so that

$$I_{23} = I_{32} = \lambda\sum_{i=1}^{k}\frac{r_i}{F(t_i;\alpha,\beta)}\frac{\partial F(t_i;\alpha,\beta)}{\partial\alpha}\frac{\partial F(t_i;\alpha,\beta)}{\partial\beta}$$

From $\dfrac{\partial \ln L(\lambda,\alpha,\beta)}{\partial\beta} = \displaystyle\sum_{i=1}^{k} s_i \frac{\partial F(t_i;\alpha,\beta)}{\partial\beta}\frac{1}{F(t_i;\alpha,\beta)} - \lambda\sum_{i=1}^{k} r_i \frac{\partial F(t_i;\alpha,\beta)}{\partial\beta}$

we find;

i) $\dfrac{\partial^2 \ln L(\lambda,\alpha,\beta)}{\partial\beta^2} = \sum\limits_{i=1}^{k} S_i\left(\dfrac{\partial^2 F(t_i;\alpha,\beta)}{\partial\beta^2} \dfrac{1}{F(t_i;\alpha,\beta)}\right.$

$\left.-\dfrac{\dfrac{\partial F(t_i;\alpha,\beta)}{\partial\beta}^2}{\quad} \dfrac{1}{F^2(t_i;\alpha,\beta)}\right) - \lambda\sum\limits_{i=1}^{k} r_i \dfrac{\partial^2 F(t_i;\alpha,\beta)}{\partial\beta^2}$

so that

$$I_{23} = \lambda\sum\limits_{i=1}^{k}\dfrac{r_i}{F(t_i;\alpha,\beta)}\dfrac{\partial F(t_i;\alpha,\beta)}{\partial\beta}^2$$

Inverting the information matrix yields the desired result.

In type B experiments, we are also interested in constructing a confidence region for the actual value (λ,μ). We approximate the variance-covariance matrix for the estimator $\hat{\underline{X}}^t=(\hat\lambda,\hat\mu)$ using the same method as in Theorem 3.4.

Theorem 3.6   Assume a type B experiment. If λ,α,β are known, then the approximate variance-covariance matrix, $\Sigma_{\hat{\underline{X}}}$, for the maximum likelihood estimator $\hat{\underline{X}}^t=(\hat\lambda,\hat\mu)$ is

$\Sigma_{\hat{\underline{X}}}= \dfrac{1}{|I|}$ G   where $|I|$ is the determinant of the information matrix for $\hat{\underline{Y}}$ and the matrix G has as entries g_{ij} , $i,j=1,2$ given by

$$g_{11} = v_{11} \qquad g_{12} = g_{21} = \beta v_{12} + \alpha v_{13}$$

$$g_{22} = \beta^2 v_{22} + 2\alpha\beta v_{23} + \alpha^2 v_{33}$$

and the v_{ij}'s are defined in Theorem 3.5.

Proof   See Theorem 3.4.

The maximum likelihood estimator $\hat{\underline{X}}^t=(\hat\lambda,\hat\mu)$ in both type A and type B experiments is, asymptotically, a bivariate normal random variable with approximate variance-covariance matrix

$\Sigma_{\hat{X}}$. Therefore, a confidence region for the true value (λ, μ)

can be described and simple hypothesis tests can be conducted.

However, the variance-covariance matrix $\Sigma_{\hat{X}}$ contains λ, α, β,

the unknown values of the parameters we are trying to estimate.

When the sample size n, the number of experimental animals

per group, is large, the variance-covariance matrix for $\hat{X}$

may be approximated by substituting the maximum likelihood

estimates $\hat{\lambda}, \hat{\alpha}, \hat{\beta}$ into $\Sigma_{\hat{X}}$ for λ, α, β. This approximate variance-

covariance matrix for $\hat{X}$ will be denoted $\hat{\Sigma}_{\hat{X}}$.

Confidence Regions

The vector of parameters that we wish to estimate, (λ, μ),

has a set of possible values in $\mathbb{R}^+ \times \mathbb{R}^+$. Therefore, we

would like to determine an entire region of plausible values

in addition to the maximum likelihood point estimate. This

leads to the need for a confidence region. We will need the

following result from multivariate statistics.

Theorem 3.7 Let $\underline{X}$ be p-variate normal with mean $\underline{\mu}$ and

nonsingular variance-covariance matrix Σ. Define

$Q = (\underline{X} - \underline{\mu})^t \Sigma^{-1} (\underline{X} - \underline{\mu})$. Then Q is distributed as a chi-square,

χ^2, random variable on p degrees of freedom.

Proof Tatsuoka [12].

We apply this result to the maximum likelihood estimator $\hat{X}$.

Corollary 3.8 Let $\hat{X}$ be the maximum likelihood estimator

for $\underline{\mu} = (\lambda, \mu)$. Then $(\hat{\underline{X}} - \underline{\mu})^t \hat{\Sigma}_{\hat{X}}^{-1} (\underline{X} - \underline{\mu})$ has approximately a χ^2

distribution on two degrees of freedom.

<u>Proof</u>   We have already seen that $\hat{\underline{X}}$ is approximately

bivariate normal with mean $\underline{\mu}=(\lambda,\mu)$, and approximate variance-

covariance matrix $\hat{\Sigma}_{\underline{X}}$.   Applying Theorem 3.7 yields the

desired result.

Let R be a region in two-space defined by

$$R = \{\ \underline{x}\epsilon R^2\ |\ (\underline{x}-\underline{\mu})^t\hat{\Sigma}_{\underline{X}}^{-1}(\underline{x}-\underline{\mu}) \le \chi_2^2(\alpha')\ \} \tag{24}$$

where $\chi_2^2(\alpha')$ is the $100(1-\alpha')$ percent point of the chi-square

distribution with two degrees of freedom.   Then the following

is true;

1) the probability that $\hat{\underline{X}}\epsilon R$ is equal to $1-\alpha'$,

2) the region R contains $\underline{\mu}$, since the quadratic form in (24)

   is equal to 0 at $\underline{\mu}$,

3) R is a region bounded by an ellipse.

In practice we obtain a point estimate $\hat{\underline{x}}=(\hat{\lambda},\hat{\mu})$, but would

like a region of possible values for (λ,μ).   The objective is

to manipulate the quadratic form in Corollary 3.8 to produce

a region centered at the maximum likelihood estimate $(\hat{\lambda},\hat{\mu})$.

Rearranging the first and last vector of this quadratic form

yields $(\underline{\mu}-\hat{\underline{X}})^t\hat{\Sigma}_{\underline{X}}^{-1}(\underline{\mu}-\hat{\underline{X}})$ which is also approximately a chi-square

distribution on two degrees of freedom.

Thus, we define a region E in two space to be

$$E = \{\ \underline{\mu}\epsilon R^2\ |\ (\underline{\mu}-\hat{\underline{x}})^t\hat{\Sigma}_{\underline{X}}^{-1}(\underline{\mu}-\hat{\underline{x}}) \le \chi_2^2(\alpha')\ \}.$$

This defines a random region having an ellipse as boundary.

The location of the center $\hat{\underline{x}}$ is random.   The probability is

$1-\alpha'$ that this random ellipse contains the true value $\underline{\mu}$, so

that in the long run, $100(1-\alpha')\%$ of the samples will generate an ellipse that will cover $\underline{\mu}$.

Hypothesis Tests

Next, we consider the general hypothesis test concerning the population mean $\underline{\mu}=(\lambda,\mu)$. It is based on the multivariate result in Theorem 3.7. Recall that the maximum likelihood estimator $\underline{\hat{X}}$ is asymptotically bivariate normal, is an asymptotically unbiased estimator for $\underline{\mu}$, and has approximate variance-covariance matrix $\hat{\Sigma}_{\underline{\hat{X}}}$. The null hypothesis can be stated in the form $H_o: \underline{\mu} = \underline{\mu}_o = (\lambda_o,\mu_o)$ where $\underline{\mu}_o$ is a fixed point in $\mathbb{R}^2$, the null value of the parameter. Under the null hypothesis $(\underline{\hat{X}}-\underline{\mu}_o)^t \hat{\Sigma}_{\underline{\hat{X}}}^{-1}(\underline{\hat{X}}-\underline{\mu}_o)$ has approximately a chi-square distribution with two degrees of freedom. This leads to the following test procedure.

$H_o: \underline{\mu} = \underline{\mu}_o$

$H_a: \underline{\mu} \neq \underline{\mu}_o$

Test statistic: $\chi^2 = (\underline{\hat{x}}-\underline{\mu}_o)^t \hat{\Sigma}_{\underline{\hat{X}}}^{-1} (\underline{\hat{x}}-\underline{\mu}_o)$

Rejection region: $\chi^2 \geq \chi_2^2(\alpha')$

In chapter V we apply the theoretical results obtained here to a specific experimental tumor system of type A.

CHAPTER IV

TWO SAMPLE PROBLEM

Introduction

In this chapter we discuss a method for comparing the effects of two different treatments (diets) on the mean number of promoted tumors and the mean time to tumor detection. A diet supplemented by an effective chemopreventative agent should decrease the number of promoted tumors per animal and/or decrease the rate of tumor development, therefore increasing μ, the mean time to detection. Given two treatment groups we first obtain, through the estimation procedure described in Chapter II, the maximum likelihood estimates $\hat{\lambda}, \hat{\alpha}, \hat{\beta}$ for each group. Then we determine if there is a statistically significant difference between the two vectors $(\hat{\lambda}_1, \hat{\alpha}_1, \hat{\beta}_1)$ and $(\hat{\lambda}_2, \hat{\alpha}_2, \hat{\beta}_2)$.

The initial test procedure described here is designed to detect overall group differences. Once a difference between groups is established, further statistical analysis is used to detail which factor(s) (number of promoted tumors, mean time to tumor detection) is(are) contributing to the overall difference. The test procedures are based on the Likelihood Ratio Principle.

The Likelihood Ratio Test

Suppose group 1 and group 2 are random samples of sizes

n_1 and n_2, respectively, and that both samples arise from a density function $f(y; \underline{\theta})$, where $\underline{\theta}$ may be a vector of unknown parameters. The null hypothesis (H_o) specifies that all observations come from the same population. The alternate hypothesis (H_a) specifies that these two groups come from different populations. We must calculate the likelihood of the observed sample under the null and alternate hypotheses. The observed data is used to estimate all unknown parameters by the method of maximum likelihood under either hypothesis. Under H_o, the likelihood of the sample is a pooled likelihood assuming the two groups are actually observations from the same population. Under H_a, the likelihood of the sample becomes the product of the separately estimated group likelihoods.

Let $L(\hat{H}_o)$ denote the likelihood function assuming H_o is true, with all unknown parameters replaced by maximum likelihood estimates. Similarly, let $L(\hat{H}_a)$ denote the likelihood function assuming H_a is true, with all unknown parameters replaced by maximum likelihood estimates. A likelihood ratio test is based on the ratio $L(\hat{H}_o)/L(\hat{H}_a)$.

We define $r = L(\hat{H}_o)/L(\hat{H}_a)$. A likelihood ratio test of H_o: the two groups arise from the same population (or $\underline{\theta}_1 = \underline{\theta}_2$) versus H_a: the two groups arise from different populations (or $\underline{\theta}_1 \neq \underline{\theta}_2$) uses r as a test statistic. The rejection region is determined by $r \leq k$, where k is selected so that the significance level is at a predetermined value. The value of r is bounded by $0 \leq r \leq 1$. If r is close to zero, this indicates that the likelihood of the sample appears to be very small under H_o. Equivalently, this suggests that the observations

arise from different populations.

The likelihood ratio test does not always produce a test statistic with a known probability distribution. However, if the sample sizes n_1 and n_2 are large, then the distribution of r can be approximated.

If we assume the density function $f(y;\underline{\theta})$ satisfies the regularity conditions, then $-2\ln r$ converges in distribution, as $n \to \infty$ to a chi-square random variable [2]. In general, the number of degrees of freedom associated with this chi-square random variable is equal to the number of parameters or functions of parameters assigned specific values under H_o. In this case of comparing two groups, we have $\underline{\theta}_1 = \underline{\theta}_2$ under H_o. Hence the number of degrees of freedom equals the dimension of the vector $\underline{\theta}$. Therefore, when the sample sizes are large, we use a χ^2 table for finding regions with a fixed significance level. In this hypothesis test for comparing two groups, the test statistic is $\chi^2 = -2\ln r$ where $r = L(\hat{H}_o)/L(\hat{H}_a)$, and the rejection region is given by $\chi^2 \geq \chi^2_p(\alpha')$.

Type A Experiments - Likelihood of the Sample

In this section we apply the previous results about the likelihood ratio principle to type A experiments involving two treatment groups. Let m_{ij} , $i=1,2$, $j=1,\ldots,n_i$ be the number of observed tumors for animal j in group i. Let t_{ijk} $k=1,\ldots,m_{ij}$ be the time of detection for tumor k in animal j in group i. On the basis of these observed values we wish to test H_o: the vector of parameters (λ,α,β) is the same for

both groups; versus H_a: the vectors are different.

