Tatra Mt. Math. Publ. **39** (2008), 165–174

TATRA MOUNTAINS
Mathematical Publications

ASYMPTOTIC DISTRIBUTIONS FOR ESTIMATORS AND STATISTICS IN MIXED POISSON PROCESSES

Vippal Savani — Anatoly A. Zhigljavsky

ABSTRACT. Mixed Poisson processes provide natural and popular models in many applied areas including market research, insurance and disease modeling. This paper derives the joint asymptotic distributions of statistics and parameter estimates which are computed in different time intervals from data generated by mixed Poisson processes. These distributions can be used, for example, to test the adequacy of the mixed Poisson process against data.

1. Introduction

The class of mixed Poisson processes has been used as a natural model for events occurring in continuous or discrete time in many fields. The mixed Poisson process has been successfully applied in the modeling of accidents and sickness, consumer buying, non-life related insurance and risk theory, to name a few examples (see, e.g., [2]–[4] and references therein).

In this paper we consider the joint asymptotic distributions between statistics and estimators which are computed in different time intervals when the underlying process is mixed Poisson.

The fitting of mixed Poisson processes to observed data has mainly focussed on fitting the one-dimensional mixed Poisson distribution when considering data observed over fixed time intervals. Fitting the one-dimensional distribution does not fully justify the adequacy of the process; in particular, the dynamical behavior of the mixed Poisson process is not validated. The derivation of the joint asymptotic distributions of statistics and estimators allows testing the hypothesis as to whether parameter estimates computed in the two different time intervals could have been generated from the same process. This will allow us to verify the dynamical properties of the underlying model against data.

2000 Mathematics Subject Classification: Primary 62F12, 62M10, 91B70.
Keywords: mixed Poisson process, negative binomial distribution, maximum likelihood, moment estimators.

Note that it is easy to construct methods of testing the adequacy of mixed Poisson processes which are based on testing whether each individual realization follows the standard Poisson process; the distribution of the intensities of the individual Poisson processes can then be checked against a specified structure distribution. However, in practice the individual behavior basically never follows the pure Poisson model (see, e.g., [1], [2]) and therefore the related tests would almost certainly reject the Poisson process assumption. At the same time, it is widely known that the Poisson and mixed Poisson models often work fairly well when the data is aggregated over either time or realizations, or both.

The asymptotic distributions derived in this paper allow us to test the mixed Poisson model hypotheses using the aggregated data. We are not aware of any other procedure of testing the dynamics of the mixed Poisson models, except those based on testing individual realizations (which is not practical). Furthermore, when only a few events are registered in each individual realization, testing the pure Poisson hypothesis is meaningless as there is not enough data. However, the methodology described in this paper can be perfectly suitable for testing the mixed Poisson model if there are enough realizations in the multiple realization scheme and a suitable aggregation is made.

The structure of the paper is as follows. We introduce mixed Poisson processes in this section. In Section 2 we formulate a general scheme of parameter estimation and derive an expression for the asymptotic covariance matrix of the estimators. In Sections 3 and 4 we derive the main results of the paper, the asymptotic distributions between different statistics and estimators computed in different time intervals.

Mixed Poisson processes

Let $\boldsymbol{Z} = \big\{Z(t_1), Z(t_2), \dots, Z(t_n)\big\}$ be a random vector, let $\boldsymbol{x} = \{x_1, x_2, \dots, x_n\}$ be a set of non-negative integers with $0 = x_0 \leqslant x_1 \leqslant \dots \leqslant x_n$ and let $0 = t_0 \leqslant t_1 \leqslant \dots \leqslant t_n$ represent an increasing sequence of time points; the multivariate Poisson distribution is defined as

$$\mathbb{P}\left(\boldsymbol{Z} = \boldsymbol{x} | \Lambda = \lambda\right) = \prod_{i=0}^{n-1} \frac{\big[\lambda(t_{i+1} - t_i)\big]^{x_{i+1}-x_i}}{(x_{i+1} - x_i)!} \exp\big(-\lambda(t_{i+1} - t_i)\big), \tag{1}$$

where $\lambda > 0$ is the intensity. The mixed Poisson process is then defined as a process whose finite-dimensional distributions are

$$\mathbb{P}\left(\boldsymbol{Z} = \boldsymbol{x}\right) = \int_{0-}^{\infty} \mathbb{P}\left(\boldsymbol{Z} = \boldsymbol{x} | \Lambda = \lambda\right) \mathrm{d}U_\Lambda(\lambda; \boldsymbol{\theta}). \tag{2}$$

Here $U_\Lambda(\lambda;\boldsymbol{\theta})$ is the distribution function for the random variable Λ and $\boldsymbol{\theta}$ is a vector of unknown parameters. The function $U_\Lambda(\lambda;\boldsymbol{\theta})$ is commonly known as the structure distribution of the mixed Poisson process. A common distribution for Λ is the Gamma distribution, in which case $\boldsymbol{Z}$ has a multivariate negative binomial distribution. Other distributions include the beta, shifted-Gamma, generalized inverse Gaussian and lognormal distributions (see, e.g., [3, p. 27]).

