*The Annals of Statistics*
1993, Vol. 21, No. 3, 1621–1639

# CHI-SQUARE GOODNESS-OF-FIT TESTS FOR RANDOMLY CENSORED DATA[1]

BY JOO HAN KIM

*Chungnam National University*

We consider general chi-square goodness-of-fit test statistics for randomly censored data—call these generalized Pearson statistics—which are nonnegative definite quadratic forms in the cell frequencies obtained from the product-limit estimator, allowing random cells and general estimators of nuisance parameters. This class of statistics generalizes the class studied by Moore and Spruill in the no censoring case. The large sample behavior of these statistics under the null hypothesis and local alternatives is presented. The chi-square type statistics based on the observed cell frequencies obtained from the product-limit estimator are members of this class for which the quadratic form is selected to produce a chi-square asymptotic null distribution. The generalized Pearson statistic and the statistic by Akritas for a simple hypothesis are compared on the basis of asymptotic relative Pitman efficiency. It is shown that neither statistic dominates the other. The efficiencies are shown to depend on the degree of censoring and the number of cells. For heavily censored data, the Akritas statistic is superior to the generalized Pearson statistic. In the uncensored case, the Akritas statistic, which does not reduce to the Pearson statistic, is not as good as the Pearson statistic in the sense of Pitman efficiency.

**1. Introduction.** Under the random censorship model, we assume that the responses $X_1, \ldots, X_n$ are independent nonnegative random variables with continuous distribution function $F$. The censoring variables $Y_1, \ldots, Y_n$ are also nonnegative and are assumed to be a random sample, drawn independently of the $X_j$'s, from a population with continuous distribution function $G$. We say that the $X_j$'s are censored on the right by the $Y_j$'s since we can only observe $Z_j = \min(X_j, Y_j)$ and $\delta_j = I[Z_j = X_j]$, which indicates whether $Z_j$ is an uncensored observation or not. The problem of goodness-of-fit for censored data is to test the null hypothesis that $F$ is a member of a family $\{F(\cdot|\theta)\}$ of distribution functions indexed by a parameter $\theta$ running over a parameter space $\Omega$.

If there exists no censoring, the first step of the standard procedure for a chi-square test of fit is to partition the range of the response variable into $k + 1$ cells $A_i$, $i = 1, \ldots, k + 1$. After the sample is taken, the observed cell frequencies $n_i$ are obtained by counting the numbers of observations falling into $A_i$. These observed cell frequencies can be expressed as $n_i = n\int_{A_i} dF_e$

Received May 1989; revised December 1991.
[1]Research supported in part by NSF Grant DMS-87-40401.
*AMS* 1991 *subject classifications*. Primary 62E20; secondary 62G20.
*Key words and phrases*. Chi-square tests, goodness-of-fit, limiting distributions, random censoring, Pitman efficiency.

where $F_e$ is the empirical distribution function. Then, using an estimate $\theta_n$ for $\theta$, the expected cell probabilities $p_i(\theta) = \int_{A_i} dF(x|\theta)$ are estimated under the null hypothesis by $p_i(\theta_n)$. A general chi-square statistic is a nonnegative definite quadratic form in the vector of standardized cell frequencies $W_n = (w_{n1}, \ldots, w_{n(k+1)})'$, where $w_{ni} = (n_i - np_i(\theta_n))/\sqrt{np_i(\theta_n)}$. A general asymptotic theory for these statistics appears in Moore and Spruill (1975). The most useful such statistics have a chi-square limiting null distribution.

If we have censored observations in the sample, the empirical distribution function is is no longer a consistent estimator of $F$. The product-limit estimator of $F$ introduced by Kaplan and Meier (1958) is the most commonly used estimator for censored data. It is a consistent estimator of $F$ and also reduces to the usual empirical distribution function in the case of no censoring.

Let $\hat{F}$ be the product-limit estimators of $F$. The observed cell probabilities are $u_i = \int_{A_i} d\hat{F}$. The expected cell probabilities can be estimated under the null hypothesis by $p_i(\theta_n) = \int_{A_i} dF(x|\theta_n)$, where $\theta_n$ is an estimate of unknown $\theta$. The vector of observed minus expected cell probabilities $V_n(\theta) = (v_{n1}(\theta), \ldots, v_{nk}(\theta))'$, where $v_{ni}(\theta) = \sqrt{n}(u_i - p_i(\theta))$, has a normal limiting null distribution for fixed $\theta$. Such asymptotic properties of the product-limit estimator were established initially by Breslow and Crowley (1974) and more generally by Gill (1983). Using this fact we can get chi-square statistics—call these generalized Pearson statistics—as nonnegative definite quadratic forms in $V_n(\theta)$ with suitable centering matrices.

Chi-square tests for type II censored data were developed by Mihalko and Moore (1980). For randomly censored data, Chen (1975) proposed generalized Pearson type chi-square tests for simple and composite null hypotheses using the product-limit estimator. He used a modified minimum chi-square estimator for composite null hypotheses. Habib and Thomas (1986) proposed a statistic for composite hypotheses using the maximum likelihood estimator. Turnbull and Weiss (1978) considered a likelihood ratio statistic applicable for discrete or grouped censored data with finite support. Hjort (1984, 1990) proposed goodness-of-fit tests based on a weighted version of the cumulative hazard process. Akritas (1988) introduced Pearson type goodness-of-fit test statistics based on the number of the uncensored observations in each cell.

Necessary notation and assumptions are introduced in Section 2. In Section 3, we present asymptotic theory for general chi-square statistics, which are nonnegative definite quadratic forms in the cell frequencies obtained from the product-limit estimator, allowing random cells and general estimators of nuisance parameters. This class of statistics generalizes the class studied by Moore and Spruill (1975) in the uncensored case. This general theory includes tests based on nonnegative definite quadratic forms in cell frequencies obtained from the product-limit estimator such as the statistics by Chen (1975), Habib and Thomas (1986) and Kim (1988), but it cannot include other types of statistics. For example, the chi-square statistic by Akritas (1988) cannot be studied in our framework.

Akritas (1988) introduced chi-square statistics for randomly censored data based on the number of uncensored observations in each cell. The resulting

chi-square statistics have one more degree of freedom than the corresponding generalized Pearson chi-square statistics and they do not reduce to the Pearson statistics in the no censoring case. The Akritas statistic and the Pitman efficiency are discussed in Section 4 and Section 5. For the Pitman efficiency, we will use the general definition of Rothe (1981) which can be used to derive the Pitman efficiency for two chi-square distributed test statistics with different degrees of freedom. In Section 6, the generalized Pearson statistic and the statistic by Akritas (1988) for a simple hypothesis are compared on the basis of the asymptotic relative Pitman efficiency. It is shown that neither statistic dominates the other. The efficiency is shown to depend on the degree of censoring and the number of cells. The Akritas statistic is superior to the generalized Pearson statistic if we have heavily censored data. In the uncensored case, the Akritas statistic, which does not reduce to the Pearson statistic, is not as good as the Pearson statistic in the sense of the Pitman efficiency. These facts are illustrated by computing the efficiencies of two statistics for testing fit to the family of exponential distributions.

**2. Notation and assumptions.** We observe $Z_j = \min(X_j, Y_j)$ and $\delta_j = I[Z_j = X_j]$, $j = 1, \ldots, n$, where $X_j$'s are i.i.d. with continuous distribution function $F(x|\theta, \eta)$ and $Y_j$'s are also i.i.d. with continuous distribution function $G(y)$. Both $X_j$'s and $Y_j$'s are nonnegative and they are independent. The parameter $\theta$ ranges over an open set $\Omega_1$ in $R^s$ and $\eta$ ranges over a neighborhood of a point $\eta_0$ in $R^m$. We write $F(x|\theta, \eta_0) = F(x|\theta)$, so that the composite null hypothesis that the $X_i$ have a distribution function in the family $F(x|\theta)$ becomes $H_0$: $\eta = \eta_0$. We will present the large sample behavior of tests for $H_0$ under the sequence of parameter values $(\theta_0, \eta_n)$ where $\theta_0 \in \Omega_1$ and $\eta_n = \eta_0 + n^{-1/2}\gamma$ for fixed $\gamma \in R^m$. This model was used by Moore and Spruill (1975), Durbin (1973) and Chibisov (1971) and covers many common alternatives such as the *contamination alternative* under which

$$F(x|\theta, \eta) = (1 - \eta)F(x|\theta) + \eta K(x),$$

where $0 \leq \eta \leq 1$ and $K$ is a fixed distribution function.

