The Annals of Statistics
1988, Vol. 16, No. 4, 1490–1502

RANK TESTS WITH INTERVAL-CENSORED DATA

BY MICHAEL G. AKRITAS

Pennsylvania State University

An interval censoring model is carefully defined and the range of its applicability is illustrated. A class of rank tests for the two- and k-sample problems is proposed. The statistic is based on the exact ranks of the observed responses and resembles the rank statistic in the usual uncensored case. This statistic is shown to be asymptotically equivalent to the efficient scores test statistic under an assumed parametric model. An alternative rank statistic, based on the estimated ranks of the unobserved variables of interest and applicable under a more general interval censoring model, is also proposed.

1. Introduction. Unlike the random censoring model, relatively little attention is paid to the statistical analysis of interval-censored data. Such data arise quite naturally in medical follow-up studies or in industrial life-testing and thus more work on testing procedures with such data is warranted.

Peto (1973), Turnbull (1976) and Chang and Yang (1987) deal with the problem of estimating the underlying survival distribution; Mantel (1967) gives a very simple rank test for arbitrarily censored data; Schemper (1983) extended Mantel's test to k samples; Schemper (1984) extended Friedman's test to data defined by intervals; Kariya (1981) considers properties of the maximum likelihood estimator and Abel (1986) considers testing against order alternatives for such data. It should be mentioned that the above rank tests do not possess any special optimality properties.

In this paper we consider a more restricted model than is usually assumed and deal with the problem of constructing efficient rank tests for the two-sample problem. The restricted model essentially requires equally spaced inspection intervals. It is remarked that this restriction is necessary in order to show the full Pitman efficiency of the proposed class of rank tests. A different class of rank tests that does not require this restriction is also proposed in Remark 2.5. The discussion in Section 2.3 shows that even the restricted model has a wide range of applicability.

The method for constructing an efficient rank test statistic is based on an appropriate choice of the score function so that the rank statistic, when expressed as a functional, resembles the efficient scores test statistic. The asymptotic efficiency of the rank test is then established by proving the asymptotic equivalence of the two statistics. The proof makes essential use of certain linear bounds in the sense of Shorack (1972).

Received May 1986; revised February 1988.
AMS 1980 *subject classifications*. 62G10, 62N05.
Key words and phrases. Efficient scores statistic, $C(\alpha)$ tests, adaptive inference, empirical processes.

In Section 2 we give a careful description of the interval censorship model and we propose a class of rank statistics; the candidate optimal score function for a given parametric model is derived. The statistic employs the exact ranks of the observed responses. An alternative rank statistic which uses the estimated ranks of the unobserved variables of interest is also proposed in Remark 2.5 but proving its efficiency would require a number of theoretical results which are presently unavailable. In Section 3 we state the assumptions, two preliminary results and present a number of examples. The two-sample case is considered in detail in Section 4.

2. The model and the rank statistic.

2.1. *The model.* Suppose that in some medical or industrial set up, inspection occurs at times $k\Delta$, $k \geq 1$, where Δ is some positive constant, and that the time a subject enters the study is recorded exactly. Assume without loss of generality that $\Delta = 1$. The data will consist of independent and identically distributed (iid) pairs of random variables (A_i, X_i), $i = 1, \ldots, n$, where A denotes the time mod (1) of entry to study (so $0 \leq A \leq 1$ a.s.) and X is the time from entry until the response is observed. The true response time will be denoted by Y; Figure 1 helps clarify the relationship between the variables A, X and Y (see also Remark 2.2).

REMARK 2.1. In the interval censoring model considered by other authors [cf. Kariya (1981)] the inspection intervals can be different for each subject and do not have to be of the same length. However, the tests we will consider use the ranks of the observable X's and in order to achieve asymptotic efficiency for such tests we found it necessary to consider the more special model. If the inspection intervals are not the same, efficient tests based on the ranks of X cannot be constructed; Remark 2.5 discusses the possibility of constructing different rank tests in this case.

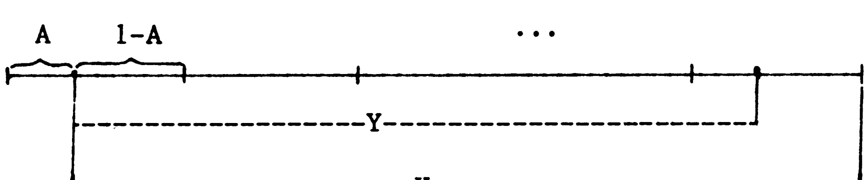

FIG. 1.

