AN ABSTRACT OF THE THESIS OF

JIIN-CHIANG CHEN for the degree DOCTOR OF PHILOSOPHY
(Name) (Degree)

in STATISTICS presented on August 9, 1974
(Major Department) (Date)

Title: GOODNESS OF FIT TESTS UNDER RANDOM CENSORSHIP

Abstract approved: Redacted for privacy
David R. Thomas

The classical chi-square goodness of fit test and the Neyman smooth tests are generalized for arbitrarily right-censored samples. The test statistics are shown to have limiting chi-square distributions under completely specified (simple) null hypotheses and random censorship. Limiting distributions of the generalized chi-square goodness of fit tests are also established when using minimum chi-square estimation for composite parametric null hypotheses. Tests of fit for negative-exponential and Weibull survival distributions are illustrated for a group of breast cancer patients.

Goodness of Fit Tests Under Random Censorship

by

Jiin-Chiang Chen

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

June 1975

APPROVED:

Redacted for privacy

Associate Professor of Statistics

in charge of major

Redacted for privacy

Chairman of Department of Statistics

Redacted for privacy

Dean of Graduate School

Date thesis is presented August 9, 1974

Typed by Clover Redfern for Jiin-Chiang Chen

ACKNOWLEDGMENT

The author, in particular, wishes to thank Dr. David Thomas for his guidance and constant help during the preparation of this thesis. Also, the author is grateful to Dr. Justus Seely, and Dr. David Birkes for their very useful suggestions, to Billy Chou's help in computation and to his wife, Chris, for her typing and encouragement.

TABLE OF CONTENTS

LIST OF TABLES

GOODNESS OF FIT TESTS UNDER RANDOM CENSORSHIP

1. INTRODUCTION

In life-testing and medical survival studies it is often inconvenient or impossible to make complete measurements of the life-times for all experimental units in the sample. In such situations the observations on life-time may be right-censored. The simplest kind of censoring is that of single censoring which occurs when all observations are censored at the same time. There are two types of single censoring: for Type I censoring the censoring time is predetermined and for Type II censoring the censoring time corresponds to a predetermined ordered failure and is therefore random. In many studies observations are not censored at the same time, which is frequently referred to as the case of arbitrarily censored data.

In single censoring cases, Barr and Davidson (1973) developed appropriate tables for the Kolmogorov-Smirnov statistic for both types of single censoring. Lurie and Hartley (1972) have also generalized the Hartley and Pfaffenberger (1972) criterion for the case of Type II censoring and a simple hypothesis.

In this thesis, the problem of testing goodness of fit for arbitrarily censored data is considered. We generalize the classical goodness of fit test statistic for both the simple and composite hypothesis cases and the Neyman smooth goodness of fit test statistic for the simple hypothesis case. For studying large sample distribution properties of the generalized goodness of fit test statistics, we assume random censorship.

For the random censorship model the random sample of failure times $X_1, \dots, X_n$ are censored by a corresponding random sample of censoring times $T_1, \dots, T_n$. That is, if we define $Y_j = \min(X_j, T_j)$ and $\delta_j = 1(0)$ for $Y_j = X_j (Y_j = T_j)$, $j = 1, \dots, n$, then only the random variables $(Y_1, \dots, Y_n, \delta_1, \dots, \delta_n)$ may be observed. If we let F_θ and H denote the right-sided distribution functions for random variables X and T, respectively, and $0 = a_0 < a_1 < \dots < a_r < a_{r+1} = \infty$ the partition points on the real line, then the generalized statistic

$$n \sum_{i=1}^{r} \frac{\left(\frac{\hat{F}(a_{i-1})}{F_\theta(a_{i-1})} - \frac{\hat{F}(a_i)}{F_\theta(a_i)} \right)^2}{\hat{D}_i}$$

is shown to have a limiting chi-square distribution with r degrees of freedom, where

$$\hat{D}_i = \int_{a_{i-1}}^{a_i} \frac{1}{\hat{H}(t) F_\theta^2(t)} (-dF_\theta(t)),$$

$\hat{F}$ and $\hat{H}$ are product-limit estimators developed by Kaplan and Meier (1958). In the composite hypothesis case, we show that if θ

is replaced by a minimum chi-square estimator $\hat{\theta}$ in the statistic given above, then it will have a limiting chi-square distribution with $r-s$ degrees of freedom, where s is the number of parameters estimated from the sample.

In Chapter 2, some basic probability relationships are given and certain lemmas are proved for later reference. In Chapter 3, the product-limit estimation method is illustrated and the chi-square goodness of fit test statistics are developed. The generalized goodness of fit tests for exponential and Weibull distributions are illustrated for a sample of survival times for breast cancer patients. In Chapter 4, Neyman's smooth goodness of fit test for a simple hypothesis is generalized for the random censorship case. The limiting distribution of the generalized statistic is also considered for a sequence of alternative hypotheses.

2. PROBABILITY BACKGROUND AND NOTATION

We shall use some notation and results given by Mann and Wald (1943) for stochastic limits. Let $\{a_n\}$ be a sequence of k-dimensional vectors, $\{Z_n\}$ a sequence of k-dimensional random vectors, $\{q_n\}$ a sequence of positive functions of n, and define

$a_n = o(q_n)$ if $\lim_{n \to \infty} \|a_n/q_n\| = 0$

$a_n = O(q_n)$ if $\|a_n/q_n\| < M$ for all n, where M is some positive constant

$Z_n = o_p(q_n)$ if Z_n/q_n converges to zero in probability. More precisely, given $\epsilon > 0$ and $\delta > 0$, there exists $N_{\epsilon,\delta}$ such that for $n > N_{\epsilon,\delta}$,
$P(\|Z_n/q_n\| \leq \delta) > 1 - \epsilon$.

$Z_n = O_p(q_n)$ if for given $\epsilon > 0$, there exists M_ϵ such that for all n, $P(\|Z_n/q_n\| \leq M_\epsilon) > 1 - \epsilon$.

Mann and Wald (1943) have shown that all the ordinary operation rules regarding O and o are also applicable to O_p and o_p. For example, if $Z_n = O_p(n^{1/2})$ and $Y_n = o_p(1)$, or $Z_n = O_p(1)$ and $Y_n = o_p(n^{1/2})$, then $Z_n' Y_n = o_p(n^{1/2})$.

Some basic probability results are listed below for later reference.

If Z is a random variable and independent of n, then $Z = O_p(1)$. (2.1)

If Z_n converges in distribution, then $Z_n + o_p(1)$ converges to the same distribution as Z_n. (2.2)

If Z_n converges to Z in probability ($Z_n \xrightarrow{p} Z$) and g is a continuous function, then $g(Z_n)$ converges to $g(Z)$ in probability. In other words, $Z_n - Z = o_p(1)$ and g continuous imply that $g(Z_n) - g(Z) = o_p(1)$. (2.3)

If Z_n converges to Z in distribution ($Z_n \xrightarrow{d} Z$) and g is a continuous function, then $g(Z_n)$ converges to $g(Z)$ in distribution. (2.4)

Relation (2.1) is immediate, (2.2) was established by Mann and Wald (1943), and (2.3) and (2.4) can be found in Chung (1968, p. 66) and Burrill (1972, p. 291), respectively.

In addition, we need the following lemmas:

Lemma 2.1. If $\{Z_n\}$ is a sequence of k-dimensional random variables which converges in distribution to Z, then $Z_n = O_p(1)$.

Proof. It is sufficient to show for given $\epsilon > 0$, there exist $M_\epsilon > 0$ and N_ϵ such that for $n > N_\epsilon$, $P(\|Z_n\| \leq M_\epsilon) > 1 - \epsilon$.

Because B_n can be chosen such that $P(\|Z_n\| \leq B_n) > 1 - \epsilon$ for $n = 1, \ldots, N_\epsilon$, and $M'_\epsilon = \max(M_\epsilon, B_1, \ldots, B_{N_\epsilon})$, it follows that $P(\|Z_n\| \leq M'_\epsilon) > 1 - \epsilon$ for all n.

Let $\{F_n\}$ and F be the distribution functions corresponding to random variables $\{Y_n = \|Z_n\|\}$ and $Y = \|Z\|$. Since the norm is a continuous function, result (2.3) and the assumption of the Lemma gives $\lim_{n \to \infty} F_n(y) = F(y)$ for all y such that $F(y)$ is continuous. That is, for given $\epsilon > 0$, there exists N_ϵ such that for $n > N_\epsilon$,

$$F(y) - \frac{\epsilon}{2} \leq F_n(y) \leq F(y) + \frac{\epsilon}{2} \quad . \tag{2.5}$$

Also for given $\epsilon > 0$, there exists M_ϵ such that $P(\|Z\| < M_\epsilon) > 1 - \frac{\epsilon}{2}$. Without loss of generality, we can choose M_ϵ to be a continuity point of F. Using (2.5) we than have for $n > N_\epsilon$

$$\begin{aligned} P(\|Z_n\| \leq M_\epsilon) &= F_n(M_\epsilon) \\ &\geq F(M_\epsilon) - \frac{\epsilon}{2} \\ &= P(\|Z\| \leq M_\epsilon) - \frac{\epsilon}{2} \\ &> 1 - \frac{\epsilon}{2} - \frac{\epsilon}{2} \\ &= 1 - \epsilon \; . \end{aligned}$$

<u>Lemma 2.2.</u> Let $\{Z_n\}$ and Z be defined as in Lemma 2.1 and $\{A_n\}$ be a sequence of $s \times s$ random matrices such that $A_n \xrightarrow{p} 0$, then $Z'_n A_n Z_n \xrightarrow{p} 0$.

Proof. Applying Lemma 2.1.,

$Z_n' A_n Z_n = O_p(1) \cdot o_p(1) \cdot O_p(1) = o_p(1)$, the Lemma follows.

Lemma 2.3. Let $\{Z_n\}$ be a sequence of k-dimensional random vectors such that $Z_n \xrightarrow{d} N_k(0, V)$, where V is a positive semi-definite symmetric constant matrix, and V^- be a generalized inverse of V. Then $Z_n' V^- Z_n \xrightarrow{d} \chi^2(s)$, the chi-square distribution with $s = \text{rank}(V)$ degrees of freedom.

Proof. Let the random vector Z be distributed according to $N_k(0, V)$ $(Z \sim N_k(0, V))$, and define the function f from R^k to R^1 by $f(Z) = Z' V^- Z$. Since $f(Z)$ is a polynomial it is a continuous function of Z, which implies, from (2.4), that $f(Z_n) \xrightarrow{d} f(Z)$. A well known result is that if $Z \sim N_k(0, V)$, then $Z' V^- Z \sim \chi^2(s)$, where $s = \text{rank}(V)$. Hence, $f(Z_n) = Z_n' V^- Z_n \xrightarrow{d} \chi^2(s)$.

Lemma 2.4. Let $\{V_n\}$ be a sequence of $s \times s$ random matrices, and V be some $s \times s$ nonsingular constant matrix. If $V_n \xrightarrow{p} V$, then for given $\epsilon > 0$, there exists N_ϵ such that for $n > N_\epsilon$, $P(V_n$ is nonsingular$) > 1 - \epsilon$.

Proof. The $\det(V)$ is a polynomial in the elements of V and is, therefore, a continuous function.

From (2.3), we have $\det(V_n) \xrightarrow{p} \det(V)$. Let $a = \det(V) \neq 0$,

$\delta = \frac{|a|}{2}$ and $\epsilon > 0$ be arbitrarily chosen; then there exists N_ϵ such that for $n > N_\epsilon$,

$$P(|\det(V)-\det(V_n)| \leq \delta) > 1 - \epsilon.$$

Since for $n > N_\epsilon$,

$$\begin{aligned}\{V_n: |\det(V)-\det(V_n)| \leq \delta\} &\subseteq \{V_n: |\det(V)|-|\det(V_n)| \leq \delta\} \\ &= \{V_n: |\det(V_n)| \geq |a| - \frac{|a|}{2}\} \\ &= \{V_n: |\det(V_n)| \geq \frac{|a|}{2}\},\end{aligned}$$

it follows that $P(|\det(V_n)| \geq \frac{|a|}{2}) > 1 - \epsilon$, i.e., $P(V_n$ is nonsingular$) > 1 - \epsilon$.

<u>Lemma 2.5.</u> Let $\{V_n\}$ and V be defined as in Lemma 2.4, then $V_n \xrightarrow{p} V$ implies that $V_n^{-1} \xrightarrow{p} V^{-1}$.

<u>Proof</u>. Since $V_n \xrightarrow{p} V$, for any chosen $\epsilon > 0$ and $\eta > 0$ there exists $N'_{\epsilon,\eta}$ such that for $n > N'_{\epsilon,\eta}$, $P(A_n) > 1 - \frac{\epsilon}{2}$, where $A_n = \{V_n: \|V_n - V\| \leq \eta\}$. Let $B_n = \{V_n: V_n$ is nonsingular$\}$, then from Lemma 2.4, for given ϵ, there exists N''_ϵ such that for $n > N''_\epsilon$, we have $P(B_n) > 1 - \frac{\epsilon}{2}$. Hence, for $n > \max(N'_{\epsilon,\eta}, N''_\epsilon)$, we obtain

$$P(A_n \cap B_n) \geq P(A_n) - P(B_n^c)$$

$$\geq 1 - \frac{\epsilon}{2} - \frac{\epsilon}{2}$$

$$= 1 - \epsilon ,$$

where the first inequality can be easily shown by using the relation

$$P(A_n) = P(A_n \cap B_n) + P(A_n \cap B_n^c) .$$

Since V_n^{-1} is a continuous function of V_n (provided the inverse exists), for given $\delta > 0$ there will exist η such that $\|V_n - V\| \leq \eta$ implies $\|V_n^{-1} - V^{-1}\| \leq \delta$. Hence

$$A_n \cap B_n = \{V: \|V_n - V\| \leq \eta \text{ and } V_n \text{ is nonsingular}\}$$

$$\subseteq \{V: \|V_n^{-1} - V^{-1}\| \leq \delta \text{ and } V_n \text{ is nonsingular}\} ,$$

which gives

$$P(\|V_n^{-1} - V^{-1}\| \leq \delta) \geq P(\|V_n^{-1} - V^{-1}\| \leq \delta \text{ and } V_n \text{ is nonsingular})$$

$$\geq P(\|V_n - V\| \leq \eta \text{ and } V_n \text{ is nonsingular})$$

$$= P(A_n \cap B_n)$$

$$\geq 1 - \epsilon .$$

In summary, for given $\epsilon > 0$ and $\delta > 0$ there exists $N_{\epsilon,\delta}$ such that for $n > N_{\epsilon,\delta}$, $P(\|V_n^{-1} - V^{-1}\| \leq \delta) > 1 - \epsilon$. This completes the proof.

Lemma 2.6. Let $\{G_n(y)\}$ be a sequence of monotone random functions such that $G_n(y) \xrightarrow{p} G(y)$, where $G(y)$ is continuous over the closed interval $[0, a]$. Then for $\epsilon > 0$ and $\delta > 0$ there exists $N_{\epsilon, \delta}$ such that for $n > N_{\epsilon, \delta}$,

$$P\left(\sup_{0 \le y \le a} |G_n(y) - G(y)| \le \delta \right) > 1 - \epsilon .$$

Proof. The proof will be given for the case where the sequence $\{G_n(y)\}$ is nonincreasing. The proof for the nondecreasing case is similar.

Since the continuous function $G(y)$ is defined on a compact set $[0, a]$, it is uniformly continuous. That is, for given $\delta > 0$ there exists η such that for any two points c, d in $[0, a]$ with $|c-d| \le \eta$, $|G(c)-G(d)| \le \frac{\delta}{2}$. If we define k by $k-1 < \frac{a}{\eta} \le k$, and divide the interval $[0, a]$ evenly into k subintervals with corresponding partition points $0 = b_0 < b_1 < \dots < b_k = a$, then

$$|G(b_{i-1}) - G(b_i)| \le \frac{\delta}{2} \quad \text{for} \quad i = 1, \dots, k.$$

For the assumption $G_n(y) \xrightarrow{p} G(y)$, for given $\epsilon > 0$ and $\delta > 0$ there exists $N_{\epsilon, \delta}$ such that for $n > N_{\epsilon, \delta}$,

$$P\left(|G_n(b_i) - G(b_i)| \le \frac{\delta}{2}, \quad i = 0, \dots, k \right) > 1 - \epsilon .$$

Defining the set $B_{n,\delta} = \{G_n(y): |G_n(b_i)-G(b_i)| < \frac{\delta}{2}, i = 0, \ldots, k\}$, we then obtain for $G_n(y) \epsilon B_{n,\delta}$ and $b_{i-1} \leq y \leq b_i$,

$$\begin{aligned} G_n(y) - G(y) &\leq G_n(b_{i-1}) - G(b_i) \\ &= G_n(b_{i-1}) - G(b_{i-1}) + G(b_{i-1}) - G(b_i) \\ &\leq \frac{\delta}{2} + \frac{\delta}{2} = \delta \end{aligned}$$

and

$$\begin{aligned} G(y) - G_n(y) &\leq G(b_{i-1}) - G_n(b_i) \\ &= G(b_{i-1}) - G(b_i) + G(b_i) - G_n(b_i) \\ &\leq \frac{\delta}{2} + \frac{\delta}{2} = \delta. \end{aligned}$$

Hence, $\sup\limits_{b_{i-1} \leq y \leq b_i} |G_n(y)-G(y)| \leq \delta$ for $i = 1, \ldots, k$, gives

$$\sup_{0 \leq y \leq a} |G_n(y)-G(y)| \leq \max_i \sup_{b_{i-1} \leq y \leq b_i} |G(y)-G(y)| \leq \delta.$$

Therefore, for given $\epsilon > 0$ and $\delta > 0$ there exists $N_{\epsilon,\delta}$ such that for $n > N_{\epsilon,\delta}$, $P(\sup\limits_{0 \leq y \leq a} |G_n(y)-G(y)| \leq \delta) > 1 - \epsilon$.

<u>Lemma 2.7.</u> Let $\{Q_n(y)\}$ and $\{G_n(y)\}$ be sequences of monotone random functions such that $\{G_n(y)\}$ are uniformly bounded, $Q_n(y) \xrightarrow{p} Q(y)$ and $G_n(y) \xrightarrow{p} G(y)$, where $Q(y)$ and $G(y)$ are bounded and continuous. Then, for arbitrary $0 \leq a < b < \infty$,

$$\int_a^b Q_n(y)dG_n(y) \xrightarrow{p} \int_a^b Q(y)dG(y) .$$

Proof. The proof will be given for the case where $Q_n(y)$ is nondecreasing and $G_n(y)$ is nonincreasing. For the other cases, the proof will be similar.

