SOME RESULTS ON TESTS OF SEPARATE FAMILIES OF HYPOTHESES

by

Basilio de Bragança Pereira

Thesis submitted for the
Diploma of Membership of Imperial College
and the degree of
Doctor of Philosophy in the University of London

June 1976

To Iracema and Marcel

ABSTRACT

Chapter 1 contains an introduction to the problem of discriminating between alternative statistical models, and reviews previous work.

Chapter 2 is devoted to a comparison in the single sample case between the asymptotic procedures proposed by Cox and by Atkinson. General results are obtained on the consistency of the tests derived from the two methods. The adequacy of the asymptotic results for finite samples is investigated and some conclusions reached, through examination of the terms which differentiate the two procedures. Empirical results are also discussed. The two methods are used to derive tests and for these, empirical simulated results are obtained for the first four moments, the power and the significance level attained. From the analytical and empirical results general conclusions are given.

In Chapter 3 a generalization of Cox's method is used to derive tests for regression models. The tests developed are generalizations of those given in Chapter 2. The efficiency of the estimators of the regression coefficients when using a false model in relation to the true model is investigated. An example of the choice of a survival model for patients with a brain tumour is given. Finally, it is shown that Cox's method can be generalized for dependent observations forming a Markov process and some related applications are suggested.

ACKNOWLEDGMENTS

I am very grateful to my supervisor Professor D.R. Cox for his guidance and encouragement throughout the period of this work and to whom I owe much of my way of looking at statistics. I wish to thank Mr. V.T. Farewell for helpful discussions and Dr. A.M. Herzberg for careful comments on the form of this thesis.

This research was carried out while I was on leave from Universidade Federal do Rio de Janeiro and on receipt of a scholarship from Coordenação de Aperfeiçoamento do Pessoal de Nível Superior, Brasil. Both are acknowledged most gratefully.

TABLE OF CONTENTS

Chapter 1

INTRODUCTION AND BACKGROUND

1.1 Preliminaries

Let $\underset{\sim}{y} = (y_1,\dots,y_n)$ be independent observations from some unknown distribution F. Suppose that it is desired to test the null hypothesis $H_f : F \in \mathfrak{F}_f$, where $\mathfrak{F}_f$ is a family of probability distributions having density $f(\underset{\sim}{y},\underset{\sim}{\alpha})$ against the alternative hypothesis $H_g : F \in \mathfrak{F}_g$, where $\mathfrak{F}_g$ is another family of probability distributions having density $g(\underset{\sim}{y},\underset{\sim}{\beta})$.

The families $\mathfrak{F}_f$ and $\mathfrak{F}_g$, are assumed separate, i.e. an arbitrary member of one family cannot be obtained as the limit of members of the other. Here $\underset{\sim}{\alpha}$ and $\underset{\sim}{\beta}$ are unknown vector parameters indexing the members of the families. This problem was first considered by Cox (1961, 1962) who developed an asymptotic test for this situation based on the maximum likelihood ratio.

If H_f is the null hypothesis and H_g the alternative the test statistic considered was

$$T_f^*(C) = \ell_f(\hat{\underset{\sim}{\alpha}}) - \ell_g(\hat{\underset{\sim}{\beta}}) - E_{\hat{\underset{\sim}{\alpha}}}\{\ell_f(\underset{\sim}{\alpha}) - \ell_g(\underset{\sim}{\beta}_{\underset{\sim}{\alpha}})\} \quad , \qquad (1.1.1)$$

where $\hat{\underset{\sim}{\alpha}}$ and $\hat{\underset{\sim}{\beta}}$ are respectively the maximum likelihood estimators of $\underset{\sim}{\alpha}$ and $\underset{\sim}{\beta}$, $\ell_f(\hat{\underset{\sim}{\alpha}}) = \log f(\underset{\sim}{y},\hat{\underset{\sim}{\alpha}})$, $\ell_g(\hat{\underset{\sim}{\beta}}) = \log g(\underset{\sim}{y},\hat{\underset{\sim}{\beta}})$, $\underset{\sim}{\beta}_{\underset{\sim}{\alpha}}$ is the probability limit of $\hat{\underset{\sim}{\beta}}$ under H_f, $E_{\underset{\sim}{\alpha}}$ denotes the expected value under H_f and

$$E_{\underset{\sim}{\alpha}}\left\{\frac{\partial \log g(\underset{\sim}{y},\underset{\sim}{\beta}_{\underset{\sim}{\alpha}})}{\partial \underset{\sim}{\beta}'}\right\} = \underset{\sim}{0} \quad . \qquad (1.1.2)$$

Cox showed that under certain conditions $T_f^*(C)$ is asymptotically normally distributed with mean zero and variance

$$V_{\underset{\sim}{\alpha}}\{T_f^*\} = V_{\underset{\sim}{\alpha}}\{\ell_f(\underset{\sim}{\alpha}) - \ell_g(\underset{\sim}{\beta}_{\underset{\sim}{\alpha}})\} - \underset{\sim}{C}_f' \, \underset{\sim}{V}_f \, \underset{\sim}{C}_f \quad , \qquad (1.1.3)$$

where

$$\underset{\sim}{C}_f = \operatorname{Cov}_{\underset{\sim}{\alpha}}\left\{\ell_f(\underset{\sim}{\alpha}) - \ell_g(\underset{\sim}{\beta}_{\alpha});\ \frac{\partial \ell_f(\underset{\sim}{\alpha})}{\partial \underset{\sim}{\alpha}}\right\},\quad V_f^{-1} = V_{\underset{\sim}{\alpha}}\left\{\frac{\partial \ell_f(\underset{\sim}{\alpha})}{\partial \underset{\sim}{\alpha}}\right\}.$$

When H_g is the null hypothesis and H_f is the alternative hypothesis, the test statistic is, in an analogous notation,

$$T_g^*(C) = \ell_g(\hat{\underset{\sim}{\beta}}) - \ell_f(\hat{\underset{\sim}{\alpha}}) - E_{\hat{\underset{\sim}{\beta}}}\{\ell_g(\underset{\sim}{\beta}) - \ell_f(\underset{\sim}{\alpha}_{\beta})\} \tag{1.1.4}$$

which is asymptotically normally distributed with mean zero and variance

$$V_{\underset{\sim}{\beta}}\{T_g^*\} = V_{\underset{\sim}{\beta}}\{\ell_g(\underset{\sim}{\beta}) - \ell_f(\underset{\sim}{\alpha}_{\beta})\} - \underset{\sim}{C}_g'\, \underset{\sim}{V}_g\, \underset{\sim}{C}_g \quad , \tag{1.1.5}$$

where

$$\underset{\sim}{C}_g = \operatorname{Cov}_{\underset{\sim}{\beta}}\left\{\ell_g(\underset{\sim}{\beta}) - \ell_f(\underset{\sim}{\alpha}_{\beta}),\ \frac{\partial \ell_g(\underset{\sim}{\beta})}{\partial \underset{\sim}{\beta}}\right\},\ V_g^{-1} = V_{\underset{\sim}{\beta}}\left\{\frac{\partial \ell_g(\underset{\sim}{\beta})}{\partial \underset{\sim}{\beta}}\right\}.$$

Here $\underset{\sim}{\alpha}_{\beta}$ is the probability limit of $\hat{\alpha}$ under H_g.

Another approach suggested by Cox was based on the comprehensive family of density functions which are proportional to

$$\{f(\underset{\sim}{y},\underset{\sim}{\alpha})\}^{\lambda}\ \{g(\underset{\sim}{y},\underset{\sim}{\beta})\}^{1-\lambda}$$

which reduces to H_f and H_g in the special cases when $\lambda = 1,0$. This approach was developed by Atkinson (1970). He derived a test based on the score function for λ. The distribution of the test statistic was derived under the null hypothesis $\lambda=1$ (or $\lambda=0$) and for this a consistent estimator for $\underset{\sim}{\beta}$ (or $\underset{\sim}{\alpha}$) was chosen. He has shown that under the null hypothesis these tests statistics are asymptotically equivalent to Cox's test statistics. The resulting test statistic is

$$T_f^*(A) = \ell_f(\hat{\underset{\sim}{\alpha}}) - \ell_g(\underset{\sim}{\beta}_{\hat{\alpha}}) - E_{\hat{\underset{\sim}{\alpha}}}\{\ell_f(\underset{\sim}{\alpha}) - \ell_g(\underset{\sim}{\beta}_{\alpha})\}\ , \tag{1.1.6}$$

which under H_f is also asymptotically normally distributed with mean zero and variance again given by (1.1.3). Here $\underset{\sim}{\beta}_{\hat{\alpha}}$ is a consistent estimator for $\underset{\sim}{\beta}_{\alpha}$.

When H_g is the null hypothesis and H_f the alternative, the test statistic is

$$T_g^*(A) = \ell_g(\hat{\underset{\sim}{\beta}}) - \ell_f(\alpha_{\hat{\underset{\sim}{\beta}}}) - E_{\hat{\underset{\sim}{\beta}}}\{\ell_g(\underset{\sim}{\beta}) - \ell_f(\alpha_{\underset{\sim}{\beta}})\} , \tag{1.1.7}$$

which is asymptotically normally distributed with mean zero and variance given by (1.1.5). Here $\alpha_{\hat{\underset{\sim}{\beta}}}$ is the estimator for $\alpha_{\underset{\sim}{\beta}}$.

We can, therefore, consider

$$T_f(j) = T_f^*(j)\ [V_{\underset{\sim}{\alpha}}\{T_f^*\}]^{-\frac{1}{2}} , \quad T_g(j) = T_g^*(j)\ [V_{\underset{\sim}{\beta}}\{T_g^*\}]^{-\frac{1}{2}} , \tag{1.1.8}$$

for $j = A,C$, as approximately standard normal variates and perform the tests in the following way. A large negative value of $T_f(.)$ indicates a departure from H_f in the direction of H_g. A large negative value of $T_g(.)$ indicates a departure from H_g in the direction of H_f. Large negative values or large positive values for both T_f and T_g would indicate that the sample is inconsistent with both H_f and H_g. A large negative value of one of $T_f(.)$ and $T_g(.)$ together with a large positive value of the other would also indicate departure from both models.

It is assumed that observations are to be used to test the null hypothesis H_f and that it is required to have high power for the particular alternative hypothesis H_g. In addition to the answer to the tests it is also useful to look at the numerical value of the log likelihood ratio $\ell_f(\hat{\underset{\sim}{\alpha}}) - \ell_g(\hat{\underset{\sim}{\beta}})$, which is of direct interest in a pure discrimination problem.

For the remainder of Chapter 1 some properties of the models frequently used in later chapters will be considered. At the end of the chapter some related work is reviewed.

In Chapter 2 the tests of separate families of hypothesis are considered in the case of independent identically distributed observations and a comparison is made between the procedures proposed by Cox and by Atkinson. General results are obtained on the consistency of the tests derived from the two procedures. It is shown that under the alternative hypothesis Atkinson's test is not always consistent. The adequacy of the asymptotic results for finite samples are investigated and some conclusions reached, through examination of the terms which differentiates the two procedures.

Empirical results are also discussed. Cox derived test statistics in the case of the lognormal distribution versus the exponential distribution and for the complementary problem. Jackson (1968) used Cox's method and derived tests for the case of the lognormal distribution versus the gamma distribution and conversely. Atkinson used his method and derived a test for the case of the exponential distribution versus the lognormal distribution. Atkinson's method is used to derive new tests for the cases given by Jackson and for the case of the lognormal versus the exponential distributions. Further new tests are developed using both Cox's and Atkinson's methods for the lognormal distribution versus the Weibull distribution and conversely, and for the case of the gamma distribution versus the Weibull distribution and conversely. For the tests presented, empirical simulated results are obtained for the first four moments, the power and the significance level attained.

From the analytical and empirical comparisons it is concluded that generally Cox's method is expected to perform rather better than Atkinson's method.

In Chapter 3 a generalization of Cox's method is usedto derive tests for independent but not identically distributed observations. The tests developed in this part are generalizations of those given in Chapter 2 for the case in which the models contain regression covariates. The efficiency of the estimators of the regression coefficient when using a false model in relation to the true model is investigated. It is found that asymptotically the test statistics do not depend on the design matrix and the design problem is separated from distributional assumptions.

An example of the choice of a survival model for patients with a brain tumour is given.

Finally, it is shown that Cox's method can be generalized to the case of dependent observations forming a Markov process and some applications are suggested.

1.2 Maximum likelihood estimation for survival models

In this section the distributions and the regression models are presented for which tests are developed later in Chapter 2 and 3. Some results and properties of the maximum likelihood estimators are given briefly. For a concise presentation the results for the survival models are derived from those for the generalized gamma regression model.

The generalized gamma regression model can be written

$$f(y_i;a,b,k,\underset{\sim}{\theta}') = \left\{\frac{b}{\Gamma(k)}\right\}\left\{\frac{a\ e^{\underset{\sim}{z}_i\underset{\sim}{\theta}}}{k}\right\}^{-bk} y_i^{\,bk-1} \exp -\left\{\left(\frac{ky_i}{a\ e^{\underset{\sim}{z}_i\underset{\sim}{\theta}}}\right)^b\right\} \quad (1.2.1)$$

for $y_i>0$, $a,b,k>0$ and $\underset{\sim}{\theta}' = (\theta_2,\ldots,\theta_m)$. It would be possible to generalize the dependence on $\underset{\sim}{\theta}$ in (1.2.1), for example by replacing $a\ e^{\underset{\sim}{z}_i\underset{\sim}{\theta}}$ by $h(z_i,\underset{\sim}{\theta})$ for some known function $h(\ .\ )$. The properties and fitting of such models will not be explored here.

For n independent observations $(y_1,\ldots,y_n)$ we assume $\sum_{i=1}^{n} z_{ij} = 0$ $(j = 1,\ldots,n)$ and that $\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \underset{\sim}{z}_i' \underset{\sim}{z}_i = \lim_{n\to\infty} \frac{1}{n} Z'Z$ is a bounded positive definite matrix. Model (1.2.1) is log-linear in that $x = \log y$ can be written

$$x = \log \frac{a}{k} + \underset{\sim}{z}\underset{\sim}{\theta} + \psi(k)b^{-1} + k^{-\frac{1}{2}}b^{-1}k^{\frac{1}{2}}\ \{w - \psi(k)\}\ ,$$

with

$$f(w,k) = \frac{1}{\Gamma(k)} \exp\{kw\} \exp\{-e^w\}\ , \quad (1.2.2)$$

where

$$\psi(x) = d\{\log\Gamma(x)\}/dx \quad \text{etc.}$$

Let $\alpha = \log \frac{a}{k} + \psi(k)$, $\sigma = k^{-\frac{1}{2}} b^{-1}$, $q = k^{-\frac{1}{2}}$ and $e = k^{\frac{1}{2}} \{w - \psi(k)\}$ (Prentice 1974). This parameterization allows the limiting case as $k \to \infty$ to be mapped to the origin ($q = 0$) and the class to be extended to negative q, still maintaining a regular estimation problem. The model can then be written

$$x = \alpha + \underset{\sim}{z}\underset{\sim}{\theta} + \sigma e,$$

with

$$f(e;q) = \begin{cases} \dfrac{|q|}{\Gamma(q^{-2})} \exp[q^{-2}\{qe + \psi(q^{-2})\} - \exp\{qe + \psi(q^{-2})\}] & (q \neq 0), \\ (2\pi)^{-\frac{1}{2}} \exp\{-\frac{1}{2} e^2\} & (q = 0). \end{cases} \tag{1.2.3}$$

The model (1.2.3) is in the form of a conditional structural model with an additional quantity q (Fraser, 1968, Ch. 4). Fraser's structural analysis could be used for inferences about (σ, q). For example the marginal likelihood function for q is formally proportional to

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \prod_{i=1}^{n} f\left(\frac{x_i - \alpha - \underset{\sim}{z}_i \underset{\sim}{\theta}}{\sigma} ; q \right) \frac{d\alpha d\theta' d}{\sigma^{(n+1)}}$$

and generally only the integral over α can be performed analytically. Even for the simple case when q is known, approximations to simplify the calculations were used by Prentice (1973). An alternative approach would be via the maximized relative likelihood function obtained by maximizing the likelihood function over $(\alpha, \underline{\theta})$ at specified values of (σ, q) but this does not take account of the uncertainty in (α, θ). Here instead in view of the purposes of this section the classical maximum likelihood results obtained by Prentice (1974) for (1.2.3) are used.

From (1.2.3) the log likelihood function for data $\underset{\sim}{y} = (y_1, \ldots, y_n)$ is

$$\ell(\alpha,\sigma,q,\underset{\sim}{\theta}';\underset{\sim}{y}) = \begin{cases} n\log|q| - n\log\sigma - n\log\Gamma(q^{-2}) - \sum_{i=1}^{n}\log y_i + \sum_{i=1}^{n}\left[\dfrac{\log y_i-\alpha-\underset{\sim}{z}_i\underset{\sim}{\theta}}{\sigma}\right]q^{-1} \\ \qquad +n\psi(q^{-2})q^{-2} - \sum_{i=1}^{n}\exp\left\{\left(\dfrac{\log y_i-\alpha-\underset{\sim}{z}_i\underset{\sim}{\theta}}{\sigma}\right)q+\psi(q^{-2})\right\} & (q\neq 0), \\ -n\log\sigma^2 - n\log\sqrt{2\pi} - \sum_{i=1}^{n}\log y_i - \dfrac{1}{2\sigma^2}\sum_{i=1}^{n}(\log y-\alpha-\underset{\sim}{z}_i\underset{\sim}{\theta})^2 & (q=0). \end{cases}$$

(1.2.4)

The expression for $q\neq 0$ is differentiable with respect to $\alpha,\sigma,\underset{\sim}{\theta}',q$ at $q=0$, and the maximum likelihood estimators of $\alpha,\sigma,\underset{\sim}{\theta}'$ for $q=0$ can be obtained from any of the two expressions in (1.2.4).

The information matrix corresponding to the maximum likelihood estimators of $(\alpha,\sigma,q,\underset{\sim}{\theta}')$ is

$$\begin{bmatrix} I(\alpha,\sigma,q) & \\ & \frac{1}{\sigma^2}Z'Z \end{bmatrix}, \qquad (1.2.5)$$

where

$$I(\alpha,\sigma,q) = \begin{bmatrix} I_{11} & I_{12} & I_{13} \\ I_{12} & I_{22} & I_{23} \\ I_{13} & I_{23} & I_{33} \end{bmatrix},$$

$$I_{11} = \frac{1}{\sigma},\ I_{22} = \frac{1}{\sigma^2}\left(\frac{\psi'(q^{-2})}{q^2}+1\right),\ I_{33} = \frac{1}{q^2} - \frac{3\psi'(q^{-2})}{q^4} + \frac{4\psi'(q^{-2})}{q^8}(\psi(q^{-2})-q^2)$$

$$I_{12} = \frac{q}{\sigma^2},\ I_{13} = \frac{2}{\sigma}\left(\frac{\psi'(q^{-2})}{q^4} - \frac{1}{q^2}\right) - \frac{1}{\sigma},\ I_{23} = \frac{1}{\sigma q}\left(\frac{\psi'(q^{-2})}{q^2} - 1\right). \qquad (1.2.6)$$

In Section 1.1 it was mentioned that in the later chapters tests are derived for a null hypothesis H_f when high power is required for the particular alternative H_g. The results (1.2.4), (1.2.5) and (1.2.6) are all that is needed to present the models to be used later. The parameterization most commonly used will be chosen and for these new parameters the corresponding information matrix is found by a straightforward application of the chain rule for derivatives to (1.2.5). Substitution of these new parameters in (1.2.4) will give the log likelihood functions of the models of interest. From these the maximum likelihood estimators are obtained.

A Lognormal survival models

(i) For q=0 and $\underset{\sim}{\theta}=\underset{\sim}{0}$, (1.2.5) becomes

$$I(\alpha,\sigma;q=0) = n\begin{bmatrix} 1/\sigma^2 & 0 \\ 0 & 2/\sigma^2 \end{bmatrix}$$

and, for the transformation $\alpha=\alpha_1$ and $\sigma = \sqrt{\alpha_2}$,

$$I(\alpha_1,\alpha_2) = n\begin{bmatrix} 1/\alpha_2 & 0 \\ 0 & 1/(2\alpha_2^2) \end{bmatrix}. \tag{1.2.7}$$

The log likelihood function obtained from (1.2.4) and the maximum likelihood estimates of (α_1,α_2) are

$$\ell_L(\alpha_1,\alpha_2;y) = -\frac{n}{2}\log\alpha_2 - n\log\sqrt{2\pi} - \sum_{i=1}^{n}\log y_i - \frac{1}{2\alpha_2}\sum_{i=1}^{n}(\log y_i-\alpha_1)^2,$$

$$\hat{\alpha}_1 = \frac{\sum_{i=1}^{n}\log y_i}{n}, \qquad \hat{\alpha}_2 = \frac{\sum_{i=1}^{n}(\log y_i-\hat{\alpha}_i)^2}{n}. \tag{1.2.8}$$

The corresponding density function will be denoted by $f_L(y;\alpha_1,\alpha_2)$.

(ii) For q=0 and $\underset{\sim}{\theta}$ arbitrary, (1.2.5) becomes

$$I(\alpha,\sigma,\theta;q=0) = \begin{bmatrix} I(\alpha,\sigma;q=0) & 0 \\ 0 & \frac{1}{\sigma^2}Z'Z \end{bmatrix}$$

and, for the transformation $\alpha=\alpha_1$, $\sigma = \sqrt{\alpha_2}$ and $\underset{\sim}{\theta}' = \underset{\sim}{a}'$,

$$I(\alpha_1,\alpha_2,\underset{\sim}{a}') = \begin{bmatrix} I(\alpha_1,\alpha_2) & 0 \\ 0 & \frac{1}{\alpha_2}Z'Z \end{bmatrix}. \tag{1.2.9}$$

By writing $\underset{\sim}{L}' = (\log y_i,\ldots,\log y_n)$, the log likelihood function obtained from (1.2.4) and the maximum likelihood estimators of $(\alpha_1,\alpha_2,\underset{\sim}{a}')$ are

$$\ell_L(\alpha_1,\alpha_2,\underset{\sim}{a}';\underset{\sim}{y}) = -\frac{n}{2}\log\alpha_2 - n\log\sqrt{2\pi} - \sum_{i=1}^{n}\log y_i - \frac{1}{2\alpha_2}\sum_{i=1}^{n}(\log y_i-\alpha_1-\underset{\sim}{z}_i\underset{\sim}{a})^2,$$

$$\hat{\alpha}_1 = \frac{\sum_{i=1}^{n}\log y_i}{n}, \quad \hat{\underset{\sim}{a}} = (Z'Z)^{-1}\underset{\sim}{Z}\,L, \quad \hat{\alpha}_2 = \frac{1}{n}(\underset{\sim}{L}-\alpha_1\underset{\sim}{1}-\underset{\sim}{Z}\hat{\underset{\sim}{a}})^1(\underset{\sim}{L}-\hat{\alpha}_1\underset{\sim}{1}-\underset{\sim}{Z}\hat{\underset{\sim}{a}}). \tag{1.2.10}$$

The corresponding density function will be denoted by $f_L(y_i;\alpha_1,\alpha_2,\underset{\sim}{a}')$.

B Weibull survival models

(i) For q=1 and $\underset{\sim}{\theta}=0$, (1.2.5) becomes

$$I(\alpha,\sigma,q=1) = n\begin{bmatrix} 1/\sigma^2 & 1/\sigma^2 \\ 1/\sigma^2 & 1/\sigma^2\{\psi'(1)+1\}\end{bmatrix}$$

and, for the transformation $\alpha = \log\beta_1 + \frac{\psi(1)}{\beta_2}$ and $\sigma = \beta_2^{-1}$,

$$I(\beta_1,\beta_2) = n\begin{bmatrix} \left(\frac{\beta_2^2}{\beta_1}\right) & -\frac{\psi(2)}{\beta} \\ -\frac{\psi(2)}{\beta_1} & \frac{\psi'(1)+\{\psi(2)\}^2}{\beta_2^2}\end{bmatrix} . \tag{1.2.11}$$

Here, the log likelihood function and the maximum likelihood estimators of (β_1,β_2) are

$$\ell_W(\beta_1,\beta_2,\underset{\sim}{y}) = n\log\beta_2 - n\beta_2\log\beta_1 + (\beta_2-1)\sum_{i=1}^{n}\log y_i - \sum_{i=1}^{n}\left(\frac{y_i}{\beta_1}\right)^{\beta_2},$$

$$\hat{\beta}_1^{\hat{\beta}_2} = \frac{\sum_{i=1}^{n} y_i^{\hat{\beta}1}}{n}, \qquad \hat{\beta}_2 = \left[\frac{\sum_{i=1}^{n} y_i^{\hat{\beta}_2}\log y_i}{\sum_{i=1}^{n} y_i^{\hat{\beta}_2}} - \frac{\sum_{i=1}^{n}\log y_i}{n}\right]^{-1}. \tag{1.2.12}$$

The corresponding density function will be denoted by $f_W(y_,;\beta_1,\beta_2)$.

(ii) For q=1 and $\underset{\sim}{\theta}$ arbitrary, (1.2.5) becomes

$$I(\alpha,\sigma,\theta;q=1) = \begin{bmatrix} I(\alpha,\sigma;q=1) & 0 \\ 0 & \frac{1}{\sigma^2}Z'Z\end{bmatrix}$$

and, for the transformation $\alpha=\beta_1+\frac{\psi(1)}{\beta_2}$, $\sigma=\beta_2^{-1}$ and $\underset{\sim}{\theta}=\underset{\sim}{b}'$,

$$I(\beta_1,\beta_2,\underset{\sim}{b}') = \begin{bmatrix} n\beta_2^2 & -n\psi(2) & 0 \\ -n\psi(2) & n\frac{\psi'(1)+\{\psi(2)\}^2}{\beta_2^2} & 0 \\ 0 & 0 & \beta_2^2 Z'Z \end{bmatrix} . \tag{1.2.13}$$

The log likelihood function and the maximum likelihood estimators of $(\beta_1,\beta_2,\underset{\sim}{b}')$ are

$$\ell_W(\beta_1,\beta_2,\underset{\sim}{b}';\underset{\sim}{y}) = n\log\beta_2 - n\beta_1\beta_2 + (\beta_2-1)\sum_{i=1}^{n}\log y_i - \sum_{i=1}^{n}\left(\frac{y_i}{e^{\beta_1+\underset{\sim}{z}\underset{\sim}{b}}}\right)^{\beta_2},$$

$$\sum_{i=1}^{n}\underset{\sim}{z}_i'\left(\frac{y_i}{e^{\underset{\sim}{z}_i\hat{\underset{\sim}{b}}}}\right)^{\hat\beta_2} = \underset{\sim}{0}, \qquad \hat\beta_2^{-1} = \frac{\sum_{i=1}^{n}\left(\frac{y_i}{e^{\underset{\sim}{z}_i\hat{\underset{\sim}{b}}}}\right)^{\beta_2}\log y_i}{\sum_{i=1}^{n}\left(\frac{y_i}{e^{\underset{\sim}{z}_i\hat{\underset{\sim}{b}}}}\right)^{\beta_2}} - \frac{\sum_{i=1}^{n}\log y_i}{n}$$

$$\sum_{i=1}^{n}\left(\frac{y_i}{e^{\underset{\sim}{z}_i\hat{\underset{\sim}{b}}}}\right)^{\hat\beta_2} - ne^{\hat\beta_1\hat\beta_2} = 0, \tag{1.2.14}$$

The corresponding density function will be denoted by $f_W(y_i;\beta_1,\beta_2,\underset{\sim}{b}')$.

It may often be convenient both in interpretation and in computation to diagonalize the information matrix by a suitable parametrization. If c and d are location and scale parameters of a distribution, respectively, one possible way of obtaining a diagonal information matrix (Huzurbazar, 1950) is to take the transformation $d=\pi_2$ and $c=\pi_1-\frac{I_{12}}{I_{11}}\pi_2$, where I_{ij} denotes the (i,j)th element of the information matrix $I(c,d)$.

For the Weibull distribution, the transformation to obtain orthogonal parameters and the resulting information matrix are:

$$\sigma=\pi_2 , \qquad \alpha=\pi_1-\psi(2)\pi_2 ,$$

$$I(\pi_1,\pi_2) = \begin{bmatrix} \frac{1}{\pi_2^2} & 0 \\ 0 & \frac{\psi'(1)}{\pi_2^2} \end{bmatrix} .$$

The relation of (π_1,π_2) with the more usual parameterization (β_1,β_2) is

$\beta_1 = \exp\{\pi_1 - \psi(2)\pi_2\}$ and $\beta_2 = \frac{1}{\pi 2}$. With this new parameterization log y has mean $\pi_1 - \pi_2$ and variance $\pi_2^2\psi'(1)$ while with the usual parameterization it has mean $\log \beta_1 + \frac{\psi(1)}{\beta_2}$ and variance $\frac{\psi'(1)}{\beta_2^2}$. Similar results can be obtained for the Weibull regression model.

C Gamma survival models

(i) For b=1 and $\underset{\sim}{\theta}=\underset{\sim}{0}$ we first make the transformation

$\alpha = \log \gamma_1 - \log \gamma_2 + \psi(\gamma_2)\gamma_3^{-1}$, $\sigma = \gamma_2^{-\frac{1}{2}}\gamma_3^{-1}$ and $q = \gamma_2^{-\frac{1}{2}}$ because α, σ and q are functions of k. Then (1.2.5) becomes

$$I(\gamma_1,\gamma_2,\gamma_3) = n\begin{bmatrix} \frac{\gamma_2\gamma_3^2}{\gamma_1^2} & \frac{\gamma_3}{\gamma_1}(1-\gamma_3) & -\frac{\gamma_2}{\gamma_1}\{\psi(\gamma_2+1)\} \\ \frac{\gamma_3}{\gamma_1}(1-\gamma_3) & \psi(\gamma_2) - 2\frac{\gamma_3}{\gamma_2} + \frac{\gamma_3^2}{\gamma_2} & \psi(\gamma_2+1) - \frac{\psi(\gamma_2)}{\gamma_3} \\ -\frac{\gamma_2}{\gamma_1}\{\psi(\gamma_2+1)\} & \psi(\gamma_2+1) - \frac{\psi(\gamma_2)}{\gamma_3} & \frac{1}{\gamma_3^2}[1+\gamma_2\psi'(\gamma_2+1)+\gamma_2\{\psi(\gamma_2+1)\}^2] \end{bmatrix}$$

and, for $b=\gamma_3=1$

$$I(\gamma_1,\gamma_2) = \begin{bmatrix} \frac{\gamma_2}{\gamma_1^2} & 0 \\ 0 & \psi(\gamma_2)+\frac{1}{\gamma_2} \end{bmatrix} . \qquad (1.2.15)$$

The log likelihood function from (1.2.4) and the maximum likelihood estimators of (γ_1,γ_2) are

$$\ell_G(\gamma_1,\gamma_2,\underset{\sim}{y}) = -n \log \Gamma(\gamma_2) + n\,\gamma_2 \log \frac{\gamma_2}{\gamma_1} + (\gamma_2-1)\sum_{i=1}^{n} \log y_i - \frac{\gamma_2}{\gamma_1}\sum_{i=1}^{n} y_i,$$

$$\hat{\gamma}_1 = \frac{\sum_{i=1}^{n} y_i}{n}, \qquad \log \hat{\gamma}_2 - \psi(\hat{\gamma}_2) = \log \hat{\gamma}_1 - \frac{\sum_{i=1}^{n} \log y_i}{n} . \qquad (1.2.16)$$

The corresponding density function will be denoted by $f_G(y_i;\gamma_1,\gamma_2)$.

