THE GENERALIZED WARING PROCESS

Mimoza Zografi and Evdokia Xekalaki *

Abstract – The Generalized Waring Distribution is a discrete distribution with a wide spectrum of applications in areas such as accident statistics, income analysis, environmental statistics, etc. It has been used as a model that better describes such practical situations as opposed to the Poisson Distribution or the Negative Binomial Distribution. Associated to both the Poisson and the Negative Binomial distributions are the well known Poisson and Pólya Processes. In this paper, the Generalized Waring Process is defined. Two models have been shown to lead to the Generalized Waring process. One is related to a Cox Process while the other is a Compound Poisson Process. The defined Generalized Waring Process is shown to be a stationary but non-homogenous Markov process. Several properties are studied and the intensity, the individual intensity, the Chapman-Kolmogorov differential equations of it, are obtained. Moreover, the Poisson and the Pólya processes are shown to arise as special cases of the Generalized Waring Process. Using this fact, some known results and some properties of them are obtained.

Keywords and Phrases – Pólya process, accident proneness, accident liability, Markovian property, stationary increments, Cox process, transition probabilities, Chapman-Kolmogorov equations, individual intensity.

I. INTRODUCTION-BASIC CONCEPTS

The Poisson and the Pólya processes have been used in Accident Theory to describe the accident pattern. Under the hypotheses of pure chance, the Poisson process with intensity λ has been proposed as a model that can describe the number of accidents sustained by an individual during several years. The Pólya process which is of Negative Binomial form, is defined by starting from a Poisson process, which then, is mixed with a Gamma distribution. It has been obtained as a model which can describe the accident pattern of a population of individuals during several years, under the hypotheses of *"accident proneness"* i.e that individuals differ in their probabilities of having an accident which remain constant in time [7]. Both of these processes satisfy the Markovian property because this is a property of the accident pattern, i.e. the number of accidents during the 'next' period $(t, t + h]$ depends only on the number of accidents at the present time t.

In this paper, a new process is defined and studied. This process is associated with a discrete distribution with a wide spectrum of applications known in the literature as the generalized Waring distribution (see, e.g. [6], [10]). This new process is termed in the sequel as the *generalized*

Waring process. Analogously to the case of Poisson and Pólya process, the generalized Waring process is postulated as a Markov process, as shown in section II. The starting point is a process of Negative binomial form but different from a Pólya process. This process is then mixed with a Betall distribution of the second type. This scheme is shown to lead to the generalized Waring process. A proof that this process is a stationary but non-homogenous Markov process is also provided. Further, an alternative genesis scheme referring to Cresswell and Froggatt's [2] *spells model* is proposed in the framework considered by Xekalaki [10]. This scheme, too, allows for stationary increments and the validity of the Markovian property, as also shown in section II.

Section III indicates how the above considerations formulate the framework for the definition of the generalized Waring process as a stationary but non-homogenous Markov process. Expressions for the first two moments of this process, as well as results on the intensity and the individual intensity of it, are also given in section III and its transition probabilities and the associated forward and the backward Chapman-Kolmogorov differential equations are derived.

The Poisson process and the Pólya process are special cases of the Generalized Waring process. Using this fact, some known theoretical results concerning these processes are presented and their transition probabilities and the associated Chapman-Kolmogorov differential equations, are derived in this context (section IV). Finally, two further genesis schemes are considered in section V. The results are based on Zografi and Xekalaki [14] and have been obtained in the context of models that have widely been considered for the interpretation of accident data. However, the concepts and terminology used can easily be modified so that the obtained results can be applied in several other fields ranging from economics, inventory control and insurance through to demometry, biometry and psychometry.

II. THE BASIC HYPOTHESES OF THE GENERALIZED WARING PROCESS

A. The description of the accident pattern by a Cox Process.

In this section, we consider first the assumptions of a *Pólya Process*, developed by Newbold [7]. This model considers several individuals exposed to the same external risk (e.g. drivers all driving about the same distance in a similar traffic environment) and that there are intrinsic differences among different individuals (e.g. differences in accident

* ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS, Patission 76, 104 34, ATHENS, GREECE

proneness). Supposing that, the number of accidents up to time t, for each individual, conforms with a Poisson process with a "personal rate λ" (λ stands for the respective accident proneness), and regarding λ as the outcome of a random variable Λ with a Gamma distribution with parameters k and $1/v$, the number of accidents $N(t)$ at time t, $t = 0, 1, 2, \ldots$ defines the Pólya Process with parameters k and $1/v$ thus: (I) $N(0) = 0$, (II) N is a birth process, (III) $N(t + h) - N(t)$ has a distribution defined by the probability function

$$P\{N(t + h) - N(t) = m\} = E\left[\frac{(\Lambda h)^n}{n!}e^{-\Lambda h}\right]$$

$$= \binom{k + m - 1}{m}\left(\frac{1}{1 + vh}\right)^k\left(\frac{vh}{1 + vh}\right)^m$$

$$m = 0, 1, \ldots \qquad (2.1.1)$$

where Λ is $\Gamma(k, 1/v)$-distributed. It is clear that $N(t)$ has a Negative Binomial distribution with parameters k and $1/vt$, i.e. $N(t) \sim NB(k, 1/vt)$.

