ON SELECTION PROCEDURES BASED ON RANKS:
COUNTEREXAMPLES CONCERNING LEAST FAVORABLE CONFIGURATIONS

by
M. Haseeb Rizvi\(^1\) and George G. Woodworth

TECHNICAL REPORT NO. 6

October 28, 1968

Work supported in part by the U.S. Department of Transportation, Federal Highway Administration, Bureau of Public Roads, under Contract No. FH 11-6667 with Stanford University

The opinions, findings and conclusions expressed in this publication are those of the authors and not necessarily those of the Bureau of Public Roads.

1. On leave of absence from the Ohio State University.

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
The Exponential and the Erlang (Gamma) distributions appear often in traffic flow problems: for example, the former as the distribution of gaps on a highway or a road and the latter as the gap acceptance probability distribution for pedestrians crossing a road and various service time distributions. The scale parameter in these distributions is generally unknown. Several types of roads will give rise to different scale parameters and one may like to select roads or highways on the basis of a ranking of these scale parameters. For instance, one might be interested in selecting a road which has the largest (smallest) mean gap and this will correspond to the selection of the smallest (largest) scale parameter of several exponential distributions. Such a problem will arise if one has a choice between several alternative roads or highways; also the above considerations can lead to effective planning if the construction of a new highway is desired. There is a wealth of literature available for the problem of ranking and selection in the parametric case, see for instance, Barr and Rizvi [1]. If, however, the form of the distributions is unknown and all that is known is that they involve a scale (or a location) parameter, then the use of non-parametric selection procedures is in order. This area of research is relatively of recent origin. Bartlett and Govindarajulu [2], Lehmann [4], Puri and Puri [6,7] and Woodworth [8] have only recently considered some procedures in this area. However, these authors have made certain incorrect assertions about the least favorable configurations
of the parameters and this renders the use of their procedures somewhat undesirable. The purpose of the present paper is to give counterexamples pointing to the inadequacies of the procedures studied in the above papers.

Specifically this paper is concerned with certain multiple-decision procedures based on rank sums which have been proposed for analyzing data in a one-way layout:

\[ X_{ij} = \theta_i + \epsilon_{ij}, \quad i=1, \ldots, k, \quad j=1, \ldots, n \]

where the errors \((\epsilon_{ij})\) are independent, have the same known cumulative distribution function \((\text{cdf}) F\) and where \(\theta = (\theta_1, \ldots, \theta_k)\) is unknown. The problems involving scale parameters can be suitably transformed to involve the location parameters of the above set-up. Two problems are considered:

I. Select the indices of the \(t\) largest \(\theta\)-values.

II. Select a subset containing the index of the largest \(\theta\)-value.

In problem I the experimenter sets a preassigned separation threshold \(\delta^* > 0\) and a preassigned probability threshold \(P^* < 1\) and requires that the procedure he uses have the property that the probability of correct selection is greater than or equal to \(P^*\) whenever the \(t\) largest \(\theta\)-values are at least \(8^*\) larger than the rest of the \(\theta\)-values. This problem might arise if there were \(k\) different batches of raw materials available for purchase and one wanted to select the \(t\) best batches.
In problem II the experimenter sets only the $P^*$ value and requires that, with probability greater than or equal to $P^*$, the selected subset contains the index of the largest $\theta$-value. This problem might arise in the first stage of screening certain types of roads; one would want to reduce the number of types of roads which are to be submitted to further tests but at the same time be reasonably sure of not eliminating any type which has great potentiality as to its use.

In this paper we examine certain procedures which have been claimed in [2], [4], [6], [7], and [8] to be solutions to these problems. We show by means of specific examples that these procedures are in fact not solutions and should be used with caution if they are used at all.
On Selection Procedures Based on Ranks:

Counterexamples Concerning Least Favorable Configurations

By

M. Haseeb Rizvi and George G. Woodworth

1. Introduction.

Let \( \pi_1, \pi_2, \ldots, \pi_k \) denote \( k > 2 \) univariate populations differing only in location; that is, an observation \( x_i \) drawn from \( \pi_i \) has cumulative distribution function (cdf) \( F(x - \theta_i) \) where \( F \) is a known continuous cdf with square integrable density \( f \) but the location parameter vector \( \theta = (\theta_1, \ldots, \theta_k) \) is unknown. Let the ordered values of the location parameters be denoted by \( \theta[1] \leq \theta[2] \leq \ldots \leq \theta[k] \).

