# Early detection of change-point in hazard rate with small sample size

Laurent Bordes, Frédéric Ibled, Christian Paroissin, Jean-Christophe Turlot

**Noname manuscript No.**
(will be inserted by the editor)

# Early detection of change-point in hazard rate with small sample size

**Laurent Bordes · Frédéric Ibled · Christian Paroissin · Jean-Christophe Turlot**

Received: date / Accepted: date

**Abstract** In this paper we address the problem of deciding if either $n$ consecutive independent failure times have the same failure rate or if there exists some $k \in \{1, \ldots, n\}$ such that the common failure rate of the first $k$ failure times is different from the common failure rate of the last $n-k$ failure times. The statistical tests we propose are based on mean ratio under the assumption that lifetimes are exponentially distributed and on Mann-Whitney or precedence test statistics when no parametric assumption on the underlying distribution is done. Our statistics are free of the unknown underlying distribution under the null hypothesis of homogeneity of the $n$ failure times which allows to compute critical values of our tests by Monte Carlo methods for small sample size. These tests have been developed in order to perform sequential change-point analysis of small sets of feedback data. They are applied to feedback data from Alstom Transport in order to detect an early change-point of the failure rate of an industrial equipment.

**Keywords** Exponential distribution · Mann-Whitney test · Precedence test · Statistical tables

L. Bordes
Université de Pau et des Pays de l'Adour, Laboratoire de Mathématiques et de leurs Applications - UMR CNRS 5142, Avenue de l'Université, 64013 Pau cedex, France, E-mail: laurent.bordes@univ-pau.fr

F. Ibled
Alstom Transport, Rue du Docteur Guinier, BP4, 65600 Séméac, France, E-mail: frederic.ibled@transport.alstom.com

Ch. Paroissin
Université de Pau et des Pays de l'Adour, Laboratoire de Mathématiques et de leurs Applications - UMR CNRS 5142, Avenue de l'Université, 64013 Pau cedex, France, E-mail: cparoiss@univ-pau.fr

J.-C. Turlot
Université de Pau et des Pays de l'Adour, Laboratoire de Mathématiques et de leurs Applications - UMR CNRS 5142, Avenue de l'Université, 64013 Pau cedex, France, E-mail: turlot@univ-pau.fr

**Mathematics Subject Classification (2000)** 62N05 · 62Q05 · 62P30

## 1 Introduction

Most of manufacturer companies have to supervise materials in use. Generally feedback data about components failure times are available. These data allow to compute the mean time to failure (MTTF) of the component lifetime. However high reliable components have relatively small failure rates, leading to small or moderate sample size. Hence one major problem is the on-line monitoring and the early detection of a change-point in the hazard rate.

Figure 1 represents typical feedback data from Alstom Transport. Circles represent failure times. The time elapsed between two consecutive failure times will be called inter-event duration. At calendar time $t$ (in days), the failure rate is estimated by $\lambda_t = N(t)/t$ where $N(t)$ is the number of observed failures over the time interval $[0, t]$. At first glance, one might legitimate ask if a change in the hazard rate occurs from some dates (vertical lines).

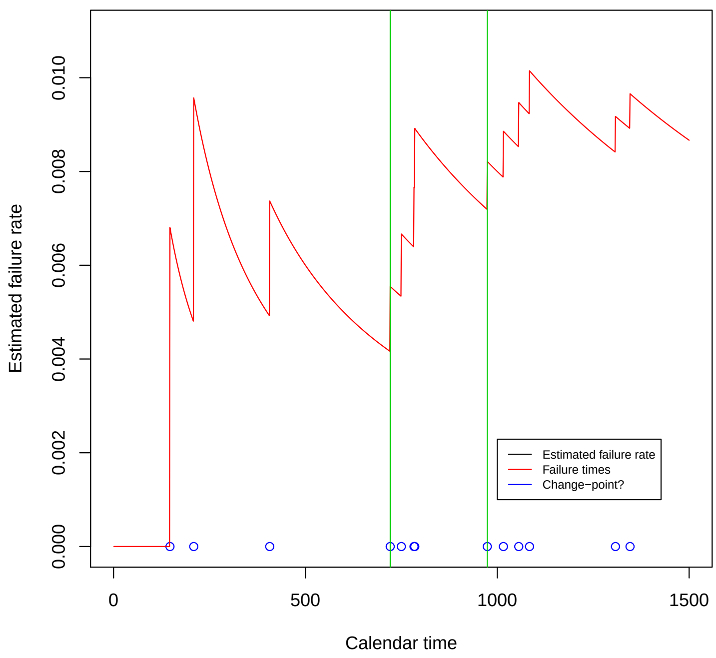


**Fig. 1** Typical feedback data from Alstom Transport

The aim of this paper is to provide some statistical tools to point out a potential change in the hazard rate, or equivalently in the lifetime distribution of inter-event durations. Since we will consider the case of small sample sizes, we will assume that at most a single change-point may occur. For Alstom Transport a change-point may occur due to two main reasons:

– environmental change for a system component: software modification, new operating constraints, excessive exploitation of one component, exceptional climatic conditions;
– change of the manufacturing process: assembly quality problem, minor modification of the manufacturing process (not mentioned by the supplier) with unexpected effect on the component reliability, quality slippage of the manufacturing process.

In practice these statistical tools are used only as warning of a change in failure rate.

This statistical problem has been studied from several decades. It was first treated half a century ago in a series of papers by Page [7–9]. Bayesian analysis was then considered ten years later by Chernoff and Zacks [4]. Early papers deal mainly with parametric distributions. The nonparametric case was first considered by Bhattacharya and Johnson [2] and by Sen and Srivastava [12]. Large sample studies of change-point detection methods have been paid attention, see for instance the book by Csörgö and Horváth [5].

Many change-point detection methods are based on classical maximum type statistic. To detect a change-point on a sample having small size we have to face two problems. The first one is that it is difficult to base a decision on large sample properties of involved statistics since it is well known that the convergence rates of maximum type statistics is rather low. The second problem is that statistics are generally not free of the underlying distribution of the sample (under the null hypothesis of "no change-point") which prevent to determine test critical values through Monte Carlo methods. Here we propose several methods that overcome the later problem and that do not require necessarily a parametric assumption on the underlying distribution. The first one assumes that inter-event durations are exponentially distributed which is equivalent to assume that failure times occur according to a homogeneous Poisson process. The two other methods do not rely on any parametric assumption. They are based respectively on Mann-Whitney test and precedence test statistics. The next section is devoted to the description of these methodologies. In addition statistical tables associated to our change-point tests are provided and computed using Monte Carlo simulations. Numerical illustrations and power comparisons are provided in Section 3. Section 4 deals with an application to a real data set from Alstom Transport.

## 2 Change-point detection methods

Let us denote by $X_1, \dots, X_n$ the $n$ inter-event durations available at the calendar time $t$. These random variables are assumed to be independent, but one wants to test whether they are identically distributed or not. In particular, one can be interested in detecting a change-point in this sequence of random variables. One says that a single change-point occurs if there exists an integer $k \in \{1, \dots, n\}$ such that $X_1, \dots, X_k$ are identically distributed according to $F$ and $X_{k+1}, \dots, X_n$ are identically distributed according to $G$ with $F \neq G$. The main difficulty is that we do not know if there is a change-point and, if it exists, where it occurs. Thus one has to consider all possible subsamples by splitting the whole sample into two consecutive parts. The general methodology for a procedure based on a homogeneity test will be detailed in the first subsection. Then three methods will be presented. The first one assumes that the inter-event lifetimes are exponentially distributed. The two other methods do not require any parametric assumption. They are based respectively on Mann-Whitney statistics and precedence tests [1].

