Ann. Inst. Statist. Math.
Vol. 44, No. 4, 745–753 (1992)

SYMMETRIZED APPROXIMATE SCORE RANK TESTS FOR THE TWO-SAMPLE CASE

MICHAEL G. AKRITAS[1] AND RICHARD A. JOHNSON[2]

[1] *Department of Statistics, Pennsylvania State University, 219 Pond Laboratory, University Park, PA 16802, U.S.A.*
[2] *Department of Statistics, University of Wisconsin, 1210 West Dayton Street, Madison, WI 53706, U.S.A.*

(Received September 11, 1990; revised November 11, 1991)

Abstract. Rank test statistics for the two-sample problem are based on the sum of the rank scores from either sample. However, a critical difference can occur when approximate scores are used since the sum of the rank scores from sample 1 is not equal to minus the sum of the rank scores from sample 2. By centering and scaling as described in Hajek and Sidak (1967, *Theory of Rank Tests*, Academic Press, New York) for the uncensored data case the statistics computed from each sample become identical. However such symmetrized approximate scores rank statistics have not been proposed in the censored data case. We propose a statistic that treats the two approximate scores rank statistics in a symmetric manner. Under equal censoring distributions the symmetric rank tests are efficient when the score function corresponds to the underlying model distribution. For unequal censoring distributions we derive a useable expression for the asymptotic variance of our symmetric rank statistics.

Key words and phrases: Two-sample problem, approximate scores, Pitman efficiency, unequal censoring, Skorokhod construction.

1. Introduction

Let X_{1i}, $i = 1, \dots, n_1$ and X_{2j}, $j = 1, \dots, n_2$ be independent random samples from populations with continuous distribution functions (df) F_1, F_2 respectively. The two-sample problem tests the null hypothesis $H_0 : F_1 = F_2$ versus location or scale alternatives, i.e. $H_1 : F_2(x) = F_1(x - \theta)$, $\theta \in (-\infty, \infty)$ or $H_1 : F_2(x) = F_1(x/\theta)$, $\theta > 0$. A rank test statistic for this problem is of the form

$$(1.1) \qquad T_{1N} = \sum_{i=1}^{N} a_N(i) W_{Ni} \quad \text{or} \quad T_{2N} = \sum_{i=1}^{N} a_N(i)(1 - W_{Ni})$$

where $N = n_1 + n_2$, W_{Ni} is 1 (0) as the i-th smallest value in the combined sample is an X_{1i} (X_{2j}), and $a_N(i) = J(i/(N+1))$ (approximate scores) or $a_N(i) = EJ(U_{Ni})$

(exact scores) for some function $J(s)$ on $[0,1]$. Exact scores are usually hard to compute, the exponential and logistic scores being notable exceptions; also most statistical packages produce good approximations to exact normal scores. Approximate scores rank statistics are asymptotically equivalent to statistics with exact scores and are computationally simpler. Examples of practically useful rank statistics with approximate scores are the squared-ranks statistic (Taha (1964), Duran and Mielke (1968)), the quantile based statistics (Johnson *et al.* (1987)), the class of powers of ranks statistics (Mielke (1972)) and the class of generalized F distribution scores statistics (McKean and Sievers (1989)). When exact scores are used the two samples are treated symmetrically in the sense that $T_{1N} = -T_{2N}$. However, this is not the case when approximate scores are used. As illustrated below, T_{1N} can differ considerably from $-T_{2N}$ and the two statistics will often lead to contradictory results.

