CHI-SQUARE TESTS OF FIT FOR TYPE II CENSORED DATA[1]

Censored Data Chi-Square Tests

By Daniel P. Mihalko[2] and David S. Moore
Purdue University

Department of Statistics

Division of Mathematical Sciences

Mimeograph Series #505

August, 1977

AMS 1970 subject classifications. Primary 62G10; Secondary 62E20.

Key words and phrases. Goodness of fit, chi-square tests, censored data.

[1]Research supported by the Air Force Office of Scientific Research, AFSC, USAF, under Grant AFOSR-77-3291. The United States Government is authorized to reproduce and distribute reprints for governmental purposes.

[2]Now at the University of Nebraska.

CHI-SQUARE TESTS OF FIT FOR TYPE II CENSORED DATA[1]

Censored Data Chi-Square Tests

By Daniel P. Mihalko[2] and David S. Moore
Purdue University

The theory of general chi-square statistics for testing fit to parametric families of distributions is extended to samples censored at sample quantiles. Data-dependent cells with sample quantiles as cell boundaries are employed. Asymptotic distribution theory is given for statistics in which unknown parameters are estimated by estimators asymptotically equivalent to linear combinations of functions of order statistics. Emphasis is placed on obtaining statistics having a chi-square limiting null distribution. Examples of such statistics for testing the fit of Type II censored samples to the negative exponential, normal, two-parameter uniform and two-parameter Weibull families are given.

AMS 1970 subject classifications. Primary 62G10; Secondary 62E20.

Key words and phrases. Goodness of fit, chi-square tests, censored data.

[1]Research supported by the Air Force Office of Scientific Research, AFSC, USAF, under Grant AFOSR-77-3291. The United States Government is authorized to reproduce and distribute reprints for governmental purposes.

[2]Now at the University of Nebraska

1. Introduction. Randomly censored data occur frequently in engineering (life testing and reliability) and medical studies. Since inference procedures for such data commonly make distributional assumptions, a considerable body of recent literature concerns tests of fit for censored samples. In this paper we extend the applicability of chi-square tests of fit to data censored (on one or both sides) at sample percentiles. This is commonly called Type II censoring. Data censored in this way arise in engineering settings. Medical data typically display more complex random censoring, as when studies of survival after treatment encounter dropouts and deaths from other causes. Such general random censoring is not considered here. Our goal is to provide tests of fit to parametric families of distributions using test statistics having chi-square limiting null distributions. Such tests can be used with standard tables and, in many cases, the test statistic itself requires only simple computations.

Much of the literature on tests of fit for censored data considers the special case of testing fit to a completely specified distribution (e.g., Lurie, Hartley and Stroud (1974), Kozial and Byar (1975), Kozial and Green (1976)). Tests for the more useful composite hypothesis case encounter the dependence of the large sample distribution on the family tested, so that separate tables of critical points are required for each hypothesized family. This circumstance is familiar in the full sample case to users of (among others) tests based on the empirical distribution function. For censored samples, the distribution of most available tests of fit also depends on the degree of censoring, as evidenced by the tables of critical points in Pettit (1976) and Smith and Bain (1976). Thus a test of fit having a standard tabled distribution is even more desirable than in the full sample case. Turnbull and Weiss (1976) offer a generalized likelihood ratio test which not only has a chi-square distribution in large samples but also applies to some types of more general random censoring.

However, they assume that the observed variables are discrete with finite range, and their test statistics must usually be obtained by numerical solution of equations.

The extension of chi-square tests to Type II censored data depends on several recent advances in the theory of such tests. The first, due to A. R. Roy (1956) and G. S. Watson (see [23] for references), is the use of data-dependent cells. Type II censored data lend themselves naturally to the use of sample quantiles as cell boundaries, with the censored observations falling in the extreme cells. Suppose, then, that of a random sample $X_1,\ldots,X_n$ we observe only the order statistics

(1.1) $$X_{([n\alpha]+1)} < X_{([n\alpha]+2)} < \ldots < X_{([n\beta])}$$

where $0 \leq \alpha < \beta \leq 1$ and $[x]$ is the greatest integer in x. We form M cells having boundaries

$$-\infty = \xi_{0n} < \xi_{1n} < \ldots < \xi_{M-1,n} < \xi_{Mn} = \infty$$

where $\xi_{in} = X_{([n\delta_i])}$ is the sample δ_i-quantile from $X_1,\ldots,X_n$ and $0 = \delta_0 < \delta_1 < \ldots < \delta_{M-1} < \delta_M = 1$. To accommodate nontrivial left censoring ($\alpha>0$), right censoring ($\beta<1$), or both, with a single notation, we adopt the convention that $\alpha=\delta_1$ when $\alpha>0$ and otherwise $\alpha = \delta_0 = 0$; similarly, $\beta = \delta_{M-1}$ when $\beta<1$ and otherwise $\beta = \delta_M = 1$. The observed frequency N_{in} in the ith cell $E_i = (\xi_{i-1,n}, \xi_{in}]$ is nonrandom, $N_{in} = [n\delta_i]-[n\delta_{i-1}]$. In particular, the left-censored $[n\alpha]$ observations and the right-censored $n-[n\beta]$ observations occupy the extreme cells.

We wish to test the composite null hypothesis that the distribution of the X_i is a member of the family of continuous distribution functions $\{F(x,\theta): \theta \text{ in } \Omega\}$, where Ω is an open set in Euclidean m-space R^m. The parameter θ must be estimated by an estimator θ_n which is a function of the observed order statistics (1.1). Chi-square statistics for data-dependent cells are formed by "forgetting" that the cells are functions of the data. The ith "estimated cell probability" under H_0 is therefore

(1.2) $$p_{in} = F(\xi_{in},\theta_n)-F(\xi_{i-1,n},\theta_n)$$

These are random, unlike the cell frequencies. Chi-square statistics are nonnegative definite quadratic forms in the standardized cell frequencies $(N_{in}-np_{in})/(np_{in})^{\frac{1}{2}}$. The second development in chi-square tests that we exploit here is the use of quadratic forms other than the Pearson sum of squares. The goal of this generalization is to find a quadratic form having a chi-square limiting null distribution for general estimators θ_n, just as the Pearson statistic does for minimum chi-square estimation. In the full sample case, the use of appropriate quadratic forms to obtain a chi-square distribution was initiated by Robson and his students ([7], [16]) and treated in some generality by Moore (1977).

The development in this paper follows the pattern laid down for general chi-square statistics for full samples by Moore and Spruill (1975). Their results require independent observations, and so do not include the results given here. Section 2 contains results on the asymptotic multivariate normality of the vector of standardized cell frequencies, for several classes of estimators θ_n. Based on these results, Section 3 discusses the large-sample behavior of several chi-square statistics for Type II censored data. The specific statistics are censored-sample analogs of the classical Pearson-Fisher statistic, the Pearson statistic using maximum likelihood estimation (studied in the full sample case by Chernoff and Lehmann (1954)), the Rao-Robson (1974) statistic for maximum likelihood estimation, and the Dzhaparidze-Nikulin (1974) statistic for arbitrary $n^{\frac{1}{2}}$-consistent estimators. As might be expected, the behavior of these statistics for censored samples parallels that of their full sample analogs. Section 4 applies the general results to obtain tests of fit for censored samples to the negative exponential, normal, Weibull and uniform families of distributions.

We remark that the approach taken here applies also to "multiple Type II censoring", in which observations between several sets of sample percentiles are unavailable. It is necessary only to take each unobserved inter-percentile group as a cell. This is conceptually quite similar to the generality of censoring allowed in the procedures of Turnbull and Weiss (1976).

2. Asymptotic normality of standardized cell frequencies. We will be concerned with the large sample behavior of chi-square statistics when $X_1,\ldots,X_n$ have distribution function $F(x,\theta_0)$, so that θ_0 is the "true" parameter value. Our major conclusions will not depend on the particular θ_0, as when statistics have the same limiting chi-square distribution for all θ_0 in Ω. Similarly, assumptions made locally at θ_0 must in practice hold everywhere in Ω. Denote by x_i the population δ_i-quantile of $F(x,\theta_0)$, so that $x_0 = -\infty$, $x_M = \infty$ and

$$x_i = \min \{x: \ F(x,\theta_0) = \delta_i\} \qquad i=1,\ldots,M-1.$$

For any vector of cell boundaries $\xi = (\xi_1,\ldots,\xi_{M-1})^T$, define the cell probabilities

$$p_i(\xi,\theta) = F(\xi_i,\theta)-F(\xi_{i-1},\theta)$$

and the M x m matrix $B(\xi,\theta)$ having (i,j)th entry

$$p_i(\xi,\theta)^{-\frac{1}{2}} \frac{\partial p_i(\xi,\theta)}{\partial \theta_j} .$$

Denote the vector of ξ_{in} (the cell boundaries actually used) by ξ_n, and the vector of x_i (their limits in probability under $F(x,\theta_0)$) by ξ_0. By convention, the arguments ξ,θ will be suppressed whenever $\xi=\xi_0$ and $\theta=\theta_0$. Thus $B = B(\xi_0,\theta_0)$ and $p_i = p_i(\xi_0,\theta_0) = \delta_i-\delta_{i-1}$. In particular, all derivatives and expected values not otherwise identified are evaluated at (ξ_0,θ_0). All vectors are column vectors, and derivatives and integrals of vectors are understood componentwise.

