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: 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 $\chi^2$ test and a generalized $\chi^2$ 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, 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

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>2. CONSTRUCTION OF THE GENERALIZED SMOOTH TESTS</td>
<td>5</td>
</tr>
<tr>
<td>2.1. Univariate Models</td>
<td>5</td>
</tr>
<tr>
<td>2.1.1. General Case</td>
<td>5</td>
</tr>
<tr>
<td>2.1.2. Scale-Location Parameter Families</td>
<td>11</td>
</tr>
<tr>
<td>2.1.3. Special Cases of Scale-Location Parameter Families</td>
<td>13</td>
</tr>
<tr>
<td>2.2. Linear Models</td>
<td>27</td>
</tr>
<tr>
<td>3. MONTE CARLO STUDY OF PERFORMANCE</td>
<td>36</td>
</tr>
<tr>
<td>3.1. Description of the Test Statistics</td>
<td>36</td>
</tr>
<tr>
<td>3.1.1. Generalized Chi-Square ($G_X^2$)</td>
<td>36</td>
</tr>
<tr>
<td>3.1.2. Locally Most Powerful (LMP) Test</td>
<td>42</td>
</tr>
<tr>
<td>3.1.3. List of All Test Statistics</td>
<td>46</td>
</tr>
<tr>
<td>3.2. Alternatives Distributions</td>
<td>52</td>
</tr>
<tr>
<td>3.3. Random Number Generation</td>
<td>54</td>
</tr>
<tr>
<td>3.4. Size of Tests</td>
<td>54</td>
</tr>
<tr>
<td>3.5. Power of Tests</td>
<td>56</td>
</tr>
<tr>
<td>3.5.1. Normal Case</td>
<td>58</td>
</tr>
<tr>
<td>3.5.2. Extreme Value Case</td>
<td>65</td>
</tr>
<tr>
<td>4. APPLICATIONS</td>
<td>66</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td>72</td>
</tr>
<tr>
<td>Table</td>
<td>Page</td>
</tr>
<tr>
<td>-------</td>
<td>------</td>
</tr>
<tr>
<td>2.1.</td>
<td>16</td>
</tr>
<tr>
<td>2.2.</td>
<td>20</td>
</tr>
<tr>
<td>2.3.</td>
<td>25</td>
</tr>
<tr>
<td>2.4.</td>
<td>26</td>
</tr>
<tr>
<td>3.1.</td>
<td>43</td>
</tr>
<tr>
<td>3.2.</td>
<td>55</td>
</tr>
<tr>
<td>3.3.</td>
<td>57</td>
</tr>
<tr>
<td>3.4.</td>
<td>59</td>
</tr>
<tr>
<td>3.5.</td>
<td>61</td>
</tr>
<tr>
<td>3.6.</td>
<td>63</td>
</tr>
</tbody>
</table>

2.1. The $I^{\theta\theta}$ matrices for testing normality with $K \leq 4$.  
2.2. The $I^{\theta\theta}$ matrices for testing the logistic distribution with $K \leq 4$.  
2.3. The $I^{\theta\theta}$ matrices for testing extreme value distribution with $K \leq 4$.  
2.4. The $I^{\theta\theta}$ matrices for testing the negative exponentiality with $K \leq 4$.  
3.1. Numerical values for $I^{\theta\theta}$.  
3.2. Empirical size of the generalized Neyman smooth, classical chi-square, and generalized chi-square tests based on 2000 simulations.  
3.3. Empirical critical value based on 2000 simulations from normal and extreme value distributions.  
3.4. Empirical power of several tests of normality for symmetrically distributed alternatives.  
3.5. Empirical power of several tests of normality for asymmetrically distributed alternatives.  
3.6. Empirical powers of several tests of the extreme value distribution for several alternative distributions.
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,

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

where \( \pi_i \) are functions that describe the smoothness of the alternatives.
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 \( \chi^2_k \) 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 \( \chi^2 \) distribution of \( \psi_k^2 \) under local alternatives. He recommended that \( k \) generally need not be greater than four.

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

\[ g(y \mid \theta, \hat{\lambda}) = \left(1 + \sum_{i=1}^{k} \theta_i \pi_i[F(y \mid \hat{\lambda})]\right) C(\theta) \hat{r}(y \mid \hat{\lambda}) \]

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 \( \psi_k^2 \) when the smooth alternatives are of various orders and used a simulation study to compare the small sample powers of his tests with those 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 \( \lambda \) 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 \( \lambda \) assuming \( \theta = \varnothing \). The Javitz (1975) and Thomas and Pierce (1977) generalized smooth tests possess the property that their statistics have limiting \( \chi^2 \) 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,

\[
H_0 : Y \sim F(y|\lambda), \quad \lambda = (\lambda_1, \ldots, \lambda_p)^T \in \Lambda,
\]

(2.1)

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

\[
H_a : Y \sim G(y|\lambda, \theta), \quad \theta = (\theta_1, \ldots, \theta_k)^T \in \Theta,
\]

(2.2)

let \( G(y|\lambda, \theta) \) denote the cdf whose density function is

\[
g(y|\lambda, \theta) = e^{\sum_{i=1}^{k} \theta_i F_i(y|\lambda) - R(\theta)} \cdot f(y|\lambda)
\]

where

\[
R(\theta) = \log \left( \int_{-\infty}^{\infty} e^{\sum_{i=1}^{k} \theta_i F_i(y|\lambda)} f(y|\lambda) dy \right)
\]

is the normalizing constant. Consider the following hypotheses
\[ H_0 : \theta = \theta (\equiv \theta_0) \] vs. \[ H_a : \theta \neq \theta (2.3) \]

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, \( \ell = \ell (\theta, \lambda) \), is

\[
\ell = \sum_{j=1}^{N} \text{Log}[g(y_j | \lambda, \theta)] \\
= \sum_{j=1}^{N} \left( \sum_{i=1}^{k} \theta_i F^i(y_j | \lambda) - R(\theta) + \text{Log } f(y_j | \lambda) \right). \tag{2.5}
\]

Define the score vectors \( U_\theta \) and \( U_\lambda \) by

\[
U_\theta = \frac{\partial \ell}{\partial \theta} \bigg|_{\theta_0, \lambda} \quad \text{and} \quad U_\lambda = \frac{\partial \ell}{\partial \lambda} \bigg|_{\theta_0, \lambda}
\]

with \( i \)th components,

\[
U_{\theta}^{(i)} = \frac{\partial \ell}{\partial \theta_i} \bigg|_{\theta_0, \lambda} = \sum_{j=1}^{N} \left[ F^i(y_j | \lambda) - \frac{\partial R(\theta)}{\partial \theta_i} \right] \theta_0 = \sum_{j=1}^{N} F^i(y_j | \lambda) - \frac{N}{i+1}. \tag{2.6}
\]
since

\[
\frac{\partial R(\theta)}{\partial \theta_i} \bigg|_{\theta_0} = \frac{\partial}{\partial \theta_i} e^{R(\theta)} \bigg|_{\theta_0}
\]

\[
= \frac{\sum_{t=1}^{k} \theta_t F^t(y|\lambda) f(y|\lambda) dy}{\sum_{t=1}^{\infty} \sum_{e} \theta_t F^t(y|\lambda) f(y|\lambda) dy} \bigg|_{\theta_0}
\]

\[
= \frac{\int_{0}^{1} \sum_{i} \theta_t u^t e^{-t} du}{\int_{0}^{1} \sum_{i} \theta_t u^t e^{-t} du} \bigg|_{\theta_0}
\]

\[
= \frac{1}{i+1} \quad \text{for} \quad i = 1, \ldots, k,
\]

and

\[
U_{(i)} = \frac{\partial f}{\partial \lambda_i} \bigg|_{\hat{\lambda}, \hat{\theta}_0} = 0 \quad \text{for} \quad i = 1, \ldots, p+1
\]

where \(\hat{\lambda}\) is MLE of \(\lambda\).

The elements of the partitioned information matrix of dimension \((k+p+1) \times (k+p+1)\) are denoted by
\[
I(\theta_0, \hat{\lambda}) = \begin{pmatrix}
I_{\theta\theta}(\theta_0) & I_{\theta\lambda}(\theta_0, \hat{\lambda}) \\
I_{\theta\lambda}(\theta_0, \hat{\lambda}) & I_{\lambda\lambda}(\hat{\lambda}) \\
\end{pmatrix}
\]

\[
= \begin{pmatrix}
I_{\theta\theta} & I_{\theta\lambda} \\
I_{\theta\lambda} & I_{\lambda\lambda} \\
\end{pmatrix}
\quad \text{for } \theta = \theta_0
\quad (2.7)
\]

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

\[
I_{\theta\theta}(i, j) = \mathbb{E}\left(-\frac{\partial^2 \ell}{\partial \theta_i \partial \theta_j}\right)_{\theta_0}
\]

\[
= N \left( \frac{2\bar{R}(\theta)}{\partial \theta_i \partial \theta_j} \right)_{\theta_0}
\]

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

\[
= \frac{ijN}{(i+1)(j+1)(i+j+1)}, \quad (2.8)
\]

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

\[
I_{\theta\lambda}(i, j) = \mathbb{E}\left(-\frac{\partial^2 \ell}{\partial \theta_i \partial \lambda_j}\right)_{\theta_0, \hat{\lambda}}
\]

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

\[
= -NE(i \bar{F}^i(Y|\lambda) \frac{\partial}{\partial \lambda_j} F(Y|\lambda))_{\hat{\lambda}} \quad (2.9)
\]
And for \( i, j = 1, \ldots, p+1 \),

\[
I_{\lambda \lambda}(i, j) = E(-\frac{\partial^2 f}{\partial \lambda_i \partial \lambda_j})_{\hat{\theta}} \\
= -NE\left( \frac{\partial^2}{\partial \lambda_i \partial \lambda_j} \log(f(Y|\lambda)) \right)_{\lambda} \\
= -N \int_{-\infty}^{\infty} \left( \frac{\partial^2 f(Y|\lambda)}{\partial \lambda_i \partial \lambda_j} - \frac{\frac{\partial f(Y|\lambda)}{\lambda_i} \frac{\partial f(Y|\lambda)}{\lambda_j}}{f(Y|\lambda)} \right) dy.
\]

(2.10)

From (2.8) the matrix \( I_{\theta \theta} \) with \( K = 5 \) is

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

(2.11)

Denote the partitioned matrix in the inverse of the full information matrix (2.7) corresponding to \( I_{\theta \theta} \) by \( I_{\theta \theta}^{-1} \) and

\[
I_{\theta \theta}|\lambda = I_{\theta \theta} - I_{\theta \lambda} I_{\lambda \lambda}^{-1} I_{\lambda \theta}.
\]

