Forecasting a class of doubly stochastic Poisson processes

P. R. Bouzas¹, A. M. Aguilera², M. J. Valderrama¹

¹ Department of Statistics and Operations Research, Faculty of Pharmacy, University of Granada, Campus of Cartuja, 18071 Granada, Spain
² Department of Statistics and Operations Research, Faculty of Sciences, University of Granada, Campus of Fuentenueva, 18071 Granada, Spain

Received: July 21, 2000; revised version: April 24, 2001

This paper deals with the doubly stochastic Poisson process (DSPP) with mean a truncated Gaussian distribution at any instant of time. The expression of its probability mass function is derived in this paper and it is also proved that the value of the process with maximum probability can be found in a known bounded interval. Furthermore, this paper also presents two methods to forecast the evolution of this kind of DSPP. The first one consists in modelling the mean process and then the probability mass function of the DSPP. The second method uses Multivariate Principal Component Regression to forecast the sample mean in the future instant and then the mass function. Both methods are applied to the real process of number of unpaid bank bills in Spain, forecasting the mass function of this process in 1997 and also its mode.

Keywords: Doubly Stochastic Poisson Process, Truncated Normal Distribution, Multivariate Principal Component Regression Models, Bank Bill

1 Introduction

A random point process is a mathematical model for numerous phenomena characterized by localized events randomly distributed in a continuous space. In practice, it can be of interest to count the number of points in subsets of the space given arise to the counting processes. The Poisson processes are suitable models for many real counting phenomena but they are insufficient in some cases because of the deterministic character of its intensity function.
The doubly stochastic Poisson process (DSPP) is a generalization of the Poisson process when the intensity of the occurrence of the points is influenced by an external process called information process such that the intensity becomes a random process. This process was introduced by Cox (1955). A previous research about this process was deeply developed by Snyder and Miller (1991) and Valderrama et al. (1995), among others.

Studying different data series (e.g. number of mortgaged properties, number of unpaid bank bills) we observed that the mean of the data for various paths collected in a point time could be modeled by a Normal distribution. However, this distribution can not be the parameter of a DSPP as it can take negative values. In order to avoid this problem without changing significantly the appearance of the distribution we have chosen the truncated Normal distribution. In fact, as we show in Section 4, this distribution is a good model adjusted from the data and gives no technical problem. So, the DSPP with mean a truncated Normal seems to be useful and common when we model real phenomena.

In this paper, we study the DSPP, whose parametric function (its mean) has become a truncated Normal variable at every instant of time, and give the expression of the probability mass function, in Section 2. Then, we also prove that the most probable value of the process in a fixed instant is within a known bounded interval.

Two methods for forecasting the DSPP in a future time point are given in Section 3. The first one estimates the mean of the process in that time and then the process. So, its mode can also be estimated using Proposition 1 given in Section 2. The second one estimates a sample of the process in the future by means of Multivariate Principal Component Regression (MPCR) models. Then, the mass function of the DSPP is estimated from the predicted mean, so that the mode can also be estimated. The behavior of MPCR models can be seen in an application developed in Aguilera et al. (1999).

Although, there exist other powerful techniques for relating two sets of variables such as Partial Least Square Regression (see Schmidly, 1995); this paper develops a previous PCA for the variables in order to reduce the dimension mainly if we take into account that the framework of our application is Functional Data Analysis.

Finally, in Section 4, we apply the two methods to the real DSPP of the number of unpaid bank bills in Spain and discuss the results.
2 DSPP with random mean a truncated Normal

A DSPP \( \{N(t) : t \geq t_0\} \) with intensity process \( \{\lambda(t, x(t)) : t \geq t_0\} \) is defined as a conditioned Poisson process with intensity function \( \{\lambda(t, x(t)) : t \geq t_0\} \) given the information process \( \{x(t) : t \geq t_0\} \).

Therefore, the probability that the number of points occurring in \([t_0, t)\) is \(n\) is given by

\[
P[N(t) = n] = E\{P[N(t) = n/x(\sigma) : t_0 \leq \sigma < t]\} =
E \left\{ \frac{1}{n!} \left( \int_{t_0}^{t} \lambda(\sigma, x(\sigma))d\sigma \right)^n \exp \left( -\int_{t_0}^{t} \lambda(\sigma, x(\sigma))d\sigma \right) \right\}
\]

for \( n = 0, 1, 2, \ldots \).

Let \( \Lambda(t, x(t)) \) be the parametric function of the conditionally Poisson process. Then, \( \Lambda(t, x(t)) = \int_{t_0}^{t} \lambda(\sigma, x(\sigma))d\sigma \) and is obviously also influenced by the information process, so it is also a stochastic process. Therefore, we can write

\[
P[N(t) = n] = \frac{1}{n!} E \{\Lambda(t, x(t))^n e^{-\Lambda(t,x(t))}\} = \frac{1}{n!} G_{\Lambda(t,x(t))}^{(n)}(-1) \tag{1}
\]

where \( G_{\Lambda(t,x(t))}(s) \equiv G_{\Lambda}(s) \) is the moment generating function of \( \Lambda(t, x(t)) \).