Let $\hat{\lambda}_i, \hat{\alpha}_i, \hat{\beta}_i$ be the separately estimated maximum likelihood estimates for λ, α, β in group i. Let $\hat{\lambda}_p, \hat{\alpha}_p, \hat{\beta}_p$ be the pooled maximum likelihood estimates. The likelihood of the sample under H_o is given by

$$L(\hat{H}_o) = \prod_{i=1}^{2} \frac{e^{-n_i \hat{\lambda}_p F(t^*; \hat{\alpha}_p, \hat{\beta}_p)} \hat{\lambda}_p^{s_{il}}}{K_i \hat{\beta}_p^{\hat{\alpha}_p s_{il}} \Gamma(\hat{\alpha}_p)^{s_{il}}} \prod_{j=1}^{n_i} \prod_{k=1}^{m_{ij}} t_{ijk}^{\hat{\alpha}_p - 1} e^{-t_{ijk}/\hat{\beta}_p}$$

where $K_i = m_{il}! \cdots m_{ij}!$

and $s_{il} = \sum_{j=1}^{n_i} m_{ij}$.

The likelihood of the sample under H_a is the product of the separately estimated group likelihoods. Therefore

$$L(\hat{H}_a) = [\frac{e^{-n_1 \hat{\lambda}_1 F(t^*; \hat{\alpha}_1, \hat{\beta}_1)} \hat{\lambda}_1^{s_{11}}}{K_1 \hat{\beta}_1^{\hat{\alpha}_1 s_{11}} \Gamma(\hat{\alpha}_1)^{s_{11}}} \prod_{j=1}^{n_1} \prod_{k=1}^{m_{1j}} t_{1jk}^{\hat{\alpha}_1 - 1} e^{-t_{1jk}/\hat{\beta}_1}]$$

$$\cdot [\frac{e^{-n_2 \hat{\lambda}_2 F(t^*; \hat{\alpha}_2, \hat{\beta}_2)} \hat{\lambda}_2^{s_{21}}}{K_2 \hat{\beta}_2^{\hat{\alpha}_2 s_{21}} \Gamma(\hat{\alpha}_2)^{s_{21}}} \prod_{j=1}^{n_2} \prod_{k=1}^{m_{2j}} t_{2jk}^{\hat{\alpha}_2 - 1} e^{-t_{2jk}/\hat{\beta}_2}]$$

Type B Experiments - Likelihood of the Sample

In this section we derive the likelihood of the sample under H_o and H_a for a type B experiment. Let r_{ij} , i=1,2 , j=1,...,k_i be the number of animals sacrificed in group i at time t_{ij}. Let m_{ijk} , k=1,...,r_{ij} be the number of observed

tumors in rat k sacrificed at time t_{ij} in group i.

Again, let $\hat{\lambda}_i, \hat{\alpha}_i, \hat{\beta}_i$ be the separately estimated maximum likelihood estimates for λ, α, β in group i and $\hat{\lambda}_p, \hat{\alpha}_p, \hat{\beta}_p$ be the pooled maximum likelihood estimates. The likelihood of the sample under H_o is given by

$$L(\hat{H}_o) = \prod_{i=1}^{2} \prod_{j=1}^{k_i} \prod_{k=1}^{r_{ij}} \frac{e^{-\hat{\lambda}_p F(t_{ijk}; \hat{\alpha}_p, \hat{\beta}_p)} (\hat{\lambda}_p F(t_{ijk}; \hat{\alpha}_p, \hat{\beta}_p))^{m_{ijk}}}{m_{ijk}!}$$

The likelihood of the sample under H_a is the product of the separately estimated group likelihoods. Therefore

$$L(\hat{H}_a) = \left[\prod_{j=1}^{k_1} \prod_{k=1}^{r_{1j}} \frac{e^{-\hat{\lambda}_1 F(t_{1jk}; \hat{\alpha}_1, \hat{\beta}_1)} (\hat{\lambda}_1 F(t_{1jk}; \hat{\alpha}_1, \hat{\beta}_1))^{m_{1jk}}}{m_{1jk}!} \right]$$

$$\cdot \left[\prod_{j=1}^{k_2} \prod_{k=1}^{r_{2j}} \frac{e^{-\hat{\lambda}_2 F(t_{2jk}; \hat{\alpha}_2, \hat{\beta}_2)} (\hat{\lambda}_2 F(t_{2jk}; \hat{\alpha}_2, \hat{\beta}_2))^{m_{2jk}}}{m_{2jk}!} \right]$$

Comparing Treatments

On the basis of the observed values in a type A or type B experiment, we wish to compare two treatments. The general test procedure is;

H_o: the two groups arise from the same population

   or $(\lambda_1, \alpha_1, \beta_1) = (\lambda_2, \alpha_2, \beta_2)$

H_a: the two groups arise from different populations

   or $(\lambda_1, \alpha_1, \beta_1) \neq (\lambda_2, \alpha_2, \beta_2)$

Test statistic: $\chi^2 = -2 \ln r$ where $r = L(\hat{H}_o)/L(\hat{H}_a)$

Rejection region: $\chi^2 \geq \chi_3^2(\alpha')$

If $\chi^2 \geq \chi_3^2(\alpha')$ we reject the null hypothesis at the α' level of significance and conclude that the two groups are probably different. This statistical test serves as a starting point for comparing treatment groups. Once a group difference is established a technique, discussed later in this chapter, is needed to determine along what dimensions the differences occur.

Asymptotic Power of the Likelihood Ratio Test

In this section we outline a general method for estimating the power of a likelihood ratio test, then apply these results to the likelihood ratio test for comparing two groups. Recall, under the regularity conditions, the maximum likelihood estimator $\hat{\underline{\theta}}$ is asymptotically multivariate normal, is an asymptotically unbiased estimator for $\underline{\mu}$ and has asymptotic variance-covariance matrix Σ defined by

$$\Sigma^{-1} = -E\left(\frac{\partial^2 \ln L(\underline{x}|\underline{\theta})}{\partial\theta_i \partial\theta_j}\right) \quad .$$

Suppose we conduct a likelihood ratio test as described previously with $H_o: \underline{\theta}=\underline{\theta}_o$ and $H_a: \underline{\theta}\neq\underline{\theta}_o$. Kendall and Stuart [9] discuss the asymptotic distribution of the likelihood ratio statistic associated with this test. Under the regularity conditions we have

$$r = \frac{L(\hat{H}_o)}{L(\hat{H}_a)} = \exp\left\{ -\tfrac{1}{2}(\hat{\underline{\theta}}-\underline{\theta}_o)^t \Sigma^{-1}(\hat{\underline{\theta}}-\underline{\theta}_o) \right\}$$

so that

$$-2\ln r = (\hat{\underline{\theta}}-\underline{\theta}_o)^t \Sigma^{-1} (\hat{\underline{\theta}}-\underline{\theta}_o)$$

In the likelihood ratio test, if there are p constraints in the null hypothesis, then $-2\ln r$ is asymptotically distributed as a non-central χ^2 distribution with p degrees of freedom and non-centrality parameter

$$\ell = (\underline{\theta}-\underline{\theta}_o)^t \Sigma^{-1} (\underline{\theta}-\underline{\theta}_o) \quad .$$

When H_o is true, $\ell=0$ and $-2\ln r$ reduces to a central χ^2 distribution with p degrees of freedom.

This result enables us to calculate the asymptotic power function of the likelihood ratio test. Suppose there are p constraints in the null hypothesis and we calculate the non-centrality parameter ℓ. The asymptotic power of the likelihood ratio test is given by

$$\pi = \int_{\chi'^2_{(p,0)}(\alpha')}^{\infty} d\chi'^2_{(p,\ell)}$$

where $\chi'^2_{(p,\ell)}$ is the non-central χ^2 distribution with p degrees of freedom and non-centrality parameter ℓ, and $\chi'^2_{(p,0)}(\alpha')$ is the $100(1-\alpha')$ percent point of the central χ^2 distribution. Kendall and Stuart use an approximation to the non-central χ^2 distribution so that the asymptotic power becomes

$$\pi \approx \int_{(\frac{p+\ell}{p+2\ell})\chi^2_p(\alpha')}^{\infty} d\chi^2_{(p+\frac{\ell^2}{p+2\ell})}$$

where χ^2_p is the central χ^2 distribution with p degrees of freedom and $\chi^2_p(\alpha')$ is the $100(1-\alpha')$ percent point. The degrees of freedom in this approximation can be fractional. Therefore

interpolation in the tables for χ^2 is often necessary.

We assume a type A or a type B experiment is conducted. Let $\hat{\underline{Y}}_1^t=(\hat{\lambda}_1,\hat{\alpha}_1,\hat{\beta}_1)$ and $\hat{\underline{Y}}_2^t=(\hat{\lambda}_2,\hat{\alpha}_2,\hat{\beta}_2)$ be the vectors of maximum likelihood estimators for the mean number of promoted tumors per animal and the time to detection parameters in groups 1 and 2 respectively. Assume n_1 and n_2, the number of animals per group, are large. We have that $\hat{\underline{Y}}_1$ and $\hat{\underline{Y}}_2$ are approximately multivariate normal, approximately unbiased estimators, with approximate variance-covariance matrices equal to the inverse of their respective information matrices.

In order to calculate the approximate power of the likelihood ratio test for comparing these two groups, we first define the vector $\hat{\underline{\delta}}=\hat{\underline{Y}}_1 - \hat{\underline{Y}}_2$. Then the following is true;

1) the random variable $\hat{\underline{\delta}}$ is approximately multivariate normal and is approximately an unbiased estimator for the difference

$$\underline{\delta} = \underline{\Theta}_1 - \underline{\Theta}_2 = \begin{pmatrix} \lambda_1 - \lambda_2 \\ \alpha_1 - \alpha_2 \\ \beta_1 - \beta_2 \end{pmatrix}$$

2) since $\hat{\underline{Y}}_1$ and $\hat{\underline{Y}}_2$ are independent, the variance-covariance matrix for $\hat{\underline{\delta}}$ is given by $\Sigma_{\hat{\underline{\delta}}} = \Sigma_1 + \Sigma_2$.

The null hypothesis and alternate hypothesis in the likelihood ratio test for comparing two groups can now be rewritten as

$H_o: \underline{\delta} = \underline{\Theta}_1 - \underline{\Theta}_2 = \underline{\delta}_o = \underline{0}$

$H_a: \underline{\delta} \neq \underline{0}$

Since there are three constraints in the null hypothesis, the asymptotic power of this test is given by

$$\pi = \int_{\chi'^2_{(3,0)}(\alpha')}^{\infty} d\chi'^2_{(3,\ell)}$$

where

$$\ell = (\hat{\underline{\delta}} - \underline{\delta}_o)^t \Sigma_{\hat{\delta}}^{-1} (\hat{\underline{\delta}} - \underline{\delta}_o) = (\hat{\underline{\delta}})^t \Sigma_{\hat{\delta}}^{-1} (\hat{\underline{\delta}})$$

and can be approximated by

$$\pi \approx \int_{(\frac{3+\ell}{3+2\ell})\chi_3^2(\alpha')}^{\infty} d\chi^2_{(3+\frac{\ell^2}{3+2\ell})} \quad .$$

Using these theoretical results, power tables corresponding to a specific type A experimental system are presented in Chapter V.

Isolating Group Differences

Once the hypothesis test for comparing two treatment groups is conducted, there are two courses of action depending on the conclusion. If we fail to reject H_o we conclude there is no support for a difference in tumor number or detection time parameters. However, if there is statistical evidence to support the conclusion of an overall group difference, then we would like to know if this difference is due to a change in the mean number of promoted tumors and/or a change in the rate of tumor development as manifest in the mean time to tumor detection.

We assume a type A or a type B experiment is conducted

and let $\hat{\underline{X}}_1^t = (\hat{\lambda}_1, \hat{\mu}_1)$ and $\underline{X}_2^t = (\hat{\lambda}_2, \hat{\mu}_2)$ be the vectors of maximum likelihood estimators of the mean number of promoted tumors per animal and the mean time to tumor detection in groups 1 and 2 respectively. If n_1 and n_2, the number of animals per group, are large, then we have that $\hat{\underline{X}}_1$ and $\hat{\underline{X}}_2$ are approximately bivariate normal, approximately unbiased estimators with approximate variance-covariance matrices $\hat{\Sigma}_1$ and $\hat{\Sigma}_2$.

