https://doi.org/10.26219/heal.aueb.5966

ATHENS UNIVERSITY
OF ECONOMICS AND BUSINESS

DEPARTMENT OF STATISTICS

POSTGRADUATE PROGRAM

GENERALIZED WARING PROCESS

By

Mimoza Zografi

A THESIS
Submitted to the Department of Statistics
of the Athens University of Economics and Business
in partial fulfilment of the requirements for
the degree of Master of Science in Statistics

Athens, Greece
2003

ISBN: 960-8287-28-6

https://doi.org/10.26219/heal.aueb.5966

ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

DEPARTMENT OF STATISTICS

POSTGRADUATE PROGRAM

GENERALIZED WARING PROCESS

By

Mimoza Zografi

A THESIS
Submitted to the Department of Statistics
of the Athens University of Economics and Business
in partial fulfilment of the requirements for
the degree of Master of Science in Statistics

Athens, Greece
2003

ISBN: 960-8287-28-6

https://doi.org/10.26219/heal.aueb.5966

ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

DEPARTMENT OF STATISTICS

POSTGRADUATE PROGRAM

GENERALIZED WARING PROCESS

By

Mimoza Zografi

A THESIS

Submitted to the Department of Statistics
of the Athens University of Economics and Business
in partial fulfillment of the requirements for
the degree of Master of Science in Statistics

Athens, Greece
September, 2002

ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ

ΓΕΝΙΚΕΥΜΕΝΗ ΑΝΕΛΥΞΗ WARING

Μιμόζα Ζωγράφη

ΔΙΑΤΡΙΒΗ

Που υποβλήθηκε στο Τμήμα Στατιστικής
του Οικονομικού Πανεπιστημίου Αθηνών
ως μέρος των απαιτήσεων για την απόκτηση
Μεταπτυχιακού Διπλώματος Ειδίκευσης στη Στατιστική

Αθήνα
Σεπτέμβριος 2002

https://doi.org/10.26219/heal.aueb.5966

ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

DEPARTMENT OF STATISTICS

A Thesis submitted in partial fulfillment of
the requirements for the degree of
Master of Science

Generalized Waring Process

Mimoza Zografi

Supervisor:
E. Xekalaki
Professor

External examiner:
Professor Ildar A. Ibragimov
Department of Staklov
Institute of Mathematics,
St. Petersburg,
Russia

Approved by the Graduate Committee
Athens, October 2003

Michael Zazanis, Associate Professor
Director of the Graduate Program

https://doi.org/10.26219/heal.aueb.5966

DEDICATION

To my husband and my son.

https://doi.org/10.26219/heal.aueb.5966

ACKNOWLEDGEMENTS

I gratefully acknowledge the ideas, the instructions, the suggestions and many valuable comments from my supervisor Dr. Professor Evdokia Xekalaki.

https://doi.org/10.26219/heal.aueb.5966

https://doi.org/10.26219/heal.aueb.5966

VITA

Name	:	Spata – Zografi Mimoza
Nationality	:	Greek
Marital Status	:	Married
Birth date	:	22.10.1963

Education

1978-1982: Graduated from General Middle School with Excellent results. I am two times winner of the National Mathematics Olympiads.

1982-1987: BS degree in Mathematics (the Special Course) at the Faculty of Natural Sciences, University of Tirana. Graduated with very good results.

1993: University of Amsterdam, Scholarship of Tempus Program for Specialization in Econometrics.

1994: University of Macedonia, Greece - Scholarship of Tempus Program for Specialization in Statistics.

1998-today: Athens University of Economics and Business, Greece - Postgraduate studies in Statistics.

Professional Status

1987-1998: University of Korcha – Professor of Mathematics Probability, Statistics at the Economic Faculty.

Computer Skills

Languages	:	Fortran, Visual Basic
Database Skills	:	SQL, Server Oracle, Oracle Developer 2000
Statistics Analysis	:	SPSS, Minitab, S-Plus, Winbugs,

Languages

Greek (native), English (very good).

https://doi.org/10.26219/heal.aueb.5966

Publications

- ***"A study of Job Separation Probability in Korcha District"***

Bulletin of Natural Sciences, Tirana, September 1996

- ***" A new parametric model for Job Separation Probability"***

Bulletin of "Fan S. Noli" University, Korcha, June 1998.

- ***"The Generalized Waring Process"***

To appear in the Proceedings of the Hellenic-Europian Conference on Computer Mathematics and its Applications (HERCMA), 2001 Athens, Greece. (with E. Xekalaki)

https://doi.org/10.26219/heal.aueb.5966

ABSTRACT

Mimoza Zografi

THE GENERALIZED WARING PROCESS

Septembre 2002

In this thesis, a new process, termed as the *generalized Waring process* is defined and studied. This process is associated with the generalized Waring distribution which is a discrete distribution with a wide spectrum of applications, such as accident statistics, income analysis, inventory control, environmental statistics etc. Two models are shown to lead to the Generalized Waring process. One is related to a Cox Process while the other is a Compound Poisson Process.

The obtained Generalized Waring process is shown to be a stationary but non-homogenous Markov process. Several properties are studied and the intensity, the individual intensity, the Chapman-Kolmogorov differential equations of it, are obtained.

A Bayesian interpretation of the Generalized Waring Distribution is used to present some estimates of the "accident proneness".

Moreover, the Poisson and the Pólya processes are shown to arise as special cases of the Generalized Waring process. Using this fact, some known results and some properties of them are obtained. The moments, the transition probabilities and the Chapman-Kolmogorov differential equations of those processes, are derived in this context.

Two other genesis schemes are considered too.

https://doi.org/10.26219/heal.aueb.5966

https://doi.org/10.26219/heal.aueb.5966

ΠΕΡΙΛΗΨΗ

Μιμόζα Ζωγράφη

ΓΕΝΙΚΕΥΜΕΝΗ ΑΝΕΛΙΞΗ WARING

Σεπτέμβριος 2002

Μία νέα ανέλιξη, που ονομάζεται Γενικευμένη Ανέλιξη Waring, έχει ορισθεί και μελετηθεί στην διατριβή αυτή. Η ανέλιξη αυτή είναι συνδεδεμένη με την Γενικευμένη κατανομή Waring η οποία είναι μία διακριτική κατανομή με ένα ευρύ φάσμα εφαρμογών σε διάφορους τομείς, όπως στην στατιστική ανάλυση των ατυχημάτων, του εισοδημάτος, στατιστικά που αφορούν το περιβάλλον, κτλ. Δύο μοντέλα έχουν αποδειχτεί να οδηγούν στην Γενικευμένη Ανέλιξη Waring. Το ένα σχετίζεται με την Ανέλιξη Cox, ενώ το άλλο με την Σύνθετη Ανέλιξη Poisson.

Αποδεικνύεται ότι η ορισμένη Γενικευμένη Ανέλιξη Waring είναι μία ανέλιξη Markov, σταθερή αλλά όχι ομογενής και μελετούνται μερικές ιδιότητές της. Επίσης παρεστάθηκαν εκφράσεις για την ένταση και την ιδιαίτερη ένταση της ανελίξης και κατασκευάστηκαν οι διαφορικές εξισώσεις Chapman-Kolmogorov.

Επίσης θεορόντας μία Bayesian ερμηνεία της Γενικευμένης κατανομής Waring παρεστάθηκαν μερικές εκτιμήσεις της "ροπής ατυχημάτων".

Επιπλέων, αποδεικνύεται ότι οι Poisson και Polya ανελίξεις προκύπτουν σαν ειδικές περιπτώσεις της Γενικευμένης Ανέλιξης Waring. Χρησιμοποιώντας το γεγονός αυτό, μελετούνται μερικά γνωστά αποτελέσματα και λαμβάνωνται μερικές ιδιότητές τους.

Τέλος μελετούνται, δύο εναλλακτικά σχήματα γένεσις.

https://doi.org/10.26219/heal.aueb.5966

https://doi.org/10.26219/heal.aueb.5966

TABLE OF CONTENTS

1. Introduction- The basic concepts Page 1

2. The basic hypothesis of the Generalized Waring Process Page 5
- 2.1 The description of the accident pattern by a Cox Process. Page 5
- 2.2 A particular case. Page 7
- 2.3 An extension of Irwin's accident model Page 8
- 2.4 The Markovian Property. Page 10
- 2.5 The Spells' Model Page 11

3. The Generalized Waring Distribution. Page 15
- 3.1 The definition of the Generalized Waring distribution. Page 15
- 3.2 A Bayesian interpretation of the Genralized Waring distribution and some estimators of ν Page 16
- 3.3 A comparison through diagrams. Page 19

4. The Generalized Waring Process. Page 23
- 4.1 The definition of the Generalized Waring Process. Page 23
- 4.2 The moments and some other properties. Page 24
- 4.3 The transition probabilities and the Chapman-Kolmogorov equations of the Generalized Waring process Page 26

5. Some particular cases of the Generalized Waring Process. Page 31
- 5.1 The Pólya and the Poisson processes as special cases of the Generalized Waring process. Page 31
- 5.2 The moments and some other properties. Page 33
- 5.3 The transition probabilities of the Pólya and Poisson processes. Page 35
- 5.4 The Chapman-Kolmogorov equations for the Pólya and Poisson processes. Page 37
- 5.5 Continuous analogues. Two other limiting forms of the Generalized Waring Process Page 39

6. Some alternative genesis schemes Page 41
- 6.1 A mixed Pólya process Page 41
- 6.2 The Markovian property of the mixed Pólya process Page 42
- 6.3 A non-markovian Stochastic process of the Generalized Waring form Page 43

7. References Page 45

https://doi.org/10.26219/heal.aueb.5966

https://doi.org/10.26219/heal.aueb.5966

LIST OF FIGURES

Figure 2.2.1 The probability density functions of proneness $\Lambda(h)$, for different values of h. Page 8

Figure 3.3.1 The probability density function of the accident number N in two cases (Generalized Waring and Negative Binomial). The values of the parameters are: $a = 6$, $k = 2$, $\rho = 4$, $\nu = 2$, and $t = 10$. Page 19

Figure 3.3.2 The probability density function of the accident number N in two cases (Generalized Waring and Negative Binomial). The values of the parameters are: $a = 6$, $k = 0{,}8$, $\rho = 4$, $\nu = 2$, and $t = 10$. Page 21

Figure 3.3.3 The probability density function of the accident number N in two cases (Generalized Waring and Negative Binomial). The values of the parameters are: $a = 6$, $k = 0{,}2$, $\rho = 4$, $\nu = 2$, and $t = 10$. Page 21

https://doi.org/10.26219/heal.aueb.5966

CHAPTER 1

INTRODUCTION-THE BASIC CONCEPTS

The Generalized Waring Distribution is a discrete distribution that has many applications in areas such as accident statistics, income analysis, inventory control, environmental statistics etc. It has been proposed as an alternative to the Poisson and Negative Binomial Distribution to fit data that manifest a long tailed frequency distribution. Associated to both the Poisson and the Negative Binomial distributions are the well known Poisson and Polya Processes. Both of them are counting processes with independent and stationary increments. As it is known, the Generalized Waring distribution has been used as a model that better describes several practical situations as opposed to the Poisson Distribution or the Negative Binomial Distribution. Considering that, in this dissertation, the Generalized Waring Process has been defined. The results have been obtained in the context of models that have widely been considered for the interpretation of accident data. However, the concepts and terminology used can easily be modified so that the obtained results can be applied in several other fields ranging from economics, inventory control and insurance through to demometry, biometry and psychometry.

There are a lot of unpredictable events in every person's life. One such event is to be involved in a traffic accident. Various theories have been developed concerning the interpretation of accident situations. One of them is the theory of the pure chance which assumes that the probabilities of having an accident are solely the result of random factors. A natural model for this situation (see, for instance, Breiman (1963) is that the number of accidents is Poisson distributed with mean λ, i.e.

$$P\{N=n\}=\frac{\lambda^n}{n!}e^{-\lambda}, \qquad n=0,1,\ldots$$

where N is the random variable which describes the number of accidents of a single person.