The cells for chi-square tests are intervals in $R^1$ whose boundaries are functions of a variable $\varphi$ defined on an open set $\Omega_2 \in R^r$. The resulting cells are denoted by $A_i(\varphi) = [a_{i-1}(\varphi), a_i(\varphi))$, where $0 = a_0(\varphi) < a_1(\varphi) < \cdots < a_k(\varphi) < a_{k+1}(\varphi) = \infty$. In common cases $r = s$ and $\varphi$ is replaced by an estimator of $\theta$. In general, since the parameter $\theta$ is unknown, $\theta$ has to be estimated by an estimator $\theta_n = \theta_n(Z_1, \ldots, Z_n, \delta_1, \ldots, \delta_n)$ satisfying $\theta_n - \theta_0 = O_p(n^{-1/2})$. For the cells, we will use $\varphi_n = \varphi_n(Z_1, \ldots, Z_n, \delta_1, \ldots, \delta_n)$ which satisfies $\varphi_n - \varphi_0 = o_p(1)$. The arguments $\theta, \varphi, \eta$ will usually be suppressed when they take the values $\theta_0, \varphi_0, \eta_0$, respectively. Expected values and derivatives are computed under $(\theta_0, \eta_0)$ unless otherwise stated.

Define the observed cell probability $u_{ni}(\varphi)$ for each cell by $u_{ni}(\varphi) = \int_{A_i(\varphi)} d\hat{F}(x)$ where $\hat{F}$ is the product-limit estimator of $F$. The expected cell

probability for each cell under $(\theta, \eta)$ is

$$p_i(\theta, \eta, \varphi) = \int_{A_i(\varphi)} dF(x|\theta, \eta), \qquad i = 1, \ldots, k. \tag{2.1}$$

Each $p_i(\theta, \eta, \varphi)$ is estimated by $p_i(\theta_n, \varphi_n)$ under the null hypothesis. Let

$$v_{ni}(\theta, \eta, \varphi) = \sqrt{n}\,(u_{ni}(\varphi) - p_i(\theta, \eta, \varphi)), \qquad i = 1, \ldots, k \tag{2.2}$$

and $V_n(\theta, \eta, \varphi)$ be the $k$-vector whose $i$th component is $v_{ni}(\theta, \eta, \varphi)$.

A general chi-square statistic has the form

$$T_n = V_n'(\theta_n, \varphi_n) K_n V_n(\theta_n, \varphi_n), \tag{2.3}$$

where $K_n$ is a nonnegative definite, possibly random, symmetric $k \times k$ matrix converging to a fixed nonnegative definite matrix $K$. In some cases, $K_n$ will be a generalized inverse of a consistent estimator of the asymptotic variance-covariance matrix of the random vector $V_n(\theta_n, \varphi_n)$.

We impose the following assumptions, which are almost the same as the ones used by Moore and Spruill (1975), slightly changed to fit the random censoring case.

ASSUMPTION A1. Under $(\theta_0, \eta_n)$, $\theta_n - \theta_0 = O_p(n^{-1/2})$ and $\varphi_n - \varphi_0 = o_p(1)$. The cell boundaries $a_i(\varphi)$ are real valued continuous functions of $\varphi$ in a neighborhood of $\varphi_0$.

ASSUMPTION A2. For each $i$, $p_i(\theta, \eta, \varphi)$ is continuous in $(\theta, \eta, \varphi)$ and continuously differentiable in $(\theta, \eta)$ in a neighborhood of $(\theta_0, \eta_0, \varphi_0)$. Moreover, $\Sigma p_i = 1$ and $p_i > 0$ for each $i$.

ASSUMPTION A3. $F(x)$ is continuous and $\sup_x |F(x|\eta_n) - F(x)| \to 0$ as $n \to \infty$.

ASSUMPTION A4. $K_n$ is a nonnegative definite, possibly random, $k \times k$ matrix which converges to a fixed nonnegative definite $k \times k$ matrix $K$ as $n \to \infty$.

ASSUMPTION A5. Under $(\theta_0, \eta_n)$,

$$\sqrt{n}\,(\theta_n - \theta_0) = n^{-1/2} \sum_{j=1}^{n} h(z_j, \delta_j, \eta_n) + A\gamma + o_p(1)$$

for some $s \times m$ matrix $A$ and measurable function $h(z, \delta, \eta)$ from $R \times \{0, 1\} \times R^m$ to $R^s$ satisfying

$$E_{(\theta_0, \eta_n)}[h(z, \delta, \eta_n)] = 0 \quad \text{and} \quad E_{(\theta_0, \eta_n)}[h(z, \delta, \eta_n) h(z, \delta, \eta_n)'] = L(\eta_n),$$

where $L(\eta_n)$ is a nonnegative definite matrix converging to the finite nonnegative definite matrix $L = E[h(z, \delta) h(z, \delta)']$ as $n \to \infty$.

ASSUMPTION A6. The distribution functions $F(x|\eta)$ and $G(x)$ possess the probability density functions $f(x|\eta)$ and $g(x)$ with respect to a $\sigma$-finite

dominating measure $\zeta$. As $n \to \infty$, $f(x|\eta_n) \to f(x|\eta_0)$ and $h(z, \delta, \eta_n) \to h(z, \delta, \eta_0)$ a.e. $(\zeta)$.

Assumption A1 is a usual assumption about $\theta_n$ and random cell boundaries. Assumptions A2 and A4 are familiar assumptions for a chi-square statistic. Assumption A3 is needed to handle the alternative case. In the null case we only need the continuity of $F(x)$.

Assumption A5 specifies the asymptotic behavior of the estimator $\theta_n$ under the sequence of alternatives. The raw data MLE $\hat{\theta}_n$ and the minimum chi-square estimator $\bar{\theta}_n$, which minimizes the chi-square statistic $V_n'(\theta)\Psi_n(\theta)V_n(\theta)$, where $\Psi_n(\theta)$ is a consistent estimator of the inverse of the asymptotic covariance matrix of $V_n(\theta)$, satisfy Assumption A5 in most cases. This extends Theorem 2.4.1 of Kim (1988) and results of Borgan (1984). Arguments of Davidson and Lever (1970) can be used to obtain the asymptotic forms (2.5) and (2.7) of estimators $\hat{\theta}_n$ and $\bar{\theta}_n$ as in Assumption A5 in regular cases. [See Kim (1988).] Assumption A5 is in fact satisfied in many cases in which the regularity conditions of Davidson and Lever (1970) do not hold.

Let $l(z, \delta|\theta, \eta) = \delta \log \alpha(z|\theta, \eta) - \int_0^z \alpha(u|\theta, \eta)\, du$, where $\alpha(z|\theta, \eta) = f(z|\theta, \eta)/(1 - F(z|\theta, \eta))$. Define an $s \times s$ matrix $J$ and an $s \times m$ matrix $J_{12}$ by

$$J = \left[-E\left(\frac{\partial^2 l(z, \delta|\theta, \eta)}{\partial\theta_i\, \partial\theta_j}\right)\right], \qquad J_{12} = \left[-E\left(\frac{\partial^2 l(z, \delta|\theta, \eta)}{\partial\theta_i\, \partial\eta_j}\right)\right]. \tag{2.4}$$

Then in regular cases we have under $(\theta_0, \eta_n)$,

$$\sqrt{n}\left(\hat{\theta}_n - \theta_0\right) = n^{-1/2} \sum_{j=1}^{n} J^{-1}\left(\frac{\partial l\left(z_j, \delta_j|\theta, \eta_n\right)}{\partial\theta_i}\right)_{\theta=\theta_0} + J^{-1}J_{12}\gamma + o_p(1). \tag{2.5}$$

In the case of minimum chi-square estimators $\bar{\theta}_n$ based on the random cells $A_i(\varphi_n)$ where $\varphi_n - \varphi_0 = o_p(1)$, under suitable regularity conditions, these estimators in the random cell case have the same limiting behavior as in the fixed cell case under the null hypothesis.

