SOME RESULTS IN STATISTICAL INFERENCE

by

DAVID CHARLES CHANT

A THESIS SUBMITTED TO THE
AUSTRALIAN NATIONAL UNIVERSITY
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
IN THE
DEPARTMENT OF STATISTICS
RESEARCH SCHOOL OF SOCIAL SCIENCES

FEBRUARY, 1974

This thesis contains the original work of the author except where specific reference is made in the text.

D. C. Chant.

D.C.Chant

ACKNOWLEDGEMENTS

To my supervisor Professor P.A.P.Moran, who suggested many of the problems considered here, I extend my most sincere thanks.

I also wish to thank the Australian Government for their support under the Commonwealth Scholarship and Fellowship Plan, and Mrs.J.A.Radley for her excellent typing of this thesis.

CONTENTS

Corrigenda.

The expression for $f(\underset{\sim}{x};\underset{\sim}{\theta})$ on page 18 has a singularity at each data point (put $\mu = x_i$, any i, $\alpha > 0$ and let $\sigma \to 0$). An immediate consequence of this fact is that the usual asymptotic theory of maximum likelihood estimation cannot strictly speaking be applied to the full maximum likelihood treatment used in section (3.4). A similar stricture applies to the $C(\alpha)$ test of section (3.4) where the hypothesis $H_{o1} : \mu = \mu_o$ is considered. The $C(\alpha)$ tests constructed in sections (3.5), (3.6) and (3.7) are valid as the likelihoods used are non singular under the hypotheses considered.

These singularities may be the source of the failure of some of the Monte Carlo work in Appendix A.1, though I feel a greater cause of failure in the cases considered is the lack of bimodality in the underlying density. The program used to find the maximum likelihood estimates did, for numerical convenience, truncate σ slightly away from zero and this may, on occasion, have obscured the fact that the estimation procedure was moving towards a singularity.

SUMMARY

The chapters of this thesis compare the optimal $C(\alpha)$ tests of composite statistical hypotheses and tests based on maximum likelihood estimators when the parameter under test is interior to open sets in parameter space and in chapters 5 and 6 when the parameter under test lies on the boundary of a closed parameter space.

The performance of optimal $C(\alpha)$ tests and tests based on maximum likelihood estimators are compared for the problem of a mixture of two normal distributions with one component known and for the two sample problem with gamma random variables. The optimal $C(\alpha)$ tests for a class of regression problems and a general testing situation are discussed.

In the non standard situation where the parameter under test lies on a boundary of closed parameter space we find the asymptotic joint distribution function of the maximum likelihood estimators and give some examples. Attention is drawn to the related problem of truncating multivariate normal populations.

Key words: Optimal $C(\alpha)$ tests. Maximum likelihood.
Boundary problems.

1. INTRODUCTION

This thesis presents two main bodies of work. First a comparison of tests of composite hypotheses using mainly the optimal $C(\alpha)$ tests and the Wald statistic based on maximum likelihood estimators when the parameter under test is interior to open sets in parameter space. Second, we discuss the form of the joint asymptotic distribution function of the maximum likelihood estimator when the parameter under test lies on the boundary of a closed parameter space.

The study of the properties of maximum likelihood estimators and properties of tests based on maximum likelihood estimators in standard conditions is pursued in Wald (1943, 1949), Wolfowitz (1949), Mann and Wald (1943) and Rao (1948, 1961, 1965). A very general discussion may be found in Le Cam (1956). Wald's 1943 paper considers independent and identically distributed random variables. The extension to independent but not necessarily identically distributed random variables may be found in Hoadley (1971). The testing procedures involved, both in general and with particular reference to maximum likelihood, are dealt with in Lehmann (1959), Cramér (1946), Davies, R. (1969), Hodges and Lehmann (1970), Isaacson (1950) and Schaafsma and Smid (1966).

The optimal $C(\alpha)$ test procedures are developed in Bartlett (1953a, b, 1956) and in a major exposition by Neyman (1959). Neyman's paper covers the case of independently and identically distributed random variables with a scalar parameter under test and a vector nuisance parameter. The extension of the theory to independent but not necessarily identically distributed random variables with a scalar parameter under test is found in Bartoo and Puri (1967). The further generalisation of the theory to independent but not necessarily identically distributed random variables with a vector parameter under test is found in Buhler and Puri (1966). The literature on the application of optimal $C(\alpha)$ tests includes Neyman and Scott (1965), Kulkarni (1968) where particular reference is made to randomized experiments, Moran (1970a, b) with reference to rain-making experiments, Klonecki (1973) considers Poisson homogeneity tests, as does Moran (1973a).

In section (2.2) we demonstrate the asymptotic equivalence of optimal $C(\alpha)$ tests and tests based on maximum likelihood estimators when the vector parameter under test is interior to open sets in parameter space. This is a generalisation of a result in Moran (1970b) concerning tests for a scalar parameter. The optimal $C(\alpha)$ tests and tests based on maximum likelihood estimators are compared for some particular cases. The problem of a mixture of two normal components with one component known is treated in detail. This problem arises out of studies of Down's syndrome in Penrose and Smith (1966) and Moran (1973b). The associated five parameter problem when neither normal component is assumed known and the four parameter problem when each component is assumed to have the same unknown variance is covered in Blischke (1963), Hill (1963) and Day (1969). In our study we find the optimal $C(\alpha)$ tests to be more successful than the tests based on maximum likelihood estimators in the cases where optimal $C(\alpha)$ tests are applicable. Day (1969) points out that maximum likelihood fails in the five parameter situation. A similar failure of $C(\alpha)$ tests in this situation is shown, plus the failure of certain optimal $C(\alpha)$ tests in the four parameter case, which is amenable to treatment by maximum likelihood. However, as will be shown, optimal $C(\alpha)$ tests of certain relevant hypotheses concerning mixtures can be constructed if we have two samples.

In a comparism of methods using samples of gamma random variables, the test based on maximum likelihood estimators fares better than the optimal $C(\alpha)$ test, as shown in section (4.4), in the sense of rejecting the null hypotheses more consistently when the null hypotheses are in fact false.

The optimal $C(\alpha)$ tests are also shown to generate similar easily computable test statistics for regressions in three situations: regression with Cauchy errors and regression in Poisson and binary data. The performance of the test statistic for a simple linear regression is evaluated in section (4.2).

The study of maximum likelihood under the non-standard condition that the parameter lies on the boundary of a closed parameter space has been discussed in Moran (1971a, 1971b). We extend and modify the results of these papers in a study of the joint asymptotic distribution function of the maximum likelihood estimators when a vector parameter under test lies on the plane boundary of a closed Euclidean space. It will be shown that the equivalence between optimal $C(\alpha)$ tests and tests based on maximum likelihood estimators is lost apart from the case of a scalar parameter under test and one sided alternatives. In this case both tests have asymptotically the same rejection regions. The asymptotic joint distribution function is shown to be a sum of normal distributions, the components of which are dependent on certain conditionings on the maximum likelihood estimator of the parameter fixed under the null hypothesis. The problem is akin asymptotically to that of truncating a multivariate normal population, work on which is found in Birnbaum (1950), Brunk (1958) and Perlman (1969). The non-standard problem arising when the likelihood possesses a singularity, for example, may not admit a derivative for certain parameter values, is discussed in Daniells (1961) and Huber (1967).

As will be shown in section (2.2), the construction of $C(\alpha)$ tests requires the estimation of the nuisance parameter, and we may employ the so-called locally root-n consistent estimators. Such estimators are often readily available, and much easier to compute than some maximum likelihood estimators requiring an iterative procedure to determine them.

We also consider other inefficient estimators such as moment estimators, in order to find initial estimates for some of the maximum likelihood iterations performed, apart from the intrinsic interest of such estimates.

A condensed version of chapter 2 plus the work contained in chapter 5 and sections 6.1, 6.2, 6.3 and 6.5 of chapter 6 is to appear in the second part of *Biometrika*, 61, 1974, entitled "On optimal asymptotic tests of composite hypotheses in non-standard conditions". The work in chapter 3 is being prepared for publication.

2. THE EQUIVALENCE OF OPTIMAL $C(\alpha)$ TESTS BASED ON MAXIMUM LIKELIHOOD ESTIMATORS WHEN THE PARAMETER $\underset{\sim}{\theta}$ IS INTERIOR TO OPEN SETS IN PARAMETER SPACE

2.1 INTRODUCTION

Various asymptotic procedures have been constructed to test a composite hypothesis concerning a point $\underset{\sim}{\theta}$, say, in an s dimensional Euclidean parameter space Θ. For simplicity we shall refer to $\underset{\sim}{\theta}$ as a vector parameter.

Let

$$\underset{\sim}{\theta}' = [\underset{\sim}{\theta}_1' \; \underset{\sim}{\theta}_2'] \quad , \quad \underset{\sim}{\theta}_1' = [\theta_1 \ldots \theta_t] \quad , \quad \underset{\sim}{\theta}_2' = [\theta_{t+1} \ldots \theta_s]$$

where ' denotes transpose of a matrix or vector throughout this work.

The null hypotheses considered specify that $\underset{\sim}{\theta}_1 = \underset{\sim}{\theta}_{o1}$, say, where

$$\underset{\sim}{\theta}_{o1}' = [\theta_{o1} \ldots \theta_{ot}]$$

is a given vector, whilst $\underset{\sim}{\theta}_2$ is a nuisance parameter. Prominent among these procedures are tests based upon maximum likelihood estimators (Wald, 1943) and the optimal $C(\alpha)$ tests developed by Bartlett (1953), Neyman (1959), Bartoo and Puri (1967) and Buhler and Puri (1966).

Bartlett and Neyman cover tests based on independent and identically distributed random variables in which the parameter under test, $\underset{\sim}{\theta}_1$, is a scalar, θ_1, say.

Bartoo and Puri extend this case to independently but not necessarily identically distributed random variables with a scalar parameter θ_1 under test. Buhler and Puri consider the most general case of tests of the null hypothesis based upon independently distributed random variables with both $\underset{\sim}{\theta}_1$ and $\underset{\sim}{\theta}_2$ vector parameters.

Moran (1970) showed that optimal $C(\alpha)$ tests are equivalent to tests based on maximum likelihood estimators when dealing with independently and identically distributed random variables and the parameter under test, θ_1, is a scalar. The tests are equivalent in the sense that they asymptotically lead to the same rejection regions. This equivalence also requires that $\underset{\sim}{\theta}' = [\underset{\sim}{\theta}_1' \; \underset{\sim}{\theta}_2']$ is interior to an open set in parameter space Θ.

We now introduce some notation and show that optimal $C(\alpha)$ tests and tests based upon maximum likelihood estimators are still equivalent in the Buhler and Puri framework if $\underset{\sim}{\theta}$ is interior to an open set in Θ.

2.2 RESULTS FOR OPTIMAL $C(\alpha)$ TESTS IN THE BUHLER-PURI FRAMEWORK

Let $\underset{\sim}{\theta}_1' = [\theta_1 \ldots \theta_t]$ lie in an arbitrary open set Θ_1 in t dimensional Euclidean space and $\underset{\sim}{\theta}_2' = [\theta_{t+1} \ldots \theta_s]$ lie in an arbitrary open set Θ_2 in $(s-t)$ dimensional Euclidean space. Let $\Theta = \Theta_1 \times \Theta_2$.

For every $[\underset{\sim}{\theta}_1' \; \underset{\sim}{\theta}_2'] \varepsilon \Theta$, let $\{X_k(\underset{\sim}{\theta}_1, \underset{\sim}{\theta}_2)\}$, $k = 1,2,\ldots.$ denote a sequence of independent but not necessarily identically distributed random variables. The sample space, W_k^*, of $X_k(\underset{\sim}{\theta}_1, \underset{\sim}{\theta}_2)$ is assumed to be independent of $(\underset{\sim}{\theta}_1, \underset{\sim}{\theta}_2) \varepsilon \Theta$.

Let $p_k(x; \underset{\sim}{\theta}_1, \underset{\sim}{\theta}_2)$ be the probability density function of $X_k(\underset{\sim}{\theta}_1, \underset{\sim}{\theta}_2)$ with respect to some σ-finite measure, itself independent of the parameters.

Let

$$\underset{\sim}{X}_n'(\underset{\sim}{\theta}_1, \underset{\sim}{\theta}_2) = [X_1(\underset{\sim}{\theta}_1, \underset{\sim}{\theta}_2) \ldots X_n(\underset{\sim}{\theta}_1, \underset{\sim}{\theta}_2)]$$

be a vector random variable with sample space

$$W_n = W_1^* \times W_2^* \times \ldots \times W_n^*$$

Let w_n be a measurable subset of W_n, and α be an arbitrary number with $0 < \alpha < 1$.

The test constructed is to be an optimal asymptotic test of the hypothesis $H_o : \theta_1 = \theta_{o1}$ against alternatives $\theta_1 \neq \theta_{o1}$. The definitions of asymptotic tests and the optimality criterion are given in Neyman (1959), Bartoo and Puri (1967) and Buhler and Puri (1966). For convenience some of these definitions in the Buhler and Puri framework are given here.

(i) If a sequence $\{w_n\}$ of measurable sets has the property that for every $\theta_2 \in \Theta_2$,

$$\lim_{n\to\infty} P\{X_n(\theta_{o1},\theta_2) \in w_n\} = \alpha ,$$

then we say that $\{w_n\}$ defines an asymptotic test of the hypothesis $H_o : \theta_1 = \theta_{o1}$ corresponding to the level of significance α.

Let $K(\alpha)$ be a class of asymptotic tests of the hypothesis H_o, all corresponding to the same level α. Let $\theta_1^* = \{\theta_{n1}\}$, say, be a sequence of points belonging to Θ_1 and converging to θ_{o1}. Let Γ denote a certain class of sequences θ_1^* and let $\{w_n^{(o)}\} \in K(\alpha)$.

(ii) With reference to Γ, the test $\{w_n^{(o)}\}$ is optimal within the class $K(\alpha)$, if whatever be the sequence $\{\theta_{n1}\} \in \Theta_1$ and whatever be the fixed $\theta_2 \in \Theta_2$, the lower limit of the differences between the powers of $\{w_n^{(o)}\}$ and $\{w_n\}$, calculated at $[\theta_{n1}' \ \theta_2']$ is at least equal to zero. That is,

$$\lim_{n\to\infty} \inf[P\{X_n(\theta_{n1},\theta_2) \in w_n^{(o)}\}-P\{X_n(\theta_{n1},\theta_2) \in w_n\}] \geq 0.$$

(iii) Consider a random vector

$$\tilde{\theta}'_{n2} = [\tilde{\theta}_{n,t+1} \cdots \tilde{\theta}_{ns}] .$$

If there exists a non zero vector $a'_j = [a_{t+1} \cdots a_s]$ of real numbers such that for each $j = t+1, \ldots, s$

$$n^{\frac{1}{2}} | \tilde{\theta}_{nj} - \theta_j - a'_j (\theta_1 - \theta_{o1}) |$$

remains bounded in probability as $n \to \infty$ for all θ_1, θ_2, then we say that the sequence $\{\tilde{\theta}_{n2}\}$ represents a locally root-n consistent estimator of θ_2. In particular we may take $\tilde{\theta}_{n2}$ to be the maximum likelihood estimator of θ_2 under H_o.

Let

$$\phi_{ki}(X_k;\theta_2) = \frac{\partial}{\partial \theta_i} \log p_k(X_k;\theta_1,\theta_2) \Big|_{\theta_1 = \theta_{o1}}$$

for $i = 1, \ldots, s$; $k = 1, \ldots, n$.

Let

$$\phi_{k i}(\theta_2) = \sum_{k=1}^{n} \phi_{ki}(X_{ki};\theta_2).$$

We assume that $p_k(x;\theta_1,\theta_2)$ is at least twice differentiable with respect to all s parameters and that these differentiations are permissible under the integral sign extending over the sample space W_k^*. We also assume that the variance of each of the random variables $\phi_{ki}(X_k;\theta_2)$, $i = 1, \ldots, s$, exists.

Let

$$f_i(\theta_2) = \phi_i(\theta_2) - \sum_{j=t+1}^{s} a_{ij}^{(o)}(\theta_2)\phi_j(\theta_2) \ , \quad j = 1, \ldots, t. \tag{2.2.1}$$

The coefficients $\{a_{ij}^{(o)}(\underset{\sim}{\theta}_2)\}$ are chosen to minimize the variance of $f_i(\underset{\sim}{\theta}_2)$ as defined in (2.2.1), for each $i = 1,\ldots,t$.

Let

$$\{\underset{\sim}{f}(\underset{\sim}{\theta}_2)\}' = [f_1(\underset{\sim}{\theta}_2)\ldots f_t(\underset{\sim}{\theta}_2)]$$

and

$$\{\underset{\sim}{Y}_n(\underset{\sim}{\theta}_2)\}' = n^{-\frac{1}{2}}\ \underset{\sim}{f}(\underset{\sim}{\theta}_2) .$$

Let $\underset{\sim}{\Sigma}_{\underset{\sim}{f}}$ be the variance-covariance matrix of $\underset{\sim}{Y}_n(\underset{\sim}{\theta}_2)$. The detailed assumptions concerning the probability density functions and the $f_i(\underset{\sim}{\theta})$ are set out in Buhler and Puri (1966, pp.74-76).

Let

$$T_n(\underset{\sim}{\theta}_2) = \underset{\sim}{Y}_n'(\underset{\sim}{\theta}_2)\underset{\sim}{\Sigma}_{\underset{\sim}{f}}^{-1}\underset{\sim}{Y}_n(\underset{\sim}{\theta}_2) , \qquad (2.2.2)$$

we then have (Buhler and Puri, 1966),

(iv) $T_n(\tilde{\underset{\sim}{\theta}}_{n2})-T_n(\underset{\sim}{\theta}_2) \to 0$ in probability.

$\tilde{\underset{\sim}{\theta}}_{n2}$ is a locally root-n consistent estimator of $\underset{\sim}{\theta}_2$ as defined in (iii) above.

(v) $T_n(\tilde{\underset{\sim}{\theta}}_{n2})$ is asymptotically distributed as a chi-squared random variable with t degrees of freedom.

(vi) Let Γ^* be the class of all sequences $\{\underset{\sim}{\theta}_{n1}\} \varepsilon\ \Theta_1$ such that $\underset{\sim}{\theta}_{n1} \to \underset{\sim}{\theta}_{o1}$ with $n^{\frac{1}{2}}(\underset{\sim}{\theta}_{n1}-\underset{\sim}{\theta}_{o1})$ remaining bounded.

Let $\nu(\alpha)$ be determined by the condition

$$\int_{\nu(\alpha)}^{\infty} h_t(v)dv = \alpha$$

where $h_t(v)$ is the density function of a chi-squared random variable with t degrees of freedom. Let $B(\alpha)$ be the set of real numbers greater than $\nu(\alpha)$. We then have that the optimal $C(\alpha)$ test of the hypothesis $H_o : \underset{\sim}{\theta}_1 = \underset{\sim}{\theta}_{o1}$ against alternatives $\underset{\sim}{\theta}_1 \neq \underset{\sim}{\theta}_{o1}$ is based on the sequence $\{w_n^\dagger\}$, where $\{w_n^\dagger\}$ is defined by

$$T_n(\underset{\sim}{\tilde{\theta}}_{n2}) > \nu(\alpha) .$$

The test is optimal with respect to sequences in the family Γ^* in the sense of (i) and (ii) above.

2.3 PROOF OF THE EQUIVALENCE OF $C(\alpha)$ TESTS AND TESTS BASED ON MAXIMUM LIKELIHOOD ESTIMATORS

Let $\underset{\sim}{\hat{\theta}}_n$ be the maximum likelihood estimator of $\underset{\sim}{\theta}$ based on a sample of size n of independent random variables.

Then $\underset{\sim}{\hat{\theta}}_n$ satisfies

$$\sum_{k=1}^{n} \frac{\partial}{\partial \theta_i} \log p_k(X_k; \underset{\sim}{\hat{\theta}}_{n1}, \underset{\sim}{\hat{\theta}}_{n2}) = 0 \quad , \quad i = 1,\ldots,s, \tag{2.3.1}$$

where $\underset{\sim}{\hat{\theta}}'_n = [\underset{\sim}{\hat{\theta}}'_{n1}\ \underset{\sim}{\hat{\theta}}'_{n2}]$, $\underset{\sim}{\hat{\theta}}_{ni}$ being the maximum likelihood estimator of $\underset{\sim}{\theta}_i$, $i = 1,2.$

Let

$$c_{i\ell} = \lim_{n\to\infty} E\{-\frac{1}{n} \sum_{k=1}^{n} \frac{\partial^2}{\partial\theta_i \partial\theta_\ell} \log p_k(X_k; \underset{\sim}{\theta}_1, \underset{\sim}{\theta}_2)\} . \tag{2.3.2}$$

Put $\underset{\sim}{c} = (c_{i\ell})$ and partition $\underset{\sim}{c}$ so that

$$\underset{\sim}{c} = \begin{bmatrix} \underset{\sim}{c}_{11} & \underset{\sim}{c}'_{12} \\ \underset{\sim}{c}_{12} & \underset{\sim}{c}_{22} \end{bmatrix}$$

where $\underset{\sim}{c}_{11}$ is $t \times t$, $\underset{\sim}{c}_{12}$ is $t \times (s-t)$ and $\underset{\sim}{c}_{22}$ is $(s-t) \times (s-t)$. Put

$$\underset{\sim}{d} \equiv \underset{\sim}{c}^{-1} = \begin{bmatrix} \underset{\sim}{d}_{11} & \underset{\sim}{d}'_{12} \\ \underset{\sim}{d}_{12} & \underset{\sim}{d}_{22} \end{bmatrix},$$

and

$$T_n^{(m)}(\underset{\sim}{\theta}_2) = (\hat{\underset{\sim}{\theta}}_{n1}-\underset{\sim}{\theta}_1)' n \underset{\sim}{d}_{11}^{-1} (\hat{\underset{\sim}{\theta}}_{n1}-\underset{\sim}{\theta}_1)\Big|_{\underset{\sim}{\theta}_1=\underset{\sim}{\theta}_{o1}} \tag{2.3.3}$$

The maximum likelihood estimator $\hat{\underset{\sim}{\theta}}_n$, of $\underset{\sim}{\theta}$, is asymptotically distributed as $N_s(\underset{\sim}{\theta}, n^{-1}\underset{\sim}{c}^{-1})$, where $N_s(\underset{\sim}{\mu}, \underset{\sim}{\Sigma})$ denotes the s dimensional multivariate normal distribution, means vector $\underset{\sim}{\mu}$ and variance-covariance matrix $\underset{\sim}{\Sigma}$. In the scalar case, where $s = 1$, we drop the subscript s from the notation.

We now have

Theorem 1

With $T_n(\underset{\sim}{\theta}_2)$ and $T_n^{(m)}(\underset{\sim}{\theta}_2)$ as defined in (2.2.2) and (2.3.3) respectively, we have

$$T_n(\underset{\sim}{\theta}_2)-T_n^{(m)}(\underset{\sim}{\theta}_2) \to 0$$

in probability.

In a neighbourhood of $\underset{\sim}{\theta}$, the true parameter point, we may expand (2.3.1) as

$$0 = \sum_{k=1}^{n} \frac{\partial}{\partial\theta_i} \log p_k(X_k; \underset{\sim}{\theta}_1, \underset{\sim}{\theta}_2)$$

$$+ \sum_{k=1}^{n} \sum_{\ell=1}^{s} (\hat{\theta}_{n\ell}-\theta_\ell) \frac{\partial^2}{\partial\theta_i \partial\theta_\ell} \log p_k(X_k; \underset{\sim}{\theta}_1, \underset{\sim}{\theta}_2) + \varepsilon \tag{2.3.4}$$

where ε denotes higher order terms which converge to zero in probability. Also $n^{-1}\times$ (second derivatives of the log likelihood) converges in probability to the limit of its expectation. Hence from (2.3.4)

$$\sum_{k=1}^{n} \frac{\partial}{\partial\theta_i} \log p_k(X_k;\underset{\sim}{\theta}_1,\underset{\sim}{\theta}_2) \sim n \sum_{\ell=1}^{s} (\hat{\theta}_{n\ell}-\theta_\ell)c_{i\ell} , \quad i = 1,\ldots,s.$$

If H_o is true, then $[\underset{\sim}{\theta}'_{o1}\ \underset{\sim}{\theta}'_2]$ is the true parameter point and asymptotically (2.2.1) becomes

$$f_i \sim n \sum_{\ell=1}^{s} (\hat{\theta}_{n\ell}-\theta_\ell)\{c_{i\ell}-\sum_{j=t+1}^{s} a_{ij}^{(o)}(\underset{\sim}{\theta}_2)c_{j\ell}\}\Big|_{\underset{\sim}{\theta}_1=\underset{\sim}{\theta}_{o1}} . \tag{2.3.5}$$

The $\{a_{ij}^{(o)}(\underset{\sim}{\theta}_2)\}$ are chosen to minimize

$$E[\{\phi_i(\underset{\sim}{\theta}_2)-\sum_{j=t+1}^{s} a_{ij}(\underset{\sim}{\theta}_2)\phi_j(\underset{\sim}{\theta}_2)\}^2] , \quad i = 1,\ldots,t.$$

Hence

$$c_{i\ell}-\sum_{j=t+1}^{s} a_{ij}^{(o)}(\underset{\sim}{\theta}_2)c_{j\ell} = 0 , \quad \ell = t+1,\ldots,s; \; i = 1,\ldots,t. \tag{2.3.6}$$

Substituting (2.3.6) into (2.3.5) we have

$$f_i \sim n \sum_{\ell=1}^{t} (\hat{\theta}_{n\ell}-\theta_\ell)\{c_{i\ell}-\sum_{j=t+1}^{s} a_{ij}^{(o)}(\underset{\sim}{\theta}_2)c_{j\ell}\} . \tag{2.3.7}$$

We may write (2.3.6) in vector notation as

$$\underset{\sim}{a}^{(o)}(\underset{\sim}{\theta}_2) = \underset{\sim}{c}'_{12}\underset{\sim}{c}_{22}^{-1}\Big|_{\underset{\sim}{\theta}_1=\underset{\sim}{\theta}_{o1}}$$

where $\underset{\sim}{a}^{(o)}(\underset{\sim}{\theta}_2) = (a_{ij}^{(o)}(\underset{\sim}{\theta}_2))$, $i = 1,\ldots,t$; $j = t+1,\ldots,s$.

By independence of the observed random variables and the minimization property of the $\{a_{ij}^{(o)}(\underset{\sim}{\theta}_2)\}$, we have

$$(\underset{\sim}{\Sigma}_f)_{i\ell} = \frac{1}{n}\,\mathrm{cov}\{\phi_i(\underset{\sim}{\theta}_2) - \sum_{j=t+1}^{s} a_{ij}^{(o)}(\underset{\sim}{\theta}_2)\phi_j(\underset{\sim}{\theta}_2)\ ,$$

$$\phi_\ell(\underset{\sim}{\theta}_2) - \sum_{m=t+1}^{s} a_{\ell m}^{(o)}(\underset{\sim}{\theta}_2)\phi_m(\underset{\sim}{\theta}_2)\}$$

$$= c_{i\ell} - \sum_{j=t+1}^{s} a_{ij}^{(o)}(\underset{\sim}{\theta}_2)c_{\ell j} \quad , \quad i = \ell+1,\ldots,t,$$

or, in matrix notation,

$$\underset{\sim}{\Sigma}_f = \underset{\sim}{c}_{11} - \underset{\sim}{c}_{12}'\underset{\sim}{d}_{22}\underset{\sim}{c}_{12}\Big|_{\underset{\sim}{\theta}_1=\underset{\sim}{\theta}_{o1}} \ .$$

Since $\underset{\sim}{c}^{-1}\underset{\sim}{c} = \underset{\sim}{I}$, the identity matrix, it easily follows that

$$\underset{\sim}{d}_{11} = (\underset{\sim}{c}_{11} - \underset{\sim}{c}_{12}'\underset{\sim}{d}_{22}\underset{\sim}{c}_{12})^{-1}$$

(see, for example, Rao, 1965, p.29).

Hence (2.3.7) becomes, in vector notation,

$$\underset{\sim}{f} \sim n\underset{\sim}{d}_{11}^{-1}(\hat{\underset{\sim}{\theta}}_{n1} - \underset{\sim}{\theta}_1)\Big|_{\underset{\sim}{\theta}_1=\underset{\sim}{\theta}_{o1}} \qquad (2.3.8)$$

and so, by (2.3.8) we have

$$T_n(\underset{\sim}{\theta}_2) \sim (\hat{\underset{\sim}{\theta}}_{n1} - \underset{\sim}{\theta}_1)'n\underset{\sim}{d}_{11}^{-1}(\hat{\underset{\sim}{\theta}}_{n1} - \underset{\sim}{\theta}_1)\Big|_{\underset{\sim}{\theta}_1=\underset{\sim}{\theta}_{o1}}$$

$$= T_n^{(m)}(\underset{\sim}{\theta}_2)\ .$$

We may replace $\underset{\sim}{\theta}_2$ by a consistent estimator $\tilde{\underset{\sim}{\theta}}_{n2}$, say, and the rejection regions for tests based on $T_n(\tilde{\underset{\sim}{\theta}}_{n2})$ and $T_n^{(m)}(\tilde{\underset{\sim}{\theta}}_{n2})$ are asymptotically the same. The tests are also equivalent to a likelihood ratio test (see, for example, Lehmann (1959), Chernoff (1952)).

In applications, it is worth noting that whilst determining $\underset{\sim}{d}_{11}$ involves inverting the $s \times s$ matrix $\underset{\sim}{c}$, the evaluation of the equivalent matrix $(\underset{\sim}{c}_{11}-\underset{\sim}{c}'_{12}\underset{\sim}{d}_{22}\underset{\sim}{c}_{12})^{-1}$ involves two inversions of matrices of smaller dimensions, which is often computationally easier. Also, the fact that $\underset{\sim}{c}$ is evaluated under the null hypothesis when constructing the optimal $C(\alpha)$ test often affords further reduction in computation.

In chapter 5 we shall show that the conventional maximum likelihood theory must be refined if the parameter under test lies on the boundary of a closed parameter space, and that the equivalence between optimal $C(\alpha)$ tests and tests based on maximum likelihood estimators established in this chapter only holds if $\underset{\sim}{\theta}_1$ is a scalar θ_1, say, and the alternative hypothesis is one sided.

3. A COMPARISM OF SOME OPTIMAL $C(\alpha)$ TESTS AND TESTS BASED ON MAXIMUM LIKELIHOOD ESTIMATORS FOR A MIXTURE OF NORMAL COMPONENTS WITH ONE COMPONENT KNOWN

3.1 INTRODUCTION

We now introduce some notation which will be used throughout this chapter and chapter 4.

The log likelihood of a sample of observations for a particular model is written as $\ell(\underset{\sim}{\theta};\underset{\sim}{X})$, where $\underset{\sim}{X}$ represents the sample.

We write

$$\underset{\sim}{\phi}' = [\underset{\sim}{\phi}_1' \; \underset{\sim}{\phi}_2'] ,$$

where

$$\underset{\sim}{\phi}_1' = [\frac{\partial \ell}{\partial \theta_1}(\underset{\sim}{\theta};\underset{\sim}{X}) \ldots \frac{\partial \ell}{\partial \theta_t}(\underset{\sim}{\theta};\underset{\sim}{X})] , \quad \underset{\sim}{\phi}_2' = [\frac{\partial \ell}{\partial \theta_{t+1}}(\underset{\sim}{\theta};\underset{\sim}{X}) \ldots \frac{\partial \ell}{\partial \theta_s}(\underset{\sim}{\theta};\underset{\sim}{X})] ,$$

and

$$\underset{\sim}{c} = \begin{bmatrix} \underset{\sim}{c}_{11} & \underset{\sim}{c}_{12}' \\ \underset{\sim}{c}_{12} & \underset{\sim}{c}_{22} \end{bmatrix} , \quad \text{where} \quad c_{ij} = -E\{ \frac{\partial^2}{\partial \theta_i \partial \theta_j} \ell(\underset{\sim}{\theta};\underset{\sim}{X}) \} .$$

When considering different possible hypotheses concerning a particular model, we will use the same notations for $\underset{\sim}{c}$ and $\underset{\sim}{\phi}$ throughout, although the actual quantities may change. The exact meanings of $\underset{\sim}{c}$ and $\underset{\sim}{\phi}$ will be clear in the context of the problem.

Two main test statistics will be constructed.

(i) The optimal $C(\alpha)$ test statistic which may be written

(a) with a scalar parameter under test as

$$T^{(C)} = \left. \frac{\phi_1 - \underset{\sim}{c}'_{12} \underset{\sim}{c}^{-1}_{22} \underset{\sim}{\phi}_2}{(c_{11} - \underset{\sim}{c}'_{12} \underset{\sim}{c}^{-1}_{22} \underset{\sim}{c}_{12})^{\frac{1}{2}}} \right|_{H_o}$$

$T^{(C)}$ is asymptotically distributed as $N(0,1)$,

(b) with a vector parameter under test

$$T^{(C)} = (\underset{\sim}{\phi}_1 - \underset{\sim}{c}'_{12} \underset{\sim}{c}^{-1}_{22} \underset{\sim}{\phi}_2)' (\underset{\sim}{c}_{11} - \underset{\sim}{c}'_{12} \underset{\sim}{c}^{-1}_{22} \underset{\sim}{c}_{12})^{-1} (\underset{\sim}{\phi}_1 - \underset{\sim}{c}'_{12} \underset{\sim}{c}^{-1}_{22} \underset{\sim}{\phi}_2) \Big|_{H_o}$$

and $T^{(C)}$ is asymptotically distributed as a chi-squared random variable on t degrees of freedom.