Applying Lemma 2.6, for given $\epsilon > 0$ and $\delta > 0$ there exists a $N'_{\epsilon,\delta}$ such that for $n > N'_{\epsilon,\delta}$, $P(B'_{n,\delta}) > 1 - \frac{\epsilon}{2}$, where

$$B'_{n,\delta} = \{Q_n(y): \sup_{a \leq y \leq b} |Q_n(y) - Q(y)| \leq \frac{\delta}{4M}\} ,$$

and M is some constant such that $|G_n(a) - G_n(b)| < M$, $Q(a) < M$ and $Q(b) < M$. Similarly, for given $\epsilon > 0$ and $\delta > 0$ there exists $N''_{\epsilon,\delta}$ such that for $n > N''_{\epsilon,\delta}$, $P(B''_{n,\delta}) > 1 - \frac{\epsilon}{2}$, where

$$B''_{n,\delta} = \{G_n(y): \sup_{a \leq y \leq b} |G_n(y) - G(y)| \leq \frac{\delta}{8M}\} .$$

Hence, for $n > \max(N'_{\epsilon,\delta}, N''_{\epsilon,\delta})$, we have

$$\begin{aligned} P(B'_{n,\delta} \text{ and } B''_{n,\delta}) &\geq P(B'_{n,\delta}) - P((B''_{n,\delta})^c) \\ &> 1 - \frac{\epsilon}{2} - \frac{\epsilon}{2} \\ &= 1 - \epsilon . \end{aligned}$$

Taking $Q_n(y) \in B'_{n,\delta}$ and $G_n(y) \in B''_{n,\delta}$, we obtain

$$\left|\int_a^b Q_n(y)(-dG_n(y)) - \int_a^b Q(y)(-dG(y))\right|$$

$$\leq \left|\int_a^b (Q_n(y)-Q(y))(-dG_n(y))\right| + \left|\int_a^b Q(y)(-d(G_n(y)-G(y)))\right|$$

$$\leq \int_a^b |Q_n(y)-Q(y)|\,(-dG_n(y))$$

$$+ \left|\left[-Q(y)(G_n(y)-G(y))\right]_a^b + \int_a^b (G_n(y)-G(y))dQ(y)\right|$$

$$\leq \sup_{a \leq y \leq b} |Q_n(y)-Q(y)|\;|G_n(a)-G_n(b)|$$

$$+ |Q(a)|\,|G_n(a)-G(a)| + |Q(b)|\,|G_n(b)-G(b)|$$

$$+ \sup_{a \leq y \leq b} |G_n(y)-G(y)|\,|Q(b)-Q(a)|$$

$$\leq \frac{\delta}{4M}\; M + 2M\frac{\delta}{8M} + 2M\frac{\delta}{8M} + \frac{\delta}{8M}\; 2M$$

$$= \frac{\delta}{4} + \frac{\delta}{4} + \frac{\delta}{4} + \frac{\delta}{4}$$

$$= \delta.$$

That is, for given $\epsilon > 0$ and $\delta > 0$ there exists $N_{\epsilon,\delta}$ such that for $n > N_{\epsilon,\delta}$,

$$P\left(\left|\int_a^b Q_n(y)(-dG_n(y)) - \int_a^b Q(y)(-dG(y))\right| \leq \delta\right) > 1 - \epsilon.$$

This concludes the Lemma.

<u>Lemma 2.8.</u> Let $\{Y_n\}$ be a sequence of k-dimensional random vectors and $t = (t_1, \ldots, t_k) \in R^k$. Suppose $E(e^{t'Y_n})$ exists for all n in the neighborhood of $t = 0$, $\lim_{n\to\infty} E(e^{t'Y_n})$ exists, and R_n is some uniformly bounded random function which almost surely converges to zero, then

$$\lim_{n\to\infty} E(e^{t'Y_n+R_n}) = \lim_{n\to\infty} E(e^{t'Y_n}) .$$

<u>Proof</u>. Since $E(e^{t'Y_n+R_n}) = E(e^{t'Y_n+R_n}-e^{t'Y_n}) + E(e^{t'Y_n})$, it is sufficient to show $\lim_{n\to\infty} E(e^{t'Y_n}(e^{R_n}-1)) = 0$.

Applying the Cauchy-Schwarz Inequality, we have

$$\begin{aligned} E(e^{t'Y_n}(e^{R_n}-1)) &\le E(|e^{t'Y_n}(e^{R_n}-1)|) \\ &\le (E(e^{2t'Y_n}))^{1/2}(E(e^{R_n}-1)^2)^{1/2} \end{aligned}$$

and

$$\begin{aligned} E(e^{t'Y_n}(e^{R_n}-1)) &\ge -E(|e^{t'Y_n}(e^{R_n}-1)|) \\ &\ge -(E(e^{2t'Y_n}))^{1/2}(E(e^{R_n}-1)^2)^{1/2} . \end{aligned}$$

Hence, by Dominated Convergence Theorem (Chung, 1968, p. 40),

$$
\begin{aligned}
0 &= -\lim_{n\to\infty} (E(e^{2t'Y_n}))^{1/2} \lim_{n\to\infty} (E(e^{R_n}-1)^2)^{1/2} \\
&\leq \lim_{n\to\infty} (e^{t'Y_n}(e^{R_n}-1)) \\
&\leq \lim_{n\to\infty} (E(e^{2t'Y_n}))^{1/2} \lim_{n\to\infty} (E(e^{R_n}-1)^2)^{1/2} \\
&= 0,
\end{aligned}
$$

which implies $\lim_{n\to\infty} E(e^{t'Y_n}(e^{R_n}-1)) = 0.$

3. CHI-SQUARE GOODNESS OF FIT TESTS UNDER RANDOM CENSORSHIP

3.1. Distribution Structure

Let $X_1, \ldots, X_n$ represent a random sample of n failure times from some unknown right-sided cumulative distribution $F_\theta(x) = \Pr(X \geq x)$ with density $f_\theta(x)$, $x \geq 0$, where θ is the parameter in some s-dimensional parameter space Ω, i.e., $\theta \in \Omega \subseteq R^s$. Note that $F_\theta(x)$ is a left continuous and nonincreasing function of x, $F_\theta(0) = 1$ and $F_\theta(\infty) = 0$. We also assume that a corresponding random sample of censoring times $T_1, \ldots, T_n$, which is independent of $X_1, \ldots, X_n$, come from some other unknown right-sided cumulative distribution function $H(t) = \Pr(T \geq t)$ with density $h(t)$, $t \geq 0$. For the random censorship model, the actual sample observations are $Y_1, \ldots, Y_n$, $\delta_1, \ldots, \delta_n$ where, for $i = 1, \ldots, n$, $Y_i = \min(X_i, T_i)$ and

$$\delta_i = \begin{cases} 0 & \text{if} \quad Y_i = T_i \quad (X_i > T_i) \\ 1 & \text{if} \quad Y_i = X_i \quad (X_i \leq T_i) . \end{cases}$$

The joint distribution of Y and δ can be derived easily as

$$\ell = (f_\theta(y)H(y))^\delta (F_\theta(y)h(y))^{1-\delta} .$$

Hence, the likelihood function for a sample size of n is

$$L = \prod_{i=1}^{n} \ell_i = \prod_{i=1}^{n} (f_\theta(y_i)H(y_i))^{\delta_i}(F_\theta(y_i)h(y_i))^{1-\delta_i} .$$

3.2. The Classical Chi-Square Goodness of Fit Test for Uncensored Data

We consider the problem of testing the hypothesis

$$H_0: X \sim F_\theta(x) , \tag{3.1}$$

where $F_\theta(x)$ belongs to some specified family of distributions $\{F_\theta(x): \theta \in \Omega\}$ and $\theta = (\theta_1, \ldots, \theta_s) \in \Omega \subseteq R^s$. For the uncensored data case, failure times of all sample units are observed, hence the classical chi-square goodness of fit test can be applied to test the hypothesis (3.1).

Let $0 = a_0 < \ldots < a_{r+1} = \infty$ be the partition points on the real line, and the random variables V_i be the number of $X_1, \ldots, X_n$ that fall in the ith interval $(a_{i-1}, a_i]$ for $i = 1, \ldots, r+1$; then the joint density function of $V_1, \ldots, V_{r+1}$, $f_\theta(v_1, \ldots, v_{r+1})$ will be

$$f_\theta(v_1, \ldots, v_{r+1}) = \binom{n}{v_1, \ldots, v_{r+1}} P_1^{v_1}(\theta) \ldots P_{r+1}^{v_{r+1}}(\theta) , \tag{3.2}$$

where

$$\sum_{i=1}^{r+1} P_i(\theta) = 1, \quad \sum_{i=1}^{r+1} v_i = n$$

and $P_i(\theta) = F_\theta(a_{i-1}) - F_\theta(a_i)$, $i = 1, \ldots, r+1$, i.e., $V_1, \ldots, V_{r+1}$ are multinomially distributed. Using the asymptotic normality property of the multinomial distribution of $V_1, \ldots, V_{r+1}$, and supposing the null hypothesis (3.1) is simple (θ simplified), Karl Pearson (1900) introduced the statistic

$$\sum_{i=1}^{r+1} \frac{(V_i - nP_i(\theta))^2}{nP_i(\theta)} \tag{3.3}$$

and established that it has a limiting chi-square distribution with r degrees of freedom. In the case when the null hypothesis (3.1) is composite, θ may be replaced in (3.3) by an estimator $\hat{\theta}$, such as a modified minimum chi-square or equivalently the maximum likelihood estimator derived from (3.2), which is a solution of the equations

$$\sum_{i=1}^{r+1} \frac{V_i - nP_i(\theta)}{P_i(\theta)} \frac{\partial P_i(\theta)}{\partial \theta_j} = 0, \quad j = 1, \ldots, s.$$

Cramer (1966, p. 426) has shown that $\sum_{i=1}^{r+1} \frac{(V_i - nP_i(\hat{\theta}))^2}{nP_i(\hat{\theta})}$ is asymptotical chi-square with $r-s$ degrees of freedom. However, in the case when θ in (3.3) is replaced by the maximum likelihood estimator $\tilde{\theta}$ based on the original random sample $X_1, \ldots, X_n$, Chernoff and Lehmann (1954) proved that

$$\sum_{i=1}^{r+1} \frac{(V_i - nP_i(\tilde{\theta}))^2}{nP_i(\tilde{\theta})} \xrightarrow{d} \chi^2(r-s) + \sum_{j=1}^{s} \lambda_j Z_j^2$$

where $0 < \lambda_j < 1$ and the Z_j's are independent and $N(0, 1)$ distributed for $j = 1, \ldots, s$.

In the next section, a nonparametric estimator for $F_\theta(x)$ is introduced, and in Section 3.4 the chi-square goodness of fit statistics are generalized to random censorship. For the case of a composite null hypothesis we only consider the minimum chi-square parameter estimation procedure for θ.

3.3. The Product-Limit Estimator for Censored Data

The product-limit estimator $\hat{F}$ for the distribution function F_θ was derived by Kaplan and Meier (1958) and was reaffirmed by Efron (1967) as a self-consistent estimator. If we define the sequences of random variables $Y_1, \ldots, Y_n$ and $\delta_1, \ldots, \delta_n$ as in Section 3.1, and assume, without loss of generality, that $Y_1 < \ldots < Y_n$, then the product-limit estimator for the failure distribution $F_\theta(s)$ is defined by

$$\hat{F}(s) = \prod_{r=1}^{k-1} \left(\frac{n-r}{n-r+1}\right)^{\delta_r} \quad \text{if} \quad s \in (Y_{k-1}, Y_k]$$

$$= 0 \quad \text{if} \quad s > Y_n \tag{3.4}$$

$$= 1 \quad \text{if} \quad s \leq Y_1 .$$

The product-limit estimator for the censoring distribution $H(s)$ can be expressed in a similar form

$$\hat{H}(s) = \prod_{r=1}^{k-1} \left(\frac{n-r}{n-r+1}\right)^{1-\delta_r} \quad \text{if} \quad s \in (Y_{k-1}, Y_k]$$

$$= 0 \quad \text{if} \quad s > Y_n \tag{3.5}$$

$$= 1 \quad \text{if} \quad s \leq Y_1 .$$

The following example is given to illustrate the product-limit estimation procedure. For a sample size $n = 10$, suppose five experimental units failed at times 0.5, 1.2, 3.7, 7.8, 11.0 months, and the other five units were censored at 1.0, 3.8, 8.1, 9.2, 11.4 months. From the table below the product-limit estimates of F_θ

i	1	2	3	4	5	6	7	8	9	10
Y_i	0.5	1.0	1.2	3.7	3.8	7.8	8.1	9.2	11.0	11.4
δ_i	1	0	1	1	0	1	0	0	1	0

and H at 4 and 10 months are

$$\hat{F}(4) = \prod_{r=1}^{5} \left(\frac{10-r}{10-r+1} \right)^{\delta_r} = \left(\frac{9}{10}\right)^1 \left(\frac{8}{9}\right)^0 \left(\frac{7}{8}\right)^1 \left(\frac{6}{7}\right)^1 \left(\frac{5}{6}\right)^0 = \frac{27}{40},$$

$$\hat{F}(10) = \prod_{r=1}^{8} \left(\frac{10-r}{10-r+1} \right)^{\delta_r} = \left(\frac{27}{40}\right) \left(\frac{4}{5}\right)^1 \left(\frac{3}{4}\right)^0 \left(\frac{2}{3}\right)^0 = \frac{27}{50},$$

$$\hat{H}(4) = \prod_{r=1}^{5} \left(\frac{10-r}{10-r+1} \right)^{1-\delta_r} = \left(\frac{9}{10}\right)^0 \left(\frac{8}{9}\right)^1 \left(\frac{7}{8}\right)^0 \left(\frac{6}{7}\right)^0 \left(\frac{5}{6}\right)^1 = \frac{20}{27},$$

$$\hat{H}(10) = \prod_{r=1}^{8} \left(\frac{10-r}{10-r+1} \right)^{1-\delta_r} = \left(\frac{20}{27}\right) \left(\frac{4}{5}\right)^0 \left(\frac{3}{4}\right)^1 \left(\frac{2}{3}\right)^1 = \frac{10}{27}.$$

Kaplan and Meier (1958) have shown that $\hat{F}(s)$ is a consistent estimator of $F_\theta(s)$; in addition Efron (1967) stated that under random censorship $n^{1/2}(\hat{F}(s)-F_\theta(s))$, considered as a stochastic process in s, has a limiting normal process, as $n \to \infty$, with mean vector 0 and covariance kernel

$$\Gamma(s,t) = F_\theta(s)F_\theta(t) \int_0^{\min(s,t)} \frac{1}{H(z)F_\theta^2(z)} (-dF_\theta(z)). \tag{3.6}$$

The formal proof of this property for continuous F_θ and H, was given recently by Breslow and Crowley (1974). Hence, defining $\hat{F}' = (\hat{F}(a_1), \ldots, \hat{F}(a_r))$, $F_\theta' = (F_\theta(a_1), \ldots, F_\theta(a_r))$ at the partition points $0 < a_1 < \ldots < a_r < \infty$, and assuming F_θ and H are continuous, we have

$$Z_n = n^{1/2}(\hat{F} - F_\theta) \xrightarrow{d} N_r(0, \Sigma) \tag{3.7}$$

i.e., Z_n has a limiting r-dimensional normal distribution with mean vector 0 and covariance matrix $\Sigma = [\Gamma(a_i, a_j)]$. This is the underlying distribution theory on which the goodness of fit test statistic is developed in the next section.

3.4. Derivation of the Generalized Chi-Square Goodness of Fit Test Statistic

Corresponding to the partition points $0 < a_1 < \ldots < a_r < \infty$, we define the $r \times r$ matrices

$$E = \begin{pmatrix} \frac{1}{F_\theta(a_1)} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{F_\theta(a_r)} \end{pmatrix}, \quad C = \begin{pmatrix} -1 & & & 0 \\ 1 & -1 & & \\ & \ddots & \ddots & \\ 0 & & 1 & -1 \end{pmatrix}$$

and let $\bar{Z}_n = CEZ_n$, $\Sigma^* = CE\Sigma E'C'$, where Z_n and Σ are defined in the last section. From the result (3.7), we have $\bar{Z}_n \xrightarrow{d} N_r(0, \Sigma^*)$, where

$$\overline{Z}_n = \left(n^{1/2}\left(\frac{\hat{F}(a_0)}{F_\theta(a_0)} - \frac{\hat{F}(a_1)}{F_\theta(a_1)}\right), \ldots, n^{1/2}\left(\frac{\hat{F}(a_{r-1})}{F_\theta(a_{r-1})} - \frac{\hat{F}(a_r)}{F_\theta(a_r)}\right)\right)$$

$$\Sigma^* = \begin{pmatrix} D_1 & & 0 \\ & \ddots & \\ 0 & & D_r \end{pmatrix}, \tag{3.8}$$

$$D_i = \int_{a_{i-1}}^{a_i} \frac{1}{H(z)F_\theta^2(z)}(-dF_\theta(z))$$

$$\text{for} \quad i = 1, \ldots, r.$$

We can assume that Σ is a nonsingular matrix; from Lemma 2.3 we obtain $Z_n'\Sigma^{-1}Z_n \xrightarrow{d} \chi^2(r)$. Further, from the following relation

$$\begin{aligned} Z_n'\Sigma^{-1}Z_n &= Z_n'E'C'C'^{-1}E'^{-1}\Sigma^{-1}E^{-1}C^{-1}CEZ_n \\ &= (CEZ_n)'(CE\Sigma EC)^{-1}(CEZ_n) \\ &= \overline{Z}_n'\Sigma^{*-1}\overline{Z}_n \\ &= n\sum_{i=1}^{r} \frac{\left(\frac{\hat{F}(a_{i-1})}{F_\theta(a_{i-1})} - \frac{\hat{F}(a_i)}{F_\theta(a_i)}\right)^2}{D_i}, \end{aligned}$$

we have

$$Z_n'\Sigma^{-1}Z_n = \overline{Z}_n'\Sigma^{*-1}\overline{Z}_n = n\sum_{i=1}^{r} \frac{\left(\frac{\hat{F}(a_{i-1})}{F_\theta(a_{i-1})} - \frac{\hat{F}(a_i)}{F_\theta(a_i)}\right)^2}{D_i} \xrightarrow{d} \chi^2(r). \tag{3.9}$$

If we estimate H by the product-limit estimator $\hat{H}$ defined in (3.5) and assume H is not equal to 0 at all sample points, then by using the consistent property of H and applying Lemma 2.7 (set $Q_n = \frac{1}{\hat{H}F_\theta^2}$, $Q = \frac{1}{HF_\theta^2}$, $G_n = G = F_\theta$) we have

$$\hat{D}_i = \int_{a_{i-1}}^{a_i} \frac{1}{\hat{H}(z)F_\theta^2(z)}(-dF_\theta(z)) \xrightarrow{p} \int_{a_{i-1}}^{a_i} \frac{1}{\hat{H}(z)F_\theta^2(z)}(-dF_\theta(z)) = D_i \tag{3.10}$$

If we let

$$\hat{\Sigma}* = \begin{pmatrix} \hat{D}_1 & & 0 \\ & \ddots & \\ 0 & & \hat{D}_r \end{pmatrix}, \tag{3.11}$$

then (3.10) gives $\hat{\Sigma}* \xrightarrow{p} \Sigma*$. Therefore, by Lemma 2.5, $\hat{\Sigma}*^{-1} = \Sigma*^{-1} + o_p(1)$. The following relations

$$\begin{aligned} \bar{Z}_n' \hat{\Sigma}*^{-1} \bar{Z}_n &= \bar{Z}_n'(\Sigma*^{-1} + o_p(1)) \bar{Z}_n \\ &= \bar{Z}_n' \Sigma*^{-1} \bar{Z}_n + o_p(1)\, O_p(1) \\ &= \bar{Z}_n' \Sigma*^{-1} \bar{Z}_n + o_p(1), \end{aligned}$$

with (2.2) and (3.9) imply that

$$\bar{Z}_n' \Sigma^{*-1} \bar{Z}_n = n \sum_{i=1}^{r} \frac{\left(\frac{\hat{F}(a_{i-1})}{F_\theta(a_{i-1})} - \frac{\hat{F}(a_i)}{F_\theta(a_i)}\right)^2}{\hat{D}_i} \xrightarrow{d} \chi^2(r). \tag{3.12}$$

Hence, for large sample sizes, we can use the statistic specified in (3.12) to test hypothesis (3.1) when θ is known.