### 2.1 General description of the methodology

The main scheme of the proposed methodology is as follows:

1. split the sample into two subsamples : $(X_1, \dots, X_k)$ and $(X_{k+1}, \dots, X_n)$ for $k \in \{m, \dots, n-m\}$;
2. compute the homogeneity test statistic for each splitting;
3. use all homogeneity test statistics to take a decision.

Decision can be either that no change-point occurs or that a change-point occurs at $k^* \in \{m, \dots, n-m\}$.

For the two subsamples $(X_1, \dots, X_k)$ and $(X_{k+1}, \dots, X_n)$, assume that one can apply a given homogeneity test. We denote by $S_{n,k}$ the corresponding statistic that aims to measure the "distance" between the two subsamples parent distributions. From all these statistics, we suggest three types of global test based on the $n - 2m + 1$ statistics:

1. maximum-type:
$$M_n = \max_{m \leqslant k \leqslant n-m} \frac{|S_{n,k}|}{\sqrt{\mathrm{var}(S_{n,k})}};$$
2. $\chi^2$-type:
$$\chi_n^2 = \sum_{k=m}^{n-m} \frac{S_{n,k}^2}{\mathrm{var}(S_{n,k})};$$
3. quadratic-type:
$$Q_n = \mathbf{S_n}^T \Sigma^{-1} \mathbf{S_n},$$
where $\mathbf{S_n} = (S_{n,m}, \dots, S_{n,n-m})$ and $\Sigma$ is the covariance matrix of $\mathbf{S_n}$.

One can observe that we consider that subsamples should contain a minimal number of observations. The choice of the value of $m$ will be discussed later. Quite similar statistics have been proposed in the literature, especially in the nonparametric case when using Mann-Whitney tests (see below).

### 2.2 Exponential distribution case

In this first case, one assumes that $X_1, \ldots, X_n$ are exponentially distributed. For any subsamples $\{X_1, \ldots, X_k\}$ and $\{X_{k+1}, \ldots, X_n\}$, one would like to test whether failure rates are equal ($H_0$: no change-point) or not ($H_1$: a single change-point). To do so, one can compare the empirical average of the two subsamples, i.e. the following statistics:

$$\frac{T_k}{k} = \frac{\sum_{i=1}^{k} X_i}{k} \quad \text{and} \quad \frac{T_n - T_k}{n-k} = \frac{\sum_{i=k+1}^{n} X_i}{k}.$$

In order to remove the unknown parameter under the null, it rather is convenient to consider the ratio of these two statistics:

$$T_{n,k} = \frac{n-k}{k} \frac{T_k}{T_n - T_k}.$$

Notice that considering the ratio of the empirical averages is equivalent to consider the ratio of the hazard rates. One can easily prove that (see appendix for details):

$$\mathbb{E}\left(T_{n,k}\right) = \frac{n-k}{k} \mathbb{E}(T_k) \mathbb{E}\left(\frac{1}{T_n - T_k}\right) = (n-k)\mathbb{E}\left(\frac{1}{T_n - T_k}\right) = \frac{n-k}{n-k-1}.$$

Therefore we can deduce from $T_{n,k}$ an unbiased statistic $S_{n,k}$ for the ratio of hazard rates:

$$S_{n,k} = \frac{n-k-1}{k} \frac{T_k}{T_n - T_k}.$$

It follows that $\mathbb{E}(S_{n,k}) = 1$ under the null hypothesis. This statistic tends to be larger than 1 when the failure frequency increases. Notice that according to the alternative hypothesis one could prefer to use a symetrized version of the above statistic:

$$S^*_{n,k} = \frac{n-k-1}{k} \frac{T_k}{T_n - T_k} - \frac{k-1}{n-k} \frac{T_n - T_k}{T_k}.$$

The covariance matrix of $\mathbf{S_n}$ is given in the following proposition (details for the proof are postponed in the appendix):

**Proposition 1.** *For any* $k \in \{m, \ldots, n-m\}$ *with* $m \geqslant 2$,

$$\mathrm{var}(S_{n,k}) = \frac{(k+1)(n-k-1)}{k(n-k-2)} - 1$$

*and for any* $(k,k') \in \{m,\dots,n-m\}^2$ *such that* $k' > k$ *and* $m \geqslant 3$,

$$\operatorname{cov}(S_{n,k},S_{n,k'}) = -\frac{k}{k'} + \frac{(k+1)(n-k-1)(n-k'-1)}{k'} \times \sum_{j=0}^{k'-k-1} \frac{(-1)^{k'-k-j-1}\Gamma(n-k-2)}{\Gamma(j+1)\Gamma(n-k')\Gamma(k'-k-j)(n-k-j-2)}.$$

Hence the value of $m$ will be set to 3 in order to make the above expressions accurate.

Using these results, one can use Monte Carlo simulations to compute any quantile of the three statistics and then obtain the critical values that are summarized in statistical tables. Tables 1–3 give critical values for sample sizes between 7 and 30, and for $\alpha \in \{20\%, 10\%, 5\%\}$.

| $n$ | $\alpha = 0.2$ | $\alpha = 0.1$ | $\alpha = 0.05$ |
|---|---|---|---|
| 7 | 1.59 | 2.26 | 3.04 |
| 8 | 1.96 | 2.70 | 3.56 |
| 9 | 2.27 | 3.07 | 3.95 |
| 10 | 2.53 | 3.34 | 4.26 |
| 11 | 2.71 | 3.55 | 4.48 |
| 12 | 2.88 | 3.72 | 4.63 |
| 13 | 3.05 | 3.89 | 4.79 |
| 14 | 3.19 | 4.05 | 4.98 |
| 15 | 3.30 | 4.15 | 5.06 |
| 16 | 3.41 | 4.26 | 5.14 |
| 17 | 3.51 | 4.35 | 5.23 |
| 18 | 3.62 | 4.45 | 5.32 |
| 19 | 3.71 | 4.56 | 5.45 |
| 20 | 3.77 | 4.59 | 5.43 |
| 21 | 3.86 | 4.69 | 5.55 |
| 22 | 3.93 | 4.77 | 5.62 |
| 23 | 4.01 | 4.82 | 5.68 |
| 24 | 4.08 | 4.90 | 5.75 |
| 25 | 4.15 | 4.96 | 5.79 |
| 26 | 4.20 | 5.00 | 5.82 |
| 27 | 4.26 | 5.07 | 5.90 |
| 28 | 4.32 | 5.11 | 5.89 |
| 29 | 4.38 | 5.17 | 5.99 |
| 30 | 4.44 | 5.23 | 6.03 |