(2.12)
In computing \( I^{\theta \theta} \), depending on convenience, one of the following two formulas can be used (Rao, 1973, p. 33):

\[
I^{\theta \theta} = I^{-1}_{\theta \theta} + I^{-1}_{\theta \theta} I_\lambda (I_{\lambda \lambda} - I_{\lambda \theta} I^{-1}_{\theta \theta} I_{\theta \lambda})^{-1} I_{\lambda \theta} I^{-1}_{\theta \theta} \tag{2.13}
\]

\[
I^{\theta \theta} = I^{-1}_{\theta \theta | \lambda} \tag{2.14}
\]

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

\[
W_k = \begin{bmatrix} U_{\theta} - I_{\theta \lambda} & I_{\lambda \lambda}^{-1} \end{bmatrix} I^{\theta \theta} \begin{bmatrix} U_{\theta} - I_{\theta \lambda} & I_{\lambda \lambda}^{-1} \end{bmatrix} = U^{\dagger}_{\theta} I^{\theta \theta} U_{\theta} \tag{2.15}
\]

From the large sample theory, under suitable regularity conditions (Kopecky, 1977), \( W_k \) has a limiting \( \chi^2 \) 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

\[ F(y|\lambda) = F_0\left(\frac{Y-\mu}{\sigma}\right), \]

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

\( \lambda = (\lambda_1, \lambda_2)' = (\mu, \sigma)' \)

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

\[ U_\theta(i) = \sum_{j=1}^{N} F_0^i\left(\frac{Y_j-\mu}{\sigma}\right) \cdot \frac{N}{i+1} \]

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

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

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

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

\[ = iN \int_{-\infty}^{\infty} F_0^{i-1}(z) f_0^2(z) zdz, \quad (2.16) \]
\[
\sigma^2 I_{\lambda \lambda}(1, 1) = -\sigma N \int_{-\infty}^{\infty} \left( \frac{\partial^2 f_0(y - \frac{\mu}{\sigma})}{(\partial \mu)^2} - \frac{\partial}{\partial \mu} \frac{f_0(y - \frac{\mu}{\sigma})}{f_0(y - \frac{\mu}{\sigma})} \right) \, dy \\
= -N \int_{-\infty}^{\infty} \left( f''_0(z) - \frac{[f'_0(z)]^2}{f'_0(z)} \right) \, dz,
\]

\[
\sigma^2 I_{\lambda \lambda}(1, 2) = -N \sigma \int_{-\infty}^{\infty} \left( \frac{\partial^2 f_0(y - \frac{\mu}{\sigma})}{(\partial \mu \partial \sigma)} - \frac{\partial}{\partial \mu} \frac{f_0(y - \frac{\mu}{\sigma}) \partial f_0(y - \frac{\mu}{\sigma})}{f_0(y - \frac{\mu}{\sigma})} \right) \, dy \\
= -N \int_{-\infty}^{\infty} \left( z f''_0(z) + 2 f'_0(z) - \frac{[f'_0(z)]^2}{f'_0(z)} \right) \, dz,
\]

and

\[
\sigma^2 I_{\lambda \lambda}(2, 2) = -N \sigma \int_{-\infty}^{\infty} \left( \frac{\partial^2 f_0(y - \frac{\mu}{\sigma})}{(\partial \sigma)^2} - \frac{\partial}{\partial \sigma} \frac{f_0(y - \frac{\mu}{\sigma})}{f_0(y - \frac{\mu}{\sigma})} \right) \, dy \\
= -N \int_{-\infty}^{\infty} \left( 2 f'_0(z) + 4z f'_0(z) + z^2 f''_0(z) - \frac{(f_0(z) + zf'_0(z))^2}{f'_0(z)} \right) \, dz
\]

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 \( \mu \) and \( \sigma \).
Proof. This result can be seen from relation (2.12), since
\[ I_{\theta} \] is independent of \( \mu \) and \( \sigma \) and the relation

\[ I_{\Theta \lambda} I_{\lambda \lambda}^{-1} I_{\lambda \theta} = [\sigma I_{\Theta \lambda}] [\sigma^2 I_{\lambda \lambda}]^{-1} [\sigma I_{\lambda \theta}] \]  

(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 \( \mu \) and \( \sigma \).

2.1.3. Special Cases of Scale-Location Parameter Families

In this section we evaluate the \( I_{\theta} \) 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),

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

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

\[
= (1-i)\sigma I_{\Theta \lambda}(i, 1) + iN \int_{-\infty}^{\infty} z \Phi^{i}(z)\phi(z) dz
\]
\[
\sigma^2 I_{\lambda\lambda}(i, 1) = \sigma^2 I_{\lambda\lambda}(1, 1) = -N \int_{-\infty}^{\infty} \left[ \phi''(z) - \frac{(\phi'(z))^2}{\phi(z)} \right] \, dz
\]

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

\[= N \]

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

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

\[= 0 \]
\[
\sigma^2 I_{\lambda\lambda}(z, 2) = -N \int_{-\infty}^{\infty} \left[ 2\phi(z) + 4z\phi(z) + z\phi''(z) - \frac{[\phi(z) + z\phi'(z)]^2}{\phi(z)} \right] dz
\]

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

\[
= 2N
\]

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

\[
\sigma I_{\theta\lambda} = N \begin{pmatrix}
.2820947918 & 0. \\
.2820947918 & .09188814923 \\
.2573438433 & .137832239 \\
.2325928947 & .1600040872
\end{pmatrix},
\]

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

\[
I_{\theta\theta} \mid \lambda = I_{\theta\theta} - I_{\theta\lambda} I_{\lambda\lambda}^{-1} I_{\lambda\theta}
\]

\[
I_{\theta\theta} \mid \lambda = \frac{N}{1000} \begin{pmatrix}
3.755861773 & 3.755861773 & 2.404642103 & 1.053422462 \\
3.755861773 & 5.089701344 & 4.405401457 & 3.225992265 \\
2.404642103 & 4.405401457 & 4.632428200 & 4.116790968 \\
1.053422462 & 3.225992265 & 4.116790968 & 4.211002486
\end{pmatrix}
\]

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 $I_0$ matrices for testing normality with $K \leq 4$.

| $K$ = 1 | $\frac{1}{N}$ [ 266.2504801 ] |
| $K$ = 2 | $\frac{1}{N}$ \[
\begin{pmatrix}
1015.965829 & -749.7153493 \\
749.7153493 &
\end{pmatrix}
\]
| $K$ = 3 | $\frac{1}{N}$ \[
\begin{pmatrix}
9072.727849 & -14806.08280 & 9370.911645 \\
25273.29688 & -16349.12104 &
\end{pmatrix}
\]
| $K$ = 3 | $\frac{1}{N}$ \[
\begin{pmatrix}
23058.99425 & -66471.76917 & 84729.75129 & -37679.41831 \\
216127.9875 & -294726.9285 & 139188.8981 &
\end{pmatrix}
\]
| $K$ = 3 | $\frac{1}{N}$ \[
\begin{pmatrix}
216127.9875 & -294726.9285 & 139188.8981 &
\end{pmatrix}
\]
| $K$ = 3 | $\frac{1}{N}$ [ 101509.4753 ] |
b) Logistic Distribution

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

\[ F_0(z) = \frac{e^z}{1+e^z} \]

and

\[ f_0(z) = \frac{e^z}{(1+e^z)^2}, \quad -\infty < z < \infty \]

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

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

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

and

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

\[ = \frac{iN}{(i+1)} \left\{ \int_{0}^{1} u^{i+1} \log \left( \frac{u}{1-u} \right) du - \frac{1}{i+1} \right\} \]

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

From equation (2.17), it is found that
\[ \sigma^2 I_{\lambda \lambda} (1, 1) = N \int_{-\infty}^{\infty} \frac{2e^{2z}}{(1+e^z)^4} \, dz = \frac{N}{3} \]

\[ \sigma^2 I_{\lambda \lambda} (1, 2) = N \int_{-\infty}^{\infty} \frac{e^{3z} - e^{-z} + 2ze^{2z}}{(1+e^z)^4} \, dz = 0, \]

and

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

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

\[ + 2E(Q_3) \frac{\Gamma(3) \Gamma(1)}{\Gamma(4)} - 1 \right] \]

\[ = N \left[ \frac{2}{3} \psi'(2) - \frac{4}{3} (\psi(1) - \psi(3)) - 1 \right] \]

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

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,
\begin{align*}
\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},
\end{align*}

and

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

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 $I^{th}$ matrices for testing the Logistic distribution with $R \leq 4$.

<table>
<thead>
<tr>
<th>R = 1</th>
<th>$\frac{1}{N} \begin{bmatrix} 1430.305318 \end{bmatrix}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>R = 2</td>
<td>$\frac{1}{N} \begin{bmatrix} 7730.305318 &amp; -4200 \ 2800 \end{bmatrix}$</td>
</tr>
<tr>
<td></td>
<td>$\begin{bmatrix} 239522.5025 &amp; -331123.3367 &amp; 163461.6684 \ 463897.7823 &amp; -230548.8911 \ 115274.4456 \end{bmatrix}$</td>
</tr>
<tr>
<td>R = 3</td>
<td>$\frac{1}{N} \begin{bmatrix} 724622.5025 &amp; -1624723.337 &amp; 1618761.668 &amp; -582120 \ 3913497.782 &amp; -4111348.891 &amp; 1552320 \end{bmatrix}$</td>
</tr>
<tr>
<td>R = 4</td>
<td>$\frac{1}{N} \begin{bmatrix} 4481174.446 &amp; -1746360 \ 698544 \end{bmatrix}$</td>
</tr>
</tbody>
</table>
c) Extreme Value Distribution

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

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

and

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

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

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

while \( Y \) has an extreme value distribution (EVD). Consequently, the matrices \( I^0 \theta \) 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,

\[ \bar{F}(y|\lambda) = 1 - F(y|\lambda) \]

is used in place of \( F(y|\lambda) \). The evaluation of equations (2.16), give, for \( i = 1, \ldots, k \)
\[ \sigma I_{\theta \lambda}(i, 1) = -N \int_{-\infty}^{\infty} i e^{2z-(i+1)e^z} \, dz \]
\[ = -iN \int_{0}^{\infty} u e^{-(i+1)u} \, du \]
\[ = -\frac{iN}{(i+1)^2} \]

and

\[ \sigma I_{\theta \lambda}(i, 2) = -N \int_{-\infty}^{\infty} iz e^{2z-(i+1)e^z} \, dz \]
\[ = -iN \int_{0}^{\infty} u [\log u] e^{-(i+1)u} \, du \]
\[ = -iN \int_{0}^{\infty} \left[ \log \left( \frac{w}{i+1} \right) \frac{w}{i+1} e^{-w} \, dw \right] \]
\[ = -\frac{iN}{(i+1)^2} \left[ \int_{0}^{\infty} w \log w e^{-w} \, dw - \log(i+1) \int_{0}^{\infty} w e^{-w} \, dw \right] \]
\[ = -\frac{iN}{(i+1)^2} [1 - \gamma - \log(i+1)], \]

where the Euler's constant,

\[ \gamma = -\int_{0}^{\infty} e^{x} \log x \, dx = .577215664... . \]