One other theory which has attracted much interest is the "accident proneness" theory, in the context of which, the individuals differ from each other in their probabilities of having an accident or in their accident proneness and that the individual's accident proneness remains constant in time. This theory assumes that the

https://doi.org/10.26219/heal.aueb.5966

factors contributing to the happening of the accidents are of two kinds, random and non-random, where the non-random ones referred to the individual's psychology, explaining in this way, more or less, why the individuals have unequal accident proneness. Under this type of assumptions Greenwood and Wood (1919) obtained the number of accidents N has Negative Binomial distribution with parameters k and $\frac{1}{\nu}$, i.e.

$$P\{N=n\}=\binom{k+n-1}{n}\nu^{n}(1+\nu)^{-(n+k)}, \quad n=0,1,\ldots$$

The Irwin's (1968) "proneness-liability" model which assumes also that the non-random factors can be further split into psychological and external factors provides more explanation as to why some individuals in the population tend to have more accidents than others. In the context of this model, the individual accident proneness does not remain constant, because the population is exposed to a variable risk. In his model, Irwin 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 the accident and he derived the univariate generalized Waring distribution as the distribution of the number of accidents. This distribution was applied by Irwin (1968, 1975b) to data on accidents sustained by men in a soap factory, providing an improved fit as compared to the Negative binomial. It is one interesting member of the family of mixed Poisson distribution. In Irwin's model the conditional distribution of the number of accidents N given λ is a Poisson distribution with parameter λ, where λ for a given ν follows itself a Gamma distribution with parameters k and $\frac{1}{\nu}$, i.e. λ for a given ν is a random variable with density u given by

$$u(l)=\frac{\nu^{-k}}{\Gamma(k)}l^{k-1}e^{-\left(l\nu^{-1}\right)}, \quad l\geq 0,$$

resulting thus in a Negative Binomial distribution with parameters k and $\frac{1}{\nu}$. He, also allowed the parameter ν of the Negative Binomial to follow a *BetaII* distribution of the second type (see for instance Johnson and Kotz 1969) with parameters a and ρ, i.e. ν is a random variable with density φ given by

https://doi.org/10.26219/heal.aueb.5966

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

obtaining thus the Generalized Waring distribution with parameters a, k and ρ.

The Poisson and the Polya processes have also been used in the Accident Theory to describe the accident pattern. Under the hypotheses of the pure chance, the Poisson process with intensity λ has been proposed as a model which can describe the number of accidents sustained by an individual during several years. The Polya 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 (Newbold 1927) 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. (We have to denote that this is not the only scheme which leads to the Polya process, see for instance Grandell (1997, p.5-7) for the Polya scheme and the discussions of Bates and Neyman (1952) and Cane(1974) for the differencies between these two schemes). Both of these processes satisfy the Markovian property because this is a property of the accident pattern, i.e. the number of accidents during the 'next' period $(t,t+h]$ depends only on the number of accidents at the present time t.

In this disertation, a new process is defined and studied. This process is associated with the generalized Waring distribution which, as it is mentioned above, is a discrete distribution with a wide spectrum of applications (see, e.g. Irwin (1975), Xekalaki (1983b). This new process is termed in the sequel as the *generalized Waring process.* Analogously to the case of Poisson and Polya process, the generalized Waring process is postulated as a Markov process. The starting point is a process of Negative binomial form but different from a Polya process. We consider a group of individuals under the hypotheses of the Irwin's "proneness-liability" model during several years. We also assume that "accident liability", in an interval of time $(t,t+h]$, is the outcome of a random variable $\Lambda(h)$ with distribution $U(h)$ which depends on the interval size h, and in two non-overlapping time periods the respective random variables are independent. If these distributions for each h are Gamma, the resulting process is of Negative binomial form, but different of a Polya process. This is a Cox

https://doi.org/10.26219/heal.aueb.5966

process and we call it the Negative binomial process. This process is then mixed with a BetaII distribution of the second type. This scheme is shown to lead to the generalized Waring process. We treat it in the second chapter. A proof that this process is a stationary but non-homogenous Markov process is provided in this chapter, too. We use the fourth structural property of a multivariate generalized Waring distributed random vector, treated by Xekalaki E. (1986), in order to prove it.

We have considered, another scheme which leads to the Generalized Waring process, referring to Cresswell and Froggatt's (1963) *spells model* is proposed in the framework considered by Xekalaki (1983b). This is another model that describes the number of accidents, which rejects the concepts of accident proneness and contagion. This scheme, too, allows for stationary increments and the validity of the Markovian property, as also shown in chapter 2.

In the sequel the definitions of the Negative binomial and the Generalized Waring distributions and a Bayesian interpretation of the second one, have been used to present some properties and to interpret some estimates of the "accident proneness" ν.

The definition of the Generalized Waring process is the subject of the fourth chapter. We define it as a stationary but non-homogenous Markov process. Expressions for the first two moments of this process, as well as the results of the intensity and the individual intensity of it, are given in the fourth chapter. The definition of the Negative binomial process and an equivalent definition of the Generalized Waring process are used here in order to find the individual intensity of it. In the sequel we present the transition probabilities derived in the second chapter, and we calculate the forward and the backward Chapman-Kolmogorov differential equations of the Generalized Waring process.

The Poisson process and the Polya process are special cases of the Generalized Waring process. Using this fact, in the fifth chapter, some new definitions of them as limiting cases of the Generalized Waring process have been proposed, and the derivations of some known theoretical results about them are presented. The moments, the transition probabilities and the Chapman-Kolmogorov differential equations of those processes, are derived in this context.

Finally, two further genesis schemes are considered in chapter 6. The results are based on Zografi and Xekalaki (2001).

https://doi.org/10.26219/heal.aueb.5966

CHAPTER 2

THE BASIC HYPOTHESES OF THE GENERALIZED WARING PROCESS

Two different generating schemes which give rise to a process of Generalized Waring form are considered in this section. The first is an extension of the "proneness-liability" model developed by Irwin. It considers a population of individuals exposed to a random accident risk whose distribution varies during the time. The second scheme is a variant of the "spells" model due to Cresswell and Froggatt (1963), treated also in the paper by Xekalaki (1984). This model assumes that each person is liable to spells (periods of time) and that all of his accidents occur within these spells. The hypotheses of proneness and contagion are not present here.

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

In this section we consider first the assumptions for a **Polya 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 in 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 in an interval of time $(t,t+h)$ for each individual is a Poisson process with a "personal rate λ", where this λ stands for the respective accident proneness, and regarding a personal λ as the outcome of a random variable Λ with a Gamma distribution with parameters k and $\frac{1}{\nu}$, i.e. $\Lambda \sim \Gamma\left(k,\frac{1}{\nu}\right)$ [1], it follows that the number of accidents $N(t+h)-N(t)$ in the interval of time $(t,t+h)$, defines a Polya Process with parameters k and $\frac{1}{\nu}$, i.e.:

(I) $N(0)=0$

---

[1] $\Gamma\left(k,\frac{1}{\nu}\right)$ stands for Gamma distribution with parameters k and $\frac{1}{\nu}$

(II) N is a birth process

(III) $N(t+h)-N(t)$ has the following distribution

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

where Λ is $\Gamma\left(k,\frac{1}{\nu}\right)$-distributed.

So,

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

$$= \int_0^{+\infty} \frac{(lt)^n}{n!}e^{-lt}\frac{\nu^{-k}}{\Gamma(k)}l^{(k-1)}e^{-l/\nu}dl$$

$$= \binom{k+n-1}{n}\left(\frac{1}{1+\nu t}\right)^k\left(\frac{\nu t}{1+\nu t}\right)^n \quad n=0,1,\ldots \qquad \textbf{(2.1.1)}$$

i.e. $N(t)$ has a Negative Binomial distribution with parameters k and $\frac{1}{\nu t}$, i.e. $N(t) \sim NB\left(k,\frac{1}{\nu t}\right)$ [2].

The distribution of the random variable Λ explains here the variation of the accident proneness from individual to individual. As mentioned above, the term accident proneness here refers to the individual's psychology. It seems more natural to assume that this variation in an interval of time $(t,t+h]$ depends on the size h of the interval and that in two non-overlapping time periods the respective variations are independent. So, we regard now a personal λ, in an interval of time $(t,t+h]$, as the outcome of a random variable $\Lambda(h)$ with distribution $U(h)$ which depends on the interval size h. If we assume that $U(h)$ is $\Gamma\left(k(h),\frac{1}{\nu(h)}\right)$, where $k(h)$ and $\nu(h)$ are in general some functions of h, then it is clear that the number of accidents $N(t)$ will

[2] $NB\left(k,\frac{1}{\nu t}\right)$ stands for Negative Binomial distribution with parameters k and $\frac{1}{\nu t}$

be a stochastic process of Negative binomial form which fulfills the following assumptions

(I) $N(0)=0$

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

$$P\{N(t+h)-N(t)=n\} = \int_0^{+\infty} \frac{(\lambda h)^n}{n!} e^{-\lambda h} \frac{\nu(h)^{-k(h)}}{\Gamma(k(h))} \lambda^{(k(h)-1)} e^{(-\lambda/\nu(h))} d\lambda \qquad (2.1.2)$$

$$n=0,1,\ldots$$

We find,

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

$$n=0,1,\ldots$$

and using the first assumption we find also,

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

$$n=0,1,\ldots$$

This does mean that $N(t)$ for any t has Negative Binomial distribution with parameters $k(t)$ and $\frac{1}{\nu(t)t}$.

Hence, we can confirm that the following relation holds

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

$$n=0,1,\ldots$$

This tells us precisely that $N(t)$ is a *Cox Process* (see e.g. Grandell, J. 1997, p.83).

2.2 A particular case

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

https://doi.org/10.26219/heal.aueb.5966

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

$$n=0,1,\ldots$$

and also

$$P_n(t)=P\{N(t)=n\}=\binom{kt+n-1}{n}\left(\frac{1}{1+v}\right)^{kt}\left(\frac{v}{1+v}\right)^{n} \qquad (2.2.1)$$

$$n=0,1,\ldots$$

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

The following diagram illustrates how the distribution of of the random variable $\Lambda(h)$ depends on the length h of the time interval.

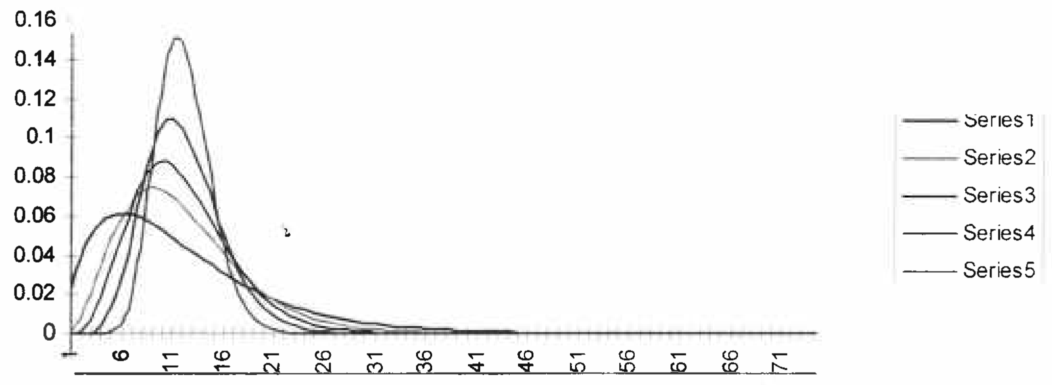

Figure 2.2.1. The probability density functions of $\Lambda(h)$, for different values of h, when $\nu=6$ and $k=2$ (The mean is 12). The series 1, 2, 3, 4, 5, correspond to the values h =1, 2, 3, 5, 10.