**Table 1** Exponential case: critical values for the test based on $M_n$

### 2.3 Nonparametric approach based on Mann-Whitney test

Hereafter we still assume that random variables under consideration are independent but we no longer do any parametric assumption on their distributions. A way to compare nonparametrically the distribution of two subsamples is the

| $n$ | $\alpha = 0.2$ | $\alpha = 0.1$ | $\alpha = 0.05$ |
|---|---|---|---|
| 7 | 3.46 | 6.88 | 12.45 |
| 8 | 6.64 | 12.31 | 21.08 |
| 9 | 10.66 | 18.97 | 30.69 |
| 10 | 15.31 | 26.24 | 41.57 |
| 11 | 20.51 | 33.92 | 52.39 |
| 12 | 26.07 | 42.03 | 63.36 |
| 13 | 32.50 | 51.61 | 75.52 |
| 14 | 39.46 | 61.19 | 89.24 |
| 15 | 46.18 | 70.25 | 100.26 |
| 16 | 53.74 | 80.51 | 113.86 |
| 17 | 61.76 | 91.94 | 127.51 |
| 18 | 70.46 | 103.24 | 141.83 |
| 19 | 79.94 | 116.23 | 159.23 |
| 20 | 88.75 | 126.45 | 170.43 |
| 21 | 98.65 | 140.37 | 189.31 |
| 22 | 108.73 | 153.79 | 204.83 |
| 23 | 119.67 | 167.11 | 222.74 |
| 24 | 131.63 | 182.67 | 239.56 |
| 25 | 142.98 | 196.76 | 255.60 |
| 26 | 154.31 | 211.07 | 273.69 |
| 27 | 167.02 | 226.40 | 292.55 |
| 28 | 179.88 | 241.44 | 309.45 |
| 29 | 192.57 | 260.23 | 333.01 |
| 30 | 206.57 | 276.58 | 351.37 |

**Table 2** Exponential case: critical values for the test based on $\chi_n^2$

| $n$ | $\alpha = 0.2$ | $\alpha = 0.1$ | $\alpha = 0.05$ |
|---|---|---|---|
| 7 | 2.66 | 5.37 | 9.86 |
| 8 | 4.32 | 8.22 | 14.48 |
| 9 | 6.15 | 11.40 | 19.27 |
| 10 | 8.08 | 14.35 | 24.04 |
| 11 | 9.88 | 17.21 | 27.74 |
| 12 | 11.74 | 19.97 | 31.56 |
| 13 | 13.82 | 22.98 | 35.40 |
| 14 | 15.87 | 25.96 | 39.89 |
| 15 | 17.57 | 28.22 | 42.51 |
| 16 | 19.53 | 30.80 | 46.19 |
| 17 | 21.34 | 33.55 | 48.90 |
| 18 | 23.37 | 35.80 | 52.44 |
| 19 | 25.36 | 39.09 | 56.52 |
| 20 | 27.07 | 40.52 | 58.03 |
| 21 | 29.10 | 43.68 | 62.13 |
| 22 | 30.96 | 45.80 | 64.43 |
| 23 | 32.72 | 48.18 | 67.71 |
| 24 | 34.91 | 50.85 | 71.11 |
| 25 | 36.80 | 53.13 | 73.51 |
| 26 | 38.49 | 55.22 | 75.99 |
| 27 | 40.18 | 57.75 | 78.99 |
| 28 | 41.85 | 59.50 | 80.37 |
| 29 | 43.95 | 62.19 | 83.92 |
| 30 | 45.94 | 64.59 | 86.54 |

**Table 3** Exponential case: critical values for the test based on $Q_n$

so-called Mann-Whitney test. Let us recall that the Mann-Whitney statistic $S_{n,k}$ computed over the two subsamples $X_1, \ldots, X_k$ and $X_{k+1}, \ldots, X_n$ is defined by

$$S_{n,k} = \sum_{i=1}^{k} \sum_{j=k+1}^{n} \mathbb{I}_{X_j < X_i},$$

where $\mathbb{I}_A = 1$ if $A$ is true and 0 otherwise. Let us recall the expressions of the expectation and the variance of $S_{n,k}$:

$$\mathbb{E}[S_{n,k}] = \frac{k(n-k)}{2},$$

and

$$\mathrm{var}[S_{n,k}] = \frac{k(n-k)(n+1)}{12}.$$

The distribution of $S_{n,k}$ is not tractable explicitly, but recurrence formulae can be derived [3]. From these statistics, one can compute the three global statistics described previously: $M_n$, $\chi_n^2$ and $Q_n$. For the last one, the covariances of $S_{n,k}$ and $S_{n,k'}$ is required.

**Proposition 2.** *For any $k \in \{1, \ldots, n\}$ such that $k' > k$,*

$$\mathrm{cov}(S_{n,k}, S_{n,k'}) = \frac{k(n-k')(n+1)}{12}.$$

This result was early mentioned by Pettitt [10]. Such statistical tests were already considered in the literature. A maximum-like statistic have been suggested briefly by Sen and Srivastava [12]. Their statistic is slightly different from ours. Indeed they consider the maximum of the standardized the Mann-Whitney statistics while here we do not center Mann-Whitney statistics. Maximum of non standardized Mann-Whitney statistics was preferred by Pettitt [10] for which he derived the asymptotic distribution (see also [11]) where the presence of ties is discussed. Less frequently, other functionals of Mann-Whitney statistics have been considered in the literature. For instance, Bhattacharya and Johnson [2] considered the sum of these statistics. Sen and Srivastava [12] discussed briefly the power of their test and the one suggested by Bhattacharya and Johnson [2] using simulations. It appears that the power of these tests depend on the location of change-point in the whole sample. Lombard [6] proposed a kind of quadratic-like statistic (moreover he considered both the case of an abrupt change or a smooth one).

As for the previous case, we use Monte Carlo simulations to obtain a statistical table for the three types of statistics. Tables 4–6 give critical values for sample sizes between 7 and 30, and for $\alpha \in \{20\%, 10\%, 5\%\}$.

| $n$ | $\alpha = 0.2$ | $\alpha = 0.1$ | $\alpha = 0.05$ |
|---|---|---|---|
| 7 | 3.18 | 3.54 | 3.89 |
| 8 | 3.58 | 3.88 | 4.33 |
| 9 | 3.87 | 4.16 | 4.41 |
| 10 | 3.97 | 4.33 | 4.60 |
| 11 | 4.16 | 4.54 | 4.75 |
| 12 | 4.25 | 4.64 | 4.93 |
| 13 | 4.39 | 4.78 | 5.09 |
| 14 | 4.53 | 4.93 | 5.20 |
| 15 | 4.65 | 5.05 | 5.32 |
| 16 | 4.77 | 5.15 | 5.44 |
| 17 | 4.91 | 5.27 | 5.56 |
| 18 | 4.98 | 5.34 | 5.69 |
| 19 | 5.10 | 5.47 | 5.78 |
| 20 | 5.22 | 5.59 | 5.86 |
| 21 | 5.30 | 5.69 | 5.99 |
| 22 | 5.41 | 5.79 | 6.08 |
| 23 | 5.49 | 5.88 | 6.20 |
| 24 | 5.60 | 5.97 | 6.27 |
| 25 | 5.66 | 6.05 | 6.35 |
| 26 | 5.77 | 6.15 | 6.46 |
| 27 | 5.86 | 6.23 | 6.54 |
| 28 | 5.94 | 6.31 | 6.63 |
| 29 | 6.02 | 6.41 | 6.72 |
| 30 | 6.11 | 6.49 | 6.79 |

