AN ABSTRACT OF THE THESIS OF

SOO-IL KANG for the degree of DOCTOR OF PHILOSOPHY

in STATISTICS presented on October 6, 1978

Title: PERFORMANCE OF GENERALIZED NEYMAN SMOOTH GOODNESS OF FIT TESTS

Abstract approved: Redacted for Privacy

Dr. David R. Thomas

The small sample performance of generalized Neyman smooth goodness of fit tests for univariate distributions is investigated. The general form of such tests was developed by Javitz (1975) and Thomas and Pierce (1977) as extensions to composite hypotheses of the test proposed by Neyman (1937) for a simple hypothesis against "smooth" alternatives. The performance study is based on simulated random samples of size $N = 20$ and 50 from several distributions. The generalized smooth tests of fit for normal and extreme value distributions perform well with regard to power when compared to several other tests, including the classical Pearson χ^2 test and a generalized χ^2 test.

The smooth tests of fit are further generalized for linear models, including normal regression and extreme value regression models.

Performance of Generalized Neyman Smooth Goodness of Fit Tests

by

Soo-Il Kang

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

June 1979

APPROVED:

Redacted for Privacy

Associate Professor of Statistics

in charge of major

Redacted for Privacy

Head of Department of Statistics

Redacted for Privacy

Dean of Graduate School

Date thesis is presented October 6, 1968

Typed by Clover Redfern for Soo-Il Kang

ACKNOWLEDGMENT

The author greatly wishes to express his deepest appreciation to Dr. David Thomas who served as the author's major professor. Without his steady technical guidance, patience, and encouragement, this thesis would never have been written.

The author wishes to thank you, the many members of the Department of Statistics, which whom he has had many valuable discussions. Special thanks go to Dr. Donald Pierce and Dr. David Birkes. Thanks, also, to the O. S. U. Computer Center for a generous allocation of unsponsored research funds.

Actually, it was the encouragement given by the author's wife, Jung-Hee (Julia), that prompted him to attempt the doctoral program. To her and to our children--Bella and Barry--the author expresses his deepest love and appreciation.

TABLE OF CONTENTS

LIST OF TABLES

PERFORMANCE OF GENERALIZED NEYMAN SMOOTH GOODNESS OF FIT TESTS

1. INTRODUCTION

A problem that has received much attention is that of testing goodness of fit (GOF) of a distribution model. Such models may either completely specify the distribution (simple hypothesis) or specify a parametric family of distributions (composite hypothesis). Models that are frequently used include the normal or lognormal, Weibull or extreme value, negative exponential, and gamma distributions. The classical chi square GOF test, originated by K. Pearson (1900), and the Kolmogorov-Smirnov GOF test appropriate for composite hypotheses (Stephens, 1969, 1970, 1974; Green and Hegazy, 1967) are widely used.

Neyman (1937) proposed a GOF test for a completely specified distribution with smooth alternatives. He adopted the word "smooth" to indicate edgeless alternative pdf curves which are not much different from the null distribution pdf curve. He considered the simple hypothesis that a random variable Y has a pdf f(y) and cdf F(y) against the class of smooth alternative pdf's,

$$g(y|\underset{\sim}{\theta}) = C(\underset{\sim}{\theta}) \exp\left(\sum_{i=1}^{k} \theta_i \pi_i(F(y)) \right) f(y),$$

where $\theta = (\theta_1, \ldots, \theta_k)'$ is a real vector belonging to an open set, and $\pi_i(\cdot)$ is the i^{th} order Legendre orthogonal polynomial. Let $y_1, \ldots, y_N$ be a random sample from a family with hypothetical distribution function $F(y)$. Then for the simple null hypothesis,

$$H_0: \theta_1 = \ldots = \theta_k = 0,$$

$$\psi_k^2 = \sum_{i=1}^{k} u_i^2$$

has a limiting χ_k^2 distribution, where

$$u_i = \frac{1}{\sqrt{N}} \sum_{j=1}^{N} \pi_i[F(y_j)], \quad i = 1, \ldots, k.$$

He derived directly the asymptotic noncentral χ^2 distribution of ψ_k^2 under local alternatives. He recommended that k generally need not be greater than four.

Barton (1956) subsequently studied the limiting distribution of the ψ_k^2 statistics for composite null hypotheses. He used the same form for ψ_k^2, only with $F(y)$ replaced by $F(y|\hat{\underset{\sim}{\lambda}})$, where $\hat{\underset{\sim}{\lambda}}$ is an estimator of a parameter vector, $\underset{\sim}{\lambda}$. Barton's generalized test statistic does not have a limiting χ^2 distribution. However, he approximated the limiting distribution of the test statistic by a linear combination of χ^2. He considered the power of ψ_k^2 for some local alternatives whose forms are expressed by

$$g(y|\underset{\sim}{\theta}, \hat{\underset{\sim}{\lambda}}) = \left(1 + \sum_{i=1}^{k} \theta_i \pi_i [F(y|\hat{\underset{\sim}{\lambda}})]\right) C(\underset{\sim}{\theta}) f(y|\hat{\underset{\sim}{\lambda}})$$

where $\pi_i(y)$ is the i^{th} order Legendre polynomial in y.

Javitz (1975) applied Neyman's $C(\alpha)$ theory approach to obtain a generalized Neyman smooth test for the composite case. For testing normality, he also examined the noncentrality parameter(s) of ψ_k^2 when the smooth alternatives are of various orders and used a simulation study to compare the small sample powers of his tests with these of the classical chi square test.

Thomas and Pierce (1977) used $F^i(\cdot)$ instead of $\pi_i[F(\cdot)]$ and the large sample theory for score statistics with maximum likelihood estimator (MLE) for $\underset{\sim}{\lambda}$ as discussed in Cox and Hinkley (1974, Ch. 9.1-9.3). Their MLE of nuisance parameters is equivalent to Javitz's generalization for MLE of $\underset{\sim}{\lambda}$ assuming $\underset{\sim}{\theta} = \underset{\sim}{0}$. The Javitz (1975) and Thomas and Pierce (1977) generalized smooth tests possess the property that their statistics have limiting χ^2 distributions under the null hypothesis.

Kopecky (1977) investigated the asymptotic efficiency of generalized smooth tests (GST) and considered regularity conditions to assure limiting chi square distributions.

The main purpose of this thesis is to study the small sample performance of the GST. A simulation study is discussed in Chapter 3 for comparing the power of the GST with several alternative test statistics. First, the GST are constructed in Chapter 2 for univariate distributions and linear models. Applications are illustrated by numerical examples in Chapter 4.

2. CONSTRUCTION OF THE GENERALIZED SMOOTH TESTS

2.1. Univariate Models

2.1.1. General Case

Let $y_1, y_2, \ldots, y_N$ be a random sample from some distribution. For a goodness-of-fit hypothesis,

$$H_0 : Y \sim F(y|\underset{\sim}{\lambda}), \quad \underset{\sim}{\lambda} = (\lambda_1, \ldots, \lambda_{p+1})' \in \Lambda, \tag{2.1}$$

the cumulative distribution function (cdf) $F(y|\underset{\sim}{\lambda})$ is a member of some specified parametric family. For the smooth alternatives,

$$H_a : Y \sim G(y|\underset{\sim}{\lambda}, \underset{\sim}{\theta}), \quad \underset{\sim}{\theta} = (\theta_1, \ldots, \theta_k)' \in \Theta, \tag{2.2}$$

let $G(y|\underset{\sim}{\lambda}, \underset{\sim}{\theta})$ denote the cdf whose density function is

$$g(y|\underset{\sim}{\lambda}, \underset{\sim}{\theta}) = e^{\Sigma_{i=1}^{k} \theta_i F^i(y|\underset{\sim}{\lambda}) - R(\underset{\sim}{\theta})} \cdot f(y|\underset{\sim}{\lambda})$$

where

$$R(\underset{\sim}{\theta}) = \text{Log}\left(\int_{-\infty}^{\infty} e^{\Sigma_{i=1}^{k} \theta_i F^i(y|\underset{\sim}{\lambda})} f(y|\underset{\sim}{\lambda}) dy\right)$$

is the normalizing constant. Consider the following hypotheses

$$H_0: \underset{\sim}{\theta} = \underset{\sim}{0} \; (\equiv \underset{\sim}{\theta}_0) \quad \text{vs.} \tag{2.3}$$

$$H_a: \underset{\sim}{\theta} \neq \underset{\sim}{0} \tag{2.4}$$

Then testing hypotheses (2.1) vs. (2.2) is equivalent to testing (2.3) vs. (2.4).

The generalized Neyman smooth test (GST) is constructed by employing the general large sample theory described in Cox and Hinkley (1974, pp. 279-284). The log-likelihood function, $\ell \equiv \ell(\underset{\sim}{\theta}, \underset{\sim}{\lambda})$, is

$$\ell = \sum_{j=1}^{N} \text{Log}[g(y_j | \underset{\sim}{\lambda}, \underset{\sim}{\theta})]$$

$$= \sum_{j=1}^{N} \left(\sum_{i=1}^{k} \theta_i F^i(y_j | \underset{\sim}{\lambda}) - R(\underset{\sim}{\theta}) + \text{Log } f(y_j | \underset{\sim}{\lambda}) \right). \tag{2.5}$$

Define the score vectors $U_{\underset{\sim}{\theta}}$ and $U_{\underset{\sim}{\lambda}}$ by

$$U_{\underset{\sim}{\theta}} = \frac{\partial \ell}{\partial \underset{\sim}{\theta}} \bigg|_{\underset{\sim}{\theta}_0, \hat{\underset{\sim}{\lambda}}} \quad \text{and} \quad U_{\underset{\sim}{\lambda}} = \frac{\partial \ell}{\partial \underset{\sim}{\lambda}} \bigg|_{\underset{\sim}{\theta}_0, \hat{\underset{\sim}{\lambda}}}$$

with i^{th} components,

$$U_{\underset{\sim}{\theta}}(i) = \frac{\partial \ell}{\partial \theta_i} \bigg|_{\underset{\sim}{\theta}_0, \hat{\underset{\sim}{\lambda}}} = \sum_{j=1}^{N} [F^i(y_j | \hat{\underset{\sim}{\lambda}}) - \frac{\partial R(\underset{\sim}{\theta})}{\partial \theta_i}]_{\underset{\sim}{\theta}_0} = \sum_{j=1}^{N} F^i(y_j | \hat{\underset{\sim}{\lambda}}) - \frac{N}{i+1}, \tag{2.6}$$

since

$$\frac{\partial R(\underset{\sim}{\theta})}{\partial \theta_i}\Big|_{\underset{\sim}{\theta}_0} = \frac{\frac{\partial}{\partial \theta_i} e^{R(\underset{\sim}{\theta})}}{e^{R(\underset{\sim}{\theta})}}\Bigg|_{\underset{\sim}{\theta}_0}$$

$$= \frac{\int_{-\infty}^{\infty} F^i(y|\underset{\sim}{\lambda}) e^{\Sigma_{t=1}^{k} \theta_t F^t(y|\underset{\sim}{\lambda})} f(y|\underset{\sim}{\lambda})dy}{\int_{-\infty}^{\infty} e^{\Sigma_{t=1}^{k} \theta_t F^t(y|\underset{\sim}{\lambda})} f(y|\underset{\sim}{\lambda})dy}\Bigg|_{\underset{\sim}{\theta}_0}$$

$$= \frac{\int_0^1 u^i e^{\Sigma\theta_t u^t} du}{\int_0^1 e^{\Sigma\theta_t u^t} du}\Bigg|_{\underset{\sim}{\theta}_0}$$

$$= \frac{1}{i+1} \qquad \text{for} \quad i = 1, \ldots, k,$$

and

$$U_{\underset{\sim}{\lambda}}(i) = \frac{\partial \ell}{\partial \lambda_i}\Big|_{\underset{\sim}{\theta}_0, \hat{\underset{\sim}{\lambda}}} = 0 \qquad \text{for} \quad i = 1 \ \ldots, p+1$$

where $\hat{\underset{\sim}{\lambda}}$ is MLE of $\underset{\sim}{\lambda}$.

The elements of the partitioned information matrix of dimension $(k+p+1) \times (k+p+1)$ are denoted by

$$I(\underset{\sim}{\theta}_0, \hat{\underset{\sim}{\lambda}}) = \begin{bmatrix} I_{\underset{\sim}{\theta}\underset{\sim}{\theta}}(\underset{\sim}{\theta}_0) & I_{\underset{\sim}{\theta}\underset{\sim}{\lambda}}(\underset{\sim}{\theta}_0, \hat{\underset{\sim}{\lambda}}) \\ I'_{\underset{\sim}{\theta}\underset{\sim}{\lambda}}(\underset{\sim}{\theta}_0, \hat{\underset{\sim}{\lambda}}) & I_{\underset{\sim}{\lambda}\underset{\sim}{\lambda}}(\hat{\underset{\sim}{\lambda}}) \end{bmatrix}$$

$$\equiv \begin{bmatrix} I_{\theta\theta} & I_{\theta\lambda} \\ I'_{\theta\lambda} & I_{\lambda\lambda} \end{bmatrix} \qquad \text{for} \quad \underset{\sim}{\theta} = \underset{\sim}{\theta}_0 \tag{2.7}$$

and evaluated as follows: For $i, j = 1, \ldots, k,$

$$I_{\theta\theta}(i,j) = E\left(- \frac{\partial^2 \ell}{\partial\theta_i \partial\theta_j}\right)_{\underset{\sim}{\theta}_0}$$

$$= N\left(\frac{\partial^2 R(\underset{\sim}{\theta})}{\partial\theta_i \partial\theta_j}\right)_{\underset{\sim}{\theta}_0}$$

$$= N\left[\int_{-\infty}^{\infty} F^{i+j}(z) f(z)dz - \left(\int_{-\infty}^{\infty} F^{j}(z) f(z)dz\right)\left(\int_{-\infty}^{\infty} F^{i}(z) f(z)dz\right)\right]$$

$$= \frac{ijN}{(i+1)(j+1)(i+j+1)}\,, \tag{2.8}$$

which is independent of $F(y|\underset{\sim}{\lambda})$. For $i = 1, \ldots, k,\ j = 1, \ldots, p+1,$

$$I_{\theta\lambda}(i,j) = E\left(- \frac{\partial^2 \ell}{\partial\theta_i \partial\lambda_j}\right)_{\underset{\sim}{\theta}_0, \hat{\underset{\sim}{\lambda}}}$$

$$= E\left(\sum_{t=1}^{N} \left[- \frac{\partial F^i(Y_t|\underset{\sim}{\lambda})}{\partial\lambda_j}\right]\right)_{\hat{\underset{\sim}{\lambda}}}$$

$$= -NE\left(iF^{i-1}(Y|\underset{\sim}{\lambda}) \frac{\partial}{\partial\lambda_j} F(Y|\underset{\sim}{\lambda})\right)_{\hat{\underset{\sim}{\lambda}}} \tag{2.9}$$

And for $i, j = 1, \ldots, p+1$,

$$I_{\lambda\lambda}(i,j) = E\left(-\frac{\partial^2 \ell}{\partial\lambda_i \partial\lambda_j}\right)_{\underset{\sim}{\theta}_0, \hat{\underset{\sim}{\lambda}}}$$

$$= -NE\left(\frac{\partial^2}{\partial\lambda_i \partial\lambda_j} \operatorname{Log}(f(Y|\underset{\sim}{\lambda}))\right)_{\hat{\underset{\sim}{\lambda}}}$$

$$= -N\int_{-\infty}^{\infty}\left(\frac{\partial^2 f(y|\underset{\sim}{\lambda})}{\partial\lambda_i \partial\lambda_j} - \frac{\left(\frac{\partial}{\partial\lambda_i} f(y|\underset{\sim}{\lambda})\right) \frac{\partial}{\lambda_j} f(y|\underset{\sim}{\lambda})}{f(y|\underset{\sim}{\lambda})}\right)_{\hat{\underset{\sim}{\lambda}}} dy. \tag{2.10}$$

From (2. 8) the matrix $I_{\theta\theta}$ with $K = 5$ is

$$I_{\theta\theta} = N \begin{pmatrix} \frac{1}{12} & \frac{1}{12} & \frac{3}{40} & \frac{1}{15} & \frac{5}{84} \\ & \frac{4}{45} & \frac{1}{12} & \frac{8}{105} & \frac{5}{72} \\ & & \frac{9}{112} & \frac{3}{40} & \frac{5}{72} \\ & & & \frac{16}{225} & \frac{1}{15} \\ & & & & \frac{25}{396} \end{pmatrix}. \tag{2.11}$$

Denote the partitioned matrix in the inverse of the full information matrix (2. 7) corresponding to $I_{\theta\theta}$ by $I^{\theta\theta}$ and

$$I_{\theta\theta|\lambda} \equiv I_{\theta\theta} - I_{\theta\lambda} I_{\lambda\lambda}^{-1} I_{\lambda\theta} \, . \tag{2.12}$$

In computing $I^{\theta\theta}$, depending on convenience, one of the following two formulas can be used (Rao, 1973, p. 33):

$$I^{\theta\theta} = I_{\theta\theta}^{-1} + I_{\theta\theta}^{-1} I_{\theta\lambda} (I_{\lambda\lambda} - I_{\lambda\theta} I_{\theta\theta}^{-1} I_{\theta\lambda})^{-1} I_{\lambda\theta} I_{\theta\theta}^{-1} \quad (2.13)$$

$$I^{\theta\theta} = I_{\theta\theta|\lambda}^{-1} \quad (2.14)$$

For any m when $m \leq k$, the matrix $I^{\theta\theta}(m)$ is obtained by taking the inverse of the left upper corner square matrix $(m \times m)$ of $I_{\theta\theta|\lambda}$.

