THE GENERALIZED WARING PROCESS
STATISTICAL INFERENCE AND APPLICATIONS

A Thesis
Presented to
The Academic Faculty

by

Mimoza S. Zografi

In Partial Fulfillment
of the Requirements for the Degree
PHD

Department of Statistics
Athens University for Economy and Business
July 2021

THE GENERALIZED WARING PROCESS
STATISTICAL INFERENCE AND APPLICATIONS

Approved by:

____

Professor Evdokia Xekalaki, Advisor

Date Approved ________________________

ACKNOWLEDGEMENTS

I would like to thank my supervisor Professor Evdokia Xekalaki for her valuable help, her valuable guidance and advice

PREFACE

In this thesis we develop a theory of the generalized Waring Process which is relevant to a wide variety of applications. In particular, we first define the generalized Waring process in real line as a stationary, but non-homogenous Markov process. An application in a web access modelling context has been given and applied to real data. We then construct the generalized Waring Process on a complete separable metric space. The Generalized Waring process in $\mathbb{R}^d$ is defined. By deriving a number of its properties like additivity stationarity, ergodicity and orderliness we demonstrate that the defined process is completely satisfactory for statistical applications.

Table of contents

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

The Poisson process is one of the most important random processes in probability theory. It is widely used to model random "points" in time and space such as the accident pattern in accident theory, the arrival times of customers at a service center, the times of radioactive emissions, the incidence of coal-mining disasters , the positions of flaws in a piece of material, in plant ecology position of plants of a particular species along a line transect taken in a field, etc (see e.g. Cox and Lewis (1966) , Cox and Isham (1980), Snyder and Miller (1991), Ross (1995), Ross (2007) etc.). It is a too simplistic model for real data but it can be successfully used for constructing more flexible models such as Cox Processes and Markov Point processes.

A Cox process is a natural extension of a Poisson process, obtained by considering the intensity function of the Poisson process as a realisation of a random field. The mixed Poisson processes are Cox processes in which the intensity function of the Poisson process is considered as a realisation of a random variable (see e.g. Daley and Vere-Jones (1988), Grandell (1997), Muller and Waagepetersen (2004)). They also have the Markovian property. The processes of Negative Binomial form are important examples of the Mixed Poisson processes.

Both, the Poisson processes and the processes of Negative Binomial form are associated to the Poisson and the Negative Binomial distributions which have a wide spectrum of applications in areas such as accident statistics, income analysis, environmental statistics, etc. The univariate generalized Waring distribution is also a discrete distribution which has been used as a model that better describes such practical situations as opposed to the Poisson distribution or the Negative Binomial distribution. It was derived by Irwin (see Irwin (1968), Irwin (1975)) as the distribution of the accidents of an 'accident prone' population exposed to variable risk. For certain values of the parameters, the $UGWD(a, k; \rho)$

can be very long - tailed and so it was shown (Irwin (1963), Irwin (1968)) to be a suitable theoretical form for the description of biological distributions. An enormous number of problems in areas such as biology, economics, accident theory, linguistics, reliability and bibliographical analysis have been linked with the generalized Waring distribution (see e.g. Xekalaki (1981), Xekalaki (1983a), Xekalaki (1983c), Xekalaki and Panaretos (1983)).

In her paper, Xekalaki (1981) presents a number of results concerning the genesis schemata that give rise to the UGWD and suggests some new ones. Some important results are the mixed models. In these models the UGWD has been obtained as a mixture of Negative Binomial, Poisson and generalised Poisson distributions. The Negative Binomial and generalised Poisson distributions, can be derived as mixtures of the Poisson distribution respectively with gamma and logarithmic series distributions. Two further derivations of the UGWD in the context of accidents, which are based on a 'contagion' hypothesis and a 'spells' hypothesis, respectively assuming that individuals are exposed to varying envorinmental risk have been considered by Xekalaki (1983b). She demonstrates there, that, while the UGWD is a plausible model, if accident proneness is accepted as an established fact, a satisfactory fit of this model is not be regarded as evidence for the validity of the proneness hypothesis.

The UGWD was shown by Irwin (1968) and Irwin (1975) to provide useful accident model which enables one to split the variance into three additive components due to randomness, proneness and liability. The two non-random variance components, however, can not be separately estimated. Xekalaki (1984) suggests a way to overcome this problem by defining a bivariate extension of the generalized Waring distribution. A multivariate version of the generalized Waring distribution has also been defined by Xekalaki (1986). The structure of this multivariate distribution was studied and shown among other results that it allows for the marginal distributions and their convolution to be UGW distributions.

The aim of this thesis is to define the generalized Waring process as a process associated to the generalized Waring distribution intending to have a model that better describes some practical situations mentioned above as opposed to the Poisson and Negative Binomial processes.

As mentioned earlier, 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 (Newbold (1927)). Both of these processes satisfy the Markovian property as 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 what follows, we develop a theory of the generalized Waring Process which is relevant to a wide variety of applications. We start by defining the generalized Waring process on the real line as a stationary, but non-homogenous Markov process and then we construct the generalized Waring Process on a complete separable metric space.

Most of the results have been obtained in the context of models that have been used for the description of accident data but can be adjusted so as to be fit for other practical frameworks with appropriate modifications of concepts and terminology.

The thesis is organized in six chapters. In particular, in chapter 2, the generalized Waring process is defined and studied first in an accident theory context. The starting point is a process of negative binomial form, but different from a Pólya process. This process is then mixed with a beta distribution of the second type (beta II). Further, in section 2.5, an alternative genesis scheme referring to Cresswell and Froggatt's (1963) spells model is proposed in the framework considered by Xekalaki (1983b). Moreover, it is demonstrated how the above considerations formulate the framework for the definition of the generalized Waring process as a stationary, but non-homogenous Markov process. Some inferential aspects connected with the mixed negative binomial derivation of the generalized Waring processare are also discussed in section 2.6. Further, an application in a web access modeling context is provided and discussed in section 2.4. The results stem from Xekalaki and Zografi

(2008) 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, psychometry and web access modeling as the case is with the application discussed in section 2.4.

Focus is subsequently turned to the problem of modeling spatial data in cases these are prone to exhibit overdispersion, however, it may be challenging to specify a point process model that simultaneously features additivity, stationarity, ergodicity, and orderliness.

An early process with such properties (stationarity, ergodicity, and orderliness) was introduced by Neyman & Scott (1958) as a statistical approach to problems of cosmology. It is a stochastic process of clustering of the second order and, in particular, a special case of a Poisson cluster process with daughter clusters assumed to be Poisson. However, constructing point processes with finite dimensional laws of the negative binomial form has become a very popular modeling strategy because of its tractability, elegant closed form and interpretability of its parameters in applied contexts, where factors other than pure chance play a role in the happening of an event. Such processes are known as negative binomial processes, and have been defined and studied on general state spaces (Gregoire (1983)). Owing to their combination of flexibility and mathematical tractability, they have been employed in many practical situations (see for example Bates (1955) , Boswell & Patil (1977) , Cliff & Ord (1973), Ramakrishnan (1951) etc.). However, they have been shown to fail in simultaneously accommodating the three properties listed above. As a matter of fact, it has been conjectured by Diggle & Milne (1983), that additive, stationary, ergodic, orderly spatial point processes with negative binomial finite-dimensional distributions may not even exist. In their words, it would seem that one is "unable to exhibit a negative binomial point process that is statistically interesting according to the criteria we laid down" [these criteria being additivity, stationarity, ergodicity, orderliness].

For this purpose, in chapter 3, using the GWD as a building block, we construct an additive, stationary, ergodic, and orderly spatial point process, and study its basic properties. We develop our results on a general separable metric state space, before focussing on the practically relevant case of $\mathbb{R}^d$ in chapter 4. The process is seen to satisfy several useful closure properties (under projection, marginalization, and superposition) and to be easy to simulate. We further show that, in the limit as certain parameters of the process diverge, the Generalized Waring Point Process approximates a negative binomial process. In doing so, we give an approximate positive solution to the task set out by Diggle and Milne (1983) by demonstrating that a spatial non- negative binomial point process that is simultaneously a stationary, ergodic and orderly spatial point process does exist and has one dimensional distributions that can take a negative binomial form depending on parameter choice.

In section 3.1, we provide some background highlighting various frameworks giving rise to the univariate and multivariate cases of the generalized Waring distributions in the univariate and multivariate cases, the moments and some of their properties such as the finite and countable additivity that will be utilized in subsequent sections.

Then, the definition of the generalized Waring process in a complete separable metric space is given in section 3.2 . It is shown that the finite dimensional distributions of the process defined are of the multivariate generalized Waring form and that the process defined fulfills the Kolmogorov consistency conditions for the finite dimensional distributions and the measure requirements given by the basic existence Theorem of a point process (see Daley and Vere-Jones (1988)). The generalized Waring process in $\mathbb{R}^d$ with the Lebesgue measure as parameter measure $\mu(\cdot)$ is then defined in section 4.1 and proved to be orderly, stationary, ergodic and nth-order stationary.

Some limiting forms of the generalized Waring process are also obtained for various limiting values of its parameters. In particular, it is shown that the generalized Waring process can take the form of a negative binomial and a Poisson process.

Furthermore, the multivariate generalized Waring process is defined as a special case of the generalized Waring process on the product space $S \times \{1, 2, ...m\}$.

In chapter 5, the generalized Waring process on the Real Line is defined either as a special case of the generalized Waring process in $\mathbb{R}^d$ or as a projection of the generalized Waring process in $\mathbb{R}^2$ and is shown to be orderly (and a simple point process as well), stationary, ergodic and nth-order stationary. The generalized Waring process on the postive half-line

$\mathbb{R}^+$ is also examined and proved that it posesses the Markovian property.

CHAPTER 2

THE GENERALIZED WARING PROCESS ON THE REAL LINE AND ITS APPLICATIONS

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 (Newbold, 1927). Both of these processes satisfy the Markovian property as 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 chapter, a new process is defined and studied (see Xekalaki and Zografi (2008)). 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. Irwin, 1975; Xekalaki, 1983b). Analogously to the case of Poisson and Pólya process, this new process, termed in the sequel as the generalized Waring process, is postulated to be a Markov process, as shown in section 2.3. The starting point is a process of negative binomial form, but different from a Pólya process. This process is then mixed with a beta distribution of the second type (beta II). Further, an alternative genesis scheme referring to Cresswell and Froggatt's (1963) spells model is proposed in the framework considered by Xekalaki (1983b). Section 2.3 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 2.3 and its transition probabilities are derived as well. An application in a web access modeling context is provided in section 2.4. Two further genesis schemes considered by Zografi and Xekalaki (2001) are presented in section 2.5. Finally, some inferential aspects connected with the mixed negative binomial derivation of the generalized Waring process are discussed in section 2.6. The results obtained are in the context of models that have widely been considered for the interpretation of accident data. However, the concepts and the 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, psychometry and web access modeling as the case is with the application discussed in section 2.4.

2.1 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 (1927). This model considers several individuals exposed to the same external risk (e.g. drivers all driving about the same distance within a similar traffic environment) and that there are intrinsic differences among different individuals (e.g. differences in accident 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 $\Lambda$with a gamma distribution with parameters k and ν, the number of accidents $N(t)$ at time t, $t = 0,\ 1,\ 2,$ ... defines the Pólya process with parameters k and ν as follows:

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

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

(III)$N(t+h) - N(t)$ has a distribution defined by the probability function

$$P\left\{N(t+h) - N(t) = m\right\} =$$
$$E\left[\frac{(\Lambda h)^n}{n!}e^{-\Lambda h}\right]\binom{k+m-1}{m}\left(\frac{1}{1+\nu h}\right)^k\left(\frac{\nu h}{1+\nu h}\right)^m, m = 0, 1, ... \tag{1}$$

where Λ is a random variable with density u given by

$$u(l) = \frac{\nu^{-k}}{\Gamma(k)} l^{k-1} e^{-(l\nu^{-1})}, l > 0$$

It is clear that $N(t)$ has a negative binomial distribution with parameters k and $\dfrac{1}{1+\nu t}$, i.e. $N(t) \sim NB\left(k, \dfrac{1}{1+\nu t}\right)$.

The distribution of the random variable Λ explains here the variation of the accident proneness from individual to individual. As noted by Irwin (1968) and Xekalaki (1984), the term accident proneness here refers to both, the external and the 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 $\Gamma(k(h), 1/\nu(h))$, where $k(h)$ and $\nu(h)$ are in general some functions of h, then, clearly, the number of accidents $N(t)$ forms 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:

$$P\{N(t+h) - N(t) = n\} = \int_0^{+\infty} \frac{(\lambda h)^n}{n!} e^{-\lambda h} \frac{\nu(h)^{-k(h)}}{\Gamma(k(h))} \lambda^{(k(h)-1)} e^{(-\lambda/\nu(h))} d\lambda, \qquad (2)$$

$$n = 0, 1, ...$$

It can be shown that

$$P\{N(t+h) - N(t) = n\} = \binom{kh+n-1}{n} \left(\frac{1}{1+v(h)h}\right)^{k(h)} \left(\frac{v(h)h}{1+v(h)h}\right)^n .$$

Then, using the first assumption, it follows that for any t, $N(t)$ has a negative binomial distribution with parameters $k(t)$ and $\frac{1}{1+\nu(t)t}$. Hence, one can verify that

$$P\{N(t) = n\} = \int_0^{+\infty} \frac{(\lambda t)^n}{n!} e^{-\lambda t} \frac{\nu(t)^{-k(t)}}{\Gamma(k(t))} \lambda^{(k(t)-1)} e^{-\lambda/\nu(t)} d\lambda,$$

$$n = 0, 1, ...$$

This tells us precisely that $N(t)$ is a Cox Process (see e.g. Grandell (1997), 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 $(k(h),\ \nu(h))$with

$k(h) \cdot \nu(h)$ =constant. So, letting $\nu(h) = \nu/h$, and $k(h) = kh$, i.e., allowing $\Lambda(h)$having a gamma distribution that changes with time so that its expectation remains constantly equal to νk, we obtain

$$P\{N(t+h) - N(t) = n\} = \binom{kh+n-1}{n} \left(\frac{1}{1+v}\right)^{kh} \left(\frac{v}{1+v}\right)^{n}, \tag{3}$$

$$n = 0, 1, ...$$

and that $N(t)$ is $NB\left(kt, \frac{1}{1+\nu}\right)$-distributed.

2.2 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 (1968) and Irwin (1975), used the term "accident proneness" ν to refer to a person's predisposition to accidents, and the term "accident liability" ($\lambda|\ \nu$, i.e. λ for given ν) to refer to a person's exposure to external risk of accident.

The conditional distribution of the random variable Λ given ν 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 $\Gamma(kh, 1/\nu h)$ distribution for $\Lambda|\nu$. Moreover, assuming independence in two non-overlapping time periods, the number of accidents $N(t)$ given ν will be a stochastic process of a negative binomial form with parameters kt and $\dfrac{1}{1+\nu}$. This starts at 0 and has stationary increments with a distribution given by 3. Let us further allow the parameter ν of the negative binomial to follow a beta distribution of the second kind with parameters a and ρ, i.e. ν is a random variable with density $\varphi(\nu)$ given by

$$\varphi(\nu) = \frac{\Gamma(a+\rho)}{\Gamma(a)\Gamma(\rho)} v^{(a-1)} (1+v)^{-(a+\rho)} \qquad a, \rho \geq 0$$

obtaining thus for the distribution of the number of accidents $N(t)$:

$$N\{(t+h) - N(t) = n\} = \frac{\rho_{(kh)}}{(a+\rho)_{(kh)}} \frac{a_{(n)} (kh)_{(n)}}{(a+\rho+kh)_{(n)}} \frac{1}{n!}$$

and

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

$$n=0,\ 1,\ ...$$

It is clear that $N\left(t\right)$ is a generalized Waring process on positive half-line $\mathbb{R}^{+}$.

Remark 1 *If we consider individuals of proneness ν and liability $\lambda_i|\nu\ i=1,2$ respectively in each of two non-overlapping intervals of time, it follows by the model's assumptions that the numbers $N_1,\ N_2$ of accidents incurred by these individuals are jointly distributed according to a double Poisson distribution with parameter $(\lambda_1|\nu,\ \lambda_2|\nu)$. Then, for individuals with the same proneness but varying liabilities, the joint distribution of accidents over the two intervals, is the double negative binomial with parameters $\left(\left(kh_{1,}\ \frac{1}{1+\nu}\right);\left(kh_{2,}\ \frac{1}{1+\nu}\right)\right)$, where $h_1,\ h_2$ are the respective sizes of these intervals. If, further, the proneness parameter ν is allowed to follow a beta distribution of the second kind with parameters a and ρ, the joint distribution of the numbers of accidents over the two intervals is a bivariate generalized Waring distribution with parameter $\left((a,kh_{1,}\rho);(a,kh_{2,}\rho)\right)$(see Xekalaki (1984)). Now, it is clear that, if a number of non-overlapping intervals greater than two is considered, the joint distribution of the numbers of accidents over those intervals, will follow a multivariate generalized Waring distribution (see Xekalaki (1986)).*

In the sequel, we use the above remark (Remark 1) to show that the generalized Waring process resulting from the above generating scheme is a Markov process, i.e. that

$$P\left(N\left(t+h\right)=n\ \mid N\left(t\right)=m,\ N\left(s\right)=n_s\ ,\ 0\leq s<t\right)$$

coincides with

$$P\left(N\left(t+h\right)=n\mid N\left(t\right)=m\right)$$

for every non-negative integer n,m,n_s $0\leq s<t$.

