POISSON LIMIT OF SUMS OF INDEPENDENT NON-NEGATIVE INTEGER RANDOM VARIABLES

By Stig Rosenlund

2011-03-23

We deduce the limit distribution for sums of independent non-negative integer random variables, such that the sequence of sum expectations converges to a positive number and such that in the limit each summand is infinitesimally small and at most 1, in a certain sense. This limit distribution is Poisson. The well-known rare events theorem is a special case. Our results are compared to the Poisson process limit result of Grigelionis (1963). Applications to mixed Poisson distributions and to renewal processes are given. Implications for insurance claim number distributions are stated. *

1. Introduction. We consider here a triangular array in the sense of Feller (1971), VI.3, Definition 2, i.e. a double sequence of random variables N_{Ki} $(i = 1, \ldots, K;\ K = 1, 2, \ldots)$ such that the $N_{K1}, \ldots, N_{KK}$ of the Kth row are independent. The variables need not be defined on the same probability space for different K. When we write P, E and Var it is implicit that these functionals are defined on the probability space where N_{Ki} is defined. This is unambiguous since in any formula below only one K is present. We assume N_{Ki} to be non-negative and integer-valued. We study weak convergence of the row sums to the Poisson distribution as $K \to \infty$.

Here N_{Ki} can be thought of as counting some kind of events, e.g. claim occurrences on an insurance policy, in a certain time interval, whose length will be fixed initially. Later we will discuss increasing time lengths with special reference to insurance claims.

In section 5 implications are stated for applications, especially insurance claim number distributions. The point is – for many sources, each contributing little to the total – that the mixed Poisson distribution (customarily with a gamma claim frequency giving a negative binomial claim number) is an unnecessary complication and that the simple Poisson distribution suffices.

In Theorem 2.1 of section 2 we give a rather short self-contained proof of the limiting Poisson distribution of a sum with expectation converging to a positive number, under two simple conditions implying infinitesimally small

AMS 2000 subject classifications: Primary 60F05; secondary 91B30

Keywords and phrases: GLM log link, Mixed Poisson, Poisson limit, Rare events theorem, Renewal process.

*These results are to a large extent known already, so the paper cannot be published. See Barbour et. al. (1992)

independent summands each being, in the limit, at most 1. Theorem 2.1 is a generalization of the 'rare events' result for independent 0/1 Bernouilli variables with possibly unequal probabilities for the value 1. See the comment after the proof. Koopman (1950) gave necessary and sufficient conditions for the result. The subject is classical and the 'rare events' result has been generalized in different directions. See e.g. Meyer (1973), where further references can be found. The present result, where the variables can take any non-negative values, is new. It is simple, strong and useful. In Theorem 3.1 of section 3.2 the two conditions are verified to hold for mixed Poisson distributions of N_{Ki} under boundedness conditions. A result with the weaker condition that variance/mean converges to 1, using the properties of mixed Poisson, is also given as Theorem 3.2 of section 3.3. In Theorem 4.1 we use Theorem 2.1 to give conditions for superpositions of renewal point processes to give a Poisson distribution in the limit for a fixed time interval.

Grigelionis (1963) showed weak convergence of superpositions of counting processes to a Poisson process under various conditions, which imply that in the limit each process contributes infinitesimally little to the total and at most one point, if at all, in any finite time interval. The scopes of these results are broader than ours in that we do not consider a whole process. On the other hand we can make do with weaker conditions. Take processes over a fixed time interval, e.g. a year, where our conditions and results hold, but where the bulk (e.g. 80 percent) of the total number of points occur alternatingly in the first half and the second half of the interval as the number K of processes increases. Then the finite-dimensional distributions of the sum process for collections of time points do not all converge.

2. Poisson limit of a sum of small summands.

Define a triangular array of independent variables, for example claim numbers on policies,

N_{Ki} = non-negative integer-valued random variable for $i \in \{1, \ldots, K\}$

The sum of these and its mean, assumed to converge, are

$$N_K = \sum_{i=1}^{K} N_{Ki}$$

$$E[N_K] \to \lambda \ (K \to \infty) \quad \text{for a fixed number } \lambda > 0.$$

Theorem 2.1. *Assume that uniformly in $i \in \{1, \ldots, K\}$*

(i) $\lim_{K\to\infty} P(N_{Ki} = 0) = 1$

(ii) $\lim_{K\to\infty} P(N_{Ki} = 1)/E[N_{Ki}] = 1$ *(define the ratio to be 1 if $E[N_{Ki}] = 0$)*

Then the distribution of $N_K \to$ Poisson(λ) as $K \to \infty$.