It can be seen from the figure 2.2.1 that, as h increases, the shape of the curves becomes sharper around the mean. In fact for a big h the value of the variance $\frac{k\nu^2}{h}$ is small. Hence, what we have assumed is that the accident's proneness in an interval of time $(t,t+h)$ varies in such a way that the values of it for a big size h are too near to the constant mean νk.

2.3 An extension of Irwin's accident model.

In the sequel, we consider the hypotheses of Irwin's (1963) accident model.

https://doi.org/10.26219/heal.aueb.5966

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

In this model the conditional distribution of the number of accidents $N(t)$ given λ is a Poisson distribution with parameter λ, where λ for a given ν follows itself a Gamma distribution with parameters k and $\frac{1}{\nu}$. The conditional distribution of the random variable Λ for a given ν explains here the variation of the accidents' liability from individual to individual. As mentioned above, the term accidents' liability, here, refers to the individual's psychology. The conditional distribution of the random variable Λ for a given ν describes differences in external risk factors among individuals. As before we suppose that, 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 $N(t)$, for a given ν, will be a stochastic process of Negative binomial form with parameters kt and $1/\nu$. This starts at 0 and has stationary increments with a distribution given by (2.1.2). Let us further allow the parameter ν of the Negative Binomial to follow a BetaII distribution with parameters a and ρ, obtaining thus for the distribution of the number of accidents $N(t)$:

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

$$=\int_0^\infty \binom{kh+n-1}{n}\nu^{n}(1+\nu)^{-(kh+n)}\frac{\Gamma(p+a)}{\Gamma(a)\Gamma(p)}\nu^{a-1}[1+\nu]^{-(a+p)}$$

$$=\frac{p_{(kh)}}{(a+p)_{(kh)}}\frac{a_{(n)}(kh)_{(n)}}{(a+p+kh)_{(n)}}\frac{1}{n!}\qquad n=0,1,\ldots$$

and

$$P\big(N(t)=n\big)=P_n(t)=\frac{\rho_{(kt)}}{(a+\rho)_{(kt)}}\frac{a_{(n)}(kt)_{(n)}}{(a+\rho+kt)_{(n)}}\frac{1}{n!}\quad n=0,\ 1,\ \ldots\qquad(2.3.1)$$

https://doi.org/10.26219/heal.aueb.5966

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

Remark 2.3.1

If we consider individuals of proneness ν and liability $\lambda_i | \nu \quad i=1,2$ respectively for each of the two non-overlapping intervals of time, by the model's assumptions the numbers N_1, N_2 of accidents incurred by these individuals follow a double Poisson distribution with parameters $(\lambda_1 | \nu, \lambda_2 | \nu)$. Then for individuals with the same proneness but varying liabilities the joint distribution of accidents over the two intervals is double Negative Binomial with parameters $\left(\left(kh_{1,} \frac{1}{\nu}\right);\left(kh_{2,} \frac{1}{\nu}\right)\right)$, where h_1, h_2 are the respective sizes of these intervals. If we allow now the proneness parameter ν to follow a a *BetaII* distribution with parameters a and ρ then the joint distribution of the numbers of accidents over the two intervals is a bivariate Generalized Waring distribution with parameter $\left(\left(a, kh_{1,}\rho\right);\left(a, kh_{2,}\rho\right)\right)$ (see Xekalaki, E. 1984). Now, it is clear that, if we consider a number of intervals greater than two, then the joint distribution of the numbers of accidents over those intervals, will follow a multivariate Generalized Waring distribution.

In the sequel we are going to use the above remark to prove that the resulting process of this first generating scheme is a Markov process. a multivariate Generalized Waring distribution.

In the sequel we are going to use the above remark to prove that the resulting process of this first generating scheme is a Markov process i.e. to prove that $P\left(N(t+h)=n \mid N(t)=m, N(s)=n_s, 0 \le s < t\right)$ coincides with $P\left(N(t+h)=n \mid N(t)=m\right)$ for every non-negative integers n,m,n_s $0 \le s < t$.

First we denote that

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

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

Consider, now, the random vector

$$\left(N(t+h)-N(t), N(t)-N(s), N(s)-N(0)\right), \; 0 \le s < t$$

https://doi.org/10.26219/heal.aueb.5966

It follows from the remark 2.3.1 that this vector has $MGWD(a;\underline{k};\rho)$ where $\underline{k}=(kh,k(h-s),ks)$. We refer now to the third structural property of the multivariate Generalized Waring distribution proved by Xekalaki E. in her paper 'The multivariate Generalized Waring distribution', (p.1054) and we find $\left(N(t+h)-N(t)\mid N(t)-N(s),N(s)-N(0)\right)\sim MGWD(a+n(t);kh;\rho+kt)$, where *n(t)* is the value of *N(t)*.

Hence

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

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

$$=\frac{\Gamma(a+n)}{\Gamma(a+m)}\frac{(kh)_{(n-m)}}{(n-m)!}\frac{(\rho+kt)_{(a+m)}}{(\rho+kt+kh)_{(a+n)}}= \tag{2.3.1}$$

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

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

2.4 The Spells' Model

In the sequel, an alternative scheme generating a process of a Generalized Waring form is considered. This is a variant of Cresswell and Froggatt's (1963) *Spells* model that has been considered in the paper of Xekalaki (1984). According to this model, each person is liable to spells, (a spell is a period of time during which the person's performance is weak). 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 $S(t)$ is a homogeneous Poisson process with rate $\frac{k}{m}$, $t\geq 0$, $k>0$, 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

https://doi.org/10.26219/heal.aueb.5966

$$X(t)=\sum_{k=1}^{S(t)} X_k,$$

where $S(t)$ is a homogenous Poisson process with rate $\frac{k}{m}$ and $\{X_k\}_1^{\infty}$ are identical independent 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.

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

and

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

the random variable $X(t)$, is a Negative Binomial random variable with parameters $\left(kt,\frac{1}{\nu}\right)$ (see Kemp, 1967 and Chatfield and Theobald, 1973). Here ν is regarded as the external risk parameter, too. Then, if the differences in this external risk can be described by a $BetaII(a,\rho)$ distribution, the resulting accident distribution is of a Generalized Waring form with parameters a, kt, and ρ.

Let us consider, now, the counting process $\{N(t),t\geq 0\}$ where $N(t)$ can be represented, for $t\geq 0$, by

$$\sum_{k=1}^{S(t)} X_k,\quad \left(\sum_{k=1}^{0} X_k=0\right),$$

where $S(t)$ is a homogenous Poisson process with rate $\frac{k}{m}$, $\{X_k\}_1^{\infty}$ is a logarithmic series distributions with parameters (m,ν) that is also independent of the process $S(t)$, and ν is a non negative random variable with $BetaII(a,\rho)$ distribution.

Theorem 2.4.1

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

(I) $N(0)=0$

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

https://doi.org/10.26219/heal.aueb.5966

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

Proof

The proof of (I) is straightforward.

To prove condition (II), denote by φ the probability distribution function (p.d.f) of the random variable ν. Then we can write:

$$p\big(N(t+h)-N(t)=n\big)=\int_0^{+\infty} P\big(N(t+h)-N(t)=n/\nu\big)\varphi(\nu)d\nu$$

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

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

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

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

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

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

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

https://doi.org/10.26219/heal.aueb.5966

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

$$P_\nu(N(t+h)=n, N(t)=m, N(s)=n(s) \ for \ 0\leq s\leq t)$$

$$=p_h(m-n)p_{t-s}(m-n_s)p_s(n(s))$$

$$=\binom{kh+m-n-1}{m-n}\binom{k(t-s)+m-n(s)-1}{m-n(s)}\binom{ks+n(s)-1}{n(s)}\nu^n(1+n)^{-(n+kh+kt)}$$

In the same way we find that

$$P(N_\nu(t)=m, N_\nu(s)=n(s), 0\leq s\leq t)=p_{t-s}(m-n(s))p_s(n(s))$$

$$=\binom{k(t-s)+m-n(s)-1}{m-n(s)}\binom{ks+n(s)-1}{n(s)}\nu^m(1+\nu)^{-(m+kt)}$$

Hence, we take

$$P(N(t+h)=n, N(t)=m, N(s)=n(s) \ for \ 0\leq s\leq t)=\frac{\Gamma(a+n)}{\Gamma(a+m)}\frac{(kh)_{(n-m)}}{(n-m)!}\frac{(\rho+kt)_{(a+m)}}{(\rho+kt+kh)_{(a+n)}} \quad (2.4.1)$$

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

CHAPTER 3

THE GENERALIZED WARING DISTRIBUTION

This chapter is devoted to the Generalized Waring distribution Most of the results obtained refer to the ideas that Grandell (1997) have used to obtain similar results in the case of a Pólya process and in general of a mixed Poisson process. In the first part the definition of this distribution, a Bayesian interpretation, some properties of it, and some estimates of the accidents' proneness, are provided, while the relation between the Generalized Waring and the Negative binomial distributions with the same mean is illustrated by some diagrams in the second part.

3.1 The definition of **the Generalized Waring distribution.**

Definition 3.1.1 A discrete random variable N is said to be Negative Binomial distributed with parameters k and $\frac{1}{\nu}$, $k>0$, $\nu>0$ i.e. $N \sim NB\left(k,\frac{1}{\nu}\right)$, if

$$P\{N=n\}=\binom{k+n-1}{n}\left(\frac{1}{\nu+1}\right)^{k}\left(\frac{\nu}{1+\nu}\right)^{n}, \quad n=0,1,\ldots \tag{3.1.1}$$

Definition 3.1.2 A discrete random variable N is said to be Generalized Waring distributed with parameters a, k and ρ, $a>0$, $k>0$, $\rho>0$, i.e. $GW(a,k,\rho)$ if

$$P_n=P\{N=n\}=E\left[\binom{k+n-1}{n}\left(\frac{1}{\nu+1}\right)^{k}\left(\frac{\nu}{1+\nu}\right)^{n}\right] \tag{3.1.2}$$

$$n=0,1,\ldots$$

where $\nu \sim BetaII(a,\rho)$.

So we can write for these probabilities:

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

$$=\frac{\Gamma(a+p)}{\Gamma(a)\Gamma(p)}\int_{0}^{+\infty}\binom{k+n-1}{n}\left(\frac{1}{\nu+1}\right)^{k}\left(\frac{\nu}{1+\nu}\right)^{n}v^{a-1}(1+v)^{-(a+p)}\,dv$$

$$=\frac{\rho_{(k)}}{(\rho+a)_{(k)}}\frac{a_{(n)}k_{(n)}}{(\rho+a+k)_{(n)}}\frac{1}{n!} \qquad n=0,1,\ldots$$

https://doi.org/10.26219/heal.aueb.5966

In both the above definitions k need not to be integer.

Let us consider a random variable $N \sim NB\left(kt, \frac{1}{\nu}\right)$, i.e.

$$P\{N=n\} = \binom{kt+n-1}{n}\left(\frac{1}{\nu+1}\right)^{kt}\left(\frac{\nu}{1+\nu}\right)^{n}, \quad n=0,1,\ldots \tag{3.1.3}$$

If we assume that the parameter ν is the outcome of a random variable that is $BetaII(a,\rho)$-distributed, the distribution of N is given by

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

$$= \frac{\rho_{(kt)}}{(\rho+a)_{(kt)}}\frac{a_{(n)}(kt)_{(n)}}{(\rho+a+kt)_{(n)}}\frac{1}{n!} \qquad n=0,1,\ldots \tag{3.1.4}$$

So, $N \sim GW(a,kt,\rho)$.

It follows from (3.1.4) that

$$P_{n+1}(t) = P_n(t)\frac{(a+n)(kt+n)}{(n+1)(a+\rho+kt+n)} \tag{3.1.5}$$

Since $P_0(t) = \dfrac{\rho_{(a)}}{(\rho+kt)_{(a)}}$ we get by (3.1.5) a recursive algorithm for the distribution of N.