3.5. The Composite Null Hypothesis Case

The case of a completely specified life time distribution is very limited in application. More generally, the null hypothesis is composite and the parameter needs to be estimated. In this section we use the minimum chi-square method to estimate θ and will generalize R.A. Fisher's theory (1922, 1924), which gives a reduction of one degree of freedom in the limiting distribution for each parameter estimated from the sample, to the statistic

$$n \sum_{i=1}^{r} \left(\frac{\left(\frac{\hat{F}(a_{i-1})}{F_\theta(a_{i-1})} - \frac{\hat{F}(a_i)}{F_\theta(a_i)}\right)^2}{\hat{D}_i} \right)_\theta . \tag{3.13}$$

We assume the number of parameters to be estimated in the null hypothesis (3.1) is less than the number of partition points, i.e., $s < r$, and let θ^o represent the true value of θ and $\hat{\theta}$ the minimum chi-square estimator for θ. We also use superscripts,

j, k, ℓ to represent partial derivatives with respect to $\theta_j, \theta_k, \theta_\ell$; $j, k, \ell = 1, \ldots, s$. Some further notation is now summarized for later reference. For $i = 1, \ldots, r$ and $j, k, \ell = 1, \ldots, s$;

$$\hat{F}_i = \hat{F}(a_i), \quad \text{with } F(a_i) \text{ defined as in (3.4)} \tag{3.14}$$

$$F_i = F_\theta(a_i) \tag{3.15}$$

$$S_i = \frac{\hat{F}_i}{F_i} \tag{3.16}$$

$$g_n(\theta) = n\left(\sum_{i=1}^{r} \frac{(S_{i-1}-S_i)^2}{\hat{D}_i}\right)_\theta \quad \text{with } \hat{D}_i \text{ defined in (3.10)} \tag{3.17}$$

$$\overline{G}_n'(\theta) = \left(\frac{1}{\sqrt{n}}\, g_n^1(\theta), \ldots, \frac{1}{\sqrt{n}}\, g_n^s(\theta)\right) \tag{3.18}$$

$$G_n(\theta) = \left[\frac{1}{n}\, g_n^{jk}(\theta)\right] \tag{3.19}$$

$$A^j_{\theta,i} = \left(\frac{F_i^j}{F_i} - \frac{F_{i-1}^j}{F_{i-1}}\right)_\theta \tag{3.20}$$

$$A_i = \frac{1}{(D_i^{1/2})_{\theta^o}}\left(A^1_{\theta^o,i}, \ldots, A^s_{\theta^o,i}\right) \tag{3.21}$$

$$A' = (A_1', \ldots, A_r') \tag{3.22}$$

$$K_\theta^{jk} = 2\sum_{i=1}^{r} \frac{1}{(D_i)_\theta}\left(A^j_{\theta,i}\, A^k_{\theta,i}\right) \tag{3.23}$$

$$K_{\theta^o} = [K^{jk}_{\theta^o}] = 2\sum_{i=1}^{r} A_i' A_i = 2A'A \tag{3.24}$$

$$\overline{X}_n' = (\overline{X}_{n1}, \ldots, \overline{X}_{nr}) = \left(\frac{\sqrt{n}(S_0 - S_1)}{D_1^{1/2}}, \ldots, \frac{\sqrt{n}(S_{r-1} - S_r)}{D_r^{1/2}}\right)_{\theta^o} \tag{3.25}$$

$$\overline{Y}_n' = (\overline{Y}_{n1}, \ldots, \overline{Y}_{nr}) = \left(\frac{\sqrt{n}(S_0 - S_1)}{\hat{D}_1^{1/2}}, \ldots, \frac{\sqrt{n}(S_{r-1} - S_r)}{\hat{D}_r^{1/2}}\right)_{\hat{\theta}} . \tag{3.26}$$

From the definition (3.17), we can represent the test statistic (3.13) by $g_n(\theta)$. The minimum chi-square estimator for θ, which minimizes $g_n(\theta)$, will satisfy the equations

$$g_n^j(\theta) = \frac{\partial}{\partial \theta_j} g_n(\theta) = 0, \quad j = 1, \ldots, s. \tag{3.27}$$

In the theorem given below we are going to show, under certain regular conditions, Equation (3.27) possesses a solution which converges to the true parameter θ^o in probability and is asymptotic normally distributed.

<u>Condition R.</u> The distribution function $F_\theta(y)$ is continuous with density $f_\theta(y)$ and satisfies

1) For almost all y, $F_\theta^j(y)$, $F_\theta^{jk}(y)$, $F_\theta^{jk\ell}(y)$, and $f_\theta^j(y)$, $f_\theta^{jk}(y)$, $f_\theta^{jk\ell}(y)$ exist for every θ in the closure of some neighborhood $\omega \subseteq \Omega$ of θ^o for $j, k, \ell = 1, \ldots, s$.

2) F_i^j, F_i^{jk}, $F_i^{jk\ell}$ are bounded for all $\theta \in \omega$ for $j, k, \ell = 1, \ldots, s$, $i = 1, \ldots, r$.

3) The $r \times s$ matrix A defined in (3.22) is of rank s.

Theorem 3.1. Under the condition R Equation (3.27) has a solution $\hat{\theta}$ which converges to true parameter θ^o in probability. In addition

$$\sqrt{n}(\theta-\theta^o) = -(A'A)^{-1}A'\overline{X}_n + o_p(1) ,$$

where $\overline{X}_n$ is defined in (3.25) and has a limiting $N_r(0, I)$ distribution. Hence $\sqrt{n}(\theta-\theta^o)$ has a limiting $N_r(0, (A'A)^{-1})$ distribution.

Proof. The approach used for this proof is in some ways similar to that used by Cramer (1954, p. 500) to show that a likelihood function possesses a solution. The detail of the proof is divided into the following three steps.

Step 1. Auxiliary derivations.

a) $$(S_{i-1}-S_i)^j = \frac{\partial}{\partial\theta_j}\left(\frac{\hat{F}_{i-1}}{F_{i-1}} - \frac{\hat{F}_i}{F_i}\right)$$

$$= -\frac{\hat{F}_{i-1}}{F_{i-1}^2} F_{i-1}^j + \frac{\hat{F}_i}{F_i^2} F_i^j$$

$$= \left(-\frac{F_{i-1}^j}{F_{i-1}} - (\hat{F}_{i-1}-F_{i-1})\frac{F_{i-1}^j}{F_{i-1}^2}\right) + \left(\frac{F_i^j}{F_i} + (\hat{F}_i-F_i)\frac{F_i^j}{F_i^2}\right) .$$

By $((\hat{F}_i - F_i)\frac{F_i^j}{F_i^2})_{\theta^o} = o_p(1)\,O(1) = o_p(1)$, and

$((\hat{F}_i - F_i)\frac{F_i^j}{F_i^2})_{\theta} = O(1)$ for all $\theta \epsilon \omega$, $i = 1, \ldots, r$;

$j = 1, \ldots, s$, we have

$$(S_{i-1} - S_i)^j_\theta = \begin{cases} A^j_{\theta^o, i} + o_p(1) & \text{if } \theta = \theta^o \\ \\ A^j_{\theta, i} + O(1) & \text{if } \theta \epsilon \omega , \end{cases} \tag{3.28}$$

where $A^j_{\theta, i}$ is defined in (3.20).

b) $$(S_{i-1} - S_i)^{jk} = \frac{\partial^2}{\partial\theta_j \partial\theta_k} \left(\frac{\hat{F}_{i-1}}{F_{i-1}} - \frac{\hat{F}_i}{F_i}\right)$$

$$= \frac{\partial}{\partial\theta_k}\left(- \frac{\hat{F}_{i-1}}{F^2_{i-1}} F^j_{i-1} + \frac{\hat{F}_i}{F^2_i} F^j_i\right)$$

$$= \left(2\frac{\hat{F}_{i-1}}{F^3_{i-1}} F^j_{i-1} F^k_{i-1} - \frac{\hat{F}_{i-1}}{F^2_{i-1}} F^{jk}_{i-1}\right)$$

$$+ \left(-2\frac{\hat{F}_i}{F^3_i} F^j_i F^k_i + \frac{\hat{F}_i}{F^2_i} F^{jk}_i\right) .$$

Now, applying the same technique as for $\hat{F}_{i-1}$ and $\hat{F}_i$ in a), we can write

$$(S_{i-1}-S_i)_\theta^{jk} = \begin{cases} B^{jk}_{\theta^o, i} + o_p(1) & \text{if} \quad \theta = \theta^o \\ \\ B^{jk}_{\theta, i} + O(1) & \text{if} \quad \theta \in \omega , \end{cases} \tag{3.29}$$

where $B^{jk}_{\theta, i} = 2\left(\dfrac{F^j_{i-1}F^k_{i-1}}{F^2_{i-1}} - \dfrac{F^j_i F^k_i}{F^2_i} \right)_\theta + \left(\dfrac{F^{jk}_i}{F_i} - \dfrac{F^{jk}_{i-1}}{F_{i-1}} \right)_\theta ,$

for $i = 1, \ldots, r; \quad j, k = 1, \ldots, s.$

c) $$(S_{i-1}-S_i)^{jk\ell} = \frac{\partial^3}{\partial\theta_\ell \partial\theta_k \partial\theta_j} (S_{i-1}-S_i)$$

$$= \frac{\partial^3}{\partial\theta_\ell \partial\theta_k \partial\theta_j} (S_{i-1}) - \frac{\partial^3}{\partial\theta_\ell \partial\theta_k \partial\theta_j} (S_i) ,$$

where $$\frac{\partial^3}{\partial\theta_\ell \partial\theta_k \partial\theta_j} S_i = \Big(-6 \frac{\hat{F}_i}{F^4_i} F^j_i F^k_i F^\ell_i + 2 \frac{\hat{F}_i}{F^3_i} F^{j\ell}_i F^k_i + 2 \frac{\hat{F}_i}{F^3_i} F^j_i F^{k\ell}_i$$

$$+ 2 \frac{\hat{F}_i}{F^3_i} F^{jk}_i F^\ell_i - \frac{\hat{F}_i}{F^2_i} F^{jk\ell}_i \Big) ,$$

for $i = 1, \ldots, r$. By the same approach used in a) and b), we have

$$(S_{i-1}-S_i)_\theta^{jk\ell} = C^{jk\ell}_{\theta, i} + O(1), \quad \text{if} \quad \theta \in \omega , \tag{3.30}$$

where

$$C^{jk\ell}_{\theta, i} = 6\left(\frac{F_i^j F_i^k F_i^\ell}{F_i^3} - \frac{F_{i-1}^j F_{i-1}^k F_{i-1}^\ell}{F_{i-1}^3}\right)_\theta$$

$$+ 2\left(\frac{F_{i-1}^{j\ell} F_{i-1}^k + F_{i-1}^{k\ell} F_{i-1}^j + F_{i-1}^{jk} F_{i-1}^\ell}{F_{i-1}^2} - \frac{F_i^{j\ell} F_i^k + F_i^{k\ell} F_i^j + F_i^{jk} F_i^\ell}{F_i^2}\right)_\theta$$

$$+ \left(\frac{F_i^{jk\ell}}{F_i} - \frac{F_{i-1}^{jk\ell}}{F_{i-1}}\right)_\theta,$$

$$i = 1, \ldots, r; \quad j, k, \ell = 1, \ldots, s.$$

From the condition R it is easy to establish that $(\hat{D}_i)^j$, $(\hat{D}_i)^{jk}$ and $(\hat{D}_i)^{jk\ell}$ are all bounded in probability. Hence, applying (3.28), (3.29) and (3.30), we obtain for $j, k, \ell = 1, \ldots, s$,

a') $$\frac{1}{n} g_n^j(\theta^o) = \frac{1}{n} \frac{\partial}{\partial \theta_j} \left(n \sum_{i=1}^r \frac{(S_{i-1} - S_i)^2}{\hat{D}_i}\right)_{\theta^o}$$

$$= \sum_{i=1}^r \left(2 \frac{(S_{i-1} - S_i)}{\hat{D}_i} (S_{i-1} - S_i)^j - \frac{(S_{i-1} - S_i)^2}{\hat{D}_i^2} (\hat{D}_i)^j\right)_{\theta^o}$$

$$= \sum_{i=1}^r \left(2 \frac{o_p(1)}{(D_i)_{\theta^o} + o_p(1)} (A^j_{\theta^o, i} + o_p(1)) - \frac{o_p(1)}{(D_i)^2_{\theta^o} + o_p(1)} (O_p(1))\right)$$

$$= o_p(1).$$

b') $$\frac{1}{n} g_n^{jk}(\theta^o) = \sum_{i=1}^{r} \Big(\Big[2 \frac{1}{\hat{D}_i} ((S_{i-1}-S_i)^j (S_{i-1}-S_i)^k + (S_{i-1}-S_i)(S_{i-1}-S_i)^{jk})$$

$$- \frac{1}{\hat{D}_i^2} ((S_{i-1}-S_i)(S_{i-1}-S_i)^j (\hat{D}_i)^k) \Big]$$

$$- \Big[2 \frac{1}{\hat{D}_i^2} ((S_{i-1}-S_i)(S_{i-1}-S_i)^k (\hat{D}_i)^j)$$

$$+ \frac{1}{\hat{D}_i^2} ((S_{i-1}-S_i)^2 (\hat{D}_i)^{jk})$$

$$-2 \frac{1}{\hat{D}_i^3} ((S_{i-1}-S_i)^2 (\hat{D}_i)^j (\hat{D}_i)^k) \Big] \Big)_{\theta^o}$$

$$= 2 \sum_{i=1}^{r} \Big(\frac{1}{\hat{D}_i} (S_{i-1}-S_i)^j (S_{i-1}-S_i)^k \Big)_{\theta^o} + o_p(1)$$

$$= 2 \sum_{i=1}^{r} \frac{1}{(D_i)_{\theta^o}} (A^j_{\theta^o, i} \cdot A^k_{\theta^o, i}) + o_p(1)$$

$$= K^{jk}_{\theta^o} + o_p(1) ,$$

where $K^{jk}_{\theta^o}$ is defined in (3.23).

c') By applying the above results and condition R it is straightforward to show

$$\frac{1}{n} g_n^{jk\ell}(\theta) = \frac{1}{n} \frac{\partial}{\partial \theta_\ell} (g_n^{jk}(\theta))$$

$$= O_p(1) \quad \text{for} \quad \theta \in \omega .$$

Hence, in summary, we have

$$\frac{1}{n} g_n^{j}(\theta^o) = o_p(1)$$

$$\frac{1}{n} g_n^{jk}(\theta^o) = K_{\theta^o}^{jk} + o_p(1) \qquad j, k, \ell = 1, \ldots, s. \tag{3.31}$$

$$\frac{1}{n} g_n^{jk\ell}(\theta) = O_p(1) \quad \text{for} \quad \theta \in \omega$$

Step 2. Show there exists a solution $\hat{\theta}$ for Equation (3.27) which converges to θ^o in probability.