PROOF. We will show that for $s > 0$ the Laplace transform $E[e^{-sN_K}] \to \exp\{\lambda(e^{-s} - 1)\}$. See Feller (1971), XIII.1 Theorem 2, for the continuity theorem for Laplace transforms. Let $\epsilon_{Ki}^{(j)}$ be numbers that $\to 0$ and $\delta_{Ki}^{(j)}$ be numbers that $\to 1$, uniformly in i. The assumptions

$$P(N_{Ki} = 0) = \delta_{Ki}^{(1)} \quad \text{and} \quad P(N_{Ki} = 1) = E[N_{Ki}]\delta_{Ki}^{(2)}$$

imply

$$\begin{aligned}
E[N_{Ki}] &= P(N_{Ki} = 1)/\delta_{Ki}^{(2)} \leq P(N_{Ki} \geq 1)/\delta_{Ki}^{(2)} \\
&= (1 - \delta_{Ki}^{(1)})/\delta_{Ki}^{(2)} = \epsilon_{Ki}^{(1)}
\end{aligned}$$

$$\begin{aligned}
P(N_{Ki} \geq 2) &\leq \sum_{r=2}^{\infty} rP(N_{Ki} = r) = E[N_{Ki}] - P(N_{Ki} = 1) \\
&= E[N_{Ki}](1 - \delta_{Ki}^{(2)}) = E[N_{Ki}]\epsilon_{Ki}^{(2)}
\end{aligned}$$

$$\begin{aligned}
P(N_{Ki} = 0) &= 1 - P(N_{Ki} = 1) - P(N_{Ki} \geq 2) \\
&= 1 - E[N_{Ki}]\delta_{Ki}^{(2)} - E[N_{Ki}]\epsilon_{Ki}^{(3)} = 1 - E[N_{Ki}]\delta_{Ki}^{(3)}
\end{aligned}$$

and thus

$$\begin{aligned}
E[e^{-sN_{Ki}}] &= P(N_{Ki} = 0) + P(N_{Ki} = 1)e^{-s} + \sum_{r=2}^{\infty} e^{-sr} P(N_{Ki} = r) \\
&= 1 - E[N_{Ki}]\delta_{Ki}^{(3)} + E[N_{Ki}]\delta_{Ki}^{(2)} e^{-s} + E[N_{Ki}]\epsilon_{Ki}^{(4)} \to 1 \\
&\quad \text{uniformly in } i
\end{aligned}$$

which means

$$\begin{aligned}
\log E[e^{-sN_{Ki}}] &= E[N_{Ki}]\delta_{Ki}^{(4)} \left(-\delta_{Ki}^{(3)} + \delta_{Ki}^{(2)} e^{-s} + \epsilon_{Ki}^{(4)}\right) \\
&\quad \text{since } \log(1 + x)/x \to 1 \text{ as } x \to 0
\end{aligned}$$

so for the sum

$$\begin{aligned}
\log E[e^{-sN_K}] &= \log E\left[\exp\{-s\sum_{i=1}^{K} N_{Ki}\}\right] = \log E\left[\prod_{i=1}^{K} e^{-sN_{Ki}}\right] \\
&= \log \prod_{i=1}^{K} E[e^{-sN_{Ki}}] = \sum_{i=1}^{K} \log E[e^{-sN_{Ki}}] \\
&= \sum_{i=1}^{K} E[N_{Ki}]\delta_{Ki}^{(4)} \left(-\delta_{Ki}^{(3)} + \delta_{Ki}^{(2)} e^{-s} + \epsilon_{Ki}^{(4)}\right) \to \lambda(e^{-s} - 1) \quad \square
\end{aligned}$$

The uniform convergence in (i) and (ii) is crucial. As a counter-example for (i), consider the case $P(N_{KK} = 1) = 1 - P(N_{KK} = 0) = 0.5$ and $P(N_{Ki} = 1) = 1 - P(N_{Ki} = 0) = 0.5/(K-1)$ for $i < K$, so that $E[N_K] = 1$. Then $P(N_{Ki} = 0) \to 1$ for any i, but Poisson convergence does not hold.