Assume that

$$P_\eta(A \leq a, Y \leq y) = G(a)F(y; \eta), \tag{2.1}$$

where η is some unknown Euclidean parameter, $F(\cdot\,; \eta)$ is a distribution function of known form and $G(\cdot)$ is, in general, an unknown distribution function on $[0,1]$. Noting that $X = [A + Y] + 1 - A$, where $[\cdot]$ denotes integer part, relation (2.1), and some straightforward calculations imply

$$P_\eta(X \le x) \equiv H_\eta(x) = \int_0^1 F([x + a - 1] + 1 - a; \eta)\, dG(a). \tag{2.2}$$

Differentiating this with respect to x [note that the integrand on the right-hand side of (2.2) is not differentiable with respect to x], we see that H_η has density

$$h_\eta(x) = \{F(x; \eta) - F(x - 1; \eta)\} g([x] + 1 - x), \tag{2.3}$$

with respect to the Lebesgue measure, where g is the density of G. From (2.3) we have that the efficient scores statistic for testing H_0: $\eta = \eta_0$ based on X_i, $i = 1, \dots, n$, is

$$\sum_{i=1}^{n} \frac{\partial/\partial\eta\{F(X_i; \eta) - F(X_i - 1; \eta)\}}{F(X_i; \eta) - F(X_i - 1; \eta)}\Bigg|_{\eta=\eta_0}. \tag{2.4}$$

REMARK 2.2. The random variable A and the assumption (2.1) were introduced merely as a means for deriving the density expression in (2.3). In fact we have $1 - A = X - [X]$ so that observing the pair (A, X) is equivalent to observing just X.

2.2. *The rank statistic for the two-sample problem.* In later sections we will be dealing with the two-sample problem and in this setting it is customary to consider classes of probability models $F(\cdot; \eta)$ for which the parameter η bears some easily understood impact on the distribution. In this paper we will consider testing equality of a scale parameter. Take $\eta = (\theta, \sigma)$ and assume that, for the two-sample case $Y_{11}, \dots, Y_{1n_1}$ are iid $F_1(x; \theta, \sigma) = F(x\sigma \exp(-q\theta))$ and $Y_{21}, \dots, Y_{2n_2}$ are iid $F_2(x; \theta, \sigma) = F(x\sigma \exp(p\theta))$, where $-\infty < \theta < \infty$, $\sigma > 0$ and $q = n_2/N$, $p = n_1/N$, where $N = n_1 + n_2$. Thus $\sigma e^{-q\theta}$ and $\sigma e^{p\theta}$ are the true values of the scale parameter whose equality ($\theta = 0$) is being tested. Then, letting f denote the density of F, relation (2.4) implies that the efficient score test statistic for testing H_0: $\theta = 0$ is proportional to

$$\begin{aligned} &-q\sum_{i=1}^{n_1} \frac{X_{1i} f(\hat\sigma X_{1i}) - (X_{1i} - 1) f(\hat\sigma X_{1i} - \hat\sigma)}{F(\hat\sigma X_{1i}) - F(\hat\sigma X_{1i} - \hat\sigma)} \\ &+p\sum_{i=1}^{n_2} \frac{X_{2i} f(\hat\sigma X_{2i}) - (X_{2i} - 1) f(\hat\sigma X_{2i} - \hat\sigma)}{F(\hat\sigma X_{2i}) - F(\hat\sigma X_{2i} - \hat\sigma)} \\ &\equiv q n_1 \int \Phi_{\hat\sigma}(x)\, d\hat H_1(x) - p n_2 \int \Phi_{\hat\sigma}(x)\, d\hat H_2(x), \end{aligned} \tag{2.5}$$

where $\hat\sigma$ is some root n consistent estimator of σ obtained from both samples under the null hypothesis, $\hat H_t$ denotes the empirical distribution function (edf) corresponding to $X_{t1}, \dots, X_{tn_t}$, $t = 1, 2$; note that (2.5) also defines Φ_σ.

REMARK 2.3. Root n consistency of nuisance parameters is required in the theory of $C(\alpha)$ tests [Neyman (1959)]; according to Section 2.4, the statistic in (2.5) is a $C(\alpha)$ test statistic and thus $\hat{\sigma}$ is required to be root n consistent. For the rest of the paper we assume that such a $\hat{\sigma}$ is available.

For $\eta = (0, \sigma)$ let $H_\eta(\cdot)$ as defined by (2.2) be denoted by $H_\sigma(\cdot)$ and define the score function

$$\hat{J}(u) \equiv J_{\hat{\sigma}}(u) = \Phi_{\hat{\sigma}}\big(H_{\hat{\sigma}}^{-1}(u)\big), \qquad 0 < u < 1. \tag{2.6}$$

Intuitively then the rank statistic, expressed in a functional form as

$$qn_1 \int \hat{J}\left(\frac{N\hat{H}(x)}{N+1}\right) d\hat{H}_1(x) - pn_2 \int \hat{J}\left(\frac{N\hat{H}(x)}{N+1}\right) d\hat{H}_2(x), \tag{2.7}$$

where $\hat{H}$ is the edf corresponding to X_{ti}, $i = 1, \dots, n_t$, $t = 1, 2$, will be asymptotically equivalent to the efficient scores statistic in (2.5); this will be established in Section 4. Note, however, that the distribution G [which is involved in the definition of the score function in (2.6)] is unknown. The proofs in Section 4 hold for any distribution function G that satisfies assumptions (E2) and (E3) of Section 3, but for the rest of the paper we will assume that $G(a) = a$, $0 \le a \le 1$. In Section 2.3 below we include a discussion on this assumption.

REMARK 2.4. Unlike the score function in the usual iid case, the present score function in (2.6) is random since it depends on $\hat{\sigma}$. Conditionally, however, on the estimate $\hat{\sigma}$ (recall that $\hat{\sigma}$ is obtained under the null hypothesis) the usual finite sample properties of rank statistics hold. In particular, for small sample sizes, the cut off points of the test may be obtained from the permutation distribution of the statistic.