2. Asymptotic properties of a general estimator

This section considers the asymptotic distribution of a general class of estimators. The main result can be considered as a reformulation of known results on M- and Z-estimators, see [6, Chapters 3–5]. We need this reformulation to unify notation and simplify exposition in the next sections.

2.1. General estimation scheme

Let ζ be a random variable taking values in some set $\mathcal{Z}$ and let ζ have probability mass function $p(z;\boldsymbol{\theta}),\ z \in \mathcal{Z}$, where $\boldsymbol{\theta} = (\theta_1,\dots,\theta_d)^T$ (d ≥ 1) is a vector of parameters taking values in some set $\Theta \subseteq \mathbb{R}^d$ with non-empty interior $\mathrm{int}(\Theta)$. For the purpose of this paper we only need to consider $\mathcal{Z} = \{0,1,\dots\}$, but the results of this section can be extended to arbitrary sets $\mathcal{Z}$; in the case of continuous distributions, $p(z;\boldsymbol{\theta})$ is a density. We now define a general method of estimating $\boldsymbol{\theta}^* = (\theta_1^*,\dots,\theta_d^*)^T \in \mathrm{int}(\Theta)$, the true parameter values of the sampling distribution, by using an i.i.d. sample $\{z_1,\dots,z_N\}$ of values of ζ.

Let $\boldsymbol{f} = (f_1,\dots,f_d)^T \in \mathbb{R}^d$ where $f_i\colon \mathcal{Z}\times\Theta \to \mathbb{R}$ $(i = 1,\dots,d)$ are some functions which are smooth enough and possibly depend on $\boldsymbol{\theta}$; set $\bar{\boldsymbol{f}} = (\bar{f}_1,\dots,\bar{f}_d)^T \in \mathbb{R}^d$ with $\overline{f_i} = \frac{1}{N}\sum\limits_{l=1}^{N} f_i(z_l;\boldsymbol{\theta})$. Since $\{z_1,\dots,z_N\}$ form an i.i.d. sample of values of ζ, we have $\mathbb{E}\overline{f_i} = \mathbb{E}f_i(\zeta;\boldsymbol{\theta})$. The estimator $\hat{\boldsymbol{\theta}} = (\hat{\theta}_1,\dots,\hat{\theta}_d)^T$ is then defined to be the solution to the equations

$$G_i(\boldsymbol{\theta},\overline{f_i}) = 0, \qquad i = 1,\dots,d\,, \tag{3}$$

where $G_i(\boldsymbol{\theta},\overline{f_i}) = \mathbb{E}f_i(\zeta;\boldsymbol{\theta}) - \overline{f_i}$.

Let us give examples of possible functions f_i $(i = 1,\dots,d)$:

Example 2.1.1. $f_i(z;\boldsymbol{\theta}) = \partial \log(p(z;\boldsymbol{\theta}))/\partial\theta_i$ implying $\mathbb{E}f_i(\zeta;\boldsymbol{\theta}) = 0$;

Example 2.1.2. $f_i(z;\boldsymbol{\theta}) = f_i(z)$ so that the functions f_i do not depend on $\boldsymbol{\theta}$.

The system of equations (3) may be represented in vector form as

$$\boldsymbol{G}(\boldsymbol{\theta},\bar{\boldsymbol{f}}) = \big(G_1(\boldsymbol{\theta},\overline{f_1}),\dots,G_d(\boldsymbol{\theta},\overline{f_d})\big)^T = \mathbf{0}\,. \tag{4}$$

For each i, we can represent $G_i(\boldsymbol{\theta}, \overline{f_i})$ as $G_i(\boldsymbol{\theta}, \overline{f_i}) = \frac{1}{N} \sum\limits_{l=1}^{N} g_i(z_l, \boldsymbol{\theta}) = \bar{g}_i$, where $g_i(z, \boldsymbol{\theta}) = \mathbb{E} f_i(\zeta; \boldsymbol{\theta}) - f_i(z; \boldsymbol{\theta})$.