The distribution of the random variable Λ explains here the variation of the accident proneness from individual to individual. As noted by Irwin [5] and Xekalaki [11], the term accident proneness here refers to both, external and internal risk of accident. It seems more natural to assume that this variation in an interval of time $(t, t + h]$ depends on the length h of the interval, while, in two non-overlapping time periods, the respective variations are independent. So, now, a personal λ, in an interval of time $(t, t + h]$, is regarded as the outcome of a random variable $\Lambda(h)$ with distribution $U(h)$ which depends on the interval length h. If $U(h)$ is assumed to be $\Gamma(k(h), 1/v(h))$, where $k(h)$ and $v(h)$ are in general some functions of h, then, clearly the number of accidents $N(t)$ will form a stochastic process of a Negative binomial form satisfying the assumptions

$$(I)\, N(0) = 0$$

and

(II) $N(t + h) - N(t)$ has the distribution:

$$P\{N(t + h) - N(t) = n\} =$$

$$\int_0^{+\infty}\frac{(\lambda h)^n}{n!}e^{-\lambda h}\frac{v(h)^{-k(h)}}{\Gamma(k(h))}\lambda^{(k(h)-1)}e^{(-\lambda/v(h))}d\lambda, \quad n = 0,1,\ldots$$

$$(2.1.2)$$

It can be shown that

$$P\{N(t + h) - N(t) = n\} =$$

$$\binom{k(h) + n - 1}{n}\left(\frac{1}{1 + v(h)h}\right)^{k(h)}\left(\frac{v(h)h}{1 + v(h)h}\right)^n$$

Then, using the first assumption, it follows that for any t, $N(t)$ has Negative Binomial distribution with parameters $k(t)$ and $1/tv(t)$. Hence, one can confirm that the following relation stands

$$P\{N(t) = n\} =$$

$$\int_0^{+\infty}\frac{(\lambda t)^n}{n!}e^{-\lambda t}\frac{v(t)^{-k(t)}}{\Gamma(k(t))}\lambda^{(k(t)-1)}e^{-\lambda/v(t)}d\lambda, \quad n = 0,1,\ldots$$

This tells us precisely that $N(t)$ is a *Cox Process* (see e.g. Grandell, [3], p.83).

Assume that the accident proneness varies from individual to individual with a mean that does not depend on time. This is equivalent to considering a parameter pair $(k(h), v(h))$ with $k(h) \cdot v(h) = \text{constant}$. So, letting $v(h) = v/h$, and $k(h) = kh$, i.e., allowing $\Lambda(h)$ to have a gamma distribution that changes with time so that its expectation remains constantly equal to vk, we obtain

$$P\{N(t + h) - N(t) = n\} =$$

$$\binom{kh + n - 1}{n}\left(\frac{1}{1 + v}\right)^{kh}\left(\frac{v}{1 + v}\right)^n, \quad n = 0,1,\ldots \qquad (2.1.3)$$

and that $N(t)$ is $NB(kt, 1/v)$-distributed.

B. An extension of Irwin's accident model.

This model considers a population which is not homogeneous with respect to personal and environmental attributes that affect the occurrence of accidents. In his model, Irwin used the term "accident proneness" v to refer to a person's predisposition to accidents, and the term "accident liability" $(\lambda| v$, i.e λ for given $v)$ to refer to a person's exposure to external risk of accident.

The conditional distribution of the random variable Λ for a given v describes differences in external risk factors among individuals. As before, liability fluctuations over a time interval $(t, t + h)$ depend on the length h of the interval and are described by a $\Gamma(kh, 1/vh)$ distribution for $\Lambda| v$. Moreover, assuming independence in two non-overlapping time periods, the number of accidents $N(t)$, for a given v, will be a stochastic process of Negative binomial form with parameters kt and $1/v$. This starts at 0 and has stationary increments with a distribution given by (2.1.3). Let us further allow the parameter v of the Negative Binomial to follow a BetaII distribution with parameters a and ρ, obtaining thus for the distribution of the number of accidents $N(t)$:

$$P\big(N(t+h)-N(t)=n\big)=$$

$$\frac{\rho_{(kh)}}{(a+\rho)_{(kh)}}\frac{a_{(n)}(kh)_{(n)}}{(a+\rho+kh)_{(n)}}\frac{1}{n!}$$

and

$$P\big(N(t)=n\big)=P_n(t)=\frac{\rho_{(kt)}}{(a+\rho)_{(kt)}}\frac{a_{(n)}(kt)_{(n)}}{(a+\rho+kt)_{(n)}}\frac{1}{n!}$$

$$n=0,\ 1,\ \dots \qquad (2.2.1)$$

In the sequel, we refer to the process defined by $N(t)$ as the *Generalized Waring Process*.