A similar situation exists under random censoring. In this case the data consist of (Z_{ti}, δ_{ti}), $i = 1, \ldots, n_t$, $t = 1, 2$, with $Z_{ti} = \min(X_{ti}, Y_{ti})$, $\delta_{ti} = I(X_{ti} \le Y_{ti})$ where X_{ti}, $i = 1, \ldots, n_t$, $t = 1, 2$ are as described above and Y_{1i}, $i = 1, \ldots, n_1$, Y_{2j}, $j = 1, \ldots, n_2$ are two independent samples distributed according to G_1, G_2, respectively, independently of the X-samples. Further, let $T_{(1)} < \cdots < T_{(k)}$ denote the k ordered uncensored observations from the combined sample and let T_{ij}, $j = 1, \ldots, m_i$ denote the censored times in $[T_{(i)}, T_{(i+1)})$, and let W_{Ni} (W_{Nij}) be 1 or 0 according to whether $T_{(i)}$ (T_{ij}) is a Z_{1j} or a Z_{2j}. In this notation, a rank statistic for the two-sample problem with censored data is of the form

$$S_{1N} = \sum_{i=1}^{k} \Big(a_N(i) W_{Ni} + A_N(i) \sum_{j=1}^{m_i} W_{Nij} \Big) \quad \text{or}$$

$$(1.2) \qquad S_{2N} = \sum_{i=1}^{k} \Big(a_N(i)(1 - W_{Ni}) + A_N(i) \sum_{j=1}^{m_i} (1 - W_{Nij}) \Big)$$

where the scores $a_N(i)$, $A_N(i)$ are defined through score functions $J_u(s)$, and $J_c(s) = (1-s)^{-1} \int_s^1 J_u(v)dv$. The exact scores are given in Kalbfleisch and Prentice ((1980), p. 154); the approximate scores are $a_N(i) = J_u(N\hat{F}(T_{(i)})/(N+1))$, $A_N(i) = J_c(N\hat{F}(T_{(i)})/(N+1))$, where $\hat{F}$ is a combined Kaplan-Meier estimator from the two samples. Again, the two samples are treated symmetrically only if exact scores are used. Now, however, exact exponential scores are the only widely available exact scores. Exact scores for a family of distributions related to the logistic have also been obtained by Pettitt (1983).

To illustrate the difference between the two statistics with approximate scores consider the data in Kalbfleisch and Prentice ((1980), Table 6.1, p. 147). In this data set $n_1 = 5$ with one censored observation and $n_2 = 4$ with one censored observation. When exact exponential scores are used, $S_{1N} = -2.26 = -S_{2N}$; however with approximate exponential scores, $S_{1N} = -2.36$ and $S_{2N} = 1.96$. If all the data were treated as uncensored and exact exponential scores are used, $T_{1N} = -2.53 = -T_{2N}$; however with approximate exponential scores $T_{1N} = -2.705$ and $T_{2N} = 1.627$.

The purpose of this paper is to propose a symmetric approximate scores rank statistic defined by

$$S_N = \frac{n_2}{N} S_{1N} - \frac{n_1}{N} S_{2N}. \tag{1.3}$$

In the next section it will be seen that this combination is suggested by the efficient scores test statistic. More insight into the statistic in (1.3) is gained by noting that it equals the conventional rank statistic with centered score functions. In the uncensored case, use of centered score functions is suggested by Hajek and Sidak ((1967), p. 61).

Remark 1.1. In the censored data case, Cuzick (1985) showed the asymptotic equivalence of the approximate scores rank statistic S_{1N} (or S_{2N}) to the corresponding statistic with exact scores but did not address the issue of choosing between S_{1N} and S_{2N}. It should be noted that our Theorem 2.1 is different in nature than Cuzick's Theorem 1 as it deals with the (parametric) efficient scores statistic. Thus, in addition to motivating the symmetric approximate scores statistic, Theorem 2.1 allows a direct proof of the efficiency of S_N (see Corollary 2.1).

For simplicity the effect of using uncentered score function is demonstrated with uncensored data. We performed a simulation study with exponential samples to evaluate the nominal significance level of six scaled rank statistics. The exact and approximate log-rank statistics are

$$\begin{aligned}
ELR &= \sum_{i=1}^{N} c_N(i) W_{Ni} \Big/ \left[\frac{n_1 n_2}{N(N-1)} \sum_{i=1}^{N} c_N(i)^2\right]^{1/2}, \\
ALR1 &= T_{N1} \Big/ \left[\frac{n_1 n_2}{N(N-1)} \sum_{i=1}^{N} a_N(i)^2\right]^{1/2}, \\
ALR2 &= T_{N2} \Big/ \left[\frac{n_1 n_2}{N(N-1)} \sum_{i=1}^{N} a_N(i)^2\right]^{1/2}, \\
SALR &= T_N \Big/ \left[\frac{n_1 n_2}{N(N-1)} \sum_{i=1}^{N} a_N(i)^2\right]^{1/2},
\end{aligned}$$

