A monotonicity property of weighted log-rank tests

Tahani Coolen-Maturi*, Frank P.A. Coolen

Department of Mathematical Sciences, Durham University, Durham, DH1 3LE, UK

Abstract

The logrank test is a well-known nonparametric test which is often used to compare the survival distributions of two samples including right-censored observations, it is also known as the Mantel-Haenszel test. The G^ρ family of tests, introduced by Harrington and Fleming (1982), generalizes the logrank test by using weights assigned to observations. In this paper, we present a switch monotonicity property for the G^ρ family of tests, which was motivated by the need to derive bounds for the test statistic in case of imprecise data observations. This property states that, when all observations from two independent groups are ranked together, the value of the z-test statistic is monotonically increasing after switching a pair of adjacent values from the two groups. Two examples are provided to motivate and illustrate the result presented in this paper.

Keywords: Logrank, Monotonicity, Imprecise probability, survival distribution, monotonicity property.

---

1. Introduction

The logrank test is a well-known nonparametric test which is often used to compare the survival distributions of two groups containing right-censored observations. It generalizes the Wilcoxon test, for data without right-censored observations, and is also known as Mantel-Haenszel test (Mantel, 1966). Several variations to this test have been introduced in the literature, e.g. Gehan's Generalized Wilcoxon test (Gehan, 1965; Lou and Lan, 1998), Weighted Logrank tests (Latta, 1977) and Wilcoxon-Peto test (Peto and Peto, 1972). The Mantel-Haenszel test (Mantel, 1966) gives equal weights to observations regardless of the time at which an event occurs. On the other hand, the Wilcoxon-Peto test statistic assigns more weights to earlier event times (Peto and Peto, 1972). Harrington and Fleming (1982) introduced a class of tests, the G^ρ family, which can be used to test the null hypothesis $H_0 : S_0(t) = S_1(t)$ for all $t > 0$ against the alternative hypothesis $H_1 : S_0(t) \neq S_1(t)$ for some $t > 0$.

---

*Corresponding author

Email addresses: `tahani.maturi@durham.ac.uk` (Tahani Coolen-Maturi), `frank.coolen@durham.ac.uk` (Frank P.A. Coolen)

In this paper, we consider the G^ρ family of tests for right-censored data introduced by Harrington and Fleming (1982), in which the weight assigned to each observed failure time t is of the form $[\hat{S}(t)]^\rho$ for fixed $\rho \geq 0$, where $\hat{S}(t)$ is the well-known Kaplan-Meier estimate of the survival function. Note that the use of 'failure time' does not restrict the test applications and could be interpreted as time of any event of interest, as long as each individual (or 'item') has only one event associated with it, which may either be observed (failure time) or only known to be greater than an observed right censoring time. Throughout this paper it is assumed that right censoring is non-informative, which means that the residual lifetime of the censored observation is independent of the censoring process. As special cases, $\rho = 0$ gives the log-rank Mantel-Haenszel test (Mantel, 1966) and $\rho = 1$ gives the Peto-Prentice extension of the Wilcoxon statistic (Peto and Peto, 1972; Prentice, 1978). Several R packages are available to perform these tests, e.g. the function `survdiff` within the **survival** package (Therneau, 2015) and the comprehensive **FHtest** package (Oller and Langohr, 2017).

In this paper we prove a monotonicity property of the G^ρ class family of tests for right-censored data introduced by Harrington and Fleming (1982). Formally, a function f is called monotonically non-decreasing if it preserves the order, that is if for all a and b with $a \leq b$ we have $f(a) \leq f(b)$. This research was motivated by possible applications of such tests in case of imprecise data (Augustin et al, 2014; Coolen, Ahmadini and Coolen-Maturi, 2021), where the ordering of observations per group is known but where the ranking of observations between the groups may not be precisely determined due to lack of precise values for some or all of the observations, it is most natural to assume that each observation is only known to belong to an interval. In such cases, when intervals are overlapping, different combined rankings of the data from different groups may be possible and one typically would like to find the minimum and maximum values of the test statistic corresponding to all possible combined rankings. The result in this paper makes the derivation of these minimum and maximum values straightforward, it has been applied to develop robust statistical inference for accelerated life testing by Coolen, Ahmadini and Coolen-Maturi (2021).

This paper is organised as follows: Section 2 introduces the notation and the setting, while the main results are presented in Section 3. Two examples to motivate and illustrate the result presented in this paper are provided in Section 4. The paper ends with concluding remarks in Section 5.

2. Notation and Setting

Let $\tau_1 < \tau_2 < \ldots < \tau_k$ denote k times of observed failures. Let $Y_i(\tau_j)$ be the number of individuals in group i who are at risk at τ_j $(i = 0, 1)$, i.e. the number of individuals from both groups at risk at τ_j is $Y(\tau_j) = Y_0(\tau_j) + Y_1(\tau_j)$, $j = 1, 2, \ldots, k$. Let d_{ij} be the number of individuals in group i who fail at τ_j $(i = 0, 1)$, so the total number of failures at τ_j from both groups is $d_j = d_{0j} + d_{1j}$, $j = 1, 2, \ldots, k$. The information at time τ_j can be summarised in the following 2×2 table:

	Fail at τ_j	Do not fail	at risk at τ_j
Group 0	d_{0j}	$Y_0(\tau_j) - d_{0j}$	$Y_0(\tau_j)$
Group 1	d_{1j}	$Y_1(\tau_j) - d_{1j}$	$Y_1(\tau_j)$
	d_j	$Y(\tau_j) - d_j$	$Y(\tau_j)$

Consider the test statistic

$$Z = \frac{O - E}{\sqrt{V}} = \frac{\sum_j O_j - \sum_j E_j}{\sqrt{\sum_j V_j}} \tag{1}$$

with

$$O_j = \left[\hat{S}(\tau_j)\right]^{\rho} d_{1j} \tag{2}$$

$$E_j = \left[\hat{S}(\tau_j)\right]^{\rho} \left(\frac{Y_1(\tau_j)}{Y(\tau_j)}\right) d_j \tag{3}$$

$$V_j = \left[\hat{S}(\tau_j)\right]^{2\rho} \frac{Y_0(\tau_j)Y_1(\tau_j)}{(Y(\tau_j))^2} \left(\frac{Y(\tau_j) - d_j}{Y(\tau_j) - 1}\right) d_j \tag{4}$$

where $\rho \geq 0$ and $\hat{S}(\tau_j)$ is the Kaplan-Meier estimate at time τ_j (Kaplan and Meier, 1958). Then under the null hypothesis $H_0 : S_0(t) = S_1(t)$ for all $t > 0$, the test statistic Z follows the standard normal distribution, i.e. $Z \sim N(0, 1)$, so $Z^2 \sim \chi^2_1$.

