On locally most powerful sequential rank tests

Jan Kalina

To cite this article: Jan Kalina (2017) On locally most powerful sequential rank tests, Sequential Analysis, 36:1, 111-125, DOI: 10.1080/07474946.2016.1275501

To link to this article: http://dx.doi.org/10.1080/07474946.2016.1275501

Published online: 07 Mar 2017.

Submit your article to this journal

Article views: 5

View related articles

View Crossmark data
On locally most powerful sequential rank tests

Jan Kalina

Abstract
Sequential ranks are defined as ranks of such observations, which have been observed so far in a sequential design. This article studies hypotheses tests based on sequential ranks for different situations. The locally most powerful sequential rank test is derived for the hypothesis of randomness against a general alternative, including the two-sample difference in location or regression in location as special cases for the alternative hypothesis. Further, the locally most powerful sequential rank tests are derived for the one-sample problem and for independence of two samples in an analogous spirit as the classical results of Hájek and Šidák (1967) for (classical) ranks. The locally most powerful tests are derived for a fixed sample size and the results bring arguments in favor of existing tests. In addition, we propose a sequential testing procedure based on these statistics of the locally most powerful tests. Principles of such sequential testing are explained on the two-sample Wilcoxon test based on sequential ranks.

1. Introduction

The history of nonparametric sequential tests was initiated by Miller, who proposed a sequential signed-rank test (Miller, 1970) and a two-sample sequential Wilcoxon test (Miller, 1973). Berk (1975) derived a locally most powerful sequential rank test, and its structure was recently characterized by Novikov and Novikov (2010). Lombard investigated a sequential test based on Kendall's tau statistic (Lombard, 1977), invariance principles for a sequential test based on classical ranks (Lombard, 1981), and rank tests for the change-point problem (Lombard, 1983). Inverting sequential tests based on ranks allows defining R-estimators investigated by Jurečková and Sen (1982). Other nonparametric tools for sequential analysis were overviewed by Hušková (1991), and recently they found applications in nonparametric tests in the context of control charts (Zhou et al., 2016).

The simple linear rank statistic (say $M_n$) based on sequential ranks first appeared in the pioneering work of Mason (1981), who considered $M_n$ as a useful tool for deriving limit theorems for tests based on classical ranks and derived the asymptotic normality of $M_n$ using martingale theory for $n \to \infty$. He proved that $M_n$ and the simple linear rank statistic $T_n$ (based on classical ranks) are asymptotically equivalent in the quadratic mean under the null hypothesis that $X_1, \ldots, X_n$ are independent and identically distributed (i.i.d.). Further, he followed Hájek and Šidák (1967) to state that they are asymptotically equivalent also under contiguous alternatives of regression in location.
Mason (1984) proved the Pitman efficiency of $T_n$ and $M_n$ to be the same, and he showed a test based on $T_n$ to be more powerful in terms of the Bahadur efficiency. Lombard and Mason (1985) also derived invariance principles for $M_n$. Though $M_n$ is constructed in these papers in a natural and intuitive way (but used entirely for a fixed sample size), there have been no efforts to show any optimal property of $M_n$ among all possible test statistics based on sequential ranks.

This article fills the gap of the optimality results for various hypothesis tests based on sequential ranks. It has the following structure. After reviewing basic properties of sequential ranks in Section 2, the locally most powerful tests among all tests based on sequential ranks are derived. A rather general first result of Section 3 is formulated for various special cases throughout Sections 4–7. Independent of these results, a proof for the locally most powerful test based on sequential ranks is given for the one-sample test (Section 8) and for the test of independence of two random samples (Section 9). Though these results are valid for a fixed number of observations $n$, we describe a sequential construction of tests in Section 10. Finally, a discussion in Section 11 concludes the article.

2. Sequential ranks

Let $X = (X_1, \ldots, X_n)^T$ represent a random vector with values in $(\mathbb{R}^n, B^n)$, where $B$ is the system of Borel sets on $\mathbb{R}$. Arranging $X$ in ascending order, we obtain the vector of order statistics

$$X_{(1)}^n \leq X_{(2)}^n \leq \cdots \leq X_{(n)}^n,$$

(2.1)

where the upper index stresses that these are computed from $n$ random variables. If no two observations are equal, the rank $R_i$ of the $i$th observation is defined by $X_i = X_{(R_i)}$ for $i = 1, \ldots, n$ and the vector of ranks will be denoted by $R = (R_1, \ldots, R_n)^T$.

Sequential ranks $R^* = (R_{11}, \ldots, R_{nn})^T$ are defined as ranks computed from the data observed so far, denoting $R_{ik}$ the rank of $X_i$ among the values $X_1, \ldots, X_k$ for any $i = 1, \ldots, n$ (Khmaladze, 2011). Throughout the article, we use the notation $x_i$ (or $r_{ii}$) in case we want to stress that the values of $X_i$ (or $R_{ii}$) are considered as fixed.