Selecting the \( t \) best populations.

The decision problem here is to select the populations corresponding to the \( t < k \) largest \( \theta \)-values. The goal of the decision maker is to find a procedure, say \( R \), and a sample size \( n \) such that the probability of a correct selection using rule \( R \), \( P[CS|R, \theta] \), has the property that

\[
\inf_{\theta \in D(\delta^*)} P[CS|R, \theta] \geq P^*,
\]

where

\[
D(\delta^*) = \{ \theta; \theta[k-t+1] - \theta[k-t] \geq \delta^* \},
\]

and \( \left( \frac{k}{t} \right)^{-1} < \delta^* < 1 \) and \( \delta^* > 0 \) are preassigned constants.
Selecting a subset containing the best population.

The decision problem here is to select a subset of the $k$ populations containing the population associated with $\theta[k]$. The goal of the decision maker is to find for fixed $n$ and preassigned $P^* < 1$ a procedure, say $R'$, such that

$$(1.3) \quad \inf_{\theta} P\{CS|R', \theta \geq P^*\}. $$

We consider two procedures (proposed elsewhere) based on rank sums and show by counterexamples in sections 2 and 3 that they do not satisfy (1.1) (or (1.3)).

2. A procedure based on rank sums for selecting the $t$ best populations.

Let $\{X_{ij}: i = 1, \ldots, k, j = 1, \ldots, n\}$ be $k$ samples each of size $n$ ($n$ is to be determined by (1.1)), $X_{ij}$ being the $j$th observation from $\pi_i$, and let $R_{ij}$ be the rank of $X_{ij}$ among all the observations.

Define the rank sums

$$(2.1) \quad T_{in} = \frac{1}{n^2} \sum_{j=1}^{n} R_{ij}, \quad i = 1, \ldots, k$$

$$(2.2) \quad = \frac{1}{n^2} \sum_{j=1}^{n} \sum_{s=1}^{n} \sum_{r=1}^{k} I(X_{ij} > X_{rs}) + \frac{1}{n},$$

where $I(\cdot)$ is the indicator of the event in parentheses.

The proposed selection rule, call it $R(n)$, is as follows:

i) Draw samples of size $n$ from each population and compute $T_{in}$ for $i = 1, \ldots, k$.

ii) Select the $t$ populations having the largest $T_{in}$-values, resolving ties by the obvious randomization.
The problem now is to find a value \( n = n(5^*, P^*, k, t, F) \) such that \( R(n) \) satisfies (1.1).

In solving this problem a crucial role is played by the slippage configuration \( \theta_0 \):

\[
(2.3) \quad \theta_1 = \ldots = \theta_{[k-t]} = \theta_{[k-t+1]} - 5^* = \ldots = \theta_k - 5^*.
\]

Many selection rules, for example the rule based on the sample means, have the property that the infimum in (1.1) is attained when \( \theta \) is in the slippage configuration; in other words for many rules the slippage configuration is the least favorable configuration. For such rules it is a relatively easy task to find the appropriate value of \( n \) (see, for instance, Example 1 of [1]). The following counterexample, kindly communicated to the authors by E. L. Lehmann, shows that for the rank-sum rule \( R(n) \) the slippage configuration is not least favorable.

**Counterexample 1 (E. L. Lehmann).**

Let \( k = 3, t = 1 \) and let \( F \) be a continuous cdf which places probability \( q \) and \( p = 1 - q \) respectively on the intervals \((0, \varepsilon)\) and \((1, 1 + \varepsilon)\); \( \varepsilon < 1/3 \) is a constant. Let \( 5^* = \varepsilon \) and consider two parameter values:

\[
\theta_0 = (0, 0, 5^*), \quad \theta_1 = (0, 5^*, 25^*).
\]

For \( n = 2 \), we show that

\[
(2.4) \quad P[CS|R(2), \theta_0] > P[CS|R(2), \theta_1].
\]

Since \( \theta_0 \) is in the slippage configuration and \( \theta_0, \theta_1 \in D(5^*) \),
defined by (1.2), this provides the required counterexample.