The GST statistics, W_k, are then constructed as

$$W_k = [U_{\underset{\sim}{\theta}} - I_{\theta\lambda} I_{\lambda\lambda}^{-1} U_{\underset{\sim}{\lambda}}]' I^{\theta\theta} [U_{\underset{\sim}{\theta}} - I_{\theta\lambda} I_{\lambda\lambda}^{-1} U_{\underset{\sim}{\lambda}}]$$

$$= U_{\underset{\sim}{\theta}}' I^{\theta\theta} U_{\underset{\sim}{\theta}} \quad (2.15)$$

From the large sample theory, under suitable regularity conditions (Kopecky, 1977), W_k has a limiting χ^2 distribution with k degrees of freedom when H_0 is true. For some distributions, for instance, the Logistic distribution, $I_{\theta\theta|\lambda}$ is a singular matrix. In such cases certain score(s) can be omitted to obtain nonsingularity. The procedure is that row(s) and column(s) of $I_{\theta\theta}$, $I_{\theta\lambda}$, and $I_{\lambda\theta}$, corresponding to the omitted score(s), are deleted. Then the statistic, W_k in (2.15) has a reduced degree of freedom (df), $k - r$ rather than k, where r is the number of the omitted score(s).

2.1.2. Scale-Location Parameter Families

For scale and location parameter families denote the cdf by

$$F(y|\underset{\sim}{\lambda}) = F_0\left(\frac{y-\mu}{\sigma}\right),$$

where $F_0(\cdot)$ is some specified cdf and

$$\underset{\sim}{\lambda} = (\lambda_1, \lambda_2)' = (\mu, \sigma)'$$

Then the i^{th} component of the score vector $U_{\underset{\sim}{\theta}}$, $U_{\underset{\sim}{\theta}}(i)$ is

$$U_{\underset{\sim}{\theta}}(i) = \sum_{j=1}^{N} F_0^i\left(\frac{y_j-\hat{\mu}}{\hat{\sigma}}\right) - \frac{N}{i+1}$$

Evaluation of equations (2.9) and (2.10) for $\lambda_1 = \mu$ and $\lambda_2 = \sigma$ gives respectively, for $i = 1, \ldots, k$,

$$\sigma I_{\theta\lambda}(i, 1) = -iN\int_{-\infty}^{\infty} F_0^{i-1}\left(\frac{y-\mu}{\sigma}\right) f_0\left(\frac{y-\mu}{\sigma}\right) \frac{\partial}{\partial\mu} F_0\left(\frac{y-\mu}{\sigma}\right) dy$$

$$= iN\int_{-\infty}^{\infty} F_0^{i-1}(z)\, f_0^2(z)\, dz,$$

$$\sigma I_{\theta\lambda}(i, 2) = -iN\int_{-\infty}^{\infty} F_0^{i-1}\left(\frac{y-\mu}{\sigma}\right) f_0\left(\frac{y-\mu}{\sigma}\right) \frac{\partial}{\partial\sigma} F_0\left(\frac{y-\mu}{\sigma}\right) dy$$

$$= iN\int_{-\infty}^{\infty} F_0^{i-1}(z)\, f_0^2(z)\, z\, dz, \qquad (2.16)$$

$$\sigma^2 I_{\lambda\lambda}(1,1) = -\sigma\ N\int_{-\infty}^{\infty}\left(\frac{\partial^2 f_0(\frac{y-\mu}{\sigma})}{(\partial\mu)^2} - \frac{[\frac{\partial}{\partial\mu} f_0(\frac{y-\mu}{\sigma})]^2}{f_0(\frac{y-\mu}{\sigma})}\right) dy$$

$$= -N\int_{-\infty}^{\infty}\left(f_0''(z) - \frac{[f_0'(z)]^2}{f_0(z)}\right) dz,$$

$$\sigma^2 I_{\lambda\lambda}(1,2) = -N\sigma\int_{-\infty}^{\infty}\left(\frac{\partial^2 f_0(\frac{y-\mu}{\sigma})}{\partial\mu\partial\sigma} - \frac{\frac{\partial}{\partial\mu} f_0(\frac{y-\mu}{\sigma})\frac{\partial}{\partial\sigma} f_0(\frac{y-\mu}{\sigma})}{f_0(\frac{y-\mu}{\sigma})}\right) dy$$

$$= -N\int_{-\infty}^{\infty}\left(z\ f_0''(z)+2\ f_0'(z) - \frac{[f_0'(z)]^2}{f_0(z)}\right) dz, \qquad (2.17)$$

and

$$\sigma^2 I_{\lambda\lambda}(2,2) = -N\sigma\int_{-\infty}^{\infty}\left(\frac{\partial^2 f_0(\frac{y-\mu}{\sigma})}{(\partial\sigma)^2} - \frac{[\frac{\partial}{\partial\sigma} f_0(\frac{y-\mu}{\sigma})]^2}{f_0(\frac{y-\mu}{\sigma})}\right) dy$$

$$= -N\int_{-\infty}^{\infty}\left(2f_0(z)+4z\ f_0'(z)+z^2 f_0''(z) - \frac{(f_0(z)+zf_0'(z))^2}{f_0(z)}\right) dz$$

The matrix, $I^{\theta\theta}$ is then computed by using either (2.13) or (2.14).

Theorem 1. The matrix $I^{\theta\theta}$ is independent of the nuisance parameters μ and σ.

Proof. This result can be seen from relation (2.12), since $I_{\theta\theta}$ is independent of μ and σ and the relation

$$I_{\theta\lambda} I_{\lambda\lambda}^{-1} I_{\lambda\theta} = [\sigma I_{\theta\lambda}][\sigma^2 I_{\lambda\lambda}]^{-1}[\sigma I_{\lambda\theta}] \tag{2.18}$$

where the matrices on the right hand side of the equation are shown in (2.14), (2.16), and (2.17) to be independent of μ and σ.

2.1.3. Special Cases of Scale-Location Parameter Families

In this section we evaluate the $I^{\theta\theta}$ matrices for normal, extreme value, and logistic distributions.

a) Normal Distribution

Let $F_0(z) = \Phi(z)$, $-\infty < z < \infty$, denote the standard normal cdf and let $Z_{(m),n}$ be the m^{th} smallest order statistic in a random sample of size n from the standard normal distribution. From (2.16),

$$\sigma I_{\theta\lambda}(i,1) = iN \int_{-\infty}^{\infty} \Phi^{i-1}(z)\, \phi^2(z)dz$$

$$= -iN \int_{-\infty}^{\infty} [(i-1)\Phi^{i-2}(z)\phi^2(z) - \Phi^{i-1}(z)\phi(z)z]\Phi(z)dz$$

$$= (1-i)\sigma I_{\theta\lambda}(i,1) + iN \int_{-\infty}^{\infty} z\Phi^{i}(z)\phi(z)dz \quad =$$

$$= NE(Z\Phi^i(Z))$$

$$= \frac{N}{(i+1)} E(Z_{(i+1), i+1}), \qquad i = 1, \ldots, k,$$

$$\sigma I_{\theta\lambda}(i, 2) = iN \int_{-\infty}^{\infty} z\Phi^{i-1}(z)\, \phi^2(z)dz$$

$$= N \left(-\frac{1}{i+1} + \int_{-\infty}^{\infty} \Phi^i(z)\, z^2\, \phi(z)dz \right)$$

$$= \frac{N}{i+1} [E(Z^2_{(i+1), i+1}) - 1]$$

$$= \frac{N}{i+1} [-1 + Var(Z_{(i+1), i+1}) + \{E(Z_{(i+1), i+1})\}^2],$$

$$i = 1, \ldots, k$$

From (2. 17),

$$\sigma^2 I_{\lambda\lambda}(1, 1) = -N \int_{-\infty}^{\infty} [\phi''(z) - \frac{(\phi'(z))^2}{\phi(z)}]dz$$

$$= -N \int_{-\infty}^{\infty} -\phi(z)dz$$

$$= N$$

$$\sigma^2 I_{\lambda\lambda}(1, 2) = -N \int_{-\infty}^{\infty} [z\phi''(z) + 2\phi'(z) - \frac{[\phi'(z)]^2}{\phi(z)}]dz$$

$$= -N \int_{-\infty}^{\infty} 3z\phi(z)dz$$

$$= 0$$

$$\sigma^2 I_{\lambda\lambda}(2,2) = -N\int_{-\infty}^{\infty}[2\phi(z)+4z\cdot\phi(z)+z\phi''(z)-\frac{[\phi(z)+z\phi'(z)]^2}{\phi(z)}]dz$$

$$= -N\int_{-\infty}^{\infty}[\phi(z)-3z^2\phi(z)]dz$$

$$= 2N$$

For $K = 4$, numerical values of the matrices $I_{\theta\lambda}$, $I_{\lambda\lambda}$, and $I_{\theta\theta|\lambda}$, are

$$\sigma I_{\theta\lambda} = N \begin{pmatrix} .2820947918 & 0. \\ .2820947918 & .09188814923 \\ .2573438433 & .1378322239 \\ .2325928947 & .1600040872 \end{pmatrix},$$

$$\sigma^2 I_{\lambda\lambda} = N \begin{pmatrix} 1 & 0 \\ & 2 \end{pmatrix},$$

$$I_{\theta\theta|\lambda} = I_{\theta\theta} - I_{\theta\lambda} I_{\lambda\lambda}^{-1} I_{\lambda\theta}$$

$$= \frac{N}{1000} \begin{pmatrix} 3.755861773 & 3.755861773 & 2.404642103 & 1.053422462 \\ & 5.089701344 & 4.405401457 & 3.225992265 \\ & & 4.632428200 & 4.116790968 \\ & & & 4.211002486 \end{pmatrix}$$

The $I^{\theta\theta}$ matrices, with $K = 1,2,3,4$, for testing goodness of fit of normal, extreme value, logistic, and negative exponential distributions are in Tables 2.1-2.4, respectively.

Table 2.1. The $I^{\theta\theta}$ matrices for testing normality with $K \leq 4$.

K = 1

$$\frac{1}{N}\begin{bmatrix} 266.2504801 \end{bmatrix}$$

K = 2

$$\frac{1}{N}\begin{bmatrix} 1015.965829 & -749.7153493 \\ & 749.7153493 \end{bmatrix}$$

K = 3

$$\frac{1}{N}\begin{bmatrix} 9072.727849 & -14806.08280 & 9370.911645 \\ & 25273.29688 & -16349.12104 \\ & & 10899.41404 \end{bmatrix}$$

K = 3

$$\frac{1}{N}\begin{bmatrix} 23058.99425 & -66471.76917 & 84729.75129 & -37679.41831 \\ & 216127.9875 & -294726.9285 & 139188.8981 \\ & & 416937.3481 & -203018.9588 \\ & & & 101509.4753 \end{bmatrix}$$

b) Logistic Distribution

The cdf and pdf for the standard logistic distribution are respectively,

$$F_0(z) = e^z/(1+e^z)$$

and

$$f_0(z) = e^z/(1+e^z)^2, \qquad -\infty < z < \infty$$

Then for i=1, . . . , k evaluation of equations (2.16) give

$$\sigma I_{\theta\lambda}(i, 1) = \frac{iN}{(i+1)} \int_{-\infty}^{\infty} [\frac{e^z}{1+e^z}]^{i+1} \frac{e^z}{(1+e^z)^2} dz$$

$$= \frac{iN}{(i+1)(i+2)}$$

and

$$\sigma I_{\theta\lambda}(i, 2) = \frac{iN}{(i+1)} \int_{-\infty}^{\infty} [\frac{e^z}{1+e^z}]^{i+1} [\frac{ze^z}{(1+e^z)^2} - \frac{1}{1+e^z}]dz$$

$$= \frac{iN}{(i+1)} \left(\int_0^1 u^{i+1} \operatorname{Log}(\frac{u}{1-u})du - \frac{1}{i+1} \right)$$

$$= \frac{iN}{(i+1)} \left[-\frac{1}{(i+2)^2} + \sum_{s=0}^{i+1} \binom{i+1}{s}(-1)^s \frac{1}{(s+1)^2} - \frac{1}{i+1} \right]$$

From equation (2.17), it is found that

$$\sigma^2 I_{\lambda\lambda}(1,1) = N\int_{-\infty}^{\infty} 2e^{2z}/(1+e^z)^4 dz$$

$$= \frac{N}{3}$$

$$\sigma^2 I_{\lambda\lambda}(1,2) = N\int_{-\infty}^{\infty} \frac{e^{3z}-e^z+2ze^{2z}}{(1+e^z)^4}\, dz$$

$$= 0,$$

and

$$\sigma^2 I_{\lambda\lambda}(2,2) = N\left[\int_{-\infty}^{\infty} \frac{2e^{2z}-2ze^z+2ze^{3z}}{(1+e^z)^4}\, dz - 1\right]$$

$$= N\left[2E(Q_2)\,\frac{[\Gamma(2)]^2}{\Gamma(4)} - 2E(Q_1)\,\frac{\Gamma(1)\,\Gamma(3)}{\Gamma(4)} + 2E(Q_3)\,\frac{\Gamma(3)\,\Gamma(1)}{\Gamma(4)} - 1\right]$$

$$= N[\tfrac{2}{3}\psi'(2) - \tfrac{4}{3}(\psi(1)-\psi(3)) - 1]$$

$$= \frac{3+\pi^2}{9}\, N$$

where Q_i, $i = 1, 2, 3$, denote the i^{th} order statistics of an independent sample of size 3 from the standard logistic distribution and $\psi'(\cdot)$ is the first derivative of the digamma function, $\psi(\cdot)$ (Johnson and Kotz, 1970, Chap. 22.3-22.7).

For $K = 5$, numerical values of $I_{\theta\lambda}$, $I_{\lambda\lambda}$, and $I_{\theta\theta|\lambda}$ are respectively,

$$\sigma I_{\theta\lambda} = N \begin{pmatrix} \frac{1}{6} & 0 \\ \frac{1}{6} & \frac{1}{12} \\ \frac{3}{20} & \frac{1}{8} \\ \frac{2}{15} & \frac{13}{90} \\ \frac{5}{42} & \frac{11}{72} \end{pmatrix},$$

$$\sigma^2 I_{\lambda\lambda} = N \begin{pmatrix} \frac{1}{3} & 0 \\ & 1.429956045 \end{pmatrix},$$

and

$$I_{\theta\theta|\lambda} = \frac{N}{1000} \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ & .6991514243 & 1.048727137 & 1.106042363 & 1.017227347 \\ & & 1.930233562 & 2.373349259 & 2.517904512 \\ & & & 3.186981366 & 3.615045919 \\ & & & & 4.291392443 \end{pmatrix}$$

Of course, $I_{\theta\theta|\lambda}$ is a singular matrix. As mentioned above, omit the first row and column of $I_{\theta\theta}$ and the first row of $I_{\theta\lambda}$. Let $I_{\theta\theta}$ and $I_{\theta\lambda}$ also denote the corresponding matrices after deletion. The rank of $I_{\theta\theta|\lambda}$, evaluated from (2.12), is thus $R = k-1$.