where $c_N(i) = n^{-1} + (n-1)^{-1} + \cdots + (n-i+1)^{-1} - 1$ are the exact scores, T_{N1}, T_{N2} employ $a_N(i) = -\log(1 - i/(n+1)) - 1$, and in analogy to (1.3) the symmetrized statistic $T_N = (n_2/N)T_{1N} - (n_1/N)T_{2N}$. The other two statistics are $AMIN = \min(|ALR1|, |ALR2|)$ and $AMAX = \max(|ALR1|, |ALR2|)$.

These simulation results indicate that, with small samples, the attained level of the symmetric test is close to the nominal level whereas the attained level of each of the two conventional tests is very discrepant. The practice of choosing the smaller or the larger of $ALR1$, $ALR2$ should clearly be avoided.

Our main results, pertaining to randomly censored data, are stated in the next section as Theorems 2.1, 2.2. The proofs are given in the Appendix.

Table 1. Achieved level and power of exact and approximate log-rank statistics with exponential samples using 1000 replications and nominal level $\alpha = 0.1$ (uncensored samples).

(n_1, n_2)	ELR^a	$ALR1^a$	$ALR2^a$	$ASLR^a$	$AMIN^a$	$AMAX^a$
			Null Hypothesis			
(8, 8)	0.95	0.136	0.130	0.100	0.025	0.241
(8, 15)	0.112	0.129	0.156	0.112	0.045	0.240
(15, 15)	0.112	0.136	0.124	0.111	0.047	0.213
(10, 20)	0.107	0.124	0.133	0.105	0.049	0.208
		Alternative Hypothesis : Scale Ratio = $\exp(1.6N^{-1/2})$				
(8, 8)	0.171	0.299	0.103	0.168	0.071	0.331
(8, 15)	0.155	0.245	0.053	0.156	0.030	0.268
(15, 15)	0.195	0.306	0.123	0.195	0.106	0.323
(10, 20)	0.162	0.235	0.075	0.165	0.05	0.260
		Alternative Hypothesis : Location Difference = $2N^{-1/2}$				
(8, 8)	0.285	0.425	0.164	0.292	0.148	0.441
(8, 15)	0.249	0.364	0.125	0.254	0.107	0.382
(15, 15)	0.237	0.370	0.145	0.247	0.136	0.379
(10, 20)	0.254	0.356	0.112	0.263	0.101	0.367

[a] The statistic is explained in the main text.

2. The main results

We begin by deriving the efficient scores test statistic for the two-sample scale problem. This will be used both for motivating the symmetrized rank statistic and for establishing its asymptotic efficiency.

Assume we observe (Z_{ti}, δ_{ti}), $i = 1, \ldots, n_t$, $t = 1, 2$, where

$$Z_{ti} = \min(X_{ti}, Y_{ti}), \qquad \delta_{ti} = I(X_{ti} \leq Y_{ti}) \tag{2.1}$$

as described in Section 1. We will consider the scale alternative model $F_1(x) = F(x\sigma_1)$, $F_2(x) = F(x\sigma_2)$ where F is a specified distribution function. However location alternatives can be treated similarly. We will employ the parameter transformation $(\sigma_1, \sigma_2) \leftrightarrow (\sigma, \theta)$, where

$$\sigma_1 = \sigma e^{-q\theta}, \qquad \sigma_2 = \sigma e^{p\theta} \tag{2.2}$$

where $p = n_1/N$ and $q = 1 - p$ with $N = n_1 + n_2$. This reparametrization is due to Neyman and Scott (1967). Note that $\sigma_2/\sigma_1 = \exp(\theta)$ so the hypothesis of equal scales can be written as $\theta = 0$. Covariance calculations under $\theta = 0$ show that in this two-sample model the parameters σ and θ are orthogonal (Moran (1970)). This result implies that the efficient scores statistic for testing $\theta = 0$ can be obtained by replacing a square root n consistent estimator for σ in the efficient scores statistic (Neyman (1959)). It should be remarked that the above choice of p and q is crucial for the orthogonality.