2.2. Asymptotic normality of estimators

Theorem 2.1. *Assume that the function $\boldsymbol{G}$ is invertible as a function of $\boldsymbol{\theta}$ in some neighbourhood of $(\boldsymbol{\theta}^*, \mathbb{E}\boldsymbol{f})$ and let $\hat{\boldsymbol{\theta}}$ be the solution of $\boldsymbol{G}(\boldsymbol{\theta}, \bar{\boldsymbol{f}}) = \boldsymbol{0}$. Assume that $\mathbb{E}\,|\partial g_i(\zeta, \boldsymbol{\theta})/\partial \theta_j| < \infty$ for all i, j. Additionally, assume that the estimator $\hat{\boldsymbol{\theta}}$ is a consistent estimator of $\boldsymbol{\theta}$ and $\sqrt{N}(\bar{\boldsymbol{f}} - \mathbb{E}\boldsymbol{f})$ is asymptotically normally distributed $\mathcal{N}(0, \mathbb{D}\boldsymbol{f})$, where*

$$\mathbb{D}\boldsymbol{f} = \mathbb{E}(\boldsymbol{f} - \mathbb{E}\boldsymbol{f})(\boldsymbol{f} - \mathbb{E}\boldsymbol{f})^T = \left\|\mathrm{Cov}\big(f_i(\zeta; \boldsymbol{\theta}), f_j(\zeta; \boldsymbol{\theta})\big)\right\|_{i,j=1}^{d}.$$

Then asymptotically as $N \to \infty$,

$$\sqrt{N}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^*) \overset{\mathcal{D}}{\Rightarrow} \mathcal{N}\big(0, \boldsymbol{V}\,(\mathbb{D}\boldsymbol{f})\,\boldsymbol{V}^T\big) \tag{5}$$

where

$$\boldsymbol{V} = \left[\lim_{N\to\infty} \frac{\partial \boldsymbol{G}(\boldsymbol{\theta}, \bar{\boldsymbol{f}})}{\partial \boldsymbol{\theta}}\bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}^*}\right]^{-1}. \tag{6}$$

P r o o f. According to the weak law of large numbers as $N \to \infty$, $\bar{\boldsymbol{f}} \to \mathbb{E}\boldsymbol{f}$ in probability and for any $\boldsymbol{\theta}$ there exists the weak limit

$$\lim_{N\to\infty} \left\|\frac{\partial \boldsymbol{G}(\boldsymbol{\theta}, \bar{\boldsymbol{f}})}{\partial \boldsymbol{\theta}}\right\| = \lim_{N\to\infty} \left\|\frac{1}{N} \sum_{l=1}^{N} \frac{\partial g_i(z_l, \boldsymbol{\theta})}{\partial \theta_j}\right\|$$

which is a non-random matrix.

Since $\boldsymbol{G}$ is invertible as a function of $\boldsymbol{\theta}$ in the neighbourhood of $(\boldsymbol{\theta}^*, \mathbb{E}\boldsymbol{f})$, for N large enough the inverse $\left(\frac{\partial \boldsymbol{G}(\boldsymbol{\theta}, \bar{\boldsymbol{f}})}{\partial \boldsymbol{\theta}}\right)^{-1}$ exists in the neighbourhood of $\boldsymbol{\theta}^*$. We approximate equation (4) using the first order Taylor expansion:

$$\boldsymbol{G}(\boldsymbol{\theta}, \bar{\boldsymbol{f}}) \simeq \boldsymbol{G}(\boldsymbol{\theta}^*, \bar{\boldsymbol{f}}) + \frac{\partial \boldsymbol{G}(\boldsymbol{\theta}, \bar{\boldsymbol{f}})}{\partial \boldsymbol{\theta}}\bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}^*} (\boldsymbol{\theta} - \boldsymbol{\theta}^*) = 0. \tag{7}$$

According to the well-known δ-method (see, e.g., [5]) the asymptotic distribution of $\hat{\boldsymbol{\theta}}$ is the same as the asymptotic distribution of $\tilde{\boldsymbol{\theta}}$, which is the solution to equation (7). Solving equation (7) we obtain

$$\boldsymbol{\theta}^* - \tilde{\boldsymbol{\theta}} = \left(\frac{\partial \boldsymbol{G}(\boldsymbol{\theta}, \bar{\boldsymbol{f}})}{\partial \boldsymbol{\theta}}\right)^{-1}\Bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}^*} \boldsymbol{G}(\boldsymbol{\theta}^*, \bar{\boldsymbol{f}})\,.$$

