Lithuanian Journal of Statistics *2017, vol. 56, No 1, pp. 72–76*
Lietuvos statistikos darbai *2017, 56 t., Nr. 1, 72–76 p.*
www.statisticsjournal.lt

PREDICTION OF THE GEOMETRIC RENEWAL PROCESS

Vaidotas Kanišauskas[1], Karolina Piaseckienė[2]

Faculty of Technology, Physical and Biomedical Sciences, Šiauliai University.
Address: Vilniaus St. 141, Šiauliai, Lithuania.
E-mail: [1]vaidotaska@yahoo.com, [2]k.piaseckiene@gmail.com

Received: August 2017 Revised: October 2017 Published: December 2017

Abstract. The first part of the paper presents major concepts and theoretical statements on prediction of processes. The second part presents the obtained results on the geometric renewal process by indicating its distribution which has a binomial distribution and is a process with independent and stationary increments. Further, having applied the theory introduced in the first part to the geometric renewal process, the sufficient and unbiased prediction with the minimum-variance has been found.

Keywords: renewal process, binomial distribution, prediction, unbiased prediction.

1. Introduction

The concept of prediction sufficiency was introduced by K. Takeuchi and M. Akahira (1975). The primary application of prediction sufficiency was demonstrated by E. N. Torgersen (1977). More comprehensive applications of this concept were demonstrated by B. Johansson (1990).

It is shown that much of the classical theory of unbiased parameter estimation can be transferred to a predictive setting. The main object of the present papers [7, 3] is to develop these ideas further and, in particular, to study a close connection which exists between unbiased prediction and time reversal of Markov processes (Björk & Johansson, 1992). Johansson (1990) replaced the usual sufficiency concept by that of prediction sufficiency, so the Rao-Blackwell and Lehmann-Scheffé theorem can be rephrased to suit the above context.

The return from prediction to the parameter estimation theory, enriching the latter by the new findings obtained after prediction, was demonstrated by T. Björk and B. Johansson (1996). These studies investigated Poisson processes, the Yule model, a Wiener process with the unknown drift, diffusion with the unknown drift, and the geometric Brownian motion.

The aim of this research is to find the minimum variance unbiased predictor of the geometric renewal process N_t, $t > s$, based on observations $\{N_u, 0 \le u \le s\}$.

Major concepts and results of prediction of processes introduced in the second section of the paper have been mostly based on the research study [3]. This is displayed in a similar manner in [1], too. In the third section, using paper [6], we introduce a definition of the geometric renewal process, demonstrate that it has a binomial distribution and is a process with independent and stationary increments. The geometric renewal process is called by some authors the discrete Poisson process [9] which, together with the continuous Poisson process, is considered to be classical in the theory of renewal processes. Therefore they are often investigated in monographs dealing with this theory. At the end of the section, the form of the process of local density (Radon-Nikodym derivative) of the geometric renewal process, taken from [5], is presented. Basic concepts of renewal processes are presented in [4]. The fourth section displays the found UMSEUP ("Uniformly Minimum Squared Error Unbiased Predictor") predictions of the renewal process, both when a parameter is unknown and when it is known.

2. Unbiased prediction

We now recall the definitions of a prediction sufficient statistic and main theorems (see e. g. [3, 7]). We consider some sample space Ω and two σ-algebras $\mathcal{F}_1$ and $\mathcal{F}_2$, where $\mathcal{F}_1$ is generated by some set of random variables which we observe, and $\mathcal{F}_2$ is generated by a set of (yet) unobserved variables. We also have a family $\mathcal{P}$ of probability measures on

$(\Omega, \mathcal{F}_1 \vee \mathcal{F}_2)$. The objective is to predict some square integrable, $\mathcal{F}_2$-measurable random variable (r.v.) W. A predictor is any square-integrable, $\mathcal{F}_1$-measurable r.v. X. The performance of the predictor X is evaluated by its quadratic loss function $P \to \mathbb{E}_\theta[(X - W)^2]$, $P \in \mathcal{P}$. The predictor X is called unbiased, if $\mathbb{E}_P[X] = \mathbb{E}_P[W], \forall P \in \mathcal{P}$. The predictor X is said to be complete for $\mathcal{P}$ if, for every fixed Borel-function g, the condition