For the special case $P(N_{Ki} \geq 2) = 0$, where (ii) holds automatically, Theorem 2.1 is the 'rare events' Poisson convergence result for 0/1 variables.

In the following we will simplify the setting by letting the sum have a fixed expectation $E[N_K] = \lambda$. All results below will hold for $E[N_K] \to \lambda$ with obvious modifications, such as replacing λ with $\lambda(1+\epsilon_K)$ where $\epsilon_K \to 0$.

3. Independent mixed Poisson variables. Conditional on independent variables $\xi_1, \ldots, \xi_K$ with $E[\xi_i] = \mu_i$ and $\mathrm{Var}[\xi_i] = \sigma_i^2$, let N_{Ki} be independent Poisson variables with $E[N_{Ki}] = h_K \xi_i$. Let N_K be defined as before and let its mean be fixed to λ. We have

$$\lambda = E[N_K] = E\left[\sum_{i=1}^{K} N_{Ki}\right] = h_K \sum_{i=1}^{K} \mu_i$$

$$h_K = \lambda \left(\sum_{i=1}^{K} \mu_i\right)^{-1} \tag{3.1}$$

We compute the variance of N_K and hence its dispersion parameter. Then we give two derivations of the Poisson limit as $K \to \infty$. First we derive Theorem 3.1 using Theorem 2.1 under two boundedness conditions. Then we give Theorem 3.2, using the special properties of the mixed Poisson distribution under a weaker condition. Theorem 3.1 thus follows from Theorem 3.2, but is justified since its proof illustrates the use of Theorem 2.1.

3.1. *Dispersion parameter.* Define the dispersion parameter ϕ_K as the ratio $\mathrm{Var}[N_K]/E[N_K]$. Its excess over 1 measures the degree of overdispersion compared to the pure Poisson distribution, which has $\phi_K = 1$. It holds

$$\begin{aligned} \mathrm{Var}[N_{Ki}] &= E[\mathrm{Var}[N_{Ki}|\xi_i]] + \mathrm{Var}[E[N_{Ki}|\xi_i]] \\ &= E[\xi_i h_K] + \mathrm{Var}[\xi_i h_K] = \mu_i h_K + \sigma_i^2 h_K^2 \end{aligned} \tag{3.2}$$

$$\phi_K = \frac{\mathrm{Var}[N_K]}{E[N_K]} = \frac{1}{\lambda} \sum_{i=1}^{K} \mathrm{Var}[N_{Ki}] = \frac{1}{\lambda} \sum_{i=1}^{K} (\mu_i h_K + \sigma_i^2 h_K^2) \tag{3.3}$$

$$= 1 + \frac{h_K^2}{\lambda}\sum_{i=1}^{K}\sigma_i^2 = 1 + \lambda\left(\sum_{i=1}^{K}\sigma_i^2\right)\left(\sum_{i=1}^{K}\mu_i\right)^{-2}$$

Example. Take $K = 100{,}000$ insurance policies in one year with claim frequency (mean claim number per policy and year) 0.07, so that $\lambda = 7{,}000$. Take all $\mu_i = 10$ and all $\sigma_i^2 = 50$. Then ϕ_K is close to 1, although σ_i^2 is large.

$\phi_K = 1 + 7000 \times 100000 \times 50/(100000 \times 10)^2 = 1.035$

3.2. *Poisson limit result using Theorem 2.1.* We shall prove that the conditions for Theorem 2.1 are true if the ξ_is' means are bounded away from 0 and ∞ and their variances are bounded away from ∞. These assumptions are

3-A1. $0 < a_1 \le \mu_i \le a_2 < \infty$

3-A2. $0 \le b_1 \le \sigma_i^2 \le b_2 < \infty$

and imply

$$\frac{\lambda}{Ka_2} \le h_K \le \frac{\lambda}{Ka_1} \tag{3.4}$$

For the dispersion parameter **3-A1** and **3-A2** imply

$$1 + \frac{\lambda b_1}{Ka_2^2} \le \phi_K \le 1 + \frac{\lambda b_2}{Ka_1^2} \tag{3.5}$$

where the inequalities are equalities if all $\mu_i = a_1 = a_2$ and all $\sigma_i^2 = b_1 = b_2$.

Theorem 3.1. *If* **3-A1** *and* **3-A2** *hold, the distribution of* $N_K \to$ *Poisson*(λ) *as* $K \to \infty$.