The asymptotic ($N \to \infty$) distribution of $\sqrt{N}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^*)$ can be related to the asymptotic distribution of $\sqrt{N}\boldsymbol{G}(\boldsymbol{\theta}^*, \bar{\boldsymbol{f}})$ using Slutsky's theorem, which allows the replacement of

$$\left(\frac{\partial \boldsymbol{G}(\boldsymbol{\theta}, \bar{\boldsymbol{f}})}{\partial \boldsymbol{\theta}}\right)^{-1}\Bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}^*} \quad \text{with} \quad \boldsymbol{V} = \left[\lim_{N\to\infty} \frac{\partial \boldsymbol{G}(\boldsymbol{\theta}, \bar{\boldsymbol{f}})}{\partial \boldsymbol{\theta}}\Bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}^*}\right]^{-1}$$

and we obtain that the asymptotic distributions of $\sqrt{N}(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}^*)$ and $\sqrt{N}\boldsymbol{V}\boldsymbol{G}(\boldsymbol{\theta}^*, \bar{\boldsymbol{f}})$ coincide. Note that $\sqrt{N}(\bar{\boldsymbol{f}} - \mathbb{E}\boldsymbol{f}) = \sqrt{N}\boldsymbol{G}(\boldsymbol{\theta}, \bar{\boldsymbol{f}})$ and therefore $\sqrt{N}\boldsymbol{G}(\boldsymbol{\theta}, \bar{\boldsymbol{f}})$ is asymptotically normally distributed $\mathcal{N}(0, \mathbb{D}\boldsymbol{f})$. This implies that $\sqrt{N}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^*)$ is asymptotically normally distributed $\mathcal{N}\big(0, \boldsymbol{V}(\mathbb{D}\boldsymbol{f})\boldsymbol{V}^T\big)$. □

Example 2.2.1. **Maximum likelihood.** We have $f_i(z;\boldsymbol{\theta}) = \partial \log\big(p(z;\boldsymbol{\theta})\big)/\partial\theta_i$ $(i = 1, \dots, d)$ so that

$$\mathbb{E}\overline{f_i} = \mathbb{E}f_i(\zeta;\boldsymbol{\theta}) = 0$$

and

$$\mathbb{D}\boldsymbol{f} = \left\|\mathbb{E}\frac{\partial}{\partial\theta_i}\log p(\zeta;\boldsymbol{\theta})\frac{\partial}{\partial\theta_j}\log p(\zeta;\boldsymbol{\theta})\right\| = I(\boldsymbol{\theta}),$$

$$V^{-1} = -\lim_{N\to\infty}\left\|\frac{\partial}{\partial\theta_j}\frac{1}{N}\sum_{l=1}^{N}\frac{\partial}{\partial\theta_i}\log p(z_l;\boldsymbol{\theta})\right\| = I(\boldsymbol{\theta}),$$

where $I(\boldsymbol{\theta})$ is the Fisher information matrix. The covariance matrix of the maximum likelihood estimators is therefore $\mathbb{D}\hat{\theta} = I(\boldsymbol{\theta})^{-1}I(\boldsymbol{\theta})I(\boldsymbol{\theta})^{-1} = I(\boldsymbol{\theta})^{-1}$.

Example 2.2.2. **General method of moments.** We have $f_i(z, \boldsymbol{\theta}) = f_i(z)$ $(i = 1, \dots, d)$ so that the functions f_i do not depend on the unknown parameters $\boldsymbol{\theta}$. This implies

$$\mathbb{D}\boldsymbol{f} = \big\|\mathrm{Cov}\big(f_i(\zeta), f_j(\zeta)\big)\big\| \quad \text{and} \quad V^{-1} = \|\partial\mathbb{E}\boldsymbol{f}(\zeta)/\partial\boldsymbol{\theta}\|.$$

3. Covariances between statistics

In this section we derive the asymptotic distributions between different statistics $\bar{\boldsymbol{f}}$, as defined in Section 2, computed in two different time intervals. All the results can be easily generalized to any number of intervals.

3.1. Non-overlapping intervals

Note that since the Poisson process is a stationary and homogenous process, considering covariances of two statistics computed over the intervals $[t_1, t_2)$ and $[t_3, t_4)$ with $0 \le t_1 < t_2 \le t_3 < t_4$ is equivalent to considering covariances of the same statistics over the time intervals $[0, t)$ and $[t, t + s)$, so that $t_1 = 0,\ t_2 = t_3 = t$ and $t_4 = t + s$.