It holds that $X_i = X_{(R_{ii})}$ for any $i = 1, \ldots, n$; in other words, $R_{ii}$ denotes the number of values among $r_{11}, \ldots, r_{nm}$ smaller or equal to $x_i$ for $i = 1, \ldots, n$. Barndorff-Nielsen (1963) proved the favorable property that $R_{11}, \ldots, R_{nn}$ are independent and

$$P(R_{ii} = r_{ii}) = \frac{1}{i}, \quad i = 1, \ldots, n.$$  

(2.2)

Another useful property is devoted to the conditional expectation of a measurable function $t$ of random variables $X_1, \ldots, X_n$. Conditioning on the sequential ranks $r^* = (r_{11}, \ldots, r_{nm})^T$, it holds that

$$E \left[ t(X_1, \ldots, X_n) | R_{11} = r_{11}, \ldots, R_{nn} = r_{nn} \right] = E t(X_{(R_{11})}, \ldots, X_{(R_{nn})}).$$  

(2.3)

If the sample size $n$ is fixed, sequential ranks can be uniquely determined from the vector of classical ranks in a trivial way. The reversed procedure is also possible but perhaps less obvious. Thus, sequential ranks allow us to reconstruct the original observations conditionally on the knowledge of (2.1). The computations corresponding to these one-to-one mappings are formulated as Algorithms 2.1 and 2.2.
Algorithm 2.1. Computation of classical ranks from sequential ranks.

Require: $R_{11}, \ldots, R_{nn}$
Ensure: $R_1, \ldots, R_n$

$I = \{1, \ldots, n\}$

for $j = 1$ to $n$

$m := \max \{R_{ij}; i \in I\}$
$p := \min \{i; i \in I, R_{ii} = m\}$
$R_p := j$
$I := I \setminus \{p\}$
end for

Algorithm 2.2. Computation of the original observations from arranged observations and sequential ranks.

Require: $X^n_{(1)} \leq \cdots \leq X^n_{(n)}$ and $R_{11}, \ldots, R_{nn}$
Ensure: $X_1, \ldots, X_n$

$I = \{1, \ldots, n\}$

for $j = 1$ to $n$

$m := \max \{R_{ij}; i \in I\}$
$p := \min \{i; i \in I, R_{ii} = m\}$
$X_p := X^n_{(j)}$
$I := I \setminus \{p\}$
end for

The classical theory of rank tests was developed by Hájek and Šidák (1967) and extended in Hájek et al. (1999). Standard monographs on sequential nonparametrics (Sen [1981] or Ghosh and Sen [1981]) studied sequential hypotheses tests based on statistics computed from classical ranks. These are recalculated after each new observation is added. However, hypotheses tests based on sequential ranks were not considered in these monographs.

In this article, we study hypotheses tests based on sequential ranks for various situations, which are shown to be locally most powerful among all tests based on sequential ranks. These tests will be called locally most powerful sequential ranks tests.

**Definition 2.1.** Let $m(Q)$ be a measure of distance of alternative $Q \in H_1$ from the null hypothesis $H_0$. The $\alpha$—test $\Phi_0$ depending on the data only through the sequential ranks is called the locally most powerful sequential rank $\alpha$—test of $H_0$ against $H_1$, if for any other test $\Phi$ depending on the data only through sequential ranks there exists $\epsilon > 0$ such that the powers of the test satisfy $\beta_{\Phi_0} \geq \beta_\Phi$ for each $Q$ with the property $0 < m(Q) < \epsilon$.

3. **Test against a general alternative**

Let us assume i.i.d. random variables $X_1, \ldots, X_n$ to be observed for $n < \infty$. Let $X_1$ have a density with respect to the Lebesgue measure, so that there is a zero probability of two observations attaining the same value. Let $c_1, c_2, \ldots$, denote a known sequence of regression constants. The joint density of the random vector $(X_1, \ldots, X_n)^T$ with fixed $n$ under the null
hypothesis will be denoted by \( p(x_1, \ldots, x_n) \) and under the alternative hypothesis depending on a parameter \( \Delta > 0 \) by \( q^\Delta_n (x_1, \ldots, x_n) \).

For a fixed \( n < \infty \), the null hypothesis of interest formulated in a very general way,

\[
H_0: p(x_1, \ldots, x_n) = \prod_{i=1}^{n} d(x_i, 0),
\]

(3.1)

for a family of densities \( d(x, \theta) \), where \( \theta \in \Theta, \Theta \) is an open interval containing 0 and \( d(x, \theta) \) is absolutely continuous in \( \theta \) for almost every \( x \). \( H_0 \) will be tested against the alternative hypothesis

\[
H_1: q^\Delta_n (x_1, \ldots, x_n) = \prod_{i=1}^{n} d(x_i, \Delta c_i), \quad \Delta > 0.
\]

(3.2)

The null hypothesis can be also formulated as \( H_0 : \Delta = 0 \). We now derive the locally most powerful sequential rank test of \( H_0 \) against the general alternative \( H_1 \).

**Theorem 3.1.** Let the condition II.4.8.A1 of Hájek and Šidák (1967) be fulfilled. Then the test with the critical region

\[
\sum_{i=1}^{n} c_i \mathbb{E} \frac{\hat{d}(X^i_{(r_i)}, 0)}{d(X^i_{(r_i)}, 0)} \geq k_\alpha,
\]

(3.3)

where

\[
\hat{d}(x, \theta) = \frac{\partial}{\partial \theta} d(x, \theta),
\]

(3.4)

is the locally most powerful sequential rank test of \( H_0 \) against (3.2) at level \( \alpha \).