For a proof of this, observe that

$$P\left\{N\left(t+h\right)=n\mid N\left(t\right)=m,N\left(s\right)=n_s,\ 0\leq s<t\right\}=$$

$$P\left\{N\left(t+h\right)-N\left(t\right)=n-m\mid N\left(t\right)-N\left(s\right)=m-n,\quad N\left(s\right)-N\left(0\right)=n_s,0\leq s<t\right\}$$

and consider the random vector

$$\left(N\left(t+h\right)-N\left(t\right),N\left(t\right)-N\left(s\right),N\left(s\right)-N\left(0\right)\right),\ 0\leq s<t.$$

It follows from Remark 1 that this vector has a trivariate generalized Waring distribution with parameters α, $\underline{k}$, and ρ, where $\underline{k}=\left(kh,k\left(t-s\right),ks\right)$. This is a three dimensional special case of Xekalaki's (1986) multivariate generalized Waring distribution whose structural properties imply that the random vector

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

has a univariate generalized Waring distribution with parameters $\alpha+n\left(t\right)$, kh and $\rho+kt$, where $n\left(t\right)$ is the value of $N\left(t\right)$. Hence,

$$P\left\{N\left(t+h\right)-N\left(t\right)=n-m\mid N\left(t\right)-N\left(s\right)=m-n_s,N\left(s\right)-N\left(0\right)=n_s\right\}$$

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

$$=\frac{\Gamma\left(a+n\right)}{\Gamma\left(a+m\right)}\frac{\left(kh\right)_{(n-m)}}{(n-m)!}\frac{\left(\rho+kt\right)_{(a+m)}}{\left(\rho+kt+kh\right)_{(a+n)}} \tag{4}$$

$$=P\left\{N\left(t+h\right)-N\left(t\right)=n-m\mid N\left(t\right)-N\left(0\right)=m\right\}$$

$$=P\ \left[N\left(t+h\right)=n\mid N\left(t\right)=m\right]$$

which proves that the generalized Waring process has the Markovian property, i.e. the conditional distribution of the future state $N\left(t+h\right)$ given the present state $N\left(t\right)$ and the past state $N\left(s\right)$, $0\leq s\leq t$, depends only on the present state.

2.3 The Spells Model

In the sequel, an alternative scheme generating a process of a generalized Waring form, on positive half-line $\mathbb{R}^{+}$, is considered. This is a variant of Cresswell and Froggatt's (1963) spells model that has been considered in the paper of Xekalaki (1984). 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 $S(t)$, $t = 0, 1, 2, ...$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, ν), i.e.

$$P(X_i = 0) = 1 - m\log(1 + v)$$

and

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

the random variable $X(t)$, is a negative binomial random variable with parameters $\left(kt, \frac{1}{1+\nu}\right)$ for each t (Chatfield and Theobald (1973)). Here ν is regarded as the external risk parameter, too. Then, if the differences in the external risk can be described by a *beta* (a, ρ) distribution of the second kind, the resulting accident distribution is of a generalized Waring form with parameters a, kt, and ρ.

Let us consider, now, the counting process $\{N(t), t \geq 0\}$ with $N(t)$ represented, for $t \geq 0$, 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, ν) and is independent of the process $S(t)$, and ν is a non-negative random variable with a *Beta* (a, ρ) distribution of the second kind.

Theorem 1 *For the process* $\{N(t), \; t \geq 0\}$ *defined as above the following conditions hold: (I)* $N(0) = 0$ *(II)* $\{N(t), \; t \geq 0\}$ *possesses stationary increments (III)* $\{N(t), \; t \geq 0\}$ *is a Markov process.*

The proof of (I) is straightforward.

To prove condition (II), denote by φ the probability distribution function (p.d.f.) of the random variable ν. 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 $N_\nu(t)=\sum\limits_{k=1}^{S(t)} X_k$ for a given ν. The process $N_\nu=\{N_\nu(t),\ t\geq 0\}$ is a compound Poisson process. Hence, it is a Markov process.

We now note that:

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

$$=\frac{\int_0^{+\infty} P_\nu(N(t+h)=n,N(t)=m,N(s)=n_s\quad for\quad 0\leq s\leq t)\,\varphi(\nu)\,d\nu}{\int_0^{+\infty} P_\nu(N(t)=m,N(s)=n_s\quad for\quad 0\leq s\leq t)\,\varphi(\nu)\,d\nu},$$

where $P_\nu(A)$ stands for the conditional probability of an event A given the value ν of the random variable ν.

Then, $P_\nu(N(t+h)=n,\ N(t)=m,\ N(s)=n(s),\ 0\leq s\leq t)$ is equal to $p_h(m-n)\cdot p_{t-s}(m-n_s)\cdot p_s(n(s))$ and $P_\nu(\ N(t)=m,\ \ N(s)=n(s),\ \ 0\leq s\leq t\ )$ is equal to $p_{t-s}(m-n_s)\cdot p_s(n(s))$.

Therefore,

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

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

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

2.4 *An Application in the Context of Modeling Web Access Patterns*

As mentioned above, the concepts and terminology used in this thesis can easily be modified so that the obtained results can be implemented in several other fields. As an example, we present here an application of the generalized Waring process on positive half-line $\mathbb{R}^+$ in the context of modeling web access patterns. The results are an adaptation of those obtained by Xekalaki and Zografi (2008) in the context of models that have widely been used for the interpretation of accident occurrence with appropriate modifications of concepts and terminology.

Consider in particular, modeling the whole counting process $\{N(s), s > 0\}$ associated with the access pattern of a web site, where, for any $t > 0$, the variable $N(t)$ denotes the number of visits that the web pages on this particular site get within the interval $(0, t)$. (Note that the generalized Waring distribution has been cited in Ajiferuke, Wolfram and Xie (2004) as used by them to fit an observed website visitation frequency distribution for a given period, i.e, to model counts $N(t_0)$ of web visits on a given fixed time interval $(0, t_0)$)

Except for chance, visits to a web site can be regarded as affected by the intrinsic appeal of the particular site to web users (corresponding to proneness) as well as by exogenous factors (corresponding to external factors) such as, links provided by other sites to the particular site, how well the site is advertised etc.

Let us denote by ν the intrinsic factors and by $\lambda|\nu$the exogenous factors. Assume that $N(t)|\lambda$ follows a $Poisson(\Lambda(t))$ distribution, where $\Lambda(t) = \lambda t$ with $\lambda|\nu$ following a $Gamma\left(kt, \frac{1}{\nu t}\right)$ distribution.

Then, the conditional distribution of $N(t)|\nu$ is a $NB\left(kt, \frac{\nu}{1+\nu}\right)$ distribution with ν following a $Beta(\alpha, \rho)$ distribution of the second kind, while the unconditional distribution of $N(t)$ is the $GWD(a, kt; \rho)$ distribution, i.e.$\{N(t),\ t \geq 0\}$is a generalized Waring process

Table 1
Visits made by a given IP address to an e-shop site per date and time

Date	Day	Hour	Minute	Second	Visit duration (in days)
12/04/206:16:15:27	12	16	15	27	0
13/04/206:01:30:57	13	1	30	57	0.385763889
13/04/206:09:38:04	13	9	38	4	0.724039352
13/04/206:14:44:41	13	14	44	41	0.936967593
13/04/206:20:39:53	13	20	39	53	1.183634259
15/04/206:21:28:53	15	21	28	53	3.217662037
16/04/206:11:59:50	16	11	59	50	3.822388426
16/04/206:19:27:24	16	19	27	24	4.133298611
17/04/206:02:13:47	17	2	13	47	4.415509259
18/04/206:17:41:12	18	17	41	12	6.059548611
24/04/206:06:00:26	24	6	0	26	11.57290509
24/04/206:12:36:52	24	12	36	52	11.64959491
24/04/206:18:27:59	24	18	27	59	12.09203704
25/04/206:00:17:51	25	0	17	51	12.335
25/04/206:06:35:20	25	6	35	20	12.5971412
26/04/206:21:05:30	26	21	5	30	14.20142361
29/04/206:09:09:02	29	9	9	2	16.70387731
29/04/206:09:17:15	29	9	17	15	16.70958333
29/04/206:10:33:00	29	10	33	0	16.7621875

on $\mathbb{R}^+$.

The log files representing the hits on an e-shop site for the period 31/03/2006−30/04/2006, have been used to fit this model. A log file typically contains information on the times of visits per IP address per day. On the basis of such log files, the visits per day made by each of 468 IP addresses to the particular site have been enumerated for the above-mentioned one-month period yielding the corresponding observed paths

$$\{N_i(t_j)\,,\ i = 1,\ 2,\ ...,\ 468,\ j = 1,\ 2,\ ...,\ 31\}$$

of the numbers of visits $N_i(t_j)$ made by IP address i up to and including time t_j. A sample of one thus obtained path corresponding to one of the IP addresses considered is presented in Table 1.

The observed paths were compared to the corresponding time series of simulated realizations of the generalized Waring process over the same time segment.

Estimates of the parameters of the generalized Waring process have been obtained employing the centered reduced moment estimation procedure for spatial point process data (see, e.g., Ripley (1988), Daley and Vere-Jones(1988), and Chetwynd and Diggle (1998) among others) which do exist for the generalized Waring process. This procedure utilizes the moment estimators

$$E\left(N\left(s\right)\right) = \hat{\mu}_1 = \hat{\eta} \cdot s = \frac{n \cdot s}{h},$$

$$E\left(\overset{\wedge}{N^2}\left(s\right)\right) = \hat{\mu}_2 = \frac{X}{n_{(2)}}$$

$$E\left(\overset{\wedge}{N^3}\left(s\right)\right) = \hat{\mu}_3 = \frac{Z - X}{n_{(3)}}$$

with

$$X = \sum_{i=1}^{n} \sum_{i \neq j} \phi_s^2\left(x_i, x_j\right),$$

$$Z = \sum_{i=1}^{n} \left(\sum_{j \neq i} \phi_s\left(x_i, x_j\right) \right) \left(\sum_{k \neq i} \phi_s\left(x_i, x_k\right) \right),$$

where the quantities involved in the above equations represent weights defined, for each value x_i in the collection of points $\{x_i : i = 1, 2, ..., n\}$ of the process within a time interval of length h, defined as follows:

For each x_i in $\{x_i : i = 1, 2, ..., n\}$ and a given $s > 0$, consider the interval of center x_i and length s and assign to every point $x_j, j \neq i$ in this interval the weight $\phi_s\left(x_i, x_j\right) = \omega\left(x_i, x_j\right)^{-1}$, where $\omega\left(x_i, x_j\right)$ is the number of other points $\{x_k, k \neq i, k \neq j\}$ of the process that are included in the interval of length $|x_i - x_j|$ and center x_i. Within the setting of our example, the set $\{x_i : i = 1, 2, ..., n\}$ represents, for each IP address, the visits made by the particular IP address for the entire duration of the period of time $h = x_n - x_1$ considered, n = 31, $\hat{\eta}$ denotes an estimator of the process intensity η, i.e. of the expected number of visits in an interval of unit length, while the value set for the constants s was $s = 0, 5$.

Using The Shedler – Lewis thinning technique for a point process with bounded conditional intensity, described in the end of this section, for each of the IP addresses, one hundred simulated realizations of the generalized Waring process with the above estimated parameter values were obtained by using and each of the observed time series paths was

Table 2
Centered reduced moment estimates for the parameters of the $UGWD(a, kt; \rho)$

IP address	$\overset{\wedge}{\alpha}$	$\hat{k}$	$\overset{\wedge}{\rho}$
1	5.6888256	0.594154	3.86463
2	4.139105	0.929841	3.521098
3	3.8695139	0.8293397	4.2061137

compared to the corresponding simulated ones. The comparison showed that, on average, the realizations of a generalized Waring process with the obtained parameter values notably 'resembled' the observed paths of the observed time series, in the sense that they had recognizable similar structural characteristics.

For illustration purposes, the paths of the observed time series associated with a sample of three of the IP addresses are presented (Figures 1, 2). Each of these paths is superimposed by a sample of three of the 100 corresponding simulated realizations of the generalized Waring process with parameter estimates obtained as above and given in Table 2.

Inspection of the graphs depicted by Figures 1, 2 provides a visual appreciation of the degree of similarity in the structural characteristics of the paths of the observed and the realized time series.

2.4.1 Simulation Methods

2.4.1.1 ***Conditionally Bounded Property***

The following Lemma proves that the conditional intensity function of a generalized Waring process in $\mathbb{R}^+$ is conditionally bounded and suggests a limit function for the conditional intensity which is used in the *The Shedler – Lewis* algorithm for simulating the points of generalized Waring process in $\mathbb{R}^+$.

Lemma 1 *The conditional intensity function of a generalized Waring process in* $\mathbb{R}^+$ *is conditionally bounded.*

Proof According to 29 the conditional intensity function is

$$\lambda^*(t) = \begin{cases} k\left[\Psi\left(\rho + kt + a\right) - \Psi\left(\rho + kt\right)\right] & (0 < t \leq t_1) \\ k\left[\Psi\left(\rho + k\left(t + t_{n-1}\right) + a + n - 1\right) - \Psi\left(\rho + k\left(t + t_{n-1}\right)\right)\right] & (t_{n-1} < t \leq t_n,\ n \geq 2) \end{cases}$$

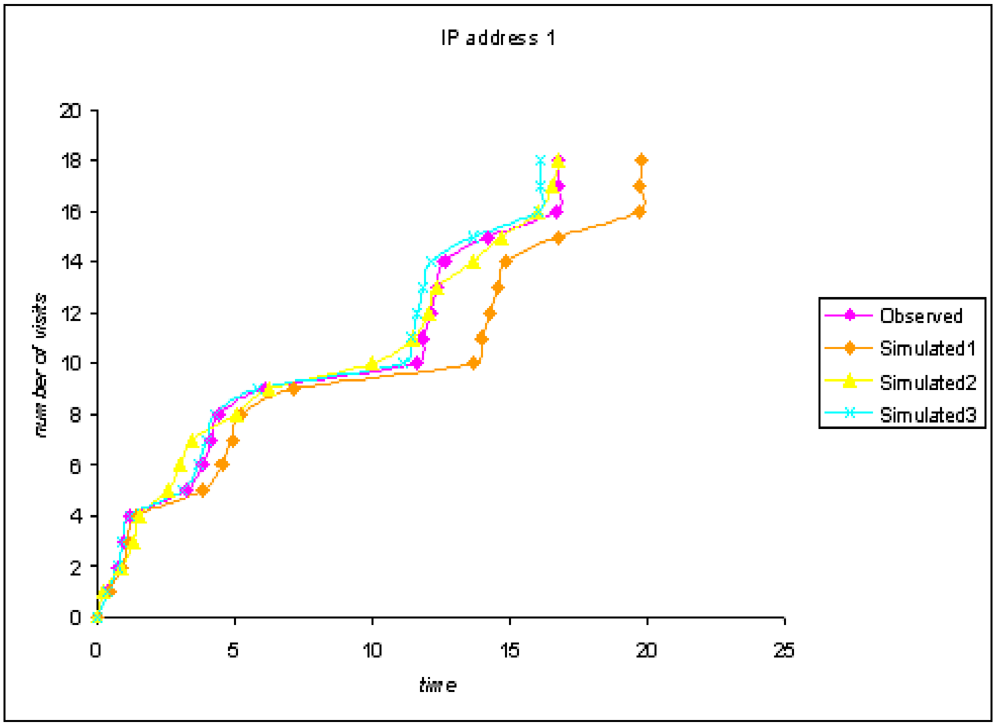

Figure 1
Observed and simulated paths corresponding to IP address 1

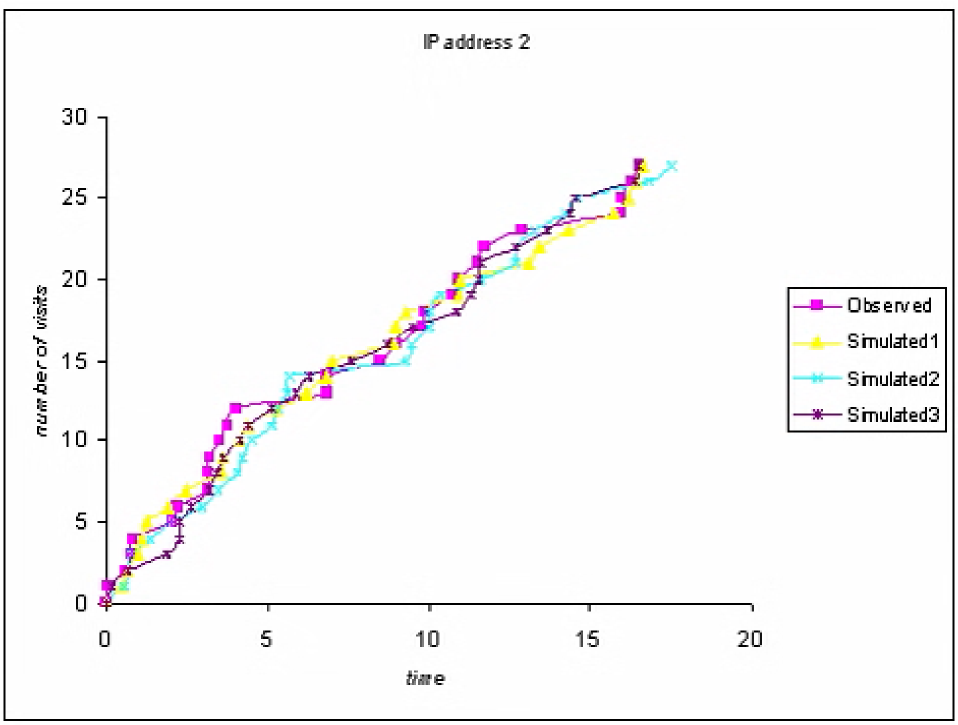

Figure 2
Observed and simulated paths corresponding to IP address 2

Knowing that $\forall x > \frac{1}{2}$, $\log\left(x - \frac{1}{2}\right) \leq \Psi(x) \leq \log(x) - \dfrac{1}{2x}$ (see Janardan, K. G. & Patil, G. P. (1972)) we obtain for $t > \dfrac{1-\rho}{2k}$

$$\Psi(a+\rho+kt+c) \leq \log(a+\rho+kt+c) - \frac{1}{2(a+\rho+kt+c)}$$

$$-\Psi(\rho+kt) \leq -\log\left(\rho+kt-\frac{1}{2}\right)$$

hence if we denote $\varphi_{a,\rho,c}(t) = \log(a+\rho+kt+c) - \log\left(\rho+kt-\frac{1}{2}\right) - \dfrac{1}{2(a+\rho+kt+c)}$

we obtain $\Psi(a+\rho+kt+c) - \Psi(\rho+kt) \leq \varphi_{a,\rho,c}(t)$. Let us consider

$$\begin{aligned}
\frac{d\varphi_{a,\rho,c}(t)}{dt} &= \frac{k}{a+\rho+kt+c} - \frac{k}{\left(\rho+kt-\frac{1}{2}\right)} + \frac{2k}{4(a+\rho+kt+c)^2} \\
&= \frac{-k\left(4kt(a+c+2) + 4\rho(a+c) + 4(a+c+1)^2 + 6\left(a+c+\frac{1}{2}\right)\right)}{4(a+\rho+kt+c)^2\left(\rho+kt-\frac{1}{2}\right)}
\end{aligned}$$

Clearly, $\dfrac{d\varphi_{a,\rho,c}(t)}{dt} < 0$, which means that the function $\varphi_{a,\rho,c}(t)$ is monotonically decreasing. Hence we can write $\Psi(a+\rho+k(t+u)+c) - \Psi(\rho+k(t+u)) < \varphi_{a,\rho,c}(t+u) < \varphi_{a,\rho,c}(t)$ (all $u > 0$), which proves the lemma.