From equation (2.17), it is found that
\[ \sigma^2 I_{\lambda\lambda}(1, 1) = N \int_{-\infty}^{\infty} e^z e^{-e^z} \, dz \]
\[ = N \int_{0}^{\infty} u e^{-u} \, du \]
\[ = N, \]

\[ \sigma^2 I_{\lambda\lambda}(1, 2) = N \int_{-\infty}^{\infty} [(1+z)e^z - 1] e^{-e^z} \, dz \]
\[ = N \int_{0}^{\infty} u \left( \log u \right) e^{-u} \, du \]
\[ = N(1 - \gamma), \]

and

\[ \sigma^2 I_{\lambda\lambda}(2, 2) = N \int_{-\infty}^{\infty} [2z(e^z - 1) + z^2 e^z - 1] e^{-e^z} \, dz \]
\[ = N \left[ 2 \int_{0}^{\infty} u e^{-u} (\log u) \, du - 2 \int_{0}^{\infty} e^{-u} (\log u) \, du \right. \]
\[ + \left. \int_{0}^{\infty} u e^{-u} (\log u)^2 \, du - 1 \right] \]
\[ = N[(1 - \gamma)^2 + \psi'(1)] \quad (2.19) \]

where \( \gamma \) 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,
\[
\sigma^2 I_{\theta \theta} = N \begin{pmatrix}
\frac{1}{4} & 0.06759071137 \\
\frac{2}{9} & 0.1501839897 \\
\frac{3}{16} & 0.1806581299 \\
\frac{4}{25} & 0.1898647524
\end{pmatrix},
\]

\[
\sigma^2 I_{\lambda \lambda} = N \begin{pmatrix}
1 & 0.4227843351 \\
1.823680661
\end{pmatrix},
\]

and

\[
I_{\theta \theta | \lambda} = I_{\theta \theta} - I_{\theta \lambda} I_{\lambda \lambda}^{-1} I_{\lambda \theta}
\]

\[
= \frac{N}{1000} \begin{pmatrix}
2.578307813 & 2.059083644 & 0.7424617186 & -0.4609215771 \\
3.272247710 & 3.088625464 & 2.416067210 \\
4.127085433 & 4.308617632 & 5.198593163
\end{pmatrix}
\]

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,

\[
f(x | \beta) = \beta e^{-\beta x} \quad 0 < x < \infty, \ 0 < \beta < \infty
\]

can be treated as a scale parameter family (\(\mu = 0, \sigma = 1/\beta\)) and by the transformation,
Table 2.3. The $I^0$ matrices for testing extreme value distribution with $K \leq 4$.

<table>
<thead>
<tr>
<th>$K$</th>
<th>$\frac{1}{N}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$[387.8512856]$</td>
</tr>
<tr>
<td>2</td>
<td>$\begin{pmatrix} 779.6567427 &amp; -490.6041943 \ 614.3162862 &amp; \end{pmatrix}$</td>
</tr>
<tr>
<td>3</td>
<td>$\begin{pmatrix} 10852.26625 &amp; -16981.68805 &amp; 10756.42426 \ 27613.85861 &amp; -17610.63946 &amp; \ 11486.66220 &amp; \end{pmatrix}$</td>
</tr>
<tr>
<td>4</td>
<td>$\begin{pmatrix} 19720.32418 &amp; -54593.34281 &amp; 66758.06618 &amp; -28208.45334 \ 187134.3066 &amp; -255127.6289 &amp; 119639.1157 &amp; \ 365165.1221 &amp; -178135.9251 &amp; \end{pmatrix}$</td>
</tr>
</tbody>
</table>
\[ Y = \log X, \quad \mu = \log \beta \]

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

\[
\begin{pmatrix}
\frac{1}{48} & \frac{1}{36} & \frac{9}{320} & \frac{2}{75} \\
\frac{16}{405} & \frac{1}{24} & \frac{64}{1575} & \\
\frac{81}{1792} & \frac{9}{200} & \\
\frac{256}{5625} &
\end{pmatrix}
\]

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

<table>
<thead>
<tr>
<th>K = 1</th>
<th>( \frac{1}{N} ) [48]</th>
</tr>
</thead>
<tbody>
<tr>
<td>K = 2</td>
<td>( \frac{1}{N} \begin{pmatrix} 768 &amp; -540 \ 405 &amp; \end{pmatrix} )</td>
</tr>
<tr>
<td>K = 3</td>
<td>( \frac{1}{N} \begin{pmatrix} 480 &amp; -8100 &amp; 4480 \ 14580 &amp; -8400 &amp; 4977.7 \end{pmatrix} )</td>
</tr>
<tr>
<td>K = 4</td>
<td>( \frac{1}{N} \begin{pmatrix} 19200 &amp; -56700 &amp; 71680 &amp; -31500 \ 178605 &amp; -235200 &amp; 106312.5 &amp; \ 318577.7 &amp; -147000 &amp; \ &amp; &amp; 68906.25 \end{pmatrix} )</td>
</tr>
</tbody>
</table>
2.2. Linear Models

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

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

(2.20)

with $\mu(x) = \beta_0 + \beta_1 x_1 + \ldots + \beta_p x_p = \beta' x$ and smooth alternatives,

$$ H_a : Y \sim G \left( \frac{Y - \mu(x)}{\sigma} \right) $$

(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 $\left( Y_j - \mu(x_j) \right)/\sigma$, given $X_j = x_j$. Let $\lambda = (\beta/\sigma)$, and $F(y|\lambda, x)$ and $G(y|\theta, \lambda, x)$ denote the cdf's in (2.20) and (2.21), respectively. Then the pdf corresponding to $G(y|\theta, \lambda, x)$

$$ g(y|\theta, \lambda, x) = e^{\sum_{i=1}^{k} \theta_i F_i(y|\lambda, x)} R(\theta) $$

$$ f(y|\lambda, x) $$

where

$$ R(\theta) = \log \left[ \int_{-\infty}^{\infty} e^{\sum_{i=1}^{k} \theta_i F_i(y|\lambda, x)} f(y|\lambda, x) dy \right] $$
The log-likelihood function, \( l \equiv l(\theta, \lambda) \), is

\[
l = \sum_{j=1}^{N} \log \left[ g(y_j | \lambda, \lambda) \right] \\
= \sum_{j=1}^{N} \left( \sum_{i=1}^{k} \theta_i F_i^i(y_j | \lambda, \lambda) - R(\theta) + \log \left[ f(y_j | \lambda, \lambda) \right] \right)
\]  

(2.22)

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

\[
U_{\theta}^{(i)} = \sum_{j=1}^{N} F_i^i(y_j | \lambda, \lambda) - \frac{N}{i+1},
\]

and, for \( i = 1, \ldots, p+1, \)

\[
U_{\lambda}^{(i)} = 0,
\]

respectively.

Let \( x_0 \equiv x \equiv 1 \). Then for \( i = 1, \ldots, k, \)

\[
I_{\theta \lambda}^{(i, j)} = \begin{cases} 
\overline{x}_{j-1} I_{\theta \lambda}^{(i, 1)}, & j = 1, \ldots, p \\
I_{\theta \lambda}^{(i, 2)}, & j = p+1
\end{cases}
\]

(2.23)

And

\[
I_{\lambda \lambda}^{(j, i)} = I_{\lambda \lambda}^{(i, j)} = \begin{cases} 
\overline{x}_{j-1} I_{\lambda \lambda}^{(1, 1)}, & 1 \leq i \leq j \leq p \\
\overline{x}_{i-1} I_{\lambda \lambda}^{(1, 2)}, & 1 \leq i < j = p+1 \\
I_{\lambda \lambda}^{(2, 2)}, & i = j = p+1
\end{cases}
\]

(2.24)
where \( \bar{x}_l \) and \( \bar{x}_l \bar{x}_m \) denote means of \( x_l \) and \( x_l x_m \), 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 \( \theta \) and \( \sigma \).

**Proof.** The matrices \( \sigma I_{\theta \lambda} \) and \( \sigma^2 I_{\lambda \lambda} \) are independent of \( \theta \) and \( \sigma \). 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 \( \theta \) and \( \sigma \).

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

**Proof.** Let \( p \) be an arbitrary positive integer \( \geq 1 \), 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) \( P = 2_\gamma [\lambda = (\beta_0, \beta_1, \sigma)'] \). From equations (2.23) and (2.24), we may define \( \sigma I_{\theta\lambda}(3) \) and \( \sigma^2 I_{\lambda\lambda}(3) \) as

\[
\sigma I_{\theta\lambda}(3) = (C, x_1 C, D)
\]

and

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

So let
\[
\frac{1}{N} I(3) = I_{\theta\lambda}(3) I_{\lambda\lambda}^{-1}(3) I_{\lambda\theta}(3)
\]

\[
= (C, x_1 C, D)
\]

\[
\begin{pmatrix}
a_{11} & \overline{x_1 a_{11}} & a_{12} \\
\overline{x_1 a_{11}} & \overline{x_1 a_{11}} & \overline{x_1 a_{12}} \\
a_{12} & \overline{x_1 a_{12}} & a_{22}
\end{pmatrix}^{-1}
\]

\[
\begin{pmatrix}
C' \\
\overline{x_1 C'} \\
D'
\end{pmatrix}
\]

\[
= (C, D, \overline{x_1 C})
\]

\[
\begin{pmatrix}
a_{11} & a_{12} & \overline{x_1 a_{11}} \\
a_{12} & a_{22} & \overline{x_1 a_{12}} \\
\overline{x_1 a_{11}} & \overline{x_1 a_{12}} & \overline{x_1 a_{11}}
\end{pmatrix}^{-1}
\]

\[
\begin{pmatrix}
C' \\
D' \\
\overline{x_1 C'}
\end{pmatrix}
\]
\[ \begin{aligned}
&= (\sigma I_{\theta \lambda}(2), \overline{x}_1 C) \left( \begin{array}{cc}
\sigma^2 I_{\lambda \lambda}(2) & B \\
B' & \frac{x_1}{x_1}a_{11}
\end{array} \right)^{-1} \left( \begin{array}{c}
\sigma I_{\lambda \theta}(2) \\
\overline{x}_1 C
\end{array} \right) \\
&= (\sigma I_{\theta \lambda}(2), \overline{x}_1 C) \left( \begin{array}{cc}
[\sigma^2 I_{\lambda \lambda}(2)]^{-1} + FE^{-1} & -FE^{-1} \\
-FE^{-1} & E^{-1}
\end{array} \right) \left( \begin{array}{c}
\sigma I_{\lambda \theta}(2) \\
\overline{x}_1 C
\end{array} \right) \\
&= \frac{1}{N} I(2), \\
\end{aligned} \]

(2.25)

