Tests for an Increasing Trend in the Intensity of a Poisson Process: A Power Study

Maxime Baba, Mathile Ferreira, Félix Foucher de Brandois

A Nonhomogeneous Poisson Process (NHPP) is a stochastic process often used to model phenomena where the rate of occurrence of events changes over time. The rate, or intensity function $\lambda(t)$, represents the expected number of events per unit time at a given time t. Understanding and analyzing the behavior of this rate is crucial in diverse fields such as reliability engineering, healthcare, and environmental studies, as it helps identify patterns, predict future events, and optimize interventions [1]. Detecting increasing trends in the rate can be particularly important, for example, in monitoring system deterioration or identifying escalating risks in processes. Although various statistical methods have been proposed to identify increasing trends in NHPPs, existing treatments—such as those by Bain, Engelhardt, and Wright [2]—lack clarity and precision in both theoretical explanation and practical application. This study addresses these shortcomings by providing a detailed exploration of the Laplace test and Boswell's likelihood ratio test. The paper is structured in three parts: a theoretical discussion of the selected tests, numerical simulations comparing their performance, and an application to real-world data.

1 Theoretical Framework for Trend Detection Tests

A statistical test evaluates two competing hypotheses: the null hypothesis H_0 (representing the status quo) and the alternative hypothesis H_1 (indicating a deviation from H_0). Based on the sample data, a test statistic is computed and compared to a critical threshold. If the test statistic exceeds this threshold, H_0 is rejected in favor of H_1. In our case, we test whether the intensity function $\lambda(t)$ of a Poisson process is constant (H_0) or increasing (H_1).

1.1 Laplace Test

1.1.1 Theoretical Basis

Let $(N(t))_{t\geq 0}$ be a Poisson process with intensity function $\lambda(t)$. We observe $N(t)$ in the interval $[0, T^*]$. Let $0 < T_1 < T_2 < \ldots < T_n < T^*$ be the ordered observation times.

Test Statistic

Under H_0, the arrival times $T_1, \ldots, T_n$ (conditioned on $N_{T^*} = n$) behave like order statistics from a uniform distribution. Specifically:

$(T_1, \ldots, T_n)|\{N_{T^*} = n\} \overset{(d)}{=} (U_1, \ldots, U_n)$, where $U_1, \ldots, U_n \underset{i.i.d.}{\sim} \mathcal{U}([0, T^*])$.

Therefore, $(\frac{T_1}{T^*}, \ldots, \frac{T_n}{T^*})|\{N_{T^*} = n\} \overset{(d)}{=} (V_1, \ldots, V_n)$, where $V_1, \ldots, V_n \underset{i.i.d.}{\sim} \mathcal{U}([0, 1])$.

Define the Laplace test statistic as:

$$F = \frac{1}{T^*} \sum_{i=1}^{n} T_i. \tag{1}$$

By the Central Limit Theorem, under H_0, can be standardized as: $Z = \frac{F - \mathbb{E}[F]}{\sqrt{\mathrm{Var}(F)}} = \frac{F - \frac{n}{2}}{\sqrt{\frac{n}{12}}} \sim \mathcal{N}(0, 1)$ for large n.

Decision Rule

If the intensity $\lambda(t)$ is increasing, the arrival times T_i are expected to cluster towards the end of the interval $[0, T^*]$, making F larger. Therefore, we reject H_0 if $F \geq s_\alpha$, where s_α is the critical threshold determined as: $s_\alpha = \frac{n}{2} + z_{1-\alpha}\sqrt{\frac{n}{12}}$.

with $z_{1-\alpha}$ being the $(1-\alpha)$-th quantile of the standard normal distribution.

Type I Error
The type I error of the test is the probability of rejecting H_0 when it is true.
$e_1 = \mathbb{P}_{H_0}(F \geq s_\alpha) = \mathbb{P}(Z \geq z_{1-\alpha}) = \alpha$.

P-value
The p-value measures the probability of obtaining a test statistic at least as extreme as the observed F_{obs}, under H_0. It is given by:
$\hat{\alpha} = \mathbb{P}_{H_0}(F \geq F_{\text{obs}}) = 1 - \Phi(\frac{F_{\text{obs}} - \frac{n}{2}}{\sqrt{\frac{n}{12}}})$.