2.4.1.2 ***Simulation of a point process with bounded conditional intensity***

The Shedler – Lewis thinning technique

This technique can be carried over to the point process context when the conditional intensity λ^* is known explicitly as a function of past variables. The thinning technique is particularly useful when λ^* is conditionally bounded, by which we mean that for every $n = 1, 2, ...$ and all sequences with the hazard function satisfies $h_n(t+u|t_1, ..., t_{n-1}) \leq M^*(t)$ (all $u > 0$) for some $M^*(t) = M^*(t; t_1, ..., t_{n-1}) < \infty$. In case of the generalized Waring process, as $M^*(t)$ can be used the function

$$\begin{aligned}
M^*(t) &= k\varphi_{a,\rho,kt_{n-1}+a+n-1}(t) \\
&= k\left[\log(a+\rho+kt+i-1) - \log\left(\rho+kt-\frac{1}{2}\right) - \frac{1}{2(a+\rho+kt+i-1)}\right]
\end{aligned}$$

according to the above Lemma.

Algorithm 1 *The algorithm for simulating the points* $\{t_i : i = 1, 2, ...\}$ *of such a process on* $(0, \infty)$ *is specified as follows:*

(1) *set* $t = 0, i = 1$;

(2) *calculate* $M^*(t)$;

(3) *generate an exponential r.v.* T *with mean* $\frac{1}{M^*(t)}$ *and a r.v.* U *uniformly distributed on* $(0, 1)$;

(4) *if* $\dfrac{\lambda^*(t+T)}{M^*(t)} > 0$, *replace* t *by* $t + T$ *and return to step* (2);
while otherwise

(5) *set* $t_i = t + T$, *advance* i *by* 1, *replace* t *by* t_i , *and return to step* (2).

2.5 Some Alternative Genesis Schemes

The generalized Waring process on positive half-line $\mathbb{R}^+$ 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 this section, we consider two further genesis schemes where the underlying mechanism is indicative of contagion rather than proneness in the sense of Irwin (1941) and Xekalaki (1983b) .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 $f_\lambda(x, t)$, with λ referring to the individual's risk exposure.

The two different schemes assume different form for this function. in the first case the resulting process is a birth process, but not of a generalized Waring form, while in the second case, the increments of the process have a generalized Waring distribution but it is not a Markov process.

1. Assuming that $f_\lambda(x,\ t) = \dfrac{k + x}{(1/\lambda) + t} = \lambda \cdot \dfrac{k + x}{1 + \lambda t}$, the distribution of accidents for each t (λ fixed) is negative binomial with parameters $\left(k,\ \frac{1}{1+\lambda t}\right)$ (the accident pattern is described in that case by a Pólya process). As shown by Xekalaki (1983b), 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., $\lambda \sim ae^{-a\lambda}, \quad a > 0$ for $t = 1$.

Adopting a similar approach, one may obtain

$$P_n(t) = P(N = n) =$$

$$= \frac{p_{(\gamma)}}{(a+p)_{(\gamma)}} \frac{a_{(n)}\gamma_{(n)}}{(a+p+\gamma)_{(n)}} \frac{(1/t)^a}{n!} \; {}_1F_2\left(a+p,\ a+n,\ a+p+\gamma+n;\ 1-\frac{1}{t}\right) \qquad (5)$$

$${}_1F_2(a,b,\gamma;z) = \sum_{m=0}^{\infty} \frac{a_{(m)}b_{(m)}}{\gamma_{(m)}} \frac{z^m}{m!}.$$

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, is a birth process, but not of a generalized Waring form. Also, calculating the values of the function $\frac{1}{t^a}{}_1F_2(\ a+p,\ a+n,\ a+p+\gamma+n;\ 1-\frac{1}{t})$ and the respective probabilities is quite involved.

2. Assuming that $f_\lambda(x,t) = \lambda(k+mx)$, the distribution of accidents for each t is negative binomial with parameters $\left(-\frac{k}{m},\ \frac{1}{1-e^{-\lambda mt}}\right)$, when λ is fixed (see Irwin (1941)) and generalized Waring with parameters $\left(\frac{k}{m},\ 1,\ \frac{a}{mt}\right)$, when $\lambda\ ae^{-a\lambda}, \quad a > 0$ (Xekalaki (1981)).

Further, following Irwin (1941), 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 $\left(-\frac{k}{m}+x,\ \frac{1}{1-e^{-\lambda mt}}\right)$ when λ is fixed, and a generalized Waring distribution with parameters $\left(\frac{k}{m}+x,\ 1,\ \frac{a}{mt}\right)$, when $\lambda\ ae^{-a\lambda}, \quad a > 0$. Hence, in this case,

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

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

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.

2.6 Inferential Aspects Connected to a Process of GW Form

In this section, we discuss some inferential aspects connected with the mixed negative binomial derivation of the generalized Waring process.

Let $M(t)$ be associated with a negative binomial process specified by 1 and $N(t)$ be associated with a generalized Waring process as defined by the Definition 3. The derivation of the latter implies that regarding the parameter ν in

$$P\{M(t)=n\}=\binom{kt+n-1}{n}\left(\frac{1}{\nu+1}\right)^{kt}\left(\frac{\nu}{1+\nu}\right)^{n}, \quad n=0.1,... \tag{6}$$

as the outcome of a random variable having the *beta* (a,ρ)distribution of the second kind, we can interpret $\{P(M(t)=n)\,;\ n=0,1,...\}$ as the conditional distribution of $N(t)$ given the value ν. Hence, the unconditional distribution of $N(t)$ can be represented by

$$P_n(t)=P\{N(t)=n\}=E\left[\binom{kt+n-1}{n}\left(\frac{1}{\nu+1}\right)^{kt}\left(\frac{\nu}{1+\nu}\right)^{n}\right]$$

$$=\frac{\rho_{(kt)}}{(\rho+a)_{(kt)}}\frac{a_{(n)}\,(kt)_{(n)}}{(\rho+a+kt)_{(n)}}\frac{1}{n!}\ n=0,1,... \tag{7}$$

Using this interpretation of the generalized Waring distribution we can, for any event B, regard the probability

$$U^{B}(x)=P\{\nu\leq x|N(s)\in B\}=\frac{\int_0^x P(N(s)\in B|\nu)\,dU(\nu)}{\int_0^{+\infty} P(N(s)\in B|\nu)\,dU(\nu)},$$

with U denoting the probability function of the random variable ν, as the posterior distribution of ν given B or, more precisely, given$\{N(s)\in B\}$, provided that

$$P\{N(s)\in B\}=\int_0^{+\infty}P(N(s)\in B|l)\,dU(l)>0.$$

Proposition 2 *Let $N(s)$ be defined as above. Then*

$$P\{\nu\leq x|N(s)=n\}=\frac{\int_0^x \nu^{n+a-1}(1+\nu)^{-(n+a+\rho+ks)}\,d\nu}{\int_0^{+\infty} \nu^{n+a-1}(1+\nu)^{-(n+a+\rho+ks)}\,d\nu}$$

Proof

Using an argument similar to that used by Xekalaki (1983b) for the case of the generalized Waring distribution, we obtain

$$P\{\nu \le x|N(s)=n\} = \frac{P\{\nu \le x,\ N(s)=n\}}{P\{\ N(s)=n\}}$$

$$= \frac{\int_0^x \frac{\Gamma(a+\rho)}{\Gamma(a)\Gamma(\rho)}\nu^{a-1}(1+\nu)^{-(a+\rho)}\,d\nu \int_0^x \frac{(ls)^n \exp(-ls)}{n!}\frac{(\nu/s)^{-ks}}{\Gamma(ks)} l^{ks-1}\exp\left(-\frac{ls}{\nu}\right)dl.}{\int_0^{+\infty} \frac{\Gamma(a+\rho)}{\Gamma(a)\Gamma(\rho)}\nu^{a-1}(1+\nu)^{-(a+\rho)}\,d\nu \int_0^{+\infty} \frac{(ls)^n \exp(-ls)}{n!}\frac{(\nu/s)^{-ks}}{\Gamma(ks)} l^{ks-1}\exp\left(-\frac{ls}{\nu}\right)dl.}$$

$$= \frac{\binom{ks+n-1}{n}\int_0^x \nu^{n+a-1}(1+\nu)^{-(a+n+\rho+ks)}\,d\nu}{\binom{ks+n-1}{n}\int_0^{+\infty} \nu^{n+a-1}(1+\nu)^{-(a+n+\rho+ks)}\,d\nu} = \frac{\int_0^x \nu^{n+a-1}(1+\nu)^{-(a+n+\rho+ks)}\,d\nu}{\int_0^{+\infty} \nu^{n+a-1}(1+\nu)^{-(a+n+\rho+ks)}\,d\nu}$$

This proposition implies that $\left(\frac{\rho+ks}{a+n}\right)\nu|\left(N(s)=n\right)$ has the F distribution with $2(a+n)$ and $2(\rho+ks)$ degrees of freedom. Following Xekalaki (1983b), this result can be used to construct confidence intervals for $\nu|\left(N(s)=n\right)$, 'estimating' in this way a person's proneness on the basis of the incurred number of accidents. In particular, a $100(1-\alpha)\%$ interval for $\nu|\left(N(t)=n\right)$ is $\left(\frac{a+n}{\rho+kt}F_{\frac{1-\alpha}{2}}, \frac{a+n}{\rho+kt}F_{\frac{1+\alpha}{2}}\right)$, where $F_{\frac{1-\alpha}{2}}(\ 2(a+n), 2(\rho+kt)\ ) = F_{\frac{1-\alpha}{2}}$, $F_{\frac{1+\alpha}{2}}(\ 2(a+n), 2(\rho+kt)\ ) = F_{\frac{1+\alpha}{2}}$.

Corollary 1 *If $N(s)$ is defined as above, then*

$$E\{\nu|N(s)=n\} = \frac{\int_0^{+\infty} \nu^{n+a}(1+\nu)^{-(n+a+\rho+ks)}\,d\nu}{\int_0^{+\infty} \nu^{n+a-1}(1+\nu)^{-(n+a+\rho+ks)}\,d\nu} = \frac{a+n}{\rho+ks}.$$

Proof

Using the result of the proposition and the relation

$$E\{\nu|N(s)=n\} = \int_0^{+\infty} x\,dP\{\nu \le x|N(s)=n\},$$

we obtain

$$E\{\nu|N(s)=n\} = \frac{\int_0^{+\infty} \nu^{n+a}(1+\nu)^{-(n+a+\rho+ks)}\,d\nu}{\int_0^{+\infty} \nu^{n+a-1}(1+\nu)^{-(n+a+\rho+ks)}\,d\nu}$$

$$= \frac{\Gamma(n+a+1)\,\Gamma(\rho+ks-1)}{\Gamma(a+n+\rho+ks)} \frac{\Gamma(a+n+\rho+ks)}{\Gamma(n+a)\,\Gamma(\rho+ks)},$$

which leads to the result.

CHAPTER 3

MODELING SPATIAL OVERDISPERSION WITH THE GW PROCESS

The definition of an appropriate probability model for clustered spatial patterns typically requires the determination of an additive point process on the domain in question. Additivity is a minimal requirement, requiring that when the region of observation changes, or when non-overlapping regions are aggregated in a systematic manner, the corresponding count distribution remains in the same family. Additional assumptions that are often made for convenience include stationarity, ergodicity (allowing estimation of the model based on a single realization) and orderliness (to avoid the apparition of coincident events).

However, point processes known as negative binomial processes have been employed in many practical situations, they have been shown to fail in simultaneously accommodating the three properties listed above. As a matter of fact, it has been conjectured by Diggle & Milne (1983), that additive, stationary, ergodic, orderly spatial point processes with negative binomial finite-dimensional distributions may not even exist.

To elaborate, the construction of a negative binomial process N usually hinges on one of two schemes. The first scheme is based on compounding Poisson processes by means of the logarithmic distribution (see Feller (1968). Let B be a member of the family $\mathcal{B}$ of Borel sets of $\mathbb{R}^2$ and let $N(B)$ denote the number of x_i in B. Then, $N(B) = \sum_{k=1}^{M(B)} X_k$ denotes the count of points corresponding to M disjoint subsets of B, where M is a stationary Poisson process with mean (intensity) measure $E\left(M(B)\right) = \lambda{\cdot}\mu\left(B\right)$, with $\mu\left(B\right)$ denoting the area (Lebesgue measure) of of B. Given M, the random variables X_i are taken to be independently and identically distributed (i.i.d) according to the logarithmic series distribution with parameter δ, having probability generating function (p.g.f.) $\dfrac{-\ln(1-\frac{\delta z}{1+\delta})}{\ln(1+\delta)}$, $\delta > 0$. The resulting process can be seen to be of negative binomial form with p.g.f. $E\{z^{N(B)}\} = \{1+\delta(1-$

$z)\}^{\frac{-\lambda \cdot \mu(B)}{\ln(1+\delta)}}$. This is a Poisson cluster scheme (e.g. Daley & Vere-Jones (1972, Example 2.4 B), Cox & Isham (1980), Fisher (1972, Example 5.6), Burnett & Wasan (1980)), and, as remarked by Diggle & Milne (1983), is always stationary and ergodic (since any stationary Poisson cluster process is known to be mixing and hence ergodic (Westcott (1971, p.300))), but clearly non-orderly.

A second scheme is based on mixing Poisson processes, generating so-called Polya processes (see e.g. Matern (1971), Daley & Vere-Jones (1972, Example 2.1 C), Fisher (1972, p.500)). Here, one samples a gamma random variable Λ with parameters α and $\beta > 0$, and conditionally specifies $N(B)$ to be Poisson given Λ, with intensity $\Lambda \cdot \mu(B)$. The resulting process is again of the negative binomial type, with p.g.f. $E\{z^{N(B)}\} = \{1+\beta(1-z)\}^{-\alpha \cdot \mu(B)}$. Polya processes on the real line are well-established in the literature on accident proneness e.g. Cane (1972) . As mentioned again by Diggle & Milne (1983), they are stationary by construction. However, the only stationary mixed Poisson processes which are ergodic are those for which the mixing distribution is concentrated at a single point, thus giving an (ordinary) Poisson process (e.g. Westcott (1972, p. 464)). It follows that non-trivial processes of this type can be orderly but never ergodic.

In summary, the first approach yields ergodic but non-orderly processes, whereas the second approach yields orderly but non-ergodic processes. In this thesis, therefore, rather than make a new attempt at finding a point process with precisely negative binomial one-dimensional distributions (which may not even be possible), we change strategy, and consider a different choice of over-disperse one-dimensional distributions. An established competitor to the negative binomial distribution is the Generalized Waring Distribution (GWD; see, e.g. Irwin (1975), Xekalaki (1983b, 1984)). This has long been used to fit overdisperse count data, particularly in the field of accident studies, providing a more plausible model for the interpretation of the data generating mechanism by allowing for the distinction of the non-random factors that contribute to the occurrence of an event (e.g. an accident) into intrinsic (inherent, endogenous) and extrinsic (external, exogenous) factors. Moreover, it can approximate the negative binomial and the Poisson distribution as limiting cases.

In this chapter, using the GWD as a building block, we construct an additive, stationary, ergodic, and orderly spatial point process, and study its basic properties (see Zografi, M. & Xekalaki, E. (2019)). We develop our results on a general separable metric state space. The process is seen to satisfy several useful closure properties (under projection, marginalization, and superposition) and to be easy to simulate. We further show that, in the limit as certain parameters of the process diverge, this Generalized Waring Point Process approximates a negative binomial process. In doing so, we give an approximate positive solution to the task set out by Diggle & Milne: while a stationary, ergodic and orderly point process with one-dimensional negative binomial distributions may not exist, there exists a point process that is stationary, ergodic and orderly and has one dimensional distributions that are approximately negative binomial (depending on parameter choice).

The chapter is organised as follows. In section 3.1, we provide some necessary background notions related to the generalized Waring distribution, its moments, and properties that will be used in subsequent chapters. Specifically, it is shown that the generalized Waring distribution posesses the property of countable additivity, which is fundamental to our later construction. The definition and existence of the generalized Waring process in a complete separable metric space is given in section 3.2. In particular, the process is shown to be orderly, and to be characterised by the property that $N(A)$ follows a Univariate Generalized Waring Distribution (UGWD) with parameters $(a, k\mu(A), \rho)$ for all bounded sets A in a dissecting ring $\mathcal{A}$ of the complete separable metric space. A conditional property of the generalized Waring process is also shown enabling simulation of it, and exact expressions for the corresponding intensity measure, factorial moment measures and the n^{th} order moment measures are provided.

3.1 The Generalized Waring Distribution and Additivity

In this section we provide some background on the generalized Waring distribution and discuss some of its structural properties that will be essential in what follows. In particular, we extend the previously established finite additivity property to countable additivity, as a first important step in the construction of the GW point process.

3.1.1 The Univariate Case

A random variable X is said to have the generalized Waring distribution with parameters a, k and ρ, denoted by GWD$(a, k; \rho)$, if

$$P\{X = n\} = \pi_n(a, k; \rho) = \frac{\rho_{(k)}}{(\rho + a)_{(k)}} \frac{a_{(n)} k_{(n)}}{(\rho + a + k)_{(n)}} \frac{1}{n!} \quad {}_{n=0,1,...} \tag{8}$$

where $P(X = x) = 0$, $x \in \{0, 1, 2, ...\}^C$ and the symbol $s_{(t)}$ stands for $\frac{\Gamma(s+t)}{\Gamma(s)}$ for $s > 0$, $t \in \mathbb{R}$ (see e.g. Irwin (1975), Xekalaki (1981), Xekalaki (1983b)). Here $a > 0$, $k > 0$, $\rho > 0$ and k need not to be integers. The distribution is symmetric in a and k.