(ii) For b=1 and $\underset{\sim}{\theta}$ arbitrary, again we first make the transformations

$\alpha = \gamma_1 - \log \gamma_2 + \psi(\gamma_2)\gamma_3^{-1}$, $\sigma = \gamma_2^{-\frac{1}{2}}\gamma_3^{-1}$, $\underset{\sim}{q} = \gamma_2^{-\frac{1}{2}}$ and $\underset{\sim}{\theta}=\underset{\sim}{q}$. Then

(1.2.5) becomes

$$I(\gamma_1,\gamma_2,\gamma_3,\underset{\sim}{g}') = n\begin{bmatrix} \gamma_2\gamma_3^2 & \gamma_3(1-\gamma_3) & -\gamma_2\{\psi(\gamma_2+1)\} & \underset{\sim}{0} \\ \gamma_3(1-\gamma_3) & \psi(\gamma_2)-\frac{2\gamma_3}{\gamma_2}+\frac{\gamma_3^2}{\gamma_2} & \psi(\gamma_2+1)-\frac{\psi(\gamma_2)}{\gamma_3} & \underset{\sim}{0} \\ -\gamma_2\{\psi(\gamma_2+1)\} & \psi(\gamma_2+1)-\frac{\psi(\gamma_2)}{\gamma_3} & \frac{1}{\gamma_3^2}[1+\gamma_2\psi'(\gamma_2)+\gamma_2\{\psi(\gamma_2+1)\}^2] & \underset{\sim}{0} \\ \underset{\sim}{0}' & \underset{\sim}{0}' & \underset{\sim}{0}' & \frac{\gamma_2\gamma_3^2}{n}Z'Z \end{bmatrix}$$

and, for $\gamma_3=1$

$$I(\gamma_1,\gamma_2,\underset{\sim}{g}') = \begin{bmatrix} n\gamma_2 & 0 & \\ 0 & n\{\psi(\gamma_2)-\frac{1}{\gamma_2}\} & \underset{\sim}{0} \\ & \underset{\sim}{0}' & \gamma_2\, Z'Z \end{bmatrix} . \qquad (1.2.17)$$

The log likelihood function and the maximum likelihood estimators of $(\gamma_1,\gamma_2,\underset{\sim}{g}')$ are

$$\ell_G(\gamma_1,\gamma_2,\underset{\sim}{g}';\underset{\sim}{y}) = -\,n\log\Gamma(\gamma_2) + n\,\gamma_2\log\gamma_2 - n\,\gamma_1\gamma_2 + (\gamma_2-1)\sum_{i=1}^{n}\log y_i - \gamma_2\sum_{i=1}^{n}\frac{y_i}{e^{\gamma_1+\underset{\sim}{z}_i\underset{\sim}{g}}} ,$$

$$\sum_{i=1}^{n}\frac{y_i}{e^{\underset{\sim}{z}_i\hat{\underset{\sim}{g}}}} - ne^{\hat{\gamma}_1} = 0 , \qquad \sum_{i=1}^{n}\underset{\sim}{z}_i'\frac{y_i}{e^{\underset{\sim}{z}_i\hat{\underset{\sim}{g}}}} = \underset{\sim}{0}' ,$$

$$\log\hat{\gamma}_2 - \psi(\hat{\gamma}_2) = \hat{\gamma}_1 - \frac{\sum_{i=1}^{n}\log y_i}{n} . \qquad (1.2.18)$$

The corresponding density function will be denoted by $f_G(y_i;\gamma_1,\gamma_2,\underset{\sim}{g})$.

D Exponential survival models

The exponential models are special cases of the Weibull ($\beta_2 = 1$) and of the gamma ($\gamma_2 = 1$) models, therefore the results could be obtained from either of these.

(i) For $\sigma=q=1$ and $\underset{\sim}{\theta}=0$, (1.2.5) becomes $I(\alpha;\ \sigma=q=1) = n$ and for $\alpha = \log\delta + \psi(1)$, we have

$$I(\delta) = \frac{n}{\delta^2}\ , \tag{1.2.19}$$

The log likelihood and the maximum likelihood estimator of δ are

$$\ell_E(\delta,\underset{\sim}{y}) = -\,n\log\delta - \frac{1}{\delta}\sum_{i=1}^{n} y_i \quad , \quad \hat{\delta} = \frac{\sum_{i=1}^{n} y_i}{n}\ ; \tag{1.2.20}$$

the corresponding density function will be denoted by $f_E(y_i;\delta)$.

(ii) For $\sigma=q=1$ and $\underset{\sim}{\theta}'$arbitrary, by taking the transformation $\alpha = \delta + \psi(1)$ and $\underset{\sim}{\theta}'=\underset{\sim}{d}$, (1.2.5) becomes

$$I(\delta,\underset{\sim}{d}') = \begin{bmatrix} n & \underset{\sim}{0} \\ \underset{\sim}{0}' & Z'Z \end{bmatrix}. \tag{1.2.21}$$

The log likelihood function and the maximum likelihood estimators of $(\delta,\underset{\sim}{d}')$ are

$$\ell_E(\delta,\underset{\sim}{d}';\ \underset{\sim}{y}) = -\,n\,\delta - \sum_{i=1}^{n}\frac{y_i}{e^{\delta+\underset{\sim}{z}_i\underset{\sim}{d}}}\ ,$$

$$\sum_{i=1}^{n}\frac{y_i}{e^{\underset{\sim}{z}_i\underset{\sim}{d}}} - ne^{\hat{\delta}} = 0, \quad \sum_{i=1}^{n}\underset{\sim}{z}_i'\,\frac{y_i}{e^{\underset{\sim}{z}_i\hat{\underset{\sim}{d}}}} = \underset{\sim}{0}\ . \tag{1.2.22}$$

The corresponding density function will be denoted by $f_E(y;\delta,\ \underset{\sim}{d}')$.

Note from (1.2.18) and (1.2.22) that the estimators $(\hat{\gamma}_1,\hat{\underset{\sim}{g}})$ and $(\hat{\delta},\hat{\underset{\sim}{d}})$ are the same.

Finally, there is a further property of the maximum likelihood estimator which will also be used frequently later. This result is useful in identifying the parameters on which the distribution of the tests depends and therefore in determining the parameters to be changed in the simulations of Chapter 2. From the considerations leading to

(1.2.3), for $\underset{\sim}{\theta}=\underset{\sim}{0}$ the model can also be written in the forms

$$\frac{1}{\sigma} f(\frac{x-\alpha}{\sigma}; q) , \tag{1.2.23}$$

$$f(x-\alpha; \sigma,q) . \tag{1.2.24}$$

It can be shown that for models of the form (1.2.23) and (1.2.24) the distribution of the maximum likelihood ratio depends only on q and (σ,q) respectively.

1.3 Some related literature

The problem of testing separate families of hypothesis as mentioned in Section 1.1 was first considered by Cox (1961, 1962). He developed the large sample procedure based on the likelihood ratio and also described other approaches that could be used such as a Bayesian approach and the use of more comprehensive models. In subsequent papers, Walker (1967) applied these ideas to some time-series problems; Jackson (1968, 1969) investigated the adequacy of Cox's asymptotic results for the tests involving the exponential and the lognormal distributions and gave further tests involving the gamma and the lognormal distribution. Atkinson (1969, 1970) derived a general method based on the score function for the parameter of a mixed model including both hypotherized distributions. This mixed model has also been used by Cox and Brandwood (1959) and by Selby (1968) who obtained results similar to Atkinson's using the Lagrange multiplier test.

Thomas (1972) gives a computer program for one of Cox's examples. A simulation procedure useful when analytical results are cumbersome or impossible is given by Williams (1970 and in his discussion of Atkinson's 1970 paper).

Invariant and equivalent tests for some problems of separate families are given in Uthoff (1970, 1973), Starbuck (1975) and Quesenberry and Starbuck (1975). Results treating more than two families is provided by Hogg, Uthoff, Randles and Davenport (1972). Also, for location-scale models, simulated results on the likelihood ratio test and other statistics are given by Weibull (1971), Dumonceaux, Antle and Haas (1973), Dumonceaux and Antle (1973) and Antle and Klimko (1975). An empirical comparison of several procedures for discrimination and of testing separate families is reported by Dyer (1971, 1973, 1974).

For a likelihood approach to the discrimination problem, see Lindsey (1974a, 1974b) and for a Bayesian approach with reference to normal regression theory see Lampers (1971), Zellner (1971, p.306) and Box and Kanemasu (1973).

Estimation procedures and economic applications for the multiplicative models of Section 1.2 was studied by Teekens (1972). References to applications in survival studies are Prentice (1973) and Holt and Prentice (1974) and further references can be found in Gross and Clark (1975). The log-gamma and extensions were studied by Prentice (1974), Farewell and Prentice (1974) and Prentice (1975).

Chapter 2

SINGLE SAMPLE CASE

2.1 Introduction

It has been emphasised in Chapter 1 that the problem of interest is that of testing a hypothesis H_f against a hypothesis H_g which specifies the type of departure from H_f thought to be of particular importance. In this chapter, for $\underset{\sim}{y} = (y_1,\ldots,y_n)$, where the y_i are independent and identically distributed observations, the general procedures of Cox and Atkinson are compared. Under the alternative hypothesis the behaviour of the tests is compared through the concept of consistency. The approach of the distribution of the test statistics to the limiting normal distribution is investigated through examination of the terms which differentiate the two procedures.

Tests of separate families of hypothesis involving the probability density functions of Section 1.2, are developed. Empirical simulation is then performed on these cases to investigate the adequacy of the asymptotic theory for finite samples. The sample mean, variance, coefficients of skewness and kurtosis are compared with those of a standard normal distribution. Values are given of the power function and significance level attained at values $t = -1.64$ and $t = -1.28$, i.e. corresponding to 5% and 10% one-sided probability of a standard normal distribution. Comparison of power is made for values in which the A and C statistics attained approximately same significance level.

Histograms of the test statistics under the null hypothesis are presented to show the approach to normality.

2.2 Consistency of the tests

In the general discussion of Section 1.1, it was shown that under the alternative hypothesis the statistics leading to (1.1.8) are expected to have a negative mean. This is closely related to the notion of <u>consistency of a test</u>. A test of a hypothesis H_f against a class of alternatives H_g is said to be <u>consistent</u> if, when any member of H_g holds, the probability of rejecting H_f tends to 1 as the sample size tends to infinity (Cox and Hinkley, 1974, p.317).

Throughout, only the case of independent and identically distributed observations and α and β scalar unknown parameters is dealt with; the same argument applys to the non-homogeneous multi-parameter case. Further, let $f(y,\alpha) > 0$ and $g(y,\beta) > 0$ in the same region, assume the usual conditions for limits and integration to be interchanged, and finally that the expectations involved in what follows are defined. For n observations, $(\hat{\alpha},\hat{\beta})$ are the maximum likelihood estimators of (α,β), β_α is the probability limit of $\hat{\beta}$ when H_f is true. The log likelihood ratio is

$$R(\alpha,\beta;\underset{\sim}{y}) = \log L_f(\alpha,\underset{\sim}{y}) - \log L_g(\beta,\underset{\sim}{y}) ,$$

where $L_f(\alpha;\underset{\sim}{y})$ and $L_g(\beta;\underset{\sim}{y})$ are the likelihood functions for the separate models.

Suppose the null hypothesis is H_f and that H_g is the alternative; from (1.1.1) and (1.1.6) we then have

$$T_f^*(C)/n = \frac{1}{n}\left[R(\hat{\alpha},\hat{\beta};\underset{\sim}{y}) - \{\int R(\alpha,\beta_\alpha;\underset{\sim}{y})\ L_f(\alpha,\underset{\sim}{y})d\underset{\sim}{y}\}_{\hat{\alpha}}\right] , \quad (2.2.1)$$

$$T_f^*(A)/n = \frac{1}{n}\left[R(\hat{\alpha},\beta_{\hat{\alpha}};\underset{\sim}{y}) - \{\int R(\alpha,\beta_\alpha;\underset{\sim}{y})\ L_f(\alpha,\underset{\sim}{y})d\underset{\sim}{y}\}_{\hat{\alpha}}\right] . \quad (2.2.2)$$

Under H_g we have, plim $\hat{\alpha}=\alpha_\beta$, plim $\hat{\beta}=\beta$ and plim $\beta_{\hat{\alpha}}=\beta_{\alpha_\beta}$, where in general $\beta\neq\beta_{\alpha_\beta}$; plim denotes limit in probability and we assume α_β and β_α to be continuous functions. Considering only the terms of order n in probability, in the expansion of the likelihood function, that is

$$L_f(\hat{\alpha},\underset{\sim}{y}) = L_f(\alpha,\underset{\sim}{y}) + O_p(1),$$

$$L_g(\hat{\beta},\underset{\sim}{y}) = L_g(\beta,\underset{\sim}{y}) + o_p(1),$$

$$L_g(\beta_{\hat{\alpha}},\underset{\sim}{y}) = L_g(\beta_{\alpha_\beta};\underset{\sim}{y}) + o_p(1).$$

We have that the test statistics are asymptotically equivalent to

$$T_f^+(C) = -\frac{1}{n}\left[\{-R(\alpha_\beta,\beta;\underset{\sim}{y})\} - \int\{-R(\alpha_\beta,\beta_{\alpha_\beta};\underset{\sim}{y})\}\ L_f(\alpha_\beta;\underset{\sim}{y})d\underset{\sim}{y}\right], \quad (2.2.3)$$

$$T_f^+(A) = -\frac{1}{n}\left[\{-R(\alpha_\beta,\beta_{\alpha_\beta};\underset{\sim}{y})\} - \int\{-R(\alpha_\beta,\beta_{\alpha_\beta};\underset{\sim}{y})\}\ L_f(\alpha_\beta,\underset{\sim}{y})d\underset{\sim}{y}\right]. \quad (2.2.4)$$

Since $\hat{\beta}$ is a consistent estimator of β the true parameter value we also have for n large

$$\frac{L_g(\hat{\beta};\underset{\sim}{y})}{L_g(\beta_{\hat{\alpha}};\underset{\sim}{y})} = \frac{L_g(\beta;\underset{\sim}{y})}{L_g(\beta_{\alpha_\beta};\underset{\sim}{y})} \geq 1 . \quad (2.2.5)$$

Further, the following relations hold:

$$\int\{-R(\alpha_\beta,\beta;\underset{\sim}{y})\}\ L_g(\beta,\underset{\sim}{y})d\underset{\sim}{y} > 0 > \int\{-R(\alpha_\beta,\beta;\underset{\sim}{y})\}\ L_f(\alpha_\beta;\underset{\sim}{y})d\underset{\sim}{y} , \quad (2.2.6)$$

$$\int\{-R(\alpha_\beta,\beta;\underset{\sim}{y})\}\ L_f(\alpha_\beta;\underset{\sim}{y})d\underset{\sim}{y} \geq \int\{-R(\alpha_\beta,\beta_{\alpha_\beta};\underset{\sim}{y})\}\ L_f(\alpha_\beta;\underset{\sim}{y})d\underset{\sim}{y} , \quad (2.2.7)$$

$$\text{plim}\ \frac{1}{n}\{-R(\alpha_\beta,\beta;\underset{\sim}{y})\} = \int\log\frac{g(z,\beta)}{f(z,\alpha_\beta)}\ g(z,\beta)dz = \frac{1}{n}\int\{-R(\alpha_\beta,\beta;\underset{\sim}{y})\}\ L_g(\beta,\underset{\sim}{y})d\underset{\sim}{y} . \quad (2.2.8)$$

We then have, from (2.2.3) and (2.2.7),

$$T_f^+(C) \leq -\frac{1}{n}\left[\{-R(\alpha_\beta,\beta;\underset{\sim}{y})\} - \int\{-R(\alpha_\beta,\beta;\underset{\sim}{y})\}\ L_f(\alpha_\beta;\underset{\sim}{y})d\underset{\sim}{y}\right]. (2.2.9)$$

Inside the square brackets in (2.2.9) the first term has a positive mean and combining (2.2.6) and (2.2.8) we see the full expression in brackets to be always positive and so $T_f^+(C)$ will always converge in probability to a negative value under any member of H_g.

Now, applying the same argument to $T_f^+(A)$, we need an inequality analogous to (2.2.6) stating

$$\int\{-R(\alpha_\beta,\beta_{\alpha_\beta};\underset{\sim}{y})\}\, L_g(\beta,\underset{\sim}{y})d\underset{\sim}{y} > \int\{-R(\alpha_\beta,\beta_{\alpha_\beta};\underset{\sim}{y})\}\, L_f(\alpha_\beta,\underset{\sim}{y})d\underset{\sim}{y} \qquad (2.2.10)$$

but this does not necessarily hold, since the left hand side is not always positive. We then can conclude that for some parameter values, $T_f^+(A)$ may converge to a positive value and in this case it will provide an inconsistent test statistic.

If the roles of H_f and H_g are interchanged, analogous conclusions are obtained for the statistics $T_g(.)$.

Example (2.2.1)

Consider the test of the hypothesis that the observations are from an exponential distribution against the alternative hypothesis that they are from a lognormal distribution. Thus we have

$$f(y,\alpha) = \alpha^{-1}\, e^{-y/\alpha}\ ,\quad g(y,\underline{\beta}) = \frac{1}{y(2\pi\beta_2)}\exp\left\{-\frac{(\log y-\beta_1)^2}{\beta_2}\right\}; \qquad (2.2.11)$$

the test statistics are (Cox, 1961 p.117; Atkinson, 1970, p.337)

$$T_f^+(C) = \hat{\beta}_1 - \beta_{1\hat{\alpha}} + \frac{1}{2}\log\frac{\hat{\beta}_2}{\beta_{2\hat{\alpha}}}\ , \qquad (2.2.12)$$

$$T_f^+(A) = \hat{\beta}_1-\beta_{1\hat{\alpha}} + \frac{1}{2\beta_{2\hat{\alpha}}}\{\hat{\beta}_2-\beta_{2\hat{\alpha}} + (\hat{\beta}_1-\beta_{1\hat{\alpha}})^2\}, \qquad (2.2.13)$$

where

$$\hat{\alpha} = \frac{1}{n}\sum_{i=1}^{n} y_i,\quad \hat{\beta}_1 = \frac{1}{n}\sum_{i=1}^{n}\log y_i,\quad \hat{\beta}_2 = \sum_{i=1}^{n}(\log y_i-\hat{\beta}_1)^2,$$

$$\beta_{1\alpha} = \log\alpha + \psi(1),\quad \beta_{2\alpha} = \psi'(1),\quad \alpha_{(\beta_1,\beta_2)} = e^{\beta_1+\frac{1}{2}\beta_2},$$

$$\psi(x) = \{d\log\Gamma(x)\}/dx,\quad \text{etc.}$$

If the alternative H_g, i.e. the lognormal holds, we have

$$\text{plim } \hat{\beta}_1=\beta_1, \quad \text{plim } \hat{\beta}_2=\beta_2, \quad \text{plim } \hat{\alpha}=e^{\beta_1+\frac{1}{2}\beta_2}, \qquad (2.2.14)$$

$$\text{plim } \beta_{2\hat{\alpha}} = \psi'(1), \quad \text{plim } \beta_{1,\hat{\alpha}} = \text{plim}\{\psi(1)+\log \hat{\alpha}\} = \psi(1)+\beta_1+\tfrac{1}{2}\beta_2.$$

By substituting (2.2.14) in (2.2.12) and (2.2.13), a simple calculation gives

$$\text{plim } T_f^+(C) = \frac{1}{2}(\log \beta_2 - \beta_2 + 0.6567), \qquad (2.2.15)$$

$$\text{plim } T_g^+(A) = \frac{\beta_2^2}{8\psi'(1)} + \frac{1+\psi(1)-\psi'(1)}{2\psi'(1)}\beta_2 + \left\{\frac{\psi^2(1)}{2\psi'(1)} - \frac{1}{2} - \psi(1)\right\}$$

$$= 0.0759\,\beta_2^2 - 0.3714\,\beta_2 + 0.1784 . \qquad (2.2.16)$$

The expression (2.2.15) is negative for all β_2 while (2.2.16) is negative only for β_2 in the interval (0.5401, 4.3484). Table 2.2.1 gives some simulations confirming the second result empirically.

Table 2.2.1 Probability limits and mean of $T_f^+(A)$ under H_g

n	$\mu(T_f^+(A)/H_g)$				
	β_2=0.2	β_2=0.5	β_2=0.8	β_2=4.0	β_2=5.0
20	0.1134	0.0292	-0.0432	0.0205	0.3127
100	0.1092	0.0170	-0.0641	-0.0359	0.2600
200	0.1084	0.0140	-0.0644	-0.0580	0.2281
plim $T_f^+(A)$	0.1072	0.0117	-0.0701	-0.0916	0.2208
plim $T_f^+(C)$	-0.5764	-0.5182	-0.3364	-1.9570	-2.7338

Results from 500 trials. Lognormal deviates obtained using the Box-Muller transformation from uniform variates.

It is interesting to note that

$$\text{plim } T_f^+(C) \le \text{plim } T_f^+(A).$$

This is a general result and follows from (2.2.5). However, this alone does not imply that $T_f(C)$ has higher power than $T_g(A)$ since the variances under the alternative hypothesis are not equal.

2.3 Finite sample comparisons : general discussion

The usefulness of any large sample result is to be assessed by its application to the real problem of finite samples. It is common practice in statistics to use a technique which has well understood asymptotic properties, in the hope that the technique will yield reasonable approximations for finite samples. Explicit small sample results are usually presented by performing simulations on the asymptotic theory, or by analytical methods when the underlying distribution has some simple form.

The purpose of this section is to give a general, although very qualitative, explanation of the simulation results on the behaviour of the A and the C statistics, obtained in the next sections. First, the approach to normality is investigated. For simplicity of notation α and β are assumed to be scalar. The statistics (1.1.1) and (1.1.6) can be approximated by expansion of $E_{\hat{\alpha}}\{\ell_f(\alpha)\}$ and $E_{\hat{\alpha}}\{\ell_f(\beta_\alpha)$ around α, $\ell_f(\alpha)$ around $\hat{\alpha}$ and of $\ell_g(\beta_\alpha)$ around $\hat{\beta}$ and $\beta_{\hat{\alpha}}$ to give

$$T_f^*(C) = T_f + U_n , \tag{2.3.1}$$

$$T_f^*(A) = T_f + U_n + (\beta_\alpha - \beta_{\hat{\alpha}}) \frac{\partial \, lg \, (\beta_{\hat{\alpha}})}{\partial \beta} , \tag{2.3.2}$$

where T_f, [Cox, 1962, eq. (16)] is the sum of deviations of $\log f(y_i;\alpha) - \log g(y_i;\beta_\alpha)$ from its regression on $\partial \log f(y_i;\alpha)/\partial\alpha$, and is of order $\sqrt{n}$ in probability, whereas the other terms are of order 1 in probability.

Now, T_f is a sum of independent and identically distributed random variables of zero mean and therefore quite generally a strong central limit effect can be expected to operate, unless of course, the individual components have a markedly badly behaved distribution. The properties of U_n depend on the particular application but often it also will approach its limiting form quite rapidly. In any case it affects both $T_f(A)$ and $T_f(C)$. The last term in (2.3.2), at least in some applications, may have a markedly nonnormal distribution in samples of moderate size and it is the poor behaviour of this term that accounts for the slower convergence of the distribution of $T_f(A)$. In particular for some of the distributions investigated in this chapter $\partial \ell_g(\beta_{\hat\alpha})/\partial\beta$ requires a large sample size to become relatively small.

The previous discussion was concerned with the approach to normality of the distributions of $T_f(C)$ and $T_f(A)$; this is related to the third and fourth order central moments. To comment on the lower order moments a different argument will be used. The statistics (1.1.1) and (1.1.6) can be written respectively as

$$T_f^*(C) = \ell_f(\hat\alpha) - \ell_g(\hat\beta) - E_{\hat\alpha}\{\ell_f(\hat\alpha) - \ell_g(\beta_{\hat\alpha})\},$$

$$T_f^*(A) = \ell_f(\hat\alpha) - \ell_g(\beta_{\hat\alpha}) - E_{\hat\alpha}\{\ell_f(\hat\alpha) - \ell_g(\beta_{\hat\alpha})\}. \qquad (2.3.4)$$

It has already been pointed out by Atkinson (1970, p.335) that when α is estimated, both statistics in (2.3.4) will be biased, but that $T_f^*(A)$ will be less biased. It then follows that the asymptotic variance (1.1.3) is expected to be approached more rapidly for $T_f^*(A)$ than for $T_f^*(C)$

since in the theory the variance was calculated as if both statistics were unbiased.

There is a final comment on the adequacy of the normal approximations for the distribution of $T_f(.)$. The moments of the test statistics were evaluated from expansions leading to (2.3.1) and (2.3.2); where judged necessary, this can be refined by taking further terms on the expansion. This can happen when for example some terms deleted were not negligible.

If the roles of H_f and H_g are interchanged analogous conclusions are obtained for statistics $T_g(.)$.

2.4 Tests for the lognormal and exponential distributions

A Test statistics and their distributions

The null hypothesis, H_L is that the distribution is lognormal and the alternative H_E that it is exponential, that is $H_L : f_L(\underset{\sim}{y};\alpha_1,\alpha_2)$ against $H_E : f_E(\underset{\sim}{y};\delta)$; see Section 1.2. Under H_L, the estimator $\hat{\delta}$ converges in probability to

$$\delta_L = \exp\left\{\alpha_1 + \tfrac{1}{2}\alpha_2\right\}, \qquad (2.4.1)$$

that is δ_L is the mean of the lognormal distribution. Further, for H_L we have (Cox, 1961, 1962)

$$T^*_{LE}(C) = n \log \frac{\hat{\delta}}{\delta_{\hat{L}}},$$

$$V_L\{T^*_{LE}\} = n\left(e^{\alpha_2} - 1 - \alpha_2 - \frac{\alpha_2^2}{2}\right), \qquad (2.4.2)$$

and after some calculation

$$T^*_{LE}(A) = n\left(\frac{\hat{\delta}}{\delta_{\hat{L}}} - 1\right), \qquad (2.4.3)$$

where $$\delta_{\hat{L}} = \exp\left\{\hat{\alpha}_1 + \frac{\hat{\alpha}_2}{2}\right\} .$$

Now, suppose that H_L and H_E change roles so that the null distribution is exponential and the alternative is lognormal. Under H_E, the estimators $\hat{\alpha}_1$ and $\hat{\alpha}_2$ converge in probability respectively to

$$\alpha_{1E} = \psi(1) + \log\delta , \qquad \alpha_{2E} = \psi'(1) , \tag{2.4.4}$$

that is α_{1E} and α_{2E} are the mean and variance of the logarithm of a random variable with an exponential distribution, where $\psi(x) = \{d \log \Gamma(x)\}/dx$, etc. For H_E we obtain (Cox 1961, 1962)

$$T^*_{EL}(C) = n\left\{\hat{\alpha}_1 - \alpha_{1\hat{E}} + \frac{1}{2}\log\frac{\hat{\alpha}_2}{\alpha_{2\hat{E}}}\right\} ,$$

$$V_E\{T^*_{EL}\} = n\left\{\psi'(1) - \frac{1}{2} + \frac{\psi''(1)}{\psi'(1)} + \frac{\psi'''(1)}{4\{\psi'(1)\}^2}\right\} \ 0.2834n, \tag{2.4.5}$$

and similarly (Atkinson, 1970)

$$T^*_{EL}(A) = n\ \{\hat{\alpha}_1 - \alpha_{1\hat{E}} + \frac{1}{2\alpha_{2\hat{E}}}\ [\hat{\alpha}_2 - \alpha_{2\hat{E}} + (\hat{\alpha}_1 - \alpha_{1\hat{E}})^2]\ \} , \tag{2.4.6}$$

where $\alpha_{1\hat{E}} = \psi(1) + \log\hat{\delta}$ and $\alpha_{2\hat{E}} = \psi'(1)$.

Then, asimptotically the statistics,

$$T_{LE}(j) = T^*_{LE}(j)\quad V_L[\{T^*_{LE}\}]^{-\frac{1}{2}} \qquad (j = A,C) , \tag{2.4.7}$$

$$T_{EL}(j) = T^*_{EL}(j)\quad V_E[\{T^*_{EL}\}]^{-\frac{1}{2}} \qquad (j = A,C) , \tag{2.4.8}$$

have a standard normal distribution, (2.4.7) under H_L and (2.4.8) under H_E.

B Empirical results

Now the empirical investigations for comparison between $T_{LE}(C)$ and $T_{LE}(A)$ and between $T_{EL}(C)$ and $T_{EL}(A)$ and on the adequacy of the asymptotic results are discussed.

Results on the null distribution of $T_{LE}(C)$ and $T_{LE}(A)$ and on the distribution of $T_{EL}(C)$ and $T_{EL}(A)$ under the alternative were obtained as follows. Random independent variates u_i from a uniform (0,1) distribution were generated. Then the Box-Muller transformation was applied to obtain independent variates t_i from a standard normal distribution. Taking $y_i = \exp\{\alpha_1 + \sqrt{\alpha_2}\, t_i\}$ gave independent variates from a lognormal distribution. From the comments on (1.2.24) of Section 1.2 only $\alpha_1 = 0$ needed be considered since it follows that the distribution of the test statistics in this case depends only on α_2. Some different values of α_2 were considered. Then $T_{LE}(C)$, $T_{LE}(A)$, $T_{EL}(C)$ and $T_{EL}(A)$ were calculated under the lognormal hypothesis H_L. For various sample sizes n, 500 trials were obtained and from these were calculated (i) the first four moments of all tests, (ii) the significance level attained by $T_{LE}(C)$ and $T_{LE}(A)$ at $t = -1.64$ and $t = -1.28$, (iii) the power of $T_{EL}(C)$ and $T_{EL}(A)$ at $t = -1.64$ and $t = -1.28$.

Results on the null distribution of $T_{EL}(C)$ and $T_{EL}(A)$ and on the distribution of $T_{LE}(C)$ and $T_{LE}(A)$ under the alternative were obtained in an analogous way. Here the transformation $y = -\delta \log y_i$ gave independent variates from an exponential distribution. From the comments on (1.2.24), it follows that the distribution of the tests is independent of the parameter δ. For various sample sizes n, 1000 trials were obtained with $\delta=1$.

The results are summarized in Tables 2.4.1 to 2.4.8.

The sampling moments of $T_{LE}(C)$ and $T_{EL}(C)$ are in agreement with those calculated by Jackson (1968). Also, results of Table 2.4.2 are in agreement with Atkinson (1970, Table 4).

Results of Tables 2.4.1 and 2.4.2 show that the mean and variance of the A statistics are in closer agreement with the asymptotic values than are those of the C statistics. The measures of skewness and of kurtosis of the C statistics are however in closer agreement with the asymptotic values than are those of the A statistics. This is to be expected in view of the discussion of Section 2.3.

For $\alpha_2 = 0.2$ in Table 2.4.4 the statistic $T_{EL}(A)$ shows a positive mean under the alternative hypothesis, which agrees with the results of Section 2.2 about consistency of the test.

Two further points can be noticed from Table 2.4.1. For α_2 increasing it seems that the approach to normality becomes slower for both statistics and that it affects $T_{LE}(A)$ more than $T_{LE}(C)$. For the latter case, the term which differentiates $T_{LE}(A)$ from $T_{LE}(C)$ is

$$\frac{\partial}{\partial\delta}\,\ell_E(\hat{\delta}_L,\underset{\sim}{y}) = \frac{\sum_{i=1}^{n}(y_i - e^{\hat{\alpha}_1+\frac{1}{2}\hat{\alpha}_2})}{(e^{\hat{\alpha}_1+\frac{1}{2}\hat{\alpha}_2})^2} = \frac{n(\hat{\delta} - e^{\hat{\alpha}_1+\frac{1}{2}\hat{\alpha}_2})}{(e^{\hat{\alpha}_1+\frac{1}{2}\hat{\alpha}_2})^2}\,. \tag{2.4.9}$$

It is well known that the sample mean is an inefficient estimator of the mean of the lognormal distribution. The variance of $\hat{\delta}$ is of order $(e^{\alpha_2})^3$ and, for large α_2, the numerator in (2.4.9) will then require a large sample size to become small.