For simplicity of notation, we assume throughout this paper that there are no ties, therefore $d_j = d_{0j} + d_{1j} = 1$, and $O = \sum_j \left[\hat{S}(\tau_j)\right]^{\rho} d_{1j}$ is the weighted number of failures from group G_1. The expected value formula (3) and the variance formula (4) can be simplified (as $d_j = 1$) as

$$E_j = \left[\hat{S}(\tau_j)\right]^{\rho} \frac{Y_1(\tau_j)}{Y(\tau_j)} \tag{5}$$

$$V_j = \left[\hat{S}(\tau_j)\right]^{2\rho} \frac{Y_0(\tau_j)Y_1(\tau_j)}{(Y(\tau_j))^2} \tag{6}$$

Now let $\tau_{k_j-1} \leq x_{j_i} < y_j \leq \tau_{k_j}$, and let u_0 (u_1) be the number of censored observations from group G_0 (G_1) between τ_{k_j-1} and τ_{k_j}, thus $u = u_0 + u_1$. Define $\delta_i(\tau_j)$ to be equal to 1 if τ_j is a failure from group G_i, and zero otherwise, $i = 0, 1$. Let $Y_i(\tau_{k_j})$ be the number of individuals in group G_i who are at risk at τ_{k_j}, $i = 0, 1$, and let $Y(\tau_{k_j}) = Y_0(\tau_{k_j}) + Y_1(\tau_{k_j})$ be the number of individuals from both groups at risk at τ_{k_j}, $k_j \in \{1, 2, \ldots, k\}$. This is illustrated in the first row of Figure 1. The next section introduces the main results of this paper.

3. Main Results

In this section, we consider the following setting. For a particular data set, with fixed failure-censored status, suppose that all observations from group G_0, $x_1 < x_2 < \ldots < x_{n_0}$, precede all observations from group G_1, $y_1 < y_2 < \ldots < y_{n_1}$. We would like to swap between neighbouring data observations, one pair of a G_0 observation and a G_1 observation at a time, where the latter is the smallest G_1 observation greater than the G_0 observation, until we have all observations from group G_1 preceding all observations from group G_0. In total we can do that in $n_0 n_1$ steps (switches), where n_0 and n_1 are the sample sizes of group G_0 and group G_1, respectively. For example, if we have 3 observations from each group, the number of switches from $x_1 < x_2 < x_3 < y_1 < y_2 < y_3$ to $y_1 < y_2 < y_3 < x_1 < x_2 < x_3$ is 9. The property presented in this paper is that, under the null hypothesis, the z-test statistic behaves monotonically throughout this switching process. This is stated in the following theorem.

Theorem 3.1 (Switch monotonicity). *Suppose we have data observations from two independent groups, their ordered values are denoted by $x_1 < x_2 < \ldots < x_{n_0}$ and $y_1 < y_2 < \ldots < y_{n_1}$. Let Z_B be the z-test statistic value, obtained from* (1), *corresponding to these data sets and Z_A be the value of the z-statistic after a switch of two adjacent values $x_i < y_j$, with all observations ranked together, then $Z_B \leq Z_A$.*

Note that for the special case where there are no right-censored observations, this test is the same as the Wilcoxon rank-sum test, for which this theorem trivially hold as the sum of the ranks for one group clearly changes monotonically with such switches. The remainder of this section consists of the proof of this theorem. To this end, suppose we swap x_{j_i} and y_j, that is now $x_{j_i} > y_j$, then we have four different scenarios we need to consider:

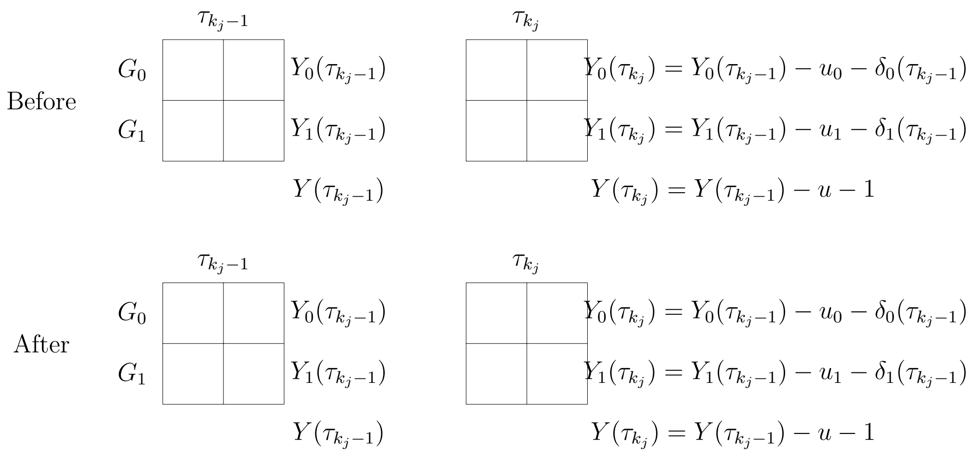


Figure 1: Setting and Scenario S1

Scenario 1 (S1): when both x_{j_i} and y_j are censoring times
In this case, nothing will change to the 2×2 tables in Figure 1, where the first row is corresponding to *before* the swap and the second row to *after* the swap. As the value of $\hat{S}(\tau_j)$ is a step function that change value only at the time of observed failure, therefore the expected

value and the variance formula are the same before and after the swap. That is if we swap any two censored observations between τ_{k_j-1} and τ_{k_j} this will not affect the expected value and the variance, as it does not affect the margins in the 2×2 tables in Figure 1. Thus $Z_B = \frac{(O-E_B)}{\sqrt{V_B}}$ and $Z_A = \frac{(O-E_A)}{\sqrt{V_A}}$ are equal, $Z_A = Z_B$, where we use B as subscript for the case before the swap and A for the case after the swap.

The proofs for the next three cases are very similar, yet for the sake of completeness full details are given.