Assume for the moment that σ is known. The efficient scores test statistic for $\theta = 0$ has the following functional representation

$$\begin{aligned}(2.3)\qquad L_N &= \left\{qn_1 \int J_u[F(\sigma x)]d\hat{H}_{1u}(x) - pn_2 \int J_u[F(\sigma x)]d\hat{H}_{2u}(x)\right\} \\ &+ \left\{qn_1 \int J_c[F(\sigma x)]d\hat{H}_{1c}(x) - pn_2 \int J_c[F(\sigma x)]d\hat{H}_{2c}(x)\right\} \\ &=: L_{Nu} + L_{Nc}\end{aligned}$$

where $\hat{H}_{tu}(x) = n_t^{-1} \sum_{i=1}^{n_t} I(X_{ti} \le x, \delta_{ti} = 1)$, $\hat{H}_{tc}(x) = n_t^{-1} \sum_{i=1}^{n_t} I(X_{ti} \le x, \delta_{ti} = 0)$, $t = 1, 2$, and

$$(2.4)\qquad J_u(s) = -1 - F^{-1}(s)\frac{f'}{f}(F^{-1}(s)), \qquad J_c(s) = \frac{1}{1-s}\int_s^1 J_u(s)ds.$$

The functional expression (2.3) suggests a rank statistic of the form

$$\begin{aligned}(2.5)\qquad S_N &= \left\{qn_1 \int J_u[\hat{F}(x)]d\hat{H}_{1u}(x) - pn_2 \int J_u[\hat{F}(x)]d\hat{H}_{2u}(x)\right\} \\ &+ \left\{qn_1 \int J_c[\hat{F}(x)]d\hat{H}_{1c}(x) - pn_2 \int J_c[\hat{F}(x)]d\hat{H}_{2c}(x)\right\} \\ &=: S_{Nu} + S_{Nc}\end{aligned}$$

where $\hat{F}_t$, $t = 1, 2$ is the Kaplan-Meier estimator for the survival distribution obtained from sample t and

$$(2.6)\qquad \hat{F} = p\hat{F}_1 + q\hat{F}_2.$$

(In practice one actually uses $\hat{F}^* = \hat{F}N/(N+1)$ but, for notational simplicity, this will not be made explicit.)

In the above discussion we assumed that σ was known. However, S_N is free of σ since, under the null hypothesis, $\hat{F}$ estimates $F(\sigma x)$ without requiring σ to be known. Since the proof of the asymptotic equivalence of S_N and L_N is the same for all values of σ, without loss of generality, we take $\sigma = 1$.

The results of this paper will be shown under the following assumptions

(E1) (i) F has a differentiable density f.

(ii) p remains bounded away from 0 and 1.

(E2) For $\alpha = u$ or c the derivative $J'_\alpha(\cdot)$ exists on $(0, 1)$ is continuous and we have $|J_\alpha(s)| \le K[s(1-s)]^{-.5+\beta}$, $|J'_\alpha(s)| \le K[s(1-s)]^{-1.5+\beta}$ for all $s \in (0, 1)$ where K, β are positive constants.

Assumption (E2) has been checked for the exponential scores, log-normal scores, half-logistic scores and generalized F-distribution scores.

THEOREM 2.1. *Let* (E1), (E2) *hold and assume further that* $G_1 = G_2$. *Then*

$$(2.7)\qquad N^{-1/2}(L_N - S_N) \to 0, \qquad \text{as} \quad N \to \infty,$$

in probability under $\theta = 0$.