If \( \Lambda(t, x(t)) \) is a truncated Gaussian variable for each \( t \), (see Johnson and Kotz, 1970) the \( n \)-th derivative of the generating function of \( \Lambda \) is given by

\[
G_{\Lambda}^{(n)}(s) = E \left[ \Lambda^n e^{s\Lambda} \right] =
\frac{1}{\sigma\sqrt{2\pi}} \left( P \left[ \frac{A - \mu}{\sigma} < Z < \frac{B - \mu}{\sigma} \right] \right)^{-1} \int_{A}^{B} \Lambda^n e^{s\Lambda} e^{-\frac{1}{2} \left( \frac{\Lambda - \mu}{\sigma} \right)^2} d\Lambda \tag{2}
\]

where \( A \) and \( B \) are the truncation points, with \( A \geq 0 \) for the parametric function to be nonnegative, and \( A \leq \mu \leq B \) and \( \sigma > 0 \) and \( Z \) is the \( N(0, 1) \). We also notice that \( A, B, \mu \) and \( \sigma \) are functions depending upon \( t \), \( A(t), B(t), \mu(t) \) and \( \sigma(t) \), although we have omitted to write such dependence in order not to complicate the notation in the formula. Thus, we will denote this truncated Normal by \( N_T(A(t), B(t), \mu(t), \sigma(t)) \equiv N_T(A, B, \mu, \sigma) \).
Then, we have that the counting statistic for the DSPP whose mean is a truncated Normal results

\[ P[N(t) = n] = \frac{1}{n! \sigma \sqrt{2\pi}} \left( \Phi \left[ \frac{A - \mu}{\sigma} < Z < \frac{B - \mu}{\sigma} \right] \right)^{-1} \int_{A}^{B} \Lambda^n e^{-\Lambda} e^{-\frac{1}{2} (\frac{\Lambda - \mu}{\sigma})^2} d\Lambda \]  

for \( n = 0, 1, 2, \ldots \) and \( 0 < A < \Lambda < B \) where \( A, B, \mu \) and \( \sigma \) are functions of time and \( Z \) is the \( N(0, 1) \) distribution.

For any fixed \( t = t_1 \), a DSPP with mean the truncated Normal distribution \( N_T(A(t), B(t), \mu(t), \sigma(t)) \), would have constant values for the parameters \( A, B, \mu, \sigma \). Therefore, the probability \( P[N(t_1) = n] \) would be just a function of \( n \) (see equation 3); let us denote it by \( \varphi(n) \).

We will try to find the integer with maximum probability at instant \( t_1 \) for the DSPP with mean a truncated Normal. We can not find the extremes of \( \varphi(n) \) by differential calculus because this function is not continuous. So, the following proposition gives a solution for this problem.

**Proposition 1** The integer in which \( P[N(t_1) = n], n = 0, 1, 2, \ldots \) is maximum belongs to the bounded interval \([A - 1, B - 1]\).

**Proof.** Given a fixed moment of time \( t = t_1 \), equation (3) can be written

\[ \varphi(n) = P[N(t_1) = n] = \frac{c}{n!} \int_{A}^{B} \Lambda^n e^{-\Lambda} e^{-\frac{1}{2} (\frac{\Lambda - \mu}{\sigma})^2} d\Lambda, \]

where \( A < \Lambda < B \) and \( c = \text{constant} \).

Let us consider \( f(n) = \frac{\Lambda^n}{n!} e^{-\Lambda} e^{-\frac{1}{2} (\frac{\Lambda - \mu}{\sigma})^2} \), \( A < \Lambda < B \). Maximizing \( \varphi(n) \) is equivalent to maximize \( f(n) \) as they just differ in a constant and the integral preserves the monotonicity. As \( e^{-\Lambda} e^{-\frac{1}{2} (\frac{\Lambda - \mu}{\sigma})^2} \) does not depend on \( n \), let us study \( \frac{\Lambda^n}{n!} \).

Let us observe that \( \frac{\Lambda^n}{n!} \rightarrow 0 \) as \( n \rightarrow \infty \), so that there exists an integer \( n_0 \) such that for all \( n \geq n_0 \), \( \frac{\Lambda^n}{n!} \) is decreasing in \( n \), for all \( \Lambda \in \mathbb{R} \). Taking a fixed \( \Lambda \in (A, B) \), we have

\[ n \geq n_0(\Lambda) : \frac{\Lambda^n}{n!} \geq \frac{\Lambda^{n+1}}{(n+1)!} \text{ if and only if } n \geq \Lambda - 1, \]
so the integer $n$ where $f(n)$ is maximum for every $\Lambda \in (A, B)$ verifies to be less or equal to $B - 1$. The same way, $\frac{A_n}{n!}$ is increasing for $n \leq \Lambda - 1$ and therefore, for every $\Lambda \in (A, B)$, the integer in which $f(n)$ attains its maximum should be greater or equal to $A - 1$.