**Proof.** The condition of the assumptions, which is also denoted as 3.4.8.A1 of Hájek et al. (1999), ensures the existence of

\[
\hat{d}(x, 0) = \lim_{\theta \to 0} \frac{1}{\theta} [d(x, \theta) - d(x, 0)]
\]

(3.5)

for every \( x \). Replacing now the data by the sequential ranks, the joint likelihood of the data under \( H_0 \) equals

\[
P(R_{11} = r_{11}, \ldots, R_{nn} = r_{nn}) = \frac{1}{n!}
\]

(3.6)

for the vector \((r_{11}, \ldots, r_{nn})^T\) fulfilling \( r_{ii} \in \{1, \ldots, i\} \) for \( i = 1, \ldots, n \). Under the alternative hypothesis with a fixed \( \Delta \), \( Q^\Delta_n \) will denote the probability distribution corresponding to the density (3.2). This can be expressed as

\[
Q^\Delta_n (\mathbf{r}^* = \mathbf{r}^*) = \int_{\mathbf{r}^* = \mathbf{r}^*} \cdots \int q^\Delta_n (x_1, \ldots, x_n) dx_1 \cdots dx_n.
\]

(3.7)

Neyman-Pearson’s lemma gives the most powerful test of \( H_0 \) against (3.2) based on sequential ranks, which rejects \( H_0 \) if \( n! Q^\Delta_n (\mathbf{r}^* = \mathbf{r}^*) \) exceeds the critical value. Therefore, we study \( Q^\Delta_n (\mathbf{r}^* = \mathbf{r}^*) \) for the system of alternatives \( \{Q^\Delta_n\} \), where \( \Delta \to \infty \).
From (3.2) it follows that (3.7) equals
\[
\frac{1}{n!} + \Delta \sum_{i=1}^{n} \int_{\mathbb{R}^* = \mathbb{R}^n} \left[ \frac{d(x_i, \Delta c_i) - d(x_i, 0)}{\Delta} \prod_{j=i+1}^{n} d(x_j, 0) \prod_{k=1}^{i-1} d(x_k, \Delta c_k) \right] dx_1 \cdots dx_n. \tag{3.8}
\]
Because
\[
\limsup_{\Delta \to 0} \int_{\mathbb{R}^* = \mathbb{R}^n} \left[ \frac{d(x_i, \Delta c_i) - d(x_i, 0)}{\Delta} \prod_{j=i+1}^{n} d(x_j, 0) \prod_{k=1}^{i-1} d(x_k, \Delta c_k) \right] dx_1 \cdots dx_n
\leq |c_i| \int_{-\infty}^{\infty} |d(x_i, 0)| dx_i \tag{3.9}
\]
for \( i = 1, \ldots, n \), the convergence theorem II.4.2 of Hájek and Šidák (1967) can be applied to obtain
\[
\lim_{\Delta \to 0} \frac{1}{n!} \sum_{i=1}^{n} \int_{\mathbb{R}^* = \mathbb{R}^n} \left[ \frac{d(x_i, \Delta c_i) - d(x_i, 0)}{\Delta} \prod_{j=i+1}^{n} d(x_j, 0) \prod_{k=1}^{i-1} d(x_k, \Delta c_k) \right] dx_1 \cdots dx_n
\]
\[
= \sum_{i=1}^{n} \int_{\mathbb{R}^* = \mathbb{R}^n} \left[ c_i d(x_i, 0) \prod_{j \neq i} d(x_j, 0) \right] dx_1 \cdots dx_n
\]
\[
= \sum_{i=1}^{n} c_i \int_{\mathbb{R}^* = \mathbb{R}^n} \left[ \frac{d(x_i, 0)}{d(x_i, 0)} \prod_{j=1}^{n} d(x_j, 0) \right] dx_1 \cdots dx_n
\]
\[
= \frac{1}{n!} \sum_{i=1}^{n} c_i \left[ \frac{\dot{d}(X_i, 0)}{d(X_i, 0)} \bigg| R_{11} = r_{11}, \ldots, R_{nn} = r_{nn} \right], \tag{3.10}
\]
using properties of the conditional expectation.

Now using (2.3) we obtain that (3.9) equals
\[
\frac{1}{n!} \sum_{i=1}^{n} c_i E \left[ \frac{\dot{d}(X_i, 0)}{d(X_i, 0)} \bigg| R_{11} = r_{11}, \ldots, R_{nn} = r_{nn} \right], \tag{3.11}
\]
where the \( i \)-th summand depends on \( X_1, \ldots, X_i \). (3.11) can be interpreted as the linear sequential-rank statistic. So we get an approximation for \( Q^\Delta_n (\mathbb{R}^* = \mathbb{R}^n) \) by
\[
\frac{1}{n!} + \Delta \frac{1}{n!} \sum_{i=1}^{n} c_i E \left[ \frac{\dot{d}(X_i, 0)}{d(X_i, 0)} \right] \tag{3.12}
\]
for \( \Delta \to \infty \). For this limit behavior the linear sequential-rank statistic (3.11) arranges the sequential ranks \( r_{11}, \ldots, r_{nn} \) in the same way as \( Q^\Delta_n \), which is at the same time the locally most powerful sequential rank test of \( H_0 \) against \( H_1 \). This completes the proof. \( \square \)