PROOF. Relation (2.7) will follow from relations

$$N^{-1/2}(S_{Nu} - L_{Nu}) \xrightarrow{P} 0 \quad \text{and} \tag{2.8}$$

$$N^{-1/2}(S_{Nc} - L_{Nc}) \xrightarrow{P} 0. \tag{2.9}$$

For $t = 1, 2$, let H_{tu} (H_{tc}) denote the sub-distribution functions $P(X_{ti} \le x, \delta = 1)$ ($P(X_{ti} \le x, \delta = 0)$). Noting that under $\theta = 0$ and the assumption that $G_1 = G_2$ we have $H_{1u} = H_{2u} =: H_u$ it follows that (2.8) will be established by showing

$$n_t^{1/2} \int [J_u(\hat{F}) - J_u(F)] d(\hat{H}_{tu} - H_u) \xrightarrow{P} 0, \qquad t = 1, 2. \tag{2.10}$$

Intuitively this is justified by noting that $(\hat{H}_{tu} - H_u)$ can absorb $n_t^{1/2}$ without "blowing up" while $J_u(\hat{F}) - J_u(F) \to 0$ pointwise; for a rigorous proof of (2.10) see Akritas and Johnson (1990). The proof of relation (2.9) follows by similar arguments and this completes the proof of the theorem.

A standard contiguity argument yields the followig corollary.

COROLLARY 2.1. *The rank statistic S_N has Pitman efficiency one with respect to the efficient scores statistic L_N.*

Theorem 2.2 below gives a variance formula for S_N under unequal censoring distributions. First we need the following

LEMMA 2.1. *Let $\tau = \max(\tau_{G_1}, \tau_{G_2})$ and assume that $\tau < \tau_F$, where $\tau_G = \sup\{x : G(x) < 1\}$ for any d.f. G. Set $Q(x) = \int_0^x [J_c(F(y)) - J_u(F(x))] d[G_1(y) - G_2(y)]$. Then*

$$N^{-1/2} S_N \simeq pqN^{1/2} \Big\{ \int J_u(F) d\hat{H}_{1,u} - \int J_u(F) d\hat{H}_{2,u} + \int J_c(F) d\hat{H}_{1,c} - \int J_c(F) d\hat{H}_{2,c} + \int (\hat{F} - F) dQ \Big\} \tag{2.11}$$

where $\simeq$ denotes asymptotic equivalence.

The proof is given in the appendix.

Remark 2.1. Under equal censoring distributions $Q(x) = 0$ and the right-hand side of (2.11) is a scaled difference of two averages. Each average is formed from a set of iid random variables and the two sets are independent. Thus, under the null hypothesis, the asymptotic variance of S_N is given by

$$N^{-1}\operatorname{Var}(S_N) = pq\operatorname{Var}[J(X,\delta)] \tag{2.12}$$

where $J(X,\delta) = \delta J_u(F(X)) + (1-\delta)J_c(F(X))$. An estimate of $\operatorname{Var}[J(X,\delta)]$ is $N^{-1}[\sum_{j=1}^{n_1}\hat{J}(X_{1j},\delta_{1j})^2 + \sum_{j=1}^{n_2}\hat{J}(X_{ij},\delta_{ij})^2]$ where $\hat{J}(X,\delta) = \delta J_u(\hat{F}(X)) + (1-\delta)J_c(\hat{F}(X))$. In the uncensored case (2.12) reduces to the familiar expression

$$N^{-1}\operatorname{Var}(S_N) = pq\operatorname{Var}[J_u(F(X))] \tag{2.13}$$

and $\operatorname{Var}[J_u(F(X))]$ can be estimated from $N^{-1}\sum_{i=1}^N J_u^2(i/(N+1))$.

The next lemma is essentially contained in Breslow and Crowley (1974) and can also be obtained from the results in Peterson (1977).

LEMMA 2.2. *For $t = 1,2$, $n_t^{1/2}(\hat{F}_t(x) - F(x))$ is asymptotically equivalent to*

$$[1-F(x)]\left[\int_0^x \frac{dn_t^{1/2}[\hat{H}_{tu} - H_{tu}]}{1-H_t} + \int_0^x n_t^{1/2}(\hat{H}_t - H_t)dC_t\right]$$

where $\hat{H}_t(x) = n_t^{-1}\sum_{i=1}^{n_t} I(X_{ti} \le x)$, $H_t = 1-(1-F)(1-G_t)$ and $C_t(x) = \int_0^x (1-H_t)^{-2}dH_{tu}$.