Then it is proved that the integer in which $\varphi(n)$ is maximum is in the bounded interval $[A - 1, B - 1]$. ■

The importance of this proposition is that it allows us to look for the most probable value of the Poisson process in instant $t_1$ just evaluating $P[N(t_1) = n]$ in the integers in $[A-1, B-1]$ although it is defined for every $n = 0, 1, 2, \ldots$

3 Forecasting the DSPP with mean a truncated Normal

Once the counting statistics for a DSPP have been stated by means of equation (3), the present section deals with the problem of modelling and forecasting such processes from the knowledgment of $r$ sample-paths. In order to be able to forecast the distribution of the DSPP in that future instant of time $t_1$, we propose two methods for estimating the distribution of $\Lambda$ at $t_1$ known the process in a past interval $[t_0, T]$, $T < t_1$. This will allow us to study characteristics of the process in the future; for example, it is possible to estimate the mode of the process in $t_1$ using proposition 1.

3.1 Forecasting by modelling the mean process

This method for estimating the distribution of the DSPP in $t_1$ is based on the idea of modelling the evolution of $\Lambda(t)$ in time.

In the first step of the method, we will need to estimate a truncated Normal distribution from a sample. So, let us explain how to found the estimates for its parameters.

It is shown in Mittal and Dahiya (1987) that the maximum likelihood estimates (MLE) for the parameters of a truncated Normal distribution, $N_T(A, B, \mu, \sigma)$, become infinite with positive probability. For this reason, they proposed another way of estimation, the Bayes modal estimation, using a prior density for the parameter $\frac{1}{\sigma^2}$. In Bayesian literature, it is commonly used a chi-square density as prior density when there is no information about
the parameter, so then this density would be

\[ f(\sigma^2) = f(\theta) = c(\nu) \theta^{-\frac{\nu+2}{2}} \exp \left( -\frac{1}{2\theta} \right), \theta > 0 \]

where \( c(\nu) = 2^{-\frac{\nu}{2}} \left[ \Gamma\left(\frac{\nu}{2}\right) \right]^{-1} \) is the non-informative prior for \( \mu \). The modified likelihood function is then as follows:

\[
L_m(\mu, \sigma; \Lambda_1, \ldots, \Lambda_r) = \frac{c_2 \sigma^{2-\nu} \exp \left[ -\left( \sum_{i=1}^{r} (\Lambda_i - \mu) + 1 \right) / 2\sigma^2 \right]}{\left[ \int_A \exp \left( - \frac{(y - \mu)^2}{2\sigma^2} \right) dy \right]^r}
\]

where \( c_2 \) is a function of \( \nu \). In the paper of Mittal and Dahiya (1987), the authors proposed \( \nu = 4 \) as optimal value to use.

The Bayes modal estimation provides us the modified maximum likelihood estimators (MMLE) that are proved to exist with probability one and be remarkably better than the MLE which are expanded in the following proposition.

**Proposition 2** Let \( \Lambda \) be a random variable distributed by the truncated Normal \( N_T(A, B, \mu, \sigma) \) with \( \mu \) and \( \sigma \) unknown.

Then, the MMLE of the truncated Normal verify the following system of equations:

\[
\overline{\Lambda} = \frac{\int_A^B y \exp\left[ -\frac{1}{2} \left( \frac{y-\mu}{\sigma} \right)^2 \right] dy}{\int_A^B \exp\left[ -\frac{1}{2} \left( \frac{y-\mu}{\sigma} \right)^2 \right] dy}
\]

\[
S^2 + \frac{1}{r} = \frac{2\sigma^2}{r} + \frac{\int_A^B y^2 \exp\left[ -\frac{1}{2} \left( \frac{y-\mu}{\sigma} \right)^2 \right] dy}{\int_A^B \exp\left[ -\frac{1}{2} \left( \frac{y-\mu}{\sigma} \right)^2 \right] dy}
\]

\[
- \left( \frac{\int_A^B y \exp\left[ -\frac{1}{2} \left( \frac{y-\mu}{\sigma} \right)^2 \right] dy}{\int_A^B \exp\left[ -\frac{1}{2} \left( \frac{y-\mu}{\sigma} \right)^2 \right] dy} \right)^2
\]

where \( \overline{\Lambda} \) and \( S \) are the sample mean and standard deviation, respectively and \( r \) the sample size.

We have implemented the resolution of this system of non-linear equations, using the Newton-Raphson method.