4. Test of \( H_0 \) against regression in location

The null hypothesis of randomness
\[
H_0 : p(x_1, \ldots, x_n) = \prod_{i=1}^{n} f(x_i) \tag{4.1}
\]
and the alternative hypothesis of regression in location in the form
\[ H_1 : q^n_\Delta(x_1, \ldots, x_n) = \prod_{i=1}^{n} f(x_i - \Delta c_i), \quad \Delta > 0, \quad (4.2) \]
are special cases of (3.1) and (3.2) of the previous section. Let us define
\[ F(x) = \int_{-\infty}^{x} f(y)dy, \quad (4.3) \]
its inverse \( F^{-1}(u) = \inf\{x; F(x) \geq u\} \), and the score function \( \varphi \) by
\[ \varphi(u, f) = \frac{f'(F^{-1}(u))}{f(F^{-1}(u))}, \quad 0 < u < 1. \quad (4.4) \]
Let us assume a random sample of \( n \) random variables \( U_1^n, \ldots, U_n^n \) uniform on \([0, 1]\) and denote their \( i \)-th order statistic by \( U_{(i)}^n \).
Let us assume \( f \) to be absolutely continuous with
\[ \int_{-\infty}^{\infty} |f'(x)|dx < \infty \quad (4.5) \]
to have well-defined scores for a fixed \( n < \infty \)
\[ a_n(i, f) = E[\varphi(U_{(i)}^n, f)] = E\left[-\frac{f'}{f}(F^{-1}(U_{(i)}^n))\right] = E\left[-\frac{f'}{f}(X_{(i)}^n)\right], \quad i = 1, \ldots, n. \quad (4.6) \]

**Theorem 4.1.** The locally most powerful sequential rank test of \( H_0 \) against the system \( \{q^n_\Delta, \Delta > 0\} \) with \( q^n_\Delta \) defined by (4.2) at level \( \alpha \) has the critical region
\[ \sum_{i=1}^{n} c_i a_i(R_{ii}, f) \geq k_\alpha \quad (4.7) \]
assuming (4.5).

This is a consequence of Theorem 3.1. The test statistic (3.3) for the general test can be expressed as
\[ \sum_{i=1}^{n} c_i E\left[-\frac{f'}{f}(X_{(i)}^n)\right]. \quad (4.8) \]
The statistic (4.8) can be expressed as
\[ \sum_{i=1}^{n} c_i E\left[-\frac{f'}{f}(F^{-1}(U_{(i)}^n))\right] = \sum_{i=1}^{n} c_i E\left[\varphi(U_{(i)}^n, f)\right] = \sum_{i=1}^{n} c_i a_i(R_{ii}, f) \quad (4.9) \]
by applying (4.6), because \( X_{(R_{ii})}^i \) has the same distribution as \( F^{-1}(U_{(R_{ii})}^i) \) for each \( i \).

**5. Test of \( H_0 \) against two samples differing in location**

Let \( Z_1, Z_2, \ldots, Z_n \) denote a random sample, which is obtained as a pooled sample of two independent sequences \( X_1, \ldots, X_m \) and \( Y_1, \ldots, Y_{n-m} \). The null hypothesis
\[ H_0 : p(z_1, \ldots, z_n) = \prod_{i=1}^{n} f(z_i) \quad (5.1) \]
is tested against the alternative

\[ H_1 : q^n_{\Delta} = \prod_{i=1}^{n} f(z_i - \Delta c_i), \quad \Delta > 0, \] (5.2)

with a special choice of constants, namely, \( c_j = 1 \) if \( Z_i \) comes from the second sample and \( c_j = 0 \) if \( Z_i \) comes from the first sample. This alternative corresponds to the variables \( X_1, X_2, \ldots \), each with density \( f(x) \) and variables \( Y_1, Y_2, \ldots \), each with density \( f(y - \Delta) \).

Then for a fixed \( n < \infty \) and random sample \( Z_1, \ldots, Z_n \), the test statistic of Theorem 4.1 is used with the special choice of \( c_1, \ldots, c_n \). The sums in (4.8) and (4.9) are evaluated only over the second sample.

The locally most powerful sequential rank test is a special case of (4.7) and the test statistic is equal to

\[ \sum_{i=1; Y} n \mathbb{E} \left[ \frac{f'(F^{-1}(U_{(R_{ii})}^i))}{f(F^{-1}(U_{(R_{ii})}^i))} \right] = \sum_{i=1; Y} n \mathbb{E} \phi(U_{(R_{ii})}^i, f) = \sum_{i=1; Y} a_i(R_{ii}, f). \] (5.3)

It is convenient to replace the scores (4.6) by approximate scores \( \phi(EU_{(i)}^n, f) \). Then by applying that

\[ \mathbb{E} U_{(i)}^n = \frac{i}{n + 1}, \] (5.4)

we obtain an approximation for the test statistic (5.3) in the form

\[ \sum_{i=1; Y} \phi(EU_{(i)}^n, f) = \sum_{i=1; Y} \phi \left( \frac{R_{ii}}{i + 1}, f \right) = \sum_{i=1; Y} \left[ \frac{f'(F^{-1})}{f(F^{-1})} \left( \frac{R_{ii}}{i + 1} \right) \right]. \] (5.5)

Special cases include the two-sample Wilcoxon test

\[ W = \sum_{i=1; Y} \frac{R_{ii}}{i + 1} \] (5.6)

with the approximate scores corresponding to the logistic distribution and the median test

\[ \sum_{i=1; Y} \text{sgn} \left( R_{ii} - \frac{i + 1}{2} \right) \] (5.7)

derived with the approximate scores corresponding to the Laplace distribution.