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

\[ \begin{aligned}
\sigma I_{\theta \lambda}(2) F - \overline{x}_1 C &= (C, D) [\sigma^2 I_{\lambda \lambda}(2)]^{-1} \overline{x}_1 \begin{pmatrix} a_{11} \\ a_{12} \end{pmatrix} - \overline{x}_1 C \\
&= (C, D) \begin{pmatrix} a_{22} & -a_{12} \\ -a_{12} & a_{11} \end{pmatrix} \frac{\overline{x}_1}{a_{11}a_{22}(a_{12})^2} \begin{pmatrix} a_{11} \\ a_{12} \end{pmatrix} - \overline{x}_1 C \\
&= (C, D) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \frac{\overline{x}_1}{x_1} - \overline{x}_1 C \\
&= 0
\end{aligned} \]

ii) Show the theorem is true for \( p = \ell + 1 \) assuming it is true when \( p = \ell \ (\geq 2) \).
\[
\frac{1}{N} I_{\ell+1} = I_{\theta \lambda}(\ell+1) I_{\lambda \lambda}^{-1}(\ell+1) I_{\lambda \theta}(\ell+1)
\]

\[
= (C, x_1 C, \ldots, x_{\ell-1} C, D)
\]

\[
\begin{pmatrix}
  a_{11} & x_1 a_{11} & \ldots & x_{\ell-1} a_{11} & a_{12} \\
x_1^2 a_{11} & \ldots & x_{\ell-1} x_{\ell-1} a_{11} & \ldots & x_{\ell-1} a_{12} \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
x_{\ell-1}^2 a_{11} & \ldots & x_{\ell-1} a_{11} & \ldots & a_{22}
\end{pmatrix}^{-1}
\]

\[
\times \begin{pmatrix}
  C' \\
x_1 C' \\
\vdots \\
x_{\ell-1} C' \\
D'
\end{pmatrix}
\]

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

\[
|A| = \sum_{s=1}^{\ell+1} (-1)^{s+1} A(s, 1) A_s, 1
\]

\[
= \sum_{s=1}^{\ell+1} (-1)^{s+1} a_{11} \left( \frac{a_{12}}{a_{11}} \right)^{\delta} t_{s, 1} A_s, 1
\]

where

\[
t_{i, j} = \frac{x_{i-1} x_{j-1}}{x_{i-1} x_{j-1}} = \frac{1}{N} \sum_{s=1}^{N} x_{i-1, s} x_{j-1, s} \quad \text{for} \quad i, j = 1, \ldots, \ell+1
\]
and \( x_0, s = x_l, s = 1 \) 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 \( \sigma H = I_{\ell+1} \lambda(\ell+1) I_{\lambda(\ell+1)}^{-1} \) and \( H_{:,j} \) is the \( j \)th column of \( H \), \( j = 1, \ldots, \ell+1 \). Then

\[
H_{:,1} = \begin{bmatrix} C(1, \overline{1}, \ldots, \overline{1}_{\ell-1}), D \end{bmatrix} \begin{pmatrix} q(1, 1) \\ \vdots \\ q(\ell+1, 1) \end{pmatrix} = C(t_1, 1', \ldots, t_{\ell}, 1) \begin{pmatrix} A_{1,1} \\ -A_{2,1} \\ \vdots \\ (-1)^{\ell+1} A_{\ell,1} \end{pmatrix} \frac{1}{|A|} + D (-1)^{\ell+2} A_{\ell+1,1} \frac{1}{|A|} =
\]
\[
\begin{align*}
&= \frac{C}{a_{11} |A|} \left\{ \sum_{s=1}^{t} (-1)^{s+1} a_{11} t_s, 1 A_{1, 1} + (-1)^{t+2} a_{12} t_{t+1}, 1 A_{t+1, 1} - (-1)^{t+2} a_{12} t_{t+1}, 1 A_{t+1, 1} \right\} + \frac{D A_{t+1, 1}}{|A|} \\
&= \frac{C}{a_{11}} - \frac{C}{a_{11}} a_{12} \frac{a_{12}}{a_{11} a_{22} - (a_{12})^2} - \frac{a_{12} D}{a_{11} a_{22} - (a_{12})^2} \\
&= \frac{a_{22} C - a_{12} D}{a_{11} a_{22} - (a_{12})^2}.
\end{align*}
\]

For \( j = 2, \ldots, t \),

\[
H_{., j} = [C(1, x_1, \ldots, x_{t-1}), D] \begin{pmatrix} (-1)^{j+1} A_{1, j} \\ \vdots \\ (-1)^{j+\ell} A_{\ell+1, j} \end{pmatrix} \frac{1}{|A|}
\]

\[
= C \sum_{s=1}^{t} (-1)^{j+s} t_s, 1 A_{s, j} \frac{1}{|A|}
\]

\[
= 0, \quad \text{since} \quad j \neq 1.
\]

And
\begin{equation}
H_{+, \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}
\end{equation}

\begin{align*}
&= 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}
\end{align*}

Therefore,

\begin{equation}
\frac{1}{N} I(\ell+1) = H \lambda \theta (\ell+1)
\end{equation}

\begin{align*}
&= \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) \\
&= \left( \begin{array}{c}
\frac{C'}{\overline{x_1} C'} \\
\vdots \\
\frac{\overline{x_{\ell-1}} C'}{D'}
\end{array} \right)
\end{align*}

\begin{align*}
&= \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{bmatrix}
C' \\
D'
\end{bmatrix}
\frac{1}{a_{11}a_{22}-(a_{12})^2}
\end{align*}

\begin{equation}
\frac{1}{N} I(2)
\end{equation}

Therefore, for any positive integer \( p (\geq 2) \), \( I(p) = I(2) \). 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 \( \chi^2 \)-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 \( \chi^2 \) 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 $\theta_1 = \ldots = \theta_k = 0$:

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

where

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

$R(\theta)$ is the normalizing constant and $K$ is the number of intervals.

The log-likelihood function, $\ell = \ell(\theta, \mu, \sigma)$, is

$$\ell = \sum_{j=1}^{N} \sum_{i=1}^{k} \theta_i \varphi_i \left[ F_0 \left( \frac{y_j - \mu}{\sigma} \right) \right] - N R(\theta) - N \log \sigma + \sum_{j=1}^{N} \log f_0 \left( \frac{y_j - \mu}{\sigma} \right).$$

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, $T_0 = -\infty$, $T_i = \mu + \sigma F_0^{-1}(i/K)$, $i = 1, \ldots, k-1$, 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.

$$\frac{\partial \ell}{\partial \theta_i} \bigg|_{\theta = 0, \hat{\mu}, \hat{\sigma}} = \sum_{j=1}^{N} \varphi_i \left[ F_0 \left( \frac{y_i - \hat{\mu}}{\hat{\sigma}} \right) \right] - N \frac{\partial R(\theta)}{\partial \theta_i} \bigg|_{\theta = 0}$$

$$= O_i - N P_i$$

where $Q = (O_1, \ldots, O_K)'$ 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 $\mu$ and $\sigma$ are equal to zero when evaluated at the maximum likelihood estimates:

$$\frac{\partial \ell}{\partial \mu} \bigg|_{\theta = 0, \hat{\mu}, \hat{\sigma}} = -\frac{1}{\hat{\sigma}} \sum_{j=1}^{N} \frac{f'_0 \left( \frac{y_i - \hat{\mu}}{\hat{\sigma}} \right)}{f_0 \left( \frac{y_i - \hat{\mu}}{\hat{\sigma}} \right)}$$

$$\frac{\partial \ell}{\partial \sigma} \bigg|_{\theta = 0, \hat{\mu}, \hat{\sigma}} = -\frac{1}{\hat{\sigma}} \sum_{j=1}^{N} \frac{\left( \frac{y_i - \hat{\mu}}{\hat{\sigma}} \right) f'_0 \left( \frac{y_i - \hat{\mu}}{\hat{\sigma}} \right)}{f_0 \left( \frac{y_i - \hat{\mu}}{\hat{\sigma}} \right)}$$

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 \( i, j = 1, \ldots, K \),

\[
I_{\theta \theta}(i, j) = E\left( -\frac{\partial^2 l}{\partial \theta_i \partial \theta_j} \mid \theta = \theta \right) \\
= N \left( \int_{-\infty}^{\infty} \varphi_i[F_0(z)] \varphi_j[F_0(z)] f_0(z) dz \right) \\
- \left( \int_{-\infty}^{\infty} \varphi_i[F_0(z)] f_0(z) dz \right) \left( \int_{-\infty}^{\infty} \varphi_j[F_0(z)] f_0(z) dz \right) \\
= N \left\{ \begin{array}{ll}
P_i(1-P_i) & \text{if } i = j \\
-P_i P_j & \text{if } i \neq j \\
\end{array} \right.
\]

\[= N \left\{ \begin{array}{ll}
\frac{1}{K} (1 - \frac{1}{K}) & \text{if } i = j \\
- \frac{1}{K^2} & \text{if } i \neq j \\
\end{array} \right. \]  \hfill (3.1)

For \( i = 1, \ldots, K \),

\[
I_{\theta \lambda}(i, 1) = E\left[ \frac{\partial l}{\partial \theta_i} \frac{\partial l}{\partial \mu} \mid \theta = \theta \right] \\
= - \frac{1}{\sigma} E \left( \sum_{j=1}^{N} \varphi_i[F_0(\frac{Y_j - \mu}{\sigma})] - \frac{N}{K} \right) \sum_{k=1}^{N} \frac{Y_k - \mu}{\sigma} f_0\left( \frac{Y_k - \mu}{\sigma} \right)
\]
N \sum_{j=1}^{N} \sum_{k=1}^{N} \left\{ \varphi_i \left[ F_0(Z_j) \right] - \frac{1}{K} \right\} \frac{f_0'(Z_k)}{f_0(Z_k)} \\
= -N \left\{ \varphi_i \left[ F_0(Z) \right] - \frac{1}{K} \right\} \frac{f_0'(Z)}{f_0(Z)} - N(N-1)E \left\{ \varphi_i \left[ F_0(Z) \right] - \frac{1}{K} \right\} E \left\{ \frac{f_0'(Z)}{f_0(Z)} \right\} \\
= -N \left[ \int \frac{F_0^{-1}(i)}{K} f_0'(z)dz - \frac{1}{K} \int_{-\infty}^{\infty} f_0'(z)dz \right] \\
= -N \left[ f_0 \left\{ F_0^{-1}(\frac{i}{K}) \right\} - f_0 \left\{ F_0^{-1}(\frac{i-1}{K}) \right\} \right]. \\
(3.2) \\
For \ i = 1, \ldots, K. \\
I_{\theta_{\lambda}}(i, 2) = E \left[ \frac{\partial \theta_i}{\partial \theta} \frac{\partial \theta_i}{\partial \sigma} \right]_{\theta=\theta} \\
= -\frac{1}{\sigma} \left( \sum_{j=1}^{N} \left( \varphi_i \left[ F_0 \left( \frac{Y_i-\mu}{\sigma} \right) \right] - \frac{1}{K} \right) \sum_{k=1}^{N} \left( 1 + \left( \frac{Y_i-\mu}{\sigma} \right) \frac{f_0'(\frac{Y_i-\mu}{\sigma})}{\frac{Y_i-\mu}{\sigma}} \right) \right) \\
= -E \left\{ \sum_{j=1}^{N} \sum_{k=1}^{N} \left( \varphi_i \left[ F_0(Z_j) \right] - \frac{1}{K} \right) \left( 1 + Z_k \frac{f_0'(Z_k)}{f_0(Z_k)} \right) \right\} \\
= -N \left\{ \varphi_i \left[ F_0(Z) \right] - \frac{1}{K} \right\} \left\{ 1 + Z \frac{f_0'(Z)}{f_0(Z)} \right\} \\
= -N \left[ \int \frac{F_0^{-1}(\frac{i}{K})}{K} \left\{ f_0(z) + z f_0'(z) \right\}dz - \frac{1}{K} \int_{-\infty}^{\infty} \left\{ f_0(z) + z f_0'(z) \right\}dz \right].
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.