3.2 A Bayesian interpretation of the Genralized Waring distribution and some estimators of V

Interpreting the $BetaII(a,\rho)$-distribution of the random variable V as a prior distribution, we can consider a Generalized Waring variable N as generated in the following way. First a realization ν of the random variable V is generated. Conditional upon that realization, N is a Negative Binomial random variable with parameters $\left(kt, \frac{1}{\nu}\right)$. Formally we can give the following definition:

Definition 3.2.2 A discrete random variable N is said to be mixed Negative Binomial distributed if its distribution $\Pi_U(B)$ is given by

$$f_U(B) = \int_0^{\infty} f_{\nu}(B)\, dU(\nu)$$

https://doi.org/10.26219/heal.aueb.5966

for any event B, where f stands for the probability density function of the Negative Binomial distribution.

If f is assumed to have the parameters $\left(kt, \frac{1}{\nu}\right)$ and U stands for the probability function of the random variable $\nu \sim BetaII(a,\rho)$, N is Generalized Waring distributed with parameters a, kt and ρ, $a>0$, $k>0$, $\rho>0$.

Using this interpretation of the Generalized Waring distribution we can obtain a Bayesian estimate of proneness:

For any event B, we can regard

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

where U stands for the probability function of the random variable $\nu \sim BetaII(a,\rho)$, as the posterior distribution of ν given B or, more precisely, given $\{N \in B\}$, provided that $P\{N \in B\} = \int_0^{+\infty} P(N \in B| \, l)dU(l) > 0$.

Proposition 3.2.1 Let N be $GW(a,kt,\rho)$-distributed. Then

$$P\{\nu \le x | \, N(s)=n\} = \frac{\int_0^x \nu^{n+a-1}(1+\nu)^{-(n+a+\rho+ks)}d\nu}{\int_0^{+\infty} \nu^{n+a-1}(1+\nu)^{-(n+a+\rho+ks)}d\nu} \tag{3.2.1}$$

Proof

Using an argument similar to that used by Xekalaki (1983b) for the case of the generalised 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 $\left(\frac{\rho+ks}{a+n}\right)\nu \,\Big|\, (N(s)=n)$ has the F distribution with $2(a+n)$ and $2(\rho+ks)$ degrees of freedom. Following Xekalaki (1983b), this result can be used to obtain a confidence interval for $\nu \mid (N(s)=n)$, 'estimating' in this way a person's proneness on the basis of the incurred number of accidents.

Corollary

1. Let N be $GW(a,kt;\rho)$-distributed. Then

$$E\{\nu \mid N(s)=n\} = \frac{\int_0^{+\infty} \nu^{n+a}(1+\nu)^{-(n+a+\rho+ks)} d\nu}{\int_0^{+\infty} \nu^{n+a-1}(1+\nu)^{-(n+a+\rho+ks)} d\nu} = \frac{a+n}{\rho+ks}. \qquad (3.2.2)$$

Proof

Combining the above property with the relation

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

we have

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

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

From (3.1.7) it seems natural to interpret $\nu_B^* \equiv E\{\nu \mid N(s)\} = \frac{a+N(s)}{\rho+ks}$ as an estimator, or the Bayes estimate of ν.

Another simple and natural estimator of ν is $\frac{1}{kt}N(t)$. We find

https://doi.org/10.26219/heal.aueb.5966

$$E\left(\frac{1}{kt}N(t)\,\Big|\,\nu\right)=\frac{1}{kt}E\left(N(t)\mid\nu\right)=\frac{1}{kt}\nu kt=\nu$$

and we have also

$$E\left[\left(\frac{1}{kt}N(t)-\nu\right)^2\right]=E\left[E\left[\left(\frac{1}{kt}N(t)-\nu\right)^2\Big|\,\nu\right]\right]$$

$$=E\left[\frac{1}{(kt)^2}E\left[\left(N(t)-\nu kt\right)^2\Big|\,\nu\right]\right]$$

$$=E\left[\frac{1}{(kt)^2}kt\left(\nu(\nu+1)\right)\right]=\frac{E\left[\nu(\nu+1)\right]}{kt}$$

Thus, for large values of t, we have $\frac{1}{kt}N(t)\approx\nu$.

Considering linear combinations of this estimator, we can construct a number of other useful estimators. So, for example, a simple such estimator is $\nu_L^*=b+c\frac{1}{kt}N(t)$, where b and c are chosen so that $E\left[\left(\nu_L^*-\nu\right)^2\right]$ is minimized.

The estimator thus obtained can be referred to as the best linear estimator is given by:

$$\nu_L^*=\frac{E\left[\nu(\nu+1)\right]}{kt\operatorname{var}\nu+E\left[\nu(\nu+1)\right]}+\frac{kt\operatorname{var}\nu}{kt\operatorname{var}\nu+E\left[\nu(\nu+1)\right]}\frac{1}{kt}N(t)$$

$$=\frac{E\left[\nu(\nu+1)\right]+N(t)\operatorname{var}\nu}{kt\operatorname{var}\nu+E\left[\nu(\nu+1)\right]}.$$

3.3 A comparison through diagrams

In figures 3.3.1-3.3.3 we illustrate the relation between the distributions of N in the cases where (i) N is $GW(a,kt,\rho)$-distributed (series 1) and (ii) N is $NB\left(kt,\frac{1}{\nu}\right)$-distributed (series 2), but with the same mean. We have chosen the relatively large value $t=10$, in order to be able to see how $P_n(t)$ behaves for small values of n. In figures 3.3.1, 3.3.2 the Generalized Waring variation differs from the Negative binomial variation, while, in figure 3.3.3, they are almost the same.

https://doi.org/10.26219/heal.aueb.5966

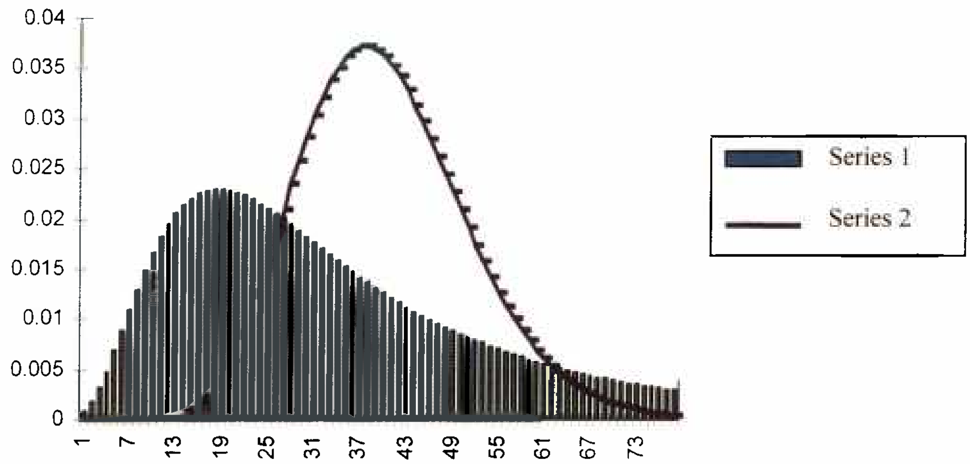

Figure 3.3.1 The probability density function of N in two cases. The values of the parameters are: $a = 6,\ k = 2,\ \rho = 4,\ \nu = 2,$ and $t = 10$. The mean is 40.

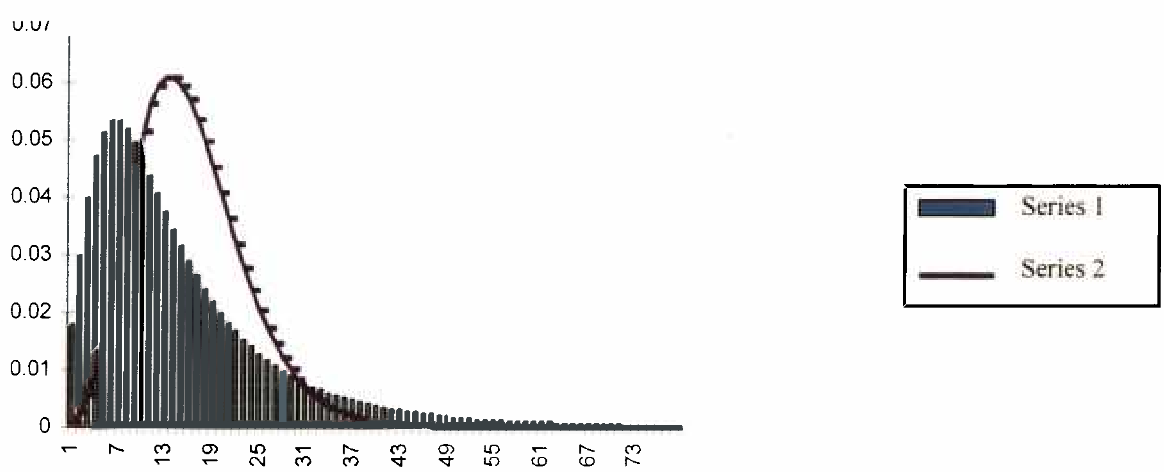

Figure 3.3.2 The probability density function of N in two cases. The values of the parameters are: $a = 6,\ k = 0{,}8,\ \rho = 4,\ \nu = 2,$ and $t = 10$. The mean is 16.

https://doi.org/10.26219/heal.aueb.5966

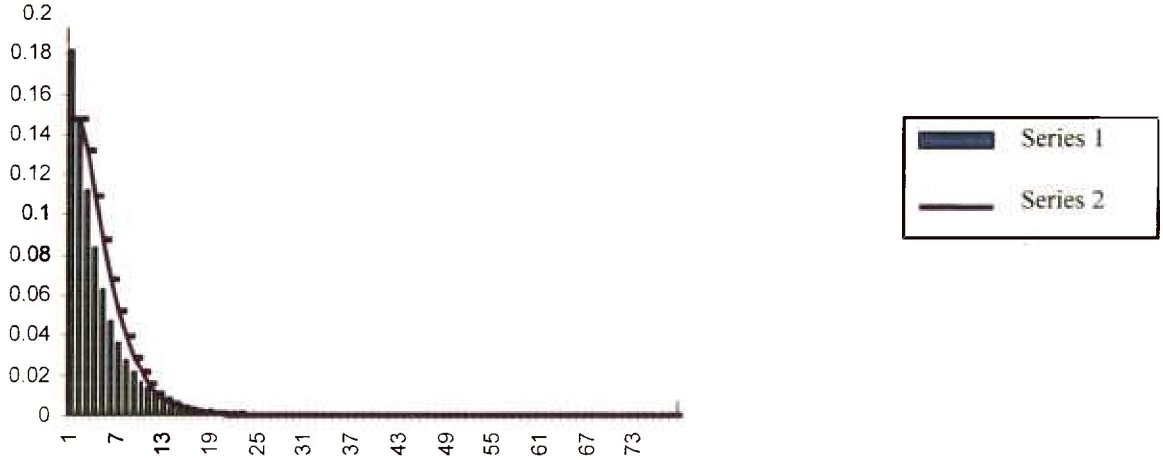

Figure 3.3.3 The probability density function of N in two cases. The values of the parameters are: $a = 6,\ k = 0{,}2,\ \rho = 4,\ \nu = 2,$ and $t = 10$. The mean is 4.

https://doi.org/10.26219/heal.aueb.5966

CHAPTER 4

GENERALIZED WARING PROCESS

In this chapter we shall give the definition of the Generalized Waring process. This definition is based on the two schemes treated in the second chapter. The two first moments the intensity and the individual intensity of this process are presented in the sequel. Here we need to give the definition of the Negative binomial process and an equivalent definition of the Generalized Waring process, too. The last part of the chapter treats the transition probabilities and the Chapman-Kolmogorov equations of this process.