REMARK 2.5. The statistic in (2.7) employs the (exact) ranks of X_{ti}, $i = 1, \dots, n_t$, $t = 1, 2$. One can also construct a statistic which employs the (estimated) ranks of the actual response times, namely,

$$qn_1 \int \phi\left(\frac{N\hat{F}(x)}{N+1}\right) d\hat{H}_1(x) - pn_2 \int \phi\left(\frac{N\hat{F}(x)}{N+1}\right) d\hat{H}_2(x), \tag{2.8}$$

where $\phi(u) = \Phi_{\hat{\sigma}}(F_{\hat{\sigma}}^{-1}(u))$ with $F_{\hat{\sigma}}(x) = F(\hat{\sigma}x)$, and $\hat{F}$ is (under the null hypothesis) some nonparametric estimator of $F(\sigma x)$ computed from both samples. The advantage of this statistic over the one in (2.7) is that the score function ϕ does not involve G and thus we can avoid the design considerations discussed in Section 2.3. However, in order to show efficiency of the test statistic in (2.8) with the present methodology one needs certain theoretical results about $\hat{F}$ (such as linear bounds) which, at the moment, are not available for any nonparametric estimator of $F(\sigma x)$. Tsai and Crowley (1985), Gill (1988) and Chang (1987) present a general methodology for obtaining large sample results for generalized maximum likelihood estimators but further work is needed [the recent preprint by Groeneboom (1987) shows that asymptotics for nonparametric estimation with interval censored data are definitely harder than with ordinary uncensored data]. If such results were available for some estimator it would also be possible to show that asymptotic equivalence of a rank statistic similar to that in (2.8) with the statistic that is used in the more general model described in Remark 2.1.

2.3. *Discussion of the model and its assumptions.* The requirement that the inspection intervals be of equal length is a design restriction which is easier to implement in industrial reliability studies than in medical follow-up studies; in particular the larger the medical follow-up study the harder it will be to adhere to this design. However, for smaller (pilot) medical studies the restriction of equal length inspection intervals is not unreasonable.

Similarly, the assumption that the random variables A are uniform in $(0, 1)$ is easier to implement (as a design requirement) in an industrial study than in a medical one. For instance consider a comparative study of the life length of two different brands of tubes in a large company that inspects equipment at intervals of length 1. Let n_1, n_2 denote the number of tubes of brands 1, 2, respectively, that are used in the study. One can then generate $N = n_1 + n_2$ uniform in $(0, 1)$ random numbers and use these as the times mod (1) of entry to the study (see Figure 1). (Clearly, in the setup just described, one can easily implement not only uniformity but any other distributional requirement on the random variables A; some possible reasons for using distributions other than the uniform are given in Section 5).

Next we will present a result which shows that for a large number of retrospective studies the assumption that the A's constitute a sample from the uniform in $(0, 1)$ distribution is satisfied. In particular suppose that data are kept on subjects as they become available and assume that the interarrival times $Z_1, \dots, Z_n$ are iid Q. Then $A_n = (\sum_{i=1}^{n} Z_i) \bmod (1)$, $n \geq 1$, forms a Markov process and it is easy to show that if Q is absolutely continuous the n-step transition probabilities converge to the unique stationary initial distribution which is the uniform. [Related results are given by Diaconis and Engel (1986) in their discussion of Good (1986).] Thus, the random sample which will be taken from the collected data for the retrospective study will satisfy the assumption that the A's are iid uniform in $(0, 1)$.

In cases where distributional requirements on the random variable A cannot be implemented in the design and the above probabilistic argument does not apply, the choice of G is not obvious. More elaboration on this problem is provided in Section 5.

2.4. *A comment on the parametric test.* The efficient scores statistic of relationship (2.5) involves an estimator of the nuisance parameter σ and thus its optimality needs to be justified. In particular we need to determine whether, in testing H_0: $\theta = 0$, it is possible to "adapt" for σ (meaning whether we can do as well asymptotically as if the true value of σ were known). In cases where nuisance parameters are present Neyman (1959) introduced a statistic constructed by projecting the efficient score and showed that the test based on such a statistic is asymptotically optimal (even when the testing problem cannot be adapted). For a modern approach to adaptive inference see Bickel (1982). A necessary condition for adaptation is that the scores corresponding to the two parameters are orthogonal (uncorrelated) and this can easily be shown to be the case here. It should be mentioned that even though Neyman's paper helps justify the statistic in (2.5) the conditions under which his results were shown will not concern us.

3. Assumptions and preliminary results. For $t = 1, 2$, let $(A_{t1}, Y_{t1}), \dots, (A_{tn_t}, Y_{tn_t})$ be iid $G(\cdot)F_t(\cdot; \theta, \sigma)$, where $G(\cdot)$ is a distribution on $[0, 1]$, $F_t(\cdot; \theta, \sigma)$ a distribution on $[0, \infty)$, and assume we observe (A_{ti}, X_{ti}), $i = 1, \dots, n_t$, $t = 1, 2$, where $X_{ti} = [A_{ti} + Y_{ti}] + 1 - A_{ti}$.

ASSUMPTIONS E.