\[
I^{\theta\theta} = I_{\theta\theta}^{-1} + I_{\theta\theta}^{-1} I_{\theta\lambda} (I_{\lambda\lambda} - I_{\lambda\theta} I_{\theta\theta}^{-1} I_{\theta\lambda})^{-1} I_{\lambda\theta} I_{\theta\theta}^{-1}
\]

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 \( \chi^2 \) statistic may be evaluated as

\[
G_{\chi^2} = (\mathcal{Q} - \mathcal{E})' I^{\theta\theta} (\mathcal{Q} - \mathcal{E}) \tag{3.4}
\]

where \( \mathcal{E} = (E_1, \ldots, E_K)' \) 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 \( \chi^2 \).
distribution with \( K-1 \) df. Table 3.1 gives the evaluation of \( 1^{\theta \theta} \) 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

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

which include the extreme value densities when \( \theta = 0 \). A test of

\[ H_0: \theta = 0 \text{ vs. } H_a: \theta \neq 0 \]

based on the scores statistic is developed below. Since

\[
\frac{y-\mu}{\sigma}
\]

equation (3.5) can be written as

\[
g_\theta(y; \mu, \sigma) = e^{\frac{\theta \log(-\log(F_0(\frac{y-\mu}{\sigma})))) - \log(\Gamma(\theta+1))}{f_0(\frac{y-\mu}{\sigma})}}
\]

The loglikelihood function, \( l \equiv l(\theta, \mu, \sigma) \) is
Table 3.1 Numerical values for $I_{00}$.

a) Normal with $K = 5$

\[
\begin{bmatrix}
  25.3208 & 6.46276 & -1.08428 & -7.98856 & -17.7106 \\
  8.19162 & .633865 & -2.29960 & -7.98856 \\
  5.90083 & .633865 & -1.08428 \\
  8.19162 & 6.46276 & \\
  25.3208 &
\end{bmatrix}
\]

b) EV with $K = 5$

\[
\begin{bmatrix}
  8.55649 & .168470 & -3.38135 & -8.97018 \\
  5.75376 & .904801 & -.412283 \\
  9.73575 & 8.54678 & \\
  27.6255 &
\end{bmatrix}
\]

c) Logistic with $K = 5$

\[
\begin{bmatrix}
  55.0721 & 22.8370 & -1.81810 & -25.1630 & -45.9279 \\
  17.6528 & 1.02046 & -11.3472 & -25.1630 \\
  6.59529 & 1.02046 & -1.81810 \\
  17.6528 & 22.8370 & \\
  55.0721 &
\end{bmatrix}
\]

d) Normal with $K = 10$

\[
\begin{bmatrix}
  11.9980 & 2.57891 & 2.96156 & 3.03903 & 1.70158 \\
  21.6856 & 17.1894 & 36.6767 & 45.1892 \\
  93.7395 &
\end{bmatrix}
\]
Table 3.1 (Continued)

e) EV with $K = 10$

<table>
<thead>
<tr>
<th>$\frac{1}{N}$</th>
<th>116.151</th>
<th>61.3047</th>
<th>38.4436</th>
<th>20.2480</th>
<th>4.00210</th>
<th>-11.4830</th>
<th>-27.0162</th>
<th>-43.4395</th>
<th>-62.1240</th>
<th>-86.0864</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>17.2342</td>
<td>11.0462</td>
<td>15.1227</td>
<td>18.9880</td>
<td>27.7768</td>
<td>25.4410</td>
<td>35.3063</td>
<td>47.6802</td>
<td>55.9942</td>
<td>103.636</td>
</tr>
</tbody>
</table>

f) Logistic with $K = 10$

<table>
<thead>
<tr>
<th>$\frac{1}{N}$</th>
<th>248.74</th>
<th>183.71</th>
<th>135.94</th>
<th>82.735</th>
<th>29.094</th>
<th>-24.906</th>
<th>-79.265</th>
<th>-134.06</th>
<th>-194.29</th>
<th>-247.26</th>
</tr>
</thead>
<tbody>
<tr>
<td>84.791</td>
<td>36.293</td>
<td>8.1463</td>
<td>14.535</td>
<td>23.603</td>
<td>29.094</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11.970</td>
<td>36.293</td>
<td>44.615</td>
<td>65.157</td>
<td>82.735</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>84.791</td>
<td>106.17</td>
<td>135.94</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>150.42</td>
<td>183.71</td>
<td>248.74</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
\[
\ell = \theta \sum_{i=1}^{N} \log[-\log\left(F_0\left(\frac{y_i - \mu}{\sigma}\right)\right)] - N \log \Gamma(\theta + 1) + \sum_{i=1}^{N} \log[f_0\left(\frac{y_i - \mu}{\sigma}\right)]
\]

where \(\psi(\cdot)\) is a digamma function and \(\gamma\) 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:

\[
I_{\theta\theta} = \mathbb{E}\left(-\frac{\partial^2 \ell}{\partial \theta^2}\right)_{\theta=0} = N \psi'(1)
\]

\[
= 1.644934N
\]

\[
I_{\theta\lambda}(1, 1) = \frac{\partial^2 \ell}{\partial \theta \partial \mu} |_{\theta=0} = \frac{N}{\sigma}
\]

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

\[
= -\frac{N}{\sigma} \gamma
\]

From (2.19)
\[ I_{\lambda \lambda} = \frac{N}{\sigma^2} \left( \begin{array}{c} 1 \\ (1-\gamma)^2 + \psi'(1) \end{array} \right)^{-1} \]

\[ I_{\theta \theta}^{-1} = (I_{\theta \theta} - I_{\lambda \lambda}^{-1} I_{\lambda \theta} I_{\lambda \theta})^{-1} \]

\[ = \left( N \psi'(1) - N(1-\gamma) \left( 1 - \frac{1-\gamma}{(1-\gamma)^2 + \psi'(1)} \right) \right)^{-1} \left( 1 \right) \]

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

The LMP test statistic is defined by

\[ \text{LMP} = \left( \frac{\partial f}{\partial \theta}, \frac{\partial f}{\partial \mu}, \frac{\partial f}{\partial \sigma} \right) \left( I_{\theta \theta}^{-1} I_{\theta \lambda} \right)^{-1} \left( \begin{array}{c} \frac{\partial f}{\partial \theta} \\ \frac{\partial f}{\partial \mu} \\ \frac{\partial f}{\partial \sigma} \end{array} \right) \theta = 0, \hat{\mu}, \hat{\sigma} \]

\[ = \frac{\partial f}{\partial \theta} I_{\theta \theta}^{-1} \frac{\partial f}{\partial \theta} \bigg|_{\theta = 0, \hat{\mu}, \hat{\sigma}} \]

\[ = I_{\theta \theta}^{-1} \theta \ U^2 \theta \quad (3.7) \]

### 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 \( \chi^2 (C \chi^2) \)
\[
C \chi^2 = \frac{(Q-E)'(Q-E)}{E_1}
= \frac{K}{N} \sum_{i=1}^{K} \frac{Q_i^2}{E_i} - N
\]
where \( Q, E, E_1 \) are described in Section 3.1.1 and \( K \) is number of intervals. As mentioned before, even for fixed intervals, the \( \chi^2 \) statistics do not, in general, have limiting \( \chi^2 \) 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 $\chi^2(Gx^2)$

From (3.4),

$$Gx^2 = (Q-E)^T \theta (Q-E)$$

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.

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

In the KS statistic, and other statistics below, $y_1 \leq y_2 \leq \ldots \leq y_N$ 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.

\[
SW = \frac{\sum_{i=1}^{N} a_i y_i}{\sum_{i=1}^{N} (y_i - \bar{y})^2}
\]

where the coefficients

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

are such that \( \sum_{i=1}^{N} a_i y_i \) is the minimum variance unbiased linear combination order statistic estimator for \( \sigma \). 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) $SK = \frac{\sqrt{N} \sum_{i=1}^{N} (y_i - \bar{y})^3}{\left[ \sum_{i=1}^{N} (y_i - \bar{y})^2 \right]^{3/2}}$

and

8) $KU = \frac{N \sum_{i=1}^{N} (y_i - \bar{y})^4}{\left[ \sum_{i=1}^{N} (y_i - \bar{y})^2 \right]^2}$

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 $N > 25$ and the KU statistic with $N > 50$.

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

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

$$v_i \approx K_i = \log(-\log(1 - \frac{i}{N+1}))$$

9) Mann (MN)

Mann and Fertig (1975) proposed the following test statistic.

For even $N$, that is, $N = 2R$, $R = 1, 2, \ldots$, 

$$MN = \frac{N A_2}{(N-2)A_1}$$
where

\[ A_1 = \sum_{i=R+1}^{N} \frac{Y_i - Y_{i-1}}{Y_i - Y_{i-1}/\sigma} \text{ and } A_2 = \sum_{i=2}^{R} \frac{Y_i - Y_{i-1}}{Y_i - Y_{i-1}/\sigma} \]

We use the approximation

\[ \frac{Y_i - \mu}{\sigma} \approx K_i = \log[\log(1 - \frac{i}{N+1})] \]

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

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

11) Cramér-von Mises

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

\[ CM = \frac{1}{12N} + \sum_{i=1}^{N} \left[ F_0 \left( \frac{y_i - \overline{A}}{\sigma} \right) - \frac{2i-1}{2N} \right]^2 \]
12) LMP

From (3.7),

\[ \text{LMP} = I^{-1}_\theta \left\{ \frac{\partial}{\partial \theta} \right\} \theta^2. \]

3.2. Alternative Distributions

The following alternatives are considered for testing normality

1) Weibull (\( \beta \)), \( \beta = .5 \) and 2

\[ f(x) \propto x^{\beta-1} \exp(-x^\beta) \quad x \geq 0 \]

2) Extreme value

\[ f(x) \propto \exp(x-\exp(x)) \quad -\infty < x < \infty \]

3) Gamma (\( \beta \)), \( \beta = .5, 1, 2, \) and 5

\[ f(x) \propto x^{\beta-1} \exp(-x) \quad x \geq 0 \]

4) Lognormal (\( \beta \)), \( \beta = 1 \)

\[ f(x) \propto x^{-1} \exp\left( -\frac{1}{2} \left( \frac{\log x}{\beta} \right)^2 \right) \quad x > 0 \]

5) Uniform

\[ f(x) = 1 \quad 0 \leq x \leq 1 \]

6) Logistic

\[ f(x) \propto \frac{\exp(x)}{[1+\exp(x)]^2} \quad -\infty < x < \infty \]
7) Double $\chi^2(\beta)$, $\beta = -0.5$ and 1

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

8) Cauchy

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

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

$$f(x) \propto \exp\left(-\frac{1}{2} x^2\right) \quad -\infty < x < \infty$$

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

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

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

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

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

$$f(x) \propto \alpha \exp(-x) + 3\beta \exp(-3x) \quad 0 < x < \infty$$
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 ($\beta$), $\beta = .5$ and 2; Logistic; Cauchy; Double $\chi^2(\beta)$, $\beta = .5$, 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-\(\alpha\) 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.