**Table 4** Nonparametric tests using Mann-Whitney statistics: critical values for the test based on $M_n$

## 2.4 Nonparametric approach using precedence tests

Precedence tests have been proposed to compare the distribution of a small sample to the distribution of a sample of moderate size (control group). A booklength account of these developments is due to Balakrishnan and Ng [1]. Hence, these tests can be used to detect an early change-point in a sequence of random variables. For convenience of notations, we set $(Y_1, \ldots, Y_{n-k}) = (X_{k+1}, \ldots, X_n)$. Then we define the random variables counting the number of observations of the second sample in each interval induced by the partition coming from the order statistics $X_{1:k}, \ldots, X_{k:k}$ of the first part of the sample

$$\forall j \in \{1, \ldots, k+1\}, \quad N_j = \sum_{l=1}^{n-k} \mathbb{I}_{Y_l \in [X_{j-1:k}, X_{j:k}[},$$

with $X_{0:k} = 0$ and $X_{k+1:k} = \infty$. Precedence test statistics are based on the sum of the $r$-th first random variables among $N_1, \ldots, N_{k+1}$, say

$$P_{n,k}(r) = \sum_{j=1}^{r} N_j = \sum_{j=1}^{r} \sum_{l=1}^{n-k} \mathbb{I}_{Y_l \in [X_{j-1:k}, X_{j:k}[}.$$

There is no heuristic for the best choice for $r$, but simulations studies in [1] prove that it should be rather small than large. Under the null hypothesis (i.e.

| $n$ | $\alpha = 0.2$ | $\alpha = 0.1$ | $\alpha = 0.05$ |
|---|---|---|---|
| 7 | 18.13 | 23.13 | 28.13 |
| 8 | 28.44 | 35.98 | 42.76 |
| 9 | 41.03 | 51.33 | 59.63 |
| 10 | 54.58 | 67.41 | 77.93 |
| 11 | 69.23 | 84.57 | 97.81 |
| 12 | 85.16 | 103.36 | 118.93 |
| 13 | 102.65 | 124.11 | 142.54 |
| 14 | 121.14 | 145.33 | 166.2 |
| 15 | 140.97 | 168.1 | 191.54 |
| 16 | 162.13 | 192.66 | 218.55 |
| 17 | 182.88 | 216.55 | 245.5 |
| 18 | 206.19 | 242.72 | 273.8 |
| 19 | 230.16 | 270.94 | 305.72 |
| 20 | 256.29 | 300.31 | 337.12 |
| 21 | 282.56 | 329.59 | 369.8 |
| 22 | 310.82 | 361.17 | 403.69 |
| 23 | 339.49 | 394.09 | 440.29 |
| 24 | 370.06 | 427.85 | 477.88 |
| 25 | 401.32 | 463.33 | 515.37 |
| 26 | 431.48 | 498.02 | 554.69 |
| 27 | 466.62 | 535.93 | 594.51 |
| 28 | 501.75 | 574.15 | 638.45 |
| 29 | 535.03 | 611.49 | 677.84 |
| 30 | 574.44 | 656.16 | 728.56 |

**Table 5** Nonparametric tests using Mann-Whitney statistics: critical values for the test based on $\chi_n^2$

when the underlying distributions of the two subsamples are identical), the distribution of $P_{n,k}(r)$ is known (see [1]) and given by

$$\forall x \in \{0, \dots, n-k\}, \quad \mathbb{P}(P_{n,k}(r) = x) = \frac{\binom{r+x-1}{x}\binom{n-r-x}{k-x}}{\binom{n}{n-k}}.$$

In order to compute the two first global statistics described above, one has to compute the expectation and the variance of $P_{n,k}(r)$. Let us set

$$M_n(r) = \max_{m \leqslant k \leqslant n-m} \frac{P_{n,k}(r)}{\sqrt{\mathrm{var}(P_{n,k}(r))}}$$

and

$$\chi_n^2(r) = \sum_{k=m}^{n-m} \frac{P_{n,k}(r)^2}{\mathrm{var}(P_{n,k}(r))}.$$

**Proposition 3.** *Under the null hypothesis, for $1 \leqslant r \leqslant k \leqslant n$, we have*

$$\mathbb{E}(P_{n,k}(r)) = \frac{(n-k)r}{k+1},$$

*and*

$$\mathrm{var}(P_{n,k}(r)) = \frac{r(k+1-r)(n-k)(n+1)}{(k+1)^2(k+2)}.$$

| $n$ | $\alpha = 0.2$ | $\alpha = 0.1$ | $\alpha = 0.05$ |
|---|---|---|---|
| 7 | 10.57 | 13.14 | 18.00 |
| 8 | 13.78 | 16.78 | 19.56 |
| 9 | 16.18 | 19.56 | 22.40 |
| 10 | 18.69 | 22.11 | 25.02 |
| 11 | 21.15 | 24.67 | 27.58 |
| 12 | 23.44 | 27.03 | 30.05 |
| 13 | 25.80 | 29.45 | 32.53 |
| 14 | 28.08 | 31.92 | 34.97 |
| 15 | 30.33 | 34.17 | 37.37 |
| 16 | 32.59 | 36.53 | 39.68 |
| 17 | 34.77 | 38.77 | 42.07 |
| 18 | 37.03 | 41.17 | 44.51 |
| 19 | 39.24 | 43.43 | 46.97 |
| 20 | 41.47 | 45.75 | 49.28 |
| 21 | 43.62 | 47.98 | 51.57 |
| 22 | 45.83 | 50.28 | 53.96 |
| 23 | 48.06 | 52.59 | 56.39 |
| 24 | 50.25 | 54.77 | 58.57 |
| 25 | 52.41 | 57.14 | 60.97 |
| 26 | 54.52 | 59.28 | 63.24 |
| 27 | 56.66 | 61.42 | 65.47 |
| 28 | 58.98 | 63.84 | 67.88 |
| 29 | 61.15 | 66.13 | 70.19 |
| 30 | 63.22 | 68.29 | 72.47 |

**Table 6** Nonparametric tests using Mann-Whitney statistics: critical values for the test based on $Q_n$

In particular, for $r = 1$, we have

$$\mathrm{var}(P_{n,k}(1)) = \frac{k(n-k)(n+1)}{(k+1)^2(k+2)} .$$

As for the previous tests, one can use Monte Carlo simulations to obtain statistical tables for the two first global statistics for $r = 1$ and $r = 2$. Tables 7–8 give critical values for sample sizes between 7 and 30, and for $\alpha \in \{20\%, 10\%, 5\%\}$.

## 3 Numerical illustrations

For the two first approaches considered above, we have computed the power of the three tests under several alternative hypothesis. Two samples, with respective length $n_1$ and $n_2$ such that $n = n_1 + n_2 = 20$, exponentially distributed with respective failure rates $\lambda_1$ and $\lambda_2$ were simulated. Three couples of parameters were considered: (i) $(\lambda_1, \lambda_2) = (1/2, 1)$, (ii) $(\lambda_1, \lambda_2) = (1/3, 1)$ and (iii) $(\lambda_1, \lambda_2) = (1/5, 1)$.