Table 2.2. The $I^{\theta\theta}$ matrices for testing the Logistic distribution with $R \leq 4$.

$R = 1$

$$\frac{1}{N}\begin{bmatrix} 1430.305318 \end{bmatrix}$$

$R = 2$

$$\frac{1}{N}\begin{bmatrix} 7730.305318 & -4200 \\ & 2800 \end{bmatrix}$$

$R = 3$

$$\frac{1}{N}\begin{bmatrix} 239522.5025 & -331123.3367 & 163461.6684 \\ & 463897.7823 & -230548.8911 \\ & & 115274.4456 \end{bmatrix}$$

$R = 4$

$$\frac{1}{N}\begin{bmatrix} 724622.5025 & -1624723.337 & 1618761.668 & -582120 \\ & 3913497.782 & -4111348.891 & 1552320 \\ & & 4481174.446 & -1746360 \\ & & & 698544 \end{bmatrix}$$

c) Extreme Value Distribution

The cdf and pdf for the standard extreme value distribution are respectively,

$$F_0(z) = 1 - \exp[-\exp(z)]$$

and

$$f_0(z) = \exp[z - \exp(z)], \qquad -\infty < z < \infty$$

Denote $\underset{\sim}{\lambda} = (\mu, \sigma)'$ and $\overline{F}(\cdot) = 1 - F(\cdot)$. Under the transformations, $Y = \mu + \sigma Z = \text{Log}(X)$, $\mu_1 = e^{\mu}$, and $\sigma_1 = \sigma^{-1}$, then X has a Weibull distribution (WD) with pdf

$$f(x|\mu_1, \sigma_1) = \frac{\sigma_1}{\mu_1} \left(\frac{x}{\mu_1}\right)^{\sigma_1 - 1} e^{-(x/\mu_1)^{\sigma_1}}, \quad x \geq 0,\ \mu_1, \sigma_1 > 0,$$

while Y has an extreme value distribution (EVD). Consequently, the matrices $I^{\theta\theta}$ in EVD and those in WD are identical. Therefore testing EVD is the same as testing WD. Hereafter, the former is considered. For convenience, without losing generality, $\overline{F}(y|\underset{\sim}{\lambda}) = 1 - F(y|\underset{\sim}{\lambda})$ is used in place of $F(y|\underset{\sim}{\lambda})$. The evaluation of equations (2.16), give, for $i = 1, \ldots, k$

$$\sigma I_{\theta\lambda}(i, 1) = -N\int_{-\infty}^{\infty} i\, e^{2z-(i+1)e^{z}}\, dz$$

$$= -iN\int_{0}^{\infty} ue^{-(i+1)u}du$$

$$= -\frac{iN}{(i+1)^2}$$

and

$$\sigma I_{\theta\lambda}(i, 2) = -N\int_{-\infty}^{\infty} iz\, e^{2z-(i+1)e^{z}}\, dz$$

$$= -iN\int_{0}^{\infty} u\left[\text{Log } u\right]e^{-(i+1)u}du$$

$$= -iN\int_{0}^{\infty}\left[\text{Log}\frac{w}{i+1}\right]\frac{w}{i+1}\, e^{-w}\,\frac{dw}{i+1}$$

$$= -\frac{iN}{(i+1)^2}\left[\int_{0}^{\infty} w\left[\text{Log } w\right]e^{-w}dw - \text{Log}(i+1)\int_{0}^{\infty} we^{-w}dw\right]$$

$$= -\frac{iN}{(i+1)^2}\,[1-\gamma-\text{Log}(i+1)],$$

where the Euler's constant,

$$\gamma = -\int_{0}^{\infty} e^{x}\,\text{Log}\, x\, dx = .577215664\ldots\ .$$

From equation (2. 17), it is found that

$$\sigma^2 I_{\lambda\lambda}(1, 1) = N \int_{-\infty}^{\infty} e^z e^{z-e^z} dz$$

$$= N \int_0^{\infty} ue^{-u} du$$

$$= N,$$

$$\sigma^2 I_{\lambda\lambda}(1, 2) = N \int_{-\infty}^{\infty} [(1+z)e^z - 1]e^{z-e^z} dz$$

$$= N \int_0^{\infty} u\left(\text{Log } u\right) e^{-u} du$$

$$= N(1-\gamma),$$

and

$$\sigma^2 I_{\lambda\lambda}(2, 2) = N \int_{-\infty}^{\infty} [2z(e^z - 1) + z^2 e^z - 1]e^{z-e^z} dz$$

$$= N\left\{2\int_0^{\infty} ue^{-u}\left(\text{Log } u\right) du - 2\int_0^{\infty} e^{-u}\left(\text{Log } u\right) du + \int_0^{\infty} ue^{-u}(\text{Log } u)^2 du - 1\right\}$$

$$= N[(1-\gamma)^2 + \psi'(1)] \qquad (2.19)$$

where γ and $\psi'(\cdot)$ are defined above. For $K = 4$, numerical values of $I_{\theta\lambda}$, $I_{\lambda\lambda}$, and $I_{\theta\theta|\lambda}$ are respectively,

$$\sigma I_{\theta\lambda} = N \begin{bmatrix} -\frac{1}{4} & .06759071137 \\ -\frac{2}{9} & .1501839897 \\ -\frac{3}{16} & .1806581299 \\ -\frac{4}{25} & .1898647524 \end{bmatrix},$$

$$\sigma^2 I_{\lambda\lambda} = N \begin{bmatrix} 1 & .4227843351 \\ & 1.823680661 \end{bmatrix},$$

and

$$I_{\theta\theta|\lambda} = I_{\theta\theta} - I_{\theta\lambda} I_{\lambda\lambda}^{-1} I_{\lambda\theta}$$

$$= \frac{N}{1000} \begin{bmatrix} 2.578307813 & 2.059083644 & .7424617186 & -.4609215771 \\ & 3.272247710 & 3.088625464 & 2.416067210 \\ & & 4.127085433 & 4.308617632 \\ & & & 5.198593163 \end{bmatrix}$$

From a location parameter family $(\sigma = 1)$ or scale parameter family $(\mu = 0)$, $I_{\theta\lambda}$ is a column vector and $I_{\lambda\lambda}$ is a scalar. For example, the family of negative exponential distribution with probability density function,

$$f(x|\beta) = \beta e^{-\beta x} \qquad 0 < x < \infty,\ 0 < \beta < \infty$$

can be treated as a scale parameter family $(\mu = 0, \sigma = 1/\beta)$ and by the transformation,

Table 2.3. The $I^{\theta\theta}$ matrices for testing extreme value distribution with $K \leq 4$.

$K = 1$

$$\frac{1}{N}\left[387.8512856\right]$$

$K = 2$

$$\frac{1}{N}\begin{bmatrix} 779.6567427 & -490.6041943 \\ & 614.3162862 \end{bmatrix}$$

$K = 3$

$$\frac{1}{N}\begin{bmatrix} 10852.26625 & -16981.68805 & 10756.42426 \\ & 27613.85861 & -17610.63946 \\ & & 11486.66220 \end{bmatrix}$$

$K = 4$

$$\frac{1}{N}\begin{bmatrix} 19720.32418 & -54593.34281 & 66758.06618 & -28208.45334 \\ & 187134.3066 & -255127.6289 & 119639.1157 \\ & & 365165.1221 & -178135.9251 \\ & & & 89728.42152 \end{bmatrix}$$

$$Y = \text{Log } X, \qquad \mu = \text{Log } \beta$$

it can be considered as the extreme location parameter family with $\sigma = 1$. Thus consider the single parameter negative exponential distribution, with $K = 4$, then

$$I_{\theta\theta|\lambda} = N \begin{bmatrix} \frac{1}{48} & \frac{1}{36} & \frac{9}{320} & \frac{2}{75} \\ & \frac{16}{405} & \frac{1}{24} & \frac{64}{1575} \\ & & \frac{81}{1792} & \frac{9}{200} \\ & & & \frac{256}{5625} \end{bmatrix}$$

Table 2.4. The $I^{\theta\theta}$ matrices for testing negative exponentiality with $K \leq 4$.

$K = 1$	$\frac{1}{N}[48]$
$K = 2$	$\frac{1}{N}\begin{bmatrix} 768 & -540 \\ & 405 \end{bmatrix}$
$K = 3$	$\frac{1}{N}\begin{bmatrix} 480 & -8100 & 4480 \\ & 14580 & -8400 \\ & & 4977.\dot{7} \end{bmatrix}$
$K = 4$	$\frac{1}{N}\begin{bmatrix} 19200 & -56700 & 71680 & -31500 \\ & 178605 & -235200 & 106312.5 \\ & & 318577.\dot{7} & -147000 \\ & & & 68906.25 \end{bmatrix}$

2.2. Linear Models

Now we generalize smooth tests to the linear model. Consider the following null hypothesis

$$H_0 : Y \sim F_0\left(\frac{y-\mu(\underset{\sim}{x})}{\sigma}\right) \tag{2.20}$$

with $\mu(\underset{\sim}{x}) = \beta_0 + \beta_1 x_1 + \ldots + \beta_{p-1} x_{p-1} \equiv \underset{\sim}{\beta}'\underset{\sim}{x}$ and smooth alternatives,

$$H_a : Y \sim G\left(\frac{y-\mu(\underset{\sim}{x})}{\sigma} \middle| \underset{\sim}{\theta}\right) \tag{2.21}$$

In this generalized case the observations, $y_1, y_2, \ldots, y_N$, are statistically independent, but not identically distributed. Further, F and G are cdf's for the conditional distribution of $\{Y_j - \mu(x_j)\}/\sigma$, given $\underset{\sim}{X}_j = \underset{\sim}{x}_j$. Let $\underset{\sim}{\lambda} = \left(\begin{smallmatrix}\underset{\sim}{\beta} \\ \sigma\end{smallmatrix}\right)$, and $F\ (y|\underset{\sim}{\lambda}, \underset{\sim}{x})$ and $G(y|\underset{\sim}{\theta}, \underset{\sim}{\lambda}, \underset{\sim}{x})$ denote the cdf's in (2.20) and (2.21), respectively. Then the pdf corresponding to $G(y|\underset{\sim}{\theta}, \underset{\sim}{\lambda}, \underset{\sim}{x})$

$$g(y|\underset{\sim}{\theta}, \underset{\sim}{\lambda}, \underset{\sim}{x}) = e^{\Sigma_{i=1}^{k} \theta_i F^i (y|\underset{\sim}{\lambda}, \underset{\sim}{x}) - R(\underset{\sim}{\theta})}\ f\ (y|\underset{\sim}{\lambda}, \underset{\sim}{x})$$

where

$$R(\underset{\sim}{\theta}) = \text{Log}\left[\int_{-\infty}^{\infty} e^{\Sigma_{i=1}^{k} \theta_i F^i (y|\underset{\sim}{\lambda}, \underset{\sim}{x})}\ f\ (y|\underset{\sim}{\lambda}, \underset{\sim}{x})dy\right]$$

The log-likelihood function, $\ell \equiv \ell(\underset{\sim}{\theta}, \underset{\sim}{\lambda})$, is

$$\ell = \sum_{j=1}^{N} \text{Log}[g(y_j|\underset{\sim}{\lambda}, \underset{\sim}{x})]$$

$$= \sum_{j=1}^{N} \left\{ \sum_{i=1}^{k} \theta_i F^i(y_j|\underset{\sim}{\lambda}, \underset{\sim}{x}) - R(\underset{\sim}{\theta}) + \text{Log}[f\ (y_j|\underset{\sim}{\lambda}, \underset{\sim}{x})] \right\} \qquad (2.22)$$

Then the i^{th} components of the scores $U_{\underset{\sim}{\theta}}(i)$ and $U_{\underset{\sim}{\lambda}}(i)$ are, for $i = 1, \ldots, k$,

$$U_{\underset{\sim}{\theta}}(i) = \sum_{j=1}^{N} F^i(y_j|\hat{\underset{\sim}{\lambda}}, \underset{\sim}{x}) - \frac{N}{i+1},$$

and, for $i = 1, \ldots, p+1$,

$$U_{\underset{\sim}{\lambda}}(i) = 0, \qquad \text{respectively.}$$

Let $x_0 \equiv x_p \equiv 1$. Then for $i = 1, \ldots, k$,

$$I_{\theta\lambda}(i, j) = \begin{cases} \overline{x_{j-1}}\ I^{\#}_{\theta\lambda}(i, 1), & j = 1, \ldots, p \\ I^{\#}_{\theta\lambda}(i, 2), & j = p+1\ . \end{cases} \qquad (2.23)$$

And

$$I_{\lambda\lambda}(j, i) = I_{\lambda\lambda}(i, j) = \begin{cases} \overline{x_{i-1}x_{j-1}}\ I^{\#}_{\lambda\lambda}(1, 1), & 1 \le i \le j \le p \\ \overline{x_{i-1}}\ I^{\#}_{\lambda\lambda}(1, 2), & 1 \le i < j = p+1 \\ I^{\#}_{\lambda\lambda}(2, 2), & i = j = p+1 \end{cases} \qquad (2.24)$$

where $\overline{x_\ell}$ and $\overline{x_\ell x_m}$ denote means of x_ℓ and $x_\ell x_m$, respectively, and the superscripts # denote the matrices defined in (2. 16) or (2. 17). The $I^{\theta\theta}$ matrix is obtained from either (2. 13) or (2. 14) with (2.8), (2.23), and (2.24). The results given in the following theorems were also proved independently by Pierce and Kopecky (1978).

Theorem 2. For a general linear model, the matrix $I^{\theta\theta}$ is independent of β and σ.

Proof. The matrices $\sigma I_{\theta\lambda}$ and $\sigma^2 I_{\lambda\lambda}$ are independent of β and σ. This fact follows from equations (2. 16), (2. 17), (2.23), and (2.24). Hence, from relations (2. 13) and (2. 18), the matrix $I^{\theta\theta}$ is independent of β and σ.

Theorem 3. For any given p (≥ 1), the matrix $I^{\theta\theta}$, corresponding to the GST of the general linear model (2.20) is independent of x.