In order to determine which factor, tumor number and/or detection time, is contributing to an overall group difference, we begin be defining the difference vector $\hat{\underline{D}} = \hat{\underline{X}}_1 - \hat{\underline{X}}_2$. Then the following is true;

1) the random variable $\hat{\underline{D}}$ is approximately bivariate normal and is approximately an unbiased estimator for the difference in means

$$\underline{\mu}_1 - \underline{\mu}_2 = \begin{pmatrix} \lambda_1 - \lambda_2 \\ \mu_1 - \mu_2 \end{pmatrix} = \underline{\mu}_{\hat{\underline{D}}}$$

2) since $\hat{\underline{X}}_1$ and $\hat{\underline{X}}_2$ are independent, the approximate variance-covariance matrix for $\underline{D}$ is given by $\hat{\Sigma}_{\hat{\underline{D}}} = \hat{\Sigma}_1 + \hat{\Sigma}_2$.

Next we apply Theorem 3.7 from multivariate statistics and we have

$$(\hat{\underline{D}} - \underline{\mu}_{\hat{\underline{D}}})^t \hat{\Sigma}_{\hat{\underline{D}}}^{-1} (\hat{\underline{D}} - \underline{\mu}_{\hat{\underline{D}}}) \tag{25}$$

has approximately a chi-square distribution on two degrees of freedom. This result leads to one method for obtaining simultaneous confidence intervals for the differences $\lambda_1 - \lambda_2$ and $\mu_1 - \mu_2$. The inclusion or exclusion of zero in these

confidence intervals indicates which response(s) has(have)
probably led to the rejection of the vector hypothesis.
The following theorem is used to illustrate one method for
constructing simultaneous confidence intervals.

<u>Theorem 4.1</u>

$$P\left[\left(\frac{\underline{a}^t\hat{\underline{D}} - \underline{a}^t\underline{\mu}_{\hat{\underline{D}}}}{\sqrt{\underline{a}^t\hat{\Sigma}_{\hat{\underline{D}}}\underline{a}}}\right)^2 \leq \chi_2^2(\alpha'), \text{ for all vectors } \underline{a}\right] = 1-\alpha' \qquad (26)$$

In order to prove this theorem we will need the following
lemma.

<u>Lemma 4.2</u>

$$\text{Let} \quad \ell(\underline{a}) = \left(\frac{\underline{a}^t\hat{\underline{D}} - \underline{a}^t\underline{\mu}_{\hat{\underline{D}}}}{\sqrt{\underline{a}^t\hat{\Sigma}_{\hat{\underline{D}}}\underline{a}}}\right)^2$$

then $\sup\limits_{\underline{a}} \ell(\underline{a}) = c_1$ where c_1 is the largest eigenvalue

of the matrix $\hat{\Sigma}_{\hat{\underline{D}}}^{-1}(\hat{\underline{D}}-\underline{\mu}_{\hat{\underline{D}}})(\hat{\underline{D}}-\underline{\mu}_{\hat{\underline{D}}})^t$.

<u>Proof</u> $\ell(\underline{a})$ may be written as

$$\frac{\underline{a}^t(\hat{\underline{D}}-\underline{\mu}_{\hat{\underline{D}}})(\hat{\underline{D}}-\underline{\mu}_{\hat{\underline{D}}})^t\underline{a}}{\underline{a}^t\hat{\Sigma}_{\hat{\underline{D}}}\underline{a}} .$$

To find the supremum of this expression, we take the derivative
with respect to the vector $\underline{a}$ and set the result equal to zero.

$$\frac{(\underline{a}^t\hat{\Sigma}_{\hat{\underline{D}}}\underline{a})(2(\hat{\underline{D}}-\underline{\mu}_{\hat{\underline{D}}})(\hat{\underline{D}}-\underline{\mu}_{\hat{\underline{D}}})^t\underline{a})-(\underline{a}^t(\hat{\underline{D}}-\underline{\mu}_{\hat{\underline{D}}})(\hat{\underline{D}}-\underline{\mu}_{\hat{\underline{D}}})^t\underline{a})(2\hat{\Sigma}_{\hat{\underline{D}}}\underline{a})}{(\underline{a}^t\hat{\Sigma}_{\hat{\underline{D}}}\underline{a})^2} = 0 \qquad (27)$$

Dividing both numerator and denominator by $\underline{a}^t \hat{\Sigma}_{\wedge \underline{D}} \underline{a}$ and using

the definition of ℓ, equation (27) reduces to

$$\frac{2[(\hat{\underline{D}}-\mu_{\wedge \underline{D}})(\hat{\underline{D}}-\mu_{\wedge \underline{D}})^t \underline{a} - \ell \hat{\Sigma}_{\wedge \underline{D}} \underline{a}]}{\underline{a}^t \hat{\Sigma}_{\wedge \underline{D}} \underline{a}} = 0 \qquad .$$

This is equivalent to solving

$$[(\hat{\underline{D}}-\mu_{\wedge \underline{D}})(\hat{\underline{D}}-\mu_{\wedge \underline{D}})^t - \ell \hat{\Sigma}_{\wedge \underline{D}}]\underline{a} = 0 \qquad . \tag{28}$$

Assuming $\hat{\Sigma}_{\wedge \underline{D}}^{-1}$ exists, premultiply both sides of equation (28)

by it to obtain

$$[\hat{\Sigma}_{\wedge \underline{D}}^{-1}(\hat{\underline{D}}-\mu_{\wedge \underline{D}})(\hat{\underline{D}}-\mu_{\wedge \underline{D}})^t - \ell I]\underline{a} = 0 \tag{29}$$

where I is the identity matrix.

Equation (29) is of the form $(A - \ell I)\underline{v} = 0$ and has as a

solution the eigenvalues c_m and associated eigenvectors $\underline{v}_m$

of the matrix A. Therefore, let c_1 be the largest eigenvalue

for $\hat{\Sigma}_{\wedge \underline{D}}^{-1}(\hat{\underline{D}}-\mu_{\wedge \underline{D}})(\hat{\underline{D}}-\mu_{\wedge \underline{D}})^t$ then for any eigenvector $\underline{a}_m$ corresponding

to c_m we have

$$\ell(\underline{a}_m) = [\frac{\underline{a}_m^t(\hat{\underline{D}}-\mu_{\wedge \underline{D}})(\hat{\underline{D}}-\mu_{\wedge \underline{D}})^t \underline{a}_m}{\underline{a}_m^t \hat{\Sigma}_{\wedge \underline{D}} \underline{a}_m}]$$

$$= \frac{\underline{a}_m^t \hat{\Sigma}_{\wedge \underline{D}} c_m \underline{a}_m}{\underline{a}_m^t \hat{\Sigma}_{\wedge \underline{D}} \underline{a}_m} = c_m \leq c_1 \qquad .$$

Strictly speaking, equation (29) represents only a necessary condition for maximizing $\ell(\underline{a})$. We should examine the second order derivative of $\ell(\underline{a})$ with respect to $\underline{a}$ in order to prove that the solution to (29) yields a maximum. However it is a common multivariate result that equation (29) is a sufficient as well as necessary condition for maximizing $\ell(\underline{a})$ [12].

<u>Proof of Theorem 4.1</u> Rewriting equation (26) we have

$$P[\sup_{\underline{a}} \left\{ \frac{\underline{a}^t(\hat{\underline{D}}-\mu_{\hat{\underline{D}}})(\hat{\underline{D}}-\mu_{\hat{\underline{D}}})^t\underline{a}}{\underline{a}^t\hat{\Sigma}_{\hat{\underline{D}}}\underline{a}} \right\} \leq \chi_2^2(\alpha')] = 1-\alpha'$$

which is true if and only if

$$P[c_1 \leq \chi_2^2(\alpha')] = 1-\alpha' \qquad \text{where } c_1 \text{ is defined in the Lemma,}$$

$$\leftrightarrow P[\text{tr } \hat{\Sigma}_{\hat{\underline{D}}}^{-1}(\hat{\underline{D}}-\mu_{\hat{\underline{D}}})(\hat{\underline{D}}-\mu_{\hat{\underline{D}}})^t \leq \chi_2^2(\alpha')] = 1-\alpha'$$

$$\leftrightarrow P[\text{tr } (\hat{\underline{D}}-\mu_{\hat{\underline{D}}})^t\hat{\Sigma}_{\hat{\underline{D}}}^{-1}(\hat{\underline{D}}-\mu_{\hat{\underline{D}}}) \leq \chi_2^2(\alpha')] = 1-\alpha'$$

$$\leftrightarrow P[(\hat{\underline{D}}-\mu_{\hat{\underline{D}}})^t\hat{\Sigma}_{\hat{\underline{D}}}^{-1}(\hat{\underline{D}}-\mu_{\hat{\underline{D}}}) \leq \chi_2^2(\alpha')] = 1-\alpha'$$

which is true from equation (25). Since these statements are if and only if, working in the other direction proves the theorem.

Manipulating the inequality inside the probability statement in equation (26) we have that

$$\underline{a}^t\hat{\underline{D}} \pm \sqrt{\chi_2^2(\alpha')}\sqrt{\underline{a}^t\hat{\Sigma}_{\hat{\underline{D}}}\underline{a}}$$

are endpoints for $100(1-\alpha')\%$ confidence intervals generated

by different choices of the vector $\underline{a}$. The probability that
all such intervals are simultaneously true is $1-\alpha'$. In
practice, we are interested in the two vectors $\underline{a}_1^t = (1,0)$
and $\underline{a}_2^t = (0,1)$. The endpoints of the two simultaneous
confidence intervals for $\lambda_1-\lambda_2$ and $\mu_1-\mu_2$ reduce to

$$(\hat{\lambda}_1-\hat{\lambda}_2) \pm \sqrt{\chi_2^2(\alpha')}\ \hat{\sigma}_{11} \tag{30}$$

and

$$(\hat{\mu}_1-\hat{\mu}_2) \pm \sqrt{\chi_2^2(\alpha')}\ \hat{\sigma}_{22} \tag{31}$$

where $\hat{\sigma}_{11}$ and $\hat{\sigma}_{22}$ are the square roots of their respective
entries in the matrix $\hat{\Sigma}_{\hat{\underline{D}}}$. Since only two confidence

intervals are calculated, the full power of Theorem 4.1 is
not being utilized. Thus, the probability that the two
confidence intervals are simultaneously true might be
greater than $1-\alpha'$.

An alternate method for constructing simultaneous
confidence intervals for $\lambda_1-\lambda_2$ and $\mu_1-\mu_2$ is described below.
For any vector $\underline{a}$ we can standardize the random variable $\underline{a}^t\hat{\underline{D}}$
and make the following probability statement

$$P\left(\left|\frac{\underline{a}^t\hat{\underline{D}} - \underline{a}^t\underline{\mu}_{\hat{\underline{D}}}}{\sqrt{\underline{a}^t\hat{\Sigma}_{\hat{\underline{D}}}\underline{a}}}\right| \le z_{\alpha'/2}\right) = 1-\alpha' \tag{32}$$

where $z_{\alpha'/2}$ denotes the $100(1-\alpha')$ percent point of the
standard normal distribution. The inequality within the
probability statement in equation (32) can be expanded so
that a $100(1-\alpha')\%$ confidence interval for $\underline{a}^t\underline{\mu}_{\hat{\underline{D}}}$ has as
endpoints

$$\underline{a}^t \hat{\underline{D}} \pm z_{\alpha'/2} \sqrt{\underline{a}^t \hat{\Sigma}_{\hat{\underline{D}}} \underline{a}} \quad .$$

The probability that this random interval contains the true value $\underline{a}^t \mu_{\hat{\underline{D}}}$ is $1-\alpha'$. We are interested in the random variables generated by the vectors $\underline{a}_1$ and $\underline{a}_2$. Thus, the simultaneous $100(1-\alpha')\%$ confidence (Bonferroni) intervals for $\lambda_1 - \lambda_2$ and $\mu_1 - \mu_2$ have as endpoints

$$(\hat{\lambda}_1 - \hat{\lambda}_2) \pm z_{\alpha'/4} \; \hat{\sigma}_{11} \tag{33}$$

and

$$(\hat{\mu}_1 - \hat{\mu}_2) \pm z_{\alpha'/4} \; \hat{\sigma}_{22} \quad . \tag{34}$$

These two methods produce joint confidence intervals for the random variables of interest. The only difference is the length of the intervals, determined by $\sqrt{\chi_2^2(\alpha')}$ versus $z_{\alpha'/4}$. The smaller of these two quantites dictates which set of confidence intervals is used. A comparison is shown in Table 1 for common critical values.

Table 1

α'	$z_{\alpha'/4}$	$\sqrt{\chi_2^2(\alpha')}$
.1	1.96	2.15
.05	2.24	2.45
.025	2.495	2.72
.01	2.81	3.03
.005	3.025	3.26

Critical values for given α'

Table 1 indicates that the second method described for constructing joint confidence intervals should be used in this case where we are interested in estimating intervals

for only two values, $\lambda_1 - \lambda_2$ and $\mu_1 - \mu_2$.