We can now expose the method for estimating the DSPP in a future time point. The steps we should follow are:
1.- Estimate the parameters of the truncated Normal of the mean of the DSPP for every $t \in [t_0, T]$.

Having $r$ sample-paths of the process and the data of their means of occurrences for every $t \in [t_0, T]$; we will have a sample $\Lambda_1, \ldots, \Lambda_r$ for each $t$.

Applying the program for estimating the parameters of a truncated Normal mentioned above to the data of any known instant we will have an estimated truncated Normal for every $t \in [t_0, T]$.

2.- Test that the mean data can be modelled by those truncated Normals.

It has also been implemented a Kolmogorov-Smirnov test for goodness of fit of the truncated Normal distribution. Applying it to the data for $t \in [t_0, T]$, we can know if the estimated distributions in last step fit well.

3.- Once accepted, in step 2, that $\Lambda(t)$ is a particular truncated Normal for any $t \in [t_0, T]$, we have a different $A, B, \mu$ and $\sigma$ for every $t$. Now, we have to study how the parameters develop with $t$ and so, establish the functions $A(t), B(t), \mu(t)$ and $\sigma(t)$. After that, we can extrapolate the parameters of $N_T(A(t_1), B(t_1), \mu(t_1), \sigma(t_1))$, that is to say the suitable distribution of $\Lambda(t_1)$. So, we have the estimation of the mass probability function of the DSPP in $t_1$.

4.- If we are interested in looking for the mode we would have to complete this $4^{th}$ step as follows:

Find $n$ so that $P[N(t_1) = n]$ be maximum.

Once steps 1, 2 and 3 have been completed, and using Proposition 1, the number of occurrences of the DSPP with maximum probability at instant $t_1$ is an integer in $[A(t_1) - 1, B(t_1) - 1]$. We have also implemented the searching of this value for a DSPP with mean a truncated Normal.

The four steps explained above have been implemented in notebooks of Mathematica 3.0, which uses the adaptive gaussian method to calculate numerically the integrals that appear in the density function of $\Lambda(t)$, the counting statistic of the DSPP, the nonlinear system of equations, etc.
3.2 Forecasting by means of MPCR models

Before giving the forecasting method, we will give a brief introduction to MPCR models and its adaptation to counting processes.

The Principal Component Analysis (PCA) is a multivariate analysis technique whose final purpose is to reduce the dimension of the problem by setting up uncorrelated linear transformations of maximum variance of the variables. See Jackson, (1991) for a deeper study.

The Multivariate Principal Component Regression (MPCR) solves the problem of estimating a set of output variables represented by the random vector $\mathbf{Y} = (Y_1, \ldots, Y_s)'$, from another set of predictor variables, denoted by $\mathbf{X} = (X_1, \ldots, X_m)'$ ($m, s \in \mathbb{Z}^+$). Due to the uncorrelation between the principal components, this technique allows us to avoid the problem of multicollinearity of the multivariate multiple linear regression. Furthermore, the reduction of dimension provided by PCA leads very simple models, such that if the principal components had an easy interpretation the regression equations would be more significant and easier of estimating.

The principal characteristic of this technique is that it is necessary to make a double PCA, that is, a classic PCA for each set a variables.

Let $\{\xi_1, \ldots, \xi_m\}$ be the principal components (p.c.’s) associated to the random vector $\mathbf{X}$, given by

$$\xi_i = \sum_{j=1}^{m} \phi_{ji} (X_j - \mu_X) = \phi_i' (\mathbf{X} - \mu_X),$$

where $\phi_i$ is the $i$-th eigenvector of the covariance matrix of $\mathbf{X}$, and $\mu_X$ is its mean vector.

Likewise, let us denote by $\{\eta_1, \ldots, \eta_s\}$, the principal components associated to the random vector $\mathbf{Y}$, and defined by

$$\eta_k = \sum_{j=1}^{s} \delta_{jk} (Y_j - \mu_Y) = \delta_k' (\mathbf{Y} - \mu_Y),$$

where $\delta_k$ is the $k$-th eigenvector of the matrix of covariances of $\mathbf{Y}$ and $\mu_Y$ is the corresponding mean vector.

Then, a MPCR model for predicting the output variable $Y_j$ ($j = 1, \ldots, s$) is given by

$$\tilde{Y}_j = \mu_{Y_j} + \sum_{k=1}^{q} \eta_k^{p_k} \delta_{jk}, \quad q \leq s, \quad \text{(5)}$$
where $\mu_{Y_j}$ is the mean of the variable $Y_j$ and $\hat{\eta}_k^{p_k} = \sum_{i=1}^{p_k} \frac{E[\hat{\eta}_k \xi_i]}{\lambda_i} \xi_i \ (k = 1, \ldots, s)$ is the linear least-squares estimator of the p.c. $\eta_k$ on a set of $p_k$ p.c.'s $\{\xi_i\}_{i=1}^{m}$, with $\lambda_i$ denotes the $i$-th eigenvalue of the covariance matrix of $X$ that represents the variance of the $i$-th p.c. $\xi_i$. This model will be denoted by MPCR($q_1; p_1, \ldots, p_k$).