Proof. Condition (i) that $P(N_{Ki} = 0) \to 1$ uniformly in i is verified by

$$1 > P(N_{Ki} = 0) = 1 - P(N_{Ki} \ge 1) \ge 1 - E[N_{Ki}] = 1 - h_K\mu_i \ge 1 - \frac{\lambda a_2}{Ka_1}$$

For condition (ii) that $P(N_{Ki} = 1)/E[N_{Ki}] \to 1$ uniformly in i, observe that $e^{-x} \ge 1 - x$. Thus

$$\begin{aligned} 1 \ge \frac{P(N_{Ki} = 1)}{E[N_{Ki}]} &= \frac{E[h_K\xi_i e^{-h_K\xi_i}]}{h_K\mu_i} \ge \mu_i^{-1}E[\xi_i(1 - h_K\xi_i)] \\ &= \mu_i^{-1}(E[\xi_i] - h_K E[\xi_i^2]) = \mu_i^{-1}(\mu_i - h_K[\mu_i^2 + \sigma_i^2]) \\ &= 1 - h_K\left(\mu_i + \frac{\sigma_i^2}{\mu_i}\right) \ge 1 - \frac{\lambda}{Ka_1}\left(a_2 + \frac{b_2}{a_1}\right) \end{aligned}$$

□

3.3. *Poisson limit result using mixed Poisson properties.* The conditions for Theorem 3.1 are unnecessarily retrictive. Take the case $\mu_1 = \lambda$, $\sigma_1^2 = 0$ and $\mu_i = 0$ for $i > 1$. For ξ_i that are not stochastic N_{Ki} will be Poisson, so those μ_i can be large without decreasing the degree of Poisson approximation.

We derive the Poisson limit under the weaker condition that the dispersion parameter converges to 1. By (3.5) the conditions **3-A1** and **3-A2** together imply this condition, so Theorem 3.2 is stronger than Theorem 3.1.

Theorem 3.2. *If ϕ_K in (3.3) $\to 1$ as $K \to \infty$, the distribution of $N_K \to$ Poisson(λ) as $K \to \infty$.*

Proof. Set $\eta_K = h_K \sum_{i=1}^{K} \xi_i =$ the sum of all random Poisson means. Then $N_K | \eta_K \sim \text{Poisson}(\eta_K)$. Now

$$E[\eta_K] = h_K \sum_{i=1}^{K} \mu_i = \lambda \text{ and } \mathrm{Var}[\eta_K] = h_K^2 \sum_{i=1}^{K} \sigma_i^2 = \lambda(\phi_K - 1) \to 0 \Rightarrow \eta_K \xrightarrow{\text{P}} \lambda$$

Hence

$$E[e^{-sN_K}] = E[E[e^{-sN_K} | \eta_K]] = E[\exp\{\eta_K(e^{-s} - 1)\}] \to \exp\{\lambda(e^{-s} - 1)\} \quad \square$$

3.4. *Insurance claim number distribution for long time length.* Consider the insurance claim number application. Assuming the uniform boundedness conditions **3-A1** and **3-A2**, $\phi_K - 1$ is essentially proportional to the expected number of claims per policy, i.e. λ/K, as is seen from (3.5). Consider a time interval $(0, t)$ and let $\lambda = \nu t$, where t increases while K is kept fixed. Here we drop the condition that λ is fixed. Let us preserve the construction of variables ξ_i with fixed means and variances and assume that exposure of every policy increases linearly with time. E. g. a customer insured for half the year will continue so. Then $h_K = \nu t / \sum_{i=1}^{K} \mu_i$ where the μ_i do not change. If $b_1 > 0$ then $\lim_{t\to\infty} \phi_K = \infty$, so that the distribution of N_K deviates increasingly from Poisson. The more the expected number of claims per policy increases, the more violated both conditions (i) and (ii) for Theorem 2.1 will be.

How should the mixed Poisson model be applied to forecast the future when statistics collected in a long time period are used?

In e. g. automobile insurance consumers mostly change insurer within a couple of years. The number K in our model is the number of distinct customers, which thus would tend to increase approximately proportionally to t, even if the total portfolio at any given time point stays the same.