<table>
<thead>
<tr>
<th>N</th>
<th>$W_1$</th>
<th>$W_2$</th>
<th>$W_3$</th>
<th>$W_4$</th>
<th>$C\chi^2$</th>
<th>$G\chi^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>4.40</td>
<td>3.40*</td>
<td>3.80*</td>
<td>3.55*</td>
<td>7.85*</td>
<td>4.50</td>
</tr>
<tr>
<td>50</td>
<td>4.65</td>
<td>4.45</td>
<td>4.35</td>
<td>4.30</td>
<td>5.05</td>
<td>5.25</td>
</tr>
<tr>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>4.65</td>
<td>3.30*</td>
<td>4.00*</td>
<td>3.70*</td>
<td>7.30*</td>
<td>4.95</td>
</tr>
<tr>
<td>50</td>
<td>4.95</td>
<td>4.25</td>
<td>4.25</td>
<td>4.25</td>
<td>5.10</td>
<td>4.20</td>
</tr>
<tr>
<td>100</td>
<td>4.95</td>
<td>4.75</td>
<td>5.35</td>
<td>4.65</td>
<td>5.05</td>
<td>4.55</td>
</tr>
<tr>
<td>20</td>
<td>4.00*</td>
<td>4.90</td>
<td>4.70</td>
<td>4.10</td>
<td>7.40*</td>
<td>4.65</td>
</tr>
<tr>
<td>50</td>
<td>4.75</td>
<td>4.40</td>
<td>5.15</td>
<td>5.15</td>
<td>5.10</td>
<td>4.85</td>
</tr>
</tbody>
</table>

$\alpha = .05$

<table>
<thead>
<tr>
<th>N</th>
<th>$W_1$</th>
<th>$W_2$</th>
<th>$W_3$</th>
<th>$W_4$</th>
<th>$C\chi^2$</th>
<th>$G\chi^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>9.85</td>
<td>6.85*</td>
<td>7.20*</td>
<td>7.15*</td>
<td>13.90*</td>
<td>8.15*</td>
</tr>
<tr>
<td>50</td>
<td>9.70</td>
<td>9.40</td>
<td>8.95</td>
<td>8.00*</td>
<td>10.10</td>
<td>9.75</td>
</tr>
<tr>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>9.40</td>
<td>7.40*</td>
<td>7.80*</td>
<td>6.65*</td>
<td>14.40*</td>
<td>10.15</td>
</tr>
<tr>
<td>50</td>
<td>9.40</td>
<td>9.25</td>
<td>8.75</td>
<td>7.65*</td>
<td>10.40</td>
<td>9.15</td>
</tr>
<tr>
<td>100</td>
<td>9.65</td>
<td>9.50</td>
<td>9.50</td>
<td>9.15</td>
<td>10.55</td>
<td>10.20</td>
</tr>
<tr>
<td>20</td>
<td>8.75</td>
<td>9.60</td>
<td>9.15</td>
<td>8.45*</td>
<td>11.50*</td>
<td>10.20</td>
</tr>
<tr>
<td>50</td>
<td>9.45</td>
<td>9.55</td>
<td>8.95</td>
<td>9.55</td>
<td>11.00</td>
<td>9.40</td>
</tr>
</tbody>
</table>

$\alpha = .10$

$\alpha = .10$

$\alpha = .10$

$\alpha = .10$

\*The estimates are not within $\pm 2$ standard errors of the nominal values.
test statistics: generalized Neyman smooth \( (W_k) \), classical chi-square \( (\chi^2) \), and generalized chi square \( (G\chi^2) \). All estimates are based on 2000 simulations. The estimates which are not within \( \pm 2 \) 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 \( \pm 2 \) 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 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 

\[ \text{For any symmetric pdf's, } h(x) \text{ and } g(x), \text{ } h(x) \text{ is said to have a heavier tail than } g(x) \text{ if there exists a constant } c > 0 \text{ such that } h(d) > g(d) \text{ for any } d \in (c, \infty). \]
Table 3.3. Empirical critical* value based on 2000 simulations from normal and extreme value distributions.

<table>
<thead>
<tr>
<th></th>
<th>Normal</th>
<th></th>
<th>Extreme Value</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N = 20</td>
<td>N = 50</td>
<td></td>
<td>N = 20</td>
</tr>
<tr>
<td></td>
<td>α = .05</td>
<td>.10</td>
<td>α = .05</td>
<td>.10</td>
</tr>
<tr>
<td>W_1</td>
<td>3.668</td>
<td>2.683</td>
<td>3.643</td>
<td>2.560</td>
</tr>
<tr>
<td>W_2</td>
<td>5.160</td>
<td>4.091</td>
<td>5.601</td>
<td>4.234</td>
</tr>
<tr>
<td>W_3</td>
<td>7.122</td>
<td>5.683</td>
<td>7.588</td>
<td>5.820</td>
</tr>
<tr>
<td>BS_1</td>
<td>.165</td>
<td>.121</td>
<td>.164</td>
<td>.115</td>
</tr>
<tr>
<td>BS_2</td>
<td>.868</td>
<td>.625</td>
<td>.946</td>
<td>.685</td>
</tr>
<tr>
<td>BS_3</td>
<td>4.207</td>
<td>3.188</td>
<td>4.153</td>
<td>2.976</td>
</tr>
<tr>
<td>BS_4</td>
<td>6.408</td>
<td>4.873</td>
<td>6.125</td>
<td>4.729</td>
</tr>
<tr>
<td>Cx^2</td>
<td>6.5</td>
<td>5.5</td>
<td>14.4</td>
<td>12.4</td>
</tr>
<tr>
<td></td>
<td>(.3876)</td>
<td>(.6364)</td>
<td>(.95)</td>
<td>(.8571)</td>
</tr>
<tr>
<td>KS</td>
<td>.193</td>
<td>.177</td>
<td>.124</td>
<td>.115</td>
</tr>
<tr>
<td>SK L</td>
<td>-.968</td>
<td>-.802</td>
<td>-.650</td>
<td>-.536</td>
</tr>
<tr>
<td>SK U</td>
<td>.926</td>
<td>.759</td>
<td>.621</td>
<td>.532</td>
</tr>
<tr>
<td>KU L</td>
<td>1.745</td>
<td>1.836</td>
<td>2.050</td>
<td>2.135</td>
</tr>
<tr>
<td>KU U</td>
<td>4.666</td>
<td>4.141</td>
<td>4.419</td>
<td>3.951</td>
</tr>
<tr>
<td>SW L</td>
<td>.906</td>
<td>.920</td>
<td>.948</td>
<td>.956</td>
</tr>
<tr>
<td>SW U</td>
<td>.906</td>
<td>.920</td>
<td>.948</td>
<td>.956</td>
</tr>
</tbody>
</table>

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

**Rejection with certainty if Cx^2 is greater than critical value and rejection with probability (·) if Cx^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 ($\beta$), $\beta = 2$ 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 alterna-
tives.
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.