Proof. Let p be an arbitrary positive integer (≥ 1), and let the number of nuisance parameters be denoted by the argument $p+1$ in matrices such as $I_{\theta\lambda}(p+1)$ and $I_{\lambda\lambda}(p+1)$. Define $\sigma I_{\theta\lambda}(2) = (C, D)$ and

$$\sigma^2 I_{\lambda\lambda}(2) = \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix},$$

where $C = (c_1, \ldots, c_k)'$, $D = (d_1, \ldots, d_k)'$, and $\frac{1}{N} I(2) = I_{\theta\lambda}(2)\, I_{\lambda\lambda}^{-1}(2)\, I_{\lambda\theta}(2)$.

i) $P = 2$, $[\lambda = (\beta_0, \beta_1, \sigma)']$. From equations (2.23) and (2.24), we may define $\sigma I_{\theta\lambda}(3)$ and $\sigma^2 I_{\lambda\lambda}(3)$ as

$$\sigma I_{\theta\lambda}(3) = (C, \overline{x}_1 C, D)$$

and

$$\sigma^2 I_{\lambda\lambda}(3) = \begin{pmatrix} a_{11} & \overline{x}_1 a_{11} & a_{12} \\ \overline{x}_1 a_{11} & \overline{x_1^2} a_{11} & \overline{x}_1 a_{12} \\ a_{12} & \overline{x}_1 a_{12} & a_{22} \end{pmatrix}.$$

So let

$$\frac{1}{N} I(3) = I_{\theta\lambda}(3)\, I_{\lambda\lambda}(3)^{-1}\, I_{\lambda\theta}(3)$$

$$= (C, \overline{x}_1 C, D) \begin{pmatrix} a_{11} & \overline{x}_1 a_{11} & a_{12} \\ \overline{x}_1 a_{11} & \overline{x_1^2} a_{11} & \overline{x}_1 a_{12} \\ a_{12} & \overline{x}_1 a_{12} & a_{22} \end{pmatrix}^{-1} \begin{pmatrix} C' \\ \overline{x}_1 C' \\ D' \end{pmatrix}$$

$$= (C, D, \overline{x}_1 C) \begin{pmatrix} a_{11} & a_{12} & \overline{x}_1 a_{11} \\ a_{12} & a_{22} & \overline{x}_1 a_{12} \\ \overline{x}_1 a_{11} & \overline{x}_1 a_{12} & \overline{x_1^2} a_{11} \end{pmatrix}^{-1} \begin{pmatrix} C' \\ D' \\ \overline{x}_1 C' \end{pmatrix} =$$

$$= (\sigma I_{\theta\lambda}(2), \overline{x}_1 C) \begin{pmatrix} \sigma^2 I_{\lambda\lambda}(2) & B \\ B' & \overline{x_1^2} a_{11} \end{pmatrix}^{-1} \begin{pmatrix} \sigma I_{\lambda\theta}(2) \\ \overline{x}_1 C' \end{pmatrix}$$

$$= (\sigma I_{\theta\lambda}(2), \overline{x}_1 C) \begin{pmatrix} [\sigma^2 I_{\lambda\lambda}(2)]^{-1} + FE^{-1}F' & -FE^{-1} \\ -E^{-1}F' & E^{-1} \end{pmatrix} \begin{pmatrix} \sigma I_{\lambda\theta}(2) \\ \overline{x}_1 C' \end{pmatrix}$$

$$= I_{\theta\lambda}(2)\, I_{\lambda\lambda}^{-1}(2)\, I_{\lambda\theta}(2) + [\sigma I_{\theta\lambda}(2)\, F - \overline{x}_1 C] E^{-1} [\lambda I_{\theta\lambda}(2) F - \overline{x}_1 C]'$$

$$= \frac{1}{N} I(2), \qquad (2.25)$$

where $B = \overline{x}_1 (a_{11} \ \ a_{12})'$, $E = \overline{x_1^2} a_{11} - B'[\sigma^2 I_{\lambda\lambda}(2)]^{-1} B$ and $F = [\sigma^2 I_{\lambda\lambda}(2)]^{-1} B$. The last term in (2.25) follows:

$$\sigma I_{\theta\lambda}(2)\, F - \overline{x}_1 C = (C, D)\, [\sigma^2 I_{\lambda\lambda}(2)]^{-1}\, \overline{x}_1 \begin{pmatrix} a_{11} \\ a_{12} \end{pmatrix} - \overline{x}_1 C$$

$$= (C, D) \begin{pmatrix} a_{22} & -a_{12} \\ -a_{12} & a_{11} \end{pmatrix} \frac{\overline{x}_1}{a_{11} a_{22} - (a_{12})^2} \begin{pmatrix} a_{11} \\ a_{12} \end{pmatrix} - \overline{x}_1 C$$

$$= (C, D) \binom{1}{0} \overline{x}_1 - \overline{x}_1 C$$

$$= 0$$

ii) Show the theorem is true for $p = \ell + 1$ assuming it is true when $p = \ell \ (\geq 2)$.

$$\frac{1}{N} I(\ell+1) \equiv I_{\theta\lambda}(\ell+1)\, I_{\lambda\lambda}^{-1}(\ell+1)\, I_{\lambda\theta}(\ell+1)$$

$$= (C, \overline{x}_1 C, \ldots, \overline{x}_{\ell-1} C, D) \begin{bmatrix} a_{11} & \overline{x}_1 a_{11} & \cdots & \overline{x}_{\ell-1} a_{11} & a_{12} \\ & \overline{x_1^2} a_{11} & \cdots & \overline{x_1 x_{\ell-1}} a_{11} & \overline{x}_1 a_{12} \\ & & & \vdots & \vdots \\ & & & \overline{x_{\ell-1}^2} a_{11} & \overline{x}_{\ell-1} a_{12} \\ & & & & a_{22} \end{bmatrix}^{-1}$$

$$\times \begin{pmatrix} C' \\ \overline{x}_1 C' \\ \vdots \\ \overline{x}_{\ell-1} C' \\ D' \end{pmatrix}$$

Let $A = \sigma^2 I_{\lambda\lambda}(\ell+1)$, $Q = A^{-1} = \{q(i, j)\}$, and $A_{i,j}$ = the cofactor of $A(i, j)$ in the determinant of A. Hence,

$$|A| = \sum_{s=1}^{\ell+1} (-1)^{s+1} A(s, 1)\, A_{s,1}$$

$$= \sum_{s=1}^{\ell+1} (-1)^{s+1} a_{11} \left(\frac{a_{12}}{a_{11}} \right)^{\delta_{s,\ell+1}} t_{s,1} A_{s,1}$$

where

$$t_{i,j} = \overline{x_{i-1} x_{j-1}} = \frac{1}{N} \sum_{s=1}^{N} x_{i-1,s} x_{j-1,s} \quad \text{for} \quad i, j = 1, \ldots, \ell+1$$

and $x_{0,s} = x_{\ell,s} = 1$ for $s = 1, \ldots, N$. Since, for $i, j = 1, \ldots, \ell+1$,

$$q(i,j) = (-1)^{i+j} \frac{A_{i,j}}{|A|},$$

it follows that

$$q(1, \ell+1) = q(\ell+1, 1) = -\frac{a_{12}}{a_{11}a_{22} - (a_{12})^2}$$

$$q(i, \ell+1) = q(\ell+1, i) = 0, \quad i = 2, \ldots, \ell$$

and

$$q(\ell+1, \ell+1) = \frac{a_{11}}{a_{11}a_{22} - (a_{12})^2}$$

Denote $\sigma H = I_{\theta\lambda}(\ell+1)\, I_{\lambda\lambda}(\ell+1)^{-1}$ and $H_{\cdot,j}$ = the j^{th} column of H, $j = 1, \ldots, \ell+1$. Then

$$H_{\cdot,1} = [C(1, \overline{x}_1, \ldots, \overline{x}_{\ell-1}), D] \begin{pmatrix} q(1,1) \\ \vdots \\ q(\ell+1, 1) \end{pmatrix}$$

$$= C(t_{1,1}, \ldots, t_{\ell,1}) \begin{bmatrix} A_{1,1} \\ -A_{2,1} \\ \vdots \\ (-1)^{\ell+1} A_{\ell,1} \end{bmatrix} \frac{1}{|A|}$$

$$+ D\,(-1)^{\ell+2} A_{\ell+1,1} \frac{1}{|A|} =$$

$$= \frac{C}{a_{11}\,|A|}\left[\sum_{s=1}^{\ell} (-1)^{s+1}\, a_{11} t_{s,1}\, A_{1,1} + (-1)^{\ell+2}\, a_{12} t_{\ell+1,1} A_{\ell+1,1} - (-1)^{\ell+2}\, a_{12} t_{\ell+1,1}\, A_{\ell+1,1}\right] + \frac{D\, A_{\ell+1,1}}{|A|}$$

$$= \frac{C}{a_{11}} - \frac{C}{a_{11}}\, a_{12}\, \frac{a_{12}}{a_{11}a_{22}-(a_{12})^2} - \frac{a_{12}\, D}{a_{11}a_{22}-(a_{12})^2}$$

$$= \frac{a_{22}C - a_{12}D}{a_{11}a_{22}-(a_{12})^2}\,.$$

For $j = 2, \ldots, \ell$,

$$H_{\cdot,j} = [C(1, \overline{x}_1, \ldots, \overline{x}_{\ell-1}), D] \begin{bmatrix} (-1)^{j+1}\, A_{1,j} \\ \vdots \\ (-1)^{j+\ell+1}\, A_{\ell+1,j} \end{bmatrix} \frac{1}{|A|}$$

$$= C \sum_{s=1}^{\ell} (-1)^{j+s}\, t_{s,1}\, A_{s,j}\, \frac{1}{|A|}$$

$$= 0, \qquad \text{since} \quad j \neq 1.$$

And

$$H_{\cdot,\ell+1} = [C(1,\overline{x}_1,\ldots,\overline{x}_{\ell-1}), D] \begin{bmatrix} q(1,\ell+1) \\ \vdots \\ q_{(\ell+1,\ell+1)} \end{bmatrix}$$

$$= C\left[-\frac{a_{12}}{a_{11}a_{22}-(a_{12})^2}\right] + D\frac{a_{11}}{a_{11}a_{22}-(a_{12})^2}$$

$$= \frac{-a_{12}C+a_{11}D}{a_{11}a_{22}-(a_{12})^2}$$

Therefore,

$$\frac{1}{N} I(\ell+1) = H\, I_{\lambda\theta}(\ell+1)$$

$$= \left(\frac{a_{22}C-a_{12}D}{a_{11}a_{22}-(a_{12})^2}, 0, \ldots, 0, \frac{-a_{12}C+a_{11}D}{a_{11}a_{22}-(a_{12})^2}\right) \begin{pmatrix} C' \\ \overline{x}_1 C' \\ \vdots \\ \overline{x}_{\ell-1} C' \\ D' \end{pmatrix}$$

$$= \frac{a_{22}CC'-a_{12}DC'-a_{12}CD'+a_{11}DD'}{a_{11}a_{22}-(a_{12})^2}$$

$$= (C, D)\begin{pmatrix} a_{22} & -a_{12} \\ -a_{12} & a_{11} \end{pmatrix}\begin{pmatrix} C' \\ D' \end{pmatrix} \frac{1}{a_{11}a_{22}-(a_{12})^2}$$

$$= \frac{1}{N} I(2)$$

Therefore, for any positive integer p (≥ 2), $I(p) = I(2)$. Apply equations (2.8), (2.12), and (2.14) with the above result. QED.

3. MONTE CARLO STUDY OF PERFORMANCE

A Monte Carlo study is used to compare the power performance of the generalized Neyman smooth (GST) tests with several other tests for the following scale and location parameter distributions: (1) normal and (2) extreme-value. Maximum likelihood estimates for the scale and location parameters are used in each case. All tests described below are invariant under linear transforms of the data. Samples of size $n = 20$ and 50 are employed.

3.1. Description of the Test Statistics

Two of the test statistics are developed in the next two sections. A complete listing of all 12 tests statistics considered is given in the third section.

3.1.1. Generalized Chi-Square ($G\chi^2$)

It is well known (Chernoff and Lehmann, 1954) that the Pearson chi-square goodness of fit statistic (see 3.8) with unknown parameters estimated by maximum likelihood from ungrouped data does not have a limiting χ^2-distribution but instead is asymptotically distributed as a linear function of chi-square variables. Several authors have considered modifying the Pearson statistic so as to have a limiting χ^2 distribution. Nikulin (1973) considered scale and location parameter families, Rao and Robson (1974) studied members within the exponential family, and Moore (1977) developed the modification in more generality.

For scale and location parameter families a different derivation from that of Nikulin is given here. The generalized chi-square statistic (3.4) agrees with that of Nikulin (1973). Consider the following family of density functions which include the hypothesized densities with $\theta_1 = \ldots = \theta_k = 0$:

$$g_\theta(y|\mu,\sigma) = \exp\left[\sum_{i=1}^{k} \theta_i \varphi_i\{F_0(\tfrac{y-\mu}{\sigma})\} - R(\underset{\sim}{\theta})\right] \frac{1}{\sigma} f_0(\tfrac{y-\mu}{\sigma})$$

where

$$\varphi_i(u) = \begin{cases} 1 & \text{if } (i-1)/K < u \leq i/K \\ 0 & \text{otherwise,} \end{cases}$$

$R(\underset{\sim}{\theta})$ is the normalizing constant and K is the number of intervals. The log-likelihood function, $\ell \equiv \ell(\underset{\sim}{\theta}, \mu, \sigma)$, is

$$\ell = \sum_{j=1}^{N} \sum_{i=1}^{k} \theta_i \varphi_i[F_0(\frac{y_j-\mu}{\sigma})] - N\, R(\underset{\sim}{\theta}) - N \operatorname{Log} \sigma + \sum_{j=1}^{N} \operatorname{Log} f_0(\frac{y_j-\mu}{\sigma}) .$$

Random intervals $[T_{i-1}, T_i)$ are chosen such that the fitted null distribution has equal probable classes for $i = 1, \ldots, K$ that is, $T_0 = -\infty$, $T_i = \hat{\mu} + \hat{\sigma} F_0^{-1}(i/K)$, $i = 1, \ldots, k-1$, and $T_K = \infty$. Let

P_i, $i = 1, \ldots, K$, be the estimated probability of an observation belonging to the i^{th} interval $[T_{i-1}, T_i)$. The generalized chi-square statistic is based on the scores.

$$\frac{\partial \ell}{\partial \theta_i}\Big|_{\underset{\sim}{\theta}=\underset{\sim}{0},\hat{\mu},\hat{\sigma}} = \sum_{j=1}^{N} \varphi_i[F_0(\frac{y_j-\hat{\mu}}{\hat{\sigma}})] - N\frac{\partial R(\underset{\sim}{\theta})}{\partial \theta_i}\Big|_{\underset{\sim}{\theta}=\underset{\sim}{0}}$$

$$\equiv O_i - N P_i$$

where $\underset{\sim}{O} = (O_1, \ldots, O_K)'$ is the vector of observed class frequencies in the corresponding intervals $[T_{i-1}, T_i)$, $i = 1, \ldots, K$. The scores corresponding to the nuisance parameters μ and σ are equal to zero when evaluated at the maximum likelihood estimates:

$$\frac{\partial \ell}{\partial \mu}\Big|_{\underset{\sim}{\theta}=\underset{\sim}{0},\hat{\mu},\hat{\sigma}} = -\frac{1}{\hat{\sigma}}\sum_{j=1}^{N}\frac{f_0'(\frac{y_j-\hat{\mu}}{\hat{\sigma}})}{f_0(\frac{y_j-\hat{\mu}}{\hat{\sigma}})}$$

$$\frac{\partial \ell}{\partial \sigma}\Big|_{\underset{\sim}{\theta}=\underset{\sim}{0},\hat{\mu},\hat{\sigma}} = -\frac{1}{\hat{\sigma}}\sum_{j=1}^{N}\frac{(\frac{y_j-\hat{\mu}}{\hat{\sigma}})f_0'(\frac{y_j-\hat{\mu}}{\hat{\sigma}})}{f_0(\frac{y_j-\hat{\mu}}{\hat{\sigma}})}$$

The elements of the partitioned information matrix,

$$I = \begin{pmatrix} I_{\theta\theta} & I_{\theta\lambda} \\ I_{\lambda\theta} & I_{\lambda\lambda} \end{pmatrix}$$

are evaluated as follows:

For $i, j = 1, \ldots, K$,

$$I_{\theta\theta}(i,j) = E\left(-\frac{\partial^2 \ell}{\partial\theta_i \partial\theta_j}\right)_{\underset{\sim}{\theta}=\underset{\sim}{0}}$$

$$= N \frac{\left[\int_{-\infty}^{\infty} \varphi_i[F_0(z)]\varphi_j[F_0(z)]f_0(z)dz - \int_{-\infty}^{\infty} \varphi_i[F_0(z)]f_0(z)]dz \int_{-\infty}^{\infty} \varphi_j[F_0(z)]f_0(z)dz\right]}{\left(\int_{-\infty}^{\infty} f_0(z)dz\right)^2}$$

$$= N \begin{cases} P_i(1-P_i) & \text{if} \quad i = j \\ -P_iP_j & \text{if} \quad i \neq j \end{cases}$$

$$= N \begin{cases} \frac{1}{K}(1-\frac{1}{K}) & \text{if} \quad i = j \\ -\frac{1}{K^2} & \text{if} \quad i \neq j \end{cases} \quad . \tag{3.1}$$

For $i = 1, \ldots, K$,

$$I_{\theta\lambda}(i,1) = E\left[\frac{\partial \ell}{\partial\theta_i}\frac{\partial \ell}{\partial\mu}\right]_{\underset{\sim}{\theta}=\underset{\sim}{0}}$$

$$= -\frac{1}{\sigma}E\left[\left(\sum_{j=1}^{N} \varphi_i[F_0(\frac{Y_j-\mu}{\sigma})] - \frac{N}{K}\right)\sum_{\ell=1}^{N}\frac{f_0'(\frac{Y_\ell-\mu}{\sigma})}{f_0(\frac{Y_\ell-\mu}{\sigma})}\right] =$$

$$= -E\left[\sum_{j=1}^{N}\sum_{\ell=1}^{N}\{\varphi_i[F_0(Z_j)]-\frac{1}{K}\}\ \frac{f_0'(Z_\ell)}{f_0(Z_\ell)}\right]$$

$$= -NE\left[\{\varphi_i[F_0(Z)]-\frac{1}{K}\}\ \frac{f_0'(Z)}{f_0(Z)}\right]-N(N-1)E[\varphi_i[F_0(Z)]-\frac{1}{K}]\ E[\frac{f_0'(Z)}{f_0(Z)}]$$

$$= -N\left[\int_{F_0^{-1}(\frac{i-1}{K})}^{F_0^{-1}(\frac{i}{K})} f_0'(z)dz - \frac{1}{K}\int_{-\infty}^{\infty} f_0'(z)dz\right]$$

$$= -N\left[f_0\{F_0^{-1}(\frac{i}{K})\} - f_0\{F_0^{-1}(\frac{i-1}{K})\}\right]. \qquad (3.2)$$