On the other hand, consider a faithful customer. We can model its claims to follow mixed Poisson distributions based on a sequence of random intensities for time periods (years) $\in \{1, 2, \ldots\}$, with non-negative autocorrelation. Or we can assume a single random intensity. Either way, whether a faithful customer included in the statistics increases or decreases the Poisson closeness of the claim number distribution, for practical purposes, depends on whether the customer will or will not be part of the portfolio in the future that we forecast. A single random intensity for a customer that stays in the future would obviously best be considered to be a non-stochastic unknown claim frequency, making its claims follow a pure Poisson distribution. In our model we can set its $\sigma_i^2 = 0$ and hence it contributes 0 to $(\phi_K - 1)$ in Theorem 3.2. Bonus/malus calculation for an individual customer is another matter and requires the mixed Poisson model.

The worst case for the Poisson approximation is a customer that was faithful in the past but will not be part of the future portfolio. For instance, if the claim statistics for one insurer is used to forecast the future or set the prices for another insurer.

4. Superpositions of renewal processes. Feller (1971), XI.4 Example (a), treats the waiting time for the first renewal following epoch 0, in a sum of many independent renewal processes, and shows that it is approximately exponentially distributed. There are also results showing asymptotically exponential distributions for waiting times between points in renewal processes under thinning mechanisms, such that the expected number of points in any finite interval tends to zero. See Rosenlund and Råde (1974), Theorems 1 and 2.

We will now use Theorem 2.1 to show a Poisson approximation for the number of points in a finite time interval.

Define random variables X_0 with distribution function G and X with distribution function F. For process $i \in \{1, \ldots, K\}$, let the waiting time to the first point be distributed as $X_0/(h_K \mu_i)$ and the following waiting times as $X/(h_K \mu_i)$, where all waiting times are independent. That is, the points come in K independent delayed renewal processes as defined in Feller (1971), VI.6 Definition 3. We shall let $h_k \to 0$ and $K \to \infty$, i.e. we add successively more processes each of which has successively fewer points per time interval.

If we let G be the limiting residual waiting time distribution function for a renewal process with inter-arrival distribution function F

$$G(x) = E[X]^{-1} \int_0^x [1 - F(\tau)]\, d\tau$$

the expected number of points in the time interval $(0, t)$ will be linear in t for every i, if F is non-arithmetic. See Feller (1971), XI.4. But we will consider arbitrary G, such as $G = F$.

Let N_{Ki} and their sum N_K be defined as before and pertain to the number of points in the interval $[0, t]$ for some $t > 0$. We assume the following for our limit result, meaning that we could as well write $(0, t]$.

4-A1. $F(0) = 0$

4-A2. $G(x) \sim c\, x^p$ as $x \downarrow 0$ for some $c > 0$ and $p > 0$

4-A3. $\mu_i \le a_2 < \infty$

4-A4. $\sum_{i=1}^{\infty} \mu_i^p = \infty$

Theorem 4.1. *Assume* **4-A1**, **4-A2**, **4-A3** *and* **4-A4**. *Set h_K if possible so that $E[N_K] = \lambda$, with λ fixed, for a time interval $[0, t]$. It is possible for large enough K. Then as $K \to \infty$ the distribution of $N_K \to$ Poisson(λ) and*

$$h_K \sim \frac{1}{t}(\lambda/c)^{\frac{1}{p}} \left(\sum_{i=1}^{K} \mu_i^p \right)^{\frac{-1}{p}}$$

Proof. Let $\star$ denote convolution. We denote by $F^{r\star}$ the r-fold convolution of F, i.e. the distribution function of r independent random variables distributed as X. By $G\star F(x)$ we mean $H(x)$ where $H = G\star F$, etc. By Feller (1971), XI.4 (4.2), and from the definitions

$$\begin{aligned}
E[N_{Ki}] &= \sum_{r=0}^{\infty} G\star F^{r\star}(h_K \mu_i t) \\
P(N_{Ki} > r) &= G\star F^{r\star}(h_K \mu_i t) \\
P(N_{Ki} = r) &= P(N_{Ki} > r-1) - P(N_{Ki} > r) \\
&= G\star F^{(r-1)\star}(h_K \mu_i t) - G\star F^{r\star}(h_K \mu_i t),\ r \ge 1 \\
P(N_{Ki} = 0) &= 1 - G(h_K \mu_i t) \\
P(N_{Ki} = 1) &= G(h_K \mu_i t) - G\star F(h_K \mu_i t)
\end{aligned}$$