(ii) The Wald statistic, which may be written

(a) with a scalar parameter under test as

$$T^{(W)} = \left. \frac{\hat{\theta}_{n1} - \theta_{o1}}{d_{11}^{\frac{1}{2}}} \right|_{H_o}$$

and $T^{(W)}$ is asymptotically distributed as $N(0,1)$,

(b) when a vector parameter is under test

$$T^{(W)} = (\hat{\theta}_{n1} - \theta_{o1})' \underset{\sim}{d}^{-1}_{11} (\underset{\sim}{\hat{\theta}}_{n1} - \underset{\sim}{\theta}_{o1}) \Big|_{H_o} ,$$

and $T^{(W)}$ is asymptotically distributed as a chi-squared random variable on t degrees of freedom.

3.2 A MIXTURE OF NORMAL COMPONENTS, WITH ONE COMPONENT KNOWN

In chapters 10 and 11 of Penrose and Smith's (1966) study of Down's anomaly they discuss the frequency distribution of the ages of mothers and patients, compared with that of the general population at the corresponding place and time.

The remarkable feature of the distribution is that there are two bumps in the curve. Moran (1973b) discusses this phenomenon and says that

> "any statistician confronted with these figures (for maternal age in Down's syndrome) would consider it highly likely that the distribution is a mixture of two different distributions, one identical with that in the general population and having no mean age shift, and one which is displaced well to the right. These two distributions are probably not very well fitted by the normal distribution and a better fit might be obtained by supposing that the logarithms of the mothers' ages are normally distributed. The resulting distribution would then be a mixture of two normal distributions and would have five parameters, the fifth being the relative proportions, α and $1-\alpha$ say, which each normal component contributes to the mixture. Since the mean and variance in the general population is known, only three parameters have to be estimated and this would not be very difficult".

The final comment is not borne out by results of simulation experiments with a small mean shift but is true for large shift in the mean.

The problem of estimating the components of a mixture of normal distributions, multivariate or scalar, but with the assumption of a common variance-covariance matrix has been examined elsewhere (see in particular, Blischke (1963) and Day (1969)).

In the univariate case, Day points out that if we assume that each component has a different unknown variance, then each sample point generates a singularity in the likelihood function. In this five parameter case maximum likelihood clearly breaks down. It is also impossible, as would be expected, to construct optimal $C(\alpha)$ tests based on a single sample of the three most relevant hypotheses

(i) equality of means

(ii) equality of variances

(iii) a joint test of (i) and (ii), being in effect a test for a univariate normal distribution against a normal mixture alternative.

If we assume one component known, as Moran suggests, we have three unknown parameters and the density function $f(x;\underset{\sim}{\theta})$ say, of the underlying random variable X, say, is given by

$$f(x;\underset{\sim}{\theta}) = \frac{\alpha}{\sigma}\,\phi^*\left(\frac{x-\mu}{\sigma}\right) + \frac{(1-\alpha)}{\sigma_o}\,\phi^*\left(\frac{x-\mu_o}{\sigma_o}\right) \qquad (3.2.1)$$

where

$$\underset{\sim}{\theta}' = [\mu\ \alpha\ \sigma] \quad \text{and} \quad \phi^*(z) = (2\pi)^{-\frac{1}{2}} \exp(-\tfrac{1}{2}z^2).$$

α is the proportion of the "unknown" component, with unknown mean μ and variance σ^2. μ_o and σ_o^2 are assumed known.

We may write the likelihood of a sample $x_1,\ldots,x_n$ as $f(\underset{\sim}{x};\underset{\sim}{\theta})$, where

$$f(\underset{\sim}{x};\underset{\sim}{\theta}) = (2\pi)^{-\frac{n}{2}} \left(\frac{\alpha}{\sigma}\right)^n \exp[-\frac{n}{2\sigma^2}\{s_{xx}+(\bar{x}-\mu)^2\}] \times$$

$$\prod_{i=1}^{n} \left[1 + \frac{\sigma}{\alpha}\cdot\frac{(1-\alpha)}{\sigma_0} \exp\left\{ \frac{1}{2\sigma^2}(x_i-\mu)^2 - \frac{1}{2\sigma_0^2}(x_i-\mu_0)^2 \right\} \right]$$

and

$$\bar{x} = n^{-1}\sum_{i=1}^{n} x_i \,, \qquad s_{xx} = n^{-1}\sum_{i=1}^{n}(x_i-\bar{x})^2.$$

Clearly this likelihood function possesses singularities.* We may consider the following three cases:-

(iv) Full maximum likelihood treatment, leading to estimators $\hat{\mu}, \hat{\alpha}, \hat{\sigma}$ of μ, α, σ respectively.

* see corrigenda

(v) A null hypothesis $H_{o1} : \mu = \mu_o$; α, σ^2 unknown.

(vi) A null hypothesis $H_{o2} : \mu = \mu_o$; $\sigma^2 = \sigma_o^2$, α unknown.

The latter is a test for a known normal population against an alternative of a normal mixture with one component known.

We also look at moment estimators in the above cases, as they yield initial values for the iterative schemes necessary to find maximum likelihood estimators of the parameters. We also briefly consider modified minimum chi-squared estimators.

3.3 MOMENT ESTIMATORS FOR THE MIXTURE OF NORMAL COMPONENTS

The first four moments of X are:-

$$E(X) \equiv m = \alpha\mu + (1-\alpha)\mu_o$$

$$\mu_2 = \sigma_o^2 + \alpha(\sigma_o^2-\sigma^2)+\alpha(1-\alpha)(\mu-\mu_o)^2$$

$$\mu_3 = 3\alpha(1-\alpha)(\mu-\mu_o)(\sigma^2-\sigma_o^2)+\alpha(1-\alpha)(1-2\alpha)(\mu-\mu_o)^3$$

$$\mu_4 = 3\sigma_o^4+3\alpha(\sigma^4-\sigma_o^4)+6\alpha(1-\alpha)(\mu-\mu_o)^2\{\sigma^2-\alpha(\sigma^2-\sigma_o^2)\}$$
$$+\alpha(1-\alpha)(1-3\alpha+3\alpha^2)(\mu-\mu_o)^4$$

where μ_r denotes the r^{th} central moment. Let $S_i = \frac{1}{n}\sum_{k=1}^{n}(X_k-\overline{X})^i$, $i = 2,3,4$. Then the moment estimators, $\tilde{\alpha}, \tilde{\mu}, \tilde{\sigma}^2$, say, in a full three parameter treatment are given by

$$a\tilde{\alpha}^2+b\tilde{\alpha}+c = 0 \qquad (3.3.1)$$

$$\tilde{\mu} = \mu_o + \frac{(S_1-\tilde{\alpha})}{\tilde{\alpha}} , \quad S_1 = n^{-1}\sum_{i=1}^{n} X_i$$

$$\tilde{\sigma}^2 = \sigma_o^2 + \frac{(S_2-\sigma_o^2)}{\tilde{\alpha}} - (1-\tilde{\alpha})(\frac{S_1-\mu_o}{\tilde{\alpha}})^2 \qquad (3.3.2)$$

where

$$a = S_3+3(S_1-\mu_o)(S_2-\sigma_o^2)+(S_1-\mu_o)^3$$

$$b = -3(S_1-\mu_o)\{(S_1-\mu_o)^2+(S_2-\sigma_o^2)\}$$

$$c = 2(S_1-\mu_o)^3.$$

Under the null hypothesis $H_{o1} : \mu = \mu_o$, the estimators $\tilde{\tilde{\alpha}}$, $\tilde{\tilde{\sigma}}^2$ of α and σ^2 respectively, are given by

$$\tilde{\tilde{\alpha}} = \frac{3(S_2-\sigma_o^2)^2}{(S_4-3\sigma_o^4)-6\sigma_o^2(S_2-\sigma_o^2)}$$

$$\tilde{\tilde{\sigma}}^2 = \left(\frac{S_2-\sigma_o^2}{\tilde{\tilde{\alpha}}}\right) + \sigma_o^2 . \tag{3.3.3}$$

Under the null hypothesis $H_{o2} : \mu = \mu_o;\ \sigma^2 = \sigma_o^2$, it will be seen that no estimator of α is required. This is intuitively reasonable, as there is no mixture under this hypothesis.

The moment estimators behave reasonably well if the true value of α is away from zero, that is, a large proportion of the mixture is formed by the unknown component. However, as appendix A.1 shows, the estimates are occasionally wild.

It often arises that the discriminant of (3.3.1) is negative or both roots lie outside (0,1). In this case, the order statistics $x_{(1)},\ldots,x_{(n)}$, say, of the sample may be used to construct a new estimate of α by taking every r^{th} order statistic as a sample point for the recalculation of the moments and of $\tilde{\alpha}$. This seems to generate estimators in (0,1), though not necessarily close to the true value of α. The values of $r = 5$ and 10 were used when needed in the simulation results.

The estimator $\tilde{\sigma}^2$ in (3.3.2) behaves badly if $\tilde{\alpha}$ is close to zero, and much the same is true for the estimator $\tilde{\tilde{\sigma}}^2$ in (3.3.3). Both estimators frequently give rise to negative values, particularly if the means are not much separated. In this case a new estimate $\tilde{\sigma}^2$ is found by calculating the usual estimate of variance based on the lower $100\,\tilde{\alpha}\,\%$ of the order statistics if $\tilde{\mu} < \mu_o$, and on the upper $100\,\tilde{\alpha}\,\%$ of the order statistics if $\tilde{\mu} > \mu_o$. This procedure produces good estimates of σ^2 if the true value of $\mu-\mu_o$ is appreciable, as the components of the mixture are then well separated and few observations "stray" into the wrong component. However if $\mu-\mu_o$ is small, in particular if $f(x;\underset{\sim}{\theta})$ is not bimodal, the procedure does little more than guarantee positive estimates of variance. If $\tilde{\tilde{\sigma}}^2 < 0$, we replace $\tilde{\tilde{\sigma}}^2$ by $\tilde{\sigma}^2$, the estimator of variance with three unknown parameters.

3.4 MAXIMUM LIKELIHOOD ESTIMATORS FOR THE COMPONENTS OF A MIXTURE

The log likelihood of a sample of observations $\underset{\sim}{x}' = [x_1 \ldots x_n]$ is

$$\ell(\underset{\sim}{\theta};\underset{\sim}{x}) = \sum_{i=1}^{n} \log\left\{ \frac{\alpha}{\sigma}\,\phi^*\left(\frac{x_i-\mu}{\sigma}\right) + \frac{(1-\alpha)}{\sigma_o}\,\phi^*\left(\frac{x_i-\mu_o}{\sigma_o}\right)\right\} .$$

The derivatives are

$$\frac{\partial \ell}{\partial \mu}(\underset{\sim}{\theta};\underset{\sim}{x}) = \frac{\alpha}{\sigma^2}\sum_{i=1}^{n} \left(\frac{x_i-\mu}{\sigma}\right)\phi^*\left(\frac{x_i-\mu}{\sigma}\right) \Big/$$

$$\left\{ \frac{\alpha}{\sigma}\,\phi^*\left(\frac{x_i-\mu}{\sigma}\right) + \frac{(1-\alpha)}{\sigma_o}\,\phi^*\left(\frac{x_i-\mu_o}{\sigma_o}\right)\right\}$$

$$\frac{\partial \ell}{\partial \alpha}(\underset{\sim}{\theta};\underset{\sim}{x}) = \sum_{i=1}^{n} \left\{ \frac{1}{\sigma}\,\phi^*\left(\frac{x_i-\mu}{\sigma}\right) - \frac{1}{\sigma_o}\,\phi^*\left(\frac{x_i-\mu_o}{\sigma_o}\right)\right\} \Big/$$

$$\left\{ \frac{\alpha}{\sigma}\,\phi^*\left(\frac{x_i-\mu}{\sigma}\right) + \frac{(1-\alpha)}{\sigma_o}\,\phi^*\left(\frac{x_i-\mu_o}{\sigma_o}\right)\right\}$$

$$\frac{\partial \ell}{\partial \sigma}(\underset{\sim}{\theta};\underset{\sim}{x}) = \frac{-\alpha}{\sigma^2}\sum_{i=1}^{n}[\{1-(\frac{x_i-\mu}{\sigma})^2\}\ \phi^*(\frac{x_i-\mu}{\sigma})]/$$

$$\{\frac{\alpha}{\sigma}\ \phi^*(\frac{x_i-\mu}{\sigma}) + \frac{(1-\alpha)}{\sigma_o}\ \phi^*(\frac{x_i-\mu_o}{\sigma_o})\}\ .$$

Let

$$P_i(\alpha,\mu,\sigma) = \frac{\alpha}{\sigma}\ \phi^*(\frac{x_i-\mu}{\sigma})/$$

$$\{\ \frac{\alpha}{\sigma}\ \phi^*(\frac{x_i-\mu}{\sigma}) + \frac{(1-\alpha)}{\sigma_o}\ \phi^*(\frac{x_i-\mu_o}{\sigma_o})\}\ .$$

Then $P_i(\alpha,\mu,\sigma)$ is the probability that the i^{th} observation is a member of the first (unknown) component of the mixture. The maximum likelihood estimators of α, μ and σ^2 in a full maximum likelihood treatment satisfy

$$\hat{\alpha} = \frac{1}{n}\sum_{i=1}^{n} P_i(\hat{\alpha},\hat{\mu},\hat{\sigma}) \tag{3.4.1}$$

$$\hat{\mu} = \sum_{i=1}^{n} x_i P_i(\hat{\alpha},\hat{\mu},\hat{\sigma}) / \sum_{i=1}^{n} P_i(\hat{\alpha},\hat{\mu},\hat{\sigma}) \tag{3.4.2}$$

$$\hat{\sigma}^2 = \sum_{i=1}^{n} (x_i-\hat{\mu})^2 P_i(\hat{\alpha},\hat{\mu},\hat{\sigma}) / \sum_{i=1}^{n} P_i(\hat{\alpha},\hat{\mu},\hat{\sigma})\ . \tag{3.4.3}$$

The equations (3.4.1)-(3.4.3) have the general character

$$\underset{\sim}{\hat{\theta}} = \underset{\sim}{\xi}(\underset{\sim}{\hat{\theta}};\underset{\sim}{x})\ .$$

For given initial conditions $\underset{\sim}{\tilde{\theta}} \equiv \underset{\sim}{\hat{\theta}}_o$, say, we may employ the iterative scheme

$$\underset{\sim}{\hat{\theta}}_i = \underset{\sim}{\xi}(\underset{\sim}{\hat{\theta}}_{i-1};\underset{\sim}{x})\ . \tag{3.4.4}$$

In an obvious notation we could alternatively employ the iterative scheme

$$\hat{\alpha}_i = \xi_1(\hat{\alpha}_{i-1},\ \hat{\mu}_{i-1},\ \hat{\sigma}_{i-1};\underset{\sim}{x})$$

$$\hat{\mu}_i = \xi_2(\hat{\alpha}_i,\ \hat{\mu}_{i-1},\ \hat{\sigma}_{i-1}; \underset{\sim}{x})$$

$$\hat{\sigma}_i^2 = \xi_3(\hat{\alpha}_i,\ \hat{\mu}_i,\ \hat{\sigma}_{i-1}; \underset{\sim}{x}) \ . \tag{3.4.5}$$

However the scheme (3.4.5) showed no advantage over (3.4.4) in reducing the number of iterations required for convergence to $\hat{\underset{\sim}{\theta}}$. In view of this (3.4.4) is a preferable scheme as only one pass through the data is required per iteration, effecting a considerable saving in computing time.

The speed of convergence, if any, depends on the initial values and on the underlying separation $\mu-\mu_o$. In fact the scheme often fails when $\mu-\mu_o$ is small. When (3.4.4) fails to converge to $\hat{\underset{\sim}{\theta}}$, it is possible to solve for $\hat{\underset{\sim}{\theta}}$ by the usual methods involving the second derivatives of the log likelihood function. This is also highly dependent for success on the accuracy of the initial estimates and the separation $\mu-\mu_o$.

Under the null hypothesis $H_{o1} : \mu = \mu_o$, the maximum likelihood estimates of α and σ^2 are given by

$$\hat{\hat{\alpha}} = \frac{1}{n} \sum_{i=1}^{n} P_i(\hat{\hat{\alpha}},\ \mu_o,\ \hat{\hat{\sigma}}) \tag{3.4.6}$$

$$\hat{\hat{\sigma}}^2 = \sum_{i=1}^{n} (x_i-\mu_o)^2 P_i(\hat{\hat{\alpha}},\ \mu_o,\ \hat{\hat{\sigma}}) / \sum_{i=1}^{n} P_i(\hat{\hat{\alpha}},\ \mu_o,\ \hat{\hat{\sigma}}) \ . \tag{3.4.7}$$

Again we may employ an iterative scheme similar to (3.4.4). As the success of the convergence of (3.4.4) depends on the separation of the components, it seems reasonable that we cannot expect the similar scheme based on (3.4.6) and (3.4.7) to behave any better when we have assumed zero separation. If the underlying model has a significantly non-zero value of $\mu-\mu_o$, the bias introduced into $f(x;\underset{\sim}{\theta})$ by assuming $\mu = \mu_o$ also hinders the convergence of the iterative scheme to $\hat{\hat{\underset{\sim}{\theta}}}$.

It is therefore reasonable to compare the $C(\alpha)$ test and the test based on the Wald statistic for values of μ close to μ_o.

For large underlying values of $\mu-\mu_o$, we deal only with the maximum likelihood estimates of the unknown parameters, which are of intrinsic interest.

The second derivatives of the log likelihood are given by

$$\frac{\partial^2 \ell}{\partial \alpha^2}(\underset{\sim}{\theta};\underset{\sim}{x}) = -\sum_{i=1}^{n} \{\frac{1}{\sigma}\phi^*(\frac{x_i-\mu}{\sigma}) - \frac{1}{\sigma_o}\phi^*(\frac{x_i-\mu_o}{\sigma_o})\}^2 / \{\frac{\alpha}{\sigma}\phi^*(\frac{x_i-\mu}{\sigma}) + \frac{(1-\alpha)}{\sigma_o}\phi^*(\frac{x_i-\mu_o}{\sigma_o})\}^2$$

$$\frac{\partial^2 \ell}{\partial \alpha \partial \mu}(\underset{\sim}{\theta};\underset{\sim}{x}) = \frac{1}{\sigma_o\sigma^2}\sum_{i=1}^{n} (\frac{x_i-\mu}{\sigma})\,\phi^*(\frac{x_i-\mu}{\sigma})\,\phi^*(\frac{x_i-\mu_o}{\sigma_o}) / \{\frac{\alpha}{\sigma}\phi^*(\frac{x_i-\mu}{\sigma}) + \frac{(1-\alpha)}{\sigma_o}\phi^*(\frac{x_i-\mu_o}{\sigma_o})\}^2$$

$$\frac{\partial^2 \ell}{\partial \alpha \partial \sigma}(\underset{\sim}{\theta};\underset{\sim}{x}) = \frac{-1}{\sigma_o\sigma^2}\sum_{i=1}^{n} \{1-(\frac{x_i-\mu}{\sigma})^2\}\,\phi^*(\frac{x_i-\mu}{\sigma})\,\phi^*(\frac{x_i-\mu_o}{\sigma_o})\} / \{\frac{\alpha}{\sigma}\phi^*(\frac{x_i-\mu}{\sigma}) + \frac{(1-\alpha)}{\sigma_o}\phi^*(\frac{x_i-\mu_o}{\sigma_o})\}^2$$

$$\frac{\partial^2 \ell}{\partial \mu^2}(\underset{\sim}{\theta};\underset{\sim}{x}) = -\sum_{i=1}^{n} [\frac{\alpha^2}{\sigma^4}\{\phi^*(\frac{x_i-\mu}{\sigma})\}^2 + \frac{\alpha(1-\alpha)}{\sigma_o\sigma^3}\phi^*(\frac{x_i-\mu}{\sigma})\,\phi^*(\frac{x_i-\mu_o}{\sigma_o})] / \{\frac{\alpha}{\sigma}\phi^*(\frac{x_i-\mu}{\sigma}) + \frac{(1-\alpha)}{\sigma_o}\phi^*(\frac{x_i-\mu_o}{\sigma_o})\}^2$$

$$\frac{\partial^2 \ell}{\partial\mu\partial\sigma}(\underset{\sim}{\theta};\underset{\sim}{x}) = -\sum_{i=1}^{n}\Big[\frac{2\alpha^2}{\sigma^4}\Big(\frac{x_i-\mu}{\sigma}\Big)\Big\{\phi^*\Big(\frac{x_i-\mu}{\sigma}\Big)\Big\}^2$$

$$+\frac{\alpha(1-\alpha)}{\sigma_o\sigma^3}\Big\{3-\Big(\frac{x_i-\mu}{\sigma}\Big)^2\Big\}\Big]\Big/$$

$$\Big\{\frac{\alpha}{\sigma}\phi^*\Big(\frac{x_i-\mu}{\sigma}\Big)+\frac{(1-\alpha)}{\sigma_o}\phi^*\Big(\frac{x_i-\mu_o}{\sigma_o}\Big)\Big\}^2$$

$$\frac{\partial^2 \ell}{\partial\sigma^2}(\underset{\sim}{\theta};\underset{\sim}{x}) = \sum_{i=1}^{n}\Big[\frac{\alpha^2}{\sigma^4}\Big\{1-3\Big(\frac{x_i-\mu}{\sigma}\Big)^2\Big\}\Big\{\phi^*\Big(\frac{x_i-\mu}{\sigma}\Big)\Big\}^2$$

$$+\frac{\alpha(1-\alpha)}{\sigma_o\sigma^3}\Big\{2-5\Big(\frac{x_i-\mu}{\sigma}\Big)^2+\Big(\frac{x_i-\mu}{\sigma}\Big)^4\Big\}\phi^*\Big(\frac{x_i-\mu}{\sigma}\Big)\phi^*\Big(\frac{x_i-\mu_o}{\sigma_o}\Big)\Big]\Big/$$

$$\Big\{\frac{\alpha}{\sigma}\phi^*\Big(\frac{x_i-\mu}{\sigma}\Big)+\frac{(1-\alpha)}{\sigma_o}\phi^*\Big(\frac{x_i-\mu_o}{\sigma_o}\Big)\Big\}^2. \qquad (3.4.8)$$

In estimating the asymptotic variance-covariance matrix of the maximum likelihood estimators, we may use either $E\{-\frac{\partial^2\ell}{\partial\alpha^2}(\underset{\sim}{\theta};\underset{\sim}{X})\}$ or $E[\{\frac{\partial\ell}{\partial\alpha}(\underset{\sim}{\theta};\underset{\sim}{X})\}^2]$ etc. The latter are computationally easier and lead to

$$c_{11} \equiv c_{\mu\mu} = \frac{n\alpha^2}{\sigma^4}F_2 \qquad (3.4.9)$$

$$c_{21} \equiv c_{\mu\alpha} = \frac{n\alpha}{\sigma^2}\Big(\frac{F_1}{\sigma}-\frac{G_1}{\sigma_o}\Big) \qquad (3.4.10)$$

$$c_{22} \equiv c_{\alpha\alpha} = n\Big(\frac{F_o}{\sigma^2}-\frac{2G_o}{\sigma\sigma_o}+\frac{H_o}{\sigma_o^2}\Big) \qquad (3.4.11)$$

$$c_{31} \equiv c_{\sigma\mu} = -\frac{n\alpha^2}{\sigma^4}(F_1-F_3)$$

$$c_{32} \equiv c_{\sigma\alpha} = -\frac{n\alpha}{\sigma^2}\Big\{\frac{1}{\sigma}(F_o-F_2)-\frac{1}{\sigma_o}(G_o-G_2)\Big\}$$

$$c_{33} \equiv c_{\sigma\sigma} = \frac{n\alpha^2}{\sigma^4} (F_o - 2F_2 + F_4)$$

where

$$F_i = \int_{-\infty}^{\infty} [(\frac{x-\mu}{\sigma})^i \{\phi^*(\frac{x_i-\mu}{\sigma})\}^2 /$$

$$\{ \frac{\alpha}{\sigma} \phi^*(\frac{x-\mu}{\sigma}) + \frac{(1-\alpha)}{\sigma_o} \phi^*(\frac{x-\mu}{\sigma})\}]dx \quad , \quad i = 0,\ldots,4.$$

$$G_i = \int_{-\infty}^{\infty} [(\frac{x-\mu}{\sigma})^i \phi^*(\frac{x-\mu}{\sigma}) \phi^*(\frac{x-\mu_o}{\sigma_o}) /$$

$$\{ \frac{\alpha}{\sigma} \phi^*(\frac{x-\mu}{\sigma}) + \frac{(1-\alpha)}{\sigma_o} \phi^*(\frac{x-\mu_o}{\sigma_o})\}]dx \quad , \quad i = 0,1,2.$$

$$H_o = \int_{-\infty}^{\infty} [\{\phi(\frac{x-\mu_o}{\sigma_o})\}^2 / \{ \frac{\alpha}{\sigma} \phi^*(\frac{x-\mu}{\sigma}) + \frac{(1-\alpha)}{\sigma_o} \phi^*(\frac{x-\mu_o}{\sigma_o})\}]dx. \qquad (3.4.12)$$

Under $H_{o1} : \mu = \mu_o$ the expressions (3.4.8) and the integrals (3.4.12) are evaluated with $\mu = \mu_o$. The integrals (3.4.12) are quickly evaluated numerically by the trapezium rule over the range $\{\min(\tilde{\mu},\mu_o) - d\max(\tilde{\sigma},\sigma_o), \ \max(\tilde{\mu},\mu_o) + d\max(\tilde{\sigma},\sigma_o)\}$. The value $d = 4$ is adequate to ensure against loss of accuracy by replacing the infinite range with a finite one.

The results in the appendix are given in terms of α, μ and σ^2. The asymptotic variance-covariance matrix of $(\hat{\alpha},\hat{\mu},\hat{\sigma}^2)$ is formed from (3.4.9), (3.4.10) and (3.4.11) with in addition,

$$c_{31} = -\frac{n\alpha^2}{2\sigma^5} (F_1 - F_3)$$

$$c_{32} = -\frac{n\alpha^2}{2\sigma^3} \{ \frac{1}{\sigma} (F_o - F_2) - \frac{1}{\sigma_o} (G_o - G_2)\}$$

$$c_{33} = \frac{n\alpha^2}{4\sigma^6} (F_o - 2F_2 + F_4).$$

The accuracy of the maximum likelihood estimators depends on the value of $(\mu-\mu_o)$ and the precision of the variance of the unspecified component of the mixture, i.e., σ^2. With $\mu = 2$, $\mu_o = 0$, the procedure worked in all but one of the trials with $\sigma^2 = 0{\cdot}25$. With $\sigma^2 = 1$, the procedure does not appear to work well until at least half the mixture is formed by the unspecified component. This last comment is more valid with $\sigma^2 = 2{\cdot}25$, in which case the components of the mixture are poorly separated and large sample sizes are required for the convergence of (3.4.4) to the maximum likelihood estimators.

With $\mu = 4$, $\mu_o = 0$, (3.4.4) behaves well for $\sigma^2 = 0{\cdot}25$ and $1{\cdot}0$. However, with $\sigma^2 = 2{\cdot}25$, the components of the mixture are again poorly separated and although (3.4.4) does lead to maximum likelihood estimates, the accuracy of the estimates, measured by the terms of the asymptotic variance-covariance matrix, is reduced considerably.

The results for $\mu = 2$ and 4; $\mu_o = 0$ are contained in appendix A.1. Only results for a full maximum likelihood analysis are given. The resultant test statistics are not quoted, as they would obviously be significant, given the achieved values of the estimates of μ and their variances.

With $\mu = 0$ and $0{\cdot}2$ a comparism of the optimal $C(\alpha)$ test of $H_{o1} : \mu = \mu_o$, with a test based on maximum likelihood estimators is given in appendix A.1.2. The procedure (3.4.4) behaves poorly, in most cases not leading to a maximum likelihood estimate at all. When (3.4.4) does work, it seems important that the greater proportion of the mixture be made up of the unspecified component. Even when a full maximum likelihood treatment fails, it is possible to construct the optimal $C(\alpha)$ test statistic using the moment estimates of the nuisance parameter.

Surprisingly the test statistics generated, treated as standard normal random variables, lead to consistent acceptance of H_{o1} when $\mu = 0$ and consistent rejection of H_{o1} when $\mu = 0{\cdot}2$. This is true even in the many cases when the moment estimates were far removed from the true underlying values.

3.5 A TEST OF A KNOWN NORMAL DISTRIBUTION AGAINST A NORMAL MIXTURE ALTERNATIVE WITH ONE COMPONENT KNOWN

Under $H_{o2} : \mu = \mu_o, \ \sigma^2 = \sigma_o^2$, the log likelihoods become

$$\frac{\partial \ell}{\partial \mu}(\underset{\sim}{\theta};\underset{\sim}{x}) = \frac{n\alpha}{\sigma_o^2}(\bar{x}-\mu_o)$$

$$\frac{\partial \ell}{\partial \sigma}(\underset{\sim}{\theta};\underset{\sim}{x}) = -\frac{\alpha}{\sigma_o^2}\sum_{i=1}^{n}\{1-(\frac{x_i-\mu_o}{\sigma_o})^2\}$$

$$\frac{\partial \ell}{\partial \alpha}(\underset{\sim}{\theta};\underset{\sim}{x}) \equiv 0.$$

We ignore all contributions associated with α. Clearly

$$\underset{\sim}{C} = \begin{array}{c} \mu \\ \sigma \end{array}\begin{bmatrix} \frac{n\alpha^2}{\sigma_o^2} & 0 \\ 0 & \frac{2n\alpha^2}{\sigma_o^2} \end{bmatrix}.$$

We have

$$T^{(C)} = \{n^{\frac{1}{2}}(\frac{\bar{X}-\mu_o}{\sigma_o})\}^2 + \frac{1}{2n}\{\sum_{i=1}^{n}(\frac{X_i-\mu_o}{\sigma_o})^2 - n\}^2$$

which is asymptotically distributed as a chi-squared random variable on 2 degrees of freedom. Under H_{o2} we have exactly that

$$E\{T^{(C)}\} = 2$$

$$\operatorname{var}\{T^{(C)}\} = 4 + \frac{72}{n} \to V(\chi_2^2) \qquad (3.5.1)$$

as $n \to \infty$ where χ_2^2 denotes a chi-squared random variable on 2 degrees of freedom. We may write

$$T^{(C)} = U + \frac{1}{2n}(V_{n-1}+U-n)^2$$

$$= \frac{(U+V_{n-1})^2}{2n} - V_{n-1} + \frac{n}{2}$$

where U is distributed as a chi-squared random variable on 1 degree of freedom, independently of V_{n-1}, which is distributed as a chi-squared random variable on $(n-1)$ degrees of freedom.

The distribution function $F_{T^{(C)}}(t)$ of $T^{(C)}$ is given by

$$F_{T^{(C)}}(t) = \frac{n^{\frac{1}{2}}}{2^{\frac{n+1}{2}}\pi^{\frac{1}{2}}\Gamma(\frac{n-1}{2})} \iint_{D^*} \{2^{\frac{1}{2}} n^{\frac{1}{2}} (y+x-\tfrac{n}{2})^{\frac{1}{2}} - x\}^{-\frac{1}{2}}$$

$$\times (y+x-\tfrac{n}{2})^{-\frac{1}{2}} x^{\frac{n-1}{2}-1} \exp\{-\frac{n^{\frac{1}{2}}}{2^{\frac{1}{2}}}(y+x-\tfrac{n}{2})^{\frac{1}{2}}\}dxdy \qquad (3.5.2)$$

where

$$D^* = [(x,y)\,|\max\{0,n-(2n)^{\frac{1}{2}}y\} < x < n+(2n)^{\frac{1}{2}}y^{\frac{1}{2}};\quad 0 < y \le t].$$

Alternatively, the exact distribution of $T^{(C)}$ under H_{o2} may be evaluated by integrating the joint distribution of (U,V_{n-1}) over

$$D^{\dagger} = [(u,v)\,|0 < u < (2n)^{\frac{1}{2}}(t+v-\tfrac{n}{2})^{\frac{1}{2}};\ \max\{0,n-(\tfrac{nt}{2})^{\frac{1}{2}}\} \le v \le n+(\tfrac{nt}{2})^{\frac{1}{2}}].$$

This approach leads to

$$F_{T^{(C)}}(t) = \frac{1}{\pi^{\frac{1}{2}}\,\Gamma(\frac{n-1}{2})} \int_{\max\{0,\frac{n}{4} - \frac{1}{2}(\frac{nt}{2})^{\frac{1}{2}}\}}^{\frac{n}{4} + \frac{1}{2}(\frac{nt}{2})^{\frac{1}{2}}} v^{\frac{n}{2} - \frac{3}{2}} \times$$

$$\gamma\{n^{\frac{1}{2}} (\frac{t}{2} + v - \frac{n}{4})^{\frac{1}{2}} - v,\ \frac{1}{2}\}\ e^{-v} dv \qquad (3.5.3)$$

where

$$\gamma(x,t) = \int_{o}^{x} v^{t-1} e^{-v} dv$$

is the incomplete gamma function. The numerical integration of either (3.5.2) or (3.5.3) by Simpson's rule or an adaptive quadrature algorithm based on a closed five point Newton Cotes formula proved to take excessive computer time. Percentage points of the distribution of $T^{(C)}$ are reasonably quickly found using Monte-Carlo techniques. Values of t such that

$$\mathrm{pr}\{T^{(C)} \leq t\ ;\ H_o\} = \gamma \qquad (3.5.4)$$

are given in appendix 2 for $\gamma = \cdot 9(\cdot 05)\cdot 95$, based on sample sizes $n = 20(20)100$.

The distribution of $T^{(C)}$ is close to its asymptotic value for sample sizes 80 and 100, but is removed from its asymptotic value for smaller sample sizes, a result in line with the variance of $T^{(C)}$ given by (3.5.1).

3.6 THE FAILURE OF AN OPTIMAL $C(\alpha)$ TEST FOR A FIVE PARAMETER MIXTURE

Using the density (3.2.1) but treating μ_o and σ_o as unknown we may test a null hypothesis H_{o4} : $\mu = \mu_o$, $\sigma = \sigma_o$ by employing the reparameterisation $\mu_o = \mu+\Delta$, $\sigma_o = \sigma e^{\xi}$, the latter being employed to ensure positive estimates of variance.

H_{o4} then becomes H_{o4} : $\Delta = \xi = 0$.