C. The Markovian property
Theorem 2.3.1

Let $N(t)$ be defined as above. Then,

$$P\big(N(t+h)=n \mid N(t)=m,\ N(s)=n_s,\ 0\le s<t\big)$$

coincides with $P\big(N(t+h)=n \mid N(t)=m\big)$ for every non-negative integers n,m,n_s $0\le s<t$.

Proof: First we denote that

$$P\Big\{N(t+h)=n \mid N(t)=m, N\big(s\big)=n_s,\ 0\le s<t\Big\}$$

is equal to

$$P\{N(t+h)-N(t)=$$

$$n-m\mid N(t)-N(s)=m-n_sN\big(s\big)-N(0)=n_s,0\le s<t\}$$

Consider, now, the random vector

$$\Big(N(t+h)-N(t),\ N(t)-N(s),\ N\big(s\big)-N(0)\Big),\ 0\le s<t$$

As is known (see [11], [13]), this vector has a multivariate generalized Waring distribution $\big(MGWD(a;\underline{k};\rho)\big)$ where $\underline{k}=\big(kh,\ k(h-s),\ ks\big)$, and

$$\big(N(t+h)-N(t)\mid N(t)-N(s),N(s)-N(0)\big)$$

$$\sim MGWD\big(a+n(t),\ kh;\ \rho+kt\big),$$

where $n(t)$ is the value of $N(t)$.

Hence,

$$P\{N(t+h)-N(t)=$$

$$n-m\mid N(t)-N\big(s\big)=m-n_s,\ N\big(s\big)-N(0)=n_s=$$

$$=\frac{(\rho+kt)_{(a+m)}}{(\rho+kt+kh)_{(a+m)}}\frac{(a+m)_{(n-m)}(kh)_{(n-m)}}{(\rho+kt+kh+a+m)_{(n-m)}}\frac{1}{(n-m)!}$$

$$=\frac{\Gamma(a+n)}{\Gamma(a+m)}\frac{(kh)_{(n-m)}}{(n-m)!}\frac{(\rho+kt)_{(a+m)}}{(\rho+kt+kh)_{(a+n)}}$$

$$=P\big[N(t+h)-N(t)=n-m/N(t)-N(0)=m\big]=$$

$$P\big[N(t+h)=n \mid N(t)=m\big], \qquad (2.3.1)$$

which proves the theorem.

This result tells that the generalized Waring process has the Markovian property.

D. The Spells Model

In the sequel, an alternative scheme generating a process of a Generalized Waring form is considered. This is a variant of Cresswell and Froggatt's [2] *Spells* model that has been considered in the paper of Xekalaki [11]. According to this model, each person is liable to spells. For each person, no accidents can occur outside spells. Let $S(t)$ denote the number of spells up to a given moment t.

It is assumed that $S(t)$, $t=0,\ 1,\ 2,\ \dots$ is a homogeneous Poisson process with rate k/m, $k>0$, the number of accidents within a spell is a random variable with a given distribution F and that the number of accidents arising out of different spells are independent and also independent of the number of spells. So, the total number of accidents at time t is $X(t)=\sum_{k=1}^{S(t)}X_k$, where $S(t)$ is a homogenous Poisson process with rate k/m and $\{X_k\}_1^{\infty}$ are identically and independently distributed (i.i.d) random variables from the distribution F.

When $\{X_k\}_1^{\infty}$ is a logarithmic series distribution with parameters (m,v), i.e.

$$P(X_i=0)=1-m\log(1+v)\ \text{and}$$

$$P(X_i=n)=\frac{m}{n}\left(\frac{v}{1+v}\right)^n,\ n\ge1,\ m>0,\ v>0,\ \text{the}$$

random variable $X(t)$, is a Negative Binomial random variable with parameters $(kt,1/v)$ for each t [1]. Here v is regarded as the external risk parameter, too. Then, if the differences in this external risk can be described by a $BetaII(a,\rho)$ distribution, the resulting accident distribution is of a Generalized Waring form with parameters a, kt, and ρ.

Let us consider, now, the counting process $\{N(t),t\ge0\}$ where $N(t)$ can be represented, for $t\ge0$,

by $\sum_{k=1}^{S(t)}X_k$, $\left(\sum_{k=1}^{0}X_k=0\right)$, where $S(t)$ is a

homogenous Poisson process with rate k/m, $\{X_k\}_1^{\infty}$ has a logarithmic series distribution with parameters (m,v)

888

and is independent of the process $S(t)$, and v is a non negative random variable with a $BetaII(a, \rho)$ distribution.