Let $\Psi$ be the inverse of the asymptotic covariance matrix of $V_n$. The elements of $\Psi$ are listed in (3.1). Define a $k \times s$ matrix $B$ and a $k \times m$ matrix $B_{12}$ by

$$B = \left(\frac{\partial p_i(\theta, \eta)}{\partial\theta_j}\right), \qquad B_{12} = \left(\frac{\partial p_i(\theta, \eta)}{\partial\eta_j}\right). \tag{2.6}$$

Then, under $(\theta_0, \eta_n)$, we have the following form for $\bar{\theta}_n$ which satisfies Assumption A5:

$$\begin{aligned}\sqrt{n}\left(\bar{\theta}_n - \theta_0\right) = n^{-1/2} \sum_{j=1}^{n} &\left(B'\Psi B\right)^{-1} B'\Psi W\left(z_j, \delta_j|\eta_n\right) \\ &+ \left(B'\Psi B\right)^{-1} B'\Psi B_{12}\gamma + o_p(1).\end{aligned} \tag{2.7}$$

Here $W(z, \delta|\eta)$ is the function defined in Lemma 3.2.

Assumption A6 is needed to obtain the limiting distribution of $T_n$ with estimators $\theta_n$ satisfying Assumption A5 under $(\theta_0, \eta_n)$. It is not needed when limiting null distributions are being studied. In general, the assumptions become less restrictive if only the null case is of interest.

**3. General chi-square statistics for randomly censored data.** Let $y_n(t)$ be the stochastic process $y_n(t) = \sqrt{n}\,(\hat{F}(t) - F(t|\eta_n))$. The weak convergence of $y_n(t)$ to a process with continuous sample paths can be proved by using Theorem 4.1.1 and Theorem 4.2.1 of Gill (1981).

When we use random cells, the differences between the random cells actually used and the fixed cells which they approach produce error terms in the large sample form of chi-square statistics. The following lemma, that can be proved easily using the usually random change of time argument as in Billingsley [(1968), page 145], shows that those terms converge to zero in probability.

LEMMA 3.1. *Suppose Assumptions* A1, A2 *and* A3 *hold. Then under* $(\theta_0, \eta_n)$,

$$y_n(a_i(\varphi_n)) - y_n(a_i(\varphi_0)) = o_p(1).$$

Now we can describe the limiting behavior of the vector $V_n$ defined in (2.2). Let $B$ and $B_{12}$ be the matrices defined in (2.6). The following result extends Theorem 4.1 of Moore and Spruill (1975) to the case of censored data. This can be proved by using arguments similar to those employed in Theorem 4.1 of Moore and Spruill. [See Kim (1988).]

THEOREM 3.1. *If Assumptions* A1, A2 *and* A3 *hold, then under* $(\theta_0, \eta_n)$,

$$V_n(\theta_n, \varphi_n) = V_n(\eta_n) - B\sqrt{n}\,(\theta_n - \theta_0) + B_{12}\gamma + o_p(1).$$

The following lemma states that the vector $V_n$ can be expressed as a normalized sum of continuous functions of $(Z_1, \delta_1), \ldots, (Z_n, \delta_n)$. This can be obtained from Breslow and Crowley's (1974) results. [See Kim (1988).]

LEMMA 3.2. *Let* $q(t|\eta) = 1 - F(t|\eta)$ *and* $H$ *be the distribution function of* $Z_i$. *Define a continuous, nonnegative, nondecreasing function* $C(t|\eta)$ *by*

$$C(t|\eta) = \int_0^t \frac{dF(s|\eta)}{(1 - F(s|\eta))^2(1 - G(s))}.$$

*Then, for fixed* $\eta$,

$$V_n(\eta) = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} W(Z_j, \delta_j|\eta) + o_p(1).$$

*The ith component of the k-vector* $W(Z_j, \delta_j|\eta)$ *is given by*

$$w_i(Z_j, \delta_j|\eta) = k_i(Z_j, \delta_j|\eta) - k_{i-1}(Z_j, \delta_j|\eta),$$

*where*

$$k_i(Z_j, \delta_j|\eta) = q(a_i|\eta)\Big(C(a_i|\eta) - I[Z_j < a_i]\int_{Z_j}^{a_i} dC(u|\eta) - I[Z_j < a_i, \delta_j = 1]\big(1 - H(Z_j|\eta)\big)^{-1}\Big).$$

It requires more notation to describe the limiting distribution of a general chi-square statistic of the form $T_n$ given in (2.3). $A$, $h$ and $L$ are as in Assumption A5, $S$ is a $k \times k$ matrix such that $K = SS'$, $B$ and $B_{12}$ are as in (2.6). Let $W(z, \delta) = W(z, \delta|\eta_0)$ which is defined in Lemma 3.2. Now define

$$\mu = (B_{12} - BA)\gamma,$$

$$\mu_0 = S'\mu,$$

$$\Sigma = \Gamma + BLB' - BE[h(Z, \delta)W(Z, \delta)'] - E[W(Z, \delta)h(Z, \delta)']B',$$

$$\Sigma_0 = S'\Sigma S,$$

where $\Gamma = E[WW']$ is the asymptotic covariance matrix of $V_n$. The following form of the elements of $\Psi = \Gamma^{-1}$ are derived in Kim (1988):

$$\psi_{ii} = \frac{1}{r_i q_i^2} + \sum_{m=i+1}^{k} \frac{p_m^2}{r_m q_m^2 q_{m-1}^2}, \qquad i = 1, \ldots, k-1,$$

$$\psi_{kk} = \frac{1}{r_k q_k^2}, \tag{3.1}$$

$$\psi_{ij} = \psi_{ji} = \frac{p_j}{r_j q_{j-1} q_j^2} + \sum_{m=j+1}^{k} \frac{p_m^2}{r_m q_m^2 q_{m-1}^2}, \qquad i < j,$$

where

$$q_i = 1 - F(a_i) \quad \text{for } i = 0, 1, \ldots, k, \tag{3.2}$$

$$r_i = \int_{A_i} \frac{dF(x)}{(1 - F(x))^2(1 - G(x))} \quad \text{for } i = 1, \ldots, k. \tag{3.3}$$

The limiting distributions of $T_n$ under the null hypothesis and local alternatives are presented in the next theorem. This extends Theorem 4.2 of Moore and Spruill (1975) to the censored data case.

THEOREM 3.2. *If Assumptions* A1–A5 *hold with* $\eta = \eta_0$ *and* $\gamma = 0$, *then under* $(\theta_0, \eta_0)$ *the limiting distribution of* $T_n$ *is the distribution of*

$$\sum_{j=1}^{k} \lambda_j \chi_{1j}^2,$$

*where* $\lambda_j$ *are the characteristic roots of* $\Sigma_0$ *and the* $\chi_{1j}^2$ *are independent* $\chi^2$ *random variables with one degree of freedom.*

*If Assumptions* A1–A6 *hold,* $T_n$ *has as its limiting distribution under* $(\theta_0, \eta_n)$ *the distribution of*

$$\sum_{\lambda_j \neq 0} \lambda_j \chi_{1j}^2\left(\nu_j^2/\lambda_j\right) + \sum_{\lambda_j = 0} \nu_j^2,$$

*where* $\chi_{1j}^2(\nu_j^2/\lambda_j)$ *are independent noncentral* $\chi^2$ *random variables with one degree of freedom and noncentrality parameter* $\nu_j^2/\lambda_j$, *and* $\nu_j$ *are the components of the vector* $\nu = P'\mu_0$, *where* $P$ *is an orthogonal matrix such that* $P'\Sigma_0 P$ *is a diagonal matrix with diagonal elements* $\lambda_1, \ldots, \lambda_k$.

The proof of this theorem is similar to the proof of Theorem 4.2 of Moore and Spruill. See Kim (1988) for details. Theorem 3.2 seems to be hard to apply because of the complicated form of $\Sigma$. But we will see in the next example that for the usual chi-square statistics $\Sigma$ simplifies immediately.

EXAMPLE 3.1. In this example we will show that the limiting distribution of the chi-square statistics introduced by Chen (1975), Habib and Thomas (1986) and Kim (1988) can be derived using Theorem 3.2. Since we derive the limiting distributions of test statistics under the null hypothesis, $\mu$ and $\mu_0$ equal zero in this example.

CASE 1 (Simple hypothesis case). For a simple hypothesis, since there is no parameters to be estimated, $\Sigma = \Gamma$. Let $\hat{r}_i$ be an estimator of $r_i$ defined in (3.3) which is obtained by replacing the unknown distribution function $G$ in $r_i$ by the product-limit estimator $\hat{G}$. Let $K_n = \Psi_n$ where $\Psi_n$ is the result of replacing the $r_i$ in $\Psi$ by the $\hat{r}_i$. Since $\Psi_n$ is a consistent estimator of $\Psi$ and $\Psi = \Gamma^{-1}$, $K = \Psi$ and $\Sigma_0 = I_k$. Applying Theorem 3.2 establishes

$$Q = V_n'\Psi_n V_n = n \sum_{i=1}^{k} \frac{(d_i - p_i)^2}{r_i q_i^2 q_{i-1}^2} \to_d \chi_k^2, \tag{3.4}$$

where $q_i$ is as in (3.9) and $d_i = q_{i-1}\hat{F}(a_i) - q_i\hat{F}(a_{i-1})$. This chi-square statistic reduces to the Pearson statistic if there is no censoring.