If in equation (5) $q$ is equal to $s$ and all the p.c.'s $\{\xi_i\}_{i=1}^{m}$ are used for estimating $\eta_k$, then the MPCR gives the same results as the least multivariate linear squares regression but more accurate if the covariance matrix of $X$ has problems of inversion.

The main problem of MPCR is to choose the optimum $p_k$ p.c.'s $\xi_i$ to be introduced in the model as the best predictors for each of the first $q$ p.c.'s $\eta_k$. The usual practice consists of automatically dropping, as predictors, those p.c.'s with the smallest variances (eigenvalues). However, some authors (see Jackson, 1991) have given examples where some of the smallest p.c.'s are highly correlated with the response variable so that there is no guarantee for the p.c.'s with the largest variances to be the best predictors. In order to avoid this problem, we will choose those p.c.'s $\xi_i$ having the highest correlations with each of the p.c.'s $\eta_k$. That means that the p.c.'s $\xi_i$ will be entered into the regression model (5) in the order of magnitude of the square of their correlations with the response variables $\eta_k$, by following a stepwise regression procedure.

Forecasting a counting process in an interval of amplitude $T$ starting from its former evolution with prediction models based on PCA needs a special disposal of the data to be modelled.

Let $N(t)$ be a counting process with known data of an observed path $n(t)$ in $p \times q$ instants of time, $t_j$, denoted by

$$\{n_j : j = 1, \ldots, p, p+1, \ldots, p \times q\}$$

In order to use a MPCR model to forecast $n(t)$ in $p$ future instants of time, denoted by $t_{p \times q+1}, \ldots, t_{p \times (q+1)}$, we propose to cut the path $n(t)$ in $q - 1$ periods of $2p$ instants of time obtaining a new process in the following way:

$$X_{wj} = \begin{cases} n_{(w-1) \times p + j} - n_{(w-1) \times p} & j = 1, \ldots, p \\ n_{(w-1) \times p + j} - n_{w \times p} & j = p+1, \ldots, 2p \end{cases}$$

(6)

where $w = 1, \ldots, q - 1, n_0 = 0$ and $X_j$ is the accumulated number of points of the DSPP in each period of $p$ observations.
Then, we derive a MPCR model in order to estimate the random vector
\( Y = (X_{p+1}, \ldots, X_{2p})' \) by means of \( X = (X_1, \ldots, X_p)' \), which will be used
to estimate \( X_{nj} \) for \( j = p + 1, \ldots, 2p \) and as a consequence \( n(t) \) in future
instants \( t_{p+1}, \ldots, t_{p+q-1} \).

Let us expose the prediction method using MPCR models. The funda-
mental idea of this method is to forecast the \( r \) sample-paths of the process
in a future set of time points using the powerful tool of MPCR models so we
can obtain \( r \) new data that conform an estimated sample of the process in
the future time \( t_1 \) and then, model the DSPP in that instant of time.

The steps we should follow are:

1.- Estimate the suitable MPCR model for each of the \( r \) paths of the
process \( N(t) \), denoted by \( n_1(t), \ldots, n_r(t) \).

First, we have to dispose the data of each path as in equation (6),
obtaining the processes \( X_{w1}, \ldots, X_{wq} \); \( w = 1, \ldots, q - 1 \) in order to
estimate a MPCR model for each random vector \( X'' = (X'_{p+1}, \ldots, X'_{2p})' \)
from each \( X'' = (X'_1, \ldots, X'_p)' \) and forecast each sample path \( n_r(t) \) at
a desired set of future times.

2.- Make estimations of the mean process for each path in the future instant
\( t_1 \) from the predictions given by its corresponding MPCR model. Then,
we will have a sample \( \Lambda_1, \ldots, \Lambda_r \) at time \( t_1 \).

3.- Estimate the parameters of the truncated Normal distribution suitable
for the new sample as shown in equation (4) so that we will have the
probability distribution of \( \Lambda(t_1) \).

4.- Test that the truncated Normal fits well using the implementation of
the Kolmogorov-Smirnov test mentioned in the other method. Once the
distribution estimated for \( \Lambda(t_1) \) is accepted, we have the mass function
of the DSPP in \( t_1 \).

5.- If we are interested in the mode, we have to find the integer value \( n \) so
that \( P[N(t_1) = n] \) be maximum as explained in the other method.

4 Modelling and forecasting unpaid bank bills

In this section, we are going to apply the methodologies exposed in Section
4 to a real case. As a representative tool of the Spanish Financial System,
Table 1: Unpaid bank bills per 10000 inhabitants by provinces

we consider the number of returned or unpaid bank bills.