When α_2 is increased, the adequacy of the asymptotic results for both $T_{LE}(A)$ and $T_{LE}(C)$ is now investigated. For this, higher order terms are examined as explained at the end of Section 2.3. The term

$$\frac{1}{2}\,(\hat{\delta} - \delta_L)^2\,\frac{\partial^2\ell_E(\hat{\delta};\underset{\sim}{y})}{\partial\delta^2} = \frac{n}{2}\left(\frac{\hat{\delta} - \delta_L}{\hat{\delta}}\right)^2 , \tag{2.4.10}$$

has mean of order e^{α_2}/n and variance of order $(e^{\alpha_2})^6/n^2$ and as before will not be negligible for large α_2. Further terms could be investigated, but (2.4.10) shows the magnitude and the importance as α_2 increases of the deleted terms. Fortunately values which arise in practice seem to be quite often in the neighbourhood of $\alpha_2 = 0.5$, and for these the results seem adequate.

For the purpose of power comparisons Table 2.4.7 shows that except for $\alpha_2 = 2$, the significance levels for both tests are of about the same order. Thus, it is meaningful to compare the power in Table 2.4.5 and it then follows that $T_{LE}(C)$ should be recommended. When the hypothesis H_E changes roles with H_L, Table 2.4.8 shows that the significance levels do not permit comparison of the results for $T_{EL}(A)$ and $T_{EL}(C)$ in Table 2.4.6. However, from the results on the inconsistency of $T_{EL}(A)$ in Section 2.2 only for certain values of α_2 $T_{EL}(A)$ could be recommended. It seems reasonable, therefore, in practice to use $T_{EL}(C)$.

From a more practical point of view the statistics C are also to be recommended because the significance levels attained agree more closely to those of the standard normal, and this is what would be hoped in a specific application.

Figures 2.4.1 to 2.4.4 present the histograms of the data of Table 2.4.1 and 2.4.2 showing clearly the approach to normality and the effects of increasing α_2.

Table 2.4.1 Null distribution of $T_{LE}(C)$ and $T_{LE}(A)$.

n	$T_{LE}(.)$	$\mu_1\{T_{LE}(.)/H_L\}$			$\mu_2\{T_{LE}(.)/H_L\}$			$\gamma_1\{T_{LE}(.)/H_L\}$			$\beta_2\{T_{LE}(.)/H_L\}$		
		$\alpha_2=0.2$	$\alpha_2=1.0$	$\alpha_2=2.0$	$\alpha_2=0.2$	$\alpha_2=1.0$	$\alpha_2=2.0$	$\alpha_2=0.2$	$\alpha_2=1.0$	$\alpha_2=2.0$	$\alpha_2=0.2$	$\alpha_2=1.0$	$\alpha_2=2.0$
20	C	-0.116	-0.179	-0.210	0.629	0.428	0.253	0.481	0.856	0.971	3.601	4.283	4.717
	A	-0.113	-0.156	-0.164	0.631	0.448	0.257	0.507	1.157	1.603	3.658	5.373	7.171
50	C	-0.108	-0.144	-0.173	0.955	0.729	0.478	0.332	0.799	1.018	3.301	4.248	4.893
	A	-0.105	-0.120	-0.119	0.956	0.755	0.542	0.349	1.027	1.714	3.324	5.049	8.318
100	C	-0.071	-0.123	-0.155	0.931	0.757	0.554	0.373	0.756	1.046	3.279	4.009	4.892
	A	-0.069	-0.104	-0.108	0.932	0.777	0.622	0.384	0.916	1.630	3.296	4.514	7.761
150	C	-0.083	-0.100	-0.104	0.903	0.771	0.686	0.266	0.671	1.053	3.119	4.025	4.910
	A	-0.081	-0.085	-0.057	0.903	0.783	0.788	0.275	0.798	1.645	3.129	4.440	7.795
200	C	-0.035	-0.086	-0.125	0.917	0.811	0.643	0.120	0.494	0.883	3.056	3.419	4.454
	A	-0.034	-0.072	-0.088	0.917	0.820	0.684	0.127	0.592	1.252	3.062	3.618	5.852

Results from 500 trials.

Table 2.4.2 Null distribution of $T_{EL}(C)$ and $T_{EL}(A)$.

n	$T_{EL}(.)$	$\mu_1\{T_{EL}(.)/H_E\}$	$\mu_2\{T_{EL}(.)/H_E\}$	$\gamma_1\{T_{EL}(.)/H_E\}$	$\beta_2\{T_{EL}(.)/H_E\}$
20	C	-0.441	0.708	0.374	3.845
	A	-0.057	0.934	2.479	15.528
50	C	-0.250	0.859	0.618	4.430
	A	0.092	1.092	2.430	17.708
100	C	-0.167	0.933	0.374	3.198
	A	0.069	1.025	0.964	5.086
150	C	-0.193	1.032	0.477	3.736
	A	0.003	1.080	1.081	6.020
200	C	-0.129	0.984	0.388	3.340
	A	0.037	1.009	0.775	4.095

Results from 1000 trials.

Table 2.4.3 Distribution of $T_{LE}(C)$ and $T_{LE}(A)$ under alternative H_E.

n	$T_{LE}(.)$	$\mu_1\{T_{LE}(.)/H_E\}$	$\mu_2\{T_{LE}(.)/H_E\}$	$\gamma_1\{T_{LE}(.)/H_E\}$	$\beta_2\{T_{LE}(.)/H_E\}$
20	C	-0.823	0.157	0.804	5.254
	A	-0.729	0.125	0.913	6.262
50	C	-1.464	0.142	0.530	4.098
	A	-1.292	0.100	0.434	4.593
100	C	-2.156	0.151	0.331	3.886
	A	-1.904	0.096	0.310	4.260
150	C	-2.668	0.171	0.207	3.998
	A	-2.367	0.105	0.236	4.301
200	C	-3.117	0.144	0.261	3.200
	A	-2.762	0.083	0.218	3.373

Results from 1000 trials.

Table 2.4.4 Distribution of $T_{EL}(C)$ and $T_{EL}(A)$ under alternative H_L.

n	$T_{LE}(.)$	$\mu_1\{T_{EL}(.)/H_L\}$			$\mu_2\{T_{EL}(.)/H_L\}$			$\gamma_1\{T_{LE}(.)/H_L\}$			$\beta_2\{T_{LE}(.)/H_L\}$		
		$\alpha_2=0.2$	$\alpha_2=1.0$	$\alpha_2=2.0$	$\alpha_2=0.2$	$\alpha_2=1.0$	$\alpha_2=2.0$	$\alpha_2=0.2$	$\alpha_2=1.0$	$\alpha_2=2.0$	$\alpha_2=0.2$	$\alpha_2=1.0$	$\alpha_2=2.0$
20	C	- 5.283	-1.514	-2.171	1.157	0.488	2.661	-0.518	-1.001	-1.730	3.408	5.668	7.062
	A	- 0.953	-0.644	-1.382	0.031	0.567	0.986	-1.026	-1.330	-0.611	4.713	5.722	4.221
50	C	- 7.875	-2.275	-3.858	1.210	0.531	4.570	-0.352	-0.894	-1.514	3.061	5.312	6.316
	A	1.449	-1.344	-2.781	0.043	0.760	1.630	-0.840	-0.940	-0.843	4.821	4.327	4.670
100	C	-11.093	-3.218	-5.692	1.130	0.641	6.471	-0.288	-0.862	-1.425	3.201	4.620	6.103
	A	2.052	-2.019	-4.342	0.039	0.978	2.483	-0.767	-0.856	-0.835	4.732	4.187	4.056
150	C	-13.410	-3.930	-7.068	1.202	0.609	6.256	-0.153	-0.535	-1.038	3.095	3.644	4.437
	A	2.488	-2.535	-5.518	0.042	0.917	2.682	-0.583	-0.590	-0.748	3.757	3.231	3.955
200	C	-15.505	-4.549	-8.226	1.142	0.657	6.963	-0.218	-0.437	-1.061	2.923	3.556	4.572
	A	2.880	-2.952	-6.474	0.042	0.984	2.934	-0.328	-0.618	-0.660	2.833	3.444	3.653

Results from 500 trials.

Table 2.4.5 Null : Lognormal; Alternative:exponential. Tests: $T_{LE}(C)$, $T_{LE}(A)$. Power at t = -1.64; t = -1.28.

n	$T_{LE}(.)$	Power function	
		SL=0.05	SL=0.10
20	C	0.011	0.105
	A	0.003	0.036
50	C	0.341	0.717
	A	0.117	0.536
100	C	0.914	0.982
	A	0.826	0.969
150	C	0.987	0.998
	A	0.981	0.998
200	C	1.000	1.000
	A	1.000	1.000

Results from 1000 trials.

Table 2.4.6 Null : Exponential; Alternative : Lognormal.
Tests : $T_{EL}(C)$, $T_{EL}(A)$.
Power at t = -1.64; -1.28.

n	$T_{EL}(.)$	Power function					
		SL=0.05			SL=0.10		
		α_2=0.2	α_2=1.0	α_2=2.0	α_2=0.2	α_2=1.0	α_2=2.0
20	C	1.000	0.372	0.556	1.000	0.598	0.674
	A	0	0.086	0.388	0	0.156	0.530
50	C	1.000	0.826	0.906	1.000	0.924	0.960
	A	0	0.318	0.826	0	0.470	0.900
100	C	1.000	0.994	0.986	1.000	0.998	0.998
	A	0	0.638	0.982	0	0.758	0.992
150	C	1.000	1.000	1.000	1.000	1.000	1.000
	A	0	0.830	1.000	0	0.928	1.000
200	C	1.000	1.000	1.000	1.000	1.000	1.000
	A	0	0.928	1.000	0	0.972	1.000

Results from 500 trials.

Table 2.4.7 Null: Lognormal; Alternative exponential.
Tests: $T_{LE}(C)$, $T_{LE}(A)$.
One side significant levels at t = -1.64; t = -1.28.

n	$T_{LE}(.)$	Significance Level					
		SL=0.05			SL=0.10		
		α_2=0.2	α_2=1.0	α_2=2.0	α_2=0.2	α_2=1.0	α_2=2.0
20	C	0.016	0.004	0	0.056	0.018	0.004
	A	0.016	0.002	0	0.054	0.018	0
50	C	0.052	0.014	0	0.116	0.064	0.024
	A	0.050	0.010	0	0.114	0.054	0.008
100	C	0.028	0.018	0.004	0.096	0.048	0.020
	A	0.028	0.018	0	0.094	0.042	0.014
150	C	0.048	0.018	0.006	0.100	0.054	0.034
	A	0.048	0.016	0	0.100	0.048	0.020
200	C	0.036	0.024	0.006	0.102	0.076	0.030
	A	0.036	0.022	0.002	0.102	0.062	0.026

Results from 500 trials.

Table 2.4.8 Null: Exponential; Alternative lognormal.
Tests: $T_{EL}(C)$, $T_{EL}(A)$.
One sided significance levels at t = -1.64; t = -1.28.

n	$T_{EL}(.)$	Significance Level	
		SL=0.05	SL=0.10
20	C	0.059	0.134
	A	0.009	0.030
50	C	0.049	0.132
	A	0.007	0.039
100	C	0.049	0.108
	A	0.019	0.052
150	C	0.066	0.125
	A	0.026	0.078
200	C	0.056	0.112
	A	0.024	0.066

Results from 1000 trials.

FIGURE 2.4.1 - NULL DISTRIBUTION OF $T_{LE}(\cdot)$ - LOGNORMAL vs EXPONENTIAL, $\alpha_2 = 0.2$, 500 trials, (see Appendix B).

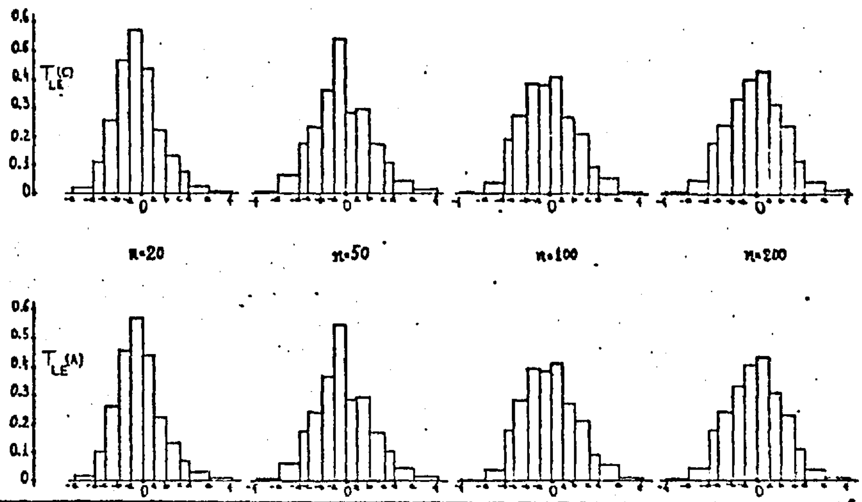


FIGURE 2.4.2 - NULL DISTRIBUTION OF $T_{LE}(\cdot)$ - LOGNORMAL vs EXPONENTIAL, $\alpha_2 = 1.0$, 500 trials, (see Appendix B)

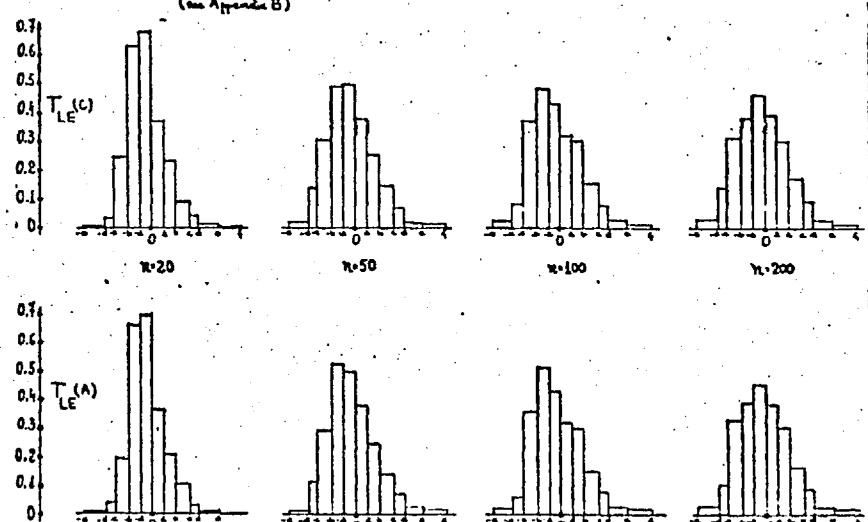


FIGURE 2.4.3 - NULL DISTRIBUTION OF $T_{LE}(\cdot)$ - LOGNORMAL vs EXPONENTIAL, $\alpha_2 = 2.0$, 500 trials, (see Appendix B)

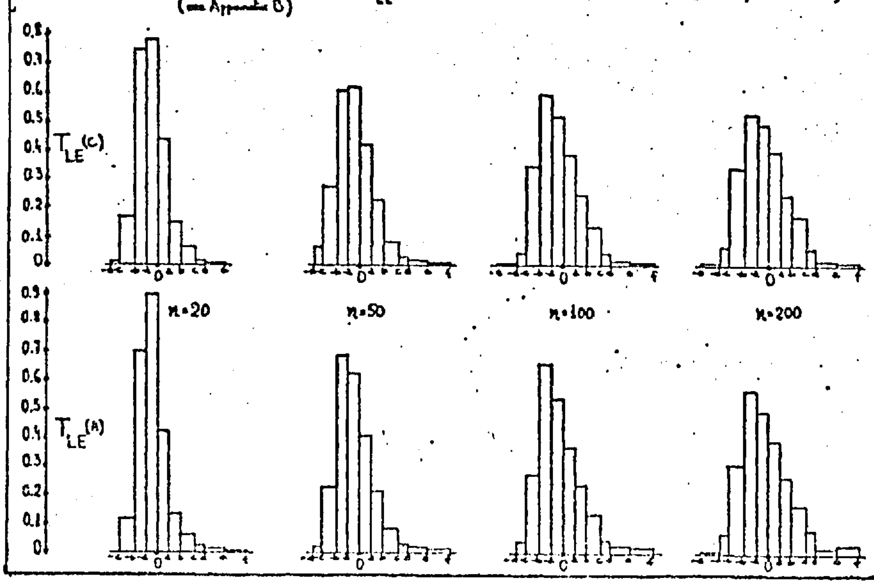
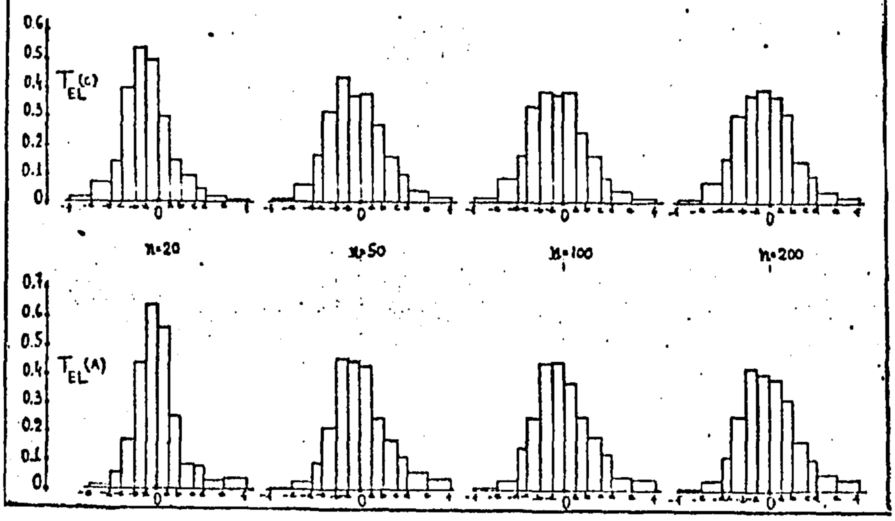


FIGURE 2.4.4 - NULL DISTRIBUTION OF $T_{EL}(\cdot)$ - EXPONENTIAL vs LOGNORMAL, 1000 trials, (see Appendix B)



2.5 Tests for the lognormal and gamma distributions

A Test statistics and their distributions

The null hypothesis, H_L is that the distribution is lognormal and the alternative H_G that it is gamma, that is $H_L : f_L(\underset{\sim}{y}, \alpha_1, \alpha_2)$ against $H_G : f_G(\underset{\sim}{y}, \gamma_1, \gamma_2)$; see Section 1.2. Under H_L, the estimators $\hat{\gamma}_1$ and $\hat{\gamma}_2$ converge in probability to γ_{1L} and γ_{2L} respectively, where

$$\gamma_{1L} = \exp\{\alpha_1 + \tfrac{1}{2}\alpha_2\} \ , \ \log \gamma_{2L} - \psi(\gamma_{2L}) = \log \gamma_{1L} - \alpha_1 = \tfrac{1}{2}\alpha_2 . \tag{2.5.1}$$

Thus, $\hat{\gamma}_1$ converges to the mean of the lognormal distribution and the right hand side of the equation for γ_{2L} is the logarithm of the ratio of the arithmetic mean to the geometric mean of the lognormal distribution. Further, for H_L we have (Jackson 1968)

$$T^*_{LG}(C) = n\{\log \Gamma(\hat{\gamma}_2) - \hat{\gamma}_2\psi(\hat{\gamma}_2) + \hat{\gamma}_2 - \log \Gamma(\gamma_{2\hat{L}}) + \gamma_{2\hat{L}}\psi(\gamma_{2\hat{L}}) - \gamma_{2\hat{L}}\},$$
$$V_L\{T^*_{LG}\} = n\ \gamma_{2L}^2 \left(e^{\alpha_2} - 1 - \alpha_2 - \frac{\alpha_2^2}{2}\right) , \tag{2.5.2}$$

and after some calculation,

$$T^*_{LG}(A) = n\ \gamma_{2\hat{L}} \left\{\frac{\hat{\gamma}_1}{\gamma_{1\hat{L}}} - 1\right\} , \tag{2.5.3}$$

where $\gamma_{1\hat{L}} = \exp\{\hat{\alpha}_1 + \tfrac{1}{2}\hat{\alpha}_2\}$ and $\gamma_{2\hat{L}}$ given by $\log \gamma_{2\hat{L}} - \psi(\gamma_{2\hat{L}}) = \tfrac{1}{2}\hat{\alpha}_2$.

Now, suppose that H_L and H_G change roles so that the null distribution is gamma and the alternative is lognormal. Under H_G, the estimators $\hat{\alpha}_1$ and $\hat{\alpha}_2$ converge in probability to

$$\alpha_{1G} = \psi(\gamma_2) - \log \frac{\gamma_2}{\gamma_1} \ , \ \alpha_{2G} = \psi'(\gamma_2) \ , \tag{2.5.4}$$

respectively. That is α_{1G} and α_{2G} are respectively the mean and variance of the logarithm of a random variable with a gamma distribution. For H_G we have (Jackson 1968)

$$T^*_{GL}(C) = \frac{n}{2} \log \frac{\hat{\alpha}_2}{\alpha_{2\hat{G}}} ,$$

$$V_G\{T^*_{GL}\} = n \left\{ \frac{\psi'''(\gamma_2)}{4\{\psi'(\gamma_2)\}^2} - \frac{\gamma_2\{\psi''(\gamma_2)\}^2}{4\{\psi'(\gamma_2)\}^2\{\gamma_2\psi'(\gamma_2)-1\}} - \tfrac{1}{2} \right\}, \quad (2.5.5)$$

and after some calculations

$$T^*_{GL}(A) = n \left(\frac{\hat{\alpha}_2}{\alpha_{2\hat{G}}} - 1 \right) , \quad (2.5.6)$$

where $\alpha_{1\hat{G}} = \psi(\hat{\gamma}_2) - \log \frac{\hat{\gamma}_2}{\hat{\gamma}_1}$, $\alpha_{2\hat{G}} = \psi'(\hat{\gamma}_2)$.

It should be noted from the relation

$$\hat{\alpha}_1 = \psi(\hat{\gamma}_2) - \log \frac{\hat{\gamma}_2}{\hat{\gamma}_1} = \psi(\gamma_{2\hat{L}}) - \log \frac{\gamma_{2\hat{L}}}{\gamma_{1\hat{L}}} = \alpha_{1\hat{G}}$$

that for $\gamma_2 = 1$ we obtain $\gamma_{2L} = 1$ and that the expressions (2.5.1) to (2.5.6) recover the corresponding expression of Section 2.4.

Finally, asymptotically the statistics,

$$T_{LG}(\gamma) = T^*_{LG}(\gamma) \; [V_L\{T^*_{LG}\}]^{-\frac{1}{2}} \qquad (\gamma = A,C) \quad , \quad (2.5.7)$$

$$T_{GL}(j) = T^*_{GL}(j) \; [V_G\{T_{GL}\}]^{-\frac{1}{2}} \qquad (j = A,C) \quad , \quad (2.5.8)$$

have a standard normal distribution, (2.5.7) under H_L and (2.5.8) under H_G.

B Empirical results

The empirical results for comparison between $T_{LG}(C)$ and $T_{LG}(A)$ and between $T_{GL}(C)$ and $T_{GL}(A)$ and on the adequacy of the asymptotic results are now discussed.

Results on the null distribution of $T_{LG}(C)$ and $T_{LG}(A)$ and on the distribution of $T_{GL}(C)$ and $T_{GL}(A)$ under the alternative, that is the lognormal distribution, were obtained in a manner similar to Section 2.4. Here again from (1.2.24) it follows that the distribution of the test statistics depends only on α_2. For $\alpha_1 = 0$ and each different value of α_2, 500 trials for various sample sizes were obtained.

In a similar way the results on the distribution of $T_{GL}(C)$ and $T_{GL}(A)$ and on the distribution of $T_{LG}(C)$ and $T_{LG}(A)$ under the alternative, that is the gamma distribution, were obtained. Here random variates from a gamma distribution were obtained from u_i independent uniform (0,1) random variates, as follows. For γ_2 integer the transformation $y_i = \sum_{i=1}^{\gamma_2} (-\gamma_1) \log u_i$ gave independent variates from a gamma distribution with parameters γ_1 and γ_2. For γ_2 non-integer the method described by Whittaker (1974) was used. Again, from the comments on (1.2.24) it follows that the distribution of the test statistics depends only on γ_2. For $\gamma_1 = 1$ and each different value of γ_2, 500 trials for several sample sizes were obtained.

For calculating the test statistics the functions $\Gamma(z)$, $\psi(z)$, $\psi'(z)$, $\psi''(z)$ and $\psi'''(z)$ are needed. For these the approximations given in Abramowitz & Stegun [1972, eq.(6.1.41), (6.3.18), (6.4.12), (6.4.13) and (6.4.14)] were used. Further, for any z the approximations were used for z + 8 and $\Gamma(z)$ and $\psi^{(n)}(z)$ obtained from the relations $\Gamma(z+1) = z\Gamma(z)$ and $\psi^{(n)}(z+1) = \psi^{(n)}(z) + (-1)^n n! z^{-n-1}$. The approximations get better as z increases and for values as small as $z = 0.2$, $\psi(z)$ is correct up to four decimal places and all others are correct up to at least nine decimal places.

To solve the maximum likelihood equations and other equations for calculating the test statistics, Newton's method was used; the iterations were stopped when the equations differed from zero by less than 0.001. No problem of convergence was encountered.

The results are summarized in Tables 2.5.1 to 2.5.9.

Results of Table 2.5.1 and 2.5.2 generally agrees with the discussion of Section 2.3 on the behaviour of the A and the C statistics. The A statistics have a better agreement for the two first moments while the C statistics have a better agreement for the skewness and kurtosis coefficients.

Two further points can again be noticed from Table 2.5.1. Similarly to Section 2.4, for α_2 increasing it seems that the approach to normality becomes slower for both statistics and it affects $T_{LG}(A)$ more than $T_{LG}(C)$. Here the terms which differentiate $T_{LG}(A)$ from $T_{LG}(C)$ are

$$\frac{\partial}{\partial\gamma_1} \ell_G(\gamma_{1\hat{L}},\gamma_{2\hat{L}};\underset{\sim}{y}) = \gamma_{2\hat{L}} \frac{n(\hat{\gamma}_1 - e^{\hat{\alpha}_1+\frac{1}{2}\hat{\alpha}_2})}{(e^{\hat{\alpha}_1+\frac{1}{2}\hat{\alpha}_2})^2} , \qquad (2.5.9)$$

$$\frac{\partial}{\partial\gamma_2} \ell_G(\gamma_{1\hat{L}},\gamma_{2\hat{L}};\underset{\sim}{y}) = \frac{n(e^{\hat{\alpha}_1+\frac{1}{2}\hat{\alpha}_2}-\hat{\gamma}_1)}{e^{\hat{\alpha}_1+\frac{1}{2}\hat{\alpha}_2}} , \qquad (2.5.10)$$

and one of the higher order terms is

$$\tfrac{1}{2}(\hat{\gamma}_1-\gamma_{1L})^2 \frac{\partial^2\ell_G(\hat{\gamma}_1,\hat{\gamma}_2;\underset{\sim}{y})}{\partial\gamma_1^2} = \frac{n}{2}\hat{\gamma}_2^2 \left[\frac{\hat{\gamma}_1-\gamma_{1L}}{\hat{\gamma}_1}\right]^2 . \qquad (2.5.11)$$

For the same reason given for (2.4.9), it is required a large sample size for (2.5.9) and (2.5.10) to become relatively small. The mean and the variance of (2.5.11) is of the same order as that of (2.4.10) and similarly shows the magnitude and importance of the neglected terms.

For the parameter values of Tables 2.5.3 and 2.5.4 the means of the tests $T_{LG}(A)$ and $T_{GL}(A)$ are negative and the tests are then consistent. A general investigation on the consistency of these tests is not simple and for $T_{GL}(A)$ it does not seem possible since the estimates are obtained by iterative processes.

Exact comparison of the power of the A and the C statistics would require the same significance level on both statistics for all parameter values. Here instead an approximate argument was used. The power and the corresponding significance level were compared at that parameter values for which both distributions have a similar shape. Although no conclusion can be inferred for values not used in the simulations, it would be expected that for values corresponding to shapes which are more dissimilar between the two distributions, a higher power would be attained and a closer agreement to the asymptotic significance level obtained.

For small values of α_2 for the lognormal density function and large values of γ_2 for the gamma density function both have shapes similar to that of a normal density function. For the power of $T_{LG}(A)$ and $T_{LG}(C)$, Tables 2.5.7 shows that for $\alpha_2 = 0.1$ and $\alpha_2 = 0.25$ the significance levels are about the same for A and C. Table 2.5.5 gives values corresponding to $\gamma_2 = 5.0$ and $\gamma_2 = 10.0$ and it follows that there is not much difference in the power of the two statistics. The difference could well be due to the slight difference in the significance levels. Similarly, for the power of $T_{GL}(A)$ and $T_{GL}(C)$, Table 2.5.8 shows that for $\gamma_2 = 5.0$ and $\gamma_2 = 10.0$ the significance levels are about the same for A and C. It follows from Table 2.5.6 for values $\alpha_2 = 0.1$ and $\alpha_2 = 0.25$ that again there is not much difference of power between the two statistics and the difference could be due to the slight difference in significance levels.

For other values of the parameter, the difficulties are overcome by defining closeness in another way. Consider as the nearest alternative to a particular member of H_f say, that member of H_g with parameter value given by the probability limit of its maximum likelihood estimator when that particular member of H_f is true. For example if $\alpha_2 = 0.21$, we would expect $T_{LG}(.)$ to have lower power for a gamma distribution with $\gamma_{2L} = 5.0$ the solution of $\log \gamma_{2L} - \psi(\gamma_{2L}) = \frac{0.21}{2}$, that is equation (2.5.1). Similarly, for $\gamma_2 = 5.0$ we would expect $T_{GL}(.)$ to have lower power for a lognormal distribution with $\alpha_{2G} = 0.22$ the solution of $\alpha_{2G} = \psi'(5.0)$, equation (2.5.4). The example shows that the method agrees with the comparisons of power previously made using the normal shape.

Consider a further comparison using this argument. For $\gamma_2 = 2.0$, the 10% significance levels in Table 2.5.8 are not very different for the A and the C statistics. The corresponding values for power comparisons of $T_{GL}(.)$ is $\alpha_{2G} = \psi'(2.0) \simeq 0.64$ in Table 2.5.6 and allowing for the slight difference in Table 2.5.8 the power in Table 2.5.6 does not seem to be very different for A and C. Similarly, for $\alpha_2 = 0.54$, the significance levels in Table 2.5.7, except for n = 20 and n = 50, are not very different. The corresponding value to look at in Table 2.5.5 is $\gamma_{2L} = 2.0$ and the same conclusion is reached. For these cases the further results of Table 2.5.9 seem to confirm the assumption of equal power.

Another point should be observed from Table 2.5.7 and 2.5.8, generally the significance levels of the C statistics agrees more closely to those of the standard normal. This is related to the faster approach to normality of the statistics C. From a practical viewpoint this provides an argument for C to be preferable.

Figures 2.5.1 to 2.5.6 present histograms of the data of Tables 2.5.1 and 2.5.2. They show the approach to normality and the effects of increasing α_2. It is interesting to note that changes in γ_2 does not seem to have much effect on the approach to normality of $T_{GL}(.)$.

TABLE 2.5.1 Null distribution of $T_{LG}(C)$ and $T_{LG}(A)$.

n	$T_{LG}(.)$	$\mu_1(T_{LG}(.)/H_L)$			$\mu_2(T_{LG}(.)/H_L)$			$\gamma_1(T_{LG}(.)/H_L)$			$\beta_2(T_{LG}(.)/H_L)$		
		$\alpha_2=0.1$	$\alpha_2=0.5$	$\alpha_2=2.0$	$\alpha_2=0.1$	$\alpha_2=0.5$	$\alpha_2=2.0$	$\alpha_2=0.1$	$\alpha_2=0.5$	$\alpha_2=2.0$	$\alpha_2=0.1$	$\alpha_2=0.5$	$\alpha_2=2.0$
20	C	-0.002	-0.083	-0.253	0.710	0.592	0.264	0.458	0.776	0.478	5.336	5.728	3.633
	A	0.020	-0.032	-0.164	0.725	0.686	0.278	0.813	1.838	1.603	6.288	11.048	7.171
50	C	-0.114	-0.159	-0.220	0.892	0.863	0.471	0.304	0.316	0.567	3.211	3.292	3.673
	A	-0.097	-0.114	-0.119	0.888	0.876	0.542	0.415	0.637	1.714	3.334	3.860	8.318
100	C	-0.080	-0.105	-0.193	0.951	0.864	0.536	0.173	0.548	0.680	2.975	3.465	3.782
	A	-0.067	-0.074	-0.108	0.947	0.890	0.622	0.248	0.778	1.630	3.023	3.931	7.761
150	C	-0.049	-0.094	-0.141	0.890	0.836	0.652	0.078	0.263	0.708	2.725	2.970	3.833
	A	-0.039	-0.068	-0.057	0.886	0.841	0.788	0.132	0.420	1.644	2.750	3.198	7.795
200	C	-0.047	-0.121	-0.155	0.981	0.899	0.622	0.397	0.427	0.609	3.358	3.464	3.718
	A	-0.037	-0.098	-0.088	0.984	0.909	0.791	0.460	0.596	1.252	3.464	3.787	5.852

Results from 500 trials.