Scenario 2 (S2): when x_{j_i} is a failure time and y_j is a censoring time

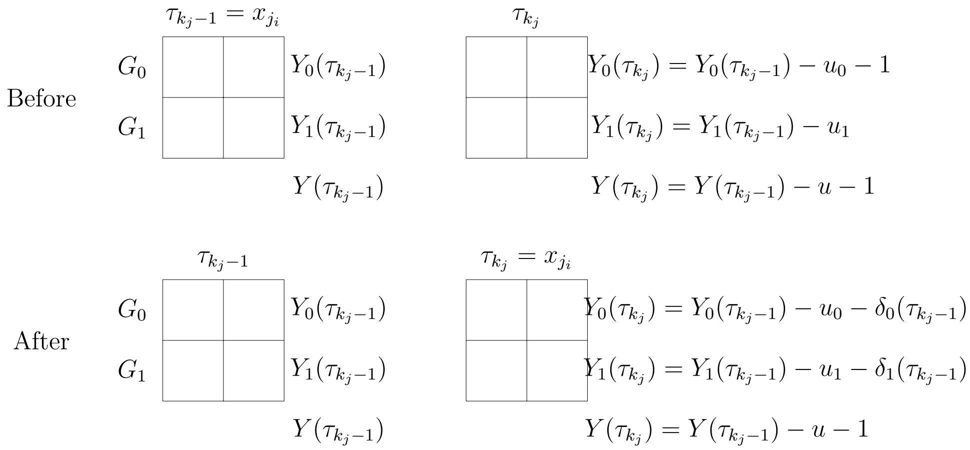


Figure 2: Scenario S2

This second scenario is illustrated in Figure 2, and the expect values for before and after the swap are given as

$$E_B = \Sigma_{j \neq k_j} E_j + \left[\hat{S}(\tau_j)\right]^{\rho} \frac{Y_1(\tau_{k_j-1}) - u_1}{Y(\tau_{k_j-1}) - u - 1}$$

$$E_A = \Sigma_{j \neq k_j} E_j + \left[\hat{S}(\tau_j)\right]^{\rho} \frac{Y_1(\tau_{k_j-1}) - u_1 - \delta_1(\tau_{k_j-1})}{Y(\tau_{k_j-1}) - u - 1}$$

Clearly $E_B \geq E_A$ thus $(O - E_B) \leq (O - E_A)$. And the variances

$$V_B = \Sigma_{j \neq k_j} V_j + \left[\hat{S}(\tau_j)\right]^{2\rho} \frac{(Y_0(\tau_{k_j-1}) - u_0 - 1)(Y_1(\tau_{k_j-1}) - u_1)}{(Y(\tau_{k_j-1}) - u - 1)^2}$$

$$V_A = \Sigma_{j \neq k_j} V_j + \left[\hat{S}(\tau_j)\right]^{2\rho} \frac{(Y_0(\tau_{k_j-1}) - u_0 - \delta_0(\tau_{k_j-1}))(Y_1(\tau_{k_j-1}) - u_1 - \delta_1(\tau_{k_j-1}))}{(Y(\tau_{k_j-1}) - u - 1)^2}$$

We have two main cases:

(i) If $\delta_0(\tau_{k_j-1}) = 1$, that is τ_{k_j-1} is a failure from group G_0 and thus $\delta_1(\tau_{k_j-1}) = 0$, then $V_B = V_A$. Thus $Z_B \leq Z_A$, as from above $(O - E_B) \leq (O - E_A)$.

(ii) If $\delta_1(\tau_{k_j-1}) = 1$, that is τ_{k_j-1} is a failure from group G_1 and thus $\delta_0(\tau_{k_j-1}) = 0$, then

$$V_A = \Sigma_{j \neq k_j} V_j + \left[\hat{S}(\tau_j)\right]^{2\rho} \frac{(Y_0(\tau_{k_j-1}) - u_0)(Y_1(\tau_{k_j-1}) - u_1 - 1)}{(Y(\tau_{k_j-1}) - u - 1)^2}$$

and thus we have two sub-cases:

(iia) If $Y_1(\tau_{k_j-1}) - u_1 > Y_0(\tau_{k_j-1}) - u_0$ then $V_B < V_A$, i.e. $\frac{1}{V_B} > \frac{1}{V_A}$.

- If $(O - E_B) > 0$, then $(O - E_A)$ also has to be positive.
We multiply $(O - E_B) \leq (O - E_A)$ by $\frac{1}{\sqrt{V_B}}$ and by $\frac{1}{\sqrt{V_A}}$, then we have

$$\frac{(O - E_B)}{\sqrt{V_A}} \leq \frac{(O - E_A)}{\sqrt{V_A}}$$
$$\frac{(O - E_B)}{\sqrt{V_B}} \leq \frac{(O - E_A)}{\sqrt{V_B}}$$

Now we multiply $\frac{1}{\sqrt{V_A}} < \frac{1}{\sqrt{V_B}}$ by $(O - E_A)$ and by $(O - E_B)$, then we have

$$\frac{(O - E_A)}{\sqrt{V_A}} < \frac{(O - E_A)}{\sqrt{V_B}}$$
$$\frac{(O - E_B)}{\sqrt{V_A}} < \frac{(O - E_B)}{\sqrt{V_B}}$$

thus we obtain the following inequalities

$$\frac{(O - E_B)}{\sqrt{V_A}} \leq \frac{(O - E_A)}{\sqrt{V_A}} < \frac{(O - E_A)}{\sqrt{V_B}}$$
$$\frac{(O - E_B)}{\sqrt{V_A}} < \frac{(O - E_B)}{\sqrt{V_B}} \leq \frac{(O - E_A)}{\sqrt{V_B}}$$

Then the proof follows the same argument given in the appendix, and indeed $Z_B \leq Z_A$.

- If $(O - E_B) < 0$, then

$$\frac{1}{\sqrt{V_B}} > \frac{1}{\sqrt{V_A}}$$
$$\frac{(O - E_B)}{\sqrt{V_B}} < \frac{(O - E_B)}{\sqrt{V_A}} \leq \frac{(O - E_A)}{\sqrt{V_A}}$$

thus $Z_B \leq Z_A$.

(iib) If $Y_1(\tau_{k_j-1}) - u_1 < Y_0(\tau_{k_j-1}) - u_0$ then $V_B > V_A$, i.e. $\frac{1}{V_B} < \frac{1}{V_A}$.