(E1) (i) $F_1(x; \theta, \sigma) = F(xe^{-q\theta}\sigma)$, $F_2(x; \theta, \sigma) = F(xe^{p\theta}\sigma)$, where F is a specified distribution function, $p = n_1/N$, $q = 1 - p$, $N = n_1 + n_2$; (ii) p remains bounded away from 0 and 1 as $N \to \infty$; (iii) F has a differentiable density f; and (iv) the support of F is an interval in $[0, \infty)$.

(E2) Let $J_\sigma(u) = \Phi_\sigma(H_\sigma^{-1}(u))$, $0 < u < 1$, where H_σ is defined in connection with (2.6). Then the derivative $J_\sigma'(\cdot)$ exists on $(0, 1)$, is continuous, and we have that for σ in any compact set C,

$$|J_\sigma(u)| \le K\{u(1-u)\}^{-0.5+\varepsilon}, \qquad |J_\sigma'(u)| \le K\{u(1-u)\}^{-1.5+\varepsilon},$$

for all $u \in (0, 1)$ and some K, ε positive constants possibly depending on C.

(E3) For any σ_0 and $\varepsilon^* > 0$ there exists a neighborhood of σ_0 such that, for any σ_1 in that neighborhood the ratios

$$\left[H_{\sigma_1}(x)\right]^{1+\varepsilon^*}/H_{\sigma_0}(x) \quad \text{and} \quad \left[1 - H_{\sigma_1}(x)\right]^{1+\varepsilon^*}\Big/\left[1 - H_{\sigma_0}(x)\right]$$

remain bounded uniformly in x.

REMARK 3.1. Assumption (E1)(iv) was included in order to avoid difficulties in the definition of Φ_σ. Indeed if the support of F is split in two sets by an interval of length greater than 1, Φ_σ is not well defined.

LEMMA 3.1. *Let Y_n, $n \ge 1$, be a sequence of random variables and assume that for each $\varepsilon > 0$ we can find a sequence of measurable sets $B_{n,\varepsilon}$, $n \ge 1$, such that $P(B_{n,\varepsilon}) \ge 1 - \varepsilon$ for all $n \ge N(\varepsilon)$. Then if $I_{B_{n,\varepsilon}} Y_n \to 0$ in probability for each $\varepsilon > 0$, where I_B denotes the indicator function of the set B, it follows that $Y_n \to 0$ in probability.*

PROOF. The proof is straightforward; it also follows from Theorem 4.2 of Billingsley (1968). □

LEMMA 3.2. *Let μ_n, $n \geq 1$, and μ be functions on the real line such that*

1. $\mu_n(x) \to \mu(x)$ *for all* $x \in (-\infty, \infty)$ *and*
2. μ_n, $n \geq 1$, *and* μ *are uniformly of bounded variation.*

It then follows that $\int g(x)\, d\mu_n(x) \to \int g(x)\, d\mu(x)$ *for all* g *continuous and bounded.*

PROOF. The proof is similar to the proof of the Helly–Bray lemma (cf. Loève 1963, page 180). □

We close this section by presenting the distribution function H_σ and the score function J_σ in a number of scale models for $G(a) = a$, $0 \leq a \leq 1$; for all these examples assumptions (E2) and (E3) have been verified.

EXAMPLE 1 (Exponential distribution). Here $F(\sigma x) = 1 - e^{-\sigma x}$, $\sigma > 0$, $x > 0$ and

$$H_\sigma(x) = \{x - \sigma^{-1}(1 - e^{-\sigma x})\} I(0 < x < 1) + \{1 - \sigma^{-1}(e^\sigma - 1)e^{-\sigma x}\} I(x > 1),$$

$$J_\sigma(H_\sigma(x)) = -x + e^\sigma(e^\sigma - 1)^{-1}.$$

REMARK 3.2. In the usual uncensored case the family of Lehmann alternatives $F(\sigma x) = 1 - (1 - F_0(x))^\sigma$ leads to the same score function for any continuous distribution function F_0. This is not true in the present interval censoring model and thus a rank test using the score function of Example 1 will be optimal only if the exponential distribution holds.

EXAMPLE 2 (Half-logistic distribution). Let $F(\sigma x) = 2/(1 + e^{-\sigma x}) - 1$, $\sigma > 0$, $x > 0$. Here

$$H_\sigma(x) = 2\sigma^{-1}\log\frac{1 + e^{-\sigma x}}{2} I(0 < x < 1) + \left\{1 + 2\sigma^{-1}\log\frac{1 + e^{-\sigma x}}{1 + e^{-\sigma(x-1)}}\right\} I(x > 1),$$

$$J_\sigma(H_\sigma(x)) = \frac{xf(\sigma x) - (x-1)f(\sigma x - \sigma)}{F(\sigma x) - F(\sigma x - \sigma)}.$$

EXAMPLE 3 [One-parameter lognormal ($\mu = 0$) distribution]. Let $F(\sigma x) = \Phi(\log \sigma x)$, $\sigma > 0$, $x > 0$, where Φ is the $N(0,1)$ distribution function. Here

$$H_\sigma(x) = \int_0^x \Phi(\log \sigma y)\, dy\, I(0 < x < 1) + \int_{x-1}^x \Phi(\log \sigma y)\, dy\, I(x > 1),$$

$$J_\sigma(H_\sigma(x)) = \sigma^{-1}\frac{\phi(\log \sigma x) - \phi(\log(\sigma x - \sigma))}{\Phi(\log \sigma x) - \Phi(\log(\sigma x - \sigma))},$$

where ϕ is the density of Φ.