6. Test of \( H_0 \) against regression in scale

Regression in scale

\[ H_1 : q^n_{\Delta}(x_1, \ldots, x_n) = \exp \left( -\Delta \sum_{i=1}^{n} c_i \right) \prod_{i=1}^{n} f[(x_i - \mu)e^{-\Delta c_i}], \quad \Delta > 0 \] (6.1)

is a special case of the general alternative (3.2) with known regression constants \( c_1, \ldots, c_n \) and an arbitrary shift parameter \( \mu \in \mathbb{R} \). The locally most powerful sequential rank test of \( H_0 \) against (6.1) follows from Theorem 3.1.
The score function $\varphi_1$ will be defined by means of the distribution function $F$ corresponding to the density $f$ in the form

$$\varphi_1(u,f) = -1 - F^{-1}(u) \frac{f'(F^{-1}(u))}{f(F^{-1}(u))}$$

(6.2)

and the corresponding scores are defined as

$$a_{1n}(i,f) = \mathbb{E}\varphi_1(U_{ni}^{(i)},f).$$

(6.3)

**Theorem 6.1.** The locally most powerful sequential rank test of $H_0$ against the system $\{q^{n}_{\Delta}, \Delta > 0\}$ with $q^{n}_{\Delta}$ defined by (6.1) at level $\alpha$ has the critical region

$$\sum_{i=1}^{n} c_i a_{1i}(R_{ii},f) \geq k_\alpha$$

(6.4)

assuming

$$\int_{-\infty}^{\infty} |xf'(x)|dx < \infty.$$  

(6.5)

**Proof.** Based on Theorem 3.1, the test statistic is a special case of (3.3), which can be expressed as

$$\sum_{i=1}^{n} c_i \mathbb{E}\left[ -1 - F^{-1}(U_{ni}^{(i)}) \frac{f'(F^{-1}(U_{ni}^{(i)}))}{f(F^{-1}(U_{ni}^{(i)}))} \right] = \sum_{i=1}^{n} c_i \mathbb{E}\varphi_1(U_{ni}^{(i)},f) = \sum_{i=1}^{n} c_i a_{1i}(R_{ii},f),$$

(6.6)

which concludes the proof.

By approximating the scores (6.3) by $\varphi_1(E U_{ni}^{(i)},f)$, we obtain an approximation to the test statistic of (6.4) in the form

$$\sum_{i=1}^{n} c_i \varphi_1(E U_{ni}^{(i)},f) = \sum_{i=1}^{n} c_i \varphi_1\left( \frac{R_{ii}}{i+1},f \right) = \sum_{i=1}^{n} c_i \left[ -1 - F^{-1}\left( \frac{R_{ii}}{i+1} \right) \frac{f'(F^{-1}(R_{ii}/(i+1)))}{f(F^{-1}(R_{ii}/(i+1)))} \right].$$

(6.7)

**7. Test of $H_0$ against two samples differing in scale**

Let $Z_1, Z_2, \ldots, Z_n$ denote a random sample, which is obtained as a pooled sample of two independent sequences $X_1, \ldots, X_m$ and $Y_1, \ldots, Y_{n-m}$. The joint density is assumed in the form

$$H_1 : q^{n}_{\Delta}(z_1, \ldots, z_n) = \exp(-m\Delta) \prod_{i=1}^{n} f[(z_{i} - \mu)e^{-\Delta}] \prod_{i=1}^{n} f[(z_{i} - \mu)], \quad \Delta > 0,$$

(7.1)

where the products are considered over the first and second sample, respectively. Here, $f$ is a known density and $\mu$ an arbitrary shift parameter.

The test statistic of the locally most powerful sequential rank test is expressed in Theorem 6.1 assuming (6.5) and is equal to

$$\sum_{i=1}^{n} a_{1i}(R_{ii},f).$$

(7.2)
An approximation based on (6.7) replaces the test statistic by
\[ \sum_{i=1}^{n} Y_i \left[ -1 - F^{-1} \left( \frac{R_{ii}}{i+1} \right) \frac{f'(F^{-1})}{f(F^{-1})} \left( \frac{R_{ii}}{i+1} \right) \right]. \] (7.3)

8. Test for the one-sample problem

We assume i.i.d. random variables \( X_1, \ldots, X_n \). Their sequential ranks will be denoted by \((R_{11}^+, \ldots, R_{nn}^+)\)^T, where \( R_{ii}^+ \) is the rank of \(|X_i|\) among \(|X_1|, \ldots, |X_i|\) for \( i = 1, \ldots, n \). The vector \( v = (v_1, \ldots, v_n)^T \) will stand for the vector of signs of \( X_1, \ldots, X_n \).

Both the null hypothesis
\[ H_0^+ : p(x_1, \ldots, x_n) = \prod_{i=1}^{n} f(x_i) \] (8.1) and the alternative
\[ H_1 : q_{\Delta}^n(x_1, \ldots, x_n) = \prod_{i=1}^{n} f(x_i - \Delta), \quad \Delta > 0, \] (8.2)
require \( f \) to be a known symmetric density satisfying \( f(x) = f(-x), x \in \mathbb{R} \).

The score function \( \varphi^+ \) will be defined as
\[ \varphi^+(u,f) = \varphi \left( \frac{1}{2} + \frac{1}{2} u, f \right) \] (8.3)
and the corresponding scores
\[ a_i^+(i,f) = E\varphi^+(U_{ij}^n, f). \] (8.4)

**Theorem 8.1.** The locally most powerful sequential rank test of \( H_0^+ \) against the system \( \{q^n_{\Delta}, \Delta > 0\} \) with \( q^n_{\Delta} \) defined by (8.2) at level \( \alpha \) has the critical region
\[ \sum_{i=1}^{n} v_i a_i^+(R_{ii}^+, f) \geq k_{\alpha} \] (8.5)
assumed (6.5).