CASE 2 (Composite hypothesis case using the minimum chi-square estimator). In this case, the chi-square statistic is $\bar{Q} = V_n'(\bar{\theta}_n)\Psi_n(\bar{\theta}_n)V_n(\bar{\theta}_n)$. Here $\bar{\theta}_n$ is the minimum chi-square estimator which minimizes $Q(\theta) = V_n'(\theta)\Psi_n(\theta)V_n(\theta)$. The centering matrix is $K_n = \Psi_n(\bar{\theta}_n)$ and it converges to $K = \Psi$. Therefore, by (2.7), $\Sigma = \Psi^{-1} - B(B'\Psi B)^{-1}B'$ and $\Sigma_0 = I_k - \Psi^{1/2}B(B'\Psi B)^{-1}B'\Psi^{1/2}$, where

$\Psi^{1/2}$ is a symmetric square root of $\Psi$ and $B$ is the matrix defined in (2.6). $\Sigma_0$ is an idempotent matrix with rank $k - s$. Hence $\overline{Q} \to_d \chi^2_{k-s}$ by Theorem 3.2. This generalizes the Pearson–Fischer statistic in the uncensored case.

CASE 3 (Composite hypothesis case using the maximum likelihood estimator).

(i) Let $\hat{\theta}_n$ be the maximum likelihood estimator that has the asymptotic form (2.5). Replacing the unknown parameter $\theta$ in $Q(\theta)$ by $\hat{\theta}_n$ yields $\hat{Q}_1 = V_n'(\hat{\theta}_n)\Psi_n(\hat{\theta}_n)V_n(\hat{\theta}_n)$. $K_n = \Psi_n(\theta_n)$ converges to $K = \Psi$ and $\Sigma = \Psi^{-1} - BJ^{-1}B'$ where $J$ is the matrix defined in (2.4). So $\Sigma_0 = I_k - \Psi^{1/2}BJ^{-1}B'\Psi^{1/2}$. Therefore $Q_1 \to_d \sum_{j=1}^k \lambda_j \chi^2_{1j}$ where $\lambda_j$ are the characteristic roots of $\Sigma_0$ and $\chi^2_{1j}$ are independent $\chi^2_1$ random variables. This extends the result of Chernoff and Lehmann (1954) to the censored data case. The statistic $\hat{Q}_1$ is not useful because the limiting distribution is not a chi-square distribution and it depends on the true parameter value $\theta_0$.

(ii) To get a chi-square limiting distribution, let $\hat{\Sigma}(\theta) = \Psi_n^{-1}(\theta) - B(\theta)J_n^{-1}(\theta)B'(\theta)$, where

$$J_n(\theta) = -\frac{1}{n}\left[\int \frac{\partial^2}{\partial\theta_i\,\partial\theta_j}\log\alpha(u|\theta)\,dN_1(u) - \int \frac{\partial^2}{\partial\theta_i\,\partial\theta_j}\alpha(u|\theta)N(u)\,du\right].$$

Here $N_1(u) = \sum_{j=1}^n I[Z_j \le u, \delta_j = 1]$ and $N(u) = \sum_{j=1}^n I[Z_j \ge u]$. Borgan (1984) showed that $J_n(\theta)$ is a consistent estimator of $J(\theta)$ for fixed $\theta$. So $K_n = \hat{\Sigma}^{-1}(\hat{\theta}_n)$ converges to $K = \Sigma^{-1}$. Now $\hat{Q} = V_n'(\hat{\theta}_n)\hat{\Sigma}^{-1}(\hat{\theta}_n)V_n(\hat{\theta}_n) \to_d \chi^2_k$ follows from Theorem 3.2 together with $\Sigma_0 = I_k$. This statistic reduces to the chi-square statistic by Rao and Robson (1975) when no censoring is present.

**4. Akritas statistics.** Akritas (1988) introduced chi-square statistics for randomly censored data based on the number of uncensored observations in each cell. The model in Section 2 will be used without the parameter $\theta$. Suppose that the cell boundaries are fixed. Let $H$ be the distribution function of $Z_j$ and $\hat{H}$ be the empirical distribution function defined by $\hat{H}(t) = (1/n)\sum_{j=1}^n I[Z_j \le t]$. Note that $1 - H = (1 - F)(1 - G)$. Define an estimator $\tilde{G}$ of $G$ under the null hypothesis, that is $\eta = \eta_0$, by $\tilde{G}(x) = 1 - (1 - \hat{H}(x))/(1 - F(x))$. This $\tilde{G}$ will replace the unknown censoring distribution function $G$ in the Akritas statistic.

Let $n_{1i}$ be the number of uncensored observations in each cell, that is $n_{1i} = \sum_{j=1}^n I[Z_j \in A_i, \delta_j = 1]$. The expectation of $n_{1i}/n$ is given by

$$\pi_{1i}(\eta) = \int_{A_i} (1 - G(x))\,dF(x|\eta). \tag{4.1}$$

This is the $i$th component of the $(k + 1)$-vector $\pi_1(\eta)$. Denote by $\tilde{\pi}_1(\eta)$ the result of replacing the unknown censoring distribution function $G$ in $\pi_1(\eta)$ by $\tilde{G}$. Define $W_n(\eta) = (w_{n1}(\eta), \ldots, w_{n(k+1)}(\eta))'$ where $w_{ni}(\eta) = \sqrt{n}\,(n_{1i}/n - \tilde{\pi}_{1i}(\eta))$. Let $D_{\tilde{\pi}}(\eta)$ be the diagonal matrix with elements $\tilde{\pi}_{1i}(\eta)$. The Akritas

statistic [see Akritas (1988)] for a simple hypothesis is given by

$$Q_A = W_n' D_{\tilde{\pi}}^{-1} W_n = \sum_{i=1}^{k+1} \frac{(n_{1i} - n\tilde{\pi}_{1i})^2}{n\tilde{\pi}_{1i}}, \tag{4.2}$$

which has asymptotically the $\chi^2_{k+1}$ distribution under the null hypothesis.

**5. Pitman efficiencies.** Rothe (1981) introduced a general approach to Pitman efficiency as a limit of sample sizes. His results can be used to compute the Pitman efficiency of tests based on asymptotically $\chi^2$-distributed test statistics with different degrees of freedom.

For $\eta_0 \in \Omega$, let $\{\phi_n\}$ be a sequence of level $\alpha$ tests based on $n$ observations for $H_0$: $\eta = \eta_0$. We assume $E_\eta(\phi_n) \geq \alpha$, $\lim_{n\to\infty} E_\eta(\phi_n) = 1$, and $\{\eta_0\} \neq C(\eta_0)$ where $C(\eta_0)$ is the connected component of $\eta_0$. Let $\Omega' = \Omega - \{\eta_0\}$. Let $\Pi$ be the set of all sequences $\{\eta_n\}$ satisfying $\eta_n \in \Omega'$ and $\eta_n \to \eta_0$. Rothe considered sequences of tests which satisfy Conditions P1–P3 given below.

CONDITION P1. There are functions $\rho$: $\Omega \to (0, \infty)$ and $\tau$: $(0, \infty) \to (\alpha, 1)$ such that:

(a) $\rho$ is continuous and $\rho(\eta) = 0$ if and only if $\eta = \eta_0$.
(b) $\tau$ is strictly increasing and bijective.
(c) For sequences $\{\eta_n\}$ in $\Omega$ satisfying $\rho(\eta_n)n \to \zeta \geq 0$ as $n \to \infty$ we have $\lim_{n\to\infty} E_{\eta_n}(\phi_n) = \tau(\zeta)$.

CONDITION P2. For every $n$, the function $\psi_n$: $\eta \to E_\eta(\phi_n)$ is continuous at $\eta = \eta_0$.

CONDITION P3. For every sequence $\{\eta_n\} \in \Pi$ such that $\rho(\eta_n)n \to \infty$, $E_{\eta_n}(\phi_n) \to 1$.