Lemma 3.1. For given $\epsilon > 0$ and $\delta > 0$, define the sets $U = \{u : u \in R^s \text{ and } \|u\| = 1\}$, $B = \{\theta = \theta^o + \lambda u : \|\theta^o - (\theta^o + \lambda u)\| \leq \delta\} \subseteq \omega$, and the sequence of functions $f_n(\lambda, u) = \frac{1}{n} g_n(\theta^o + \lambda u)$, where $|\lambda| \leq \delta$. Then under the condition R, there exists $N_{\epsilon,\delta}$ such that for $n > N_{\epsilon,\delta}$

$$P\left(\frac{\partial}{\partial \lambda} f_n(\delta, u) > 0, \ \frac{\partial}{\partial \lambda} f_n(-\delta, u) < 0 \text{ for all } u \in U\right) > 1 - \epsilon .$$

Proof. First we need to know the limiting structure of the 1st, 2nd, 3rd derivatives of the function $f_n(\lambda, u)$.

a'') $$\frac{\partial}{\partial\lambda} f_n(\lambda, u) = \frac{1}{n}\frac{\partial}{\partial\lambda} g_n(\theta^o+\lambda u)$$

$$= \frac{1}{n}\sum_{j=1}^{s} \frac{\partial g_n(\theta^0+\lambda u)}{\partial(\theta^o+\lambda u)} \frac{\partial(\theta^o+\lambda u)}{\partial\lambda}$$

$$= \frac{1}{n}\sum_{j=1}^{s} u_j g_n^j(\theta^o+\lambda u) ,$$

and from (3.31) and $|u_j| \leq 1$ for $j = 1, \ldots, s$, we have for all $u \in U$

$$\frac{\partial}{\partial\lambda} f_n(0, u) = \sum_{j=1}^{s} u_j \frac{1}{n} g_n^j(\theta^o)$$

$$= o_p(1) . \tag{3.32}$$

b'') $$\frac{\partial^2}{\partial\lambda^2} f_n(\lambda, u) = \frac{1}{n}\frac{\partial^2}{\partial\lambda^2} g_n(\theta^o+\lambda u)$$

$$= \frac{1}{n}\sum_{j,k=1}^{s} u_j u_k g_n^{jk}(\theta^o+\lambda u)$$

$$= u' G_n(\theta^o+\lambda u)u$$

where $u = (u_1, \ldots u_s)$ and $G_n(\theta)$ is given in (3.19). Applying the second equation in (3.31) and equation (3.24), we can write

$$G_n(\theta^o) = K_{\theta^o} + o_p(1)$$

$$= 2A'A + o_p(1) , \qquad (3.33)$$

where K_{θ^o} and A are defined from (3.20) through (3.24). Applying condition R 3), we have that $K_{\theta^o} = 2A'A$ is positive definite. Hence,

$$\frac{\partial^2}{\partial\lambda^2} f_n(0, u) = u'G_n(\theta^o)u$$

$$= u'(K_{\theta^o}+o_p(1))u$$

$$= u'K_{\theta^o}u+o_p(1) ,$$

where $u'K_{\theta^o}u > 0$. Let the function h from U to R^1 be defined by $h(u) = u'K_{\theta^o}u$. Since the set U is compact and $h(u)$ is continuous, the set $\{u'K_{\theta^o}u : u \in U\}$ is compact. In particular, we have $u'K_{\theta^o}u > c > 0$ for all $u \in U$. Hence for given $\epsilon > 0$, there exists N_ϵ such that for $n > N_\epsilon$

$$P(\frac{\partial^2}{\partial\lambda^2} f_n(0, u) > \frac{c}{2} \text{ for all } u \in U) > 1 - \epsilon . \qquad (3.34)$$

c'')
$$\frac{\partial^3}{\partial\lambda^3} f(\lambda, u) = \sum_{j,k,\ell=1}^{s} u_j u_k u_\ell \frac{1}{n} g_n^{jk\ell}(\theta^o+\lambda u)$$

$$= O_p(1) . \qquad (3.35)$$

Applying the results (3.32), (3.34) and (3.35), for given $\epsilon > 0$ and $\delta > 0$ there will exist $N_{\epsilon,\delta}$ such that for $n > N_{\epsilon,\delta}$ the following probability statements will hold:

a*) $P(|\frac{\partial}{\partial\lambda} f_n(0,u)| < \delta^2$ for all $u \in U) > 1 - \frac{\epsilon}{3}$ or

$P_1 = P(|\frac{\partial}{\partial\lambda} f_n(0,u)| \geq \delta^2$ for some $u) < \frac{\epsilon}{3}$.

b*) $P(\frac{\partial^2}{\partial\lambda^2} f_n(0,u) > \frac{c}{2}$ for all $u \in U) > 1 - \frac{\epsilon}{3}$ or

$P_2 = P(\frac{\partial^2}{\partial\lambda^2} f_n(0,u) \leq \frac{c}{2}$ for some $u) < \frac{\epsilon}{3}$.

c*) $P(|\frac{\partial^3}{\partial\lambda^3} f_n(\lambda,u)| < 2M_\epsilon$ for all $u \in U$ and $|\lambda| \leq \delta) > 1 - \frac{\epsilon}{3}$

or $P_3 = P(|\frac{\partial^3}{\partial\lambda^3} f_n(\lambda,u)| \geq 2M_\epsilon$ for some u and $|\lambda| \leq \delta) < \frac{\epsilon}{3}$.

Hence, by defining the set S

$$S = \{y : |\frac{\partial}{\partial\lambda} f_n(0,u)| < \delta^2, \frac{\partial^2}{\partial\lambda^2} f_n(0,u) > \frac{c}{2}, |\frac{\partial^3}{\partial\lambda^3} f_n(\lambda,u)| < 2M_\epsilon$$

$$\text{for all } u \in U \text{ and } |\lambda| \leq \delta\},$$

and letting S^c denote the complement of S, then we have

$$P(S^c) \leq P_1 + P_2 + P_3 < \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon,$$

which implies $P(S) > 1 - \epsilon$.

Now we expand $\frac{\partial}{\partial\lambda} f_n(\lambda, u)$ at the point $\lambda = 0$,

$$\frac{\partial}{\partial\lambda} f_n(\lambda, u) = \frac{\partial}{\partial\lambda} f_n(0, u) + (\lambda - 0) \frac{\partial^2}{\partial\lambda^2} f_n(0, u)$$

$$+ \frac{1}{2} (\lambda - 0)^2 \frac{\partial^3}{\partial\lambda^3} f_n(\lambda^*, u) , \tag{3.36}$$

where $0 < \lambda^* < \lambda$. For $y \in S$, $\delta < \frac{c}{2(1+M_\epsilon)}$ and all $u \in U$ the Equation (3.36) will give, for $\lambda = \delta$

$$\frac{\partial}{\partial\lambda} f_n(\delta, u) \geq -\delta^2 + \delta \frac{\partial^2}{\partial\lambda^2} f_n(0, u) + \frac{1}{2} \delta^2(-2M_\epsilon)$$

$$= -(1+M_\epsilon)\delta^2 + \delta \frac{\partial^2}{\partial\lambda^2} f_n(0, u)$$

$$\geq -\frac{c}{2} \delta + \delta \frac{\partial^2}{\partial\lambda^2} f_n(0, u) > 0 ,$$

and for $\lambda = -\delta$,

$$\frac{\partial}{\partial\lambda} f_n(-\delta, u) \leq \delta^2 - \delta \frac{\partial^2}{\partial\lambda^2} f_n(0, u) + \frac{1}{2} \delta^2(2M_\epsilon)$$

$$= (1+M_\epsilon)\delta^2 - \delta \frac{\partial^2}{\partial\lambda^2} f_n(0, u)$$

$$\leq \frac{c}{2} \delta - \delta \frac{\partial^2}{\partial\lambda^2} f_n(0, u) < 0.$$

Hence for $\delta > 0$ and $\epsilon > 0$ there exists $N_{\delta, \epsilon}$ such that for $n > N_{\delta, \epsilon}$,

$$P\left(\frac{\partial}{\partial\lambda} f_n(\delta, u) > 0,\ \frac{\partial}{\partial\lambda} f_n(-\delta, u) < 0 \text{ for all } u \in U\right) > 1 - \epsilon .$$

This completes the proof for Lemma 3.1.

Now we are going to find a point in the ball $B = \{\theta = \theta^o + \lambda u: \|\theta^o - \theta\| = \|\lambda u\| \leq \delta,\ u \in U\}$ which is a solution for Equation (3.27). From the assumption that $g_n(\theta)$ is continuous for $\theta \in \omega$ and $B \subseteq \omega$, it follows $g_n(\theta)$ is continuous in B. This implies there exists $\hat{\theta} = \theta^o + \hat{\lambda}\hat{u}$ such that

$$\frac{1}{n} g_n(\theta^o + \hat{\lambda}\hat{u}) \leq \frac{1}{n} g_n(\theta^o + \lambda u) , \tag{3.37}$$

for all $|\lambda| < \delta$ and $u \in U$. In particular

$$\frac{1}{n} g_n(\theta^o + \hat{\lambda}\hat{u}) \leq \frac{1}{n} g(\theta^o \pm \delta u) = f_n(\pm\delta, u) ,$$

for all $u \in U$. Applying Lemma 3.1, with probability greater than $1 - \epsilon$ there exists $\lambda^*(u)$ such that

$$\frac{1}{n} g_n(\theta^o + \lambda^*(u) u) = f_n(\lambda^*(u), u) < f_n(\pm\delta, u) = \frac{1}{n} g_n(\theta^o \pm \delta u)$$

for every $u \in U$. Hence, by defining the set

$$S^* = \{y: \frac{\partial}{\partial\lambda} f_n(\delta, u) > 0,\ \frac{\partial}{\partial\lambda} f_n(-\delta, u) < 0 \text{ for all } u \in U\} ,$$

and applying the relation (3.37) for $y \in S^*$, we have

$$\frac{1}{n} g_n(\theta^o+\hat{\lambda}\hat{u}) \leq \frac{1}{n} g_n(\theta^o+\lambda^*(u)\ u) < \frac{1}{n} g_n(\theta^o \pm \delta\ u)$$

for all $u \epsilon U$. This implies $\hat{\theta} = \theta^o + \hat{\lambda}\hat{u}$ is an interior point of B and is a local minimum. Therefore, we have $g_n^j(\hat{\theta}) = 0$ for $j = 1, \ldots, s$, and $\|\theta^o - \hat{\theta}\| < \delta$. Thus for arbitrarily small $\delta > 0$ and $\epsilon > 0$ there exists $N_{\delta,\epsilon}$ such that for $n > N_{\delta,\epsilon}$,

$$\begin{aligned} &P(\text{exist a } \hat{\theta} \text{ such that } g_n^j(\hat{\theta}) = 0,\ j = 1, \ldots, s, \text{ and} \\ &\quad \|\theta^o - \theta\| < \delta) > 1 - \epsilon , \end{aligned} \tag{3.38}$$

which implies $\hat{\theta} \xrightarrow{P} \theta^o$. This establishes the first part of Theorem 3.1.

<u>Step 3.</u> Show that $\sqrt{n}(\hat{\theta}-\theta^o)$ has a limiting s-dimensional normal distribution.

From (3.38) we can write $g_n^j(\hat{\theta}) = o_p(1)$ for $j = 1, \ldots, s$. Now expand $\frac{1}{\sqrt{n}} g_n(\hat{\theta})$ at the true point θ^o

$$o_p(1) = \frac{1}{\sqrt{n}} g_n^j(\hat{\theta}) = \frac{1}{\sqrt{n}} g_n^j(\theta^o) + \frac{1}{\sqrt{n}} \sum_{k=1}^{s} (\hat{\theta}_k - \theta_k^o) g_n^{jk}(\bar{\theta})$$

$$\text{for} \quad j = 1, \ldots, s,$$

where $\bar{\theta} = \theta^o + \xi(\hat{\theta}-\theta^o)$, $|\xi| < 1$, or in the matrix form

$$\bar{G}_n(\theta^o) = -G_n(\bar{\theta})\ (\sqrt{n}(\hat{\theta}-\theta^o)) + o_p(1) ,$$

where $\overline{G}_n(\theta)$ and $G_n(\theta)$ are defined respectively in (3.18) and (3.19). By assuming $G_n(\overline{\theta})$ is nonsingular, we can rewrite the preceding equation

$$\sqrt{n}(\hat{\theta}-\theta^o) = -G_n(\overline{\theta})^{-1}\overline{G}_n(\theta^o) + o_p(1) . \qquad (3.39)$$

The fact that $\hat{\theta} \xrightarrow{p} \theta^o$ implies $\overline{\theta} \xrightarrow{p} \theta^o$ for $\overline{\theta} = \theta^o + \xi(\hat{\theta}-\theta^o)$, $|\xi| < 1$, and (3.31) will give

$$\frac{1}{n} g_n^{jk}(\overline{\theta}) = \frac{1}{n} g_n^{jk}(\theta^o) + \sum_{\ell=1}^{s} (\overline{\theta}_\ell - \theta_\ell^o) \frac{1}{n} g_n^{jk\ell}(\bar{\bar{\theta}})$$

$$= \frac{1}{n} g_n^{jk}(\theta^o) + o_p(1) \quad \text{for} \quad j, k = 1, \ldots, s,$$

where $\bar{\bar{\theta}} = \theta^o + \eta(\overline{\theta}-\theta^o)$, $|\eta| < 1$, or in matrix form

$$\begin{aligned} G_n(\overline{\theta}) &= G_n(\theta^o) + o_p(1) \\ &= (K_{\theta^o} + o_p(1)) + o_p(1) \\ &= 2A'A + o_p(1) , \end{aligned}$$

where the last two equalities are given in (3.33). Now applying Lemma 2.5, we can write

$$G_n^{-1}(\overline{\theta}) = \frac{1}{2}(A'A)^{-1} + o_p(1) . \qquad (3.40)$$

In addition, from the following relations

$$\frac{1}{\sqrt{n}} g_n^j(\theta^o) = 2 \sum_{i=1}^{r} \left(\frac{\sqrt{n}(S_{i-1}-S_i)}{\hat{D}_i}\right)_{\theta^o} \left(\frac{F_i^j}{F_i} - \frac{F_{i-1}^j}{F_{i-1}}\right)_{\theta^o} + o_p(1)$$

$$= 2 \sum_{i=1}^{r} \left(\frac{\sqrt{n}(S_{i-1}-S_i)}{D_i^{1/2}}\right)_{\theta^o} \frac{1}{D_i^{1/2}}\left(\frac{F_i^j}{F_i} - \frac{F_{i-1}^j}{F_{i-1}}\right)_{\theta^o} + o_p(1)$$

$$\text{for} \quad j = 1, \ldots, s$$

and (3.20), (3.21), (3.22) and (3.25) we obtain

$$\bar{G}_n(\theta^o) = 2A'\bar{X}_n + o_p(1) . \tag{3.41}$$

Applying (3.40) and (3.41), Equation (3.39) can be rewritten as

$$\sqrt{n}(\hat{\theta}-\theta^o) = -\left(\frac{1}{2}(A'A)^{-1} + o_p(1)\right)\left(2A'\bar{X}_n + o_p(1)\right) + o_p(1)$$

$$= -(A'A)^{-1}A'\bar{X}_n + o_p(1) . \tag{3.42}$$

It is easy to show that $\bar{X}_n \xrightarrow{d} N_r(0, I_r)$, which implies

$$(A'A)^{-1}A'\bar{X}_n \xrightarrow{d} N_s(0, (A'A)^{-1}A'I_rA(A'A)^{-1})$$

$$= N_s(0, (A'A)^{-1}) .$$

Hence

$$\sqrt{n}(\hat{\theta}-\theta^o) \xrightarrow{d} N_s(0, (A'A)^{-1}) .$$

This completes the proof for Theorem 3.1.

Now, we are ready to determine the limiting distribution of the statistic

$$g_n(\hat{\theta}) = n \sum_{i=1}^{r} \left(\frac{(S_{i-1}-S_i)^2}{\hat{D}_i} \right)_{\hat{\theta}}$$

where $\hat{\theta}$ is a solution for (3.27). As defined in (3.26), we expand $\overline{Y}_{ni}$ at the point θ^o,

$$\overline{Y}_{ni} = \left(\frac{\sqrt{n}(S_{i-1}-S_i)}{\hat{D}_i^{1/2}} \right)_{\hat{\theta}}$$

$$= \left(\frac{\sqrt{n}(S_{i-1}-S_i)}{\hat{D}_i^{1/2}} \right)_{\theta^o} + \sum_{j=1}^{s} (\hat{\theta}_j - \theta_j^o) \frac{\partial}{\partial \theta_j} \left(\frac{\sqrt{n}(S_{i-1}-S_i)}{\hat{D}_i^{1/2}} \right)_{\theta^o}$$

$$+ \sum_{j,k=1}^{s} (\hat{\theta}_j - \theta_j^o)(\hat{\theta}_k - \theta_k^o) \frac{\partial^2}{\partial \theta_k \partial \theta_j} \left(\frac{\sqrt{n}(S_{i-1}-S_i)}{\hat{D}_i^{1/2}} \right)_{\overline{\theta}}$$

$$= \left(\frac{\sqrt{n}(S_{i-1}-S_i)}{\hat{D}_i^{1/2}} \right)_{\theta^o} + \sum_{j=1}^{s} (\hat{\theta}_j - \theta_j^o) \left[\left(\frac{\sqrt{n}(S_{i-1}-S_i)^j}{\hat{D}_i^{1/2}} \right)_{\theta^o} - \left(\frac{\sqrt{n}(S_{i-1}-S_i)}{\hat{D}_i^{3/2}} (\hat{D}_i)^j \right)_{\theta^o} \right]$$

$$+ \sum_{j,k=1}^{s} \sqrt{n}(\hat{\theta}_j - \theta_j^o)(\hat{\theta}_k - \theta_k^o) \left[\left(\frac{(S_{i-1}-S_i)^{jk}}{\hat{D}_i^{1/2}} \right)_{\overline{\theta}} - \left(\frac{(S_{i-1}-S_i)^j}{\hat{D}_i^{3/2}} (\hat{D}_i)^k \right)_{\overline{\theta}} \right.$$

$$\left. - \left(\frac{(S_{i-1}-S_i)^k}{\hat{D}_i^{3/2}} (\hat{D}_i)^j \right)_{\overline{\theta}} + \frac{3}{2} \left(\frac{(S_{i-1}-S_i)}{\hat{D}_i^{5/2}} (\hat{D}_i)^j (\hat{D}_i)^k \right)_{\overline{\theta}} - \left(\frac{(S_{i-1}-S_i)}{\hat{D}_i^{3/2}} (\hat{D}_i)^{jk} \right)_{\overline{\theta}} \right] =$$

$$= \left(\frac{\sqrt{n}(S_{i-1}-S_i)}{D_i^{1/2}+o_p(1)}\right)_{\theta^o} + \sum_{j=1}^{s} (\hat{\theta}_j-\theta_j^o)\left[\left(\frac{\sqrt{n}(S_{i-1}-S_i)^j}{D_i^{1/2}+o_p(1)}\right)_{\theta^o} - O_p(1)\right] + o_p(1)$$

$$= \left(\frac{\sqrt{n}(S_{i-1}-S_i)}{D_i^{1/2}}\right)_{\theta^o} + \sqrt{n}\sum_{j=1}^{s} (\hat{\theta}_j-\theta_j^o)\left(\frac{1}{D_i^{1/2}}\left(\frac{F_i^j}{F_i} - \frac{F_{i-1}^j}{F_{i-1}}\right)\right)_{\theta^o} + o_p(1),$$

$$i = 1, \ldots, r$$

where $\bar{\theta} = \theta^o + \pi(\theta^o - \hat{\theta})$, $|\pi| < 1$. The above equation can be written in matrix form

$$\bar{Y}_n = \bar{X}_n + A\sqrt{n}(\hat{\theta}-\theta^o) + o_p(1). \tag{3.43}$$

Replacing $\sqrt{n}(\hat{\theta}-\theta^o)$ in (3.43) by (3.42) gives

$$\bar{Y}_n = \bar{X}_n + A(-(A'A)^{-1}A'\bar{X}_n)+o_p(1)) + o_p(1)$$

$$= (I_r - A(A'A)^{-1}A')\bar{X}_n + o_p(1).$$

It is simple to check that $(I_r - A(A'A)^{-1}A')$ is symmetric and idempotent, hence

$$\bar{Y}_n'\bar{Y}_n = \bar{X}_n'(I_r - A(A'A)^{-1}A')\bar{X}_n + o_p(1). \tag{3.44}$$

We can choose the generalized inverse of $(I_r - A(A'A)^{-1}A')$ equal to itself and apply Lemma 2.3 to give

$$\bar{X}_n'(I_r - A(A'A)^{-1}A')\bar{X}_n \xrightarrow{d} \chi^2(r-s),$$

where the $r-s$ degrees of freedom is obtained by

$$\begin{aligned} \text{Rank } (I_r - A(A'A)^{-1}A') &= \text{trace}(I_r - A(A'A)^{-1}A') \\ &= \text{trace}(I_r) - \text{trace}((A'A)^{-1}A'A) \\ &= \text{trace}(I_r) - \text{trace}(I_s) \\ &= r-s . \end{aligned}$$

From (2.2) and (3.44) we finally obtain

$$\bar{Y}_n'\bar{Y}_n = n \sum_{i=1}^{r} \left(\frac{(S_{i-1}-S_i)^2}{\hat{D}_i}\right)_{\hat{\theta}} \xrightarrow{d} \chi^2(r-s). \tag{3.45}$$

This suggests that under the random censorship assumption, when the null hypothesis is composite and sample size is large, the statistic specified in (3.45) can be used to perform the goodness of fit test.