If there is statistical evidence in support of an overall group difference, we use these joint confidence intervals to isolate the parameters (tumor number and/or detection time) contributing to the difference. If zero is contained in either interval, (33) or (34), then we conclude at the α'% joint significance level that the mean number of promoted tumors and/or the mean time to tumor detection are not different. If zero lies outside an interval, this is evidence of different means, probably contributing to the overall group difference. If there is an overall group difference, but both confidence intervals contain zero, then we conclude that neither parameter alone is causing the difference. The overall difference between the two treatment groups may be the result of the simultaneous action of both parameters or a difference in the variance of the time to detection distributions.

CHAPTER V

NUMERICAL RESULTS - TYPE A EXPERIMENTS

Sensitivity of the Estimation Procedure

The experimenter controls several variables which directly affect the cost of an investigation and influence the estimation procedure. The frequency of examination intervals, the duration of the experiment, and the number of animals per group are all determined prior to the experiment. In this section we examine the effects of these variables on the accuracy of the estimation procedure for type A experiments in order to design more efficient experiments and better utilize resources. Type B experiments will be considered in a later work.

We specialize to the particular type A experiment with which we are familiar, the mammary tumor system as developed by Gullino [15] and modified by Thompson and Meeker. Dr. Thompson's experiments involve the chemical induction of mammary tumors in Sprague-Dawley rats by the carcinogens 1-methyl-1-nitrosourea (MNU) or 7,12 dimethylbenz (α) anthracene (DMBA). After exposure to a carcinogen, the mammary glands of the experimental animals are palpated regularly and the times of appearance and locations of the induced tumors are noted. At the termination of the experiment (usually 120-180 days after carcinogen exposure) the animals are sacrificed and all tumors are removed and classified

histopathologically.

After exposure to a direct acting carcinogen, experimental animals in this mammary system typically begin to exhibit palpable tumors 20-50 days later. Once the onset of tumors begins, animals are examined regularly for the detection of new tumors. Substantial time and resources are expended in this process in order to accurately determine the time to appearance and location of each cancer.

Clearly more frequent examination yield more accurate estimates of the underlying distribution parameters. However, examining animals every day (or ideally - continuously) is physically (and financially) impossible. In the mammary tumor system described above, animals are examined twice weekly, or approximately once every three days. This frequency of examinations makes it more difficult to estimate the under-lying continuous (Gamma) distribution characterizing the time to appearance of tumors.

Since we assume that the detection time distribution is continuous, the probability of detecting two or more tumors at the same time is zero. However, animals are not examined continuously, therefore several tumors are often detected at each examination time. Preliminary computer simulations suggest that examining animals every three days still provides adequate data for accurately estimating the underlying continuous distribution.

The analytical method we will use to evaluate the sensitivity of the estimation procedure subject to changes in the experimenter controlled variables involves calculating

the area of the 95% confidence ellipse centered at $(\hat{\lambda}, \hat{\mu})$.
This area is not directly dependent on the frequency of
examination intervals. However, computer simulations indicate
that, as examination intervals increase, the accuracy of the
estimation procedure decreases.

If the termination date of an experiment is extended, more
data can be collected, therefore the underlying distributions
can be more accurately estimated. However, lengthy experiments
are costly. Thus, we seek a balance between the length of the
experiment and accuracy of estimation. We will examine the
sensitivity of the estimation procedure to duration of
experiments. For a fixed set of distribution parameters, if
the estimation procedure continues to be as accurate even when
the duration of the experiment is shortened, savings in time
and resources will result.

The number of animals per group is also set by the
investigator prior to the experiment. The care and feeding
of extra animals is expensive while additional animals may
increase the accuracy of the estimates of the mean number of
promoted tumors and mean time to detection only marginally.
Dr. Thompson's experiments typically have 20-25 animals per
group. We will examine the effect of number of animals per
group on the accuracy of the estimation procedure.

Several of the theorems stated in this study are based
on asymptotic properties. In experiments of type A, we also
note the direct dependence of the variance-covariance matrix
of the estimator $\underline{\hat{X}}^t = (\hat{\lambda}, \hat{\mu})$ on n - the number of animals in the

group. Therefore, fewer animals per group may result in a breakdown of our model assumptions and certainly larger confidence regions for the true value $\underline{\mu}^t=(\lambda,\mu)$. Once again we seek a balance between the number of animals per group and the accuracy of the estimation procedure.

Method of Evaluation

In this section the method used to test the sensitivity of the estimation procedure is presented. Suppose $\hat{\underline{\theta}}$ is a p-variate normal unbiased estimator for the parameter $\underline{\theta}$, and suppose $\hat{\underline{\theta}}$ has variance-covariance matrix Σ. As we have seen before, the region defined by

$$R = \{ \underline{\theta}\epsilon \, \mathbb{R}^p \mid (\underline{\theta}-\hat{\underline{\theta}})^t\Sigma^{-1}(\underline{\theta}-\hat{\underline{\theta}}) \leq \chi_p^2(\alpha') \}$$

is a $100(1-\alpha')\%$ confidence region for the true value of $\underline{\theta}$. We would like to have the volume of this confidence region, an indication of accuracy, as small as possible. To calculate the volume, we let $c^2=\chi_p^2(\alpha')$, then the volume of the region is given by

$$\text{volume} = k_p c^p |\Sigma|^{1/2}$$

where $k_p = 2\pi^{p/2}/(p\Gamma(p/2))$,

$\Gamma(z)$ denotes the gamma function evaluated at z, and

$|\Sigma|$ is the determinant of the matrix Σ [13].

If there are any experimental parameters that effect Σ we would like to design the experiment so that the volume, or equivalently $|\Sigma|$, is a minimum.

Next, we apply this result to obtain the area of a confidence ellipse about the point estimate $\hat{\underline{x}}^t=(\hat{\lambda},\hat{\mu})$.

Theorem 5.1 Assume a type A experiment, λ, α, β known, and a $100(1-\alpha')\%$ confidence ellipse about the point $\hat{\underline{x}}^t = (\hat{\lambda}, \hat{\mu})$ is defined by

$$E = \{ \underline{x} \in \mathbb{R}^2 \mid (\underline{x}-\hat{\underline{x}})^t \Sigma_{\hat{\underline{X}}}^{-1} (\underline{x}-\hat{\underline{x}}) \leq \chi_2^2(\alpha') \} .$$

Then the area of this confidence ellipse is given by

$$\text{area} = \frac{1}{n} \frac{\pi \chi_2^2(\alpha')}{|I|} |G|^{1/2}$$

where $\Sigma_{\hat{\underline{X}}} = \frac{1}{n|I|} G$ is the variance-covariance matrix for $\hat{\underline{X}}$ and $|I|$ is the determinant of the information matrix for $\hat{\underline{Y}}$ as defined in Theorem 3.3.

Proof From the previous result we have

$$\text{area} = k_2 (\sqrt{\chi_2^2(\alpha')})^2 \left| \frac{1}{n|I|} G \right|^{1/2} .$$

Since $k_2 = 2\pi^{2/2}/(2\Gamma(2/2)) = \pi$ and G is a 2x2 matrix, then

$$\text{area} = \pi \chi_2^2(\alpha') ((\frac{1}{n} \frac{1}{|I|})^2)^{1/2} |G|^{1/2}$$

$$= \frac{1}{n} \frac{\pi \chi_2^2(\alpha')}{|I|} |G|^{1/2} .$$

Since the area of the confidence ellipse is inversely proportional to n, we can attain a desired level of accuracy by selecting a large enough n. We will examine the gain in accuracy for each additional animal.

Theorem 5.2 Assume a type A experiment, λ, α, β known, and a $100(1-\alpha')\%$ confidence ellipse about the point $\hat{\underline{x}}^t = (\hat{\lambda}, \hat{\mu})$ as as defined in the previous theorem. Then the area of the

confidence ellipse is independent of λ.

Proof   The area of the confidence ellipse is given by

$$\text{area} = \frac{1}{n} \frac{\pi \chi_2^2(\alpha')}{|I|} |G|^{1/2} .$$

To show this is independent of λ, we need only consider the terms that contain λ, α, β - $|I|$ and $|G|^{1/2}$. From Theorem 3.3 and Theorem 3.4 we have the following;

1) $|I| = \lambda h(\alpha, \beta, t^*)$ where h is a function of only α, β, and t^*,

2)
$$G = \begin{pmatrix} \lambda^2 g_{11} & \lambda g_{12} \\ \lambda g_{21} & g_{22} \end{pmatrix}$$

where the g_{ij}'s are functions of only α, β, and t^*. Therefore,

$$\frac{|G|^{1/2}}{|I|} = \frac{(\lambda^2 g_{11} g_{22} - \lambda^2 g_{12} g_{21})^{1/2}}{\lambda h(\alpha, \beta, t^*)}$$

$$= \frac{(g_{11} g_{22} - g_{12} g_{21})^{1/2}}{h(\alpha, \beta, t^*)}$$

which is independent of λ. Therefore the area of the confidence ellipse is determined by α, β, t^*, and n.

From the factorization of $|I|$, we can write the area of the confidence ellipse as

$$\text{area} = \frac{1}{n} \pi \chi_2^2(\alpha') \begin{vmatrix} \lambda g'_{11} & g'_{12} \\ g'_{21} & g'_{22}/\lambda \end{vmatrix}^{1/2} \tag{35}$$

where g'_{ij}, $i,j=1,2$ is a function of α, β, and t^* only. If the parameters α, β, and t^* remain fixed, as λ increases the shape of the confidence ellipse changes but the area is

constant. The variance of $\hat{\lambda}$ is proportional to λ and the variance of $\hat{\mu}$ is inversely proportional to λ. As λ increases, we expect more variability in the estimator $\hat{\lambda}$, however, more tumors per animal allow a better estimate of μ.

As a consequence of the previous theorem and equation (35) we can write the area of the confidence ellipse in the following way. If α and β, and hence μ and σ, are known, we have

$$\text{area} = \frac{1}{n} \, \Phi(\mu, \sigma, t^*)$$

where $\quad \Phi(\mu, \sigma, t^*) = \pi \chi_2^2(\alpha')(g'_{11}g'_{22} - g'_{12}g'_{21})$,   (36)

n is the number of animals in the group and t^* is the length of the experiment.

Suppose t^* is fixed and the experimenter has prior knowledge about μ and σ. Then the number of animals per group can be selected to specify the area of the confidence ellipse. If we choose area=A_1, then selecting $n \geq \Phi(\mu, \sigma, t^*)/A_1$ ensures a confidence region no larger than A_1.

In order to examine the sensitivity of the estimation procedure subject to changes in the experimental parameters n and t^* we need a reference point of accuracy.

Definition Assume a type A experiment, α and β known. If the length of the experiment is t^*, then the normalized number of animals in the group for a $100(1-\alpha')\%$ confidence ellipse is defined to be

$$n_o = \begin{cases} \sqrt{\Phi(\mu, \sigma, t^*)} & \text{if } \Phi(\mu, \sigma, t^*) \text{ is a perfect square} \\[2em] [\sqrt{\Phi(\mu, \sigma, t^*)}] + 1 & \text{otherwise} \end{cases}$$

where Φ is defined in equation (36) and $[x]$ is the greatest integer function.

   Definition   The normalized area is given by

$$A_o = \frac{1}{n_o} \, \Phi(\mu,\sigma,t^*) \qquad .$$

The reasoning for these definitions is as follows. Assuming α and β are known and t^* is fixed, the area of the confidence ellipse is a decreasing function of n, Therefore, each additional animal increases the accuracy of the estimation procedure. However, the gain in accuracy reaches a point of diminished returns when adding an additional animal to the group decreases the area by an amount less than one. This is the point on the curve $A(n) = \frac{1}{n} \, \Phi(\mu,\sigma,t^*)$ where the slope is equal to -1. Taking the first derivative of $A(n)$ and setting it equal to -1 yields

$$A'(n) = -\frac{1}{n^2} \, \Phi(\mu,\sigma,t^*) = -1$$

   so that $n = \sqrt{\Phi(\mu,\sigma,t^*)}$.

Since we want the normalized number of animals to be a whole number, we use the greatest integer function in the definition. We round up to ensure the slope of $A(n)$ at n_o is greater than or equal to -1.

   In order to discuss the sensitivity of the estimation procedure, we assume the length of the experiment to be 180 days. For a fixed μ,σ, we select as a reference point A_o^*, which is the normalized area for t^* equal to 180 days. Additionally we assume the significance level is .05.