**Proof:** The supports of the distributions of the populations under the two parameter configurations can be depicted as shown in Figure 1.

![Diagram of distributions](image)

Figure 1: Supports of Distributions.

Let $B_i$ be 0, 1 or 2 according as 0, 1 or 2 observations from $\pi_i$ are in the upper interval of the support of its distribution, $B = (B_1, B_2, B_3)$ and $b = (b_1, b_2, b_3)$ is a realization of $B$. Clearly

$$P[B = b|\theta] = \prod_{i=1}^{3} \left(2_{b_i}\right)^{b_i} q^{2-b_1}$$

for $\theta = \theta_0$ or $\theta_1$.

$R = (R_{ij}: i = 1, 2, 3; j = 1, 2)$ is the vector of ranks and $r = (r_{ij})$ is a realization of $R$. Given $R = r$ a correct selection (selection of $\pi_3$) occurs with probability 1 if $r_{31} + r_{32} > \max(r_{21} + r_{22}, r_{11} + r_{12})$, with probability $\frac{1}{3}$ if $r_{31} + r_{32} = r_{21} + r_{22} > r_{11} + r_{12}$ or $r_{31} + r_{32} = r_{11} + r_{12} > r_{21} + r_{22}$, and with probability $\frac{1}{3}$ if $r_{31} + r_{32} \neq r_{21} + r_{22} = r_{11} + r_{12}$. The conditional probability that $R = r$ given $B = b$ is easy to compute, for example
\[ P[R = (1, 2; 3, 4; 5, 6) \mid B = (0, 0, 0), \theta_i] = \begin{cases} 1/48 & i = 0 \\ 1/8 & i = 1. \end{cases} \]

Thus, for each of the 27 values of \( b \) one can determine the conditional probability of a correct selection given \( B = b \) under \( \theta_0 \) and \( \theta_1 \). For most of the \( b \) the probability is the same under \( \theta_0 \) and \( \theta_1 \) but in the six cases listed in Table 1 there is a difference.

| \( b \)   | \( P[B = b] \) | \( P[CS|B = b, \theta_0] \) | \( P[CS|B = b, \theta_1] \) |
|----------|----------------|---------------------------|---------------------------|
| (0, 1, 0) | \( 2pq^5 \)    | 5/6                       | 1/2                       |
| (1, 0, 0) | \( 2pq^5 \)    | 5/6                       | 1                         |
| (1, 1, 0) | \( 4p^2q^4 \)  | 1/6                       | 0                         |
| (1, 2, 1) | \( 4p^2q^4 \)  | 1/2                       | 0                         |
| (2, 1, 1) | \( 4p^2q^4 \)  | 1/2                       | 1                         |
| (2, 2, 1) | \( 2p^5q \)    | 1/9                       | 0                         |

Thus

\[ P[CS|R(2), \theta_0] - P[CS|R(2), \theta_1] = \frac{1}{3} pq^5 + \frac{2}{3} p^2q^4 + \frac{2}{9} p^5q > 0, \]

which establishes counterexample 1.

The possibility still remains that the slippage configuration is asymptotically \( (\delta^* \to 0) \) least favorable; an asymptotic solution based on this assumption has been claimed by various authors ([4], [7] and [8]).

This solution is as follows:

Let \( A(P^*; k, t) \) be the solution of
(2.5) \[ f \phi^{k-t}(x + A) d\phi^t(x) = P^* \]

where \( \phi \) is the standard normal cdf, and define \( n(8^*, P^*; k, t, F) \) to be the smallest integer larger than

(2.6) \[ A^2(P^*; k, t) / [2(8^* f^2(x) dx)]^2, \]

where \( f \) is the derivative of \( F \). The selection rule

\( R(8^*, P^*; k, t, F) = R(8^*, P^*) \) is the rule \( R(n) \) with \( n \) set equal to \( n(8^*, P^*; k, t, F) \). The natural inclination to call \( R(8^*, P^*) \) "distribution-free" must be resisted; obviously one needs to know \( F \) to carry out this procedure.