4.1 The definition of the Generalized Waring process.

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

Definition 4.1.1 The counting process $\{N(t), t \geq 0\}$ is said to be a Generalized Waring Process with parameters (a, k, ρ), $a > 0, k > 0, \rho > 0$ if:

(I) $N(0) = 0$

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

(III) $N(t+h) - N(t)$ is $GW(a, kh, \rho)$-distributed for each $h > 0$, $t \geq 0$.

Conditions (I), (II) and (III) tell us that this process starts at 0, it has stationary increments and $P(N(t) = n) = \frac{\rho_{(kt)}}{(\rho + a)_{(kt)}} \frac{a_{(n)} (kt)_{(n)}}{(a + \rho + kt)_{(n)}} \frac{1}{n!}$, i.e. $N(t)$ is $GW(a,\ kt,\ \rho)$-distributed.

Given that the defined Generalized Waring process is a Markov process the relations (2.3.1) and (2.4.1) give the transition probabilities of it.

$$P\{N(s+t) = n \mid N(s) = m\} = \frac{\Gamma(a+n)}{\Gamma(a+m)} \frac{(kt)_{(n-m)}}{(n-m)!} \frac{(\rho + ks)_{(a+m)}}{(\rho + ks + kt)_{(a+n)}}$$

We denote

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

https://doi.org/10.26219/heal.aueb.5966

and recall that it represents the probability that a process presently in state m will be in state n a time t later. This probability in this case depends on the present time so the defined Generalized Waring process is a non-homogenous Markov process.
It is clear that $p_{0,n}(0,t)=P(N(t)=n \mid N(0)=0)=P(N(t)=n)=p_n(t)$

In order to show that such a process does exist it is sufficient to prove that the transition probabilities satisfy the Chapman-Kolmogorov equations i.e.

$$p_{m,n}(s,t)=\sum_{i=m}^{n} p_{m,i}(s,\tau)p_{i,n}(\tau,t), \qquad \text{for } s\le\tau\le t,\ m\le n \qquad (4.2.1)$$

We calculate first:

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

$$=\left[\sum_{i=m}^{n}\frac{\Gamma(a+i)}{\Gamma(a+m)}\frac{(k(\tau-s))_{(n-i)}}{(i-m)!}\frac{(\rho+ks)_{(a+m)}}{(\rho+k\tau)_{(a+i)}}\cdot\frac{\Gamma(a+n)}{\Gamma(a+i)}\frac{(k(t-\tau))_{(i-m)}}{(n-i)!}\frac{(\rho+k\tau)_{(a+i)}}{(\rho+kt)_{(a+n)}}\right]=$$

$$=\frac{\Gamma(a+n)}{\Gamma(a+m)}\frac{(\rho+ks)_{(a+m)}}{(\rho+kt)_{(a+n)}}\sum_{i=m}^{n}\frac{(k(t-\tau))_{(n-i)}}{(n-i)!}\frac{(k(\tau-s))_{(i-m)}}{(i-m)!}$$

Using the identity

$$\binom{k(t-s)+(n-m)}{(n-m)}=\sum_{i=m}^{n}\binom{k(t-\tau)+(n-i)}{(n-i)}\cdot\binom{k(\tau-s)+(i-m)}{(i-m)}$$

we obtain that

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

$$=\frac{\Gamma(a+n)}{\Gamma(a+m)}\frac{(\rho+ks)_{(a+m)}}{(\rho+kt)_{(a+n)}}\binom{k(t-s)+(n-m)}{(n-m)}=p_{m,n}(s,t)$$

which proves 4.2.1.

In order to determine if an arbitrary counting process is actually a Generalized Waring process, we must show that the conditions (I), (II), (III) are satisfied. The first two conditions can usually be directly verified from our knowledge of the process. However, it is not all clear how we would determine that the condition (III) is satisfied.

https://doi.org/10.26219/heal.aueb.5966

4.2 The moments and some other properties

Let N be a Generalized Waring process with parameters (a,k,ρ). For any t, $N(t)$ is a Generalized Waring distributed random variable, Hence, (see for instance Xekalaki 1994), for any t, $E[N(t)]=\dfrac{akt}{\rho-1}$, $Var[N(t)]=\dfrac{akt(\rho+kt-1)(\rho+a-1)}{(\rho-1)^2(\rho-2)}$.

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

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

where,

$\sigma^2_{\Lambda(t)}=akt(a+1)(\rho-1)^{-1}(\rho-2)^{-1}$ is the component due to liability

$\sigma^2_{\nu}=a(a+\rho-1)(\rho-1)^{-2}(\rho-2)^{-1}$ is the component due to proneness

and

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

The Generalized Waring process is a stationary process. For a stationary process N, $E[N(t)]=\eta\cdot t$, 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 $\eta=\dfrac{ak}{\rho-1}$. For this process (like for all stationary processes), there always exists, a random variable $\overline{N}$ with $E(\overline{N})=\eta$, called the *individual intensity*, such that $\dfrac{N(t)}{t}\xrightarrow{p.}\overline{N}$[c] as $t\to+\infty$ (see e.g. Grandell, 1997, p.53). The intensity η is finite. Hence, it follows that the individual intensity $\overline{N}$ is finite $a.s.$[d]

We give now a definition of the Negative binomial process in order to give an equivalent definition of the Generalized Waring process which shall be useful to

[c] The symbol $p.$ here stands for the convergence in probability of a random variable

[d] The symbol $a.s.$ (almost sure) implies that $P(\omega\in\Omega|\overline{N}(\omega)\text{ is finite})=1$

https://doi.org/10.26219/heal.aueb.5966

prove the theorem 4.2.1 below, which gives the individual intensity of the Generalized Waring Process.

Definition 4.2.1. The counting process $\{N(t), t \geq 0\}$ is said to be a Negative binomial Process with parameters $\left(k\ ,\ \frac{1}{\nu}\right)$ $k>0,\ \nu>0$, if

(I) $N(0)=0$

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

(III) $N(t+h)-N(t)$ is $NB\left(kh\ ,\ \frac{1}{\nu}\right)$-distributed for each $h>0,\ t \geq 0$.

The first condition together with the condition (III) allow us to take also that $N(t)$ is $NB\left(kt\ ,\ \frac{1}{\nu}\right)$-distributed.

Definition 4.2.2 A Negative binomial process with parameters $k=1$ and $\nu=1$ is called a *Standard Negative Binomial Process.*

Definition 4.2.3 Let ν be a random variable $BetaII(a,\rho)$-distributed and a standard Negative binomial process $\tilde{N}$ independent of it. Let $k>0$ be a constant.

The point process $N=\tilde{N}\circ\left(k,\ \frac{1}{\nu}\right)$, where $\tilde{N}\circ\left(kt,\frac{1}{\nu}\right)\overset{def}{=}\tilde{N}\left(kt,\frac{1}{\nu}\right)$ and for every t, $\tilde{N}\left(kt,\frac{1}{\nu}\right)\sim NB\left(kt,\frac{1}{\nu}\right)$, is called the *Generalized Waring Process.*

It is already clear that definition 4.2.3 is equivalent to definition 4.1.1. By definition 4.2.3 one can prove the following property

Theorem 4.2.1 Let N be a Generalized Waring process. Then

$$\frac{1}{t}N(t)\xrightarrow[t\to\infty]{p.}\nu k\ .$$

Proof

$$\lim_{t\to\infty}\frac{1}{t}N(t)=\lim_{t\to\infty}\frac{1}{t}\tilde{N}\left(kt,\frac{1}{\nu}\right)=\nu k\lim_{t\to\infty}\frac{1}{\nu kt}\tilde{N}\left(kt,\frac{1}{\nu}\right)=\nu k\lim_{t\to\infty}\frac{\tilde{N}(kt,1/\nu)}{E\left[\tilde{N}(kt,1/\nu)\right]}$$

https://doi.org/10.26219/heal.aueb.5966

We use now the Chebishev Inequality in order to find $\lim\limits_{t\to\infty} \dfrac{\tilde{N}(kt,1/\nu)}{E\left[\tilde{N}(kt,1/\nu)\right]}$.

We have

$$E\left\{\frac{\tilde{N}(kt,1/\nu)}{E\left[\tilde{N}(kt,1/\nu)\right]}\right\}=1 .$$

$$\operatorname{var}\left\{\frac{\tilde{N}(kt,1/\nu)}{E\left[\tilde{N}(kt,1/\nu)\right]}\right\}=\frac{\operatorname{var}\left\{\tilde{N}(kt,1/\nu)\right\}}{E^2\left[\tilde{N}(kt,1/\nu)\right]}=\frac{\nu kt(1+\nu)}{(\nu kt)^2}=\frac{1+\nu}{\nu kt}\xrightarrow[t\to\infty]{}0$$

Hence

$$p\left\{\left|\frac{\tilde{N}(kt,1/\nu)}{E\left[\tilde{N}(kt,1/\nu)\right]}-1\right|\geq\varepsilon\right\}\leq\frac{1+\nu}{\varepsilon\nu kt}.\xrightarrow[t\to\infty]{}0$$

which tells that

$$\frac{\tilde{N}(kt,1/\nu)}{E\left[\tilde{N}(kt,1/\nu)\right]}\xrightarrow[t\to\infty]{p.}1$$

and

$$\frac{1}{t}N(t)\xrightarrow[t\to\infty]{p.}\nu k$$

Considering this result and the fact that if ν is $BetaII(a,\rho)$-distributed then $E(\nu)=\dfrac{a}{\rho-1}$, we take $E(\nu k)=\dfrac{ak}{\rho-1}$. Hence, the random variable $\overline{N}=\nu k$ is the individual intensity of the Generalized Waring process.

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

The relations (2.3.1) and (2.4.1) give the transition probabilities of the Generalized Waring process.

$$P\{N(s+t)=n\mid N(s)=m\}=\frac{\Gamma(a+n)}{\Gamma(a+m)}\frac{(kt)_{(n-m)}}{(n-m)!}\frac{(\rho+ks)_{(a+m)}}{(\rho+ks+kt)_{(a+n)}}$$

We denote

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

https://doi.org/10.26219/heal.aueb.5966

and recall that it represents the probability that a process presently in state m will be in state n a time t later. This probability in this case depends on the present time because the Generalized Waring process is a non-homogenous Markov process.

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

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

In order to find the *forward* and *backward* Kolmogorov differential equations we have used the ideas used by Ross (1996, p. 240-242) to obtain the respective equations of an homogenous Markov process.