We are thus testing for a normal distribution with unknown mean and variance against a five parameter normal mixture.

We have

$$\frac{\partial \ell}{\partial \Delta}(\underset{\sim}{\theta};\underset{\sim}{x}) = \frac{(1-\alpha)}{\sigma^2}\exp(-3\xi)\sum_{i=1}^{n}\left(\frac{x_i-\mu-\Delta}{\sigma}\right)\left\{\phi^*\left(\frac{x_i-\mu-\Delta}{\sigma}\right)\exp(-\xi)\right\}/$$

$$\left[\frac{\alpha}{\sigma}\phi^*\left(\frac{x_i-\mu}{\sigma}\right)+\frac{(1-\alpha)}{\sigma}\exp(-\xi)\;\phi^*\left\{\left(\frac{x_i-\mu-\Delta}{\sigma}\right)\exp(-\xi)\right\}\right] \quad (3.6.1)$$

$$\frac{\partial \ell}{\partial \xi}(\underset{\sim}{\theta};\underset{\sim}{x}) = \frac{(1-\alpha)}{\sigma}\exp(-\xi)\sum_{i=1}^{n}\left\{\left(\frac{x_i-\mu-\Delta}{\sigma}\right)^2\exp(-2\xi)-1\right\}$$

$$\times\;\phi^*\left\{\left(\frac{x_i-\mu-\Delta}{\sigma}\right)\exp(-\xi)\right\}/\left[\frac{\alpha}{\sigma}\phi^*\left(\frac{x_i-\mu}{\sigma}\right)\right.$$

$$\left.+\frac{(1-\alpha)}{\sigma}\exp(-\xi)\;\phi^*\left\{\left(\frac{x_i-\mu-\Delta}{\sigma}\right)\exp(-\xi)\right\}\right] \quad (3.6.2)$$

$$\frac{\partial \ell}{\partial \alpha}(\underset{\sim}{\theta};\underset{\sim}{x}) = \sum_{i=1}^{n}\left[\frac{1}{\sigma}\phi^*\left(\frac{x_i-\mu}{\sigma}\right)-\frac{\exp(-\xi)}{\sigma}\phi^*\left\{\left(\frac{x_i-\mu-\Delta}{\sigma}\right)\exp(-\xi)\right\}\right]/$$

$$\left[\frac{\alpha}{\sigma}\phi^*\left(\frac{x_i-\mu}{\sigma}\right)+\frac{(1-\alpha)}{\sigma}\exp(-\xi)\;\phi^*\left\{\left(\frac{x_i-\mu-\Delta}{\sigma}\right)\exp(-\xi)\right\}\right] \quad (3.6.3)$$

$$\frac{\partial \ell}{\partial \mu}(\underset{\sim}{\theta};\underset{\sim}{x}) = \sum_{i=1}^{n} \left[\frac{\alpha}{\sigma^2}\left(\frac{x_i-\mu}{\sigma}\right)\phi^*\left(\frac{x_i-\mu}{\sigma}\right) + \frac{(1-\alpha)}{\sigma^2}\exp(-3\xi) \right.$$

$$\times \left(\frac{x_i-\mu-\Delta}{\sigma}\right)\phi^*\{\left(\frac{x_i-\mu-\Delta}{\sigma}\right)\exp(-\xi)\}]/$$

$$[\frac{\alpha}{\sigma}\phi^*\left(\frac{x_i-\mu}{\sigma}\right) + \frac{(1-\alpha)}{\sigma}\exp(-\xi)\ \phi^*\{\left(\frac{x_i-\mu-\Delta}{\sigma}\right)\exp(-\xi)\}]$$

$$\frac{\partial \ell}{\partial \sigma}(\underset{\sim}{\theta};\underset{\sim}{x}) = \sum_{i=1}^{n} \left[\frac{\alpha}{\sigma^2}\{\left(\frac{x_i-\mu}{\sigma}\right)^2-1\}\phi^*\left(\frac{x_i-\mu}{\sigma}\right) \right.$$

$$+ \frac{(1-\alpha)e^{-\xi}}{\sigma^2}\{\left(\frac{x_i-\mu-\Delta}{\sigma}\right)^2\exp(-2\xi)-1\}$$

$$\times\ \phi^*\{\left(\frac{x_i-\mu-\Delta}{\sigma}\right)\exp(-\xi)\}]/$$

$$[\frac{\alpha}{\sigma}\phi^*\left(\frac{x_i-\mu}{\sigma}\right) + \frac{(1-\alpha)}{\sigma}\exp(-\xi)\ \phi^*\{\left(\frac{x_i-\mu-\Delta}{\sigma}\right)\exp(-\xi)\}].$$

Hence under $H_{o4} : \Delta = \xi = 0$

$$\frac{\partial \ell}{\partial \Delta}(\underset{\sim}{\theta};\underset{\sim}{x}) = \frac{n}{\sigma^2}(1-\alpha)(\overline{x}-\mu) \qquad (3.6.4)$$

$$\frac{\partial \ell}{\partial \xi}(\underset{\sim}{\theta};\underset{\sim}{x}) = (1-\alpha)\{\sum_{i=1}^{n}\left(\frac{x_i-\mu}{\sigma}\right)^2-n\} \qquad (3.6.5)$$

$$\frac{\partial \ell}{\partial \alpha}(\underset{\sim}{\theta};\underset{\sim}{x}) \equiv 0.$$

We ignore all contributions associated with α. In effect we treat the problem as having two parameters under test, and two nuisance parameters.

$$\frac{\partial \ell}{\partial \mu}(\underset{\sim}{\theta};\underset{\sim}{x}) = \frac{n}{\sigma^2}(\overline{x}-\mu) \qquad (3.6.6)$$

$$\frac{\partial \ell}{\partial \sigma}(\underset{\sim}{\theta};\underset{\sim}{x}) = \frac{1}{\sigma}\{\sum_{i=1}^{n}\left(\frac{x_i-\mu}{\sigma}\right)^2-n\}. \qquad (3.6.7)$$

Note that (3.6.6) is a linear function of (3.6.4) and (3.6.7) is a linear function of (3.6.5).

The information matrix becomes

$$\underset{\sim}{c} = \begin{array}{c} \Delta \\ \xi \\ \mu \\ \sigma \end{array} \begin{bmatrix} \frac{n(1-\alpha)^2}{\sigma^2} & 0 & \frac{n(1-\alpha)}{\sigma^2} & 0 \\ 0 & 2n(1-\alpha)^2 & 0 & \frac{2n(1-\alpha)}{\sigma} \\ \frac{n(1-\alpha)}{\sigma^2} & 0 & \frac{n}{\sigma^2} & 0 \\ 0 & \frac{2n(1-\alpha)}{\sigma} & 0 & \frac{2n}{\sigma^2} \end{bmatrix} \qquad (3.6.8)$$

This leads to

$$\underset{\sim}{c}_{11}-\underset{\sim}{c}'_{12}\underset{\sim}{c}_{22}^{-1}\underset{\sim}{c}_{12} = \begin{bmatrix} \frac{n(1-\alpha)^2}{\sigma^2} & 0 \\ 0 & 2n(1-\alpha)^2 \end{bmatrix}$$

and

$$(\underset{\sim}{c}'_{12}\underset{\sim}{c}_{22}^{-1}\underset{\sim}{\phi}'_2)' = \begin{bmatrix} \frac{n(1-\alpha)}{\sigma^2}(\bar{x}-\mu) & \frac{(1-\alpha)}{\sigma}\{\sum_{i=1}^{n}(\frac{x_i-\mu}{\sigma})^2-n\} \end{bmatrix}$$

$$= \phi_1' \ .$$

Hence the optimal $C(\alpha)$ test of H_{o4} cannot be constructed.

An optimal $C(\alpha)$ test of H_{o4} can however be constructed if we perform a two sample experiment in which the first sample $Y_1,\ldots,Y_m$ say consists of independent $N(\mu,\sigma^2)$ random variables and the second sample $X_1,\ldots,X_n$, say has density (3.2.1) with μ_o and σ_o^2 treated as unknown. In this case (3.6.1), (3.6.2) and (3.6.3) are unchanged, whilst under the null hypothesis

$$\frac{\partial \ell}{\partial \mu}(\underset{\sim}{\theta};\underset{\sim}{x}) = \frac{m}{\sigma^2}(\bar{y}-\mu) + \frac{n}{\sigma^2}(\bar{x}-\mu)$$

$$\frac{\partial \ell}{\partial \sigma}(\underset{\sim}{\theta};\underset{\sim}{x}) = -\frac{m}{\sigma} - \frac{1}{\sigma}\sum_{i=1}^{m}\left(\frac{y_i-\mu}{\sigma}\right)^2 + \frac{1}{\sigma}\left\{\sum_{i=1}^{n}\left(\frac{x_i-\mu}{\sigma}\right)^2 - n\right\}.$$

(3.6.8) remains unchanged apart from the third and fourth diagonal elements, which become $\frac{(m+n)}{\sigma^2}$, $\frac{2(m+n)}{\sigma^2}$ respectively. The optimal $C(\alpha)$ test statistic for $H_{o4} : \xi = \Delta = 0$ now becomes

$$T^{(C)} = \left[\frac{(\bar{x}-\bar{y})^2}{\tilde{\sigma}^2_{m,n}} + \frac{1}{2}\left\{\frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i-\tilde{\mu}_{m,n}}{\tilde{\sigma}_{m,n}}\right)^2 - \frac{1}{m}\sum_{i=1}^{m}\left(\frac{y_i-\tilde{\mu}_{m,n}}{\tilde{\sigma}_{m,n}}\right)^2\right\}\right]^2 \frac{mn}{m+n}$$

$T^{(C)}$ is asymptotically distributed as a chi-squared random variable on 2 degrees of freedom. $\tilde{\mu}_{m,n}$ may be taken as the overall mean of the combined sample. $\tilde{\sigma}^2_{m,n}$ may be taken as a pooled estimate of variance or the sample variance of the combined sample. Any other locally root-n consistent estimators of μ and σ^2 will also suffice.

3.7 THE FAILURE OF AN OPTIMAL $C(\alpha)$ TEST FOR A FOUR PARAMETER MIXTURE

Consider the density (3.2.1) and assume μ_o unknown, $\sigma^2 = \sigma_o^2$ (σ^2 unknown). This is the mixture considered by Day (1969). The problem is now amenable to a maximum likelihood treatment.

However, the optimal $C(\alpha)$ test for a common unknown mean μ cannot be constructed for a single sample problem. We reparameterise so that $\mu_o = \mu+\Delta$ and pose the null hypothesis $H_{o5} : \Delta = 0$.

Under H_{o5} we obtain the information matrix

$$\underset{\sim}{c} = \begin{matrix} \Delta \\ \mu \\ \sigma \end{matrix} \begin{bmatrix} \frac{n(1-\alpha)^2}{\sigma^2} & \frac{n(1-\alpha)}{\sigma^2} & 0 \\ \frac{n(1-\alpha)}{\sigma^2} & \frac{n}{\sigma^2} & 0 \\ 0 & 0 & \frac{2n}{\sigma^2} \end{bmatrix}$$

leading immediately to $c_{11} - \underset{\sim}{c}'_{12} \underset{\sim}{c}_{22}^{-1} \underset{\sim}{c}_{12} \equiv 0$.

The two sample problem does lead to an optimal $C(\alpha)$ test. In addition to the sample $X_1, \ldots, X_n$ we take a sample $Y_1, \ldots, Y_m$ of independent $N(\mu, \sigma^2)$ random variables. The test statistic for $H_{o5} : \Delta = 0$ becomes

$$T^{(C)} = \frac{(\bar{x} - \bar{y})}{(\frac{1}{m} + \frac{1}{n})^{\frac{1}{2}} \tilde{\sigma}_{m,n}}$$

$T^{(C)}$ is asymptotically distributed as $N(0,1)$. If $\tilde{\sigma}^2_{m,n}$ is taken as the usual pooled estimator of variance then $T^{(C)}$ has the t distribution on $(m+n-2)$ degrees of freedom. The usual t statistic thus provides a locally asymptotically most powerful test of a normal density against a mixture alternative with common unknown variance.

3.8 MODIFIED MINIMUM CHI-SQUARED ESTIMATORS

Instead of the raw observations, we now observe whether a sample point falls into one of k cells (y_i, y_{i+1}), $i = 1, \ldots, k$, say, with observed totals $n_1, \ldots, n_k$ for each cell.

The probability $\pi_i(\underset{\sim}{\theta})$, that an observation falls in the i^{th} cell is

$$\pi_i(\underset{\sim}{\theta}) = \alpha p_i(\underset{\sim}{\theta}) + (1-\alpha)p_i^{(o)}$$

where

$$p_i(\underset{\sim}{\theta}) = \Phi(\frac{y_{i+1}-\mu}{\sigma}) - \Phi(\frac{y_i-\mu}{\sigma})$$

$$p_i^{(o)} = \Phi(\frac{y_{i+1}-\mu_o}{\sigma_o}) - \Phi(\frac{y_i-\mu_o}{\sigma_o})$$

and

$$\Phi(z) = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{z} \exp(-w^2/2)dw .$$

The modified minimum chi-squared estimator $\underset{\sim}{\theta}^{(C)}$, say, is the value of $\underset{\sim}{\theta}$ that minimises

$$\chi^2_{(C)} = \sum_{i=1}^{k} \frac{\{n_i - n\pi_i(\underset{\sim}{\theta})\}^2}{n_i} \quad , \quad n = \sum_{i=1}^{k} n_i .$$

Hence

$$\sum_{i=1}^{k} \frac{\partial}{\partial\theta_j} \pi_i\{\underset{\sim}{\theta}^{(C)}\} = \sum_{i=1}^{k} \frac{n}{n_i} \pi_i\{\underset{\sim}{\theta}^{(C)}\} \frac{\partial}{\partial\theta_j} \pi_i\{\underset{\sim}{\theta}^{(C)}\} , \quad j = 1,2,3 \tag{3.8.1}$$

Let

$$\left.\begin{aligned} q_i(\underset{\sim}{\theta}) &= \phi^*(\frac{y_{i+1}-\mu}{\sigma}) - \phi^*(\frac{y_i-\mu}{\sigma}) \\ r_i(\underset{\sim}{\theta}) &= y_{i+1}\ \phi^*(\frac{y_{i+1}-\mu}{\sigma}) - y_i\ \phi^*(\frac{y_i-\mu}{\sigma}) . \end{aligned}\right\} \tag{3.8.2}$$

We have that

$$\frac{\partial}{\partial\alpha} \pi_i(\underset{\sim}{\theta}) = p_i(\underset{\sim}{\theta}) - p_i^{(o)} \tag{3.8.3}$$

$$\frac{\partial}{\partial\mu}\pi_i(\underset{\sim}{\theta}) = -\frac{\alpha}{\sigma}q_i(\underset{\sim}{\theta}) \tag{3.8.4}$$

$$\frac{\partial}{\partial\sigma}\pi_i(\underset{\sim}{\theta}) = -\frac{\alpha}{\sigma^2}\{r_i(\underset{\sim}{\theta})-\mu q_i(\underset{\sim}{\theta})\} . \tag{3.8.5}$$

(3.8.1), (3.8.2) and (3.8.3) lead to

$$\left\{\alpha^{(C)}\right\}^{-1} = -\sum_{i=1}^{k}\frac{1}{n_i}[p_i\{\underset{\sim}{\theta}^{(C)}\}-p_i^{(o)}]^2/$$

$$\sum_{k=1}^{n} p_i^{(o)}[p_i\{\underset{\sim}{\theta}^{(C)}\}-p_i^{(o)}] .$$

(3.8.1), (3.8.2) and (3.8.5) lead to

$$\mu^{(C)} = \sum_{i=1}^{k}\frac{1}{n_i}[\alpha^{(C)}p_i\{\underset{\sim}{\theta}^{(C)}\}+\{1-\alpha^{(C)}\}p_i^{(o)}]r_i\{\underset{\sim}{\theta}^{(C)}\}/$$

$$\sum_{i=1}^{k}\frac{1}{n_i}[\alpha^{(C)}p_i\{\underset{\sim}{\theta}^{(C)}\}+\{1-\alpha^{(C)}\}p_i^{(o)}]q_i\{\underset{\sim}{\theta}^{(C)}\}.$$

From (3.8.1), (3.8.2) and (3.8.4) we have

$$0 = \sum_{i=1}^{k}\frac{1}{n_i}[\alpha^{(C)}p_i\{\underset{\sim}{\theta}^{(C)}\}+\{1-\alpha^{(C)}\}p_i^{(o)}]q_{i+1}\{\underset{\sim}{\theta}^{(C)}\} \tag{3.8.6}$$

and hence an iterative procedure for $\sigma^{(C)}$ is given by

$$\sigma_{new}^{(C)} = \sigma^{(C)}-h\{\underset{\sim}{\theta}^{(C)}\}/h'\{\underset{\sim}{\theta}^{(C)}\}$$

where $h\{\underset{\sim}{\theta}^{(C)}\}$ is the expression on the right hand side of (3.8.6) and

$$h'\{\underset{\sim}{\theta}^{(C)}\} = \sum_{i=1}^{k} \frac{1}{n_i} \frac{\alpha^{(C)}}{\{\sigma^{(C)}\}^2} [\mu^{(C)} q_i\{\underset{\sim}{\theta}^{(C)}\} - r_i\{\underset{\sim}{\theta}^{(C)}\}] q_i\{\underset{\sim}{\theta}^{(C)}\}$$

$$- \sum_{i=1}^{k} \frac{2}{n_i\{\sigma^{(C)}\}^3} [\alpha^{(C)} p_i\{\underset{\sim}{\theta}^{(C)}\} + \{1-\alpha^{(C)}\} p_i^{(o)}]$$

$$\times [s_i\{\underset{\sim}{\theta}^{(C)}\} - 2\mu^{(C)} r_i\{\underset{\sim}{\theta}^{(C)}\} + \{\mu^{(C)}\}^2 q_i\{\underset{\sim}{\theta}^{(C)}\}] ,$$

where

$$s_i(\underset{\sim}{\theta}) = y_{i+1}^2 \ \phi^*(\frac{y_{i+1}-\mu}{\sigma}) - y_i^2 \ \phi^*(\frac{y_i-\mu}{\sigma}) .$$

4. SOME MORE COMPARISMS OF OPTIMAL $C(\alpha)$ TESTS AND TESTS BASED ON MAXIMUM LIKELIHOOD ESTIMATORS

4.1 SIMPLE LINEAR REGRESSION WITH CAUCHY ERRORS

The problem of estimating the location parameter for the Cauchy distribution

$$f(y;\beta_o,\xi) = \frac{\xi}{\pi\{\xi^2+(y-\beta_o)^2\}}$$

$-\infty < y < \infty$, $-\infty < \beta_o < \infty$, $\xi > 0$ has been discussed in detail in Barnett (1966) and Haas, Bain and Antle (1970). Barnett concludes that

> "the extensive calculations needed to ensure the correct evaluation of the maximum likelihood estimator of the location parameter of a Cauchy distribution, and the poor efficiency of (this) estimator for small samples render the utility of the maximum likelihood method rather doubtful in this context".

In this section we consider a test for simple regression with Cauchy errors. The density of the i^{th} member of a sample of independent observations (x_i,y_i), $i = 1,\ldots,n$ is given by

$$f_{Y_i}(y;\underset{\sim}{\theta}) = \frac{\xi}{\pi\{\xi^2+(y-\beta_o-\beta_1x_i)^2\}} \qquad i = 1,\ldots,n \tag{4.1.1}$$

where $\underset{\sim}{\theta}' = [\beta_o\beta_1\xi]$.

Under the hypothesis $H_o : \beta_1 = 0$ the problem of estimating the nuisance parameters is exactly the problem considered above. In view of the above remarks and Barnett's further conclusions that

> "fuller use of order statistics in the estimation procedure will inevitably reduce the advantage of the maximum likelihood estimator for little increase in effort"

we will use estimators based on the observed quartiles of the sample. The derivatives of the log likelihood of the sample are given by

$$\frac{\partial \ell}{\partial \beta_1}(\underset{\sim}{\theta};\underset{\sim}{y}) = \sum_{i=1}^{n} \frac{2(y_i-\beta_o-\beta_1 x_i)x_i}{\xi^2+(y_i-\beta_o-\beta_1 x_i)^2}$$

$$\frac{\partial \ell}{\partial \beta_o}(\underset{\sim}{\theta};\underset{\sim}{y}) = \sum_{i=1}^{n} \frac{2(y_i-\beta_o-\beta_1 x_i)}{\xi^2+(y_i-\beta_o-\beta_1 x_i)^2} \qquad (4.1.2)$$

$$\frac{\partial \ell}{\partial \xi}(\underset{\sim}{\theta};\underset{\sim}{y}) = \sum_{i=1}^{n} \frac{-2\xi}{\xi^2+(y_i-\beta_o-\beta_1 x_i)^2} + \frac{n}{\xi} . \qquad (4.1.3)$$

The information matrix is

$$\underset{\sim}{c} = \begin{array}{c} \beta_1 \\ \beta_o \\ \xi \end{array} \begin{bmatrix} \frac{1}{2\xi^2}\sum_{i=1}^{n} x_i^2 & \frac{n\bar{x}}{2\xi^2} & 0 \\ \frac{n\bar{x}}{2\xi^2} & \frac{n}{2\xi^2} & 0 \\ 0 & 0 & \frac{n}{2\xi^2} \end{bmatrix} .$$

Hence

$$\underset{\sim}{c}'_{12}\underset{\sim}{c}_{22}^{-1} = [\bar{x}\ 0] , \quad c_{11}-\underset{\sim}{c}'_{12}\underset{\sim}{c}_{22}^{-1}\underset{\sim}{c}_{12} = \frac{s_{xx}}{2\xi^2} , \quad \text{where } s_{xx} = \sum_{i=1}^{n}(x_i-\bar{x})^2 .$$

The optimal $C(\alpha)$ test of $H_o : \beta_1 = 0$ is based upon

$$T^{(C)} = \left\{ \frac{\partial \ell}{\partial \beta_1}(\underset{\sim}{\theta};\underset{\sim}{y}) - \bar{x}\frac{\partial \ell}{\partial \beta_o}(\underset{\sim}{\theta};\underset{\sim}{y}) \right\}\left(\frac{s_{xx}}{2\xi^2}\right)^{-\frac{1}{2}} \Bigg|_{\substack{\beta_1=0;\ \beta_o=\tilde{\beta}_o \\ \xi=\tilde{\xi}}}$$

$\tilde{\beta}_0$, $\tilde{\xi}$ are locally root n consistent estimators of β_0, ξ respectively. The maximum likelihood estimators of α, ξ, found from (4.1.2) and (4.1.3) in the usual manner with $\beta_1 \equiv 0$, will not be used. Consider

$$\tilde{\beta}_o = \frac{1}{2}\{y_{(n,\frac{3n}{4})} + y_{(n,\frac{n}{4})}\} \tag{4.1.4}$$

$$\tilde{\xi} = \frac{1}{2}\{y_{(n,\frac{3n}{4})} - y_{(n,\frac{n}{4})}\} \tag{4.1.5}$$

where $y_{(n,r)}$ denotes the r^{th} order statistic from a sample of n observations. It is easily shown that (4.1.4) and (4.1.5) are locally root-n consistent estimators.

We now have that

$$T^{(C)} = \{\sum_{i=1}^{n} \frac{2(y_i-\tilde{\beta}_o)x_i}{\tilde{\xi}^2+(y_i-\tilde{\beta}_o)^2} - \bar{x}\sum_{i=1}^{n} \frac{(y_i-\tilde{\beta}_o)}{\tilde{\xi}^2+(y_i-\tilde{\beta}_o)^2}\}/(\frac{s_{xx}}{2\tilde{\xi}^2})^{\frac{1}{2}}$$

$$= \frac{2^{3/2}\tilde{\xi}}{s_{xx}^{\frac{1}{2}}}\sum_{i=1}^{n} \frac{(x_i-\bar{x})(y_i-\tilde{\beta}_o)}{\tilde{\xi}^2+(y_i-\tilde{\beta}_o)^2} . \tag{4.1.6}$$

$T^{(C)}$ is asymptotically distributed as $N(0,1)$. Note that (4.1.6) requires no iterative technique, save sorting the data, in its derivation. In appendix A.3, values of this test statistic are given for various sample sizes with $\beta_o = 0$, $\xi = 1$, $x_i \in [-q,q]$ for various q and $\beta_1 = 0(0{\cdot}1)0{\cdot}5$.

When H_o is true, the test statistic, treated as a normal random variable, leads to acceptance of the null hypothesis for sample sizes greater than 500, and in 90% of the smaller sample sizes considered. This was so even if the estimators (4.1.4) and (4.1.5) yielded poor results. With the true value of $\beta_1 = 0{\cdot}1$, a test based on $T^{(C)}$ continually rejected $H_o : \beta_1 = 0$ for sample sizes of 11 upwards, with highly non-significant values of $T^{(C)}$ being generated for sample sizes greater than 51. Similar results were found for larger true values of β_1, the test in all cases leading to the rejection of H_o for sample sizes greater than 31. It would appear that the test statistic (4.1.6) is quite efficient in detecting regression and is robust to the inadequacies of (4.1.4) and (4.1.5) as estimators.

If we use values of $\{x_i\}$ such that $x_i = -q+i-1$, $i = 1,\ldots,2q+1$, then we have that

$$E(\tilde{\beta}_o;\beta_o,\beta_1) = \beta_o + \frac{\beta_1}{2}\{x_{(n,\frac{3n}{4})}+x_{(n,\frac{n}{4})}\} + \frac{\xi}{2}[E\{Z_{(n,\frac{3n}{4})}+Z_{(n,\frac{n}{4})}\}]$$

and

$$E(\tilde{\beta}_1;\beta_o,\beta_1) = \frac{\beta_1}{2}\{x_{(n,\frac{3n}{4})}-x_{(n,\frac{n}{4})}\} + \frac{\xi}{2}[E\{Z_{(n,\frac{3n}{4})}-Z_{(n,\frac{n}{4})}\}] ,$$

where $Z_1,\ldots,Z_n$ represents a random sample with density (4.1.1) with $\beta_o = \beta_1 = 0$; $\xi = 1$. Expectations are taken with respect to the true values of the parameters β_o and β_1.

We then have

$$E(\tilde{\beta}_o;\beta_o,\beta_1) = \beta_o$$

$$E(\tilde{\beta}_1;\beta_o,\beta_1) = \frac{\beta_1}{2} + \xi h(q)$$

where

$$h(q) = \frac{\Gamma\{2(q+1)\}}{\Gamma(\frac{q+1}{2})\Gamma\{\frac{3}{2}(q+1)\}}\int_o^1 u^{\frac{q-1}{2}}(1-u)^{\frac{3q+1}{2}}\tan\{\pi(u-\tfrac{1}{2})\}du .$$

Hence $\tilde{\beta}_o$ is unbiased for β_o, whilst $\tilde{\beta}_1$ has bias $b(\beta_1)$ given by

$$b(\beta_1) = \xi\{h(q)-1\} + \frac{\beta_1}{2}(q+1) \rightarrow \frac{\beta_1}{2}(q+1) \quad \text{as} \quad q \rightarrow \infty \,.$$

This is clearly demonstrated in appendix A.3.

Several multiple regression problems may be treated by the method of optimal $C(\alpha)$ tests leading to similar intuitively reasonable test statistics. This is pursued in sections (4.2), (4.3) and (4.4).

4.2 MULTIPLE REGRESSION WITH CAUCHY ERRORS

The example of section (4.1) may be easily extended to construct a test of regression with several independent variables. The density of the i^{th} observation y_i may be taken as

$$f_{Y_i}(y_i;\underset{\sim}{\theta}) = \frac{\xi}{\pi\{\xi^2+(y_i-\beta_o-\sum_{j=1}^{p}\beta_j x_{ij})^2\}}\,, \qquad i = 1,\ldots,n.$$

where $\underset{\sim}{\theta}' = [\beta_1 \ldots . \beta_p \; \beta_o \; \xi]$.

The derivatives of the log likelihood are

$$\frac{\partial \ell}{\partial \beta_k}(\underset{\sim}{\theta};\underset{\sim}{y}) = 2\sum_{i=1}^{n} \frac{(y_i-\beta_o-\sum_{j=1}^{p}\beta_j x_{ij})x_{ik}}{\xi^2+(y_i-\beta_o-\sum_{j=1}^{p}\beta_j x_{ij})^2}\,, \qquad k = 1,\ldots,p.$$

$$\frac{\partial \ell}{\partial \beta_o}(\underset{\sim}{\theta};\underset{\sim}{y}) = 2\sum_{i=1}^{n} \frac{y_i-\beta_o-\sum_{j=1}^{p}\beta_j x_{ij}}{\xi^2+(y_i-\beta_o-\sum_{j=1}^{p}\beta_j x_{ij})^2}$$

$$\frac{\partial \ell}{\partial \xi}(\underset{\sim}{\theta};\underset{\sim}{y}) = \frac{n}{\xi} - \sum_{i=1}^{n} \frac{2\xi}{\xi^2+(y_i-\beta_o-\sum_{j=1}^{p}\beta_j x_{ij})^2}\,.$$

The information matrix, $\underset{\sim}{c}$, is given by

$$\underset{\sim}{c} = \begin{matrix} \beta_1 \\ \vdots \\ \beta_p \\ \beta_o \\ \xi \end{matrix} \left[\begin{array}{c|cc} (\sum_{i=1}^{n} x_{ik}x_{i\ell}) & n\underset{\sim}{\bar{x}} & 0 \\ \hline n\underset{\sim}{\bar{x}}' & n & 0 \\ 0 & 0 & \frac{n}{2} \end{array} \right] \times \frac{1}{2\xi^2}$$

where $\underset{\sim}{\bar{x}}' = [\bar{x}_{.1} \ldots \bar{x}_{.p}]$, $\bar{x}_{.k} = \frac{1}{n} \sum_{i=1}^{n} x_{ik}$, $k = 1,\ldots,p$.

Hence

$$\underset{\sim}{c}_{11} - \underset{\sim}{c}'_{12}\underset{\sim}{c}_{22}^{-1}\underset{\sim}{c}_{12} = \frac{1}{2\xi^2}\underset{\sim}{s}_{xx} , \quad \underset{\sim}{s}_{xx} = (\sum_{i=1}^{n} (x_{ik}-\bar{x}_{.k})(x_{i\ell}-\bar{x}_{.\ell}))$$

and the test statistic $T^{(C)}$ for a test of $H_o : \beta_k = 0, k = 1,\ldots,p$ is given by

$$T^{(C)} = 4\tilde{\xi}^2 \underset{\sim}{s}'_{xy}\underset{\sim}{s}_{xx}^{-1}\underset{\sim}{s}_{xy} \tag{4.2.1}$$

where

$$\underset{\sim}{s}'_{xy} = [\sum_{i=1}^{n} \frac{(x_{ik}-\bar{x}_{.k})(y_i-\tilde{\beta}_o)}{\tilde{\xi}^2+(y_i-\tilde{\beta}_o)^2} \ldots \sum_{i=1}^{n} \frac{(x_{ip}-\bar{x}_{.p})(y_i-\tilde{\beta}_o)}{\tilde{\xi}^2+(y_i-\tilde{\beta}_o)^2}] .$$

$T^{(C)}$ is asymptotically distributed as a chi-squared random variable on p degrees of freedom. $\tilde{\beta}_o$ and $\tilde{\xi}$ are given by (4.1.4) and (4.1.5).

4.3 TESTS FOR THE SIGNIFICANCE OF A REGRESSION WITH POISSON AND BINARY DATA

We have a sample of independent observations $y_1,\ldots,y_n$, the i^{th} observation y_i having a Poisson distribution with mean μ_i, where

$$\mu_i = \exp(\beta_o + \sum_{j=1}^{p} x_{ij}\beta_j) \quad , \quad i = 1,\ldots,n.$$

The log likelihood of the sample is

$$\ell(\underset{\sim}{\beta};\underset{\sim}{y}) = n\beta_o\bar{y} + \sum_{j=1}^{p}\beta_j\sum_{i=1}^{n}x_{ij}y_i - \sum_{i=1}^{n}\exp(\beta_o + \sum_{j=1}^{p}x_{ij}\beta_j)$$

where $\bar{y} = n^{-1}\sum_{i=1}^{n}y_i$; $\underset{\sim}{\beta}' = [\beta_1 \ldots \beta_p\ \beta_o]$.

Methods of testing a hypothesis $H_{o1} : \beta_j = 0,\ j = 1,\ldots,p$ have been discussed in Cox (1966). We now construct an optimal $C(\alpha)$ test of H_{o1}. We have

$$\frac{\partial\ell}{\partial\beta_k}(\underset{\sim}{\beta};\underset{\sim}{y}) = \sum_{i=1}^{n}x_{ik}y_i - \exp(\beta_o)\sum_{i=1}^{n}x_{ik}\exp(\sum_{j=1}^{p}x_{ij}\beta_j)\ , \quad k = 1,\ldots,p.$$

$$\frac{\partial\ell}{\partial\beta_o}(\underset{\sim}{\beta};\underset{\sim}{y}) = n\bar{y} - \exp(\beta_o)\sum_{i=1}^{n}\exp(\sum_{j=1}^{p}x_{ij}\beta_j).$$

Hence, under H_{o1};

$$\frac{\partial\ell}{\partial\beta_k}(\underset{\sim}{\beta};\underset{\sim}{y}) = \sum_{i=1}^{n}x_{ik}y_i - n\bar{x}_{.k}\exp(\beta_o)\ , \quad k = 1,\ldots,n.$$

$$\frac{\partial\ell}{\partial\beta_o}(\underset{\sim}{\beta};\underset{\sim}{y}) = n\{\bar{y} - \exp(\beta_o)\}.$$

The information matrix, $\underset{\sim}{c}$, is given under H_{o1} by

$$\underset{\sim}{c} = \begin{matrix}\beta_1\\ \beta_p\\ \beta_o\end{matrix}\left[\begin{array}{c|c} (\sum_{i=1}^{n}x_{ik}x_{i\ell}) & n\bar{\underset{\sim}{x}}_{.} \\ \hline n\bar{\underset{\sim}{x}}'_{.} & n\end{array}\right] \times \exp(\beta_o)\ .$$

Hence $\underset{\sim}{c}_{11} - \underset{\sim}{c}'_{12}\underset{\sim}{c}_{22}^{-1}\underset{\sim}{c}_{12} = \exp(\beta_o)\underset{\sim}{s}_{xx}$, where $\underset{\sim}{s}_{xx}$ and $\bar{\underset{\sim}{x}}_{.}$ are as defined in section (4.2).