$$\mathbb{E}_P[g(X)] = 0,\ P \in \mathcal{P}$$

implies

$$g(X) = 0,\ P \text{ – almost surely (a.s.).}$$

Definition 1. An $\mathcal{F}_1$-measurable statistic Y is said to be prediction-sufficient with respect to (w.r.t.) $(\mathcal{F}_1, \mathcal{F}_2, \mathcal{P})$, if Y is sufficient w.r.t. $\mathcal{P}$ restricted on $\mathcal{F}_1$, i.e. for every bounded $\mathcal{F}_1$-measurable r.v. Z, there exists a common version of $\mathbb{E}_P[Z|Y]$, $P \in \mathcal{P}$; and for every $P \in \mathcal{P}$, $\mathcal{F}_1$ and $\mathcal{F}_2$ are conditionally independent given Y.

Theorem 1. (Rao–Blackwell) [3]. Suppose that Y is prediction-sufficient w.r.t. $(\mathcal{F}_1, \mathcal{F}_2, \mathcal{P})$. Let X be an $\mathcal{F}_1$-measurable predictor of the $\mathcal{F}_2$-measurable variable W. Then the predictor $\varphi(Y) = \mathbb{E}_P[X|Y] = \mathbb{E}[X|Y]$ satisfies

$$\mathbb{E}_P[(\varphi(Y) - W)^2] \le \mathbb{E}_P[(X - W)^2],\ \forall P \in \mathcal{P}.$$

Theorem 2. (Lehmann–Scheffé) [3]. Assume that the statistic Y is prediction-sufficient w.r.t. $(\mathcal{F}_1, \mathcal{F}_2, \mathcal{P})$ and complete w.r.t. $\mathcal{P}$. Also assume that there exists some $\mathcal{F}_1$-measurable unbiased predictor X of W. The predictor $\varphi(Y) = \mathbb{E}[X|Y]$ then satisfies

$$\mathbb{E}_P[(\varphi(Y) - W)^2] \le \mathbb{E}_P[(Z - W)^2],\ \forall P \in \mathcal{P},$$

for every $\mathcal{F}_1$-measurable unbiased predictor Z. It is also unique, P-a.s. unique with this property.

Definition 2. An unbiased predictor X of W is UMSEUP if, for every other unbiased predictor Z,

$$\mathbb{E}_P[(X - W)^2] \le \mathbb{E}_P[(Z - W)^2],\ \forall P \in \mathcal{P}.$$

Corollary (Theorem 2)**.** If we have a complete and prediction sufficient statistic (prediction) Y and can find the function $f(Y)$ such that $\mathbb{E}_P[f(Y)] = \mathbb{E}_P[W]$, $\forall P \in \mathcal{P}$, then $f(Y)$ is UMSEUP.

3. Geometric renewal process

On a stochastic basis $(\Omega, \mathcal{F}, \mathbb{F}, P_\theta, \theta \in \Theta)$, where Θ is an abstract space, let there be given a counting process $N_t = \sum_{n=1}^{\infty} \mathbf{1}(T_n \le t)$, $t \ge 0$, such that the random moments $X_i = T_i - T_{i-1}$, $i = 1, 2, \ldots$ $(T_0 = 0)$ are independent and identically distributed. Such a process is called a renewal process.

We consider a geometric renewal process $N_t = \sum_{n=1}^{\infty} \mathbf{1}(T_n \le t)$, $t \ge 0$, with random variables $X_i = T_i - T_{i-1}, i = 1, 2, \ldots$, having the geometric distribution $P_\theta(X_i = k) = \theta(1-\theta)^{k-1}$, $\theta \in \Theta = (0,1)$, $k = 1, 2, \ldots$.

We can note that by the renewal process definition

$$P_\theta\{N_t = 0\} = P\{t < X_1\}$$

and

$$P_\theta\{N_t = k\} = P\{X_1 + \cdots + X_k \le t < X_1 + \cdots + X_k + X_{k+1}\},\ k = 1, 2, \ldots .$$

Theorem 3. [6]. Suppose $T_n = \sum_{i=1}^{n} X_i$ for $n \ge 1, T_0 = 0$ and $F_n(t) = P_\theta(T_n \le t)$. Then

$$F_n(x) = 1 - (1-\theta)^x \sum_{i=0}^{n-1} C_x^i \left(\frac{\theta}{1-\theta}\right)^i,\ \ x = n, n+1, \ldots,$$

and

$$P_\theta\{N_t = n\} = C_{[t]}^n \theta^n (1-\theta)^{[t]-n}, n = 0, 1, 2, \ldots, [t].$$

It is a formula of binomial distribution with parameters $([t], \theta)$. We shall characterize this renewal process.