For K large enough we can set h_K so that $E[N_K] = \lambda$ for any $\lambda > 0$. Namely, $V_K(h) = E[N_K] = \sum_{i=1}^{K} \sum_{r=0}^{\infty} G\star F^{r\star}(h\mu_i t)$ is a non-decreasing and right-continuous function of h with $V_K(0) = 0$ and $V_K(\infty) = \infty$. If there exists h such that $V_K(h) = \lambda$, then set $h_K = h$. If not, i.e. if $V_K(\cdot)$ jumps from below to above λ at some point h, which then must be positive, then set $h_K = 0.9h$. Hence h_K is well defined and $E[N_K] \le \lambda$. We shall show that $h_K \to 0$. Then for large enough K the continuity of $G\star F^{r\star}$ at 0 implies $E[N_K] = \lambda$.

Assumption **4-A2** implies that $G(x) \ge 0.9c\, x^p$ for $x < x_0$ where $x_0 > 0$, hence $G(x) \ge 0.9c[\min(x, x_0)]^p$. We proceed by contradiction. Assume that $h_K > h > 0$ for infinitely many K. Now for those K we have $E[N_{Ki}] \ge G(h_K\mu_i t) \ge G(h\mu_i t) \ge 0.9c[\min(h\mu_i t, x_0)]^p = 0.9c(ht)^p[\min(\mu_i, \frac{x_0}{ht})]^p$. With $x_1 = \frac{x_0}{ht} > 0$ we have for those K

$$\lambda \ge \sum_{i=1}^{K} E[N_{Ki}] \ge 0.9c(ht)^p \left(\sum_{\{i:1\le i\le K,\ \mu_i\le x_1\}} \mu_i^p + \sum_{\{i:1\le i\le K,\ \mu_i>x_1\}} x_1^p \right)$$

Either $\mu_i > x_1$ for infinitely many i, in which case the right sum $\to \infty$ for the K in question. Or there is K_0 such that $\mu_i \le x_1$ for $i \ge K_0$, in which case the left sum $\to \infty$ by assumption **4-A4**. Hence by contradiction $h_K \to 0$.

By **4-A3** we have $P(N_{Ki} = 0) \ge 1 - G(h_K a_2 t) \to 1$, hence condition (i) for Theorem 2.1 that $P(N_{Ki} = 0) \to 1$ uniformly in i is satisfied.

Now we have to assert condition (ii) that $P(N_{Ki} = 1)/E[N_{Ki}] \to 1$ uniformly in i. It holds $\{X + Y \le x\} \subset \{X \le x\} \cap \{Y \le x\}$ for non-negative random variables, so if they are independent we have $P(X + Y \le x) \le P(X \le x)P(Y \le x)$. Thus $G\star F(x) \le G(x)F(x)$ and $F^{r\star}(x) \le F(x)^r$, for example.

Let K be so large that $F(h_K a_2 t) < 1$. Using **4-A1** and **4-A3** we obtain

$$G(h_K\mu_i t) \le E[N_{Ki}] \le G(h_K\mu_i t)\sum_{r=0}^{\infty} F(h_K\mu_i t)^r = \frac{G(h_K\mu_i t)}{1 - F(h_K\mu_i t)}$$

$$\begin{aligned} 1 \ge \frac{P(N_{Ki}=1)}{E[N_{Ki}]} &\ge [G(h_K\mu_i t) - G(h_K\mu_i t)F(h_K\mu_i t)]\left(\frac{G(h_K\mu_i t)}{1 - F(h_K\mu_i t)}\right)^{-1} \\ &= [1 - F(h_K\mu_i t)]^2 \ge [1 - F(h_K a_2 t)]^2 \to 1 \end{aligned}$$

which proves that condition (ii) holds.

As before, let $\delta_{Ki}^{(j)}$ be numbers that $\to 1$, uniformly in i. By **4-A2** we get the following, from which the asymptotic expression for h_K is deduced.

$$\sum_{i=1}^{K} E[N_{Ki}] = \sum_{i=1}^{K} G(h_K\mu_i t)\delta_{Ki}^{(1)} = \sum_{i=1}^{K} c\,(h_K\mu_i t)^p\,\delta_{Ki}^{(2)} = c\,h_K^p t^p \sum_{i=1}^{K} \mu_i^p\,\delta_{Ki}^{(2)} = \lambda \quad \square$$

If in **4-A3** we make the stronger assumption that $0 < a_1 \le \mu_i \le a_2 < \infty$, as in **3-A1**, then **4-A4** is implied.