The probability generating function of the generalized Waring distribution is given by

$$E\left(z^X\right) = \sum_{n=0}^{\infty} z^n \pi_n(a, k; \rho) = \frac{\rho_{(k)}}{(\rho + a)_{(k)}} \, {}_2F_1(a, k; \rho + a + k; z) \tag{9}$$

where

$${}_2F_1(a, \beta; \gamma; z) = \sum_{n=0}^{\infty} \frac{a_{(n)} \beta_{(n)}}{\gamma_{(n)}} \frac{z^n}{n!}.$$

The rth factorial moments are

$$\mu_{[r]} = \frac{a_{[r]} k_{[r]}}{(\rho - 1)(\rho - 2) \ldots (\rho - r)} \tag{10}$$

where $x_{[r]} = x(x+1) \ldots (x+r)$ for each x and r.

From (10), it follows immediately that all of the rth order moments, ordinary moments about any origin, central moments as well as factorial moments are infinite if $\rho \leq r$. Moments about any origin, including central moments, can be obtained from (10) by the usual transformation formula (see Irwin (1975), PartI). In particular, the mean is given by

$$E(X) = \frac{ak}{\rho - 1}, \ \rho > 1 \tag{11}$$

while the variance is

$$\sigma^2 = \mu_2 = \frac{ka(\rho + a - 1)(\rho + k - 1)}{(\rho - 1)^2 (\rho - 2)}, \ \rho > 2. \tag{12}$$

It is worth noting that the variance is splittable into additive components associated to the effect of random factors (σ_R^2) and to non-random factors further distinguished into intrinsic factors (σ_ν^2) and extrinsic factors (σ_λ^2).

$$\sigma^2 = \sigma_\lambda^2 + k^2\sigma_\nu^2 + \sigma_R^2$$

where $\sigma_\lambda^2 = ak\,(a+1)\,(\rho-1)^{-1}\,(\rho-2)^{-1}$, $\sigma_\nu^2 = a\,(a+\rho-1)\,(\rho-1)^{-2}\,(\rho-2)^{-1}$ and $\sigma_R^2 = ak\,(\rho-1)^{-1}$.

3.1.2 The Multivariate Generalized Waring Distribution

The multivariate generalized Waring distribution with parameter vector (α, k_1, ..., k_s; ρ), denoted by MGWD(a; $\mathbf{k}$; ρ), is the probability distribution of a random vector (X_i, $i = 1,\ 2,\ ...,\ s$) of nonnegative integer-valued components, with probability function given by

$$P_{x_1,\dots,x_s} = P\left(X_i = x_i, i = 1, 2, ..., s\right) = \frac{\rho_{\left(\sum_{i=1}^{s} k_i\right)}\, a_{\left(\sum_{i=1}^{s} x_i\right)}}{(\rho+a)_{\left(\sum_{i=1}^{s} k_i + \sum_{i=1}^{s} x_i\right)}} \prod_{i=1}^{s} \frac{k_{i(x_i)}}{x_i!} \tag{13}$$

(see Xekalaki (1986)). The special case for $s = 2$ is known in the literature as the bivariate generalized Waring distribution, denoted by BGWD$(a; k_1, k_2; \rho)$.

The probability generating function of the multivariate Generalized Waring distribution can be expressed in terms of Lauricella's hypergeometric function of type D as

$$G\left(\underline{z}\right) = \frac{\rho_{\left(\sum_{i=1}^{s} k_i\right)}}{(\rho+a)_{\left(\sum_{i=1}^{s} k_i\right)}}\; F_D(a, k_1, k_2, ..., k_s; \rho + a + \sum_{i=1}^{s} k_i; \underline{z})$$

where

$$F_D(a, \beta_1, \beta_2, ..., \beta_s; \gamma; \underline{z}) = \sum_{r_1, r_2, \dots r_s} \frac{a_{\left(\sum r_i\right)}}{\gamma_{\left(\sum r_i\right)}} \prod_{i=1}^{s} \frac{(\beta_i)_{(r_i)}\quad (z_i)^{r_i}}{r_i!},$$

$$\underline{z} = (z_1,\ z_2,\ ...,\ z_s)$$

The factorial moments of the MGWD(a; $\mathbf{k}$; ρ) (see Xekalaki (1985a), Xekalaki (1986)) are then given by

$$\mu_{(r_1,r_2,...r_l)} = E\left[(X_1)_{[r_1]}(X_2)_{[r_2]}...(X_s)_{[r_s]}\right] \tag{14}$$

$$= \frac{a_{\left(\sum r_i\right)}\prod_{i=1}^{s}(k_i)_{(r_i)}}{(\rho-1)(\rho-2)...(\rho-\sum r_i)},\ r_i = 0,1,...;\ i=1,2,...,s \tag{15}$$

and are finite for $\rho > \sum r_i$, the latter being a necessary condition for the series $F_D(a, k_1 + r_1, k_2 + r_2, ..., k_s + r_s; \rho + a + \sum_{i=1}^{s}(k_i + r_i); \underline{1})$ to converge. Moments of order n can be derived from these factorial moments.

The marginal means and marginal variances are respectively given by

$$\mu_{X_i} = E(X_i) = \frac{ak_i}{\rho - 1},\ \rho > 1 \tag{16}$$

$$\sigma^2_{X_i} = \frac{k_i a(\rho + a - 1)(\rho + k_i - 1)}{(\rho-1)^2(\rho-2)}, \rho > 2 \tag{17}$$

$i = 1, 2, ..., s$.(see Xekalaki (1986)).

The second moment and the pairwise covariances are

$$\mu_{X_iX_j} = E(X_iX_j) = \frac{a(a+1)k_ik_j}{(\rho-1)(\rho-2)},\ i,j = 1,2,...,s;\ \rho > 2 \tag{18}$$

$$\sigma_{X_iX_j} = \frac{a(\rho+a-1)k_ik_j}{(\rho-1)^2(\rho-2)}, i,j = 1,2,...,s; \rho > 2 \tag{19}$$

One of the most important features of the GWD is **additivity.** Specifically, if X and Y are random variables with marginal distributions UGWD(a, k_1; ρ) and UGWD(a, k_2; ρ), respectively, and with joint distribution BGWD(a; k_1, k_2; ρ), then $X + Y$ is a UGWD(a, $k_1 + k_2$; ρ) random variable. More generally, letting X_j be UGWD(a, k_j; ρ) or each j, $j = 1, 2, ..., n$ and jointly distributed as MGWD(a; k_1, k_2, ..., k_n; ρ), then, if we denote $m = \sum_{j=1}^{n} k_j$, we have that $S = \sum_{j=1}^{n} X_j$ also has a UGWD(a, m; ρ) distribution.

These last two properties hint at the possibility of using the GWD as a basis for the construction of overdisperse point processes. This requires extending additivity to countable additivity, which we do in the form of the next theorem:

3.1.3 Countable Additivity Theorem

Theorem 3 *Let X_j be UGWD$(a, k_j; \rho)$ variables for each j, $j = 1, 2, ...$ and for each $n \geq 3$ let their joint distribution be the MGWD$(a, k_1, k_2, ..., k_n; \rho)$. If $m = \sum_{j=1}^{\infty} k_j$ converges, then $S = \sum_{j=1}^{\infty} X_j$ converges with probability 1, and S has a UGWD$(a, m; \rho)$ distribution. If on the other hand, $\sum_{j=1}^{\infty} k_j$ diverges, then S diverges with probability 1.*

Proof By induction on n, the random variable $S_n = \sum_{j=0}^{n} X_j$ has a UGWD$(a,\ m_n;\ \rho)$ distribution, where $m_n = \sum_{j=1}^{n} k_j$.Thus, for any r,

$$P\left\{S_n \leq r\right\} = \sum_{i=0}^{r} \pi_i \left(a, m_n; \rho\right) \tag{20}$$

The sequence $\{S_n \leq r\}$ is a decreasing sequence of events for fixed r, and their intersection is $\{S \leq r\}$.Thus, using continuity from above,

$$\begin{aligned} P\left\{S \leq r\right\} &= \lim_{n \to \infty} P\left\{S_n \leq r\right\} \\ &= \lim_{n \to \infty} \sum_{i=0}^{r} \pi_i \left(a, m_n; \rho\right). \end{aligned}$$

If m_n converges to a finite limit m, the continuity of π_j implies that

$$P\left\{S \leq r\right\} = \sum_{i=0}^{r} \pi_i \left(a, m; \rho\right) \tag{21}$$

leading to

$$P\left\{S = r\right\} = \pi_r \left(a, m; \rho\right). \tag{22}$$

This in turn implies that S is finite and distributed as generalized Waring with parameters $a, m; \rho$ (UGWD$(a, m; \rho)$).

On the other hand, if $m_n \to \infty$,

$$\sum_{i=0}^{r} \pi_i \left(a, m_n; \rho\right) = \sum_{i=0}^{r} \frac{\rho_{(m_n)}}{(\rho + a)_{(m_n)}} \frac{a_{(i)} m_{n(i)}}{(\rho + a + m_n)_{(i)}} \frac{1}{i!} =$$

$$\left[\frac{\rho_{(a)}}{(\rho + m_n)_{(a)}} \frac{a_{(i)}}{i!} \frac{m_n}{(\rho + a + m_n)} \frac{(m_n + 1)}{(\rho + a + m_n + 1)} \cdots \frac{(m_n + i - 1)}{(\rho + a + m_n + i - 1)} \to 0 \right]$$

so that $P\left\{S > r\right\} = 1$. Since this holds for all r, S diverges with probability 1.

3.2 The Generalized Waring Process in a Complete Separable Metric Space

We now proceed to the definition of the generalized Waring process on a complete separable metric space and the investigation of some of its basic properties. The construction starts from postulating the existence of a point process with finite dimensional distributions of the generalized Waring form (subsection 3.2.1), and then demonstrating the existence and uniqueness of such a process (subsection 3.2.2). Basic features of the process such as a conditional property useful for simulation, as well as its intensity measure, factorial moment measures and n^{th} order moment measures are then derived in subsection 3.2.3.

3.2.1 Definition and Basic Properties

Let $\mathcal{S}$ be a complete separable metric space, $\mathcal{A}$ a semiring of bounded Borel sets generating the Borel σ-algebra $\mathcal{B}_S$ of subsets of $\mathcal{S}$ (Appendix2. Lemma A2.I.III, Daley and Vere-Jones (1988)) and $\mu(\cdot)$ a boundedly finite Borel measure. The distribution of a random measure is completely determined by its finite dimensional (fidi) distributions, i.e. the joint distribution of arbitrary finite families $\{A_i,\ i = 1,\ ...,\ s\}$ of disjont sets from $\mathcal{A}$ (Proposition 6.2.III, Daley & Vere-Jones (1988)). Now consider the space of all boundedly finite, integer-valued measures $(\overset{\wedge}{\mathcal{N}}_S,\ \mathcal{B}\left(\overset{\wedge}{\mathcal{N}}_S\right))$ and let $(\Omega, \mathcal{F}, \mathcal{P})$ be some probability space.

Definition 1 *Let*

$$N : (\Omega, \mathcal{F}, \mathcal{P}) \to \left(\overset{\wedge}{\mathcal{N}}_S, \mathcal{B}\left(\overset{\wedge}{\mathcal{N}}_S\right)\right)$$

be a point process for whose finite dimensional distributions over disjoint bounded Borel sets $\{A_i,\ i = 1,\ ...,\ l\ \}$ *are given by*

$$P\left\{N\left(A_i\right) = n_i; i = 1, ..., l\right\} = \frac{\rho_{\left(k\sum_{i=1}^{l}\mu(A_i)\right)} a_{\left(\sum_{i=1}^{l} n_i\right)}}{(\rho + a)_{\left(k\sum_{i=1}^{l}\mu(A_i) + \sum_{i=1}^{l} n_i\right)}} \prod_{i=1}^{l} \frac{\left[k\mu\left(A_i\right)\right]_{(n_i)}}{n_i!}. \tag{23}$$

Then N is called a generalized Waring process with parameters a, ρ, $k > 0$ and parameter measure $\mu(\cdot)$.

In other words, for every finite family of disjoint bounded Borel sets $\{A_i,\ i = 1,\ ...,\ l\}$ the joint distribution of $\{N(A_i) = n_i, i = 1, ..., l\}$ is the MGWD($a,\ k\mu(A_1),\ k\mu(A_2),\ ...,\ k\mu(A_l);\ \rho$). As usual, the process $\{N(A);\ A \in \mathcal{B}_S\}$ is to be thought of as a random measure. In particular, for any $A \in \mathcal{B}_S$, $N(A)$ is a Z^t-valued random variable, while for any $\omega \in \Omega$, $N(\omega, \cdot)$ is a discrete Radon measure.

We remark that, if such a process exists, it will necessarily be countably additive. To see this, let $\{A_i, i = 1, 2, ...\}$ be disjoint and have union A. Using Theorem 1, and the fact that $\mu(A) = \sum_{i=1}^{\infty} \mu(A_i)$ converges, we immediately obtain that $N(A) = \sum_{i=1}^{\infty} N(A_i)$ is distributed as UGWD$(a, k\mu(A); \rho)$. Furthermore, such a process will be orderly provided the parameter measure is diffuse:

Theorem 4 *A process as in Definition 1 is an orderly point process if and only if its parameter measure has no fixed atoms.*

Proof A point process is orderly when given any bounded $A \in \mathcal{B}_S$, there is a dissecting system $\mathcal{T} = \{\mathcal{T}_n\} = \{\ \{\ A_{ni}\ :\ i\ =\ 1,\ ...,\ k_n\ \}\ \}$ such that $\inf\limits_{\mathcal{T}_n} \ \sum_{i-1}^{k_n} P\{N(A_{ni}) > 2\} = 0$. (see Daley & Vere-Jones (1988)). Hence it is sufficient to examine when the ratio $\dfrac{P\{N(A_{\varepsilon,x}) > 1\}}{P\{N(A_{\varepsilon,x}) > 0\}}$ tends to 0, where $A_{\varepsilon,x}$ is the open sphere of radius ε and center $x \in A$. In the case of a GW process, $N(A_{\varepsilon,x})$ has a generalized Waring distribution with parameters $a > 0$, $\rho > 0$ and $\mu(A_{\varepsilon,x}) = \mu_\varepsilon$, so that

$$P\{N(A_{\varepsilon,x}) > 0\} = 1 - P\{N(A_{\varepsilon,x}) = 0\} = 1 - \frac{\rho_{(k\mu_\varepsilon)}}{(\rho + a)_{(k\mu_\varepsilon)}},$$

$$P\{N(A_{\varepsilon,x}) > 1\} = 1 - \frac{\rho_{(k\mu_\varepsilon)}}{(\rho + a)_{(k\mu_\varepsilon)}} - \frac{\rho_{(k\mu_\varepsilon)}}{(\rho + a)_{(k\mu_\varepsilon)}} \frac{a \cdot k\mu_\varepsilon}{(\rho + a + k\mu_\varepsilon)}.$$

If x is a fixed atom of μ, then $\mu_\varepsilon \to \mu_0 = \mu\{x\} > 0$ as $\varepsilon \to 0$, while if x is not a fixed atom, then $\mu(A_{\varepsilon,x}) \to 0$.

In the first case, the ratio $\dfrac{P\{N(A_{\varepsilon,x}) > 1\}}{P\{N(A_{\varepsilon,x}) > 0\}}$ tends to the constant $1 - \dfrac{\rho_{(k\mu_0)} \cdot a \cdot k\mu_0}{(\rho + a)_{(k\mu_0+1)} - \rho_{(k\mu_0)}}$, while in the second case it tends to 0, and the proof is complete.

From this point on, we will consider only orderly generalized Waring processes. Indeed, any orderly point process with finite dimensional distributions of the generalized Waring type is necessarily a GWP with a non-atomic parameter measure:

Theorem 5 *Let $N(\cdot)$ be an orderly point process. For $N(\cdot)$ to be a generalized Waring process with parameters $a > 0$, $\rho > 0$, $k > 0$ and parameter measure $\mu(\cdot)$, it is necessary and sufficient that there exist a boundedly finite nonatomic measure μ on the Borel sets $\mathcal{B}_s$ such that $N(A)$ has generalized Waring distribution with parameters $a, k\mu(A), \rho$ for each bounded set A of a dissecting ring $\mathcal{A}$ of the complete separable metric space $\mathcal{S}$.*

Proof We begin with necessity. Let $N(\cdot)$ be a generalized Waring Process and A a bounded set of a dissecting ring $\mathcal{A}$ (A is also a Borel set). Then, by definition, there exists a boundedly finite Borel measure $\mu(\cdot)$ such that for every finite family of disjoint bounded Borel sets $\{A_i, i = 1, ..., s\}$, $P\{N(A_i) = n_i, i = 1, ..., s\}$ is given by 23. From this, it follows that the distribution of $N(A)$ is the GWD($a,\ k\mu(A);\ \rho$).

To prove sufficiency, suppose that there exists a boundedly finite nonatomic measure μ on the Borel sets $\mathcal{B}_s$ such that $N(A)$ has generalized Waring distribution with parameter $a, k\mu(A), \rho$ for each bounded set A of a dissecting ring. According to Theorem 7.3.II of Daley & Vere-Jones (1988), the values of the avoidance function $P_0(A) = P\{N(A) = 0\}$ $= \dfrac{\rho_{(k\mu(A))}}{(\rho + a)_{(k\mu(A))}}$ on the bounded sets of a dissecting ring for the complete separable metric space, determine the distribution of a simple point process $N(\cdot)$ on this space.

3.2.2 Existence Lemma

The following Lemma proves that the equation $P_0(A) = \dfrac{\rho_{(k\mu(A))}}{(\rho + a)_{(k\mu(A))}}$ has always a solution. This result wil be used to prove the measure requirements in the Existence Theorem of the generalized Waring process.