The following conditions on $F(x,\theta)$ will be assumed to hold throughout this paper.

(F-1) $F(x,\theta)$ has density function $f(x,\theta)$ which is continuous in (x,θ) in a neighborhood of (x_i,θ_0), i=1,...,M-1.

(F-2) $\partial F(x,\theta)/\partial\theta_j$ exists and is continuous in a neighborhood of (x_i,θ_0), i=1,...,M-1.

(F-3) $f(x_i) > 0$, i=1,...,M-1.

These conditions are sufficient for joint asymptotic normality of $n^{\frac{1}{2}}(\xi_{in}-x_i)$, for convergence in probability of the cells $(\xi_{i-1,n},\xi_{in}]$ to the fixed cells $(x_{i-1},x_i]$, and for convergence in probability of the p_{in} of (1.2) to p_i whenever $\{\theta_n\}$ is a consistent sequence of estimators of θ. Finally, let $V_n(\theta_n)$ be the M-vector of standardized cell frequencies $(N_{in}-np_{in})/(np_{in})^{\frac{1}{2}}$ using random cells and estimating θ by θ_n.

The following basic lemma relates the large sample behavior of $V_n(\theta_n)$ to that of the estimator θ_n and the sample quantiles ξ_{in}. As the proof of the lemma shows, the ith standardized cell frequency when θ_0 is known is (up to $o_p(1)$)

$$-n^{\frac{1}{2}}\{\frac{f(x_i)}{p_i^{\frac{1}{2}}}(\xi_{in}-x_i) - \frac{f(x_{i-1})}{p_i^{\frac{1}{2}}}(\xi_{i-1,n}-x_{i-1})\}.$$

This is the ith component of the M-vector $V_n = V_n(\theta_0)$. The statement of the lemma is thus identical to that of the null case of Theorem 4.1 of Moore and Spruill (1975) for full samples.

LEMMA 2.1. _If_ $n^{\frac{1}{2}}(\theta_n - \theta_0) = O_p(1)$ _under_ $F(x,\theta_0)$, _then_

(2.1) $$V_n(\theta_n) = V_n - Bn^{\frac{1}{2}}(\theta_n - \theta_0) + o_p(1).$$

PROOF: Write

$$n^{-\frac{1}{2}}(N_{in} - p_{in}) = n^{\frac{1}{2}}\left(\frac{[n\delta_i]}{n} - \frac{[n\delta_{i-1}]}{n}\right) - n^{\frac{1}{2}}(F(\xi_{in},\theta_n) - F(\xi_{i-1,n},\theta_n))$$

$$= n^{\frac{1}{2}}\left(\frac{[n\delta_i]}{n} - \frac{[n\delta_{i-1}]}{n}\right) - n^{\frac{1}{2}}(\delta_i - \delta_{i-1})$$

$$-n^{\frac{1}{2}}(F(\xi_{in},\theta_n) - F(x_i)) + n^{\frac{1}{2}}(F(\xi_{i-1,n},\theta_n) - F(x_{i-1})).$$

The first line of the last expression is o(1), and the mean value theorem with (F-1) and (F-2) reduces the second line to

$$-n^{\frac{1}{2}}(f(x_i)(\xi_{in} - x_i) - f(x_{i-1})(\xi_{i-1,n} - x_{i-1}))$$

$$-n^{\frac{1}{2}} \sum_{j=1}^{m} \frac{\partial p_i}{\partial \theta_j} (\theta_n - \theta_0)_i + o_p(1).$$

Since $p_{in} = p_i + o_p(1)$, the lemma follows from this.

When θ_n itself is a finite linear combination of sample quantiles, as are many short-cut estimators, asymptotic normality of $V_n(\theta_n)$ follows at once from (2.1) and the joint asymptotic normality of sample quantiles. Suppose then that the jth component of θ_n has the form

(2.2) $$\theta_{jn} = \sum_{k=1}^{s_j} a_{jkn}\xi_{jkn} + c_{jn} \quad j=1,\ldots,m$$

where ξ_{jkn} is the sample Δ_{jk}-quantile from $X_1,\ldots,X_n$ for $\alpha \leq \Delta_{j1} < \ldots < \Delta_{js_j} \leq \beta$, and the a_{jkn} and c_{jn} are real numbers. Denote the population Δ_{jk}-quantile of $F(x,\theta_0)$ by x_{jk}. The following theorem is proved by direct computation from the limiting law of the sample quantiles.

THEOREM 2.1. *Suppose that* θ_n *satisfies* (2.2), *that there exist constants* a_{jk} *and* c_j *such that* $a_{jkn}-a_{jk} = o(n^{-\frac{1}{2}})$ *and* $c_{jn}-c_j = o(n^{-\frac{1}{2}})$, $j=1,\ldots,m$, *that the* j*th component of* θ_0 *is*

$$\theta_{j0} = \sum_{k=1}^{s_j} a_{jk}x_{jk} + c_j \qquad j=1,\ldots,m$$

and that all $f(x_{jk}) > 0$. *Then*

$$\mathcal{L}\{V_n(\theta_n)\} \to N_M(0,\Sigma)$$

where

$$\Sigma = I_M - qq^T + BA^T + AB^T + BCB^T$$

and $q = (p_1^{\frac{1}{2}},\ldots,p_M^{\frac{1}{2}})^T$, A *is the* $M \times m$ *matrix with* (i,j)th *entry*

$$\sum_{k=1}^{s_j} \frac{a_{jk}}{p_i^{\frac{1}{2}} f(x_{jk})} (\min(\delta_i,\Delta_{jk})-\delta_i\Delta_{jk}),$$

and C *is the* $m \times m$ *matrix with* (i,j)th *entry*

$$\sum_{k=1}^{s_i} \sum_{r=1}^{s_j} \frac{a_{ik}a_{jr}}{f(x_{ik})f(x_{jr})} (\min(\Delta_{ik},\Delta_{jr})-\Delta_{ik}\Delta_{jr}).$$

An estimator of particular interest which can be asymptotically expressed as a linear function of sample quantiles is the grouped data maximum likelihood estimator $\overline{\theta}_n$ (or the asymptotically equivalent minimum chi-square estimator) obtained by "forgetting" the dependence of the cells on the data and solving the usual multinomial likelihood equations

$$\sum_{i=1}^{M} \frac{N_{in}}{p_i(\xi_n,\theta)} \frac{\partial p_i(\xi_n,\theta)}{\partial\theta} = 0.$$

Watson (1958) observed that in the full sample case the use of random cells does not alter the asymptotic form of $\overline{\theta}_n$. Similar methods show easily that under suitable regularity conditions in the censored sample case it remains true that

$$(2.3) \qquad n^{\frac{1}{2}}(\overline{\theta}_n-\theta_0) = (B^TB)^{-1}B^TV_n + o_p(1).$$

From (2.3) and (2.1) it follows that

$$V_n(\bar{\theta}_n) = (I_M - B(B^TB)^{-1}B^T)V_n + o_p(1).$$

Since

$$\mathcal{L}\{V_n\} \to N_M(0, I - qq^T)$$

and qq^T and $B(B^TB)^{-1}B^T$ are projections orthogonal to each other,

(2.4) $$\mathcal{L}\{V_n(\bar{\theta}_n)\} \to N_M(0, I - qq^T - B(B^TB)^{-1}B^T)$$

which is the same result obtained in the full sample fixed cell case. Here, as often, it was convenient to use (2.1) directly rather than apply the general formulation in Theorem 2.1.

A natural class of general estimators of θ from the observations (1.1) are θ_n which are asymptotically equivalent to linear combinations of functions of order statistics. For such θ_n, asymptotic normality of $V_n(\theta_n)$ will follow from Lemma 2.1 via any theorem on asymptotic normality of linear combinations of functions of order statistics which allows special weight to be given to a finite number of sample quantiles. Such theorems appear in, e.g. Chernoff, Gastwirth and Johns (1967) and Shorack (1972). A result more useful for our purposes is obtained by appealing to the proof of Theorem 1 of Shorack (1972) rather than to the statement of that theorem. This we now do.

Shorack shows the existence of a particular probability space $(\Omega, \mathcal{A}, P)$ with "very special random quantities" defined on it. These are independent Uniform (0,1) rv's having order statistics $0 < t_{1n} < \ldots < t_{nn} < 1$ and a Brownian bridge process U such that every sample path of the empiric df process of the t_{in} converges uniformly to the corresponding sample path of U. He (and we) operate in $(\Omega, \mathcal{A}, P)$ to draw conclusions about convergence in probability, from which there follow conclusions about convergence in distribution for functions not necessarily defined in this space.