**Proof.** Under \( H_0^+ \) it holds that
\[ P[\text{sgn } X = v, (R_{11}^+, \ldots, R_{nn}^+)^T = r^+] = \frac{1}{2^n n!} \] (8.6)
due to the independence between \( \text{sgn } X \) and \((R_{11}^+, \ldots, R_{nn}^+)\)^T. The Neyman-Pearson lemma allows to construct the most powerful test based on \( \text{sgn } X \) and \((R_{11}^+, \ldots, R_{nn}^+)\)^T as
\[ \frac{Q^n_{\Delta}[\text{sgn } X = v, (R_{11}^+, \ldots, R_{nn}^+)^T = r^+]}{P[\text{sgn } X = v, (R_{11}^+, \ldots, R_{nn}^+)^T = r^+] = 2^n n! Q^n_{\Delta}[\text{sgn } X = v, (R_{11}^+, \ldots, R_{nn}^+)^T = r^+],} \] (8.7)
where \( Q^*_\Delta \) denotes the probability distribution corresponding to the density (8.2). Following the steps of Hájek and Šidák (1967), we compute

\[
\lim_{\Delta \to 0} \frac{1}{\Delta} \left[ 2^n n! Q^*_\Delta \left\{ \text{sgn } X = v, (R^+_{11}, \ldots, R^+_{nn})^T = (r^+_1, \ldots, r^+_n)^T \right\} - 1 \right]
\]

\[
= \sum_{i=1}^n v_i E \left[ -\frac{f''(|X_i|)}{f(|X_i|)} \left| R^+_{11} = r^+_1, \ldots, R^+_{nm} = r^+_n \right| \right]
\]

\[
= \sum_{i=1}^n v_i E \left[ -\frac{f''(\{X^i_{(k)}\})}{f(\{X^i_{(k)}\})} \right] = \sum_{i=1}^n v_i E \varphi^+ (U^{(R^+_i)}_{(k)}, f) \quad (8.8)
\]

where \(|X^i_{(k)}|\) denotes the \(k\)th-order statistic among \(|X_1|, \ldots, |X_i|\). Therefore, the test statistic in (8.5) is for \( \Delta \to 0 \) equivalent to the most powerful sequential rank test, so (8.5) is the locally most powerful sequential rank test, which was to be proven.

The scores (8.4) can be approximated by \( \varphi^+ (E \cup_{(i)} U^n_{(R^+_i)}, f) \). Then the test statistic of (8.5) can be approximated by

\[
\sum_{i=1}^n v_i \varphi^+ (E \cup_{(R^+_i)}, f) = \sum_{i=1}^n v_i \varphi^+ \left( \frac{R^+_{ii}}{i + 1}, f \right) = \sum_{i=1}^n v_i \varphi \left( \frac{1}{2} + \frac{R_{ii}}{2(i + 1)}, f \right) \quad (8.9)
\]

### 9. Test of independence

We consider i.i.d. random variables \( X_1, \ldots, X_n \) and \( Y_1, \ldots, Y_n \). We denote the sequential ranks in the first and second samples by \( R^* = (R_{11}, \ldots, R_{nn})^T \) and \( Q^* = (Q_{11}, \ldots, Q_{nn})^T \), respectively.

The null hypothesis

\[
H_0^* : p(x_1, y_1, \ldots, x_n, y_n) = \prod_{i=1}^n f(x_i)g(y_i) \quad (9.1)
\]

will be tested against the alternative that \( X_i = X^*_i + \Delta Z_i \) and \( Y_i = Y^*_i + \Delta Z_i \), where \( X^*_i, Y^*_i, \) and \( Z_i \) are mutually independent, \( \Delta > 0 \), and \( X^*_i \) and \( Y^*_i \) have a specified distribution. This alternative can be formally expressed as

\[
H_1^* : q^n_{\Delta}(x_1, y_1, \ldots, x_n, y_n) = \prod_{i=1}^n h_{\Delta}(x_i, y_i), \quad \Delta > 0, \quad (9.2)
\]

where

\[
h_{\Delta}(x, y) = \int_{-\infty}^{\infty} f(x - \Delta z)g(y - \Delta z) dM(z) \quad (9.3)
\]

with an arbitrary distribution function \( M \). The score function (4.4) and scores (4.6) are used also in this context.

**Theorem 9.1.** The locally most powerful sequential rank test of \( H_0^* \) against the system \( \{ q^n_{\Delta}, \Delta > 0 \} \) with \( q^n_{\Delta} \) defined by (9.2) at level \( \alpha \) has the critical region

\[
\sum_{i=1}^n a_i(R_{ii}, f) a_i(Q_{ii}, g) \geq k_\alpha. \quad (9.4)
\]
Proof. Starting to express the most powerful sequential rank test based on the Neyman-Pearson lemma, we obtain

$$\lim_{\Delta \to 0} \frac{1}{\Delta^2} \left[ (n!)^2 Q^n (R_{11}, \ldots, R_{nn}, Q_{11}, \ldots, Q_{nn}) - 1 \right]$$

$$= \sigma^2 (n!)^2 \sum_{i=1}^{n} \int_{\mathbf{R}^* = \mathbf{r}^*} \int_{\mathbf{Q}^* = \mathbf{q}^*} \frac{f'(x_i) g'(y_i)}{f(x_i) g(y_i)} \prod_{k=1}^{n} f(x_k) g(y_k) \, dx_1 \cdots dx_n \, dy_1 \cdots dy_n, \quad (9.5)$$

where $\sigma^2$ is a constant corresponding to the arbitrary variables $Z_1, \ldots, Z_n$ and $\mathbf{r}^* = (r_{11}, \ldots, r_{nn})^T$ and $\mathbf{q}^* = (q_{11}, \ldots, q_{nn})^T$ are observed values of the sequential ranks of the two samples.