For $i = 1, \ldots, K$,

$$I_{\theta\lambda}(i,2) = E[\ \frac{\partial \ell}{\partial \theta_i}\ \frac{\partial \ell}{\partial \sigma}\ ]_{\underset{\sim}{\theta}=\underset{\sim}{0}}$$

$$= -\frac{1}{\sigma}E\left[\sum_{j=1}^{N}\left(\varphi_i[F_0(\frac{Y_j-\mu}{\sigma})-\frac{1}{K}\right)\sum_{\ell=1}^{N}\left(1+(\frac{Y_j-\mu}{\sigma})\ \frac{f_0'(\frac{Y_j-\mu}{\sigma})}{f_0(\frac{Y_j-\mu}{\sigma})}\right)\right]$$

$$= -E\left[\sum_{j=1}^{N}\sum_{\ell=1}^{N}\left(\varphi_i[F_0(Z_j)]-\frac{1}{K}\right)\left(1+Z_\ell\ \frac{f_0'(Z_\ell)}{f_0(Z_\ell)}\right)\right]$$

$$= -NE\left[\{\varphi_i[F_0(Z)]-\frac{1}{K}\}\{1+Z\ \frac{f_0'(Z)}{f_0(Z)}\}\right]$$

$$= -N\left[\int_{F_0^{-1}(\frac{i-1}{K})}^{F_0^{-1}(\frac{i}{K})} \{f_0(z)+z\ f_0'(z)\}dz - \frac{1}{K}\int_{-\infty}^{\infty}\ \{f_0(z)+z\ f_0'(z)\}dz\ \right] =$$

$$= -N\left[\left\{z\, f_0(z)\right\}_{F_0^{-1}(\frac{i-1}{K})}^{F_0^{-1}(\frac{i}{K})} - \frac{1}{K}\{z\, f_0(z)\}_{-\infty}^{\infty}\right]$$

$$= -N\left[F_0^{-1}(\tfrac{i}{K})\, f_0\{F_0^{-1}(\tfrac{i}{K})\} - F_0^{-1}(\tfrac{i-1}{K})\, f_0\{F_0^{-1}(\tfrac{i-1}{K})\}\right]. \qquad (3.3)$$

The matrix $I_{\lambda\lambda}$ is the same as (2.10). Denote the partitioned matrix in the inverse of the full information matrix corresponding to $I_{\theta\theta}$ by $I^{\theta\theta}$. The matrix $I^{\theta\theta}$ may be evaluated from the identity.

$$I^{\theta\theta} = I_{\theta\theta}^{-} + I_{\theta\theta}^{-}\, I_{\theta\lambda}\, (I_{\lambda\lambda} - I_{\lambda\theta}\, I_{\theta\theta}^{-}\, I_{\theta\lambda})^{-1}\, I_{\lambda\theta}\, I_{\theta\theta}^{-}$$

when a generalized inverse of the matrix $I_{\theta\theta}$, which is of rank K-1, is given by the diagonal matrix,

$$I_{\theta\theta}^{-} = \frac{K}{N} I_K .$$

Then the generalized χ^2 statistic may be evaluated as

$$G\chi^2 = (\underset{\sim}{O}-\underset{\sim}{E})' I^{\theta\theta} (\underset{\sim}{O}-\underset{\sim}{E}) \qquad (3.4)$$

where $\underset{\sim}{E} = (E_1, \ldots, E_K)'$ with $E_1 = \ldots = E_K = \frac{N}{K}$ are the expected frequency vector under the null distribution in the corresponding interval $[T_{i-1}, T_i)$, $i = 1, \ldots, K$. As Nikulin (1973) shows, under certain regularity conditions which are satisfied by normal, logistic, and extreme value distributions, $G\chi^2$ has a limiting χ^2

distribution with K-1 df. Table 3.1 gives the evaluation of $I^{\theta\theta}$ for normal, extreme value, and logistic distributions with $K = 5$ and 10 equally probable intervals, respectively.

3.1.2. Locally Most Powerful (LMP) Test

Consider the transformed generalized gamma density functions

$$g_\theta(y;\mu,\sigma) = \frac{1}{\Gamma(\theta+1)\sigma} e^{(\theta+1)(\frac{y-\mu}{\sigma}) - e^{\frac{y-\mu}{\sigma}}} \tag{3.5}$$

which include the extreme value densities when $\theta = 0$. A test of $H_0: \theta = 0$ vs. $H_a: \theta \neq 0$ based on the scores statistic is developed below. Since

$$\bar{F}_0(\frac{y-\mu}{\sigma}) = e^{-e^{\frac{y-\mu}{\sigma}}}$$

equation (3.5) can be written as

$$g_\theta(y;\mu,\sigma) = e^{\theta \operatorname{Log}[-\operatorname{Log}\{\bar{F}_0(\frac{y-\mu}{\sigma})\}] - \operatorname{Log}[\Gamma(\theta+1)]} f_0(\frac{y-\mu}{\sigma})$$

The loglikelihood function, $\ell \equiv \ell(\theta,\mu,\sigma)$ is

Table 3.1 Numerical values for $I^{\theta\theta}$.

a) Normal with K = 5

$\frac{1}{N}$

25.3208	6.46276	-1.08428	-7.98856	-17.7106
	8.19162	.633865	-2.29960	-7.98856
		5.90083	.633865	-1.08428
			8.19162	6.46276
				25.3208

b) EV with K = 5

$\frac{1}{N}$

30.3847	8.62661	-1.41474	-10.8059	-21.7898
	8.55649	.168470	-3.38135	-8.97018
		5.75376	.904801	-.412283
			9.73575	8.54678
				27.6255

c) Logistic with K = 5

$\frac{1}{N}$

55.0721	22.8370	-1.81810	-25.1630	-45.9279
	17.6528	1.02046	-11.3472	-25.1630
		6.59529	1.02046	-1.81810
			17.6528	22.8370
				55.0721

d) Normal with K = 10

$\frac{1}{N}$

93.7395	45.1892	27.0248	13.3771	1.70158	-9.10163	-19.7619	-31.0539	-44.3794	-66.7354
	36.6767	17.1894	9.71078	3.03903	-3.39147	-10.0149	-17.3813	-26.6381	-44.3794
		21.6856	7.15176	2.96156	-1.20815	-5.63888	-10.7310	-17.3813	-31.0539
			14.8478	2.57891	.199727	-2.45037	-5.63888	-10.0149	-19.7619
				11.9980	1.22241	.199727	-1.20815	-3.39147	-9.10163
					11.9980	2.57891	2.96156	3.03903	1.70158
						14.8478	7.15176	9.71078	13.3771
							21.6856	17.1894	27.0248
								36.6767	45.1892
									93.7395

Table 3.1 (Continued)

e) EV with K = 10

$\frac{1}{N}$ ×

116.151	61.3047	38.4436	20.2480	4.00210	-11.4830	-27.0162	-43.4395	-62.1240	-86.0864
	46.2525	23.2808	12.8287	3.38088	-5.74689	-15.0514	-25.1020	-36.9369	-54.2103
		25.2956	8.77746	2.81063	-3.03252	-9.08279	-15.7507	-23.8463	-36.8958
			15.3820	2.19557	-1.00548	-4.41532	-8.30559	-13.2677	-22.4376
				11.5005	.683657	-.323162	-1.65628	-3.67823	-8.91579
					12.1660	3.49788	4.68385	5.61507	4.62146
						17.2342	11.0462	15.1227	18.9880
							27.7768	25.4410	35.3063
								47.6802	55.9942
									103.636

f) Logistic with K = 10

$\frac{1}{N}$ ×

248.74	183.71	135.94	82.735	29.094	-24.906	-79.265	-134.06	-194.29	-247.26
	150.42	106.17	65.157	23.603	-18.397	-60.843	-103.83	-153.58	-194.29
		84.791	44.615	14.535	-15.465	-45.385	-75.209	-103.83	-134.06
			36.293	8.1463	-9.8537	-27.707	-45.385	-60.843	-79.265
				11.970	-4.0304	-9.8537	-15.465	-18.397	-24.906
					11.970	8.1463	14.535	23.603	29.094
						36.293	44.615	65.157	82.735
							84.791	106.17	135.94
								150.42	183.71
									248.74

$$\ell = \theta \sum_{i=1}^{N} \text{Log}[-\text{Log}\{\overline{F}_0(\frac{y_i-\mu}{\sigma})\}]-N \text{ Log } \Gamma(\theta+1)+\sum_{i=1}^{N} \text{Log}[f_0(\frac{y_i-\mu}{\sigma})]$$

$$U_\theta = \frac{\partial \ell}{\partial \theta}\Big|_{\substack{\theta=0\\ \mu=\hat{\mu}\\ \sigma=\hat{\sigma}}} = \sum_{i=1}^{N} (\frac{y_i-\hat{\mu}}{\hat{\sigma}}) - N\,\psi(1)$$

$$= \sum_{i=1}^{N} (\frac{y_i-\hat{\mu}}{\hat{\sigma}}) + N\,\gamma \qquad (3.6)$$

where $\psi(\cdot)$ is a digamma function and γ is the Euler's constant and $\hat{\mu}$ and $\hat{\sigma}$ denote the ML estimators for parameters of the extreme value distribution. The elements of the partitioned information matrix are evaluated as follows:

$$I_{\theta\theta} = E(-\frac{\partial^2 \ell}{\partial \theta^2})_{\theta=0} = N\,\psi'(1)$$

$$\approx 1.644934N$$

$$I_{\theta\lambda}(1,1) = E(-\frac{\partial^2 \ell}{\partial\theta\partial\mu})_{\theta=0} = \frac{N}{\sigma}$$

$$I_{\theta\lambda}(1,2) = E(-\frac{\partial^2 \ell}{\partial\theta\partial\sigma})_{\theta=0} = \frac{N}{\sigma} E(\frac{Y-\mu}{\sigma})$$

$$= -\frac{N}{\sigma}\,\gamma$$

From (2.19)

$$I_{\lambda\lambda} = \frac{N}{\sigma^2}\begin{bmatrix} 1 & 1-\gamma \\ & (1-\gamma)^2+\psi'(1)\end{bmatrix}$$

$$I^{-1}_{\theta\theta|\lambda} = (I_{\theta\theta}-I_{\theta\lambda}\, I^{-1}_{\lambda\lambda}\, I_{\lambda\theta})^{-1}$$

$$= \left[N\,\psi'(1)-N(1-\gamma)\begin{pmatrix} 1 & 1-\gamma \\ & (1-\gamma)^2+\psi'(1)\end{pmatrix}^{-1}\begin{pmatrix}1\\ -\gamma\end{pmatrix}\right]^{-1}$$

$$= N^{-1}[\psi'(1)-\frac{1+\psi'(1)}{\psi'(1)}]^{-1}$$

$$= \frac{1}{N}\,[\frac{\psi'(1)}{[\psi'(1)]^2-\psi'(1)-1}]$$

The LMP test statistic is defined by

$$\mathrm{LMP} = (\frac{\partial \ell}{\partial\theta}, \frac{\partial \ell}{\partial\mu}, \frac{\partial \ell}{\partial\sigma})\begin{pmatrix} I_{\theta\theta} & I_{\theta\lambda} \\ & I_{\lambda\lambda}\end{pmatrix}^{-1}\begin{bmatrix}\frac{\partial \ell}{\partial\theta}\\ \frac{\partial \ell}{\partial\mu}\\ \frac{\partial \ell}{\partial\sigma}\end{bmatrix}_{\theta=0,\hat{\mu},\hat{\sigma}}$$

$$= \frac{\partial \ell}{\partial\theta}\, I^{-1}_{\theta\theta|\lambda}\, \frac{\partial \ell}{\partial\theta}\Big|_{\theta=0,\hat{\mu},\hat{\sigma}}$$

$$= I^{-1}_{\theta\theta|\lambda}\, U^2_\theta \qquad (3.7)$$

3.1.3. List of All Test Statistics

The next five tests are considered for normal and extreme value distributions.

1) Generalized Neyman Smooth Tests (W_k)

$$W_k = U_k' I^{\theta\theta} U_k$$

where U_k and $I^{\theta\theta}$ are defined in (2.6) and (2.13), respectively.

2) Barton's Smooth Tests (BS_k)

$$BS_k = U_k' I_{\theta\theta}^{-1} U_k$$

where U_k and $I_{\theta\theta}$ are defined in (2.6) and (2.8), respectively. BS_1 and W_1 are identical.

3) Classical $\chi^2 (C\chi^2)$

$$C\chi^2 = \frac{(\underset{\sim}{O}-\underset{\sim}{E})'(\underset{\sim}{O}-\underset{\sim}{E})}{E_1}$$

$$= \frac{K}{N} \sum_{i=1}^{K} O_i^2 - N \qquad (3.8)$$

where $\underset{\sim}{O}$, $\underset{\sim}{E}$, E_1 are described in Section 3.1.1 and K is number of intervals. As mentioned before, even for fixed intervals, the χ^2 statistics do not, in general, have limiting χ^2 distributions, when MLE formed from the raw data are used (Kendall and Stuart, 1973, p. 443). Moreover, when fully efficient estimators are used the limiting distribution of (3.8) under H_0 is known to be bounded between distributions of $\chi^2(K-3)$ and $\chi^2(K-1)$ variables.

4) Generalized $\chi^2(G\chi^2)$

From (3.4),

$$G\chi^2 = (\underset{\sim}{Q}-\underset{\sim}{E})'I^{\theta\theta}(\underset{\sim}{Q}-\underset{\sim}{E})$$

5) Kolmogorov-Smirnov

The power performance of Kolmogorov-Smirnov (KS) tests for composite hypotheses have been investigated by several authors, including Green and Hegazy (1976), Lilliefors (1967), Stephens (1970, 1974), Durbin (1975), Locke (1976), and Pettitt and Stephens (1977). Durbin and Lilliefors apply KS statistics for testing exponentiality and normality, respectively.

$$KS = \max_{1 \leq i \leq N} \left[\left| F_N(y_i) - F_0\left(\frac{y_i - \hat{\mu}}{\hat{\sigma}}\right) \right|, \left| F_N(y_{i-1}) - F_0\left(\frac{y_i - \hat{\mu}}{\hat{\sigma}}\right) \right| \right]$$

In the KS statistic, and other statistics below, $y_1 \leq y_2 \leq \cdots \leq y_N$ denote the order statistics and $F_N(y)$ the sample distribution function, i.e., $F_N(y_i) = \frac{i}{N}$ for $i = 1, 2, \ldots, N$.

The next three statistics are used only for testing normality.

6) Shapiro-Wilk (SW)

Shapiro-Wilk (1965) proposed the following test statistic for normality, which is obtained by dividing the square of an appropriate linear combination of order statistic by the usual symmetric estimate of variance. This ratio is both scale and location invariant and hence the statistic is appropriate for a test of the composite hypothesis of normality.

$$SW = \frac{(\Sigma_{i=1}^{N} a_i y_i)^2}{\Sigma_{i=1}^{N} (y_i - \bar{y})^2}$$

where the coefficients

$$(a_1, \ldots, a_N)' = \frac{m'V^{-1}}{(m'V^{-1}V^{-1}m)^{1/2}}$$

are such that $\Sigma_{i=1}^{N} a_i y_i$ is the minimum variance unbiased linear combination order statistic estimator for σ. The coefficients a_i are tabled by Shapiro-Wilk (1965) for $N = 2(1)50$. Here m and V denote respectively the mean vector and variance-covariance matrix of the standard normal order statistics. Shapiro and Francia (1972) consider a large sample approximation of the coefficients a_i. Several studies comparing the power of various tests for normality have included the univariate Skewness (SK) and Kurtosis (KU) test statistics. Such studies include Bowman (1973), D'Agostino and Pearson (1973), and Kopecky (1977).

$$7)\ SK = \frac{\sqrt{N}\,\Sigma_{i=1}^{N}(y_i-\overline{y})^3}{[\Sigma_{i=1}^{N}(y_i-\overline{y})^2]^{3/2}}$$

and

$$8)\ KU = \frac{N\,\Sigma_{i=1}^{N}(y_i-\overline{y})^4}{[\Sigma_{i=1}^{N}(y_i-\overline{y})^2]^2}$$

Malkovich and Afifi (1973) generalized the SK and KU to tests of multivariate normality. Pearson (1965) gave tables for the 5 and 1% points for the SK statistic with $N \geq 25$ and the KU statistic with $N \geq 50$.