With F non-arithmetic and G the limiting residual waiting time distribution we have $c = E[X]^{-1}$ and $p = 1$. Then $E[N_{Ki}] = h_K \mu_i E[X]^{-1} t$. The asymptotic expression for h_K in Theorem 4.1 will then be exact. If we scale X so that $E[X] = t$, then the numbers h_K and μ_i will have the same meaning for $E[N_{Ki}]$ as in section 3.

5. Conclusions for applications. Much actuarial literature concerns generalizations of the simple Poisson model for claim number distributions. It has been thought that independent random Poisson intensities generate the so called Overdispersed Poisson model, where $\phi = \text{Var}[N]/E[N] > 1$ for a claim number N and ϕ is the same number for all claim numbers regardless of time length and background variables. See Renshaw (1994). This misconception was laid to rest in Rosenlund (2010).

A more general notion is that the mixed Poisson model has to be used in order to allow for overdispersion. See e.g. Ohlsson and Johansson (2010), Remark 2.2 and section 3.4. However, our results here show that this is an unnecessary complication for an insurance line with moderately many policies and a small claim frequency. The simple Poisson model will suffice here. Estimating overdispersion in the mixed Poisson model, assuming independent policies and using Pearson's χ^2, may be misleading. What will be estimated will, for these insurance lines, likely be a measure of fluctuations affecting many policies in the same way. It can be seasonal variations, in which case the variability of the claim numbers between calendar years tends to be overestimated. And/or it can be a measure of other macroscopic influences such as business cycles and crime waves, which should be analyzed by other methods.

These conclusions can be applied to other models for occurrences of events, where many sources contribute to the total and each source contributes little and almost at most one event, in the sense of this paper. For instance customer arrivals at queueing systems such as telephone exchanges. Letting each source have a random Poisson intensity, independently of other sources, will still result in an approximative Poisson distribution for the total. Either mixed Poisson with a mixing variable common to all sources, or else pure Poisson – possibly after conditioning with respect to a mixing variable – should be used.

Finally, consider GLM log link analysis of claim numbers for an insurance multiplicative tariff. See Ohlsson and Johansson (2010), section 2.3. With many tariff cells most will contain at most one policy, to which our results are not applicable. But the GLM Poisson log link claim frequency parameter point estimates are computed only from the marginal numbers of claims in the different arguments. They are the same as the marginal totals method estimates, which depend only on claim statistics summed to the marginals. So in order to justify the pure Poisson model, it suffices that the marginals contain claims from sufficiently many independent policies, of which each one contributes little to the distribution of the marginal number of claims and has at most one claim in the sense of Theorem 2.1 . It is not neccessary that all marginals satisfy this condition. A marginal with only a few policies will have a small effect on the parameter estimates of those with many policies. In a typical mass consumer application there will be many arguments and tariff cells, but most marginals will have many policies.

References.

Barbour, A. D., Holst, L. and Janson, S. (1992) *Poisson Approximation.* Oxford University Press, Oxford.

Feller, W. (1971). *An Introduction to Probability Theory and Its Applications, Volume II.* 2nd ed. Wiley, New York.

Grigelionis, B. (1963). On the convergence of sums of random step processes to a Poisson process. *Theory Probab. Appl.* **8(2)** 177–182.

Koopman, B. O. (1950). Necessary and sufficient conditions for Poisson's distribution. *Proc. Amer. Math. Soc.* **1950** 813–823.

Meyer, R. M. (1973). A Poisson-type limit theorem for mixing sequences of dependent 'rare' events. *Ann. Probab.* **1(3)** 480–483.

Ohlsson, E. and Johansson, B. (2010) *Non-Life Insurance Pricing with Generalized Linear Models.* Springer, Berlin.

Renshaw, A. E. (1994). Modelling the claims process in the presence of covariates. *Astin Bull.* **24:2** 265–285.

Rosenlund, S. I. and Råde, L. (1974). Waiting for clusters in inhibited renewal processes. *Studia Sci. Math. Hungar.* **9** 341–346.

Rosenlund, S. (2010). Dispersion estimates for Poisson and Tweedie models. *Astin Bull.* **40:1** 271–279.

Stockholm
Sweden
E-mail: stig.ingvar.rosenlund@gmail.com