If \( \theta \) is in the slippage configuration (2.3), then it can be shown ([7] or [8]) that

\[ \lim_{8^* \to 0} P[C_S|R(8^*, P^*), \theta_0] = P^* \]

The authors of [4] and [8] have incorrectly asserted that the slippage configuration is least favorable (this was also asserted in earlier versions of [7]) from which it would follow that \( R(8^*, P^*) \) satisfies (1.1) asymptotically as \( 8^* \to 0 \); i.e. for fixed \( P^* \), it has been claimed that

(2.7) \[ \lim_{8^* \to 0} \inf_{\theta \in D(8^*)} P[C_S|R(8^*, P^*), \theta] = P^*. \]
The next counterexample shows that (2.7) is false; and it seems to us that this invalidates \( R(\delta^*, P^*) \) as a reasonable procedure since the infimum of \( P[CS] \) is not controlled even asymptotically.

The expedient of the authors of the latest version of [7] of considering only that part of the parameter space where

\[ \theta_k - \theta_1 = O(n^{-\frac{1}{2}}) \]

is difficult to translate into practice.

Does it mean that one should use \( R(\delta^*, P^*) \) only when one is convinced that \( \theta_k - \theta_1 = O(n^{-\frac{1}{2}}) \)?

Counterexample 2.

Consider the logistic cdf \( P(x) = (1 + e^{-x})^{-1} \) and let \( \delta(\delta^*) \in D(\delta^*) \) be a sequence of \( \delta \)-values depending on \( \delta^* \) as follows:

\[
\theta_k = \cdots = \theta_{k-t-1} = -\theta_0, \quad \theta_k = 0, \quad \theta_{k-t+1} = \delta^*, \quad \theta_{k-t+2} = \cdots = \theta_k = \theta_0,
\]

where \( \theta_0 \) is a fixed positive constant and \( \delta^* < \theta_0 \).

We now prove the following assertion: For each \( k \geq 3 \) and each \( t < k \), there exists a value of \( P^* \), say \( P^*_0 \), \( \left(\frac{k}{t}\right)^{-1} < P^*_0 < 1 \), such that
\[(2.9) \quad \lim_{\delta^* \to 0} P[CS|R(\delta^*, P^*), \theta(\delta^*)) < P^*|_{\theta(\delta^*)}] < P^*|_{\theta(\delta^*)}, \]

which clearly contradicts (2.7).

**Lemma 1.**

\[(2.10) \quad \lim_{\delta^* \to 0} P[CS|R(\delta^*, P^*), \theta(\delta^*))]
\leq \Phi(2^{-\frac{1}{2}} A^* p(\theta_0)),\]

where

\[(2.11) \quad A^* = A(F^*; k, t),\]

\[(2.12) \quad p(\theta_0) = \frac{2}{\delta^*} \int_{H_{\theta_0}} (2F - 1) dF/[\int_{H_{\theta_0}} dF - (\int_{H_{\theta_0}} dF)^2]^{\frac{1}{2}}\]

and

\[(2.13) \quad H_{\theta_0}(x) = k^{-1}[(k - t - 1)F(x + \theta_0) + 2F(x) + (t - 1)F(x - \theta_0)].\]

**Proof:** Notice first that if \[\theta_1 \leq \theta_2 \leq \ldots \leq \theta_k,\] then

\[(2.14) \quad P[CS|R(\delta^*, P^*), \theta]\]

\[\leq P[\max_{1 \leq i < k-t} T_{in} \leq \min_{k-t < j < k} T_{jn}|\theta] \]

\[\leq P[T_{k-t+1,n} - T_{k-t,n} \geq 0|\theta],\]
where \( n \) is the smallest integer greater than (2.6). From (2.2) one has, with probability one when \( \theta = \theta(\varepsilon*) \),

\[
T_{k-t+1, n} - T_{k-t, n} = \frac{1}{n^2} \sum_{j=1}^{n} \sum_{s=1}^{n} \{2I(X_{k-t+1, j} > X_{k-t, s}) - 1
\]

\[
+ \sum_{i \neq k-t \text{ or } k-t+1} [I(X_{k-t+1, j} > X_{is})
\]