Let us consider the covariance between the statistics

$$\bar{\phi}_{0,t} = \frac{1}{N}\sum_{l=1}^{N}\phi\big(z_l(0,t)\big) \quad \text{and} \quad \bar{\psi}_{t,t+s} = \frac{1}{N}\sum_{l=1}^{N}\psi\big(z_l(t,t+s)\big),$$

where $\{z_1(0,t),\ldots,z_N(0,t)\}$ and $\{z_1(t,t+s),\ldots,z_N(t,t+s)\}$ are i.i.d. data from a mixed Poisson process observed over two adjacent time intervals $[0,t)$ and $[t,t+s)$ respectively $(t,s>0)$. Here ϕ and ψ are some functions possibly dependent upon the vector of parameters $\boldsymbol{\theta}$.

We note that for fixed u and v the observations $z_l(u,u+v)$ $(l=1,\ldots,N)$ are mutually independent. For fixed l, the observations $z_l(0,t)$ and $z_l(t,t+s)$ are conditionally independent and Poisson distributed with means $\lambda_l t$ and $\lambda_l s$, respectively. Here, λ_l is random for $l=1,\ldots,N$, but is the same for fixed l as time varies. The samples $\{z_1(0,t),\ldots,z_N(0,t)\}$ and $\{z_1(t,t+s),\ldots,z_N(t,t+s)\}$ are dependent since, for each l, $z_l(0,t)$ and $z_l(t,t+s)$ are Poisson distributed with common λ_l. Let $\zeta_{u,v}$ be a random variable whose distribution is identical to the distribution of the i.i.d. random variables $z_l(u,v)$ $(l=1,\ldots,N)$, the number of events occurring in the time interval $[u,v)$. Then

$$N\mathrm{Cov}\left[\bar{\phi}_{0,t},\bar{\psi}_{t,t+s}\right] = \mathrm{Cov}\big[\phi(\zeta_{0,t}),\psi(\zeta_{t,t+s})\big]. \tag{8}$$

Let $\mathcal{L}(c) = \mathbb{E}e^{-c\Lambda}$ be the Laplace transform of the random variable Λ. Then $\mathcal{L}'(c) = \frac{\partial}{\partial c}\mathbb{E}e^{-c\Lambda} = -\mathbb{E}[\Lambda e^{-c\Lambda}]$. Additionally, let $p_{[u,v)}(z;\boldsymbol{\theta})$ denote the mixed Poisson distribution over the time interval $[u,v)$. We have the following cases:

Case 1: $\phi(z) = z^{\alpha}, \psi(z) = z^{\beta}$:

$$\mathrm{Cov}\big[\phi(\zeta_{0,t}),\psi(\zeta_{t,t+s})\big] = \mathbb{E}\mu_{\alpha}(\lambda t)\mu_{\beta}(\lambda s) - \mathbb{E}\mu_{\alpha}(\lambda t)\mathbb{E}\mu_{\beta}(\lambda s),$$

where $\mu_{\alpha}(\nu) = \mathbb{E}\kappa_{\nu}^{\alpha}$ and κ_{ν} is a Poisson random variable with intensity ν.

Case 1a: $\phi(z) = z, \psi(z) = z$:

$$\mathrm{Cov}\big[\phi(\zeta_{0,t}),\psi(\zeta_{t,t+s})\big] = \mathbb{E}\Lambda^2 ts - \mathbb{E}\Lambda t\mathbb{E}\Lambda s = ts\mathrm{Var}\,\Lambda.$$

Case 1b: $\phi(z) = z, \psi(z) = z^2$:

$$\mathrm{Cov}\big[\phi(\zeta_{0,t}),\psi(\zeta_{t,t+s})\big] = ts^2\mathrm{Cov}[\Lambda,\Lambda^2] + ts\mathrm{Var}\Lambda.$$

Case 2: $\phi(z) = \frac{\partial}{\partial\theta_i}\log p_{[0,t)}(z;\boldsymbol{\theta}), \psi(z) = \frac{\partial}{\partial\theta_j}\log p_{[t,t+s)}(z;\boldsymbol{\theta})$:

$$\begin{aligned}
&\mathrm{Cov}\big[\phi(\zeta_{0,t}),\psi(\zeta_{t,t+s})\big]\\
&\quad= \mathbb{E}\frac{\partial}{\partial\theta_i}\log p_{[0,t)}(\zeta_{0,t};\boldsymbol{\theta})\frac{\partial}{\partial\theta_j}\log p_{[t,t+s)}(\zeta_{t,t+s};\boldsymbol{\theta})\\
&\quad= \mathbb{E}\frac{1}{p_{[0,t)}(\zeta_{0,t};\boldsymbol{\theta})p_{[t,t+s)}(\zeta_{t,t+s};\boldsymbol{\theta})}\frac{\partial}{\partial\theta_i}p_{[0,t)}(\zeta_{0,t};\boldsymbol{\theta})\frac{\partial}{\partial\theta_j}p_{[t,t+s)}(\zeta_{t,t+s};\boldsymbol{\theta}).
\end{aligned}$$