Also

$$\phi_1 - c_{12}' c_{22}^{-1} \phi_2 = s_{xy} \ , \qquad (s_{xy})_k = \sum_{i=1}^{n} (x_{ik} - \bar{x}_{.k})(y_i - \bar{y}) .$$

Hence the optimal $C(\alpha)$ test statistic of $H_{o1} : \beta_j = 0, \ j = 1,\ldots,p$ is given by

$$T^{(C)} = \frac{s_{xy}' S_{xx}^{-1} s_{xy}}{\bar{y}} \tag{4.3.1}$$

as $\bar{y}$ provides a locally root-n consistent estimate of e^{β_o}. $T^{(C)}$ is asymptotically distributed as a chi-squared random variable on p degrees of freedom. In particular, when $p = 1$ we have that

$$T^{(C)} = \{\sum_{i=1}^{n} (x_{i1} - \bar{x}_{.1})(y_i - \bar{y})\}^2 / \{\bar{y} \Sigma (x_{i1} - \bar{x}_{.1})^2\}$$

whose square-root tends to the test statistic for simple linear regression with Poisson data considered in Cox (1966).

A similar problem with discrete data is to construct a test for regression with binary data. We take a sample $y_1,\ldots,y_m$ of independent observations, the i^{th} observation having the binomial distribution with parameters n_i, θ_i, where

$$\log\{\theta_i/(1-\theta_i)\} = \beta_o + \sum_{j=1}^{p} x_{ij}\beta_j \ , \qquad i = 1,\ldots,m.$$

Methods of testing the hypothesis $H_{o2} : \beta_j = 0, \ j = 1,\ldots,p$ have been discussed in Cox (1970). We now construct the optimal $C(\alpha)$ test of H_{o2}.

The log likelihood of the sample is given by

$$\ell(\beta;y) = m\beta_o\bar{y} + \sum_{j=1}^{p} \beta_j \sum_{i=1}^{m} x_{ij} y_i$$

$$- \sum_{i=1}^{m} n_i \log\{1+\exp(\beta_o + \sum_{j=1}^{p} x_{ij}\beta_j)\}$$

where $\underset{\sim}{\beta}' = [\beta_1 \ldots \beta_p\ \beta_o]$. We have

$$\frac{\partial \ell}{\partial \beta_k}(\underset{\sim}{\beta};\underset{\sim}{y}) = \sum_{i=1}^{m} x_{ik}y_i - \sum_{i=1}^{m} \frac{n_i x_{ik} \exp(\beta_o + \sum_{j=1}^{p} x_{ij}\beta_j)}{1+\exp(\beta_o + \sum_{j=1}^{p} x_{ij}\beta_j)}$$

$$\frac{\partial \ell}{\partial \beta_o}(\underset{\sim}{\beta};\underset{\sim}{y}) = m\bar{y} - \sum_{i=1}^{m} \frac{n_i \exp(\beta_o + \sum_{j=1}^{p} x_{ij}\beta_j)}{1+\exp(\beta_o + \sum_{j=1}^{p} x_{ij}\beta_j)} .$$

Hence, under H_{o2},

$$\frac{\partial \ell}{\partial \beta_k}(\underset{\sim}{\beta};\underset{\sim}{y}) = \sum_{i=1}^{m} x_{ik}y_i - \frac{\exp(\beta_o)}{1+\exp(\beta_o)} \sum_{i=1}^{m} n_i x_{ik}$$

$$\frac{\partial \ell}{\partial \beta_o}(\underset{\sim}{\beta};\underset{\sim}{y}) = m\bar{y} - \frac{n_{.}\exp(\beta_o)}{1+\exp(\beta_o)} , \quad n_{.} = \sum_{i=1}^{m} n_i .$$

The information matrix, $\underset{\sim}{c}$, is given under H_{o2} by

$$\underset{\sim}{c} = \begin{matrix} \beta_1 \\ \beta_p \\ \\ \beta_o \end{matrix} \left[\begin{array}{c|c} (\sum_{i=1}^{m} n_i x_{ik} x_{i\ell}) & (\sum_{i=1}^{m} n_i x_{ik}) \\ \hline (\sum_{i=1}^{m} n_i x_{ik})' & n_{.} \end{array} \right] \frac{e^{\beta_o}}{(1+e^{\beta_o})^2}$$

Hence

$$\underset{\sim}{c}_{11} - \underset{\sim}{c}'_{12}\underset{\sim}{c}_{22}^{-1}\underset{\sim}{c}_{12} = \frac{\exp(\beta_o)}{\{1+\exp(\beta_o)\}^2} \underset{\sim}{s}_{xx} ,$$

where

$$(\underset{\sim}{s}_{xx})_{k\ell} = \sum_{i=1}^{m} n_i(x_{ik}-\bar{x}_{.k})(x_{i\ell}-\bar{x}_{.\ell}) \quad , \quad \bar{x}_{.k} = \frac{1}{n_{.}} \sum_{i=1}^{m} n_i x_{ik}$$

$$\underset{\sim}{\phi}_1 - \underset{\sim}{c}'_{12}\underset{\sim}{c}_{22}^{-1}\underset{\sim}{\phi}_2 = \underset{\sim}{s}_{xy} \ , \quad (\underset{\sim}{s}_{xy})_k = \sum_{i=1}^{m} n_i(x_{ik}-\bar{x}_{.k})(y_i-\bar{y}) .$$

The optimal $C(\alpha)$ test statistic of $H_{o2} : \beta_j = 0, j = 1,\ldots,p,$ is given by

$$T^{(C)} = \frac{1}{\bar{y}(1-\bar{y})} \underset{\sim}{s}'_{xy}\underset{\sim}{s}_{xx}^{-1}\underset{\sim}{s}_{xy} \ . \tag{4.3.2}$$

$T^{(C)}$ is asymptotically distributed as a chi-squared random variable on p degrees of freedom.

Note that (4.2.1), (4.3.1) and (4.3.2) although differing in detail, have the same general form. There is a severe drawback to these methods in that we must have a genuine regression situation. If the $\{x_{ij}\}$ are such that $\underset{\sim}{c}$ is singular the method will fail as the model is underidentified. This will be the case where we wish to test for the absence of some effect in the usual analysis of variance framework.

4.4 THE TWO SAMPLE PROBLEM FOR THE GAMMA DISTRIBUTION

We have two independent random samples with gamma densities and construct tests of the hypothesis that the densities differ in some respect.

A suitable parameterisation is to let $x_1,\ldots,x_m$ be a random sample of $\Gamma(\alpha,\beta)$ random variables, with density

$$f(x) = \frac{\alpha^{\beta}x^{\beta-1}e^{-\alpha x}}{\Gamma(\beta)} \ , \quad x > 0 \ , \quad \alpha,\beta > 0.$$

We take a second sample $y_1,\ldots,y_n$ of $\Gamma(\alpha e^{\xi},\beta+\Delta)$ random variables. The problem of testing $H_o^{(o)} : \Delta = 0$ when $\xi \equiv 0$ has been considered in Moran (1970b). We consider a null hypothesis $H_o : \Delta = \xi = 0$.

The log likelihood of the observations is

$$\ell(\underset{\sim}{\theta};\underset{\sim}{x}) = m\beta\log\alpha+(\beta-1)\sum_{i=1}^{m}\log x_i-m\alpha\bar{x}-m\log\Gamma(\beta)$$

$$+n(\beta+\Delta)\log\alpha+n(\beta+\Delta)\xi+(\beta+\Delta-1)\sum_{i=1}^{n}\log y_i$$

$$-n\alpha e^{\xi}\bar{y}-n\log\Gamma(\beta+\Delta)$$

where $\underset{\sim}{\theta}' = [\Delta\ \xi\ \beta\ \alpha]$, $\underset{\sim}{x}' = [x_1\ldots x_m y_1\ldots y_n]$.

$$\frac{\partial\ell}{\partial\Delta}(\underset{\sim}{\theta};\underset{\sim}{x}) = n\log\alpha+n\xi+\sum_{i=1}^{n}\log y_i-n\psi(\beta+\Delta) \tag{4.4.1}$$

$$\frac{\partial\ell}{\partial\xi}(\underset{\sim}{\theta};\underset{\sim}{x}) = n(\beta+\Delta)-n\alpha\bar{y}\exp(\xi) \tag{4.4.2}$$

$$\frac{\partial\ell}{\partial\beta}(\underset{\sim}{\theta};\underset{\sim}{x}) = \frac{\partial\ell}{\partial\Delta}(\underset{\sim}{\theta};\underset{\sim}{x}) + m\log\alpha+\sum_{i=1}^{m}\log x-m\psi(\beta) \tag{4.4.3}$$

$$\frac{\partial\ell}{\partial\alpha}(\underset{\sim}{\theta};\underset{\sim}{x}) = \frac{1}{\alpha}\frac{\partial\ell}{\partial\xi}(\underset{\sim}{\theta};\underset{\sim}{x}) + \frac{m\beta}{\alpha} - m\bar{x} \tag{4.4.4}$$

where

$$\psi(z) = \frac{d}{dz}\log\Gamma(z)$$

is the digamma function.

The full maximum likelihood solution is, from (4.4.1)-(4.4.4)

$$\psi(\hat{\beta})-\log\hat{\beta} = \frac{1}{m}\sum_{i=1}^{m}\log x_i-\log\bar{x} \tag{4.4.5}$$

$$\hat{\alpha} = \frac{\hat{\beta}}{\bar{x}} \tag{4.4.6}$$

$$\psi(\hat{\beta}+\hat{\Delta})-\log(\hat{\beta}+\hat{\Delta}) = \frac{1}{n}\sum_{i=1}^{n}\log y_i-\log\bar{y} \tag{4.4.7}$$

$$\hat{\xi} = \log\left(\frac{\hat{\beta}+\hat{\Delta}}{\hat{\alpha}\bar{y}}\right). \tag{4.4.8}$$

The information matrix, $\underset{\sim}{c}$, is given under H_o by

$$\underset{\sim}{c} = \begin{array}{c} \Delta \\ \xi \\ \beta \\ \alpha \end{array}\begin{bmatrix} n\psi'(\beta) & -n & n\psi'(\beta) & \frac{-n}{\alpha} \\ -n & n\beta & -n & \frac{n\beta}{\alpha} \\ n\psi'(\beta) & -n & (m+n)\psi'(\beta) & \frac{-(m+n)}{\alpha} \\ \frac{-n}{\alpha} & \frac{n\beta}{\alpha} & \frac{-(m+n)}{\alpha} & \frac{(m+n)\beta}{\alpha^2} \end{bmatrix}.$$

The Wald statistic, $T^{(W)}$, for testing $H_o : \xi = \Delta = 0$ is given by

$$T^{(W)} = \frac{1}{(\frac{1}{m}+\frac{1}{n})}\{\hat{\Delta}^2\psi'(\hat{\beta})-2\hat{\xi}\hat{\Delta}+\hat{\beta}\hat{\xi}^2\}. \tag{4.4.9}$$

The optimal $C(\alpha)$ test statistic is given by

$$T^{(C)} = \frac{1}{(\frac{1}{m}+\frac{1}{n})}\{\hat{\hat{\beta}}(Z_y-Z_x)^2-2\hat{\hat{\alpha}}(\bar{y}-\bar{x})(Z_y-Z_x) + \psi'(\hat{\hat{\beta}})\hat{\hat{\alpha}}^2(\bar{x}-\bar{y})^2\} \times \frac{1}{\{\hat{\hat{\beta}}\psi'(\hat{\hat{\beta}})-1\}} \tag{4.4.10}$$

where

$$Z_y = \frac{1}{m}\sum_{i=1}^{n}\log y_i \quad\text{and}\quad Z_x = \frac{1}{n}\sum_{i=1}^{m}\log x_i.$$

The estimators $\hat{\hat{\alpha}}$, $\hat{\hat{\beta}}$ in (4.4.10) need only be locally root-n consistent. In particular we may take $\hat{\hat{\alpha}}$, $\hat{\hat{\beta}}$ to be the maximum likelihood estimators of α and β under H_o. These estimators satisfy

$$\psi(\hat{\hat{\beta}})-\log\hat{\hat{\beta}} = T_1-\log T_2 \tag{4.4.11}$$

$$\hat{\hat{\alpha}} = \hat{\hat{\beta}}/T_2 \tag{4.4.12}$$

where

$$T_1 = \frac{1}{m+n}\left(\sum_{i=1}^{m} \log x_i + \sum_{i=1}^{n} \log y_i\right)$$

$$T_2 = \frac{1}{m+n}\left(\sum_{i=1}^{m} x_i + \sum_{i=1}^{n} y_i\right).$$

The equations (4.4.5), (4.4.7) and (4.4.11) need initial values to start an iterative procedure leading to their solution.

The moment estimators of β and Δ in the full maximum likelihood framework are

$$\tilde{\beta} = \frac{\bar{x}^2}{s_{xx}}, \qquad s_{xx} = \frac{1}{m}\sum_{i=1}^{m} (x_i-\bar{x})^2$$

$$\tilde{\Delta} = \frac{\bar{y}^2}{s_{yy}} - \tilde{\beta}, \qquad s_{yy} = \frac{1}{n}\sum_{i=1}^{n} (y_i-\bar{y})^2.$$

Under H_o, we have as moment estimator for β

$$\tilde{\tilde{\beta}} = \frac{T_2^2}{s_p}, \qquad s_p = \frac{1}{m+n}\left\{\sum_{i=1}^{m} (x_i-T_2)^2 + \sum_{i=1}^{n} (y_i-T_2)^2\right\}.$$

Both $T^{(C)}$ and $T^{(W)}$ are asymptotically distributed as a chi-squared random variable on 2 degrees of freedom.

In appendix A.4, the results of simulations on the model of this section are given. The solutions of (4.4.5)-(4.4.8) and test statistic (4.4.9) are given along with the solutions of (4.4.11), (4.4.12) and the optimal $C(\alpha)$ test statistic (4.4.10).

The simulations were performed with $m = n$, $n = 20(20)100(100)300$, 500; $\alpha = 1$; $\beta = 0\cdot5, 1\cdot0$ and $6\cdot0$; $\xi = -0\cdot1(0\cdot1)0\cdot1$; $\Delta = -0\cdot1(0\cdot1)0\cdot1$.

For $\Delta \neq 0$ the gamma distributed random variables were generated using either the results of Stuart (1962) or Bankövi (1964) and Jöhnck (1964). Stuart's result is that if $X \sim \Gamma(1,p)$ and $Y \sim B(q,p-q)$ independently of X, where $B(r,s)$ is the beta distribution, then $Z = XY$ is distributed as $\Gamma(1,q)$. Bankövi and Jöhnck show that

$$\Gamma(1,\beta) \sim \Gamma(1,[\beta])+\Gamma(1,1)B(\beta-[\beta],[\beta]+1-\beta)$$

where the symbols are taken to represent random variables whose distributions have the given parameters. The beta random variables are generated by the usual rejection technique, see for example, Newman and Odell (1971, pp. 30-31).

The results indicate that point estimation of α and β is good in moderate sample sizes, less than 100, say. For larger sample sizes estimation improves considerably. Estimation of ξ and Δ is poor throughout, more precision being gained with larger sample size, but for the true value $\beta = 6$, the estimation of β is imprecise, leading to grosser errors in the estimators of ξ and Δ. Larger sample sizes still may produce markedly better estimates, but the computer time involved in generating the gamma random variables with non integral β becomes excessive.

Treating the test statistics as observations on a chi-squared random variable with two degrees of freedom, with 95% confidence interval $(0\cdot0506, 7\cdot377)$, neither the optimal $C(\alpha)$ test or the Wald statistic leads to rejection of the null hypothesis when $H_0 : \xi = \Delta = 0$ is actually true. For the (ξ,Δ) pair $(-0\cdot1, -0\cdot1)$ the optimal $C(\alpha)$ test rejected H_o for large n when the true value of β was either $\frac{1}{2}$ or 6. The Wald statistic did not lead to rejection of H_o. For the (ξ,Δ) pair (-0·1, 0) the optimal $C(\alpha)$ test failed to reject H_o whilst the maximum likelihood test did so in half the cases when $\beta = 6$. With the (ξ,Δ) pair (-0·1, 0·1) both tests lead to rejection of H_o for $\beta = \frac{1}{2}$ and $\beta = 1$ with large sample sizes. Only the maximum likelihood test lead to rejection of H_o when $\beta = 6$. For the (ξ,Δ) pair (0·0, -0·1) the maximum likelihood test lead to rejection of H_o only when $n = 500$. For the remaining samplesand with $\beta = \frac{1}{2}$ or 1 both tests failed to reject H_o. When $\beta = 6$ the maximum likelihood test lead to consistent rejection of H_o whilst the optimal $C(\alpha)$ test failed to do so.

For the (ξ,Δ) pair (0·0, 0·1) both tests lead to rejection of H_o for sample sizes greater than 100 with $\beta = \frac{1}{2}$ or 1. For $\beta = 6$ the maximum likelihood test consistently rejected H_o whilst the optimal $C(\alpha)$ test failed to do so.

For the (ξ,Δ) pair (0·1, 0·0) the maximum likelihood test rejected H_o with a sample size of 500 and β either 1 or 6. Otherwise both tests failed to reject H_o.

For the (ξ,Δ) pair (0·1, 0·1) both tests lead to rejection of H_o for sample sizes greater than 80 when $\beta = \frac{1}{2}$ or 1. The maximum likelihood test lead to rejection of H_o for large sample sizes when $\beta = 6$, whilst the optimal $C(\alpha)$ test did not reject H_o for any of the sample sizes used.

On the whole, maximum likelihood appears the better procedure, though it would appear that sample sizes larger than 500 (in each of the two samples) are needed for the test to work consistently if H_o is false. If the null hypothesis is true both tests lead to acceptance of H_o.

For the values $\beta = \frac{1}{2}$ and 1 the two test statistics approximate to each other, as is foreshadowed in the asymptotic result of section (2.3). However for $\beta = 6$ this was not so for the sample sizes considered.

4.5 A GENERAL TESTING SITUATION GIVING RISE TO AN OPTIMAL $C(\alpha)$ TEST

In this section we consider a test of a null hypothesis specifying that $X_1,\ldots,X_n$ are independently distributed random variables with probability density $f_i(x;\underset{\sim}{\theta})$, $i = 1,\ldots,n$; $\underset{\sim}{\theta}' = [\theta_1\ldots\theta_s]$ unknown. The alternative hypothesis specifies that the density takes the form

$$f_i(x;\psi,\underset{\sim}{\theta}) = \frac{f_i(x;\underset{\sim}{\theta})\exp\{\psi g_i(x;\underset{\sim}{\theta})\}}{h_i(\psi;\underset{\sim}{\theta})} \tag{4.5.1}$$

where

$$h_i(\psi;\underset{\sim}{\theta}) = \int f_i(u;\underset{\sim}{\theta})\exp\{\psi g_i(u;\underset{\sim}{\theta})\}du \ .$$

In this framework the hypothesis tested is $H_o : \psi = 0$. The log likelihood of the sample $X_1,\ldots,X_n$ is given by

$$\ell(\psi;\underset{\sim}{\theta}) = \sum_{i=1}^{n} \log f_i(X_i;\underset{\sim}{\theta}) + \psi \sum_{i=1}^{n} g_i(X_i;\underset{\sim}{\theta})$$

$$- \sum_{i=1}^{n} \log h_i(\psi;\underset{\sim}{\theta}) \ ,$$

with derivatives

$$\frac{\partial\ell}{\partial\psi}(\psi;\underset{\sim}{\theta}) = \sum_{i=1}^{n} g_i(X_i;\underset{\sim}{\theta}) - \sum_{i=1}^{n} \left[\int g_i(u;\underset{\sim}{\theta}) f_i(u;\underset{\sim}{\theta})\right.$$

$$\left. \times \exp\{\psi g_i(u;\underset{\sim}{\theta})\}du/h_i(\psi;\underset{\sim}{\theta})\right]$$

and hence

$$\frac{\partial \ell}{\partial \psi}(0;\underset{\sim}{\theta}) = \sum_{i=1}^{n} [g_i(X_i;\underset{\sim}{\theta}) - E_o\{g_i(X_i;\underset{\sim}{\theta})\}] \tag{4.5.2}$$

where E_o denotes expectation taken with respect to the density $f_i(x;\underset{\sim}{\theta})$.

$$\frac{\partial \ell}{\partial \theta_j}(\psi;\underset{\sim}{\theta}) = \sum_{i=1}^{n} \frac{\partial}{\partial \theta_j} \log f_i(X_i;\underset{\sim}{\theta}) + \psi \sum_{i=1}^{n} \frac{\partial}{\partial \theta_j} g_i(X_i;\underset{\sim}{\theta})$$

$$- \sum_{i=1}^{n} [\int \frac{\partial}{\partial \theta_j} \log f_i(u;\underset{\sim}{\theta}) f_i(u;\underset{\sim}{\theta}) \exp\{\psi g_i(u;\underset{\sim}{\theta})\} du /$$

$$h_i(\psi;\underset{\sim}{\theta})] + \psi \sum_{i=1}^{n} [\int \frac{\partial}{\partial \theta_j} g_i(u;\underset{\sim}{\theta}) f_i(u;\underset{\sim}{\theta})$$

$$\times \exp\{\psi g_i(u;\underset{\sim}{\theta})\} du / h_i(\psi;\underset{\sim}{\theta})] \ .$$

Hence

$$\frac{\partial \ell}{\partial \theta_j}(0;\underset{\sim}{\theta}) = \sum_{i=1}^{n} \frac{\partial}{\partial \theta_j} \log f_i(X_i;\underset{\sim}{\theta}) \quad , \quad j = 1,\ldots,s.$$

The second derivatives of the log likelihood are given by

$$\frac{\partial^2 \ell}{\partial \psi^2}(\psi;\underset{\sim}{\theta}) = - \sum_{i=1}^{n} [\int \{g_i(u;\underset{\sim}{\theta})\}^2 f_i(u;\underset{\sim}{\theta}) \exp\{\psi g_i'(u;\underset{\sim}{\theta})\} du /$$

$$h_i(\psi;\underset{\sim}{\theta})] + \sum_{i=1}^{n} [\int g_i(u;\underset{\sim}{\theta}) f_i(u;\underset{\sim}{\theta}) \ \times$$

$$\exp\{\psi g_i(u;\underset{\sim}{\theta})\} du]^2 / \{h_i(\psi;\underset{\sim}{\theta})\}^2$$

$$\frac{\partial^2 \ell}{\partial\psi\partial\theta_j}(\psi;\underset{\sim}{\theta}) = \sum_{i=1}^{n} \frac{\partial}{\partial\theta_j} g_i(X_i;\underset{\sim}{\theta})$$

$$- \sum_{i=1}^{n} [\int \frac{\partial}{\partial\theta_j} g_i(u;\underset{\sim}{\theta}) f_i(u;\underset{\sim}{\theta}) \exp\{\psi g_i(u;\underset{\sim}{\theta})\} du]/h(\psi;\underset{\sim}{\theta})$$

$$- \sum_{i=1}^{n} [\int g_i(u;\underset{\sim}{\theta}) \frac{\partial}{\partial\theta_j} \log f_i(u;\underset{\sim}{\theta}) f_i(u;\underset{\sim}{\theta}) \exp\{\psi g_i(u;\underset{\sim}{\theta})\} du]/$$

$$h(\psi;\underset{\sim}{\theta}) - \psi \sum_{i=1}^{n} [\int g_i(u;\underset{\sim}{\theta}) \frac{\partial}{\partial\theta_j} g_i(u;\underset{\sim}{\theta}) f_i(u;\underset{\sim}{\theta})$$

$$\times \exp\{\psi g_i(u;\underset{\sim}{\theta})\} du]/h(\psi;\underset{\sim}{\theta})$$

$$+ \sum_{i=1}^{n} [\int g_i(u;\underset{\sim}{\theta}) f_i(u;\underset{\sim}{\theta}) \exp\{\psi g_i(u;\underset{\sim}{\theta})\} du]$$

$$\times [\int \frac{\partial}{\partial\theta_j} \log f_i(u;\underset{\sim}{\theta}) f_i(u;\underset{\sim}{\theta}) \exp\{\psi g_i(u;\underset{\sim}{\theta})\} du]/\{h(\psi;\underset{\sim}{\theta})\}^2$$

$$+ \psi \sum_{i=1}^{n} [\int f_i(u;\underset{\sim}{\theta}) \frac{\partial}{\partial\theta_j} g_i(u;\underset{\sim}{\theta}) \exp\{\psi g_i(u;\underset{\sim}{\theta})\} du]$$

$$\times [\int g_i(u;\underset{\sim}{\theta}) f_i(u;\underset{\sim}{\theta}) \exp\{\psi g_i(u;\underset{\sim}{\theta})\} du]/\{h(\psi;\underset{\sim}{\theta})\}^2$$

$$\frac{\partial^2 \ell}{\partial\theta_j\partial\theta_k}(\psi;\underset{\sim}{\theta}) = \sum_{i=1}^{n} \frac{\partial^2}{\partial\theta_j\partial\theta_k} \log f_i(X_i;\underset{\sim}{\theta})$$

$$+ \psi \sum_{i=1}^{n} \frac{\partial^2}{\partial\theta_j\partial\theta_k} g_i(X_i;\underset{\sim}{\theta})$$

$$-\sum_{i=1}^{n} [\int \frac{\partial^2}{\partial\theta_j\partial\theta_k} \log f_i(u;\underset{\sim}{\theta}) f_i(u;\underset{\sim}{\theta}) \exp\{\psi g_i(u;\underset{\sim}{\theta})\}du]/$$

$$h(\psi;\underset{\sim}{\theta}) - \sum_{i=1}^{n} \int \frac{\partial}{\partial\theta_j} \log f_i(u;\underset{\sim}{\theta}) \frac{\partial}{\partial\theta_k} \log f_i(u;\underset{\sim}{\theta}) f_i(u;\theta) \exp\{\psi g_i(u;\underset{\sim}{\theta})\}du/$$

$$h(\psi;\underset{\sim}{\theta}) - \psi \sum_{i=1}^{n} [\int \frac{\partial}{\partial\theta_k} g_i(u;\underset{\sim}{\theta}) \frac{\partial}{\partial\theta_j} \log f_i(u;\underset{\sim}{\theta}) f_i(u;\underset{\sim}{\theta})$$

$$\times \exp\{\psi g_i(u;\underset{\sim}{\theta})\}du]/h(\psi;\underset{\sim}{\theta}) - \psi \sum_{i=1}^{n} [\int \frac{\partial^2}{\partial\theta_j\partial\theta_k} g_i(u;\underset{\sim}{\theta})$$

$$\times f_i(u;\underset{\sim}{\theta}) \exp\{\psi g_i(u;\underset{\sim}{\theta})\}du]/h(\psi;\underset{\sim}{\theta})$$

$$-\psi^2 \sum_{i=1}^{n} [\int \frac{\partial}{\partial\theta_j} g_i(u;\underset{\sim}{\theta}) \frac{\partial}{\partial\theta_k} g_i(u;\underset{\sim}{\theta}) f_i(u;\underset{\sim}{\theta}) \exp\{\psi g_i(u;\underset{\sim}{\theta})\}du]/$$

$$h(\psi;\underset{\sim}{\theta}) + \psi \sum_{i=1}^{n} [\int \frac{\partial}{\partial\theta_j} g_i(u;\underset{\sim}{\theta}) f_i(u;\underset{\sim}{\theta}) \exp\{\psi g_i(u;\underset{\sim}{\theta})\}du]$$

$$\times [\int \frac{\partial}{\partial\theta_k} \log f_i(u;\underset{\sim}{\theta}) \exp\{\psi g_i(u;\underset{\sim}{\theta})\}du + \psi \int \frac{\partial}{\partial\theta_k} g_i(u;\underset{\sim}{\theta})$$

$$\times f_i(u;\underset{\sim}{\theta}) \exp\{\psi g_i(u;\underset{\sim}{\theta})\}du]/\{h(\psi;\underset{\sim}{\theta})\}^2.$$

Hence, as may be seen directly from (4.5.2),

$$c_{\psi\psi}\Big|_{H_o} = E_o \{ \frac{-\partial^2 \ell}{\partial\psi^2} (\psi;\underset{\sim}{\theta})\}$$

$$= \sum_{i=1}^{n} \mathrm{var}\{g_i(X;\underset{\sim}{\theta})\} \tag{4.5.3}$$

$$c_{\psi\theta_j}\Big|_{H_o} = E_o \left\{ \frac{-\partial^2 \ell}{\partial\psi\partial\theta_j} (\psi;\underset{\sim}{\theta}) \right\}$$

$$= -\sum_{i=1}^{n} E_o \left[\frac{\partial}{\partial\theta_j} g_i(X_i;\underset{\sim}{\theta}) - E_o \left\{ \frac{\partial}{\partial\theta_j} g_i(X;\underset{\sim}{\theta}) \right\} \right]$$

$$+ \sum_{i=1}^{n} \text{cov}\{g_i(X;\underset{\sim}{\theta}) \ , \ \frac{\partial}{\partial\theta_j} \log f_i(X;\underset{\sim}{\theta})\}$$

$$= \sum_{i=1}^{n} \text{cov}\{g_i(X;\underset{\sim}{\theta}) \ , \ \frac{\partial}{\partial\theta_j} \log f_i(X;\underset{\sim}{\theta})\} \qquad (4.5.4)$$

$$c_{\theta_j\theta_k}\Big|_{H_o} = -\sum_{i=1}^{n} E_o \left\{ \frac{\partial^2}{\partial\theta_j\partial\theta_k} \log f_i(X;\underset{\sim}{\theta}) \right\}$$

$$+ \sum_{i=1}^{n} \left[E_o \left\{ \frac{\partial^2}{\partial\theta_j\partial\theta_k} \log f_i(X;\underset{\sim}{\theta}) \right\} + E_o \left\{ \frac{\partial}{\partial\theta_j} \log f_i(X;\underset{\sim}{\theta}) \right. \right.$$

$$\left. \left. \times \frac{\partial}{\partial\theta_k} \log f_i(X;\underset{\sim}{\theta}) \right\} \right]$$

$$= -\sum_{i=1}^{n} E_o \left\{ \frac{\partial^2}{\partial\theta_j\partial\theta_k} \log f_i(X;\underset{\sim}{\theta}) \right\} . \qquad (4.5.5)$$

Let

$$\underset{\sim}{c}'_{12} = [c_{\psi\theta_1} \cdots c_{\psi\theta_s}] \ , \qquad \underset{\sim}{c}_{22} = \begin{bmatrix} c_{\theta_1\theta_1} & \cdots & c_{\theta_s\theta_1} \\ \vdots & & \\ c_{\theta_s\theta_1} & \cdots & c_{\theta_s\theta_s} \end{bmatrix}$$

$$\phi_1 = \frac{\partial \ell}{\partial \psi}(\psi;\underset{\sim}{\theta}) \quad \text{and} \quad \underset{\sim}{\phi}_2' = [\frac{\partial \ell}{\partial \theta_1}(\psi;\underset{\sim}{\theta}) \;\ldots\; \frac{\partial \ell}{\partial \theta_s}(\psi;\underset{\sim}{\theta})].$$

Then the test statistic for $H_o : \psi = 0$ is given by

$$T^{(C)} = (\phi_1 - \underset{\sim}{c}_{12}' \underset{\sim}{c}_{22}^{-1} \underset{\sim}{\phi}_2)(c_{\psi\psi} - \underset{\sim}{c}_{12}' \underset{\sim}{c}_{22}^{-1} \underset{\sim}{c}_{12})^{-\frac{1}{2}} \Bigg|_{\substack{\psi=0 \\ \underset{\sim}{\theta}=\underset{\sim}{\tilde{\theta}}_n}} \qquad (4.5.6)$$

where $\underset{\sim}{\tilde{\theta}}_n$ is a locally root-n consistent estimator of $\underset{\sim}{\theta}$.

Restricting ourselves to independently and identically distributed random variables, and $g_i(x;\underset{\sim}{\theta}) = g(x;\underset{\sim}{\theta})$ all i, say, the condition that $c_{\psi\psi} - \underset{\sim}{c}_{12}' \underset{\sim}{c}_{22}^{-1} \underset{\sim}{c}_{12} \neq 0$ imposes restrictions on the form of $g(x;\underset{\sim}{\theta})$ if we have a single sample of observations.