Lemma 2 *For each $a > 0$, $\rho > 0, 0 \le P_0 \le 1$, there exists one and only one root $x > 0$ of the equation $\frac{\Gamma(\rho+x+a)}{\Gamma(\rho+x)} = \frac{\Gamma(\rho+a)}{P_0\Gamma(\rho)}$*

Proof It has been proved (see Bai-ni, Ying-jie and Feng[[2]] Theorem 3) that for $y > x \ge 1$, $\dfrac{\Gamma(y)}{\Gamma(x)} > \dfrac{y^{y-\gamma}}{x^{x-\gamma}} e^{x-y}$where γ is Euler-Mascheroni's constant

Hence we can obtain

$$\frac{\Gamma(\rho + x + a)}{\Gamma(\rho + x)} > \frac{1}{e^a} \frac{(\rho + x + a)^{\rho + x + a - \gamma}}{(\rho + x)^{\rho + x - \gamma}}$$

Let us consider the function

$$f(x) = \frac{(\rho + x + a)^{\rho + x + a - \gamma}}{(\rho + x)^{\rho + x - \gamma}} e^{-a}$$

We will show that it is an increasing function of x for x>1. To this aim, we need to examine its derivative

$$\begin{aligned} f'(x) &= e^{-a} \frac{\left[\ln(\rho + x + a) + 1 - \frac{\gamma}{\rho + x + a}\right](\rho + x + a)^{\rho + x + a - \gamma}}{(\rho + x)^{2(\rho + x - \gamma)}} \\ & e^{-a} \frac{-\left[\ln(\rho + x) + 1 - \frac{\gamma}{\rho + x}\right](\rho + x)^{\rho + x - \gamma}}{(\rho + x)^{2(\rho + x - \gamma)}} \end{aligned}$$

Observe that the functions $\varphi(x) = \left(1 + \ln x - \frac{\gamma}{x}\right)$ and $\omega(x) = x^x$ are increasing for $x > 1$, (since $\varphi'(x) = \left(\frac{1}{x} + \frac{\gamma}{x^2}\right) > 0$ if $x > -\gamma$, $\omega'(x) = (1 + \ln x)\, x^x > 0$ if $x > 1$). Hence

the function $g(x) = \varphi(x)\,\omega(x) = \left(1 + \ln x - \dfrac{\gamma}{x}\right) x^x$ increases for $x > 1$.

So, $\left[\ln(\rho + x + a) + 1 - \dfrac{\gamma}{\rho + x + a}\right] (\rho + x + a)^{\rho + x + a - \gamma} - \left[\ln(\rho + x) + 1 - \dfrac{\gamma}{\rho + x}\right]$ $(\rho + x)^{\rho + x - \gamma} > 0$ for $x > 1$ and the $f'(x) > 0$ which proves that the function $f(x)$ increases for $x > 1$.

So we can state that $\forall b \in R$, $\exists x > 1$, such that $f(x) > b$.

Let us consider $b = \dfrac{\Gamma(\rho + a)}{P_0 \Gamma(\rho)}$. For that value there exists an $x > 0$, such that $f(x) > b$.

Clearly for $x = 0$, $\dfrac{\Gamma(\rho + x + a)}{\Gamma(\rho + x)} = \dfrac{\Gamma(\rho + a)}{\Gamma(\rho)} < b$. The function $\dfrac{\Gamma(\rho + x + a)}{\Gamma(\rho + x)}$ is continuous for $x > 0$ as a ratio of two continuous functions ($\Gamma(x)$ is continuous $x > 0$). So, using Bolzano's Theorem for the function $\dfrac{\Gamma(\rho + x + a)}{\Gamma(\rho + x)} - b$, we obtain that $\exists x > 0$ such that $\dfrac{\Gamma(\rho + x + a)}{\Gamma(\rho + x)} - b = 0$.

On the other hand,

$$\frac{d}{dx}\left[\frac{\Gamma(\rho + x + a)}{\Gamma(\rho + x)}\right] = \frac{d}{dx}\left[\exp\left(\ln \frac{\Gamma(\rho + x + a)}{\Gamma(\rho + x)}\right)\right]$$

$$= \frac{\Gamma(\rho + x + a)}{\Gamma(\rho + x)} \frac{d}{dx}\left[\ln \Gamma(\rho + x + a) - \ln \Gamma(\rho + x)\right]$$

$$= \frac{\Gamma(\rho + x + a)}{\Gamma(\rho + x)} \left[\Psi(\rho + x + a) - \Psi(\rho + x)\right]$$

and using the relation $\Psi(t) = -\gamma + \sum_{i=0}^{\infty} \left(\frac{1}{i+1} - \frac{1}{i+t} \right)$ where γ is the Euler-Mascheroni constant, we obtain

$$\begin{aligned} \frac{d}{dx}\left[\frac{\Gamma(\rho + x + a)}{\Gamma(\rho + x)}\right] &= \frac{\Gamma(\rho + x + a)}{\Gamma(\rho + x)} \\ \sum_{i=0}^{\infty} \left(\frac{1}{i + \rho + x} - \frac{1}{i + \rho + x + a} \right) &> 0 \end{aligned}$$

This proves the Lemma.

3.2.3 Existence and Uniqueness

To prove that the point process stipulated in the previous section does indeed exist, it is sufficient to establish that the finite dimensional distributions given by (<ref>1a</ref>) fulfill Kolmogorov's consistency conditions, combined with the measure requirements given by the basic existence theorem for point processes (Theorem 7.I.XI Daley & Vere-Jones (1988)).

Theorem 6 *(Kolmogorov's Consistency Conditions) A collection of finite dimensional distributions as defined via Definition 2 satisfies Kolmogorov's consistency conditions. That is, for every finite family of disjoint bounded Borel sets* $\{A_i, i = 1, ..., l\}$*,*

(I) for any permutation $i_1, ..., i_l$ *of the indexes* $1, ..., l$

$$P_l(A_{i_1}, ..., A_{i_l}; n_{i_1}, ..., n_{i_l}) = P_l(A_1, ..., A_l; n_1, ..., n_l) \tag{24}$$

(II) $\sum_{r=0}^{\infty} P_l(A_1, ..., A_l, n_1, ..., n_{l-1}, r) = P_{l-1}(A_1, ..., A_{l-1}, n_1, ..., n_{l-1})$

Proof To show (I), we notice that one can write $\sum_{j=1}^{l} \mu = \sum_{j=1}^{l} \mu(A_j)$, $\sum_{j=1}^{l} n_{i_j} = \sum_{j=1}^{l} n_j$, $\prod_{j=1}^{l} \frac{\left[k\mu(A_{i_j})\right]_{\left(n_{i_j}\right)}}{n_i!} = \prod_{j=1}^{l} \frac{[k\mu(A_j)]_{(n_j)}}{n_j!}$ which proves (24).

To show (II), we write $\sum_{r=0}^{\infty} P_l(\ A_1,\ ...,\ A_l;\ n_1,\ ...,\ n_{l-1},\ r\ ) = P_{l-1}(\ A_1,\ ...,\ A_{l-1};\ n_1,\ ...,$ $n_{l-1})\ \sum_{r=0}^{\infty} \frac{\left(\rho + k\sum_{i=1}^{l-1}\mu(A_i)\right)_{(k\mu(A_r))}\left(a+\sum_{i=1}^{l-1}n_i\right)_{(r)}}{\left(\rho + a + k\sum_{i=1}^{l-1}\mu(A_i)+\sum_{i=1}^{l-1}n_i\right)_{(k\mu(A_r)+r)}}\ \frac{[k\mu(A_i)]_{(r)}}{r!} = P_{l-1}(\ A_1,\ ...,\ A_{l-1}; n_1,\ ...,$ $n_{l-1})$

Theorem 7 *(Measure Requirements) Suppose that*

(I) N is bounded finite a.s. and has no fixed atoms.

(II) N satisfies Definition 2.

Then, there exists a boundedly finite nonatomic Borel measure $\mu(\cdot)$ such that $P_0(A) = \Pr\{N(A)=0\} = \frac{\rho_{(k\mu(A))}}{(\rho+a)_{(k\mu(A))}}$ *for all bounded borel sets A and $\forall i, i=1,...,s$ $\mu(A_i) = \mu_i$.*

Proof Let $A \in \mathcal{B}_s$ and let $\mu(A) > 0$ be the root of the equation $P_0(A) = \frac{\rho_{(k\mu(A))}}{(\rho+a)_{(k\mu(A))}}$ which does exist (see Lemma 2).

a) We first prove that $\mu(\cdot)$ is a measure. To show finite additivity, we observe that

$$P_0(A) = \Pr\{N(A)=0\} = \frac{\rho_{(k\mu(A))}}{(\rho+a)_{(k\mu(A))}}.$$

Hence for each family of bounded, disjoint, Borel sets $\{A_i, i=1,...,s\}$, the joint distribution of $\{N(A_i)=n_i, i=1,...,s\}$ is the MGWD($a, k\mu(A_1), k\mu(A_2), ..., k\mu(A_s); \rho$), and if $A = \sum_{i=1}^{s} A_i$ then $N(A) = \sum_{i=1}^{s} N(A_i)$ has distribution GWD$(a, k\mu(A); \rho)$. So $\mu(A) = \sum_{i=1}^{s}\mu(A_i)$ which establishes finite additivity of $\mu(\cdot)$. To extend this to countable additivity, it suffices to prove that $\mu(A_i) \to 0$ for any decreasing sequence $\{A_i\}$ of bounded Borel sets for which $\mu(A_i) < \infty$ and $A_i \downarrow \emptyset$. For $A_i \downarrow \emptyset$ $N(A_i) \to 0$ a.s. and thus $P_0(A_i) = \Pr\{N(A_i)=0\} \to 1$ a.s. hence $\mu(A_i) = \frac{\rho(1-P_0(A_i))}{kP_0(A_i)} \to 0$ a.s.

b)To show that $\mu(\cdot)$ is non-atomic, we can consider by (I) that for every x that $\Pr\{\ N(\{x\}) > 0\ \} = (1-P_0(\{x\})) = 0$.So $\mu(\{x\}) = \frac{\rho(1-P_0(\{x\}))}{kP_0(\{x\})} = 0$

c)To show that $\mu(\cdot)$ is boundedly finite it is enough to prove that $P_0(A) > 0$ for every bounded borel set A. By supposing the contrary that for some set A, $P_0(A) = 0$, one, following Daley & Vere-Jones (1988), Lemma 2.4.VI, can find that in this case there exists a fixed atom of the process, contradicting (I) which proves that $P_0(A) > 0$ for every bounded borel set A.

3.2.4 Conditional Property and Moment Measures

In this subsection, we prove a key structural property of the generalized Waring process, that enables one to simulate segments of the process. The algorithm constructed for this purpose, reveals that using the generalized Waring process one is enabled not only to obtain a description of the counts, but also acquire knowledge of the exact locations where spatial clustering occurs. Some simulation results are also provided. Exact expressions of the n^{th} order moment measures of the generalized Waring process that will be needed to establish n^{th} order stationarity in chapter 4, are also derived.

Theorem 8 *(Conditional Property). Consider a Generalized Waring point process in* Ω *with parameters* $a > 0, \rho > 0, k > 0$ *. Let* $W \subset \Omega$ *be any region with* $0 < \mu(W) < +\infty$*. Given that* $N(W) = n$*, the conditional distribution of* $N(B)$ *for* $B \subset W$ *is the beta-binomial distribution with parameters* $\mu(B)$, $\mu(W) - \mu(B)$ *and* n *:*

$$p(N(B) = k \,|N(W) = n) = \binom{n}{k} \frac{(\mu(B))_{(k)} (\mu(W) - \mu(B))_{(n-k)}}{(\mu(W))_{(n)}}$$

Proof

$$p(N(B) = k \,|N(W) = n) = \frac{p(N(B) = k, N(W-B) = n-k)}{p(N(W) = n)}$$

$$= \frac{\frac{\rho_{(a)}}{(\rho + \mu(W))_{(a)}} \frac{a_{(n)} (\mu(B))_{(k)} (\mu(W-B))_{(n-k)}}{(\rho + \mu(W) + a)_{(n)}} \frac{1}{k!} \frac{1}{(n-k)!}}{\frac{\rho_{(a)}}{(\rho + \mu(W))_{(a)}} \frac{a_{(n)} (\mu(W))_{(n)}}{(\rho + \mu(W) + a)_{(n)}} \frac{1}{n!}}$$

$$= \frac{n!}{k!\,(n-k)!} \frac{(\mu(B))_{(k)} (\mu(W) - \mu(B))_{(n-k)}}{(\mu(W))_{(n)}}$$

$$= \binom{n}{k} \frac{(\mu(B))_{(k)} (\mu(W) - \mu(B))_{(n-k)}}{(\mu(W))_{(n)}}$$

Furthermore, the conditional joint distribution of $N(B_1)$, ..., $N(B_m)$ for any B_1, ..., $B_m \subseteq W$ is the Multinomial Dirichlet distribution with parameters $\mu(B_1)$, ..., $\mu(B_m)$, $\mu(W) - \sum_{i=1}^{m} \mu(B_i)$ and n.

The following algorithm describes how a Generalized Waring process with parameters $a > 0, \rho > 0, k > 0$ can be generated using the conditional property:

Algorithm 2 *(Generalized Waring Process Generator in a Quadrat W using Multinomial Dirichlet Distribution)*

1. Generate a random variable M with a generalized Waring distribution with parameters $a, \rho, k \cdot \mu(W)$ where $\mu(W) = \mu$.

2. Let B_1, ..., $B_s \subseteq W$ a partition of W in s equal areas, i.e. $\mu(B_i) = \mu/s$ for i in $\{1 : s\}$. We then draw a random vector (M_1, ..., M_s) from a Multinomial Dirichlet Distribution with parameters $(\frac{\mu}{S}, ..., \frac{\mu}{S}, M)$.

3. For each i in $(1, 2, ..., s)$, if $M_i > 1$ do the step 2 for the respective B_i, otherwise stop.

Since $E(M_i) = \frac{M}{s}$ the algorithm converges.

As is clear from the algorithm, no more than one point can occupy the same position. This indicates that contrary to the negative binomial process, the generalized Waring process can not only describe the observed counts, but can also pinpoint the exact locations where spatial clustering occurs. Figure 3 depicts three different realizations of the algorithm applied in a unit square for the same set of parameters.

We now turn our attention to determining the n^{th} order moment measures of the process, needed to establish n^{th} order stationarity, as discussed in the next chapter.

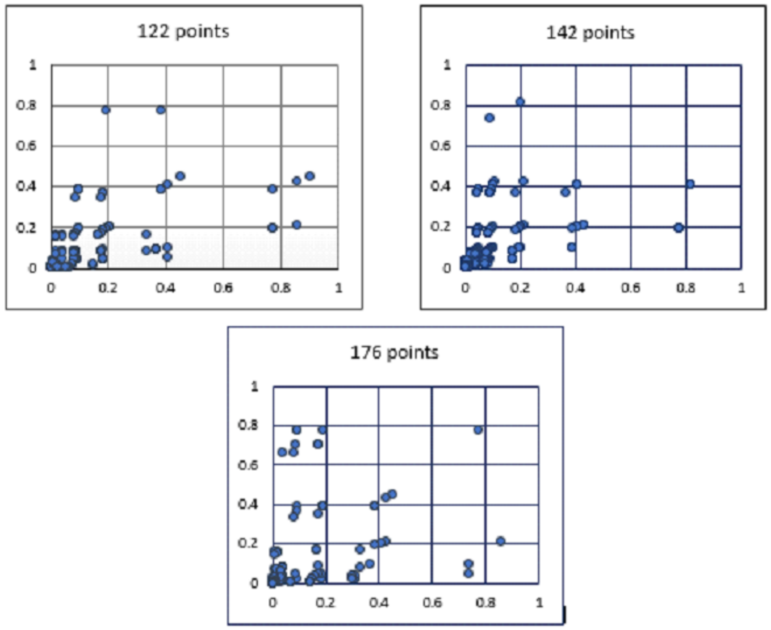

Figure 3
Realizations of a GWP (20,30,5) with mean=150 in the unit square

Let N be a generalized Waring process with parameters $(a, k; \rho)$ and parameter measure $\mu(\cdot)$. For A a Borel set, the distribution of $N(A)$ is the GWD$(a, k\mu(A); \rho)$. Therefore, its first moment measure is

$$\lambda(A) = E(N(A)) = \frac{ak\mu(A)}{\rho - 1}, \ \rho > 1$$

and its intensity rate is the Radon-Nikodym derivative

$$\eta(A) = \frac{d\lambda}{d\mu} = \frac{ak}{\rho - 1}, \ \rho > 1.$$

For A, B two Borel sets the joint distribution of $(N(A), N(B))$ is the BGWD($a, k\mu(A), k\mu(B); \rho$), hence the second order moment measure of the process is

$$M_2(A \times B) = E(N(A)N(B)) = \frac{a(a+1)k^2\mu(A)\mu(B)}{(\rho - 1)(\rho - 2)}, \ \rho > 2 \tag{25}$$

Given a finite family of disjoint bounded Borel sets $\{A_i, i = 1, ..., s\}$ the joint distribution of $\{N(A_i) = n_i, i = 1, ..., s\}$ is the MGWD(a, $k\mu(A_1), k\mu(A_2), ..., k\mu(A_s); \rho$), hence the factorial moment measure, $E\,[N(A_1)_{[r_1]}\ N(A_2)_{[r_2]}\ ...\ N(A_s)_{[r_s]}]$, of the process is

$$\mu_{(r_1, r_2, ... r_l)}(A_1 \times A_2 \times ... \times A_s) = \frac{a^{\left(\sum r_i\right)} k^s \prod_{i=1}^{s} (\mu(A_i))_{(r_i)}}{(\rho - 1)(\rho - 2)...(\rho - \sum r_i)}, \tag{26}$$

$r_i = 0, 1, ...; i = 1, 2, ..., s$. The n^{th} order moment measures can now be obtained from (26)

$$\begin{aligned} M_n\left(A_1 \times A_2 \times ... \times A_n\right) &= E\left[\left(N\left(A_1\right)\right)_{[1]}\left(N\left(A_2\right)\right)_{[1]}...\left(N\left(A_s\right)\right)_{[1]}\right] \\ &= \frac{a_{(n)}k^n \prod_{i=1}^{n}\left(\mu\left(A_i\right)\right)_{(r_i)}}{\left(\rho-1\right)\left(\rho-2\right)...\left(\rho-n\right)}, \text{ for } \rho > n. \end{aligned} \tag{27}$$

for $\rho > \sum r_i$.

CHAPTER 4

MODELING OVERDISPERSION IN $\mathbb{R}^d$

In this chapter we develop a practically relevant case of generalized Waring process in $\mathbb{R}^d$.