If two sequences of tests satisfy Conditions P1–P3 with functions $\rho_1, \rho_2$ and $\tau_1, \tau_2$, respectively, and also

$$\rho_{12} = \inf_{\Pi} \liminf_{n\to\infty} \frac{\rho_1(\eta_n)}{\rho_2(\eta_n)} = \sup_{\Pi} \limsup_{n\to\infty} \frac{\rho_1(\eta_n)}{\rho_2(\eta_n)},$$

then Rothe showed that the asymptotic relative Pitman efficiency of two tests exists and is given by $\rho_{12}(\tau_2^{-1}(\beta)/\tau_1^{-1}(\beta))$, where $\beta$ is the power of the test.

Therefore, to calculate asymptotic relative Pitman efficiencies, we need to verify Conditions P1–P3 and find suitable $\rho$ and $\tau$. In most cases we can easily find $\rho$ and verify Conditions P2 and P3. The following theorem by Rothe is a useful tool in finding $\tau$ as in Condition P1 for an upper level $\alpha$ test based on a statistic which has a limiting normal or chi-square distribution.

THEOREM 5.1. *Let $\phi_n$ be an upper level $\alpha$ test based on a test statistic $T_n$ which has one of the following asymptotic properties $C_k$ for $k \geq 0$.*

$C_0$. *There is a* $u > 0$ *such that* $\rho(\eta_n)n \to \zeta$ *implies* $\mathscr{L}(T_n|\eta_n) \to N(\zeta^u, 1)$ *for every* $\zeta \geq 0$.

$C_k$. *There is a* $u > 0$ *such that* $\rho(\eta_n)n \to \zeta$ *implies* $\mathscr{L}(T_n|\eta_n) \to \chi^2(k, \zeta^{2u})$, *where* $\chi^2(k, \delta)$ *is a noncentral* $\chi^2$ *distribution with* $k$ *degrees of freedom and noncentrality parameter* $\delta$.

*Then, for* $0 < \alpha < \beta < 1$, *Condition* P1 *holds with* $\tau^{-1}(\beta) = (d(\alpha, \beta, k))^{1/u}$. *Here* $d(\alpha, \beta, 0) = \Phi^{-1}(\beta) - \Phi^{-1}(\alpha)$, *where* $\Phi$ *is the distribution function of the standard normal distribution and, for* $k \geq 1$, $d(\alpha, \beta, k) = \sqrt{\delta}$ *where* $\delta$ *is the uniquely determined noncentrality parameter such that the* $(1 - \beta)$*th quantile of* $\chi_k^2(\delta)$ *and the* $(1 - \alpha)$*th quantile of* $\chi_k^2$ *coincide.*

**6. Comparison of tests.** In this section we will compare the asymptotic performances of the generalized Pearson statistic $Q$ in (3.4) and the Akritas statistic $Q_A$ in (4.2). Let us consider the model used in Section 4 with fixed cell boundaries. Recall that our goodness-of-fit testing problem is equivalent to test $H_0$: $\eta = \eta_0$ under the model. Define

$$b_1 = \left( \frac{\partial p_1(\eta)}{\partial \eta}, \ldots, \frac{\partial p_k(\eta)}{\partial \eta} \right)'_{\eta = \eta_0}, \qquad b_2 = \left( \frac{\partial \pi_{11}(\eta)}{\partial \eta}, \ldots, \frac{\partial \pi_{1(k+1)}(\eta)}{\partial \eta} \right)'_{\eta = \eta_0},$$

where $p_i$'s and $\pi_{1i}$'s are as in (2.1) and (4.1).

THEOREM 6.1. *Assume that, for each* $i$, $p_i(\eta)$ *and* $\pi_{1i}(\eta)$ *are continuous in* $\eta$ *and continuously differentiable in a neighborhood of* $\eta_0$. *Then, for* $0 < \alpha < \beta < 1$,

(a) $Q$ *satisfies Conditions* P1–P3 *with*

$$\rho_1(\eta) = \tfrac{1}{2}(\eta - \eta_0)^2 b_1' \Psi b_1 \quad \text{and} \quad \tau_1^{-1}(\beta) = \delta(\alpha, \beta, k),$$

*where* $\delta(\alpha, \beta, k)$ *is the noncentrality parameter defined in Theorem* 5.1.

(b) $Q_A$ *satisfies Conditions* P1–P3 *with*

$$\rho_2(\eta) = \tfrac{1}{2}(\eta - \eta_0)^2 b_2' D_\pi^{-1} b_2 \quad \text{and} \quad \tau_2^{-1}(\beta) = \delta(\alpha, \beta, k + 1).$$

PROOF. We will only prove (a). (b) can be proved similarly. Conditions P2 and P3 can be verified easily. Suppose $n\rho_1(\eta_n) \to \zeta$, then $\sqrt{n}(\eta_n - \eta_0) \to \gamma$, where $\gamma = \pm\sqrt{\zeta/\lambda}$ and $\lambda = b_1' \Psi b_1/2$. Now, by the Taylor theorem,

$$\begin{aligned} \sqrt{n}(u_n - p(\eta_0)) &= \sqrt{n}(u_n - p(\eta_n)) + \sqrt{n}(p(\eta_n) - p(\eta_0)) \\ &= \sqrt{n}(u_n - p(\eta_n)) + \gamma b_1 + o(1). \end{aligned}$$

So, under $\eta_n$, $\sqrt{n}(u_n - p(\eta_0)) \to_d N_k(\gamma b_1, \Psi^{-1})$ and $Q \to_d \chi_k^2(\zeta)$. $Q$ therefore satisfies $C_k$ with $u = 1/2$. Hence, by Theorem 5.1, $Q$ satisfies Condition P1 with $\tau_1^{-1}(\beta) = \delta(\alpha, \beta, k)$. □

COROLLARY 6.1. *Under the assumption of Theorem* 6.1, *the asymptotic relative Pitman efficiency of* $Q$ *to* $Q_A$ *is given by*

$$e^P(Q, Q_A) = \frac{b_1'\Psi b_1}{b_2' D_\pi^{-1} b_2} \cdot \frac{\delta(\alpha, \beta, k+1)}{\delta(\alpha, \beta, k)}.$$

The Pitman efficiency in the corollary is a product of two terms, a ratio of quadratic forms and a ratio of noncentrality parameters. We present some facts about those two terms. We need the following lemma to prove that the ratio of the two noncentrality parameters is always greater than 1.

LEMMA 6.1. *Let* $G_\nu$ *be the distribution function of the* $\chi^2_\nu$ *distribution and* $g_\nu$ *be the density function. Let* $\xi_\nu(\alpha)$ *be the* $\alpha$*th quantile of the* $\chi^2_\nu$ *distribution. Then* $G_\nu(\xi_\nu(\alpha) - c)$ *is increasing with* $\nu$ *for any* $c$ *where* $0 < c < \xi_\nu(\alpha)$.

This lemma directly follows from the proof of (1.3) in Saunders and Moran (1978) via their (1.5) through (1.8). [See Kim (1988).]

THEOREM 6.2. *For all* $0 < \alpha < \beta < 1$ *and* $k$,

$$\frac{\delta(\alpha, \beta, k+1)}{\delta(\alpha, \beta, k)} > 1.$$

PROOF. Let $G_k$, $g_k$ and $\xi_k$ be defined as Lemma 6.1. Let $\Lambda(x; k, \lambda)$ be the distribution function of $\chi^2_k(\lambda)$ distribution. Then $\delta(\alpha, \beta, k+1)$ and $\delta(\alpha, \beta, k)$ satisfy

$$1 - \beta = \Lambda(\xi_{k+1}(1-\alpha); k+1, \delta(\alpha, \beta, k+1)) = \Lambda(\xi_k(1-\alpha); k, \delta(\alpha, \beta, k)).$$

Now, since $\Lambda(x; k, \lambda)$ is decreasing with $\lambda$ for fixed $x$ and $k$, it is enough to show that $\Lambda(\xi_{k+1}(1-\alpha); k+1, \lambda) > \Lambda(\xi_k(1-\alpha); k, \lambda)$ for any $\lambda$. The distribution function $\Lambda(\xi_k(1-\alpha); k, \lambda)$ can be expressed as

$$\Lambda(\xi_k(1-\alpha); k, \lambda) = \sum_{j=0}^{\infty} \left[ \frac{(\lambda/2)^j}{j!} e^{-\lambda/2} \right] G_{k+2j}(\xi_k(1-\alpha)).$$

[See Johnson and Kotz (1970), page 132.] Now, by Lemma 6.1, for all $j = 1, 2, \ldots,$

$$\begin{aligned} G_{k+2j}(\xi_k(1-\alpha)) &= \int_0^{\xi_k(1-\alpha)} G_k(\xi_k(1-\alpha) - u) g_{2j}(u)\, du \\ &< \int_0^{\xi_k(1-\alpha)} G_{k+1}(\xi_{k+1}(1-\alpha) - u) g_{2j}(u)\, du \\ &< \int_0^{\xi_{k+1}(1-\alpha)} G_{k+1}(\xi_{k+1}(1-\alpha) - u) g_{2j}(u)\, du \\ &= G_{k+2j+1}(\xi_{k+1}(1-\alpha)). \end{aligned}$$

Hence $\Lambda(\xi_{k+1}(1-\alpha); k+1, \lambda) > \Lambda(\xi_k(1-\alpha); k, \lambda)$ for all $\lambda$. This completes the proof. □

In the next theorem we will prove that the ratio of the two noncentrality parameters converges to 1 as $\alpha$ goes to 0.