For example:-

(i) if $f(x;\underset{\sim}{\theta})$ is a member of the exponential family, with general form

$$f(x;\underset{\sim}{\theta}) = C(\underset{\sim}{\theta})D(x)\exp\{\sum_{k=1}^{s} A_k(\underset{\sim}{\theta})B_k(x)\},$$

then if $\underset{\sim}{a}' = [a_1 \ldots a_s]$ is a vector of arbitrary constants with $a_i \neq 0$ for at least one i, $g(x;\underset{\sim}{\theta})$ cannot take the form $g(x;\underset{\sim}{\theta}) = \underset{\sim}{a}'\underset{\sim}{B}(x)$,

where

$$\{\underset{\sim}{B}(x)\}' = [B_1(x) \ldots B_s(x)].$$

Let

$$\{\underset{\sim}{A}(\underset{\sim}{\theta})\}' = [A_1(\underset{\sim}{\theta}) \ldots A_s(\underset{\sim}{\theta})] \quad , \quad (\underset{\sim}{A}_1(\underset{\sim}{\theta}))_{ij} = \frac{\partial}{\partial \theta_j} A_i(\underset{\sim}{\theta}) \; ; \quad i,j = 1,\ldots,s,$$

$$\{\underset{\sim}{C}_1(\underset{\sim}{\theta})\}' = [\frac{\partial}{\partial\theta_1} C(\underset{\sim}{\theta}) \ \dots \ \frac{\partial}{\partial\theta_s} C(\underset{\sim}{\theta})] ,$$

$$(\underset{\sim}{\partial\ell})' = [\frac{\partial}{\partial\theta_1} \log f(x;\underset{\sim}{\theta}) \ \dots \ \frac{\partial}{\partial\theta_s} \log f(x;\underset{\sim}{\theta})] ,$$

$$(\underset{\sim}{B}_g)' = [\text{cov}\{g(X;\underset{\sim}{\theta}),B_1(X)\},\dots,\text{cov}\{g(X;\underset{\sim}{\theta}),B_s(X)\}]$$

We have the well known results

$$E\{\underset{\sim}{B}(X)\} = -\{C(\underset{\sim}{\theta})\}^{-1}[\{A_1(\underset{\sim}{\theta})\}']^{-1}\underset{\sim}{C}_1(\underset{\sim}{\theta})$$

$$\underset{\sim}{c}(\underset{\sim}{\partial\ell}) \equiv \underset{\sim}{c}_{22} = \{\underset{\sim}{A}_1(\underset{\sim}{\theta})\}'\underset{\sim}{c}\{\underset{\sim}{B}(\underset{\sim}{X})\}\underset{\sim}{A}_1(\underset{\sim}{\theta})$$

where $c(.)$ denotes a variance-covariance matrix. Now, if $g(x;\underset{\sim}{\theta}) = \underset{\sim}{a}'\underset{\sim}{B}(x)$, then from (4.5.3)-(4.5.5)

$$c_{\psi\psi}-\underset{\sim}{c}_{12}'\underset{\sim}{c}_{22}^{-1}\underset{\sim}{c}_{12} = \underset{\sim}{a}'\underset{\sim}{c}\{\underset{\sim}{B}(X)\}\underset{\sim}{a} - \underset{\sim}{B}_g'\underset{\sim}{A}_1(\underset{\sim}{\theta})\underset{\sim}{c}_{22}^{-1}\{\underset{\sim}{A}_1(\underset{\sim}{\theta})\}'B_g$$

$$= \underset{\sim}{a}'\underset{\sim}{c}\{\underset{\sim}{B}(X)\}\underset{\sim}{a} - \underset{\sim}{a}\ \underset{\sim}{c}\{\underset{\sim}{B}(X)\}\underset{\sim}{A}_1(\underset{\sim}{\theta})\{A_1(\underset{\sim}{\theta})\}^{-1}$$

$$\times [\underset{\sim}{c}\{\underset{\sim}{B}(X)\}]^{-1}[\{\underset{\sim}{A}_1(\underset{\sim}{\theta})\}']^{-1}\{\underset{\sim}{A}_1(\underset{\sim}{\theta})\}'[\underset{\sim}{c}\{\underset{\sim}{B}(X)\}]'\underset{\sim}{a}$$

$$= 0.$$

In this case we have of course an s dimensional sufficient statistic

$$[\sum_{i=1}^{n} B_1(X_i) \ \dots \ \sum_{i=1}^{n} B_s(X_i)]$$

for an $(s+1)$ dimensional parameter $[\psi\ \underset{\sim}{\theta}']$ and the model is under-identified.

(ii) More generally, we cannot have $g(x;\underset{\sim}{\theta}) = \underset{\sim}{a}'\underset{\sim}{\partial\ell}$. In this case $(c_{\psi\theta_j}) = \underset{\sim}{c}_{22}\underset{\sim}{a}$ and immediately

$$c_{\psi\psi}-\underset{\sim}{c}_{12}'\underset{\sim}{c}_{22}^{-1}\underset{\sim}{c}_{12} = \underset{\sim}{a}'\underset{\sim}{c}_{22}\underset{\sim}{a}-\underset{\sim}{a}'\underset{\sim}{c}_{22}\underset{\sim}{c}_{22}^{-1}\underset{\sim}{c}_{22}\underset{\sim}{a} = 0.$$

The situation in (i) is, of course, a particular case of (ii).

The framework of this section often leads to standard solutions. We give two examples

(iii) Let $X_1,\ldots,X_m$, $Y_1,\ldots,Y_n$ be a sample of independent random variables whose distribution under the null hypothesis has the gamma density

$$f(x;\underset{\sim}{\theta}) = \{\Gamma(\beta)\}^{-1}\alpha^{\beta}x^{\beta-1}e^{-\alpha x} \quad , \quad i = 1,\ldots,m+n.$$

Under the alternative hypothesis the density of an observation takes the form (4.5.1) and let

$$g(y;\underset{\sim}{\theta}) = \begin{cases} 0 \ , & \text{for a random variable among } X_1,\ldots,X_m \\ \log y \ , & \text{for a random variable among } Y_1,\ldots,Y_n. \end{cases}$$

The test statistic (4.5.6) becomes

$$T^{(C)} = \frac{\frac{1}{n}\sum_{i=1}^{n} \log Y_i - \frac{1}{m}\sum_{i=1}^{m} \log X_i}{\{(\frac{1}{m}+\frac{1}{n})\psi'(\tilde{\beta}_n)\}^{\frac{1}{2}}} \tag{4.5.7}$$

where $\tilde{\beta}_n$ is a locally root-n consistent estimator of β and $\psi(z)$ is the digamma function. The statistic is the same as found in Moran (1970b) under a different parameterisation for the problem of testing for a change in the scale parameter of a gamma distribution.

(iv) Let $X_1,\ldots,X_m$, $Y_1,\ldots,Y_n$ be a sample of independent random variables with density $N(\mu,\sigma^2)$ under a null hypothesis. Under the alternative hypothesis the density of an observation takes the form (4.5.1) and let

$$g(y;\underset{\sim}{\theta}) = \begin{cases} 0 \ , & \text{for a random variable among } X_1,\ldots,X_m \\ -\frac{1}{2}(\frac{y-\mu}{\sigma})^2, & \text{for a random variable among } Y_1,\ldots,Y_n. \end{cases}$$

We have $\mathrm{var}(Y_i) = \sigma^2(1+\psi)^{-1}$, $i = 1,\ldots,n$, under the alternative hypothesis and the test of $H_o : \psi = 0$ is based upon

$$T^{(C)} = \frac{1}{2^{\frac{1}{2}}\tilde{\sigma}^2_{m,n}(\frac{1}{m}+\frac{1}{n})^{\frac{1}{2}}}\left\{\frac{1}{m}\sum_{i=1}^{m}(X_i-\tilde{\mu}_{m,n})^2 - \frac{1}{n}\sum_{i=1}^{n}(Y_i-\tilde{\mu}_{m,n})^2\right\} \qquad (4.5.8)$$

where $\tilde{\mu}_{m,n}$ and $\tilde{\sigma}^2_{m,n}$ are locally root-n consistent estimators of μ and σ^2 respectively. For example the mean of the combined samples and the pooled estimator of variance. In both (4.5.7) and (4.5.8), $T^{(C)}$ is asymptotically distributed as $N(0,1)$.

5. MAXIMUM LIKELIHOOD THEORY WHEN A SUBSET OF $\underset{\sim}{\theta}_1$ LIES ON A BOUNDARY OF PARAMETER SPACE

5.1 INTRODUCTION

The previous discussions were concerned with the equivalence of optimal $C(\alpha)$ tests and tests based on maximum likelihood estimators when the true parameter point $\underset{\sim}{\theta}$ lies in an open subset of parameter space Θ.

We now restrict $\underset{\sim}{\theta}$ to lie in a closed space $\Theta^* \subseteq \Theta$, which may or may not be a proper subset of Θ. We will be concerned with the case where the null hypothesis specifies that $\underset{\sim}{\theta}$ lies on the boundary of Θ^*. Moran (1971b) showed that if a scalar parameter is under test, and lies on the plane boundary of a closed set in Euclidean space, then the conventional maximum likelihood theory must be refined, whilst optimal $C(\alpha)$ tests may be applied as usual. The rejection regions for the two tests are, however, asymptotically the same for one-sided hypotheses. We further show that if a vector parameter is under test, and a subset of this vector parameter lies on a boundary of a closed subset of Euclidean space, then the two tests are no longer equivalent. We derive the asymptotic joint distribution of the maximum likelihood estimators under the null hypothesis.

When considering the power of a maximum likelihood test specifying that $\underset{\sim}{\theta}$ lies on the boundary of a parameter space, uniform consistency of the maximum likelihood estimator is necessary. By restricting Θ^* to be a closed subset of s dimensional Euclidean space, we ensure that the parameter space is compact and then (Moran, 1971a), the maximum likelihood estimator is uniformly consistent. Compactness is, however, not always necessary. Wald (1949) assumes that the common density $p(x;\underset{\sim}{\theta})$ of a sample of independent and identically distributed random variables has the following properties:-

(i) That it is possible to introduce a distance $\delta(\underset{\sim}{\theta}_1, \underset{\sim}{\theta}_2)$ into the space Θ^* in which the following conditions hold.

(a) The distance $\delta(\underset{\sim}{\theta}_1, \underset{\sim}{\theta}_2)$ makes Θ^* a metric space.

(b) $\lim_{i\to\infty} p(x;\underset{\sim}{\theta}_i) = p(x;\underset{\sim}{\theta})$ if $\lim_{i\to\infty} \underset{\sim}{\theta}_i = \underset{\sim}{\theta}$ for any x except a set which may depend on $\underset{\sim}{\theta}$ (but not on the sequence $\underset{\sim}{\theta}_i$) and whose probability measure is zero according to the probability distribution corresponding to the true parameter point $\underset{\sim}{\theta}_o$.

(ii) If $\underset{\sim}{\theta}_o$ is a fixed point in Θ^* and $\lim_{i\to\infty} \delta(\underset{\sim}{\theta}_i, \underset{\sim}{\theta}_o) = \infty$, then $\lim_{i\to\infty} p(x;\underset{\sim}{\theta}_i) = 0$ for any x.

(iii) Any closed and bounded subset of Θ^* is compact.

5.2 FORMULATION OF THE PROBLEM

For clarity, let Θ^* be a subset of Euclidean space given by

$0 \le \theta < a_i$, $i = 1,\dots,t$ $(a_i > 0)$, $-\infty < \theta_i < \infty$, $i = t+1,\dots,s$.

(Note that we cannot have $0 \le \theta_i \le a_i$, as then we can have boundary points in which one or more of the θ_i are equal to the corresponding a_i). We assume uniform consistency of the maximum likelihood estimator. Without loss of generality we may write the null hypothesis as $H_{o1} : \theta_i = 0$, $i = 1,\dots,p \le t$; $\theta_{p+1},\dots,\theta_s$ fixed and unknown in a closed subset of Θ^*, $\underset{\sim}{\theta}_2$ a nuisance parameter.

We take $p < t$ as we are more likely to be interested in problems where a subset of $\underset{\sim}{\theta}_1$ lies on the boundary of Θ^*, rather than fixing $\underset{\sim}{\theta}_1$ in a "corner" of Θ^*, in which case $p = t$. Hence $\theta_{p+1},\dots,\theta_s$ becomes, in effect, part of an enlarged nuisance parameter. However, the fact that $\theta_{p+1},\dots,\theta_t$ may themselves lie on the boundary of Θ^* must be taken into account.

The hypothesis H_{o2} : $\theta_1=\ldots=\theta_p=0$; $\theta_{p+1},\ldots,\theta_t,\ldots,\theta_s$ fixed and unknown in an open subset of Θ^* is a special case of H_{o1} with $p = t$, as $\hat{\theta}_{n,p+1},\ldots,\hat{\theta}_{nt}$ are almost surely positive under H_{o2}. Moran (1971b) considers tests of H_{o2} for $t = 1,2$.

Let

$$L(\underset{\sim}{\theta};\underset{\sim}{X}) = \prod_{k=1}^{n} p(X_k;\underset{\sim}{\theta})$$

be the likelihood of a sample $X_1,\ldots,X_n$, which we now restrict to consist of independent and identically distributed random variables whose common probability density satisfies the conditions set out in Moran (1971a).

$\hat{\underset{\sim}{\theta}}_n$ is defined to be the value of $\underset{\sim}{\theta}$ that maximises $L(\underset{\sim}{\theta};\underset{\sim}{X})$ for $\underset{\sim}{\theta} \in \Theta^*$.

5.3 THE ASYMPTOTIC JOINT DISTRIBUTION OF THE MAXIMUM LIKELIHOOD ESTIMATORS

At the maximum,

$$\sum_{k=1}^{n} \frac{\partial}{\partial\theta_i} \log p(X_k;\hat{\underset{\sim}{\theta}}_n) \leq 0 \quad , \quad i = 1,\ldots,t, \tag{5.3.1}$$

$$\sum_{k=1}^{n} \frac{\partial}{\partial\theta_i} \log p(X_k;\hat{\underset{\sim}{\theta}}_n) = 0 \quad , \quad i = t+1,\ldots,s, \tag{5.3.2}$$

where the derivative in (5.3.1) is taken to the right if $\hat{\theta}_{ni} = 0$, $i = 1,\ldots,t$, and is an equality if $\hat{\theta}_{ni} > 0$, $i = 1,\ldots,t$.

Let

$$Y_{ni} = n^{-\frac{1}{2}} \sum_{k=1}^{n} \frac{\partial}{\partial\theta_i} \log p(X_k;\underset{\sim}{\theta}) \quad , \quad i = 1,\ldots,s$$

$$\underset{\sim}{Y}'_n = [Y_{n1}\ldots Y_{ns}] \quad , \quad Z_{nj} = n^{\frac{1}{2}}(\hat{\theta}_{nj}-\theta_j) \quad \text{and}$$

$$\underset{\sim}{Z}'_n = [Z_{n1}\ldots Z_{ns}].$$

Now, if $\underset{\sim}{\theta}$ is the true parameter point, we have from (5.3.1) and (5.3.2) that

$$Y_{ni} - \sum_{j=1}^{s} c_{ij} Z_{nj} \leq 0 \quad , \quad i = 1,\ldots,t \tag{5.3.3}$$

$$Y_{ni} - \sum_{j=1}^{s} c_{ij} Z_{nj} = 0 \quad , \quad i = t+1,\ldots,s. \tag{5.3.4}$$

Asymptotically, $\underset{\sim}{Y}_n$ is distributed as $N_s(\underset{\sim}{0},\underset{\sim}{c})$. To find the asymptotic joint distribution of $\underset{\sim}{Z}_n$, we consider the distribution of $\underset{\sim}{Z}_n$ conditionally upon the values taken by $\hat{\underset{\sim}{\theta}}{}'_{n1} = [\hat{\theta}_{n1} \ldots \hat{\theta}_{nt}]$. We will later distinguish between the cases $t = 1$ and $t > 1$.

Let

$$\Phi^*(\underset{\sim}{u};\underset{\sim}{\theta}) = pr(\underset{\sim}{Z}_n \leq \underset{\sim}{u};\underset{\sim}{\theta})$$

be the asymptotic joint distribution function of $\underset{\sim}{Z}_n$, where $\underset{\sim}{u}' = [u_1 \ldots u_s]$. We have

$$\Phi^*(\underset{\sim}{u};\underset{\sim}{\theta}) = \sum_{i=1}^{\ell} a_i F_i(\underset{\sim}{u};\underset{\sim}{\theta})$$

where $F_i(\underset{\sim}{u};\underset{\sim}{\theta})$ is the distribution function of $\underset{\sim}{Z}_n$ conditional upon a subset of $\hat{\underset{\sim}{\theta}}_{n1}$ having all zero elements, and the complement of this subset of $\hat{\underset{\sim}{\theta}}_{n1}$ having all non-zero elements.

A precise method of constructing the $F_i(\underset{\sim}{u};\underset{\sim}{\theta})$ is to first find the distribution of $\underset{\sim}{Z}_n$ conditional upon

$$\hat{\theta}_{ni} = 0 \ , \quad \hat{\theta}_{nj} > 0 \ , \quad j = 1,\ldots,t; \ \ j \neq i \text{ for each } i = 1,\ldots,t.$$

Then to condition upon

$$\hat{\theta}_{ni} = \hat{\theta}_{nj} = 0 \ , \quad \hat{\theta}_{nk} > 0 \ , \quad k = 1,\ldots,t; \ \ k \neq i, \ k \neq j, \text{ for } j = 1,\ldots,t; \ \ i \neq j,$$

and so on until a final condition $\hat{\underset{\sim}{\theta}}_{n1} = 0$. There are

$$\ell = \sum_{i=0}^{t} \binom{t}{i} = 2^t$$

such distinct subsets and we let

$$a_i = \text{pr}\{\text{conditions placed upon } \hat{\underset{\sim}{\theta}}_{n1} \text{ hold}\}.$$

Preliminary to investigating the asymptotic form of $F_i(\underset{\sim}{u};\underset{\sim}{\theta})$, let $\underset{\sim}{c}_{22}$ be the matrix formed by striking out the rows and columns of $\underset{\sim}{c}$ corresponding to elements of $\hat{\underset{\sim}{\theta}}_{n1}$ conditionally set to zero. For example, if we set

$$\hat{\theta}_{n1} = \hat{\theta}_{n,p-1} = \hat{\theta}_{n,p+2} = 0$$

we strike out the 1^{st}, $(p-1)^{th}$ and $(p+2)^{th}$ row and columns of $\underset{\sim}{c}$.

Let

$$\{\underset{\sim}{\theta}_1^{(o)}\}' = [w_{p+1}\theta_{p+1} \cdots w_t\theta_t]$$

where

$$w_i = \begin{cases} 1, & \text{if } \hat{\theta}_{ni} \text{ is conditionally set to zero, } i = p+1,\ldots,t \\ 0, & \text{otherwise.} \end{cases}$$

Consider the matrix

$$\begin{bmatrix} c_{1,p+1} & \cdots & c_{1t} \\ \cdot & & \\ \cdot & & \\ \cdot & & \\ c_{s,p+1} & \cdots & c_{st} \end{bmatrix} \tag{5.3.5}$$

Let $\underset{\sim}{c}^{(r)}$ be the matrix formed from (5.3.5) by striking out the rows and columns of (5.3.5) corresponding to elements of $\hat{\underset{\sim}{\theta}}_{n1}$ conditionally set to zero.

Let r_i be the number of elements of $\hat{\underset{\sim}{\theta}}_{n1}$ conditionally set to zero. We are now in a position to prove

Theorem 2

$F_i(\underset{\sim}{u};\underset{\sim}{\theta})$ is asymptotically the distribution function of a

$N_{s-r_i}\{n^{\frac{1}{2}}\underset{\sim}{c}_{22}^{-1}\underset{\sim}{c}^{(r)}\underset{\sim}{\theta}_1^{(o)}, \underset{\sim}{c}_{22}^{-1}\}$ random variable.

Notes (i) If the elements of $\hat{\underset{\sim}{\theta}}_{n1}$ conditionally set to zero are all from among $\hat{\theta}_{n1},\ldots,\hat{\theta}_{np}$, then $\underset{\sim}{\theta}_1^{(o)} = \underset{\sim}{0}$ and the corresponding means vector is $\underset{\sim}{0}$.

(ii) In the case $p = t$ (that is, the hypothesis H_{o2} of section 5.2) the means vector is $\underset{\sim}{0}$ and

$$(\underset{\sim}{c}_{22})_{ij} = -E\{\frac{\partial^2}{\partial\theta_i\partial\theta_j}\log p(X;\underset{\sim}{\theta})\} \qquad i,j = t+1,\ldots,s$$

and corresponds to c_{ij} defined in (2.3.2) for the case of independently and identically distributed random variables. In general this correspondence between $\underset{\sim}{c}_{22}$ and the nuisance parameter of chapters 1-4 is lost.

proof:

First, we consider the case where we set

$$\hat{\theta}_{n1} = \ldots = \hat{\theta}_{nq} = 0 \qquad q \leq p; \quad \hat{\theta}_{ni} > 0, \quad i = q+1,\ldots,p.$$

This is the situation in note (i) above.

The equations (5.3.3) and (5.3.4) become

$$Y_{ni} - \sum_{j=q+1}^{s} c_{ij}Z_{nj} < 0\ , \qquad i = 1,\ldots,q \tag{5.3.6}$$

$$Y_{ni} - \sum_{j=q+1}^{s} c_{ij}Z_{nj} = 0\ , \qquad i = q+1,\ldots,s \tag{5.3.7}$$

as both $\hat{\theta}_{ni}$ and θ_i are zero for $i = 1,\ldots,q$.

Let

$$\underset{\sim}{Y}'_{n1} = [Y_{n1}\cdots Y_{nq}]\ , \qquad \underset{\sim}{Y}'_{n2} = [Y_{n,q+1}\cdots Y_{ns}]$$

$$\underset{\sim}{Z}'_{n1} = [Z_{n1}\cdots Z_{nq}]\ , \qquad \underset{\sim}{Z}'_{n2} = [Z_{n,q+1}\cdots Z_{ns}]$$

and

$$c = \begin{bmatrix} c_{11} & c'_{12} \\ c_{12} & c_{22} \end{bmatrix}$$

where c_{11} is $q \times q$, c'_{12} is $q \times (s-q)$ and c_{22} is $(s-q) \times (s-q)$. Note that generally c_{22} is determined according to the subset of $\hat{\theta}_{n1}$ conditionally set to zero and c_{22} changes when the conditioning changes.

(5.3.6) and (5.3.7) may now be written

$$Y_{n2} = c_{22} Z_{n2}$$

$$T_n \equiv Y_{n1} - c'_{12} c_{22}^{-1} Y_{n2} < 0 \tag{5.3.8}$$

where $\ell < m$ is taken to mean that if $\ell' = [\ell_1 \ldots \ell_r]$ and $m' = [m_1 \ldots m_r]$, then $\ell_i < m_i$, all i.

Hence $F_i(u;\theta)$, $u_1 = \ldots = u_q = 0$; $0 < u_i < \infty$, $i = q+1,\ldots,t$; $-\infty < u_i < \infty$, $i = t+1,\ldots,s$ has a probability density which is that of

$$Z_{n2} = c_{22}^{-1} Y_{n2}$$

conditional upon (5.3.8), and $0 < Z_{ni} < a_i$, $i = q+1,\ldots,t$. We must therefore find the distribution of $Y_{n2} | T_n < 0$. Now, Y_{n2} is asymptotically distributed as $N_{s-q}(0, c_{22})$. T_n is a vector of linear combinations of normal random variables and is itself multivariate normal, means vector 0.

Let Σ_T be the variance-covariance matrix of T_n and Σ_{TY} be the variance-covariance matrix

$$(\mathrm{cov}(Y_{ni}, T_{nj})) \ , \qquad i = q+1,\ldots,s; \ \ j = 1,\ldots,q.$$

Using standard results for the truncated normal distribution (see, for example Rao, 1965, pp.441-442) the density of the distribution of Y_{n2} conditional upon $T_n = u$ is

$$N_{s-q}(\underset{\sim\sim}{\Sigma_{TY}}\,\underset{\sim}{\Sigma_T^{-1}}\,\underset{\sim}{u},\ \underset{\sim}{c}_{22}-\underset{\sim\sim}{\Sigma_{TY}}\,\underset{\sim}{\Sigma_T^{-1}}\,\underset{\sim\sim}{\Sigma'_{TY}})$$

In our case

$$\underset{\sim\sim}{\Sigma_{TY}} = E(\underset{\sim}{T}_n\ \ \underset{\sim}{Y}'_{n2}) \tag{5.3.9}$$

$$= E(\underset{\sim}{Y}_{n1}\underset{\sim}{Y}'_{n2})-\underset{\sim}{c}'_{12}\underset{\sim}{c}_{22}^{-1}E(\underset{\sim}{Y}_{n2}\underset{\sim}{Y}'_{n2})$$

$$\equiv 0$$

where E refers to expectations taken with respect to the asymptotic distribution. Hence the distribution of $\underset{\sim}{Y}_{n2}|\underset{\sim}{T}_n < \underset{\sim}{u}$ is independent of the conditioning.

The distribution of $\underset{\sim}{Z}_{n2}$ conditional upon $\hat{\theta}_{n1} = \ldots = \hat{\theta}_{nq} = 0$; $\hat{\theta}_{ni} > 0$, $i = q+1,\ldots,t$ is asymptotically $N_{s-q}(\underset{\sim}{0},\underset{\sim}{c}_{22}^{-1})$ over the range $u_1 = \ldots = u_q = 0$; $0 < u_i < \infty$, $i = q+1,\ldots,t$; $-\infty < u_i < \infty$, $i = t+1,\ldots,s$.

In the general case where the subset of $\hat{\underset{\sim}{\theta}}_{n1}$ conditionally set to zero contains some of the $\hat{\theta}_{n,p+1},\ldots,\hat{\theta}_{nt}$, we must carefully consider the form taken by the equations (5.3.3) and (5.3.4).

First consider the result of the following matrix multiplication

$$\left[\begin{array}{ccc:ccc:ccc}
c_{11} & \cdots & c_{p1} & c_{p+1,1} & \cdots & c_{t1} & c_{t+1,p} & \cdots & c_{s1} \\
\cdot & & \cdot & \cdot & & \cdot & \cdot & & \cdot \\
\cdot & & \cdot & \cdot & & \cdot & \cdot & & \cdot \\
c_{p1} & \cdots & c_{pp} & c_{p+1,p} & \cdots & c_{tp} & c_{t+1,p} & \cdots & c_{sp} \\
\hdashline
c_{p+1,1} & \cdots & c_{p+1,p} & c_{p+1,p+1} & \cdots & c_{p+1,t} & c_{t+1,p+1} & \cdots & c_{s,p+1} \\
\cdot & & \cdot & \cdot & & \cdot & \cdot & & \cdot \\
\cdot & & \cdot & \cdot & & \cdot & \cdot & & \cdot \\
c_{t,1} & \cdots & c_{tp} & c_{t,p+1} & \cdots & c_{tt} & c_{t+1,t} & \cdots & c_{st} \\
\hdashline
c_{t+1,1} & \cdots & c_{t+1,p} & c_{t+1,p+1} & \cdots & c_{t+1,t} & c_{t+1,t+1} & \cdots & c_{s,t+1} \\
\cdot & & \cdot & \cdot & & \cdot & \cdot & & \cdot \\
\cdot & & \cdot & \cdot & & \cdot & \cdot & & \cdot \\
c_{s1} & \cdots & c_{sp} & c_{s,p+1} & \cdots & c_{st} & c_{s,t+1} & \cdots & c_{ss}
\end{array}\right]
\left[\begin{array}{c}
\theta_1 \\ \cdot \\ \cdot \\ \theta_p \\ \hdashline \theta_{p+1} \\ \cdot \\ \cdot \\ \theta_t \\ \hdashline \theta_{t+1} \\ \cdot \\ \cdot \\ \theta_s
\end{array}\right]$$

(5.3.10)

In the general framework, if $\hat{\theta}_{nj}$, $j > p$, is conditionally set to zero, the corresponding Z_{nj} becomes $-n^{\frac{1}{2}}\theta_j$, a non-zero term which must be taken into account when considering (5.3.3) and (5.3.4).

The equation (5.3.4) becomes, on inspecting (5.3.10)

$$\underset{\sim}{Y}_{n2}+n^{\frac{1}{2}}\underset{\sim}{c}^{(r)}\underset{\sim}{\theta}_1^{(o)}-\underset{\sim}{c}_{22}\underset{\sim}{Z}_{n2} = \underset{\sim}{0}.$$

The conditioning is again of no account, taking the same form as (5.3.9) and involving the addition of a constant to $\underset{\sim}{T}_n$ and $\underset{\sim}{Y}_{n2}$.

Hence the distribution of Z_{n2} is now

$$N_{s-r_i}(n^{\frac{1}{2}}\underset{\sim}{c}_{22}^{-1}\underset{\sim}{c}^{(r)}\underset{\sim}{\theta}_1^{(o)},\underset{\sim}{c}_{22}^{-1})$$

and the theorem is proved.

6. SOME EXAMPLES OF TESTS INVOLVING PARAMETERS ON BOUNDARIES

6.1 A SCALAR PARAMETER UNDER TEST

Moran (1971b) considers the cases $p = t = 1$ and $p = t = 2$ of the hypotheses $H_{01} : \theta_i = 0$, $i = 1,\ldots,p \leq t$; $\theta_{p+1},\ldots,\theta_s$ fixed and unknown in a closed subset of Θ^*.

He states that the component of $\Phi^*(\underset{\sim}{u};\underset{\sim}{\theta}) = \text{pr}\{\underset{\sim}{Z}_n \leq \underset{\sim}{u};\underset{\sim}{\theta}\}$ formed by conditioning on $\hat{\underset{\sim}{\theta}}_{n1} > 0$ ($\hat{\underset{\sim}{\theta}}_{n1}$ is a scalar for $t = 1$) has asymptotically the density of an s dimensional multivariate normal distribution defined on

$$u_i > 0 \quad i = 1,\ldots,p; \quad -\infty < u_i < \infty, \quad i = p+1,\ldots,s.$$

However, as we have shown, his statement that the other components of $\Phi^*(\underset{\sim}{u};\underset{\sim}{\theta})$ are non normal is incorrect. We can now write down the asymptotic joint distribution function of the maximum likelihood estimators in the case $p = t = 1$.

We have

$H_o : \theta_1 = 0$, $[\theta_2\ldots\theta_s]$ fixed and unknown in an open subset of Θ^*.

Let

$$\underset{\sim}{\theta}_o' = [0\ \theta_2\ldots\theta_s].$$

Then under H_o,

$$\Phi^*(\underset{\sim}{u};\underset{\sim}{\theta}_o) = \frac{|\underset{\sim}{c}|^{\frac{1}{2}}}{(2\pi)^{\frac{s}{2}}} \int_0^{u_1}\int_{-\infty}^{u_2}\cdots\int_{-\infty}^{u_s} \exp(-\tfrac{1}{2}\underset{\sim}{t}'\underset{\sim}{c}\underset{\sim}{t})dt_1\ldots dt_s$$

$$+ \frac{1}{2}\frac{|\underset{\sim}{c}_{22}|^{\frac{1}{2}}}{(2\pi)^{\frac{s-1}{2}}} \int_{-\infty}^{u_2}\int_{-\infty}^{u_s} \exp(-\tfrac{1}{2}\underset{\sim}{t}_1'\underset{\sim}{c}_{22}\underset{\sim}{t}_1)dt_2\ldots dt_s$$

where $\underset{\sim}{t}' = [t_1 \dots t_s]$, $\underset{\sim}{t}_1' = [t_2 \dots t_s]$ and

$$\underset{\sim}{c}^{*} = \left[\begin{array}{c|ccc} c_{11} & c_{21} & \cdots & c_{s1} \\ \hline c_{21} & c_{22} & \cdots & c_{s2} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ c_{s1} & c_{s2} & \cdots & c_{ss} \end{array}\right]^{-1} = \begin{bmatrix} c_{11} & \underset{\sim}{c}_{12}' \\ \underset{\sim}{c}_{12} & \underset{\sim}{c}_{22} \end{bmatrix}^{-1} = \begin{bmatrix} d_{11} & \underset{\sim}{d}_{12}' \\ \underset{\sim}{d}_{12} & \underset{\sim}{d}_{22} \end{bmatrix}$$

In particular

$$\mathrm{pr}\{n^{\frac{1}{2}}\hat{\theta}_{n1} < \mu_1; \underset{\sim}{\theta}_o\} = \frac{1}{2} + \frac{d_{11}^{\frac{1}{2}}}{(2\pi)^{\frac{1}{2}}} \int_o^{u_1} \exp(-\tfrac{1}{2} d_{11} t^2) dt \qquad (6.1.1)$$

where

$$d_{11}^{*} = c_{11} - \underset{\sim}{c}_{12}' \underset{\sim}{c}_{22}^{-1} \underset{\sim}{c}_{12} .$$

Writing the probability density of $n^{\frac{1}{2}}\hat{\theta}_{n1}$ on the positive real axis as $f_1(x)$, we have

$$\mathrm{pr}(n^{\frac{1}{2}}\hat{\theta}_{n1} = 0) = \frac{1}{2}$$

$$f_1(x) = \frac{d_{11}^{\frac{1}{2}}}{(2\pi)^{\frac{1}{2}}} \exp(-\tfrac{1}{2} d_{11} x^2) , \qquad x > 0. \qquad (6.1.2)$$

We note that if $\theta_2, \dots, \theta_s$ have fixed values in open intervals $-\infty < \theta_i < \infty$, $i = 2, \dots, s$, whilst $\theta_1 = an^{-\frac{1}{2}}$, $0 \le a \le \infty$, then according to Moran (1971b) we have

$$\Phi^{*}(\underset{\sim}{u}; \underset{\sim}{\theta}) = \alpha F_1(\underset{\sim}{u}; \underset{\sim}{\theta}) + (1-\alpha) F_2(\underset{\sim}{u}; \underset{\sim}{\theta})$$

where

$$F_1(\underset{\sim}{u};\underset{\sim}{\theta}) = \frac{1}{\alpha} \frac{|\underset{\sim}{c}|^{\frac{1}{2}}}{(2\pi)^{\frac{s}{2}}} \int_{-a}^{u_1} \int_{-\infty}^{u_2} \cdots \int_{-\infty}^{u_s} \exp(-\tfrac{1}{2}\underset{\sim}{t}'\underset{\sim\sim}{c}\underset{\sim}{t})dt_1\ldots dt_s$$

$$F_2(\underset{\sim}{u};\underset{\sim}{\theta}) = \frac{|\underset{\sim}{c}_{22}|^{\frac{1}{2}}}{(2\pi)^{\frac{s-1}{2}}} \int_{-\infty}^{u_2} \cdots \int_{-\infty}^{u_s} \exp\{-\tfrac{1}{2}(\underset{\sim}{t}_1+a\underset{\sim}{c}_{22}^{-1}\underset{\sim}{c}_{12})'\underset{\sim}{c}_{22}$$

$$\times (\underset{\sim}{t}_1+a\underset{\sim}{c}_{22}^{-1}\underset{\sim}{c}_{12})\}dt_2\ldots dt_s$$

and

$$\alpha = \text{pr}\{Z_{n1} > -a\} = \phi^*(ad_{11}^{-\frac{1}{2}})$$

where

$$\phi^*(x) = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{x} \exp(-\tfrac{1}{2}t^2)dt.$$

6.2 A VECTOR PARAMETER UNDER TEST

We now consider the case where $t = 2$. We take $H_o : \theta_1 = \theta_2 = 0$, $\underset{\sim}{\theta}_2' = [\theta_3\ldots\theta_s]$ fixed and unknown in an open subset of Θ^*.