The generalized Waring process in $\mathbb{R}^d$ with Lebesgue measure as parameter measure $\mu(\cdot)$ is defined in section 4.1. It is shown to be orderly, ergodic and n^{th} order stationary. The existence of the n^{th} order reduced moments of a generalized Waring process in $\mathbb{R}^d$, if $\rho > n$, useful for applications, is obtained as a corollary.

Multivariate extensions are considered in section 4.2, where we define the multivariate GWP as a special case of the GWP on the product space $S \times \{1, 2, ...m\}$, and is shown to satisfy several appealing closure properties with respect to marginalization. It is proved that the Poisson and the Pólya processes are limiting cases of the generalized Waring process on $\mathbb{R}^+$ and by utilizing this result, the moments and transition probabilities are obtained for the Poisson and the Pólya processes.

4.1 The Generalized Waring Process in $\mathbb{R}^d$

In what follows, we turn our focus to the generalized Waring process on the state-space $\mathbb{R}^d$, with Lebesgue measure as its parameter measure $\mu(\cdot)$. We show that this constitutes an orderly, stationary, ergodic and n^{th}order stationary point process.

4.1.1 The Generalized Waring Process as a Simple Point Process

Let $S = \mathbb{R}^d$ and let $\mu(\cdot)$ be the Lebesgue measure on $\mathbb{R}^d$. The Borel algebra $\mathcal{B}_{\mathbb{R}^d}$ in $\mathbb{R}^d$

is the smallest σ- algebra on $\mathbb{R}^d$ which contains all the open rectangles of $d-$dimensions. The generalized Waring process $\{N(A);\ A \in \mathcal{B}_{\mathbb{R}^d}\}$ can be defined by assuming that for every finite family of disjoint bounded Borel sets $\{A_i,\ i = 1,\ ...,\ s\}$ the joint distribution of $\{N(A_i) = n_i,\ i = 1,\ ...,\ s\}$ is the MGWD$(a,\ k\mu(A_1),\ k\mu(A_2),\ ...,\ k\mu(A_s);\ \rho)$ $, a > 0,$

$\rho > 0$, $k > 0$.

The Lebesgue measure in $\mathbb{R}^d$ has no atoms.Thus, the process is orderly.

Theorem 9 *The generalized Waring Process is a simple point process*

This follows directly from the Proposition 7.2.V, Daley & Vere-Jones (1988), since the generalized Waring Process in $\mathbb{R}^d$ is orderly.

4.1.2 Stationarity, n^{th} order Stationarity and Ergodicity

The Lebesgue measure in $\mathbb{R}^d$ is also invariant under translations, hence the following results can be proved:

Theorem 10 *Let $N(\cdot)$ be a generalized Waring process in $\mathbb{R}^d$ with parameters $a > 0$, $\rho > 0$, $k \in N$ .Then $N(\cdot)$ is stationary.*

Proof We need to prove that for each $u \in \mathbb{R}^d$ and all bounded Borel sets $A \in \mathcal{B}_{\mathbb{R}^d}$, the avoidance function $P_0(\cdot)$ of the generalized Waring process defined above satisfies $P_0(A) = P_0(A+u)$ (see Daley & Vere-Jones (1988), Theorem 10.1.III).

From the invariance of the Lebesgue measure on $\mathbb{R}^d$ one can write

$$P_0(A) = \frac{\rho_{(k\mu(A))}}{(\rho+a)_{(k\mu(A))}} = \frac{\rho_{(k\mu(A+u))}}{(\rho+a)_{(k\mu(A+u))}} = P_0(A+u)$$

which proves the theorem.

A stationary Point process for which the n^{th} order moment measure exists is n^{th} order stationary (see Daley & Vere-Jones (1988)). Hence, using Diggle & Milne the following theorem and its corollary are trivial.

Theorem 11 *The generalized Waring process in $\mathbb{R}^d$ with parameters $a > 0$, $\rho > n$, $k \in N$ is n^{th} order stationary.*

Theorem 12 *The generalized Waring process in $\mathbb{R}^d$ is ergodic*

Proof A necessary and sufficient criteria for a stationary process to be ergodic is to be metrically transitive. From Lemma 2, there exists one and only one root $x > 0$ of the equation $\frac{\Gamma(\rho+x+a)}{\Gamma(\rho+x)}$ $=b > 0$. Let us consider A a set in $\mathbb{R}^d$ and let S_x be the shift operator. If A is such that $P(S_x A \cap A) = P(A)$ then $\frac{\rho_{(k\mu(S_x A \cap A))}}{(\rho+a)_{(k\mu(S_x A\cap A))}} = \frac{\rho_{(k\mu(A))}}{(\rho+a)_{(k\mu(A))}}$. Hence we obtain that $\frac{\Gamma(\rho+k\mu(S_x A\cap A)+a)}{\Gamma(\rho+k\mu(S_x A\cap A))} = \frac{\Gamma(\rho+k\mu(A)+a)}{\Gamma(\rho+k\mu(A))}$and from Lemma 2, it follows that $\mu(S_x A\cap A) = \mu(A)$. The last relation stands if $A = \varnothing$ or $A = \mathbb{R}^d$, which does mean that $P(A) = 0$ or 1.This proves the theorem.

4.2 Special Cases of the Generalized Waring Process

In what follows, we consider three instances of Genelarized Waring Processes that may arise by multivariate extension, marginalization, projection, and limiting arguments. Specifically, we define the multivariate generalized Waring process as a special case of the generalized Waring process on the product space $S \times \{1, 2, ...m\}$ and show that marginals of a multivariate GWP, as well as their sums, are all GWP as well. We then show that generalized Waring processes are closed under projection, and finally demonstrate how negative binomial and Poisson processes can be seen as special (limiting) cases of the GWP as some parameters are allowed to suitably diverge.

4.2.1 The Multivariate Generalized Waring Process

Consider the product space $S\times\{1, 2, ...m\}$ and let $\mathcal{B}_{S\times\{1,2,...m\}}$ be the associated product Borel σ-algebra. Define the function $\nu : \mathcal{B}_{S\times\{1,2,...m\}} \to \mathbb{R}^+$ such that for each $B = \sum_{i=1}^{\infty} A_i \times C_i \in \mathcal{B}_{S\times\{1,2,...m\}}(A_i \in \mathcal{B}_S,\, C_i \in \mathcal{P}(\{1, 2, ...m\}))$, $\nu(B) = \sum_{i=1}^{\infty}\mu(A_i)$ where $\mathcal{B}_S$ is the Borel σ-algebra and $\mu(\cdot)$ some boundedly finite Borel measure. It is clear that $\nu(\cdot)$ is a boundedly finite Borel measure on $S \times \{1, 2, ...m\}$. This allows us to define:

Definition 2 *The generalized Waring process with parameters a, k, ρ and parameter measure $\nu(\cdot)$ on $S \times \{1, 2, ...m\}$ is called the multivariate generalized Waring process with parameters a, k, ρ and parameter measure $\mu(\cdot)$ on S.*

The multivariate GWP satisfies a number of convenient closure properties:

Theorem 13 *Let $N(\cdot)$ be a multivariate generalized Waring process with parameters a, k, ρ and parameter measure $\mu(\cdot)$ on S. Then the following hold:*

1. For every $i \in \{1, 2, ...m\}$, the marginal process $N_i(\cdot) = N(\cdot \times \{i\})$ is a GW process with parameters a, k, ρ and parameter measure $\mu(\cdot)$.

2. $\sum_{j=1}^{l} N_{i_j}(\cdot)$ is a generalized Waring process with parameters a, ρ and parameter measure $kl\mu(\cdot)$.

3. For every finite collection of distinct indices $i_1, i_2, ..., i_l \in \{1, 2, ...m\}$, $\{N_{i_1}(\cdot), N_{i_2}(\cdot), ..., N_{i_l}(\cdot)\}$ is a multivariate generalized Waring process with parameters a, ρ, k and parameter measure $\mu(\cdot)$.

4. $\left\{N_i(\cdot), \sum_{j \neq i} N_j(\cdot)\right\}$ is a bivariate generalized Waring process with parameters a, ρ and parameter measure $k\mu(\cdot), (m-1)k\mu(\cdot)$.

Proof For each bounded Borel set $A \in \mathcal{B}_s$, the joint distribution of $\{N_1(A), N_2(A), ..., N_m(A)\}$ is the MGWD(a; $k\mu(A_1), k\mu(A_2), ..., k\mu(A_m); \rho$). From the structural properties of the multivariate generalized Waring distribution (see Xekalaki (1986)), one has:

1. The distribution of $\{N_i(A) = x_i\}$, for i a given value on $\{1, 2, ...m\}$ is the generalized Waring distribution with parameters a, $k\mu(A), \rho$. By Theorem 5, this is a sufficient condition for the process $N_i(\cdot)$ to be a generalized Waring process.

2. The distribution of $\left\{\sum_{j=1}^{l} N_{i_j}(A) = x_{i_j}\right\}$, is the generalized Waring distribution with parameters a, $kl\mu(A), \rho$. By Theorem 5 this is a sufficient condition for the process $\sum_{j=1}^{l} N_{i_j}(\cdot)$ to be a generalized Waring process.

3. For every $\{A_{i_1}, A_{i_2}, ..., A_{i_l} \in \mathcal{B}_S\}$, let us consider $\{B_1, B_2, ..., B_m \in \mathcal{B}_S\}$ where $B_i = B$ for $i \notin i_1, i_2, ..., i_l$ and $B_i = A_{i_j}$ for $i = i_j$. The joint distribution of $\{N_1(B_1), N_2(B_2), ..., N_m(B_m)\}$ is the MGWD(a; $k\mu(B_1), k\mu(B_2), ..., k\mu(B_m); \rho$). From the structural properties of the multivariate generalized Waring distribution (see Xekalaki (1986)), it follows that the joint distribution of $\{N_{i_1}(A_{i_1}), N_{i_2}(A_{i_2}), ..., N_{i_l}(A_{i_l})\}$ is the MGWD(a; $k\mu(A_{i_1}), k\mu(A_{i_2}), ..., k\mu(A_{i_l}); \rho$) which proves part 3.

4. For every $A, B \in \mathcal{B}_S$ let us consider $\{B_1,\ B_2,\ ..., B_m\ \in \mathcal{B}_S\}$ where $B_i = A$ and $B_j = B$ for $j \neq i$. The joint distribution of $\{N_1(B_1),\ N_2(B_2),\ ...,\ N_m(B_m)\}$ is the

$$\mathrm{MGWD}\left(a;\ k_1\mu(B_1),\ k_2\mu(B_2),\ ...,\ k_m\mu(B_m);\ \rho\right).$$

From the structural properties of the multivariate generalized Waring distribution (see Xekalaki (1986)), it follows that the joint distribution of $\{N_i(A),\ \sum_{j\neq i} N_j(B)\}$ is the BGWD(a; $k\mu(A),\ (m-1)\,k\mu(B);\ \rho$) which proves part 4.

4.2.2 Projections of Generalized Waring Processes

Consider a product measurable space $(S_1 \times S_2,\ \mathcal{B}_{S_1} \otimes \mathcal{B}_{S_2},\ \mu_1 \times \mu_2)$ and let $N(\cdot)$ be a GWP on that space, with parameteres a, k, ρ. Define $N_{S_1}(\cdot)$ and $N_{S_2}(\cdot)$ to be the projections of $N(\cdot)$ onto $(S_1, \mathcal{B}_{S_1}, \mu_1)$ and $(S_2, \mathcal{B}_{S_2}, \mu_2)$, respectively, defined by $N_{S_1}(A) = N(A \times S_1)$ and $N_{S_2}(B) = N(S_2 \times B)$. These projections will also be generalized Waring processes:

Theorem 14 *The projections $N_{S_1}(\cdot)$ an(d $N_{S_2}(\cdot)$ of a GW process $N(\cdot)$with parameteres $a, k, \rho,$, onto the product measurable space $(S_1 \times S_2,\ \mathcal{B}_{S_1} \otimes \mathcal{B}_{S_2},\ \mu_1 \times \mu_2)$ are also GW processes with parameteres a, b, ρ respectively onto $(S_1, \mathcal{B}_{S_1}, \mu_1)$ and $(S_2, \mathcal{B}_{S_2}, \mu_2)$.*

Proof Let $\{A_i \in \mathcal{B}_{S_1}, i = 1, 2, ..., l\}$ be finite family of disjoint bounded Borel sets. The family $\{A_i \times S_1 \in \mathcal{B}_{S_1} \otimes \mathcal{B}_{S_2},\ i = 1,\ 2,\ ...,\ l\}$ is also a finite family of disjoint bounded sets. Hence, the joint distribution of $\{N_{S_1}(A_i) = n_i, i = 1, ..., l\}$ is the MGWD(a, $k\mu(A_1)$, $k\mu(A_2)$, ..., $k\mu(A_l)$; ρ) which proves that the $N_{S_1}(\cdot)$ is a GWP with parameteres a, k, ρ. The same argument yields the result for $N_{S_2}(\cdot)$.

4.2.3 The Poisson and the NB Processes as Limiting Cases of the Generalized Waring Process

We finally turn to demonstrate how negative binomial processes can be obtained as limiting cases of the GWP. Doing so establishes that, even though negative binomial processes cannot be orderly and stationary/ergodic simultaneously, they can be approximated by a process with these properties.

Theorem 15 *Let $N(\cdot)$ be a generalized Waring process with parameters $a > 0$, $\rho > 0$, and $k \in \mathrm{N}$, and with parameter measure $\mu(\cdot)$. Letting $k \to \infty$ and setting $\rho = c \cdot k$ for $c > 0$ a constant, the generalized Waring process converges weakly to a Negative Binomial process with parameters a and c.*

Proof Denote $N_k(\cdot)$, $k > 0$ the generalized Waring process indexed by the parameter k and $N(\cdot)$ the Negative Binomial process with parameters a and c. In order to prove that $N_k(\cdot) \underset{k\to\infty}{\to} N(\cdot)$ weakly, it is sufficient to prove (see e.g. Daley & Vere-Jones (1988), Kallenberg (2002)):

(i). $P(N_k(A) = 0) \underset{k\to\infty}{\to} P(N(A) = 0)$ for all bounded A of a dissecting ring $\mathcal{T}$ of S.

(ii) That the generalized Waring process is uniformly tight.

In order to prove (i) we consider $P(N_k(A) = 0) = \dfrac{\rho_{(k\mu(A))}}{(\rho + \alpha)_{(k\mu(A))}}$.

We calculate:

$$\frac{\rho_{(k\mu(A))}}{(\rho+\alpha)_{(k\mu(A))}} = \frac{\rho_{(a)}}{(\rho + k\mu(A))_{(a)}} = \frac{ck_{(a)}}{(ck + k\mu(A))_{(a)}}$$

$$\cdot \frac{ck(ck+1)\cdot \ldots \cdot (ck+a-1)}{(ck+k\mu(A))(ck+k\mu(A)+1)\cdot \ldots \cdot (ck+k\mu(A)+a-1)}$$

$$= \frac{k^a c\left(c+\frac{1}{k}\right)\cdot \ldots \cdot \left(c+\frac{a-1}{k}\right)}{k^a(c+\mu(A))\left(c+\mu(A)+\frac{1}{k}\right)\cdot \ldots \cdot \left(c+\mu(A)+\frac{a-1}{k}\right)}$$

$$\underset{k\longrightarrow\infty}{\longrightarrow} \frac{c^a}{(c+\mu(A))^a} = P(N(A) = 0)$$

To establish uniform tightness as required in (ii), we use two results concerning regular and tight measures in a complete separable metric space $\mathcal{S}$. A Borel measure is tight if and only if it is compact regular (see e.g Lema A2.2.IV Daley & Vere-Jones (1988)). In turn, a finite, finitely additive, and nonnegative set function defined on the Borel sets of a complete separable metric space $\mathcal{S}$ is compact regular if and only if it is countably additive (see e.g Corollary A2.2.VII Daley & Vere-Jones (1988)). Therefore (ii) follows from the countable additivity theorem (Theorem 3), proven in earlier.

In turn, a Poisson process can be approximated by a negative binomial process, so that it can also be approximated by a GWP:

Theorem 16 *Let $N(\cdot)$ be the limit proces $N(\cdot)$ of the previous Theorem, i.e. a Negative*

Binomial process with parameters a and c. If $c \to \infty$ and $a = \lambda \cdot c$ where $\lambda > 0$ is a constant, $N(\cdot)$ converges weakly to a Poisson process with parameter λ.

Proof Write $\{N_c(\cdot)\}$, $c > 0$, to highlight that the negative binomial process in question is indexed by c. We need to show that there exists a Poisson process $M(\cdot)$ such that: $N_c(\cdot) \underset{c\to\infty}{\to} M(\cdot)$ weakly. Following Daley & Vere-Jones (1988), Lemma 9.I.IV, weak convergence of the process and convergence of finite dimensional (fidi) distributions are equivalent. So, in order to prove that $N_c(\cdot) \underset{c\to\infty}{\to} M(\cdot)$ weakly, it is sufficient to prove that the fidi distributions of $N_c(\cdot)$ converge weakly to those of $M(\cdot)$.

For every $\{A_1, A_2, ..., A_n \in \mathcal{B}_S\}$ we consider the probability generating function G_c (A_1, A_2, ..., A_n; z_1, ..., z_n) of $N_c(\cdot)$ and obtain

$$G_n\ (A_1,\ A_2,\ ...,\ A_n;\ z_1,\ ...,\ z_n) = \frac{1}{c}\left(c + \sum_{i=1}^{n} (1 - z_i)\mu\,(A_i)\right)^{-\lambda\cdot c}$$

But $\frac{1}{c}\left(c + \sum_{i=1}^{n} (1 - z_i)\mu\,(A_i)\right)^{-\lambda\cdot c} \underset{c\to\infty}{\to} \exp\left(-\lambda \sum_{i=1}^{n} (1 - zi)\mu\,(Ai)\right)$, which is the probability generating function G (A_1, A_2, ..., A_n; z_1, ..., z_n) of the Poisson process with parameter λ.