To compute the derivative $\frac{\partial}{\partial\theta_i} p_{[u,v)}(z;\boldsymbol{\theta})$ we can use the formula

$$\frac{\partial}{\partial\theta_i} p_{[u,v)}(z;\boldsymbol{\theta}) = \frac{\partial}{\partial\theta_i} \int_{0-}^{\infty} \frac{\big(\lambda(v-u)\big)^z \exp\big(-\lambda(v-u)\big)}{z!} \,\mathrm{d}U_\Lambda(\lambda;\boldsymbol{\theta}).$$

3.2. Overlapping intervals

In this section we consider covariances between statistics in the most general case when the intervals are possibly overlapping. This includes the cases when the intervals do not overlap and also when the intervals coincide. To derive the results of this section we use the results of Section 3.1.

Let us consider the covariance between the statistics

$$\bar{\phi}_{t_1,t_3} = \frac{1}{N}\sum_{l=1}^{N} \phi\big(z_l(t_1,t_3)\big) \quad \text{and} \quad \bar{\psi}_{t_2,t_4} = \frac{1}{N}\sum_{l=1}^{N} \psi\big(z_l(t_2,t_4)\big),$$

where $\big\{z_1(t_1,t_3),\dots,z_N(t_1,t_3)\big\}$ and $\big\{z_1(t_2,t_4),\dots,z_N(t_2,t_4)\big\}$ are data from a mixed Poisson process observed over two possibly overlapping intervals $[t_1,t_3)$ and $[t_2,t_4)$ with $0 \le t_1 \le t_2 \le t_3 \le t_4$. Similarly to (8) we have

$$N\mathrm{Cov}\left[\bar{\phi}_{t_1,t_3}, \bar{\psi}_{t_2,t_4}\right] = \mathrm{Cov}\big[\phi(\zeta_{t_1,t_3}), \psi(\zeta_{t_2,t_4})\big].$$

Computing these covariances for different functions, ϕ and ψ can be simplified using the fact that the Poisson process has stationary and independent increments. Some covariances are given below.

Case 1a: $\phi(z) = z, \psi(z) = z$:

$$\begin{aligned}\mathrm{Cov}\big[\phi(\zeta_{t_1,t_3}), \psi(\zeta_{t_2,t_4})\big] &= \mathbb{E}\zeta_{t_1,t_3}\zeta_{t_2,t_4} - \mathbb{E}\zeta_{t_1,t_3}\mathbb{E}\zeta_{t_2,t_4}\\ &= \mathrm{Cov}(\zeta_{t_1,t_2},\zeta_{t_2,t_3}) + \mathrm{Cov}(\zeta_{t_1,t_2},\zeta_{t_3,t_4})\\ &\quad + \mathrm{Cov}(\zeta_{t_2,t_3},\zeta_{t_3,t_4}) + \mathrm{Var}(\zeta_{t_2,t_3})\end{aligned}$$

and using the results of case 1a Section 3.1 we obtain

$$\mathrm{Cov}\big[\phi(\zeta_{t_1,t_3}), \psi(\zeta_{t_2,t_4})\big] = (t_4-t_2)(t_3-t_1)\mathrm{Var}\Lambda + (t_3-t_2)\mathbb{E}\Lambda.$$

Case 1b: $\phi(z) = z, \psi(z) = z^2$:

$$\mathrm{Cov}\big[\phi(\zeta_{t_1,t_3}), \psi(\zeta_{t_2,t_4})\big] = \mathbb{E}\zeta_{t_1,t_3}\zeta^2_{t_2,t_4} - \mathbb{E}\zeta_{t_1,t_3}\mathbb{E}\zeta^2_{t_2,t_4}$$

and using the results of case 1b Section 3.1 we obtain

$$\begin{aligned}\mathrm{Cov}\big[\phi(\zeta_{t_1,t_3}), \psi(\zeta_{t_2,t_4})\big] &= (t_4-t_2)^2(t_3-t_1)\mathrm{Cov}(\Lambda,\Lambda^2)\\ &\quad + (t_4-t_2)(t_3-t_1)\mathrm{Var}\Lambda\\ &\quad + 2(t_4-t_2)(t_3-t_2)\mathbb{E}\Lambda^2\\ &\quad + (t_3-t_2)\mathbb{E}\Lambda\,.\end{aligned}$$