The final four statistics are used only for testing extreme-value distributions. For the description of the next three statistics denote:

$$v_i = E(\frac{Y_i-\mu}{\sigma})$$

$$v_i \approx K_i = \text{Log}[-\text{Log}(1-\frac{i}{N+1})]$$

9) Mann (MN)

Mann and Fertig (1975) proposed the following test statistic. For even N, that is, $N = 2R$, $R = 1, 2, \ldots,$

$$MN = \frac{N\,A_2}{(N-2)A_1} \quad .$$

where

$$A_1 = \sum_{i=R+1}^{N} \frac{y_i - y_{i-1}}{E(\frac{Y_i - Y_{i-1}}{\sigma})} \quad \text{and} \quad A_2 = \sum_{i=2}^{R} \frac{y_i - y_{i-1}}{E(\frac{Y_i - Y_{i-1}}{\sigma})}$$

We use the approximation

$$E(\frac{Y_i - \mu}{\sigma}) \approx K_i = \mathrm{Log}[-\mathrm{Log}(1 - \frac{i}{N+1})]$$

10) Smith-Bain (SB)

Smith and Bain (1976) proposed a correlation type test statistics of y_i and the expectations v_i. We use the modification where v_i is approximated by K_i as defined above

$$SB = 1 - \frac{[\Sigma_{i=1}^{N} y_i(K_i - \overline{K})]^2}{\Sigma_{i=1}^{N} (K_i - \overline{K})^2 \, \Sigma_{i=1}^{N} (y_i - \overline{y})^2}.$$

11) Cramér-von Mises

Green and Hegazy (1976) and Stephens (1970) discussed the Cramér-von Mises (CM) test for composite hypothesis.

$$CM = \frac{1}{12N} + \sum_{i=1}^{N} [F_0(\frac{y_i - \hat{\mu}}{\hat{\sigma}}) - \frac{2i-1}{2N}]^2$$

12) LMP

From (3.7),

$$LMP = I^{-1}_{\theta\theta|\lambda} U^2_\theta .$$

3.2. Alternative Distributions

The following alternatives are considered for testing normality

1) Weibull (β), β = .5 and 2

$$f(x) \propto x^{\beta-1} \exp(-x^\beta) \qquad x \geq 0$$

2) Extreme value

$$f(x) \propto \exp(x-\exp(x)) \qquad -\infty < x < \infty$$

3) Gamma (β), β = .5, 1, 2, and 5

$$f(x) \propto x^{\beta-1} \exp(-x) \qquad x \geq 0$$

4) Lognormal (β), β = 1

$$f(x) \propto x^{-1} \exp(-\frac{1}{2} (\frac{\text{Log } x}{\beta})^2) \qquad x > 0$$

5) Uniform

$$f(x) = 1 \qquad 0 \leq x \leq 1$$

6) Logistic

$$f(x) \propto \frac{\exp(x)}{[1+\exp(x)]^2} \qquad -\infty < x < \infty$$

7) Double $\chi^2(\beta)$, $\beta = -.5$ and 1

$$f(x) \propto \exp(-\frac{1}{2}|x|^{2/(\beta+1)}) \qquad -\infty < x < \infty$$

8) Cauchy

$$f(x) \propto \frac{1}{1+x^2} \qquad -\infty < x < \infty$$

The logistic, Weibull, and gamma distributions and the next four distributions are considered as alternatives for goodness of fit tests for an extreme value distribution.

9) Normal

$$f(x) \propto \exp[-\frac{1}{2}x^2] \qquad -\infty < x < \infty$$

10) Transformed gamma (θ), $\theta = 5$

$$f(x) \propto \exp[\theta x - \exp(x)] \qquad -\infty < x < \infty$$

11) Weibull mixture (α, β), $(\alpha, \beta) = (1, 3)$

$$f(x) \propto \alpha \exp(-x) + 3\beta x^2 \exp(-x^3) \qquad 0 < x < \infty$$

12) Exponential mixture (α, β), $(\alpha, \beta) = (1, 3)$

$$f(x) \propto \alpha \exp(-x) + 3\beta \exp(-3x) \qquad 0 < x < \infty$$

3.3. Random Number Generation

Random numbers were generated on the Cyber 70/73 computer at Oregon State University from the IMSL Library 3 (1975) to produce the independent sets of N pseudo-random numbers. The generators were used as follows: GGNOF for an individual Normal (0, 1) pseudo-random number, GGTMAJ for a N-vector of independent Gamma (A, B) variables which are distributed as $(\Gamma(A)\ B^{A})^{-1}\ x^{A-1}\ \exp(-x/B)$ where A, B, and x are all positive, and GGUB for a N-vector of independent Uniform (0, 1) variates which are converted by the probability transformation to extreme value; Weibull (β), β = .5 and 2; Logistic; Cauchy; Double $\chi^2(\beta)$, β = -.5, 1, 2, and 5; and mixtures of Weibull (1, 3) or Exponential (1, 3) variates. The IMSL subroutine MDNOR evaluates the cumulative standard normal distribution function and the subroutine VSORTA (A, N) rearranges a vector A of size N, by ascending order. In generating samples of size 20 and 50 from a distribution, different seed numbers are used. The null distributions of the statistics are obtained by empirical random sampling.

3.4. Size of Tests

Estimates for the true test sizes (significance levels) corresponding to nominal sizes .05 and .10 are given in Table 3.2 for the three

Table 3.2. Empirical size (%) of the generalized Neyman smooth (W_k), classical chi square ($C\chi^2$), and generalized chi square ($G\chi^2$) tests based on 2000 simulations. The number of class intervals (NC) used for the $C\chi^2$ and $G\chi^2$ tests are: NC = 5, 10, 10 corresponding to sample sizes N = 20, 50, 100, respectively. Critical values corresponding to 1-α quantiles of chi square distributions with k, NC-3, and NC-1 degrees of freedom are used for the W_k, $C\chi^2$, and $G\chi^2$ tests, respectively.

	N	W_1	W_2	W_3	W_4	$C\chi^2$	$G\chi^2$
α = .05							
Normal	20	4.40	3.40*	3.80*	3.55*	7.85*	4.50
	50	4.65	4.45	4.35	4.30	5.05	5.25
Extreme value	20	4.65	3.30*	4.00*	3.70*	7.30*	4.95
	50	4.95	4.25	4.25	4.25	5.10	4.20
	100	4.95	4.75	5.35	4.65	5.05	4.55
Logistic	20	4.00*	4.90	4.70	4.10	7.40*	4.65
	50	4.75	4.40	5.15	5.15	5.10	4.85
α = .10							
Normal	20	9.85	6.85*	7.20*	7.15*	13.90*	8.15*
	50	9.70	9.40	8.95	8.00*	10.10	9.75
Extreme value	20	9.40	7.40*	7.80*	6.65*	14.40*	10.15
	50	9.40	9.25	8.75	7.65*	10.40	9.15
	100	9.65	9.50	9.50	9.15	10.55	10.20
Logistic	20	8.75	9.60	9.15	8.45*	11.50*	10.20
	50	9.45	9.55	8.95	9.55	11.00	9.40

*The estimates are not within ±2 standard errors of the nominal values.

test statistics: generalized Neyman smooth (W_k), classical chi-square $(C\chi^2)$, and generalized chi square $(G\chi^2)$. All estimates are based on 2000 simulations. The estimates which are not within ± 2 standard errors of the nominal values are indicated in Table 3.2. The nominal level of W_1 is quite accurate. The accuracy of the nominal levels for W_k tends to decrease as k increases. The nominal levels of $G\chi^2$ are accurate, too. For $\alpha = .05$, all estimates are within ± 2 standard errors of the nominal values for the samples of size $N \geq 50$.

3.5. Power of Tests

We estimate the power of several goodness of fit tests for normal and extreme value distributions. Over several alternative distributions, the empirical critical values listed in Table 3.3 are used. The alternative distributions for testing normality are divided into two groups: those which are symmetrically distributed and those which are asymmetrically distributed. For symmetric distribution, heavy or light tailed[1] pdf relatives to the normal are separately considered. Empirical powers are given in Tables 3.4 and 3.5 for testing normality against symmetrically and asymmetrically distributed

[1]For any symmetric pdf's, $h(x)$ and $g(x)$, $h(x)$ is said to have a heavier tail than $g(x)$ if there exists a constant $c > 0$ such that $h(d) > g(d)$ for any $d \in (c, \infty)$.

Table 3.3. Empirical critical* value based on 2000 simulations from normal and extreme value distributions.

		Normal						Extreme Value			
	N =	20		50			N =	20		50	
	α =	.05	.10	.05	.10		α =	.05	.10	.05	.10
W_1		3.668	2.683	3.643	2.560	W_1		3.727	2.612	3.824	2.610
W_2		5.160	4.091	5.601	4.234	W_2		5.291	4.086	5.764	4.435
W_3		7.122	5.683	7.588	5.820	W_3		7.255	5.610	7.489	5.976
W_4		8.632	6.853	9.331	7.268	W_4		8.839	6.963	9.246	7.222
BS_1		.165	.121	.164	.115	BS_1		.115	.081	.118	.081
BS_2		.868	.625	.946	.685	BS_2		1.025	.753	1.140	.851
BS_3		4.207	3.188	4.153	2.976	BS_3		4.589	3.217	4.333	2.929
BS_4		6.408	4.873	6.125	4.729	BS_4		6.393	5.144	6.215	4.602
$G\chi^2$		9.325	7.174	17.0	14.62	$G\chi^2$		9.475	7.818	16.53	14.30
$C\chi^{2**}$		6.5	5.5	14.4	12.4	$C\chi^{2**}$		6.5	5.5	14.4	12.4
		(.3876)	(.6364)	(.95)	(.8571)			(.3514)	(.6415)	(.8333)	(.7097)
KS		.193	.177	.124	.115	KS		.192	.176	.123	.112
SK	L	-.968	-.802	-.650	-.536	MN	L	.432	.516	.608	.662
	U	.926	.759	.621	.532		U	2.303	2.075	1.704	1.552
KU	L	1.745	1.836	2.050	2.135	SB		.130	.100	.079	.059
	U	4.666	4.141	4.419	3.951	LMP		3.193	2.256	3.689	2.513
SW		.906	.920	.948	.956	CM		.127	.103	.120	.101

*For two sided tests, lower (L) and upper (U) critical values are included.

**Rejection with certainty if $C\chi^2$ is greater than critical value and rejection with probability $(\cdot)$ if $C\chi^2$ is equal to critical value.

alternatives, respectively, and in Table 3.6 testing extreme value distribution. Overall the GST's W_1 and W_2 are found to perform well in comparison with the other various tests.

3.5.1. Normal Case

a) Symmetrically distributed alternatives.

In Table 3.4 the W_2 test tends to have the highest power in heavier tailed alternatives (Logistic, Laplace, Double $\chi^2(2)$, and Cauchy). The KU and BS_2 tests tend to have the highest power among the various tests in lighter tailed distributed alternatives. Overall the powers for BS_2 and W_2 are similar. The W_2 test has the highest power among the W_1-W_4 tests over all cases.

b) Asymmetrically distributed alternatives.

In Table 3.5 either W_1 or SW has the highest power of various tests among all asymmetrically distributed alternative cases. Moreover, the powers of W_1 and SW are similar. The W_1 test has a higher power than SW for the extreme value, Weibull (2), and gamma (β), β = 2 and 5, distribution alternatives. The W_1 test has a higher power for lognormal (0, 1/2) distribution at the 10% level, and SW has the higher power in all the remaining alternatives.

Table 3.4 Empirical power * (%) of several tests of normality for symmetrically distributed alternatives. Empirical critical values based on 2000 simulations are used for each test.

	sample size	W_1	W_2	W_3	W_4	$G\chi^2$	SW	SK	KU	KS	$C\chi^2$	BS_2	BS_3	BS_4
$\alpha = .05$														
Uniform (0, 1)	20	2.2	16.3	11.4	11.0	12.3	19.5	0.0	33.1	10.2	8.2	28.8	8.6	18.4
	50	4.3	70.5	53.8	40.7	35.4	87.3	0.2	86.9	31.3	19.8	80.7	21.0	48.8
Double χ^2 (-.5)	20	2.0	6.5	5.5	5.6	8.5	6.7	0.7	12.7	7.2	6.9	11.8	6.9	10.2
	50	2.9	20.9	13.4	10.0	10.2	24.5	0.9	32.9	12.6	9.1	31.6	8.1	18.8
Logistic	20	12.1	15.1	14.2	14.0	10.0	10.2	14.9	11.8	9.3	7.7	11.6	8.4	7.8
	50	13.2	23.6	23.6	22.2	12.1	14.4	19.5	19.3	12.2	8.0	22.7	9.9	10.8
Laplace	20	21.2	32.5	30.6	28.9	20.4	27.7	26.0	24.3	21.3	15.1	28.2	17.1	18.9
	50	22.0	60.1	57.1	54.8	38.7	40.6	35.8	47.1	42.2	27.1	61.9	30.5	46.6
Double χ^2 (2)	20	33.8	58.2	56.5	56.5	45.4	52.8	40.6	45.0	46.4	40.1	54.6	36.0	47.8
	50	41.5	91.6	89.9	90.2	83.8	81.9	61.8	82.4	85.6	75.1	92.8	73.8	87.9
Cauchy	20	70.1	89.7	88.8	89.3	47.9	88.1	77.3	83.8	84.5	80.6	88.9	79.2	85.6

Table 3.4 (Continued)

	sample size	W_1	W_2	W_3	W_4	$G\chi^2$	SW	SK	KU	KS	$C\chi^2$	BS_2	BS_3	BS_4
$\alpha = .10$														
Uniform (0, 1)	20	5.8	30.6	22.3	20.1	25.0	34.8	1.7	47.0	17.8	16.9	48.6	16.2	30.1
	50	8.5	85.5	75.3	62.4	53.0	96.3	0.7	91.9	44.4	29.8	89.5	37.9	61.5
Double χ^2 (-.5)	20	5.8	14.9	12.6	12.3	16.2	14.7	2.0	21.0	14.8	13.0	24.5	12.1	18.6
	50	6.6	38.3	27.8	20.8	18.7	40.2	1.7	46.0	22.4	15.6	49.2	18.0	31.0
Logistic	20	18.4	20.8	20.2	20.8	16.2	16.0	21.0	18.4	15.2	13.2	19.7	13.1	13.6
	50	19.7	32.4	32.3	32.0	20.1	21.7	27.3	33.2	20.3	14.4	31.3	18.2	19.5
Laplace	20	28.7	39.2	37.8	38.8	29.8	35.1	33.8	34.6	30.9	25.2	36.2	25.4	29.4
	50	30.9	68.7	66.1	66.5	50.0	50.1	43.2	61.2	53.9	38.1	69.7	44.6	57.8
Double χ^2 (2)	20	42.2	67.7	64.1	67.7	57.5	60.7	50.3	56.4	58.4	51.1	62.1	46.0	59.4
	50	50.3	95.1	93.8	94.4	88.3	87.9	67.0	90.0	90.4	81.2	95.2	86.0	93.5
Cauchy	20	74.0	92.0	91.2	92.0	61.3	90.1	80.5	88.2	89.2	85.3	91.0	83.8	89.2

*2000 simulations are used for samples from the double χ^2 (-.5) and Laplace distributions and 1000 simulations are used for the other cases.