Example. In this example, the generalized chi-square goodness of fit tests for exponential and Weibull distributions are illustrated. The data used is extracted from a large group of breast cancer cases which were collected by Calvin Zippin from 12 hospitals in different parts of the United States. (These original data have been used by Cutler and Myers (1967), Koch, Johnson and Tolley (1972), and Brunk, Thomas, Elashoff and Zippin (1973) in other areas of research.) The subgroup of 175 patients we use have largest tumor size that is less than 2 cm, no skin attachment or fixation, and no lymph node involvement determined by physical examination. For this subgroup, 61 patients died and the remaining were censored after varying periods of time.

For testing exponential fit $H_0: F_\theta(x) = \exp(-x/\theta)$ and Weibull fit $H_0: F_{\theta_1,\theta_2}(x) = \exp(-x^{\theta_2}/\theta_1)$ the minimum chi-square estimates (3.45) were found to be respectively $\hat{\theta} = 369.91$ and $\hat{\theta}_1 = 306.47$, $\hat{\theta}_2 = 0.962$. The corresponding values of the generalized chi-square statistics were respectively 12.23 and 12.14 with corresponding degrees of freedom 11 and 10. In Table 1 below, columns 2, 3 and 4 include respectively the product-limit estimator, fitted exponential distribution, and fitted Weibull distribution evaluated at the selected partition points given in column 1. In Figure 1, the product-limit estimator and the fitted exponential and Weibull distributions are plotted. It is interesting to see that both fitted distributions agree quite well with the product-limit estimator.

Table 1. Product-limit estimator and exponential and Weibull fits for the breast cancer data.

Months	Product-limit Estimator	Exponential	Weibull
0	1.	1.	1.
12	0.9886	0.9681	0.9649
24	0.9371	0.9372	0.9328
36	0.8914	0.9073	0.9024
48	0.8627	0.8783	0.8733
60	0.8397	0.8503	0.8456
72	0.8165	0.8231	0.8188
84	0.7990	0.7969	0.7930
96	0.7873	0.7714	0.7683
108	0.7581	0.7468	0.7444
120	0.7283	0.7230	0.7213
180	0.6703	0.6147	0.6173
240	0.5171	0.5228	0.5293

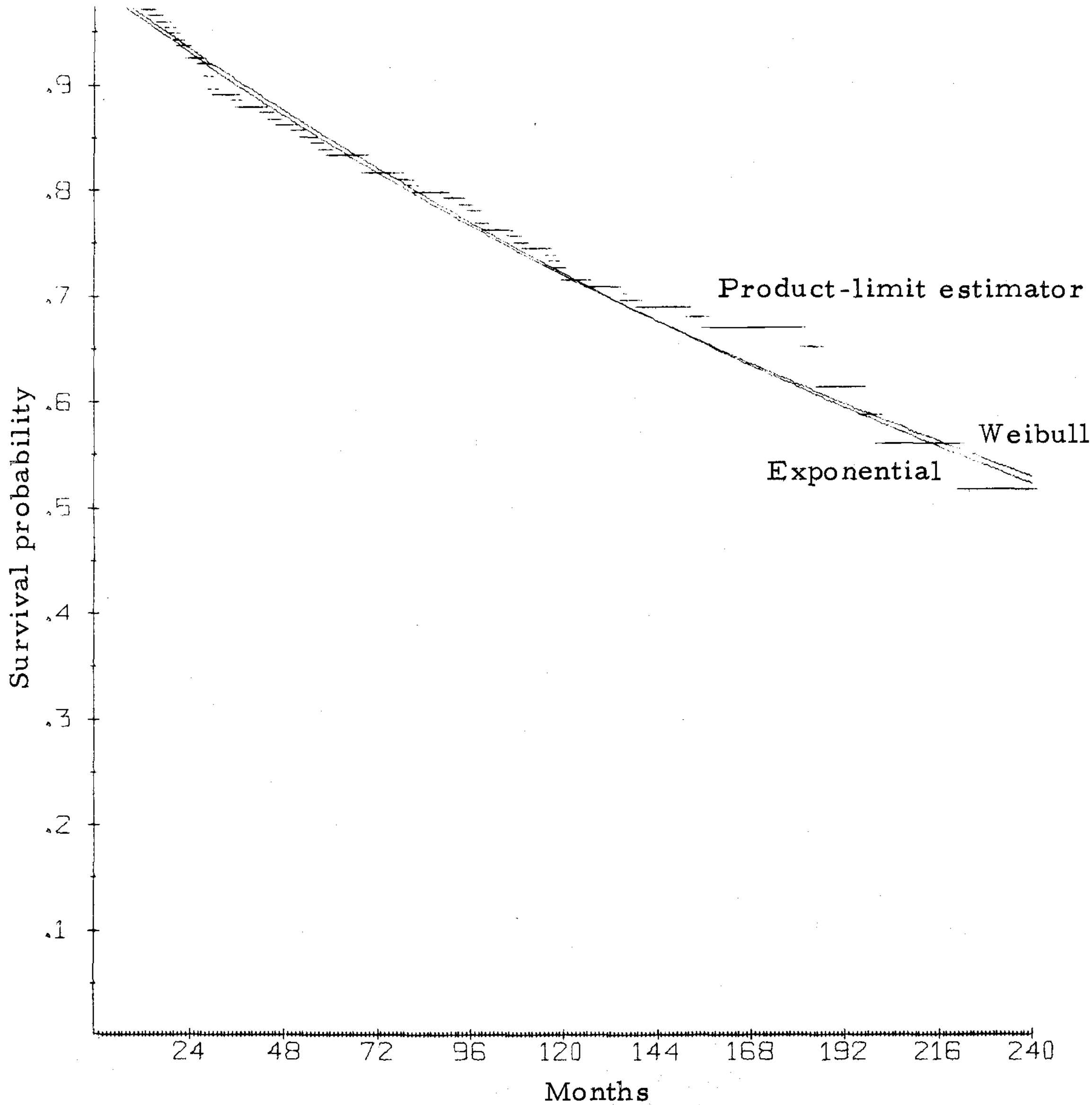

Figure 1. Product-limit estimator and the exponential and Weibull fitted distributions for the breast cancer example.

3.6. Relationship with the Classical Chi-Square Test Statistic

It is interesting to see that the goodness of fit test statistic (3.9) for arbitrarily censored data will reduce to the classical chi-square statistic (3.3) provided the sample is uncensored. In the uncensored data case, we may set $H(y) = 1$ for $y \geq 0$. Hence the covariance matrix for random vector

$$Z_n = n^{1/2}(\hat{F}-F_\theta) = (n^{1/2}(\hat{F}_1-F_1), \ldots, n^{1/2}(\hat{F}_r-F_r))'$$

is

$$\begin{aligned}
\Sigma &= [\Gamma(a_i, a_j)] \\
&= [F_i F_j \int_0^{a_i} \frac{1}{H(z)F_\theta^2(z)} (-dF_\theta(z))] \\
&= [F_i F_j \int_0^{a_i} \frac{1}{F_\theta^2(z)} (-dF_\theta(z))] \\
&= [F_i F_j \frac{1-F_i}{F_i}] \\
&= [F_j(1-F_i)] ,
\end{aligned}$$

for $a_i < a_j$, $i, j = 1, \ldots, r$. Let the $r \times r$ matrix C be defined as in the beginning of Section 3.4, then from the fact

$$Z_n = n^{1/2}(\hat{F}-F_\theta) \xrightarrow{d} N_r(0, \Sigma),$$

we obtain

$$Z_n^\Delta = CZ_n \xrightarrow{d} N_r(0, C\Sigma C') = N_r(0, \Sigma^\Delta) ,$$

where

$$Z_n^\Delta = \frac{1}{\sqrt{n}} \begin{pmatrix} n(\hat{F}_0-\hat{F}_1) - n(F_0-F_1) \\ n(\hat{F}_1-\hat{F}_2) - n(F_1-F_2) \\ \vdots \\ n(\hat{F}_{r-1}-\hat{F}_r) - n(F_{r-1}-F_r) \end{pmatrix}$$

$$= \frac{1}{\sqrt{n}} \begin{pmatrix} V_1 - nP_1 \\ V_2 - nP_2 \\ \vdots \\ V_r - nP_r \end{pmatrix} \tag{3.46}$$

and

$$\Sigma^{\Delta} = C\Sigma C'$$

$$= \begin{pmatrix} F_1(1-F_1) & (1-F_1)(F_2-F_1) & \cdots & (1-F_1)(F_{r-1}-F_{r-2}) & (1-F_1)(F_r-F_{r-1}) \\ (1-F_1)(F_2-F_1) & (F_1-F_2)(1+F_2-F_1) & \cdots & (F_1-F_2)(F_{r-1}-F_{r-2}) & (F_1-F_2)(F_r-F_{r-1}) \\ \vdots & \vdots & & \vdots & \vdots \\ (1-F_1)(F_{r-1}-F_{r-2}) & (F_1-F_2)(F_{r-1}-F_{r-2}) & \cdots & (F_{r-2}-F_{r-1})(1+F_{r-1}-F_{r-2}) & (F_{r-2}-F_{r-1})(F_r-F_{r-1}) \\ (1-F_1)(F_r-F_{r-1}) & (F_1-F_2)(F_r-F_{r-1}) & \cdots & (F_{r-2}-F_{r-1})(F_r-F_{r-1}) & (F_{r-1}-F_r)(1+F_r-F_{r-1}) \end{pmatrix}$$

$$= \begin{pmatrix} (1-P_1)P_1 & -P_1P_2 & \cdots & -P_1P_{r-1} & -P_1P_r \\ -P_1P_2 & (1-P_2)P_2 & \cdots & -P_2P_{r-1} & -P_2P_r \\ \vdots & \vdots & & \vdots & \vdots \\ -P_1P_{r-1} & -P_2P_{r-1} & \cdots & (1-P_{r-1})P_{r-1} & -P_{r-1}P_r \\ -P_1P_4 & -P_2P_r & \cdots & -P_{r-1}P_r & (1-P_r)P_r \end{pmatrix}$$

$$= \begin{pmatrix} P_1 & & & & O \\ & P_2 & & & \\ & & \ddots & & \\ O & & & P_{r-1} & \\ & & & & P_r \end{pmatrix} - \begin{pmatrix} P_1 \\ P_2 \\ \vdots \\ P_{r-1} \\ P_r \end{pmatrix} (P_1 P_2 \cdots P_{r-1} P_r), \qquad (3.47)$$

with V_i and P_i, $i = 1, \ldots, r+1$, defined as in Section 3.2. Applying the inversion method given by Graybill (1969, p. 170), we have

$$\Sigma^{\Delta -1} = \begin{pmatrix} \frac{1}{P_1} & & & & 0 \\ & \frac{1}{P_2} & & & \\ & & \ddots & & \\ & & & \frac{1}{P_{r-1}} & \\ 0 & & & & \frac{1}{P_r} \end{pmatrix} + \frac{1}{P_{r+1}} \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & & \vdots \\ 1 & \cdots & 1 \end{pmatrix} .$$

Therefore, from (3.46), (3.47),

$$\begin{aligned} Z_n'\Sigma^{-1}Z_n &= Z_n'C'C'^{-1}\Sigma^{-1}C^{-1}CZ_n \\ &= (CZ_n)'(C\Sigma C')^{-1}(CZ_n) \\ &= Z_n^{\Delta\,'}\Sigma^{\Delta -1}Z_n^{\Delta} \\ &= (\sqrt{n}\, Z_n^{\Delta})'(n\Sigma^{\Delta})^{-1}(\sqrt{n}\, Z_n^{\Delta}) \\ &= \sum_{i=1}^{r} \frac{(V_i - nP_i)^2}{nP_i} + \frac{1}{nP_{i+1}} \left(\sum_{i=1}^{r} (V_i - nP_i)\right)^2 \\ &= \sum_{i=1}^{r+1} \frac{(V_i - nP_i)^2}{nP_i} , \end{aligned}$$

which is the classical goodness of fit test statistic.

4. SMOOTH GOODNESS OF FIT TEST FOR A SIMPLE HYPOTHESIS UNDER RANDOM CENSORSHIP

4.1. Neyman's Smooth Goodness of Fit Test for a Simple Hypothesis with Uncensored Data

Neyman (1937) has developed a class of goodness of fit tests for a simple null hypothesis against a family of alternatives (4.3) which is relatively smooth compared to the null hypothesis. Barton (1955, 1956) generalized Neyman's test for the cases of a composite null hypothesis and grouped data. In this chapter, Neyman's goodness of fit test for a simple hypothesis is generalized for random censorship.

First, we give a brief development of Neyman's test for uncensored data. Let the random variable X have some continuous distribution function $F(x)$, with the null hypothesis to be tested

$$H_0'': F(x) = F_0(x) , \tag{4.1}$$

where $F_0(x) = P_0(X \geq x)$ is completely specified with density function $f_0(x)$. The probability integral transformation

$$z = \int_x^{\infty} f_0(t)dt = F_0(x) \tag{4.2}$$

is applied to the random sample $X_1, \dots, X_n$, and thus, we obtain n independent observations $Z_1, \dots, Z_n$, where Z_i is uniformly distributed over the interval $[0, 1]$ when H_0 is true. Therefore, the equivalent null hypothesis in terms of transformed variable Z is

$$H_0': \; q_0(z) = 1 \qquad 0 \leq z \leq 1,$$

where $q_0(z)$ designates the density function of Z and $Q_0(z)$ the corresponding distribution function. Neyman specified a family of distributions which allow the distributions in the alternative hypothesis to vary smoothly from H_0'

$$H_a': q_\theta(z) = c(\theta) e^{\Sigma_{i=1}^{r} \theta_i \pi_i(z)}, \qquad 0 \leq z \leq 1; \quad r = 1, 2, \ldots \tag{4.3}$$

where $\theta = (\theta_1, \ldots, \theta_r) \in R^r$ and $c(\theta)$ is the normalizing constant. More specifically

$$c(\theta)^{-1} = \int_0^1 e^{\Sigma_{i=1}^{r} \theta_i \pi_i(z)} \, dz , \tag{4.4}$$

and $\pi_i(z) = a_{i0} + a_{i1} z^1 + \ldots + a_{ii} z^i$ are transformed Legendre orthogonal polynomials (see 4.19) of z which satisfy

$$\int_0^1 \pi_i(z) \pi_j(z) dz = \begin{cases} 1 & \text{if} \quad i = j \\ 0 & \text{if} \quad i \neq j \end{cases}$$

for $i, j = 1, \ldots, r$. In terms of the alternative (4.3), the problem of testing H_0' against H_a' will be equivalent to testing

$$H_0:\ \theta_i = 0, \qquad i = 1, \ldots, r,$$

against

$$H_a:\ \text{some } \theta_i \neq 0, \qquad i = 1, \ldots, r .$$

The likelihood function of θ is

$$L = \prod_{i=1}^{n} q_\theta(z_i) = (c(\theta))^n \exp\Big(\sum_{i=1}^{r} \theta_i \sum_{j=1}^{n} \pi_i(z_j) \Big)$$

$$= (c(\theta))^n \exp\Big(\sum_{i=1}^{r} \sqrt{n}\, \theta_i \Big(\frac{1}{\sqrt{n}} \sum_{j=1}^{n} \pi_i(z_j) \Big)$$

$$= (c(\theta))^n \exp\Big(\sum_{i=1}^{r} \beta_i u_i \Big) , \tag{4.5}$$

where

$$\beta_i = \sqrt{n}\, \theta_i \quad \text{and} \quad u_i = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \pi_i(z_j) .$$

Neyman has shown that the critical region defined by the inequality

$$\sum_{i=1}^{r} u_i^2 \geq \chi_\alpha^2(r)$$

will asymptotically satisfy his definition of an unbiased critical region of type C, where $\chi^2_\alpha(r)$ is the upper αth quantile of the chi-square distribution with r degrees of freedom. The tests constructed by this procedure are generally referred to as Neyman smooth goodness of fit tests.

We now give another approach for showing that the statistic $\sum_{i=1}^{r} u_i^2$ has a limiting chi-square distribution with r degrees of freedom. This derivation will also be used in the next section to generalize the test statistic for randomly censored data.