Power
The power of the test is the probability of rejecting H_0 when H_1 is true.
$\pi(\lambda) = \mathbb{P}_{H_1}(F \geq s_\alpha)$.

1.2 Boswell's Likelihood Ratio Test

The goal of Boswell's test is to maximize the likelihood function under the constraint that the intensity function $\lambda(t)$ is non-decreasing. Since the likelihood function involves the intensity values at the observed failure times, we look for the optimal estimates $\hat{\lambda}(T_i)$ that satisfy this constraint.

1.2.1 Theoretical Basis

Likelihood Function Recap
The likelihood function for the ordered observation times $T_1, \ldots, T_n$ is given by:

$$\mathcal{L}((T_i)_{1 \leq i \leq n}; \lambda) = \left(\prod_{i=1}^{n} \lambda(T_i) \right) \exp \left(- \int_0^{T_n} \lambda(t) dt \right)$$

We assume that the intensity function $\lambda(t)$ is piecewise constant and non-decreasing between successive failure times T_i and T_{i+1}. Therefore, we need to determine the values $\hat{\lambda}(T_i)$ that maximize this likelihood.

Formulating the Optimization Problem
Under the previous assumptions:
$\int_0^{T_n} \lambda(t) dt = \sum_{i=1}^{n} (T_{i+1} - T_i) \lambda(T_i)$

We want to find $\hat{\lambda}(T_i)$ that maximizes:
$\prod_{i=1}^{n} \lambda(T_i) \exp(-(T_{i+1} - T_i)\lambda(T_i))$

This can be rewritten as:

$$\max_{\lambda} \prod_{i=1}^{n} f_i(\lambda(T_i))$$

where $f_i(x) = x \exp(-(T_{i+1} - T_i)x)$

The function $f_i(x)$ is unimodal, meaning it has a unique maximum. According to Boswell's theorem [3], following the works of Brunk and von Eeden [4], the likelihood is maximized by the non-decreasing function $\hat{\lambda}(t)$ given by:

$$\hat{\lambda}(T_i) = \max_{1 \leq \alpha \leq ii \leq \beta \leq n} \min M(\alpha, \beta)$$

where $M(\alpha, \beta)$ is the maximum of $x \mapsto \prod_{i=\alpha}^{\beta} f_i(x)$

We can find the form of $M(\alpha, \beta)$ by derviating the product of $f_i(x)$ and setting it to zero.
This leads to:

$$\hat{\lambda}(T_i) = \max_{1 \leq \alpha \leq ii \leq \beta \leq n} \min \frac{\beta - \alpha + 1}{T_{\beta+1} - T_\alpha} \tag{2}$$

Test Statistic
The likelihood ratio test (LRT) compares the likelihood of the data under the null hypothesis H_0 (constant intensity) with the likelihood under the alternative hypothesis H_1 (non-decreasing intensity). The test statistic is defined as:

$$W = -2 \log \left(\frac{\sup_{\ \lambda \in \Lambda_0} \mathcal{L}(\lambda)}{\sup_{\ \lambda \in \Lambda} \mathcal{L}(\lambda)} \right)$$

where Λ_0 is the set of constant intensity functions and Λ is the set of non-decreasing intensity functions.

Under H_0, the maximum likelihood estimate of $\lambda(t)$ is $\lambda_0(t) = \frac{n}{T_n}$.
The log-likelihood under H_0 is:

$$\log \mathcal{L}(\lambda_0) = \sum_{i=1}^{n} \log(\lambda_0(T_i)) - T_n \lambda_0(T_n) = n \log \left(\frac{n}{T_n} \right) - n$$

Under H_1, the maximum likelihood using the optimal estimate $\hat{\lambda}(t)$ is:

$$\log \mathcal{L}(\hat{\lambda}) = \sum_{i=1}^{n} \log(\hat{\lambda}(T_i)) - \int_0^{T_n} \hat{\lambda}(t) dt$$

By Grenander's theorem [5], we know that:
$\int_0^{T_n} \hat{\lambda}(t)dt = n$

Therefore, the test statistic W becomes:

$$W = 2\left(\sum_{i=1}^{n} \log(\hat{\lambda}(T_i)) + n\log\left(\frac{T_n}{n}\right)\right) \qquad (3)$$

Under H_0, by Wilks' theorem [6], the test statistic W follows a χ^2 distribution with $n-1$ degrees of freedom: $W \sim \chi^2(n-1)$.