Define then

$$T_n = n^{-1} \sum_{r=1}^{n} c_{rn} Q(t_{rn})$$

(2.5)

$$\mu_n = \int_0^1 Q(t) J_n(t)dt$$

where $J_n(t)$ is the function equal to c_{in} for $(i-1)/n < t \le i/n$ and $1 \le i \le n$, with $J_n(0) = c_{1n}$. For fixed b_1, b_2, M and $\gamma > 0$ define

$$D_1(t) = Mt^{-b_1}(1-t)^{-b_2} \qquad 0 < t < 1$$

$$D_2(t) = Mt^{-\frac{1}{2}+b_1+\gamma}(1-t)^{-\frac{1}{2}+b_2+\gamma} \qquad 0 < t < 1.$$

The version of Shorack's assumptions which we require is as follows.

(S-1) Q is left continuous on $(0,1)$, of bounded variation on $(\varepsilon, 1-\varepsilon)$ for all $\varepsilon > 0$, and for some b_1, b_2, M and $\gamma > 0$, $|Q(t)| \le D_2(t)$ on $(0,1)$.

(S-2) Let $|Q|$ denote the total variation measure associated with the signed measure induced by Q. There is a function J such that except on a set of t's of $|Q|$-measure 0 both J is continuous at t and $J_n \to J$ uniformly in some small neighborhood of t as $n \to \infty$. Moreover, $|J_n(t)| \le D_1(t)$ and $|J(t)| \le D_1(t)$ on $(0,1)$.

(S-3) Let r_n and r be constants such that $r_n - r = o(n^{-\frac{1}{2}})$. Let $0 < \Delta < 1$, and R be a function for which $R'(\Delta)$ exists.

The following result is contained in the proof of Theorem 1 of [19].

LEMMA 2.2. (Shorack [19]) _If_ (S-1) _and_ (S-2) _hold_, _then_

(2.6) $$n^{\frac{1}{2}}(T_n - \mu_n) \to -\int_0^1 JUdQ \quad (P).$$

_If_ (S-3) _holds_, _then_

(2.7) $$n^{\frac{1}{2}}(r_n R(t_{[n\Delta],n}) - rR(\Delta)) \to -rR'(\Delta)U(\Delta) \quad (P).$$

Convergence here is convergence in probability in $(\Omega, \mathcal{Q}, P)$.

We now show that Lemma 2.2 can be applied to $V_n(\theta_n)$ when θ_n has the following asymptotic form under $F(x,\theta_0)$

(2.8) $$n^{\frac{1}{2}}(\theta_n-\theta_0) = n^{-\frac{1}{2}}\{\sum_{r=[n\alpha]+1}^{[n\beta]} h(X_{(r)})$$

$$+ c_n(\alpha,\beta)g(X_{([n\beta])}) + k_n(\alpha,\beta)d(X_{([n\alpha]+1)})\} + o_p(1).$$

Here h,g and d are functions from R^1 to R^m; in accordance with our convention their dependence on θ_0 is suppressed. We require that $c_n(\alpha,1) = 0$ and $k_n(0,\beta) = 0$. Estimators putting special weight on sample quantiles other than the point(s) of censoring could be accommodated, but (2.8) covers all estimators used in the examples of Section 4. In particular, regular cases of the maximum likelihood estimator (mle) from the censored data (1.1) have this form.

Define now for $0 < t < 1$ the inverse function of $F(\cdot,\theta_0)$,

$$b(t) = \min\{x: F(x,\theta_0) = t\},$$

and $H(t) = h(b(t))$, $G(t) = g(b(t))$, $L(t) = d(b(t))$. The following assumptions will be made.

(A-1) There are numbers $c_0(\alpha,\beta)$ and $k_0(\alpha,\beta)$ such that $c_n(\alpha,\beta)/n-c_0(\alpha,\beta) = o(n^{-\frac{1}{2}})$ and $k_n(\alpha,\beta)/n-k_0(\alpha,\beta) = o(n^{-\frac{1}{2}})$.

(A-2) $-\int_\alpha^\beta H(s)ds = c_0(\alpha,\beta)G(\beta) + k_0(\alpha,\beta)L(\alpha)$.

(A-3) The jth component of H, H_j, is continuous at α and β, left continuous on (0,1), and of bounded variation on $(\varepsilon,1-\varepsilon)$ for all $\varepsilon > 0$, $j=1,\ldots,m$.

(A-4) H_j' exists a.e. on (0,1) $j=1,\ldots,m$

G_j' exists at α if $\alpha > 0$, $j=1,\ldots,m$

L_j' exists at β if $\beta < 1$, $j=1,\ldots,m$

(A-5) For some $M > 0$, $\gamma > 0$, b_1 and b_2 (where $b_1 \geq 0$ if $\alpha=0$ and $b_2 \geq 0$ if $\beta=1$), $|H_j(t)| \leq D_2(t)$ on $(0,1)$.

Assumptions 1, 3, 4 and 5 reflect the assumptions of Lemma 2.2. Assumption 2 is an asymptotic unbiasedness condition on θ_n. To express the asymptotic variance of $V_n(\theta_n)$, let $U(t)$ be a Brownian bridge process on $(0,1)$ and define the m-vector

$$S = \int_\alpha^\beta H'(t)U(t)dt + c_0(\alpha,\beta)G'(\beta)U(\beta) + k_0(\alpha,\beta)L'(\alpha)U(\alpha)$$

and the M-vector W having ith component $W_i = [U(\delta_i)-U(\delta_{i-1})]/p_i^{\frac{1}{2}}$.

THEOREM 2.2. Let θ_n satisfy (2.8) and suppose that (A-1)-(A-5) hold. Then under $F(x,\theta_0)$

$$\mathcal{L}\{V_n(\theta_n)\} \to N_M(0,\Sigma)$$

where

$$\Sigma = I-qq^T + BA^T + AB^T + BCB^T$$

$$C = E[SS^T], \qquad A = E[WS^T].$$

PROOF. Substituting (2.8) into (2.1) and using the inverse df transformation $x = b(t)$ shows that the ith component of $V_n(\theta_n)$ differs by $o_p(1)$ from a quantity having the same distribution as the sum of

$$S_n = n^{\frac{1}{2}}\{n^{-1} \sum_{r=[n\alpha]+1}^{[n\beta]} (-B_iH(t_{rn}))-c_0B_iG(\beta)-k_0(\alpha,\beta)B_iL(\alpha)\}$$

$$Z_{1n} = -n^{\frac{1}{2}}p_i^{-\frac{1}{2}}f(x_i)(b(\zeta_{in})-b(\delta_i))$$

$$Z_{2n} = n^{\frac{1}{2}}p_i^{-\frac{1}{2}}f(x_{i-1})(b(\zeta_{i-1,n})-b(\delta_{i-1}))$$

$$Z_{3n} = -n^{\frac{1}{2}}(\frac{c_n(\alpha,\beta)}{n} B_iG(\zeta_{\beta n}) - c_0(\alpha,\beta)B_iG(\beta))$$

$$Z_{4n} = -n^{\frac{1}{2}}(\frac{k_n(\alpha,\beta)}{n} B_iL(\zeta_{\alpha n}) - k_0(\alpha,\beta)B_iL(\alpha)).$$

Here B_i is the ith row of B, t_{rn} is the rth order statistic from Shorack's special uniform rv's on $(\Omega,\mathcal{G},P)$, ζ_{in} is the sample δ_i-quantile of these rv's, and $\zeta_{\alpha n},\zeta_{\beta n}$ are the sample α and β-quantiles.