Table 3.5 Empirical power* (%) of several tests of normality for asymmetrically distributed alternatives. Empirical critical values based on 2000 simulations are used for each test.

	sample size	W_1	W_2	W_3	W_4	$G\chi^2$	SW	SK	KU	KS	$C\chi^2$	BS_2	BS_3	BS_4
$\alpha = .05$														
Extreme value (.5)	20	33.8	30.4	27.1	25.0	16.2	32.1	31.3	19.1	20.6	11.5	16.4	25.8	19.5
	50	72.7	63.9	56.0	56.6	37.9	67.5	68.9	31.8	43.2	25.9	30.9	55.1	46.1
Weibull (2)	20	16.0	13.1	11.1	11.3	10.1	16.0	15.2	8.8	10.2	6.5	7.7	13.5	10.0
	50	43.1	35.9	26.7	26.9	16.7	43.8	40.6	12.5	23.6	13.1	11.5	30.1	25.9
Weibull (.5)	20	98.9	98.8	98.3	98.9	97.6	99.8	97.9	77.7	98.4	95.9	84.8	98.9	99.1
Lognormal (0, .5)	20	77.1	72.7	65.8	68.5	50.7	79.2	73.5	44.5	57.6	39.7	43.8	70.6	60.9
	50	99.5	98.7	97.8	97.5	94.0	99.6	98.7	75.3	92.8	91.6	80.9	97.0	96.9
Lognormal (0, 1)	20	92.0	90.1	85.9	87.5	74.8	94.1	87.1	57.9	78.7	65.3	61.3	87.8	84.3
Gamma (β), $\beta = .5$	20	95.0	93.3	89.2	93.4	83.6	98.8	89.7	55.5	87.3	72.8	60.1	92.3	91.5
$\beta = 1$	20	76.8	71.8	65.3	70.6	50.8	83.7	69.5	34.4	56.7	36.6	35.7	69.9	64.2
$\beta = 2$	20	54.9	47.1	40.6	42.6	28.3	53.6	48.5	25.6	42.4	18.7	24.4	45.7	35.6
$\beta = 5$	20	26.0	22.4	20.5	21.2	14.9	25.5	25.8	14.2	16.0	10.0	12.4	20.4	14.9

Table 3.5 (Continued)

	sample size	W_1	W_2	W_3	W_4	GX^2	SW	SK	KU	KS	CX^2	BS_2	BS_3	BS_4
$\alpha = .10$														
Extreme value (.5)	20	44.8	40.2	35.7	37.5	26.3	41.4	42.6	26.2	30.9	22.8	23.9	36.2	30.3
	50	82.6	75.8	68.9	68.5	48.2	77.5	81.0	41.7	55.2	34.6	41.2	67.1	58.7
Weibull (2)	20	25.6	22.3	17.9	19.5	17.5	25.4	23.8	15.1	18.4	11.9	15.2	20.8	19.0
	50	58.2	51.6	42.2	40.5	26.8	58.2	51.8	21.3	34.1	23.2	21.7	41.4	37.7
Weibull (.5)	20	99.7	99.5	99.2	99.4	98.4	100.0	98.5	84.5	99.3	96.4	90.1	99.4	99.3
Lognormal (0, .5)	20	84.9	80.4	75.6	76.9	63.3	84.8	81.6	52.0	68.9	47.7	53.1	79.0	72.3
	50	99.9	99.4	99.1	98.7	96.5	99.8	99.5	82.8	96.1	95.2	88.2	98.5	97.6
Lognormal (0, 1)	20	95.0	93.1	90.6	92.1	83.3	96.2	92.5	65.2	85.9	70.9	68.0	91.8	90.7
Gamma (β), $\beta = .5$	20	97.1	97.3	93.3	95.7	89.8	99.4	94.6	65.5	93.2	76.4	69.2	96.2	96.0
$\beta = 1$	20	84.1	81.6	75.3	80.3	65.8	89.6	80.0	42.8	69.4	43.9	46.9	78.9	76.2
$\beta = 2$	20	65.8	60.7	51.3	53.7	41.3	65.2	61.2	34.7	53.8	27.7	33.0	56.5	50.2
$\beta = 5$	20	37.6	31.6	29.0	29.9	22.8	35.5	37.1	22.4	26.0	17.1	20.0	30.2	25.3

*2000 simulations are used for sample from the Weibull (2) distribution and 1000 simulations are used for the other cases.

Table 3.6 Empirical powers* (%) of several tests of the Extreme value distribution for several alternative distributions. Empirical critical values based on 2000 simulations are used for each test.

	sample size	W_1	W_2	W_3	W_4	$G\chi^2$	SB	MN	LMP	CM	KS	$C\chi^2$	BS_2	BS_3	BS_4
$\alpha = .05$															
Normal	20	13.5	20.7	18.4	16.3	13.5	11.0	19.8	24.1	20.3	16.5	14.4	21.8	23.2	16.6
	50	51.4	56.8	49.6	49.6	34.7	24.8	54.1	64.5	53.2	38.5	27.5	47.6	57.2	49.3
Logistic	20	21.0	28.0	24.7	23.6	19.5	16.5	21.4	29.7	27.9	22.5	17.9	30.7	25.9	22.6
	50	50.2	71.7	65.6	65.6	38.4	36.5	60.2	76.9	65.6	51.6	40.1	70.5	69.1	61.4
Gamma (β),	20	10.8	9.1	8.6	7.6	5.4	8.0	9.2	10.5	5.7	5.6	5.2	5.3	6.4	6.4
$\beta = .5$	50	17.6	14.6	14.0	12.0	9.5	12.3	13.4	17.3	11.8	10.9	6.1	7.1	11.7	9.7
$\beta = 2$	20	3.8	5.4	4.8	4.5	4.7	3.8	6.0	5.6	6.1	5.4	6.6	7.3	5.8	5.1
	50	6.1	8.7	8.4	7.3	7.5	2.9	8.4	9.0	10.6	9.1	6.7	9.7	9.2	9.1
$\beta = 5$	20	5.2	7.9	7.3	6.8	6.7	3.6	7.8	9.0	8.4	6.9	7.3	10.3	9.2	7.2
	50	16.7	17.0	15.5	14.9	11.8	3.9	19.6	23.7	18.0	13.8	10.1	17.0	19.6	14.8
Weibull	20	3.2	3.7	3.3	4.2	7.1	1.5	5.8	3.8	5.1	5.2	7.1	7.7	5.1	6.5
	50	9.3	9.6	6.4	7.7	7.8	1.3	12.2	12.2	10.4	9.4	7.8	9.8	10.0	9.6
Weibull mixture	20	14 6	24.8	23.8	23.9	14.5	19.0	11.9	20.1	18.6	15.4	12.3	20.9	13.4	17.9
	50	28.2	53.4	53.4	51.2	32.9	30.6	33.3	33.4	47.0	36.4	22.9	48.5	26.5	39.8
Exponential mixture	20	3.2	4.5	4.8	5.0	5.3	4.2	4.4	4.9	4.9	5.1	5.2	6.3	4.3	4.8
	50	7.3	5.8	6.4	6.3	6.0	4.3	7.1	7.9	8.8	7.3	7.4	7.3	9.4	7.5
Transf. Gamma (5)	100	40.6	39.5	30.8	35.2	21.6	5.1	37.5	51.3	33.3	25.1	16.2	28.9	37.2	30.5

Table 3.6 (Continued)

	sample size	W_1	W_2	W_3	W_4	$G\chi^2$	SB	MN	LMP	CM	KS	$C\chi^2$	BS_2	BS_3	BS_4
$\alpha = .10$															
Normal	20	31.7	30.9	28.6	27.0	22.2	22.9	33.9	36.7	31.5	25.2	22.4	28.9	36.5	26.7
	50	69.5	69.1	61.6	62.3	47.6	44.8	66.8	76.3	63.0	53.0	39.2	54.8	70.5	63.5
Logistic	20	38.4	36.9	35.3	31.1	26.0	28.6	33.8	39.4	37.8	32.4	28.4	36.1	39.5	29.7
	50	62.9	79.6	74.8	75.3	60.6	53.5	71.4	85.2	72.9	65.9	50.5	76.0	79.6	72.4
Gamma (β), $\beta = .5$	20	18.3	15.3	14.8	13.6	11.2	15.1	15.7	18.2	12.2	11.5	10.5	11.2	13.3	11.4
	50	25.9	23.2	21.9	19.3	15.8	18.9	22.9	25.6	18.5	17.4	11.3	12.4	20.1	18.6
$\beta = 2$	20	8.5	10.5	10.3	9.1	9.4	8.6	11.9	9.9	10.6	10.3	11.9	12.1	11.0	9.7
	50	15.0	14.2	14.0	16.2	14.3	6.6	14.8	17.6	17.0	15.6	12.6	15.0	18.1	16.8
$\beta = 5$	20	15.0	13.7	13.7	12.7	11.8	10.1	16.6	15.5	14.3	12.9	13.5	15.8	17.7	13.2
	50	32.4	27.8	22.4	23.7	20.3	11.3	30.0	36.8	26.7	24.1	17.4	23.6	33.0	25.2
Weibull	20	9.4	10.5	8.7	9.0	12.6	5.2	12.1	9.4	11.6	11.2	13.9	13.8	12.1	10.6
	50	22.9	19.4	14.2	14.6	14.7	4.8	21.8	20.1	16.9	17.1	13.4	15.0	21.4	20.4
Weibull mixture	20	24.0	34.9	32.7	33.1	22.7	32.9	18.0	29.9	28.8	25.1	23.0	28.8	23.8	26.4
	50	43.7	62.7	62.2	62.6	46.3	44.6	44.3	50.9	57.4	49.1	33.3	55.4	46.2	53.6
Exponential mixture	20	9.0	9.4	8.5	8.7	9.7	7.9	9.3	9.7	9.6	10.1	11.2	11.8	10.0	8.0
	50	15.7	13.7	11.2	11.9	13.1	7.9	13.9	15.5	14.3	13.2	12.0	12.0	17.4	14.8
Transf. Gamma (5)	100	57.0	53.1	44.4	44 9	30.8	19.2	50.8	64.5	43.4	37.4	25.8	37.8	50.4	43.0

*2000 simulations are used for samples from the Gamma (β), $\beta = .5$, 2, and 5 of size 20 and normal distributions and 1000 simulations are used for the other cases.

3.5.2. Extreme Value Case

in Table 3.6 either W_1, W_2, LMP, or BS_3 tends to have the highest power among the various tests. The W_1, W_2, and LMP tests have the highest powers, respectively for gamma (1/2), Weibull mixture, and normal distributions. Moreover, the W_2, LMP, and BS_3 tests have similar powers except for the Weibull mixture distribution where W_2 has substantially higher power.

4. APPLICATIONS

Evaluation of the generalized smooth tests, W_k, are illustrated in the following six examples.

a) Scale parameter family

Ex. 1. Negative exponential distribution

b) Scale-location parameter family

Ex. 2. Extreme value distribution

Ex. 5. Normal distribution

c) Linear regression model

Ex. 3. Negative exponential distribution

(Extreme value distribution with $\sigma = 1$)

Ex. 4. Extreme value distribution

Ex. 6. Normal distribution

The common procedure consists of three steps: evaluate the MLE $\hat{\lambda}$, compute the test statistic W_k, and compare W_k with the approximate critical value from a χ^2 table. The MLE $\hat{\lambda}$, for negative exponential and normal distributions are easily evaluated. For the extreme value distribution the MLE $\hat{\lambda}$ can be found by sequential approximation (Thomas, 1977). Data A and B are employed for examples 1-3 and 4-6, respectively.

Data A: The comparative susceptibility of the Dojo fish to the poison EI-43064 was determined by immersing each individual fish in 2 liters of an emulsion of the poison and measuring its survival time in minutes, where

x_1 = Log-concentration of EI-43064 in parts/million,

x_2 = Log-grams weight of the fish, and

y = Log(1000/minutes survival). N = 45. (Bliss, 1967, p. 352).

Data B: Effect of age of oat plants in days (x) upon their N_i/Fe ratio (y). N = 30 (Bliss, 1967, p. 31).

Ex. 1. Consider the negative exponential cdf,

$$F(y|\underset{\sim}{\lambda}) = F(\frac{y}{\sigma}) = 1-e^{-y/\sigma}$$

where $\underset{\sim}{\lambda} = \sigma$. Then the MLE

$$\hat{\underset{\sim}{\lambda}} = \frac{1}{N}\sum_{i=1}^{N} y_i = 1.288608.$$

From (2.6) and (2.15) with the $I^{\theta\theta}$ matrix in Table 2.4, the test statistics, W_k, $k = 1, \ldots, 4$, are evaluated as

$$W' = (35.4,\ 43.71,\ 84.88,\ 35.41).$$

Since $W_k > \chi^2_{.995}(k)$ for $k = 1, 2, 3, 4$, the null hypothesis in (2.3) is rejected at the significant level $\alpha = .005$ for any of the GST's with $k \leq 4$.

Ex. 2. Consider the extreme value cdf,

$$F(y|\underset{\sim}{\lambda}) = 1 - \exp[-\exp(\frac{y-\mu}{\sigma})]$$

where $\lambda = (\beta_0, \sigma)' = (\mu, \sigma)'$. The MLE $\hat{\underset{\sim}{\lambda}}$ is the solution of the following equations:

$$N\hat{\sigma} = \sum_{i=1}^{N} y_i(\exp(\frac{y_i - \hat{\mu}}{\hat{\sigma}}) - 1)$$

and

$$N = \sum_{i=1}^{N} \exp(\frac{y_i - \hat{\mu}}{\hat{\sigma}})$$

The solution to these equations is found by iteration (Thomas, 1977), to give

$$\hat{\underset{\sim}{\lambda}}' = (\hat{\mu}, \hat{\sigma}) = (1.353053, .1508056)$$

Using (2.6) and (2.15) with the $I^{\theta\theta}$ matrix in Table 2.3 the test statistics, W_k, $k = 1, 2, 3, 4$, are

$$W' = (5.75, 10.01, 10.01, 10.18)$$

The null hypothesis in (2.3) is rejected at $\alpha = .05$ for $k = 1, 2, 3, 4$.

Ex. 3. Consider the extreme value distribution cdf,

$$F(y|\underset{\sim}{x}, \underset{\sim}{\lambda}) = 1 - \exp[-\exp(y-\beta_0 - \cdots - \beta_M x_M)],$$

where $\underset{\sim}{\lambda} = \underset{\sim}{\beta} = (\beta_0, \beta_1, \ldots, \beta_M)'$. The equations,

$$\sum_{i=1}^{N} \exp[y_i - \hat{\beta}_0 - \cdots - \hat{\beta}_M x_{Mi}] = N$$

and

$$\sum_{i=1}^{N} x_{ji}[\exp(y_i - \hat{\beta}_0 - \cdots - \hat{\beta}_M x_{Mi}) - 1] = 0, \quad j = 1, \ldots, M,$$

yield

$$\hat{\underset{\sim}{\lambda}}' = (.4606657, .3746265) \quad \text{for} \quad M = 1,$$

and

$$\hat{\underset{\sim}{\lambda}}' = (.7153553, .374289, -.3988452) \quad \text{for} \quad M = 2.$$

The test statistics W_k, $k = 1, 2, 3, 4$, and

$$W' = (36.03, 44.38, 87.39, 88.35) \quad \text{for} \quad M = 1$$

and

$$W' = (36.47, 44.92, 90.21, 90.63) \quad \text{for} \quad M = 2.$$

The null hypothesis in (2.20) is rejected at $\alpha = .005$ when $k = 1, 2, 3, 4$ for both $M = 1$ and 2.

Ex. 4. Consider the extreme value regression model with

$$F(y|\underset{\sim}{x}, \underset{\sim}{\lambda}) = 1 - \exp[-\exp(\frac{y-\beta_0 - \cdots - \beta_M x_M}{\sigma})],$$

where $\lambda = \binom{\beta}{\sigma}$ and $x_q = x_1^q$ for $q = 1, \ldots, M$. The equations,

$$\sum_{i=1}^{N} (y_i - \hat{\beta}_0 - \cdots - \hat{\beta}_M x_{Mi}) \left[\exp(\frac{y_i - \hat{\beta}_0 - \cdots - \hat{\beta}_M x_{Mi}}{\hat{\sigma}}) - 1 \right] = 0$$

and

$$\sum_{i=1}^{N} x_{Mi} \left[\exp(\frac{y_i - \hat{\beta}_0 - \cdots - \hat{\beta}_M x_{Mi}}{\hat{\sigma}}) - 1 \right] = 0, \quad j = 1, \ldots, M$$

yield

$$\hat{\underset{\sim}{\lambda}}' = (.3463413, .4513125, .1152703) \quad \text{for} \quad M = 1$$

and

$$\hat{\underset{\sim}{\lambda}}' = (.723785, .3859209, -.3771623, .1017292) \quad \text{for} \quad M = 2.$$

The test statistics W_k, $k = 1, 2, 3, 4$, are

$$W' = (2.13, 2.20, 2.42, 2.91) \quad \text{for} \quad M = 1$$

and

$$W' = (2.92, 3.38, 3.41, 3.45) \quad \text{for} \quad M = 2.$$

Therefore, when $M = 1$, the null hypothesis in (2.20) is not rejected at $\alpha = .10$ for $k = 1, 2, 3, 4$, and when $M = 2$, the null hypothesis in (2.20) is not rejected at $\alpha = .10$ for $k = 2, 3, 4$, and for $k = 1$, the null hypothesis in (2.20) is not rejected at $\alpha = .05$, however, it is rejected at $\alpha = .10$.