EXAMPLE 4 [One-parameter Weibull ($p = 2$) or Rayleigh distribution]. Let $F(\sigma x) = 1 - \exp(-(\sigma x)^2)$, $\sigma > 0$, $x > 0$. Then

$$H_\sigma(x) = \left\{x - \frac{\sqrt{\pi}}{\sigma}\left(\Phi\left(\frac{x}{\sqrt{2\sigma^2}}\right) - 0.5\right)\right\} I(0 < x < 1)$$

$$+ \left\{1 - \frac{\sqrt{\pi}}{\sigma}\left(\Phi\left(\frac{x}{\sqrt{2\sigma^2}}\right) - \Phi\left(\frac{x-1}{\sqrt{2\sigma^2}}\right)\right)\right\} I(x > 1),$$

$$J_\sigma(H_\sigma(x)) = \frac{xf(\sigma x) - (x-1)f(\sigma x - \sigma)}{F(\sigma x) - F(\sigma x - \sigma)}.$$

Clearly the function $J_\sigma(u)$ will have a closed form expression if H_σ^{-1} has a closed form expression. Of the above examples only the half-logistic distribution results in a closed form expression for the score function. The implementation of the score functions for the exponential, lognormal and Weibull would require the use of a computer; the score function for the exponential distribution is the simplest of the three.

REMARK 3.3. Implementation of any one of the above score functions requires a $\sqrt{n}$-consistent estimate of σ under the null hypothesis. However, under this hypothesis the combined sample comes from $F(\sigma x)$ and one may apply, for instance, the maximum likelihood method. Estimating σ this way means the test is not generally a *rank* test. One can still get exactly size α by using the H_0 permutation distribution (under which $\hat{\sigma}$ stays fixed).

4. The main results. In this section we will show that the test based on the rank statistic of relationship (2.7) has efficiency 1 with respect to the efficient scores test of relationship (2.5). Let now

$$L_N = p^{1/2} q n_1^{1/2} \int \hat{\Phi}(x)\, d(\hat{H}_1(x) - H(x))$$

$$-pq^{1/2} n_2^{1/2} \int \hat{\Phi}(x)\, d(\hat{H}_2(x) - H(x)),$$

(4.1)

$$S_N = p^{1/2} q n_1^{1/2} \int \hat{J}\left(\frac{N\hat{H}(x)}{N+1}\right) d(\hat{H}_1(x) - H(x))$$

$$-pq^{1/2} n_2^{1/2} \int \hat{J}\left(\frac{N\hat{H}(x)}{N+1}\right) d(\hat{H}_2(x) - H(x)),$$

where $\hat{\Phi}(\cdot) = \Phi_{\hat{\sigma}}(\cdot)$ and $H(\cdot) = H_{\sigma_0}(\cdot)$ with σ_0 the true underlying value of σ, be appropriately centered and scaled versions of the statistics in relations (2.5) and (2.7), respectively [L_N equals $N^{-1/2}$ times the statistic in (2.5), and similarly for S_N].

THEOREM 4.1. *Let Assumptions* (E) *hold. Then, under* $\theta = 0$,

$$S_N - L_N \to 0$$

in probability, as $N \to \infty$.

PROOF. Consider an $\varepsilon_1 > 0$. According to Lemma 3.1 we will show that there exists a sequence of sets B_{N,ε_1} so that $P(B_{N,\varepsilon_1}) > 1 - \varepsilon_1$ and

$$I_{B_{N,\varepsilon_1}}(S_N - L_N) \to 0$$

in probability. Let $B^*_{n,\varepsilon_1/2}$ be the set on which the linear bounds of Shorack (1972) for $\hat{H}$, $\hat{H}_t$ and $1 - \hat{H}, 1 - \hat{H}_t$, $t = 1, 2$, hold with probability greater than $1 - \varepsilon_1/2$ and let ε be as in Assumption (E2) corresponding to some neighborhood C_1 of σ_0 and set $q(u) = [u(1-u)]^{0.5-\delta}$, $0 < u < 1$, for some $\delta < \varepsilon$. Finally, let C_2 be a neighborhood of σ_0 corresponding to ε^*, $\delta < \varepsilon^* < \varepsilon$, according to Assumption (E3) and set $B_{N,\varepsilon_1} = B^*_{N,\varepsilon_1/2} \cap [\hat{\sigma} \in C]$, where $C = C_1 \cap C_2$. Clearly $P(B_{N,\varepsilon_1}) > 1 - \varepsilon_1$, for all $N \geq N(\varepsilon_1)$. In this notation we will show

$$(4.2)\quad I_{B_{N,\varepsilon_1}}\left\{n_1^{1/2}\int \hat{\Phi}(x)\, d(\hat{H}_1 - H) - n_1^{1/2}\int \hat{J}\left(\frac{N\hat{H}(x)}{N+1}\right) d(\hat{H}_1 - H)\right\} \to_P 0.$$