   Tables 2-11 display the accuracy of the

Table 2

	t*						
	120	135	150	165	180	195	210
n 15	2.10	1.90	1.78	1.71	1.67	1.64	1.62
16	1.97	1.78	1.67	1.60	1.56	1.54	1.52
17	1.86	1.68	1.57	1.51	1.47	1.45	1.43
18	1.75	1.58	1.48	1.42	1.39	1.37	1.35
19	1.66	1.50	1.40	1.35	1.32	1.29	1.29
20	1.58	1.43	1.33	1.28	1.25	1.23	1.21
21	1.50	1.36	1.27	1.22	1.19	1.17	1.16
22	1.43	1.30	1.21	1.17	1.14	1.12	1.10
23	1.37	1.24	1.16	1.11	1.09	1.07	1.06
24	1.31	1.19	1.11	1.07	1.04	1.02	1.01
25	1.26	1.14	1.07	1.03	1.00	.98	.97
26	1.21	1.10	1.03	.99	.96	.95	.93
27	1.17	1.06	.99	.95	.93	.91	.90
28	1.13	1.02	.95	.92	.89	.88	.87
29	1.09	.98	.92	.88	.86	.85	.84
30	1.05	.95	.89	.85	.83	.82	.81
31	1.02	.92	.86	.83	.81	.79	.78
32	.99	.89	.83	.80	.78	.77	.76
33	.96	.86	.81	.78	.76	.75	.74
34	.93	.84	.78	.75	.74	.72	.71
35	.90	.81	.76	.73	.71	.70	.69

Accuracy of the estimation procedure
Table entries are $A(n,t*)/A_0^*$
$\mu=80$, $\sigma=30$

Table 3

t*	120	135	150	165	180	195	210
n 15	2.35	2.01	1.83	1.74	1.67	1.64	1.62
16	2.21	1.89	1.72	1.63	1.56	1.53	1.52
17	2.08	1.77	1.62	1.53	1.47	1.44	1.43
18	1.96	1.68	1.53	1.45	1.39	1.36	1.35
19	1.86	1.59	1.45	1.37	1.32	1.29	1.28
20	1.76	1.51	1.38	1.30	1.25	1.23	1.21
21	1.68	1.44	1.31	1.24	1.19	1.17	1.16
22	1.60	1.37	1.25	1.18	1.14	1.12	1.10
23	1.53	1.31	1.20	1.13	1.09	1.07	1.06
24	1.47	1.26	1.15	1.08	1.04	1.02	1.01
25	1.41	1.21	1.10	1.04	1.00	.98	.97
26	1.36	1.16	1.06	1.00	.96	.94	.93
27	1.31	1.12	1.02	.96	.93	.91	.90
28	1.26	1.08	.98	.93	.89	.88	.87
29	1.22	1.04	.95	.90	.86	.85	.84
30	1.18	1.01	.92	.87	.83	.82	.81
31	1.14	.97	.89	.84	.81	.79	.78
32	1.10	.94	.86	.81	.78	.77	.76
33	1.07	.91	.83	.79	.76	.74	.74
34	1.04	.89	.81	.77	.74	.72	.71
35	1.01	.86	.79	.74	.71	.70	.69

Accuracy of the estimation procedure
Table entries are $A(n,t*)/A_0^*$
$\mu=90$, $\sigma=30$

Table 4

n	t* 120	135	150	165	180	195	210
15	2.70	2.17	1.90	1.75	1.67	1.62	1.59
16	2.53	2.03	1.78	1.64	1.56	1.52	1.49
17	2.38	1.91	1.68	1.55	1.47	1.43	1.41
18	2.25	1.81	1.58	1.46	1.39	1.35	1.33
19	2.13	1.71	1.50	1.38	1.32	1.28	1.26
20	2.02	1.63	1.43	1.31	1.25	1.21	1.19
21	1.93	1.55	1.36	1.25	1.19	1.16	1.14
22	1.84	1.48	1.30	1.20	1.14	1.10	1.09
23	1.76	1.41	1.24	1.14	1.09	1.06	1.04
24	1.69	1.36	1.19	1.10	1.04	1.01	1.00
25	1.62	1.30	1.14	1.05	1.00	.97	.96
26	1.56	1.25	1.10	1.01	.96	.93	.92
27	1.50	1.21	1.06	.97	.93	.90	.89
28	1.44	1.16	1.02	.94	.89	.87	.85
29	1.40	1.12	.98	.91	.86	.84	.82
30	1.35	1.08	.95	.88	.83	.81	.80
31	1.31	1.05	.92	.85	.81	.78	.77
32	1.26	1.02	.89	.82	.78	.76	.75
33	1.23	.99	.86	.80	.76	.74	.72
34	1.19	.96	.84	.77	.74	.71	.70
35	1.16	.93	.81	.75	.71	.69	.68

Accuracy of the estimation procedure
Table entries are $A(n,t*)/A_0^*$
$\mu=100$, $\sigma=30$

Table 5

	t*						
	120	135	150	165	180	195	210
n 15	3.13	2.38	2.00	1.79	1.67	1.60	1.56
16	2.94	2.23	1.88	1.68	1.56	1.50	1.46
17	2.76	2.10	1.77	1.58	1.47	1.41	1.38
18	2.61	1.98	1.67	1.49	1.39	1.33	1.30
19	2.47	1.88	1.58	1.42	1.32	1.26	1.23
20	2.35	1.78	1.50	1.35	1.25	1.20	1.17
21	2.24	1.70	1.43	1.28	1.19	1.14	1.12
22	2.14	1.62	1.36	1.22	1.14	1.09	1.07
23	2.04	1.55	1.31	1.17	1.09	1.04	1.02
24	1.96	1.49	1.25	1.12	1.04	1.00	.98
25	1.88	1.43	1.20	1.08	1.00	.96	.94
26	1.81	1.37	1.15	1.03	.96	.92	.90
27	1.74	1.32	1.11	1.00	.93	.89	.87
28	1.68	1.27	1.07	.96	.89	.86	.84
29	1.62	1.23	1.04	.93	.86	.83	.81
30	1.57	1.19	1.00	.90	.83	.80	.78
31	1.52	1.15	.97	.87	.81	.77	.76
32	1.47	1.11	.94	.84	.78	.75	.73
33	1.42	1.08	.91	.82	.76	.73	.71
34	1.38	1.05	.88	.79	.74	.71	.69
35	1.34	1.02	.86	.77	.71	.68	.67

Accuracy of the estimation procedure
Table entries are $A(n,t*)/A_0^*$
$\mu=110$, $\sigma=30$

Table 6

	t*						
	120	135	150	165	180	195	210
n 15	4.54	3.17	2.27	1.94	1.73	1.64	1.58
16	4.26	2.97	2.13	1.81	1.63	1.54	1.48
17	4.01	2.79	2.01	1.71	1.53	1.45	1.39
18	3.78	2.64	1.90	1.61	1.44	1.37	1.32
19	3.58	2.50	1.80	1.53	1.37	1.29	1.25
20	3.40	2.38	1.71	1.45	1.30	1.23	1.19
21	3.24	2.26	1.62	1.38	1.24	1.17	1.13
22	3.10	2.16	1.55	1.32	1.18	1.12	1.08
23	2.96	2.07	1.48	1.26	1.13	1.07	1.03
24	2.84	1.98	1.42	1.21	1.08	1.02	.99
25	2.72	1.90	1.36	1.16	1.04	.98	.95
26	2.62	1.83	1.31	1.12	1.00	.95	.91
27	2.52	1.76	1.26	1.08	.96	.91	.88
28	2.43	1.70	1.22	1.04	.93	.88	.85
29	2.35	1.64	1.18	1.00	.90	.85	.82
30	2.27	1.58	1.14	.97	.87	.82	.79
31	2.20	1.53	1.10	.94	.84	.79	.76
32	2.13	1.48	1.07	.91	.81	.77	.74
33	2.06	1.44	1.03	.88	.79	.74	.72
34	2.00	1.40	1.00	.85	.76	.72	.70
35	1.95	1.36	.97	.83	.74	.70	.68

Accuracy of the estimation procedure
Table entries are $A(n,t*)/A_0^*$
$\mu=120$, $\sigma=30$

Table 7

n	t* 120	135	150	165	180	195	210
15	2.52	2.28	2.17	2.05	1.93	1.90	1.83
16	2.36	2.14	2.03	1.92	1.81	1.78	1.71
17	2.22	2.01	1.91	1.81	1.71	1.68	1.61
18	2.10	1.90	1.81	1.71	1.61	1.58	1.52
19	1.99	1.80	1.71	1.62	1.53	1.50	1.44
20	1.89	1.71	1.63	1.54	1.45	1.43	1.37
21	1.80	1.63	1.55	1.46	1.38	1.36	1.30
22	1.72	1.55	1.48	1.40	1.32	1.30	1.25
23	1.64	1.49	1.41	1.34	1.26	1.24	1.19
24	1.57	1.42	1.35	1.28	1.21	1.19	1.14
25	1.51	1.37	1.30	1.23	1.16	1.14	1.10
26	1.45	1.31	1.25	1.18	1.12	1.10	1.05
27	1.40	1.27	1.20	1.14	1.07	1.06	1.01
28	1.35	1.22	1.16	1.10	1.04	1.02	.98
29	1.30	1.18	1.12	1.06	1.00	.98	.94
30	1.26	1.14	1.08	1.03	.97	.95	.91
31	1.22	1.10	1.05	.99	.94	.92	.88
32	1.18	1.07	1.02	.96	.91	.89	.86
33	1.14	1.04	.99	.93	.88	.86	.83
34	1.11	1.01	.96	.90	.85	.84	.81
35	1.08	.98	.93	.88	.83	.81	.78

Accuracy of the estimation procedure
Table entries are $A(n,t*)/A_0^*$
$\mu=80$, $\sigma=40$

Table 8

	t*	120	135	150	165	180	195	210
n	15	2.79	2.44	2.20	2.07	1.93	1.88	1.83
	16	2.61	2.29	2.06	1.94	1.81	1.77	1.72
	17	2.46	2.15	1.94	1.82	1.71	1.66	1.61
	18	2.32	2.03	1.83	1.72	1.61	1.57	1.52
	19	2.20	1.93	1.73	1.63	1.53	1.49	1.44
	20	2.09	1.83	1.65	1.55	1.45	1.41	1.37
	21	1.99	1.74	1.57	1.48	1.38	1.35	1.31
	22	1.90	1.66	1.50	1.41	1.32	1.28	1.25
	23	1.82	1.59	1.43	1.35	1.26	1.23	1.19
	24	1.74	1.53	1.37	1.29	1.21	1.18	1.14
	25	1.67	1.47	1.32	1.24	1.16	1.13	1.10
	26	1.61	1.41	1.27	1.19	1.12	1.09	1.06
	27	1.55	1.36	1.22	1.15	1.07	1.05	1.02
	28	1.49	1.31	1.18	1.11	1.04	1.01	.98
	29	1.44	1.26	1.14	1.07	1.00	.97	.95
	30	1.39	1.22	1.10	1.03	.97	.94	.91
	31	1.35	1.18	1.06	1.00	.94	.91	.89
	32	1.31	1.14	1.03	.97	.91	.88	.86
	33	1.27	1.11	1.00	.94	.88	.86	.83
	34	1.23	1.08	.97	.91	.85	.83	.81
	35	1.19	1.05	.94	.89	.83	.81	.78

Accuracy of the estimation procedure
Table entries are $A(n,t*)/A_0^*$
$\mu=90$, $\sigma=40$

Table 9

t*						
120	135	150	165	180	195	210

n		120	135	150	165	180	195	210
	15	3.17	2.63	2.31	2.12	2.00	1.92	1.86
	16	2.97	2.47	2.17	1.99	1.88	1.80	1.74
	17	2.79	2.32	2.04	1.87	1.76	1.69	1.64
	18	2.64	2.19	1.93	1.77	1.67	1.60	1.55
	19	2.50	2.08	1.83	1.68	1.58	1.51	1.47
	20	2.37	1.97	1.73	1.59	1.50	1.44	1.40
	21	2.26	1.88	1.65	1.52	1.43	1.37	1.33
	22	2.16	1.79	1.58	1.45	1.36	1.31	1.27
	23	2.06	1.71	1.51	1.39	1.30	1.25	1.21
	24	1.98	1.64	1.45	1.33	1.25	1.20	1.16
	25	1.90	1.58	1.39	1.27	1.20	1.15	1.12
	26	1.83	1.52	1.33	1.23	1.15	1.10	1.07
	27	1.76	1.46	1.28	1.18	1.11	1.06	1.03
	28	1.70	1.41	1.24	1.14	1.07	1.03	1.00
	29	1.64	1.36	1.20	1.10	1.03	.99	.96
	30	1.58	1.31	1.16	1.06	1.00	.96	.93
	31	1.53	1.27	1.12	1.03	.97	.93	.90
	32	1.48	1.23	1.08	1.00	.94	.90	.87
	33	1.44	1.20	1.05	.97	.91	.87	.85
	34	1.40	1.16	1.02	.94	.88	.84	.82
	35	1.36	1.13	.99	.91	.86	.82	.80

Accuracy of the estimation procedure
Table entries are $A(n,t*)/A_0^*$
$\mu=100$, $\sigma=40$