Let $\underset{\sim}{\theta}_o' = [0\ 0\ \theta_3\ldots\theta_s]$. For simplicity of exposition we take $s = 3$, i.e., the nuisance parameter is a scalar.

We have

$$\underset{\sim}{d} = \left[\begin{array}{cc|c} c_{11} & c_{12} & c_{13} \\ c_{12} & c_{22} & c_{23} \\ \hline c_{13} & c_{23} & c_{33} \end{array}\right]^{-1} \equiv \left[\begin{array}{cc|c} d_{11} & d_{12} & d_{13} \\ d_{12} & d_{22} & d_{23} \\ \hline d_{13} & d_{23} & d_{33} \end{array}\right] = \begin{bmatrix} \underset{\sim}{d}_{11} & \underset{\sim}{d}_{12}' \\ \underset{\sim}{d}_{12} & d_{22} \end{bmatrix}$$

Let

$$\sigma_1^2 = c_{33}/(c_{11}c_{33}-c_{13}^2)$$

$$\sigma_2^2 = c_{33}/(c_{22}c_{33}-c_{23}^2).$$

The results of section (5.3) lead to the following joint asymptotic distribution function for $(n^{\frac{1}{2}}\hat{\theta}_{n1}, n^{\frac{1}{2}}\hat{\theta}_{n2})$

$$\mathrm{pr}\{n^{\frac{1}{2}}\hat{\theta}_{n1} < u_1, n^{\frac{1}{2}}\hat{\theta}_{n2} < u_2\} = \mathrm{pr}(Y_{n1} - \frac{c_{13}}{c_{33}} Y_{n3} \leq 0,\ Y_{n2} - \frac{c_{23}}{c_{33}} Y_{n3} \leq 0)$$

$$+ \frac{1}{2} \cdot \frac{1}{(2\pi)^{\frac{1}{2}}} \cdot \frac{1}{\sigma_1} \int_o^{u_1} \exp(-\frac{1}{2} \frac{t_1^2}{\sigma_1^2}) dt_1$$

$$+ \frac{1}{2} \frac{1}{(2\pi)^{\frac{1}{2}}} \cdot \frac{1}{\sigma_2} \int_o^{u_2} \exp(-\frac{1}{2} \frac{t_2^2}{\sigma_2^2}) dt_2$$

$$+ \frac{1}{(2\pi)} \frac{1}{|\underset{\sim}{d}_{11}|^{\frac{1}{2}}} \int_o^{u_1} \int_o^{u_2} \exp(-\frac{1}{2} \underset{\sim}{t}' \underset{\sim}{d}_{11}^{-1} \underset{\sim}{t}) dt_1 dt_2 \qquad (6.2.1)$$

where $\underset{\sim}{t}' = [t_1 t_2]$. The random variables in the first term of (6.2.1) are asymptotically correlated with correlation coefficient ρ, say, where

$$\rho = \frac{c_{12}c_{33} - c_{13}c_{23}}{\{(c_{11}c_{33} - c_{13}^2)(c_{22}c_{33} - c_{23}^2)\}^{\frac{1}{2}}} \qquad (6.2.2)$$

and hence (Moran; 1968, pp.312-314), the probability value is

$$\mathrm{pr}\{Y_{n1} - \frac{c_{13}}{c_{33}} Y_{n3} \leq 0,\ Y_{n2} - \frac{c_{23}}{c_{33}} Y_{n3} \leq 0\} = \frac{1}{4} + \frac{1}{2\pi} \sin^{-1}\rho .$$

It may similarly be shown that the value of the last term in (6.2.1), when $u_1 = u_2 = \infty$, is given by

$$\frac{1}{4} - \frac{1}{2\pi} \sin^{-1}\rho .$$

We see from (6.1.2) that when H_o is true the asymptotic distribution of the maximum likelihood estimator of the scalar parameter under test has a normal density on the positive real axis. However for the case $t > 1$, it is easily seen from (6.2.1) that under H_o the asymptotic distribution of any one of the maximum likelihood estimators takes the form (using $\hat{\theta}_{n1}$ for example),

$$\operatorname{pr}(n^{\frac{1}{2}}\hat{\theta}_{n1} < u_1;\theta_o) = \frac{1}{2} + \frac{1}{2\pi}\sin^{-1}\rho$$

$$+ \frac{1}{2}\frac{1}{(2\pi)^{\frac{1}{2}}}\frac{1}{\sigma_1}\int_o^{u_1}\exp\left(-\frac{1}{2\sigma_1^2}\cdot t_1^2\right)dt_1$$

$$+ \frac{1}{(2\pi)^{\frac{1}{2}}}\cdot\frac{1}{d_{11}^{\frac{1}{2}}}\int_o^{u_1}\left\{1-\phi^*\left(\frac{-t_1 d_{12}}{d_{11}^{\frac{1}{2}}|\underset{\sim}{d}_{11}|^{\frac{1}{2}}}\right)\right\}\exp\left(-\frac{1}{2}\frac{t_1^2}{d_{11}}\right)dt_1 \qquad (6.2.3)$$

and (6.2.3) is obviously non-normal.

6.3 MEAN RESTRICTED MULTIVARIATE NORMAL DISTRIBUTION

The theory of sections (6.1) and (6.2) is asymptotically equivalent to testing the mean of a multivariate normal distribution subject to restrictions on the mean (see, for example, Birnbaum (1950), Brunk (1958), Perlman (1969)).

As an example, consider the following exact problem where we take a random sample $\underset{\sim}{X}_1,\ldots,\underset{\sim}{X}_n$ from a bivariate normal distribution with means μ_1 and μ_2, say, unit variances and unknown correlation coefficient ρ. We restrict the means to be positive. We have

$$\underset{\sim}{c} = \begin{matrix} \mu_1 \\ \mu_2 \\ \rho \end{matrix}\begin{bmatrix} \frac{1}{1-\rho^2} & \frac{-\rho}{1-\rho^2} & 0 \\ \frac{-\rho}{1-\rho^2} & \frac{1}{1-\rho^2} & 0 \\ 0 & 0 & \frac{1+\rho^2}{(1-\rho^2)^2} \end{bmatrix}. \qquad (6.3.1)$$

Case (i). We test $H_o : \mu_1 = 0$, μ_2, ρ fixed and unknown in an open subset of Θ^*.

The asymptotic distribution function of $n^{\frac{1}{2}}\hat{\mu}_{n1}$, where $\hat{\mu}_{n1}$ is the maximum likelihood estimator of μ_1, is given by (6.1.1). In this case we see from (6.1.1) and (6.3.1) that the relevant d_{11} is given by

$$d_{11} = \frac{+1}{1-\rho^2} - \begin{bmatrix} \frac{-\rho}{1-\rho^2} & 0 \end{bmatrix} \begin{bmatrix} 1-\rho^2 & 0 \\ 0 & \frac{(1-\rho^2)^2}{(1+\rho^2)} \end{bmatrix} \begin{bmatrix} \frac{-\rho}{1-\rho^2} \\ 0 \end{bmatrix}$$

$$= 1.$$

Hence we see that

$$\text{pr}\{n^{\frac{1}{2}}\hat{\mu}_{n1} < u_1; 0, \mu_2, \rho\} = \frac{1}{2} + \frac{1}{(2\pi)^{\frac{1}{2}}} \int_o^{u_1} \exp(-\frac{t^2}{2})dt,$$

independently of ρ and μ_2.

Case (ii). We test $H_{o2} : \mu_1 = \mu_2 = 0$, ρ arbitrary.

The asymptotic distribution function of $n^{\frac{1}{2}}\hat{\mu}_{n1}$ under H_{o2} is given by (6.2.3). The relevant terms of (6.2.3) become

$$\sigma_1^2 = 1-\rho^2$$

$$\underset{\sim}{d}_{11} = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$$

$$|\underset{\sim}{d}_{11}| = 1-\rho^2.$$

Hence we have

$$\text{pr}\{n^{\frac{1}{2}}\hat{\mu}_{n1} < u_1; 0, \rho\} = \frac{1}{2} + \frac{1}{2\pi} \sin^{-1}\rho_o$$

$$+ \frac{1}{(2\pi)^{\frac{1}{2}}} \frac{1}{(1-\rho^2)^{\frac{1}{2}}} \int_o^{u_1} \exp\{ \frac{-t_1^2}{2(1-\rho^2)} \} dt_1$$

$$+ \frac{1}{(2\pi)^{\frac{1}{2}}} \int_o^{u_1} [1-\phi^*\{ \frac{\rho t_1}{(1-\rho^2)^{\frac{1}{2}}} \}] \exp(\frac{-t_1^2}{2}) dt_1 .$$

where ρ_o is given by (6.2.2). Hence

$$\rho_o = -\rho \quad .$$

6.4 HOMOGENEITY TESTS

The Poisson homogeneity test considered in Bartoo and Puri (1967), Moran (1973a) and Klonecki (1973) concerns a null hypothesis specifying that $X_1,\ldots,X_n$ is a sample of independently and identically distributed random variables with a Poisson distribution of unknown mean λ. The alternative hypothesis specifies that X_i has a compound distribution

$$f_X(x) = \frac{1}{x!} \int v^x e^{-v} dF(v)$$

where $F(v)$ is a non degenerate distribution with mean λ. $F(v)$ is assumed to have the form $F(v) = G\{(v-\lambda)/\alpha^{\frac{1}{2}}\}$ and the test of $H_o : \alpha = 0$ treated by $C(\alpha)$ techniques or by maximum likelihood leads to the same test statistic

$$T = (2\tilde{\lambda}^2 n)^{-\frac{1}{2}} \sum_{i=1}^{n} \{(X_i-\tilde{\lambda})^2 - X_i\}$$

where $\tilde{\lambda}$ is a locally root-n consistent estimator of λ, in particular the mean of the observations. The results of chapter 5 show that rejection regions for the two tests are asymptotically the same. To ensure the validity of these tests we must have that

$$m_3 = \int (v-\lambda)^3 dF(v) = 0$$

and the test is then robust for functions $F(v)$.

The condition $m_3 = 0$ holds in a more general formulation where a set of independently and identically distributed random variables have density $f(x;\lambda)$ under a null hypothesis, and under the alternative hypothesis have the compound distribution

$$f^*(x;\alpha,\lambda) = \int f(x;v)dF(v) \tag{6.4.1}$$

with the same assumptions on $F(v)$ as above. For a particular choice of $f(x;\lambda)$ and $F(v)$ the question of the identifiability of $F(v)$ in the mixture $f^*(x;\alpha,\lambda)$ must be examined. As shown in, for example, Moran (1973a), $F(v)$ is identifiable for the Poisson homogeneity test. Maritz (1970) shows that restrictions on the family to which $F(v)$ belongs can render $F(v)$ identifiable and that for certain $f(x;\lambda)$, $F(v)$ is always identifiable. In the following we shall assume identifiability of $F(v)$ and the validity of a Taylor expansion of (6.4.1) about λ, giving

$$f^*(x;\alpha,\lambda) = f(x;\lambda)\{1 + \sum_{r=1}^{4} \frac{\alpha^{\frac{r}{2}}}{r!} \frac{f^{(r)}(x;\lambda)}{f(x;\lambda)} m_r$$

$$+ \frac{\alpha^{\frac{5}{2}}}{5!} \frac{f^{(5)}(x;\lambda^*)}{f(x;\lambda)} m_5\}$$

where

$$f^{(r)}(x;\lambda) = \frac{d^r}{d\lambda^r} f(x;\lambda) \quad \text{and} \quad m_r = \int u^r dG(u).$$

Let $\ell(\alpha,\lambda)$ be the log likelihood of the sample, then

$$\frac{\partial \ell}{\partial \alpha}(\alpha,\lambda) = \frac{\alpha^{-\frac{1}{2}}}{2} \frac{f^{(1)}(x;\lambda)}{f(x;\lambda)} m_1 + \frac{1}{2} \frac{f^{(2)}(x;\lambda)}{f(x;\lambda)} m_2 + o(1)$$

as $\alpha \to 0$.

We have specified that $m_1 = 0$ and hence $\frac{\partial \ell}{\partial \alpha}(\alpha,\lambda)$ is continuous for all α.

We have

$$\frac{\partial \ell}{\partial \alpha}(0,\lambda) = \frac{1}{2}\frac{f^{(2)}(x;\lambda)}{f(x;\lambda)} m_2 .$$

Now,

$$\mathrm{var}\{\frac{\partial \ell}{\partial \alpha}(0,\lambda)\} = E\left|\{\frac{\partial \ell}{\partial \alpha}(\alpha,\lambda)\}^2\right|_{\alpha=0} = -E\{\frac{\partial^2 \ell}{\partial \alpha^2}(\alpha,\lambda)\}\Bigg|_{\alpha=0}$$

the latter equality holding only if $\frac{\partial^2 \ell}{\partial \alpha^2}(\alpha,\lambda)$ is continuous for all (α,λ) in parameter space.

$$\frac{\partial^2 \ell}{\partial \alpha^2}(\alpha,\lambda) = \frac{\alpha^{-\frac{1}{2}}}{8}\frac{f^{(3)}(x;\lambda)}{f(x;\lambda)} m_3 + \frac{1}{12}\frac{f^{(4)}(x;\lambda)}{f(x;\lambda)} m_4$$

$$- \frac{1}{4}\{\frac{f^{(2)}(x;\lambda)}{f(x;\lambda)} m_2\}^2 + o(1)$$

as $\alpha \to 0$.

Again, we must impose the condition that $m_3 = 0$ in order that the maximum likelihood and optimal $C(\alpha)$ theories involving second derivatives of the log likelihood may be employed. Otherwise the random variable $\frac{\partial^2 \ell}{\partial \alpha^2}(\alpha;\lambda)$ becomes arbitrarily large in a neighbourhood of $\alpha = 0$ and standard arguments concerning the limiting distributional forms of test statistics break down. Under the null hypothesis $H_o : \alpha = 0$ we have

$$c_{\alpha\alpha} = \frac{nm_2^2}{4} E_o\{\frac{f^{(2)}(X;\lambda)}{f(X;\lambda)}\}^2 \qquad (6.4.2)$$

$$c_{\alpha\lambda} = \frac{nm_2}{2} E_o\{\frac{f^{(2)}(X;\lambda)}{f(X;\lambda)} \cdot \frac{f^{(1)}(X;\lambda)}{f(X;\lambda)}\}$$

$$c_{\lambda\lambda} = -n\, E_o\{ \frac{\partial^2}{\partial\lambda^2} \log f(X;\lambda)\} .$$

E_o denotes expectation taken with respect to the density $f(x;\lambda)$.
The optimal $C(\alpha)$ test of $H_o : \alpha = 0$ is based upon

$$T^{(C)} = \left. \frac{\{ \frac{\partial \ell}{\partial \alpha}(0,\lambda) - \frac{c_{\alpha\lambda}}{c_{\lambda\lambda}} \frac{\partial \ell}{\partial \lambda}(0,\lambda)\}}{(c_{\alpha\alpha} - \frac{c_{\alpha\lambda}^2}{c_{\lambda\lambda}})^{\frac{1}{2}}} \right|_{\lambda=\tilde{\lambda}_n} \tag{6.4.3}$$

and $T^{(C)}$ is independent of m_2.

With $c_{\alpha\alpha}$ as defined in (6.4.2) we may still use the test statistic (6.4.3) to test the null hypothesis even if $m_3 \neq 0$. However optimal properties in the $C(\alpha)$ case are only known to hold if $m_3 = 0$, and any optimal properties of $T^{(C)}$ when $m_3 \neq 0$ are unclear.

6.5 DISCUSSION

Restricting ourselves to a space Θ^* within which the maximum likelihood estimator is uniformly consistent, and to a null hypothesis asserting that $\underset{\sim}{\theta}$ lies on the boundary of Θ^* as described in chapter 5, then the optimal $C(\alpha)$ test is clearly unaffected by the condition that $\underset{\sim}{\theta}$ lies on a boundary and retains its asymptotic properties as described in section (2.2) and Buhler and Puri (1966).

In the maximum likelihood analysis the inference procedures are asymptotically equivalent to testing the mean of a multivariate normal distribution subject to restrictions on the means. The pair $\underset{\sim}{S}$ = (sample means vector, sample variance-covariance matrix) no longer forms a sufficient statistic for the problem and the usual test for the mean no longer has any optimal invariant properties. So the choice of tests is unclear. In the case $t = 1$, where a scalar parameter is under test, the optimal $C(\alpha)$ tests lead asymptotically to the same rejection regions as tests based on maximum likelihood estimators for one sided tests. For example, the Poisson homogeneity tests considered in Moran (1973b), Klonecki (1973) and section (6.4). However in the case $t > 1$, where the null hypothesis specifies the values of more than one parameter, this asymptotic equivalence no longer obtains, even if only one of the parameters lies on the boundary. We do not know what rejection regions to choose as we have a composite alternative. An example of this situation is furnished by the homogeneity tests for the mean of a gamma distribution where the null hypothesis also specifies the value of the scale or location parameter, see for example, Moran (1973a) and with applications to rain making experiments in Moran (1970a). A further discussion of the power of the two testing procedures is necessary to decide on which test to use in a particular case.

Moran (1971b) also considers the so-called "pseudo" maximum likelihood estimator $\underset{\sim}{V}' = [V_1, \ldots V_s]$ which is the unique solution of the set of equations

$$Y_i = \sum_{j=1}^{s} c_{ij} V_j \qquad , \qquad i = 1,\ldots,s.$$

As the $\{Y_i\}$ are asymptotically $N_s(\underset{\sim}{0},\underset{\sim}{c})$, the $\{V_{ij}\}$ are asymptotically $N_s(\underset{\sim}{0},\underset{\sim}{c}^{-1})$ and coincide asymptotically with the maximum likelihood point of $\underset{\sim}{\theta}_i + n^{-\frac{1}{2}} V_i$ lies in Θ^*, all i.

Asymptotically the $C(\alpha)$ test is equivalent to a test based on the pseudo maximum likelihood estimators as these are not restricted to lie in Θ^*, but lie in an open set Θ. We note that it is therefore possible for the optimal $C(\alpha)$ test with the usual rejection region to lead to a significant deviation even if the maximum likelihood point is equal to the true parameter value.

APPENDIX A.1

SIMULATIONS ON THE MIXTURE PROBLEM

This appendix contains the results of simulation experiments performed on the model (3.2.1). The formulae used in deriving these results are fully covered in chapter 3.

Appendix (A.1.1) gives results for a three parameter maximum likelihood analysis only, whilst appendix (A.1.2) contains a comparism of tests for the mean of the unspecified component of the mixture.

Certain abbreviations are used in the appendix as follows:-

(i) the notation $c^{\alpha\alpha}$, $c^{\sigma^2\alpha}$, etc. denotes the asymptotic variance of the estimates of α, the asymptotic covariance of the estimates of σ^2 and α, etc.

(ii) (1) denotes that the moment estimator $\tilde{\alpha}$ was recalculated using an estimator based on order statistics as described in section (3.3),

(2) denotes that the estimate $\tilde{\sigma}^2$ was recalculated using the order statistics as described in section (3.2),

(3) denotes that the set of equations (3.4.4) did not converge in the three parameter treatment,

(4) denotes that the estimator $\tilde{\tilde{\alpha}}$ was taken to be $\tilde{\alpha}$, as described in section (3.3),

(5) denotes that the estimator $\tilde{\tilde{\sigma}}^2$ was taken to be $\tilde{\sigma}^2$ as described in section (3.3),

(6) denotes that the set of equations (3.4.6) and (3.4.7) did not converge in the $C(\alpha)$ situation of section (3.4).

In appendix (A.1.1) the first line of a set of results gives the initial estimates. The second line gives the maximum likelihood estimates and the elements of the variance-covariance matrix. In appendix (A.1.2) the first and second lines give the initial and maximum likelihood estimates respectively in the three parameter case. The third and fourth line give the same for the two parameter $C(\alpha)$ case.

A.1.1 FULL MAXIMUM LIKELIHOOD RESULTS FOR μ = 2 AND 4

n	α	σ^2	μ	$c^{\alpha\alpha}$	$c^{\sigma^2\alpha}$	$c^{\sigma^2\sigma^2}$	$c^{\mu\alpha}$	$c^{\mu\sigma^2}$	$c^{\mu\mu}$
sim. val.	·1	·25	2·0						
1000	.2181	1·6728	·9649						
	·0945	·1722	2·1879	10^{-5}	·0000	·0001	·0000	·0001	·0020
750	·0514	1·2305	·7611						
	·1131	·1514	2·1763	2×10^{-5}	·0000	·0003	·0000	·0003	·0032
500	·2634	1·7296	·9243						
	·1121	·1772	2·1679	3×10^{-5}	·0000	·0007	·0000	–·0002	·0035
100	·6445	1·2780	·6919						
(3)	–	–	–						
sim. val.	·2	·25	2·0						
1000	·5516	1·5894	·7694						
	·2059	·2581	2·0059	2×10^{-5}	·0000	10^{-5}	·0000	–·0001	·0016
750	·4326	1·4691	·9893						
	·1971	·2857	2·0529	4×10^{-5}	·0000	8×10^{-5}	·0000	–·0002	·0025
500	·5847	1·5655	·7888						
	·2242	·3205	1·9754	7×10^{-5}	–·0001	10^{-4}	·0000	–·0004	·0041
100	·2829	·1842	2·0165						
	·2240	·1218	2·2476	3×10^{-4}	–·0034	·0007	·0000	–·0021	·0548
sim. val.	·3	·25	2·0						
1000	·3154	·4028	1·9241						
	·2931	·2593	2·0371	2×10^{-5}	·0000	3×10^{-5}	·0000	·0000	·0011
750	·3314	·6145	1·7809						
	·2805	·2601	2·0473	2×10^{-5}	·0000	4×10^{-5}	·0000	–·0001	·0015
500	·3264	·4443	1·8692						
	·2884	·2104	2·0537	2×10^{-5}	·0000	5×10^{-5}	·0000	–·0001	·0017
100	·4889	·5724	1·7260						
	·3603	·1828	2·1182	·0005	–·0001	10^{-4}	·0000	–·0003	·0054
sim. val.	·4	·25	2·0						
1000	·4167	·3892	1·9407						
	·3928	·2181	2·0499	10^{-5}	·0000	2×10^{-5}	·0000	·0000	·0006
750	·4045	·3089	2·0222						
	·4031	·2507	2·0464	10^{-5}	·0000	5×10^{-5}	·0000	·0000	·0009
500	·4169	·2712	1·9882						
	·4175	·2492	1·9937	2×10^{-5}	·0000	7×10^{-5}	·0000	–·0001	·0014
100	·6442	·8155	1·5635						
	·4316	·1994	2·1200	10^{-4}	–·0001	10^{-4}	·0000	–·0003	·0050

n	α	σ^2	μ	$c^{\alpha\alpha}$	$c^{\sigma^2\alpha}$	$c^{\sigma^2\sigma^2}$	$c^{\mu\alpha}$	$c^{\mu\sigma^2}$	$c^{\mu\mu}$
sim. val.	·5	·25	2·0						
1000	·4857	·2169	2·0361						
	·4949	·2392	2·0137	10^{-5}	·0000	3×10^{-5}	·0000	·0000	·0006
750	·4936	·3520	1·9187						
	·4791	·2616	1·9740	10^{-5}	·0000	5×10^{-5}	·0000	·0000	·0009
500	·5995	·5055	1·7894						
	·5263	·2666	1·9825	2×10^{-5}	·0000	7×10^{-5}	·0000	−·0001	·0012
100	·5836	·1514	2·1192						
	·5565	·2329	2·1378	6×10^{-5}	−·0001	2×10^{-5}	·0000	−·0003	·0045
sim. val.	·6	·25	2·0						
1000	·6321	·4302	1·8534						
	·5826	·2595	1·9804	10^{-5}	·0000	3×10^{-5}	·0000	·0000	·0005
750	·6047	·3063	1·9561						
	·5944	·2699	1·9825	10^{-5}	·0000	5×10^{-5}	·0000	·0000	·0007
500	·6442	·3437	1·9349						
	·6156	·2744	1·9974	2×10^{-5}	·0000	7×10^{-5}	·0000	·0000	·0011
100	·8157	·5231	1·8583						
	·6870	·2229	2·1090	3×10^{-5}	·0000	10^{-4}	·0000	−·0002	·0035
sim. val.	·7	·25	2·0						
1000	·6737	·2248	2·0289						
	·6806	·2442	2·0135	10^{-6}	·0000	2×10^{-6}	·0000	·0000	·0004
750	·7281	·2880	1·9986						
	·7186	·2456	2·0246	10^{-6}	·0000	3×10^{-6}	·0000	·0000	·0005
500	·7533	·4023	1·9256						
	·7063	·2734	2·0251	2×10^{-6}	·0000	6×10^{-6}	·0000	−·0001	·0009
100	·8901	·6237	1·7847						
	·7238	·3221	2·0679	10^{-5}	−·0002	4×10^{-5}	·0000	−·0006	·0051
sim. val.	·8	·25	2·0						
1000	·7745	·2051	2·0369						
	·7906	·2439	2·0059	10^{-6}	·0000	2×10^{-6}	·0000	·0000	·0003
750	·8071	·2946	1·9611						
	·7921	·2459	1·9933	10^{-6}	·0000	$3\ 10^{-6}$	·0000	·0000	·0005
500	·7826	·3026	2·0151						
	·7693	·2528	2·0469	10^{-6}	·0000	4×10^{-6}	·0000	−·0001	·0007
100	·7784	·0802	2·1681						
	·7985	·2355	2·0943	10^{-5}	·0000	10^{-5}	·0000	−·0002	·0031

n	α	σ^2	μ	$c^{\alpha\alpha}$	$c^{\sigma^2\alpha}$	$c^{\sigma^2\sigma^2}$	$c^{\mu\alpha}$	$c^{\mu\sigma^2}$	$c^{\mu\mu}$
sim. val.	·9	·25	2·0						
1000	·8884	·2639	1·9979						
	·8842	·2492	2·0069	10^{-6}	·0000	2×10^{-6}	·0000	·0000	·0003
750	·9252	·2916	1·9801						
	·9110	·2582	2·0043	10^{-6}	·0000	2×10^{-6}	·0000	·0000	·0004
500	·9100	·2563	1·9995						
	·9132	·2718	1·9915	10^{-6}	·0000	4×10^{-6}	·0000	–·0001	·0006
100	·9640	·2751	2·0989						
	·9329	·2128	2·1477	3×10^{-5}	·0000	9×10^{-6}	·0000	–·0002	·0023
sim. val.	·1	1·0	2·0						
1000	·1540	1·8215	1·4702						
	·1140	·9811	2·0365	·0001	–·003	·0252	·0015	·0172	·0064
750	·3890	1·5837	·4081						
(3)	–	–	–						
500	·0357	1·6778	·4549						
(3)	–	–	–						
100	·2731	·3757	1·4377						
(3)	–	–	–						
sim. val.	·2	1·0	2·0						
1000	·3006	1·9581	1·4783						
	·1945	·8597	2·2592	5×10^{-5}	–·0010	·0034	·0004	·0007	·0075
750	·0119	2·0369	·8934						
	·1599	·6229	2·5578	·0002	–·0006	·0015	·0002	–·0002	·0082
500	·3114	1·7213	·9989						
(3)	–	–	–						
100	·1561 (1)	1·0220	1·6655						
(3)	–	–	–						
sim. val.	·3	1·0	2·0						
1000	·4215	2·0135	1·3483						
	·2632	·7322	2·1980	$2\ 10^{-5}$	–·0004	·0006	·0001	–·0005	·0049
750	·2112 (1)	1·8464	1·2454						
(3)	–	–	–						
500	·2312 (1)	1·8927	1·3350						
(3)	–	–	–						
100	·5482	1·8737	1·5554						
	·3082	·5964	2·5708	·0001	–·0021	·0009	·0005	–·0040	·0286

n	α	σ^2	μ	$c^{\alpha\alpha}$	$c^{\sigma^2\alpha}$	$c^{\sigma^2\sigma^2}$	$c^{\mu\alpha}$	$c^{\mu\sigma^2}$	$c^{\mu\mu}$
sim. val.	·4	1·0	2·0						
1000	·5533	1·7850	1·4791						
(3)	–	–	–						
750	·6026	1·7325	1·3608						
	·3804	·8300	2·1004	3×10^{-5}	–·0006	·0004	·0002	–·0009	·0052
500	·5800	1·8206	1·4541						
	·3601	·7650	2·2698	4×10^{-5}	–·0007	·0005	·0002	–·0012	·0073
100	·6826	2·0766	1·6007						
	·3996	·7042	2·5909	·0001	–·0025	·0009	·0005	–·0047	·0266
sim. val.	·5	1·0	2·0						
1000	·6933	1·7698	1·3822						
	·4811	1·3012	1·9110	·0001	–·0003	·0007	·0002	–·0007	·0068
750	·5999	1·2124	1·8067						
	·5638	1·0747	1·9135	6×10^{-5}	–·0008	·0004	·0001	–·0015	·0046
500	·5158	·9314	2·2011						
	·5311	·9838	2·1525	6×10^{-5}	–·0010	·0005	·0001	–·0018	·0064
100	·4111	·2186	2·2655						
(3)	–	–	–						
sim. val.	·6	1·0	2·0						
1000	·6261	1·0942	1·9394						
	·6106	1·0133	1·9888	3×10^{-5}	–·0005	·0001	·0000	–·0009	·0028
750	·6493	1·2195	1·8296						
	·6116	1·0605	1·9388	5×10^{-5}	–·0007	·0002	·0000	–·0014	·0040
500	·6853	1·3559	1·7737						
(3)	–	–	–						
100	·5279	·8048	2·3727						
	·5412	·9604	2·3043	·0002	–·0044	·0021	·0007	–·0081	·0296
sim. val.	·7	1·0	2·0						
1000	·6754	1·0569	2·0201						
	·6988	1·1118	1·9645	4×10^{-5}	–·0005	8×10^{-5}	·0000	–·0010	·0025
750	·7415	1·1642	2·4136						
	·7312	1·1413	1·9703	3×10^{-5}	–·0007	6×10^{-5}	·0000	–·0027	·0061
500	·7304	1·0731	1·9002						
	·7113	1·1213	1·8732	3×10^{-5}	–·0011	·0009	·0000	–·0046	·0082
100	·7316	1·1882	2·2316						
	·6802	·8844	2·3978	·0002	–·0027	·0004	·0002	–·0055	·0193

n	α	σ^2	μ	$c^{\alpha\alpha}$	$c^{\sigma^2\alpha}$	$c^{\sigma^2\sigma^2}$	$c^{\mu\alpha}$	$c^{\mu\sigma^2}$	$c^{\mu\mu}$
sim. val.	·8	1·0	2·0						
1000	·8352	1·1108	1·9568						
	·8160	1·0370	2·0019	3×10^{-5}	−·0004	2×10^{-5}	·0000	−·0008	·0018
750	·8436	1·1334	1·9077						
	·8217	1·0573	1·9567	4×10^{-5}	−·0004	10^{-5}	·0000	−·0011	·0024
500	·8395	1·2143	1·9813						
	·8043	1·8029	2·0649	6×10^{-5}	−·0007	9×10^{-5}	·0000	−·0017	·0038
100	·8512	1·0488	2·1530						
	·8214	·9124	2·2284	·0002	−·0025	·0002	·0000	−·0053	·0149
sim. val.	·9	1·0	2·0						
1000	·8943	1·0298	2·0065						
	·8956	1·0235	2·0055	3×10^{-5}	−·0003	4×10^{-5}	·0000	−·0007	·0014
750	·9354	1·1065	1·8971						
	·9211	1·0671	1·9247	4×10^{-5}	−·0004	6×10^{-5}	·0000	−·0010	·0018
500	·9499	1·1681	1·9573						
	·9470	1·1635	1·9625	6×10^{-5}	−·0007	·0003	−·0001	−·0018	·0026
100	·8653	·9967	2·2797						
	·8553	·9159	2·3121	·0001	−·0023	·0004	·0000	−·0050	·0138
sim. val.	·1	2·25	2·0						
1000	·5011	1·8082	·4119						
(3)	−	−	−						
750	·2111 (1)	1·7846	·4158						
(3)	−	−	−						
500	·2069 (1)	2·2301	·5531						
(3)	−	−	−						
100	·2173 (1)	2·4240	·9372						
(3)	−	−	−						
sim. val.	·2	2·25	2·0						
1000	·7013	2·2854	·7469						
(3)	−	−	−						
750	·4112	2·2917	·7382						
(3)	−	−	−						
500	·5013 (1)	2·4594	·9244						
(3)	−	−	−						
100	·4716 (1)	1·6909	·9781						
(3)	−	−	−						

n	α	σ^2	μ	$c^{\alpha\alpha}$	$c^{\sigma^2\alpha}$	$c^{\sigma^2\sigma^2}$	$c^{\mu\alpha}$	$c^{\mu\sigma^2}$	$c^{\mu\mu}$
sim. val.	·3	2·25	2·0						
1000	·0110	2·7791	1·3269						
(3)	–	–	–						
750	·1093	2·6245	1·3083						
(3)	–	–	–						
500	·4110[1]	3·2505	1·3504						
(3)	–	–	–						
100	·3876[1]	2·7066	1·5330						
(3)	–	–	–						
sim. val.	·4	2·25	2·0						
1000	·3101	2·2459	1·5177						
	·4211	2·2563	1·9879	·0003	-·0004	·0002	·0001	-·0014	·0048
750	·5811	2·7180	1·5762						
	·3913	2·4497	1·9908	·0003	-·0069	·0388	·0013	-·0003	·0190
500	·5107[1]	2·5041	1·7975						
(3)	–	–	–						
100	·5155	1·7008	1·9239						
(3)	–	–	–						
sim. val.	·5	2·25	2·0						
1000	·2111[1]	2·0604	2·0014						
(3)	–	–	–						
750	·3156[1]	2·2411	2·1214						
(3)	–	–	–						
500	·5693	·2982	1·9369						
(3)	–	–	–						
100	·6766	2·5610	2·1236						
(3)	–	–	–						
sim. val.	·6	2·25	2·0						
1000	·0113	1·3528	2·3728						
(3)	–	–	–						
750	·4831	1·7161	2·1109						
	·6105	2·3332	2·0104	·0001	-·0009	·0023	·0001	-·0098	·0084
500	·4980	1·5071	2·5122						
	·6307	2·5493	1·9699	·0004	-·0061	·0166	·0003	-·0115	·0165
100	·7111[1]	1·3922	2·7629						
(3)	–	–	–						