That the difference of the other two terms also converges in probability to 0 can be shown similarly. The relationship in (4.2) equals

$$I_{B_{N,\varepsilon_1}} n_1^{1/2}\int\left[\hat{\Phi}(x) - \hat{J}\left(\frac{N\hat{H}(x)}{N+1}\right)\right] d\left[q(H)\frac{\hat{H}_1 - H}{q(H)}\right]$$

$$= I_{B_{N,\varepsilon_1}} n_1^{1/2}\int\left[\hat{\Phi}(x) - \hat{J}\left(\frac{N\hat{H}(x)}{N+1}\right)\right]\left[q(H)\, d\frac{\hat{H}_1 - H}{q(H)} + \frac{\hat{H}_1 - H}{q(H)}\, dq(H)\right],$$

so that (4.2) will follow from

$$(4.3)\qquad I_{B_{N,\varepsilon_1}} n_1^{1/2}\int\left[\hat{\Phi}(x) - \hat{J}\left(\frac{N\hat{H}(x)}{N+1}\right)\right]\frac{\hat{H}_1(x) - H(x)}{q(H(x))}\, dq(H) \to_P 0$$

and

$$(4.4)\qquad I_{B_{N,\varepsilon_1}} n_1^{1/2}\int\left[\hat{\Phi}(x) - \hat{J}\left(\frac{N\hat{H}(x)}{N+1}\right)\right] q(H)\, d\frac{\hat{H}_1 - H}{q(H)} \to_P 0.$$

Since $\sup\{|n_1^{1/2}[\hat{H}_1(x) - H(x)]/q(H(x))|;\ 0 < x < \infty\}$ remains bounded in probability, (4.3) follows from

$$I_{B_{N,\varepsilon_1}}\int\left|\hat{\Phi}(x) - \hat{J}\left(\frac{N\hat{H}(x)}{N+1}\right)\right| |dq(H)| \to_P 0,$$

which is true by the dominated convergence theorem which applies almost surely; indeed

$$\left|\hat{\Phi}(x) - \hat{J}\left(\frac{N\hat{H}(x)}{N+1}\right)\right| \to 0$$

for all x almost surely, and, by Assumptions (E2), (E3) and the linear bounds of Shorack (1972) we have

$$\begin{aligned} I_{B_{N,\varepsilon_1}}\left|\hat{\Phi}(x) - \hat{J}\left(\frac{N\hat{H}(x)}{N+1}\right)\right| &\le \left[|\hat{\Phi}(x)| + \left|\hat{J}\left(\frac{N\hat{H}(x)}{N+1}\right)\right|\right] I_{B_{N,\varepsilon_1}} \\ &\le \Big[K\left[H_{\hat{\sigma}}(x)(1 - H_{\hat{\sigma}}(x))\right]^{-0.5+\varepsilon} \\ &\qquad + K\left[\hat{H}(x)(1 - \hat{H}(x))\right]^{-0.5+\varepsilon}\Big] I_{B_{N,\varepsilon_1}} \\ &\le K\left[H(x)(1 - H(x))\right]^{-0.5+\varepsilon^*} \end{aligned}$$

and

$$\int [H(1 - H)]^{-0.5+\varepsilon^*} |dq(H)| < \infty.$$

Next we show (4.4). Integrate by parts to get

$$\begin{aligned} &\int \left[\hat{\Phi} - \hat{J}\left(\frac{N\hat{H}}{N+1}\right)\right] q(H)\, d\frac{n_1^{1/2}(\hat{H}_1 - H)}{q(H)} \\ &= \left[\hat{\Phi} - \hat{J}\left(\frac{N\hat{H}}{N+1}\right)\right] n_1^{1/2}(\hat{H}_1 - H)\Big|_0^\infty \\ &\quad - \int \frac{n_1^{1/2}(\hat{H}_{1-} - H)}{q(H)}\, d\left\{q(H)\left[\hat{\Phi} - \hat{J}\left(\frac{N\hat{H}}{N+1}\right)\right]\right\}. \end{aligned}$$

It is easy to show that the first term on the right-hand side above converges to 0 in probability; applying a Skorohod construction with a switch to a new probability space the second term is

$$\begin{aligned} &\int \frac{B^0(H(x))}{q(H(x))}\, d\left[\hat{\Phi} q(H) - \hat{J}\left(\frac{N\hat{H}}{N+1}\right) q(H)\right] \\ &+ \int \frac{B^0_{n_1-}(x) - B^0(H(x))}{q(H(x))}\, d\Phi\, q(H) \qquad (4.5) \\ &- \int \frac{B^0_{n_1-}(x) - B^0(H(x))}{q(H(x))}\, d\hat{J}\left(\frac{N\hat{H}}{N+1}\right) q(H), \end{aligned}$$

where we set $B^0_{n_1}(x) = n_1^{1/2}(\hat{H}_1(x) - H(x))$ and B^0 denotes the corresponding Brownian bridge on $[0, 1]$. Each of the summands in (4.5) (multiplied by $I_{B_{N,\varepsilon_1}}$)

will be shown to converge in probability to 0 provided the functions

$$I_{B_{N,\varepsilon_1}} \hat{J}\left(\frac{N\hat{H}}{N+1}\right) q(H), \qquad N \geq 1,$$