4. Covariances between parameter estimators

Let $\hat{\boldsymbol{\theta}}^{(1)}$ and $\hat{\boldsymbol{\theta}}^{(2)}$ be estimators of $\boldsymbol{\theta}$ in the intervals $[t_1, t_3)$ and $[t_2, t_4)$ constructed using the general scheme of Section 2.1 with the sets of functions $\{f_i^{(1)}(z;\boldsymbol{\theta})\}_{i=1}^d$ and $\{f_i^{(2)}(z;\boldsymbol{\theta})\}_{i=1}^d$, respectively. Assume that Theorem 2.1 applies to $\hat{\boldsymbol{\theta}}^{(1)}$ and $\hat{\boldsymbol{\theta}}^{(2)}$ so that both estimators are asymptotically normal and let $\boldsymbol{V}^{(1)}$, $\boldsymbol{V}^{(2)}$, $\mathbb{D}\boldsymbol{f}^{(1)}$ and $\mathbb{D}\boldsymbol{f}^{(2)}$ be the matrices associated with $\hat{\boldsymbol{\theta}}^{(1)}$ and $\hat{\boldsymbol{\theta}}^{(2)}$. We have $\sqrt{N}(\bar{\boldsymbol{f}} - \mathbb{E}\boldsymbol{f})$ is asymptotically normal $\mathcal{N}(0, \mathbb{D}\boldsymbol{f})$, where

$$\boldsymbol{f}(z;\boldsymbol{\theta}) = \begin{pmatrix} \boldsymbol{f}^{(1)}(z;\boldsymbol{\theta}) \\ \boldsymbol{f}^{(2)}(z;\boldsymbol{\theta}) \end{pmatrix}, \quad \bar{\boldsymbol{f}} = \begin{pmatrix} \bar{\boldsymbol{f}}^{(1)} \\ \bar{\boldsymbol{f}}^{(2)} \end{pmatrix}, \quad \mathbb{E}\boldsymbol{f} = \begin{pmatrix} \mathbb{E}\boldsymbol{f}^{(1)}(\zeta_{t_1,t_3};\boldsymbol{\theta}) \\ \mathbb{E}\boldsymbol{f}^{(2)}(\zeta_{t_2,t_4};\boldsymbol{\theta}) \end{pmatrix}$$

and

$$\mathbb{D}\boldsymbol{f} = \begin{pmatrix} \mathbb{D}\boldsymbol{f}^{(1)} & \mathbb{C}(\boldsymbol{f}^{(1)}, \boldsymbol{f}^{(2)}) \\ \mathbb{C}(\boldsymbol{f}^{(1)}, \boldsymbol{f}^{(2)})^T & \mathbb{D}\boldsymbol{f}^{(2)} \end{pmatrix} \tag{9}$$

with

$$\mathbb{C}\big(\boldsymbol{f}^{(1)}, \boldsymbol{f}^{(2)}\big) = \big\|\mathrm{Cov}\big(f_i^{(1)}(\zeta_{t_1,t_3};\boldsymbol{\theta}), f_j^{(2)}(\zeta_{t_2,t_4};\boldsymbol{\theta})\big)\big\|_{i,j=1}^d .$$

The components of the matrix $\mathbb{C}\big(\boldsymbol{f}^{(1)}, \boldsymbol{f}^{(2)}\big)$ are computed using the results of Section 3.

Consider the problem of estimating the vector $\boldsymbol{\theta}_* = \big(\boldsymbol{\theta}^{(1)}, \boldsymbol{\theta}^{(2)}\big)^T$ with $\hat{\boldsymbol{\theta}}_* = \big(\hat{\boldsymbol{\theta}}^{(1)}, \hat{\boldsymbol{\theta}}^{(2)}\big)^T$ where $\boldsymbol{\theta}^{(1)}$ and $\boldsymbol{\theta}^{(2)}$ are two different copies of $\boldsymbol{\theta}$. The fact that $\boldsymbol{\theta}^{(1)}$ and $\boldsymbol{\theta}^{(2)}$ are two different copies of $\boldsymbol{\theta}$ implies that the matrix of partial derivatives $\boldsymbol{V}$, defined by equation (6) with $\boldsymbol{\theta}_*$ substituted for $\boldsymbol{\theta}^*$, has a block diagonal structure

$$\boldsymbol{V} = \begin{pmatrix} \boldsymbol{V}^{(1)} & 0 \\ 0 & \boldsymbol{V}^{(2)} \end{pmatrix}. \tag{10}$$