n	α	σ^2	μ	$c^{\alpha\alpha}$	$c^{\sigma^2\alpha}$	$c^{\sigma^2\sigma^2}$	$c^{\mu\alpha}$	$c^{\mu\sigma^2}$	$c^{\mu\mu}$
sim. val.	·7	2·25	2·0						
1000	·5011	·2619	2·9099						
	·6351	2·1013	2·1101	·0013	·0003	·0021	·0000	–·0013	·0088
750	·4159(1)	·5124	2·7693						
	·6996	2·3255	1·9599	·0002	–·0029	·0047	·0000	–·0065	·0082
500	·6387	·6444	2·8415						
	·6967	2·4623	2·0223	·0003	–·0048	·0099	·0000	–·0108	·0132
100	·6561	1·1855	3·5331						
	·9097	2·8214	1·7968	·0014	–·0191	·0131	–·0032	–·0626	·0325
sim. val.	·8	2·25	2·0						
1000	·7169	1·0151	3·2652						
	·8814	2·3823	1·5974	·0001	–·0015	9×10^{-5}	–·0003	–·0044	·0032
750	·6993	1·0635	3·1907						
	·8563	2·2881	1·5285	·0002	–·0021	·0003	–·0003	–·0057	·0045
500	·6017	1·0921	3·3978						
	·8878	2·3055	1·6724	·0002	–·0028	·0005	–·0004	–·0083	·0061
100	·8113	2·0768	2·4335						
(3)	–	–	–						
sim. val.	·9	2·25	2·0						
1000	·8711	·9594	3·6388						
	·9498(1)	2·3024	1·7969	9×10^{-5}	–·0011	·0002	–·0002	–·0003	·0023
750	·9591(1)	·9586	3·6586						
	·9412	2·2758	1·8069	·0001	–·0015	–·0003	–·0003	–·0049	·0031
500	·8483	1·0298	3·6393						
	·9293	2·3391	1·7821	·0002	–·0025	·0004	–·0004	–·0079	·0051
100	·8087	1·3360	2·7011						
(3)	–	–	–						
sim. val.	·1	·25	4·0						
1000	·1752	3·8950	2·4602						
	·1067	·2897	3·9630	10^{-5}	·0000	·0001	·0000	·0000	·0008
750	·2087	4·1342	2·4755						
	·1273	·2643	4·0087	2×10^{-5}	·0000	·0002	·0000	·0001	·0021
500	·2092	3·5361	2·3791						
	·1136	·2414	3·9634	2×10^{-5}	·0000	·0006	·0000	–·0001	·0035
100	·1816(1)	3·9584	1·4405						
	·1394	·2983	4·3783	·0009	·0000	·0012	·0000	·0006	·0210

n	α	σ^2	μ	$c^{\alpha\alpha}$	$c^{\sigma^2\alpha}$	$c^{\sigma^2\sigma^2}$	$c^{\mu\alpha}$	$c^{\mu\sigma^2}$	$c^{\mu\mu}$
sim. val.	·2	·25	4·0						
1000	·3557	4·2309	2·4317						
	·2164	·2644	3·9892	10^{-5}	·0000	3×10^{-5}	·0000	·0000	·0007
750	·3409	4·3799	2·4427						
	·2084	·2611	4·0164	2×10^{-5}	·0000	10^{-5}	·0000	·0000	·0011
500	·3891	4·2654	2·5059						
	·2389	·3054	4·0529	10^{-5}	·0001	·0001	·0000	−·0002	·0027
100	·3872	3·3379	2·5161						
	·2004	·3016	4·1829	·0018	−·0003	·0024	−·0001	·0004	·0119
sim. val.	·3	·25	4·0						
1000	·5008	4·2662	2·4238						
	·3014	·2409	3·9984	10^{-5}	·0000	2×10^{-5}	·0000	·0000	·0006
750	·5329	4·3998	2·4316						
	·3252	·2572	4·0108	10^{-5}	·0000	·0001	·0000	·0001	·0019
500	·4652	4·4653	2·4243						
	·2771	·2267	4·0425	2×10^{-5}	·0002	·0001	·0003	·0001	·0013
100	·4457	3·2977	2·4789						
	·2208	·1440	4·1869	·0004	·0006	·0004	−·0002	−·0011	·0087
sim. val.	·4	·25	4·0						
1000	·6502	4·5518	2·4148						
	·3945	·2281	4·0253	10^{-5}	·0000	10^{-5}	·0000	·0000	·0005
750	·6639	4·4930	2·4346						
	·4044	·2383	4·0274	10^{-5}	·0000	2×10^{-5}	·0000	·0011	·0014
500	·6836	3·9881	2·4670						
	·4118	·3186	3·9942	2×10^{-5}	·0001	2×10^{-5}	·0004	−·0007	·0012
100	·6938	3·5225	2·3672						
	·3580	·1996	4·0432	·0003	−·0003	·0004	·0000	·0011	·0069
sim. val.	·5	·25	4·0						
1000	·5147	·6291	3·8255						
	·4877	·2650	3·9841	10^{-5}	·0000	10^{-5}	·0000	·0000	·0002
750	·5144	·5295	3·9219						
	·5053	·2805	4·0087	2×10^{-5}	−·0004	3×10^{-5}	−·0011	·0000	·0014
500	·4840	·5960	3·6852						
	·4598	·2819	4·0122	3×10^{-5}	·0000	4×10^{-5}	·0000	·0001	·0009
100	·5812	·8087	3·6908						
	·4978	·2196	4·0367	·0004	−·0002	·0003	·0000	·0021	·0098

n	α	σ^2	μ	$c^{\alpha\alpha}$	$c^{\sigma^2\alpha}$	$c^{\sigma^2\sigma^2}$	$c^{\mu\alpha}$	$c^{\mu\sigma^2}$	$c^{\mu\mu}$
sim. val.	·6	·25	4·0						
1000	·5789	·1935	4·0610						
	·5893	·2493	4·0230	10^{-5}	·0000	2×10^{-5}	·0000	·0000	·0005
750	·5806	·3135	3·9941						
	·5871	·2713	4·0012	10^{-5}	·0000	4×10^{-5}	·0000	·0000	·0006
500	·4840	·5960	3·8653						
	·4598	·2819	4·0122	2×10^{-5}	·0000	6×10^{-5}	·0000	·0000	·0009
100	·7301	·9876	3·6896						
	·6269	·2523	4·0693	2×10^{-5}	·0000	10^{-4}	·0000	–·0002	·0027
sim. val.	·7	·25	4·0						
1000	·6853	·1607	4·0502						
	·6944	·2419	4·0131	10^{-6}	·0000	10^{-6}	·0000	·0000	·0003
750	·7118	·3269	3·9757						
	·7065	·2536	4·0019	10^{-6}	·0000	2×10^{-6}	·0000	·0000	·0005
500	·6674	·2929	4·0339						
	·6665	·2609	4·0429	2×10^{-6}	·0000	5×10^{-6}	·0000	–·0001	·0010
100	·6723	·2149(2)	4·2119						
	·6401	·2296	4·1327	10^{-5}	–·0002	3×10^{-5}	·0000	–·0005	·0047
sim. val.	·8	·25	4·0						
1000	·7760	·2879	3·9982						
	·7751	·2518	4·0081	10^{-6}	·0000	2×10^{-6}	·0000	·0000	·0002
750	·7752	·2399	4·0249						
	·7792	·2506	4·0158	10^{-6}	·0000	3×10^{-6}	·0000	·0000	·0004
500	·8000	·2669	4·0353						
	·8021	·2580	4·0334	10^{-6}	·0000	3×10^{-6}	·0000	·0000	·0006
100	·8663	·6725	3·8689						
	·7937	·2474	4·0877	10^{-5}	·0000	10^{-5}	·0000	–·0001	·0021
sim. val.	·9	·25	4·0						
1000	·9098	·2985	3·9642						
	·9005	·2446	3·9911	10^{-6}	·0000	2×10^{-6}	·0000	·0000	·0003
750	·8963	·2555	4·0090						
	·8874	·2541	4·0214	10^{-6}	·0000	2×10^{-6}	·0000	·0000	·0003
500	·8755	·1775	4·0737						
	·8879	·2529	4·0367	10^{-6}	·0000	4×10^{-6}	·0000	–·0001	·0005
100	·9628	·4296	3·9989						
	·9274	·2490	4·0927	2×10^{-5}	·0000	9×10^{-6}	·0000	–·0001	·0016

n	α	σ^2	μ	$c^{\alpha\alpha}$	$c^{\sigma^2\alpha}$	$c^{\sigma^2\sigma^2}$	$c^{\mu\alpha}$	$c^{\mu\sigma^2}$	$c^{\mu\mu}$
sim. val.	·1	1·0	4·0						
1000	·1769	4·7471	2·8057						
	·1269	1·0715	4·0573	·0007	−·0019	·0200	·0005	·0038	·0109
750	·1434	3·9149	2·7245						
	·1035	1·2130	3·8127	·0001	−·0049	·0807	·0016	·0295	·0138
500	·1476	3·8303	2·7050						
	·1059	1·2835	3·7802	·0003	−·0079	·1351	·0027	·0518	·0195
100	·0541	3·6169	1·2255						
	·1137	2·8591	4·1490	·0041	−·2307	8·8243	·0788	3·4348	−·7898
sim. val.	·2	1·0	4·0						
1000	·3030	4·5790	2·6278						
	·2010	·8522	3·9961	3×10^{-5}	−·0006	·0015	·0000	−·0007	·0054
750	·2806	4·4857	2·6848						
	·1963	1·0541	3·9279	6×10^{-5}	−·0014	·0066	·0003	−·0003	·0101
500	·2758	4·3606	2·8115						
	·1937	1·2488	4·0082	·0001	−·0030	·0196	·0007	·0017	·0186
100	·2871[(1)]	5·2176	1·6255						
	·1527	1·2098	4·7699	·0006	−·0173	·1476	·0024	·0046	·1017
sim. val.	·3	1·0	4·0						
1000	·4414	4·5924	2·7827						
	·3055	1·1224	4·0532	3×10^{-5}	−·0007	·0016	·0000	−·0008	·0050
750	·4523	4·3206	2·6679						
	·3103	1·0863	3·9085	4×10^{-5}	−·0009	·0019	·0002	−·0011	·0065
500	·5022	4·6036	2·8026						
	·3458	1·1404	4·0806	5×10^{-5}	−·0012	·0024	·0001	−·0017	·0089
100	·4405	3·2138	2·5711						
	·2621	1·0777	3·8690	·0003	−·0080	·0235	·0013	−·0073	·0584
sim. val.	·4	1·0	4·0						
1000	·5946	4·2277	2·6848						
	·4034	1·1388	3·9193	2×10^{-5}	−·0005	·0008	·0000	−·0008	·0039
750	·6188	4·4922	2·6825						
	·4212	1·0403	3·9693	3×10^{-5}	−·0005	·0006	·0000	−·0009	·0043
500	·6107	3·9665	2·6190						
	·4058	1·0762	3·8522	5×10^{-5}	−·0010	·0013	·0001	−·0016	·0072
100	·7169	4·8116	2·8632						
	·4568	·7874	4·3784	9×10^{-5}	−·0017	·0004	·0000	−·0034	·0187

n	α	σ^2	μ	$c^{\alpha\alpha}$	$c^{\sigma^2\alpha}$	$c^{\sigma^2\sigma^2}$	$c^{\mu\alpha}$	$c^{\mu\sigma^2}$	$c^{\mu\mu}$
sim. val.	·5	1·0	4·0						
1000	·5202	·9871	4·0371						
	·5247	·9594	4·0294	10^{-5}	−·0003	·0002	·0000	−·0005	·0022
750	·7592	4·2010	2·6152						
	·5061	1·0202	3·8882	2×10^{-5}	−·0004	·0003	·0000	−·0008	·0034
500	·5408	1·2916	3·9189						
	·5269	1·0511	4·0106	3×10^{-5}	−·0006	·0005	·0000	−·0012	·0049
100	·7394	4·0175	2·8304						
	·4758	1·1640	4·1417	·0002	−·0045	·0051	·0003	−·0077	·0316
sim. val.	·6	1·0	4·0						
1000	·5835	1·0407	4·0384						
	·5862	·9998	4·0388	10^{-5}	−·0002	·0001	·0000	−·0005	·0020
750	·6312	1·1602	3·9647						
	·6246	1·0053	4·0137	2×10^{-5}	−·0003	·0002	·0000	−·0006	·0025
500	·6329	1·1784	3·9045						
	·6197	1·0264	3·9704	3×10^{-5}	−·0006	·0003	·0000	−·0010	·0040
100	·6278	1·7479	3·8428						
	·5449	·8069	4·2672	8×10^{-5}	−·0016	·0003	·0000	−·0031	·0162
sim. val.	·7	1·0	4·0						
1000	·7287	1·1295	4·0019						
	·7237	1·0091	4·0375	10^{-5}	−·0002	·0001	·0000	−·0004	·0016
750	·7001	1·1195	3·9706						
	·6939	1·0019	4·0084	2×10^{-5}	−·0003	·0002	·0000	−·0005	·0022
500	·7237	1·1080	3·9343						
	·7144	1·0115	3·9755	2×10^{-5}	−·0004	·0002	·0000	−·0008	·0033
100	·7778	1·1461	4·1303						
	·7440	·8159	4·2691	6×10^{-5}	−·0011	·0002	·0000	−·0023	·0116
sim. val.	·8	1·0	4·0						
1000	·7964	·9940	4·0233						
	·8001	·9984	4·0140	10^{-5}	−·0002	·0002	·0000	−·0004	·0014
750	·8209	1·0989	3·9514						
	·8128	1·0043	3·9861	2×10^{-5}	−·0003	·0002	·0000	−·0005	·0018
500	·8400	1·2662	3·9230						
	·8209	1·1032	3·9931	2×10^{-5}	−·0004	·0004	·0000	−·0008	·0031
100	·8370	·9182	4·1725						
	·8183	·7547	4·2423	5×10^{-5}	−·0009	·0002	·0000	−·0018	·0095

n	α	σ^2	μ	$c^{\alpha\alpha}$	$c^{\sigma^2\alpha}$	$c^{\sigma^2\sigma^2}$	$c^{\mu\alpha}$	$c^{\mu\sigma^2}$	$c^{\mu\mu}$
sim. val.	·9	1·0	4·0						
1000	·8988	1·0080	3·9968						
	·8980	·9775	4·0034	10^{-5}	–·0002	·0002	·0000	–·0003	·0012
750	·8988	1·0672	3·9888						
	·8948	1·0159	4·0051	2×10^{-5}	–·0002	·0003	·0000	–·0005	·0016
500	·9367	1·1865	3·9789						
	·9215	1·0506	4·0323	2×10^{-5}	–·0004	·0006	·0000	–·0007	·0024
100	·8545	·7454	4·4148						
	·8545	·7454	4·4147	5×10^{-5}	–·0009	·0002	·0000	–·0018	·0093
sim. val.	·1	2·25	4·0						
1000	·1344	4·3922	2·8754						
	·0973	1·5185	3·9751	·0002	–·0065	·1525	·0023	·0618	·0018
750	·1998 [1]	3·2915	·7926						
	·1093	2·6744	3·5063	·0005	–·0305	1·0359	·0115	·4560	·1219
500	·2103 [1]	4·2648	1·0070						
	·1267	2·6780	3·9379	·0006	–·0331	1·0305	·0111	·4013	·0671
100	·2119 [1]	2·6666	1·0689						
(3)	–	–	–						
sim. val.	·2	2·25	4·0						
1000	·4131 [1]	5·1482	1·6921						
	·2200	2·6056	3·8151	·0002	–·0071	·1164	·0018	·0320	·0128
750	·2783	5·2870	3·3183						
	·2262	2·6965	4·1574	·0002	–·0091	·1540	·0021	·0386	·0184
500	·3039 [1]	5·0451	1·7243						
	·2209	2·5367	3·8189	·0003	–·0132	·2068	·0033	·0563	·0262
100	·3116 [1]	5·4028	2·6680						
	·3050	3·3379	4·0994	·0021	–·0829	–1·3946	·0162	·2512	·1518
sim. val.	·3	2·25	4·0						
1000	·3902	4·3983	3·0255						
	·3006	2·2154	3·9015	10^{-5}	–·0032	·0276	·0006	·0034	·0120
750	·3632	4·3465	3·0709						
	·2848	2·3586	3·8925	·0001	–·0052	·0540	·0011	·0090	·0174
500	·4231	4·6599	3·1804						
	·3593	3·0269	3·7755	·0004	–·0113	·1336	·0020	·0158	·0285
100	·2891 [1]	4·4835	2·8863						
	·3443	2·5265	3·9702	·0011	–·0360	·3281	·0067	·0294	·1239

n	α	σ^2	μ	$c^{\alpha\alpha}$	$c^{\sigma^2\alpha}$	$c^{\sigma^2\sigma^2}$	$c^{\mu\alpha}$	$c^{\mu\sigma^2}$	$c^{\mu\mu}$
sim. val.	·4	2·25	4·0						
1000	·5584	4·8229	3·0853						
	·4270	2·1596	4·0498	7×10^{-5}	−·0019	·0099	·0002	−·0013	·0085
750	·5332	4·5968	3·1427						
	·4304	2·6416	3·8926	·0001	−·0042	·0314	·0006	−·0005	·0145
500	·5754	4·6622	3·0103						
	·4454	2·2559	3·9001	·0001	−·0042	·0224	·0006	−·0028	·0175
100	·5697	4·7983	3·3507						
	·4349	2·6460	4·2668	·0080	−·0283	·2111	·0041	−·0075	·1043
sim. val.	·5	2·25	4·0						
1000	·6405	5·0482	3·0786						
	·4895	2·1576	4·0688	6×10^{-5}	−·0016	·0071	·0002	−·0018	·0073
750	·6871	4·5598	2·9625						
	·5296	2·2059	3·8493	8×10^{-5}	−·0022	·0092	·0003	−·0027	·0094
500	·6282	4·8461	3·1418						
	·4885	2·4364	4·0320	·0002	−·0043	·0242	·0005	−·0037	·0172
100	·5398(1)	1·1779	4·2672						
	·5074	2·7204	4·0644	·0008	−·0266	·1764	·0033	−·0201	·0942
sim. val.	·6	2·25	4·0						
1000	·7926	4·8044	3·0189						
	·6073	2·2064	3·9561	5×10^{-5}	−·0014	·0052	·0001	−·0022	·0059
750	·7621	4·6398	3·0619						
	·5925	2·3354	3·9259	8×10^{-5}	−·0023	−·0098	·0002	−·0033	·0089
500	·6441	2·3934	3·9533						
	·6471	2·4096	3·9392	·0001	−·0032	·0134	·0002	−·0051	·0122
100	·6400	2·5365	4·1234						
	·5931	1·7821	4·4073	·0003	−·0082	·0195	·0006	−·0135	·0423
sim. val.	·7	2·25	4·0						
1000	·7437	2·7134	3·8548						
	·7219	2·4658	3·9552	5×10^{-5}	−·0006	·0041	·0000	−·0008	·0049
750	·7331	2·5745	3·8961						
	·7243	2·4485	3·9397	7×10^{-5}	−·0019	·0083	·0001	−·0036	·0071
500	·7207	2·6379	3·8976						
	·6982	2·3541	4·0087	·0001	−·0027	−·0109	·0001	−·0048	·0105
100	·6876	2·1505	4·1916						
	·6601	1·9077	4·3139	·0004	−·0087	·0227	·0006	−·0149	·0410

n	α	σ^2	μ	$c^{\alpha\alpha}$	$c^{\sigma^2\alpha}$	$c^{\sigma^2\sigma^2}$	$c^{\mu\alpha}$	$c^{\mu\sigma^2}$	$c^{\mu\mu}$
sim. val.	·8	2·25	4·0						
1000	·8097	2·4964	3·9054						
	·7965	2·3458	3·9628	5×10^{-5}	−·0012	·0052	·0000	−·0025	·0043
750	·8190	2·5847	3·8741						
	·7994	2·3569	3·9591	6×10^{-5}	−·0016	−·0071	·0000	−·0033	·0057
500	·7808	2·4995	4·0847						
	·7764	2·3810	4·1140	·0001	−·0025	·0108	·0000	−·0048	·0089
100	·8434	2·3417	4·1884						
	·8229	2·1874	4·2601	·0004	−·0095	·0403	·0002	−·0188	·0357
sim. val.	·9	2·25	4·0						
1000	·9118	2·4327	3·9625						
	·9059	2·2620	3·9876	5×10^{-5}	−·0010	·0066	·0000	−·0024	·0031
750	·8741	2·1400	4·0961						
	·8856	2·1651	4·0622	5×10^{-5}	−·0012	·0065	·0000	−·0027	·0041
500	·8827	2·5177	4·0442						
	·8733	2·3662	4·0889	9×10^{-5}	−·0022	·0131	·0000	−·0049	·0071
100	·8403	1·6827	4·5947						
	·8611	1·8218	4·5162	·0003	−·0060	·0210	·0000	−·0114	·0254

A.1.2 COMPARISM OF MAXIMUM LIKELIHOOD TEST AND OPTIMAL C(α) TEST FOR μ = 0, ·2, n = 1000 THROUGHOUT

n	α	σ^2	μ	$c^{\alpha\alpha}$	$c^{\sigma^2\alpha}$	$c^{\sigma^2\sigma^2}$	$c^{\mu\alpha}$	$c^{\mu\sigma^2}$	$c^{\mu\mu}$
sim.	.1	·25	0						
val.	·2089	·4114	−·0324						
(3), (6)	·2079	·4722					$T^{(C)}$=	−·3118	
sim.	·2	·25	0						
val.	·3143 (1)	·6232	·0512						
(3), (6)	·3143 (4)	·6245					$T^{(C)}$=	1·7805	
sim.	·3	·25	0						
val.	·2081	·0936	·1141						
(3), (6)	·2081 (4)	·1039					$T^{(C)}$=	1·2555	
sim.	·4	·25	0						
val.	·2001	·1933 (2)	−·0767						
	·5303	·2545	−·0001	·0033	·0017	·0016	·0000	·0000	·0011
	·2001 (4)	·1933 (5)					$T^{(W)}$=	−0·0003	
(6)	–	–					$T^{(C)}$=	−1·7868	
sim	·5	·25	0						
val.	·3001 (1)	·2629	−·0176						
(3), (6)	·3001 (4)	·2631					$T^{(C)}$=	0·0882	
sim.	·6	·25	0						
val.	·7131	·0705	·0647						
	·6041	·2435	·0228	·0027	·0012	·0010	·0000	·0000	·0008
	·7131 (4)	·0725					$T^{(W)}$=	·8124	
	·6006	·2415		·0026	·0012	·0011	$T^{(C)}$=	1·6346	
sim.	·7	·25	0						
val.	·4145 (1)	·2109 (2)	·0435						
	·6303	·1975	·0025	·0018	·0006	·0005	·0000	·0000	·0006
	·4145 (4)	·2109 (5)					$T^{(W)}$=	·1021	
(6)	–	–					$T^{(C)}$=	2·7139	

n	α	σ^2	μ	$c^{\alpha\alpha}$	$c^{\sigma^2\alpha}$	$c^{\sigma^2\sigma^2}$	$c^{\mu\alpha}$	$c^{\mu\sigma^2}$	$c^{\mu\mu}$
sim. val.	·8	·25	0						
	·5702	·1647[2]	−·0175						
	·7465	·1995	−·0091	·0014	·0005	·0004	·0000	·0000	·0004
	·5702[4]	·1647[5]					$T^{(W)}=$	−·4482	
	·8275	·2539		·0015	·0006	·0005	$T^{(C)}=$	−2·0968	
sim. val.	·9	·25	0						
	·7711[1]	·1417[2]	·0033						
	·8051	·2019	−·0202	·0011	·0004	·0003	·0000	·0000	·0004
	·7711[4]	·1417[5]					$T^{(W)}=$	−1·0803	
	·9561	·2797		·0006	·0003	·0003	$T^{(C)}=$	−·0068	
sim. val.	·1	1·0	0						
	·2161[1]	·8996	−·0089						
(3), (6)	·2161[4]	·8996[5]					$T^{(C)}=$	−·1781	
sim. val.	·2	1·0	0						
	·3901[1]	·9100	·0726						
(3), (6)	·3901[4]	·9127					$T^{(C)}=$	1·1409	
sim. val.	·3	1·0	0						
	·0110[1]	1·0957	·0042						
(3), (6)	·0010[4]	1·0957[5]					$T^{(C)}=$	−·0380	
sim. val.	·4	1·0	0						
	·0343	·1425[2]	−·5138						
	·4778	·4657	−·0006	·0179	·0133	·0128	·0000	·0000	·0025
	·0343[4]	·1425[5]					$T^{(W)}=$	−·0136	
(6)	−	−					$T^{(C)}=$	−1·1511	
sim. val.	·5	1·0	0						
	·1918	·2136[1]	−·2171						
	·4818	·5105	·0031	·0091	·0127	·0213	·0000	·0000	·0016
	·1918[4]	·2136[5]					$T^{(W)}=$	−0·4172	
(6)	−	−					$T^{(C)}=$	−1·0391	

n	α	σ^2	μ	$c^{\alpha\alpha}$	$c^{\sigma^2\alpha}$	$c^{\sigma^2\sigma^2}$	$c^{\mu\alpha}$	$c^{\mu\sigma^2}$	$c^{\mu\mu}$
sim. val.	·6	1·0	0						
	·4131	1·9102	-1·3051						
	·4518	·7023	-0·2113	·0086	-·0130	·0010	·0001	·0000	·0013
	·4131 (4)	4·9876					$T^{(W)}$=	1·2039	
(6)	-	-					$T^{(C)}$=	-0·1611	
sim. val.	·7	1·0	0						
	·8285	·8861 (2)	-0·3154						
	·7198	·8915	-·0916	·0087	-·0230	·0001	·0001	·0000	·0014
	·8285 (4)	2·1311					$T^{(W)}$=	1·1721	
(6)	-	-					$T^{(C)}$=	2·3864	
sim. val.	·8	1·0	0						
	·3141 (1)	4·6753	1·1011						
(3), (6)	·3141 (4)	4·6753 (5)	-				$T^{(C)}$=	-0·6014	
sim. val.	·9	1·0	0						
	·7163	1·2804	-0·0361						
	·8814	·9315	-0·0354	·0071	-·0011	·0001	·0011	·0000	·0009
	·7163 (4)	1·1811							
	·8763	·9471		·0069	-·0149	·0015			

In view of the previous results, the simulations were only performed with α = 0·5(0·1)(0·9) in the following section of this appendix.

sim. val.	·5	1·0	·2						
	·3341	·8434	·2720						
(3)	·3341 (4)	·8926							
	·5001	·9474		·0001	·0000	·0071	$T^{(C)}$=	2·9593	
sim. val.	·6	1·0	·2						
	·4511 (1)	·8748	·3054						
(3)	·4511 (4)	·9215							
	·6001	·9738		·0001	·0000	·0053	$T^{(C)}$=	4·9222	
sim. val.	·7	1·0	·2						
	·3111 (1)	·8991	·2855						
(3)	·3111 (4)	·9389							
	·7999	·9879		·0001	·0000	·0031	$T^{(C)}$=	4·5708	

n	α	σ^2	μ	$c^{\alpha\alpha}$	$c^{\sigma^2\alpha}$	$c^{\sigma^2\sigma^2}$	$c^{\mu\alpha}$	$c^{\mu\sigma^2}$	$c^{\mu\mu}$
sim. val.	·9	1·0	·2						
	·6101 (1)	·9246	·3458						
(3)	·6101 (4)	·9844							
	·8999	1·0246		·0001	·0000	·0026	$T^{(C)}=$	5·3788	
sim. val.	·5	2·25	0·2						
	·2773	2·7693	·7569						
(3), (6)	·2773 (4)	3·1834					$T^{(C)}=$	4·7231	
sim. val.	·6	2·25	0·2						
	·1580	6·0241	·6472						
(3), (6)	·1580 (4)	6·3768					$T^{(C)}=$	2·1304	
sim. val.	·7	2·25	·2						
	·0378	3·1258	·3554						
(3), (6)	·0378 (3)	3·227					$T^{(C)}=$	2·5410	
sim. val.	·8	2·25	·2						
	·3508	3·4536							
(3), (6)	·3508 (4)	3·7023					$T^{(C)}=$	3·4423	
sim. val.	·9	2·25	·2						
	·0162	·0787	10·748						
	·5211	1·2055	·2326	·0011	−·0013	·0011	−·0002	−·0024	·0045
	·0162 (4)	76·9394					$T^{(W)}=$	3·4589	
(6)	–	–					$T^{(C)}=$	2·1882	

APPENDIX A.2

THE NULL DISTRIBUTION OF A TEST STATISTIC FOR A MIXTURE

This appendix gives the values of t such that

$$pr\{T^{(C)} \leq t;\ H_o\} = \gamma$$

as defined in equation (3.4.4) of section (3.4).

γ	n					
	20	40	60	80	100	χ^2_2
·9	4·45	4·47	4·49	4·51	4·57	4·60
·905	4·57	4·58	4·59	4·62	4·68	4·71
·91	4·64	4·72	4·76	4·79	4·81	4·82
·915	4·75	4·84	4·86	4·89	4·91	4·93
·92	4·93	4·93	4·94	4·96	4·99	5·05
·925	5·06	5·07	5·09	5·12	5·15	5·18
·93	5·25	5·26	5·28	5·34	5·37	5·39
·935	5·40	5·41	5·42	5·44	5·46	5·47
·94	5·66	5·57	5·58	5·60	5·61	5·63
·945	5·96	5·69	5·71	5·74	5·77	5·80
·95	6·15	5·92	5·95	5·97	5·98	5·99
·955	6·47	6·14	6·16	6·18	6·19	6·20
·96	6·74	6·35	6·38	6·41	6·43	6·44
·965	7·08	6·70	6·67	6·68	6·69	6·70
·97	7·37	7·02	7·01	7·01	7·01	7·01
·975	8·00	7·48	7·45	7·42	7·41	7·38
·98	8·44	8·02	7·91	7·88	7·84	7·82
·985	9·76	8·57	8·54	8·49	8·46	8·40
·99	11·26	9·35	9·34	9·31	9·27	9·21
·95	14·29	11·28	11·16	11·01	10·78	10·60

APPENDIX A.3

SIMULATIONS OF A SIMPLE LINEAR REGRESSION WITH CAUCHY ERRORS

This appendix contains the results of simulation experiments on the model defined in (4.1.1) of section (4.1).

n denotes sample size. $\tilde{\alpha}$ and $\tilde{\xi}$ are the estimates defined by (4.1.4) and (4.1.5) respectively. β is the true value of the slope parameter used to generate the data. $T^{(C)}$ is defined by (4.1.6).