We will estimate the number of returned or unpaid bank bills per 10000 inhabitants with maximum probability in Spain in 1997. We have monthly data of these bills from 1989 to 1996 in 15 different Spanish provinces randomly chosen.

4.1 Forecasting by modelling the mean number of unpaid bank bills process

In order to use the first method explained in Subsection 3.1, we first have to calculate the mean variable in these provinces for each year as it is shown in Table 1.

The truncated Normal distribution for each year was estimated using the implementation based in Mittal and Dahiya (1987) (step 1) and showed that they fitted well using the implementation of the Kolmogorov-Smirnov test (step 2). In summary we obtained the truncated Normals shown in Table 2.

Now, the expression for the general truncated Normal distribution denoted by $N_T(A(t), B(t), \mu(t), \sigma(t))$ must be established (step3), in order to
<table>
<thead>
<tr>
<th>Year</th>
<th>( N_T ) (Average)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1989</td>
<td>90.6, 827.4, 259.9, 139.3</td>
</tr>
<tr>
<td>1990</td>
<td>181.9, 1763.6, 542.6, 299.6</td>
</tr>
<tr>
<td>1991</td>
<td>352.5, 2742.9, 812.9, 431.2</td>
</tr>
<tr>
<td>1992</td>
<td>473.4, 3747.0, 1089.4, 582.4</td>
</tr>
<tr>
<td>1993</td>
<td>599.4, 4673.7, 1343.2, 715.9</td>
</tr>
<tr>
<td>1994</td>
<td>678.9, 5364.0, 1534.5, 820.8</td>
</tr>
<tr>
<td>1995</td>
<td>753.7, 5872.6, 1691.7, 901.5</td>
</tr>
<tr>
<td>1996</td>
<td>820.0, 6316.5, 1828.1, 971.7</td>
</tr>
</tbody>
</table>

Table 2: Estimated Truncated Normals

calculate the truncated Normal for \( \Lambda(t) \) in a future point time. Fitting a regression function to the points known for each of the functional parameters (shown in Table 2), it turned out that the \( N_T(A(t), B(t), \mu(t), \sigma(t)) \) distribution has the following parameters:

\[
A(t) = 94.049383 \, t^{1.09733}, \quad B(t) = 879.58482 \, t^{0.996695},
\]

\[
\mu(t) = 275.98596 \, t^{0.949735}, \quad \text{and} \quad \sigma(t) = 148.18182 \, t^{0.943997},
\]

where \( t = t - 1988 \).

From the knowledge of the general distribution for \( \Lambda(t) \), we find that the mean of the Poisson process in 1997 is distributed by the variable \( N_T(1048.2, 7858.9, 2224.1, 1179.2) \). So, we have the estimation of the probability mass function using equation 3. The calculation of this function is part of the program implemented to find the mode mentioned above. The probability mass function of the process of unpaid bank bills estimated for 1997 can be seen in Figure 1.

Now, the number of returned or unpaid bank bills with maximum probability in 1997 has to be an integer in \([1047, 7857]\) by Proposition 1 and using the implementation mentioned in step 4, it comes out to be 2223 with probability 0.00040211147, what was our first aim in this study.

### 4.2 Forecasting the number of unpaid bank bills by MPCR models

Now, the second method of prediction explained in Subsection 4.2 will be applied to unpaid bank bills.
Figure 1: Probability mass function of the DSPP with mean the $N_T(1048.2, 7858.9, 2224.1, 1179.2)$ in 1997 in the interval $[1047, 7857]$

First of all, we will estimate a MPCR model for predicting the number of unpaid bank bills at each month of 1997 for each sample province. That is, we have to estimate $r = 15$ MPCR models. In order to do this, we take the 15 paths in the whole years from 1990 to 1996 and denote the observations of each month for each province as:

$$\{n^r_j : j = 1, \ldots, 12, 13, \ldots, 24, \ldots, 12 \times 7 = 84; r = 1, \ldots, 15\}$$

In order to use MPCR models for estimating the sample-paths in every month of 1997 in each province, we cut each of the series of data $n^r_j$ in periods of 24 months in the following way

$$X^r_{wj} = \begin{cases} n^r_{(w-1)12+j} - n^r_{(w-1)12} & j = 1, \ldots, 12 \\ n^r_{(w-1)12+j} - n^r_{w12} & j = 13, \ldots, 24 \end{cases} \quad w = 1, \ldots, 6,$$

where $X^r_{wj}$ represents now the number of accumulated bank bills per year in each month. Cutting the data of each province this way, each MPCR model will estimate the random vector $Y^r = (X^r_{13}, \ldots, X^r_{24})'$ by means of $X^r = (X^r_1, \ldots, X^r_{12})'$ and will supply predictions of the accumulated number of unpaid bank bills each month in 1997 in each province.

