# Linear Prediction of Sum of Two Poisson Process

**S. H. Raheem**
**Faculty of Human Science and Physical Education /School of Human Science,University of Garmian/Iraq**
**Received in: 17 March 2012 Accepted in: 21 May 2012**

## Abstract

Our goal from this work is to find the linear prediction of the sum of two Poisson process $Z(t)=X(t)+Y(t)$ at the future time $Z(t+\tau), \tau \geq 0$ and that is when we know the values of $Z(t)$ in the past time and the correlation function $\beta_z(\tau)$

**Key words:** Poisson process, linear Prediction, sum of two Poisson process.

## Introduction

The Poisson process has long tradition in applied probability and stochastic process theory, in (1903) thesis Fillip Lundberg already exploited it as a model for the claim number process N later on in the $1930_s$ , Gramer the famous statistician extensively developed collective risk theory by using the total claim amount process S with arrivals $T_i$ which are generated by Poisson process for historical reasons but also since it has very attractive mathematical properties the Poisson process plays a central role in insurance mathematics[1,p7].Bellow we will give a definition of Poisson process.

A Poisson process with parameter or rate $\lambda > 0$ is an integer- valued, continuous time stochastic process $\{X(t), t \geq 0\}$ satisfying :-

**1.** $X(0)=0$

**2.** for all $t_0 = t_1 < t_2 < \ldots < t_n$ the increment $X(t_1) - X(t_0)$, $X(t_2) - X(t_1)$, …, $X(t_n) - X(t_{n-1})$ are independent random variables.

**3.** for $t \geq 0$, $s > 0$ and non- negative integers k, the increments have the Poisson distribution

$$\Pr[X(t+s)-X(s)=k]=\frac{(\lambda t)^k e^{-\lambda t}}{k!}, \; k = 0, 1, 2, \ldots \qquad \ldots(1)$$

It is convenient to view the Poisson process X(t) as a special counting process of events in any interval of length t specified via condition (3), [2,p120].

We have to know that Poisson process is on of the most important examples of Marcov process that plays an important role in both theory and in variety of applications, [3,p346]. An arrival is simply an occurrence of some event-like a phone call, job offer or whatever that happens at a particular point in time [4].

Our goal from this work is to find the linear prediction of the sum of two independent homogenous Poisson process $\{X(t), t \in T\}, \{Y(t), t \in T\}$ with parameters $\lambda_1$, $\lambda_2$, (i.e.) to find the linear prediction of:

$\{Z(t), t \in T\} = \{X(t), t \in T\} + \{Y(t), t \in T\}$ at the future time $(t+\tau)$, $\tau \geq 0$.

## Theorem [2]

Let X(t), Y(t) be two independent Poisson process with parameters $\lambda_1$ and $\lambda_2$ respectively, then X(t)+Y(t) is also a Poisson process with density $\lambda_1 + \lambda_2$

**Proof:**

Since $X(t)+Y(t)$ is an additive Markov process we let:

$Z(t) = X(t)+Y(t)$ ,so

$$\Pr\{Z(s+t)\cdot Z(s)=\eta\} = \sum_{j=o}^{n} \Pr\{X(s+t)-X(s)=j\}\cdot \Pr\{Y(s+t)-Y(s)=n-j\}$$

$$= \sum_{j=0}^{n} e^{\lambda_1 t}\frac{(\lambda_1 t)^j}{j!}\cdot e^{\lambda_2 t}\frac{(\lambda_2 t)^{n-j}}{(n-j)!}$$

$$= \frac{e^{-(\lambda_1+\lambda_2)t}}{n!}\sum_{j=0}^{n}\frac{n!}{j!(n-j)!}(\lambda_1 t)^j(\lambda_2 t)^{n-j}$$

$$Z(t) = \frac{e^{-(\lambda_1+\lambda_2)t}}{n!}(\lambda_1 t+\lambda_2 t)^n \qquad \ldots(2)$$