TABLE 2.5.2 Null distribution of $T_{GL}(C)$ and $T_{GL}(A)$.

n	$T_{GL}(.)$	$\mu_1(T_{GL}(.)/H_G)$			$\mu_2(T_{GL}(.)/H_G)$			$\gamma_1(T_{GL}(.)/H_G)$			$\beta_2(T_{GL}(.)/H_G)$		
		$\gamma_2=0.5$	$\gamma_2=1.0$	$\gamma_2=10.0$	$\gamma_2=0.5$	$\gamma_2=1.0$	$\gamma_2=10.0$	$\gamma_2=0.5$	$\gamma_2=1.0$	$\gamma_2=10.0$	$\gamma_2=0.5$	$\gamma_2=1.0$	$\gamma_2=10.0$
20	C	-0.198	-0.145	-0.163	0.670	0.585	0.701	0.435	0.426	0.140	3.667	3.361	2.975
	A	-0.089	-0.077	-0.143	0.737	0.621	0.696	1.438	1.020	0.278	6.522	4.757	3.107
50	C	-0.152	-0.102	-0.186	0.772	0.801	0.845	0.306	0.683	0.111	3.129	3.731	3.343
	A	-0.073	-0.042	0.169	0.805	0.872	0.839	0.877	1.134	0.231	4.237	4.876	3.472
100	C	-0.021	-0.092	-0.091	0.865	0.916	0.990	0.298	0.395	0.156	3.016	3.650	3.463
	A	0.041	-0.044	-0.077	0.919	0.952	0.991	0.709	0.797	0.252	3.882	4.701	3.498
150	C	-0.076	-0.093	-0.144	0.918	0.962	0.989	0.363	0.684	0.126	3.184	3.734	3.476
	A	-0.022	-0.052	-0.133	0.958	1.015	0.987	0.723	0.990	0.207	3.832	4.771	3.523
200	C	-0.049	-0.002	-0.120	1.001	1.011	0.996	0.457	0.448	-0.056	3.268	3.261	2.969
	A	0.002	0.036	-0.111	1.054	1.053	0.991	0.777	0.688	-0.001	3.872	3.747	2.973

Results from 500 trials.

TABLE 2.5.3 Distribution of $T_{LG}(C)$ and $T_{LG}(A)$ under alternative H_G.

n	$T_{LG}(.)$	$\mu_1(T_{LG}(.)/H_G)$			$\mu_2(T_{LG}(.)/H_G)$			$\gamma_2(T_{LG}(.)/H_G)$			$\beta_2(T_{LG}(.)/H_G)$		
		$\gamma_2=0.5$	$\gamma_2=1.0$	$\gamma_2=10.0$	$\gamma_2=0.5$	$\gamma_2=1.0$	$\gamma_2=10.0$	$\gamma_2=0.5$	$\gamma_2=1.0$	$\gamma_2=10.0$	$\gamma_2=0.5$	$\gamma_2=1.0$	$\gamma_2=10.0$
20	C	-0.597	-0.981	-0.500	0.088	0.239	0.675	-0.101	0.417	-0.026	3.803	3.814	2.858
	A	-0.337	-0.745	-0.471	0.048	0.126	0.626	-0.159	0.914	-0.110	4.864	5.621	2.844
50	C	-0.987	-1.708	-0.814	0.111	0.258	0.791	-0.202	-0.064	-0.008	3.299	3.507	3.340
	A	-0.509	-1.291	-0.782	0.065	0.096	0.734	-0.396	0.259	-0.115	2.913	3.886	3.305
100	C	-1.414	-2.507	-1.264	0.128	0.287	0.966	0.254	0.251	-0.051	3.224	3.240	3.514
	A	-0.683	-1.906	-1.225	0.078	0.099	0.889	-0.159	0.488	-0.062	2.603	3.817	3.547
150	C	-1.717	-3.101	-1.492	0.123	0.282	0.968	0.045	-0.207	-0.052	3.518	2.866	3.366
	A	0.811	-2.353	-1.453	0.081	0.085	0.900	-0.317	-0.122	0.029	2.961	2.738	3.362
200	C	-1.979	-3.658	-1.751	0.124	0.295	0.979	0.258	-0.055	0.093	3.008	2.920	2.904
	A	-0.917	-2.767	-1.708	0.079	0.087	0.911	-0.067	0.045	0.150	2.831	3.131	2.919

Results from 500 trials.

TABLE 2.5.4 Distribution of $T_{GL}(C)$ and $T_{GL}(A)$ under alternative H_L.

n	$T_{GL}(.)$	$\mu_1(T_{GL}(.)/H_L)$			$\mu_2(T_{GL}(.)/H_L)$			$\gamma_2(T_{GL}(.)/H_L)$			$\beta_2(T_{GL}(.)/H_L)$		
		$\alpha_2=0.1$	$\alpha_2=0.5$	$\alpha_2=2.0$	$\alpha_2=0.1$	$\alpha_2=0.5$	$\alpha_2=2.0$	$\alpha_2=0.1$	$\alpha_2=0.5$	$\alpha_2=2.0$	$\alpha_2=0.1$	$\alpha_2=0.5$	$\alpha_2=2.0$
20	C	-0.656	-0.953	-1.413	0.680	0.586	0.591	-0.137	-0.150	-0.292	4.936	4.593	3.457
	A	-0.623	-0.846	-1.102	0.624	0.428	0.249	0.184	-0.590	0.658	5.014	4.927	3.725
50	C	-0.876	-1.541	-2.588	0.877	0.879	1.027	-0.213	-0.025	-0.330	3.048	3.104	3.339
	A	-0.846	-1.395	-1.992	0.815	0.649	0.357	-0.102	0.303	0.451	2.971	3.150	3.128
100	C	-1.267	-2.323	-3.880	0.952	0.649	1.160	-0.139	0.386	-0.573	2.899	3.245	3.719
	A	-1.233	-2.128	-2.991	0.885	0.598	0.353	-0.061	-0.143	0.113	2.890	2.999	3.092
150	C	-1.572	-2.919	-4.901	0.851	0.622	1.470	-0.032	-0.135	-0.693	2.752	2.866	3.992
	A	-1.536	-2.679	-3.760	0.793	0.586	0.416	-0.023	0.031	-0.025	2.742	2.815	2.907
200	C	-1.814	-3.354	-5.683	0.955	0.500	1.327	-0.367	-0.268	-0.489	3.281	3.194	3.378
	A	-1.774	-3.085	-4.377	0.885	0.638	0.393	-0.303	-0.096	-0.007	3.197	3.058	2.936

Results from 500 trials.

TABLE 2.5.5 Null: lognormal; Alternative: gamma. Tests: $T_{LG}(C)$, $T_{LG}(A)$. Power at t = -1.64; t = -1.28.

n	$T_{LG}(.)$	POWER FUNCTION											
		SL = 0.05						SL = 0.10					
		γ_2=0.5	γ_2=0.8	γ_2=1.0	γ_2=2.0	γ_2=5.0	γ_2=10.0	γ_2=0.5	γ_2=0.8	γ_2=1.0	γ_2=2.0	γ_2=5.0	γ_2=10.0
20	C	0.002	0.044	0.072	0.122	0.110	0.082	0.024	0.146	0.028	0.304	0.200	0.162
	A	0	0	0.002	0.046	0.080	0.056	0	0.022	0.040	0.186	0.180	0.146
50	C	0.028	0.484	0.558	0.444	0.266	0.156	0.176	0.824	0.802	0.644	0.414	0.300
	A	0	0.022	0.122	0.350	0.220	0.148	0.004	0.238	0.518	0.592	0.398	0.278
100	C	0.256	0.972	0.950	0.844	0.456	0.330	0.682	0.996	0.982	0.930	0.648	0.476
	A	0	0.472	0.802	0.796	0.436	0.316	0.022	0.938	0.974	0.912	0.628	0.468
150	C	0.602	0.998	1.000	0.946	0.668	0.438	0.902	1.000	1.000	0.982	0.792	0.602
	A	0.002	0.914	0.998	0.930	0.650	0.416	0.050	0.998	1.000	0.976	0.784	0.596
200	C	0.824	1.000	1.000	0.980	0.826	0.568	0.968	1.000	1.000	0.992	0.890	0.690
	A	0.008	0.996	1.000	0.978	0.812	0.542	0.100	1.000	1.000	0.990	0.890	0.674

Results from 500 trials.

TABLE 2.5.6 Null: gamma; Alternative: lognormal. Tests: $T_{GL}(C)$, $T_{GL}(A)$. Power at t = -1.64; t = -1.28.

n	$T_{GL}(.)$	POWER FUNCTION											
		SL = 0.05						SL = 0.10					
		α_2=0.1	α_2=0.25	α_2=0.5	α_2=0.64	α_2=1.0	α_2=2.0	α_2=0.1	α_2=0.25	α_2=0.5	α_2=0.64	α_2=1.0	α_2=2.0
20	C	0.106	0.112	0.166	0.188	0.218	0.360	0.194	0.226	0.306	0.304	0.418	0.544
	A	0.088	0.086	0.096	0.094	0.088	0.136	0.174	0.188	0.236	0.234	0.286	0.380
50	C	0.194	0.316	0.444	0.504	0.634	0.828	0.314	0.446	0.620	0.668	0.786	0.912
	A	0.186	0.286	0.380	0.434	0.560	0.742	0.312	0.410	0.566	0.618	0.740	0.878
100	C	0.338	0.568	0.754	0.828	0.952	0.986	0.490	0.702	0.870	0.926	0.980	0.996
	A	0.326	0.544	0.728	0.792	0.916	0.982	0.484	0.692	0.860	0.916	0.976	0.994
150	C	0.472	0.730	0.928	0.954	0.992	1.000	0.618	0.846	0.968	0.986	0.998	1.000
	A	0.462	0.720	0.918	0.946	0.938	1.000	0.612	0.838	0.966	0.980	0.998	1.000
200	C	0.552	0.846	0.970	0.988	0.998	1.000	0.688	0.910	0.998	0.998	1.000	1.000
	A	0.540	0.796	0.968	0.988	0.998	1.000	0.680	0.906	0.988	0.998	1.000	1.000

Results from 500 trials.

TABLE 2.5.7 Null: lognormal; Alternative: gamma. Tests: $T_{LG}(C)$, $T_{LG}(A)$. One-side significance level at $t = -1.64$; $t = -1.28$.

n	$T_{LG}(\cdot)$	SIGNIFICANCE LEVEL											
		SL = 0.05						SL = 0.10					
		$\alpha_2=0.1$	$\alpha_2=0.25$	$\alpha_2=0.5$	$\alpha_2=0.64$	$\alpha_2=1.0$	$\alpha_2=2.0$	$\alpha_2=0.1$	$\alpha_2=0.25$	$\alpha_2=0.5$	$\alpha_2=0.64$	$\alpha_2=1.0$	$\alpha_2=2.0$
20	C	0.022	0.020	0.012	0.014	0.008	0.002	0.054	0.052	0.052	0.058	0.032	0.016
	A	0.018	0.008	0.006	0.006	0	0	0.048	0.044	0.034	0.030	0.012	0
50	C	0.040	0.060	0.048	0.024	0.014	0.006	0.102	0.128	0.116	0.078	0.060	0.044
	A	0.034	0.050	0.032	0.012	0.004	0	0.092	0.110	0.098	0.052	0.030	0
100	C	0.042	0.036	0.028	0.024	0.020	0.014	0.104	0.092	0.086	0.084	0.068	0.032
	A	0.038	0.030	0.024	0.016	0.018	0	0.098	0.088	0.074	0.070	0.042	0.014
150	C	0.048	0.038	0.032	0.040	0.026	0.012	0.094	0.090	0.090	0.094	0.076	0.054
	A	0.040	0.034	0.026	0.028	0.016	0	0.092	0.084	0.070	0.080	0.048	0.020
200	C	0.038	0.050	0.046	0.042	0.026	0.018	0.088	0.116	0.088	0.082	0.088	0.046
	A	0.036	0.044	0.038	0.030	0.022	0.002	0.086	0.102	0.084	0.076	0.062	0.026

Results from 500 trials.

TABLE 2.5.8 Null: gamma; Alternative: lognormal. Tests: $T_{GL}(C)$, $T_{GL}(A)$. One-side significance level at $t = -1.64$; $t = -1.28$.

n	$T_{GL}(\cdot)$	SIGNIFICANCE LEVEL											
		SL = 0.05						SL = 0.10					
		$\gamma_2=0.5$	$\gamma_2=0.8$	$\gamma_2=1.0$	$\gamma_2=2.0$	$\gamma_2=5.0$	$\gamma_2=10.0$	$\gamma_2=0.5$	$\gamma_2=0.8$	$\gamma_2=1.0$	$\gamma_2=2.0$	$\gamma_2=5.0$	$\gamma_2=10.0$
20	C	0.024	0.018	0.016	0.028	0.020	0.038	0.060	0.066	0.046	0.066	0.068	0.086
	A	0.004	0	0.004	0.018	0.016	0.028	0.024	0.026	0.020	0.050	0.050	0.082
50	C	0.032	0.020	0.024	0.030	0.036	0.064	0.086	0.064	0.064	0.076	0.070	0.120
	A	0.006	0.006	0.008	0.012	0.022	0.060	0.046	0.038	0.046	0.066	0.060	0.114
100	C	0.030	0.024	0.036	0.032	0.036	0.040	0.074	0.072	0.104	0.076	0.080	0.118
	A	0.016	0.014	0.022	0.028	0.028	0.038	0.048	0.050	0.084	0.070	0.078	0.108
150	C	0.036	0.050	0.028	0.034	0.036	0.076	0.100	0.092	0.098	0.084	0.096	0.120
	A	0.022	0.032	0.020	0.028	0.032	0.068	0.072	0.072	0.076	0.076	0.088	0.116
200	C	0.048	0.044	0.028	0.046	0.032	0.060	0.086	0.096	0.084	0.090	0.086	0.138
	A	0.034	0.038	0.022	0.036	0.022	0.058	0.080	0.082	0.068	0.084	0.078	0.132

Results from 500 trials.

TABLE 2.5.9 Power and significance level at $t = -0.84$.

n	Tests	Power at 20% SL		20% SL	
		$T_{GL}(.)$	$T_{LG}(.)$	$T_{GL}(.)$	$T_{LG}(.)$
		$\alpha_2=0.64$	$\gamma_2=2.0$	$\gamma_2=2.0$	$\alpha_2=0.5$
20	C	0.562	0.556	0.160	0.124
	A	0.520	0.520	0.144	0.114
50	C	0.828	0.856	0.186	0.226
	A	0.820	0.848	0.178	0.214
100	C	0.978	0.970	0.190	0.216
	A	0.976	0.968	0.182	0.210

Results from 500 trials.

FIGURE 2.5.1 - NULL DISTRIBUTION OF $T_{LG}(\cdot)$ - LOGNORMAL vs GAMMA, $\alpha_2 = 0.1$, 500 trials, (see Appendix B).

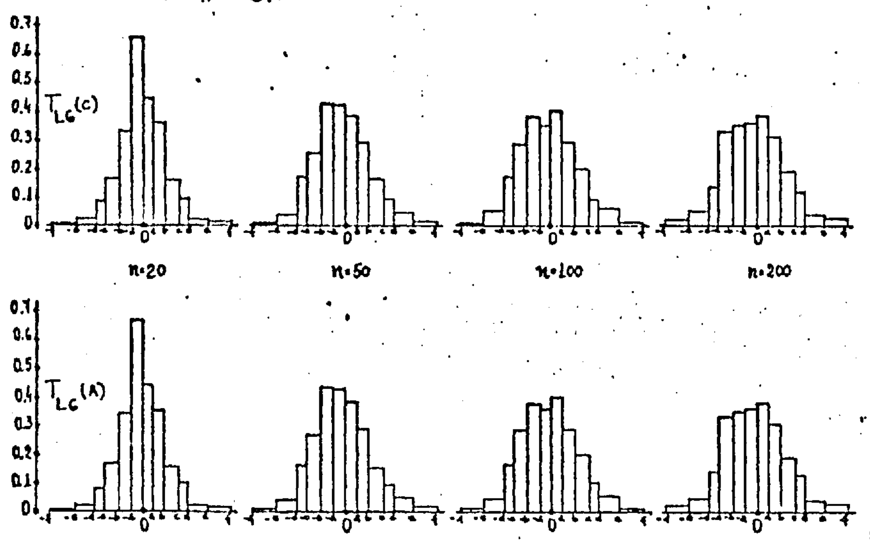


FIGURE 2.5.2 - NULL DISTRIBUTION OF $T_{LG}(\cdot)$ - LOGNORMAL vs GAMMA, $\alpha_2 = 0.5$, 500 trials, (see Appendix B).

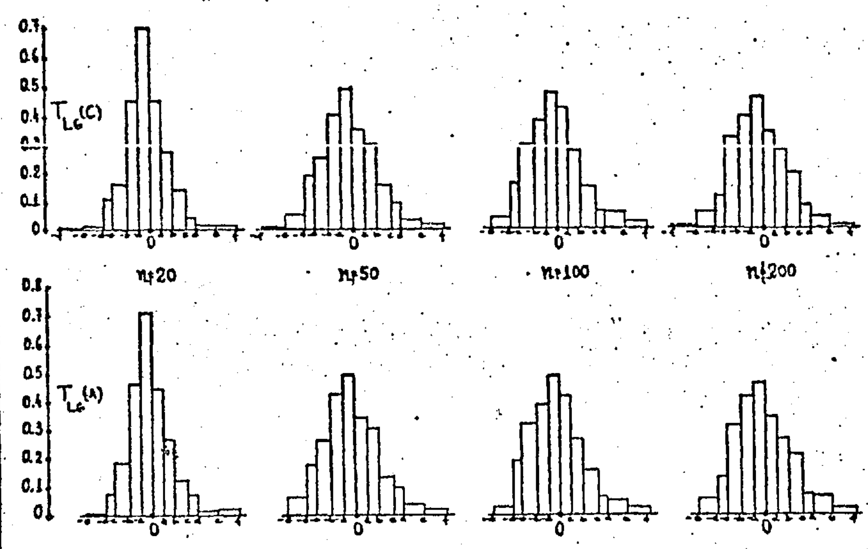


FIGURE 2.5.3 - NULL DISTRIBUTION OF $T_{LG}(\cdot)$ - LOGNORMAL vs GAMMA, $\alpha_2 = 2.0$, 500 trials. (see Appendix B)

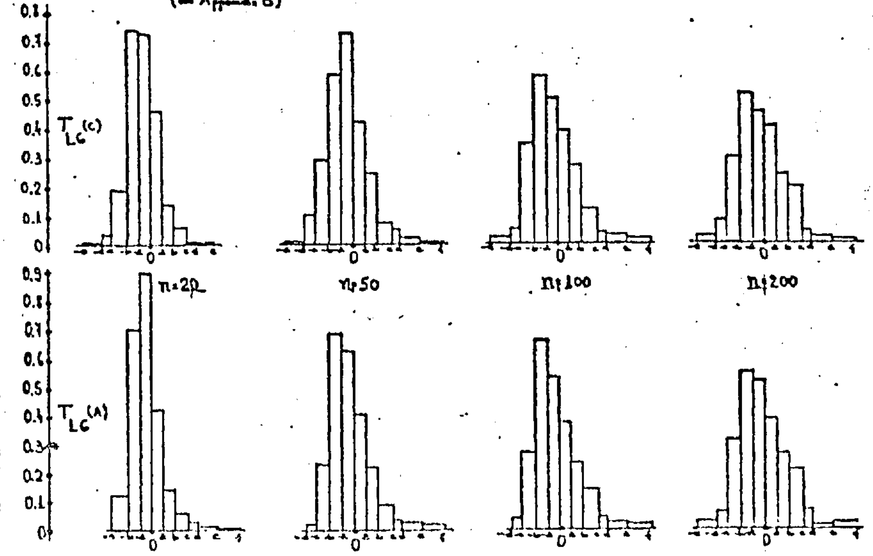


FIGURE 2.5.4 - NULL DISTRIBUTION OF $T_{GL}(\cdot)$ - GAMMA vs LOGNORMAL, $\gamma_2 = 0.6$, 500 trials, (see Appendix B)

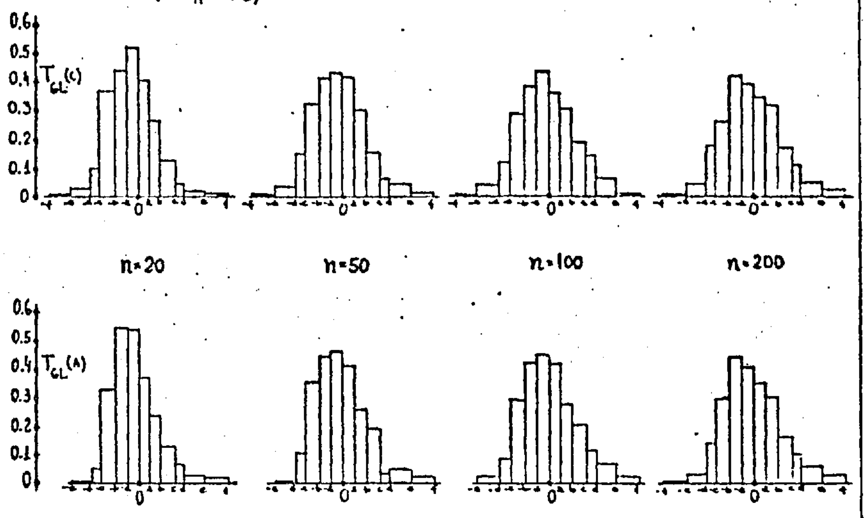


FIGURE 2.5.5 - NULL DISTRIBUTION OF $T_{GL}(\cdot)$ - GAMMA vs LOGNORMAL, $\gamma_2 = 1.0$, 500 trials, (see Appendix B).

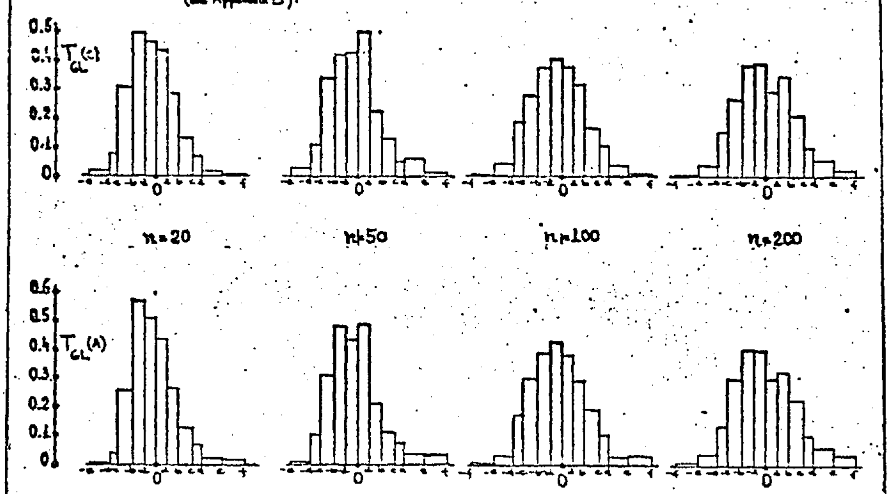
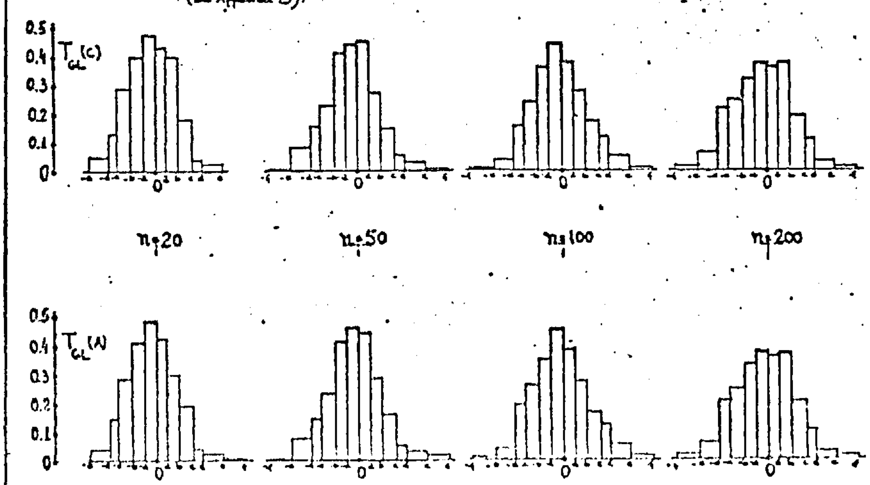


FIGURE 2.5.6 - NULL DISTRIBUTION OF $T_{GL}(\cdot)$ - GAMMA vs LOGNORMAL, $\gamma_2 = 10.0$, 500 trials, (see Appendix B).



2.6 Tests for the lognormal and Weibull distributions

A Test statistics and their distributions

Here the methods proposed by Cox and by Atkinson are used to derive tests involving the lognormal and Weibull distribution.

First suppose the null hypothesis H_L is that the distribution is lognormal and the alternative H_W that it is Weibull, that is $H_L : f_L(\underset{\sim}{y}, \alpha_1, \alpha_2)$ against $H_W : f_W(\underset{\sim}{y}; \beta_1, \beta_2)$; see Section 1.2. The expectations of the log likelihood functions in relation to the null lognormal distribution yield

$$E_L\{\ell_L(\alpha_1,\alpha_2;\underset{\sim}{y})\} = -\frac{n}{2}\log\alpha_2 - n\log\sqrt{2\pi} - n\,\alpha_1 - \frac{n}{2}\,,$$

$$E_L\{\ell_W(\beta_1,\beta_2;\underset{\sim}{y})\} = n\log\beta_2 - n\,\beta_2\log\beta_1 + (\beta_2-1)\alpha_1 - \frac{n}{\beta_1^{\beta_2}}\exp\{\beta_2\alpha_1+\frac{\beta_2^2}{2}\alpha_2\}$$

(2.6.1)

To find β_{1L} and β_{2L}, the probability limits under H_L of $\hat{\beta}_1$ and $\hat{\beta}_2$ respectively, recall Cox [1961, eq.(25)], namely

$$E_L\left\{\frac{\partial\log f_W(\underset{\sim}{y},\beta_{1L},\beta_{2L})}{\partial(\beta_1,\beta_2)'}\right\} = \frac{\partial}{\partial(\beta_1,\beta_2)'}\,E_L\{\ell_W(\beta_{1L},\beta_{2L};\underset{\sim}{y})\} = \underset{\sim}{0}. \tag{2.6.2}$$

This gives a system of equations whose unique solution is

$$\beta_{1L} = \exp\{\alpha_1 + \tfrac{1}{2}\sqrt{\alpha_2}\}\quad,\quad \beta_{2L} = \alpha_2^{-\frac{1}{2}}. \tag{2.6.3}$$

This shows that $\beta_{1L}^{\beta_{2L}}$ is the β_{2L}th moment of the lognormal distribution and β_{2L} is the inverse of the scale of the normal distribution. Writing $\hat{L} \equiv (\hat{\alpha}_1,\hat{\alpha}_2)$ and by noticing that $\ell_L(\hat{\alpha}_1,\hat{\alpha}_2,\underset{\sim}{y}) - E_{\hat{L}}\{\ell_L(\alpha_1,\alpha_2;\underset{\sim}{y})\} = 0$ we then have

$$\begin{aligned} T^*_{LW}(C) &= E_{\hat{L}}\{\ell_W(\beta_{1L},\beta_{2L};\underset{\sim}{y})\} - \ell_W(\hat{\beta}_1,\hat{\beta}_2,\underset{\sim}{y}) \\ &= n\{\hat{\beta}_2\log\hat{\beta}_1 - \beta_{2\hat{L}}\log\beta_{1\hat{L}} - \log\frac{\hat{\beta}_2}{\beta_{2\hat{L}}} - \hat{\alpha}_1(\hat{\beta}_2-\beta_{2\hat{L}})\}\,; \end{aligned} \tag{2.6.4}$$

$$T^*_{LW}(A) = E_{\hat{L}}\{\ell_W(\beta_{1L},\beta_{2L};\underset{\sim}{y})\} - \ell_W(\beta_{1\hat{L}},\beta_{2\hat{L}},\underset{\sim}{y})$$

$$= \sum_{i=1}^{n}\left(\frac{y_i}{\beta_{1\hat{L}}}\right)^{\beta_{2\hat{L}}} - n = \frac{\sum_{i=1}^{n} y_t^{(\hat{\alpha}_2)^{-\frac{1}{2}}}}{\exp\{\hat{\alpha}_1(\hat{\alpha}_2)^{-\frac{1}{2}} + \frac{1}{2}\}} - n. \qquad (2.6.5)$$

The asymptotic variance $V_L\{T^*_{LW}\}$ of these tests is required. First we evaluate

$$LW = \frac{n}{2} + n\left\{e^{\beta^2_{2L}\alpha_2} - 1\right\} - 2n\beta_{2L}\alpha_2 = 0.218281\ n,$$

$$C_{1L} = 0 \quad , \quad C_{2L} = \frac{n}{2}\left\{\beta^2_{2L} - \frac{1}{\alpha_2}\right\} = 0 \ , \qquad (2.6.6)$$

and recalling the information matrix in (1.2.7) we have

$$V_L\{T^*_{LW}\} = LW - \underset{\sim}{C}'_L\ I^{-1}\ (\alpha_1,\alpha_2)\underset{\sim}{C}_L = 0.218281n \qquad (2.6.7)$$

Now, suppose that H_L and H_W change roles, so that the null distribution is Weibull and the alternative is lognormal. The expectations of the log likelihood functions in relation to the null lognormal distribution yield

$$E_W\{\ell_W(\beta_1,\beta_2;\underset{\sim}{y})\} = n\log\beta_2 - n\beta_2\log\beta_1 + (\beta_2-1)n\left\{\frac{\psi(1)}{\beta_2} + \log\beta_1\right\} - n\ ,$$

$$E_W\{\ell_L(\alpha_1,\alpha_2;\underset{\sim}{y})\} = -\frac{n}{2}\log\alpha_2 - n\log\sqrt{2\pi} - n\left\{\frac{\psi(1)}{\beta_2} + \log\beta_1\right\}$$

$$- \frac{n}{2\alpha_2}\left[\frac{\psi'(1)}{\beta_2^2} + \left\{\frac{\psi(1)}{\beta_2} + \log\beta_1\right\}^2 - 2\alpha_1\left\{\frac{\psi(1)}{\beta_2} + \log\beta_1\right\} + \alpha_1^2\right]. \qquad (2.6.8)$$

To find α_{1W} and α_{2W}, the probability limits under H_W of $\hat{\alpha}_1$ and $\hat{\alpha}_2$, respectively, the analogue to (2.6.2) is

$$\frac{\partial}{\partial(\alpha_1,\alpha_2)'}\, E_W\{\ell_L(\alpha_{1W},\alpha_{2W},\underset{\sim}{y})\} = \underset{\sim}{0}\ ,$$

whose unique solution is

$$\alpha_{1W} = \frac{\psi(1)}{\beta_2} + \log \beta_1 \ , \qquad \alpha_{2W} = \frac{\psi'(1)}{\beta_2^2} \ . \tag{2.6.9}$$

Thus, α_{1W} and α_{2W} are respectively the mean and variance of the logarithm of a random variable with a Weibull distribution. Writing $\hat{W} \equiv (\hat{\beta}_1, \hat{\beta}_2)$, we then have

$$\begin{aligned} T^*_{WL}(C) &= \ell_W(\hat{\beta}_1, \hat{\beta}_2; \underset{\sim}{y}) - \ell_L(\hat{\alpha}_1, \hat{\alpha}_2, \underset{\sim}{y}) - E_{\hat{W}}\{\ell_W(\beta_1, \beta_2; \underset{\sim}{y}) - \ell_L(\alpha_{1W}, \alpha_{2W}; \underset{\sim}{y})\} \\ &= n\{\hat{\beta}_2(\hat{\alpha}_1 - \alpha_{1\hat{W}}) + \tfrac{1}{2} \log \frac{\hat{\alpha}_2}{\alpha_{2\hat{W}}} \ , \end{aligned} \tag{2.6.10}$$

$$\begin{aligned} T^*_{WL}(A) &= \ell_W(\hat{\beta}_1, \hat{\beta}_2; \underset{\sim}{y}) - \ell_L(\alpha_{1\hat{W}}, \alpha_{2\hat{W}}; \underset{\sim}{y}) - E_{\hat{W}}\{\ell_W(\beta_1, \beta_2; \underset{\sim}{y}) - \ell_L(\alpha_{1W}, \alpha_{2W}; \underset{\sim}{y})\} \\ &= n\left[\hat{\beta}_2(\hat{\alpha}_1 - \alpha_{1\hat{W}}) + \frac{1}{2\alpha_{2\hat{W}}} \{\hat{\alpha}_2 - \alpha_{2\hat{W}} + (\hat{\alpha}_1 - \alpha_{1\hat{W}})^2\}\right] \ . \end{aligned} \tag{2.6.11}$$

To evaluate the variance $V_W\{T^*_{WL}\}$ of these tests, we have similarly

$$WL = 0.2834 \ n \quad ; \quad C_{1W} = 0 \quad , \quad C_{2W} = 0 \quad , \tag{2.6.12}$$

and, with the information matrix in (1.2.11), we obtain

$$V_W\{T^*_{WL}\} = WL - \underset{\sim}{C}'_W \ I^{-1} \ (\beta_1, \beta_2) \ \underset{\sim}{C}_W = 0.2834 \ n \ . \tag{2.6.13}$$

It should be noted that for β_2 known and equal to 1 the previous results recover those of Section 2.4.