We obtain the *forward* Kolmogorov differential equations starting the calculations from the equations

$$p_{m,n}(s,t+h)=\sum_{i=m}^{n}p_{m,i}(s,t)p_{i,n}(t,t+h),\qquad \text{for } s\le t,\ m\le n,\ h\ge 0$$

We find

$$p'_{m,n}(s,t)=\sum p_{m,i}(s,t)\lim_{h\to 0}\frac{p_{i,n}(t,t+h)}{h}-\lim_{h\to 0}\left(1-\frac{p_{nn}(t,t+h)}{h}\right)p_{m,n}(s,t)$$

We calculate

$$\lim_{h\to 0}\frac{p_{i,n}(t,t+h)}{h}=\frac{\Gamma(a+n)}{\Gamma(a+i)}\lim_{h\to 0}\frac{(kh)_{(n-i)}}{h(n-i)!}\frac{(\rho+kt)_{(a+i)}}{(\rho+kt+kh)_{(a+n)}}$$

$$=\begin{cases}\dfrac{\Gamma(a+n)}{\Gamma(a+n-1)}k\dfrac{(\rho+kt)_{(a+n-1)}}{(\rho+kt)_{(a+n)}} & n\text{-}i=1\\[2ex] \dfrac{\Gamma(a+n)}{\Gamma(a+i)}\dfrac{k_{(n-i-1)}}{(n-i)!}\dfrac{(\rho+kt)_{(a+i)}}{(\rho+kt)_{(a+n)}} & \text{otherwise}\end{cases}$$

$$=\begin{cases}\dfrac{k(a+n-1)}{(a+\rho+kt+n-1)} & n\text{-}i=1\\[2ex] \dfrac{\Gamma(a+n)}{\Gamma(a+i)}\dfrac{k}{(n-i)(n-i-1)}\dfrac{(\rho+kt)_{(a+i)}}{(\rho+kt)_{(a+n)}} & n\text{-}i>1\end{cases}$$

and

https://doi.org/10.26219/heal.aueb.5966

$$\lim_{h\to 0}\left(\frac{1-p_{n,n}(t,t+h)}{h}\right)=\lim_{h\to 0}\frac{1}{h}\left[1-\frac{(\rho+kt)_{(a+n)}}{(\rho+kt+kh)_{(a+n)}}\right]$$

$$=\lim_{h\to 0}\frac{1-\dfrac{(\rho+kt)_{(a+n)}}{(\rho+kt+kh)_{(a+n)}}}{h}$$

$$=\lim_{h\to 0}\frac{(\rho+kt)_{(a+n)}\cdot k\cdot\sum_{i=0}^{a+n-1}\dfrac{1}{\rho+kt+kh+i}}{(\rho+kt+kh)_{(a+n)}}$$

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

If we denote

$$q_{n-1,n}(t)=\frac{k(a+n-1)}{(a+\rho+kt+n-1)}\quad , \; q_{i,n}(t)=\frac{\Gamma(a+n)}{\Gamma(a+i)}\frac{k}{(n-i)(n-i-1)}\frac{(\rho+kt)_{(a+i)}}{(\rho+kt)_{(a+n)}}\qquad i<n$$

and $v_n(t)=k\sum_{i=0}^{a+n-1}\dfrac{1}{\rho+kt+i}$ then the *forward* Chapman-Kolmogorov equations for the Generalized Waring process are:

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

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

The *backward* equations, follows from the Chapman-Kolmogorov equations with $\tau=s+h$. Hence we start the calculations from the following equations

$$p_{m,n}(s,t)=\sum_{i=m}^{n}p_{m,i}(s,s+h)p_{i,n}(s+h,t),\qquad \text{for } s+h\le t,\ m\le n,\ h\ge 0$$

We find

$$\frac{p_{m,n}(s+h,t)-p_{m,n}(s,t)}{h}=p_{m,n}(s+h,t)\left(\frac{1-p_{m,m}(s,s+h)}{h}\right)-\sum_{i=m+1}^{n}\frac{p_{m,i}(s,s+h)}{h}p_{i,n}(s+h,t)$$

We calculate similarly and find

https://doi.org/10.26219/heal.aueb.5966

$$\lim_{h\to 0}\frac{p_{m,i}(s,s+h)}{h}=\begin{cases}\dfrac{k(a+m)}{(a+\rho+ks+m)} & i-m=1\\[2ex] \dfrac{\Gamma(a+i)}{\Gamma(a+m)}\dfrac{k}{(i-m)(i-m-1)}\dfrac{(\rho+ks)_{(a+m)}}{(\rho+ks)_{(a+i)}} & \textit{otherwise}\end{cases}$$

and

$$\lim_{h\to 0}\left(\frac{1-p_{m,m}(s,s+h)}{h}\right)=k\cdot\sum_{i=0}^{a+m-1}\frac{1}{\rho+ks+i}$$

If we denote

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

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

and

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

then the *backward* equations for the Generalized Waring process are:

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

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

https://doi.org/10.26219/heal.aueb.5966

CHAPTER 5

SOME PARTICULAR CASES OF THE GENERALIZED WARING PROCESS

We refer to the relation between the Pólya, the Poisson and the Generalized Waring distributions, treated by Irwin (1968), to come to the conclusion that the Pólya and the Poisson processes are limiting cases of the Generalized Waring process. So the classical theory of those processes is a particular case of that presented here. The new definitions of them as special cases are given in this section. The Markovian property for both of those processes, the calculations of the transition probabilities and the Chapman-Kolmogorov equations, the proof that they are birth processes, the property of having independent increments of a Poisson process and a known equivalent definition of it, are provided here, using these definitions. A brief discussion about continuous analogue is intended in the last paragraph of this chapter. Two cases are considered: when $a \to \infty$ and when both $a, \rho \to \infty$. Both in these cases the continuous analogue of the Generalized Waring Process can be regarded as limiting forms of it.

5.1 The Pólya and the Poisson processes as special cases of the Generalized Waring process.

It is shown (Irwin 1968) that, if $k \to \infty$ and $\rho \to \infty$ so that $w_k = \dfrac{k}{k+\rho}$ remains constant the Generalized Waring distribution tends to the negative binomial distribution with generating function $\left\{\dfrac{1}{u_k} - \dfrac{w_k \theta}{u_k}\right\}^{-a}$ where $u_k = 1 - w_k = \dfrac{\rho}{k+\rho}$.

Analogous to the proof used by Irwin (1968) to obtain the Pólya distribution as a limiting case of the Generalized Waring distribution, the following theorem can now be prove:

Theorem 5.1.1 If $k \to \infty$ and $\rho = c \cdot k$ where $c > 0$ is a constant i.e. $w_k(t) = \dfrac{kt}{kt+\rho} = \dfrac{t}{t+c}$, then, the Generalized Waring process tends to the Negative binomial process with generating function

https://doi.org/10.26219/heal.aueb.5966

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

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

Hence, in this case

$$\lim_{k\to\infty}\frac{\rho_{(kt)}}{(\rho+a)_{(kt)}}\frac{a_{(n)}(kt)_{(n)}}{(\rho+a+kt)_{(n)}}\frac{1}{n!}=\binom{a+n-1}{n}\left(\frac{c}{t+c}\right)^{a}\left(\frac{t}{t+c}\right)^{n} \tag{5.1.1}$$

Hence, we can give the following definition of the Pólya process.

Definition 5.1.1 The Generalized Waring process with parameters $(a,\ k,\ \rho)$ in which $k\to\infty$ and $\rho=c\cdot k$ where $c>0$ is a constant, is called **Pólya process** with parameters $\left(a,\dfrac{1}{c}\right)$.

It is shown also that the Poisson process is a limiting form of a Pòlya process. Let us consider the above Pòlya process, and let us assume that $c\to\infty$ and $a=\lambda\cdot c$ where $\lambda>0$ is a constant.

In this case we find

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

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

$$=\lim_{c\to\infty}\left[\frac{(\lambda c)_{(n)}(\lambda c-1)!}{n!(\lambda c-1)!}\left(\left(1+\frac{1}{c/t}\right)^{c/t}\right)^{-\lambda t}\frac{t^n}{(t+c)^n}\right]$$

https://doi.org/10.26219/heal.aueb.5966

$$= \lim_{c\to\infty}\left[\frac{(\lambda c)_{(n)}}{(t+c)^n}\left(\left(1+\frac{1}{c/t}\right)^{c/t}\right)^{-\lambda t}\frac{t^n}{n!}\right]$$

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

which prove the following theorem

Theorem 5.1.2 Consider now the Pólya process, defined above, and let $c \to \infty$ and $a = \lambda \cdot c$, where $\lambda > 0$ is a constant. Then,

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

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

So, similar with above, we can define a Poisson process as a special case of the Pólya process, and then we can define it as a special case of a Generalized Waring process.

Definition 5.1.2 The Pólya process with parameters $\left(a, \frac{1}{c}\right)$ in which $c \to \infty$ and $a = \lambda \cdot c$ where $\lambda > 0$ is a constant, is called Poisson process with rate λ.

Definition 5.1.3 The Generalized Waring process with parameters $(a,\ k,\ \rho)$ in which $k \to \infty$, $\frac{\rho}{k} \to \infty$ and $a = \lambda \cdot \frac{\rho}{k}$ where $\lambda > 0$ is a constant is called Pólya process with parameters $\left(a, \frac{1}{c}\right)$.

It is clear that the definitions 5.1.2 and 5.1.3 are equivalent.

5.2 The moments and some other properties

In this paragraph we find the first two moments of the Pòlya process and the Poisson process, using the respective results for the Generalized Waring process.

Theorem 5.2.1 Let N be Pólya process with parameters $\left(a, \frac{1}{c}\right)$, then for any $t \geq 0$

https://doi.org/10.26219/heal.aueb.5966

1. $E[N(t)] = \frac{a}{c} t$

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

Proof

From the definition 5.1.1 the Pòlya Process $\left(a, \frac{1}{c}\right)$ is a Generalized Waring process with parameters (a, k, ρ) where $k \to \infty$ and $\rho = c \cdot k$. For this reason we only need to find the limit of the expressions of $E[N(t)]$ and $Var[N(t)]$ found for a Generalized Waring process, when $k \to \infty$ and $\rho = c \cdot k$. We find

$\lim_{k\to\infty} E[N(t)] = \lim_{k\to\infty} \frac{akt}{\rho - 1} = \frac{a}{c} t$ which proves 1.

We calculate first

$$\lim_{k\to\infty} \sigma^2_{\Lambda(t)} = \lim_{k\to\infty} \frac{akt}{(\rho-1)(\rho-2)} = 0 \ , \ \lim_{k\to\infty} (kt)^2 \sigma^2_{\nu} = \lim_{k\to\infty} \frac{(kt)^2 a(a+\rho-1)}{(\rho-1)^2(\rho-2)} = \frac{at^2}{c^2}$$

and $\lim_{k\to\infty} \sigma^2_{R} = \lim_{k\to\infty} \frac{akt}{(\rho-1)} = \frac{at}{c}$

Then from $Var[N(t)] = \sigma^2_{\alpha(t)} + (kt)^2 \sigma^2_{\nu} + \sigma^2_{R}$ we prove 2.

Theorem 5.2.2 Let N be a Poisson Process with parameter λ. Then for any $t \geq 0$

$$E[N(t)] = Var[N(t)] = t\lambda$$

Proof

From the definition 5.1.2 the Poisson Process with parameter λ is a Pòlya process with parameters $\left(a, \frac{1}{c}\right)$ in which $c \to \infty$ and $a = \lambda \cdot c$. For this reason we only need to find the limit of $E[N(t)]$ and $Var[N(t)]$, found above, when $c \to \infty$ and $a = \lambda \cdot c$.

We find

$$\lim_{a\to\infty} E[N(t)] = \lim_{a\to\infty} \frac{a}{c} t = \lambda t$$

$$\lim_{a\to\infty} Var[N(t)] = \lim_{a\to\infty} \left[\frac{a}{c} t + \frac{a}{c^2} t^2\right] = \lambda t$$

which proves the theorem.

https://doi.org/10.26219/heal.aueb.5966

5.3 The transition probabilities of the Pòlya and Poisson processes.

The Pòlya and the Poisson Processes defined by us are both stationary Markov processes. Let us find the transition probabilities of them.