$\therefore$ Z(t) is homogenous Poisson process

## Moments of Z(t)

We can find the first moment of Z(t) as follows

$$E[Z(t),t] = \sum_{Z(t)=0}^{\infty} Z(t) f_z(t)$$

$$= \sum_{Z(t)=0}^{\infty} Z(t) e^{-(\lambda_1+\lambda_2)t}\frac{[(\lambda_1+\lambda_2)t]^{Z(t)}}{Z(t)!}$$

$$= e^{-(\lambda_1+\lambda_2)t}\sum_{Z(t)=1}^{\infty}\frac{[(\lambda_1+\lambda_2)t]^{Z(t)-1+1}}{[Z(t)-1]!}$$

$$= (\lambda_1+\lambda_2)te^{-(\lambda_1+\lambda_2)t}\sum_{Z(t)=1}^{\infty}\frac{[(\lambda_1+\lambda_2)t]^{Z(t)-1}}{[Z(t)-1]!}$$

$$= (\lambda_1+\lambda_2)t$$

Know to find the second moment

$$E[Z^2(t),t] = \sum_{Z(t)=0}^{\infty} Z^2(t) f_z(t)$$

$$= \sum_{Z(t)=0}^{\infty} Z^2(t) e^{-(\lambda_1+\lambda_2)t}\frac{[(\lambda_1+\lambda_2)t]^{Z(t)}}{Z(t)!}$$

$$= (\lambda_1+\lambda_2)te^{-(\lambda_1+\lambda_2)t}\sum_{Z(t)=1}^{\infty} Z(t)\frac{[(\lambda_1+\lambda_2)t]^{Z(t)-1}}{[Z(t)-1]!}$$

Let $K(t) = Z(t)-1$ then we get

$$= (\lambda_1+\lambda_2)te^{-(\lambda_1+\lambda_2)t}\sum_{Z(t)=1}^{\infty}(K(t)+1)\frac{[(\lambda_1+\lambda_2)t]^{K(t)}}{K(t)!}$$

$$= [(\lambda_1+\lambda_2)t]e^{-(\lambda_1+\lambda_2)t}\sum_{K(t)=0}^{\infty}\frac{(\lambda_1+\lambda_2)t]^{K(t)-1+1}}{[K(t)-1]!}+(\lambda_1+\lambda_2)$$

$$\therefore E[Z^2(t),t]=[(\lambda_1+\lambda_2)t]^2+(\lambda_1+\lambda_2)t \qquad \ldots(3)$$

So that:

$$Var[Z(t),t]=E[Z^2(t),t]-[E(Z(t),t)]^2$$
$$=(\lambda_1+\lambda_2)t \qquad \cdots(4)$$

## Correlation Function of Z(t)

We have:

$$Z(t)=X(t)+Y(t)$$

The correlation function of Z(t) is:

$$\beta_z(\tau)=\beta_X(\tau)+\beta_Y(\tau)$$

Or: $\beta_z(t,t+\tau)=\beta_X(t,t+\tau)+\beta_Y(t,t+\tau)$

$$=E[X(t)X(t+\tau)]+E[Y(t)Y(t+\tau)]$$
$$=E\{X(t)X(t+\tau)-E[X(t)]E[X(t+\tau)]\}+E\{[Y(t)Y(t+\tau)]-E[Y(t)]E[Y(t+\tau)]\}$$
$$=E[X^2(t)]+E[X(t)X(t+\tau]-E[X^2(t)]-E[X(t)]E[X(t+\tau)]]+E[Y^2(t)]+ E[Y(t)Y(t+\tau)]-E[Y^2(t)]-E[Y(t)]E[Y(t+\tau)]$$
$$=E[X^2(t)]+E[X(t)]E[X(t+\tau)-X(t)]-E[X(t)]E[X(t+\tau)]+E[Y^2(t)]+ E[Y(t)E[Y(t+\tau)-Y(t)]-E[Y(t)]E[Y(t+\tau)]$$
$$=E[X^2(t)]+[E[X(t)]]^2+E[Y^2(t)]+[E[Y(t)]]^2$$
$$=VarX(t)+VarY(t)$$
$$\therefore \beta_z(\tau)=\lambda_1 t+\lambda_2 t$$
$$=(\lambda_1+\lambda_2)t \qquad \ldots(5)$$