Finally, for j = A, C, the statistics

$$T_{LW}(j) = T^*_{LW}(j) \ [V_L\{T^*_{LW}\}]^{-\frac{1}{2}} \ , \ T_{WL}(j) = T^*_{WL}(j) \ [V_W\{T^*_{WL}\}]^{-\frac{1}{2}} \tag{2.6.14}$$

are asymptotically standard normally distributed under H_L and H_W, respectively.

B Empirical results

The empirical results for comparisons between $T_{LW}(C)$ and $T_{LW}(A)$ and between $T_{WL}(C)$ and $T_{WL}(A)$ and on the adequacy of the asymptotic results are presented.

Results on the null distribution of $T_{LW}(C)$ and $T_{LW}(A)$ and on the distribution of $T_{LW}(C)$ and $T_{WL}(A)$ under the alternative, that is the lognormal distribution, was obtained as in Section 2.4. Here from the comments about (1.2.23) it follows that the distribution of the test statistics is independent of the parameter values α_1 and α_2. For various sample sizes n, 1000 trials were obtained with $\alpha_1 = 0$ and $\alpha_2 = 1$.

Similarly, results on the null distribution of $T_{WL}(C)$ and $T_{WL}(A)$ and on the distribution of $T_{LW}(C)$ and $T_{LW}(A)$ under the alternative, that is the Weibull distribution, were obtained. Again it follows from (1.2.23) that the distribution of the test statistics is independent of the parameters β_1 and β_2. For various sample sizes n, 1000 trials were obtained with $\beta_1 = 1$ and $\beta_2 = 1$, the standard exponential distribution.

The maximum likelihood estimator equation for β_2 was solved using Newton's method. The iterations stopped when the equation differed from zero by less than 0.001.

The results are summarized in Tables 2.6.1 to 2.6.9.

Results in Table 2.6.1 and 2.6.2 agree with the discussion of Section 2.3 about the first two moments and the coefficients of skewness and kurtosis of the A and C statistics. For Table 2.6.1 one of the terms which differentiate $T_{WL}(A)$ from $T_{WL}(C)$ depends on

$$\frac{\partial}{\partial\beta_1} \ell_W(\beta_{1\hat{L}},\beta_{2\hat{L}};\underset{\sim}{y}) = \frac{\beta_{2\hat{L}}}{\beta_{1\hat{L}}} \sum_{i=1}^{n} \left(\frac{y_1}{\beta_{1\hat{L}}^{\beta_{2\hat{L}}}} - 1 \right) \quad . \tag{2.6.15}$$

From the properties of the lognormal distribution $y_i^{\beta_{2\hat{L}}} \big/ \beta_{1\hat{L}}^{\beta_{2\hat{L}}}$ has a lognormal distribution with $\alpha_1 = -\frac{1}{2}$ and $\alpha_2 = 1$. Therefore, since for large α_2, the sample mean is an inefficient estimator for the mean of the lognormal distribution it will be required a large sample size for (2.6.15) to become negligible.

For Table 2.6.2 the terms which differentiate $T_{WL}(A)$ from $T_{WL}(C)$ depends on

$$\frac{\partial}{\partial\alpha_1} \ell_L(\alpha_{1\hat{W}},\alpha_{2\hat{W}};\underset{\sim}{y}) = \frac{1}{\alpha_{2\hat{W}}} \sum_{i=1}^{n} (\log y_i - \alpha_{1\hat{W}}) , \qquad (2.6.16)$$

$$\frac{\partial}{\partial\alpha_2} \ell_L(\alpha_{1\hat{W}},\alpha_{2\hat{W}};\underset{\sim}{y}) = -\frac{n}{2\alpha_{2\hat{W}}} + \frac{1}{2\alpha_{2\hat{W}}^2} \sum_{i=1}^{n} (\log y_i - \alpha_{1\hat{W}})^2. \quad (2.6.17)$$

It is known that for the extreme value distribution, the efficiency of the method of the moments in relation to maximum likelihood in estimating the location parameter is about 95% and for the scale parameter is about 55%. Therefore, at least (2.4.17) will require a large sample size to become negligible.

Tables 2.6.3 and 2.6.4 show respectively that the tests $T_{LW}(A)$ and $T_{WL}(A)$ are consistent for all parameter values. This follows from the fact mentioned earlier that the distributions of the tests are independent of the parameters.

The following relation can be observed from Tables 2.6.1 to 2.6.4:

$$\gamma_1(T_{LW}(C)/H_L) = -\gamma_1(T_{WL}(C)/H_L),\ \gamma_1(T_{WL}(C)/H_W) = -\gamma_1(T_{LW}(C)/H_W),$$

$$\beta_2(T_{LW}(C)/H_L) = \beta_2(T_{WL}(C)/H_L),\ \beta_2(T_{WL}(C)/H_W) = \beta_2(T_{LW}(C)/H_L).$$

For the significance levels in Tables 2.6.7 and 2.6.8 the C statistics show a better agreement to the asymptotic values. This is related to the approach to normality and would suggest that C is preferable. For power comparisons, Table 2.6.9 gives further results and they seem to indicate that there is not much difference of power between the A and the C statistics.

Figures 2.6.1 and 2.6.2 presents the histograms of the data of Tables 2.6.1 and 2.6.2.

TABLE 2.6.1 Null distribution of $T_{LW}(C)$ and $T_{LW}(A)$.

n	$T_{LW}(.)$	$\mu_1\{T_{LW}(.)/H_L\}$	$\mu_2\{T_{LW}(.)/H_L\}$	$\gamma_1\{T_{LW}(.)/H_L\}$	$\beta_2\{T_{LW}(.0/H_L\}$
20	C	-0.261	0.502	0.090	3.387
	A	-0.118	0.503	1.665	8.366
50	C	-0.232	0.686	0.167	3.131
	A	-0.103	0.723	1.433	8.033
100	C	-0.198	0.758	0.329	3.197
	A	-0.092	0.818	1.186	5.602
150	C	-0.163	0.769	0.298	2.867
	A	-0.072	0.832	0.880	4.000
200	C	-0.142	0.805	0.355	3.368
	A	-0.058	0.882	1.088	5.511

Results from 1000 trials.

TABLE 2.6.2 Null distribution of $T_{WL}(C)$ and $T_{WL}(A)$.

n	$T_{WL}(.)$	$\mu_1\{T_{LW}(.)/H_W\}$	$\mu_2\{T_{WL}(.)/H_W\}$	$\gamma_1\{T_{WL}(.)/H_W\}$	$\beta_2\{T_{WL}(.)/H_W\}$
20	C	-0.224	0.555	0.492	3.459
	A	-0.084	0.665	1.777	7.723
50	C	-0.094	0.918	0.512	3.480
	A	-0.043	1.089	1.406	6.059
100	C	-0.078	0.884	0.371	3.406
	A	0.011	0.957	0.984	4.481
150	C	-0.055	0.967	0.283	3.391
	A	0.023	1.018	0.824	4.335
200	C	-0.067	0.968	0.395	3.344
	A	-0.001	1.016	0.815	4.111

Results from 1000 trials.

TABLE 2.6.3 Distribution of $T_{LW}(C)$ and $T_{LW}(A)$ under alternative H_W.

n	$T_{LW}(.)$	$\mu_1\{T_{LW}(.)/H_W\}$	$\mu_2\{T_{LW}(.)/H_W\}$	$\gamma_1\{T_{LW}(.)/H_W\}$	$\beta_2\{T_{LW}(.)/H_W\}$
20	C	-1.387	0.720	-0.492	3.459
	A	-0.913	0.215	0.510	3.776
50	C	-2.419	1.003	-0.562	3.950
	A	-1.638	0.266	0.155	3.519
100	C	-3.584	1.143	-0.371	3.406
	A	-2.445	0.297	0.126	3.502
150	C	-4.436	1.256	-0.283	3.391
	A	-3.038	0.324	-0.116	3.415
200	C	-5.119	1.257	0.395	3.344
	A	-3.522	0.323	0.099	3.162

Results from 1000 trials.

TABLE 2.6.4 Distribution of $T_{WL}(C)$ and $T_{WL}(A)$ under alternative H_L.

n	$T_{WL}(.)$	$\mu_1\{T_{WL}(.)/H_L\}$	$\mu_2\{T_{WL}(.)/H_L\}$	$\gamma_1\{T_{WL}(.)/H_L\}$	$\beta_2\{T_{WL}(.)/H_L\}$
20	C	-1.213	0.387	-0.090	3.387
	A	-0.858	0.122	1.380	6.072
50	C	-2.076	0.528	-0.167	3.131
	A	-1.451	0.118	0.857	3.625
100	C	-3.050	0.584	-0.329	3.197
	A	-2.120	0.104	0.581	3.379
150	C	-3.806	0.608	-0.298	2.867
	A	-2.631	0.098	0.407	3.027
200	C	-4.433	0.670	-0.546	4.164
	A	-3.049	0.097	0.470	3.137

Results from 1000 trials.

TABLE 2.6.5 Null: lognormal; Alternative: Weibull. Tests: $T_{LW}(C)$, $T_{LW}(A)$.

Power at t = -1.64; t = -1.28.

n	$T_{LW}(.)$	POWER FUNCTION	
		SL = 0.05	SL = 0.10
20	C	0.344	0.506
	A	0.045	0.217
50	C	0.771	0.887
	A	0.511	0.756
100	C	0.974	0.986
	A	0.940	0.977
150	C	0.994	0.997
	A	0.989	0.996
200	C	1.000	1.000
	A	1.000	1.000

Results from 1000 trials.

TABLE 2.6.6 Null: Weibull; Alternative: lognormal. Tests. $T_{WL}(C)$, $T_{WL}(A)$.

Power at t = -1.64; t = -1.28

n	$T_{WL}(.)$	POWER FUNCTION	
		SL = 0.05	SL = 0.10
20	C	0.231	0.447
	A	0	0.057
50	C	0.738	0.860
	A	0.330	0.751
100	C	0.973	0.996
	A	0.925	0.986
150	C	0.999	1.000
	A	0.996	1.000
200	C	1.000	1.000
	A	1.000	1.000

Results from 1000 trials.

TABLE 2.6.7 Null: lognormal; Alternative: Weibull. Tests: $T_{LW}(C)$, $T_{LW}(A)$.

One-side significance levels at t = -1.64; t = -1.28.

n	$T_{LW}(.)$	SIGNIFICANCE LEVEL	
		SL = 0.05	SL = 0.10
20	C	0.022	0.071
	A	0	0.010
50	C	0.043	0.106
	A	0.001	0.042
100	C	0.040	0.093
	A	0.008	0.051
150	C	0.032	0.096
	A	0.009	0.053
200	C	0.041	0.101
	A	0.016	0.067

Results from 1000 trials.

TABLE 2.6.8 Null: Weibull; Alternative: lognormal. Tests: $T_{WL}(C)$, $T_{WL}(A)$.

One-side significance level at t = -1.64; t = -1.28.

n	$T_{WL}(.)$	SIGNIFICANCE LEVEL	
		SL = 0.05	SL = 0.10
20	C	0.016	0.062
	A	0	0.000
50	C	0.023	0.078
	A	0.003	0.025
100	C	0.034	0.084
	A	0.015	0.047
150	C	0.045	0.087
	A	0.020	0.060
200	C	0.043	0.103
	A	0.020	0.076

Results from 1000 trials.

TABLE 2.6.9 Power of $T_{LW}(.)$ and $T_{WL}(.)$.

(significance levels in parenthesis)

n	Tests	$T_{LW}(.)$	$T_{WL}(.)$
30	C	0.262 (0.002)	0.863 (0.188)
	A	0.187 (0.001)	0.811 (0.141)
40	C	0.415 (0.005)	0.928 (0.194)
	A	0.368 (0.004)	0.899 (0.151)
50	C	0.777 (0.043)	0.738 (0.034)
	A	0.757 (0.042)	0.751 (0.036)
80	C	0.940 (0.030)	0.996 (0.213)
	A	0.945 (0.035)	0.995 (0.186)

Results from 1000 trials.

FIGURE 2.6.1 - NULL DISTRIBUTION OF $T_{LW}(\cdot)$ - LOGNORMAL vs WEIBULL, 1000 trials, (see Appendix B).

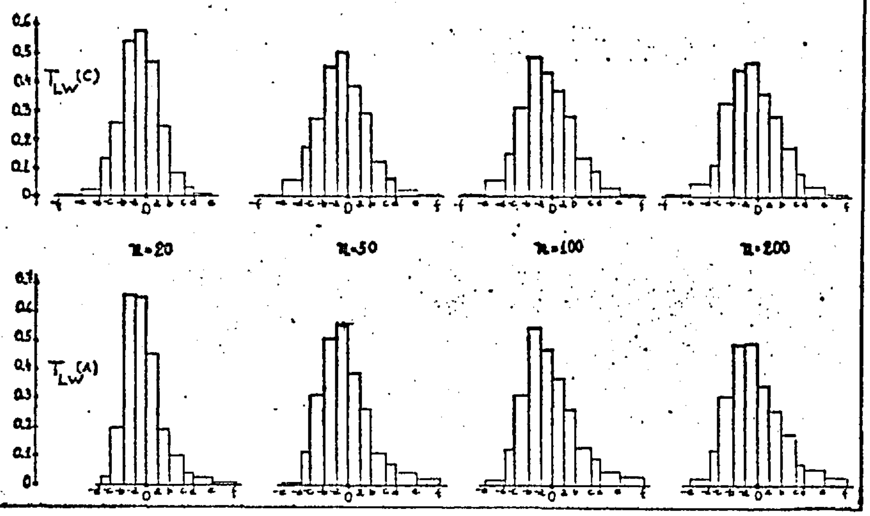
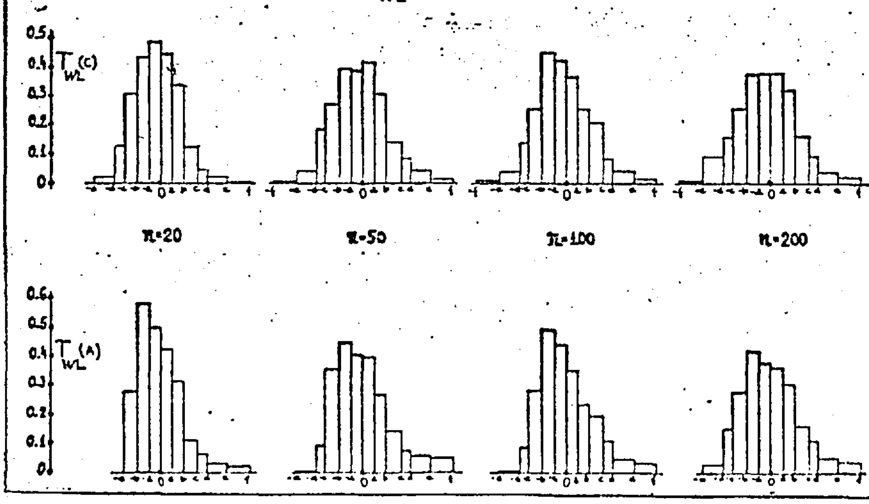


FIGURE 2.6.2 - NULL DISTRIBUTION OF $T_{WL}(\cdot)$ - WEIBULL vs LOGNORMAL, 1000 trials, (see Appendix B).



2.7 Tests for the gamma and Weibull distribution

A Test statistics and their distribution

Here again the methods of Cox and of Atkinson are used to derive tests involving the gamma and Weibull distribution. Suppose the null hypothesis H_G is that the distribution is gamma and the alternative H_W that it is Weibull, that is $H_G : f_G(\underset{\sim}{y};\, \gamma_1,\gamma_2)$ against $H_W : f_W(\underset{\sim}{y};\, \beta_1,\beta_2)$; see Section 1.2. The expectations of the log likelihoods functions in relation to the null gamma distribution yield

$$E_G\{\ell_G(\gamma_1,\gamma_2;\underset{\sim}{y})\} = -n \log \Gamma(\gamma_2) + n\, \gamma_2 \log\frac{\gamma_2}{\gamma_1} + (\gamma_2-1)\, n\left\{\psi(\gamma_2)-\log\frac{\gamma_2}{\gamma_1}\right\} - n\, \gamma_2\,, \qquad (2.7.1)$$

$$E_G\{\ell_W(\beta_1,\beta_2,\underset{\sim}{y})\} = n \log \beta_2 - n\beta_2 \log \beta_1 + (\beta_2-1)\, n\left\{\psi(\gamma_2)-\log \frac{\gamma_2}{\gamma_1}\right\} - \frac{m}{\beta_1^{\beta_2}} \left(\frac{\gamma_1}{\gamma_2}\right)^{\beta_2} \frac{\Gamma(\beta_2+\gamma_2)}{\Gamma(\gamma_2)}\,.$$

To find β_{1G} and β_{2G}, the probability limits under H_G of $\hat{\beta}_1$ and $\hat{\beta}_2$ respectively, the analogue to (2.6.2) is

$$\frac{\partial}{\partial(\beta_1,\beta_2)'} E_G\,\{\ell_W(\beta_{1G},\beta_{2G},\underset{\sim}{y})\} = \underset{\sim}{0}$$

whose unique solutions (β_{1G},β_{2G}) satisfy

$$\psi(\beta_{2G}+\gamma_2) - \frac{1}{\beta_{2G}} = \psi(\gamma_2)\,, \qquad \beta_{1G}^{\beta_{2G}} = \left(\frac{\gamma_1}{\gamma_2}\right)^{\beta_{2G}} \frac{\Gamma(\beta_{2G}+\gamma_2)}{\Gamma(\gamma_2)}\,. \qquad (2.7.2)$$

This shows that $\beta_{1G}^{\beta_{2G}}$ is the β_{2G}th moment of a gamma distribution. Writing $G \equiv (\hat{\gamma}_1,\hat{\gamma}_2)$ and by noticing that $\ell_G(\hat{\gamma}_1,\hat{\gamma}_2;\underset{\sim}{y}) - E_{\hat{G}}\{\ell_G(\gamma_1,\gamma_2,\underset{\sim}{y})\} = 0$ we then have

$$T^*_{GW}(C) = E_{\hat{G}}\{\ell_W(\beta_{1G},\beta_{2G};\underset{\sim}{y})\} - \ell_W(\hat{\beta}_1,\hat{\beta}_2;\underset{\sim}{y})$$

$$= n\left[\log \frac{\beta_{2\hat{G}}}{\hat{\beta}_2} - \left\{\beta_{2\hat{G}}\log \beta_{1\hat{G}} - \hat{\beta}_2 \log \hat{\beta}_1\right\} + \left\{\beta_{2\hat{G}}-\hat{\beta}_2\right\}\left\{\psi(\hat{\gamma}_2)-\log \frac{\hat{\gamma}_2}{\hat{\gamma}_1}\right\}\right] , \tag{2.7.3}$$

$$T^*_{GW}(A) = E_{\hat{G}}\left\{\ell_W\left(\beta_{1G},\beta_{2G};\underset{\sim}{y}\right)\right\} - \ell_W\left(\beta_{1\hat{G}},\beta_{2\hat{G}};\underset{\sim}{y}\right) = \frac{\sum_{i=1}^{n} y_i^{\beta_{2\hat{G}}}}{\beta_{1\hat{G}}^{\beta_{2\hat{G}}}} - n . \tag{2.7.4}$$

For the asymptotic variance $V_G\{T^*_{GW}\}$ of these tests we first evaluate

$$GW = n\left\{\psi'(\gamma_2)\left\{\gamma_2-\beta_{2G}\right\}^2 + \frac{\Gamma(2\beta_{2G}+\gamma_2)\Gamma(\gamma_2)}{\{\Gamma(\beta_{2G}+\gamma_2)\}^2} + 2\left\{\gamma_2-\beta_{2G}\right\}\left\{\psi(\beta_{2G}+\gamma_2)-\psi(\gamma_2)\right\} - \gamma_2-1\right\}, \tag{2.7.5}$$

$$C_{1G} = 0, \quad C_{2G} = n\left[\left\{\gamma_2-\beta_{2G}\right\}\psi'(\gamma_2) + \psi(\beta_{2G}+\gamma_2) - \psi(\gamma_2) - 1\right] .$$

and with the information matrix in (1.2.15) we obtain

$$V_G\{T^*_{GW}\} = GW - \underset{\sim}{C}'_G\, I^{-1}(\gamma_1,\gamma_2)\underset{\sim}{C}_G = n\left[\frac{\Gamma(2\beta_{2G}+\gamma_2)\Gamma(\gamma_2)}{[\Gamma(\beta_{2G}+\gamma_2)]^2}\right.$$

$$\left. + \frac{1}{\{\gamma_2\,\psi'(\gamma_2)-1\}\beta^2_{2G}}\left\{3\beta^2_{2G}-\gamma_2-\beta^4_{2G}\,\psi'(\gamma_2)-\gamma_2\,\psi'(\gamma_2)\beta^2_{2G}\right\}\right] . \tag{2.7.6}$$

Now, suppose H_G and H_W changes roles so that the null distribution is Weibull and the alternative is gamma. The expectation of the log likelihood functions in relation to the null Weibull distribution yields

$$E_W\{\ell_W(\beta_1,\beta_2;\underset{\sim}{y})\} = n \log \beta_2 - n\beta_2 \log \beta_1 + (\beta_2-1)\, n\left\{\frac{\psi(1)}{\beta_2} + \log \beta_1\right\} - n \,, \tag{2.7.7}$$

$$E_W\{\ell_G(\gamma_1,\gamma_2;\underset{\sim}{y})\} = -n \log \Gamma(\gamma_2)+n\, \gamma_2 \log \frac{\gamma_2}{\gamma_1} + (\gamma_2-1)n\left\{\frac{\psi(1)}{\beta_2}+\log \beta_1\right\}$$
$$- \frac{\gamma_2}{\gamma_1}\, n\, \beta_1\, \Gamma\left(1 + \frac{1}{\beta_2}\right) .$$

To find γ_{1W} and γ_{2W}, the probability limits under H_W of $\hat{\gamma}_1$ and $\hat{\gamma}_2$ respectively, the analogue to (2.6.2) is

$$\frac{\partial}{\partial(\gamma_1\gamma_2)'} E_W \left\{\ell_G(\gamma_{1W},\gamma_{2W},\underset{\sim}{y})\right\} = \underset{\sim}{0}$$

whose unique solution satisfy

$$\gamma_{1W} = \beta_1\, \Gamma\left(1 + \frac{1}{\beta_2}\right), \quad \log \gamma_{2W}-\psi(\gamma_{2W}) = \log \Gamma\left(1 + \frac{1}{\beta_2}\right) - \frac{\psi(1)}{\beta_2} . \tag{2.7.8}$$

This shows that $\hat{\gamma}_1$ converges to the mean of a Weibull distribution, and the righ hand side of the equation for γ_{2W} is the logarithm of the ratio of the arithmetic mean to the geometric mean of the Weibull distribution. Writing $\hat{W} \equiv (\hat{\beta}_1,\hat{\beta}_2)$, we then have

$$T^*_{WG}(C) = \ell_W(\hat{\beta}_1,\hat{\beta}_2;\underset{\sim}{y}) - \ell_G(\hat{\gamma}_1,\hat{\gamma}_2;\underset{\sim}{y}) - E_{\hat{W}}\{\ell_W(\beta_1,\beta_2;\underset{\sim}{y}) - \ell_G(\gamma_{1W},\gamma_{2W};\underset{\sim}{y})\}$$
$$= n\left[\left\{\gamma_{2\hat{W}}\left\{\psi(\gamma_{2\hat{W}})-1\right\} - \log \Gamma(\gamma_{2\hat{W}}) - \hat{\beta}_2\left\{\psi(\gamma_{2\hat{W}}) - \log \frac{\gamma_{2\hat{W}}}{\gamma_{1\hat{W}}}\right\}\right)\right.$$
$$\left. - \left\{\hat{\gamma}_2\left\{\psi(\hat{\gamma}_2)-1\right\} - \log \Gamma(\hat{\gamma}_2) - \hat{\beta}_2\left\{\psi(\hat{\gamma}_2) - \log \frac{\hat{\gamma}_2}{\hat{\gamma}_1}\right\}\right]\right. , \tag{2.7.9}$$

$$T^*_{WG}(A) = \ell_W(\hat\beta_1,\hat\beta_2;\underset{\sim}{y}) - \ell_G(\gamma_{1\hat W},\gamma_{2\hat W};\underset{\sim}{y}) - E_{\hat W}\left\{\ell_W(\beta_1,\beta_2;\underset{\sim}{y}) - \ell_G(\gamma_{1W},\gamma_{2W};\underset{\sim}{y})\right\}$$

$$= n\left[\left\{\hat\beta_2-\gamma_{2\hat W}\right\}\left\{\psi(\hat\gamma_2)-\log\frac{\hat\gamma_2}{\hat\gamma_1} - \psi(\gamma_{2\hat W}) + \log\frac{\gamma_{2\hat W}}{\gamma_{1\hat W}}\right\} + (\hat\gamma_2-\gamma_{2\hat W})\right]$$

(2.7.10)

For the asymptotic variance $V_W\{T^*_{WG}\}$ of these tests we have

$$WG = n\left[\left(\frac{\beta_2-\gamma_{2W}}{\beta_2}\right)^2 \psi'(1) + \gamma^2_{2W}\frac{\Gamma\left(1+\frac{2}{\beta_2}\right)}{\left\{\Gamma\left(1+\frac{1}{\beta_2}\right)\right\}^2} - \gamma^2_{2W} - 1 + 2\left\{\gamma_{2W}-\frac{\gamma_{2W}}{\beta_2}\right\}\left\{\psi\left(1+\frac{1}{\beta_2}\right)-\psi(1)\right\}\right],$$

(2.7.11)

$$C_{1W} = 0, \quad C_{2W} = \frac{n}{\beta_2}\left[1 - \frac{\gamma_{2W}}{\beta_2}\left\{\psi\left(1+\frac{1}{\beta_2}\right) - \psi(1)\right\}\right],$$

and with the information matrix in (1.2.11) we obtain

$$V_W\left\{T^*_{WG}\right\} = WG - \underset{\sim}{C}'_W\, I^{-1}(\beta_1,\beta_2)\underset{\sim}{C}_W = n\left[\left(\frac{\beta_2-\gamma_{2W}}{\beta_2}\right)^2\psi'(1) + \gamma^2_{2W}\frac{\Gamma\left(1+\frac{2}{\beta_2}\right)}{\left\{\Gamma\left(1+\frac{1}{\beta_2}\right)\right\}^2} - \gamma^2_{2W} - 1\right.$$

$$\left. + 2\left\{\gamma_{2W}-\frac{\gamma_{2W}}{\beta_2}\right\}\left\{\psi\left(1+\frac{1}{\beta_2}\right)-\psi(1)\right\} - \frac{1}{\psi'(1)}\left\{1-\frac{\gamma_{2W}}{\beta_2}\left\{\psi\left(1+\frac{1}{\beta_2}\right)-\psi(1)\right\}\right\}^2\right].$$

(2.7.12)

Hence, for j = A,C the statistics

$$T_{GW}(j) = T^*_{GW}(j)[V_G\{T^*_{GW}\}]^{-\frac{1}{2}}, \quad T_{WG}(j) = T^*_{WG}(j)[V_W\{T^*_{WG}\}]^{-\frac{1}{2}} \qquad (2.7.13)$$

are asymptotically standard normally distributed under H_G and H_W, respectively.

Finally, there is an observation to be made. In the application of this section there is a parameter value in H_G and H_W which gives the same probability distribution of the data. For $\gamma_2=\beta_2=1$ we have under H_G, that $\beta_{2G}=1$, $\beta_{1G}=\gamma_1$, $T^*_{GW}(.) = 0$ and $V_G\{T^*_{GW}\} = 0$ and under H_W that $\gamma_{1W} = \beta_1$, $T^*_{WG}(.) = 0$ and $V_W\{T^*_{WG}\} = 0$, therefore the asymptotic theory is not applicable. For neighbouring parametric values, the value of n required for the asymptotic theory be reasonably applicable may be large. An attempt was made to study this point when performing the simulations.

B Empirical results

Now empirical results on the tests of this section is presented. Because of the complexity to calculate the tests of this section only a small simulation study was attempted.

Results on the null distribution of $T_{GW}(C)$ and $T_{GW}(A)$ and on the distribution of $T_{WG}(C)$ and $T_{WG}(A)$ under the alternative, that is the gamma distribution, was obtained as in Section 2.5. Here, also, from the comments about (1.2.24) it follows that the distribution of the test statistics depends only on γ_2. Random variates from a gamma distribution were obtained by the methods described in Section 2.5. For $\gamma_1 = 1$ and different values of γ_2, 100 trials for sample sizes n = 50, 100, 200 were obtained.

Similarly we obtained the results on the null distribution of $T_{WG}(C)$ and $T_{WG}(A)$ and on the distribution of $T_{GW}(C)$ and $T_{GW}(A)$ under the alternative, that is the Weibull distribution. Again from the comments on (1.2.24) it follows that the distribution of the test statistics depends only on β_2. Random variates from a Weibull distribution were obtained by the transformation $y_i = \beta_1(-\log u_i)^{1/\beta_2}$ where u_i are uniform (0,1) variates. For $\beta_1 = 1$ and different values of β_2, 100 trials for sample sizes n = 50, 100, 200 were obtained.

The approximations of Section 2.5 and the accuracy of the Newton iteration described there was also used for the tests of this Section.

The results are summarized in Tables 2.7.1 to 2.7.8. In view of the small scale it is emphasized that no general conclusion will be made apart from general observations.