First consider S_n. Set in (2.5) $Q=B_iH$ and $c_{rn} = -1$ for $[n\alpha]+1 \le r \le [n\beta]$ and 0 otherwise. Then

$$\mu_n = -\int_{[n\alpha]/n}^{[n\beta]/n} B_iH(t)dt = -\int_\alpha^\beta B_iH(t)dt + o(n^{-\frac{1}{2}})$$

$$= c_0(\alpha,\beta)B_iG(\beta) + k_0(\alpha,\beta)B_iL(\alpha) + o(n^{-\frac{1}{2}})$$

by (A-3) and (A-2). Thus in the notation of (2.5), $S_n = n^{\frac{1}{2}}(T_n-\mu_n) + o(1)$. The assumptions (S-1) and (S-2) of Lemma 2.2 are satisfied. For (A-3) and (A-5) imply (S-1), and (S-2) is clearly true for the function $J(t) = -1$ for $\alpha < t \le \beta$ and 0 elsewhere. (Note that this J is bounded by D_1 for any b_1, b_2 satisfying the restrictions stated in A-5.) So by (2.6),

(2.9) $$S_n \to \int_\alpha^\beta B_iH'(t)U(t)dt \quad (P).$$

Each of the terms Z_{jn} has the form $n^{\frac{1}{2}}(r_nR(t_{[n\Delta],n})-rR(\Delta))$ of (2.7), and (A-1), (A-4) and the relation $b'(t) = 1/f(b(t))$ imply that (S-3) holds in each case. Hence the sum of the Z_{jn} converges in probability on $(\Omega,\mathcal{G},P)$ to

$$p_i^{-\frac{1}{2}}f(x_i)b'(\delta_i)U(\delta_i)-p_i^{-\frac{1}{2}}f(x_{i-1})b'(\delta_{i-1})U(\delta_{i-1})$$

$$+ c_0(\alpha,\beta)B_iG'(\beta)U(\beta) + k_0(\alpha,\beta)B_iL'(\alpha)U(\alpha).$$

Simplifying by using the fact that $f(x_i)b'(\delta_i) = 1$ and combining with (2.9), $S_n + \Sigma_1^4 Z_{jn}$ converges in probability to the rv $Y_i = B_iS + W_i$. Hence the version of $V_n(\theta_n)$ defined on $(\Omega,\mathcal{G},P)$ converges in probability to $Y = (Y_1,\ldots,Y_M)^T$. Now Y has the $N_M(0,\Sigma)$ distribution, where $\Sigma = E[YY^T]$. Computation using $E[W_iW_j] = (p_ip_j)^{\frac{1}{2}}$ for $i\neq j$ and $E[W_i^2] = 1-p_i$ reduces Σ to the form stated in the theorem. This is therefore the limiting law of any version of $V_n(\theta_n)$.

That the censored data mle has the form (2.8) is shown by Halperin (1952). He studies the right-censored ($\alpha=0$) case, but his work is easily extended to two-sided censoring. Let x_α and x_β be the population α- and β-quantiles from $F(x,\theta_0)$. The Fisher information matrix for the data (1.1) is

$$K = K(\theta_0,x_\alpha,x_\beta) = \alpha^{-1}\left(\int_{-\infty}^{x_\alpha} \frac{\partial f}{\partial\theta}\,dx\right)\left(\int_{-\infty}^{x_\alpha} \frac{\partial f}{\partial\theta}\,dx\right)^T + \int_{x_\alpha}^{x_\beta} \left(\frac{\partial \log f}{\partial\theta}\right)\left(\frac{\partial \log f}{\partial\theta}\right)^T f\,dx$$
$$+ (1-\beta)^{-1}\left(\int_{-\infty}^{x_\beta} \frac{\partial f}{\partial\theta}\,dx\right)\left(\int_{-\infty}^{x_\beta} \frac{\partial f}{\partial\theta}\,dx\right)^T.$$

(Here $\partial f/\partial\theta$ denotes the m-vector of derivatives $\partial f/\partial\theta_j$.) Then in suitably regular cases, the mle $\hat{\theta}_n$ satisfies

$$\text{(2.10)}\qquad n^{\frac{1}{2}}(\hat{\theta}_n-\theta_0) = n^{-\frac{1}{2}}\Big\{\sum_{r=[n\alpha]+1}^{[\beta n]} K^{-1}\frac{\partial \log f}{\partial\theta}(X_{(r)})$$
$$+ (n-[n\beta])K^{-1}\frac{\partial \log(1-F)}{\partial\theta}(X_{([n\beta])}) + [n\alpha]K^{-1}\frac{\partial \log F}{\partial\theta}(X_{([n\alpha]+1)})\Big\}+o_p(1)$$

under $F(x,\theta_0)$. It is easy to check that (A-1) and (A-2) are satisfied with $c_0(\alpha,\beta) = 1-\beta$ and $k_0(\alpha,\beta) = \alpha$. When (A-3)-(A-5) hold, Theorem 2.2 therefore applies to the mle $\hat{\theta}_n$. A lengthy calculation (details appear in Mihalko (1977)) shows that in this case

$$\text{(2.11)}\qquad \mathcal{L}\{V_n(\hat{\theta}_n)\} \to N_M(0,I-qq^T-BK^{-1}B^T),$$

a result identical to that for the full sample case except that K has replaced the full sample information matrix.

3. Chi-square statistics. A statistic of chi-square type is a nonnegative definite quadratic form in $V_n(\theta_n)$, $T_n = V_n(\theta_n)^T Q_n V_n(\theta_n)$, where the possibly random MxM matrices Q_n converge in probability to a nonnegative definite matrix $Q = Q(\theta_0)$. Whenever $\mathcal{L}\{V_n(\theta_n)\} \to N_M(0,\Sigma)$ under $F(x,\theta_0)$, then T_n has as its limiting null distribution that of $\sum_{j=1}^{M} \lambda_j \chi_j^2(1)$ where $\chi_j^2(1)$ are independent chi-square rv's with 1 degree of freedom and λ_j are the characteristic roots of $S'\Sigma S$ for $Q = SS'$. We are primarily concerned with three practically useful statistics within this general class.

Case 1. The Pearson statistic. The censored-sample analog of the classic Pearson-Fisher statistic is the sum of squares $T_{1n} = V_n(\bar{\theta}_n)^T\ V_n(\bar{\theta}_n)$ using the grouped data mle. Since the covariance matrix of the limiting law in (2.4) is a projection of rank M-m-1 it follows that T_{1n} has $\chi^2(M-m-1)$ as its limiting null distribution whenever $\bar{\theta}_n$ satisfies (2.3). Chernoff and Lehmann (1954) studied the full sample case of the Pearson statistic with raw-data mle's, $T_{2n} = V_n(\hat{\theta}_n)^T\ V_n(\hat{\theta}_n)$. Once again the full-sample results carry over to censored samples. These results are: (1) the matrix $K-B^TB$ is nonnegative definite in regular cases; (2) when $K-B^TB$ is positive definite and B has rank m, the limiting law of T_{2n} under $F(x,\theta_0)$ is that of

$$\chi^2(M-m-1) + \sum_{j=1}^{m} \lambda_j(\theta_0)\chi_j^2(1)$$

where $0 < \lambda_j < 1$ and the χ^2's are independent chi-square rv's with the indicated degrees of freedom. Therefore large sample critical points for T_{2n} fall between those of $\chi^2(M-m-1)$ and $\chi^2(M-1)$, bounds which make T_{2n} often useful despite the dependence of its distribution on θ_0. Proofs of (1) and (2) proceed much as in the full sample case; details appear in Mihalko (1977). Finally, we remark that if $F(x,\theta)$ is a location-scale family, then the λ_j do not depend on θ_0. This is an analog of the behavior of random-cell chi-square tests noted by Watson and Roy in the full sample case.

Case 2. The Rao-Robson statistic. Rao and Robson (1974) discovered the quadratic form in $V_n(\hat{\theta}_n)$ which has the $\chi^2(M-1)$ limiting null distribution in the full-sample case. They showed by simulation that this statistic is generally more powerful than T_{1n} or T_{2n}. Their proofs apply only in restricted situations, but a simple and general proof is given in Moore (1977). Once again the essential features of this development are unchanged in censored samples. These are (1) when $K-B^TB$ is positive definite, then the covariance matrix $\Sigma = I_M - qq^T - BK^{-1}B^T$ of (2.11) has rank M-1. (2) A generalized inverse of Σ is $\Sigma^- = I_M + B(K-B^TB)^{-1}B^T$, so that

$$\mathcal{L}\{V_n(\hat{\theta}_n)^T\Sigma^- V_n(\hat{\theta}_n)\} \to \chi^2(M-1),$$

and this is the greatest obtainable number of degrees of freedom. (3) If $B_n = B(\xi_n,\hat{\theta}_n)$ and $K_n = K(\hat{\theta}_n, X_{([n\alpha]+1)}X_{([n\beta])})$, then $I_M + B_n(K_n - B_n^TB_n)^{-1}B_n^T$

is a consistent estimator of Σ^- so that

$$T_{3n} = V_n(\hat{\theta}_n)^T\, V_n(\hat{\theta}_n) + V_n(\hat{\theta}_n)^T\, B_n(K_n-B_n^TB_n)^{-1}\, B_n^T\, V_n(\hat{\theta}_n)$$

also has the $\chi^2(M-1)$ limiting null distribution whenever $\hat{\theta}_n$ satisfies (2.11) and $K-B^TB$ is positive definite.

Note that the first term of T_{3n} is T_{2n}, the Chernoff-Lehmann statistic. The bounds on the critical points of T_{2n} given above make it unnecessary in many instances to compute the second term of T_{3n}.