<table>
<thead>
<tr>
<th>sample size</th>
<th>( W_1 )</th>
<th>( W_2 )</th>
<th>( W_3 )</th>
<th>( W_4 )</th>
<th>( CX^2 )</th>
<th>( SW )</th>
<th>( SK )</th>
<th>( KU )</th>
<th>( KS )</th>
<th>( CX^2 )</th>
<th>( BS_2 )</th>
<th>( BS_3 )</th>
<th>( BS_4 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha = .05 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uniform (0, 1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>2.2</td>
<td>16.3</td>
<td>11.4</td>
<td>11.0</td>
<td>12.3</td>
<td>19.5</td>
<td>0.0</td>
<td>33.1</td>
<td>10.2</td>
<td>8.2</td>
<td>28.8</td>
<td>8.6</td>
<td>18.4</td>
</tr>
<tr>
<td>50</td>
<td>4.3</td>
<td>70.5</td>
<td>53.8</td>
<td>40.7</td>
<td>35.4</td>
<td>87.3</td>
<td>0.2</td>
<td>86.9</td>
<td>31.3</td>
<td>19.8</td>
<td>80.7</td>
<td>21.0</td>
<td>48.8</td>
</tr>
<tr>
<td>Double ( \chi^2 ) (-.5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>2.0</td>
<td>6.5</td>
<td>5.5</td>
<td>5.6</td>
<td>8.5</td>
<td>6.7</td>
<td>0.7</td>
<td>12.7</td>
<td>7.2</td>
<td>6.9</td>
<td>11.8</td>
<td>6.9</td>
<td>10.2</td>
</tr>
<tr>
<td>50</td>
<td>2.9</td>
<td>20.9</td>
<td>13.4</td>
<td>10.0</td>
<td>10.2</td>
<td>24.5</td>
<td>0.9</td>
<td>32.9</td>
<td>12.6</td>
<td>9.1</td>
<td>31.6</td>
<td>8.1</td>
<td>18.8</td>
</tr>
<tr>
<td>Logistic</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>12.1</td>
<td>15.1</td>
<td>14.2</td>
<td>14.0</td>
<td>10.0</td>
<td>14.9</td>
<td>11.8</td>
<td>9.3</td>
<td>7.7</td>
<td>11.6</td>
<td>8.4</td>
<td>7.8</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>13.2</td>
<td>23.6</td>
<td>23.6</td>
<td>22.2</td>
<td>12.1</td>
<td>14.4</td>
<td>19.5</td>
<td>19.3</td>
<td>12.2</td>
<td>8.0</td>
<td>22.7</td>
<td>9.9</td>
<td>10.8</td>
</tr>
<tr>
<td>Laplace</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>21.2</td>
<td>32.5</td>
<td>30.6</td>
<td>28.9</td>
<td>20.4</td>
<td>27.7</td>
<td>26.0</td>
<td>24.3</td>
<td>21.3</td>
<td>15.1</td>
<td>28.2</td>
<td>17.1</td>
<td>18.9</td>
</tr>
<tr>
<td>50</td>
<td>22.0</td>
<td>60.1</td>
<td>57.1</td>
<td>54.8</td>
<td>38.7</td>
<td>40.6</td>
<td>35.8</td>
<td>47.1</td>
<td>42.2</td>
<td>27.1</td>
<td>61.9</td>
<td>30.5</td>
<td>46.6</td>
</tr>
<tr>
<td>Double ( \chi^2 ) (2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>33.8</td>
<td>58.2</td>
<td>56.5</td>
<td>56.5</td>
<td>45.4</td>
<td>52.8</td>
<td>40.6</td>
<td>45.0</td>
<td>46.4</td>
<td>40.1</td>
<td>54.6</td>
<td>36.0</td>
<td>47.8</td>
</tr>
<tr>
<td>50</td>
<td>41.5</td>
<td>91.6</td>
<td>89.9</td>
<td>90.2</td>
<td>83.8</td>
<td>81.9</td>
<td>61.8</td>
<td>82.4</td>
<td>85.6</td>
<td>75.1</td>
<td>92.8</td>
<td>73.8</td>
<td>87.9</td>
</tr>
<tr>
<td>Cauchy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>70.1</td>
<td>89.7</td>
<td>88.8</td>
<td>89.3</td>
<td>47.9</td>
<td>88.1</td>
<td>77.3</td>
<td>83.8</td>
<td>84.5</td>
<td>80.6</td>
<td>88.9</td>
<td>79.2</td>
<td>85.6</td>
</tr>
<tr>
<td>Sample Size</td>
<td>$W_1$</td>
<td>$W_2$</td>
<td>$W_3$</td>
<td>$W_4$</td>
<td>$\chi^2$</td>
<td>SW</td>
<td>SK</td>
<td>KU</td>
<td>KS</td>
<td>$\chi^2$</td>
<td>BS_2</td>
<td>BS_3</td>
<td>BS_4</td>
</tr>
<tr>
<td>-------------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>----------</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>----------</td>
<td>------</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td>Uniform (0, 1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>5.8</td>
<td>30.6</td>
<td>22.3</td>
<td>20.1</td>
<td>25.0</td>
<td>34.8</td>
<td>1.7</td>
<td>47.0</td>
<td>17.8</td>
<td>16.9</td>
<td>48.6</td>
<td>16.2</td>
<td>30.1</td>
</tr>
<tr>
<td>50</td>
<td>8.5</td>
<td>85.5</td>
<td>75.3</td>
<td>62.4</td>
<td>53.0</td>
<td>96.3</td>
<td>0.7</td>
<td>91.9</td>
<td>44.4</td>
<td>29.8</td>
<td>89.5</td>
<td>37.9</td>
<td>61.5</td>
</tr>
<tr>
<td>Double $\chi^2$ (-.5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>5.8</td>
<td>14.9</td>
<td>12.6</td>
<td>12.3</td>
<td>16.2</td>
<td>14.7</td>
<td>2.0</td>
<td>21.0</td>
<td>14.8</td>
<td>13.0</td>
<td>24.5</td>
<td>12.1</td>
<td>18.6</td>
</tr>
<tr>
<td>50</td>
<td>6.6</td>
<td>38.3</td>
<td>27.8</td>
<td>20.8</td>
<td>18.7</td>
<td>40.2</td>
<td>1.7</td>
<td>46.0</td>
<td>22.4</td>
<td>15.6</td>
<td>49.2</td>
<td>18.0</td>
<td>31.0</td>
</tr>
<tr>
<td>Logistic</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>18.4</td>
<td>20.3</td>
<td>20.2</td>
<td>20.3</td>
<td>16.2</td>
<td>16.0</td>
<td>21.0</td>
<td>18.4</td>
<td>15.2</td>
<td>13.2</td>
<td>19.7</td>
<td>13.1</td>
<td>13.6</td>
</tr>
<tr>
<td>50</td>
<td>19.7</td>
<td>32.4</td>
<td>32.3</td>
<td>32.0</td>
<td>20.1</td>
<td>21.7</td>
<td>27.3</td>
<td>33.2</td>
<td>20.3</td>
<td>14.4</td>
<td>31.3</td>
<td>18.2</td>
<td>19.5</td>
</tr>
<tr>
<td>Laplace</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>28.7</td>
<td>39.2</td>
<td>37.8</td>
<td>38.8</td>
<td>29.8</td>
<td>35.1</td>
<td>33.8</td>
<td>34.6</td>
<td>30.9</td>
<td>25.2</td>
<td>36.2</td>
<td>25.4</td>
<td>29.4</td>
</tr>
<tr>
<td>50</td>
<td>30.9</td>
<td>68.7</td>
<td>66.1</td>
<td>66.5</td>
<td>50.0</td>
<td>50.1</td>
<td>43.2</td>
<td>61.2</td>
<td>53.9</td>
<td>38.1</td>
<td>69.7</td>
<td>44.6</td>
<td>57.8</td>
</tr>
<tr>
<td>Double $\chi^2$ (2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>42.2</td>
<td>67.7</td>
<td>64.1</td>
<td>67.7</td>
<td>57.5</td>
<td>60.7</td>
<td>50.3</td>
<td>56.4</td>
<td>58.4</td>
<td>51.1</td>
<td>62.1</td>
<td>46.0</td>
<td>59.4</td>
</tr>
<tr>
<td>50</td>
<td>50.3</td>
<td>95.1</td>
<td>93.8</td>
<td>94.4</td>
<td>88.3</td>
<td>87.9</td>
<td>67.0</td>
<td>90.0</td>
<td>90.4</td>
<td>81.2</td>
<td>95.2</td>
<td>86.0</td>
<td>93.5</td>
</tr>
<tr>
<td>Cauchy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>74.0</td>
<td>92.0</td>
<td>91.2</td>
<td>92.0</td>
<td>61.3</td>
<td>90.1</td>
<td>80.5</td>
<td>88.2</td>
<td>89.2</td>
<td>85.3</td>
<td>91.0</td>
<td>83.8</td>
<td>89.2</td>
</tr>
</tbody>
</table>

*2000 simulations are used for samples from the double $\chi^2$ (-.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.

<table>
<thead>
<tr>
<th>Sample size</th>
<th>W1</th>
<th>W2</th>
<th>W3</th>
<th>W4</th>
<th>GΧ²</th>
<th>SW</th>
<th>SK</th>
<th>KU</th>
<th>KS</th>
<th>GΧ²</th>
<th>BS₂</th>
<th>BS₃</th>
<th>BS₄</th>
</tr>
</thead>
<tbody>
<tr>
<td>α = .05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extreme value (.5)</td>
<td>20</td>
<td>33.8</td>
<td>30.4</td>
<td>27.1</td>
<td>25.0</td>
<td>16.2</td>
<td>32.1</td>
<td>31.3</td>
<td>19.1</td>
<td>20.6</td>
<td>11.5</td>
<td>16.4</td>
<td>25.8</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>72.7</td>
<td>63.9</td>
<td>56.0</td>
<td>56.6</td>
<td>37.9</td>
<td>67.5</td>
<td>68.9</td>
<td>31.8</td>
<td>43.2</td>
<td>25.9</td>
<td>30.9</td>
<td>55.1</td>
</tr>
<tr>
<td>Weibull (2)</td>
<td>20</td>
<td>16.0</td>
<td>13.1</td>
<td>11.1</td>
<td>11.3</td>
<td>10.1</td>
<td>16.0</td>
<td>15.2</td>
<td>8.8</td>
<td>10.2</td>
<td>6.5</td>
<td>7.7</td>
<td>13.5</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>43.1</td>
<td>35.9</td>
<td>26.7</td>
<td>26.9</td>
<td>16.1</td>
<td>43.8</td>
<td>40.6</td>
<td>12.5</td>
<td>23.6</td>
<td>13.1</td>
<td>11.5</td>
<td>30.1</td>
</tr>
<tr>
<td>Weibull (.5)</td>
<td>20</td>
<td>98.9</td>
<td>98.8</td>
<td>98.3</td>
<td>98.9</td>
<td>97.6</td>
<td>99.8</td>
<td>97.9</td>
<td>77.7</td>
<td>98.4</td>
<td>95.9</td>
<td>84.8</td>
<td>98.9</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>99.5</td>
<td>98.7</td>
<td>97.8</td>
<td>97.5</td>
<td>94.0</td>
<td>99.6</td>
<td>98.7</td>
<td>75.3</td>
<td>92.8</td>
<td>91.6</td>
<td>80.9</td>
<td>97.0</td>
</tr>
<tr>
<td>Lognormal (0, .5)</td>
<td>20</td>
<td>77.1</td>
<td>72.7</td>
<td>65.8</td>
<td>68.5</td>
<td>50.7</td>
<td>79.2</td>
<td>73.5</td>
<td>44.5</td>
<td>57.6</td>
<td>39.7</td>
<td>43.8</td>
<td>70.6</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>99.5</td>
<td>98.7</td>
<td>97.8</td>
<td>97.5</td>
<td>94.0</td>
<td>99.6</td>
<td>98.7</td>
<td>75.3</td>
<td>92.8</td>
<td>91.6</td>
<td>80.9</td>
<td>97.0</td>
</tr>
<tr>
<td>Lognormal (0, 1)</td>
<td>20</td>
<td>92.0</td>
<td>90.1</td>
<td>85.9</td>
<td>87.5</td>
<td>74.8</td>
<td>94.1</td>
<td>87.1</td>
<td>57.9</td>
<td>78.7</td>
<td>65.3</td>
<td>61.3</td>
<td>87.8</td>
</tr>
<tr>
<td>Gamma (β), β = .5</td>
<td>20</td>
<td>95.0</td>
<td>93.3</td>
<td>89.2</td>
<td>93.4</td>
<td>83.6</td>
<td>98.8</td>
<td>89.7</td>
<td>55.5</td>
<td>87.3</td>
<td>72.8</td>
<td>60.1</td>
<td>92.3</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>76.8</td>
<td>71.8</td>
<td>65.3</td>
<td>70.6</td>
<td>50.8</td>
<td>83.7</td>
<td>69.5</td>
<td>34.4</td>
<td>56.7</td>
<td>36.6</td>
<td>35.7</td>
<td>69.9</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>54.9</td>
<td>47.1</td>
<td>40.6</td>
<td>42.6</td>
<td>28.3</td>
<td>53.6</td>
<td>48.5</td>
<td>25.6</td>
<td>42.4</td>
<td>18.7</td>
<td>24.4</td>
<td>45.7</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>26.0</td>
<td>22.4</td>
<td>20.5</td>
<td>21.2</td>
<td>14.9</td>
<td>25.5</td>
<td>25.8</td>
<td>14.2</td>
<td>16.0</td>
<td>10.0</td>
<td>12.4</td>
<td>20.4</td>
</tr>
</tbody>
</table>
Table 3.5 (Continued)

<table>
<thead>
<tr>
<th>sample size</th>
<th>$W_1$</th>
<th>$W_2$</th>
<th>$W_3$</th>
<th>$W_4$</th>
<th>$G_X^2$</th>
<th>SW</th>
<th>SK</th>
<th>KU</th>
<th>KS</th>
<th>$C_X^2$</th>
<th>BS$_2$</th>
<th>BS$_3$</th>
<th>BS$_4$</th>
</tr>
</thead>
</table>

$\alpha = .10$

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

*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.