- If $(O - E_B) > 0$ then $(O - E_A)$ also has to be positive.
  In this case, we multiply $(O - E_B) \le (O - E_A)$ by $\frac{1}{\sqrt{V_B}} < \frac{1}{\sqrt{V_A}}$, to obtain that $Z_B \le Z_A$.

- If $(O - E_A) > 0$, then $\frac{(O-E_A)}{\sqrt{V_B}} < \frac{(O-E_A)}{\sqrt{V_A}}$,
  and if we divide $(O-E_B) \le (O-E_A)$ by $\sqrt{V_B}$ then we have $\frac{(O-E_B)}{\sqrt{V_B}} \le \frac{(O-E_A)}{\sqrt{V_B}}$ and therefore we have $Z_B \le Z_A$. Note this includes the case when $(O - E_B)$ is negative but $(O - E_A)$ is positive.

- If both $(O - E_B)$ and $(O - E_A)$ are negative

  We multiply $(O - E_B) \le (O - E_A)$ by $\frac{1}{\sqrt{V_B}}$ and by $\frac{1}{\sqrt{V_A}}$ we have

$$\frac{(O - E_B)}{\sqrt{V_A}} \le \frac{(O - E_A)}{\sqrt{V_A}}$$
$$\frac{(O - E_B)}{\sqrt{V_B}} \le \frac{(O - E_A)}{\sqrt{V_B}}$$

  Now we multiply $\frac{1}{\sqrt{V_A}} > \frac{1}{\sqrt{V_B}}$ by $(O - E_A)$ and by $(O - E_B)$ we have

$$\frac{(O - E_A)}{\sqrt{V_A}} < \frac{(O - E_A)}{\sqrt{V_B}}$$
$$\frac{(O - E_B)}{\sqrt{V_A}} < \frac{(O - E_B)}{\sqrt{V_B}}$$

  thus we obtain the following inequalities

$$\frac{(O - E_B)}{\sqrt{V_A}} \le \frac{(O - E_A)}{\sqrt{V_A}} < \frac{(O - E_A)}{\sqrt{V_B}}$$
$$\frac{(O - E_B)}{\sqrt{V_A}} < \frac{(O - E_B)}{\sqrt{V_B}} \le \frac{(O - E_A)}{\sqrt{V_B}}$$

  Then the proof follows the same argument given in the appendix, and indeed $Z_B \le Z_A$.

Scenario 3 (S3): when x_{j_i} is a censoring time and y_j is a failure time
This third scenario is illustrated in Figure 3.

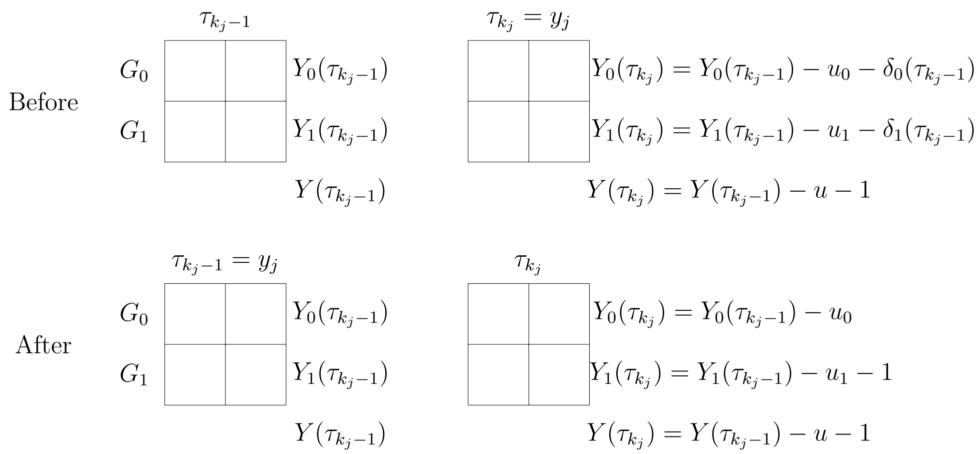


Figure 3: Scenario S3

Similarly we calculate the expected value and the variance before and after the swap as follows:

$$E_B = \Sigma_{j \neq k_j} E_j + \left[\hat{S}(\tau_j)\right]^{\rho} \frac{Y_1(\tau_{k_j-1}) - u_1 - \delta_1(\tau_{k_j-1})}{Y(\tau_{k_j-1}) - u - 1}$$

$$E_A = \Sigma_{j \neq k_j} E_j + \left[\hat{S}(\tau_j)\right]^{\rho} \frac{Y_1(\tau_{k_j-1}) - u_1 - 1}{Y(\tau_{k_j-1}) - u - 1}$$

Clearly $E_B \geq E_A$, thus $(O - E_B) \leq (O - E_A)$. And the variances are

$$V_B = \Sigma_{j \neq k_j} V_j + \left[\hat{S}(\tau_j)\right]^{2\rho} \frac{(Y_0(\tau_{k_j-1}) - u_0 - \delta_0(\tau_{k_j-1}))(Y_1(\tau_{k_j-1}) - u_1 - \delta_1(\tau_{k_j-1}))}{(Y(\tau_{k_j-1}) - u - 1)^2}$$

$$V_A = \Sigma_{j \neq k_j} V_j + \left[\hat{S}(\tau_j)\right]^{2\rho} \frac{(Y_0(\tau_{k_j-1}) - u_0)(Y_1(\tau_{k_j-1}) - u_1 - 1)}{(Y(\tau_{k_j-1}) - u - 1)^2}$$

(i) If $\delta_0(\tau_{k_j-1}) = 0$ then $\delta_1(\tau_{k_j-1}) = 1$ and $V_B = V_A$. Thus $Z_B \leq Z_A$, as from above $(O - E_B) \leq (O - E_A)$.

(ii) If $\delta_0(\tau_{k_j-1}) = 1$ then $\delta_1(\tau_{k_j-1}) = 0$,

$$V_B = \Sigma_{j \neq k_j} V_j + \left[\hat{S}(\tau_j)\right]^{2\rho} \frac{(Y_0(\tau_{k_j-1}) - u_0 - 1)(Y_1(\tau_{k_j-1}) - u_1)}{(Y(\tau_{k_j-1}) - u - 1)^2}$$

then we have two sub-cases:

(iia) if $Y_1(\tau_{k_j-1}) - u_1 > Y_0(\tau_{k_j-1}) - u_0$ then $V_B < V_A$, (iib) if $Y_1(\tau_{k_j-1}) - u_1 < Y_0(\tau_{k_j-1}) - u_0$ then $V_B > V_A$.