THEOREM 6.3. *For all $k$ and $0 < \beta < 1$,*

$$\lim_{\alpha \to 0} \frac{\delta(\alpha, \beta, k+1)}{\delta(\alpha, \beta, k)} = 1.$$

PROOF. Let $\delta_1 = \delta(\alpha, \beta, k)$ and $\xi_k(\alpha)$ be the $\alpha$th quantile of the $\chi_k^2$ distribution. A noncentral chi-square random variable with $k$ degrees of freedom and noncentrality parameter $\delta_1$ can be expressed as $Y + (Z + \sqrt{\delta_1})^2$, where $Y$ and $Z$ are independent $\chi_{k-1}^2$ and $N(0,1)$ random variables. Since $\xi_k(1-\alpha) \to \infty$ as $\alpha \to 0$, from the definition of $\delta_1$,

$$\begin{aligned}\beta &= \Pr\left[Y + \left(Z + \sqrt{\delta_1}\right)^2 \geq \xi_k(1-\alpha)\right] \\ &= \Pr\left[\left(Z + \sqrt{\delta_1}\right)^2 \geq \xi_k(1-\alpha) - Y, \xi_k(1-\alpha) > Y\right] + o(1).\end{aligned}$$

By the Taylor theorem,

$$\sqrt{\xi_k(1-\alpha) - Y} = \sqrt{\xi_k(1-\alpha)} - \frac{Y}{2\sqrt{\xi_k(1-\alpha)}} - \frac{Y^2}{8U^{3/2}}$$

for $-Y + \xi_k(1-\alpha) < U < \xi_k(1-\alpha)$. Now $Y^2/8U^{3/2}$ and $Y/2\sqrt{\xi_k(1-\alpha)}$ are $o_p(1)$ because $\xi_k(1-\alpha) \to \infty$ as $\alpha \to 0$. So we have

$$\begin{aligned}\beta &= \Pr\left[Z + \sqrt{\delta_1} \geq \sqrt{\xi_k(1-\alpha) - Y}, \xi_k(1-\alpha) > Y\right] \\ &\quad + \Pr\left[Z + \sqrt{\delta_1} \leq -\sqrt{\xi_k(1-\alpha) - Y}, \xi_k(1-\alpha) > Y\right] + o(1) \\ &= \Pr\left[Z + \sqrt{\delta_1} \geq \sqrt{\xi_k(1-\alpha)}\right] + \Pr\left[Z + \sqrt{\delta_1} \leq -\sqrt{\xi_k(1-\alpha)}\right] + o(1).\end{aligned}$$

Moreover, $\sqrt{\delta_1} \to \infty$ as $\alpha \to 0$, so that the second probability in the preceding equation converges to 0 as $\alpha \to 0$. Therefore

$$\beta = \Pr\left[Z \geq \sqrt{\xi_k(1-\alpha)} - \sqrt{\delta_1}\right] + o(1),$$

and hence as $\alpha \to 0$,

$$\sqrt{\delta_1} = \sqrt{\xi_k(1-\alpha)} - \Phi^{-1}(1 - \beta + o(1)), \tag{6.1}$$

where $\Phi$ is the distribution function of $N(0,1)$. Further, for sufficiently small $\alpha$,

$$\alpha = \int_{\xi_k(1-\alpha)}^{\infty} \frac{u^{(k-2)/2} e^{-u/2}}{2^{k/2}\Gamma(k/2)}\, du \approx C_k\left(\xi_k(1-\alpha)\right)^{(k-2)/2} e^{-\xi_k(1-\alpha)/2},$$

where $C_k = k^{-(k-3)/2}e^{k/2}/\sqrt{\pi}(k + 1/6)$. So $\log \alpha$ can be approximated by

$$\log \alpha \approx \log C_k - \frac{1}{2}\xi_k(1-\alpha)\left(1 - (k-2)\frac{\log \xi_k(1-\alpha)}{\xi_k(1-\alpha)}\right)$$

$$\approx \log C_k - \frac{1}{2}\xi_k(1-\alpha).$$

We now have an approximation of $\xi_k(1-\alpha)$ for small $\alpha$:

$$\xi_k(1-\alpha) \approx 2\log\frac{1}{\alpha} + 2\log C_k. \tag{6.2}$$

Combining (6.1) and (6.2) yields $\sqrt{\delta(\alpha,\beta,k)} \approx \sqrt{2\log(1/\alpha) + 2\log C_k} - \Phi^{-1}(1-\beta)$. Similarly

$$\sqrt{\delta(\alpha,\beta,k+1)} \approx \sqrt{2\log(1/\alpha) + 2\log C_{k+1}} - \Phi^{-1}(1-\beta).$$

Hence

$$\lim_{\alpha\to 0}\frac{\delta(\alpha,\beta,k+1)}{\delta(\alpha,\beta,k)} = \lim_{\alpha\to 0}\frac{\left(\sqrt{2\log(1/\alpha) + 2\log C_{k+1}} - \Phi^{-1}(1-\beta)\right)^2}{\left(\sqrt{2\log(1/\alpha) + 2\log C_k} - \Phi^{-1}(1-\beta)\right)^2} = 1.$$

If no observations are censored, the statistic $Q$ reduces to the Pearson statistic. The Akritas statistic $Q_A$, however, does not reduce to the Pearson statistic. In the case of no censoring $Q$, that is, the Pearson statistic, is superior to $Q_A$ in the sense of Pitman efficiency. This follows from Theorem 6.2 and the following result.

THEOREM 6.4. *If there is no censoring, then*

$$e^P(Q, Q_A) = \frac{\delta(\alpha,\beta,k+1)}{\delta(\alpha,\beta,k)}.$$

PROOF. When there is no censoring, $\pi_i(\eta) = p_i(\eta)$ and

$$\Psi = \operatorname{diag}\left(p_1^{-1}(\eta_0), \ldots, p_k^{-1}(\eta_0)\right) + \frac{1_k 1_k'}{p_{k+1}(\eta_0)},$$

where $p_{k+1}(\eta) = 1 - \sum_{i=1}^k p_i(\eta)$. Now, if we let $c_i = [\partial p_i(\eta)/\partial\eta]_{\eta=\eta_0}$, since $\sum_{i=1}^k c_i = -c_{k+1}$, $b_1'\Psi b_1 = \sum_{i=1}^{k+1} c_i^2/p_i(\eta_0) = b_2' D_\pi^{-1} b_2$. Hence $e^P(Q, Q_A) = \delta(\alpha,\beta,k+1)/\delta(\alpha,\beta,k)$. □

In the next theorem, we will consider a sequence of censoring distribution functions such that the corresponding sequence of the Pitman efficiencies of $Q$ to $Q_A$ converges to zero.

THEOREM 6.5. *Let $\{G_n\}$ be a sequence of censoring distribution functions which satisfy*

$$\lim_{n\to\infty} \frac{1 - G_n(b)}{1 - G_n(a)} = 0 \quad \text{for all } 0 < a < b < \infty. \tag{6.3}$$

*Assume that*

$$\left| \frac{\int_0^\infty (1 - G_n(x))\, dM(x|\eta_0)}{\int_0^\infty (1 - G_n(x))\, dF(x|\eta_0)} \right| \geq \varepsilon > 0, \tag{6.4}$$

*for some $\varepsilon > 0$ and for all $n$, where $M(x|\eta) = \partial F(x|\eta)/\partial\eta$. Then $\lim_{n\to\infty} e_n^P(Q, Q_A) = 0$, where $\{e_n^P(Q, Q_A)\}$ is the sequence of Pitman efficiencies corresponding to $\{G_n\}$.*

REMARK 1. Equation (6.3) specifies the tail behavior of the sequence of distribution functions $G_n$ and implies the probability of censoring converges to 1 as $n \to \infty$. Sequences of gamma and Weibull distributions satisfy the condition. [See Kim (1988).]