From (4.5) we may write

$$\frac{1}{\sqrt{n}} \ln L = \frac{1}{\sqrt{n}}\left(n \ln c(\theta) + \sum_{i=1}^{r} \theta_i \sum_{j=1}^{n} \pi_i(z_j)\right). \tag{4.6}$$

Applying

$$c(0) = 1 \quad \text{and} \quad \frac{\partial}{\partial\theta_s} c(\theta)\Big|_{\theta=0} = -\int_0^1 \pi_s(z)dz = 0 \tag{4.7}$$

and taking the first partial derivative of (4.6) with respect to θ_s and evaluating it at the point $\theta = 0$ yields

$$\frac{1}{\sqrt{n}} \frac{\partial \ln L}{\partial\theta_s}\Big|_{\theta=0} = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \pi_s(z_j) \quad \text{for} \quad s = 1, \ldots, r. \tag{4.8}$$

Hence the means of the random variables $\frac{1}{\sqrt{n}} \frac{\partial \ln L}{\partial\theta_s}\Big|_{\theta=0}$, $s = 1, \ldots, r$, under H_0 are

$$E_{\theta=0}\left(\frac{1}{\sqrt{n}}\frac{\partial \ln L}{\partial \theta_s}\Big|_{\theta=0}\right) = \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\int_0^1 \pi_s(z_j)dz_j = 0 .$$

For the evaluation of covariances, we need

$$\frac{1}{n}\frac{\partial^2 \ln L}{\partial\theta_s\partial\theta_\ell} = \frac{1}{n}\left(-n\frac{1}{c(\theta)^2}\frac{\partial}{\partial\theta_s}c(\theta)\frac{\partial}{\partial\theta_\ell}c(\theta) + n\frac{1}{c(\theta)}\frac{\partial^2}{\partial\theta_s\partial\theta_\ell}c(\theta)\right),$$

evaluated at $\theta = 0$,

$$\begin{aligned}\frac{1}{n}\frac{\partial^2 \ln L}{\partial\theta_s\partial\theta_\ell}\Big|_{\theta=0} &= \frac{1}{n}\left(0+n\frac{1}{c(\theta)}\frac{\partial^2}{\partial\theta_s\partial\theta_\ell}c(\theta)\Big|_{\theta=0}\right)\\ &= \frac{1}{n}\left(-n\int_0^1 \pi_s(z)\pi_\ell(z)dz\right)\\ &= \begin{cases}-1 & \text{if } s=\ell\\ 0 & \text{if } s\neq\ell\end{cases} .\end{aligned} \qquad (4.9)$$

Thus, under H_0 we have

$$\begin{aligned}\mathrm{Cov}\left(\frac{1}{\sqrt{n}}\frac{\partial \ln L}{\partial\theta_s}\Big|_{\theta=0}, \frac{1}{\sqrt{n}}\frac{\partial \ln L}{\partial\theta_\ell}\Big|_{\theta=0}\right) &= E_{\theta=0}\left(\frac{1}{n}\frac{\partial \ln L}{\partial\theta_s}\Big|_{\theta=0}\frac{\partial \ln L}{\partial\theta_\ell}\Big|_{\theta=0}\right)\\ &= -E_{\theta=0}\left(\frac{1}{n}\frac{\partial^2 \ln L}{\partial\theta_s\partial\theta_\ell}\Big|_{\theta=0}\right)\\ &= \begin{cases}1 & \text{if } s=\ell\\ 0 & \text{if } s\neq\ell\end{cases} \qquad s,\ell = 1,\ldots,r.\end{aligned}$$

Hence, for

$$W_n^* = \left(\frac{1}{\sqrt{n}} \frac{\partial \ln L}{\partial \theta_1} \Big|_{\theta=0}, \ldots, \frac{1}{\sqrt{n}} \frac{\partial \ln L}{\partial \theta_r} \Big|_{\theta=0} \right)'$$

$$= \left(\frac{1}{\sqrt{n}} \sum_{j=1}^{n} \pi_1(z_j), \ldots, \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \pi_r(z_j) \right)' , \tag{4.10}$$

the mean vector and covariance matrix are simply $E(W_n^*) = 0$ and $Cov(W_n^*) = I$. Therefore, by treating $(\pi_1(z_j), \ldots, \pi_r(z_j))$, $j = 1, \ldots, n$, as an r-dimensional random sample of size n, the multidimensional Central Limit Theorem (Wilks, 1963, p. 258) may be applied to show W_n^* has a limiting normal distribution with mean vector 0 and covariance matrix I. Hence, from Lemma 2.3, the statistic

$$W_n^{*'} \, Cov(W_n^*)^{-1} W_n^* = \sum_{i=1}^{r} \left(\frac{1}{\sqrt{n}} \frac{\partial \ln L}{\partial \theta_i} \Big|_{\theta=0} \right)^2$$

$$= \sum_{i=1}^{r} \left(\frac{1}{\sqrt{n}} \sum_{j=1}^{n} \pi_i(z_j) \right)^2$$

$$= \sum_{i=1}^{r} u_i^2 \xrightarrow{d} \chi^2(r) . \tag{4.11}$$

4.2. A Generalized Smooth Goodness of Fit Test

In the random censorship case, as in Section 3.1, we let $X_1, \ldots, X_n$ and $T_1, \ldots, T_n$ respectively represent the random failure and censor times with corresponding distribution functions $F_\theta(x)$ and $H(t)$ and density functions $f_\theta(x)$ and $h(t)$. Using the inverse transformation of (4.2), $x = F_0^{-1}(z)$, the family of alternative densities for X corresponding to (4.3) is

$$g_\theta(x) = c(\theta)e^{\Sigma_{i=1}^r \theta_i \pi_i(F_0(x))} f_0(x) \,, \tag{4.11}$$

for $0 \leq x < \infty$ and $r = 1, 2, \ldots,$ with corresponding distribution functions

$$\begin{aligned} G_\theta(x) &= \int_x^\infty g_\theta(t)dt \\ &= -\int_z^0 q_\theta(t)dt \\ &= \int_0^z q_\theta(t)dt = Q_\theta(z) \,. \end{aligned}$$

For the observable variables $Y_i = \min(X_i, T_i)$ and $\delta_i = 1(0)$ for $Y_i = X_i (Y_i = T_i)$, $i = 1, \ldots, n$, the likelihood function is

$$L = L_\theta = \prod_{j=1}^{n} \ell_j \tag{4.12}$$

where $\ell_j = (g_\theta(y_j)H(y_i))^{\delta_j} (G_\theta(y_j)h(y_j))^{1-\delta_j}$.

The derivation of the goodness of fit test statistic in Section 4.1 for the uncensored case will now be generalized to the random censorship case. First we note that the log-likelihood function may be written as

$$\ln L = \sum_{j=1}^{n} \ln \ell_j$$

$$= \sum_{j=1}^{n} (\delta_j \ln g_\theta(y_j) + (1-\delta_j) \ln G_\theta(y_j)) + R(y, \delta) ,$$

where $R(y, \delta)$ is independent of θ. By defining

$$W_n^* = \left(\frac{1}{\sqrt{n}} \frac{\partial \ln L}{\partial \theta_1} \Big|_{\theta=0}, \ldots, \frac{1}{\sqrt{n}} \frac{\partial \ln L}{\partial \theta_r} \Big|_{\theta=0} \right)'$$

$$= \left(\frac{1}{\sqrt{n}} \sum_{j=1}^{n} \frac{\partial \ln \ell_j}{\partial \theta_1} \Big|_{\theta=0}, \ldots, \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \frac{\partial \ln \ell_j}{\partial \theta_r} \Big|_{\theta=0} \right)'$$

$$= \frac{1}{\sqrt{n}} \sum_{j=1}^{n} W_j \tag{4.13}$$

where

$$W_j = \left(\frac{\partial \ln \ell_j}{\partial \theta_1} \Big|_{\theta=0}, \ldots, \frac{\partial \ln \ell_j}{\partial \theta_r} \Big|_{\theta=0} \right)'$$

and from (4. 7) and the following relations

$$\begin{aligned}
Q_\theta(z)\big|_{\theta=0} &= z \\
\frac{\partial}{\partial \theta_s} Q_\theta(z)\big|_{\theta=0} &= \frac{\partial}{\partial \theta_s} \int_0^z q_\theta(t)dt\big|_{\theta=0} \\
&= \int_0^z \pi_s(t)dt \\
\frac{\partial}{\partial \theta_s} \ln g_\theta(y)\big|_{\theta=0} &= \left(\frac{1}{c(\theta)} \frac{\partial}{\partial \theta_s} c(\theta) + \pi_s(F_0(y)) \right)\big|_{\theta=0} \\
&= \pi_s(z) \\
\frac{\partial}{\partial \theta_s} \ln G_\theta(y)\big|_{\theta=0} &= \frac{\partial}{\partial \theta_s} \ln Q_\theta(z)\big|_{\theta=0} \\
&= \frac{1}{Q_\theta(z)} \frac{\partial}{\partial \theta_s} Q_\theta(z)\big|_{\theta=0} \\
&= \frac{1}{z} \int_0^z \pi_s(t)dt \ ,
\end{aligned} \tag{4.14}$$

the components of W_n^* can be written as

$$\frac{1}{\sqrt{n}} \frac{\partial \ln L}{\partial \theta_s}\Big|_{\theta=0} = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \left(\delta_j \frac{\partial}{\partial \theta_s} \ln g_\theta(y_j) + (1-\delta_j) \frac{\partial}{\partial \theta_s} \ln G_\theta(y_j)\right)\Big|_{\theta=0}$$

$$= \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \left(\delta_j \pi_s(z_j) + (1-\delta_j) \frac{1}{z_j} \int_0^{z_j} \pi_s(z)dz\right) \qquad (4.15)$$

$$\text{for} \quad s = 1, \ldots, r.$$

For the derivations of the mean and covariance of W_n^* we apply the relations

$$E_{\theta=0}\left(\frac{\partial \ln \ell_j}{\partial \theta_s}\Big|_{\theta=0}\right) = 0$$

$$E_{\theta=0}\left(\frac{\partial \ln \ell_j}{\partial \theta_s}\Big|_{\theta=0} \cdot \frac{\partial \ln \ell_j}{\partial \theta_\ell}\Big|_{\theta=0}\right) = -E_{\theta=0}\left(\frac{\partial^2 \ln \ell_j}{\partial \theta_s \partial \theta_\ell}\Big|_{\theta=0}\right),$$

where $s, \ell = 1, \ldots, r$; $j = 1, \ldots, n$, which are known to hold for exponential families (Lehmann, 1959, p. 52). Hence

$$E_{\theta=0}(W_n^*) = \frac{1}{n} \sum_{j=1}^{n} E_{\theta=0}(W_j)$$

$$= E_{\theta=0}(W_j) = 0$$

and the (s, ℓ) element of $\Sigma = \mathrm{Cov}(W_n^*) = \mathrm{Cov}(W_j)$ is

$$\mathrm{Cov}_{\theta=0}\left(\frac{1}{\sqrt{n}}\frac{\partial \ln L}{\partial\theta_s}\Big|_{\theta=0}, \frac{1}{\sqrt{n}}\frac{\partial \ln L}{\partial\theta_\ell}\Big|_{\theta=0}\right)$$

$$= \mathrm{Cov}_{\theta=0}\left(\frac{\partial \ln \ell_j}{\partial\theta_s}\Big|_{\theta=0}, \frac{\partial \ln \ell_j}{\partial\theta_\ell}\Big|_{\theta=0}\right)$$

$$= E_{\theta=0}\left(\frac{\partial \ln \ell_j}{\partial\theta_s}\Big|_{\theta=0} \frac{\partial \ln \ell_j}{\partial\theta_\ell}\Big|_{\theta=0}\right)$$

$$= -E_{\theta=0}\left(\frac{\partial^2 \ln \ell_j}{\partial\theta_s\partial\theta_\ell}\Big|_{\theta=0}\right)$$

$$= -E_{\theta=0}\Big(\delta_j \frac{\partial^2}{\partial\theta_s\partial\theta_\ell} \ln g_\theta(y_j)\Big|_{\theta=0}$$

$$+ (1-\delta_j)\frac{\partial^2}{\partial\theta_s\partial\theta_\ell} \ln G_\theta(y_j)\Big|_{\theta=0}\Big) . \tag{4.16}$$

By using (4. 7), (4. 14) and the following relations

$$\frac{\partial^2}{\partial\theta_s\partial\theta_\ell} c(\theta)\Big|_{\theta=0} = -\int_0^1 \pi_s(y)\pi_\ell(y)dy$$

$$= \begin{cases} -1 & \text{if} \quad s = \ell \\ 0 & \text{if} \quad s \neq \ell \end{cases}$$

$$\frac{\partial^2}{\partial\theta_s\partial\theta_\ell} Q_\theta(z)\Big|_{\theta=0} = \left(\frac{\partial}{\partial\theta_\ell}\left(\frac{\partial}{\partial\theta_s} Q_\theta(z)\right)\right)\Big|_{\theta=0}$$

$$= \begin{cases} -z + \int_0^z \pi_s^2(z)dz & \text{if} \quad s = \ell \\ \int_0^z \pi_s(z)\pi_\ell(z)dz & \text{if} \quad s \neq \ell \end{cases}$$

$$\frac{\partial^2}{\partial\theta_s\partial\theta_\ell}\ln g_\theta(y)\Big|_{\theta=0} = \frac{\partial}{\partial\theta_\ell}\Big(\frac{1}{c(\theta)}\frac{\partial}{\partial\theta_s}c(\theta)+\pi_s(F_0(y))\Big)\Big|_{\theta=0}$$

$$= \Big(-\frac{1}{c(\theta)^2}\frac{\partial}{\partial\theta_\ell}c(\theta)\frac{\partial}{\partial\theta_s}c(\theta)+\frac{1}{c(\theta)}\frac{\partial^2}{\partial\theta_s\partial\theta_\ell}c(\theta)\Big)\Big|_{\theta=0}$$

$$= \begin{cases} -1 & \text{if} \quad s = \ell \\ 0 & \text{if} \quad s \neq \ell \end{cases}$$

$$\frac{\partial^2}{\partial\theta_s\partial\theta_\ell}\ln G_\theta(y)\Big|_{\theta=0} = \frac{\partial^2}{\partial\theta_s\partial\theta_\ell}\ln Q_\theta(z)\Big|_{\theta=0}$$

$$= \frac{\partial}{\partial\theta_\ell}\Big(\frac{1}{Q_\theta(z)}\frac{\partial}{\partial\theta_s}Q_\theta(z)\Big)\Big|_{\theta=0}$$

$$= \Big(-\frac{1}{Q_\theta^2(z)}\frac{\partial}{\partial\theta_\ell}Q_\theta(z)\frac{\partial}{\partial\theta_s}Q_\theta(z)$$

$$+\frac{1}{Q_\theta(z)}\frac{\partial^2}{\partial\theta_s\partial\theta_\ell}Q_\theta(z)\Big)\Big|_{\theta=0}$$

$$= \begin{cases} -\frac{1}{z^2}\Big(\int_0^z \pi_s(t)dt\Big)^2+\frac{1}{z}\Big(\int_0^z \pi_s^2(t)dt\Big) - 1 & \text{if} \quad s = \ell \\ -\frac{1}{z^2}\Big(\int_0^z \pi_s(t)dt\Big)\Big(\int_0^z \pi_\ell(t)dt\Big) & \\ \qquad +\frac{1}{z}\Big(\int_0^z \pi_s(t)\pi_\ell(t)dt\Big) & \text{if} \quad s \neq \ell\,, \end{cases}$$

Equation (4.16) can be rewritten as

$$\mathrm{Cov}_{\theta=0}\left(\frac{\partial \ln \ell_j}{\partial \theta_s}\Big|_{\theta=0}, \frac{\partial \ln \ell_j}{\partial \theta_\ell}\Big|_{\theta=0}\right)$$

$$= \begin{cases} E_{\theta=0}\left((1-\delta_j)\left(\frac{1}{z_j^2}\left(\int_0^{z_j} \pi_s(z)dz\right)^2 - \frac{1}{z_j}\int_0^{z_j} \pi_s^2(z)dz + 1\right)\right) & \text{if} \quad s = \ell \\ E_{\theta=0}\left((1-\delta_j)\left(\frac{1}{z_j^2}\left(\int_0^{z_j} \pi_s(z)dz\right)\left(\int_0^{z_j} \pi_\ell(z)dz\right) - \frac{1}{z_j}\int_0^{z_j} \pi_s(z)\pi_\ell(z)dz\right)\right) & \text{if} \quad s \neq \ell \,. \end{cases} \qquad (4.17)$$

Now treating W_j, $j = 1, \ldots, n$ as a random sample of size n from a distribution with mean vector 0 and covariance matrix Σ we obtain, once again from the multi-dimensional central-limit theorem, $W_n^* \xrightarrow{d} N_r(0, \Sigma)$. Applying Lemma 2.3 and assuming Σ is nonsingular, we finally have

$$W_n^{*'} \Sigma^{-1} W_n^* \xrightarrow{d} \chi^2(r) \,. \qquad (4.18)$$

In order to apply the smooth goodness of fit statistic given in (4.18), the components of W_n^* and the component of Σ should be evaluated in a simpler form. Since up to 4th order polynomial is generally considered sufficient (Neyman, 1937), we take $r = 4$ for illustration. The first four orthogonal polynomials are

$$\pi_0(z) = 1$$

$$\pi_1(z) = \sqrt{12}\,(z - \frac{1}{2})$$

$$\pi_2(z) = \sqrt{5}\,(6(z - \frac{1}{2})^2 - \frac{1}{2})$$

$$\pi_3(z) = \sqrt{7}\,(20(z - \frac{1}{2})^3 - 3(z - \frac{1}{2}))$$

$$\pi_4(z) = 210(z - \frac{1}{2})^4 - 45(z - \frac{1}{2})^2 + \frac{9}{8}\,. \tag{4.19}$$

Accordingly, the following relations can be obtained by direct integration:

$$\int_0^z \pi_1(t)dt = \sqrt{3}\, z(z-1)$$

$$\int_0^z \pi_2(t)dt = \sqrt{5}\, z(2z^2-3z+1)$$

$$\int_0^z \pi_3(t)dt = \sqrt{7}\, z(5z^3-10z^2+6z-1)$$

$$\int_0^z \pi_4(t)dt = 3z(14z^4-35z^3+30z^2-10z+1)$$

$$\int_0^z \pi_1^2(t)dt = z(4z^2-6z+3)$$

$$\int_0^z \pi_2^2(t)dt = z(36z^4-90z^3+80z^2-30z+5)$$

$$\int_0^z \pi_3^2(t)dt = z(400z^6-1400z^5+1932z^4-1330z^3+476z^2-84z+7)$$

$$\int_0^z \pi_4^2(t)dt = z(4900z^8-22050z^7+41400z^6-42000z^5+24912z^4$$
$$- 8730z^3+1740z^2-180z+9)$$

$$\int_0^z \pi_1(t)\pi_2(t)dt = \sqrt{15}\ z(3z^3-6z^2+4z-1)$$

$$\int_0^z \pi_1(t)\pi_3(t)dt = \sqrt{21}\ z(8z^4-20z^3+18z^2-7z+1)$$

$$\int_0^z \pi_1(t)\pi_4(t)dt = \sqrt{3}\ z(70z^5-210z^4+240z^3-130z^2+33z-3)$$

$$\int_0^z \pi_2(t)\pi_3(t)dt = \sqrt{35}\ z(20z^5-60z^4+68z^3-36z^2+9z-1)$$

$$\int_0^z \pi_2(t)\pi_4(t)dt = \sqrt{5}\ z(180z^6-630z^5+870z^4-600z^3+216z^2-39z+3)$$

$$\int_0^z \pi_3(t)\pi_4(t)dt = \sqrt{7}\ z(525z^7-2100z^6+3420z^5-2910z^4+1380z^3$$
$$- 360z^2+48z-3)\ .$$