Theorem 2.4.1

For the process $\{N(t), \ t \geq 0\}$ defined as above the following conditions hold:

$N(0) = 0$

$\{N(t), \ t \geq 0\}$ posseses stationary increments

$\{N(t), \ t \geq 0\}$ is a Markov process.

Proof: The proof of (I) is straightforward. To prove condition (II), denote by φ the probability distribution function (p.d.f) of the random variable v. Then we can write:

$$P(N(t+h) - N(t) = n) =$$

$$\int_0^{+\infty} P(N(t+h) - N(t) = n/v)\varphi(v)dv$$

$$= \int_0^{+\infty} P\left(\sum_{k=S(t)}^{S(t+h)} X_k = n\right)\varphi(v)dv$$

$$= \int_0^{+\infty} \left[\sum_{i=0}^{+\infty} P\left(\sum_{k=1}^{i} X_k = n\right) P(S(t+h) - S(t) = i)\right]\varphi(v)dv$$

$$= \int_0^{+\infty} \left[\sum_{i=0}^{+\infty} P\left(\sum_{k=1}^{i} X_k = n\right)\frac{1}{i!}\exp\left(-\frac{kh}{m}\right)\left(\frac{kh}{m}\right)^i\right]\varphi(v)dv$$

$$= \frac{\rho_{(kh)}}{(\rho + a)_{(kh)}}\frac{a_{(n)}(kh)_{(n)}}{(a + \rho + kh)_{(n)}}\frac{1}{n!}.$$

To prove the Markovian property, let $N_v(t) = \sum_{k=1}^{S(t)} X_k$

for a given v. The process $N_v = \{N_v(t), \ t \geq 0\}$ is a compound Poisson process. Hence, it is a Markov process. We now note that:

$$P(N(t+h) = n \,|\, N(t) = m, N(s) = n_s \ for \ 0 \leq s \leq t) =$$

$$= \frac{\int_0^{+\infty} P_v(N(t+h) = n, N(t) = m, N(s) = n_s \ for \ 0 \leq s \leq t)\varphi(v)dv}{\int_0^{+\infty} P_v(N(t) = m, N(s) = n_s \ for \ 0 \leq s \leq t)\varphi(v)dv},$$

where $P_v(A)$ stands for the conditional probability of an event A given the value v of the random variable v. Then,

$$P_v(N(t+h) = n, N(t) = m, N(s) = n(s), \ 0 \leq s \leq t)$$

is equal to $p_h(m - n) \cdot p_{t-s}(m - n_s) \cdot p_s(n(s))$ and $P_v(N(t) = m, N(s) = n(s), \ 0 \leq s \leq t)$ is equal to $p_{t-s}(m - n_s) \cdot p_s(n(s))$. Therefore

$$P(N(t+h) = n, N(t) = m, N(s) = n(s), \ 0 \leq s \leq t)$$

$$= \frac{\Gamma(a+n)}{\Gamma(a+m)}\frac{(kh)_{(n-m)}}{(n-m)!}\frac{(\rho + kt)_{(a+m)}}{(\rho + kt + kh)_{(a+n)}} \qquad (2.4.1)$$

The last result proves the Markovian property of the process and provides its transition probabilities.

III. THE GENERALIZED WARING PROCESS.

The Generalized Waring process can, now, be defined in the following way:

Definition 3.1 The counting process $\{N(t), \ t \geq 0\}$ is said to be a *Generalized Waring Process* with parameters (a, k, ρ), $a > 0$, $k > 0$, $\rho > 0$ if (I) $N(0) = 0$, (II) $N(t)$ is a Markov process, (III) $N(t+h) - N(t)$ is $GW(a, kh, \rho)$-distributed for each $h > 0$, $t \geq 0$.

It has been shown that such a process does exist. Conditions (I), (II) and (III) tell us that this process starts at 0, it has stationary increments and

$$P(N(t) = n) = \frac{\rho_{(kt)}}{(\rho + a)_{(kt)}}\frac{a_{(n)}(kt)_{(n)}}{(a + \rho + kt)_{(n)}}\frac{1}{n!}, \ i.e.$$

$N(t)$ is $GW(a, kt, \rho)$-distributed.