Decision Rule
The decision rule for the LRT is based on the critical threshold w_α. If $W \geq w_\alpha$, we reject H_0. w_α is determined by the $(1-\alpha)$-th quantile of the $\chi^2(n-1)$ distribution.

Type I Error
The probability of a Type I error (rejecting H_0 when it is true) is:
$e_1 = \mathbb{P}_{H_0}(W \geq w_\alpha) = \alpha$.

P-value
The p-value is the probability of obtaining a test statistic at least as extreme as the observed W_{obs} under H_0. It is given by:
$\hat{\alpha} = \mathbb{P}_{H_0}(W \geq W_{\text{obs}}) = 1 - F_{\chi^2(n-1)}(W_{\text{obs}})$.

Power
The power of the test is the probability of rejecting H_0 when H_1 is true:
$\pi(\lambda) = \mathbb{P}_{H_1}(W \geq w_\alpha)$.
Under H_1, the intensity function $\lambda(t)$ is non-decreasing and can potentially exhibit multiple change points. Suppose there are k change points, this implies there are $k+1$ regions where the intensity remains constant. The possible ways that the n observed failure times can be grouped into k segments are given by the Stirling numbers of the first kind $s(n,k)$.
To compute the power, we sum over all possible numbers of segments k and weigh each possibility by the probability of observing $W \geq w_\alpha$ for that configuration.

$$\pi(\lambda) = \sum_{k=1}^{n} \frac{s(n,k)}{n!}\mathbb{P}(\chi^2(k+1) \geq w_\alpha) \qquad (4)$$

2 Numerical Simulations

2.1 Simulation Methodology

Generating simulated data :
Multiple rates functions...

2.2 Power Analysis

Calculation of the power of the Laplace and Boswell tests under different conditions (e.g., exponential, Weibull, and step-function trends).

Visualization of power curves to compare test effectiveness.

Discussion of factors influencing the power of each test (sample size, intensity function shape, etc.).

Insights into when to prefer one test over the other.

3 Application to Real-World Data

3.1 Description of the Dataset

Danish Fire Insurance Claims.
Overview of the dataset containing large fire insurance claims in Denmark from 1980 to 1990.
Explanation of how the event times are extracted and processed.

3.2 Applying the Tests

Presentation of test statistics, p-values, and decisions.

3.3 Results and Discussion

Analysis of Findings.
Interpretation of the results and their implications for the intensity of fire insurance claims.
Discussion of whether an increasing trend is detected and its significance.
Comparison with Simulated Results.
Reflecting on how real-world results compare to the numerical simulations.

qsd

4 Conclusion

Ceci est une citation [7].

References

[1] John Pork. “An Example Intervention”. In: *Trouver des exemples* (2020).

[2] Max Engelhardt Lee J. Bain and F. T. Wright. “Tests for an Increasing Trend in the Intensity of a Poisson Process: A Power Study”. In: *Journal of the American Statistical Association* 80.390 (1985), pp. 419–422.

[3] M. T. Bowsell. “Estimating and Testing Trend in a Stochastic Process of Poisson Type”. In: *Annals of Mathematical Statistics* 37 (1966), pp. 1564–1573.

[4] C. Van Eeden. “Maximum Likelihood Estimation of Partially or Completely Ordered Parameters”. In: *Koninklijke Nederlandse Akademie van Wetenschappen, Proceedings, Series A / Indagationes Mathematicae* 59.4 (1956). Reprinted from *Indagationes Mathematicae*, 18, No. 4, 1956, pp. 128–136.

[5] U. Grenander. “On the Theory of Mortality Measurement”. In: *Skandinavisk Aktuarietidskrift* (1956), pp. 138–140.

[6] S. S. Wilks. “The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses”. In: *Annals of Mathematical Statistics* 9.1 (1938), pp. 60–62.

[7] John Smith. “An Example Article”. In: *Journal of Examples* (2020).