\[
- I(X_{k-t, j} > X_{is})]\}
\]

\[
(2.15) = \frac{1}{n} \sum_{j=1}^{n} \sum_{i \neq k-t, k-t+1} \{F(\bar{x}_{ij} - \varepsilon*) - F(\bar{x}_{ij})
\]

\[- \frac{1}{n} \sum_{j=1}^{n} \{2F(\bar{x}_{k-t, j} - \varepsilon*) + (k - t - 1)F(\bar{x}_{k-t, j} + \theta_0)
\]

\[+(t - 1)F(\bar{x}_{k-t, j} - \theta_0)\}
\]

\[+ \frac{1}{n} \sum_{j=1}^{n} \{2F(\bar{x}_{k-t+1, j}) + (k - t - 1)F(\bar{x}_{k-t+1, j} + \theta_0)
\]

\[+(t - 1)F(\bar{x}_{k-t+1, j} - \theta_0)\}
\]

\[+ 1 - 2F(x + \varepsilon*)dF(x) + (k - t - 1)fF(x + \theta_0)d(fF(x - \varepsilon*) - F(x))
\]

\[+ (t - 1)fF(x - \theta_0)d(F(x - \varepsilon*) - F(x))
\]

\[+ \varepsilon_n(\theta_0, \varepsilon*),\]

where \( E \varepsilon_n^2(\theta_0, \varepsilon*) \leq C/n^2 \) and \( C \) is an absolute constant. Note that (2.15) is obtained by U-statistic arguments in imitation of, say, the proof of Theorem 5.6, p. 229 of [3].
Let
\[(2.16) \quad W_n = n^{\frac{1}{2}}(T_{k-t+1,n} - T_{k-t,n}),\]
routine calculation yields
\[
EW_n = n^{\frac{1}{2}}(2F(x + \delta^*)dF(x) - 1
+ (k - t - 1)\int(F(x - \theta_0) - F(x - \theta_0 - \delta^*))dF(x)
+ (t - 1)\int(F(x + \theta_0) - F(x + \theta_0 - \delta^*))dF(x).
\]
By (2.6) and (2.11) one has \(n^{\frac{1}{2}}\delta^* \rightarrow A^*[12\int f^2]^{-\frac{3}{2}}\) as \(\delta^* \rightarrow 0\); thus, by Olshen's Lemma (p. 1766 of [5])
\[(2.17) \quad \lim_{\delta^* \rightarrow 0} EW_n = \frac{A^*}{\sqrt{12\int f^2}}\]
\[
\left(2\int f^2(x)dx + (k - t - 1)\int f(x - \theta_0)F(x)dx + (t - 1)\int f(x + \theta_0)F(x)dx.\right)
\]
Also
\[(2.18) \quad \lim_{\delta^* \rightarrow 0} \text{Var}(W_n) = 2k^2[\int_{H_{\theta_0}}^2 dF - (\int_{H_{\theta_0}}^2 dF)^2],\]
where \(H_{\theta_0}\) is defined by (2.13).
If we set \(F(x) = (1 + e^{-x})^{-1}\), then \(f(x) = F(x)(1 - F(x))\) and \(\int f^2 = 1/6\), so that (2.17) becomes, after integrating by parts,
\[
\lim_{\delta^* \rightarrow 0} EW_n = \frac{1}{\sqrt{12\int f^2}} A^* k H_{\theta_0} (2F - 1)dF.
\]
Since (2.15) is asymptotically normal by Liapunov's theorem, it follows that
\[
\lim_{\delta^* \rightarrow 0} P[CS|R(\delta^*, \theta^*), \theta(\delta^*)] \\
\leq \lim_{\delta^* \rightarrow 0} P[T_{k-t+1,n} - T_{k-t,n} \geq 0|\theta(\delta^*)]
\]
\[
\begin{align*}
&= \lim_{\delta \to 0} P \left[ (W_n - EW_n) / \sqrt{\text{Var}(W_n)} \geq -EW_n / \sqrt{\text{Var}(W_n)} \right] \\
&= \Phi(2^{-\frac{3}{2}} A^* \rho(\theta_0)),
\end{align*}
\]
which proves Lemma 1.