The power of the three global tests in the parametric case are represented on Figure 2. We observe that the higher the "difference" between the distributions is, the most powerful the three tests are. For each situation and for each

| | | $r=1$ | | | $r=2$ | |
|---|---|---|---|---|---|---|
| $n$ | $\alpha=0.2$ | $\alpha=0.1$ | $\alpha=0.05$ | $\alpha=0.2$ | $\alpha=0.1$ | $\alpha=0.05$ |
| 7 | 2.19 | 3.29 | 3.29 | 3.79 | 5.06 | 5.06 |
| 8 | 2.60 | 3.90 | 5.20 | 6.00 | 7.50 | 7.50 |
| 9 | 4.50 | 6.00 | 6.00 | 8.66 | 8.66 | 10.39 |
| 10 | 5.10 | 6.80 | 8.50 | 9.81 | 11.77 | 13.73 |
| 11 | 7.59 | 9.49 | 11.38 | 13.15 | 15.34 | 17.53 |
| 12 | 8.38 | 10.47 | 12.57 | 16.93 | 19.35 | 19.35 |
| 13 | 11.46 | 13.75 | 16.04 | 21.17 | 23.81 | 23.81 |
| 14 | 12.44 | 14.92 | 19.90 | 22.98 | 25.85 | 28.72 |
| 15 | 13.42 | 18.78 | 21.47 | 27.89 | 30.98 | 34.08 |
| 16 | 17.27 | 20.15 | 25.91 | 33.24 | 36.57 | 39.89 |
| 17 | 18.44 | 24.59 | 27.67 | 35.50 | 42.60 | 46.15 |
| 18 | 22.88 | 29.42 | 32.69 | 41.52 | 49.07 | 52.85 |
| 19 | 24.25 | 31.18 | 38.11 | 48.00 | 56.00 | 60.00 |
| 20 | 25.61 | 36.59 | 40.25 | 50.70 | 59.15 | 63.37 |
| 21 | 30.83 | 38.54 | 46.24 | 57.85 | 66.75 | 71.20 |
| 22 | 32.39 | 44.53 | 52.63 | 65.44 | 74.79 | 79.46 |
| 23 | 38.18 | 46.67 | 55.15 | 73.48 | 83.28 | 88.18 |
| 24 | 39.93 | 53.24 | 62.12 | 76.85 | 87.10 | 97.35 |
| 25 | 41.68 | 55.58 | 64.84 | 85.57 | 96.26 | 106.96 |
| 26 | 48.26 | 62.73 | 72.39 | 94.73 | 105.87 | 117.02 |
| 27 | 50.20 | 70.28 | 80.32 | 104.34 | 115.93 | 121.73 |
| 28 | 57.36 | 73.00 | 83.43 | 108.37 | 126.44 | 132.46 |
| 29 | 59.49 | 75.72 | 91.94 | 118.65 | 137.39 | 143.63 |
| 30 | 61.63 | 84.04 | 100.84 | 129.38 | 142.32 | 155.26 |

**Table 7** Nonparametric tests using precedence test statistics: critical values for the test based on $M_n(r)$ with $r \in \{1, 2\}$

global test, the power is larger when the change-point occurs at the middle of the sample (which seems quite natural). The power of the three statistics are relatively comparable. In general the quadratic-type statistic gives the best performance compared to the two others if the change-point occurs before the twelfth failure. If the change-point occurs later, then the maximum-type and the $\chi^2$-type tests are more powerful.

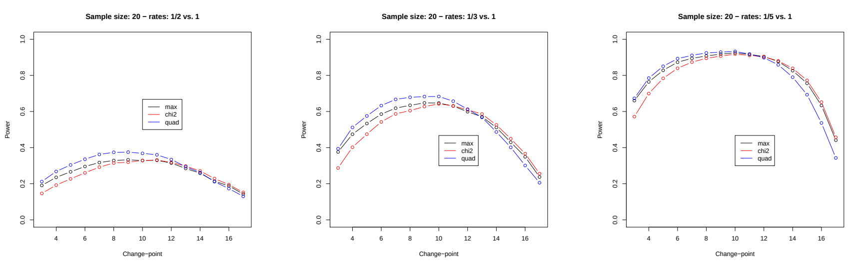


**Fig. 2** Parametric test: power comparisons of the three tests for three alternatives

| | r = 1 | | | r = 2 | | |
|---|---|---|---|---|---|---|
| $n$ | $\alpha = 0.2$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.2$ | $\alpha = 0.1$ | $\alpha = 0.05$ |
| 7 | 5.44 | 10.80 | 13.36 | 18.24 | 26.56 | 34.24 |
| 8 | 12.73 | 24.36 | 30.84 | 48.96 | 62.47 | 72.64 |
| 9 | 22.84 | 42.59 | 60.97 | 94.27 | 121.84 | 142.04 |
| 10 | 46.20 | 76.16 | 107.47 | 165.81 | 215.97 | 256.68 |
| 11 | 74.62 | 128.32 | 176.13 | 276.69 | 361.60 | 425.55 |
| 12 | 115.46 | 187.53 | 262.09 | 433.61 | 559.67 | 662.82 |
| 13 | 177.28 | 279.29 | 389.96 | 639.69 | 823.71 | 976.50 |
| 14 | 254.69 | 402.68 | 546.05 | 936.56 | 1209.00 | 1412.16 |
| 15 | 333.81 | 557.06 | 761.17 | 1321.83 | 1699.31 | 2003.80 |
| 16 | 477.44 | 764.68 | 1047.58 | 1811.42 | 2313.01 | 2732.65 |
| 17 | 615.42 | 998.40 | 1373.51 | 2340.87 | 3018.62 | 3566.47 |
| 18 | 865.69 | 1367.71 | 1839.10 | 3135.33 | 4040.79 | 4800.16 |
| 19 | 1058.24 | 1690.82 | 2318.27 | 4009.86 | 5136.13 | 6028.49 |
| 20 | 1344.06 | 2159.25 | 2854.65 | 5089.34 | 6535.66 | 7749.84 |
| 21 | 1661.28 | 2648.45 | 3630.44 | 6422.21 | 8202.64 | 9723.21 |
| 22 | 2008.16 | 3254.69 | 4355.95 | 7768.82 | 10035.01 | 12054.86 |
| 23 | 2476.97 | 3956.43 | 5491.69 | 9561.39 | 12267.25 | 14754.67 |
| 24 | 3018.32 | 4755.00 | 6381.55 | 11790.04 | 15098.63 | 17896.07 |
| 25 | 3598.05 | 5605.62 | 7840.47 | 14188.24 | 18102.13 | 21641.41 |
| 26 | 4344.18 | 6783.52 | 9011.05 | 16840.48 | 21531.24 | 25879.22 |
| 27 | 5050.47 | 8114.30 | 10669.47 | 19623.49 | 25211.01 | 29954.94 |
| 28 | 5843.74 | 9390.99 | 12465.42 | 23135.77 | 29701.95 | 34904.56 |
| 29 | 6996.68 | 10933.94 | 14985.69 | 27130.24 | 34534.58 | 40959.99 |
| 30 | 7894.21 | 12669.31 | 17181.05 | 31450.16 | 41054.62 | 48355.42 |

**Table 8** Nonparametric tests using precedence test statistics: critical values for the test based on $\chi^2_n(r)$ with $r \in \{1, 2\}$

The results for the three global tests in the nonparametric setup using Mann-Whitney statistics are represented on Figure 3. As previously, for all cases, the power is larger when the change-point occurs at the middle of the sample. Powers are naturally lower than the ones obtained with the parametric tests. If the change-point occurs at the begin or at the end (resp. middle) of the sample, then the quadratic-type (resp. maximum-type) statistic is the most powerful. At the contrary to the parametric case, we observe symmetric shapes of the power curves.

Finally we computed the power for three global statistics based on the precedence test. Figure 4 represents power computations when $r = 1$. Here the situation is quite different from the two previous cases. These tests are powerful for detecting an early change-point. Thus if the change-point occurs early, it is more powerful than the previous nonparametric test (but less powerful than the parametric test). Figure 5 represents the power computations when $r = 2$. Indeed for $r = 2$ tests are more powerful than for $r = 1$, especially when the two distributions are well separated. Again these tests are really competitive to detect an early change-point and can overcome the performance of parametric tests. The power is not increased when $r = 3$ (figures are not shown here). Similar behaviors have been already observed for the precedence test for the comparison of two samples [1].