The proof for both cases (iia) and (iib) are similar to scenario S2.

Scenario 4 (S4): when both x_{j_i} and y_j are failure times
This final scenario is illustrated in Figure 4.

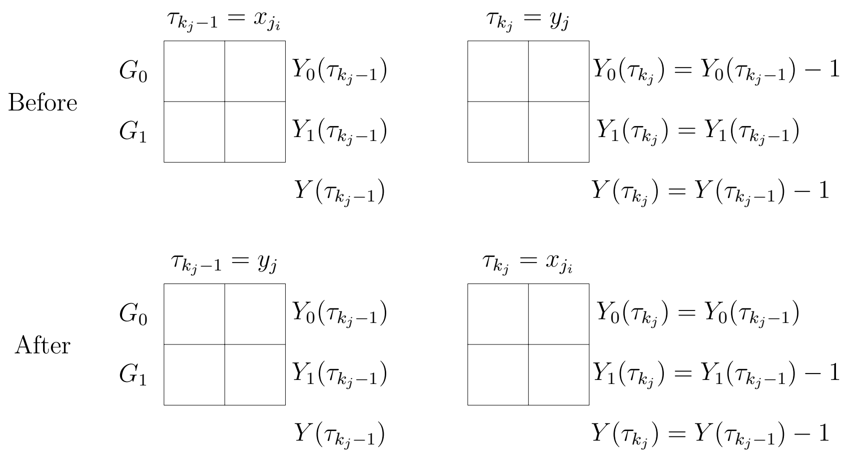


Figure 4: Scenario S4

We calculate the expected value before and after the swap as follows:

$$E_B = \Sigma_{j \neq k_j} E_j + \left[\hat{S}(\tau_j)\right]^{\rho} \frac{Y_1(\tau_{k_j-1})}{Y(\tau_{k_j-1}) - 1}$$

$$E_A = \Sigma_{j \neq k_j} E_j + \left[\hat{S}(\tau_j)\right]^{\rho} \frac{Y_1(\tau_{k_j-1}) - 1}{Y(\tau_{k_j-1}) - 1}$$

Clearly $E_B \geq E_A$, thus $(O - E_B) \leq (O - E_A)$. And the variances are

$$V_B = \Sigma_{j \neq k_j} V_j + \left[\hat{S}(\tau_j)\right]^{2\rho} \frac{(Y_0(\tau_{k_j-1}) - 1)Y_1(\tau_{k_j-1})}{(Y(\tau_{k_j-1}) - 1)^2}$$

$$V_A = \Sigma_{j \neq k_j} V_j + \left[\hat{S}(\tau_j)\right]^{2\rho} \frac{Y_0(\tau_{k_j-1})(Y_1(\tau_{k_j-1}) - 1)}{(Y(\tau_{k_j-1}) - 1)^2}$$

(i) If $Y_0(\tau_{k_j-1}) = Y_1(\tau_{k_j-1})$ then $V_B = V_A$. Thus $Z_B \leq Z_A$, as from above $(O - E_B) \leq (O - E_A)$.

(ii) As we know that both x_{j_i} and y_j are failure times, we have the two-sub cases:

(iia) If $Y_1(\tau_{k_j-1}) > Y_0(\tau_{k_j-1})$ then $V_B < V_A$, thus $\frac{1}{V_B} > \frac{1}{V_A}$

- If $(O - E_B) > 0$, then $(O - E_A)$ also has to be positive, we obtain similarly that

$$\frac{(O-E_B)}{\sqrt{V_A}} \leq \frac{(O-E_A)}{\sqrt{V_A}} < \frac{(O-E_A)}{\sqrt{V_B}}$$
$$\frac{(O-E_B)}{\sqrt{V_A}} < \frac{(O-E_B)}{\sqrt{V_B}} \leq \frac{(O-E_A)}{\sqrt{V_B}}$$

  Then the proof follows the same argument given in the appendix, and indeed $Z_B \leq Z_A$.

- If $(O - E_B) < 0$, then

$$\frac{1}{\sqrt{V_B}} > \frac{1}{\sqrt{V_A}}$$
$$\frac{(O-E_B)}{\sqrt{V_B}} < \frac{(O-E_B)}{\sqrt{V_A}} \leq \frac{(O-E_A)}{\sqrt{V_A}}$$

  then $Z_B \leq Z_A$.

(iib) if $Y_1(\tau_{k_j-1}) < Y_0(\tau_{k_j-1})$ then $V_B > V_A$, thus $\frac{1}{V_B} < \frac{1}{V_A}$.

- If $(O - E_B) > 0$ then $(O - E_A)$ also has to be positive.
  In this case, we can multiply $(O - E_B) \leq (O - E_A)$ by $\frac{1}{\sqrt{V_B}} < \frac{1}{\sqrt{V_A}}$, to obtain that $Z_B \leq Z_A$.

- If $(O - E_A) > 0$, then $\frac{(O-E_A)}{\sqrt{V_B}} < \frac{(O-E_A)}{\sqrt{V_A}}$,
  and if we divide $(O - E_B) \leq (O - E_A)$ by $\sqrt{V_B}$ then we have $\frac{(O-E_B)}{\sqrt{V_B}} \leq \frac{(O-E_A)}{\sqrt{V_B}}$ thus we obtain that $Z_B \leq Z_A$. Note this include the case when $(O - E_B)$ is negative but $(O - E_A)$ is positive.

- If both $(O - E_B)$ and $(O - E_A)$ are negative, we obtain similarly that

$$\frac{(O-E_B)}{\sqrt{V_A}} \leq \frac{(O-E_A)}{\sqrt{V_A}} < \frac{(O-E_A)}{\sqrt{V_B}}$$
$$\frac{(O-E_B)}{\sqrt{V_A}} < \frac{(O-E_B)}{\sqrt{V_B}} \leq \frac{(O-E_A)}{\sqrt{V_B}}$$

  Then the proof follows the same argument given in the appendix, and indeed $Z_B \leq Z_A$.