## The Formulation of The Status When Finite Numbers of Values of Z(t) Are Known

The idea of best linear prediction is very important linear prediction theory has important application in standard linear models and the analysis of special data the theory has traditionally been taught as part of multivariate analysis. It is important for general stochastic process, time series and it is the basis for Linear Bayesian method [5, p134]

Linear prediction is a mathematical operation where future values of discrete – time signals are estimated as a linear function of previous sample [6].

We shall express here how to predict Z(t) for the future time $(t+\tau)$ and for that we need to know the finite values of the past time and the correlation function $\beta_z(\tau)$. Suppose we know finite value of Z(t) of the past time:

$Z^{(1)}(t-s_1),Z^{(1)}(t-s_2),\ldots,Z^{(1)}(t-s_n)$; i.e.

$$\tilde{Z}(t+\tau)=g\{Z^{(1)}(t-s_1),Z^{(1)}(t-s_2),\ldots,Z^{(1)}(t-s_n)\} \qquad \ldots(6)$$

And the prediction is equivalent to the determining value of coefficients $\alpha_1, \alpha_2, \ldots, \alpha_n$ of the linear combination :

$$\alpha_1 Z(t-s_1)+\alpha_2 Z(t-s_2)+\ldots+\alpha_n Z(t-s_n) \qquad \ldots(7)$$

[7, p.97] so:

$$g\{Z^{(1)}(t-s_1),Z^{(1)}(t-s_2),\dots,Z^{(1)}(t-s_n)\}=\sum_{k=1}^{n}\alpha_k Z(t-s_k) \qquad \dots(8)$$

Substitute (8) in (6) we get :

$$\tilde{Z}(t+\tau)=\sum_{k=1}^{n}\alpha_k Z(t-s_k) \qquad \dots(9)$$

The difference between the point $Z^{(1)}(t+\tau)$ and its predicted value $\tilde{Z}(t+\tau)$ represents the prediction error and denoted by :

$$\sigma^2_{\tau,n}=Z^{(1)}(t+\tau)-\tilde{Z}(t+\tau)$$

We shall consider here only one realization $Z^{(1)}(t-s_k)$ ,k =1,2,….,n and that is because we can not take the absolute value of the prediction error $|\zeta_{\tau,n}|$ as an index of the quality of prediction formula since it will be different for different realization of Z(t) and :

$$\sigma^2_{\tau,n}=E\left|Z(t-\tau)-\sum_{k=1}^{n}\alpha_k Z(t-s_k)\right|^2 \qquad \dots(10)$$

takes its minimum value and the formula (10) called the mean square prediction error see[4, p145]

$$\sigma^2_{\tau,n}=[Z(t-\tau)-\sum_{k=1}^{n}\alpha_k Z(t-s_k)][\overline{Z(t-\tau)}-\overline{\sum_{k=1}^{n}\alpha_k Z(t-s_k)}] \qquad \dots(11)$$

$$=E[Z(t+\tau)\overline{Z(t+\tau)}]-E[Z(t+\tau)\overline{\sum_{k=1}^{n}\alpha_k Z(t-s_k)}]-E[\sum_{k=1}^{n}\alpha_k Z(t-s_k)\overline{Z(t-\tau)}+E[\sum_{k=1}^{n}\alpha_k Z(t-s_k)]$$

Hence:

$$\sigma^2_{\tau,n}=\beta[t+\tau-(t+\tau)]-E[Z(t+\tau)\sum_{k=1}^{n}\overline{\alpha Z(t-s_k)}$$

$$-E[\sum_{k=1}^{n}\alpha_k Z(t-s_k)\overline{Z(t+\tau)}+E[\sum_{k=1}^{n}\alpha_k Z(t-s_k)\sum_{k=1}^{n}\alpha_k Z(t-s_k)$$