It was pointed out in Section 2.5 and earlier in this section that when the distributions have a similar shape a large sample size is expected to be required for the asymptotic result to be adequate. Further, the power function is expected to be low. The choice of parameter values for the simulations was directed to investigate this point. For values of γ_2 and β_2 near 1 both density functions should have a similar shape. For a Weibull density function with β_2 reasonably greater than 3.6, there is no gamma density function which has a similar shape.

Only the results for parameter values near 1 are presented in Tables 2.7.1 to 2.7.4. For values less than 0.8 and greater than 1.2 the adequacy of the asymptotic results were increased.

Results of Tables 2.7.1 and 2.7.2 do not show much difference between the A and C statistics. The results for the sample mean generally agree with Section 2.3.

For the parameter values in Tables 2.7.3 and 2.7.4, $T_{GW}(A)$ and $T_{WG}(A)$ seem to be consistent, although it does not seem feasible to investigate consistency analytically. In Table 2.7.4 the large value for the kurtosis of $T_{WG}(C)$ at $\gamma_2 = 0.8$ suggests that for n = 50 the asymptotic result is not adequate.

The power of $T_{GW}(.)$ in Table 2.7.5 agrees with the comment about the shape of the densities. For β_2 near 1 the power is low as should be expected. Table 2.7.6 also shows a low power for γ_2 near 1. The further low powers in Table 2.7.6 also agree with the comments about shape, since it is always possible to approximate the true gamma distribution by a Weibull distribution.

Comparison of power between A and C could be made using the argument of nearest alternative as in Section 2.5. This was not attempted here because the complexity of the equations and also because of the small scale of the simulations. Figures are also not provided.

The simulations of this section show that the results seem adequate for samples of size greater than about 100 even for parameter values as close to 1 as 0.8 and 1.2.

TABLE 2.7.1 Null distribution of $T_{GW}(C)$ and $T_{GW}(A)$.

n	$T_{GW}(.)$	$\mu_1\{T_{GW}(.)/H_G\}$		$\mu_2\{T_{GW}(.)/H_G\}$		$\gamma_1\{T_{GW}(.)/H_G\}$		$\beta_2\{T_{GW}(.)/H_G\}$	
		γ_2=0.8	γ_2=1.2	γ_2=0.8	γ_2=1.2	γ_2=0.8	γ_2=1.2	γ_2=0.8	γ_2=1.2
50	C	-0.278	-0.215	1.345	0.758	-0.634	-0.138	3.054	2.895
	A	-0.088	-0.054	1.126	0.773	-0.185	0.435	3.028	3.029
100	C	-0.077	-0.225	0.910	0.883	-0.351	-0.669	2.742	3.772
	A	0.012	-0.065	0.938	0.717	-0.163	0.044	2.735	3.082
200	C	0.046	-0.346	1.081	1.171	-0.429	-0.217	2.760	2.757
	A	0.091	-0.221	1.072	1.038	-0.243	0.327	2.776	2.979

Results from 100 trials.

TABLE 2.7.2 Null distribution of $T_{WG}(C)$ and $T_{WG}(A)$.

n	$T_{WG}(.)$	$\mu_1\{T_{WG}(.)/H_W\}$		$\mu_2\{T_{WG}(.)/H_W\}$		$\gamma_1\{T_{WG}(.)/H_W\}$		$\beta_2\{T_{WG}(.)/H_W\}$	
		β_2=0.8	β_2=1.2	β_2=0.8	β_2=1.2	β_2=0.8	β_2=1.2	β_2=0.8	β_2=1.2
50	C	-0.157	-0.056	0.777	0.814	-0.199	-0.422	2.872	2.935
	A	-0.048	0.104	0.759	0.796	0.244	0.149	2.909	2.862
100	C	-0.078	-0.096	0.845	1.070	0.507	-1.098	3.171	4.811
	A	-0.051	0.005	0.844	0.900	0.631	0.400	3.248	3.081
200	C	-0.063	-0.117	1.203	1.004	-0.510	-0.807	2.977	3.351
	A	0.045	-0.049	1.185	0.910	-0.421	-0.474	2.797	2.787

Results from 100 trials.

TABLE 2.7.3 Distribution of $T_{GW}(C)$ and $T_{GW}(A)$ under alternative H_W.

n	$T_{GW}(.)$	$\mu_1\{T_{GW}(.)/H_W\}$		$\mu_2\{T_{GW}(.)/H_W\}$		$\gamma_1\{T_{GW}(.)/H_W\}$		$\beta_2\{T_{GW}(.)/H_W\}$	
		β_2=0.8	β_2=1.2	β_2=0.8	β_2=1.2	β_2=0.8	β_2=1.2	β_2=0.8	β_2=1.2
50	C	-0.485	-0.553	0.915	1.008	-0.335	-0.555	2.548	3.070
	A	-0.397	-0.375	0.976	0.780	-0.059	0.130	2.637	2.747
100	C	-0.733	-0.611	1.127	0.963	-0.892	0.186	3.786	2.898
	A	-0.693	-0.486	1.047	0.860	-0.819	0.775	3.726	3.875
200	C	-1.002	-0.777	1.343	0.936	0.373	0.408	2.685	2.686
	A	-1.966	-0.675	1.284	0.816	0.459	0.727	2.837	3.216

Results from 100 trials.

TABLE 2.7.4 Distribution of $T_{WG}(C)$ and $T_{WG}(A)$ under alternative H_G.

n	$T_{WG}(.)$	$\mu_1(T_{WG}(.)/H_G\}$		$\mu_2\{T_{WG}(.)/H_G\}$		$\gamma_1\{T_{WG}(.)/H_G\}$		$\beta_2\{T_{WG}(.)/H_G)$	
		γ_2=0.8	γ_2=1.2	γ_2=0.8	γ_2=1.2	γ_2=0.8	γ_2=1.2	γ_2=0.8	γ_2=1.2
50	C	-0.494	-0.359	3.280	1.040	-4.699	-0.997	36.160	3.981
	A	-0.174	-0.194	1.072	0.765	0.008	-0.190	2.562	2.874
100	C	-0.537	-0.386	1.065	0.903	-0.510	-0.473	3.591	3.049
	A	-0.407	-0.265	0.880	0.769	-0.018	0.069	2.655	2.949
200	C	-0.774	-0.392	1.117	1.334	-0.082	-0.642	2.763	3.406
	A	-0.715	-0.261	1.012	1.113	0.116	-0.045	2.495	2.716

Results from 100 trials.

TABLE 2.7.5 Null: Gamma; Alternative: Weibull. Tests: $T_{GW}(C)$, $T_{GW}(A)$ Power at t = -1.64; t = -1.28.

n	$T_{GW}(.)$	POWER FUNCTION											
		SL = 0.05						SL = 0.10					
		β_2=0.6	β_2=0.8	β_2=1.2	β_2=2.0	β_2=3.6	β_2=5.0	β_2=0.6	β_2=0.8	β_2=1.2	β_2=2.0	β_2=3.6	β_2=5.0
50	C	0.340	0.120	0.130	0.350	0.580	0.670	0.420	0.240	0.220	0.470	0.720	0.800
	A	0.330	0.120	0.080	0.220	0.400	0.430	0.420	0.200	0.140	0.400	0.610	0.680
100	C	0.460	0.170	0.130	0.620	0.800	0.890	0.590	0.260	0.260	0.710	0.910	0.960
	A	0.460	0.170	0.110	0.530	0.730	0.820	0.590	0.250	0.170	0.680	0.860	0.930
200	C	0.730	0.340	0.180	0.830	0.990	1.000	0.870	0.440	0.340	0.870	1.000	1.000
	A	0.720	0.320	0.120	0.760	0.960	1.000	0.870	0.440	0.260	0.860	0.990	1.000

Results from 100 trials.

TABLE 2.7.6 Null: Weibull; Alternative: gamma. Tests: $T_{WG}(C)$; $T_{WG}(A)$. Power at t = -1.64; t = -1.28.

n	$T_{WG}(.)$	POWER FUNCTION											
		SL = 0.05						SL = 0.10					
		γ_2=0.6	γ_2=0.8	γ_2=1.2	γ_2=2.0	γ_2=5.0	γ_2=10.0	γ_2=0.6	γ_2=0.8	γ_2=1.2	γ_2=2.0	γ_2=5.0	γ_2=10.0
50	C	0.180	0.170	0.100	0.200	0.320	0.490	0.270	0.230	0.160	0.310	0.500	0.690
	A	0.160	0.100	0.060	0.110	0.190	0.280	0.260	0.180	0.120	0.240	0.390	0.550
100	C	0.260	0.110	0.120	0.290	0.600	0.820	0.400	0.250	0.180	0.420	0.770	0.920
	A	0.260	0.080	0.070	0.220	0.520	0.760	0.380	0.190	0.140	0.370	0.740	0.860
200	C	0.410	0.220	0.140	0.520	0.930	0.980	0.550	0.340	0.210	0.640	0.960	1.000
	A	0.400	0.200	0.100	0.450	0.850	0.980	0.540	0.310	0.180	0.610	0.960	0.990

Results from 100 trials.

TABLE 2.7.7 Null: gamma; Alternative: Weibull. Tests: $T_{GW}(C)$, $T_{GW}(A)$. One-side significance level at t = -1.64; t = -1.28.

n	$T_{GW}(.)$	SIGNIFICANCE LEVELS											
		SL = 0.05						SL = 0.10					
		$\gamma_2=0.6$	$\gamma_2=0.8$	$\gamma_2=1.2$	$\gamma_2=2.0$	$\gamma_2=5.0$	$\gamma_2=10.0$	$\gamma_2=0.6$	$\gamma_2=0.8$	$\gamma_2=1.2$	$\gamma_2=2.0$	$\gamma_2=5.0$	$\gamma_2=10.0$
50	C	0.080	0.110	0.080	0.040	0.060	0.050	0.120	0.200	0.120	0.080	0.130	0.120
	A	0.060	0.070	0.020	0.020	0.010	0.020	0.110	0.120	0.090	0.040	0.070	0.060
100	C	0.030	0.060	0.070	0.070	0.070	0.020	0.070	0.110	0.100	0.080	0.150	0.140
	A	0.060	0.040	0.020	0.040	0.060	0.020	0.060	0.090	0.060	0.080	0.100	0.070
200	C	0.050	0.060	0.140	0.060	0.070	0.080	0.100	0.100	0.190	0.110	0.140	0.140
	A	0.030	0.050	0.070	0.060	0.060	0.030	0.090	0.070	0.170	0.110	0.100	0.110

Results from 100 trials.

TABLE 2.7.8 Null: Weibull; Alternative: gamma. Tests: $T_{WG}(C)$, $T_{WA}(A)$. One-side significance level at t = -1.64; t = -1.28.

n	$T_{GW}(.)$	SIGNIFICANCE LEVELS											
		SL = 0.05						SL = 0.10					
		$\beta_2=0.6$	$\beta_2=0.8$	$\beta_2=1.2$	$\beta_2=2.0$	$\beta_2=3.6$	$\beta_2=5.0$	$\beta_2=0.6$	$\beta_2=0.8$	$\beta_2=1.2$	$\beta_2=2.0$	$\beta_2=3.6$	$\beta_2=5.0$
50	C	0.050	0.070	0.060	0.030	0.030	0.030	0.110	0.100	0.120	0.100	0.120	0.110
	A	0.040	0.030	0.010	0.010	0.010	0.010	0.110	0.070	0.090	0.080	0.060	0.040
100	C	0.020	0.030	0.080	0.070	0.050	0.050	0.080	0.100	0.120	0.120	0.140	0.130
	A	0.020	0.020	0.050	0.030	0.020	0.020	0.060	0.090	0.090	0.100	0.100	0.100
200	C	0.070	0.090	0.120	0.080	0.070	0.070	0.140	0.140	0.130	0.140	0.140	0.130
	A	0.050	0.080	0.070	0.050	0.050	0.050	0.130	0.140	0.130	0.130	0.120	0.120

Results from 500 trials.

2.8 Concluding remarks

From the results about consistency, Atkinson's test should be used only after verifying that under the alternative hypothesis of interest it leads to consistent tests. It may be difficult to check this, as was the case in Sections 2.5 and 2.7.

Under the null hypothesis the C statistics should be expected to be preferable on the basis of skewness and kurtosis. Therefore, from a practical point of view, the C test are generally recommended because corrections for lower order moments are considerably more easily obtained.

Comparison of power, although very approximate, does not suggest much difference on the power between the A and C statistics, except for the test of Section 2.4. However, because of the approach to normality, the significance levels attained by the C statistics agree more closely with the asymptotic values than those for the A statistics. Again, this also recommends the C statistics.

Chapter 3

NON-HOMOGENEOUS SAMPLE CASE

3.1 Introduction

In this chapter generalizations of the test statistics to deal with non-identically distributed and with dependent observations are considered. Because of the conclusions about the comparisons made in Chapter 2 only Cox's statistics will be discussed. First, test statistics are developed for the regression models of Section 1.2. The resulting statistics are generalizations of those of Chapter 2 and the empirical results can be thought of as calculated from the regression models under the average set of covariates, that is $\bar{\underset{\sim}{z}} = \underset{\sim}{0}$. It is found that the form of the test statistics does not depend on the covariates; therefore, asymptotically the test statistic is independent of the estimators of the regression coefficients. An illustration is given of the choice of the regression model for survival data of patients with brain tumours.

An attempt is made to answer the often asked question: What are the consequences of using one model when another is true? The efficiency of the estimators of the regression coefficients when using a false model in relation to the true model is investigated. The matrix of covariances for these efficiency comparisons is always of the form $(Z'Z)^{-1}$ times a constant. Thus asymptotically the design problem is separated from distributional assumptions.

Finally, it is shown that the results on the test statistics can be extended for separate families of hypothesis about Markov processes. Some problems are suggested.

3.2 Tests for the lognormal, the gamma and the exponential regression models

First, suppose the null hypothesis H_L is that the model is the lognormal regression model and the alternative H_G is that it is the gamma regression model, that is, H_L: $f_L(y_i, \alpha_1, \alpha_2, \underset{\sim}{a}')$ against H_G: $f_G(y_i, \gamma_1, \gamma_2, \underset{\sim}{g}')$; see Section 1.2. The expectations of the log likelihood functions in relation to the null lognormal distribution yield

$$E_L\left\{\ell_L(\alpha_1, \alpha_2, \underset{\sim}{a}; \underset{\sim}{y})\right\} = -\frac{n}{2}\log \alpha_2 - n\log\sqrt{2\pi} - n\,\alpha_1 - \frac{n}{2}\,, \qquad (3.2.1)$$

$$E_L\left\{\ell_G(\gamma_1, \gamma_2, \underset{\sim}{g}'; \underset{\sim}{y})\right\} = -n\log\Gamma(\gamma_2) + n\,\gamma_2\log\gamma_2 - n\,\gamma_2\gamma_1 + n(\gamma_2 - 1)\alpha_1$$
$$- \gamma_2\sum_{i=1}^{n}\exp\left\{\alpha_1 + \underset{\sim}{z}_i\underset{\sim}{a} + \frac{\alpha_2}{2} - \gamma_1 - \underset{\sim}{z}_i\underset{\sim}{g}\right\}.$$

To find γ_{1L}, γ_{2L} and $\underset{\sim}{g}_L$ the probability limits under H_L of $\hat{\gamma}_1, \hat{\gamma}_2$ and $\hat{\underset{\sim}{g}}$ respectively, recall Cox [1961, e.g.(32) and (33)], namely

$$E_L\left\{\frac{\partial \log f_G(\underset{\sim}{y}, \gamma_{1L}, \gamma_{2L}; \underset{\sim}{g}_L')}{\partial(\gamma_1, \gamma_2, \underset{\sim}{g}')'}\right\} = \frac{\partial}{\partial(\gamma_1, \gamma_2, \underset{\sim}{g}')'}E_L\left\{\ell_G(\gamma_{1L}, \gamma_{2L}, \underset{\sim}{g}_L'; \underset{\sim}{y})\right\} = \underset{\sim}{0}. \quad (3.2.2)$$

The derivatives in relation to γ_1, γ_2 and $\underset{\sim}{g}$ respectively, gives a system of equation whose unique solution is

$$\gamma_{1L} = \alpha_1 + \frac{\alpha_2}{2}\,, \quad \log\gamma_{2L} - \psi(\gamma_{2L}) = \frac{\alpha_2}{2}\,, \quad \underset{\sim}{g}_L' = \underset{\sim}{a}'\,. \qquad (3.2.3)$$

These results show that $\hat{\underset{\sim}{g}}$ is a consistent estimator of $\underset{\sim}{a}$, while γ_{1L} and γ_{2L} is similar to the single sample case. Writing $\hat{L} \equiv (\hat{\alpha}_1, \hat{\alpha}_2, \hat{\underset{\sim}{a}})$ and by noticing that $\ell_L(\hat{\alpha}_1, \hat{\alpha}_2, \hat{\underset{\sim}{a}}; \underset{\sim}{y}) - E_{\hat{L}}\{\ell_L(\alpha_1, \alpha_2, \underset{\sim}{a}', y)\} = 0$, we then have

$$T^*_{LG} = E_{\hat{L}}\left\{\ell_G(\gamma_{1L}, \gamma_{2L}, \underset{\sim}{g}_L'; \underset{\sim}{y})\right\} - \ell_G(\hat{\gamma}_1, \hat{\gamma}_2, \hat{\underset{\sim}{g}}; \underset{\sim}{y})$$
$$= n\left\{\log\Gamma(\hat{\gamma}_2) - \hat{\gamma}_2\,\psi(\hat{\gamma}_2) + \hat{\gamma}_2 - \log\Gamma(\gamma_{\hat{2L}}) + \gamma_{\hat{2L}}\,\psi(\gamma_{\hat{2L}}) - \gamma_{\hat{2L}}\right\}. \qquad (3.2.4)$$

Now, the asymptotic variance $V_L\{T^*_{LG}\}$ of this test is required. First we evaluate

$$LG = n\left\{\frac{1}{2} + \left(e^{\alpha_2} - \alpha_2 - 1\right) \gamma_{2L}^2 - \gamma_{2L} \quad \alpha_2\right\} ,$$

(3.2.5)

$$C_{1L} = 0, \quad C_{2L} = n\left\{\frac{\gamma_{2L}}{2} - \frac{1}{2\alpha_2}\right\} , \quad \underset{\sim}{C}_{3L} = \underset{\sim}{0} \quad ,$$

and recalling the information matrix in (1.2.9) and on writing $\underset{\sim}{C}_L = (C_{1L}, C_{2L}, \underset{\sim}{C}'_{3L})$, we have

$$V_L\{T^*_{LG}\} = LG - \underset{\sim}{C}'_L \; I^{-1}(\alpha_1,\alpha_2,\underset{\sim}{a}')\underset{\sim}{C}_L = n\; \gamma_{2L}^2\left(e^{\alpha_2} - 1 - \alpha_2 - \frac{\alpha_2^2}{2}\right) . \qquad (3.2.6)$$

Now, suppose that H_L and H_G change roles so that the null model is the gamma regression model and the alternative is the lognormal regression model. The expectations of the log likelihood functions in relation to the null gamma distribution yield

$$E_G\left\{\ell_G(\gamma_1,\gamma_2,\underset{\sim}{g}';\underset{\sim}{y})\right\} = -n \log \Gamma(\gamma_2) + n\,\gamma_2\,\psi(\gamma_2) - n\left\{\psi(\gamma_2) - \log\gamma_2 + \gamma_1\right\} - n\,\gamma_2,$$

$$E_G\left\{\ell_L(\alpha_1,\alpha_2,\underset{\sim}{a}';\underset{\sim}{y})\right\} = -\frac{n}{2}\log\alpha_2 - n\log\sqrt{2\pi} - n\left\{\psi(\gamma_2) - \log\gamma_2 + \gamma_1\right\} - \frac{n}{2}\,\psi'(\gamma_2)$$

(3.2.7)

$$-\frac{1}{2\alpha_2}\sum_{i=1}^{n}\left[\alpha_1 + \underset{\sim}{z}_i\underset{\sim}{a} - \left\{\psi(\gamma_2) - \log\gamma_2 + \gamma_1 + \underset{\sim}{z}_i\underset{\sim}{g}\right\}\right]^2 .$$

To find α_{1G}, α_{2G} and $\underset{\sim}{a}_G$, the probability limits under H_G of $\hat{\alpha}_1$, $\hat{\alpha}_2$ and $\hat{\underset{\sim}{a}}$ respectively, the analogue to (3.2.2) is

$$\frac{\partial}{\partial(\alpha_1,\alpha_2,\underset{\sim}{a}')}\, E_G\{\ell_L(\alpha_{1G},\alpha_{2G},\underset{\sim}{a}'_G;\underset{\sim}{y})\} = \underset{\sim}{0}$$

whose unique solution is

$$\alpha_{1G} = \gamma_1 + \psi(\gamma_2) - \log \gamma_2, \quad \alpha_{2G} = \psi'(\gamma_2), \quad \underset{\sim}{a}_G = \underset{\sim}{g} . \tag{3.2.8}$$

Here, $\hat{\underset{\sim}{a}}$ is seen to be a consistent estimator of $\underset{\sim}{g}$ and the result on α_{1G} and α_{2G} is similar to the single sample case. Writing $\hat{G} \equiv (\hat{\gamma}_1, \hat{\gamma}_2, \hat{g})$ and by noticing that $\ell_G(\hat{\gamma}_1, \hat{\gamma}_2, \underset{\sim}{g}, \underset{\sim}{y})\ E_{\hat{G}}\{\ell_G(\gamma_{1G}, \gamma_{2G}, \underset{\sim}{a}_G; \underset{\sim}{y})\} = 0$, we have

$$T^*_{GL} = E_{\hat{G}}\{\ell_L(\alpha_{1G}, \alpha_{2G}, \underset{\sim}{a}'_G; \underset{\sim}{y})\} - \ell_L(\hat{\alpha}_1, \hat{\alpha}_2, \hat{\underset{\sim}{a}}; \underset{\sim}{y}) = \frac{n}{2} \log \frac{\hat{\alpha}_2}{\alpha_{2\hat{G}}} . \tag{3.2.9}$$

Similarly, for the variance $V_G\{T^*_{GL}\}$, we first evaluate

$$GL = n \left\{ \gamma_2^2\ \psi'(\gamma_2) + \frac{\psi'''(\gamma_2)}{4\{\psi'(\gamma_2)\}^2} + \frac{1}{2} + \frac{\gamma_2 \psi''(\gamma_2)}{\psi'(\gamma_2)} - \gamma_2 \right\} , \tag{3.2.10}$$

$$C_{1G} = 0, \quad C_{2G} = n \left\{ \gamma_2\ \psi'(\gamma_2) - 1 + \frac{\psi''(\gamma_2)}{\psi'(\gamma_2)} \right\} , \quad \underset{\sim}{C}_{3G} = \underset{\sim}{0} ,$$

and recalling the information matrix in (1.2.17) and by denoting $\underset{\sim}{C}'_G = (C_{1G}, C_{2G}, \underset{\sim}{C}'_{3G})$, we have

$$V_G\{T^*_{GL}\} = GL - \underset{\sim}{C}'_G I^{-1}(\gamma_1, \gamma_2, \underset{\sim}{g}')\underset{\sim}{C}_G = n \left\{ \frac{\psi'''(\gamma_2)}{4\{\psi'(\gamma_2)\}^2} - \frac{\gamma_2\{\psi''(\gamma_2)\}^2}{4\{\psi'(\gamma_2)\}^2\{\gamma_2\psi'(\gamma_2)-1\}} + \frac{1}{2} \right\} . \tag{3.2.11}$$

Finally, the statistics T^*_{LG} and T^*_{GL} standardized by their variances, are asymptotically standard normally distributed under H_L and H_G, respectively.

Special case - exponential regression model

Now, the tests involving the lognormal and the exponential regression model are presented. It is useful to recall the relation

$$\hat{\alpha}_1 = \hat{\gamma}_1 + \psi(\hat{\gamma}_2) - \log \hat{\gamma}_2 = \gamma_{1\hat{L}} + \psi(\gamma_{2\hat{L}}) - \log \gamma_{2\hat{L}} = \alpha_{1\hat{G}} .$$

First, suppose one wants to test the null hypothesis H_L: $f(y_i;\alpha_1,\alpha_2,\underset{\sim}{a}')$ against the alternative hypothesis H_E: $f_E(y_i;\delta,\underset{\sim}{d}')$. The expressions (3.2.3), (3.2.4) and (3.2.6) become respectively

$$\delta_L = \alpha_1 + \frac{\alpha_2}{2} , \quad \underset{\sim}{d}_L = \underset{\sim}{a} , \tag{3.2.12}$$

$$T^*_{LE} = n\left\{\hat{\delta} - \hat{\alpha}_1 - \frac{\hat{\alpha}_2}{2}\right\} , \tag{3.2.13}$$

$$V_L\{T^*_{LE}\} = n\left\{e^{\alpha_2} - 1 - \alpha_2 - \frac{\alpha_2^2}{2}\right\} . \tag{3.2.14}$$

When H_E is the null hypothesis and H_L is the alternative, we have similarly that expressions (3.2.8), and (3.2.11) become

$$\alpha_{1E} = \delta + \psi(1) , \quad \alpha_{2E}' = \psi'(1) , \quad \underset{\sim}{a}_E = \underset{\sim}{d} , \tag{3.2.15}$$

$$T^*_{EL} = n\left[\hat{\alpha}_1 - \{\hat{\delta} + \psi(1)\} + \frac{1}{2}\log\frac{\hat{\alpha}_2}{\alpha_{2\hat{E}}}\right] , \quad \left(\hat{E} \equiv (\hat{\delta},\hat{\underset{\sim}{d}}')\right) , \tag{3.2.16}$$

$$V_E\{T^*_{EL}\} = 0.2834n . \tag{3.2.17}$$

Again, T^*_{LE} and T^*_{EL}, standardized by their variances, are asymptotically standard normally distributed under H_L and H_E, respectively.

3.3 Tests for the lognormal and the Weibull regression models

Here, the null hypothesis H_L is that the model is the lognormal regression model and the alternative H_W is that it is the Weibull regression model, that is H_L: $f_L(y_i;\alpha_1,\alpha_2,\underset{\sim}{a}')$ against H_W: $f_W(y_i;\beta_1,\beta_2,\underset{\sim}{b}')$; see Section 1.2. The expectations of the log likelihood functions in relation to the null lognormal distribution yield

$$E_L\{\ell_L(\alpha_1,\alpha_2,\underset{\sim}{a}';\underset{\sim}{y})\} = -\frac{n}{2}\log\alpha_2 - n\log\sqrt{2\pi} - n\alpha_1 - \frac{n}{2}\,, \tag{3.3.1}$$

$$E_L\{\ell_W(\beta_1,\beta_2,\underset{\sim}{b}';\underset{\sim}{y})\} = n\log\beta_2 - n\beta_1\beta_2 + n(\beta_2-1)\alpha_1 - \sum_{i=1}^{n}\left[\exp\left\{\alpha_1+\underset{\sim}{z}_i\underset{\sim}{a} + \frac{\beta_2\alpha_2}{2} - \beta_1 - \underset{\sim}{z}_i\underset{\sim}{b}\right\}\right]^{\beta_2}\,.$$

To find β_{1L}, β_{2L} and $\underset{\sim}{b}_L$, the probability limits under H_L of $\hat{\beta}_1$, $\hat{\beta}_2$ and $\hat{\underset{\sim}{b}}$, respectively, the analogue to (3.2.2) is

$$\frac{\partial}{\partial(\beta_1,\beta_2,\underset{\sim}{b}')}E\{\ell_W(\beta_{1L},\beta_{2L},\underset{\sim}{b}_L';\underset{\sim}{y})\} = \underset{\sim}{0}\,,$$

whose unique solution is

$$\beta_{1L} = \alpha_1 + \frac{\sqrt{\alpha_2}}{2}\,,\quad \beta_{2L} = \frac{1}{\sqrt{\alpha_2}}\,,\quad \underset{\sim}{b}_L = \underset{\sim}{a}\,. \tag{3.3.2}$$

These show that $\hat{\underset{\sim}{b}}$ is a consistent estimator of $\underset{\sim}{a}$ and β_{1L} and β_{2L} are similar to the single sample case. Writing $\hat{L} \equiv (\hat{\alpha}_1,\hat{\alpha}_2,\hat{\underset{\sim}{a}})$, we have

$$\begin{aligned} T^*_{LW} &= E_{\hat{L}}\{\ell_W(\beta_{1L},\beta_{2L},\underset{\sim}{b}_L';\underset{\sim}{y})\} - \ell_W(\hat{\beta}_1,\hat{\beta}_2,\hat{\underset{\sim}{b}}';\underset{\sim}{y}) \\ &= n\left\{\hat{\beta}_2\hat{\beta}_1 - \beta_{2\hat{L}}\,\beta_{1\hat{L}} - \log\frac{\hat{\beta}_2}{\beta_{2\hat{L}}} - \hat{\alpha}_1\left(\hat{\beta}_2-\beta_{2\hat{L}}\right)\right\}. \end{aligned} \tag{3.3.3}$$

For the variance $V_L\{T^*_{LW}\}$ we first have similarly to (2.6.6) that

$$LW = 0.218281n\,,\quad (C_{1L},C_{2L},\underset{\sim}{C}_{3L}') = \underset{\sim}{0}' \tag{3.3.4}$$

and similarly to (2.6.7)

$$V_L\{T^*_{LW}\} = 0.218281n\,. \tag{3.3.5}$$

Now, H_L and H_W changes roles so that the null model is the Weibull regression model and the alternative is the lognormal regression model. The expectations of the log likelihood functions in relation to the null Weibull distribution yield

$$E_W\{\ell_W(\beta_1,\beta_2,\underset{\sim}{b}';\underset{\sim}{y})\} = n \log \beta_2 - n\beta_2\beta_1 + n(\beta_2-1)\left\{\beta_1 + \frac{\psi(1)}{\beta_2}\right\} - n \;,$$

(3.3.6)

$$E_W\{\ell_L(\alpha_1,\alpha_2,\underset{\sim}{a}';\underset{\sim}{y})\} = -\frac{n}{2}\log \alpha_2 - n \log\sqrt{2\pi} - n\left\{\beta_1 + \frac{\psi(1)}{\beta_2}\right\} - \frac{n\psi'(1)}{2\alpha_2\beta_2^2}$$

$$- \frac{1}{2\alpha_2}\sum_{i=1}^{n}\left\{\alpha_1 + \underset{\sim}{z}_i\underset{\sim}{a} - \left(\beta_1 + \frac{\psi(1)}{\beta_2} + \underset{\sim}{z}_i\underset{\sim}{b}\right)\right\}^2 .$$

To find α_{1W}, α_{2W} and $\underset{\sim}{a}_W$, the probability limits under H_W of $\hat{\alpha}_1$, $\hat{\alpha}_2$ and $\hat{\underset{\sim}{a}}$ respectively, the analogue to (3.2.2) is

$$\frac{\partial}{\partial(\alpha_1,\alpha_2,\underset{\sim}{a}')'} E_W\{\ell_L(\alpha_{1W},\alpha_{2W},\underset{\sim}{a}_W';\underset{\sim}{y})\} = \underset{\sim}{0}$$

whose unique solution is

$$\alpha_{1W} = \beta_1 + \frac{\psi(1)}{\beta_2} \;, \quad \alpha_{2W} = \frac{\psi'(1)}{\beta_2^2} \;, \quad \underset{\sim}{a}_W = \underset{\sim}{b} \;. \qquad (3.3.7)$$

Again, $\hat{\underset{\sim}{a}}$ is a consistent estimator of $\underset{\sim}{b}$ and the result on α_{1W} and α_{2W} is similar to the single sample case. Writing $\hat{W} \equiv (\hat{\beta}_1,\hat{\beta}_2,\hat{\underset{\sim}{b}}')$ we then have

$$T^*_{WL} = \ell_W(\hat{\beta}_1,\hat{\beta}_2,\hat{\underset{\sim}{b}}';\underset{\sim}{y}) - \ell_L(\hat{\alpha}_1,\hat{\alpha}_2,\hat{\underset{\sim}{a}}';y) - E_{\hat{W}}\{\ell_W(\beta_1,\beta_2,\underset{\sim}{b}';\underset{\sim}{y}) - \ell_L(\alpha_{1W},\alpha_{2W},\underset{\sim}{a}_W';\underset{\sim}{y})\}$$

$$= n\left\{\hat{\beta}_2\left(\hat{\alpha}_1 - \alpha_{1\hat{W}}\right) + \frac{1}{2}\log\frac{\hat{\alpha}_2}{\alpha_{2\hat{W}}}\right\} . \qquad (3.3.8)$$

For the variance $V_W\{T^*_{WL}\}$ we have similarly to (2.6.12) that

$$WL = 0.2834n \ ; \quad (C_{1W}, C_{2W}, \underset{\sim}{C}'_W) = \underset{\sim}{0} \ . \tag{3.3.8}$$

and similarly to (2.6.13)

$$V_W\{T^*_{WL}\} = 0.2834n \ . \tag{3.3.10}$$

Finally, the statistics T^*_{LW} and T^*_{WL} standardized by their variances are asymptotically standard normally distributed under H_L and H_W, respectively.