We start from the Pólya process with parameters $\left(a, \frac{1}{c}\right)$. We need only to calculate

$$\lim_{k\to\infty} \frac{\Gamma(a+n)}{\Gamma(a+m)} \frac{(kh)_{(n-m)}}{(n-m)!} \frac{(\rho+kt)_{(a+m)}}{(\rho+kt+kh)_{(a+n)}}$$

We find

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

$$= \lim_{k\to\infty} \begin{cases} \dfrac{(\rho+kt)_{(a+n)}}{(\rho+kt+kh)_{(a+n)}} & n = m \\ (a+m)\dfrac{kh(\rho+kt)_{(a+m)}}{(\rho+kt+kh)_{(a+m+1)}} & n = m+1 \\ \dfrac{\Gamma(a+n)}{\Gamma(a+m)(n-m)!} \dfrac{(kh)_{(n-m)}}{(\rho+kt+kh+a+m)_{(n-m)}} \dfrac{(\rho+kt)_{(a+m)}}{(\rho+kt+kh)_{(a+m)}} & n > m+1 \end{cases}$$

$$= \begin{cases} \left(\dfrac{c+t}{c+t+h}\right)^{(a+n)} & n = m \\ (\mathrm{a+m})\dfrac{h(c+t)^{(a+m)}}{(c+t+h)^{(a+n1)}} & n = m+1 \\ \dfrac{\Gamma(a+n)}{\Gamma(a+m)(n-m)!} \dfrac{h^{(n-m)}(c+t)^{(a+m)}}{(c+t+h)^{(a+n)}} & n > m+1 \end{cases}$$

Now, for the Poisson process with parameter λ, we need to calculate

$$\lim_{c\to\infty} \begin{cases} \left(\dfrac{c+t}{c+t+h}\right)^{(a+n)} & n = m \\ (a+m)\dfrac{h(c+t)^{(a+m)}}{(c+t+h)^{(a+n1)}} & n = m+1 \\ \dfrac{\Gamma(a+n)}{\Gamma(a+m)} \dfrac{1}{(n-m)!} \dfrac{h^{(n-m)}(c+t)^{(a+m)}}{(c+t+h)^{(a+n)}} & n > m+1 \end{cases}$$

https://doi.org/10.26219/heal.aueb.5966

We find

$$\lim_{c\to\infty}\left(\frac{c+t}{c+t+h}\right)^{(a+n)}=\lim_{c\to\infty}\left[\left(1\Big/\left(1-\left(-\frac{h}{c+t}\right)\right)\right)^{\frac{c+t}{h}}\right]^{-\lambda h}\left(\frac{c+t}{c+t+h}\right)^{-\lambda h}\left(\frac{c+t}{c+t+h}\right)^{t}=\exp(-\lambda h)$$

$$\lim_{c\to\infty}(a+m)\frac{h(c+t)^{(a+m)}}{(c+t+h)^{(a+n)}}=\lim_{c\to\infty}\frac{h(a+m)}{(c+t+h)}\left(\frac{c+t}{c+t+h}\right)^{(a+m)}=\lambda h\exp(-\lambda h)$$

$$\lim_{c\to\infty}\frac{h^{(n-m)}(a+m)_{(n-m)}}{(n-m)!(c+t+h)^{(n-m)}}\left(\frac{c+t}{c+t+h}\right)^{(a+m)}=\frac{(\lambda h)^{(n-m)}}{(n-m)!}\exp(-\lambda h)$$

5.4 The Chapman-Kolmogorov equation for the Pólya and Poisson processes.

We can find the forward Kolmogorov differential equations of the Pòlya process from the respective equations of the generalized Waring process

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

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

We need to calculate the $\lim_{k\to\infty}q_{i,n}(t),\ m\le i<n$ and $\lim_{k\to\infty}\nu_n(t)$ where $\rho=c\cdot k$. We find

$$\lim_{k\to\infty}q_{n-1,n}(t)=\lim_{k\to\infty}\frac{k(a+n-1)}{(a+\rho+kt+n-1)}=\frac{(a+n-1)}{c+t}$$

$$\lim_{k\to\infty}q_{i,n}(t)=\lim_{k\to\infty}\left[\frac{\Gamma(a+n)}{\Gamma(a+i)}\frac{k}{(n-i)(n-i-1)}\frac{(\rho+kt)_{(a+i)}}{(\rho+kt)_{(a+n)}}\right]=0\quad i<n\text{-}1$$

and

$$\lim_{k\to\infty}\nu_n(t)=\lim_{k\to\infty}\left[k\sum_{i=0}^{a+n-1}\frac{1}{\rho+kt+i}\right]=\frac{a+n}{c+t}$$

If we denote $k_n(t)=\lim_{k\to\infty}\nu_n(t)=\dfrac{a+n}{c+t}$ we see that $\lim_{k\to\infty}q_{n-1,n}(t)=k_{n-1}(t)$

Hence, the forward Chapman-Kolmogorov equations for the Pòlya Process with parameters $\left(a,\dfrac{1}{c}\right)$ are

https://doi.org/10.26219/heal.aueb.5966

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

$$\frac{\partial p_{m,n}(s,t)}{\partial t} = -k_n(t)p_{m,n}(s,t) + k_{n-1}(t)p_{m,n-1}(s,t), \quad n > m$$

In the same way we find the backward Chapman-Kolmogorov differential equations of the Pòlya process:

$$\lim_{k\to\infty} q_{m,m+1}(s) = \lim_{k\to\infty} \frac{k(a+m)}{(a+\rho+ks+m)} = \frac{a+m}{c+s} = k_m(s)$$

$$\lim_{k\to\infty} \nu_m(s) = \lim_{k\to\infty}\left(k\sum_{i=0}^{a+m-1}\frac{1}{\rho+ks+i}\right) = \frac{a+m}{c+s} = k_m(s)$$

So, these equations are:

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

$$\frac{\partial p_{m,n}(s,t)}{\partial s} = k_m(s)\left[p_{m,n}(s,t) - p_{m+1,n1}(s,t)\right], \quad n > m$$

So, we have proved the following:

Theorem 5.4.1 The Pòlya Process is a stationary non-homogenous birth process with the transition intensities $k_n(t) = \frac{a+n}{c+t}$ where $\left(a, \frac{1}{c}\right)$ are the parameters of it.

Remark 5.4.1

The transition intensities of the Pólya process can be presented in the following form:

$$P_{m,n}(t,t+h) = \begin{cases} 1 - k_m(t)h + o_1(h) & n = m \\ k_m(t)h + o_2(h) & n = m+1 \\ o_3(h) & n > m+1 \end{cases}$$

where $o_1(h),\ o_2(h),\ o_3(h)$[e] are respectively

[e] The symbol $o(h)$ means that $\lim_{h\to 0}\frac{o(h)}{h} = 0$

https://doi.org/10.26219/heal.aueb.5966

$$o_1(h) = \left(\frac{c+t}{c+t+h}\right)^{(a+m)} + \frac{a+m}{c+t}h - 1$$

$$o_2(h) = \frac{(a+m)h}{c+t}\left[\left(\frac{c+t}{c+t+h}\right)^{(a+m)} - 1\right]$$

$$o_3(h) = \frac{G(a+n)}{G(a+m)}\frac{h^{(n-m)}(c+t)^{(a+m)}}{(c+t+h)^{(a+n)}}$$

In the same way, we can find the *forward* and the *backward* Chapman-Kolmogorov differential equations of the Poisson Process with parameter λ, using the respective equations of the Pòlya Process. We only need to find now, the $\lim_{a\to\infty} k_n(t)$ where $c = \lambda \cdot a$:

$$\lim_{a\to\infty} k_n(t) = \lim_{a\to\infty} \frac{a+n}{c+t} = \lambda$$

So, we have proved the following:

Theorem 5.4.2 The Poisson Process is a stationary homogenous birth process with the transition intensities $k_n(t) = \lambda$ where λ is the rate of it.

Remark 5.4.2

The transition intensities of the Poisson process can be presented in the following form:

$$P_{m,n}(t,t+h) = \begin{cases} 1 - \lambda h + o_1(h) & n = m \\ \lambda h + o_2(h) & n = m+1 \\ o_3(h) & n > m+1 \end{cases}$$

where $o_1(h)$, $o_2(h)$, $o_3(h)$, are respectively

$$o_1(h) = \exp(-\lambda h) + \lambda h - 1$$
$$o_2(h) = \lambda h[\exp(-\lambda h) - 1]$$
$$o_3(h) = \frac{(\lambda h)^{(n-m)}}{(n-m)!}\exp(-\lambda h)$$

Theorem 5.4.3 The Poisson process has independent increments.

Proof

Let $Y = \{Y(t),\ t \geq 0\}$ be a Poisson process. We need only to prove that

https://doi.org/10.26219/heal.aueb.5966

$$P\left(Y(t+h)-Y(t)=n/Y(s)=n_s, s\le t\right)=P\left(Y(t+h)-Y(t)=n\right)$$

It follows from the Markovian property that

$$P\left(Y(t+h)-Y(t)=n/Y(s)=n_s, s\le t\right)=P\left(Y(t+h)-Y(t)=n/Y(t)=n_t\right)$$

But $\quad P\left(Y(t+h)-Y(t)=n/Y(t)=n_t\right)=P\left(Y(t+h)=n+n_s/Y(t)=n_t\right)$

$$=\frac{(\lambda h)^n\exp(-\lambda h)}{n!}=P\left(Y(t+h)-Y(t)=n\right)$$

We are now in a position to give an alternative definition of the Poisson process which is the same to the definition 3.1 of a Poisson Process, given in the book of Ross (1993, p.210-211).

Definition 5.4.1 The counting process $\{Y(t), t\ge 0\}$ is said to be a Poisson process having rate λ, $\lambda>0$, if

(I) $N(0)=0$

(II) It has independent increments

(III) For all $s.t\ge 0$

$$p\{N(t+s)-N(s)=n\}=\frac{(\lambda t)^n\exp(-\lambda t)}{n!},\quad n=0,1,\ldots$$

5.5 Continuous analogues. Two other limiting forms of the Generalized Waring Process.

We refer here to the continuous analogue of the Generalized Waring Distribution obtained by Irwin(1975c). He has shown that the continuous analogue of the Generalized Waring Distribution is Pearson's Type VI. Following him we tried to obtain the continuous analogue of the Generalized Waring Process. If we denote it by $X=\{X(t),\ t\ge 0\}$, we find that for every t the probability density function of $X(t)$ is

$$y_t(X)=C_t\left(X+a_{2_t}\right)^{q_{2_t}}\left(X+a_{1_t}\right)^{-q_{1_t}} \tag{5.5.1}$$

where a_{1_t}, a_{2_t} are the roots of the equation

$$X^2+\frac{1}{2}\{\rho+2(\alpha+kt)+1\}X+\frac{1}{2}\alpha kt=0$$

$$C_t = \frac{\rho_{(kt)}}{(a+\rho)_{(kt)}}$$

and

$$q_{2_t} = \frac{\alpha kt + (\rho+1)a_{2_t}}{a_{1_t} - a_{2_t}},\; q_{1_t} = \frac{\alpha kt + (\rho+1)a_{1_t}}{a_{1_t} - a_{2_t}}$$

In the particular case when $\alpha \to \infty$, by a change of the origin and a change of scale which makes the interval between successive values of the variate tend to zero, (writing ξ instead of $\dfrac{X + a_{2_t}}{a_{1_t} - a_{2_t}}$) the continuous analogue becomes

$$y_t(\xi) = \frac{\Gamma(\rho + kt)}{\Gamma(kt)\Gamma(\rho)}\xi^{kt-1}(1+\xi)^{-(\rho+kt)} \tag{5.5.2}$$

This form can be regarded as a limiting form of the Generalized Waring process. We have to mention here that the above transformation of the variate depends on t. Hence, each realization of this particular case of the Generalized Waring process is a random variable whose distribution can take the form of a *BetaII* one with parameters kt and ρ.

If in (5.5.2) we write $\xi = X/c\rho$ and let $\rho \to \infty$, c being finite, following Irwin(1975c) again, we obtain

$$y_t = \frac{1}{\Gamma(kt)}\left(\frac{X}{c}\right)^{kt-1} e^{-X/c} \tag{5.5.3}$$

It can be considered as a limiting form of the Generalized Waring process, too.This is also the limiting form of the Negative Binomial $\left\{\left(1-q_a A^{\lambda}\right)/p_a\right\}^{-kt}$ when $\rho \to \infty$, $a \to \infty$, $q_a \to 1$ (See Irwin 1975c).

CHAPTER 6

SOME ALTERNATIVE GENESIS SCHEMES

The Generalized Waring process has been defined as a non-homogenous stationary Markov Process arising as a Beta mixture of the Negative Binomial process in a "proneness" context. In the sequel we consider two further genesis schemes where the underlying mechanism is indicative of *contagion* rather than proneness in the sense of Irwin (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 $f_{\nu}(x,t)$, with ν referring to the individual's risk exposure.