4. Examples

In this section, two examples are presented to motivate the theory presented in this paper and to illustrate the application of the main result.

Example 1

Suppose that there are two groups with 5 observations each, where all observations from group G_0 are smaller than all observations from group G_1. Let the censoring status for G_0 be (1, 0, 1, 1, 0) and for G_1 be (1, 1, 0, 1, 0). Because only the ranks of the observations play a role in this paper, we denote the observations by there initial ranks, adding superscript + for right-censored observations, so observations for group G_0 are denoted by 1, 2^+, 3, 4, 5^+ and for group G_1 by 6, 7, 8^+, 9, 10^+. Setting $\rho = 0.5$, the z-test value for this initial data case, obtained using Equation (1), is -2.1901, it is given in the first row of Table 1. In the following rows, 25 switches of neighbouring pairs in the ordering are presented with corresponding z-test values, note that the observations remain indicated by their initial ranks. These switches are such that, at each switch, one G_1 observation becomes smaller than one G_0 observation. For example, the second row in this table, indicated by switch 1 in the first column, presents the case where the largest G_0 observation, 5^+, and the smallest G_1 observation, 6, have swapped in the overall ranking. After 25 switches, the total reversal of the observations of the two groups has been achieved, with all G_1 observations smaller than all G_0 observations. Of course, different specific pairwise switches could have been chosen to get from the initial ranking to this final ranking, all such possibilities lead to similarly monotonically changing z-test values.

Figure 5 shows the z-test values for all 25 switches in Table 1, for different values of ρ, $\rho = \{0, 0.1, 0.2, \ldots, 1\}$. The different colour lines in this figure are in the order of the different values of ρ, with the lowest line corresponding to $\rho = 0$ and the highest line corresponding to $\rho = 1$. Clearly, the z-test values are in ascending order regardless of the values of ρ, this illustrates the monotonicity property presented in this paper. Table 1 and Figure 5 also show that the z-test value does not change when a switching occurs between two right-censored observations, which happens at switches 11, 14, 21 and 24.

switch											z-test
-	1	2^+	3	4	5^+	6	7	8^+	9	10^+	−2.1901
1	1	2^+	3	4	6	5^+	7	8^+	9	10^+	−1.8797
2	1	2^+	3	6	4	5^+	7	8^+	9	10^+	−1.5791
3	1	2^+	6	3	4	5^+	7	8^+	9	10^+	−1.3374
4	1	6	2^+	3	4	5^+	7	8^+	9	10^+	−1.2602
5	6	1	2^+	3	4	5^+	7	8^+	9	10^+	−1.0872
6	6	1	2^+	3	4	7	5^+	8^+	9	10^+	−0.8729
7	6	1	2^+	3	7	4	5^+	8^+	9	10^+	−0.6236
8	6	1	2^+	7	3	4	5^+	8^+	9	10^+	−0.4107
9	6	1	7	2^+	3	4	5^+	8^+	9	10^+	−0.3533
10	6	7	1	2^+	3	4	5^+	8^+	9	10^+	−0.1873
11	6	7	1	2^+	3	4	8^+	5^+	9	10^+	−0.1873
12	6	7	1	2^+	3	8^+	4	5^+	9	10^+	−0.1096
13	6	7	1	2^+	8^+	3	4	5^+	9	10^+	−0.0175
14	6	7	1	8^+	2^+	3	4	5^+	9	10^+	−0.0175
15	6	7	8^+	1	2^+	3	4	5^+	9	10^+	0.0722
16	6	7	8^+	1	2^+	3	4	9	5^+	10^+	0.2798
17	6	7	8^+	1	2^+	3	9	4	5^+	10^+	0.5884
18	6	7	8^+	1	2^+	9	3	4	5^+	10^+	0.8718
19	6	7	8^+	1	9	2^+	3	4	5^+	10^+	0.9208
20	6	7	8^+	9	1	2^+	3	4	5^+	10^+	1.1602
21	6	7	8^+	9	1	2^+	3	4	10^+	5^+	1.1602
22	6	7	8^+	9	1	2^+	3	10^+	4	5^+	1.4627
23	6	7	8^+	9	1	2^+	10^+	3	4	5^+	1.7874
24	6	7	8^+	9	1	10^+	2^+	3	4	5^+	1.7874
25	6	7	8^+	9	10^+	1	2^+	3	4	5^+	2.0898

Table 1: z-test values and $\rho = 0.5$, Example 1.

Example 2

The data set used in this example concerns the survival of 30 patients with cervical cancer, where 16 patients received control treatment A, which was the use of radiotherapy alone, and 14 patients received new treatment B, where radiosensitizer is added to radiotherapy. The data, given in Table 2, represents the number of days between the start of the study and death of the patients caused by this cancer or a right-censoring event (Machin et al, 2006, p. 53). We use this data set to illustrate the use of the monotonicity result presented in this paper, by assuming imprecision in the recording of the events such that an observed time t_i would actually imply that the event occurred during interval $[t_i - d, t_i]$. For example, a hospital may only record events once per week, neglecting the precise day it occurred, or there may be some vagueness about the start date of the study and the actual start of recording of individual patients.

Figure 6 shows the z-test values for increasing values of d, so for increasing imprecision in the data, obtained using Equation (1), for $\rho = 0$, so for the log-rank Mantel-Haenszel test, and Figure 7 shows the z-test values for $\rho = 0.5$. The red horizontal line is the z critical value at 10% and the blue horizontal line is the z critical value at 5%. The green horizontal line is the value of the test statistic of this data set, without any imprecision in the data