3.4 Tests for the gamma and the Weibull regression models

Suppose the null hypothesis H_G is that the model is the gamma regression model and the alternative that it is the Weibull regression model, that is H_G: $f(y_i;\gamma_1,\gamma_2,\underset{\sim}{g}')$ against $f(y_i;\beta_1,\beta_2,\underset{\sim}{b}')$; see Section 1.2. The expectations of the log likelihood functions in relation to the null gamma distributions yield

$$E_G\{\ell_G(\gamma_1,\gamma_2,\underset{\sim}{g}';\underset{\sim}{y})\} = -n \log \Gamma(\gamma_2) + n\,\gamma_2\,\psi(\gamma_2) - n\{\psi(\gamma_2)-\log\gamma_2+\gamma_1\} - n\,\gamma_2,$$

$$E_G\{\ell_W(\beta_1,\beta_2,\underset{\sim}{b}';\underset{\sim}{y})\} = n \log \beta_2 - n\,\beta_2\beta_1 + (\beta_2-1)\,n\{\psi(\gamma_2)-\log\gamma_2+\gamma_1\}$$

$$\frac{\Gamma(\beta_2+\gamma_2)}{\Gamma(\gamma_2)\gamma_2^{\beta_2}} \sum_{i=1}^{n} \left[\exp\{\gamma_1+\underset{\sim}{z}_i\underset{\sim}{g}-\beta_1-\underset{\sim}{z}_i\underset{\sim}{b}\}\right]^{\beta_2} \ . \tag{3.4.1}$$

To find β_{1G}, β_{2G} and $\underset{\sim}{b}_G$, the probability limits under H_G of $\hat{\beta}_1$, $\hat{\beta}_2$ and $\hat{\underset{\sim}{b}}$ respectively, the analogue to (3.2.2) is

$$\frac{\partial}{\partial(\beta_1,\beta_2,\underset{\sim}{b}')'}\, E_G\left\{\ell_W(\beta_{1G},\beta_{2G},\underset{\sim}{b}'_G;\underset{\sim}{y})\right\} = \underset{\sim}{0}$$

whose unique solution is given by

$$\beta_{1G} = \gamma_1 - \log \gamma_2 - \frac{1}{\beta_{2G}} \log \frac{\Gamma(\beta_{2G}+\gamma_2)}{\Gamma(\gamma_2)} , \quad \psi(\beta_{2G}+\gamma_2) - \frac{1}{\beta_{2G}} = \psi(\gamma_2) , \quad \underset{\sim}{b}_G \equiv \underset{\sim}{g} . \tag{3.4.2}$$

Again, $\hat{\underset{\sim}{b}}$ is a consistent estimator of $\underset{\sim}{g}$ and the results on β_{1G} and β_{2G} are similar to the single sample case.

Writing $\hat{G} \equiv (\hat{\gamma}_1, \hat{\gamma}_2, \hat{\underset{\sim}{g}})$ we have

$$T^*_{GW} = E_{\hat{G}}\left\{\ell_W(\beta_{1G}, \beta_{2G}, \underset{\sim}{b}'_G; \underset{\sim}{y})\right\} - \ell_W(\hat{\beta}_1, \hat{\beta}_2, \underset{\sim}{b}'; \underset{\sim}{y})$$

$$= n\left[\log \frac{\beta_{2\hat{G}}}{\hat{\beta}_2} - \left(\beta_{2\hat{G}}\beta_{1\hat{G}} - \hat{\beta}_2\hat{\beta}_1\right) + \left(\beta_{2\hat{G}} - \hat{\beta}_2\right)\left\{\psi(\hat{\gamma}_2) - \log \hat{\gamma}_2 - \hat{\gamma}_1\right\}\right] . \tag{3.4.3}$$

For the variance $V_G\{T^*_{GW}\}$ we have similarly to (2.7.5) that

$$GW = n\left\{\psi'(\gamma_2)(\gamma_2-\beta_{2G})^2 + \frac{\Gamma(2\beta_{2G}+\gamma_2)\Gamma(\gamma_2)}{\{\Gamma(\beta_{2G}+\gamma_2)\}^2} + 2(\gamma_2-\beta_{2G})\{\psi(\beta_{2G}+\gamma_2)-\psi(\gamma_2)\} - \gamma_2 - 1\right\}, \tag{3.4.4}$$

$$C_{1G} = 0, \quad C_{2G} = n\left[(\gamma_2-\beta_{2G})\psi'(\gamma_2) + \psi(\beta_{2G}+\gamma_2) - \psi(\gamma_2) - 1\right] , \quad \underset{\sim}{C}_{3G} = \underset{\sim}{0} .$$

and recalling the information matrix in (1.2.17) and by denoting $\underset{\sim}{C}'_G = (C_{1G}, C_{2G}, \underset{\sim}{C}'_{3G})$ we have similarly to (2.7.6)

$$V_G\{T^*_{GW}\} = n\left[\frac{\Gamma(2\beta_{2G}+\gamma_2)\Gamma(\gamma_2)}{\{\Gamma(\beta_{2G})\}^2} + \frac{1}{\{\gamma_2\psi'(\gamma_2)-1\}\beta^2_{2G}}\left\{3\beta^2_{2G} - \gamma_2 - \beta^4_{2G}\,\psi'(\gamma_2) - \gamma_2\psi'(\gamma_2)\beta^2_{2G}\right\}\right] . \tag{3.4.5}$$

Now H_G and H_W changes roles so that the null model is the Weibull regression model and the alternative is the lognormal regression model.

The expectations of the log likelihood functions in relation to the null Weibull distributions yields

$$E_W\{\ell_W(\beta_1,\beta_2,\underset{\sim}{b}';\underset{\sim}{y})\} = n \log \beta_2 - n\,\beta_2\beta_1 + n(\beta_2-1)\left\{\beta_1 + \frac{\psi(1)}{\beta_2}\right\} - n,$$

$$E_W\{\ell_G(\gamma_1,\gamma_2,\underset{\sim}{g}';\underset{\sim}{y})\} = -n \log \Gamma(\gamma_2) + n\gamma_2 \log \gamma_2 - n\,\gamma_2\gamma_1 + (\gamma_2-1)\, n\left\{\beta_1 + \frac{\psi(1)}{\beta_2}\right\}$$

$$- \gamma_2\, \Gamma\left\{\frac{1}{\beta_2} + 1\right\} \sum_{i=1}^{n} \exp\left\{\beta_1 + \underset{\sim}{z}_i\underset{\sim}{b} - \gamma_1 - \underset{\sim}{z}_i\underset{\sim}{g}\right\} . \qquad (3.4.6)$$

To find γ_{1W}, γ_{2W} and $\underset{\sim}{g}_W$, the probability limits under H_W of $\hat{\gamma}_1$, $\hat{\gamma}_2$ and $\hat{\underset{\sim}{g}}$ respectively, the analogue to (3.2.2) is

$$\frac{\partial}{\partial(\gamma_1,\gamma_2,\underset{\sim}{g})'} E_W\left\{\ell_G(\gamma_{1W},\gamma_{2W},\underset{\sim}{g}_W';\underset{\sim}{y})\right\} = \underset{\sim}{0}$$

whose unique solution is given by

$$\gamma_{1W} = \beta_1 + \log \Gamma\left(\frac{1}{\beta_2} + 1\right), \quad \psi(\gamma_{2W}) - \log \gamma_{2W} = \frac{\psi(1)}{\beta_2} - \log \Gamma\left(\frac{1}{\beta_2} + 1\right), \quad \underset{\sim}{g}_W = \underset{\sim}{b} . \qquad (3.4.7)$$

Here also, $\hat{\underset{\sim}{g}}$ is a consistent estimator of $\underset{\sim}{b}$ and the results on γ_{1W} and γ_{2W} are similar to the single sample case.
Writing $\hat{W} \equiv (\hat{\beta}_1,\hat{\beta}_2,\hat{\underset{\sim}{b}})$ we have

$$T_{WG} = \ell_W(\hat{\beta}_1,\hat{\beta}_2,\hat{\underset{\sim}{b}}';\underset{\sim}{y}) - \ell_G(\hat{\gamma}_1,\hat{\gamma}_2,\hat{\underset{\sim}{g}}',\underset{\sim}{y}) - E_{\hat{W}}\{\ell_W(\beta_1,\beta_2,\underset{\sim}{b}';\underset{\sim}{y}) - \ell_G(\gamma_{1W},\gamma_{2W},\underset{\sim}{g}_W';\underset{\sim}{y})\}$$

$$= n\Big(\Big[\gamma_{2\hat{W}}\{\psi(\gamma_{2W})-1\} - \log \Gamma(\gamma_{2\hat{W}}) - \hat{\beta}_2\{\psi(\gamma_{2\hat{W}}) - \log \gamma_{2\hat{W}} + \gamma_{1\hat{W}}\}\Big]$$

$$- \Big[\hat{\gamma}_2\{\psi(\hat{\gamma}_2)-1\} - \log \Gamma(\hat{\gamma}_2) - \hat{\beta}_2\{\psi(\hat{\gamma}_2) - \log \hat{\gamma}_2 + \hat{\gamma}_1\}\Big]\Big) . \qquad (3.4.8)$$

For the variance $V_W\{T^*_{WL}\}$ we have similarly to (2.7.11) that

$$WG = n\left[\left(\frac{\beta_2-\gamma_{2W}}{\beta_2}\right)^2 \psi'(1) + \gamma_{2W}^2 \frac{\Gamma\left(\frac{2}{\beta_2}+1\right)}{\left\{\Gamma\left(\frac{1}{\beta_2}+1\right)\right\}^2} - \gamma_{2W}^2 - 1 + 2\left(\gamma_{2W}-\frac{\gamma_{2W}}{\beta_2}\right)\left\{\psi\left(1+\frac{1}{\beta_2}\right)-\psi(1)\right\}\right],$$

(3.4.9)

$$C_{1W} = 0, \quad C_{2W} = \frac{n}{\beta_2}\left[1-\frac{\gamma_{2W}}{\beta_2}\left\{\psi\left(\frac{1}{\beta_2}+1\right) - \psi(1)\right\}\right], \quad \underset{\sim}{C}_{3W} = \underset{\sim}{0} ,$$

and recalling the information matrix in (1.2.12) and by denoting $\underset{\sim}{C}'_W = (C_{1W}, C_{2W}, \underset{\sim}{C}'_{3W})$ we have similarly to (2.7.13)

$$V_W\{T^*_{WG}\} = n\left[\left(\frac{\beta_2-\gamma_{2W}}{\beta_2}\right)^2 \psi'(1) + \gamma_{2W}^2 \frac{\Gamma\left(1+\frac{2}{\beta_2}\right)}{\left\{\Gamma\left(1+\frac{1}{\beta_2}\right)\right\}^2} - \gamma_{2W}^2 - 1\right.$$
$$\left. + 2\left(\gamma_{2W}-\frac{\gamma_{2W}}{\beta_2}\right)\left\{\psi\left(1+\frac{1}{\beta_2}\right) - \psi(1)\right\} - \frac{1}{\psi'(1)}\left\{1-\frac{\gamma_{2W}}{\beta_2}\left\{\psi\left(1+\frac{1}{\beta_2}\right)-\psi(1)\right\}\right\}^2\right] .$$

(3.4.10)

Again, the statistics T^*_{GW} and T^*_{WG} standardized by their variances are asymptotically standard normally distributed under H_G and H_W respectively.

3.5 Example

An illustration of the previous results will now be given. Table 3.5.6 presents survival data on 93 malignant tumour patients as collected by the Brain Tumor Study Group at the M.D. Anderson Hospital and Tumor Institute, University of Texas. All patients received surgery and were randomized according to a chemotherapeutic agent (Mithramycin) and conventional care (Control) during the recovery period from surgery. The tumours were classified by their position in the brain. Other covariates recorded were, age, duration of symptoms (headache, personality change, motor deficit, etc.), sex, level of radiation. A further description is given by Walker, Gehan, Laventhal, Norrel & Mahaley (1969).

Corresponding to each patient a vector of covariates $\underset{\sim}{z} = (z_1, \ldots, z_{10})$ was defined, where z_1, z_2, z_3, z_4 and z_5 represent age, duration of symptoms, sex, treatment and radiation, respectively. The remaining z_6, z_7, z_8, z_9 and z_{10} are indicators of the position of the cancer cells with a one variate corresponding to each of frontal, temporal, parietal, occipital and deep BG/T.

For the choice of a suitable model the simplest models were first tried, that is, the exponential regression and the lognormal regression. For these it was found that $T_{LE} = -2.813$ which is significant at a level $\alpha = 0.0025$, which points to a departure from the lognormal regression in the direction of the exponential regression. Interchanging the roles of H_L and H_E, $T_{EL} = -2.909$ was found, which is significant at the level $\alpha = 0.0019$ and points to a departure from the exponential regression in the direction of the lognormal regression. This would mean that neither model fit the data well. To verify this point these models were tested against some alternative simple models.

First, departures from the exponential regression in the direction of the gamma and the Weibull regression were tested. For this, the asymptotic normal distribution of the maximum likelihood estimator of the shape parameter of the gamma and the Weibull models or equivalently the asymptotic χ^2 distribution of the maximum likelihood ratio were used. The results are summarized in Table 3.5.1 and show that assuming a Weibull or a Gamma model the null hypothesis of an exponential regression model is rejected.

TABLE 3.5.1 Testing for exponential regression

Alternative	M L E		Likelihood Ratio	
	Normal Deviate	Significance Level	$-2\log \lambda = \chi^2$	Significance Level
Gamma	3.982	0.000035	26.765	<0.00001
Weibull	5.084	<0.00001	31.267	<0.00001

Note the rough agreement of the square of the normal deviates in the first column with the χ^2 deviates in the third column. Also, note that the null hypothesis of exponential model was rejected more strongly by the Weibull test.

Following this the lognormal regression against the gamma and the Weibull regression was tested. It was found that $T_{LG} = -3.119$ which is significant at $\alpha = 0.0009$ and $T_{GL} = 1.016$ which is significant at $\alpha = 0.1539$. The first test rejected the lognormal in favour of the gamma model and the second suggested a reasonable agreement with the gamma model. For the Weibull regression the results were, $T_{LW} = -3.699$ with significance $\alpha = 0.00011$ and $T_{WL} = 0.137$ with $\alpha = 0.4443$, the former rejected the lognormal in favour of the Weibull and the latter suggested a good agreement with the Weibull model. Again, it can be seen that the lognormal regression was rejected more strongly when compared with the Weibull regression.

The tests between the gamma model and the Weibull model gave $T_{GW} = -2.436$ with $\alpha = 0.0073$ which points a departure from the gamma in the direction of the Weibull model. The converse $T_{WG} = 0.967$ with $\alpha = 0.166$ suggests a good agreement of the Weibull model with these data.

Finally, in view of the above results it is concluded that the Weibull model should be used for further analysis of the data. The result of these test statistics are summarized in Table 3.5.2.

TABLE 3.5.2 Results of the tests of separate families of hypothesis

Test	Observed		Estimates of Probability Limits		
	Normal Deviate	Significance Level	Constant Term	Shape	Regression Coefficients
T_{LE}	-2.813	0.00248	$\delta_L = 5.196$		$\underset{\sim}{d}_{\hat{L}} = \underset{\sim}{\hat{a}}$
T_{EL}	-2.909	0.00191	$\alpha_{1E} = 4.557$	$\alpha_{2E} = 1.645$	$\underset{\sim}{a}_{\hat{E}} = \underset{\sim}{\hat{d}}$
T_{LG}	-3.119	0.00090	$\gamma_{1L} = 5.196$	$\gamma_{2L} = 1.777$	$\underset{\sim}{g}_{\hat{L}} = \hat{a}$
T_{GL}	1.016	0.15386	$\alpha_{1G} = 4.890$	$\alpha_{2G} = 0.533$	$\underset{\sim}{a}_{\hat{G}} = \underset{\sim}{\hat{g}}$
T_{LW}	-3.699	0.00011	$\beta_{2L} = 5.281$	$\beta_{2L} = 1.277$	$\underset{\sim}{b}_{\hat{L}} = \underset{\sim}{\hat{a}}$
T_{WL}	0.137	0.44433	$\alpha_{1W} = 4.906$	$\alpha_{2W} = 0.570$	$\underset{\sim}{a}_{\hat{W}} = \underset{\sim}{\hat{b}}$
T_{GW}	-2.436	0.00734	$\beta_{1G} = 5.244$	$\beta_{2G} = 1.560$	$\underset{\sim}{b}_{\hat{G}} = \underset{\sim}{\hat{g}}$
T_{WG}	0.967	0.16602	$\gamma_{1W} = 5.132$	$\gamma_{2W} = 2.367$	$\underset{\sim}{g}_{\hat{W}} = \underset{\sim}{\hat{b}}$

For an ordering of the models according to their goodness in fitting the data, first comes the Weibull and then successively the gamma, the lognormal and lastly the exponential regression model. This is also the ordering from the maximum of the log likelihood functions. Table 3.5.3 gives the maximum of the log likelihood functions, Table 3.5.4 the log likelihood ratios and Table 3.5.5 the results of the maximum likelihood estimation.

TABLE 3.5.3 Maximum of the log likelihood functions – ℓ.

Model	ℓ
Lognormal – ℓ_L	-563.9347
Exponential – ℓ_E	-570.4434
Gamma – ℓ_G	-557.0608
Weibull – ℓ_W	-554.8098

TABLE 3.5.4 Log likelihood ratios – $\lambda_{(.)}$.

Log likelihood ratios	Observed $\lambda_{(.)}$
$\ell_E-\ell_L = \lambda_{EL}$	- 6.5083
$\ell_E-\ell_G = \lambda_{EG}$	-13.3826
$\ell_E-\ell_W = \lambda_{EW}$	-15.6337
$\ell_L-\ell_G = \lambda_{LG}$	- 6.8739
$\ell_L-\ell_W = \lambda_{LW}$	- 9.1250
$\ell_G-\ell_W = \lambda_{GW}$	- 2.2511

TABLE 3.5.5 Maximum Likelihood Estimates for the Models.

Model	Constant Term	Shape	Regression Coefficients										
	$\hat{\alpha}_1$	$\hat{\alpha}_2$											
Lognormal	4.8896	0.6137	$\hat{a}'=$	[-0.0065	0.0057	0.1141	0.0583	0.2883	0.3778	0.5877	0.3419	0.8351	-0.4575]
	(0.081)	(0.090)		(0.008)	(0.002)	(0.197)	(0.172)	(0.064)	(0.357)	(0.354)	(0.371)	(0.438)	(0.509)
	$\hat{\delta}$												
Exponential	5.1338	-	$\hat{d}'=$	[-0.0019	0.0075	-0.0588	0.1164	0.2556	0.5001	0.5875	0.7472	0.7474	-0.5218]
	(0.104)			(0.011)	(0.002)	(0.252)	(0.220)	(0.082)	(0.456)	(0.452)	(0.473)	(0.559)	(0.649)
	$\hat{\gamma}_1$	$\hat{\gamma}_2$											
Gamma	5.1338	2.1999	$\hat{g}'=$	[-0.0019	0.0075	-0.0588	0.1164	0.2556	0.5001	0.5875	0.7472	0.7474	-0.5218]
	(0.070)	(0.301)		(0.007)	(0.001)	(0.170)	(0.148)	(0.055)	(0.307)	(0.305)	(0.319)	(0.377)	(0.438)
	$\hat{\beta}_1$	$\hat{\beta}_2$											
Weibull	5.2461	1.6989	$\hat{b}'=$	[0.0017	0.0085	-0.1125	-0.1231	0.2305	0.5185	0.5312	0.4349	0.6551	-0.5581]
	(0.061)	(0.137)		(0.006)	(0.001)	(0.148)	(0.130)	(0.048)	(0.268)	(0.266)	(0.278)	(0.329)	(0.382)

Numbers in parentheses are the standard error given by <u>each</u> model.

TABLE 3.5.6 Data for Clinical Trial collected by the "Brain Tumor Study Group", M.D. Anderson Hospital and Tumor Institute, University of Texas:
(Y)-days of survival; (z_1)-age in years; (z_2)-duration of symptoms in weeks; (z_3)-sex*; (z_4)-treatment**; (z_5)- X-rays***

Frontal

Y	z_1	z_2	z_3	z_4	z_5
15	57	9	0	1	1
20	60	9	1	1	0
22	60	32	0	1	1
25	53	50	1	0	0
32	57	8	1	1	2
41	67	27	1	1	0
49	57	8	0	0	0
51	56	60	1	1	0
56	68	37	0	1	3
59	36	15	1	1	0
71	60	22	1	1	1
97	48	23	0	0	0
119	48	187	1	0	0
121	57	23	1	0	0
131	59	19	1	1	0
162	50	37	1	0	3
181	45	41	1	0	1
214	42	39	1	1	2
231	53	38	1	0	3
259	44	99	1	0	0
264	53	42	1	1	2
281	66	43	1	1	2
336	58	153	1	1	3
347	57	80	1	1	2
359	55	57	1	0	1
410	68	14	1	1	2
484	50	84	1	0	0
522	52	86	1	1	3
1760	27	253	1	1	3

Parietal

Y	z_1	z_2	z_3	z_4	z_5
12	56	56	0	0	2
18	39	3	1	1	1
34	72	13	1	0	0
37	55	19	1	0	0
57	59	14	0	1	3
64	56	19	1	1	0
82	45	25	1	1	1
107	71	43	1	0	0
108	50	18	1	0	0
132	49	42	1	0	0
134	60	44	1	0	2
136	22	31	0	1	0
143	52	28	1	1	0
234	60	40	0	0	2
248	51	39	0	1	3
255	53	40	1	1	3
275	56	44	0	1	0
298	51	49	1	0	0
408	36	89	1	0	3

Temporal

Y	z_1	z_2	z_3	z_4	z_5
10	46	7	1	0	0
46	52	14	1	1	0
61	71	12	1	1	0
62	46	21	0	1	2
85	43	22	0	1	0
129	59	24	1	1	3
135	35	82	0	1	0
144	41	22	1	1	0
145	55	23	1	1	3
162	66	26	0	0	0
164	53	78	0	1	0
177	48	49	1	1	0
194	57	347	1	0	0
200	42	312	1	0	3
204	57	43	1	1	3
210	65	50	1	0	0
252	31	39	1	0	3
253	47	48	1	0	3
255	70	48	1	0	2
272	55	41	0	1	0
274	47	42	1	1	2
297	54	50	1	0	3
325	56	59	1	0	3
345	52	76	1	1	3
385	59	59	1	0	3
466	40	140	0	1	3
495	53	90	0	0	1
526	59	87	1	0	2
669	47	121	1	1	2

Occiptal

Y	z_1	z_2	z_3	z_4	z_5
79	40	15	1	1	0
102	52	47	1	0	1
147	45	23	1	1	0
162	64	30	0	1	0
272	30	41	1	0	3
479	58	71	1	0	3
475	57	71	1	0	2

Deep Basal Ganglia/Thalamus

Y	z_1	z_2	z_3	z_4	z_5
42	40	73	1	0	0
51	60	76	1	0	3
54	40	10	1	0	0
72	53	13	0	1	2

Others

Y	z_1	z_2	z_3	z_4	z_5
21	52	24	1	0	0
30	54	14	1	1	0
135	40	32	1	0	3
253	73	62	1	1	3
387	45	54	0	0	3

* 0-Female, 1-male
** 0-Control, 1-treatment
*** 0-none, 1-less than 3000 rads, 2-less than 5000 rads
3-5000 or more rads.

Another 3 patients still alive and were not included in this table and in the analysis.

3.6 Efficiency of the false regression model

In previous sections of this chapter tests of separate families of hypothesis for models containing regression covariates were considered. It can be seen [e.g. expression (3.2.3)] that the estimators of the regression coefficients are always consistent, independently of distribution assumptions. Here, the consequences of using the wrong model are investigated by comparing the properties of these estimators. First some general results and the notation are presented.

Let $\underset{\sim}{y} = (y_1, \ldots, y_n)$ be independent but not identically distributed observations, each with probability density function $f(y_i; \underset{\sim}{\alpha}', \underset{\sim}{z}_i)$ under H_f, and $g(y_i, \underset{\sim}{\beta}', \underset{\sim}{z}_i)$ under H_g where $\underset{\sim}{z}_i$ are known covariates and $\underset{\sim}{\alpha}$, $\underset{\sim}{\beta}$ are vector of unknown parameters, with p and q components, respectively. Assume that for each model, α_1 and β_1 are the constant terms, α_2 and β_2 the shape and the remainder of the elements of $\underset{\sim}{\alpha}$ and $\underset{\sim}{\beta}$ are the regression coefficients. It can be seen that for the cases presented earlier in the chapter usually p = q but in one case p = q − 1. Let $\hat{\underset{\sim}{\alpha}}$ and $\hat{\underset{\sim}{\beta}}$ denote the maximum likelihood estimator of $\underset{\sim}{\alpha}$ and $\underset{\sim}{\beta}$, respectively. Recall that under H_f, $\hat{\underset{\sim}{\beta}}$ converges in probability to $\underset{\sim}{\beta}_{\underset{\sim}{\alpha}}$ and write

$$F = \sum_{i=1}^{n} \log f(y_i; \underset{\sim}{\alpha}', \underset{\sim}{z}_i) \ , \ F_{\underset{\sim}{\alpha}} = \frac{\partial}{\partial_{\underset{\sim}{\alpha}}} F \ , \ F_{\underset{\sim}{\alpha}'\underset{\sim}{\alpha}} = \frac{\partial^2}{\partial_{\underset{\sim}{\alpha}}'\partial_{\underset{\sim}{\alpha}}} F$$

with an analogous interpretation for G, $G_{\underset{\sim}{\beta}}$ and $G_{\underset{\sim}{\beta}'\underset{\sim}{\beta}}$. Here $F_{\underset{\sim}{\alpha}}$ is a (p × 1) vector, $F_{\underset{\sim}{\alpha}'\underset{\sim}{\alpha}}$ is a (p × p) matrix and further

$$\left(\frac{\partial \underset{\sim}{\beta}_{\underset{\sim}{\alpha}}}{\partial \underset{\sim}{\alpha}}\right) = \begin{pmatrix} \frac{\partial \beta_{1\underset{\sim}{\alpha}}}{\partial \alpha_1} & \cdots & \frac{\partial \beta_{q\underset{\sim}{\alpha}}}{\partial \alpha_1} \\ \cdots & \cdots & \cdots \\ \frac{\partial \beta_{1\underset{\sim}{\alpha}}}{\partial \alpha_p} & \cdots & \frac{\partial \beta_{q\underset{\sim}{\alpha}}}{\partial \alpha_p} \end{pmatrix}$$

is a (p × q) matrix.

Under H_f, $(\hat{\underset{\sim}{\alpha}}, \hat{\underset{\sim}{\beta}})$ is asymptotically multivariate normally distributed with variance-covariance matrix, given by Cox [1961, expressions(40) to (43)], namely

$$E_f(F_{\underset{\sim}{\alpha}} G_{\underset{\sim}{\beta}'}) = - \left(\frac{\partial \underset{\sim}{\beta}_{\underset{\sim}{\alpha}}}{\partial \underset{\sim}{\alpha}}\right) E_f(G_{\underset{\sim}{\beta}'\underset{\sim}{\beta}}) \, , \tag{3.6.1}$$

$$\mathrm{Cov}_f(\hat{\underset{\sim}{\alpha}}, \hat{\underset{\sim}{\beta}}) = \{E_f(F_{\underset{\sim}{\alpha}'\underset{\sim}{\alpha}})\}^{-1} E_f(F_{\underset{\sim}{\alpha}} G_{\underset{\sim}{\beta}'}) \{E_f(F_{\underset{\sim}{\alpha}'\underset{\sim}{\alpha}})\}^{-1} = - \{E_f(F_{\alpha'\alpha})\}^{-1} \left(\frac{\partial \underset{\sim}{\beta}_{\alpha}}{\partial \underset{\sim}{\alpha}}\right), \tag{3.6.2}$$

$$V_f(\hat{\alpha}) = - \{E_f(F_{\underset{\sim}{\alpha}'\underset{\sim}{\alpha}})\}^{-1} \, , \tag{3.6.3}$$

$$V_f(\hat{\beta}) = \{E_f(G_{\underset{\sim}{\beta}'\underset{\sim}{\beta}})\}^{-1} E_f(G_{\underset{\sim}{\beta}} G_{\underset{\sim}{\beta}'}) \{E_f(G_{\underset{\sim}{\beta}'\underset{\sim}{\beta}})\}^{-1} \tag{3.6.4}$$

These expressions are calculated at $(\underset{\sim}{\alpha}, \underset{\sim}{\beta}_{\underset{\sim}{\alpha}})$ the mean vector of the asymptotic normal distribution. The subscripts f, mean that the expectations, etc. are claculated under H_f.

Now, the true model is $f(y_i; \underset{\sim}{\alpha}, \underset{\sim}{z}_i)$ but $g(y_i, \underset{\sim}{\beta}, \underset{\sim}{z}_i)$ was supposedly used. In the problems considered, for a regression coefficient α_j say, it can be seen from previous sections, that the following relation holds

$$\beta_{j\underset{\sim}{\alpha}} = \alpha_j \qquad (j \neq 1, 2) \, .$$

Therefore (3.6.3) and 3.6.4) are of primary interest for comparison between $\hat{\beta}_j$ and $\hat{\alpha}_j$. Similarly it would be useful to comment on the corresponding elements of

$$p(\hat{\underset{\sim}{\beta}}/H_f) = \{E_g(G_{\underset{\sim}{\beta}'\underset{\sim}{\beta}})\}^{-1}_{\underset{\sim}{\beta}_{\underset{\sim}{\alpha}}} \, , \tag{3.6.5}$$

the probability limit of the <u>false estimator</u> of the variance-covariance matrix of $\hat{\underset{\sim}{\beta}}$, which is used when it is not known that the model is wrong.

The <u>efficiency of the false model</u> will be measured by the ratio of the determinants

$$\mathrm{Eff}(\hat{\underset{\sim}{\beta}}^*/H_f) = \frac{|V_f(\hat{\underset{\sim}{\alpha}}^*)|^{1/m}}{|V_f(\hat{\underset{\sim}{\beta}}^*)|^{1/m}} \tag{3.6.6}$$

and will provide insight into the result of using a false model. Here $\underset{\sim}{\alpha}^*$ and $\underset{\sim}{\beta}^*$ are vectors of regression coefficient estimator with m (= p − 2 or p − 1) components. The efficiency (3.6.6) is defined for $m \geq 1$.

Finally, a simplification brought about by our parametrization of the z_i's is pointed out. All models studied are log-linear (Section 1.2); it then follows (Cox & Hinkley, 1968) that

$$E_f(G_{\beta'\beta}) = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}, \quad E_f(G_\beta G_{\beta'}) = \begin{bmatrix} C & 0 \\ 0 & D \end{bmatrix},$$

where A and C are square matrices correpsonding to expected value of derivatives corresponding to the general mean and the shape or scale of $\log y_i$. The submatrices B and D are the corresponding matrices for the regression coefficients. Consequently for the elements of (3.6.4) corresponding to regression coefficients, only B and D need be determined.

