Inhomogeneous Poisson: Maximum Likelihood Estimator

Roberto Baviera[‡]

March 6, 2023

(‡) Politecnico di Milano, Department of Mathematics, 32 p.zza L. da Vinci, Milano

Abstract

We briefly describe the MLE estimators for a simple inhomogeneous Poisson process, deriving the asymptotic estimates of confidence intervals and confidence regions using the Fisher information matrix.

Keywords: Poisson process, MLE estimation.

JEL Classification: C51,

Address for correspondence:
Roberto Baviera
Department of Mathematics
Politecnico di Milano
32 p.zza Leonardo da Vinci
I-20133 Milano, Italy
Tel. +39-02-2399 4575
Fax. +39-02-2399 4621
roberto.baviera@polimi.it

1 MLE default times

Consider a default random variable with survival probability $e^{-\int_0^t \lambda_s ds}$ with $\lambda_t = \lambda_1 I_{t<\theta} + \lambda_2 I_{t\geq\theta}$ and $\lambda_1,\ \lambda_2,\ \theta > 0$. We write the likelihood of the default

$$l(t) = \lambda_t e^{-\int_0^t \lambda_s ds} \quad .$$

For an N-dimensional sample

$$L(t, N) := \prod_N l(\tau_i) \quad .$$

The log-likelihood is:

$$\log(L(t, N)) = n_1 \log(\lambda_1) + n_2 \log(\lambda_2) - \sum_{n_1} \tau_i \lambda_1 - \sum_{n_2} ((\tau_i - \theta)\lambda_2 + \theta\lambda_1) \quad ,$$

where n_1 is the number of defaults between 0 and θ and n_2 the number of defaults after θ. The MLE estimators are

$$\lambda_1^{MLE} = \frac{n_1}{\sum_{n_1} \tau_i + n_2\theta} \quad ; \quad \lambda_2^{MLE} = \frac{n_2}{\sum_{n_2} (\tau_i - \theta)} \quad .$$

We can construct the Fisher information matrix $\boldsymbol{\mathcal{I}}$ for the two estimators (see, e.g., Lehmann and Casella 2006 Lemma 5.3 p.116 and p.125). Let us define the vector $\underline{\lambda} = [\lambda_1, \lambda_2]^T$: we obtain

$$\boldsymbol{\mathcal{I}}_{11}(\underline{\lambda}) = -\mathbb{E}\left[\frac{\partial^2 \log l}{\partial \lambda_1^2}\right] = \frac{1 - e^{-\lambda_1\theta}}{\lambda_1^2} \qquad \boldsymbol{\mathcal{I}}_{22}(\underline{\lambda}) = -\mathbb{E}\left[\frac{\partial^2 \log l}{\partial \lambda_2^2}\right] = \frac{e^{-\lambda_1\theta}}{\lambda_2^2}$$

and

$$\boldsymbol{\mathcal{I}}_{12}(\underline{\lambda}) = -\mathbb{E}\left[\frac{\partial^2 \log l}{\partial \lambda_1 \partial \lambda_2}\right] = 0 \ .$$

The information matrix is diagonal.

We know that, being $\underline{\lambda}^*$ the true vector of the parameters, $\underline{\lambda}^{MLE}$ its maximum likelihood estimate and N the cardinality of the sample, the quantity $\sqrt{N}(\underline{\lambda}^{MLE} - \underline{\lambda}^*)$ is asymptotically normal with vector mean zero and covariance matrix $\boldsymbol{\mathcal{I}}^{-1}(\underline{\lambda}^*)$, i.e.

$$\sqrt{N}(\underline{\lambda}^{MLE} - \underline{\lambda}^*) \xrightarrow{\mathcal{L}} \boldsymbol{\mathcal{N}}(\underline{0}, \boldsymbol{\mathcal{I}}^{-1}(\underline{\lambda}^*))$$

(see, e.g., Lehmann and Casella 2006 Theorem 5.1, p.463).

This result allows obtaining MLE confidence intervals (CIs) of level α (e.g., 95%). For each of the true parameters λ_i^*, we have

$$\mathrm{CI}_i(\alpha) = \left[\lambda_i^{MLE} - \mathcal{N}^{-1}\Big(\frac{1+\alpha}{2}\Big)\frac{1}{\sqrt{N\boldsymbol{\mathcal{I}}_{ii}(\underline{\lambda}^{MLE})}},\ \lambda_i^{MLE} + \mathcal{N}^{-1}\Big(\frac{1+\alpha}{2}\Big)\frac{1}{\sqrt{N\boldsymbol{\mathcal{I}}_{ii}(\underline{\lambda}^{MLE})}}\right] \ .$$

Moreover, one can also deduce the confidence region (CR) of level α for the vector $\underline{\lambda}^* \in \mathbb{R}^2$. This can be done in two ways:

(1) exploiting independence only. The resulting confidence region of level α is rectangluar, obtained as

$$\mathrm{CR}(\alpha) = \mathrm{CI}_1(\sqrt{\alpha}) \times \mathrm{CI}_2(\sqrt{\alpha}) \ .$$

Indeed

$$\begin{aligned}\mathbb{P}\big(\underline{\lambda}^* \in \mathrm{CR}(\alpha)\big) &= \mathbb{P}\big(\lambda_1^* \in \mathrm{CI}_1(\sqrt{\alpha}), \lambda_2^* \in \mathrm{CI}_2(\sqrt{\alpha})\big) = \\ &= \mathbb{P}\big(\lambda_1^* \in \mathrm{CI}_1(\sqrt{\alpha})\big)\ \mathbb{P}\big(\lambda_2^* \in \mathrm{CI}_2(\sqrt{\alpha})\big) \ .\end{aligned}$$

(2) exploiting asymptotic Gaussianity. The resulting confidence region of level α is elliptical:

$$\mathrm{CR}(\alpha) = \left\{ \underline{\lambda} \in \mathbb{R}^2 \ s.t.\ (\underline{\lambda} - \underline{\lambda}^{MLE})^T \boldsymbol{\mathcal{I}}(\underline{\lambda}^{MLE})(\underline{\lambda} - \underline{\lambda}^{MLE}) \leq \frac{1}{N} X_2^2\ (\alpha) \right\} ,$$

where $X_2^2(\cdot)$ denotes the CDF of a chi-squared distribution with 2 degrees of freedom. Indeed it is well-known that, if $\underline{Y} \in \mathbb{R}^d$ is a (multivariate) normal random vector with mean $\underline{\mu}$ and covariance matrix Σ, i.e.

$$\underline{Y} \sim \boldsymbol{\mathcal{N}}\Big(\underline{\mu}, \Sigma\Big) ,$$

then

$$(\underline{Y} - \underline{\mu})^T \Sigma^{-1} (\underline{Y} - \underline{\mu}) \sim \chi_d^2$$

being χ_d^2 a chi-squared distribution with d degrees of freedom.

References

E. L. Lehmann and G. Casella. *Theory of point estimation.* Springer Science & Business Media, Second edition, 2006.