A. The moments and some other properties

Let N be a Generalized Waring process with parameters (a, k, ρ). Then for any t, $E[N(t)] = \frac{akt}{\rho - 1}$,

$$Var[N(t)] = \frac{akt(\rho + kt - 1)(\rho + a - 1)}{(\rho - 1)^2(\rho - 2)}.$$

Following Irwin [6], one may show that the variance can be divided into three additive components, thus

$$Var[N(t)] = \sigma^2_{\Lambda(t)} + (kt)^2\sigma^2_v + \sigma^2_R,$$

where

$$\sigma^2_{\Lambda(t)} = akt(a + 1)(\rho - 1)^{-1}(\rho - 2)^{-1}$$

is the component due to liability

$$\sigma^2_v = a(a + \rho - 1)(\rho - 1)^{-2}(\rho - 2)^{-1}$$

is the component due to proneness

and $\sigma^2_R = akt(\rho - 1)^{-1}$ is the component due to randomness.

The Generalized Waring process is a stationary process. For a stationary process N, $E[N(t)] = \eta \cdot t$, where η is

889

termed the *intensity* of N (see e.g. [3], p.53). It is clear that the intensity of the Generalized Waring process is

$\eta = \dfrac{ak}{\rho - 1}$. For this process (like for all stationary

processes), there always exists, a random variable $\overline{N}$ with

$E(\overline{N}) = \eta$, called the *individual intensity*, such that

$\dfrac{N(t)}{t} \xrightarrow{\ p.\ } \overline{N}$ [1] as $t \to +\infty$ (see e.g. [3], p.53). The

intensity η is finite. Hence, it follows that the individual intensity $\overline{N}$ is finite *a.s.* [2]

Definition 3.1.1. The counting process $\{N(t),\ t \ge 0\}$ is said to be a Negative binomial Process with parameters $(k,\ 1/v)$ $k > 0$, $v > 0$, if

(I) $N(0) = 0$,

(II) $N(t)$ is a Markov process,

(III) $N(t + h) - N(t)$ is $NB(kh,\ 1/v)$-distributed for each $h > 0$, $t \ge 0$.

The first condition together with the condition (III) lead to the conclusion that $N(t)$ is $NB(kt,\ 1/v)$-distributed.

Definition 3.1.2 A Negative Binomial process with parameters $k = 1$ and $v = 1$ is called a *standard Negative Binomial process*.

Definition 3.1.3 Let v be a $BetaII(a,\rho)$-distributed random variable and consider a standard negative binomial process $\widetilde{N}$ independent of it. Let $k > 0$ be a constant. The point process $N = \widetilde{N} \circ (k,\ 1/v)$,

where $\widetilde{N} \circ (kt,\ 1/v) \overset{def}{=} \widetilde{N}(kt,\ 1/v)$ and,

for every t, $\widetilde{N}(kt,\ 1/v) \sim NB(kt,\ 1/v)$, is called the *Generalized Waring Process*.

It is already clear that definition 3.1.3 is equivalent to definition 3.1. By definition 3.1.3 one can prove the following property:

Theorem 3.1.1 Let N be a Generalized Waring process. Then

$\dfrac{1}{t} N(t) \xrightarrow[t \to \infty]{\ p.\ } vk$.

Proof

---

[1] The symbol $p.$ here stands for the convergence in probability of a random variable
[2] The symbol $a.s.$ (almost sure) implies that

$P(\omega \in \Omega | \overline{N}(\omega)$ is finite$) = 1$

$\lim\limits_{t \to \infty} \dfrac{1}{t} N(t) = \lim\limits_{t \to \infty} \dfrac{1}{t} \widetilde{N}\left(kt, \dfrac{1}{v}\right) =$

$vk \lim\limits_{t \to \infty} \dfrac{1}{vkt} \widetilde{N}\left(kt, \dfrac{1}{v}\right) = vk \lim\limits_{t \to \infty} \dfrac{\widetilde{N}(kt, 1/v)}{E[\widetilde{N}(kt, 1/v)]}$

We now use the Chebishev Inequality in order to determine

$\lim\limits_{t \to \infty} \dfrac{\widetilde{N}(kt, 1/v)}{E[\widetilde{N}(kt, 1/v)]}$.

We have $E\left\{\dfrac{\widetilde{N}(kt, 1/v)}{E[\widetilde{N}(kt, 1/v)]}\right\} = 1$ and

$var\left\{\dfrac{\widetilde{N}(kt, 1/v)}{E[\widetilde{N}(kt, 1/v)]}\right\} = \dfrac{var\{\widetilde{N}(kt, 1/v)\}}{E^2[\widetilde{N}(kt, 1/v)]} =$

$\dfrac{vkt(1 + v)}{(vkt)^2} = \dfrac{1 + v}{vkt} \xrightarrow[t \to \infty]{} 0$

Hence,

$P\left\{\left|\dfrac{\widetilde{N}(kt, 1/v)}{E[\widetilde{N}(kt, 1/v)]} - 1\right| \ge \varepsilon\right\} \le \dfrac{1 + v}{\varepsilon vkt} \xrightarrow[t \to \infty]{} 0$

which implies that

$\dfrac{\widetilde{N}(kt, 1/v)}{E[\widetilde{N}(kt, 1/v)]} \xrightarrow[t \to \infty]{\ p.\ } 1$

and

$\dfrac{1}{t} N(t) \xrightarrow[t \to \infty]{\ p.\ } vk.$

Combining this result and the fact that, since v is

$BetaII(a,\rho)$-distributed, $E(v) = \dfrac{a}{\rho - 1}$, we obtain

$E(vk) = \dfrac{ak}{\rho - 1}$. Hence, the random variable $\overline{N} = vk$ is

the individual intensity of the Generalized Waring process.