and $\hat{\Phi} q(H)$ are uniformly of bounded variation on $[0, \infty)$ (use Lemma 3.2 for the first term and straightforward arguments for the other two terms). The convergence in probability to 0 will also hold in the original probability space and thus the proof of the theorem will be complete. Let $a \equiv a_N = \sup\{x\colon N\hat{H}(x)/(N+1) \leq 1/2\}$ and, on B_{N,ε_1}, consider the variation of the function $\hat{J}(N\hat{H}/(N+1))q(H)$ in the intervals $[0, a), [a, \infty)$. For $[0, a)$ we have

$$\begin{aligned}
&\sum \left| \hat{J}\left(\frac{N\hat{H}(x_1)}{N+1}\right) q(H(x_i)) - \hat{J}\left(\frac{N\hat{H}(x_{i-1})}{N+1}\right) q(H(x_{i-1})) \right| \\
&\quad = \sum \left| \hat{J}\left(\frac{N\hat{H}(x_i)}{N+1}\right) q(H(x_i)) \right. \\
&\qquad \left. - \left[\hat{J}\left(\frac{N\hat{H}(x_i)}{N+1}\right) + \frac{N}{N+1}\left(\hat{H}(x_{i-1}) - \hat{H}(x_i)\right)\hat{J}'(z) \right] q(H(x_{i-1})) \right| \\
&\qquad \left[\text{where } z \text{ lies between } N/(N+1)\hat{H}(x_{i-1}) \text{ and } N/(N+1)\hat{H}(x_i)\right] \qquad (4.6)\\
&\quad \leq \sum \left| \hat{J}\left(\frac{N\hat{H}(x_i)}{N+1}\right) \left[q(H(x_i)) - q(H(x_{i-1})) \right] \right| \\
&\qquad + \frac{N}{N+1} \sum \left| \left(\hat{H}(x_i) - \hat{H}(x_{i-1})\right) \hat{J}'(z) q(H(x_{i-1})) \right| \\
&\quad \leq K \sum \left[\frac{N\hat{H}(x_i)}{N+1}\left(1 - \frac{N\hat{H}(x_i)}{N+1}\right)\right]^{-0.5+\varepsilon} \left| q(H(x_i)) - q(H(x_{i-1})) \right| \\
&\qquad + K \sum \left| \left(\hat{H}(x_i) - \hat{H}(x_{i-1})\right) \right. \\
&\qquad \left. \times \left[\frac{N\hat{H}(x_{i-1})}{N+1}\left(1 - \frac{N\hat{H}(x_{i-1})}{N+1}\right)\right]^{-1.5+\varepsilon} q(H(x_{i-1})) \right|,
\end{aligned}$$

where in the last equality both bounds are true on B_{N,ε_1} by Assumption (E2) and the fact that the x's are in $[0, a)$ and $z \geq N\hat{H}(x_{i-1})/(N+1)$. Thus by the linear bounds of Shorack (1972), the above is

$$\begin{aligned}
&\leq K \sum \left[H(x_i)(1 - H(x_i)) \right]^{-0.5+\varepsilon} \left| q(H(x_i)) - q(H(x_{i-1})) \right| \\
&\quad + K \sum \left| \left(\hat{H}(x_i) - \hat{H}(x_{i-1})\right) \left[H(x_{i-1})(1 - H(x_{i-1})) \right]^{-0.5+\varepsilon} q(H(x_{i-1})) \right|.
\end{aligned}$$

Thus the total variation of $\hat{J}(N\hat{H}/(N+1))q(H)$ in $[0, a)$ (for $\omega \in B_{N,\varepsilon_1}$) is bounded by

$$K \int_0^a \left[H(1-H) \right]^{-0.5+\varepsilon} dq(H) + K \int_0^a \left[H(1-H) \right]^{-1.5+\varepsilon} q(H)\, d\hat{H}. \tag{4.7}$$

Working similarly, but expanding $\hat{J}(N\hat{H}(x_i)/(N+1))$ [instead of $\hat{J}(N\hat{H}(x_{i-1})/(N+1))$] in (4.6) we obtain that the total variation of $\hat{J}(N\hat{H}/(N+1))q(H)$ in $[a,\infty)$ is bounded by

$$(4.8)\quad K\int_a^{\infty}[H(1-H)]^{-0.5+\varepsilon}\,dq(H) + K\int_a^{\infty}[H(1-H)]^{-1.5+\varepsilon}q(H)\,d\hat{H}.$$

From (4.7) and (4.8) it is easily seen that the total variation of $I_{B_{N,\varepsilon_1}}\hat{J}(N\hat{H}/(N+1))q(H)$ in $[0,\infty)$ will be bounded uniformly in N almost surely provided the integral $\int[H(1-H)]^{-1.5+\varepsilon}q(H)\,dH = \int[H(1-H)]^{-1+\varepsilon-\delta}\,dH$ is finite; but this is easily seen to be true. Finally, that $\hat{\Phi}q(H)$ is uniformly of bounded variation can be seen easily. □

THEOREM 4.2. *Under Assumptions* (E), *the rank test based on the statistic* S_N *has Pitman efficiency* 1 *with respect to the test based on the statistic* L_N.