Hence, the first four components of the W_n^* vector, according to (4.15), can be rewritten as

$$\frac{1}{\sqrt{n}}\ \frac{\partial \ln L}{\partial \theta_1}\Big|_{\theta=0} = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} (\sqrt{3}\ (\delta_j(2z_j-1)+(1-\delta_j)(z_j-1))$$

$$\frac{1}{\sqrt{n}}\ \frac{\partial \ln L}{\partial \theta_2}\Big|_{\theta=0} = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} (\delta_j(\sqrt{5}\ (6(z_j-\frac{1}{2})^2-\frac{1}{2}))+(1-\delta_j)\sqrt{5}\ (2z_j^2-3z_j+1))$$

$$\frac{1}{\sqrt{n}}\ \frac{\partial \ln L}{\partial \theta_3}\Big|_{\theta=0} = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} (\delta_j(\sqrt{7}\ (20(z_j-\frac{1}{2})^3-3(z_j-\frac{1}{2})))$$

$$+\ (1-\delta_j)\sqrt{7}\ (5z_j^3-10z_j^2+6z_j-1))$$

$$\frac{1}{\sqrt{n}}\ \frac{\partial \ln L}{\partial \theta_4}\Big|_{\theta=0} = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} (\delta_j(210(z_j-\frac{1}{2})^4-45(z_j-\frac{1}{2})^2+\frac{9}{8})$$

$$+\ (1-\delta_j)3\cdot(14z_j^4-35z_j^3+30z_j^2-10z_j+1))\ .$$

Let $V_{s\ell}$ denote the (s,ℓ) element of the covariance matrix Σ. Then, from (4. 17) we have for $s = \ell$

$$V_{ss} = \frac{1}{n} \mathrm{Cov}_{\theta=0}\left(\frac{\partial \ln L}{\partial \theta_s}\Big|_{\theta=0},\ \frac{\partial \ln L}{\partial \theta_\ell}\Big|_{\theta=0}\right)$$

$$= E_{\theta=0}((1-\delta)(\frac{1}{z^2}(\int_0^z \pi_s(t)dt)^2 - \frac{1}{z}\int_0^z \pi_s^2(t)dt)) + 1$$

$$= E_{\theta=0}((1-\delta)A_{ss}(z)) + 1$$

$$= \sum_{\delta=0}^{1} \int_0^\infty (1-\delta) A_{ss}(F_0(y)) \ell(y, \delta) dy + 1$$

$$= \int_0^\infty A_{ss}(F_0(y)) \ell(y, 0) dy + 1$$

$$= \int_0^\infty A_{ss}(F_0(y))\, F_0(y) \;(-dH(y)) + 1,$$

$s = 1, 2, 3, 4,$ where $\ell(y, \delta)$ is the joint density of Y, δ and

$$A_{ss}(z) = (\frac{1}{z^2} (\int_0^z \pi_s(t)dt)^2 - \frac{1}{z} \int_0^z \pi_s^2(t)dt) \quad .$$

Using these relations and (4. 19) for $s = 1, 2, 3$ and 4 gives

$$V_{11} = 1 + \int_0^\infty (F_0(y))^3 dH(y)$$

$$V_{22} = 1 + \int_0^\infty (16(F_0(y))^5 - 30(F_0(y))^4 + 15(F_0(y))^3) dH(y)$$

$$V_{33} = 1 + \int_0^\infty (225(F_0(y))^7 - 700(F_0(y))^6 + 812(F_0(y))^5 - 420(F_0(y))^4$$
$$+ 84(F_0(y))^3) dH(y)$$

$$V_{44} = 1 + \int_0^\infty (3136(F_0(y))^9 - 13230(F_0(y))^8 + 22815(F_0(y))^7 - 20580(F_0(y))^6$$

$$+ 10260(F_0(y))^5 - 2700(F_0(y))^4 + 300(F_0(y))^3)dH(y).$$

Similarly, the covariance elements

$$V_{s\ell} = \mathrm{Cov}_{\theta=0}\left(\frac{\partial \ln \ell}{\partial \theta_s}\Big|_{\theta=0}, \frac{\partial \ln \ell}{\partial \theta_\ell}\Big|_{\theta=0}\right)$$

$$= E_{\theta=0}((1-\delta)(\frac{1}{z^2}\int_0^z \pi_\ell(t)dt \int_0^z \pi_s(t)dt - \frac{1}{z}\int_0^z \pi_s(t)\pi_\ell(t)dt))$$

$$= E_{\theta=0}((1-\delta)(A_{s\ell}(z))$$

$$= \int_0^\infty A_{s\ell}(F_0(y))\ell(y, 0)dy$$

$$= \int_0^\infty A_{s\ell}(F_0(y))F_0(y)(-dH(y)) ,$$

where

$$A_{s\ell}(z) = \frac{1}{z^2}\int_0^z \pi_s(t)dt \int_0^z \pi_\ell(t)dt - \frac{1}{z}\int_0^z \pi_s(t)\pi_\ell(t)dt ,$$

reduce, for $1 \leq s < \ell \leq 4$, to

$$V_{12} = \int_0^\infty \sqrt{15}\,((F_0(y))^4 - (F_0(y))^3)dH(y)$$

$$V_{13} = \int_0^\infty \sqrt{21}\,(3(F_0(y))^5 - 5(F_0(y))^4 + 2(F_0(y))^3)\,dH(y)$$

$$V_{14} = \int_0^\infty \sqrt{3}\,(28(F_0(y))^6 - 63(F_0(y))^5 + 45(F_0(y))^4 - 10(F_0(y))^3)\,dH(y)$$

$$V_{23} = \int_0^\infty \sqrt{35}\,(10(F_0(y))^6 - 25(F_0(y))^5 + 21(F_0(y))^4 - 6(F_0(y))^3)\,dH(y)$$

$$V_{24} = \int_0^\infty \sqrt{5}\,(96(F_0(y))^7 - 294(F_0(y))^6 + 333(F_0(y))^5 - 165(F_0(y))^4 + 30(F_0(y))^3)\,dH(y)$$

$$V_{34} = \int_0^\infty \sqrt{7}\,(315(F_0(y))^8 - 1155(F_0(y))^7 + 1668(F_0(y))^6 - 1188(F_0(y))^5 + 420(F_0(y))^4 - 60(F_0(y))^3)\,dH(y).$$

From the equations above, the covariance matrix $\Sigma = [V_{s\ell}]$ is seen to depend on the unknown censoring distribution $H(y)$. Hence, $H(y)$ must be estimated in order to estimate the covariance matrix Σ. Let $H_n(y)$ be any estimator for $H(y)$ such that $H_n(y) \xrightarrow{p} H(y)$; for example, the product-limit estimator defined in Section 3.3 may be used. Define the covariance estimator $\hat{\Sigma} = [\hat{V}_{s\ell}]$ where $\hat{V}_{s\ell}$ is the same as $V_{s\ell}$ except with $H(y)$ replaced by the estimator $H_n(y)$. Applying the Lemma 2.7, it is straightforward to show

$$\int_0^{\infty} F^k(y)(-dH_n(y)) \xrightarrow{p} \int F^k(y)(-dH(y))$$

which implies $\hat{\Sigma} \xrightarrow{p} \Sigma$. Hence $\hat{\Sigma}^{-1} \xrightarrow{p} \Sigma^{-1}$. Therefore, from (4.18),

$$W_n^{*'}\hat{\Sigma}^{-1}W_n^{*} = W_n^{*'}(\Sigma^{-1}+o_p(1))W_n^{*}$$

$$= W_n^{*'}\Sigma^{-1}W_n^{*} + o_p(1) \xrightarrow{d} \chi^2(r) . \qquad (4.20)$$

Thus, under H_0 and random censorship the statistic $W_n^{*'}\hat{\Sigma}^{-1}W_n^{*}$ has been shown to have a limiting chi-square distribution with r degrees of freedom.

Example. For the sample described at the end of Section 3.5, the generalized smooth tests of fit for the exponential and Weibull distributions are evaluated. Since the null hypothesis must be completely specified in our generalization, specified values have been assigned to the parameters of the exponential and Weibull distributions. These specified values include the minimum chi-square estimates given in Section 3.5:

$$\hat{\theta} = 369.91, \quad \hat{\theta}_1 = 306.47 \quad \text{and} \quad \hat{\theta}_2 = 0.062.$$

In Table 2 below, evaluations of the generalized smooth tests of fit for specified exponential distributions and a Weibull distribution are summarized using the $r = 1, 2, 3$, and 4 order transformed Legendre polynomials. The upper $\alpha = 0.05$ quantiles of the chi-square distributions are given in the last column in Table 2. If the null distributions had been specified a priori, those cases where the test statistic exceeds the corresponding chi-square value would be rejected at the $\alpha = 0.05$ significance level. As expected, the distributions fitted by minimum chi-square estimation yield relatively small values for the test statistics.

Table 2. Smooth goodness of fit tests for specified exponential and Weibull distributions for the breast cancer sample.

	$H_0: F_\theta(x) = e^{-\frac{x}{\theta}}$				$H_0: F_{\theta_1,\theta_2}(x) = e^{-\frac{x^{\theta_2}}{\theta_1}}$	Chi-square Value
r	$\theta = 369.91$	$\theta = 300$	$\theta = 250$	$\theta = 200$	$\theta_1 = 306.47$, $\theta_2 = 0.962$	(0.05)
1	0.0973	3.8435	11.9001	28.6240	0.1089	3.841
2	0.3360	4.5327	13.6946	33.9522	0.1401	5.991
3	0.3384	4.5567	13.8949	35.0533	0.1491	7.815
4	1.0875	5.5073	14.9572	35.9745	1.2120	9.488

4.3. Limiting Distribution for a Sequence of Alternative Hypotheses

For $\theta = 0$ (null hypothesis H_0) we have seen from (4.11) and (4.18) that in either the uncensored or the random censorship cases the statistic $W_n^{*'} \Sigma^{-1} W_n^*$ has a limiting chi-square distribution with r degrees of freedom. It is of interest to find the limiting distribution of $W_n^{*'} \Sigma^{-1} W_n^*$ for the sequence of alternatives

$$H_a : \theta = \frac{\Delta}{\sqrt{n}} \tag{4.21}$$

where $\Delta = (\Delta_1, \ldots, \Delta_r)'$. But first, we determine the limiting distribution of W_n^* by using the moment generating function approach for the sequence of alternatives $\{\frac{\Delta}{\sqrt{n}}\}$.

Theorem 4.1. Let L_θ and W_n^* be defined as in (4.12) and (4.13), and

$$\Sigma = \mathrm{Cov}(W_n^*) = [-E_{\theta=0}(\frac{1}{n} \frac{\partial^2 \ln L_\theta}{\partial\theta_s \partial\theta_\ell}|_{\theta=0})]$$

$$= [-E_{\theta=0}(\frac{\partial^2 \ln \ell_i}{\partial\theta_s \partial\theta_\ell}|_{\theta=0})] .$$

Also let $M_{W_n^*}(t, \theta)$ represent the moment generating function of W_n^*, where $t = (t_1, \ldots, t_r)$ and θ is defined in (4.21). Assume

$M_{W_n^*}(t, \theta)$ exists for all n in some neighborhood $t = 0$ and $\lim_{n \to \infty} M_{W_n^*}(t, \theta)$ exists. If

$$\frac{\partial \ln \ell_j}{\partial \theta_s}\Big|_{\theta=0}, \quad j = 1, \ldots, n, \quad \frac{1}{n} \frac{\partial^2 \ln L_\theta}{\partial \theta_s \partial \theta_\ell}\Big|_{\theta=0},$$

and

$$\frac{1}{n} \frac{\partial^3 \ln L_\theta}{\partial \theta_s \partial \theta_\ell \partial \theta_m}\Big|_{\theta},$$

where θ is in the closure of some neighborhood of $\theta = 0$, $s, \ell, m = 1, \ldots, r$, are all uniformly bounded, then

$$\lim_{n \to \infty} M_{W_n^*}(t, \theta) = e^{t'\Sigma\Delta + \frac{1}{2} t'\Sigma t} .$$

That is, W_n^* has a limiting multivariate normal distribution with mean vector $\Sigma\Delta$ and covariance matrix Σ under the sequence of alternatives (4.21).

<u>Proof</u>. For $\theta = \frac{\Delta}{\sqrt{n}}$ the moment generating function can be written

$$M_{W_n^*}(t,\theta) = E_\theta(e^{t'W_n^*})$$

$$= \int e^{t'W_n^*} L_\theta dW_n^*$$

$$= \int e^{t'W_n^* + \ln L_\theta - \ln L_0} L_0 dW_n^*$$

$$= E_{\theta=0}(e^{t'W_n^* + \ln L_\theta - \ln L_0}) ,$$

where $L_0 = L_{\theta=0}$. Expanding the function $\ln(L_\theta)$ at the point $\theta = 0$ gives

$$\ln L_\theta = \ln L_0 + \sum_{s=1}^{r} \frac{\Delta_s}{\sqrt{n}} \frac{\partial \ln L_\theta}{\partial \theta_s}\Big|_{\theta=0} + \frac{1}{2}\sum_{s=1}^{r}\sum_{\ell=1}^{r} \frac{\Delta_s \Delta_\ell}{n} \frac{\partial^2 \ln L_\theta}{\partial\theta_s \partial\theta_\ell}\Big|_{\theta=0}$$

$$+ \frac{1}{6}\sum_{s=1}^{r}\sum_{\ell=1}^{r}\sum_{m=1}^{r} \frac{\Delta_s\Delta_\ell\Delta_m}{n^{3/2}} \frac{\partial^3 \ln L_\theta}{\partial\theta_s\partial\theta_\ell\partial\theta_m}\Big|_{\theta=\theta*}$$

$$= \ln L_0 + \Delta' W_n^* - \frac{1}{2}\Delta'\Sigma^*\Delta + o_a(1)$$

$$= \ln L_0 + \Delta' W_n^* - \frac{1}{2}\Delta'\Sigma\Delta - \frac{1}{2}\Delta'(\Sigma^* - \Sigma)\Delta + o_a(1)$$

$$= \ln L_0 + \Delta' W_n^* - \frac{1}{2}\Delta'\Sigma\Delta + o_a(1) ,$$

where $\theta^* = 0 + \xi(\theta - 0)$ for $|\xi| < 1$, $\Sigma^* = [-\frac{1}{n}\frac{\partial^2 \ln L_\theta}{\partial\theta_s\partial\theta_\ell}\Big|_{\theta=0}]$, and $o_a(1)$ is a sequence of sample functions which almost surely converge to 0 (from the strong law of large numbers we have $\Sigma^* - \Sigma = o_a(1)$). Hence

$$M_{W_n^*}(t,\theta) = E_{\theta=0}\left(e^{-\frac{1}{2}\Delta'\Sigma\Delta + t'W_n^* + \Delta'W_n^* + o_a(1)}\right)$$

$$= e^{-\frac{1}{2}\Delta'\Sigma\Delta} E_{\theta=0}\left(e^{(t+\Delta)'W_n^* + o_a(1)}\right).$$

Applying Lemma 2.8, we find

$$\lim_{n\to\infty} M_{W_n^*}(t,\theta) = e^{-\frac{1}{2}\Delta'\Sigma\Delta} \lim_{n\to\infty} E_{\theta=0}\left(e^{(t+\Delta')W_n^*}\right)$$

$$= e^{-\frac{1}{2}\Delta'\Sigma\Delta} \lim_{n\to\infty} E_{\theta=0}\left(e^{(t+\Delta)'(\frac{1}{\sqrt{n}}\Sigma_{j=1}^n W_j)}\right)$$

$$= e^{-\frac{1}{2}\Delta'\Sigma\Delta} \lim_{n\to\infty} E_{\theta=0}\left(\prod_{j=1}^n e^{(t+\Delta)'\frac{W_j}{\sqrt{n}}}\right)$$

$$= e^{-\frac{1}{2}\Delta'\Sigma\Delta} \lim_{n\to\infty} \left(E_{\theta=0}\left(e^{(t+\Delta)'W}\right)\right)^n ,$$

where $W = \frac{1}{\sqrt{n}} W_j$ with W_j defined in (4.13). Using a Taylor series expansion of $e^{(t+\Delta)'W}$,

$$e^{(t+\Delta)'W} = 1 + (t+\Delta)'W + \frac{1}{2}((t+\Delta)'W)^2 + \int_0^1 \frac{(1-u)^2}{2} \frac{d^3}{du^3}(e^{u\,(t+\Delta)'W})du,$$

yields

$$E_{\theta=0}(e^{(t+\Delta)'W}) = 1 + E_{\theta=0}((t+\Delta)'W) + E_{\theta=0}(\frac{1}{2}((t+\Delta)'W)^2)$$

$$+ E_{\theta=0}(\int_0^1 \frac{(1-u)^2}{2} \frac{d^3}{du^3} (e^{u\ (t+\Delta)'W})du)$$

$$= 1 + 0 + \frac{1}{2n}(t+\Delta)'\Sigma(t+\Delta) + o(n^{-1}) ,$$

which follows from

$$E_{\theta=0}((t+\Delta)'W) = \sum_{s=1}^{r} (t_j+\Delta_j) \frac{1}{\sqrt{n}} E_{\theta=0}(\frac{\partial \ln \ell_j}{\partial \theta_s}|_{\theta=0})$$

$$= 0 ,$$

$$E_{\theta=0}(((t+\Delta)'W)^2) = E_{\theta=0}(\frac{1}{\sqrt{n}} \sum_{s=1}^{r} (t_s+\Delta_s)(\frac{\partial \ln \ell_j}{\partial \theta_s}|_{\theta=0}))^2$$

$$= \frac{1}{n} E_{\theta=0}(\sum_{s=1}^{r} \sum_{\ell=1}^{r} (t_s+\Delta_s) \frac{\partial \ln \ell_j}{\partial \theta_s}|_{\theta=0}$$

$$\times \frac{\partial \ln \ell_j}{\partial \theta_\ell}|_{\theta=0}(t_\ell+\Delta_\ell))$$

$$= \frac{1}{n} \sum_{s=1}^{r} \sum_{\ell=1}^{r} (t_s+\Delta_s)E_{\theta=0}(-\frac{\partial^2 \ln \ell_j}{\partial \theta_s \partial \theta_\ell}|_{\theta=0})(t_\ell+\Delta_\ell)$$

$$= \frac{1}{n}(t+\Delta)'\Sigma(t+\Delta) ,$$

and

$$E_{\theta=0}\left(\int_0^1 \frac{(1-u)^2}{2}\,\frac{d^3}{du^3}\,e^{u(t+\Delta)'W}\,du\right)$$

$$= E_{\theta=0}\left(\int_0^1 \frac{(1-u)^2}{2}\,((t+\Delta)'W)^3 e^{u(t+\Delta)'W}\,du\right)$$

$$= \int_0^1 \frac{(1-u)^2}{2}\,E_{\theta=0}\left(\left(\sum_{s=1}^{r}(t_s+\Delta_s)\frac{1}{\sqrt{n}}\,\frac{\partial \ln \ell_j}{\partial\theta_s}\Big|_{\theta=0}\right)^3 e^{u(t+\Delta)'W}\right)du$$

$$= \frac{1}{n^{3/2}}\,O(1)$$

$$= o(n^{-1})\ .$$

Accordingly,

$$\begin{aligned}
\lim_{n\to\infty} M_{W_n^*}(t,\theta) &= e^{-\frac{1}{2}\Delta'\Sigma\Delta}\lim_{n\to\infty}\left(E_{\theta=0}\left(e^{(t+\Delta)'W}\right)\right)^n \\
&= e^{-\frac{1}{2}\Delta'\Sigma\Delta}\lim_{n\to\infty}\left(1+\frac{1}{n}\left(\frac{1}{2}(t+\Delta)'\Sigma(t+\Delta)+o(n^{-1})\right)\right)^n \\
&= e^{-\frac{1}{2}\Delta'\Sigma\Delta}\,e^{\frac{1}{2}(t+\Delta)'\Sigma(t+\Delta)} \\
&= e^{t'\Sigma\Delta+\frac{1}{2}t'\Sigma t}\ .
\end{aligned}$$

This concludes the proof.