CHAPTER 5

THE GENERALIZED WARING PROCESS IN $\mathbb{R}$

The generalized Waring process on the Real Line is defined in this chapter, both as a special case of the generalized Waring process in $\mathbb{R}^d$ and as a projection of the generalized Waring process in $\mathbb{R}^2$. It is shown to be orderly (and a simple point process as well), stationary, ergodic and nth-order stationary. Further, it is demonstrated that for each finite union of finite intervals $A \in \mathbb{R}$, the distribution of $N(A)$ beeing the $GWD(\ a, k\mu(A)\,;\ \rho\ )$ is a necessary and sufficient condition for $N(\cdot)$ to be a generalized Waring process on $\mathbb{R}$, since it is a simple point process.

The generalized Waring process on the postive half-line $\mathbb{R}^+$ is examined as well. It is proved that the generalized Waring process on the postive half-line $\mathbb{R}^+$ has the Markovian property. The moments, the three additive components of the variance, the individual intensity as well as the transition probabilities and the Chapman-Kolmogorov Equations of this stationary Markov generalized Waring process defined on $\mathbb{R}^+$ are derived. Furthermore, it is demonstrated that the generalized Waring process on $\mathbb{R}^+$ is a regular point process, since the deriving Janossy densities for it, are absolutely continuous with respect to Lebesgue measure. The conditional probabilities and the associated survivor functions, conditional intensity, and the likelihood are derived as well. Finally, it is proved that the Poisson and the Pólya processes can arise as limiting cases of the generalized Waring process on $\mathbb{R}^+$ and by utilizing this result, the moments and transition probabilities are obtained for the Poisson and the Pólya processes.

5.1 Definiton of the Generalized Waring Process in $\mathbb{R}$

Let $\mathbb{R}$ be the real line and let $\mu(\cdot)$ be the Lebesgue measure on the Real Line. The Borel algebra $\mathcal{B}_{\mathbb{R}}$ on the Real Line is the smallest $\sigma - algebra$ on $\mathbb{R}$ which contains all the intervals.

The generalized Waring process on the Real Line can be defined as a special case of the generalized Waring process in $\mathbb{R}^d$ for $d = 1$ or as a projection of the generalized Waring process in $\mathbb{R}^2$ since the Borel algebra $\mathcal{B}_{\mathbb{R}^2}$ in $\mathbb{R}^2$ is equal to $\mathcal{B}_{\mathbb{R}} \otimes \mathcal{B}_{\mathbb{R}}$ (see Theorem 14). In both of the cases, the generalized Waring process on $\mathbb{R}$ can be defined as in the following:

Definition 3 *The generalized Waring process on the Real Line is a process* $\{N(A);\ A \in \mathcal{B}_{\mathbb{R}}\}$ *for which, the joint distribution of* $\{N(A_i) = n_i,\ i = 1,\ ...,\ s\}$ *of every finite family of disjoint bounded Borel sets* $\{A_i \in \mathcal{B}_{\mathbb{R}},\ i = 1,\ ...,\ s\}$ *is* $MGWD$ $(a,\ k\mu(A_1),\ k\mu(A_2),\ ...,\ k\mu(A_s);\ \rho\ )$, $a > 0,\ \rho > 0,\ k > 0$.

Thus, the generalized Waring process on the Real Line is orderly (and a simple point process as well), stationary, ergodic and nth-order stationary.

Theorem 17 *Let* $N(\cdot)$ *be a simple point process on* $\mathbb{R}$. *For* $N(\cdot)$ *to be a generalized Waring process* $a > 0,\ \rho > 0,\ k > 0$ *it is necessary and sufficient that for each finite union of finite intervals* A *the distribution of* $N(A)$ *is the* GWD $(a,\ k\mu(A);\ \rho\ )$.

Proof *Necessity* Let $N(\cdot)$ be a generalized Waring process and A finite union of finite intervals. Clearly, $A \in B_R$,then from the Theorem 5 the distribution of $N(A)$ is the GWD $(a,\ k\mu(A);\ \rho\ )$.

Sufficiency For each $n \in \mathbb{Z}$ and $i \in \{1,\ 2,\ ...,\ n\}$, we describe $I_{ni} = (x_{ni},\ y_{ni}]$ where $x_{ni} = n \cdot 2^{-i}$, $y_{ni} = (n+1) \cdot 2^{-i}$ as $i-$interval and let us denote $\mathcal{T}_n = \{\ I_{ni} : i = 1,\ ...,\ n\}$ the set of $i-$intervals for a given n. Then $\mathcal{T} = \{\mathcal{T}_n,\ n \in \mathbb{Z}\} = \{\ \{\ I_{ni} : i = 1,\ ...,\ n\ \},\ n \in \mathbb{Z}\}$ forms a dissection system of $\mathbb{R}$. The dissecting ring generated by finitely many intersections and unions of elements of this dissecting system is the family of all finite union of finite intervals in $\mathbb{R}$. Hence from the Theorem 5 for $N(\cdot)$ to be a generalized Waring process is sufficient that $N(A)$ follows a generalized Waring distribution with parameters $a > 0,\ \rho > 0,\ k > 0$ for each A from this family.

Let us consider now the generalized Waring process on the postive half-line $\mathbb{R}^+$. The Borel algebra $\mathcal{B}_{\mathbb{R}+}$ is the smallest $\sigma - algebra$ which contains all the intervals $(x, y]$, $x, y \in$

$\mathbb{R}^+$. In this case, the process can be taken as modelling the occurrences of some phenomenon at the time epoches $\{t, t \in \mathbb{R}^+\}$. In this case we denote for any $t > 0$, $N(0,t) = N(t)$. and then $p(N(0) = 0) = 1$, almost surely.

Theorem 18 *The generalized Waring process on the postive half-line $\mathbb{R}^+$ has the Markovian property.*

Proof Let us consider the disjoint intervals $\{ (t,\, t+h),\, (s,\, t),\ (0,\, s), 0 \leq s < t\}$. The joint distribution of $\{ N(t,\, t+h),\, N(s,t),\, N(0,s) \}$ is the $MGWD(\ a;\, kh,\, k(h-s),\, ks;\ \rho\ )$ while for the intervals $\{(t,\, t+h),\, (0,\, t)$, the joint distribution of $\{ N(t, t+h), N(0,t) \}$ is the $MGWD\ (a;\, kh,\, kt;\, \rho)$.

We refer now to the third structural property of the multivariate generalized Waring distribution proved by Xekalaki (1986) (p.1054) and we find that $(N(t,t+h) \mid N(s,t),$ $N(s))$ ˜ $UGWD\ (a+n(t);\, kh;\, \rho+kt)$ and $(N(t,t+h) \mid N(t)$ ˜ $UGWD\ (a+n(t);\, kh;$ $\rho+kt)$, where $n(t)$ is the value of $N(t)$, which proves that the generalized Waring process has the Markovian property, i.e. the conditional distribution of the future $N\ (t,\, t+h)$ given the present state $N(t)$ and the past $N(s),\ 0 \leq s \leq t$ dependsonly on the present and is independent of the past.

From the the definiton of the generalized Waring process and the above theorem we can conclude that a generalized Waring process with parameters $(a,\, k,\, \rho)$, $a > 0,\ k > 0,\ \rho > 0$ on the postive half-line $\mathbb{R}^+$ is a Markov point process that starts at zero and has stationary increments $N(t+h)\ - N(t)$ following $GW\ (a,\, kh,\, \rho)$ distribution for each $h > 0,\ t \geq 0$. These three conditions are underlined in the definiton of the generalized Waring process on $\mathbb{R}^+$ given by Zografi and Xekalaki (2001) (see also, Xekalaki and Zografi (2008)) as in the following:

Definition 4 *The counting process $\{N(t),\, t \geq 0\}$ is said to be a generalized Waring process with parameters $(a,\, k,\, \rho)$, $a > 0,\ k > 0,\ \rho > 0$ if (I) $p\{N(0) = 0\} = 1$, (II) $N(t)$ is a Markov process, (III) $N(t+h)\ - N(t)$ is $GW\ (a,\, kh,\, \rho)$-distributed for each $h > 0,\ t \geq 0$.*

5.2 The Moments, Intensity and Individual Intensity

Let $N(t)$ define a generalized Waring process with parameters $(a,\ k,\ \rho)$. Then, for any t,

$$E\left[N\left(t\right)\right] = \frac{akt}{\rho - 1},$$

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

Following Irwin (1975), one may show that the variance can be divided into three additive components, thus

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

where

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

is the component due to liability

$$\sigma^2_{\nu} = a\left(a + \rho - 1\right)\left(\rho - 1\right)^{-2}\left(\rho - 2\right)^{-1}$$

is the component due to proneness and

$$\sigma^2_R = akt\left(\rho - 1\right)^{-1}$$

is the component due to randomness.

The generalized Waring process is a stationary process. For a stationary process N, $E\left[N\left(t\right)\right] = \eta \cdot t$, where η is termed the intensity of N (see e.g. Grandell (1997), p.53). It is clear that the intensity of the generalized Waring process is $\eta = \frac{ak}{\rho - 1}$. For this process (like for all stationary processes), there always exists, a random variable $\bar{N}$ with $E\left(\bar{N}\right) = \eta$, called the individual intensity, such that $\frac{N\left(t\right)}{t} \xrightarrow[t\to\infty]{p} \bar{N}$ (see, e.g. Grandell (1997), p.53). The intensity η is finite. Hence, it follows that the individual intensity $\bar{N}$ is finite with probability 1.

In order to find it we consider that if $N\left(\cdot\right)$ be a generalized Waring process then for each $t \in \mathbb{R}^{+}$, the conditional distribution of $N\left(t\right)|\nu$ is a $NB\left(kt\ ,\ \frac{1}{1+\nu}\right)$ distribution with ν following a $Beta\left(\alpha,\rho\right)$distribution of the second kind (see e.g. Irwin (1975), Xekalaki (1981) , Xekalaki (1983b)). for every t, while $N\left(t\right)$ can be obtained in this case as a mixture of $NB\left(kt\ ,\ \frac{1}{1+\nu}\right)$ with ν. In order to remark this fact we denote $N\left(t\right) = \tilde{N}\left(kt,\ \frac{1}{1+\nu}\right)$ for each t. Then the following can be proved:

Theorem 19 *Let $N\left(\cdot\right)$ be a generalized Waring process on positive half-line $\mathbb{R}^{+}$. Then,* $\frac{1}{t}N\left(t\right) \xrightarrow[t\to\infty]{p} \nu k$.

Proof

$$\lim_{t\to\infty}\frac{1}{t}N\left(t\right) = \nu k\lim_{t\to\infty}\frac{N\left(t\right)}{\nu kt} = \nu k\lim_{t\to\infty}\frac{\tilde{N}\left(kt, \frac{1}{1+\nu}\right)}{\nu kt} \quad .$$

Taking into account that $E\left[\tilde{N}\left(kt, \frac{1}{1+\nu}\right)\right] = \nu kt$ and $var\left\{\frac{\tilde{N}\left(kt, \frac{1}{1+\nu}\right)}{\nu kt}\right\} = \frac{1+\nu}{\nu kt} \xrightarrow[t\to\infty]{} 0$, and using Chebyshev's inequality, we have that $\frac{\tilde{N}\left(kt, \frac{1}{1+\nu}\right)}{\nu kt} \xrightarrow[t\to\infty]{p} 1$, which implies that $\frac{1}{t}N\left(t\right) \xrightarrow[t\to\infty]{p} \upsilon k$.

Corollary 2 *The random variable $\bar{N} = \nu k$, where ν is random variable following a $Beta\left(\alpha,\rho\right)$ distribution of the second kind, is the individual intensity of the generalized Waring process on $\mathbb{R}^{+}$ with parameters $\left(a,\ k,\ \rho\right)$.*

Proof Since ν is $betaII\left(a,\rho\right)$-distributed, $E\left(\nu\right) = \frac{a}{\rho-1}$ i.e. $E\left(\bar{N}\right) = \frac{ak}{\rho-1}$. Hence, the random variable $\bar{N} = \nu k$ is the individual intensity of the generalized Waring process.

5.3 Transition Probabilities & the Chapman - Kolmogorov Equations of the Generalized Waring Process

Since the generalized Waring process on $\mathbb{R}^{+}$ is a Markov process, the transition probabilities can be derived:

$$\begin{aligned} p_{m,n}(t,t+h) &= P\{N(t+h)=n|N(t)=m\} \\ &= P\{N(t+h)-N(t)=n-m|N(t)=m\} \\ &= \frac{(\rho+kt)_{(kh)}}{(\rho+kt+a+m)_{(kh)}}\frac{(a+m)_{(n)}(kh)_{(n)}}{(\rho+kt+a+m+kh)_{(n)}}\frac{1}{n!} \end{aligned}$$

The above probability $p_{m,n}(t,t+h)$, represents the probability that a process presently in state m will be in state n a later time h. This probability in this case depends on the present time so the defined generalized Waring process on the Real line is a non-homogenous Markov process.

It is clear that

$$p_{0,n}(0,t)=P(N(t)=n|N(0)=0)=P(N(t)=n)=p_n(t).$$

The transition probabilities satisfy the Chapman-Kolmogorov equations, i.e.

$$p_{m,n}(s,t)=\sum_{i=m}^{n}p_{m,i}(s,\tau)\,p_{i,n}(\tau,t), \qquad for\ s\le\tau\le t,\ m\le n. \tag{28}$$

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 $s\le\tau\le t$, $m\le n$, $h\ge 0$, we obtain

$$\frac{\partial}{\partial T}p_{m,\,n}(s,t)=\sum_{i=m}^{n}p_{m,\,i}(s,\ t)\lim_{h\to 0}\frac{p_{i,\,n}(t,\ t+h)}{h}-\lim_{h\to 0}\left(1-\frac{p_{nn}(t,\ t+h)}{h}\right)p_{m,\ n}(s,\ t),$$

$$\lim_{h\to 0}\frac{p_{i,\,n}(t,\ t+h)}{h}=\begin{cases} q_{n-1,\,n}(t)=\frac{k(a+n-1)}{(a+\rho+kt+n-1)}, & n-i=1 \\ q_{i,\,n}(t)\,\frac{\Gamma(a+n)}{\Gamma(a+i)}\frac{k}{(n-i)(n-i-1)}\frac{(\rho+kt)_{(a+i)}}{(\rho+kt)_{(a+n)}}, & n-i>1 \end{cases}$$

and

$$\lim_{h\to 0}\left(\frac{1-p_{n,\,n}(t,t+h)}{h}\right)=\nu_n(t)=k\cdot\sum_{i=0}^{a+n-1}\frac{1}{\rho+kt+i}.$$

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

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

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

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

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

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

where

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

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

and

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

5.4 Conditional intensity and Likelihood of the Generalized Waring Process on the Real Line

In this section, we derive the conditional intensitiy function and the Likelihood of the generalized Waring process, needed for simulation as well as statistical analysis of the different dataset tha can be modelled by a generalized Waring process on $\mathbb{R}^+$. In subsection 2.4.1, a prove that the generalized Waring process is conditionally bound (see Lemma 1) and a simulation algorithm that can be used in this case, are presented. We use these results in the sequel, to simulate data from the generalized Waring process. These derivations are based on the regularity property of the generalized Waring process. If this is the case, the conditional intensity function is defined picewise by the hazard functions, i.e. from the conditional densities and the survival functions, while the Likelihood is nothing other than a Janossy density.

As is known in literature (e.g. Daley and Vere-Jones (1988), Proposition 13.I.IV.), we can prove the regularity property of a point process on $\mathbb{R}^+$, if we prove that there exists a uniquely determined family of conditional probabilities $p_n(t|\, t_1,\ t_2,\ ...,\ t_{n-1})$ and associated survivor functions $S_n(t\,|t_1,\ t_2,\ ...,\ t_{n-1}) = 1 - \int_{t_{n-1}}^{t} p_n(u|\, t_1,\ t_2,\ ...,\ t_{n-1})\,du$, $(t > \ t_{n-1})$ defined on $0 < t_1 < t_2 < ... < t_{n-1} < t$, for it.

We start from the survivor functions. Denote $\tau_i = t_i - t_{i-1}$, $i \geq 1$, and for the generalized Waring process with parameters $(a,\ k,\ \rho)$ on $\mathbb{R}^+$ we obtain:

$$S_1(t) = \Pr(\tau_1 > t) = \ p(N(t) = 0) = \frac{\rho_{(a)}}{(\rho + kt)_{(a)}}$$

and

$$\begin{aligned} S_n(t\,|t_1,\ t_2,\ ...,\ t_{n-1}) &= \Pr(\tau_n > t\,|t_1,\ t_2,\ ...,\ t_{n-1}) \\ &= p(\,N(t + t_{n-1}) = \ 1 \mid N\left(t_{t_{n-1}}\right) = 1) \\ &= \frac{(\rho + kt_{n-1})_{(a+n-1)}}{(\rho + k(t + t_{n-1}))_{(a+n-1)}} \text{ for each } n \geq 2 \end{aligned}$$

and for the conditional probabilities

$$p_1(t) = \frac{d}{dt} S_1(t) = \frac{k(\rho)_{(a)}}{(\rho + kt)_{(a)}} \left[\Psi(\rho + kt + a) - \Psi(\rho)\right],$$

and

$$\begin{aligned} p_n(t\,|t_1,\ t_2,\ ...,\ t_{n-1}) &= \frac{d}{dt} S_n(t\,|t_1,\ t_2,\ ...,\ t_{n-1}) \\ &= \frac{(\rho + kt_{n-1})_{(a+n-1)}}{(\rho + k(t + t_{n-1}))_{(a+n-1)}} \\ &\quad \left[\Psi(\rho + k(t + t_{n-1}) + a + n - 1) - \Psi(\rho + k(t + t_{n-1}))\right] \end{aligned}$$

Deriving then, the hazard functions we obtain

$$h_1(t) = \frac{p_1(t)}{S_1(t)} = k[\Psi(\rho + kt + a) - \ \Psi(\rho + kt)]$$

and

$$\begin{aligned} h_n(t\,|t_1,\ t_2,\ ...,\ t_{n-1}) &= \frac{p_n(t\,|t_1,\ t_2,\ ...,\ t_{n-1})}{S_n(u\,|t_1,\ t_2,\ ...,\ t_{n-1})} \\ &= k[\Psi(\rho + k(t + t_{n-1}) + a + n - 1) - \Psi(\rho + k(t + t_{n-1}))]. \end{aligned}$$

Hence the conditional intensity function is

$$\lambda^*(t) = \begin{cases} k\left[\Psi(\rho+kt+a)-\Psi(\rho+kt)\right] \\ \hfill (0<t\le t_1) \\ k\left[\Psi(\rho+k(t+t_{n-1})+a+n-1)-\Psi(\rho+k(t+t_{n-1}))\right] \\ \hfill (t_{n-1}<t\le t_n,\ n\ge 2) \end{cases} \tag{29}$$

The same property suggests a method for evaluating the process likelihood by calculating the local Janossy densities $j_n(t_1,\ t_2,\ ...,\ t_n\,|T)$ for all finite intervals $[0,T]$ with $T>0$, which in the case of a regular point process exist and are absolutely continuous with respect to Lebesgue measure. We use it to obtain

$$J_0(T) = S_1(T) = \frac{\rho_{(a)}}{(\rho+kT)_{(a)}}$$

and

$$\begin{aligned} j_n(t_1,\ t_2,\ ...,\ t_n\,|T) &= \frac{(\rho+kt_n)_{(a+n)}}{(\rho+k(T+t_n))_{(a+n)}} p_n(t_l\,|t_1,\ t_2,\ ...,\ t_{n-1}) \\ &= \frac{(\rho+kt_n)_{(a+n)}}{(\rho+k(T+t_n))_{(a+n)}} \frac{(\rho+kt_{n-1})_{(a+n-1)}}{(\rho+k(t_n+t_{n-1}))_{(a+n-1)}} \\ &\quad \left[\Psi(\rho+k(t+t_{n-1})+a+n-1)-\Psi(\rho+k(t+t_{n-1}))\right] \\ &\hfill n\ge 2. \end{aligned} \tag{30}$$

The likelihood L of a realization $t_1,\ t_2,\ ...,\ t_n$ of a regular point process on $\mathbb{R}^+$ over the interval $[0,T]$ for any$0<T<\infty$ is the local Janossy density $j_n(t_1,\ t_2,\ ...,\ t_n\,|T)$. Hence, the likelihood L of the generalized Waring process with parameters $(a,\ k,\ \rho)$ on $\mathbb{R}^+$ over the interval $[0,T]$ is expressed by relationship (30).