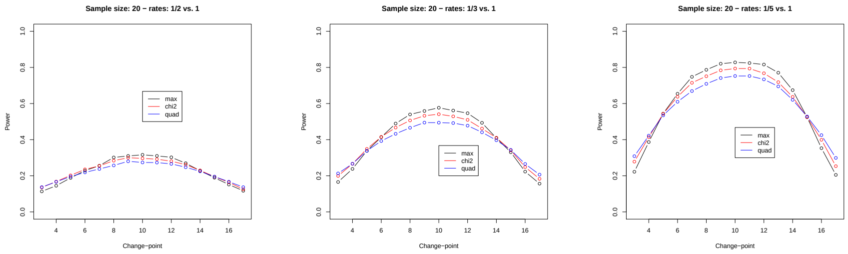


**Fig. 3** Nonparametric test using Mann-Whitney statistics: power comparisons of the three tests for three alternatives

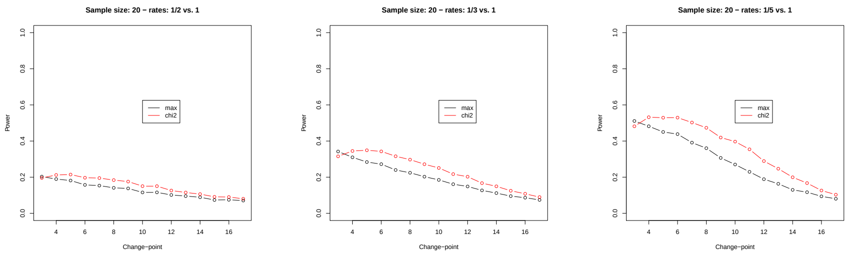


**Fig. 4** Nonparametric test using precedence tests with $r = 1$: power comparisons of the two tests for three alternatives

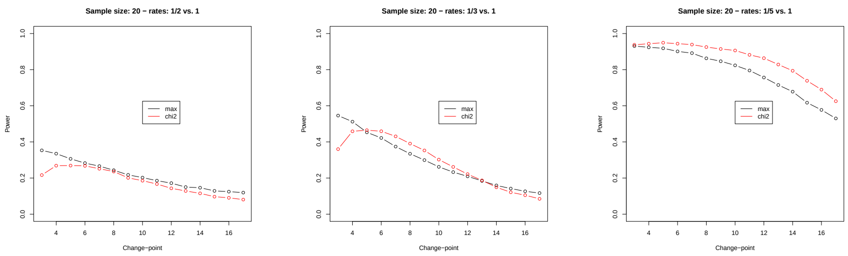


**Fig. 5** Nonparametric test using precedence tests with $r = 2$: power comparisons of the two tests for three alternatives

## 4 Application to real data

Here we come back to the application example from Alstom Transport. Table 9 contains inter-event durations (in days).

The sample size is thus equal to $n = 13$. It corresponds to the data plotted on Figure 1 in the introduction. In Table 10 the values of all the test statistics applied to the above dataset are reported (within parenthesis are the critical values from the previous tables).

| Failure number | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| Inter-events duration | 147 | 62 | 198 | 314 | 29 | 33 | 2 |
| Failure number | 8 | 9 | 10 | 11 | 12 | 13 | |
| Inter-events duration | 189 | 42 | 40 | 28 | 224 | 38 | |

**Table 9** Alstom Transport dataset

| | Parametric tests | MW test | Preced. test ($r = 1$) | Preced. test ($r = 2$) |
|---|---|---|---|---|
| $M_n$ | 3.52 | 4.78 | 16.04 | 18.52 |
| | (4.79) | (5.09) | (16.04) | (23.81) |
| $\chi^2_n$ | 36.18 | 108.7 | 436.78 | 732.26 |
| | (75.52) | (142.54) | (389.96) | (976.50) |
| $Q_n$ | 17.03 | 26.81 | | |
| | (35.40) | (32.53) | | |

**Table 10** Test statistics applied to Alstom Transport dataset

Except for precedence test with the $\chi^2_n$-type statistic and $r = 1$, the null hypothesis is never rejected with a first-type error equal to 5%.

## 5 Conclusion

In this paper we proposed several statistical methods to detect a single change-point in the failure rate of a renewal process when the number of available failure times is small. The statistical tests we developed are applied by Alstom Transport in order to detect components whose the reliability becomes weaker (due for example to a change of technology). The main advantage of our tests is that they are easy to implement and that they can be applied even if the sample size is small which is quite common in industrial situations. Alternative methods to the usual maximum type statistics are developed and simulation results show that tests based on quadratic type statistics can overcome the performances of tests based on maximum type statistics. Our methods could also be applied to propose new nonparametric control charts and the study of the distribution of quadratic-type tests for moderate or large sample sizes is also of interest.

## A Proofs

### A.1 Proof of Proposition 1

In the case of the exponential distribution, the statistic $S_{n,k}$ involves ratio of independent Erlang distributed random variables. Recall that random variable $X$ follows the Erlang distribution with parameters $k \in \mathbb{N}^*$ and $\alpha > 0$ if its probability density function $f$ is defined by

$$f(x; k, \alpha) = \frac{x^{k-1}}{\alpha^k \Gamma(k)} \exp(-x/\alpha)\mathbb{I}_{x>0},$$

where $\Gamma(k) = (k-1)!$. We note $X \sim Erl(k, \alpha)$. In particular notice that $Erl(k) \equiv Erl(k, 1)$. Calculation of the covariance matrix of $\mathbf{S_n}$ requires the calculation of $\mathbb{E}[1/(X(X+Y)]$ where $X$ and $Y$ are two independent Erlang distributed random variables.

**Lemma 1.** *Let $X \sim Erl(l)$ and $Y \sim Erl(k)$ being independent. If $l \geqslant 2$ and $k \geqslant 1$, then*

$$\begin{aligned}\mathbb{E}\left[\frac{1}{X(X+Y)}\right] &= \sum_{j=0}^{k-1}(-1)^{k-j-1}C_{k-1}^{j}\frac{1}{k+l-j-2}\frac{\Gamma(k+l-2)}{\Gamma(l)\Gamma(k)} \\ &= \sum_{j=0}^{k-1}(-1)^{k-j-1}\frac{\Gamma(k+l-2)}{\Gamma(l)\Gamma(k-j)\Gamma(j+1)(k+l-j-2)}\ .\end{aligned}$$

*Proof.*

$$\begin{aligned}&\mathbb{E}\left[\frac{1}{X(X+Y)}\right] \\ &= \frac{1}{\Gamma(l)\Gamma(k)}\int_0^{+\infty}\int_0^{+\infty}\frac{1}{x(x+y)}x^{l-1}y^{k-1}\exp(-(x+y))dxdy \\ &= \frac{1}{\Gamma(l)\Gamma(k)}\int_0^{+\infty}x^{l-2}\left(\int_x^{+\infty}\frac{(z-x)^{k-1}}{z}\exp(-z)dz\right)dx \\ &= \frac{1}{\Gamma(l)\Gamma(k)}\sum_{j=0}^{k-1}(-1)^{k-1-j}C_{k-1}^{j}\int_0^{+\infty}x^{k+l-j-3}\left(\int_x^{+\infty}z^{j-1}\exp(-z)dz\right)dx \\ &= \frac{1}{\Gamma(l)\Gamma(k)}\sum_{j=0}^{k-1}(-1)^{k-1-j}C_{k-1}^{j}\left(\left[\frac{x^{k+l-j-2}}{k+l-j-2}\int_x^{+\infty}z^{j-1}\exp(-z)dz\right]_0^{+\infty}\right. \\ &\qquad\qquad\left.+\int_0^{+\infty}\frac{x^{k+l-j-2}}{k+l-j-2}x^{j-1}\exp(-x)dx\right) \\ &= \sum_{j=0}^{k-1}(-1)^{k-j-1}C_{k-1}^{j}\frac{1}{k+l-j-2}\frac{\Gamma(k+l-2)}{\Gamma(l)\Gamma(k)}.\end{aligned}$$