Based on Theorem 4.1, we can write for $\theta = \frac{\Delta}{\sqrt{n}}$, $W_n^* \xrightarrow{d} W^*$, where $W^* \sim N_r(\Sigma\Delta, \Sigma)$. It is well known that $W^{*'}\Sigma^{-1}W^*$ is non-central chi-square distributed with $r = \text{rank}\,(\Sigma)$ degrees of freedom and noncentrality parameter $(\Sigma\Delta)'\Sigma^{-1}(\Sigma\Delta) = \Delta'\Sigma\Delta$. The fact that

$W_n^{*'}\Sigma^{-1}W_n^*$ is a continuous function of W_n^* implies $W_n^{*'}\Sigma^{-1}W_n^* \xrightarrow{d} W^{*'}\Sigma^{-1}W^*$. That is, $W_n^{*'}\Sigma^{-1}W_n^*$ has a limiting non-central chi-square distribution with r degrees of freedom and noncentrality parameter $\Delta'\Sigma\Delta$. In addition, if $H(y)$ is replaced by a consistent estimator $H_n(y)$ in Σ, the corresponding statistic $W_n^{*'}\hat{\Sigma}^{-1}W_n^*$ will have the same limiting noncentral chi-square distribution as the statistic $W_n^{*'}\Sigma^{-1}W_n^*$.

The remainder of this section is devoted to verification of the assumptions in the Theorem 4.1; that is,

$$\frac{\partial \ln \ell_j}{\partial \theta_s}\Big|_{\theta=0}, \quad j = 1, \ldots, n, \quad \frac{1}{n}\,\frac{\partial^2 \ln L_\theta}{\partial\theta_s\partial\theta_\ell}\Big|_{\theta=0},$$

and

$$\frac{1}{n}\,\frac{\partial^3 \ln L_\theta}{\partial\theta_s\partial\theta_\ell\partial\theta_m}\Big|_{\theta},$$

where θ is in the closure of some neighborhood $\theta = 0$, are uniformly bounded. We will use superscripts s, ℓ, m to represent derivatives with respect to $\theta_s, \theta_\ell, \theta_m$ and write $z = F(y)$.

First, we derive some results for later reference. By writing

$$A(\theta) = c^{-1}(\theta) = \int_0^1 e^{\Sigma_{i=1}^r \theta_i \pi_i(z)}\, dz,$$

where $c(\theta)$ is defined in Section 4.1, we may express

$$c^{s}(\theta) = -\frac{1}{A^{2}(\theta)} A^{s}(\theta)$$

$$c^{s\ell}(\theta) = \frac{2}{A^{3}(\theta)} A^{s}(\theta)A^{\ell}(\theta) - \frac{1}{A^{2}(\theta)} A^{s\ell}(\theta)$$

$$c^{s\ell m}(\theta) = -\frac{6}{A^{4}(\theta)} A^{s}(\theta)A^{\ell}(\theta)A^{m}(\theta) + \frac{2}{A^{3}(\theta)} (A^{s}(\theta)A^{\ell m}(\theta)$$

$$+ A^{s\ell}(\theta)A^{m}(\theta) + A^{sm}(\theta)A^{\ell}(\theta)) - \frac{1}{A^{2}(\theta)} A^{s\ell m}(\theta) ,$$

where, for $s, \ell, m = 1, \ldots, r$

$$A^{s}(\theta) = \int_0^1 e^{\Sigma_{i=1}^{r} \theta_i \pi_i(z)} \pi_s(z)dz$$

$$A^{s\ell}(\theta) = \int_0^1 e^{\Sigma_{i=1}^{r} \theta_i \pi_i(z)} \pi_s(z)\pi_\ell(z)dz$$

$$A^{s\ell m}(\theta) = \int_0^1 e^{\Sigma_{i=1}^{r} \theta_i \pi_i(z)} \pi_s(z)\pi_\ell(z)\pi_m(z)da.$$

Application of the mean value theorem for integrals gives

$$\frac{\partial \ln \ell_j}{\partial \theta_s}\Big|_{\theta=0} = \frac{\partial}{\partial \theta_s} (\delta_j \ln g_\theta(y_j)+(1-\delta_j) \ln G_\theta(y_j))\Big|_{\theta=0}$$

$$= (\delta_j(\frac{1}{c(\theta)} c^{s}(\theta)+\pi_s(F(y_j)))+(1-\delta_j)\frac{1}{G_\theta(y_j)} \frac{\partial}{\partial \theta} G(y_j))\Big|_{\theta=0}$$

$$= \delta_j \pi_s(z_j) + (1-\delta_j) \frac{1}{z_j} \int_0^{z_j} \pi_s(z)dz$$

$$= \delta_j \pi_s(z_j) + (1-\delta_j) \frac{1}{z_j} z_j \pi_s(z^*)$$

$$= \delta_j \pi_s(z_j) + (1-\delta_j) \pi_s(z_j^*) ,$$

where $0 < z_j^* < z_j \leq 1$, $j = 1, \ldots, n$, and hence are bounded.

Similarly, we have for $s \neq \ell$

$$\frac{1}{n} \frac{\partial^2 \ln L_\theta}{\partial\theta_s \partial\theta_\ell} \Big|_{\theta=0} = \frac{1}{n} \sum_{j=1}^{n} 0+(1-\delta_j)(-\frac{1}{z_j^2} \int_0^{z_j} \pi_s(z)dz \int_0^{z_j} \pi_\ell(z)dz$$

$$+ \frac{1}{z_j} \int_0^{z_j} \pi_s(z)\pi_\ell(z)dz)$$

$$= \frac{1}{n} \sum_{j=1}^{n} (1-\delta_j)(-\frac{1}{z_j^2} z_j \pi_s(z_j^1) z_j \pi_s(z_j^2)$$

$$+ \frac{1}{z_j} z_j \pi_s(z_j^3)\pi_\ell(z_j^3))$$

$$= \frac{1}{n} \sum_{j=1}^{n} (1-\delta_j)(-\pi_s(z_j^1)\pi_\ell(z_j^2)+\pi_s(z_j^3)\pi_\ell(z_j^3)) ,$$

where z_j^1, z_j^2, z_j^3 represent corresponding mean value points.

Hence,

$$\left|\frac{1}{n}\frac{\partial^2 \ln L_\theta}{\partial\theta_s \partial\theta_\ell}\Big|_{\theta=0}\right| \le \frac{1}{2}\sum_{j=1}^{n}\left(|\pi_s(z_j^1)|\,|\pi_\ell(z_j^2)| + |\pi_s(z_j^3)|\,|\pi_\ell(z_j^3)|\right)$$

$$\le \frac{1}{n}\sum_{j=1}^{n}(2|\pi_s(z^*)|\,|\pi_\ell(z^\Delta)|$$

$$= M ,$$

where $\pi_s(z^*) = \sup_{0 \le z \le 1} \pi_s(z)$, $\pi_\ell(z^\Delta) = \sup_{0 \le z \le 1} \pi_\ell(z)$ and $M = 2|\pi_s(z^*)|\ |\pi_\ell(z^\Delta)|$. In the case of $s = \ell$,

$$\frac{1}{n}\frac{\partial^2 \ln L_\theta}{\partial\theta_s^2}\Big|_{\theta=0} = \frac{1}{n}\sum_{j=1}^{n}\ (1-\delta_j)(\frac{-1}{z_j^2}(\int_0^{z_j}\pi_s(z)dz)^2 + \frac{1}{z_j}\int_0^{z_j}\pi_s^2(z)dz) \quad + 1$$

and using a similar argument as in the case $s \neq \ell$, we can find some constant M' such that

$$\left|\frac{1}{n}\frac{\partial^2 \ln L_\theta}{\partial\theta_s^2}\Big|_{\theta=0}\right| \le M' .$$

Hence in general,

$$\frac{1}{n}\frac{\partial^2 \ln L_\theta}{\partial\theta_s\partial\theta_\ell}\Big|_{\theta=0}$$

is uniformly bounded for $s, \ell = 1, \ldots, r$.

Finally, for θ in the closure of some neighborhood $\theta = 0$,

$$\frac{1}{n} \frac{\partial^3 \ln L_\theta}{\partial\theta_s \partial\theta_\ell \partial\theta_m}\bigg|_\theta = \frac{1}{n} \sum_{j=1}^{n} \delta_j \frac{\partial^3}{\partial\theta_s \partial\theta_\ell \partial\theta_m} \ln g_\theta(y_j)$$

$$+ (1-\delta_j) \frac{\partial^3}{\partial\theta_s \partial\theta_\ell \partial\theta_m} \ln G_\theta(y_j) \quad \Big|_\theta ,$$

where

$$\frac{\partial^3}{\partial\theta_s \partial\theta_\ell \partial\theta_m} \ln g_\theta(y_j)$$

is a function of only $c(\theta)$, $c^s(\theta)$, $c^{s\ell}(\theta)$, $c^{s\ell m}(\theta)$, and is therefore constant for given θ, say M_g. For the second term

$$\frac{\partial^3}{\partial\theta_s \partial\theta_\ell \partial\theta_m} \ln G_\theta(y_j) = \frac{\partial^3}{\partial\theta_s \partial\theta_\ell \partial\theta_m} \ln Q_\theta(z_j)$$

$$= \frac{2}{Q_\theta^3(z_j)} Q_\theta^s(z_j)\, Q_\theta^\ell(z_j) Q_\theta^m(z_j)$$

$$- \frac{1}{Q_\theta^2(z_j)} (Q_\theta^{s\ell}(z_j) Q_\theta^m(z_j) + Q_\theta^{sm}(z_j) Q_\theta^\ell(z_j)$$

$$+ Q_\theta^{\ell m}(z_j) Q_\theta^s(z_j))$$

$$+ \frac{1}{Q_\theta(z_j)} Q_\theta^{s\ell m}(z_j) .$$

Again, applying the mean value theorem for integrals and writing z with superscripts corresponding to mean value points, we can write

$$Q_\theta(z_j) = \int_0^{z_j} c(\theta)\, e^{\Sigma_{i=1}^r \theta_i \pi_i(z)}\, dz$$

$$= z_j\, c(\theta)\, e^{\Sigma_{i=1}^r \theta_i \pi_i(z^1)}$$

$$= z_j\, M_j\ ,$$

where

$$M_j = c(\theta)\, e^{\Sigma_{i=1}^r \theta_i \pi_i(z^1)}$$

is bounded away from zero, and

$$Q^s(z_j) = \int_0^{z_j} c^s(\theta)\, e^{\Sigma_{i=1}^r \theta_i \pi_i(z)} + c(\theta)\, e^{\Sigma_{i=1}^r \theta_i \pi_i(z)}\, \pi_s(z)\ dz$$

$$= z_j\left(c^s(\theta)\, e^{\Sigma_{i=1}^r \theta_i \pi_i(z^2)} + c(\theta)\, e^{\Sigma_{i=1}^r \theta_i \pi_i(z^3)}\, \pi_s(z^3)\right)$$

$$= z_j\, M_{s,j} \qquad \text{for} \quad s = 1, \ldots, r,$$

where

$$M_{s,j} = c^s(\theta)\, e^{\Sigma_{i=1}^r \theta_i \pi_i(z^2)} + c(\theta)\, e^{\Sigma_{i=1}^r \theta_i \pi_i(z^3)} \cdot \pi_s(z^3)$$

are bounded. By the same technique, we obtain for $s, \ell, m = 1, \ldots, r$

$$Q^{s\ell}(z_j) = z_j M_{s\ell, j}\ ,$$

$$Q^{s\ell m}(z_j) = z_j M_{s\ell m, j}\ ,$$

where $M_{s\ell,j}$ and $M_{s\ell m,j}$ are both bounded. Hence,

$$\frac{\partial^3}{\partial\theta_s\partial\theta_\ell\partial\theta_m} \ln G_\theta(y_j)\big|_\theta = \frac{2}{(z_jM_j)^3}(z_j)^3 M_{s,j}M_{\ell,j}M_{m,j}$$

$$- \frac{1}{(z_jM_j)^2} z_j^2 (M_{s\ell,j}M_{m,j} + M_{sm,j}M_{\ell,j}$$

$$+ M_{\ell m,j}M_{s,j}) + \frac{1}{z_jM_j} z_j M_{s\ell m,j}$$

$$= \frac{1}{M_j^3} M_{s,j}M_{\ell,j}M_{m,j}$$

$$- \frac{1}{M_j^2}(M_{s\ell,j}M_{m,j} + M_{sm,j}M_{\ell,j} + M_{\ell m,j}M_{s,j})$$

$$+ \frac{1}{M_j} M_{s\ell m,j} .$$

Therefore, we can find some constant M_T which is independent of the z_j's such that

$$\left| \frac{\partial^3}{\partial\theta_s\partial\theta_\ell\partial\theta_m} \ln G_\theta(y_j)\big|_\theta \right| \le M_T .$$

Finally,

$$\left|\frac{1}{n}\frac{\partial^3 \ln L_\theta}{\partial\theta_s \partial\theta_\ell \partial\theta_m}\right|_\theta\Big|$$

$$\leq \frac{1}{n}\sum_{j=1}^{n}\left(\left|\frac{\partial^3}{\partial\theta_s\partial\theta_\ell\partial\theta_m}\ln g_\theta(y_j)\right|_\theta\Big| + \left|\frac{\partial^3}{\partial\theta_s\partial\theta_\ell\partial\theta_m}\ln G_\theta(y_j)\right|_\theta\Big|\right)$$

$$\leq \frac{1}{n}\sum_{j=1}^{n}(M_g+M_T)$$

$$= M_g + M_T ,$$

i.e.,

$$\left|\frac{1}{n}\frac{\partial^3 \ln L_\theta}{\partial\theta_s\partial\theta_\ell\partial\theta_m}\right|_\theta \quad \text{is uniformly bounded.}$$

BIBLIOGRAPHY

Barr, D. R. and Davidson, T. (1973). A Kolmogorov-Smirnov test for censored samples. Technometrics, 15:739-757.

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

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

Breslow, N. and Crowley, J. (1974). A large sample study of the life table and product limit estimates under random censorship. Annals of Statist., 2:437-453.

Brunk, H.D., Thomas, D.R., Elashoff, R.M. and Zippin, C. (1973). Computer-aided prognosis. Technical report No. 40, Department of Statistics, Oregon State University.

Burrill, C.W. (1972). Measure, Integration, and Probability. McGraw-Hill Co., New York.

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

Chung, K.L. (1968). A Course in Probability Theory. Harcourt, Brace & World, Inc., New York.

Cramer, H. (1966). Mathematical Methods of Statistics. Princeton University Press, Princeton.

Cutler, S.J. and Myers, M.H. (1967). Clinical classification of extent of disease in cancer of the breast. J. Nat. Cancer Inst., 39:193-206.

Efron, B. (1967). The two sample problem with censored data. Proceedings of the Fifth Berkeley Symposium 4:831-853.

Fisher, R.A. (1922). On the interpretation of χ^2 from contingency tables, and the calculation of P. Journal of the Royal Statistical Society, 85:87.

Fisher, R.A. (1924). The conditions under which χ^2 measures the discrepancy between observation and hypothesis. Journal of the Royal Statistical Society, 87:442.

Graybill, F.A. (1969). Introduction to Matrices with Applications in Statistics. Wadsworth Publishing Co. Inc., Belmont.

Hartley, H.O. and Pfaffenberger, R.C. (1972). On a family of lesser known goodness of fit criteria. Submitted to Biometrika.

Kaplan, E.L. and Meier, P. (1958). Nonparametric estimation from incomplete observations. J. Amer. Statist. Assoc., 53:457-480.

Koch, G.G., Johnson, W.D. and Tolley, H.D. (1972). A linear models approach to analysis of survival and extent of disease in multidimensional contingency tables. J. Amer. Statist. Assoc., 67:783-796.

Lehmann, E.L. (1959). Testing Statistical Hypotheses. John Wiley & Sons, Inc., New York.

Lurie, D. and Hartley, H.O. (1972). A goodness of fit test for censored data. The joint statistical meetings, Montreal, Canada.

Mann, H.B. and Wald, A. (1943). On stochastic limit and order relationships. Ann. Math. Statist., 14:217-226.

Neyman, J. (1937). Smooth tests for goodness of fit. Skand. Aktuaritidskr., 20:291-319.

Pearson, K. (1900). On a 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 in random sampling. Phil. Mag., 50:157-175.

Renyi, A. (1970). Foundations of Probability. Holden-Day, Inc. San Francisco.

Wilks, S.S. (1963). Mathematical Statistics. John Wiley & Sons, Inc., New York.