5.5 Limiting Cases of the Generalized Waring Process on the Real Line

In this section we provide two theorems which follows respectively by the Theorems 15 and 16 applied on the Real Line where the Lebesgue measure is the parameter measure which provide that the Poisson and the Pólya processes are limiting cases of the generalized Waring process (in the sense of weak convergence).

Theorem 20 *If* $\rho = c{\cdot}k$*, where* $c > 0$ *is a constant, the generalized Waring process tends to a Pólya process with parameters* a *and* $1/c$*, i.e.* $\{N_k(t)\} \xrightarrow[k\to\infty]{d} \{N_c(t)\}$*, where* $\{N_k(t),\ t \geq 0\}$ *is the generalized Waring process indexed by the parameterk* > 0 *and* $\{N_c(t),\ t \geq 0\}$ *is the Polya process indexed by the parameter* $c > 0$ *and with* $P\{N_c(t) = n\} = \binom{a+n-1}{n} \left(\frac{c}{t+c}\right)^{\alpha} \left(\frac{t}{t+c}\right)^{n}, \quad n = 0, 1, ...$

Theorem 21 *Consider now the Pólya process* $\{N_c(t),\ t \geq 0\}$ *defined as in the previous theorem. Then, if* $a = \lambda \cdot c$*, where* $\lambda > 0$ *is a constant,* $\{N_c(t)\} \xrightarrow[c\to\infty]{d} \{N(t)\}$*, where* $\{N(t),\ t \geq 0\}$ *is a homogeneous Poisson process with* $P\{N(t) = n\} = \frac{(\lambda t)^n \exp(-\lambda t)}{n!}$*,* $n = 0, 1, ...$

The results of these theorems tell us that the the Pólya and the Poisson processes are limiting forms of the generalized Waring process. Thus, utilizing the results holding for the generalized Waring process, one may obtain the following results for a Pólya process$\{X(t),\ t \geq 0\}$ with parameters $(a, 1/c)$and for a Poisson process $\{Y(t),\ t \geq 0\}$ with parameter λ defined as in Theorems 20 and 21, respectively.:

- For any $t \geq 0$, $E[X(t)] = \frac{a}{c}t$, $Var[X(t)] = \frac{a}{c}t + \frac{a}{c^2}t^2$ 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:

$$P(X(t+h) \mid X(t) = m) = \begin{cases} \left(\frac{c+t}{c+t+h}\right)^{(a+n)} & n = m \\ (a+m)\, \frac{h(c+t)^{(a+m)}}{(c+t+h)^{(a+n1)}} & n = m+1 \\ \frac{\Gamma(a+n)}{\Gamma(a+m)(n-m)!} \frac{h^{(n-m)}(c+t)^{(a+m)}}{(c+t+h)^{(a+n)}} & n > m+1 \end{cases}$$

and

$$P(Y(t+h) \mid Y(t) = m) = \begin{cases} \exp(-\lambda h) & n = m \\ \lambda h \exp(-\lambda h) & n = m+1 \\ \frac{(\lambda h)^{(n-m)}}{(n-m)!} \exp(-\lambda h) & n > m+1 \end{cases}$$

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

CHAPTER 6

CONCLUSION

6.1 Discussion

In this thesis we developed a theory of the generazed Waring process starting with defining it on the real line as a stationary, non-homogeneous Markov process. The moments, the three additive components of the variance, the individual intensity as well as the transition probabilities and the Chapman-Kolmogorov Equations of this stationary Markov generalized Waring process defined on $\mathbb{R}^+$ have been derived. It has also been demonstrated that the generalized Waring process on $\mathbb{R}^+$ is a regular point process. In addition, the generalized Waring process has been defined and studied in an accident theory context. It has been generated starting with a process of negative binomial form, but different from a Pólya process, mixing it with a beta distribution of the second type (beta II). Further, an alternative genesis scheme referring to Cresswell and Froggatt's (1963) spells model has been proposed in the framework considered by Xekalaki (1983b). In addition, some inferential aspects connected with the mixed negative binomial derivation of the generalized Waring process have been discussed. An application in a web access modeling context has been provided, too.

We then turned our attention to the case of spatial overdispersion, where point processes models are required with finite dimensional distributions that are overdispersed relative to the Poisson distribution.

Fitting such models usually heavily relies on the properties of stationarity, ergodicity, and orderliness and though processes based on negative binomial finite dimensional distributions have been widely considered, they typically fail to simultaneously satisfy the three required properties for fitting. Indeed, as has been conjectured by Diggle & Milne (1983), no negative binomial model can satisfy all three properties. In light of this, in the thesis, we changed perspective, and constructed a new process based on a different overdisperse count model, the Generalized Waring Distribution. While comparably tractable and flexible to negative binomial processes, the Generalized Waring process was shown to possess all required properties, and additionally span the negative binomial and Poisson processes as limiting cases. In this sense, the GW process provides an approximate resolution to Diggle & Milne's conundrum. In particular, we have been able to define a new spatial point process for phenomena characterised by over-dispersion, in great generality. It has been shown , specifically, that the notion of Generalized Waring (Point) Process generalizes naturally to a point process defined over $\mathbb{R}^d$ and even more a metric space.

It was also shown that such a process exists, and is able to simultaneously satisfy the properties that negative binomial processes fail to satisfy (orderliness, stationarity, and ergodicity).

Moreover, we have demonstrated that the new process features appealing closure properties, in the sense that projection, marginalization, and superposition all yield processes of the same GWD type, with easily determinable parameters. These properties offer advantages relative to existing competitors of the negative binomial type, both from the theoretical and the practical viewpoints, especially in terms of fitting the process on the basis of a single realization. In addition, they allow looking at overdispersion from a new perspective compared to that of existing ones, which are not of negative binomial type and have widely been used to model spatial clustering patterns (e.g. Neyman - Scott Processes) as they offer a meaningful approach in terms of parameter interpretability to understanding the mechanism underpinning the discrepancy between nominal and observed variance.

In addition, we have demonstrated how negative binomial and Poisson processes can be obtained as limiting cases of the generalized Waring process, thus giving a positive resolution to the quandary posed in the conclusion of the paper by Diggle and Milne: "Any view we adopt seems to fall in a situation from which progress looks difficult, and we conjecture that no stationary, ergodic, orderly negative binomial processes exist." So, though negative binomial processes may fail to simultaneously verify orderliness, stationarity and ergodicity, they can be well approximated by flexible and tractable processes of the GWP class that do verify these properties.

Finally, by means of a key conditional property revealing the structure of the generalized Waring process, the process was demonstrated straightforward to simulate. Interestingly, as revealed by the algorithm, using the generalized Waring process one is enabled not only to obtain a description of the counts, but also the exact locations where spatial clustering occurs, in contrast to what is permitted by the use of the negative binomial process. Some simulation results have been displayed as well. A further comparison of the generalized Waring process to the negative binomial and Poisson process on simulated data could possibly provide a deeper insight into how spatial overdispersion could be best approached.

6.2 Scope for Further Research

Potential further advantages of the generalized Waring Process relative to negative binomial processes may arise in the context of compounding (or clustering) and mixing (or heterogeneity). In particular, Cane (1974,1977) has demonstrated that one cannot distinguish between compounding and heterogeneity under a negative binomial distribution: given a total of n events, the distribution of event times is the same, whether the model arose out of mixing or compounding. In contrast, Xekalaki (1983b) demonstrated that discriminating between clustering and mixing may well be possible in the context of the Generalized Waring Distribution, by showing that the conditional distribution of the times of events given their total number is different under compounding and under mixing (see also Xekalaki (2006, 2014, 2015). This property can then be used in order to distinguish clustering, which may otherwise be confounded with compounding.

Much of spatial statistical research to date, addresses the impact of spatial autocorrelation (SA) on parameter estimates and autoregressive negative binomial models have been successfully employed to handle these situations (see, for example, Benjamin, G. J. et al. (2008)) and, as is known, these models can accommodate only negative autocorrelation. Developing an equivalent auto Generalized Waring model to handle spatial autocorrelation and a comparison with the modeling strategies based on the use of autoregressive negative binomial models would potentially contribute to a better understanding of the mechanism causing spatial overdispersion and could be the subject of future work.

REFERENCES

[1] Ajiferuke, I.,Wolfram, D. and Xie, H. (2004). Modelling website visitation and resource usage characteristics by IP address data. In: H. Julien and S. Thompson (eds), Proceedings of the 32nd Annual Conference of the Canadian Association for Information Science (2004), Available at: www.cais-acsi.ca/proceedings/2004/ajiferuke_2004.pdf

[2] Bai-ni Guo, Ying-jie Zhang, Feng Qi (2008), Refinements and Sharpenings of Some double Inequalities for Bounding the Gamma Function, Inequalities for the gamma function, J. Inequal. Pure Appl. Math., Volume 9 Issue 1, Article 17, 5 pp.

[3] Bates G.E. (1955). Joint distributions of time intervals for the ovurrence of successive accidents in a generalized Polya scheme. Annals of Mathematical Statistics 26, 705-720

[4] Cane, V. R. (1974). The concepts of accident proneness. Bulletin of Institute of Mathematics, Bulgarian Academy of sciences 15, 183-188

[5] Cane, V.R. (1977). A class of non-identifiable stochastic models. J. Appl. Probab. 14, 475-782

[6] Chatfield, C. 7and Theobald, C.M. (1973). Mixtures and Random Sums, The Statistician, 22(3), 281-287

[7] Chetwynd, A. G. & Diggle, P. J. (1998). On estimating the Reduced Second Moment Measure of a Stationary spatial Point Process, Austral. & New Zealand J.Statist. 40(1), 1998, 11-15

[8] Cox, D.R. & Lewis, P.A.W. (1966). The Statistical Analysis of Series of Events. London: Methuen

[9] Cox, D.R., Isham, V.I. (1980) Point Processes. Chapman & Hall. ISBN 0-412-21910-7

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

[11] Daley, D.J. & Vere-Jones, D. (1972). A summary of the mathematical theory of point processes. In: Stoch. Point Proc. (eds P.A.W. Lewis). 299-383

[12] Daley, D.J. & Vere-Jones, D. (1988) An Introduction to the Theory of Point Processes, Springer

[13] Diggle, P.J. & Chetwynd, A.G. (1991). Second- order Analysis of Spatial Clustering for Inhomogeneous Populations. Biometrics, 47, 1155-1163

[14] Diggle, P.J. & Milne R.K. (1983). Negative Binomial Quadrat Counts and Point Processes, Scandinavian J. of Statist. Vol. 10(4), 257-267

[15] Feller, W. (1968). An introduction to probability theory and its applications. Vol. I, 3rd ed. Wiley, New York

[16] Fisher, L. (1972). A survey of the mathematical theory of multidimensional point processes. In: Stochastic Point Processes: Statistical Analisys, Theory, and Applications (eds P. A. W. Lewis), 468-51, New York: Wiley-Interscience

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

[18] Gregoire, G. (1983). Negative binomial distributions for point processes, Stochastic Process. Appl. 16, 179-188

[19] Gurland, J. (1959). Some applications of the negative binomial and other contagious distributions. American Journal of Public Health 49, 1388-1399

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

[21] Irwin, J.O. (1963). The place of mathematics in medical and biological statistics. Journal of the Royal Statistical Society, Series A 126, 1-44

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

[23] Irwin, J.O. (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)

[24] Janardan, K. G. and Patil, G. P. (1972).A unified approach for a class of multivariate hypergeometric models. Sankhya Series A, 34, 363-376

[25] Johnson, N. L., Kotz, S., & Kemp, A. W. (1992). Univariate Discrete Distributions. John Wiley and Sons, Inc., New York

[26] Kallenberg, O. (2002) Foundations of Modern Probability, Springer

[27] Kemp, A. W., Kemp, C. D. (1975) Models for Gaussian hypergeometric distribution. In: G. P. Patil, S. Kotz, J. K. Ord (eds.). Statistical Distributions in Scientific Work. Dodrecht: Reidel 1, 31-40

[28] Kemp, C. D. (1970). 'Accident proneness' and discrete distribution theory. In Random counts in Scientific Work, Vol. 2. G.P. Patil (ed.) 41-65. State College: Pennsilvania State University Press

[29] Matern, B. (1971). Doubly stochastic Poisson processes in the plane. In: Statistical Ecology, (eds. G.P. Patil, E.C. Pielou & W.E. Waters), 195–213. University Park: Pennsylvania

[30] Milan Merkle (2001), Conditions for Convexity of a Derivative and Applications to the Gamma and Digamma Function, Facta Universitatis, Ser. Math. Inform.16 13-20

[31] Muller, J. and Waagepetersen, R.P. (2004) Statistical Inference and Simulation for Spatial Point Processes. Chapman and Hall/CRC, Boca Raton

[32] 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

[33] Neyman J., Scott L. E. (1958). Statistical Approach to Problems of Cosmology, J. Roy. Statist. Soc.,20, 1-43

[34] Ramakrishnan, A. (1951). Some simple stochastic processes, J. Roy. Statist. Soc. Ser. B 13(1), 131-140

[35] Ripley, B. D. (1988) Statistical Inference for Spatial Processes, Cambridge University Press

[36] Ross, S.M. (1995) Stochastic Processes. Wiley. ISBN 978-0471120629

[37] Ross, S.M. (2007) Introduction to probability models Amsterdam ; Boston : Elsevier/Academic Press, c2007. 9th ed.

[38] Shaw, L. and Sichel, H. S. (1971). Accident Proneness. Oxford: Pergamon Press

[39] Sibuya, M. (1979). Generalized hypergeometric, digamma and trigamma distributions. Annals of the Institute of Statistical Mathematics, 31, Part A:373–390

[40] Sibuya, M. and Shimizu, R. (1981). Classification of the generalized hypergeometric family of distributions. KEIO Science and Technology Reports, 34:1–38

[41] Snyder, D.L., Miller, M.I. (1991) Random Point Processes in Time and Space. Springer-Verlag. ISBN 0-387-97577-2

[42] Westcott, M. (1971). On existence and mixing results for cluster point processes. J. Roy. Statist. Soc. Ser. B 33, 290-300

[43] Westcott, M. (1972). The probability generating functional. J. Aust. Math. Soc. 14, 448-466

[44] Willmot, G. E. (1986) Mixed compound distributions. ASTIN Bulletin 16/S, 59-S79

[45] Xekalaki E., Panaretos J. (1983). Identifiability of Compound Poisson Distributions". Scandinavian Actuarial Journal, 66, 39-45

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

[47] 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

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

[49] Xekalaki, E. (1983c). A Property of the Yule Distribution and its Applications. Communications in Statistics, Part A (Theory and Methods). A12, No. 10, 1181-1189

[50] Xekalaki, E.(1983d) Hazard Functions and Life Distributions in Discrete Time. Communications in Statistics, Part A, (Theory and Methods), 12 (21), 2503-2509

[51] 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

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

[53] Xekalaki, E.(1985b). Some Identifiability Problems Involving Generalized Waring Distributions. Publicationes Mathematicae, 32, 1985, 75-84

[54] Xekalaki, E. (1986). The Multivariate Generalized Waring Distribution. Comm. Statist. Theory Methods,15(3) 1047-1064

[55] Xekalaki, E. (2014) On The Distribution Theory of Over-Dispersion. J. Statist. Distr. & Applications, 1(19), on line version DOI: 10.1186/s40488-014-0019-z

[56] Xekalaki, E. (2015). Under- and Overdispersion. Wiley StatsRef: Statistics Reference Online. 1–9. in: Encyclopedia of Actuarial Science. (Update based on the 2006 original article by E.Xekalaki) Available at: http://onlinelibrary.wiley.com/doi/10.1002/9781118445112.stat04407.pub2

[57] Xekalaki, E., Zografi, M. (2008). The Generalized Waring Process and its Applications – Communications in Statistics A (Theory and Methods) 37(12), 1835-1854

[58] Zografi, M. Xekalaki, E. (2001). The Generalized Waring Process. Proceedings of the 5th Hellenic-European Conference on Computer Mathematics and its Applications, Athens, Greece, E.A. Lypitakis (ed.), 886-893

[59] Zografi, M. Xekalaki, E. (2019). Modeling spatial overdispersion via the generalized Waring point process, Scandinavian Journal of Statistics. Volume 46, Isue 4, 1098-1116