PROOF. It follows by an easy contiguity argument. □

REMARK 4.1. The orthogonality result of Section 2.3 and some standard conditions together with Theorem 4.2 imply that the rank statistic S_N has Pitman efficiency 1.

REMARK 4.2. To carry out the test using the statistic in (2.7) we need to know its asymptotic variance (see also Remark 2.4 when the sample sizes are small). Suppose for the moment that the (random) score function $\hat{J}$ in (2.7) is replaced by a (nonrandom) score function J. Then the usual theory of rank statistics applies. In particular if we set $c=(q,\dots,q,-p,\dots,-p)$ then relation (3), page 159 of Hájek and Šidák (1967), holds so that Theorem a, page 163 of the same reference, implies that the asymptotic variance is $n_1n_2/N\int_0^1(J(u)-\bar{J})^2\,du$. We claim that the asymptotic variance for the statistic in (2.7) may be approximated by $n_1n_2/N\int_0^1(\hat{J}(u)-\bar{\hat{J}})^2\,du$. Conditions under which this holds, together with efficiency calculations will be presented elsewhere.

REMARK 4.3. Results similar to the two-sample case can easily be extended to k-samples and regression.

5. Discussion. The problem of constructing efficient rank tests for the two-sample scale problem under an interval censoring model has been considered. The approach adopted resulted in a rank statistic whose score function involves an estimate of the common scale σ under the null hypothesis, as well as a specification of the entry time distribution G. Thus, in contrast to the usual iid case one needs both a target distribution F for the (unobserved) variable of interest and a target distribution G before the test statistic can be specified; the distribution function F is also used in the derivation of a $\sqrt{n}$-consistent estimator of σ.

The effects of using the wrong G need to be investigated. In particular it should be examined if there are any advantages in using the rank statistic (2.7) instead of the scores statistic (2.5) when both target distributions F, G differ from the true ones.

The technique of the present paper can be safely applied when G is known or specified from the design (see Section 2.5). Improvements, however, might still be possible. In particular, a judicious choice of G (which in some way should be related to F) may result in simpler score functions as well as improved efficiency.

Acknowledgments. The author is grateful to the Associate Editor and the referees for their careful reading of the manuscript and for their constructive and insightful comments that resulted in a much improved version of the manuscript.

REFERENCES

ABEL, U. (1986). A nonparametric test against ordered alternatives for data defined by intervals. *Statist. Neerlandica* **40** 87–91.

BICKEL, P. J. (1982). On adaptive estimation. *Ann. Statist.* **10** 647–671.

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

CHANG, M. N. (1987). Weak convergence of a self-consistent estimator of the survival function with doubly censored data. Unpublished.

CHANG, M. N. and YANG, G. (1987). Strong consistency of a nonparametric estimator of the survival function with doubly censored data. *Ann. Statist.* **15** 1536–1547.

DIACONIS, P. and ENGEL, E. (1986). Discussion of "Some statistical applications of Poisson's work," by I. J. Good. *Statist. Sci.* **1** 171–174.

GILL, R. D. (1988). Non- and semi-parametric maximum likelihood estimators and the von Mises method, Part I. *Scand. J. Statist.* To appear.

GOOD, I. J. (1986). Some statistical application of Poisson's work (with discussion). *Statist. Sci.* **1** 157–180.

GROENEBOOM, P. (1987). Asymptotics for incomplete censored observations. Report 87-18, Univ. Amsterdam.

HÁJEK, J. and ŠIDÁK, Z. (1967). *Theory of Rank Tests.* Academic, New York.

KARIYA, T. (1981). Statistical analysis of the interval-censored sample. *J. Japan. Statist. Soc.* **11** 143–160.

LOÈVE, M. (1963). *Probability Theory.* Van Nostrand, Princeton, N.J.

MANTEL, N. (1967). Ranking procedures for arbitrarily restricted observations. *Biometrics* **23** 65–78.

NEYMAN, J. (1959). Optimal asymptotic tests of composite statistical hypotheses. In *Probability and Statistics, The Harold Cramér Volume* (U. Grenander, ed.) 213–234. Almqvist and Wiksell, Stockholm.

PETO, R. (1973). Experimental survival curves for interval-censored data. *Appl. Statist.* **22** 86–91.

SCHEMPER, M. (1983). A nonparametric k-sample test for data defined by intervals. *Statist. Neerlandica* **37** 69–71.

SCHEMPER, M. (1984). A generalized Friedman test for data defined by intervals. *Biometrical J.* **26** 305–308.

SHORACK, G. R. (1972). Functions of order statistics. *Ann. Math. Statist.* **43** 412–427.

TSAI, W. Y. and CROWLEY, J. J. (1985). A large sample study of generalized maximum likelihood estimators from incomplete data via self-consistency. *Ann. Statist.* **13** 1317–1334.

TURNBULL, B. W. (1976). The empirical distribution function with arbitrary grouped, censored and truncated data. *J. Roy. Statist. Soc. Ser. B* **38** 290–295.

DEPARTMENT OF STATISTICS
219 POND LABORATORY
PENNSYLVANIA STATE UNIVERSITY
UNIVERSITY PARK, PENNSYLVANIA 16802