Case 3. The Dzhaparidze-Nikulin statistic. Suppose that θ_n is any estimator satisfying $n^{\frac{1}{2}}(\theta_n-\theta_0) = O_p(1)$ under $F(x,\theta_0)$. From (2.1) we see that whenever B has rank m,

$$(I_M-B(B^TB)^{-1}B^T)V_n(\theta_n) = (I_M-B(B^TB)^{-1}B^T)V_n + o_p(1),$$

and this is exactly the asymptotic form of $V_n(\overline{\theta}_n)$ from which (2.4) followed.

Since $I_M - B(B^T B)^{-1} B^T$ is symmetric and idempotent,

$$V_n(\theta_n)^T (I_M - B(B^T B)^{-1} B^T)^T (I_M - B(B^T B)^{-1} B^T) V_n(\theta_n)$$
$$= V_n(\theta_n)^T (I_M - B(B^T B)^{-1} B^T) V_n(\theta_n)$$

and just as in Case 1 this statistic has the $\chi^2(M-m-1)$ limiting null distribution. Since Lemma 2.1 shows that $V_n(\theta_n)$ is $O_p(1)$, the statistic

$$T_{4n} = V_n(\theta_n)^T (I_M - B_n(B_n^T B_n)^{-1} B_n^T) V_n(\theta_n)$$
$$= V_n(\theta_n)^T\ V_n(\theta_n) - V_n(\theta_n)^T\ B_n(B_n^T B_n)^{-1} B_n^T V_n(\theta_n)$$

has the $\chi^2(M-m-1)$ limiting null distribution whenever B has rank m and $\theta_n - \theta_0 = O_p(n^{-\frac{1}{2}})$. The first term of T_{4n} is once again the Pearson statistic, which here is "chopped down" to $\chi^2(M-m-1)$ rather than being "built up" to $\chi^2(M-1)$ as in the Rao-Robson case. This statistic was proposed in the full sample case by Dzhaparidze and Nikulin (1974). They gave an unwieldy (and perhaps defective) proof. Use of Lemma 2.1 (or the analogous Theorem 4.1 of [14] for full samples) shows how this universal chi-square statistic is obtained by projecting orthogonal to B.

REMARK 1. The second terms of both T_{3n} and T_{4n} are quadratic forms in the m-vector $B_n^T V_n(\theta_n)$, which has jth component

(3.1) $$\sum_{i=1}^{M} \frac{N_{in} - np_{in}}{(np_{in})^{\frac{1}{2}}}\, p_{in}^{-\frac{1}{2}} \frac{\partial p_{in}}{\partial \theta_j} = n^{-\frac{1}{2}} \sum_{i=1}^{M} \frac{N_{in}}{p_{in}} \frac{\partial p_{in}}{\partial \theta_j}$$

whenever $\partial f(x,\theta)/\partial \theta_j$ is continuous so that $\sum_1^M p_i(\theta) = 1$ and differentiation of $p_i(\theta)$ under the integral imply $\sum_1^M \partial p_i(\theta)/\partial \theta_j = 0$. This holds in the examples of Section 4, where (3.1) simplifies the form of T_{3n} and T_{4n}.

REMARK 2. Suppose that $F(x,\theta) = F((x-\theta_1)/\theta_2)$ is a location-scale family. Then the matrices

$$Q_3 = I_M + B(K - B^T B)^{-1} B^T$$
$$Q_4 = I_M - B(B^T B)^{-1} B^T$$

estimated to form T_{3n} and T_{4n}, which are evaluated at θ_0 and the population quantiles x_i of $F(x,\theta_0)$, <u>do not depend on</u> θ_0 and can be evaluated at $\theta_1 = 0$, $\theta_2 = 1$. For the remainder of this section we drop the convention that θ_0 is assumed, so that $F(x)$ and $f(x)$ are the $\theta = (0,1)^T$ distribution and density functions in the location-scale case. Let z_i be the population δ_i-quantile of F (so $x_i = \theta_{01} + z_i\theta_{02}$), $p_i = \delta_i - \delta_{i-1}$, $\varphi_i = f(z_i) - f(z_{i-1})$ and $\nu_i = z_i f(z_i) - z_{i-1} f(z_{i-1})$. Then $\partial p_i/\partial\theta_1 = -\varphi_i$, $\partial p_i/\partial\theta_2 = -\nu_i$ and the ith row of $B(\xi_0,\theta_0)$ is $-(\varphi_i,\nu_i)/\theta_{02}\, p_i^{\frac{1}{2}}$. So

$$B^TB = \theta_{02}^{-2} \begin{pmatrix} \sum_1^M \varphi_i^2/p_i & \sum_1^M \varphi_i\nu_i/p_i \\ \sum_1^M \varphi_i\nu_i/p_i & \sum_1^M \nu_i^2/p_i \end{pmatrix}. \tag{3.2}$$

Similarly, letting z_α and z_β be the population α- and β-quantiles of F, $K(\theta_0,x_\alpha,x_\beta) = \theta_{02}^{-2} J$, where J has entries (see (3.17) of [2])

$$J_{11} = \int_{z_\alpha}^{z_\beta} [f'(y)]^2/f(y)dy + f^2(z_\alpha)/\alpha + f^2(z_\beta)/(1-\beta)$$

$$J_{12} = \int_{z_\alpha}^{z_\beta} f'(y)[1+yf'(y)/f(y)]dy + z_\alpha f^2(z_\alpha)/\alpha + z_\beta f^2(z_\beta)/(1-\beta) \tag{3.3}$$

$$J_{22} = \int_{z_\alpha}^{z_\beta} [1+yf'(y)/f(y)]^2 f(y)dy + z_\alpha^2 f^2(z_\alpha)/\alpha + z_\beta^2 f^2(z_\beta)/(1-\beta)$$

Relations (3.2) and (3.3) with the expression for $B(\xi_0,\theta_0)$ show that Q_3 and Q_4 are θ_0-free. They can also be used to compute both K, B^TB and $K_n, B_n^TB_n$, replacing δ_i and z_i by their estimates in the latter case.

In location-scale cases, we have therefore alternative statistics

$$T_{3n}^{*} = V_n(\hat{\theta}_n)Q_3V_n(\hat{\theta}_n)$$

$$T_{4n}^{*} = V_n(\theta_n)Q_4V_n(\theta_n)$$

which are asymptotically equivalent under H_0 to T_{3n} and T_{4n}, respectively. Note that the simplification provided by (3.1) does not apply to T_{3n}^{*} and T_{4n}^{*}, for the term $n^{\frac{1}{2}} \Sigma(p_{in}/p_i)\partial p_i/\partial\theta_j$, which vanished when p_{in} replaced p_i, is not $o_p(1)$.

4. Examples. The statistics described in Section 3 will now be applied to derive usable tests of fit for censored data to each of four parametric families of distributions. In each case, the regularity conditions required for application of our theory are met. For example, the negative exponential family satisfies Halperin's conditions [6] for the mle to have the asymptotic form (2.10), and the estimator in this form satisfies (A-3) through (A-5). This justifies the use of the Rao-Robson statistic (Case 2 in Section 3) in Example 1 below. Regularity conditions are not checked in detail here; this work can be found in Mihalko (1977).

Example 1. The negative exponential family. It is desired to test the fit of right-censored data $0 < X_{(1)} <\ldots< X_{([n\beta])}$ to the scale-parameter family $F(x,\theta) = 1-e^{-x/\theta}$ $(x>0)$, $\Omega = \{\theta: \ 0 < \theta < \infty\}$. Epstein and Sobel (1953) show that the mle is

$$\hat{\theta}_n = [n\beta]^{-1}\Big(\sum_{r=1}^{[n\beta]} X_{(r)} + (n-[n\beta])X_{([n\beta])}\Big).$$

Substituting $f(y) = e^{-y}$, $z_\alpha = 0$ and $z_\beta = -\log(1-\beta)$ into (3.3) gives $J_{22} = \beta$ and $K = \theta_0^{-2}\beta$. From (3.2) we see that $B^TB = \theta_0^{-2}\Sigma_1^M \nu_i^2/p_i$ where

$$\nu_i = -(1-\delta_i)\log(1-\delta_i) + (1-\delta_{i-1})\log(1-\delta_{i-1})$$

and $p_i = \delta_i - \delta_{i-1}$. Hence setting $\Delta = \beta - \Sigma_1^M \nu_i^2/p_i$,

$$Q_3 = I_M + \Delta^{-1}\left(\frac{\nu_i\nu_j}{(p_ip_j)^{\frac{1}{2}}}\right)_{M\times M}.$$