REMARK 2. The left-hand side of (6.4) is the absolute value of the derivative of $I_\eta(1:2)$ with respect to $\eta$ evaluated at $\eta = \eta_0$. Here, $I_\eta(1:2)$ is the Kullback–Leibler information function for the discrimination between the density of an uncensored observation under the null hypothesis $(1 - G_n(x))f(x|\eta_0)/\pi_{1\cdot}$, and the density of $X_i$ under the alternative $f(x|\eta)$. Many sequences of alternatives and censoring distributions satisfying the condition can be found. One such simple model is considered in Example 6.1.

PROOF OF THEOREM 6.5. We will use the same notations $p_i$, $q_i$, $r_i$, $\pi_{1i}$, $b_1$ and $b_2$ except the censoring distribution function $G$ is replaced by $G_n$. Let $b_{1j}$ and $b_{2j}$ be the $j$th elements of the vectors $b_1$ and $b_2$. Let

$$l_i = \frac{\sum_{j=1}^{i} b_{1j}}{q_i} - \frac{\sum_{j=1}^{i-1} b_{1j}}{q_{i-1}}.$$

Then $b_1'\Psi b_1 = \sum_{i=1}^{k} l_i^2/r_i$ and so

$$\frac{b_1'\Psi b_1}{b_2' D_\pi^{-1} b_2} = \sum_{i=1}^{k} \frac{l_i^2}{r_i \sum_{j=1}^{k+1} b_{2j}^2/\pi_{1j}}.$$

Since the $l_i$'s do not depend on $n$, it is enough to show that $r_i \sum_{j=1}^{k+1} b_{2j}^2/\pi_{1j} \to \infty$ as $n \to \infty$ for all $i = 1, \ldots, k$. Now

$$\sum_{j=1}^{k+1} \frac{b_{2j}^2}{\pi_{1j}} = \sum_{j=1}^{k+1} \left(\frac{b_{2j}}{\pi_{1j}}\right)^2 \frac{\pi_{1j}}{\pi_{1\cdot}} \pi_{1\cdot} \geq \pi_{1\cdot} \left(\sum_{j=1}^{k+1} \frac{b_{2j}}{\pi_{1j}} \frac{\pi_{1j}}{\pi_{1\cdot}}\right)^2 = \frac{b_{2\cdot}^2}{\pi_{1\cdot}},$$

where $\pi_{1\cdot} = \int_0^\infty (1 - G_n(x))\, dF(x|\eta_0)$ and $b_{2\cdot} = \int_0^\infty (1 - G_n(x))\, dM(x|\eta_0)$. So $r_i \sum_{j=1}^{k+1} b_{2j}^2/\pi_{1j} \geq r_i b_{2\cdot}^2/\pi_{1\cdot} = r_i |b_{2\cdot}| \cdot |b_{2\cdot}/\pi_{1\cdot}|$. Now $|b_{2\cdot}/\pi_{1\cdot}| \geq \varepsilon > 0$ by (6.4). To show $r_i|b_{2\cdot}| \to \infty$ let $M_1(x|\eta_0) = 1/(1 - F(x|\eta_0))$ and $c$ be a constant such

that $a_{i-1} < c < a_i$. Then

$$r_i|b_{2\cdot}| = \int_{a_{i-1}}^{a_i} \frac{1-G_n(c)}{1-G_n(x)}\, dM_1(x|\eta_0) \cdot \left| \int_0^{\infty} \frac{1-G_n(x)}{1-G_n(c)}\, dM(x|\eta_0) \right|$$

$$= \left( \int_{a_{i-1}}^{c} \frac{1-G_n(c)}{1-G_n(x)}\, dM_1(x|\eta_0) + \int_c^{a_i} \frac{1-G_n(c)}{1-G_n(x)}\, dM_1(x|\eta_0) \right)$$

$$\times \left| \int_0^{c} \frac{1-G_n(x)}{1-G_n(c)}\, dM(x|\eta_0) + \int_c^{\infty} \frac{1-G_n(x)}{1-G_n(c)}\, dM(x|\eta_0) \right|.$$

So, by (6.3), we have

$$\lim_{n\to\infty} \left( \int_{a_{i-1}}^{c} \frac{1-G_n(c)}{1-G_n(x)}\, dM_1(x|\eta_0) + \int_c^{a_i} \frac{1-G_n(c)}{1-G_n(x)}\, dM_1(x|\eta_0) \right) = \infty$$

and

$$\lim_{n\to\infty} \left| \int_0^{c} \frac{1-G_n(x)}{1-G_n(c)}\, dM(x|\eta_0) + \int_c^{\infty} \frac{1-G_n(x)}{1-G_n(c)}\, dM(x|\eta_0) \right| = \infty.$$

Hence $\lim_{n\to\infty} r_i|b_{2\cdot}| = \infty$ for all $i$. This completes the proof. □

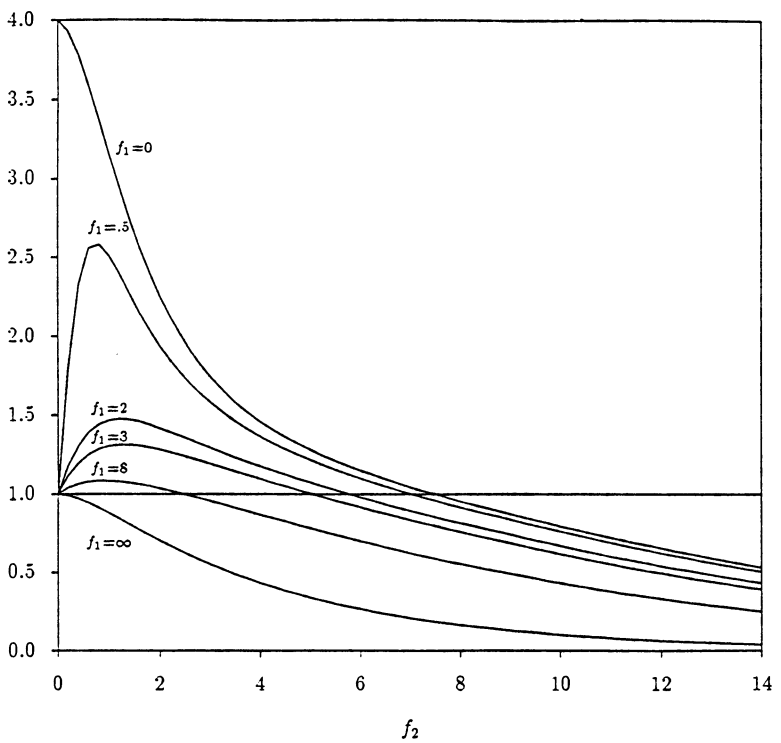


FIG. 1. $R = b_1'\Psi b_1 / b_2' D_\pi^{-1} b_2$ *as a function of* $f_2$ *when* $k = 5$.

From the results presented in this section, we can conclude that neither $Q$ nor $Q_A$ dominates the other. For heavily censored data $Q_A$ is superior to $Q$. In the uncensored case $Q_A$ is not as good as $Q$, the Pearson statistic, in the sense of Pitman efficiency. In the following example, it can be seen that $Q$ performs better than $Q_A$ when we have moderate censoring. We believe this is true in most cases with reasonable number of cells if the ratio $\min_{1 \le i \le k+1}(\pi_{1i})/\max_{1 \le i \le k}(r_i)$ is not too small. However, the proof is technically difficult in general because the efficiency function depends on the null distribution, the alternative distribution, the censoring distribution, the number of cells, and the cell boundaries. It can be done numerically case by case.

In the next example, we will illustrate the theoretical results in Theorems 6.4 and 6.5 by computing the efficiencies of $Q$ to $Q_A$ for testing the exponential distribution versus a contamination alternative.