6.1 A mixed Pólya process.

Assuming that $f_{\nu}(x,\, t) = \dfrac{k + x}{(1/\nu) + t} = \nu \cdot \dfrac{k + x}{1 + \nu t}$, the distribution of accidents for each t (λ fixed) is Negative binomial with parameters $(k,\ 1/\nu t)$ (the accident pattern is described in that case by a Pólya process). As shown by Xekalaki (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., $\lambda \sim ae^{-a\lambda}$, $a > 0$ for $t = 1$. But for $t \neq 1$, we obtained

$$P(N = n) = P_n(t) = \int_0^{\infty} \binom{\gamma + n - 1}{n} (\nu t)^n (1 + \nu t)^{-(\gamma + n)} \frac{\Gamma(p + a)}{\Gamma(a)\Gamma(p)} \nu^{a-1} [1 + \nu]^{-(a+p)} d\nu$$

$$= \frac{\Gamma(p + a)}{\Gamma(a)\Gamma(p)} \binom{\gamma + n - 1}{n} \frac{1}{t^a} \int_0^{\infty} (\nu t)^{a+n-1} (1 + \nu t)^{-(\gamma+n)} \left[1 + \nu t\left(1 - \left(1 - \frac{1}{t}\right)\right)\right]^{-(a+p)} d\nu t$$

$$= \frac{\Gamma(p + a)}{\Gamma(a)\Gamma(p)} \binom{\gamma + n - 1}{n} \frac{1}{t^a} \sum_{r=0}^{\infty} \frac{(a + p)_{(r)}}{r!} \left(1 - \frac{1}{t}\right)^r \int_0^{\infty} (\nu t)^{a+n+r-1} (1 + \nu t)^{-(a+p+\gamma+n+r)} d\nu t$$

$$= \frac{\Gamma(p + a)}{\Gamma(a)\Gamma(p)} \binom{\gamma + n - 1}{n} \frac{1}{t^a} \sum_{r=0}^{\infty} \frac{(a + p)_{(r)}}{r!} \left(1 - \frac{1}{t}\right)^r \frac{\Gamma(a + n + r)\Gamma(p + \gamma)}{\Gamma(a + p + \gamma + n + r)}$$

https://doi.org/10.26219/heal.aueb.5966

$$=\frac{\Gamma(p+a)}{\Gamma(a)\Gamma(p)}\binom{\gamma+n-1}{n}\frac{1}{t^{a}}\sum_{r=0}^{\infty}\frac{(a+p)_{(r)}}{r!}\left(1-\frac{1}{t}\right)^{r}\frac{\Gamma(a+n)(a+n)_{(r)}\Gamma(p+\gamma)}{\Gamma(a+p+\gamma+n)(a+p+\gamma+n)_{(r)}}$$

$$=\frac{\Gamma(p+a)}{\Gamma(a)\Gamma(p)}\frac{\Gamma(a+n)\Gamma(p+\gamma)}{\Gamma(a+p+\gamma+n)}\binom{\gamma+n-1}{n}\frac{1}{t^{a}}\sum_{r=0}^{\infty}\frac{(a+p)_{(r)}(a+n)_{(r)}}{(a+p+\gamma+n)_{(r)}r!}\left(1-\frac{1}{t}\right)^{r}$$

$$=\frac{\Gamma(p+a)}{\Gamma(a)\Gamma(p)}\frac{\Gamma(a+n)\Gamma(p+\gamma)}{\Gamma(a+p+\gamma+n)}\frac{\Gamma(\gamma+n)}{\Gamma(\gamma)}\frac{(1/t)^{a}}{n!}\sum_{r=0}^{\infty}\frac{(a+p)_{(r)}(a+n)_{(r)}}{(a+p+\gamma+n)_{(r)}r!}\left(1-\frac{1}{t}\right)^{r}$$

$$=\frac{p_{(\gamma)}}{(a+p)_{(\gamma)}}\frac{a_{(n)}\gamma_{(n)}}{(a+p+\gamma)_{(n)}}\frac{(1/t)^{a}}{n!}\sum_{r=0}^{\infty}\frac{(a+p)_{(r)}(a+n)_{(r)}}{(a+p+\gamma+n)_{(r)}r!}\left(1-\frac{1}{t}\right)^{r}$$

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

(6.1.1)

where

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

6.2 The Markovian property of the mixed Pólya process.

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 (6.1.1), is a birth process, but not of a Generalized Waring form.

It is clear that a birth process is determined by its intensities and we know also that the one-dimensional marginal distributions determine the intensities through the following relations (Lundberg (1964, p.18)):

$$P'_0(t)=-k_0(t)P_0(t),$$
$$P'_n(t)=-k_n(t)P_n(t)+k_{n-1}(t)P_{n-1}(t),\quad n>0$$

So we have to prove only the existence of it.

Consider the following integral

$$I_n(t)=\int_0^{\infty}(vt)^n(1+vt)^{-(\gamma+n)}v^{a-1}[1+v]^{-(a+p)}dv \qquad (6.2.1)$$

and notice that

$$P_n(t)=\frac{\Gamma(p+a)}{\Gamma(a)\Gamma(p)}\frac{(\gamma+n-1)!}{n!\Gamma(\gamma)}I_n(t)$$

We calculate

$$I'_n(t) = n\int_0^\infty (vt)^{n-1}(1+vt)^{-(\gamma+n)}v^a[1+v]^{-(a+p)}dv + (-(\gamma+n))\int_0^\infty (vt)^n(1+vt)^{-(\gamma+n+1)}v^a[1+v]^{-(a+p)}dv$$
$$= nJ_{n-1}(t) + (-(\gamma+n))J_n(t)$$

where

$$J_n(t) = \int_0^\infty (vt)^n(1+vt)^{-(\gamma+n+1)}v^a[1+v]^{-(a+p)}dv \tag{6.2.2}$$

Hence, differentiation of the probability yields

$$P'_n(t) = \frac{\Gamma(p+a)}{\Gamma(a)\Gamma(p)}\frac{(\gamma+n-2)!}{(n-1)!\Gamma(\gamma)}I_{n-1}(t)\frac{(\gamma+n-1)J_{n-1}(t)}{I_{n-1}(t)} + \frac{\Gamma(p+a)}{\Gamma(a)\Gamma(p)}\frac{(\gamma+n-1)!}{n!\Gamma(\gamma)}I_n(t)\frac{-(\gamma+n)J_n(t)}{I_n(t)}$$
$$= -P_n(t)k_n(t) + P_{n-1}(t)k_{n-1}(t)$$

It follows from the above relations, provided N is birth process, that,

$$k_n(t) = \frac{(\gamma+n)J_n(t)}{I_n(t)} \tag{6.2.3}$$

and

$$P'_0(t) = \frac{\Gamma(p+a)}{\Gamma(a)\Gamma(p)}I_0(t)\frac{-(\gamma)J_0(t)}{I_0(t)}$$
$$= -P_0(t)k_0(t)$$

where

$$k_0(t) = \frac{\gamma J_0(t)}{I_0(t)}$$

It follows from (6.2.1), (6.2.2) and (6.2.3) that the intensities are positive, continuous and differentiable for $t>0$. Thus there exists a well-defined birth process with its intensities given by (6.2.3).

We denote that it is also difficult to calculate the values of the function $(1/t)^a F\left(a+p,\ a+n,\ a+p+\gamma+n;\ 1-\frac{1}{t}\right)$, and the respective probabilities.

6.3 A non-markovian stochastic process of Generalized Waring form.

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 (Irwin, 1941) and Generalized Waring with parameters $\left(\frac{k}{m},\ 1,\ \frac{a}{mt}\right)$, when $\lambda \sim ae^{-a\lambda}$, $a>0$ (Xekalaki, 1981).

https://doi.org/10.26219/heal.aueb.5966

Following Irwin (1941) and Arbous and Kerrich (1951), and considering the increment $Y_t(h) = N(t+h) - N(t)$ at the moment t, as a first step, we verify that during the period h, $h+dh$ for an individual of liability λ having $Y_t(h) = y$, given that $N(t) = x$, this increment, can take the values:

$$\left.\begin{array}{rl} 0 & \text{with probability } 1-\varphi_\lambda(y,h)dh \\ 1 & \text{with probability } \varphi_\lambda(y,h)dh \\ >1 & \text{with probability } 0 \end{array}\right\} \qquad (6.3.1)$$

where

$$\varphi_\lambda(y,h) = f_\lambda(y+x,h) = \lambda(k+my+mx) = \lambda((k+mx)+my)$$

Hence, it is easy to verify 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 $\left(\frac{k}{m}+x,\ \frac{1}{1-e^{-\lambda mh}}\right)$ when λ is fixed, and a Generalized Waring distribution with parameters $\left(\frac{k}{m}+x,\ 1,\ \frac{a}{mh}\right)$, when $\lambda \sim ae^{-a\lambda}$, $a > 0$.

So, in this case,

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

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

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.

https://doi.org/10.26219/heal.aueb.5966

References

Arbous, A. G. and Kerrich, J. E. (1951). Accident statistics and the concept of accident proneness. *Biometrics* **7**, 340-432.

Bates, G. E. and Neyman, J. (1952). Contributions to the theory of accident proneness. *University of California. Publications on Statistics* **I**. 215-275.

Cane, V. R. (1974). The concept of the accident proneness. *Bulletin of The Institute of Mathematics. Bulgarian Academy of Sciences* **15**. 183-188.

Cane, V. R. (1977). A class of non-identifiable stochastic models *Journal of Applied Probability* **14**. 475-482.

Chatfield, C. and Theobald, C.M. (1973). Mixtures and Random Sums, *The Statistician,* **22(3)** Printed in Great Britain, 281-287.

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

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

Greenwood, M. and Woods, H. M. (1919). *On the Incidence of the Industrial Accidents upon Individuals with special Reference to Multiple Accidents.* Report of the Industrial Fatigue Research Board. **4**, 1-28. London: His Majesty's Stationery Office.

Greenwood, M. and Woods, H. M. (1920). An Inquiry into the nature of frequency distributions representative of multiple happenings with particular reference to the occurrence of multiple attack of disease or repeated accidents. *Journal of the Royal Statistical Society,* **83**, 255-279.

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

Irwin, J.I. (1975a). "The Generalized Waring Distribution. Parts I.", *Journal of the Royal Statistical Society,* **A 138**, 18-31.

Irwin, J.I. (1975b). "The Generalized Waring Distribution. Parts II.", *Journal of the Royal Statistical Society,* **A 138**, 204-227.

Irwin, J.I. (1975c). "The Generalized Waring Distribution. Parts III", *Journal of the Royal Statistical Society,* **A 138**, 374-384.

Irwin, J.I.(1968). "The Generalized Waring Distribution. Applied to accident theory", *Journal of the Royal Statistical Society,* **A 131**, 205-225.

Johnson, N.L. and Kotz, S. (1969). *Discrete Distributions.* Boston: Houghton Miflin Co.

Kemp, C. D. (1967) On a Contagious Distribution Suggested for Accident Data. *Biometrics* **23**, 241-255.

Kemp, C. D. (1970) 'Accident Proneness' and Discrete Distribution Theory. *In Random Counts in Scientific Work*, **2** G. Patiled., 41-65. State College: Pensylvania State University Press.

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.

Ross, S.M.(1993). *Introduction to the Probability Model,* Willey.

Ross, S.M.(1996). *Stochastic Processes*, Willey.

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

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.

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

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.

Xekalaki, E. (1986). The Multivariate Generalized Waring Distribution, *Communications in Statistics - Theory & Methods,* **15(3)** 1047-1064.

Zografi, M and Xekalaki, E. (2001) The Generalized Waring Process. Technical Report, **121,** Department of Statistics, Athens University of Economics and Business.

https://doi.org/10.26219/heal.aueb.5966

Δωρεά