B. The transition probabilities and the Chapman-Kolmogorov equations of the Generalized Waring process.

Using (2.3.1) and (2.4.1), we obtain for the transition probabilities of the Generalized Waring process

$P_{m,n}(t) = P\big(N(s + t) = n \mid N(s) = m\big)$

$= \dfrac{\Gamma(a + n)}{\Gamma(a + m)} \dfrac{(kh)_{(n-m)}}{(n - m)!} \dfrac{(\rho + kt)_{(a+m)}}{(\rho + kt + kh)_{(a+n)}}$

The transition probabilities of a Markov process satisfy the Chapman-Kolmogorov equations

$$p_{m,n}(s, t) = \sum_{i=m}^{n} p_{m,i}(s, \tau) p_{i,n}(\tau, t) \quad \text{for } s \leq \tau \leq t, \; m \leq n.$$

Then, for the *forward* Kolmogorov differential equations, starting from

$$p_{m,n}(s, t+h) = \sum_{i=m}^{n} p_{m,i}(s, \tau) p_{i,n}(\tau, t+h)$$

for $s \leq \tau \leq t, m \leq n, h \geq 0$, we obtain

$$p\copyright_{m,n}(s,t) = \sum_{i=m}^{n} p_{m,i}(s,t) \lim_{h \to 0} \frac{p_{i,n}(t,t+h)}{h}$$

$$- \lim_{h \to 0}\left(1 - \frac{p_{nn}(t,t+h)}{h}\right) p_{m,n}(s,t),$$

$$\lim_{h \to 0} \frac{p_{i,n}(t,t+h)}{h} =$$

$$= \begin{cases} q_{n-1,n}(t) = \dfrac{k(a+n-1)}{(a+\rho+kt+n-1)} & n-i=1 \\[3mm] q_{i,n}(t)\dfrac{\Gamma(a+n)}{\Gamma(a+i)}\dfrac{k}{(n-i)(n-i-1)}\dfrac{(\rho+kt)_{(a+i)}}{(\rho+kt)_{(a+n)}} & n-i>1 \end{cases}$$

and

$$\lim_{h \to 0}\left(\frac{1 - p_{n,n}(t,t+h)}{h}\right) = v_n(t)$$

$$= k \cdot \sum_{i=0}^{a+n-1} \frac{1}{\rho + kt + i}.$$

Hence, the *forward* Chapman-Kolmogorov equations for the generalized Waring process are:

$$\frac{\partial p_{n,n}(s,t)}{\partial t} = -v_n(t) p_{n,n}(s,t)$$

$$\frac{\partial p_{m,n}(s,t)}{\partial t} = -v_n(t) p_{m,n}(s,t) +$$

$$\sum_{i=m}^{n-1} q_{i,n}(t) p_{m,i}(s,t), \quad m < n.$$

The *backward* equations, follow from the Chapman-Kolmogorov equations with $\tau = s + h$. Then, the *backward* equations for the Generalized Waring process are:

$$\frac{\partial p_{m,m}(s,t)}{\partial t} = v_m(t) p_{m,m}(s,t)$$

$$\frac{\partial p_{m,n}(s,t)}{\partial t} = v_m(t) p_{m,n}(s,t)$$

$$- \sum_{i=m+1}^{n} q_{m,i}(t) p_{i,n}(s,t), \quad m < n,$$

where

$$q_{m,m+1}(s) = \frac{k(a+m)}{(a+\rho+ks+m)},$$

$$q_{m,i}(s) = \frac{\Gamma(a+i)}{\Gamma(a+m)} \frac{k}{(i-m)(i-m-1)} \frac{(\rho+ks)_{(a+m)}}{(\rho+ks)_{(a+i)}} \quad i > m$$

and $v_m(s) = k \cdot \sum_{i=0}^{a+m-1} \frac{1}{\rho+ks+i}$.

IV. THE POISSON AND THE PÓLYA PROCESSES AS PARTICULAR CASES OF THE GENERALIZED WARING PROCESS.

Theorem 4.1. If $k \to \infty$ and $\rho = c \cdot k$ where $c > 0$ is a constant i.e. $w_k(t) = \dfrac{kt}{kt + \rho} = \dfrac{t}{t+c}$, then, the Generalized Waring process tends to the Negative binomial process with probability generating function

$$\left\{ \frac{1}{u_k(t)} - \frac{w_k(t)\theta}{u_k(t)} \right\}^{-a}$$

where $u_k(t) = 1 - w_k(t) = \dfrac{\rho}{kt+\rho} = \dfrac{c}{t+c}$.

