Stochastic process estimation from partial observations: Poisson case

Sergey Kirgizov, François Queyroi

August 11, 2014

Abstract

Having a sequence of values $v_0, v_\Delta, v_{2\Delta}, \ldots, v_{N\Delta}$, which are measured every Δ units of time, usually we are interested in the prediction of the future outcome of this sequence at time $(N+1)\Delta$. But in some real-world cases we want to know, not the future, but rather the truth about the present: *if Maria performs more observations per unit of time than Maximilian, how can he estimates the Maria's results from his own?*

In this small note we consider the situation when the underlying process is Poissonian.

Introduction

Consider an object. The object may change at some moment. The object never changes back. Changes occur due to some Poisson process parameterized by λ. We cannot observe changes of the object directly, we have to do some measurements.

Every Δ units of time we observe the object. If we compare two successive measurements there are only two possibilities: (no changes) $\bullet \to \bullet$; (a change) $\bullet \to \bullet'$. In the case of $\bullet \to \bullet$, we know for sure that the object has not been modified between these two observations. But in the case of $\bullet \to \bullet'$ we know only that the object has been changed at least once, but we don't know the exact number of changes.

How can we infer the most likely λ having only a sequence of observations?

λ	– parameter of Poisson process
$\langle X \rangle$	– "realization" of Poisson process
$\langle Y \rangle$	– sequence of observations

$$\begin{tikzcd} \lambda \arrow[r, "r"] & \langle X \rangle \arrow[d, "o"] \\ & \langle Y \rangle \arrow[ul, dashed, "?"] \end{tikzcd}$$

Estimator

We propose an estimator $\hat{\lambda}$ and calculate the bias $\mathbb{E}\left[\hat{\lambda} - \lambda\right]$. Consider a sequence of observations $\langle Y \rangle$: $\bullet_0, \bullet_\Delta, \bullet_{2\Delta}, \ldots, \bullet_{N\Delta}$. There are $N+1$ measurements. Between any successive pair of observations there is an interval of Δ

units of time. Therefore, we have a sequence of N intervals. Denote by W the number of intervals without changes ($\bullet \to \bullet$).

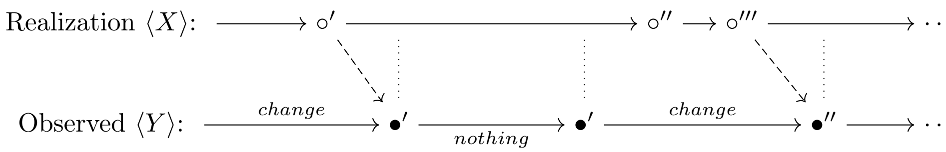

Figure 1: Schematic representation of real and observed process.

Estimator

$$\hat{\lambda} = -\Delta^{-1} \log \frac{W}{N}$$

Δ – *size of the time interval between observations*
W – *number of intervals without changes*
N – *total number of intervals*

For any Poisson process we have the following expression for the probability of exactly k changes in an interval of Δ units of time:

$$\Pr[k] = \frac{e^{-\lambda\Delta}(\lambda\Delta)^k}{k!} \qquad k = 0, 1, \ldots,$$

Due to memorylessness of Poisson process this probability does not depend of what happens in other Δ intervals. Next we write:

$$\Pr[\bullet \to \bullet] = \Pr[0] = e^{-\lambda\Delta} \tag{*}$$

We have W intervals without changes, therefore we can approximate:

$$\Pr[\bullet \to \bullet] \approx \frac{W}{N} \tag{**}$$

Combining (*) and (**) we obtain the estimator:

$$\begin{aligned} e^{-\lambda\Delta} &\approx \frac{W}{N} \\ \lambda\Delta &\approx -\log \frac{W}{N} \\ \lambda &\approx -\Delta^{-1} \log \frac{W}{N} \end{aligned}$$

Bias

$$\mathbb{E}\left[\hat{\lambda}-\lambda\right] \approx \frac{1}{2N\Delta}\left(e^{\lambda\Delta}-1\right)$$

In order to calculate the bias of $\hat{\lambda}$ we proceed as follows:

$$\mathbb{E}\left[\hat{\lambda}\right] = \mathbb{E}\left[-\Delta^{-1}\log W'\right] = -\Delta^{-1}\mathbb{E}\left[\log W'\right] \qquad W' = \frac{W}{N}$$

Using a second order Taylor expansion [1] we approximate:

$$\mathbb{E}\big[\log W'\big] \approx \log \mathbb{E}[W'] - \frac{\mathrm{Var}[W']}{2\big(\mathbb{E}[W']\big)^2} \tag{****}$$

It is easy to see that W is distributed binomially:

$$W \sim B(N, e^{-\lambda\Delta}),$$

where N is the number of trials, and $e^{-\lambda\Delta}$ is the probability of success. Now, we are able to calculate the expected value of W':

$$\mathbb{E}[W'] = \frac{\mathbb{E}[W]}{N} = \frac{Ne^{-\lambda\Delta}}{N} = e^{-\lambda\Delta},$$

and the variance of W':

$$\mathrm{Var}[W'] = \frac{\mathrm{Var}[W]}{N^2} = \frac{Ne^{-\lambda\Delta}(1-e^{-\lambda\Delta})}{N^2} = \frac{e^{-\lambda\Delta}(1-e^{-\lambda\Delta})}{N}$$

Now return to (****):

$$\begin{aligned}\mathbb{E}\big[\log W'\big] &\approx \log\left(e^{-\lambda\Delta}\right) - \frac{1}{N}\frac{e^{-\lambda\Delta}(1-e^{-\lambda\Delta})}{2e^{-2\lambda\Delta}} \\ &\approx -\lambda\Delta - \frac{1}{2N}\left(e^{\lambda\Delta}-1\right)\end{aligned}$$

And finally:

$$\begin{aligned}\mathbb{E}\left[\hat{\lambda}\right] &\approx \lambda + \frac{1}{2N\Delta}\left(e^{\lambda\Delta}-1\right) \\ \mathbb{E}\left[\hat{\lambda}-\lambda\right] &\approx \frac{1}{2N\Delta}\left(e^{\lambda\Delta}-1\right)\end{aligned}$$

We see that estimator $\hat{\lambda}$ is asymptotically unbiased in the following sense:

$$\lim_{N\to\infty}\mathbb{E}\left[\hat{\lambda}\right] = \lambda$$

Summary and discussions

Recall the question from the abstract: *if Maria performs more observations per unit of time than Maximilian, how can he estimates the Maria's results from his own?*

When the underlying process is Poissonian, Maximilian can use $\hat{\lambda}$ in order to estimate the true value of λ. Using this λ he actually knows all necessary information about the process. So, Maximilian can predict (in some sence) Maria's results.

Moreover, regardless of the delay between measurements, but with sufficient number of observations, we can estimate λ very accurately (because of $\lim_{N\to\infty} \mathbb{E}\left[\hat{\lambda}\right] = \lambda$).

The open question: what do we do when underlying prcoess is non-Poissonian?

References

[1] Y. W. Teh, D. Newman, and M. Welling, *A collapsed variational bayesian inference algorithm for latent dirichlet allocation*, in NIPS'06, 2006, pp. 1353–1360.