When sample δ_i-quantiles $0 < \xi_{1n} < \ldots < \xi_{M-1,n} = X_{([n\beta])}$ are the cell boundaries,

$$p_{in} = e^{-\xi_{i-1,n}/\hat{\theta}_n} - e^{-\xi_{in}/\hat{\theta}_n}$$

$$K_n = \hat{\theta}_n^{-2}(1-e^{-X_{([n\beta])}/\hat{\theta}_n})$$

and setting

$$\nu_{in} = \hat{\theta}_n^{-1}(\xi_{in}\, e^{-\xi_{in}/\hat{\theta}_n} - \xi_{i-1,n}\, e^{-\xi_{i-1,n}/\hat{\theta}_n})$$

gives $B_n^T B_n = \hat{\theta}_n^{-2}\Sigma_1^M \nu_{in}^2/p_{in}$. Thus the two versions of the Rao-Robson statistic are

$$T_{3n}^* = \sum_{i=1}^{M} \frac{[N_{in}-np_{in}]^2}{np_{in}} + \Delta^{-1} \sum_{i,j=1}^{M} \left(\frac{N_{in}-np_{in}}{(np_{in})^{\frac{1}{2}}}\right)\left(\frac{N_{jn}-np_{jn}}{(np_{jn})^{\frac{1}{2}}}\right) \frac{\nu_i\nu_j}{(p_ip_j)^{\frac{1}{2}}}$$

and (using (3.1))

$$T_{3n} = \sum_{i=1}^{M} \frac{[N_{in}-np_{in}]^2}{np_{in}} + (n\Delta_n)^{-1}\left(\sum_{i=1}^{M} N_{in}\nu_{in}/p_{in}\right)^2$$

$$\Delta_n = 1-e^{-X_{([n\beta])}/\hat{\theta}_n} - \sum_{i=1}^{M} \nu_{in}^2/p_{in}.$$

A slight simplification of T_{3n} is obtained by replacing Δ_n by its limit in probability Δ. Both T_{3n} and T_{3n}^* have the $\chi^2(M-1)$ limiting null distribution.

<u>Example</u> 2. <u>The Normal family</u>. We will test the fit of right-censored data to the location-scale family of normal distributions $N(\mu,\sigma^2)$. The mle's satisfy (2.10), but the likelihood equations cannot be solved in closed form.

Chernoff, Gastwirth and Johns (1967) give linear combinations of order statistics for estimating location and scale parameters from censored data which are asymptotically equivalent to the mle's. For the normal case, these estimators are $(\hat{\mu}_n, \hat{\sigma}_n) = (E_{1n}, E_{2n})J^{-1}$ where

$$E_{1n} = n^{-1} \sum_{r=1}^{[n\beta]} X_{(r)} + (-z_\beta \varphi(z_\beta) + (1-\beta)^{-1}\varphi^2(z_\beta))X_{([n\beta])}$$

$$E_{2n} = 2n^{-1} \sum_{r=1}^{[n\beta]} \Phi^{-1}(\tfrac{r}{n+1})X_{(r)}$$

$$+ (1-z_\beta^2 + (1-\beta)^{-1} z_\beta \varphi(z_\beta))X_{([n\beta])}$$

and J is the matrix of (3.3) having entries

$$J_{11} = (\beta - z_\beta \varphi(z_\beta) + (1-\beta)^{-1}\varphi^2(z_\beta))$$

$$J_{12} = (-(1+z_\beta)^2 \varphi(z_\beta) + (1-\beta)^{-1} z_\beta \varphi^2(z_\beta))$$

$$J_{22} = (2\beta - z_\beta(1+z_\beta^2)\varphi(z_\beta) + (1-\beta)^{-1} z_\beta^2 \varphi^2(z_\beta)).$$

Here φ and Φ are the standard normal density and distribution functions, and $\Phi(z_\beta) = \beta$.

We will give only the original version T_{3n} of the Rao-Robson statistic. Let $z_{in} = (\xi_{in} - \hat{\mu}_n)/\hat{\sigma}_n$, so that $p_{in} = \Phi(z_{in}) - \Phi(z_{i-1,n})$. If we set $\varphi_{in} = \varphi(z_{in}) - \varphi(z_{i-1,n})$ and $\nu_{in} = z_{in}\varphi(z_{in}) - z_{i-1,n}\varphi(z_{i-1,n})$, then $B_n^T B_n$ is given by (3.2) if $\sigma_0, \varphi_i, \nu_i, p_i$ are replaced by $\hat{\sigma}_n, \varphi_{in}, \nu_{in}, p_{in}$. Moreover, $K_n = \hat{\sigma}_n^{-2} J_n$ where J_n is J with z_β and β replaced by their estimates $z_{M-1,n}$ and $\Phi(z_{M-1,n})$. Thus $K_n - B_n^T B_n = \hat{\sigma}_n^{-2} D_n$, where the entries of D_n are

$$D_{11} = \Phi(z_{M-1,n}) - \nu_{Mn} - \sum_{i=1}^{M-1} \varphi_{in}^2 / p_{in}$$

$$D_{12} = -(1+z_{M-1,n})^2 \varphi_{Mn} - \sum_{i=1}^{M-1} \varphi_{in}\nu_{in}/p_{in}$$

$$D_{22} = 2\Phi(z_{M-1,n}) - (1+z_{M-1,n}^2)\nu_{Mn} - \sum_{i=1}^{M-1} \nu_{in}^2/p_{in}.$$

Inverting D_n and using (3.1) gives

$$T_{3n} = \sum_{i=1}^{M} \frac{[N_{in}-np_{in}]^2}{np_{in}} + (n\Delta_n)^{-1}[D_{22}(\sum_{i=1}^{M} N_{in}\varphi_{in}/p_{in})^2$$

$$- 2D_{12}(\sum_{i=1}^{M} N_{in}\varphi_{in}/p_{in})(\sum_{i=1}^{M} N_{in}\nu_{in}/p_{in})$$

$$+ D_{11}(\sum_{i=1}^{M} N_{in}\nu_{in}/p_{in})^2]$$

where $\Delta_n = D_{11}D_{22}-D_{12}^2$ is the determinant of D_n. Once again Δ_n can be replaced by its limit in probability without affecting the $\chi^2(M-1)$ limiting null distribution of T_{3n}. This statistic appears complex, but note that once the z_{in} have been obtained, the successive-difference form of p_{in}, φ_{in} and ν_{in} makes T_{3n} quite easily computable on a programmable calculator.

When the data are symmetrically doubly censored ($\alpha=1-\beta=p$), the estimators $(\hat{\mu}_n, \hat{\sigma}_n)$ and the statistic T_{3n} simplify considerably because K and B^TB are diagonal matrices. Since this case is less often met than is right censoring, we do not give the specific results here. They are easily derived from Section 3 of [2] and the general recipe for T_{3n}.

<u>Example</u> 3. <u>The two-parameter uniform family</u>. Doubly censored data (1.1) with $0 < \alpha < \beta < 1$ will be tested for fit to the family with density function

$$f(x,\theta) = \theta_2^{-1} \qquad \theta_1-\tfrac{1}{2}\theta_2 \leq x \leq \theta_1 + \tfrac{1}{2}\theta_2$$

$$\Omega = \{\theta = (\theta_1,\theta_2)^T: \ -\infty < \theta_1 < \infty, \ 0 < \theta_2 < \infty\}.$$

Sarhan (1955) derives the asymptotically best linear unbiased estimator of θ, $\theta_n = (\theta_{1n},\theta_{2n})^T$ where

$$\theta_{1n} = (b_{n(1-\beta)}X_{([n\alpha]+1)} + b_{n\alpha}X_{([n\beta])})/2\Gamma_{n\alpha\beta}$$

$$\theta_{2n} = (n+1)(X_{([n\beta])}-X_{([n\alpha]+1)})/\Gamma_{n\alpha\beta}$$

$$b_{np} = n-2[np]-1, \ \Gamma_{n\alpha\beta} = [n\beta]-[n\alpha]-1.$$

This estimator has the form (2.2). It is easy to check that the coefficients of the sample quantiles approach their limits

$$a_{11} = \frac{-(1-2\beta)}{2(\beta-\alpha)}\ ,\ a_{12} = \frac{1-2\alpha}{2(\beta-\alpha)}\ ,\ a_{21} = -a_{22} = (\beta-\alpha)^{-1}$$

at the $o(n^{-\frac{1}{2}})$ rate required by Theorem 2.1. Since the population δ-quantile under θ_0 is $x_\delta = \theta_{10} + (\delta-\frac{1}{2})\theta_{20}$, it is also easy to check that θ_0 has the required expression in terms of the a_{ij}. That $V_n(\theta_n)$ has a $N(0,\Sigma)$ limiting null distribution follows from Theorem 2.1. The covariance matrix Σ can be computed either from Theorem 2.1 or directly from Lemma 2.1 and the limiting law of $X_{([n\alpha]+1)}, X_{([n\beta])}$ and the sample quantiles ξ_{in} chosen as cell boundaries. These computations show that the MxM matrix Σ has 0's in the 1st and Mth row and column, with the central (M-2)x(M-2) matrix having entries $\delta_{ij}-(p_ip_j)^{\frac{1}{2}}/(\beta-\alpha)$ for $i,j = 2,\ldots,M-1$. (Here $\delta_{ij} = 1$ if $i=j$ and 0 otherwise.) Since Σ has rank M-3, no quadratic form in $V_n(\theta_n)$ can have a limiting chi-square distribution with more than M-3 degrees of freedom. This upper limit is attained by the Dzhaparidze-Nikulin statistic, which can therefore be employed without loss of degrees of freedom.