For convenience, some expressions needed to evaluate (3.6.4) are given. With the notation of Section 1.2 it follows

$$\frac{\partial}{\partial b} \ell_W(\beta_1,\beta_2,b';y) = \beta_2 \sum_{i=1}^{n} z_i' \left(\frac{y_i}{e^{\beta_1+z_i b}} \right), \quad \frac{\partial^2}{\partial b'\partial b} \ell_W(\beta_1,\beta_2,b';y) = -\beta_2^2 \sum_{i=1}^{n} z_i' z_i \left(\frac{y_i}{e^{\beta_i+z_i b}} \right),$$

$$\frac{\partial}{\partial g} \ell_G(\gamma_1,\gamma_2,g';y) = \gamma_2 \sum_{i=1}^{n} z_i' \frac{y_i}{e^{\gamma_1+z_i g}}, \quad \frac{\partial^2}{\partial g'\partial g} \ell_G(\gamma_1,\gamma_2,g;y) = -\gamma_2 \sum_{i=1}^{n} z_i' z_i \frac{y_i}{e^{\gamma_1+z_i g}},$$

$$\frac{\partial}{\partial \alpha} \ell_L(\alpha_1,\alpha_2,a';y) = \frac{1}{\alpha_2} \sum_{i=1}^{n} (\log y_i - \alpha_1 - z_i a), \quad \frac{\partial^2}{\partial a'\partial a} \ell_L(\alpha_1,\alpha_2,a';y) = -Z'Z \frac{1}{\alpha_2}. \qquad (3.6.7)$$

A Lognormal regression model

Suppose the correct model is $f_L(y_i;\alpha_1,\alpha_2,a')$. From (1.2.9) the asymptotic variance of $\hat{a}$ is

$$V_L(\hat{a}) = (Z'Z)^{-1}\alpha_2 . \qquad (3.6.8)$$

The consequences of using the other models is discussed

(Ai) False model - Weibull regression - $f_W(y_i;\beta_1,\beta_2;b')$

By recalling the probability limits in (3.3.2), we have

$$E_L\{\frac{\partial}{\partial \underset{\sim}{b}} \ell_W \frac{\partial}{\partial \underset{\sim}{b}'} \ell_W\} = Z'Z \frac{e-1}{\alpha_2} , \quad E_L\{\frac{\partial^2}{\partial \underset{\sim}{b}' \partial \underset{\sim}{b}} \ell_W\} = - Z'Z \frac{1}{\alpha_2} .$$

From (3.6.4), (1.2.13) and (3.3.2)

$$V_L(\hat{\underset{\sim}{b}}) = (Z'Z)^{-1} (e-1)\alpha_2 , \quad p(\hat{\underset{\sim}{b}}/H_L) = (Z'Z)^{-1}\alpha_2 , \qquad (3.6.9)$$

and the efficiency (3.6.6) becomes

$$\text{Eff}(\hat{\underset{\sim}{b}}/H_L) = \frac{1}{e-1} = 0.58 . \qquad (3.6.10)$$

It can be seen from (3.6.9) that the variance of $\hat{b}_j$ is 72% higher than its stated estimate.

(Aii) False model - gamma regression - $f_G(y_i;\gamma_1,\gamma_2,\underset{\sim}{g})$

Recalling the probability limits in (3.2.3)

$$E_L\{\frac{\partial}{\partial \underset{\sim}{g}} \ell_G \frac{\partial}{\partial \underset{\sim}{g}'} \ell_G\} = Z'Z \, (e^{\alpha_2} - 1)\gamma_{2L}^2, \quad E_L\{\frac{\partial^2}{\partial \underset{\sim}{g}' \partial \underset{\sim}{g}} \ell_G\} = - Z'Z\gamma_{2L} .$$

From (3.6.4), (1.2.17) and (3.2.3)

$$V_L(\hat{\underset{\sim}{g}}) = (Z'Z)^{-1} (e^{\alpha_2} - 1) , \quad p(\hat{\underset{\sim}{g}}/H_L) = (Z'Z)^{-1}\gamma_{2L} \qquad (3.6.11)$$

and the efficiency (3.6.6) becomes

$$\text{Eff}(\hat{\underset{\sim}{g}}/H_L) = \frac{\alpha_2}{e^{\alpha_2}-1} . \qquad (3.6.12)$$

It is easy to see that (3.6.12) is always less than one and that it decreases rapidly as α_2 increases. The values in Table 3.6.1 illustrates this point.

TABLE 3.6.1 - $\text{Eff}(\hat{\underset{\sim}{g}}/H_L)$

α_2	0.2	0.5	1.0	2.0	0.614
Eff	0.90	0.77	0.58	0.27	0.72

It is also interesting to observe that (3.6.12) approaches 1 when $\alpha_2 \to 0$. This is because as α_2 tends to zero the lognormal distribution approaches a normal distribution. For a normal distribution with mean $\exp\{\underset{\sim}{z}\underset{\sim}{a}\}$, the maximum likelihood equations for $\underset{\sim}{a}$ are the same as those for the gamma regression given in (1.2.18) or equivalently in (1.2.22).

Because the equation (3.2.3) for γ_{2L} cannot be solved analytically, the only comment that can be made from $p(\hat{\underset{\sim}{g}}/H_L)$ is that the stated estimate will only be in agreement with $V_L(\hat{\underset{\sim}{g}})$ for α_2 also satisfying $\alpha_2 = \log(\gamma_{2L} + 1)$.

(Aiii)False model - exponential regression - $f_E(y_i;\delta,\underset{\sim}{d}')$.

The same arguments could be applied to obtain the results in this case. Here instead it is simpler to recall that the maximum likelihood equations (1.2.22) for $\underset{\sim}{d}$ are the same as those in (1.2.18) for $\underset{\sim}{g}$. The expressions for this case are identical to those in (Aii) with $\gamma_2 = \gamma_{2L} = 1$.

B Weibull regression model

Now the correct model is $f_W(y_i;\beta_1,\beta_2,\underset{\sim}{b}')$. From (1.2.13) the asymptotic variance of $\underset{\sim}{b}$ is

$$V_W(\hat{\underset{\sim}{b}}) = (Z'Z)^{-1}\frac{1}{\beta_2^2}\ . \qquad (3.6.13)$$

(Bi) False model - lognormal regression-$f_L(y_i;\alpha_1,\alpha_2,\underset{\sim}{a}')$.

Recalling the probability limits in (3.3.7).

$$E_W\{\frac{\partial}{\partial \underset{\sim}{a}}\ell_L\frac{\partial}{\partial \underset{\sim}{a}'}\ell_L\} = Z'Z\frac{\beta_2^2}{\psi'(1)}\ ,\ E_W\{\frac{\partial}{\partial \underset{\sim}{a}'\partial \underset{\sim}{a}}\ell_L\} = -Z'Z\frac{\beta_2^2}{\psi'(1)}\ .$$

From (3.6.4), (1.2.9) and (3.3.7)

$$V_W(a) = (Z'Z)^{-1}\frac{\psi'(1)}{\beta_2^2}\ ,\ p(\hat{\underset{\sim}{a}}/H_W) = (Z'Z)^{-1}\frac{\psi'(1)}{\beta_2^2} \qquad (3.6.14)$$

and the efficiency (3.6.6) becomes

$$\mathrm{Eff}(\hat{\underset{\sim}{a}}/H_W) = \frac{1}{\psi'(1)} = 0.61\ . \qquad (3.6.15)$$

Here, $p(\hat{\underset{\sim}{a}}/H_W)$ shows that a correct estimate of the variance of $\hat{a}_j$, the least square estimator of a_j, is given.

(Bii) False model - gamma regression-$f_G(y_i;\gamma_1,\gamma_2;\underset{\sim}{g})$.

Recalling the probability limits in (3.4.7) we have

$$E_W\{\frac{\partial}{\partial \underset{\sim}{g}} \ell_G \frac{\partial}{\partial \underset{\sim}{g}'} \ell_G\} = Z'Z\ \gamma_{2W}^2 [CV]^2,\ E_W\{\frac{\partial^2}{\partial \underset{\sim}{g}'\partial \underset{\sim}{g}} \ell_G\} = -\ Z'Z\ \gamma_{2W}^2\ .$$

where

$$[CV]^2 = \left[\frac{\Gamma(^2/\beta_2+1)}{\{\Gamma(^1/\beta_2+1)\}^2} - 1\right]$$

is the square of the coefficient of variation of a Weibull distribution with shape parameter β_2.

From (3.6.4), (1.2.17) and (3.4.7)

$$V_W(\hat{\underset{\sim}{g}}) = (Z'Z)^{-1}\ [CV]^2\ ,\ p(\hat{\underset{\sim}{g}}/H_W) = (Z'Z)^{-1} \frac{1}{\gamma_{2W}} \qquad (3.6.16)$$

and the efficiency (3.6.6) becomes

$$Eff(\hat{\underset{\sim}{g}}/H_W) = \left\{\frac{1}{[\beta_2 CV]^2}\right\}. \qquad (3.6.17)$$

Table 3.6.2 gives the efficiency and other values of interest.

TABLE 3.6.2 - $Eff(\hat{\underset{\sim}{g}}/H_W)$

β_2	γ_{2W}	$[CV]^2$	Eff
0.4	0.266	9.865	0.63
0.6	0.468	3.091	0.90
0.8	0.712	1.589	0.98
1.2	1.333	0.699	0.99
2.0	3.131	0.273	0.92
3.6	8.931	0.094	0.82
5.0	16.612	0.052	0.76
1.699	2.365	0.365	0.95

It can be seen that the efficiency is high for β_2 near 1 as would be expected and seems to decrease for β_2 far from 1. These results on γ_{2W}, $[CV]^2$ and $p(\hat{\underset{\sim}{g}}/H_W)$ suggests that according to whether $\beta_2 < 1$ or $\beta_2 > 1$ an underestimate or an overestimate of $V_W(\hat{\underset{\sim}{g}})$ is given respectively.

(Biii) False model – exponential regression – $f_E(y_i;\delta,\underset{\sim}{d}')$.

Similarly to (Aiii), the results can be obtained by taking $\gamma_2 = \gamma_{2W}$ in the expression obtained in (Bii) for the gamma regression. Here $p(\hat{\underset{\sim}{d}}/H_W)$ also suggests that not always $V_W(\hat{d})$ is overestimated or underestimated.

C Gamma regression model

The correct model is $f_G(y_i;\gamma_1,\gamma_2,\underset{\sim}{b}')$. From (1.2.17) the asymptotic variance of $\hat{\underset{\sim}{g}}$ is

$$V_G\{\hat{\underset{\sim}{g}}\} = (Z'Z)^{-1}\frac{1}{\gamma_2} \cdot \tag{3.6.18}$$

(Ci) False model – lognormal regression – $f_L(y_i;\alpha_1,\alpha_2,\underset{\sim}{a}')$.

Recalling the probability limits in (3.2.8)

$$E_G\{\frac{\partial}{\partial \underset{\sim}{a}}\ell_L \frac{\partial}{\partial \underset{\sim}{a}'}\ell_L\} = Z'Z\frac{1}{\psi'(\gamma_2)}, \quad E\{\frac{\partial^2}{\partial \underset{\sim}{a}'\partial \underset{\sim}{a}}\ell_L\} = -Z'Z\frac{1}{\psi'(\gamma_2)} \quad .$$

From (3.6.4), (1.2.9) and (3.2.8)

$$V_G(\hat{\underset{\sim}{a}}) = (Z'Z)^{-1}\psi'(\gamma_2), \quad p(\hat{\underset{\sim}{a}}/H_G) = (Z'Z)^{-1}\psi'(\gamma_2), \tag{3.6.19}$$

and the efficiency (3.6.6) becomes

$$Eff(\hat{\underset{\sim}{a}}/H_G) = \frac{1}{[\gamma_2\psi'(\gamma_2)]} \quad . \tag{3.6.20}$$

It can be shown that the efficiency approaches 1 when γ_2 increases. This is because as γ_2 increases the gamma distribution approaches a lognormal distribution. For γ_2 tending to 0, the efficiency tends to zero. For $\gamma_2 = 2.1999$, the efficiency is 0.71; further values are presented in Cox & Hinkley (1968).

Here, $p(\hat{\underset{\sim}{a}}/H_G)$ shows that a correct estimate of the variance of $\hat{a}_j$, the least square estimate of a_j, is given.

(cii) False model – Weibull regression–$f_W(y_i,\beta_1,\beta_2,\underset{\sim}{b}')$.

Recalling the probability limits in (3.4.2)

$$E_G\{\frac{\partial}{\partial \underset{\sim}{b}} \ell_W \frac{\partial}{\partial \underset{\sim}{b}'} \ell_W\} = Z'Z\ \beta_{2G}^2\ [\overline{CV}]^2\ ,\ E_G\{\frac{\partial^2}{\partial \underset{\sim}{b}' \partial \underset{\sim}{b}} \ell_W\} = -Z'Z\ \beta_{2G}^2\ ,$$

where

$$[\overline{CV}]^2 = \left[\frac{\Gamma(2\beta_{2G}+\gamma_2)\Gamma(\gamma_2)}{\{\Gamma(\beta_{2G}+\gamma_2)\}^2} - 1\right]$$

is the square of the coefficient of variation of $Y^{\beta_{2G}}$, Y with a gamma distribution with shape parameter γ_2.

From (3.6.4), (1.2.13) and (3.4.2)

$$V_G(\hat{\underset{\sim}{b}}) = (Z'Z)^{-1}\ [\overline{CV}/\beta_{2G}]^2,\ p(\hat{\underset{\sim}{b}}/H_G) = (Z'Z)^{-1} \frac{1}{\beta_{2G}^2} \qquad (3.6.21)$$

and the efficiency (3.6.6) becomes

$$\mathrm{Eff}(\hat{\underset{\sim}{b}}/H_G) = \left\{\frac{1}{\gamma_2}\ [\beta_{2G}/\overline{CV}]^2\right\}. \qquad (3.6.22)$$

Table 3.6.3 gives the efficiency and other values of interest

TABLE 3.6.3 - $\mathrm{Eff}(\hat{\underset{\sim}{b}}/H_G)$

γ_2	β_{2G}	$[\overline{CV}]^2$	Eff
0.4	0.534	0.807	0.89
0.6	0.718	0.892	0.96
0.8	0.870	0.951	0.99
1.2	1.115	1.039	0.997
2.0	1.482	1.142	0.96
5.0	2.370	1.304	0.86
2.2	1.560	1.161	0.95

The efficiency is high for γ_2 near 1 as would be expected and seems to decrease for γ_2 far from 1. These results for $[\overline{CV}]^2$ and $p(\hat{\underset{\sim}{b}}/H_G)$ suggests that according to whether $\gamma_2 < 1$ or $\gamma_2 > 1$, an overestimate or an underestimate of $W_G(\hat{\beta})$ is given respectively.

(Ciii) False model – exponential regression – $f_E(y_i;\delta,\underset{\sim}{d})$.

Again, from the comments on the maximum likelihood equation, the efficiency is 1 for this case. Here $p(\hat{\underset{\sim}{d}}/H_G) = (Z'Z)$. It can be seen that with $\beta_2 = \beta_{2W} = 1$ in (Cii) the results for the exponential regression model are also obtained.

D Exponential regression

The correct model is $f_E(y_i;\delta,\underset{\sim}{d}')$. From (1.2.21) the asymptotic variance of $\hat{\underset{\sim}{d}}$ is $(Z'Z)^{-1}$. The results for the case of using the lognormal regression can be obtained from(3.6.14) and (3.6.15) with $\beta_2 = 1$ or from (3.6.19) and (3.6.20) with $\gamma_2 = 1$. When the model used is the gamma regression the efficiency is 1 and the other results can be obtained from (3.6.16) and (3.6.17) with $\beta_2 = \gamma_{2W} = 1$. For the Weibull case the asymptotic efficiency is 1 and the results are obtained from (3.6.21) and (3.6.22) with $\gamma_2 = \beta_{2G} = 1$.

E Concluding remarks

The last entry of the tables in this section correspond to values of the example in Section 3.5. The results show that for the true Weibull model the efficiency of the lognormal model is 0.61 and the efficiency of the gamma and the exponential regression model is 0.95.

From the results on the variances it can be seen that optimizing Z'Z, consequently optimizes the asymptotic variances of the estimators. This means that asymptotically the distributional assumption has no importance for the design problem. The small sample consequences have not been investigated.

3.7 An extension for Markov processes

A possible extension for dependent observations is now discussed. Let $\underset{\sim}{y} = (y_1, \ldots, y_{n+1})$ be an observation from a Markov process with joint probability density function under H_f and under H_g

$$f_1(y_1,\alpha) \prod_{i=1}^{n} f_1(y_{i+1}/y_i,\alpha) \quad , \quad g(y_1,\beta) \prod_{i=1}^{n} g(y_{i+1}/y_i\beta) \ ,$$

respectively, where α and β are unknown parameters. Here $f_1(y_1,\alpha)$ and $g(y_1,\beta)$ specify an initial distribution, which is assumed to be the same as the final stationary distribution, whereas $f(y_{i+1}/y_i,\alpha)$ and $g(y_{i+1}/y_i,\beta)$ are one step transition probabilities. Assuming for convenience of notation α and β to be scalar and using a notation analogous to that of Section 3.6, write

$$F^i(\alpha) = \log f(y_{i+1}/y_i,\alpha) \ , \ F^i_\alpha(\alpha) = \frac{\partial}{\partial\alpha} F^i(\alpha) \ , \ F^i_{\alpha\alpha}(\alpha) = \frac{\partial^2}{\partial\alpha^2} F^i(\alpha)$$

with a similar interpretation for $G^i(\beta)$, $G^i_\beta(\beta)$ and $G^i_{\beta\beta}(\beta)$. Also, denote the log likelihood functions under H_f and H_g respectively, by

$$\ell_f(\alpha) = \log f_1(y,\alpha) + \sum_{i=1}^{n} F^i(\alpha) \ , \ \ell_g(\beta) = \log g_1(y,\beta) + \sum_{i=1}^{n} F^i(\alpha) \ ,$$

and the maximum likelihood estimators of α and β respectively, by $\hat{\alpha}$ and $\hat{\beta}$.

The terms $\log f_1(y_1,\alpha)$ and $\log g_1(y_1,\beta)$ can be omitted (Billingsley 1961, p.4) since the initial effects are unimportant as n becomes large.

Assume that under H_f, β_α is the limit in probability of $\hat{\beta}$, that $f(y_{i+1}/y_i,\alpha)$ and $g(y_{i+1}/y_i,\beta)$ satisfy the regularity conditions given by Billingsley(1961, p.5,6) which ensures that the log likelihood functions can be expanded in the usual way. Further assume that the central limit theorem and the law of large number apply to $F^i(\alpha)$ and $G^i(\beta)$. These conditions are sufficiently general to cover autoregressive problems and Markov chain.

The test statistic of the null hypotheses H_f against the alternative hypotheses H_f is based on

$$T^*_f = \ell_f(\hat{\alpha}) - \ell_g(\hat{\beta}) - E_{\hat{\alpha}}\{\ell_f(\alpha) - \ell_g(\beta_\alpha)\} \ . \qquad (3.7.1)$$

The asymptotic variance of T_f is obtained by arguments analogous to the independent case, that is expansion of $E_{\hat{\alpha}}\{\ell_f(\alpha)\}$ and $E_{\hat{\alpha}}\{\ell_f(\beta_\alpha)\}$ around α, $\ell_f(\alpha)$ around $\hat{\alpha}$ and $\ell_f(\beta_\alpha)$ around $\hat{\beta}$ lead to

$$V\{T_f^*\} = V_\alpha\{\ell_f(\alpha)-\ell_g(\beta_\alpha)\} - \frac{Cov_\alpha^2\{\ell_f(\alpha)-\ell_g(\beta_\alpha); \frac{\partial \ell_f(\alpha)}{\partial\alpha}\}}{V_\alpha\{\ell_f(\alpha)\}} .$$

Since we have assumed the central limit theorem applies to the log likelihood functions, it follows that T_f^* is asymptotically normally distributed with mean zero and variance rewritten as

$$V_\alpha\{T_f^*\} = V_\alpha\{\sum_{i=1}^{n} [F^i(\alpha)-G^i(\alpha)]\} - \frac{E_\alpha^2\{\sum_{i=1}^{n} [F^i(\alpha)-G^i(\beta_\alpha)]\, F_\alpha^i(\alpha)\}}{V_\alpha\{\sum_{i=1}^{n} F^i(\alpha)\}} . \quad (3.7.2)$$

Apart from the fact that in (3.7.1) and (3.7.2), $F^i(\alpha)$ and $G^i(\beta_\alpha)$ are transition probabilities, these expressions differ from the independent case only by the fact that the expectations and variances are calculated with the stationary distribution as the initial distribution.

In the absence of specific application only some realistic examples were these results could be applied are mentioned. The first is a generalization of the problem with quantal response studied in Cox (1962, ¶8); see also Atkinson (1970, ¶9) and Thomas (1972).

Suppose $\underset{\sim}{Z}_i = (X_i, Y_i)$ ($i = 1, \ldots, n$) is observed, where the Y_i's take the value 0 or 1, X_i ranges over 1, ..., k and some time elapses between the observation of X_i and Y_i. Within the hypothesis that $\underset{\sim}{Z}_i$ is a Markov chain, it is desirable to test the hypothesis H_f against H_g, where each of the hypotheses specify a different form of dependence of $p(y_{j+1}/y_j)$ on the variable x_j. The only difficulty here, could be computationally since the maximum likelihood estimates would have to be obtained by iterative methods.

A second example would be for the choice of the functional form of regression models when the error is generated by an autoregressive process.

A Bayesian solution for this problem is given by Lempers (1971). Williams (1970) has done some simulations on the likelihood ratio for one such problem. Incidentally, he noticed that the distribution of the likelihood ratio was the same as in the independent case.

Another problem related to dependent variables is to know whether the results of Section 3.2, 3.3 and 3.4 with some modifications could be applied when some of the $\underset{\sim}{z}_i$'s are lagged values of the dependent variable. Properties of the least square estimators obtained by treating the models as regression models have been given by Durbin.(1960). Unfortunately few results are available on maximum likelihood estimation for autoregressive problems with non-normal errors. However it is plausible to expect maximum likelihood estimators to have better properties than those of the least square estimators when the errors are not normal. In this case perhaps, the results could be applied, but this has not been investigated yet.

APPENDIX

A Derivative of vectors and matrices

Let $\underset{\sim}{x}$ and $\underset{\sim}{a}$ be (p×1) vectors, f be a scalar function of $\underset{\sim}{x}$ and $\underset{\sim}{F}$ a (q×1) vector function of $\underset{\sim}{x}$. The following derivatives are defined:

$$\frac{\partial}{\partial \underset{\sim}{x}} \cdot (\underset{\sim}{a}'\underset{\sim}{x}) = \underset{\sim}{a} ,$$

$$\frac{\partial}{\partial \underset{\sim}{x}} \cdot f(\underset{\sim}{x}) = \begin{bmatrix} \frac{\partial}{\partial x_1} \\ \vdots \\ \frac{\partial}{\partial x_p} \end{bmatrix} \cdot f(\underset{\sim}{x}) ,$$

$$\frac{\partial}{\partial \underset{\sim}{x}' \partial \underset{\sim}{x}} \cdot f(\underset{\sim}{x}) = \begin{bmatrix} \frac{\partial^2}{\partial x_1 \partial x_1} & \cdots & \frac{\partial^2}{\partial x_1 \partial x_p} \\ \cdot & \cdots & \cdot \\ \frac{\partial^2}{\partial x_p \partial x_1} & \cdots & \frac{\partial^2}{\partial x_p \partial x_p} \end{bmatrix} \cdot f(\underset{\sim}{x}) ,$$

$$\frac{\partial}{\partial \underset{\sim}{x}} \cdot \underset{\sim}{F}(\underset{\sim}{x}) = \left[\frac{\partial}{\partial \underset{\sim}{x}} \cdot F_1(x), \ldots, \frac{\partial}{\partial \underset{\sim}{x}} \cdot F_q(\underset{\sim}{x})\right] .$$

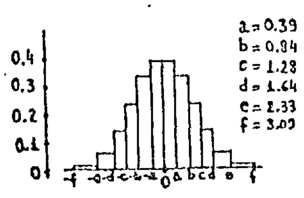
B Scale of figures

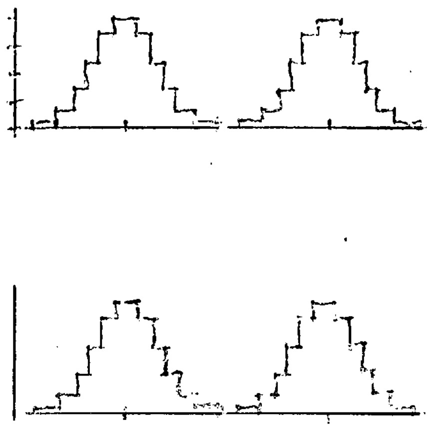
The coordinates for the graphs of Chapter 2 were chosen so that the area under each curve is one. For comparisons the corresponding results from the standard normal distribution are presented in Table B.1. For the comparison a transparency from Figure B.1 can be useful, and it is provided in the envelope.

Table B.1 Frequency and ordinate high of N(0,1)

z		Range	%	Ordinate Hight
-3.09 -	-2.33	0.76	0.009	0.01316
-2.33 -	-1.64	0.69	0.040	0.05797
-1.64 -	-1.28	0.36	0.050	0.13889
-1.28	-0.84	0.44	0.100	0.22727
-0.84	-0.39	0.45	0.150	0.33333
-0.39	0	0.39	0.150	0.38462
0	0.39	0.39	0.150	0.38462
0.39	0.84	0.45	0.150	0.33333
0.84	1.28	0.44	0.100	0.22727
1.28	1.64	0.36	0.050	0.13889
1.64	2.33	0.69	0.040	0.05797
2.33 -	-3.09	0.76	0.009	0.01316
			0.998	

FIGURE-B.1- Histogram from N(0,1)





REFERENCES

Abramowitz, M. and Stegun, I.A., eds., (1972) - Handbook of Mathematical Functions, eight printing, Dover.

Antle, C.E. and Klimko, L.A., (1975) - Choice of models for reliability studies and related topics III, ARL 75-0133, Aerospace Research Laboratories, Wright-Patterson AFB, Ohio.

Atkinson, A.C., (1969) - A test for discriminating between models. Biometrika 56, 337-347.

Atkinson, A.C., (1970) - A method for discriminating between models (with discussion), J.R. Statist.Soc. B, 32, 323-353.

Box, G.E.P. and Kanemasu, H., (1973) - Posterior probabilities of candidate models in model discrimination, Technical Report 322, Department of Statistics, University of Wisconsin.

Cox, D.R. and Brandwood, L., (1959) - On a discriminatory problem connected with the works of Plato, J.R. Statist.Soc. B, 21, 195-200.

Cox, D.R., (1961) - Tests of separate families of hypothesis, Proc. 4th Berkeley Symp., 1, 105-123.

Cox, D.R., (1962) - Further results on tests of separate families of hypothesis, J.R. Statist.Soc. B 24, 406-423.

Cox, D.R. and Hinkley, D.V., (1968) - A note on the efficiency of least-square estimates, J.R. Statist.Soc. B 30, 284-289.

Cox, D.R. and Hinkley, D.V.,(1974) - Theoretical Statistics, Chapman and Hall.

Dumonceaux, R., Antle, C.E. and Haas, G., (1973) - Likelihood ratio test for discrimination between two models with unknown location and scale parameters, Technometrics 15, 19-27.

Dumonceaux, R. and Antle, C.E., (1973) - Discrimination between the lognormal and the Weibull distribution, Technometrics 15, 923-926.

Durbin, J., (1960) - Estimation of parameters in time-series regression models, J.R. Statist.Soc. B 22, 139-153.

* Dyer, A.R., (1971) - A comparison of classification and hypothesis testing procedures for choosing between competing families of distributions, including a survey of the goodness of fit tests. Technical Memorandum No. 18, Aberdeen Research and Development Center, Aberdeen Proving Group, Maryland.

Farewell, V.T. and Prentice, R.L., (1974) - Extension and applications of the log-gamma regression model, unpublished.

Fraser, D.A.S., (1968) - The Structure of Inference, New York, Wiley.

Gross, A.J. and Clark, V.A., (1975) - Survival Distributions : Reliability Applications in the Biomedical Sciences, New York, Wiley.

Hogg, V.R., Uthoff, V.A., Randles, R.H. and Davenport, A.S., (1972) - On the selection of the underlying distribution and adaptive estimation, J. Am. Statist. Ass., 67, 597-600.

Holt, J.D. and Prentice, R.L., (1974) - Survival analysis in twin studies and matched pair experiments, Biometrika, 61, 17-30.

Huzurbazar, V.S., (1950) - Probability distributions and orthogonal parameters, Proc. Camb. Phil. Soc., 46, 281-284.

Jackson, O.A.Y., (1968) - Some results on tests of separate families of hypothesis, Biometrika, 55, 355-363.

Jackson, O.A.Y., (1969) - Fitting a gamma or log-normal distribution to fibre-diameter measurements on wool tops, Appl. Statist., 18, 70-75.

Lempers, F.B., (1971) - Posterior Probabilities for Alternative Linear Models, Rotterdam University Press.

Lindsey, J.K., (1974a) - Comparison of probability distributions, J.R. Statist. Soc., B 24, 38-47.

Lindsey, J.K., (1974b) - Construction and comparison of statistical models, J.R. Statist. Soc., B 24, 418-425.

Prentice, R.L., (1973) - Exponential survival with censoring and explanatory variables, Biometrika, 60, 279-288.

Prentice, R.L., (1974) - A log-gamma model and its maximum likelihood estimation, Biometrika, 61, 539-544.

Prentice, R.L., (1975) - Discrimination among some parametric models, Biometrika, 62, 607-614.

Quesenberry, C.R. and Starbuck, R.R., (1975) - On optimal tests for separate hypotheses and conditional probability integral transformation, Mimeograph Series No. 985, Institute of Statistics, University of North Carolina.

Selby, B., (1968) - A new approach to tests of separate families of hypothesis, Bell Telephone Technical Memorandum.

Starbuck, R.R., (1975) - Optimal tests for separable families of hypothesis, Mimeograph Series No. 978, Institute of Statistics, University of North Carolina.

Teekens, R., (1972) - Prediction Methods in Multiplicative Models, Rotterdam University Press.

Thomas, D.G., (1972) - Tests of fit for one-hit vs two-hit curve, (Statistical algorithm, AS-50), Appl. Statist. 21, 103-112.

Uthoff, V.A., (1970) - An optimum test property of two well-known statistics, J. Am. Statist. Ass., 65, 1597-1600.

Uthoff, V.A., (1974) - The most powerful scale and location invariant test of the normal versus the double-exponential, Ann. Statist., 1, 170-174.

Walker, A.M., (1967) - Some tests of separate families of hypothesis in time series problems, Biometrika, 54, 39-68. (Correction, Biometrika, 57, 226).

Walker,M.D., Gehan, E.A., Laventhal, C.M., Norrell, H.A. and Mahaley, M.S., (1969) - The evaluation of Mithramycin (NSC-24599) in the treatment of anaplastic glimas, Presented at the Fourth International Congress of Neurological Surgery, New York (September 20-27).

Weibull, W., (1971) - Outline of a theory of powerful selection of distribution functions, Technical Report - AFML-TR-71-52, Air Force Material Laboratory, Wright-Patterson, Air Force Base, Ohio.

Whittaker, J., (1974) - Generating gamma and beta random variables with non-integral shape parameters, Appl. Statist., 23, 210-214.

Williams, D.A., (1970) - Discrimination between regression models to determine the pattern of enzime syntheses in synchronous cell cultures, Biometrics, 28, 23-32.

Zellner, A., (1971) - An Introduction to Bayesian Inference in Econometrics, New York, Wiley.

* Dyer, A.R., (1973) - Discrimination procedures for separate families of hypotheses, J. Am. Statist. Ass., 68, 970-974.

* Dyer, A.R., (1974) - Hypothesis testing procedures for separate families of hypotheses, J. Am. Statist. Ass., 69, 140-145.

Billingsley, P., (1961) - Statistical Inference for Markov Processes, University of Chicago Press.