Following further steps of Hájek and Šidák (1967), we express (9.5) as

$$\sigma^2 \sum_{i=1}^{n} \left[ \frac{n! \int_{\mathbf{R}^* = \mathbf{r}^*} \int_{\mathbf{Q}^* = \mathbf{q}^*} f'(x_i) \prod_{k=1}^{n} f(x_k) \, dx_1 \cdots dx_n}{\frac{f'(x_i)}{f(x_i)}} \prod_{k=1}^{n} f(x_k) \, dx_1 \cdots dx_n \right] = \sigma^2 \sum_{i=1}^{n} \mathbb{E} \left[ \frac{f'(X_{(R_{i1})})}{f(X_{(R_{i1})})} \middle| R_{11}, \ldots, R_{nn} \right] \mathbb{E} \left[ \frac{g'(Y_{(Q_{i1})})}{g(Y_{(Q_{i1})})} \middle| Q_{11}, \ldots, Q_{nn} \right]$$

$$= \sigma^2 \sum_{i=1}^{n} \mathbb{E} \left[ \frac{-f'(F^{-1}(U_{(R_{i1})}))}{f(U_{(R_{i1})})} \middle| F^{-1}(U_{(R_{i1})}) \right] \mathbb{E} \left[ \frac{-g'(G^{-1}(U_{(Q_{i1})}))}{g(U_{(Q_{i1})})} \middle| G^{-1}(U_{(Q_{i1})}) \right], \quad (9.6)$$

where $F$ and $G$ are distribution functions corresponding to densities $f$ and $g$, respectively. This test statistic is further equivalent to

$$\sum_{i=1}^{n} \mathbb{E} \varphi(U_{(R_{i1})}, f) \mathbb{E} \varphi(U_{(Q_{i1})}, g), \quad (9.7)$$

which equals the test statistic of (9.4) and concludes the proof. \qed

As a special case we can express the test statistic of (9.4) for the linear score function $\varphi$; such a test statistic

$$\sum_{i=1}^{n} \frac{R_{ii} Q_{ii}}{(i+1)^2}, \quad (9.8)$$

can be interpreted as a sequential-rank analogy of Spearman correlation coefficient.

10. A sequential testing procedure

The tests investigated in the previous sections considered a fixed number of observations. In this section, we propose a sequential version of the two-sample test statistic $W (5.6)$ based on sequential ranks. Sequential tests for all situations in Sections 4–7 can be performed in an analogous, straightforward way.

We describe the sequential testing procedure on the example of the two-sample Wilcoxon test. In the whole section, we use the notation of Section 5 and consider $H_0 : \Delta = 0$ against the one-sided alternative $H_1 : \Delta < 0$. Because the test depends on the order of $X$s in the
pooled sample, we assume a fixed order \(X - Y - X - Y - X - Y \cdots\). In other words, \(Z_{2k-1} = X_k\) and \(Z_{2k} = Y_k\) for \(k = 1, 2, \ldots\).

The observations are obtained sequentially, and we require the maximal sample size not to exceed a given \(N_S\). Our approach allows rejecting \(H_0\) earlier compared to the test with the fixed size \(N_S\). The stopping time is a random variable required not to exceed a fixed integer \(N_S\). The computation is explained in Algorithm 10.1, using critical values of the following definition.

**Definition 10.1.** We consider a fixed level \(\alpha\) of the test and a fixed value of the maximal sample size \(N\). Let us consider an even \(n\). We define the critical value \(r_n^\alpha\) for \(n < N\) as the null critical value of the permutation test based on the test statistic

\[
\sum_{i=1}^{N_S} \frac{r_{ii}}{i+1}
\]

evaluated as

\[
\frac{R_{22}}{3} + \frac{R_{44}}{5} + \cdots + \frac{R_{nn}}{n+1} + \frac{n+2}{n+3} + \cdots + \frac{N_S}{N_S + 1}.
\]

In (10.2), the sequential ranks \(R_{22}, R_{44}, \ldots, R_{nn}\) are supplemented by the least favorable values for the sequential ranks

\[R_{22}, R_{44}, \ldots, R_{nn}, n + 2, \ldots, N_S.\]

This allows us to determine whether the test remains to reject \(H_0\) always for any really observed data \(Z_{n+1}, \ldots, Z_{N_S}\). The null critical value is found exact (or can approximated by simulations) for the exact number of observations equal to \(N_S\). In this way, the test is ensured to hold the probability of type I error on the selected level \(\alpha\).

The earliest possible stopping time for various values of \(N_S\) is shown in Figure 1 for the situation described at the beginning of this section. Let us explain how the figure was obtained.