Since the upper limits in the integrals in (2.11) can be taken to be τ, the last term on the right-hand side of (2.11) is

$$N^{1/2}\int_0^\tau (\hat{F}-F)dQ = \int_0^\tau [Q(\tau)-Q]dN^{1/2}(\hat{F}-F). \tag{2.14}$$

Lemma 2.2 and some direct calculations yield

LEMMA 2.3. *Assume $\tau < \tau_F$ and let $U(x) = Q(\tau) - Q(x)$. Then for $t = 1,2$,*

$$\begin{aligned}\int U(x)dn_t^{1/2}(\hat{F}_t(x) - F(x)) &= \int D_{tu}(x)dn_t^{1/2}(\hat{H}_{tu}(x) - H_{tu}(x))\\ &\quad + \int D_{tc}(x)dn_t^{1/2}(\hat{H}_{tc}(x) - H_{tc}(x))\end{aligned}$$

where $D_{tu}(x) = \left[U(x) + \int_x^\tau FdU\right][1-H_t(x)]^{-1} + \int_x^\tau (C_t(x) - C_t)(1-F)dU$ and $D_{tc}(x) = \int_x^\tau (C_t(x)-C_t)(1-F)dU$.

Lemma 2.1, Lemma 2.3 and relation (2.6) yield

THEOREM 2.2. *Under the assumption that $\tau < \tau_F$,*

$$\begin{aligned}N^{-1/2}S_N \simeq pqN^{1/2}\Big\{&\int V_{1u}(x)d\hat{H}_{1u}(x) - \int V_{2u}(x)d\hat{H}_{2u}(x)\\ &+ \int V_{1c}(x)d\hat{H}_{1c}(x) - \int V_{2c}(x)d\hat{H}_{2c}(x) + B\Big\}\end{aligned}$$

where $V_{1u} = J_u(F) + pD_{1u}$, $V_{2u} = J_u(F) - qD_{2u}$, $V_{1c} = J_c(F) + pD_{1c}$, $V_{2c} = J_c(F) - qD_{2c}$ *and* B *is a centering constant.*

As in Remark 2.1 it follows that the asymptotic variance of S_N is given by

$$N^{-1}\operatorname{Var}(S_N) = pq\{q\operatorname{Var}[V_1(X_1,\delta_1)] + p\operatorname{Var}[V_2(X_2,\delta_2)]\}, \tag{2.15}$$

$V_t(X,\delta) = \delta V_{tu}(X) + (1-\delta)V_{tc}(X) = J(X,\delta) + n_t N^{-1}[\delta D_{tu}(X) + (1-\delta)D_{tc}(X)]$. An estimate of $\operatorname{Var}[V_t(X_t,\delta_t)]$ is the sample variance of $\hat{V}_t(X_{ti},\delta_{ti})$, $i = 1,\ldots,n_t$.

Appendix

Proof of Lemma 2.1. For simplicity we present a proof under the assumption that $|J'_\alpha(0)| < \infty$, $\alpha = u,c$. This covers exponential scores, half-logistic scores, powers of ranks scores and generalized F-distribution scores, but not the log-normal scores. Write

$$J_\alpha(\hat{F}) = J_\alpha(F) + (\hat{F} - F)J'_\alpha(z_\alpha), \qquad \alpha = u,c \tag{A.1}$$

where z_α is between $\hat{F}$ and F. From (2.5) and (A.1) it follows that

$$\begin{aligned} N^{-1/2}S_N &= pqN^{1/2}\Big[\int J_u(\hat{F})d\hat{H}_{1u} - \int J_u(\hat{F})d\hat{H}_{2u} \\ &\qquad + \int J_c(\hat{F})d\hat{H}_{1c} - \int J_c(\hat{F})d\hat{H}_{2c}\Big] \\ &= pqN^{1/2}\Big[\int J_u(F)d\hat{H}_{1u} - \int J_u(F)d\hat{H}_{2u} \\ &\qquad + \int J_c(F)d\hat{H}_{1c} - \int J_c(F)d\hat{H}_{2c}\Big] \\ &\quad + pqN^{1/2}\int(\hat{F} - F)\sum_\alpha[J'_\alpha(z_\alpha)d\hat{H}_{1\alpha} - J'_\alpha(z_\alpha)d\hat{H}_{2\alpha}] \end{aligned} \tag{A.2}$$