The proof is analogous to that used by Irwin [5] to obtain the Pólya distribution as a limiting case of the Generalized Waring distribution. The following theorem can now be shown:

Theorem 4.2 Consider now the Pólya process defined above and let $c \to \infty$ and $a = \lambda \cdot c$, where $\lambda > 0$ is a constant. Then,

$$\lim_{c \to \infty}\left[\binom{a+n-1}{n}\left(\frac{c}{t+c}\right)^a\left(\frac{t}{t+c}\right)^n\right] =$$

$$\frac{(\lambda t)^n \exp(-\lambda t)}{n!}.$$

The result of this theorem tells us that the Generalized Waring Process tends to a homogenous Poisson process with rate λ.

So, the Pólya and the Poisson processes are limiting forms of the Generalized Waring process. Utilizing the results holding for the Generalized Waring process, one may obtain the following results for a Pólya process $X(t)$ with parameters $(a, 1/c)$ and for a Poisson process $Y(t)$ with parameter λ:

- For any $t \geq 0$, $E[X(t)] = \dfrac{a}{c}t$,

891

$$Var[X(t)] = \frac{a}{c}t + \frac{a}{c^2}t^2$$

and $E[Y(t)] = Var[Y(t)] = t\lambda$

- The Pólya and the Poisson processes are both stationary Markov processes. Their respective transition probabilities are:

$$P(X(t+h) \mid X(t) = m) =$$

$$= \begin{cases} \left(\dfrac{c+t}{c+t+h}\right)^{(a+n)} & n = m \\[2ex] (a+m)\dfrac{h(c+t)^{(a+m)}}{(c+t+h)^{(a+n+1)}} & n = m+1 \\[2ex] \dfrac{\Gamma(a+n)}{\Gamma(a+m)(n-m)!}\dfrac{h^{(n-m)}(c+t)^{(a+m)}}{(c+t+h)^{(a+n)}} & n > m+1 \end{cases}$$

and

$$P(Y(t+h) \mid Y(t) = m) =$$

$$= \begin{cases} \exp(-\lambda h) & n = m \\[1ex] \lambda h \exp(-\lambda h) & n = m+1 \\[1ex] \dfrac{(\lambda h)^{(n-m)}}{(n-m)!}\exp(-\lambda h) & n > m+1 \end{cases}$$

- The Pólya Process is a stationary non-homogenous birth process with transition intensities $k_n(t) = \dfrac{a+n}{c+t}$ and the Poisson Process is a stationary homogenous birth process with transition intensities $k_n(t) = \lambda$.

V. SOME ALTERNATIVE GENESIS SCHEMES

The Generalized Waring process has been defined as a non-homogenous stationary Markov Process arising as a Beta mixture of the Negative Binomial process in a "proneness" context. In the sequel, we consider two further genesis schemes where the underlying mechanism is indicative of *contagion* rather than proneness in the sense of Irwin [4] and Xekalaki [10].The *contagion model* assumes that, at time $t = 0$, the individuals have had no accidents and that, during a time period $(t, t + dt]$, the probability of a person having another accident depends on time t and on the number of accidents x sustained by him/her by time t. So this probability is a function $f_\nu(x,t)$, with ν referring to the individual's risk exposure.

Assuming that $f_\nu(x, t) = \dfrac{k+x}{(1/\nu)+t} = \nu \cdot \dfrac{k+x}{1+\nu t}$, the distribution of accidents for each t (λ fixed) is Negative

binomial with parameters $(k, 1/\nu t)$ (the accident pattern is described in that case by a Pólya process). As shown by Xekalaki [10], the overall distribution is the Generalized Waring with parameters (a, k, ρ), when λ varies from individual to individual, according to an exponential distribution, i.e., $\lambda \sim ae^{-a\lambda}$, $a > 0$ for $t = 1$. Adopting a similar approach, one may obtain

$$P_n(t) = P(N = n) = \frac{p_{(\gamma)}}{(a+p)_{(\gamma)}}\frac{a_{(n)}\gamma_{(n)}}{(a+p+\gamma)_{(n)}}\frac{(1/t)^a}{n!}$$

$$\cdot F\left(a+p, a+n, a+p+\gamma+n; 1-\frac{1}{t}\right)$$

(5.1)

where $F(a,b,\gamma;z) = \displaystyle\sum_{m=0}^{\infty}\frac{a_{(m)}b_{(m)}}{\gamma_{(m)}}\frac{z^m}{m!}$.