Ex. 5. For the normal distribution model with $\underset{\sim}{\lambda} = \binom{\mu}{\sigma}$,

$$\hat{\underset{\sim}{\lambda}} = \begin{pmatrix} \overline{y} \\ \sum_{i=1}^{N} (y_i - \overline{y})^2 / N \end{pmatrix}$$

$$= \begin{pmatrix} 1.2347 \\ .396154291 \end{pmatrix}.$$

The test statistics W_k, $k = 1, 2, 3, 4$, are

$$W' = (.41, .52, .73, .99).$$

The null hypothesis in (2.3) is not rejected at $\alpha = .10$ for $k = 1, 2, 3, 4$.

Ex. 6. For the normal regression model with $\underset{\sim}{\lambda} = \binom{\underset{\sim}{\beta}}{\sigma}$,

$$\hat{\underset{\sim}{\lambda}}' = (1.0147, .0050757, .3855634) \quad \text{for} \quad M = 1$$

and

$$\hat{\underset{\sim}{\lambda}}' = (-.047414, -.073549, .00086627, .2355027) \quad \text{for} \quad M = 2.$$

And the test statistics W_k, $k = 1, 2, 3, 4$, are

$$W' = (.08, 2.68, 3.15, 3.64) \quad \text{for} \quad M = 1$$

and

$$W' = (1.11, 1.89, 1.98, 2.44) \quad \text{for} \quad M = 2$$

For both $M = 1$ and 2, the null hypothesis in (2.20) is not rejected at $\alpha = .10$.

BIBLIOGRAPHY

1. Anderson, T.W. and Darling, D.A. (1954). A test of goodness of fit. Journal of the American Statistical Association 49:765-769.

2. Abramowitz, M. and Stegun, I.A. (1965). Handbook of mathematical functions. National Bureau of Standards.

3. Bain, L.J. and Antle, C.E. (1967). Estimation of parameters in the Weibull distribution. Technometrics 9:621-627.

4. Bartlett, M.S. (1953). Approximate confidence intervals II more than one unknown parameter. Biometrika 40:306-317.

5. Barton, D.E. (1953). On Neyman's smooth test of goodness of fit test and its power with respect to a particular system of alternatives. Skand. Aktuariet. 36:24-63.

6. Barton, D.E. (1955). A form of Neyman's ψ_k^2 test of goodness of fit applicable to grouped and discrete data. Skan. Aktuariet. 38:1-16.

7. Barton, D.E. (1956). Neyman's ψ_k^2 test of goodness of fit when the null hypothesis is composite. Skand. Aktuariet. 39:216-245.

8. Bliss, C.I. (1967). Statistics in biology. McGraw Hill, New York.

9. Bowman, K.O. (1973). Power of the kurtosis statistic, b_2, in tests of departured from normality. Biometrika 60:623-628.

10. Bowman, K.O. and Shenton, L.R. (1973). Notes on the distribution of $\sqrt{b_1}$ in sampling from Pearson distributions. Biometrika 60:155-167.

11. Bowman, K.O. and Shenton, L.R. (1975). Omnibus test contours for departures from normality based on $\sqrt{b_1}$ and b_2. Biometrika 62:243-251.

12. Chapman, D.G. (1958). A comparative study of several one-sided goodness of fit tests. Annals of Mathematical Statistics 29:655-674.

13. Chen, E.H. (1971). The power of the Shapiro-Wilk W test for normality in samples from contaminated normal distribution. Journal of the American Statistical Association 66:760-762.

14. Chernoff, H. and Lehmann, E. L. (1954). The use of maximum likelihood estimates in χ^2 tests for goodness of fit. Annals of Mathematical Statistics 25:579-586.

15. Cohen, A.C., Jr. (1965). Maximum likelihood estimation in the Weibull distribution based on complete and on censored samples. Technometrics 7:579-588.

16. Cochran, W.G. (1952). The χ^2 test of goodness of fit. Annals of Mathematical Statistics 23:315-345.

17. Cox, D.R. and Hinkley, D.V. (1974). Theoretical statistics. Chapman and Hall, London.

18. Cramer, H. (1946). Mathematical methods of statistics. Princeton University Press, Princeton.

19. D'Agostino, R.B. (1971). An omnibus test of normality for moderate and large sample sizes. Biometrika 58:623-628.

20. D'Agostino, R.B. and Pearson, E.S. (1973). Tests for departure from normality empirical results for the distributions of b_2 and $\sqrt{b_1}$. Biometrika 60:613-622.

21. Dahiya, R.C. and Gurland, J. (1972a). Goodness of fit tests for the gamma and exponential distribution. Technometrics 14:791-801.

22. Dahiya, R.C. and Gurland, J. (1972b). Pearson chi-square test of fit with random intervals. Biometrika 59:147-153.

23. Dahiya, R.C. and Gurland, J. (1973). How many classes in the Pearson chi-square test? Journal of the American Statistical Association 68:707-712.

24. Darling, D.A. (1957). The Kolmogorov-Smirnov, Cramer-von Mises tests. Annals of Mathematical Statistics 28:823-838.

25. David, F.N. (1939). On Neyman's "smooth" test of goodness of fit. Biometrika 31:191-199.

26. Durbin, J. (1975). Kolmogorov-Smirnov tests when parameters are estimated with applications to tests of exponentiality and tests on spacing. Biometrika 62:5-22.

27. Finkelstein, J.M. and Schafer, R.E. (1971). Improved goodness of fit tests. Biometrika 58:641-645.

28. Green, J.R. and Hegazy, Y.A.S. (1975). Some new goodness of fit tests using order statistics. Applied Statistice 24:299-308.

29. Green, J.R. and Hegazy, Y.A.S. (1976). Powerful modified EDF goodness of fit tests. Journal of the American Statistical Association 71:204-209.

30. Hager, H.W. and Bain, L.J. (1970). Inferential procedures for the generalized gamma distribution. Journal of the American Statistical Association 65:1601-1609.

31. Hamdan, M.A. (1962). The powers of certain smooth tests of goodness of fit. Australian Journal of Statistics 4:25-40.

32. Hamdan, M.A. (1964). A smooth test of goodness of fit based on the Walsh functions. Australian Journal of Statistics 6:130-136.

33. Hensler, G.L., Mehrota, K.G., and Michalek, J.E. (1977). A goodness of fit test for multivariate normality. Communications in Statistics-Theory and Methods A6:33-41.

34. IMSL Library 3, Edition 5 (1975). Vol. I and II. Reference Manual. 6th Floor, GNB Building, 7500 Bellaire Boulevard, Houston.

35. Javitz, H.S. (1975). Generalized smooth tests of goodness of fit, independence, and equality of distributions. Unpublished thesis, University of California, Berkeley.

36. Johnson, N.L. and Kotz, S. (1970). Continuous univariate distributions -1 and -2. Houghton Mifflin Company, Boston.

37. Kendall, M.G. and Stuart, A. (1973). The advanced theory of statistics. Vol. 2 (3rd edition) 419-463.

38. Kopecky, K. (1977). Efficiency of smooth goodness of fit tests. Ph.D. thesis, Dept. of Statistics, Oregon State University.

39. Kowalski, C. J. (1970). The performance of some rough tests for bivariate normality before and after coordinate transformations to normality. Technometrics 12:517-544.

40. Lehmann, E. L. (1959). Testing statistical hypotheses. John Wiley and Sons, Inc., New York.

41. Lilliefors, H. W. (1967). On the Kolmogorov-Smirnov test for the normality with mean and variance unknown. Journal of the American Statistical Association 62:399-402.

42. Lilliefors, H. W. (1969). On the Kolmogorov-Smirnov test for the exponential distribution with mean unknown. Journal of the American Statistical Association 64:387-389.

43. Littell, R.C. and Rao, P.V. (1975). Estimation based on the Kolmogorov-Smirnov goodness of fit statistic. Technical report No. 98, Dept. of Statistics, University of Florida.

44. Locke, C. (1976). A test for the composite hypothesis that a population has a gamma distribution. Communications in Statistics-Theory Methods A5(4):351-364.

45. Malkovich, J.F. and Afifi, A.A. (1973). On tests for multivariate normality. Journal of the American Statistical Association 68:176-179.

46. Mann, N.R. (1968a). Results on statistical estimation and hypothesis testing with application to the Weibull and extreme-value distribution. Aerospace Research Laboratories Report ARL 68-0068.

47. Mann, N.R. (1968b). Point and interval estimation procedure for the two-parameter Weibull and extreme-value distribution. Technometrics 10:231-256.

48. Mann, N.R. and Fertig, K.W. (1975). A goodness of fit test for the two parameter vs three parameter Weibull; confidence bounds for threshold. Technometrics 17:237-245.

49. Mann, N.R., Fertig, K.W., and Scheuer, E.M. (1971). Confidence and tolerance bounds and a new goodness of fit test for the two parameter Weibull or extreme-value distributions. Aerospace Research Laboratories Report ARL 71-0077.

50. Mann, N.R., Fertig, K.W., and Scheuer, E.M. (1973). A new goodness of fit test for the two parameter Weibull or extreme-value distribution with unknown parameters. Communications in Statistics 2:383-400.

51. Mann, H.B. and Wald, A. (1942). On the choice of the number of class intervals in the application of the chi-square test. Annals of Mathematical Statistics 13:306-317.

52. Margolin, B.H. and Maurer, W. (1976). Tests of the Kolmogorov-Smirnov type for exponential data with unknown scale, and related problems. Biometrika 63:149-160.

53. Massey, F.J., Jr. (1951). The Kolmogorov-Smirnov test for goodness of fit. Journal of the American Statistical Association 46:68-78.

54. Moore, D.S. (1977). Generalized inverse, Wald's method, and the construction of chi-squared tests of fit. Journal of the American Statistical Association 72:131-137.

55. Moore, D.S. and Spruill, M.C. (1975). Unified large sample theory of general chi-square statistic for tests of fit. Annals of Statistics 3:599-616.

56. Neyman, J. (1937). Smooth test of goodness of fit. Skand. Aktuariet. 20:149-199.

57. Neyman, J. (1940b). Empirical comparison of the "smooth" test of goodness of fit with the Pearson's χ^2 test. Annals of Mathematical statistics 11:477-478.

58. Neyman, J. (1959). Optimal asymptotic tests of composite statistical hypotheses, in Probability and Statistics (U. Grenander Edition). Almquist and Wiksell, Stockholm.

59. Nikulin, M. (1973). Chi-square tests for continuous distributions with shift and scale parameters. Theory of Probability and its Application 18:559-568.

60. O'Reilly, F.J. and Quesenberry, C.P. (1973). The conditional probability integral transformation and applications to obtain composite chi-square goodness of fit tests. Annals of Statistics 1:74-82.

61. Pearson, K. (1900). On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. Philosophy Magazine Series 5, Vol. 50:157-172.

62. Pearson, E.S. (1938). The probability integral transformation for testing goodness of fit and combining independent test of significance. Biometrika 30:134-148.

63. Pearson, E.S. (1965). Tables for percentage points of $\sqrt{b_1}$ and b_2 in normal sample; a rounding off. Biometrika 52:282-285.

64. Pearson, E.S. and Hartley, H.O., editors (1972). Biometrika tables for statisticians. Vol. II, Cambridge University Press, Cambridge.

65. Pearson, E.S. and Stephens, M.A. (1962). The goodness of fit tests based on W_N^2 and U_N^2. Biometrika 49:397-402.

66. Pettitt, A.N. and Stephens, M.A. (1977). The Kolmogorov-Smirnov goodness of fit statistics with discrete and grouped data. Technometrics 19:205-210.

67. Pierce, D. and Kopecky, K. (1978). Testing goodness of fit for the distribution of errors in regression models. To be published.

68. Rao, C.R. (1973). Linear statistical inference and its applications. John Wiley and Sons, Inc., New York.

69. Rao, K.C. and Robson, D.S. (1974). A chi-square statistic for goodness of fit tests within the exponential family. Communications in Statistics 3(12):1139-1153.

70. Sarhan, A.E. and Greenberg, B.G. (1962). Contributions to order statistics. John Wiley and Sons, Inc., New York.

71. Sarkadi, K. (1975). The consistency of the Shaprio-Francia test. Biometrika 62:445-450.

72. Schafer, R.E., Finkelstein, J.M., and Collins, J. (1972). On a goodness of fit test for the exponential distribution with mean unknown. Biometrika 59:222-224.

73. Seshadri, V., Csörgö, M., and Stephens, M.A. (1969). Tests for the exponential distribution using Kolmogorov-type statistics. Journal of Royal Statistical Society B 31:499-509.

74. Shapiro, S.S. and Francia, R.S. (1972). An approximate analysis of variance test for normality. Journal of the American Statistical Association 67:215-216.

75. Shapiro, S.S. and Wilk, M.B. (1965). An analysis of variance test for normality (complete samples). Biometrika 52:591-611.

76. Shapiro, S.S. and Wilk, M.B. (1968). Approximations for the null distribution of the W statistic. Technometrics 10:861-866.

77. Shaprio, S.S., Wilk, M.B., and Chen, H.J. (1968). A comparative study of various tests for normality. Journal of the American Statistical Association 63:1343-1372.

78. Smith, R.M. and Bain, L.J. (1976). Correlation type goodness of fit statistics with censoring sampling. Communications in Statistics-Theory Methods A5(2):119-132.

79. Spruill, M.C. (1976). A comparison of chi-square goodness of fit tests based on approximate bahadur slope. Annals of Statistics 4:409-412.

80. Srinivasan, R. (1970). Approach to testing goodness of fit of incompletely specified distribution. Biometrika 57:605-611.

81. Stephens, M.A. (1969). Results from the relation between two statistics of the Kolmogorov-Smirnov type. Annals of Mathematical Statistics 40:1833-1837.

82. Stephens, M.A. (1970). Use of Kolmogorov-Smirnov, Cramer-von Mises and related statistics without extensive tables. Journal of Royal Statistical Society B 32:115-122.

83. Stephens, M.A. (1974). EDP statistics for goodness of fit and some comparisons. Journal of the American Statistical Association 69:730-737.

84. Stephens, M.A. (1976). Asymptotic results for goodness of fit statistics with unknown parameters. Annals of Statistics 4:357-369.

85. Teichrow, D. (1956). Tables of expected values of order statistics and products of order statistics for samples of size twenty and less from the normal distribution. Annals of Mathematical Statistics 27:410-426.

86. Thoman, D.R., Bain, L.J., and Antle, C.E. (1969). Inferences on the parameters of the Weibull distribution. Technometrics 11:445-460.

87. Thomas, D.R. (1973). Some tests of fit for the two parameter Weibull distributions. Unpublished. Oregon State University.

88. Thomas, D.R. (1977). Personal communications. Unpublished.

89. Thomas, D.R. and Pierce, D.A. (1977). Neyman's smooth goodness of fit test when the hypothesis is composite. Technical report No. 59, Dept. of Statistics, Oregon State University.

90. Vasicek, O. (1976). A test for normality based on sample entropy. Journal of Royal Statistical Society B. 38:54-59.

91. Wald, A. (1943). Tests of statistical hypotheses concerning several parameters when the number of observations is large. Transactions of the American Mathematical Society 54:426-482.

92. Watson, G.S. (1957). The chi-squared goodness of fit tests for normal distributions. Biometrika 44:336-348.

93. Watson, G.S. (1958). The chi-squared goodness of fit tests for continuous distributions. Journal of Royal Statistical Society B. 20:44-61.

94. Watson, G.S. (1959). Some recent results in chi-square goodness of fit tests. Biometrics 15:440-468.

95. Wilk, M.B., Gnanadesikan, R., and Huyett, M.J. (1962). Estimation of parameters of the gamma distribution using order statistics. Biometrika 49:525-545.

96. Williams, C.A. (1950). On the choice of the number and width of classes for the chi-square test of goodness of fit. Journal of the American Statistical Association 45:77-86.

97. Zacks, S. (1971). The theory of statistical inference. John Wiley and Sons, Inc., New York.