Table 10

	t*						
	120	135	150	165	180	195	210
n 15	3.72	2.96	2.51	2.25	2.07	1.95	1.87
16	3.49	2.77	2.35	2.11	1.94	1.83	1.75
17	3.28	2.61	2.21	1.98	1.82	1.72	1.65
18	3.10	2.47	2.09	1.87	1.72	1.62	1.56
19	2.94	2.34	1.98	1.77	1.63	1.54	1.48
20	2.79	2.22	1.88	1.69	1.55	1.46	1.40
21	2.66	2.11	1.79	1.61	1.48	1.39	1.34
22	2.54	2.02	1.71	1.53	1.41	1.33	1.28
23	2.43	1.93	1.63	1.47	1.35	1.27	1.22
24	2.32	1.85	1.57	1.40	1.29	1.22	1.17
25	2.23	1.78	1.50	1.35	1.24	1.17	1.12
26	2.15	1.71	1.45	1.30	1.19	1.12	1.08
27	2.07	1.64	1.39	1.25	1.15	1.08	1.04
28	1.99	1.58	1.34	1.20	1.11	1.04	1.00
29	1.92	1.53	1.30	1.16	1.07	1.01	.97
30	1.86	1.48	1.25	1.12	1.03	.97	.94
31	1.80	1.43	1.21	1.09	1.00	.94	.91
32	1.74	1.39	1.17	1.05	.97	.91	.88
33	1.69	1.34	1.14	1.02	.94	.89	.85
34	1.64	1.31	1.11	.99	.91	.86	.83
35	1.59	1.27	1.07	.96	.89	.84	.80

Accuracy of the estimation procedure
Table entries are $A(n,t*)/A_0^*$
$\mu=110$, $\sigma=40$

Table 11

		t*						
		120	135	150	165	180	195	210
n	15	4.27	3.30	2.65	2.30	2.07	1.92	1.82
	16	4.01	3.09	2.49	2.15	1.94	1.80	1.71
	17	3.77	2.91	2.34	2.03	1.82	1.70	1.61
	18	3.56	2.75	2.21	1.91	1.72	1.60	1.52
	19	3.37	2.60	2.09	1.81	1.63	1.52	1.44
	20	3.20	2.47	1.99	1.72	1.55	1.44	1.37
	21	3.05	2.36	1.89	1.64	1.48	1.37	1.30
	22	2.91	2.25	1.81	1.57	1.41	1.31	1.24
	23	2.79	2.15	1.73	1.50	1.35	1.25	1.19
	24	2.67	2.06	1.66	1.43	1.29	1.20	1.14
	25	2.56	1.98	1.59	1.38	1.24	1.15	1.09
	26	2.47	1.90	1.53	1.32	1.19	1.11	1.05
	27	2.37	1.83	1.47	1.28	1.15	1.07	1.01
	28	2.29	1.77	1.42	1.23	1.11	1.03	.99
	29	2.21	1.71	1.37	1.19	1.07	.99	.94
	30	2.14	1.65	1.33	1.15	1.03	.96	.91
	31	2.07	1.60	1.28	1.11	1.00	.93	.88
	32	2.00	1.55	1.24	1.08	.97	.90	.85
	33	1.94	1.50	1.21	1.04	.94	.87	.83
	34	1.89	1.46	1.17	1.01	.91	.85	.80
	35	1.83	1.41	1.14	.98	.89	.82	.78

Accuracy of the estimation procedure
Table entries are $A(n,t^*)/A_0^*$
$\mu=120$, $\sigma=40$

estimation procedure subject to changes in the number of
animals in the group and the length of the experiment. For
fixed μ and σ, each table has as entries $A(n,t^*)/A_o^*$ where
$A(n,t^*)$ is the area of the 95% confidence ellipse for a
given n and t^* and A_o^* is the normalized area for a 95%
confidence ellipse for $t^*=180$ days.

Discussion

In this section, a more accurate estimate means a smaller
95% confidence ellipse for (λ,μ), and we assume $\mu<t^*$.

Suppose σ is fixed. As t^* increases we gain more in
accuracy if μ is large. For a greater mean time to detection,
we censor more data by terminating an experiment early and thus
construct larger confidence regions. If t^* is large, we observe
more tumors, and thus generate more accurate confidence ellipses.
For small μ we gain little in accuracy as the length of the
experiment is increased. Since few promoted tumors will
appear late in the experiment, after $\mu+2\sigma$ days post carcinogen,
longer experiments have less influence on the estimation
procedure.

Suppose σ and t^* are fixed. The accuracy is more
sensitive to changes in n for large μ. If the mean time
to detection is large, we are censoring observations by
ending the experiment at t^* and basing our estimates on
limited data. Additional animals in the group contribute
more data with which we form a more accurate confidence
ellipse. For smaller μ, more promoted tumors are observed
and extra animals contribute less to the estimation procedure.

Suppose μ is fixed. As t* increases, we gain more in accuracy if σ is large. A large standard deviation indicates more spread in the time to detection distribution and thus a greater number of observable tumors late in the experiment. Accuracy is increased by allowing the experiment to continue, and observing more of the tail of the time to detection distribution. For small standard deviation, most of the promoted tumors are observed for fixed t* and longer experiments add little new information. Thus we gain less in accuracy as t* increases.

Suppose μ and t* are fixed. By adding animals we gain more in accuracy if σ is large. For large standard deviation, we are censoring more promoted tumors, thus extra animals furnish more data and generate more accurate estimates. For small standard deviation, most promoted tumors are observed, so increasing n adds little new information. Thus we gain less in accuracy for additional animals.

These tables illustrate the decreasing area of the 95% confidence ellipse as n increases. They also depict the trade off between n and t* in order to maintain a constant level of accuracy. We define the total animal days for an experiment to be the number of animals in the group multiplied by the length of the experiment. This is an approximate measure of the amount of resources necessary to conduct the experiment. Suppose the experimenter believes μ=80 and σ=30. If we use 25 animals for 180 days we have 4500 animal days. In order to maintain the same level of accuracy,

Table 2 shows we could use 27 animals for 150 days. In this case we have 4050 animal days - a savings of 450 animal days. Similar calculations for fixed μ, σ, and initial n and t* reveal ways to lower the number of animal days per experiment.

Table 12 further illustrates how to select t*, the length of the experiment, and n, the number of animals per group, in order to minimize animal days while maintaining a specified level of accuracy in the estimation procedure. Given μ and σ, for each value of t* we use Tables 2-11 to select an n (thus determining animal days) in order to ensure the area of the confidence ellipse for (λ, μ) is at most A_o^*.

Robustness of the Estimation Procedure

In order to derive the likelihood function for type A experiments and develop an estimation procedure, we assume the underlying parent distribution are Poisson (tumor number) and Gamma (time to detection). In this section we consider to what extent we are justified in using this estimation procedure when the underlying distributions are not Poisson and Gamma. In some experiments, the sample variance of the observed number of tumors per animal is greater than the sample mean number of observed tumors. Perhaps the number of promoted tumors may be better approximated by a negative binomial distribution. In addition, the Weibull distribution may, in some cases, describe the tumor appearance times more accurately than the Gamma distribution. We consider two experimental groups and examine how sensitive the estimation procedure is to departure from the assumptions of Poisson

Table 12

μ, σ	t* 120	135	150	165	180	195	210
80, 30	3840(32)	3915(29)	4050(27)	4290(26)	4500(25)	4875(25)	5250(25)
90, 30	4320(36)	4185(31)	4200(28)	4290(26)	4500(25)	4875(25)	5250(25)
100, 30	4920(41)	4455(33)	4350(29)	4455(27)	4500(25)	4875(25)	5040(24)
110, 30	5640(47)	4860(36)	4500(30)	4455(27)	4500(25)	4680(24)	5040(24)
120, 30	8280(69)	6480(48)	5100(34)	4785(29)	4680(26)	4875(25)	5040(24)
80, 40	4680(39)	4725(35)	4950(33)	5115(31)	5220(29)	5655(29)	5880(28)
90, 40	5040(42)	4995(37)	4950(33)	5115(31)	5220(29)	5655(29)	5880(28)
100, 40	5760(48)	5400(40)	5250(35)	5280(32)	5400(30)	5655(29)	5880(28)
110, 40	6720(56)	6075(45)	5700(38)	5610(34)	5580(31)	5850(30)	5880(28)
120, 40	7800(65)	6750(50)	6000(40)	5775(35)	5580(31)	5655(29)	5880(28)

For each combination of μ, σ, and t*, entries in the table are animal days (animal number in parentheses) necessary to maintain the area of the confidence ellipse for (λ, μ) at A_0^*.

and Gamma distributions.

For the purpose of comparing the effects of different combinations of the underlying distributions we set the following parameters for two groups. For the type A mammary tumor system described above, a typical placebo group has parameters $\lambda_1=7$, $\mu_1=95$, and $\sigma_1=40$. Since an effective chemopreventative agent should decrease the number of promoted tumors per animal and/or slow the growth rate, we define $\lambda_2=5$, $\mu_2=110$, and $\sigma_2=40$ in our treatment group. In addition, we assume the length of the experiment is 180 days, the number of animals per group is 25, and animals are palpated once every three days.

We consider the following four combinations of tumor number and time to detection distributions;

1) Poisson - Gamma (P-G)

2) Poisson - Weibull (P-W)

3) Negative Binomial - Gamma (NB-G)

4) Negative Binomial - Weibull (NB-W).

In cases 3 and 4, where the number of promoted tumors is assumed to arise from a Negative Binomial distribution, we assume $E(M)=\lambda$ and $Var(M)=2\lambda$.

For each experimental group and each combination of underlying distributions, computer simulations are used to;

1) randomly generate data from the appropriate combination of distributions,

2) estimate λ and μ assuming the underlying distributions are actually Poisson and Gamma,

3) test to see if (λ,μ) is included in the 95% confidence

ellipse generated about the point estimate $\hat{\underline{x}}^t = (\hat{\lambda}, \hat{\mu})$.

In each of 50 trials, for a given group and a given combination of distributions, if the actual value (λ, μ) lies inside the confidence ellipse, we record a success, otherwise we record a failure.

For each group the relevant hypothesis is

H_o: $p_1 = p_2 = p_3 = p_4$ where p_i is the probability of a

success for combination i

versus

H_a: at least two of the p_i's are unequal.

In order to test this hypothesis, we construct a 4x2 contingency table for each group (the data is the result of computer simulations).

Placebo Group

	Success	Failure
P-G	48	2
	43.5	6.5
P-W	43	7
	43.5	6.5
NB-G	48	2
	43.5	6.5
NB-W	35	15
	43.5	6.5

Treatment Group

	Success	Failure
P-G	46	4
	42	8
P-W	39	11
	42	8
NB-G	41	9
	42	8
NB-W	42	8
	42	8

Expected frequencies are printed below observed frequencies.

We are now able to conduct the following statistical tests.

Placebo Group

H_o: $p_1 = p_2 = p_3 = p_4$

H_a: at least two of the p_i's are unequal

Test statistic: $\chi^2 = \sum_{i=1}^{4} \sum_{j=1}^{2} \frac{(Obs_{ij} - Exp_{ij})^2}{Exp_{ij}}$

Rejection region: $\chi^2 \geq \chi_3^2(.05) = 7.815$

$$\chi^2 = .47 + .01 + .47 + 1.66 + 3.12 + .04 + 3.12 + 11.12 \qquad (37)$$
$$= 19.98$$

Therefore, we reject H_o (p<.005) at the .05 significance level.

Treatment Group

$$\chi^2 = .38 + .21 + .02 + 0 + 2 + 1.13 + .13 + 0 \qquad (38)$$
$$= 3.87$$

There are numerous combinations of length of experiment, number of animals per group, frequency of examination intervals, and distribution parameters one could consider. However, assuming the underlying distributions are Poisson and Gamma, these statistical tests are an indication that the estimation procedure is fairly robust depending on the distribution and experimental parameters. For the treatment group above, the procedure developed produces estimates of the vector (λ, μ) with equal accuracy for the four combinations. For the placebo group, when the parent distributions are Negative Binomial and Weibull, the probability of a success decreases slightly. This combination of distributions contributes the most to the statistic χ^2 in equation (37) and thus the rejection of H_o. However, 70% of the time the

true value (λ,μ) lies within the confidence ellipse constructed about $(\hat{\lambda},\hat{\mu})$. In the placebo group with $\mu_1=95$ and $\sigma_1=40$ and the length of the experiment set at 180 days, most of the promoted tumors are observed by the experiment termination date. Therefore, the data generated from combinations 2,3, and 4 is expected to be discernably different from the data generated from Poisson and Gamma distributions. This may account in part for the decrease in the proportions of successes in combinations 2 and 4.

If in fact, it is demonstrated that a different pair (P-W, NB-G, NB-W) of distributions better describes a particular experimental system, the statistical methods developed here are still applicable. The assumptions concerning the underlying distributions can be altered and a new likelihood function can be derived. As discussed previously, even if a unique maximum likelihood estimator fails to exist, as the number of observations increases the difference of two maximum likelihood estimators converges in probability to zero.