Therefore:

$$\sigma^2_{\tau,n}=\beta(0)-\sum_{k=1}^{n}\bar{\alpha}\beta(t+\tau)\sum_{k=1}^{n}\alpha_k\beta[-(t+s_k)]+\sum_{\ell=1}^{n}\sum_{k=1}^{n}\alpha_k\alpha_\ell\beta(s_k-s_\ell)$$

Then:

$$\sigma^2_{\tau,n}=\beta(0)-2\operatorname{Re}\sum_{k=1}^{n}\alpha_k\beta(t+s_k)+\sum_{\ell=1}^{n}\sum_{k=1}^{n}\alpha_k\alpha_\ell\beta(s_k-s_\ell) \qquad \dots(12)$$

Now we have to find the value of $\alpha_1=a_1,\alpha_2=a_2,\dots,\alpha_n=a_n$ in which $\sigma^2_{\tau,n}$ takes its Minimum value, by writing (12) as follows:

$$\sigma^2_{\tau,n}=\beta(0)-\sum_{k=1}^{n}\bar{\alpha}\beta(t+\tau)-\sum_{k=1}^{n}\alpha_k\beta(-t-s_k)+\sum_{\ell=1}^{n}\sum_{k=1}^{n}\alpha_k\alpha_\ell\beta(s_k-s_\ell)$$

Then by:

$$\left.\frac{\partial\sigma_{\tau,n}}{\partial\alpha_k}\right|_{\alpha_1=a_1,\alpha_2=a_2,\dots,\alpha_n=a_n=0}=0$$

We can fined the values $\alpha_1 = a_1, \alpha_2 = a_2, \ldots, \alpha_n = \alpha_n$ where $\sigma^2_{\tau,n}$ takes its minimum value and:

$$\left.\frac{\partial \sigma_{\tau,n}}{\partial \alpha_k}\right|_{\alpha_1 = a_1, \alpha_2 = a_2, \ldots, \alpha_n = a_n = 0} = 0 - \beta(\tau + s_k) - 0\sum_{\ell=1}^{n} \alpha 1 \beta(s_k - s\ell) = 0$$

$$= -\beta(\tau + s_k) + \sum_{\ell=1}^{n} a_\ell \beta(s_k - s_\ell) = 0$$

$$\beta(\tau + s_k) = \sum_{\ell=1}^{n} a_\ell \beta(s_k - s_\ell) \qquad \ldots(13)$$

then we can write (9) by :

$$\tilde{Z}(t + \tau) = \sum_{k=1}^{n} a_k Z(t + \tau) \qquad \ldots(14)$$

From (13) and since we have the correlation function of the sum of tow Poisson process:

$\beta_{Z(\tau)} = (\lambda_1 + \lambda_2)\tau$

we get the following system of equation

$$a_1\beta(s_1 - s_1) + a_2\beta(s_1 - s_2) + \cdots + a_n\beta(s_1 - s_n) = \beta(\tau + s_1)$$
$$a_1\beta(s_2 - s_1) + a_2\beta(s_2 - s_2) + \cdots + a_n\beta(s_2 - s_n) = \beta(\tau + s_2)$$
$$\vdots$$
$$a_1\beta(s_n - s_1) + a_2\beta(s_n - s_2) + \cdots + a_n\beta(s_n - s_n) = \beta(\tau + s_n)$$