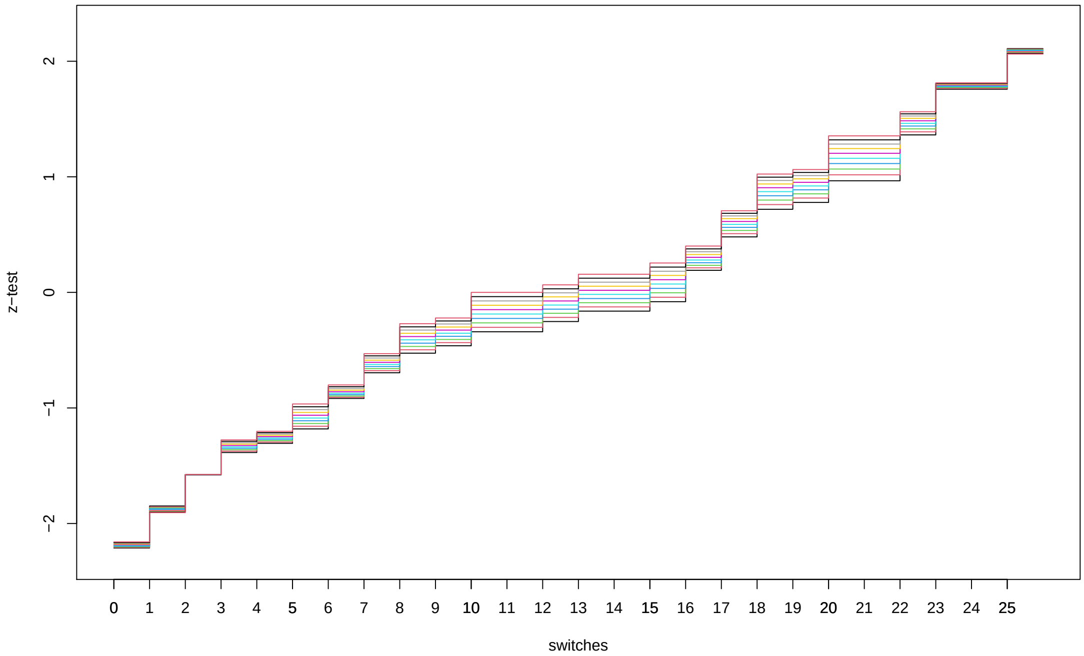


Figure 5: z-test values for different values of ρ, Example 1.

A	90	142	150	269	291	468^+	680	837	890^+	1037	1090^+
	1113^+	1153	1297	1429	1577^+						
B	272	362	373	383^+	519^+	563^+	650^+	827	919^+	978^+	1100^+
	1307	1360^+	1476^+								

Table 2: Data set, Example 2.

so with $d = 0$, which is equal to $z = -1.296818$. The black lines are the minimum and the maximum values of the z test statistic over different values of d. These are derived by applying the monotonicity result presented in this paper. The minimum value is obtained when we subtract d from all the data observations from treatment A, while keeping the data of treatment B at the original values. The maximum value is obtained when we subtract d from all the data observations from treatment B, while keeping the data of treatment B at the original values. The intersection points between the black line and the red and the blue lines are at $d = 122$ and $d = 271$ for both values of ρ. This analysis shows that, for the original data, the null hypothesis that both data sets may come from the same underlying population is not rejected, and this conclusion will remain the same with quite large imprecision added to the data, so the conclusion is very robust with regard to inaccuracies in the data.

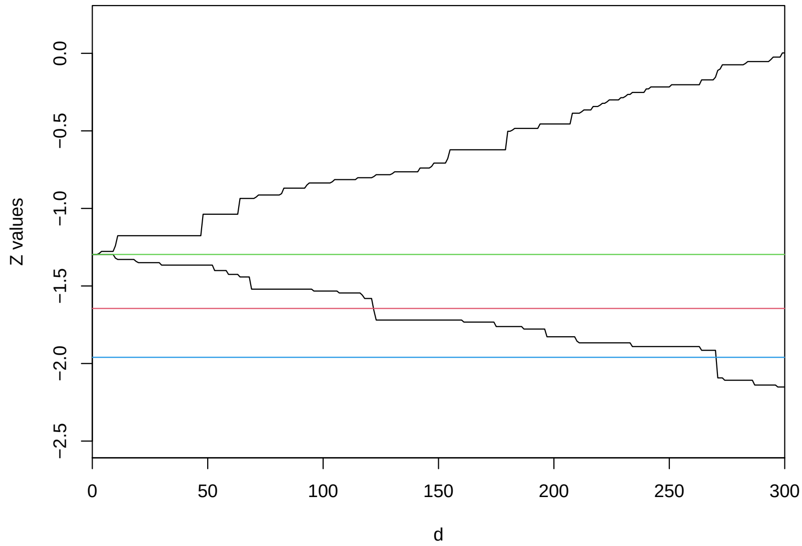


Figure 6: z-test values for different values of d and for $\rho = 0$, Example 2.

5. Concluding remarks

This paper studies the monotonicity of the G^ρ class of weighted logrank tests introduced by Harrington and Fleming (1982). We proved a convenient monotonicity property for the two-sample class of logrank tests. This property holds trivially for the special case where there are no right-censored observations (the Wilcoxon test), but, while intuitively quite clear, its proof required care due to the right censoring affecting the data. One can utilise this property to derive optimal bounds for the test statistic in case of imprecise data, as has been briefly illustrated via small examples in this paper, and it has recently been applied for robust inference with accelerated life testing data (Coolen, Ahmadini and Coolen-Maturi, 2021). Note that the form of imprecise data considered in this paper, in particular in Example 2, can also be regarded as interval-censored data. There is a huge literature on statistical inference with interval-censored data, but mostly methods based on additional assumptions are being considered. A feature of the imprecise data that is not standard in interval-censoring is the possible imprecision in a recorded right-censoring time, which was the main challenge in achieving the result in this paper.

For future research, it will be interesting to investigate the construction of statistical tests for equality of survival functions based on the number of switches, in a way that is similar to tests for perfect ranking in ranked set sampling presented by Li and Balakrishnan (2008). Possible generalization of the monotonicity property for tests with more than two groups of data is also of interest, it is left as a topic for future research.

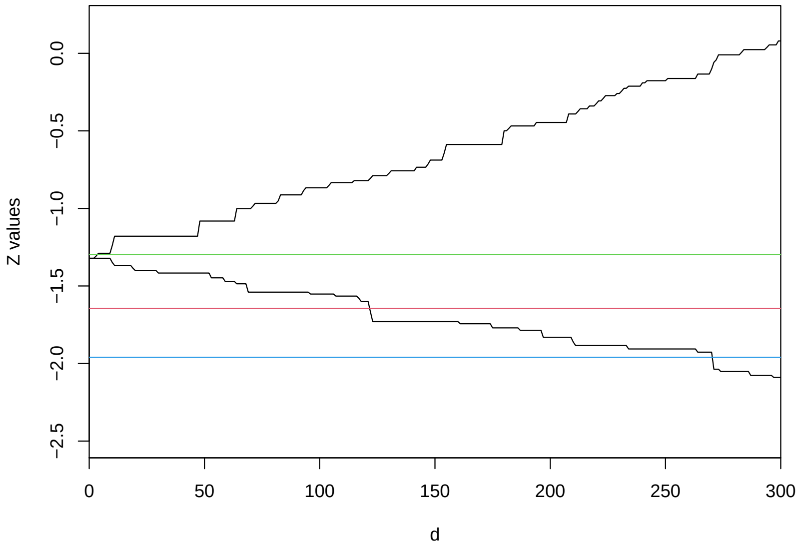


Figure 7: z-test values for different values of d and $\rho = 0.5$, Example 2.