We will summarize the construction of the MPCR model for the province of Granada (path number five in the sample) to illustrate the process of identification and estimation of these models.

First, two classic PCA's have been estimated, one of them for the vector $X^5 = (X^5_1, \ldots, X^5_{12})'$ and the other one for $Y^5 = (X^5_{13}, \ldots, X^5_{24})'$. The per-
<table>
<thead>
<tr>
<th></th>
<th>[1,12]</th>
<th></th>
<th>[12,24]</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Exp. Var</td>
<td>Acc. Var</td>
<td>Exp. Var</td>
<td>Acc. Var</td>
</tr>
<tr>
<td>1</td>
<td>99.7889</td>
<td>99.7889</td>
<td>99.8506</td>
<td>99.8506</td>
</tr>
<tr>
<td>2</td>
<td>0.1458</td>
<td>99.9348</td>
<td>0.1003</td>
<td>99.9509</td>
</tr>
<tr>
<td>3</td>
<td>0.0489</td>
<td>99.9836</td>
<td>0.0324</td>
<td>99.9833</td>
</tr>
<tr>
<td>4</td>
<td>0.0088</td>
<td>99.9924</td>
<td>0.0159</td>
<td>99.9992</td>
</tr>
<tr>
<td>5</td>
<td>0.0076</td>
<td>100.0000</td>
<td>0.0008</td>
<td>100.0000</td>
</tr>
<tr>
<td>6</td>
<td>0.0000</td>
<td>100.0000</td>
<td>0.0000</td>
<td>100.0000</td>
</tr>
<tr>
<td>Total Variance</td>
<td>62314.3634</td>
<td></td>
<td>962505.7494</td>
<td></td>
</tr>
</tbody>
</table>

Table 3: Percentages of p.c.'s explained variances in both periods

<table>
<thead>
<tr>
<th></th>
<th>(\xi_1)</th>
<th>(\xi_2)</th>
<th>(\xi_3)</th>
<th>(\xi_4)</th>
<th>(\xi_5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\hat{\eta}_1)</td>
<td>0.8617</td>
<td>-0.4820</td>
<td>0.1226</td>
<td>-0.0911</td>
<td>-0.0424</td>
</tr>
<tr>
<td>(\hat{\eta}_2)</td>
<td>-0.2538</td>
<td>-0.1430</td>
<td>0.6415</td>
<td>-0.6175</td>
<td>-0.3498</td>
</tr>
</tbody>
</table>

Table 4: Correlations between p.c.'s

Percentages of explained and accumulated explained variances appear in Table 3.

Let us observe that only the first p.c. explains more than 99.8% of the total variability. It is known that a first p.c. of this type is typical when working with Poisson processes so that a second p.c. with small variance could also be significant. Because of this and taking into account that the first two p.c.'s accumulate more than 99.9% of explained variance, the MPCR models will have the first two p.c.'s of the future interval \([13,24]\) as response variables.

In order to choose the best predictors \(\xi_i\) for each of the p.c.'s in the future \(\eta_k\), we have also calculated the linear correlations between the first two p.c.'s of the future and all the p.c.'s of the past, as it is shown in Table 4.

Due to the magnitude of the linear correlations between the p.c.'s and the application of the usual stepwise selection methods, it has been identified and estimated the following MPCR model:

\[
\tilde{X}_j^2 = \tilde{X}_j + \hat{\eta}_1^2 \delta_{j1} + \hat{\eta}_2^2 \delta_{j2} \quad j = 13, \ldots, 24
\]

where \(\tilde{X}_j\) is the sample mean of \(X_j\), \(\delta_1\) and \(\delta_2\) are the estimated first and
Table 5: Accumulated number of unpaid bank bills per month in 1997 in Granada

<table>
<thead>
<tr>
<th></th>
<th>January</th>
<th>151.5</th>
<th>July</th>
<th>869.3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>February</td>
<td>266.7</td>
<td>August</td>
<td>968.4</td>
</tr>
<tr>
<td></td>
<td>March</td>
<td>388.1</td>
<td>September</td>
<td>1067.4</td>
</tr>
<tr>
<td></td>
<td>April</td>
<td>497.5</td>
<td>October</td>
<td>1204.8</td>
</tr>
<tr>
<td></td>
<td>May</td>
<td>613.0</td>
<td>November</td>
<td>1330.5</td>
</tr>
<tr>
<td></td>
<td>June</td>
<td>724.8</td>
<td>December</td>
<td>1458.0</td>
</tr>
</tbody>
</table>