Or:

$$0 + a_2(s_1 - s_2) + \cdots + a_n(s_1 - s_n) = (\tau + s_1)$$
$$a_1(s_1 - s_1) + 0 + \cdots + a_n(s_1 - s_n) = (\tau + s_2)$$
$$\vdots$$
$$a_1(s_1 - s_1) + a_2(s_1 - s_2) + \cdots + a_{n-1}\beta(s_1 - s_{n-1}) + 0 = (\tau + s_1)$$

where this system can be solved to obtain the values of the coefficients $a_1, a_2, \ldots, a_n$ after that we can write the beast prediction formula (1.4.12) by using $a_1, a_2, \ldots, a_n$ and the know observed values $Z(t - s_1), Z(t - s_2), \cdots, Z(t - s_n)$ by

$$\tilde{Z}(t + \tau) = a_1(t - s_1) + a_2 Z(t - s_2) + \cdots + a_n Z(t - s_n) \qquad ;\tau \geq 0$$

## The Mean Square Error of Prediction

Since we have

$$\sigma^2_{\tau,n} = \beta(0) - \sum_{k=1}^{n} \overline{\alpha_k}\beta(t + s_k) - \sum_{k=1}^{n} \alpha_k \beta(-t - s_k) + \sum_{\ell=1}^{n}\sum_{k=1}^{n} \alpha_k \alpha_\ell \beta(s_k - s_\ell)$$

And by (2.2.11) we get:

$$\sum_{k=1}^{n} \overline{\alpha_k}\beta(t + s_k) = \sum_{\ell=1}^{n}\sum_{k=1}^{n} \overline{\alpha_k}\alpha_\ell \beta(s_k - s_\ell)$$ see [1,p101]

$$\text{so}: \sigma^2_{\tau,n} = \beta(0) - \sum_{\ell=1}^{n}\sum_{k=1}^{n}\alpha_k\alpha_\ell\beta(s_k - s_\ell) - \sum_{k=1}^{n}\overline{a_k}\beta(\tau + s_k) + \sum_{\ell=1}^{n}\sum_{k=1}^{n}\overline{\alpha_k}\alpha_\ell\beta(s_k - s_\ell)$$

$$\therefore \ \sigma^2_{\tau,n} = \beta(0) - \sum_{\ell=1}^{n}\sum_{k=1}^{n}\alpha_k\alpha_\ell\beta(s_k - s_\ell)$$

And by using $\beta_Z(\tau) = (\lambda_1 + \lambda_2)\tau$

$$\sigma_{\tau,n} = \sum_{k=1}^{n} a_k(\lambda_1 + \lambda_2)(\tau + s_k)$$

$$= \sum_{\ell=1}^{n}\sum_{k=1}^{n}\alpha_k\alpha_\ell\beta(s_k - s_\ell)$$

# References

1. Micosch, Th. (2009) Non–Life Insurance Mathematic, second edition-Verlag Barlin Heidelberg.
2. WWW.nas.its.tudelft.nl/people/piet/cupboockapters/pacup.
3. Gnedenko, B.V. (1962) the theory of probability, New York.
4. Yaglom, A.M, (1962) An Introduction to The Theory of Stationary Random Function, Prentice Hall.
5. Christensen,R. (2011) Plane Answers to Complex Question, fourth edition, science Business Media, LLC.
6. en.wikipedia.org/wiki/Linear_prediction
7. Winner .N. (1999) Extrapolation, Interpolation and Smoothing of Stationary Time series, New York.
8. WWW.ssc.upenn.edu/rwight/cours/poisson.

# التنبؤ الخطي لمجموع عمليتي بوايسون

سيران حمزة كريم
كلية العلوم الانسانية والرياضية الجامعة، جامعة كرميان

استلم البحث: 17 اذار 2012 قبل البحث: 21 ايار 2012

## الخلاصة

ان الهدف من عملنا هذا هو ايجاد التنبوء الخطي لمجموع عمليتي بوايسون $Z(t)=X(t)+y(t)$ في زمن المستقبل $\tau \geq 0$ , $Z(t+\tau)$ وذلك عند معرفتنا لقيم $Z(t)$ في الزمن الماضي ومعامل الارتباط لهذا المجموع $\beta_z(t)$ .

**الكلمات المفتاحية:** عمليات بوايسون، التنبؤ الخطي، جمع عمليتي بوايسون.