Using Theorem 2.1 we obtain that $\sqrt{N}(\hat{\boldsymbol{\theta}}_* - \boldsymbol{\theta}_*)$ is asymptotically normal $\mathcal{N}\big(0, \boldsymbol{V}(\mathbb{D}\boldsymbol{f})\boldsymbol{V}^T\big)$, where $\mathbb{D}\boldsymbol{f}$ and $\boldsymbol{V}$ are defined by (9) and (10). The asymptotic covariance matrix is therefore

$$\boldsymbol{V}(\mathbb{D}\boldsymbol{f})\boldsymbol{V}^T = \begin{pmatrix} \boldsymbol{V}^{(1)}\, \mathbb{D}\boldsymbol{f}^{(1)}\, \big(\boldsymbol{V}^{(1)}\big)^T & \boldsymbol{V}^{(1)}\, \mathbb{C}\big(\boldsymbol{f}^{(1)},\boldsymbol{f}^{(2)}\big)\, \big(\boldsymbol{V}^{(2)}\big)^T \\ \boldsymbol{V}^{(2)} \Big(\mathbb{C}\big(\boldsymbol{f}^{(1)},\boldsymbol{f}^{(2)}\big)\Big)^T \big(\boldsymbol{V}^{(1)}\big)^T & \boldsymbol{V}^{(2)}\, \mathbb{D}\boldsymbol{f}^{(2)}\, \big(\boldsymbol{V}^{(2)}\big)^T \end{pmatrix}. \tag{11}$$

The Figure below shows estimates of normalized means when fitting the mixed Poisson process with a Gamma structure distribution to product purchases in two different categories. The data, containing purchases of 35 000 households over a year, was kindly provided by ACNielsen BASES. Points of normalized sample means computed in two consecutive time intervals of equal length are plotted against each other. A 95% theoretical confidence ellipse constructed using (11), replacing parameters with parameter estimates obtained by standard method of moments, is also shown. In this case, the mean values computed in different intervals and the covariances between these means are exactly what one would expect from the mixed Poisson model. In some other categories we observe seasonality in the data which causes the theoretical confidence ellipses to shift to one side of the data, although the shape is still retained.

In conclusion, the main result of the paper is the derivation of the joint asymptotic distributions of statistics and parameter estimates in different time intervals in the general mixed Poisson model. These results allow construction of tests for confirming the adequacy of the mixed Poisson model without using the assumptions of individual Poisson processes which are difficult to use in practice due to the lack of data.

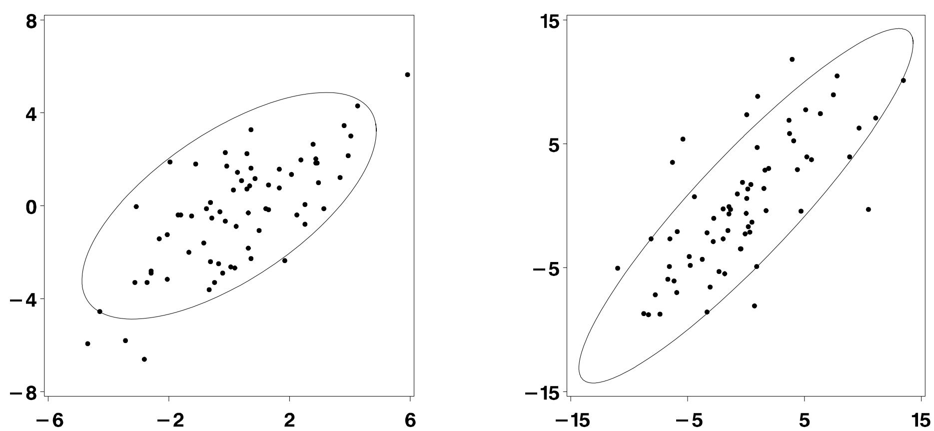

Figure 1.

REFERENCES

[1] COX, D. R.—HINKLEY, D. V.: *Problems and Solutions in Theoretical Statistics*, Chapman & Hall, London, 1978.
[2] EHRENBERG, A. S. C.: *Repeat-Buying: Facts, Theory and Applications*, Charles Griffin & Company, London, 1988.
[3] GRANDELL, J.: *Mixed Poisson Processes*, Chapman & Hall, London, 1997.

[4] MIKOSCH, T.: *Non-life Insurance Mathematics*, Springer-Verlag, Berlin, 2004.
[5] SERFLING, R. J.: *Approximation Theorems of Mathematical Statistics*, Wiley, New York, 1980.
[6] VAN DER VAART, A. W.: *Asymptotic Statistics*, Cambridge University Press, Cambridge, 1998.

Received September 29, 2006

School of Mathematics
Cardiff University
Cardiff, CF24 4AG
UNITED KINGDOM
E-mail: savaniv@cardiff.ac.uk
zhigljavskyaa@cardiff.ac.uk