Now, let N_t be a renewal process, determined by successive interarrival times $X_1, X_2, \ldots$ that are independent and identically distributed (i.i.d.) nonnegative integer r.v.s.

Condition A.

For each ω, $N_t(\omega)$ is a nonnegative integer as $t \ge 0, N_0(\omega) = 0$ and $\lim_{t\to\infty} N_t(\omega) = \infty$. Further, for each ω, $N_t(\omega)$ as a function of t is non-decreasing and right-continuous and the point of discontinuity $N_t(\omega) - \sup_{s<t} N_s(\omega)$ is exactly 1.

Condition B.

$X_1, X_2, \ldots$ are independent, having the geometric distribution with the parameter θ.

Condition C.

(a) For $0 < t_0 < t_1 < \cdots < t_k$ increments $N_{t_1}, N_{t_2} - N_{t_1}, \ldots, N_{t_k} - N_{t_{k-1}}$ are independent and

$$P_\theta\{N_{t_i} - N_{t_{i-1}} = n_i, 1 \le i \le k\} = \prod_{i=1}^{k} P_\theta\{N_{t_i - t_{i-1}} = n_i\}.$$

(b) The increments have the binomial distribution, i. e.

$$P_\theta\{N_t - N_s = k\} = C_{[t]-[s]}^k \theta^k (1-\theta)^{[t]-[s]-k}, \quad \theta \in \Theta, \ 0 \le [s] < [t], \ k = 0, 1, \dots, [t]-[s].$$

Theorem 4. [6]. Conditions B and C are equivalent in the presence of condition A.

By [9] the simplest example of the renewal process is the discrete Poisson process with a shifted geometric renewal distribution

$$p_k = P_\theta(X_i = k) = \theta(1-\theta)^{k-1}, \theta \in \Theta = (0,1), k = 1, 2, \dots .$$

However, by [9] the simplest way to define the discrete Poisson process is by introducing an i.i.d. sequence $Y_1, Y_2, \dots$ of $\{0; 1\}$ r.v.s

$$Y_i = \begin{cases} 1, & \text{with probability } \theta, \\ 0, & \text{with probability } 1-\theta, \end{cases}$$

and assuming $N_0 = 0$ and $N_t = \sum_{j=1}^{[t]} Y_j, t \ge 0$.

It follows that N_t is binomially distributed and, therefore, the discrete Poisson process is also called a binomial process.

We denote $\mathcal{F}_t^N = \sigma(N_s, s \le t)$ and suppose that $\mathbb{F} = (\mathcal{F}_t^N)_{t \ge 0}$. Let $\theta, \theta_0 \in \Theta = (0,1)$ and $\theta \ne \theta_0$. Then [5] $P_\theta^t \sim P_{\theta_0}^t$, where $P_\theta^t = P_\theta | \mathcal{F}_t^N$, and Radon-Nikodym derivative is

$$\frac{dP_\theta^t}{dP_{\theta_0}^t} = \left(\frac{\theta}{\theta_0}\right)^{N_t} \left(\frac{1-\theta}{1-\theta_0}\right)^{[t]-N_t}. \tag{1}$$

It is easy to prove that this formula is correct valid grounding on the equality

$$\mathbb{E}_{\theta_0} \frac{dP_\theta^t}{dP_{\theta_0}^t} = \int \frac{dP_\theta^t}{dP_{\theta_0}^t}(\cdot) dP_{\theta_0}^t = \sum_{k=0}^{[t]} \left(\frac{\theta}{\theta_0}\right)^k \left(\frac{1-\theta}{1-\theta_0}\right)^{[t]-k} P_{\theta_0}(N_t = k) = \sum_{k=0}^{[t]} \left(\frac{\theta}{\theta_0}\right)^k \left(\frac{1-\theta}{1-\theta_0}\right)^{[t]-k} C_{[t]}^k \theta_0^k (1-\theta)^{[t]-k}$$