Asymptotic Power

In Tables 13-24 we display the asymptotic power of the likelihood ratio test for comparing two groups in a type A experiment as shown in Chapter IV using the non-central χ^2 distribution. We assume $\alpha'=.05$ and calculate the asymptotic power as a function of the number of animals per group and the length of the experiment. These Tables can be used to determine the number of animals per group needed to attain a certain power of the likelihood ratio test. For example we assume

Table 13

t*

n		120	135	150	165	180	195	210
	15	.11	.11	.12	.12	.12	.13	.13
	16	.11	.12	.12	.13	.13	.13	.13
	17	.12	.12	.13	.13	.14	.14	.14
	18	.12	.13	.13	.14	.14	.14	.14
	19	.12	.13	.14	.14	.15	.15	.15
	20	.13	.14	.14	.15	.15	.15	.15
	21	.13	.14	.15	.15	.16	.16	.16
	22	.14	.15	.16	.16	.16	.17	.17
	23	.14	.15	.16	.17	.17	.17	.17
	24	.15	.16	.17	.17	.18	.18	.18
	25	.15	.16	.17	.18	.18	.18	.18
	26	.15	.17	.18	.18	.19	.19	.19
	27	.16	.17	.18	.19	.19	.20	.20
	28	.16	.18	.19	.19	.20	.20	.20
	29	.17	.18	.19	.20	.20	.21	.21
	30	.17	.19	.20	.21	.21	.21	.22

Asymptotic power
Placebo group parameters: $\lambda=7$, $\mu=95$, $\sigma=40$
Treatment group parameters: $\lambda=6$, $\mu=95$, $\sigma=40$

Table 14

t*

n		120	135	150	165	180	195	210
	15	.34	.38	.40	.42	.43	.43	.44
	16	.36	.40	.43	.44	.45	.46	.46
	17	.38	.42	.45	.47	.48	.48	.49
	18	.40	.45	.47	.49	.50	.51	.51
	19	.42	.47	.50	.52	.53	.53	.54
	20	.44	.49	.52	.54	.55	.56	.56
	21	.46	.51	.54	.56	.57	.58	.58
	22	.48	.53	.56	.58	.60	.60	.61
	23	.50	.55	.58	.60	.62	.62	.63
	24	.52	.57	.60	.62	.64	.64	.65
	25	.54	.59	.62	.64	.66	.66	.67
	26	.56	.61	.64	.66	.68	.68	.69
	27	.58	.63	.66	.68	.69	.70	.71
	28	.59	.65	.68	.70	.71	.72	.72
	29	.61	.66	.70	.72	.73	.74	.74
	30	.63	.68	.71	.73	.74	.75	.76

Asymptotic power
Placebo group parameters: $\lambda=7$, $\mu=95$ $\sigma=40$
Treatment group parameters: $\lambda=5$, $\mu=95$, $\sigma=40$

Table 15

t*

n		120	135	150	165	180	195	210
	15	.72	.77	.81	.82	.84	.84	.84
	16	.75	.80	.83	.85	.86	.87	.87
	17	.78	.83	.86	.87	.88	.89	.89
	18	.80	.85	.88	.89	.90	.91	.91
	19	.83	.87	.90	.91	.92	.92	.92
	20	.85	.89	.91	.92	.93	.93	.94
	21	.87	.91	.93	.94	.94	.95	.95
	22	.88	.92	.94	.95	.95	.96	.96
	23	.90	.93	.95	.96	.96	.96	.96
	24	.91	.94	.96	.96	.97	.97	.97
	25	.92	.95	.96	.97	.97	.98	.98
	26	.93	.96	.97	.98	.98	.98	.98
	27	.94	.96	.97	.98	.98	.98	.98
	28	.95	.97	.98	.98	.99	.99	.99
	29	.96	.97	.98	.99	.99	.99	.99
	30	.96	.98	.99	.99	.99	.99	.99

Asymptotic power
Placebo group parameters: $\lambda=7$, $\mu=95$, $\sigma=40$
Treatment group parameters: $\lambda=4$, $\mu=95$, $\sigma=40$

Table 16

t*

n		120	135	150	165	180	195	210
	15	.12	.12	.12	.12	.12	.12	.12
	16	.12	.13	.13	.13	.13	.13	.13
	17	.13	.13	.13	.13	.13	.13	.13
	18	.13	.14	.14	.14	.14	.14	.14
	19	.14	.14	.15	.15	.15	.14	.14
	20	.14	.15	.15	.15	.15	.15	.15
	21	.15	.15	.16	.16	.16	.15	.15
	22	.15	.16	.16	.16	.16	.16	.16
	23	.16	.17	.17	.17	.17	.17	.16
	24	.16	.17	.17	.17	.17	.17	.17
	25	.17	.18	.18	.18	.18	.18	.17
	26	.18	.18	.19	.19	.19	.18	.18
	27	.18	.19	.19	.19	.19	.19	.19
	28	.19	.19	.20	.20	.20	.19	.19
	29	.19	.20	.20	.20	.20	.20	.20
	30	.20	.21	.21	.21	.21	.21	.20

Asymptotic power
Placebo group parameters: $\lambda=7$, $\mu=95$, $\sigma=40$
Treatment group parameters: $\lambda=7$, $\mu=100$, $\sigma=40$

Table 17

t*

n		120	135	150	165	180	195	210
n	15	.22	.22	.22	.21	.21	.21	.20
	16	.23	.23	.23	.22	.22	.22	.21
	17	.25	.24	.24	.24	.23	.23	.23
	18	.26	.26	.25	.25	.24	.24	.24
	19	.27	.27	.27	.26	.26	.25	.25
	20	.28	.28	.28	.27	.27	.27	.26
	21	.30	.30	.29	.29	.28	.28	.27
	22	.31	.31	.30	.30	.29	.29	.29
	23	.32	.32	.32	.31	.31	.30	.30
	24	.34	.34	.33	.32	.32	.31	.31
	25	.35	.35	.34	.34	.33	.33	.32
	26	.36	.36	.36	.35	.34	.34	.33
	27	.38	.37	.37	.36	.36	.35	.34
	28	.39	.39	.38	.37	.37	.36	.36
	29	.40	.40	.39	.39	.38	.37	.37
	30	.41	.41	.41	.40	.39	.39	.38

Asymptotic power
Placebo group parameters: $\lambda=7$, $\mu=95$, $\sigma=40$
Treatment group parameters: $\lambda=6$, $\mu=100$, $\sigma=40$

Table 18

t*

n		120	135	150	165	180	195	210
n	15	.49	.50	.51	.51	.51	.51	.51
	16	.51	.53	.54	.54	.54	.54	.54
	17	.54	.56	.56	.57	.57	.57	.56
	18	.57	.59	.59	.59	.59	.59	.59
	19	.59	.61	.62	.62	.62	.62	.62
	20	.62	.64	.64	.64	.64	.64	.64
	21	.64	.66	.67	.67	.67	.67	.67
	22	.67	.68	.69	.69	.69	.69	.69
	23	.69	.70	.71	.71	.71	.71	.71
	24	.71	.72	.73	.73	.73	.73	.73
	25	.73	.74	.75	.75	.75	.75	.75
	26	.75	.76	.77	.77	.77	.77	.77
	27	.76	.78	.79	.79	.79	.79	.78
	28	.78	.80	.80	.80	.80	.80	.80
	29	.80	.81	.82	.82	.82	.82	.82
	30	.81	.83	.83	.83	.83	.83	.83

Asymptotic power
Placebo group parameters: $\lambda=7$, $\mu=95$, $\sigma=40$
Treatment group parameters: $\lambda=5$, $\mu=100$, $\sigma=40$

Table 19

t*

n		120	135	150	165	180	195	210
n	15	.81	.84	.85	.86	.87	.87	.87
	16	.84	.86	.88	.89	.89	.89	.89
	17	.86	.89	.90	.90	.91	.91	.91
	18	.88	.90	.92	.92	.92	.93	.93
	19	.90	.92	.93	.94	.94	.94	.94
	20	.92	.93	.94	.95	.95	.95	.95
	21	.93	.95	.95	.96	.96	.96	.96
	22	.94	.95	.96	.96	.97	.97	.97
	23	.95	.96	.97	.97	.97	.97	.97
	24	.96	.97	.97	.98	.98	.98	.98
	25	.97	.97	.98	.98	.98	.98	.98
	26	.97	.98	.98	.99	.99	.99	.99
	27	.98	.98	.99	.99	.99	.99	.99
	28	.98	.99	.99	.99	.99	.99	.99
	29	.98	.99	.99	.99	.99	.99	.99
	30	.99	.99	.99	.99	.99	.99	.99

Asymptotic power
Placebo group parameters: $\lambda=7$, $\mu=95$, $\sigma=40$
Treatment group parameters: $\lambda=4$, $\mu=100$, $\sigma=40$

Table 20

t*

n		120	135	150	165	180	195	210
n	15	.38	.38	.36	.37	.37	.35	.34
	16	.40	.40	.38	.39	.39	.37	.36
	17	.42	.42	.40	.41	.42	.39	.39
	18	.44	.45	.43	.43	.44	.41	.41
	19	.47	.47	.45	.45	.46	.43	.43
	20	.49	.49	.47	.48	.48	.45	.45
	21	.51	.51	.49	.50	.50	.47	.47
	22	.53	.53	.51	.52	.53	.49	.49
	23	.55	.55	.53	.54	.55	.51	.51
	24	.57	.57	.55	.56	.57	.53	.53
	25	.59	.59	.57	.58	.58	.55	.54
	26	.61	.61	.59	.60	.60	.57	.56
	27	.63	.63	.61	.61	.62	.59	.58
	28	.64	.65	.62	.63	.64	.61	.60
	29	.66	.67	.64	.65	.66	.62	.61
	30	.68	.68	.66	.66	.67	.64	.63

Asymptotic power
Placebo group parameters: $\lambda=7$, $\mu=95$, $\sigma=40$
Treatment group parameters: $\lambda=7$, $\mu=105$, $\sigma=40$

Table 21

	t*						
		120	135	150	165	180	195	210

n		120	135	150	165	180	195	210
n	15	.50	.49	.45	.44	.44	.42	.41
	16	.53	.51	.48	.47	.47	.44	.43
	17	.56	.54	.50	.50	.49	.47	.46
	18	.58	.57	.53	.52	.52	.49	.48
	19	.61	.59	.55	.55	.54	.51	.51
	20	.63	.62	.58	.57	.57	.54	.53
	21	.66	.64	.60	.59	.59	.56	.55
	22	.68	.66	.62	.62	.61	.58	.57
	23	.70	.69	.65	.64	.64	.60	.59
	24	.72	.71	.67	.66	.66	.62	.61
	25	.74	.73	.69	.68	.68	.64	.63
	26	.76	.74	.71	.70	.69	.66	.65
	27	.78	.76	.72	.72	.71	.68	.67
	28	.79	.78	.74	.73	.73	.70	.69
	29	.81	.79	.76	.75	.75	.72	.71
	30	.82	.81	.77	.77	.76	.73	.72

Asymptotic power
Placebo group parameters: $\lambda=7$, $\mu=95$, $\sigma=40$
Treatment group parameters: $\lambda=6$, $\mu=105$, $\sigma=40$

Table 22

n		120	135	150	165	180	195	210
n	15	.72	.71	.69	.68	.68	.66	.66
	16	.75	.74	.72	.71	.71	.69	.69
	17	.78	.77	.75	.74	.74	.72	.72
	18	.80	.80	.77	.77	.77	.75	.74
	19	.83	.82	.80	.79	.79	.77	.77
	20	.85	.84	.82	.82	.81	.80	.79
	21	.86	.86	.84	.84	.83	.82	.81
	22	.88	.88	.86	.85	.85	.84	.83
	23	.90	.89	.87	.87	.87	.85	.85
	24	.91	.90	.89	.89	.88	.87	.87
	25	.92	.92	.90	.90	.90	.88	.88
	26	.93	.93	.91	.91	.91	.90	.89
	27	.94	.94	.92	.92	.92	.91	.91
	28	.95	.95	.93	.93	.93	.92	.92
	29	.96	.95	.94	.94	.94	.93	.93
	30	.96	.96	.95	.95	.95	.94	.94

Asymptotic power
Placebo group parameters: $\lambda=7$, $\mu=95$, $\sigma=40$
Treatment group parameters: $\lambda=5$, $\mu=105$, $\sigma=40$

Table 23

t*

n		120	135	150	165	180	195	210
n	15	.71	.70	.70	.68	.68	.64	.60
	16	.74	.73	.73	.71	.71	.67	.63
	17	.77	.76	.76	.74	.74	.70	.66
	18	.80	.78	.79	.76	.76	.72	.69
	19	.82	.81	.81	.79	.79	.75	.72
	20	.84	.83	.83	.81	.81	.77	.74
	21	.86	.85	.85	.83	.83	.79	.76
	22	.88	.87	.87	.85	.85	.81	.79
	23	.89	.88	.88	.87	.87	.83	.81
	24	.91	.90	.90	.88	.88	.85	.82
	25	.92	.91	.91	.90	.90	.87	.84
	26	.93	.92	.92	.91	.91	.88	.86
	27	.94	.93	.93	.92	.92	.89	.87
	28	.95	.94	.94	.93	.93	.90	.88
	29	.95	.95	.95	.94	.94	.91	.89
	30	.96	.95	.96	.95	.95	.92	.90