From the results in Remark 2 of Section 3 it follows that

$$B = \theta_{02}^{-1}\begin{pmatrix} -p_1^{-\frac{1}{2}} & p_1^{-\frac{1}{2}}(\frac{1}{2}-\alpha) \\ 0 & -p_2^{-\frac{1}{2}} \\ \vdots & \vdots \\ 0 & -p_{M-1}^{-\frac{1}{2}} \\ p_M^{-\frac{1}{2}} & p_M^{-\frac{1}{2}}(\beta-\frac{1}{2}) \end{pmatrix}$$

and that B^TB has entries $[\theta_{02}^2\ \alpha(1-\beta)]^{-1}B_{ij}$, where

$$B_{11} = 1-\beta+\alpha \qquad B_{12} = \tfrac{1}{2}(\alpha+\beta-1)$$

$$B_{22} = \alpha(\tfrac{1}{2}-\beta)^2 + (1-\beta)(\tfrac{1}{2}-\alpha)^2 + \alpha(1-\beta)(\beta-\alpha).$$

Since Q_4 is algebraically complicated (though easy to compute numerically), we give only T_{4n}. The matrix B_n is obtained from B by substituting θ_n and $p_{in} = F(\xi_{in},\theta_n) - F(\xi_{i-1,n},\theta_n)$ for θ_0 and p_i. For $i=2,\ldots,M-1$, $p_{in} = (\xi_{in}-\xi_{i-1,n})/\theta_{2n}$. For $p_1 = \alpha$ and $p_M = 1-\beta$, this process yields $\alpha_n = \theta_{2n}^{-1}(X_{([n\alpha]+1)}-\theta_{1n}) + \frac{1}{2}$ and $\beta_n = \theta_{2n}^{-1}(X_{([n\beta])}-\theta_{1n}) + \frac{1}{2}$. From (3.1) we then obtain

$$V_n^T(\theta_n)B_n = n^{-\frac{1}{2}}\,\theta_{2n}^{-1}(A_{1n},A_{2n})$$

$$A_{1n} = (n-[n\beta])/(1-\beta_n) - [n\alpha]/\alpha_n$$

$$A_{2n} = -A_{1n}/2 - [n\beta] + \beta_n(n-[n\beta])/(1-\beta_n)$$

and finally, inverting $B_n^T B_n$,

$$T_{4n} = \sum_{i=1}^{M} \frac{[N_i - np_{in}]^2}{np_{in}} - \frac{\alpha_n(1-\beta_n)}{n\Delta_n}(A_{1n}^2 B_{22n} - 2A_{1n}A_{2n}B_{12n} + A_{2n}^2 B_{11n})$$

where B_{ijn} results from substituting α_n,β_n in B_{ij} and Δ_n is the determinant of the matrix (B_{ijn}).

The uniform family is less often encountered in censored-data situations than our other examples, but it offers an interesting contrast between full and censored samples. In a full sample, the BLUE's of θ are based on the extreme order statistics and approach θ_0 at a rate faster than $n^{\frac{1}{2}}$. Thus $V_n(\theta_n)$ is asymptotically equivalent to V_n and the Pearson statistic $V_n(\theta_n)^T V_n(\theta_n)$ has the $\chi^2(M-1)$ limiting null distribution. Censoring deletes the most informative part of the sample, leaving M-3 as the greatest obtainable number of degrees of freedom. In Examples 1 and 2, tests from censored samples attained the same number of degrees of freedom possible for full samples.

<u>Example</u> 4. <u>The Weibull family</u>. Once again the mle's for this common family are obtainable only by numerical approximation. Bain (1972) transforms the Weibull to an extreme value distribution and gives a simple but quite efficient extimator for the scale parameter of the transformed distribution.

Mann, Schafer and Singpurwalla (1976) give a corresponding estimator of the location parameter. Suppose then that (after the monotonic transformation $X = \log Y$ from the original data) we have a right-censored sample $X_{(1)} < \ldots < X_{([n\beta])}$ to be tested for fit to the extreme value family

$$F(x,\theta) = 1 - \exp\{-\exp[(x-u)/b]\} \qquad -\infty < x < \infty$$

$$\Omega = \{\theta = (u,b): \ -\infty < u < \infty, \ 0 < b < \infty\}.$$

Bain's estimator of b is

$$b_n = (nk_{\beta n})^{-1} \sum_{r=1}^{[n\beta]-1} (X_{(r)} - X_{([n\beta])})$$

where the sequence of constants $k_{\beta n}$ can be expressed in terms of order statistics $v_1 < \ldots < v_{[n\beta]}$ from the standardized distribution $F(t) = 1-e^{-e^t}$ as

$$k_{\beta n} = n^{-1} \sum_{r=1}^{[n\beta]-1} E(v_r - v_{[n\beta]}).$$

Bain gives a table of $k_{\beta n}$ for various β and n. The estimator b_n is unbiased and for the choices of β and n studied by Bain has asymptotic efficiency between 0.89 and 1 relative to the much more complicated BLUE. Mann, Schafer and Singpurwalla note that $u = E(X_{([n\beta])}) - bE(v_{[n\beta]})$, and therefore propose the estimator

$$u_n = X_{([n\beta])} - b_n E(v_{[n\beta]}).$$

We shall take $\theta_n = (u_n, b_n)$.

The asymptotic behavior of θ_n must be investigated. If $T_n = k_{\beta n} b_n$ and

(4.1) $$\mu_\beta = \int_0^\beta \log\{-\log(1-t)\}dt - \beta \log\{-\log(1-\beta)\},$$

then Bain shows that $\mathcal{L}\{n^{\frac{1}{2}}(T_n - b_0\mu_\beta)\} \to N(0, b_0^2\sigma_\beta^2)$ under $F(x,\theta_0)$, where the form of σ_β^2 does not concern us. We will show below that $k_{\beta n} - \mu_\beta = o(n^{-\frac{1}{2}})$, from which it follows that

$$n^{\frac{1}{2}}(b_n - b_0) = n^{\frac{1}{2}}(T_n - b_0\mu_\beta)/\mu_\beta + o_p(1)$$

and

$$n^{\frac{1}{2}}(u_n-u_0) = n^{\frac{1}{2}}(X_{([n\beta])}-E(X_{([n\beta])}) + n^{\frac{1}{2}}(b_n-b_0)E(v_{[n\beta]})$$

$$= n^{\frac{1}{2}}(X_{([n\beta])}-x_\beta) + n^{\frac{1}{2}}(b_n-b_0)v_\beta + o_p(1)$$

where $x_\beta(v_\beta)$ is the population β-quantile of $F(x,\theta_0)(F(x))$. Thus $n^{\frac{1}{2}}(\theta_n-\theta_0)$ is asymptotically normal and Theorem 2.2 applies. The "natural" chi-square statistic is then $V_n(\theta_n)^T\ \Sigma^-\ V_n(\theta_n)$ where Σ^- is a generalized inverse of the Σ of Theorem 2.2 or a consistent estimator of such. This statistic has the $\chi^2(k)$ limiting null distribution, where k is the rank of Σ. When $k=M-1,\Sigma^-$ is relatively easy to compute (Moore (1977), Section 4). But in this case, $k < M-1$ and we are unable to obtain Σ^-. We therefore accept the possible loss of one degree of freedom and use the Dzhaparidze-Nikulin statistic. The required fact that $n^{\frac{1}{2}}(\theta_n-\theta_0) = 0_p(1)$ follows as above from $k_{\beta n}-\mu_\beta = o(n^{-\frac{1}{2}})$. This we now establish.

LEMMA 4.1. <u>For</u> <u>any</u> β, $0 < \beta < 1$, $k_{\beta n}-\mu_\beta = o(n^{-\frac{1}{2}})$ <u>as</u> $n\to\infty$.