$$= \sum_{k=0}^{[t]} C_{[t]}^k \theta^k (1-\theta)^{[t]-k} = 1.$$

4. Model for predicting of the geometric renewal process

Let us observe the period $u \in (0, s)$ of the geometric renewal process N_u. Our aim is to obtain prediction of N_t, $t > s$, according to these observations of the process. To this end, we define a family of probability measures $\mathcal{P} = \{P_\theta, \ \theta \in \Theta =]0,1[\}$ and σ-algebra $\mathcal{F}_{(s,t)} = \sigma\{N_u, \ s < u < t\}$. We shall obtain the prediction model $(\mathcal{F}_{[0,s]}, \ \mathcal{F}_{(s,t]}, \ \mathcal{P})$.

1. Finding of a sufficient statistic.

Since the likelihood ratio of the process N_s, given observations $\{N_u, 0 \le u \le s\}$, is

$$\frac{dP_\theta^s}{dP_{\theta_0}^s} = \left(\frac{\theta}{\theta_0}\right)^{N_s} \left(\frac{1-\theta}{1-\theta_0}\right)^{[s]-N_s},$$

according to the factorisation theorem, N_s is a sufficient statistic to estimate the parameter θ.

2. Finding of a complete statistic.

For every fixed Borel function g

$$\mathbb{E}_\theta g(N_s) = \sum_{k=0}^{\infty} g(k)\, C_{[s]}^k \theta^k (1-\theta)^{[s]-k} = 0, \text{ for all } \theta \in \Theta \ =]0,1[.$$

This implies that

$$g(0) = g(1) = \dots = g([s]) = 0.$$

Therefore the statistic N_s is complete.

Corollary 1. Statistic N_s is complete and sufficient.

Corollary 2. Since the geometric renewal process N_s has the binomial distribution and is a process with independent increments, that has a valid expression

$$N_s = \sum_{j=1}^{[s]} Y_j, \ \text{ i.i.d. } Y_i = \begin{cases} 1, & \text{with probability } \theta, \\ 0, & \text{with probability } 1-\theta, \end{cases}$$

$\mathcal{F}_{[0,s]}$ and $\mathcal{F}_{(s,t]}$ are independent w.r.t. $\mathcal{P}$. Thus, the geometric renewal process N_s is prediction-sufficient w.r.t. $(\mathcal{F}_{[0,s]}, \ \mathcal{F}_{(s,t]}, \ \mathcal{P})$ and complete w.r.t. $\mathcal{P}$. Therefore by Corollary of Theorem 2, the predictor of the process N_t, $t > s$, is $f(N_s)$, where the function f is unknown as yet.

Since the process N_s has the binomial distribution with parameters θ and $[s]$, we obtain

$$\mathbb{E}_\theta N_s = \theta[s], \ \theta \in \Theta.$$

Consequently, we should find a function f such that

$$\mathbb{E}_\theta f(N_s) = \theta[t], \ \ \theta \in \Theta,$$

because $\mathbb{E}_\theta N_t = \theta[t], \ \ \theta \in \Theta$.

Case 1. Let the distribution parameter $\theta, \theta \in \Theta$ be unknown. Then, the function (predictor) f will be sought for as

$$f(N_s) = \alpha N_s,$$

where α is yet unknown constant.

Hence we obtain that

$$\mathbb{E}_\theta f(N_s) = \mathbb{E}_\theta \alpha N_s = \alpha \mathbb{E}_\theta N_s = \alpha\theta[s].$$
$$\alpha\theta[s] = \theta[t].$$

Next,

$$\alpha = \frac{[t]}{[s]}.$$

Thus, we derive:

$$f(N_s) = \frac{[t]}{[s]} N_s = N_s + \frac{N_s}{[s]}([t] - [s]).$$

Therefore, according to the theory, we obtain that

$$f(N_s) = N_s + \frac{N_s}{[s]}([t] - [s])$$

is the unbiased predictor of N_t and then UMSEUP according to Theorem 2.

Case 2. Let the parameter $\theta, \theta \in \Theta$ be known.