Table 6: Forecasted means for 1997

<table>
<thead>
<tr>
<th>Prov.</th>
<th>Mean</th>
<th>Prov.</th>
<th>Mean</th>
<th>Prov.</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barce.</td>
<td>6631.2</td>
<td>Guada.</td>
<td>823.7</td>
<td>Nava.</td>
<td>2691.3</td>
</tr>
<tr>
<td>Cant.</td>
<td>1701.3</td>
<td>Hucl.</td>
<td>857.4</td>
<td>Rioja</td>
<td>3456</td>
</tr>
<tr>
<td>Cast.</td>
<td>3250.5</td>
<td>León</td>
<td>1144.3</td>
<td>Toledo</td>
<td>1588.9</td>
</tr>
<tr>
<td>Coru.</td>
<td>1199.8</td>
<td>Lugo</td>
<td>871.9</td>
<td>Valla.</td>
<td>1397.0</td>
</tr>
<tr>
<td>Grana.</td>
<td>1554.7</td>
<td>Murcia</td>
<td>2434.3</td>
<td>Zara.</td>
<td>3533.62</td>
</tr>
</tbody>
</table>

second eigenvalues of the sample covariance matrix of $Y^5 = (X_{13}^5, \ldots, X_{24}^5)'$. The p.c.'s $\hat{\eta}_1$ and $\hat{\eta}_2$ are estimated by means of the linear models

$$\hat{\eta}_1^2 = 1.0390\hat{\xi}_1 - 15.2068\hat{\xi}_2 \quad \hat{\eta}_2^2 = 1.1077\hat{\xi}_3 - 2.5107\hat{\xi}_4.$$  

Introducing in the MPCR model the value of the p.c.'s in [1,12] for 1996, we obtain the predictions of the accumulated number of unpaid bank bills per month for 1997 that are included in Table 5.

The way to obtain forecasts for the remainder provinces is similar. Then, the estimated mean numbers of unpaid bank bills are included in Table 6, so that they can be taken into account as a sample-path of the mean of the DSPP in 1997.

Now, from system (4) and the estimated sample, the mean of the process in 1997 is distributed by the following estimated truncated Normal,

$$N_T(823.7, 6631.25, 1946.67, 1029.16).$$

Once proved the goodness of fit of this truncated Normal by the Kolmogorov-Smirnov test, we have the estimation of the distribution of the DSPP in 1997. The probability mass function can be seen in Figure 2.
Figure 2: Probability mass function of the DSPP with mean the $N_T(823.7, 6631.25, 1946.67, 1029.16)$ in 1997 in the interval $[822, 6630]$

By Proposition 1, the mode of the DSPP must be in interval $[824, 6630]$. Looking for the integer reaching the maximum value of the probability mass function of our process with the estimated mean, it becomes 1945 with probability $4.49078829729 \times 10^{-4}$.

4.3 Discussion of the results

It can be observed that both modes calculated with the two prediction methods have very small probability; this can be explained because the probability of a DSPP is distributed among the integers and more over, the range of the observed values of unpaid bank bills is very large. Even though, taking a small interval around the mode, comparing it with the large range, the accumulated probability is remarkably greater in both cases. The first method of forecasting by modelling the mean process yielded the mode equal to 2223 with probability $4.02 \times 10^{-4}$ and the interval $[2198, 2248]$ already accumulates a probability of $1.38 \times 10^{-2}$. The second method of forecasting by means of MPCR models provided the mode equal to 1945 with probability $4.49 \times 10^{-4}$ and the interval $[1920, 1970]$ accumulates a probability of $1.66 \times 10^{-2}$.

It can also be considered that the two modes, 2223 and 1945, are quite similar taking into account the differences among observed data. It also seems that a difference of 278 in accumulated number of unpaid bank bills is not large. Considering the interval $[1940, 2225]$ which includes both modes, the probability accumulated in it in the case of the DSPP in 1997 with the
mean distributed by $N_T(1048.269, 7858.952, 2224.1487, 1179.2298)$ is $8.55 \times 10^{-2}$ and with the mean distributed by $N_T(823.7, 6631.25, 1946.67, 1029.16)$ is $8.95 \times 10^{-2}$. This probability can be considered very similar and of certain importance; that indicates that the two methods have given similar results. Even though, let us analyze the two methods in order to chose between them.

The first method bases the estimation of the process in the future extrapolating the four parameters of the mean from the respective regression functions of them in the past. This idea is simple to understand and the calculations are simple after having implemented the corresponding programs. The inconvenient is that as the parameters depend on time, they may change severely from time to time. Therefore, this way of estimating the future parameters should be considered as a short-term estimation.

The second method of estimation has the inconvenient of needing the previous knowledge of the MPCR theory and involves complicated calculations, but the estimation of the sample of the mean in the future is more accurate and can be calculated for a larger period of time. Because of this, this second model can be considered suitable if the nature of data is notably changing in time and always for a medium-term estimation.

Acknowledgments

Authors are grateful with two anonymous referees whose valious comments have helped to improve the final version of this manuscript.

This research was partially supported by project BFM2000-1466 of Dirección General de Enseñanza Superior, Ministerio de Educación y Cultura.

References