Asymptotic power
Placebo group parameters: $\lambda=7$, $\mu=95$, $\sigma=40$
Treatment group parameters: $\lambda=7$, $\mu=110$, $\sigma=40$

Table 24

t*

n		120	135	150	165	180	195	210
n	15	.79	.76	.75	.71	.70	.66	.67
	16	.82	.79	.78	.74	.73	.69	.66
	17	.84	.81	.80	.77	.76	.72	.69
	18	.86	.84	.83	.79	.79	.75	.72
	19	.88	.86	.85	.82	.81	.77	.74
	20	.90	.88	.87	.84	.83	.79	.77
	21	.92	.89	.89	.86	.85	.81	.79
	22	.93	.91	.90	.88	.87	.83	.81
	23	.94	.92	.91	.89	.88	.85	.83
	24	.95	.93	.93	.90	.90	.87	.84
	25	.96	.94	.94	.92	.91	.88	.86
	26	.96	.95	.95	.93	.92	.90	.87
	27	.97	.96	.95	.94	.93	.91	.89
	28	.97	.96	.96	.94	.94	.92	.90
	29	.98	.97	.97	.95	.95	.93	.91
	30	.98	.97	.97	.96	.96	.94	.92

Asymptotic power
Placebo group parameters: $\lambda=7$, $\mu=95$, $\sigma=40$
Treatment group parameters: $\lambda=6$, $\mu=110$, $\sigma=40$

the following;

1) the length of the experiment is 180 days,

2) the number of animals per group is the same,

3) the significance level for the likelihood ratio test

   is .05.

In addition, we suppose the experimenter has prior knowledge
of the underlying group distributions. Suppose previous
experiments indicate the placebo group's parameters are
approximately $\lambda_1=7$, $\mu_1=95$, $\sigma_1=40$, and that the chemopreventative
agent which supplements the treatment group's diet slows
tumor development. We theorize the treatment group parameters
are $\lambda_2=6$, $\mu_2=105$, $\sigma_2=40$. Therefore, in order to estimate
the number of animals necessary to attain a specified power,
we consider Table 21. For the power of the test to be at
least .70, the number of animals per group must be greater
than or equal to 27.

These tables illustrate the effect n and t* have on the
asymptotic power of the likelihood ratio test for comparing
various treatment groups with a typical placebo group. In
each of the cases considered, as n increases the asymptotic
power of the test increases. One expects a similar relation-
ship between t* and power to hold. In most of the cases con-
sidered, as t* increases the asymptotic power of the test
increases, however, for some treatment groups , for example
Table 22, as t* increases the asymptotic power of the test
decreases. This is counter-intuitive and is an interesting
observation since it suggests that longer experiments do not

necessarily ensure a greater probability of detecting overall group differences. In the cases of decreasing power as a function of increasing t^*, overall group differences are more easily detected earlier in the experiment. As t^* increases the underlying distributions, especially the continuous tumor time distribution, appear increasingly similar. Thus the asymptotic power of the likelihood ratio test decreases.

Example

In the paper by Thompson [14], the effect of the chemopreventative agent retinyl acetate on the number of induced tumors and tumor growth rate was studied. The study was undertaken in order to determine if continuous treatment with retinyl acetate is necessary to sustain mammary tumor inhibition. The study is a type A experiment, specifically a mammary tumor system using rats as experimental animals. All rats received an i.v. injection of the direct-acting carcinogen MNU and animals were palpated for the detection of mammary tumors twice each week.

We apply the statistical techniques developed in this dissertation to the data compiled in that experiment. Group 1 in Thompson's experiment was the placebo group. The animals in group 5 were tumor bearing at 60 days of age and received a diet supplemented with retinyl acetate throughout the experiment. The investigation was terminated 182 days after the injection of carcinogen.

The statistical techniques developed here produce the following estimates of the underlying distribution parameters;

	Group 1	Group 5
$\hat{\lambda}$	7.632	6.234
$\hat{\alpha}$	5.367	5.589
$\hat{\beta}$	15.209	18.219
$\hat{\mu}$	81.621	101.823

We conduct the following hypothesis test to check for an overall group difference.

H_o: groups 1 and 5 arise from the same population,

$$(\lambda_1, \alpha_1, \beta_1) = (\lambda_5, \alpha_5, \beta_5)$$

H_a: groups 1 and 5 arise from different populations,

$$(\lambda_1, \alpha_1, \beta_1) \neq (\lambda_5, \alpha_5, \beta_5)$$

Test statistic: $\chi^2 = -2\ln r$ where $r = L(\hat{H}_o)/L(\hat{H}_a)$

Rejection region: $\chi^2 \geq \chi_3^2(.05) = 7.815$

Carrying out the hypothesis test yields the test statistic $\chi^2 = 11.22$. Therefore, we would reject the null hypothesis $(p \approx .01)$ at the .05 level and conclude the two groups arise from different populations.

In order to determine which factor(s) is(are) contributing to the overall group difference, we construct joint 95% confidence intervals for the differences $\lambda_1 - \lambda_5$ and $\mu_1 - \mu_5$.

$$(\hat{\lambda}_1 - \hat{\lambda}_5) \pm z_{\alpha'/4} \, \hat{\sigma}_{11} = (7.632 - 6.234) \pm (2.24)(1.0886)$$
$$= 1.398 \pm 2.4384$$

Therefore, the confidence interval for $\lambda_1 - \lambda_5$ is

(-1.04, 3.836) .

$$(\hat{\mu}_1 - \hat{\mu}_5) \pm z_{\alpha'/4} \, \hat{\sigma}_{22} = (81.621 - 101.823) \pm (2.24)(6.721)$$
$$= -20.202 \pm 15.055$$

Therefore, the confidence interval for $\mu_1 - \mu_5$ is

(-35.257, -5.147) .

Since zero is included in the first confidence interval,
we conclude at the 5 percent joint significance level that
the mean number of promoted tumors per animal in groups 1
and 5 are not different. However, the other 95% simultaneous
confidence interval indicates different mean times to tumor
detection. This analysis suggests that the mean time to
detection, related to the rate of tumor development, contributes
to the group differences. Thus, contrary to Thompson we
conclude that retinyl acetate has significantly slowed the
growth rate of promoted tumors in the tumor bearing animals
as evidenced by the increase in the mean time to tumor
detection.

CHAPTER VI

SUMMARY AND EXTENSIONS

This dissertation presents new methods for analyzing type A and type B experimental tumor systems. There are several related topics that can now be investigated. For both experiment types, the method of maximum likelihood is used to estimate the parameters characterizing the two underlying distributions - tumor number and time to detection. The likelihood equations are derived, however, they cannot be solved explicitly. An efficient computer algorithm is needed to accurately estimate the values of these parameters. A preliminary computer program is used to simulate experiments and to estimate the distribution parameters for the example in Chapter V. However, further work is necessary to provide the experimenter with an interactive program to accommodate raw data. This computer program should have the capacity to perform the following calculations;

1) summarize the experimental data, (sample means, medians, variances,...)

2) estimate the distribution parameters,

3) conduct simple one-sample hypothesis tests,

4) conduct a likelihood ratio test for the comparison of two groups, calculate the asymptotic power of the test, and isolate any group differences.

In analyzing type A experiments, one of the assumptions

is that all animals survive until the end of the experiment.
This assumption is somewhat restrictive as often a few animals
die before the experiment is terminated. Current statistical
procedures offer no method for including these animals in the
subsequent analysis. Those animals that die early still
supply valuable data and should not be discounted. The
mathematical model characterizing type A experiments can
accommodate this problem by altering one assumption; assume
animal i survives until time t_i^*. This leads to a new set of
likelihood equations, variance-covariance matrix for the
estimator, etc, which requires further study.

For each type of experiment, a better understanding of
the effect of the frequency of examination intervals on the
estimation procedure is needed. For experiments of type A,
examining animals every three days may not be necessary in
order to produce accurate estimates. Relaxing this restriction
could save valuable time and resources. For type B experiments,
there may exist a set of optimum times to sacrifice animals
and thus maximize the accuracy of the estimation procedure.
Having prior knowledge of the underlying distributions could
lead to the selection of examination intervals that ensure
accurate estimates.

The derivation of a cost function for experiments of
this type would help the experimenter realize the optimum
number of animals per group and length of the experiment.
Although we have shown how sensitive the estimation procedure
is to the number of animals per group and the length of the

experiment, a cost function would lead to a "best" choice of these experimental parameters.

Further study is also needed regarding the comparison of two treatment groups. We suppose the following;
1) a fixed number of animals are available for experimentation,
2) the experimenter has prior knowledge about the underlying distributions for each group.
These assumptions may lead to an optimum way to divide the animals between the two groups. This animal distribution procedure would be used to maximize the power of the likelihood ratio test and enable the experimenter to better detect overall group differences.

Many type A and type B experiments involve the study of more than two treatment groups simultaneously. This dissertation provides a method for comparing any two groups, and isolating group differences. Further statistical methods are needed in order to compare three or more groups simul-taneously. This procedure should establish which pairs of groups contribute to an experiment difference, and which factors are causing the differences, while maintaining a constant significance level.

It may be possible to use sequential testing in experiments of this type in order to compare treatment groups. This would involve conducting a statistical test after each examination period. At each examination interval we could accept or reject the null hypothesis that the groups arise from the same population and terminate the experiment, or

continue the experiment until the next examination time. This kind of statistical procedure provides the potential for saving additional time and resources.

LIST OF REFERENCES

1. Drinkwater, N.R., and Klotz, J.H., Cancer Res. 41: 113-119, 1981.

2. Mendenhall, W., Scheaffer, R.L., and Wackerly, D.D., Mathematical Statistics with Applications, p. 414, Boston, MA., Duxbury Press, 1981.

3. Snedecor, G.W., and Cochran, W.G., Statistical Methods, Ed. 6, pp. 235, 325, Ames, Iowa, The Iowa State University Press, 1967.

4. Mantel, N., Cancer Chemother. Rept. 50: 163, 1966.

5. Cox, D.R., J. Royal Stat. Soc. B, 34: 187, 1972.

6. Puri, M.L., and Sen, P.K., Nonparametric Methods in Multivariate Analysis, pp. 184-185, New York, John Wiley and Sons.

7. Kalbfleish, J.D., and Prentice, R.L., The Statistical Analysis of Failure Time Data, New York, John Wiley and Sons, 1980.

8. Gail, M.H., Santer, T.J., and Brown, C.C., Biometrics 36: 255-266, 1980.

9. Kendall, M.G., and Stuart, A., The Advanced Theory of Statistics, Vol.2, London, Charles Griffin and Co. Ltd., 1961.

10. Bury, K.V., Statistical Models in Applied Science, New York, John Wiley and Sons, 1975.

11. Rice, J.R., Numerical Methods, Software, and Analysis: IMSL Reference Edition, New York, McGraw-Hill Book Co., 1983.

12. Tatsuoka, M.M., Multivariate Analysis-Techniques for Educational and Psychological Research, New York, John Wiley and Sons, 1971.

13. Johnson, R.A., and Wischern, D.W., Applied Multivariate Statistical Analysis, Englewoods Cliffs, Prentice-Hall, Inc., 1982.

14. Thompson, H.J., Becci, P.J., Brown, C.C., and Moon, R.C., Effect of the duration of Retinyl Acetate Feeding on Inhibition of 1-Methyl-1-nitrosourea-induced Mammary Carcinogenesis in the Rat, Cancer Res. 39: 3977-3980, 1979.

15. Gullino, P., Pettigrew, H.M., and Grantham, F.H., Nitrosomethylurea as a Mammary Gland Carcinogen in Rats, J. Natl. Cancer Inst. 54, 1975.

16. Meeker, L.D., Thompson, H.J., and Herbst, E.J., A Multi-Stage Model of Mammary Tumorigensis, Growth and Chemoprevention, Unpublished result, University of New Hampshire, 1982.

17. Lee, E.T., Statistical Methods for Survival Data Analysis, Belmont, CA., Lifetime Learning Publications, 1980.

18. Bishop, Y.M.M., Fienberg, S.E., and Holland, P.W., Discrete Multivariate Analysis, Cambridge, MA., The MIT Press, 1975.

19. Bickel, P.J., and Doksum, K.A., Mathematical Statistics: Basic Ideas and Selected Topics, San Francisco, Holden-Day Inc., 1977.

20. Anderson, T.W., An Introduction to Multivariate Statistical Analysis, New York, John Wiley and Sons, 1962.