According to the theory, the best predictor of N_t is a conditional expectation (also see [3], Introduction):

$$\varphi(N_s) = \mathbb{E}_\theta(N_t | N_s).$$

Since the process N_s is a process with independent increments, it follows that

$$\varphi(N_s) = \mathbb{E}_\theta(N_t | N_s) = \mathbb{E}_\theta(N_s + N_t - N_s | N_s) = N_s + \theta([t] - [s])$$

is the prediction of N_t which is unbiased and UMSEUP. When basing on formula (1), it is easy to prove that, according to observations $\{N_u, u \in [0, s]\}$, the maximum likelihood estimator of the parameter θ is $\frac{N_s}{[s]}$. Hence we can understand the relation between optimal estimators and optimal predictors.

5. Conclusions

1. By applying the theorem of factorization and the definition of the complete statistic, we have found out that the statistic N_s is prediction-sufficient w.r.t. $(\mathcal{F}_{[0,s]}, \ \mathcal{F}_{(s,t]}, \ \mathcal{P})$ and complete w.r.t. $\mathcal{P}, \mathcal{P} = \{P_\theta, \ \theta \in \Theta \ =]0,1[\}$.
2. In case the model's parameter θ is unknown, the best predictor (UMSEUP) of the geometric renewal process N_t is $f(N_s) = N_s + \frac{N_s}{[s]}([t] - [s])$.
3. In case the model's parameter θ is known, the best predictor (UMSEUP) of the geometric renewal process N_t is $\varphi(N_s) = N_s + \theta([t] - [s])$.

Acknowledgements

The authors are grateful to anonymous referees for their constructive critiques and helpful comments and to the second referee for pointing out our attention to the paper by Samimi and Mohammadi (2013).

References

1. Björk T., and Johansson B. Adaptive Prediction and Reverse Martingales. *Stochastic Processes and their Applications*. 1992, 43: 191–222.
2. Björk T., and Johansson B. Parameter Estimation and Reverse Martingales. *Stochastic Processes and their Applications*. 1996, 63: 235–263.
3. Johansson B. Unbiased Prediction in the Poisson and Yule Processes. *Scandinavian Journal of Statistics*. 1990, 17(2): 135–145.
4. Karr A.F. *Point Processes and Their Statistical Inference*. 1986, New York: Macrel Dekker.
5. Pikturnaitė K. *Atstatymo proceso tikimybinių matų asimptotinis atskiriamumas*. Magistro darbas. 2011, Šiaulių universitetas. *gs.elaba.lt/object/elaba:1851549/1851549.pdf* .
6. Samimi H., and Mohammad Mohammadi. A new approach on the renewal process with Geometric interarrival times. *International Research Journal of Applied and Basic Sciences*. 2013, 4(6): 1530–1534.
7. Takeuchi K., and Akahira M. Characterizations of Prediction Sufficiency (Adequacy) in Terms of Risk Functions. *The Annals of Statistics.* 1975, 3(4): 1018–1024.

8. Torgersen E.N. Prediction Sufficiency When the Loss Function Does Not Depend on the Unknown Parameter. *The Annals of Statistics.* 1977, 5(1): 155–163.
9. Van der Weide J.A.M., Pandey M.D., van Noortwijk J.M. A conceptual interpretation of the renewal theorem with applications. *Risk, Reliability and Societal Safety*. 2007, 1: 477–484.

GEOMETRINIO ATSTATYMO PROCESO PROGNOZAVIMAS

Vaidotas Kanišauskas, Karolina Piaseckienė

Santrauka. Pirmoje straipsnio dalyje pateiktos pagrindinės sąvokos ir teoriniai teiginiai apie procesų prognozavimą. Antroje darbo dalyje pateikiami žinomi rezultatai apie nagrinėjamą geometrinį atstatymo procesą, nurodant jo skirstinį, kuris, pasirodo, turi binominį skirstinį ir yra procesas su nepriklausomais ir stacionariais pokyčiais. Geometriniam atstatymo procesui pritaikius pirmos dalies teoriją, surandama prognoziškai pakankama ir nepaslinktoji prognozė, turinti tolygiai mažiausią dispersiją.

Reikšminiai žodžiai: atstatymo procesas, binominis skirstinys, prognozė, nepaslinktoji prognozė.