This ends the proof of the lemma. □

Now let us achieve the proof of Proposition 1. We have

$$S_{n,k} = \frac{n-k-1}{k}\frac{T_k}{T_n-T_k} = \frac{n-k-1}{k}\frac{T_k}{R_k},$$

where $R_k = T_n - T_k = \sum_{j=k+1}^{n} X_j$. Let $(k, k') \in \{3, \ldots, n-3\}^2$, we obtain

$$\mathrm{cov}(S_{n,k}, S_{n,k'}) = \frac{(n-k-1)(n-k'-1)}{kk'}\mathbb{E}\left[\frac{T_kT_{k'}}{R_kR_{k'}}\right] - 1.$$

Assume that $3 \leqslant k < k' \leqslant n-3$, we write

$$\begin{aligned}R_k &= X_{k+1} + \cdots + X_{k'} + X_{k'+1} + \cdots + X_n \\ &= X_{k+1} + \cdots + X_{k'} + R_{k'} \\ &\equiv U_{k,k'} + R_{k'},\end{aligned}$$

and $T_{k'} = U_{k,k'} + T_k$. Then we need to calculate the following expectation

$$\mathbb{E}\left[\frac{T_kT_{k'}}{R_kR_{k'}}\right] = \mathbb{E}\left[\frac{T_k(T_k+U_{k,k'})}{R_{k'}(R_{k'}+U_{k,k'})}\right] = \mathbb{E}\left[\frac{X(X+Z)}{Y(Y+Z)}\right],$$

where

$$\begin{cases} X = T_k & \sim E(k) \quad \textit{with } k \geqslant 3, \\ Y = R_{k'} & \sim E(n-k') \textit{ with } l = n-k' \geqslant 3, \\ Z = U_{k,k'} & \sim E(k'-k) \textit{ with } q = k'-k \geqslant 1. \end{cases}$$

These three random variables are independent and since

$$\frac{X(X+Z)}{Y(Y+Z)} = \frac{X^2}{Y(Y+Z)} + \frac{X}{Y} - \frac{X}{Y+Z}$$

we have

$$\begin{aligned} \mathbb{E}\left[\frac{X(X+Z)}{Y(Y+Z)}\right] &= \mathbb{E}[X^2]\mathbb{E}\left[\frac{1}{Y(Y+Z)}\right] + \mathbb{E}[X]\mathbb{E}\left[\frac{1}{Y}\right] - \mathbb{E}[X]\mathbb{E}\left[\frac{1}{Y+Z}\right] \\ &= \frac{\Gamma(k+2)}{\Gamma(k)} \sum_{j=0}^{k'-k-1} \frac{(-1)^{k'-k-j-1} C_{k'-k-1}^{j} \Gamma(n-k-2)}{\Gamma(k'-k)\Gamma(n-k')(n-k-j-2)} \\ &\quad + \frac{k}{n-k'-1} - \frac{k}{n-k-1} \\ &= k(k+1) \sum_{j=0}^{k'-k-1} \frac{(-1)^{k'-k-j-1}\Gamma(n-k-2)}{\Gamma(j+1)\Gamma(k'-k-j)\Gamma(n-k')(n-k-j-2)} \\ &\quad + \frac{k(k'-k)}{(n-k-1)(n-k'-1)}. \end{aligned}$$

Then we obtain

$$\begin{aligned} \operatorname{cov}(S_{n,k}, S_{n,k'}) = -\frac{k}{k'} + \frac{(k+1)(n-k-1)(n-k'-1)}{k'} \\ \times \sum_{j=0}^{k'-k-1} \frac{(-1)^{k'-k-j-1}\Gamma(n-k-2)}{\Gamma(j+1)\Gamma(n-k')\Gamma(k'-k-j)(n-k-j-2)}. \end{aligned}$$

Similarly we calculate the variance

$$\begin{aligned} \operatorname{var}(S_{n,k}) &= \mathbb{E}((S_{n,k}-1)^2) \\ &= \left(\frac{n-k-1}{k}\right)^2 \mathbb{E}(T_k^2)\mathbb{E}\left(\frac{1}{R_k^2}\right) - 2\left(\frac{n-k-1}{k}\right)\mathbb{E}(T_k)\mathbb{E}\left(\frac{1}{R_k}\right) + 1 \\ &= \frac{(k+1)(n-k-1)}{k(n-k-2)} - 1\,. \end{aligned}$$

This ends the proof of Proposition 1.

## A.2 Proof of Proposition 3

The following calculations are carried out under the null hypothesis: $X_1, \ldots, X_n$ are assumed to be independent and identically distributed random variables. We will denote by $f$ and $F$ their common probability distribution function and cumulative distribution function. Hereinafter we need the distributions of single order statistic and couple of order statistics. For these expressions, we refer e.g. to [1].

First we have

$$\begin{aligned}\mathbb{E}(P_{n,k}(r)) &= \sum_{j=1}^{r}\sum_{l=1}^{n-k}\mathbb{E}\left(F(X_{j:k}) - F(X_{j-1:k})\right) \\ &= (n-k)\sum_{j=1}^{r}\left(\mathbb{E}\left(F(X_{j:k})\right) - \mathbb{E}\left(F(X_{j-1:k})\right)\right) \\ &= (n-k)\mathbb{E}\left(F(X_{r:k})\right), \end{aligned} \quad (1)$$

since $F(X_{0:k}) = F(0) = 0$ (we only consider positive random variables). Moreover

$$\begin{aligned}\mathbb{E}\left(F(X_{r:k})\right) &= \int_0^\infty F(x) f_{X_{r:k}}(x)dx \\ &= \frac{k!}{(r-1)!(k-r)!}\int_0^\infty F(x)^r f(x)(1-F(x))^{k-r}dx \\ &= \frac{k!}{(r-1)!(k-r)!}\int_0^1 u^r(1-u)^{k-r}du \\ &= \frac{k!}{(r-1)!(k-r)!}B(r+1,k-r+1) \\ &= \frac{k!}{(r-1)!(k-r)!}\frac{\Gamma(r+1)\Gamma(k-r+1)}{\Gamma(k+2)} \\ &= \frac{k!}{(r-1)!(k-r)!}\frac{r!(k-r)!}{(k+1)!} \\ &= \frac{r}{k+1}.\end{aligned}$$

Hence, this with (1) gives

$$\mathbb{E}(P_{n,k}(r)) = \frac{(n-k)r}{k+1}.$$

Then we can calculate the second moment of $P_{n,k}(r)$

$$\mathbb{E}\left(P_{n,k}(r)^2\right) = \sum_{j=1}^{r}\sum_{l=1}^{n-k}\sum_{j'=1}^{r}\sum_{l'=1}^{n-k}\mathbb{E}\left(\mathbb{I}_{Y_l\in[X_{j-1:k},X_{j:k}[}\mathbb{I}_{Y_{l'}\in[X_{j'-1:k},X_{j':k}[}\right).$$