$\beta = 0{\cdot}0$

n	$\tilde{\alpha}$	$\tilde{\xi}$	$T^{(C)}$
11	-·2196	·2196	-·0135
21	-·2285	·3301	-·2735
31	-·3194	·9048	-·8787
41	-·3143	·7453	-·4407
51	-2·8388	2·1609	-·3967
61	·0338	·6343	·6280
71	-·0257	·8693	·4003
81	·0645	·8509	1·3081
91	·1452	·8419	1·0455
101	-·1923	·9013	-·1179
111	-·1185	1·1056	-·1409
121	-·1739	1·0347	-·5898
131	-·1185	1·1056	-1·0027
141	-·0591	1·0858	·0516
151	-·0634	1·1298	-·1345
161	-·0491	·9816	-·2607
171	-·0152	1·0816	·2201
181	-·1032	1·0009	·0357
191	-·1282	1·0804	·7653
201	-·0872	1·0169	-1·8524
301	-·2026	1·0596	-2·9740
401	-·1532	1·0224	-1·7975
501	-·0713	1·3113	-1·1110
601	-·1723	1·0415	-1·2768
701	·0658	1·0215	-·9165
801	-·0475	1·0215	·2876
901	-·0305	1·0191	·5590
1001	-·1313	1·0144	·6776

$\beta = 0{\cdot}1$

11	-·05674	1·2086	2·6488
21	-·45099	·9454	2·8099
31	-·44595	1·0287	2·0088
41	·75989	1·0109	1·9570
51	-·4501	1·5939	3·6488
61	-·2545	2·1076	6·1269
71	·0946	2·5758	7·4505
81	-·5581	2·6031	7·3589
91	·3843	3·1151	7·8740

n	$\tilde{\alpha}$	$\tilde{\xi}$	$T^{(C)}$
101	−·3192	3·2206	8·4040
111	−2·9034	·6881	3·7695
121	−·8101	4·2373	9·6734
131	−·3380	4·0835	10·4321
141	·3881	4·4358	11·1300
151	·4537	4·7473	11·6646
161	−·5480	4·7890	11·7674
171	·0195	4·9705	12·3804
181	·2608	4·6482	12·6545
191	−·1304	5·1303	13·3859
201	−·3655	5·1971	13·5371

β = 0·2

11	−·2567	1·5086	2·7700
21	−·6690	1·3773	3·2229
31	−1·3373	1·3373	3·5562
41	−·8601	1·5512	4·1387
51	1·1307	2·0134	4·7564
61	−·5357	3·1238	6·5815
71	·2355	4·3351	7·9562
81	·5064	4·4014	8·2276
91	−·6022	4·9813	9·0991
101	−·7216	5·4844	9·3399
111	−·4906	6·5843	10·1687
121	·7435	7·2127	10·7889
131	·8622	7·4734	11·4269
141	−·3142	8·1837	12·0857
151	−·5249	9·1731	12·7265
161	·2111	9·4829	13·0387
171	−1·2837	9·8335	13·4931
181	·1035	9·1746	13·6938
191	−9·4811	·3208	·9185
201	−·0874	10·6447	14·7732

β = 0·3

11	−·2359	1·5878	2·9446
21	−5·3333	2·5824	1·0543
31	·9624	2·3613	4·7075
41	−1·0637	2·2977	5·1892
51	−·4314	3·2166	5·3664
61	−·4031	4·7448	7·0199
71	·2581	5·8415	8·2450
81	−·2256	6·3470	8·8361
91	−·4232	7·5219	9·7596
101	·1207	7·7865	10·1130
111	−1·1477	9·3636	10·7291
121	−·7667	10·5480	11·2985
131	·8904	11·2618	11·8849
141	·8825	11·7033	12·3731
151	−·6866	12·7493	12·9123
161	−1·1412	13·1718	13·3128
171	·1182	12·9990	13·6135

n	$\tilde{\alpha}$	$\tilde{\xi}$	$T^{(C)}$
181	−·2194	13·4863	13·9599
191	−·1625	14·8937	14·5591
201	·1927	15·5612	14·9959
β = 0·4			
11	·1897	1·7376	2·9847
21	−7·6757	4.2265	·9615
31	−·9693	2·9580	4·9185
41	1·0243	3·4758	5·6199
51	·3164	4·8543	5·8663
61	−·5854	6·5772	7·3988
71	·4644	7·6352	8·5150
81	−·2846	8·6457	9·2166
91	−·288	9·9820	10·0559
101	1·4542	10·5257	10·5114
111	−·9963	12·4122	11·0734
121	·8252	13·7066	11·5256
131	·1222	14·3944	12·0053
141	−1·1815	14·9543	12·4331
151	−·7649	16·7710	13·0351
161	−·1945	17·1449	13·4258
171	1·2289	17·3931	13·7866
181	·2993	17·9892	14·1429
191	·5119	19·9039	14·7570
201	·1101	20·7210	15·3287
β = 0·5			
11	−·1397	2·0876	3·0027
21	−1·2453	2·1213	3·1015
31	1·1849	3·7145	4·9463
41	1·2400	4·5101	5·7607
51	−·4958	6·0791	6·1593
61	−·2298	7·8715	7·5794
71	·7568	9·3427	8·6735
81	−·1061	10·6554	9·3488
91	−·3543	12·1569	10·1438
101	−·7199	12·9973	10·6404
111	·1742	15·2048	11·1974
121	·3389	15·9928	11·5366
131	−·0524	17·7273	12·0977
141	−1·0099	18·8369	12·5687
151	−·1849	20·3246	13·0645
161	·3049	21·0844	13·4985
171	·1789	21·7430	13·9112
181	−·3097	22·3788	14·2614
191	·4349	24·6269	14·8613
201	·0505	25·0531	15·2013

APPENDIX A.4

SIMULATIONS ON THE TWO SAMPLE PROBLEM FOR THE GAMMA DISTRIBUTION

This appendix gives the results of simulations on the model considered in section (4.4). The first line for particular sample size n gives the results of a full maximum likelihood analysis. $\hat{\beta}$ and $\hat{\Delta}$ are the solutions of (4.4.5) and (4.4.7) respectively. $\hat{\alpha}$ and $\hat{\xi}$ are given by (4.4.6) and (4.4.8). In the final column is the value of the statistic $T^{(W)}$ defined by (4.4.9).

The second line gives the results for an optimal C(α) test of the same hypothesis for the same data. $\hat{\hat{\beta}}$ is the solution of (4.4.11) and $\hat{\hat{\alpha}}$ is given by (4.4.12). The final column gives the value of the statistic $T^{(C)}$ defined by (4.4.10).

n	α	β	ξ	Δ	T
sim. val.	1·0	0·5	-0·1	-0·1	
20	·9277	·5982	·5833	-·0381	2·3353
	1·2128	·5979			2·8294
40	1·0694	·4054	·0369	-·0213	·1072
	1·0818	·3923			·2248
60	·9552	·4685	·2416	-·0519	2·0205
	1·0491	·4369			3·3750
80	1·1357	·5376	-·2633	-·1464	2·1481
	·9827	·4528			5·1389
100	·7914	·3977	·4055	·0287	2·4066
	·9498	·4092			4·5225
200	·9132	·4189	-·1614	-·0452	1·0017
	·8356	·3926			2·1431
300	1·0509	·5184	-·0864	-·0629	1·7047
	1·0159	·4906			2·1769
500	1·0299	·4907	-·0377	-·0674	4·6945
	1·0069	·4547			8·5795

n	α	β	ξ	Δ	T
sim. val.	1·0	0·5	-0·1	0·0	
20	·6006 ·7283	·3599 ·3752	·4265	·0367	·4599 ·9178
40	1·3382 ·9618	·5676 ·5447	-·5554	-·0234	3·0265 3·6815
60	·9373 1·0207	·4690 ·5555	·1594	·1773	3·8529 5·1888
80	1·0067 1·0689	·5235 ·5648	·1266	·0826	·7447 1·3392
100	·8968 ·8509	·5162 ·4919	-·1033	-·0467	·3026 ·5082
200	1·0767 1·1027	·5175 ·5509	·0463	·0674	1·6018 2·5745
300	1·0054 ·9204	·5463 ·5125	-·1779	-·0664	1·8586 3·2537
500	·9179 ·9309	·4882 ·4923	-·0206	·0083	·0682 ·1012
sim. val.	1·0	0·5	-0·1	0·1	
20	·6367 ·7374	·6040 ·6782	·4096	·2505	1·2179 1·3196
40	1·6252 ·9961	·4839 ·4975	-·6648	·Lo23	8·0905 10·5433
60	·7578 ·8510	·5484 ·5441	·2697	·0027	1·1549 1·5040
80	1·2293 ·9595	·4935 ·5037	-·3972	·0421	4·8147 6·8005
100	1·1279 ·9874	·5651 ·5787	-·1897	·0610	2·9207 3·3111
200	·8039 ·7931	·4806 ·5236	·0047	·1030	5·5094 6·3406
300	1·1239 1·0406	·4639 ·5241	-·0671	·1609	25·3495 27·3145
500	1·0983 ·9284	·5227 ·5531	-·2317	·1008	30·3021 36·1950

n	α	β	ξ	Δ	T
sim. val.	1·0	0·5	0·0	-0·1	
20	·7594 ·9206	·6282 ·5074	·8086	-·1589	7·5293 8·7034
40	2·0004 1·6389	·5675 ·5363	-·4240	-·0822	1·1853 1·3125
60	1·1214 1·0507	·6136 ·5137	-·1196	-·1850	2·5346 6·2975
80	·6724 ·7493	·4371 ·3913	·3322	-·0674	4·8557 9·0588
100	·6696 ·7592	·4447 ·4239	·3111	-·0311	3·4121 5·7916
200	1·1367 1·1329	·5031 ·4838	-·0248	-·0470	·8775 1·0471
300	1·0543 1·0518	·4874 ·4556	·0026	-·0605	2·8739 4·9976
500	·9655 ·9443	·4949 ·4462	-·0221	-·0865	8·5037 16·8124
sim. val.	1·0	0·5	0·0	0·0	
20	·8000 ·8775	·4865 ·4725	·2133	-·0225	·3433 ·5425
40	·9482 ·9046	·5167 ·4936	-·0929	-·0447	·1097 ·1842
60	1·4038 1·0517	·5226 ·4892	-·5086	-·0516	2·8454 3·8584
80	·8609 ·9402	·5077 ·5295	·1781	·0444	·3903 ·6163
100	·9381 1·0542	·4857 ·5161	·2363	·0625	·8916 1·3627
200	1·0811 ·9420	·5203 ·4843	-·2649	-·0667	2·1694 3·3492
300	1·0466 ·9794	·5078 ·4974	-·1291	-·0197	·7870 1·1398
500	1·0637 1·0581	·5054 ·5066	-·0106	·0023	·0327 ·0471

n	α	β	ξ	Δ	T
sim. val.	1·0	0·5	0·0	0·1	
20	1·2223	·6247	−·2501	·2141	3·0225
	·9639	·6710			2·8234
40	·5889	·4240	·5743	·1189	1·7932
	·7541	·4725			2·6651
60	1·0498	·6238	·0932	·0639	·2235
	1·0618	·6324			·0479
80	·7947	·3838	·0374	·2331	16·3416
	·7855	·4667			15·1125
100	1·2085	·4859	−·1368	·1754	10·8231
	1·0605	·5459			11·5306
200	1·0117	·5308	·1601	·1553	7·1428
	·9831	·5111			6·6458
300	·9753	·4803	·0675	·1339	11·8381
	·9881	·5340			12·1414
500	1·2011	·5297	−·0712	·1131	18·9901
	1·1170	·5672			20·5059
sim. val.	1·0	0·5	0·1	−0·1	
20	·5887	·3706	·2959	·0306	·2214
	·6750	·3836			·4181
40	1·1671	·5251	·2129	−·1269	3·0204
	1·2252	·4445			5·3921
60	·7227	·3476	·7929	·1182	4·8541
	1·0396	·4016			8·8133
80	1·0081	·5877	·2877	−·1559	9·1973
	1·0863	·4911			14·7368
100	·6660	·4128	·4320	−·0446	6·4611
	·7719	·3771			12·1076
200	1·2086	·5488	−·0377	−·1139	4·6935
	1·1785	·4876			8·3266
300	1·1389	·5026	·0398	−·0566	3·1425
	1·1190	·4715			4·6943
500	1·0636	·5304	−·0424	−·1140	12·3377
	1·0268	·4657			24·7545

n	α	β	ξ	Δ	T
sim. val.	1·0	0·5	0·1	0·0	
20	·8960	·4531	·5218	·2097	1·6164
	1·1739	·5545			2·3491
40	1·1407	·5431	−·1993	−·1587	1·3273
	1·0190	·4521			3·1622
60	1·1957	·6441	−·3493	−·1555	1·4479
	1·0123	·5660			3·6632
80	·9745	·4786	·3373	·0893	1·4658
	1·1450	·5216			2·2644
100	1·2664	·5360	−·0653	−·0626	·5645
	1·2252	·5037			1·0221
200	·9923	·5093	·0397	−·0065	·1505
	1·0116	·5059			·2196
300	·8563	·4639	·1955	−·0037	2·8893
	·9363	·4605			4·6105
500	1·0014	·5251	·2275	·0027	6·4975
	1·1107	·5246			8·9255
sim. val.	1·0	0·5	0·1	0·1	
20	2·3890	0·5431	−·9560	·0650	6·3889
	1·1663	·5188			7·2199
40	1·0023	·4564	−·0750	·1969	5·1175
	·9104	·5269			5·2143
60	1·6744	·5683	−·0144	·0623	·5199
	1·6211	·5848			·4758
80	·9835	·3859	·0863	·2753	21·7173
	·9531	·4809			18·0306
100	·8787	·4763	−·0632	·1077	3·8831
	·8319	·5199			4·6460
200	1·1027	·4881	·1203	·1507	8·7541
	1·1440	·5469			8·4075
300	1·1816	·5155	·5554	·1751	18·8631
	1·1662	·5767			18·0939
500	·9585	·4645	·1681	·1742	31·1255
	1·0188	·5338			28·7486

n	α	β	ξ	Δ	T
sim. val.	1·0	1·0	-0·1	-0·1	
20	1·1659 ·7773	1·4637 ·8243	-·7996	-1·0010	3·0591 6·7697
40	·9778 ·8198	1·0009 1·0338	-·2988	·0859	3·0559 1·9615
60	·9853 ·9243	·8635 ·9219	-·0985	·1350	2·1693 1·4997
80	1·1109 1·0114	·9918 ·9683	-·1790	-·0428	·7812 ·4995
100	1·0899 1·0339	1·0685 1·0003	-·1053	-·1332	·5155 ·4114
200	1·1226 1·0912	1·1089 1·1170	-·0837	·0449	·3111 ·1238
300	1·1018 1·0205	1·1343 1·0077	-·1498	-·2374	4·7480 4·6419
500	1·0127 ·9892	1·0209 ·9978	-·0533	-·0519	·4182 ·2234

sim. val.	1·0	1·0	-0·1	0·0	
20	1·1039 1·0522	1·2796 1·1648	-·1003	-·2261	1·0522 1·1659
40	·8303 ·7730	·9697 ·9241	-·1393	-·0863	·1518 ·1073
60	1·1723 ·8968	1·1438 ·9420	-·5090	-·3510	3·1824 2·3062
80	1·2054 1·0229	·9853 ·9431	-·2953	-·0640	2·1995 1·4205
100	1·0588 1·0439	1·0267 ·9780	-·0236	-·0917	·4777 ·3483
200	·9747 ·9620	·9761 1·0628	-·0116	·1876	6·4509 3·8107
300	·8731 ·8113	·9187 ·8952	-·1414	-·0434	1·4402 1·0171
500	1·0504 ·9963	1·0507 1·0278	-·1038	-·0444	1·2813 ·8004

n	α	β	ξ	Δ	T
sim. val.	1·0	1·0	-0·1	0·1	
20	1·4956	1·2836	-·8320	-·0927	7·4420
	·8299	1·1153			3·2268
40	1·1573	·8975	-·5444	·7311	41·8764
	·6205	·9644			18·9883
60	1·0699	·9911	-·4565	·5829	39·0960
	·6619	1·0749			18·5539
80	1·0022	·9611	-·2752	·6229	43·6964
	·7288	1·1079			20·0348
100	·8932	·8366	-·3374	·6739	76·2344
	·5776	·9548			34·5851
200	1·1660	1·03333	-·5820	·4889	129·3892
	·6565	1·0579			63·0621
300	·9562	1·0305	-·2795	·5226	120·3754
	·7165	1·1540			57·3471
500	·9538	·9846	-·0789	·1447	5·0595
	·9650	1·0504			2·5833
sim. val.	1·0	1·0	0·0	-0·1	
20	1·6943	1·2417	-·2614	-·1212	·3922
	1·5060	1·1987			·1855
40	·9081	·8911	·0187	-·0097	·0173
	·9162	·8859			·0129
60	·7934	1·1330	·0260	-·2344	2·6544
	·7874	·9967			2·0488
80	1·2381	·9846	·0076	·1137	·8026
	1·2638	1·0588			·6762
100	1·1285	1·0713	-·4446	-·2819	3·9704
	·0883	·9063			2·8530
200	·8824	·9235	·1526	·0218	1·5732
	·9487	·9329			1·1185
300	1·0928	1·0798	-·1655	-·2214	4·2679
	1·0001	·9575			4·0492
500	1·0749	1·0644	-·0049	-·1817	5·7978
	1·0147	·9651			5·0652

n	α	β	ξ	Δ	T
sim. val.	1·0	1·0	0·0	0·0	
20	·7719 ·8478	·7882 ·7999	·2031	·0308	·2223 ·1882
40	1·1794 1·0096	1·2330 1·0914	-·3224	-·2793	·8652 ·7748
60	1·0373 ·9699	1·0451 ·9993	-·1329	-·0889	·2107 ·1409
80	1·3435 1·1623	1·2757 1·1624	-·2932	-·2224	1·4710 1·2158
100	·9245 ·9148	·8444 ·9004	-·0041	·1269	1·7586 1·1374
200	·9819 ·9081	·9998 ·9513	-·1519	-·0915	·9053 ·6160
300	·9887 ·9146	·8902 ·9071	-·1348	·0458	4·8912 3·5274
500	·9897 ·9213	·9856 ·9413	-·0389	-·0649	1·9953 1·3306
sim. val.	1·0	1·0	0·0	0·1	
20	·6834 ·8709	·9210 1·2414	·5006	·7411	5·6887 2·3263
40	·7866 ·8126	·8074 1·1167	·2514	·8712	26·7643 8·9498
60	·8486 ·6863	·8404 1·0284	-·1091	·6867	34·9546 8·9489
80	1·0608 ·8178	1·1000 1·2380	-·2461	·5512	·8178 1·2391
100	·9745 ·9881	·8905 1·0112	-·2131	·6967	·7257 2·0130
200	1·0188 ·7666	1·0449 1·1703	-·2742	·5249	79·1747 37·5976
300	1·0393 ·7541	1·0734 1·1732	-·3297	·4801	116·3042 55·8966
500	1·0287 ·9711	1·0404 ·0172	-·1129	·0448	1·5671 ·9737

n	α	β	ξ	Δ	T
sim. val.	1·0	1·0	0·1	-0·1	
20	1·2613 1·0431	·9025 ·5890	-·1975	-·4741	2·7829 3·5722
40	1·2410 1·0732	1·0697 ·9014	-·2644	-·2904	·9407 1·0010
60	1·0399 1·0269	·8054 ·8102	-·0212	·0124	·0371 ·0288
80	1·1549 ·9693	1·0986 ·9187	-·3269	-·3119	2·1271 2·1536
100	1·0958 1·1137	1·0582 ·9754	·0613	-·1425	2·6125 1·8322
200	1·1025 1·0309	1·0282 ·8806	-·0927	-·2486	6·0336 5·9116
300	·8217 ·9049	·8453 ·8529	·2134	·0257	4·3356 3·3892
500	1·0159 1·0551	1·0353 ·9621	·1132	-·1199	15·7325 10·6714
sim. val.	1·0	1·0	0·1	0·0	
20	·8863 1·0779	·7064 ·9277	·4154	·5349	4·7579 1·6752
40	·6167 ·7499	·8744 ·8489	·5434	·0261	4·6591 3·3876
60	1·0190 1·2357	1·0417 1·1118	·4246	·1692	2·6544 1·7939
80	1·2159 1·0408	1·1437 ·9339	-·2891	-·3659	2·6242 3·0639
100	1·0085 1·1242	·9376 ·9883	·2270	·1108	1·0106 ·7009
200	1·1998 1·0985	1·0862 ·0416	·1731	-·0859	1·3577 ·8452
300	·9177 1·0272	1·0032 1·0163	·2528	·0433	6·7996 4·3609
500	1·1041 1·1214	1·0631 1·0946	·0306	·0632	·7858 ·5723

n	α	β	ξ	Δ	T
sim. val.	1·0	1·0	0·1	0·1	
20	1·2218 1·0502	1·1988 1·3231	-·1389	·4361	3·8559 1·8275
40	1·0481 1·0205	·9729 1·1869	·0643	·5897	10·4715 4·5380
60	1·0155 ·8191	1·1390 1·1947	-·2574	·2823	9·9352 4·9524
80	1·0400 ·8613	·9424 1·1364	-·1130	·6479	36·4840 15·9709
100	1·0007 ·9407	·9378 1·1723	·0376	·6783	39·1128 16·2863
200	·9897 ·8574	1·0743 1·2277	-·1123	·4986	49·4022 23·3796
300	·8584 ·8659	·9416 1·1925	·1300	·6784	99·9390 40·8922
500	·9767 ·9843	·9666 1·0065	·0204	·0853	2·3790 1·3997
sim. val.	1·0	6·0	-0·1	-0·1	
20	1·0371 1·1020	5·9387 6·0966	·1333	·3870	·2979 ·0280
40	·9205 ·9296	5·7425 6·0705	·0385	·7839	1·2994 ·1017
60	·8825 ·8615	5·4197 5·4217	-·0416	·0369	·3817 ·0373
80	1·1880 1·0077	7·5018 6·4398	-·3443	-2·0565	3·0456 ·2593
100	1·1039 1·1019	7·1193 7·2606	-·0062	·2630	·6979 ·0548
200	·8507 ·8886	5·1835 5·6054	·1060	·9864	5·6057 ·3979
300	·9419 ·9269	5·9082 6·0385	-·0175	·3453	5·3754 ·4609
500	1·0613 ·9196	6·1256 5·6450	-·2410	-·1009	26·2886 2·0592

n	α	β	ξ	Δ	T
sim. val.	1·0	6·0	−0·1	0·0	
20	·8454 ·7531	4·8579 4·9090	−·0733	·8702	3·2652 ·3143
40	·6846 ·7438	4·1008 4·9671	·2936	2·6592	14·8678 ·6863
60	1·0038 ·9188	5·7987 5·6776	−·1268	·0291	3·0250 ·2752
80	·9842 ·8584	6·0313 5·4868	−·2478	−·9187	2·6889 ·1980
100	1·0320 1·0037	5·9827 6·2772	·0016	·9539	8·1254 ·6254
200	·9066 ·9165	5·4683 5·7391	·0407	·6626	4·3235 ·3505
300	1·1056 1·0016	6·3472 6·0634	−·1629	−·4085	9·5769 ·7659
500	·9320 ·8742	5·4997 5·5171	−·0785	·2917	24·1568 2·2735
sim. val.	1·0	6·0	−0·1	0·1	
20	·7272 1·0185	4·4137 6·3541	·6480	4·4943	11·6215 ·3389
40	·9345 ·7972	6·2840 6·0238	−·1698	·3306	6·2485 ·5153
60	1·1891 ·8435	7·3481 5·8467	−·5540	−2·0987	17·1645 ·9189
80	·9033 ·8535	5·6452 5·9423	·0114	1·3670	13·2689 ·9864
100	·9288 ·7588	5·5447 5·3604	−·1171	1·1955	31·9130 2·6575
200	·8502 ·7991	5·4667 5·5252	−·0689	·6851	21·4660 1·9135
300	1·0764 ·9167	6·5990 6·3419	−·1687	·4069	52·8453 4·1275
500	1·0068 ·9288	6·0523 5·8998	−·1259	·1113	17·5473 1·5229

n	α	β	ξ	Δ	T
sim. val.	1·0	6·0	0·0	-0·1	
20	·4245 ·8791	2·3486 4·9698	1·1779	5·6076	66·7914 ·4600
40	1·1439 ·8240	7·2103 5·1887	-·6541	-3·4713	6·7218 ·4114
60	1·1070 1·2724	6·5287 7·2272	·2743	1·4268	1·3631 ·1963
80	·9537 ·9759	5·9794 5·7924	·0843	-·1667	3·0282 ·2692
100	·7895 ·8621	5·2838 5·5884	·1920	·7151	1·3347 ·1517
200	1·0877 1·1067	6·5763 6·6655	·0305	·1521	·0639 ·0076
300	1·0493 1·0409	6·1496 6·1123	-·0167	-·0785	·0274 ·0027
500	1·0492 1·0338	6·3586 6·1033	-·0292	-·2230	1·0813 ·1037
sim. val.	1·0	6·0	0·0	0·0	
20	1·0520 1·2559	5·3472 6·5604	·3369	2·5562	2·2784 ·1132
40	·9090 1·1287	5·4143 6·9274	·3994	3·1499	7·1927 ·3443
60	1·0913 1·1208	6·6360 6·8221	·0525	·3723	·0522 ·0062
80	1·0896 1·0248	6·5175 6·1152	-·1245	-·7884	·3109 ·0342
100	1·0810 1·0360	6·5527 6·3660	-·0836	-·3609	·3467 ·0280
200	1·2190 1·1635	7·0156 6·6670	-·0956	-·6766	·4875 ·0674
300	·9378 1·0020	5·8599 6·1643	·1334	·6292	1·5122 ·1577
500	·9243 ·9783	5·5510 5·8890	·1156	·0828	1·3899 ·1190

n	α	β	ξ	Δ	T
sim. val.	1·0	6·0	0·0	0·1	
20	1·7947	8·9335	-·4326	-·7656	10·7864
	1·2538	7·5168			·6083
40	·6503	4·0556	·5042	4·0503	30·6166
	·7958	5·4471			·8838
60	1·1280	6·9614	·0195	·8346	2·3309
	1·1236	7·2746			⁻1662
80	1·0631	6·4082	-·1339	·5728	12·9525
	·9262	6·2688			1·0265
100	1·3183	7·6002	-·2724	-·1873	23·3462
	1·0590	6·9628			1·5894
200	·9861	5·8913	·0342	·9805	11·7628
	·9827	6·2445			·8781
300	1·0528	6·4396	·0198	·9379	16·9687
	1·0429	6·7714			1·2369
500	·9810	5·7978	-·0825	·3287	19·9880
	·9403	6·6246			1·9858
sim. val.	1·0	6·0	0·1	-0·1	
20	·9803	5·6166	·3671	2·4851	1·3555
	1·2019	6·8836			·0989
40	·9073	5·4711	·2949	·9774	1·8181
	1·0253	5·8040			·2064
60	1·1059	6·5595	-·1250	-1·4841	2·8229
	1·0085	5·6133			·3741
80	·8256	5·1584	·3639	1·2845	4·0421
	·9595	5·5998			·5201
100	1·0882	6·7123	·0687	-·2827	4·1690
	1·1080	6·4730			·3327
200	·8936	5·3442	·1842	·5743	3·7628
	·9691	5·5671			·4079
300	·9797	5·6871	·1443	·6107	2·0773
	1·0510	5·9746			·2204
500	·8850	5·4481	·1781	-·1802	3·0773
	·9468	5·4954			1·0190

n	α	β	ξ	Δ	T
sim. val.	1·0	6·0	0·1	0·0	
20	·9642 ·9133	5·1982 4·7519	–·0958	–·8056	·3096 ·0419
40	1·2431 1·2207	7·4753 6·9583	–·0073	–·8277	1·7269 ·1475
60	·9463 ·8767	5·9700 5·0339	–·0322	–1·2477	6·2763 ·7305
80	1·2169 1·0126	7·6139 6·1086	–·3358	–2·6539	2·5854 ·6131
100	1·0700 1·1602	6·5702 6·6873	·2076	·5255	5·2222 ·4659
200	1·0614 1·1197	6·3123 6·4540	·1179	·3523	2·5946 ·2293
300	·8692 ·9734	5·2898 5·6899	·2501	·9875	5·9706 ·7086
500	·9173 ·9426	5·6443 5·4749	·0950	–·1331	19·9199 1·8762
sim. val.	1·0	6·0	0·1	0·1	
20	1·4176 1·3299	8·2324 7·6491	–·1278	–1·1249	·1036 ·0240
40	1·2627 1·8112	7·3287 7·2124	–·1023	–·0235	1·4382 ·1026
60	1·0388 ·9487	5·8603 5·7780	–·1150	·1895	3·8350 ·3425
80	1·1084 1·1131	6·6841 6·8985	·0205	·5168	·9872 ·0745
100	·9173 1·0918	5·7303 6·8339	·3319	2·2862	5·4970 ·3366
200	1·6419 1·0624	6·1276 6·4111	·0478	·6346	2·4688 ·1867
300	·9789 1·0010	5·6220 6·1609	·1007	1·3698	21·9427 1·4601
500	1·0002 1·0332	6·0450 5·9235	·1021	·0480	18·3173 1·5966

BIBLIOGRAPHY

Bankövi, J. (1964) A note on the generation of Beta distributed and Gamma distributed random variables. *Magyar Tudomanyos Akademica, Matematikai Kutato Intezet*, 9, 555-562.

Barnett, V.D. (1966) Evaluation of the maximum likelihood estimator where the likelihood equation has multiple roots. *Biometrika*, 53, 151-166.

Bartlett, M.S. (1953a) Approximate confidence intervals. *Biometrika*, 40, 12-19.

Bartlett, M.S. (1953b) Approximate confidence intervals, II. *Biometrika*, 40, 306-317.

Bartlett, M.S. (1955) Approximate confidence intervals, III. A bias correction. *Biometrika*, 42, 201-204.

Bartoo, J.B., and Puri, P.S. (1967) On optimal asymptotic tests of composite statistical hypotheses. *Ann. Math. Statist.*, 38, 1845-1852.

Birnbaum, Z.W. (1950) Effect of linear truncation on a multinormal population. *Ann. Math. Statist.*, 21, 272-279.

Blischke, W.R. (1964) Mixtures of Distributions. *Classical and contagious discrete distributions*, ed. G.P.Patil, Calcutta Statistical Publishing Society, 351-373.

Brunk, H.D. (1958) On the estimation of parameters restricted by inequalities. *Ann. Math. Statist.*, 29, 437-453.

Buhler, J.W., and Puri, P.S. (1966) On optimal asymptotic tests of composite hypotheses with several constraints. *Zeitschrift fur Wahrscheinlichkeitstheorie*, 5, 71-88.

Chernoff, H. (1952) A measure of asymptotic efficiency of tests for a hypothesis based on a sum of observations. *Ann. Math. Statist.*, 23, 493-507.

Cramér, H. (1946) *Mathematical Methods of Statistics*, Princeton University Press, Princeton.

Cox, D.R. (1966) *Statistical analysis of series of events*, Methuen Monograph.

Cox, D.R. (1970) *Analysis of Binary Data*, Methuen Monograph.

Daniells, H.E. (1961) The asymptotic efficiency of a maximum likelihood estimator. *Proc. 4th Berkeley Symposium*, 151-163.

Davies, H.T. (1933) *Tables of the higher mathematical functions*, Vols I and II. Principia Press, Indiana.

Davies, R.B. (1969) Beta optimal tests and an application to the summary evaluation of experiments. *J. R. Statist. Soc.*, B, 31, 524-538.

Day, N.E. (1969) Estimating the components of a mixture of normal distributions. *Biometrika*, 56, 463-474.

Haas, G., Bain, L., and Antle, C. (1970) Inferences for the Cauchy distribution based on maximum likelihood estimators. *Biometrika*, 57, 403-409.

Hill, B.M. (1963) Information for estimating the proportions in mixtures of exponential and normal distributions. *J. Amer. Statist. Ass.*, 58, 918-932.

Hoadley, B. (1971) Asymptotic properties of maximum likelihood estimators for the independent but not identically distributed case. *Ann. Math. Statist.*, 42, 1977-1991.

Hodges, J.L., and Lehmann, E.J. (1970) Deficiency. *Ann. Math. Statist.*, 41, 783-801.

Huber, P.J. (1967) The behaviour of maximum likelihood estimates under non-standard conditions. *Proc. 5th Berkeley Symposium*, 221-233.

Isaacson, S.L. (1950) On the theory of unbiased tests of simple statistical hypotheses specifying the values of two or more parameters. *Report to the Office of Naval Research, Columbia University.*

Jöhnck, M.D. (1964) Erzeugung von betaverteilten und gamma-verteilten zufallszahlen. *Metrika*, 8, 5-15.

Klonecki, W. (1973) A note on optimal C(α) tests for homogeneity of the Poisson distribution. *Applicationes Mathematicae*, 8, 497-505.

Kulkarni, S.R. (1968) Tests of hypotheses about two treatment effects when the effects may be fixed or variable. *Karnatak University Journal: Science*, 8, 85-96.

Le Cam, L. (1956) On the asymptotic theory of estimation and hypothesis testing. *Proc. 3rd Berkeley Symposium*, 129-156.

Lehmann, E.L. (1959) *Testing statistical hypotheses*. Wiley, New York.

Mann, H., and Wald, A. (1943) On stochastic limits and order relationships. *Ann. Math. Statist.*, 14, 217-226.

Maritz, J.S. (1970) *Empirical Bayes Methods*. Methuen Monograph.

Moran, P.A.P. (1968) *Introduction to probability theory*. Oxford University Press.

Moran, P.A.P. (1970a) Methodology of rain making experiments. *Review of the Int. Statist. Inst.*, 38, 105-119.

Moran, P.A.P. (1970b) On asymptotically optimal tests of composite hypotheses. *Biometrika*, 57, 105-119.

Moran, P.A.P. (1971a) The uniform consistency of maximum likelihood estimators. *Proc. Camb. Phil. Soc.*, 70, 435-439.

Moran, P.A.P. (1971b) Maximum likelihood estimation in non-standard conditions. *Proc. Camb. Phil. Soc.*, 70, 441-450.

Moran, P.A.P. (1973a) Asymptotic properties of homogeneity tests. *Biometrika*, 60, 79-85.

Moran, P.A.P. (1973b) Are there two maternal age groups in Down's syndrome? *Brit. J. Psychiat.*, 123 (to appear).

Newman, T.G., and Odell, P.L. (1971) *The generation of random variables.* Griffin Statistical Monograph, London.

Neyman, J. (1959) Optimal asymptotic tests of composite hypotheses. In *Probability and Statistics (Harald Cramér vol)* Almquist and Wiksalls, Uppsala, Sweden, 416-444.

Neyman, J., and Scott, E.L. (1960) Correction for bias introduced by a transformation of variables. *Ann. Math. Statist.*, 31, 643-655.

Neyman, J., and Scott, E.L. (1965) Asymptotically optimal tests of composite hypotheses for randomized experiments with non-controlled predictor variables. *J. Amer. Statist. Ass.*, 60, 72-99.

Penrose, L.S., and Smith, G.F. (1966) *Down's Anomaly.* Churchill, London.

Perlman, M.D. (1969) One sided testing procedures in multivariate analysis. *Ann. Math. Statist.*, 40, 549-567.

Rao, C.R. (1948) Large sample tests of statistical hypotheses concerning several parameters, with applications to problems of estimation. *Proc. Camb. Phil. Soc.*, 44, 50-57.

Rao, C.R. (1961) A study of large sample test criteria through properties of efficient estimates. *Sankhya*, A, 23, 25-40.

Rao, C.R. (1965) *Linear statistical inference and its applications.* Wiley, New York.

Schaafsma, W., and Smid, L.J. (1966) Most stringent somewhere most powerful tests against alternatives restricted by a number of linear inequalities. *Ann. Math. Statist.*, 37, 1161-1172.

Stuart, A. (1962) Gamma distributed products of independent random variables. *Biometrika*, 49, 564-565.

Tocher, K.D. (1963) *The art of simulation.* English Universities Press, London.

Wald, A. (1943) Tests of statistical hypotheses concerning several parameters when the number of observations is large. *Amer. Math. Soc.*, 54, 26-82.

Wald, A. (1949) On the consistency of the maximum likelihood estimator. *Ann. Math. Statist.*, 20, 597-601.

Wolfowitz, J. (1949) On Wald's proof of the consistency of the maximum likelihood estimator. *Ann. Math. Statist.*, 20, 601-602.