---

**Algorithm 10.1.** Exact sequential two-sample test based on sequential ranks.

**Require:** Integer \(N_S\)

\(n := 0\)

**repeat**

\(n := n + 2\)

observe \(Z_{n-1}\) (from Xs) and \(Z_n\) (from Ys)

compute \(W\) (5.6)

compute \(E_W = n/4\)

compute the critical value \(r_n^\alpha\)

if \(W - E_W < r_n^\alpha\) then

reject \(H_0\) stop

**end if**

until \(n = N_S\)

accept \(H_0\)
on the particular value of $N_S = 20$. Let us say that the fixed number $n$ of observations has been already observed. First, we find the critical value of the one-sided test based on $W - EW$ for $n = 20$ to be equal to $-1.30$.

Let us now say that $n = 14$ observations have been already observed and the vector of sequential ranks of $Y$s is equal to the extreme case $1 - 1 - 1 - \cdots - 1$. Then, in the most extreme situation that observations 15 to 20 extend the vector of sequential ranks to $1 - 1 - 1 - \cdots - 1 - 16 - 18 - 20$, we will have $W = 3.86$ and $W - EW = -1.14$ is not significant. Later, if $n = 16$, we obtain $W = 2.98$ and $EW = -2.02$, thus yielding already a significant result. Thus, this most extreme situation leads us to the conclusion that $n = 16$ is the smallest possible stopping time for a given $N_S = 20$.

11. Discussion

Previous research considered such test statistics based on sequential ranks, which were constructed in an intuitive way without investigating their optimality. As we document it in Section 1, the optimal tests have been unknown and this article fills exactly this gap of the optimality results for tests based on sequential ranks.

To be specific, we derive the locally most powerful tests among all tests based on sequential ranks for a variety of situations. A general version of the test of Section 3 is formulated for two important special cases, namely, regression in location (Section 4) and regression in scale (Section 6). These results are further formulated for the special case of two-sample tests, namely, for location (Section 5) and scale (Section 7). Further tests, not using the general result of Section 3, include the proof of the locally most powerful one-sample test (Section 8) and test of independence (Section 9) based on sequential ranks.

Our main result can be summarized as follows. Classical locally most powerful rank statistics are known to be obtained as projections of the locally most powerful (parametric) statistics into the space of linear rank statistics. This does not hold for sequential rank statistics.
Thus, the current article brings arguments in favor of some of the previously used test statistic based on sequential ranks. Particularly, we prove the linear sequential rank statistic of Lombard and Mason (1985) to be locally most powerful among all tests based on sequential ranks in Section 4. On the other hand, a previous statistic of Mason (1981) uses a different standardization of the regression constants. In Section 10, we consider the Wilcoxon test based on sequential ranks, which is again equal to the previous Wilcoxon test obtained as a special case of Lombard and Mason (1985). These considerations are however derived for a fixed sample size.

If the sample size is fixed, sequential ranks are fully equivalent to classical ranks. In addition, there is no advantage in computing sequential ranks in terms of the computational complexity. The computation of \( R_{11}, \ldots, R_{nn} \) has the complexity of order of \( n \log n \), which is the same as that of \( R_1, \ldots, R_n \).

Let us recall properties of the test statistics based on sequential ranks for a fixed sample size. Though previous works considered sequential ranks as a tool allowing to derive theoretical properties of nonparametric hypothesis tests based on classical (nonsequential) ranks in a standard (nonsequential) setting, we have the impression that papers investigating sequential ranks ignored these properties as well as the unique relationship between classical and sequential ranks (Algorithms 2.1 and 2.2). Therefore, we find it worth mentioning that some of the properties may be suitable only in some particular situations. In any case, we understand the sequential ranks as a tool principally different from classical ranks requiring a different (“sequential”) way of thinking and interpreting the results. We will now discuss the properties of the test statistics based on sequential ranks and for the sake of clarity; some are explained based on the example of the two-sample Wilcoxon test statistic \( W \) (5.6):

- Locally most powerful property among all tests based on sequential ranks.
- A simplified theoretical reasoning due to independence of \( R_{11}, \ldots, R_{nn} \) (Mason, 1981).
- The computation has the same computational complexity as analogous standard tests based on classical ranks.
- \( W \) depends on the order of \( Y \)s in the pooled sample.
- \( W \) depends on the order of observations even within each of the two individual samples.
- Though \( EW \) does not not depend on the order of \( X \)s in the pooled sample,

\[
\text{var } W = \text{var } \sum_{i=1; Y}^{n} \frac{R_{ii}}{i+1} = \sum_{i=1; Y}^{n} \frac{\text{var } R_{ii}}{i+1}
\]  

(11.1)

depends on this design of the study.

In Section 10, we present a straightforward way how to use the optimal test statistics in a sequential testing. Our approach considers a random stopping time that is not allowed to exceed a given upper bound \( N_S \). The test statistics keeps the properties as described above for a fixed sample size. In addition, we find these properties specific for the sequential approach important:

- It is especially suitable if acquiring new observations is expensive.
- It is able to come to the conclusion more quickly compared to a test based on classical ranks, because the weights of the first observations are larger compared to later ones.
- It is computationally more complicated than classical approaches.

We may conclude that the advantages of the sequential rank test are revealed in the sequential testing procedure, which makes the approach favorable compared to testing with a fixed sample size.
Acknowledgments

The author would like to thank Jana Jurečková, the anonymous referees, an associate editor, and the editor-in-chief for their time and constructive advice.

Funding

This work was supported by the Neuron Fund for Support of Science and by the Czech Science Foundation project number 17-07384S.

References