It can be shown that the counting process $Y = \{Y(t); t > 0; Y(0) = 0\}$, where $Y(t)$, for each t, has the distribution given by (5.1), is a birth process, but not of a Generalized Waring form. It is also difficult to calculate the values of the function

$$(1/t)^a F\left(a+p, a+n, a+p+\gamma+n; 1-\frac{1}{t}\right), \text{ and}$$

the respective probabilities. Assuming that $f_\lambda(x,t) = \lambda(k+mx)$, the distribution of accidents for each t, is Negative binomial with parameters $\left(-\dfrac{k}{m}, \dfrac{1}{1-e^{-\lambda mt}}\right)$, when λ is fixed [4] and

Generalized Waring with parameters $\left(\dfrac{k}{m}, 1, \dfrac{a}{mt}\right)$, when $\lambda \sim ae^{-a\lambda}$, $a > 0$ [8].

Also, following Irwin [4], one may be verify in this case that the distribution of the increment $Y_t(h) = N(t+h) - N(t)$ at time t, given that $N(t) = x$, has a Negative binomial distribution with parameters $\left(-\dfrac{k}{m}+x, \dfrac{1}{1-e^{-\lambda mt}}\right)$ when λ is fixed, and a Generalized Waring distribution with parameters $\left(\dfrac{k}{m}+x, 1, \dfrac{a}{mt}\right)$, when $\lambda \sim ae^{-a\lambda}$, $a > 0$. Hence, in this case,

$$P_{i,j}(s,t) = P(N(t+s) = i \mid N(t) = j) =$$

$$P(N(t+s) - N(t) = i - j \mid N(t) = j)$$

$$= \frac{(a/ms)_{(1)} \quad \left(\dfrac{k}{m}+j\right)_{(i-j)}}{\left(\dfrac{k}{m}+j+\dfrac{a}{ms}\right)_{(1)} \left(\dfrac{k}{m}+j+\dfrac{a}{ms}+1\right)_{(i-j)}}.$$

From the last relationship, one may easily find that

$$p_{2,i}(s,\tau)\cdot p_{j,2}(\tau,t) + p_{3,i}(s,\tau)\cdot p_{j,3}(\tau,t) \neq p_{j,i}(s,t)$$

for some values of a, m, s, t, τ, i, j. This implies that this process does not satisfy the Chapman-Kolmogorov equations and thus is not a Markov Process.

REFERENCES

[1] Chatfield, C. and Theobald, C.M. (1973). Mixtures and Random Sums, *The Statistician*, 22(3), 281-287.

[2] Cresswell, W. L. and Froggatt, P. (1963). *The Causation of Bus Driver Accidents. London.* Oxford University Press.

[3] Grandell, J. (1997). *Mixed Poisson Processes.* Chapman and Hall, London

[4] Irwin, J.O. (1941). Discussion on Chambers and Yule's paper. *Journal of the Royal Statistical Society,* Supplement 7, 101-109.

[5] Irwin, J.I. (1968). "The Generalized Waring Distribution. Applied to Accident Theory". *Journal of the Royal Statistical Society,* A 131, 205-225.

[6] Irwin, J.I. (1975). "The Generalized Waring Distribution", *Journal of the Royal Statistical Society,* A 138, 18-31 (Part I), 204-227 (Part II), 374-384 (Part III)

[7] Newbold, E.M. (1927). Practical Applications of the Statistics of Repeated Events, Particularly to Industrial Accidents. *Journal of the Royal Statistical Society,* 90, 487-547.

[8] Xekalaki, E. (1981). Chance Mechanisms for the Univariate Generalized Waring Distribution and Related Characterizations. *Statistical Distributions in Scientific Work,* 4, 157-172.

[9] Xekalaki, E. (1983a). Infinite Divisibility, Completeness and Regression Properties of the Univariate Generalized Waring Distribution. *Annals of the Institute of Statistical Mathematics,* A 35, 161-171.

[10] Xekalaki, E. (1983b). The Univariate Generalized Waring Distribution in Relation to Accident Theory. Proneness, Spells or contagion?, *Biometrics* 39 (3), 39-47.

[11] Xekalaki, E. (1984). The Bivariate Generalized Waring Distribution and its Application to Accident Theory. *Journal of the Royal Statistical Society,* Series A (General), 47 (3), 448-498.

[12] Xekalaki, E. (1985). Factorial Moment Estimation for the Bivariate Generalized Waring Distribution. *Statistiche Hefte* 26, 115-129.

[13] Xekalaki, E. (1986). The Multivariate Generalized Waring Distribution. *Commun. Statist. Theor. Meth.* 15(3) 1047-1064.

[14] Zografi, M. and Xekalaki, E. (2001) The Generalized Waring Process. Technical Report no 121, Department of Statistics. *Athens University of Economics and Business.*