PROOF. Corresponding to the two terms of (4.1), write

$$k_{\beta n} = n^{-1}\sum_{r=1}^{[n\beta]} E(v_r) - \frac{[n\beta]}{n} E(v_{[n\beta]})$$

$$= n^{-1}\, E[\sum_{r=1}^{[n\beta]} q(t_r)] - \frac{[n\beta]}{n} E[q(t_{[n\beta]})$$

where $q(t) = \log\{-\log(1-t)\}$ and $t_1 <\ldots< t_{[n\beta]}$ are order statistics from n iid Uniform(0,1) rv's. We first show that

$$(4.2)\qquad n^{-1}\, E[\sum_{r=1}^{[n\beta]} q(t_r)] - \int_0^\beta q(t)dt = o(n^{-\frac{1}{2}}).$$

Using the fact that the distribution of $t_1,\ldots,t_{[n\beta]}$ conditional on $t_{[n\beta]+1}$ is that of the order statistics from $[n\beta]$ iid Uniform$(0,t_{[n\beta]+1})$ rv's (this technique was suggested by Burgess Davis),

$$n^{-1}\,E[\sum_{r=1}^{[n\beta]} q(t_r)] = n^{-1}\,E[E\{\sum_{r=1}^{[n\beta]} q(t_r)\,|\,t_{[n\beta]+1}\}]$$

$$= n^{-1}\,E[[n\beta]E\{q(T)\,|\,t_{[n\beta]+1}\}]$$

where the conditional distribution of T given $t_{[n\beta]+1}$ is Uniform$(0,t_{[n\beta]+1})$.
So

$$n^{-1}\,E[\sum_{r=1}^{[n\beta]} q(t_r)] = \frac{[n\beta]}{n}\,E[(\int_0^{t_{[n\beta]+1}} q(t)dt)/t_{[n\beta]+1}]$$

$$= E[\int_0^{Y_n} q(t)dt]$$

where Y_n is a Beta$([n\beta],n-[n\beta])$ rv. Letting $\varphi(y) = \int_0^y q(t)dt$, (4.2) now states that $E[\varphi(Y_n)-\varphi(\beta)] = o(n^{-\frac{1}{2}})$. Note that $\varphi'(y) = q(y)$ and that

$$\varphi''(y) = q'(y) = [(1-y)\log(1/(1-y))]^{-1}$$

$$\leq [y(1-y)]^{-1}$$

since $\log x \geq 1-x^{-1}$ for $x \geq 1$. We have

$$E[\varphi(Y_n)-\varphi(\beta)] = E[q(\beta)(Y_n-\beta)] + \tfrac{1}{2}\,E[q'(Y_n^*)(Y_n-\beta)^2]$$

for some Y_n^* between Y_n and β. Since $E[Y_n] = [n\beta]/n$, the first term is $o(n^{-\frac{1}{2}})$. The second term is bounded by

$$\text{(4.3)} \qquad \tfrac{1}{2}\,E[\frac{(Y_n-\beta)^2}{Y_n^*(1-Y_n^*)}] \leq \tfrac{1}{2}\max\{E[\frac{(Y_n-\beta)^2}{\beta(1-\beta)}],\ E[\frac{(Y_n-\beta)^2}{Y_n(1-Y_n)}]\}.$$

The expected values on the right in (4.3) can be explicitly computed and shown to be $o(n^{-\frac{1}{2}})$. Thus (4.2) holds.

It remains only to show that in addition

$$\text{(4.4)} \qquad \frac{[n\beta]}{n}\,E[q(t_{[n\beta]})] - \beta q(\beta) = o(n^{-\frac{1}{2}}).$$

Since $t_{[n\beta]}$ is a Beta$([n\beta]+1,\ n-[n\beta])$ rv, arguments similar to those applied to φ above demonstrate (4.4).

The statistic T_{4n} will now be computed. In the notation of Remark 2 in Section 3, $f(z) = e^z e^{-e^z}$, $z_i = \log\{-\log(1-\delta_i)\}$ and $f(z_i) = -(1-\delta_i)\log(1-\delta_i)$. From this φ_i, ν_i and $B^T B$ are easily computed. Let φ_{in} and ν_{in} have the same expressions as do φ_i and ν_i, but with δ_i replaced by

$$\delta_{in} = 1 - \exp\{-\exp[(\xi_{in}-u_n)/b_n]\}.$$

Of course, $p_{in} = \delta_{in} - \delta_{i-1,n}$. Then B_n and $B_n^T B_n$ have the expressions found at (3.2) with $\theta, p_i, \varphi_i, \nu_i$ replaced by $\theta_n, p_{in}, \varphi_{in}, \nu_{in}$. The resulting statistic is, using (3.1) once more,

$$T_{4n} = \sum_{i=1}^{M} \frac{[N_{in}-np_{in}]^2}{np_{in}} + (n\Delta_n)^{-1}\{(\sum_{i=1}^{M} \nu_{in}^2/p_{in})(\sum_{i=1}^{M} N_{in}\varphi_{in}/p_{in})^2$$

$$-2(\sum_{i=1}^{M} \varphi_{in}\nu_{in}/p_{in})(\sum_{i=1}^{M} N_{in}\varphi_{in}/p_{in})(\sum_{i=1}^{M} N_{in}\nu_{in}/p_{in})$$

$$+ (\sum_{i=1}^{M} \varphi_{in}^2/p_{in})(\sum_{i=1}^{M} N_{in}\nu_{in}/p_{in})^2\}$$

where

$$\Delta_n = (\sum_{i=1}^{M} \varphi_{in}^2/p_{in})(\sum_{i=1}^{M} \nu_{in}^2/p_{in}) - (\sum_{i=1}^{M} \varphi_{in}\nu_{in}/p_{in})^2.$$

Once again Δ_n may be replaced by its limit in probability, obtained by substituting φ_i, ν_i and p_i. And once again the successive difference form of p_{in}, φ_{in} and ν_{in} makes T_{3n} more easily computable in practice than may at first appear.

REFERENCES

[1] BAIN, L. J. (1972). Inferences based on censored sampling from the Weibull or extreme-value distribution. Technometrics 14, 693-702.

[2] CHERNOFF, H., GASTWIRTH, J. L. and JOHNS, M. V. (1967). Asymptotic distribution of linear combinations of functions of order statistics with applications to estimation. Ann. Math. Statist. 38, 52-72.

[3] CHERNOFF, H. and LEHMANN, E. L. (1954). The use of maximum likelihood estimates in χ^2 test for goodness of fit. Ann. Math. Statist. 25, 579-586.

[4] DZHAPARIDZE, K. O. and NIKULIN, M. S. (1974). On a modification of the standard statistics of Pearson. Theor. Probability Appl. 19, 851-853.

[5] EPSTEIN, B. and SOBEL, M. (1953). Life testing. J. Amer. Statist. Assoc. 48, 486-502.

[6] HALPERIN, M. (1952). Maximum likelihood estimation in truncated samples. Ann. Math. Statist. 23, 226-238.

[7] HSUAN, A. and ROBSON, D. S. (1976). The χ^2-goodness-of-fit tests with moment type estimators. Comm. Statist. A5, 1509-1519.

[8] KOZIOL, J. A. and BYAR, D. P. (1975). Percentage points of the asymptotic distribution of one and two sample K-S statistics for truncated or censored data. Technometrics 17, 507-510.

[9] KOZIOL, J. A. and GREEN, S. B. A Cramér-von Mises statistic for randomly censored data. Biometrika 63, 465-474.

[10] LURIE, D., HARTLEY, H. O. and STROUD, M. R. (1974). A goodness of fit test for censored data. Comm. Statist. 3, 745-753.

[11] MANN, N. R., SCHAFER, R. E. and SINGPURWALLA, N. D. (1976). Methods for Statistical Analysis of Reliability and Life Data. Wiley, New York.

[12] MIHALKO, D. P. (1977). Chi-square tests of fit for Type II censored data. Ph.D. thesis, Department of Statistics, Purdue University.

[13] MOORE, D. S. (1977). Generalized inverses, Wald's method, and the construction of chi-squared tests of fit. Jour. Amer. Statist. Assoc. 72, 131-137.

[14] MOORE, D. S. and SPRUILL, M. C. (1975). Unified large-sample theory of general chi-squared statistics for tests of fit. Ann. Statist. 3, 599-616.

[15] PETTITT, A. N. (1976). Cramér-von Mises statistics for testing normality with censored samples. Biometrika 63, 475-482.

[16] RAO, K. C. and ROBSON, D. S. (1974). A chi-square statistic for goodness-of-fit within the exponential family. Comm. Statist. 3, 1139-1153.

[17] ROY, A. R. (1956). On χ^2 statistics with variable intervals. Technical Report No. 1, Stanford University Department of Statistics.

[18] SARHAN, A. E. (1955). Estimation of the mean and standard deviation by order statistics. Part III. Ann. Math. Statist. 26, 572-592.

[19] SHORACK, G. R. (1972). Functions of order statistics. Ann. Math. Statist. 43, 412-427.

[20] SMITH, R. M. and BAIN, L. J. (1976). Correlation type goodness-of-fit statistics with censored samples. Comm. Statist. A5, 119-132.

[21] TURNBULL, B. W. and WEISS, L. (1976). A likelihood ratio statistic for testing goodness of fit with randomly censored data. Technical Report No. 307, School of Operations Research and Industrial Engineering, Cornell University.

[22] WATSON, G. S. (1958). On chi-square goodness-of-fit tests for continuous distributions. J. Roy. Statist. Soc. Ser. B **20**, 44-61.

[23] WATSON, G. S. (1959). Some recent results in chi-square goodness-of-fit tests. Biometrics **15**, 440-468.