EXAMPLE 6.1. Consider testing $H_0$: $F = F_0$ versus $H_1$: $(1 - \eta)F_0 + \eta F_1$. Suppose that $F_0(x) = 1 - e^{-\lambda_0 x}$, $F_1(x) = 1 - e^{-\lambda_1 x}$ and $G(x) = 1 - e^{-\lambda_2 x}$. We use $k$ equiprobable cells for $k = 3, 5, 7, 9$. Let $f_1 = \lambda_1/\lambda_0$, $f_2 = \lambda_2/\lambda_0$ and $d_i = (k - i)/k$. Then the ratio of quadratic forms in $e^P(Q, Q_A)$ can be expressed in terms of $k$, $f_1$, $f_2$ and $d_i$'s. Let $R$ be the ratio of quadratic forms.

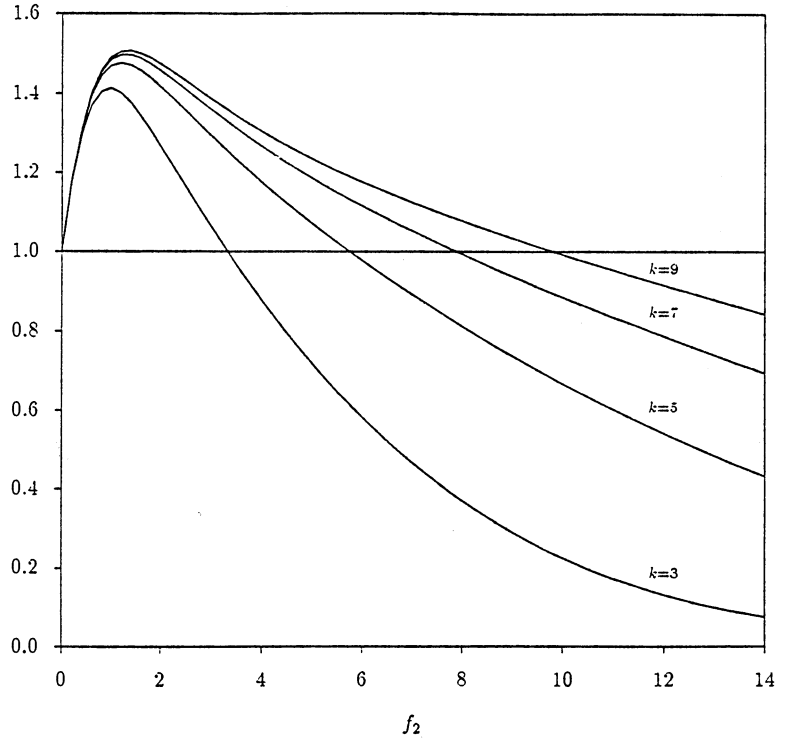


FIG. 2. $R = b_1'\Psi b_1/b_2'D_\pi^{-1}b_2$ *as a function of* $f_2$ *when* $f_1 = 2$.

We present two pictures to see the behavior of $R$ for different values of $f_1$, $f_2$ and $k$.

Figure 1 is a graph of $R$ versus $f_2$ for fixed $k = 5$, $f_1 = 0.5, 2, 3, 8$ and the limits of $R$ as $f_1 \to 0$ and $f_1 \to \infty$. The graph shows that $Q_A$ performs better in the heavy censoring case and $Q$ is superior to $Q_A$ if we have moderate censoring. For example, if we look at the graph $f_1 = 3$, $R$ equals 1 when $f_2$ is about 5. That means the probability of censoring $\Pr[\delta = 0]$ is somewhere between 5/8 and 5/6 depending on the value of $\eta$. For a fixed value of $f_2$, $Q$ gets better as $f_1$ gets smaller, that is, the mean of the contaminating distribution gets larger relative to the mean of the null distribution.

Figure 2 is a graph of $R$ versus $f_2$ for fixed $f_1 = 2$ and different numbers of cells, $k = 3, 5, 7, 9$. Again $Q_A$ performs better in the heavy censoring case. For fixed $f_2$, $Q$ gets better as the number of cells gets larger.

**Acknowledgments.** This paper is part of the author's Ph.D. thesis written at Purdue University. The author would like to express thanks to Professor David S. Moore for his guidance and valuable help during the preparation of this paper. I am also grateful to an Associate Editor and referees for many helpful suggestions in improving the presentation of this paper.

## REFERENCES

AKRITAS, M. G. (1988). Pearson-type goodness-of-fit tests: The univariate case. *J. Amer. Statist. Assoc.* **83** 222–230.

BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. Wiley, New York.

BISHOP, Y. M. M., FIENBERG, S. E. and HOLLAND, P. W. (1975). *Discrete Multivariate Analysis*. The MIT Press.

BORGAN, Ø. (1984). Maximum likelihood estimation in parametric counting process models, with applications to censored failure time data. *Scand. J. Statist.* **11** 1–16.

BRESLOW, N. and CROWLEY, J. (1974). A large sample study of the life table and product limit estimates under random censorship *Ann. Statist.* **2** 437–453.

CHEN, J. (1975). Goodness of fit tests under random censorship. Ph.D. dissertation, Dept. Statistics, Oregon State Univ.

CHERNOFF, H. and LEHMANN, E. L. (1954). The use of maximum likelihood estimates in $\chi^2$ tests for goodness-of-fit. *Ann. Math. Statist.* **25** 579–586.

CHIBISOV, D. M. (1971). Certain chi-square type tests for continuous distributions. *Theory Probab. Appl.* **16** 1–22.

D'AGOSTINO, R. B. and STEPHENS, M. A. (1986). *Goodness-of-Fit Techniques*. Dekker, New York.

DAVIDSON, R. R. and LEVER, W. E. (1970). The limiting distribution of the likelihood ratio statistic under a class of local alternatives. *Sankhyā Ser. A* **32** 209–224.

DURBIN, J. (1973). Weak convergence of the sample distribution function when parameters are estimated. *Ann. Statist.* **1** 279–290.

GILL, R. D. (1981). *Censoring and Stochastic Integrals. Math. Centre Tract* **124**. Math. Centrum, Amsterdam.

GILL, R. D. (1983). Large sample behavior of the product-limit estimator on the whole line. *Ann. Statist.* **11** 49–58.

HABIB, M. G. and THOMAS, D. R. (1986). Chi-square goodness-of-fit tests for randomly censored data. *Ann. Statist.* **14** 759–765.

HJORT, N. L. (1984). Weak convergence of cumulative intensity processes when parameters are estimated, with applications to goodness-of-fit tests in models with censoring. Research Report 763, Norwegian Computing Centre, Oslo.

HJORT, N. L. (1990). Goodness of fit tests in models for life history data based on cumulative hazard rates. *Ann. Statist.* **18** 1221–1258.

JOHNSON, N. L. and KOTZ, S. (1970). *Distributions in Statistics: Continuous Univariate Distributions 2*. Wiley, New York.

KAPLAN, E. L. and MEIER, P. (1958). Nonparametric estimation from incomplete observations. *J. Amer. Statist. Assoc.* **53** 457–481.

KIM, J. H. (1988). Chi-square tests for randomly censored data. Ph.D. dissertation, Dept. Statistics, Purdue Univ.

KULLBACK, S. (1968). *Information Theory and Statistics*. Dover, New York.

MIHALKO, D. P. and MOORE, D. S. (1980). Chi-square tests of fit for Type II censored data. *Ann. Statist.* **8** 625–644.

MILLER, R. G., JR. (1981). *Survival Analysis*. Wiley, New York.

MOORE, D. S. (1971). A chi-square statistic with random cell boundaries. *Ann. Math. Statist.* **42** 147–156.

MOORE, D. S. (1977). Generalized inverses, Wald's method, and the construction of chi-square tests of fit. *J. Amer. Statist. Assoc.* **72** 131–137.

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.

RAO, K. C. and ROBSON, D. S. (1975). A chi-square statistic for goodness-of-fit tests within the exponential family. *Comm. Statist.* **3** 1139–1153.

ROTHE, G. (1981). Some properties of the asymptotic relative Pitman efficiency. *Ann. Statist.* **9** 663–669.

SAUNDERS, I. W. and MORAN, P. A. P. (1978). On the quantiles of the gamma and $F$ distributions. *J. Appl. Probab.* **15** 426–432.

SERFLING, R. J. (1980). *Approximation Theorems of Mathematical Statistics*. Wiley, New York.

TURNBULL, B. W. and WEISS, L. (1978). A likelihood ratio statistic for testing goodness-of-fit with randomly censored data. *Biometrics* **34** 367–375.

DEPARTMENT OF STATISTICS
CHUNGNAM NATIONAL UNIVERSITY
220 GUNGDONG, YUSONG-KU
TAEJEON 305-764
SOUTH KOREA