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

<table>
<thead>
<tr>
<th>sample size</th>
<th>W1</th>
<th>W2</th>
<th>W3</th>
<th>W4</th>
<th>(G^2)</th>
<th>SB</th>
<th>MN</th>
<th>LMP</th>
<th>CM</th>
<th>KS</th>
<th>(C^2)</th>
<th>BS2</th>
<th>BS3</th>
<th>BS4</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\alpha = .10)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td>20</td>
<td>31.7</td>
<td>30.9</td>
<td>28.6</td>
<td>27.0</td>
<td>22.2</td>
<td>22.9</td>
<td>33.9</td>
<td>36.7</td>
<td>31.5</td>
<td>25.2</td>
<td>22.4</td>
<td>28.9</td>
<td>36.5</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>69.5</td>
<td>69.1</td>
<td>61.6</td>
<td>62.3</td>
<td>47.6</td>
<td>44.8</td>
<td>66.8</td>
<td>76.3</td>
<td>63.0</td>
<td>53.0</td>
<td>39.2</td>
<td>54.8</td>
<td>70.5</td>
</tr>
<tr>
<td>Logistic</td>
<td>20</td>
<td>38.4</td>
<td>36.9</td>
<td>35.3</td>
<td>31.1</td>
<td>26.0</td>
<td>28.6</td>
<td>33.8</td>
<td>39.4</td>
<td>37.8</td>
<td>32.4</td>
<td>28.4</td>
<td>36.1</td>
<td>39.5</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>62.9</td>
<td>79.6</td>
<td>74.8</td>
<td>75.3</td>
<td>60.6</td>
<td>53.5</td>
<td>71.4</td>
<td>85.2</td>
<td>72.9</td>
<td>65.9</td>
<td>50.5</td>
<td>76.0</td>
<td>79.6</td>
</tr>
<tr>
<td>Gamma ((\beta))</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\beta = .5)</td>
<td>20</td>
<td>18.3</td>
<td>15.3</td>
<td>14.8</td>
<td>13.6</td>
<td>11.2</td>
<td>15.1</td>
<td>15.7</td>
<td>18.2</td>
<td>12.2</td>
<td>11.5</td>
<td>10.5</td>
<td>11.2</td>
<td>13.3</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>25.9</td>
<td>23.2</td>
<td>21.9</td>
<td>19.3</td>
<td>15.8</td>
<td>18.9</td>
<td>22.9</td>
<td>25.6</td>
<td>18.5</td>
<td>17.4</td>
<td>11.3</td>
<td>12.4</td>
<td>20.1</td>
</tr>
<tr>
<td>(\beta = 2)</td>
<td>20</td>
<td>8.5</td>
<td>10.5</td>
<td>10.3</td>
<td>9.1</td>
<td>9.4</td>
<td>8.6</td>
<td>11.9</td>
<td>9.9</td>
<td>10.6</td>
<td>10.3</td>
<td>11.9</td>
<td>12.1</td>
<td>11.0</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>15.0</td>
<td>14.2</td>
<td>14.0</td>
<td>16.2</td>
<td>14.3</td>
<td>6.6</td>
<td>14.8</td>
<td>17.6</td>
<td>17.0</td>
<td>15.6</td>
<td>12.6</td>
<td>15.0</td>
<td>18.1</td>
</tr>
<tr>
<td>(\beta = 5)</td>
<td>20</td>
<td>15.0</td>
<td>13.7</td>
<td>13.7</td>
<td>12.7</td>
<td>11.8</td>
<td>10.1</td>
<td>16.6</td>
<td>15.5</td>
<td>14.3</td>
<td>12.9</td>
<td>13.5</td>
<td>15.8</td>
<td>17.7</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>32.4</td>
<td>27.8</td>
<td>22.4</td>
<td>23.7</td>
<td>20.3</td>
<td>11.3</td>
<td>30.0</td>
<td>26.7</td>
<td>24.1</td>
<td>17.4</td>
<td>23.6</td>
<td>33.0</td>
<td>25.2</td>
</tr>
<tr>
<td>Weibull</td>
<td>20</td>
<td>9.4</td>
<td>10.5</td>
<td>8.7</td>
<td>9.0</td>
<td>12.6</td>
<td>5.2</td>
<td>12.1</td>
<td>9.4</td>
<td>11.6</td>
<td>11.2</td>
<td>13.9</td>
<td>13.8</td>
<td>12.1</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>22.9</td>
<td>19.4</td>
<td>14.2</td>
<td>14.6</td>
<td>14.7</td>
<td>4.8</td>
<td>21.8</td>
<td>20.1</td>
<td>16.9</td>
<td>17.1</td>
<td>13.4</td>
<td>15.0</td>
<td>21.4</td>
</tr>
<tr>
<td>Weibull mixture</td>
<td>20</td>
<td>24.0</td>
<td>34.9</td>
<td>32.7</td>
<td>33.1</td>
<td>22.7</td>
<td>32.9</td>
<td>18.0</td>
<td>29.9</td>
<td>28.8</td>
<td>25.1</td>
<td>23.0</td>
<td>28.8</td>
<td>23.8</td>
</tr>
<tr>
<td>Exponential mixture</td>
<td>20</td>
<td>9.0</td>
<td>9.4</td>
<td>8.5</td>
<td>8.7</td>
<td>9.7</td>
<td>7.9</td>
<td>9.3</td>
<td>9.7</td>
<td>10.1</td>
<td>11.2</td>
<td>11.8</td>
<td>10.0</td>
<td>8.0</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>15.7</td>
<td>13.7</td>
<td>11.2</td>
<td>11.9</td>
<td>13.1</td>
<td>7.9</td>
<td>13.9</td>
<td>15.5</td>
<td>14.3</td>
<td>13.2</td>
<td>12.0</td>
<td>12.0</td>
<td>17.4</td>
</tr>
<tr>
<td>Transl. Gamma (5)</td>
<td>100</td>
<td>57.0</td>
<td>53.1</td>
<td>44.4</td>
<td>44.9</td>
<td>30.8</td>
<td>19.2</td>
<td>50.8</td>
<td>64.5</td>
<td>43.4</td>
<td>37.4</td>
<td>25.8</td>
<td>37.8</td>
<td>50.4</td>
</tr>
</tbody>
</table>

\footnote{2000 simulations are used for samples from the Gamma (\(\beta\)), \(\beta = .5, 2,\) 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 $\chi^2$ 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

\[ x_1 = \log\text{-concentration of EI-43064 in parts/million,} \]
\[ x_2 = \log\text{-grams weight of the fish, and} \]
\[ y = \log(1000/\text{minutes survival}). \quad N = 45. \quad \text{(Bliss, 1967, p. 352).} \]

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

**Ex. 1.** Consider the negative exponential cdf,

\[ F(y|\lambda) = F\left(\frac{y}{\sigma}\right) = 1 - e^{-y/\sigma} \]

where \(\lambda = \sigma\). Then the MLE

\[ \hat{\lambda} = \frac{1}{N} \sum_{i=1}^{N} y_i = 1.288608. \]

From (2.6) and (2.15) with the \(10^\theta\) matrix in Table 2.4, the test statistics, \(W_k, \quad k = 1, \ldots, 4,\) are evaluated as

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

Since \(W_k > \chi^2_{0.995}(k)\) 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,

\[ F(y|\lambda) = 1 - \exp[-\exp\left(\frac{y-\mu}{\sigma}\right)] \]

where \( \lambda = (\beta_0, \sigma)' = (\mu, \sigma)' \). The MLE \( \hat{\lambda} \) 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

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

Using (2.6) and (2.15) with the \( I^0 \) 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,

\[ F(y|x, \lambda) = 1 - \exp[-\exp(y - \beta_0 \cdots - \beta_M^M x^M)], \]

where \( \lambda = \beta = (\beta_0, \beta_1, \ldots, \beta_M)' \). The equations,

\[
\sum_{i=1}^{N} \exp[y_i - \beta_0 \cdots - \beta_M^M x^M_i] = N
\]

and

\[
\sum_{i=1}^{N} x^j_i \exp[y_i - \beta_0 \cdots - \beta_M^M x^M_i] = 0, \quad j = 1, \ldots, M
\]

yield

\( \hat{\lambda}' = (0.4606657, 0.3746265) \) for \( M = 1 \),

and

\( \hat{\lambda}' = (0.7153553, 0.374289, -0.3988452) \) for \( M = 2 \).

The test statistics \( W_k, k = 1, 2, 3, 4, \) and

\( W' = (36.03, 44.38, 87.39, 88.35) \) for \( M = 1 \)

and

\( W' = (36.47, 44.92, 90.21, 90.63) \) for \( M = 2 \).

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
\[ F(y|\mathbf{x}, \lambda) = 1 - \exp[-\exp(\frac{y - \beta_0 \cdots - \beta M^x M}{\sigma})], \]

where \( \lambda = (\frac{\hat{\beta}}{\sigma}) \) and \( x_q = x^q_1 \) for \( q = 1, \ldots, M \). The equations,

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

and

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

yield

\[
\hat{\lambda} = (0.3463413, 0.4513125, 0.1152703) \text{ for } M = 1
\]

and

\[
\hat{\lambda} = (0.723785, 0.3859209, 0.3771623, 0.1017292) \text{ for } M = 2.
\]

The test statistics \( W_k \), \( k = 1, 2, 3, 4 \), are

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

and

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

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 \( \alpha = 0.10 \).

**Ex. 5.** For the normal distribution model with \( \lambda = \begin{pmatrix} \mu \\ \sigma \end{pmatrix} \),

\[
\hat{\lambda} = \begin{pmatrix} \bar{y} \\ \frac{1}{N} \sum_{i=1}^{N} (y_i - \bar{y})^2 / N \end{pmatrix} = \begin{pmatrix} 1.2347 \\ 0.396154291 \end{pmatrix}.
\]

The test statistics \( W_k, k = 1, 2, 3, 4 \), are

\( W' = (0.41, 0.52, 0.73, 0.99) \).

The null hypothesis in (2.3) is not rejected at \( \alpha = 0.10 \) for \( k = 1, 2, 3, 4 \).

**Ex. 6.** For the normal regression model with \( \lambda = \begin{pmatrix} \beta \\ \sigma \end{pmatrix} \),

\[
\hat{\lambda}' = (1.0147, 0.0050757, 0.3855634) \text{ for } M = 1
\]

and

\[
\hat{\lambda}' = (-0.047414, -0.073549, 0.00086627, 0.2355027) \text{ for } M = 2.
\]

And the test statistics \( W_k, k = 1, 2, 3, 4 \), are

\( W' = (0.08, 2.68, 3.15, 3.64) \text{ for } M = 1 \)

and

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

For both \( M = 1 \) and \( 2 \), the null hypothesis in (2.20) is not rejected at \( \alpha = 0.10 \).
BIBLIOGRAPHY


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.


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.