where the index α takes values u,c. Note that

$$\begin{aligned} &N^{1/2}\int(\hat{F} - F)J'_\alpha(z_\alpha)d(\hat{H}_{t\alpha} - H_{t\alpha}) \\ &\quad = N^{1/2}\int[J_\alpha(\hat{F}) - J_\alpha(F)]d(\hat{H}_{t\alpha} - H_{t\alpha}) \xrightarrow{P} 0 \end{aligned} \tag{A.3}$$

by relation (2.10). Next it is easily seen that

$$N^{1/2}\int(\hat{F} - F)[J'_\alpha(z_\alpha) - J'_\alpha(F)]dH_{t,\alpha} \xrightarrow{P} 0 \text{(A.4)}$$

by the facts that $\tau < \tau_F$ and $|J'_\alpha(0)| < \infty$. Finally some calculations yield

$$Q(x) = \int_0^x J'_u(F)dH_{1u} - \int_0^x J'_u(F)dH_{2u} + \int_0^x J'_c(F)dH_{1c} - \int_0^x J'_c(F)dH_{2c}. \tag{A.5}$$

Relations (A.2)–(A.5) imply Lemma 2.1.

REFERENCES

Akritas, M. G. and Johnson, R. A. (1990). Symmetrized approximate score rank tests for the two-sample case, Tech. Report #78, Department of Statistics, Pennsylvania State University.

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.

Cuzick, J. (1985). Asymptotic properties of censored linear rank tests, *Ann. Statist.*, **13**, 133–141.

Duran, B. S. and Mielke, P. W., Jr. (1968). Robustness of sum of squared ranks test, *J. Amer. Statist. Assoc.*, **63**, 338–344.

Hajek, J. and Sidak, Z. (1967). *Theory of Rank Tests*, Academic Press, New York.

Johnson, R. A., Verrill, S. and Moore, D. H., II (1987). Two-sample rank tests for detecting changes that occur in a small proportion of the treated population, *Biometrics*, **43**, 641–655.

Kalbfleisch, J. D. and Prentice, R. L. (1980). *The Statistical Analysis of Failure Time Data*, John Wiley, New York.

McKean, J. W. and Sievers, G. L. (1989). Rank scores suitable for analyses of linear models under asymmetric error distributions, *Technometrics*, **31**, 207–219.

Mielke, P. W., Jr. (1972). Asymptotic behavior of two-sample tests based on powers of ranks for detecting scale and location alternatives, *J. Amer. Statist. Assoc.*, **67**, 850–854.

Moran, P. A. P. (1970). On asymptotically optimal tests of composite hypotheses, *Biometrika*, **57**, 47–55.

Neyman, J. (1959). Optimal asymptotic tests of composite hypotheses, *Probability and Statistics* (The Harold Cramer Volume, ed. U. Grenander), 416–444, Almquist and Wiksells, Uppsala, Sweden.

Neyman, J. and Scott, E. (1967). Note on techniques of evaluation of single rain stimulation experiments, *Proc. Fifth Berkeley Symp. on Math. Statist. Prob.*, Vol. 5, 327–350, Univ. of California Press, Berkeley.

Peterson, A. V., Jr. (1977). Expressing the Kaplan-Meier estimator as a function of empirical subdisurvival functions, *J. Amer. Statist. Assoc.*, **72**, 854–858.

Pettitt, A. N. (1983). Approximate methods using ranks for regression with censored data, *Biometrika*, **70**, 121–132.

Taha, M. A. H. (1964). Rank tests for scale parameter for asymmetrical one-sided distributions, *Publ. Inst. Statist. Univ. Paris*, **13**, 169–180.