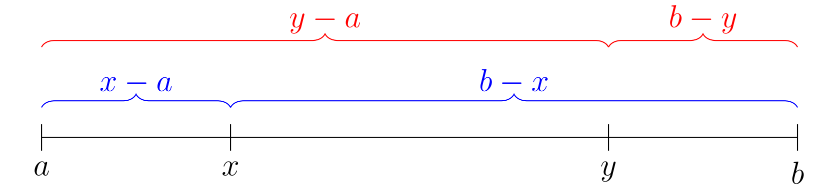
Acknowledgements

The authors are grateful to two anonymous reviewers whose supportive comments led to improved presentation of the paper.

Appendix

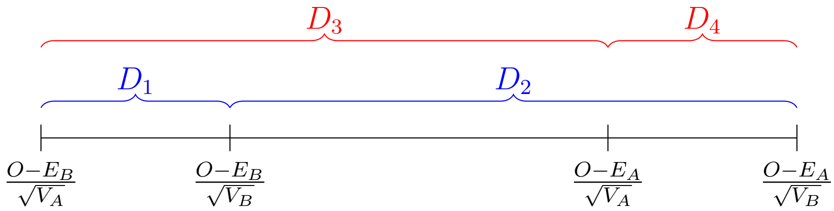
Lemma 5.1. *Let x and y be any two real numbers in an interval $[a, b]$, where $a < b$. Then if $x - a < y - a$ and $b - y < b - x$ then $x < y$.*

Proof: The setting is illustrated in the figure below.



As $y - a = (x - a) + (y - x)$ and $b - x = (b - y) + (y - x)$, then in order for both inequalities to hold, the second term in the right hand side must be positive, i.e. $y - x > 0$ thus $x < y$. $\square$

We use the lemma above to prove that $\frac{O-E_A}{\sqrt{V_A}} > \frac{O-E_B}{\sqrt{V_B}}$. First we define the 4 differences D_1, D_2, D_3 and D_4, which is illustrated in the figure below, as



$$
\begin{aligned}
D_1 &= \frac{O - E_B}{\sqrt{V_B}} - \frac{O - E_B}{\sqrt{V_A}} \\
D_2 &= \frac{O - E_A}{\sqrt{V_B}} - \frac{O - E_B}{\sqrt{V_B}} \\
D_3 &= \frac{O - E_A}{\sqrt{V_A}} - \frac{O - E_B}{\sqrt{V_A}} \\
D_4 &= \frac{O - E_A}{\sqrt{V_B}} - \frac{O - E_A}{\sqrt{V_A}}
\end{aligned}
$$

In order for the inequality $\frac{O-E_A}{\sqrt{V_A}} > \frac{O-E_B}{\sqrt{V_B}}$ to hold, both inequalities $D_3 > D_1$ and $D_2 > D_4$ must be hold. We can express D_2 and D_3 as

$$
\begin{aligned}
D_2 &= D_4 + \left[\frac{O - E_A}{\sqrt{V_A}} - \frac{O - E_B}{\sqrt{V_B}}\right] \\
D_3 &= D_1 + \left[\frac{O - E_A}{\sqrt{V_A}} - \frac{O - E_B}{\sqrt{V_B}}\right]
\end{aligned}
$$

so $\frac{O-E_A}{\sqrt{V_A}} - \frac{O-E_B}{\sqrt{V_B}}$ has to be positive, therefore $\frac{O-E_A}{\sqrt{V_A}} > \frac{O-E_B}{\sqrt{V_B}}$.

References

Augustin, T., Coolen, F.P.A., De Cooman, G., Troffaes, M.C.M., 2014. Introduction to Imprecise Probabilities. Wiley, Chichester.

Coolen, F.P.A., Ahmadini, A.A.H., Coolen-Maturi, T. 2021. Imprecise inference based on the log-rank test for accelerated life testing. Metrika, to appear, https://doi.org/10.1007/s00184-021-00807-4.

Gehan, E. A., 1965. A generalized Wilcoxon test for comparing arbitrarily singly-censored samples. Biometrika 52 (1/2), 203–223.

Harrington, D. P., Fleming, T. R., 1982. A class of rank test procedures for censored survival data. Biometrika 69 (3), 553–566.

Kaplan, E., Meier, P., 1958. Nonparametric estimation from incomplete observations. Journal of the American Statistical Association 53, 457–481.

Latta, R. B., 1977. Generalized Wilcoxon statistics for the two-sample problem with censored data. Biometrika 64 (3), 633–635.

Li, T., Balakrishnan, N., 2008. Some simple nonparametric methods to test for perfect ranking in ranked set sampling. Journal of Statistical Planning and Inference 138 (5), 1325 – 1338.

Lou, W. W., Lan, K. G., 1998. A note on the Gehan-Wilcoxon statistic. Communications in Statistics - Theory and Methods 27 (6), 1453–1459.

Mantel, N., 1966. Evaluation of survival data and two new rank order statistics arising in its consideration. Cancer chemotherapy reports 50 (3), 163–170.

Machin, D., Cheung, Y. B., Parmar, M., 2006. Survival Analysis: A Practical Approach, 2nd Edition, Wiley.

Oller, R., Langohr, K., 2017. FHtest: An R package for the comparison of survival curves with censored data. Journal of Statistical Software 81 (15), 1–25.

Peto, R., Peto, J., 1972. Asymptotically efficient rank invariant test procedures. Journal of the Royal Statistical Society. Series A (General) 135 (2), 185–207.

Prentice, R. L., 1978. Linear rank tests with right censored data. Biometrika 65 (1), 167–179.

Therneau, T. M., 2015. A Package for Survival Analysis in S. Version 2.38.