Remark. For \( \theta_0 > 0 \), \( H_{\theta_0} \) is clearly not a linear function of \( F \) and, since \( H_{\theta_0} \) and \( F \) are both monotone increasing, we have

\[ (2.19) \quad 0 \leq \rho(\theta_0) < 1. \]

Lemma 2.

For any \( k \) and \( t \)

\[ (2.20) \quad \lim_{P^* \to 1} 2^{\frac{1}{2}\cdot \Phi^{-1}(P^*)} / A^* = 1, \]

where \( A^* = A(P^*; k, t) \) and \( A \) is defined by (2.5).

Proof: Let \( Z_1, \ldots, Z_k \) be independent normal \((0,1)\) random variables. Then,

\[
1 - P^* = 1 - \int \phi^{k-t}(x + A^*) d\phi^t(x)
\]

\[ = P \left[ \max_{1 \leq i \leq k-t} Z_i > \min_{k-t \leq j \leq k} Z_j + A^* \right] \]

\[ = P \left[ \bigcup_{1 \leq i \leq k-t \leq j \leq k} [Z_i > Z_j + A^*] \right] \]

\[ \leq t(k-t) P[Z_1 > Z_k + A^*] \]

\[ = t(k-t) \left[ 1 - \Phi(2^{-\frac{1}{2}} A^*) \right]. \]

Also clearly

\[ 1 - P^* \geq [1 - \Phi(2^{-\frac{1}{2}} A^*)]. \]

\[ \therefore \]
Lemma 2 now is a consequence of the following easily verifiable fact

$$\lim_{u \to 1} \phi^{-1}(u)/[-2 \log(1 - u)]^{1/2} = 1$$

and of the well known approximation to Mills' ratio.

Counterexample 2 now follows from (2.10), (2.19) and (2.20) by selecting $P^*_0$ large enough so that

$$2^{-\frac{3}{2}}A(P_0^*, k, t)/\phi^{-1}(P_0^*) < 1/\rho(\theta_0).$$

A remark on the scale parameter case.

Suppose $\pi_i$ has cdf $F(x/\sigma_i)$ where $F(x) = 0$ for $x < 0$, $F$ is known, and $\sigma = (\sigma_1, \ldots, \sigma_k)$ is unknown (if $F(x) \neq 0$ for $x < 0$ then replace $x$ by $|x|$). $R(n)$; with $X_{ij}$ replaced by $-X_{ij}$, could be used to select the $t$ smallest $\sigma$-values; in [6] it is asserted that, for any constant $\theta^* > 1$, $F[CS|R(n), \sigma]$ attains its minimum, subject to the condition

$$\sigma^2_{[t+1]}/\sigma^2_{[t]} \geq \theta^* > 1,$$

when

$$\theta^* \sigma^2_{[1]} = \ldots = \theta^* \sigma^2_{[t]} = \sigma^2_{[t+1]} = \ldots = \sigma^2_{[k]}.$$

That this is false, even asymptotically ($\theta^* \to 1$), follows from Counterexample 2 by considering the random variable $Y = -\log(X)$, since if $X$ has cdf $F(x/\sigma)$ then $Y$ has cdf $1 - F(exp(\mu - y))$, where $\mu = -\log \sigma$, and $Y_{ij}$ has the same rank as $-X_{ij}$. 
3. A procedure based on rank sums for selecting a subset containing the best population.

The authors of [2] propose the following procedure, call it \( R'(n) \):

Put \( \pi_i \) in the selected subset iff

\[
T_{in} \geq \max_j T_{jn} - c_n
\]

where

\[
(3.1) \quad c_n = (12n)^{-\frac{1}{2}} kA^* + o(n^{-\frac{1}{2}})
\]

and \( A^* = A(P^*; k, l) \), defined by (2.5). We shall show that the slippage configuration: \( \theta[1] = \theta[2] = \ldots = \theta[k] \) is not least favorable by proving the following:

**Counterexample 3.**

Let \( \theta_{\perp} \) denote the configuration

\[
\theta_1 = \ldots = \theta_{k-2} = \perp, \quad \theta_{k-1} = \theta_k = 0
\]

and let \( \theta_{\Delta} \) denote the slippage configuration for this problem:

\[
\theta_1 = \theta_2 = \ldots = \theta_k. \quad \text{If } F(x) \text{ is as in (3.7) and } k \geq 3, \text{ then}
\]

\[
(3.2) \quad \lim_{n \to \infty} P[CS|R'(n), \theta_{\perp}] < P^* = \lim_{n \to \infty} P[CS|R'(n), \theta_{\Delta}].
\]

**Proof:** The equality is established in [2] and the inequality below.

Clearly

\[
(3.3) \quad P[CS|R'(n), \theta_{\perp}] \lesssim P[T_{kn} - T_{k-1,n} \geq c_n | \theta_{\perp}].
\]

It follows as in the proof of Lemma 1 that \( W_n = n^{\frac{3}{2}}(T_{kn} - T_{k-1,n}) \) has a limiting normal distribution with zero mean and variance.
\[ \sigma^2(H) = 2k^2 \{ \int H^2 dF - (\int H dF)^2 \}, \]

where

\begin{equation}
H(x) = k^{-1} [(k - 2)F(x + 1) + 2F(x)].
\end{equation}

Thus by (3.1) and (3.3)

\[
\lim_{n \to \infty} P[CS|R'(n), \theta, l] = \Phi(k(12)^{-\frac{1}{2}}A\ast/\sigma(H)).
\]

It follows from (2.20) that for any \( \varepsilon > 0 \) there exists \( \frac{1}{2} < P_\varepsilon < 1 \) such that

\[ A\ast = A(P_\varepsilon; k, 1) \leq (1 + \varepsilon)2^{\frac{3}{2}}\Phi^{-1}(P_\varepsilon). \]

Thus the counterexample will be proved if it can be shown that

\begin{equation}
\sigma^2(H) > k^2/6.
\end{equation}

From (3.4)

\begin{equation}
\sigma^2(H)/2 = 4/12 + 4(k-2)\text{Cov}(F(x), F(x+1)) + (k - 2)^2 \text{Var}(F(x + 1)),
\end{equation}

where \( X \) has cdf \( F \).

Now let

\begin{equation}
F(x) = \begin{cases} 
1/2 + x/2b & -b < x \leq 0 \\
1/2 & 0 < x \leq 1 \\
1/2 + (x - 1)/2a & 1 < x \leq 1 + a,
\end{cases}
\end{equation}

where \( 0 < a < 1 < b \) are constants to be determined below.
Thus,

\[ F(x + 1) = \begin{cases} 
\frac{1}{2} + (x + 1)/2b & -(b + 1) < x \leq -1 \\
\frac{1}{2} & -1 < x \leq 0 \\
\frac{1}{2} + x/2a & 0 < x \leq a \\
1 & a < x
\end{cases} \]

or, except for a set having zero \( F(x) \)-measure,

\[ (3.8) \quad F(x + 1) = \begin{cases} 
F(x) + 1/2b & 0 < F(x) \leq 1/2 - 1/2b \\
\frac{1}{2} & 1/2 - 1/2b < F(x) \leq 1/2 \\
1 & 1/2 < F(x) \leq 1 \end{cases} \]

If \( X \) has cdf \( F \) then \( F(X) \) is a uniform random variable and it follows from (3.6) and (3.8) that

\[ (3.9) \quad \sigma^2(H)/2 = k^2/12 + (13k^2 - 10)(k - 2)/492 - \beta(3k^2 - 8k + 4)/8 + 3\beta^2(k - 2)^2/8 + \beta^3(k^2 - 4)/6 - \beta^4(k - 2)^2/4 \]

where \( \beta = (2b)^{-1} \). It is clear that for sufficiently small \( \beta \) (large \( b \) the right side of (3.9) can be made larger than \( k^2/12 \) so that (3.5) is satisfied and Counterexample 3 is proved.


Procedures \( R(n) \) and \( R'(n) \) are special cases of the scores procedures proposed in [2], [4], [6], [7] and [8]. The second counterexample probably works for any scores procedure when \( F \) (instead of being logistic) is the cdf against which the scores are locally most powerful.
REFERENCES