We have to distinguish four cases: (i) $j = j'$ and $l = l'$; (ii) $j \neq j'$ and $l = l'$; (iii) $j = j'$ and $l \neq l'$; (iv) $j \neq j'$ and $l \neq l'$. Let us begin with case (i): it is exactly the same term as in the expectation. Case (ii) is also quite simple. Indeed, we have

$$\sum_{j=1}^{r}\sum_{l=1}^{n-k}\sum_{\substack{j'=1\\ j'\neq j}}^{r}\mathbb{I}_{Y_l\in[X_{j-1:k},X_{j:k}[}\mathbb{I}_{Y_l\in[X_{j'-1:k},X_{j':k}[} = 0,$$

since any $Y_l$ cannot belong to two non-overlapping intervals. Next we consider case (iii)

$$\begin{aligned}&\sum_{j=1}^{r}\sum_{l=1}^{n-k}\sum_{\substack{l'=1\\ l'\neq l}}^{n-k}\mathbb{E}\left(\mathbb{I}_{Y_l\in[X_{j-1:k},X_{j:k}[}\mathbb{I}_{Y_{l'}\in[X_{j-1:k},X_{j:k}[}\right) \\ &\quad = \sum_{j=1}^{r}\sum_{l=1}^{n-k}\sum_{\substack{l'=1\\ l'\neq l}}^{n-k}\mathbb{E}\left(\left(F(X_{j:k}) - F(X_{j-1:k})\right)^2\right) \\ &\quad = (n-k)(n-k-1)\sum_{j=1}^{r}\mathbb{E}\left(\left(F(X_{j:k}) - F(X_{j-1:k})\right)^2\right).\end{aligned}$$

We have

$$\begin{aligned}
&\mathbb{E}\left(\left(F(X_{j:k})-F(X_{j-1:k})\right)^2\right) \\
&\quad=\frac{k!}{(j-2)!(k-j)!}\int_0^\infty\int_{x_1}^\infty (F(x_1)-F(x_2))^2F(x_1)^2(1-F(x_2))^{k-j}f(x_1)f(x_2)dx_1dx_2 \\
&\quad=\frac{k!}{(j-2)!(k-j)!}\frac{(j-2)!2!(k-j)!}{(k+2)!} \\
&\quad=\frac{2k!}{(k+2)!} \\
&\quad=\frac{2}{(k+2)(k+1)}.
\end{aligned}$$

Finally we obtain

$$\sum_{j=1}^{r}\sum_{l=1}^{n-k}\sum_{\substack{l'=1\\l'\neq l}}^{n-k}\mathbb{E}\left(\mathbb{I}_{Y_l\in[X_{j-1:k},X_{j:k}[}\mathbb{I}_{Y_{l'}\in[X_{j-1:k},X_{j:k}[}\right)=\frac{2(n-k)(n-k-1)r}{(k+2)(k+1)}.$$

Last we consider the fourth case where all indices are different

$$\begin{aligned}
&\sum_{j=1}^{r}\sum_{l=1}^{n-k}\sum_{\substack{j'=1\\j'\neq j}}^{r}\sum_{\substack{l'=1\\l'\neq l}}^{n-k}\mathbb{E}\left(\mathbb{I}_{Y_l\in[X_{j-1:k},X_{j:k}[}\mathbb{I}_{Y_{l'}\in[X_{j'-1:k},X_{j':k}[}\right) \\
&\quad=(n-k)(n-k-1)\sum_{j=1}^{r}\sum_{\substack{j'=1\\j'\neq j}}^{r}\mathbb{E}\Big[\left(F(X_{j:k})-F(X_{j-1:k})\right)\left(F(X_{j':k})-F(X_{j'-1:k})\right)\Big] \\
&\quad=(n-k)(n-k-1)\sum_{j=1}^{r}\mathbb{E}\left[\left(F(X_{j:k})-F(X_{j-1:k})\right)\sum_{\substack{j'=1\\j'\neq j}}^{r}\left(F(X_{j':k})-F(X_{j'-1:k})\right)\right] \\
&\quad=(n-k)(n-k-1)\sum_{j=1}^{r}\mathbb{E}\Big[\left(F(X_{j:k})-F(X_{j-1:k})\right)\left(F(X_{r:k})-(F(X_{j:k})-F(X_{j-1:k}))\right)\Big] \\
&\quad=(n-k)(n-k-1)\sum_{j=1}^{r}\mathbb{E}\Big[\left(F(X_{j:k})-F(X_{j-1:k})\right)F(X_{r:k})\Big] \\
&\qquad\qquad-(n-k)(n-k-1)\sum_{j=1}^{r}\mathbb{E}\Big[\left(F(X_{j:k})-F(X_{j-1:k})\right)^2\Big].
\end{aligned}$$

The second term on the right hand-side is exactly the one appearing in case (iii) with opposite sign; thus it cancels. Hence we focus our attention on the calculation of the first term

$$\begin{aligned}
\sum_{j=1}^{r}\mathbb{E}[\left(F(X_{j:k})-F(X_{j-1:k})\right)F(X_{r:k})]&=\mathbb{E}[F(X_{r:k})\sum_{j=1}^{r}\left(F(X_{j:k})-F(X_{j-1:k})\right)] \\
&=\mathbb{E}[F(X_{r:k})^2] \\
&=\frac{k!}{(r-1)!(k-r)!}\int_0^1 u^{r+1}(1-u)^{k-r}du \\
&=\frac{k!}{(r-1)!(k-r)!}\frac{(r+1)!(k-r)!}{(k+2)!} \\
&=\frac{r(r+1)}{(k+2)(k+1)}.
\end{aligned}$$

Joining all the above results, we obtain the following expression for the second moment

$$\mathbb{E}\left(P_{n,k}(r)^2\right)=\frac{(n-k)r}{k+1}\left[\frac{(r+1)(n-k-1)}{k+2}+1\right].$$

Thus the variance can be derived from the two first moments.

## References

1. Balakrishnan, N., Ng, H.: Precedence-type tests and applications. Wiley, New York (2006)
2. Bhattacharya, P., Johnson, R.: Nonparametric tests for shifts at an unknown time point. Ann. Math. Statist. **39**, 1731–1734 (1968)
3. Capéraà, P., Cutsem, B.V.: Méthodes et modèles en statistique non paramétrique: exposé fondamental. Presses de l'Université de Laval - Dunod (1988)
4. Chernoff, H., Zacks, S.: Estimating the current mean of a normal distribution which is subjected to changes in time. Ann. Math. Statist. **35**, 999–1018 (1964)
5. Csörgö, M., Horváth, L.: Limit theorems in change-point analysis. Wiley, New York (1997)
6. Lombard, F.: Rank tests for change-point problems. Biometrika **74**(3), 615–624 (1987)
7. Page, E.: Continuous inspection scheme. Biometrika **41**, 100–114 (1954)
8. Page, E.: A test for a change in a parameter occurring at an unknown point. Biometrika **42**, 523–527 (1955)
9. Page, E.: On problems in which a change in parameters occurs at an unknown point. Biometrika **44**, 248–252 (1957)
10. Pettitt, A.: A non-parametric approach to the change-point problem. Appl. Statist. **28**(2), 126–135 (1979)
11. Pettitt, A.: Some results on estimating a change-point using non-parametric type statistics. Appl. Statist. **11**, 261–272 (1980)
12. Sen, A., Srivastava, M.: On tests for detecting change in mean. Ann. Statist. **3**(1), 98–108 (1975)