Instituto Tecnológico y de Estudios Superiores de Monterrey

Campus Monterrey

División de Electrónica, Computación, Información, y Comunicaciones

Programa de Graduados

URN MODELING FOR HEAVY-TAILED PHENOMENA

THESIS

Presented as a partial fulfillment of the requirements for the degree of

Master of Science in Electronic Engineering

Major in Telecommunications

Ing. Oscar Rodríguez Morales

Monterrey, N.L. Dec. 2003

© Oscar Rodríguez Morales, 2003

Instituto Tecnológico y de Estudios Superiores de Monterrey

Campus Monterrey

División de Electrónica, Computación, Información, y Comunicaciones

Programa de Graduados

The members of the thesis committee recommended the acceptance of the thesis of Oscar Rodríguez Morales as a partial fulfillment of the requirements for the degree of Master of Science in:

Electronic Engineering

Major in Telecommunications

THESIS COMMITTEE

David Muñoz Rodríguez, Ph.D.
Advisor

César Vargas Rosales, Ph.D.
Synodal

José Ramón Rodríguez Cruz, Ph.D.
Synodal

Approved

David Garza Salazar, Ph.D.
Director of the Graduate Program

Dec. 2003

This work is dedicated with love
To God for his pardon,
To Erika and Emiliano for giving me the strength to go ahead and mainly for their love,
To my parents and sister for their understanding and support,
To my grandparents for letting me cover in their family,
To the light that makes me back in the right way,
To my uncle Efrén by his unconditional help,
To all the people who did not believe in me.

I am very thankful to Ph.D. David Muñoz Rodríguez, not only for his professional advice, but also for giving me the strength to continue working for my goals and not giving up. Without his support, I would not have made it.

I also want to thank Ph.D. Cesar Vargas Rosales and Ph.D. José Ramón Rodríguez Cruz for their feedback about this work.

Also, thanks to all my friends from the "6th floor," especially Fortino, Enrique, Ulises, Miguel, Servando, Yuri, Cuitlahuac, Fernando, German, and Marco. And also, I want to thank them for their friendship to Gabriela, Aleida, Isabel, Magda, Selene, Evelia, Vladimir, Irasema, Arón, Pepe and his family.

An especial mention to my brothers of "pain", Anabel, Rodolfo and Lluvia, which companion was of great value. At their side I do not feel alone.

And I would like to give special thanks to my mates, Alberto, Roberto, Miguel Angel and Gadiel, for all their support and friendship.

Oscar Rodríguez Morales

Instituto Tecnológico y de Estudios Superiores de Monterrey
Dec. 2003

URN MODELING FOR HEAVY-TAILED PHENOMENA

Oscar Rodríguez Morales, M.Sc.
Instituto Tecnológico y de Estudios Superiores de Monterrey, 2003

During the last decade, it was clear that the use of Poisson processes for modeling network traffic underestimated certain important performance measures such as blocking or queueing delay, among others. Researches around the world agree with the presence of heavy-tail behavior in almost all the data traffic's metrics, such as connection arrivals, file sizes, central processing unit (CPU) time demands of UNIX processes, etc. As a result, during the next few years, heavy-tailed distributions will play a principal role in the modelling and developing of telecommunications systems.

Due to the nature of data traffic, researches demand a discrete heavy-tail distribution, perfectly well described, that enables them to reflect the impact of the two states present in all digital systems (on/off, successful/failed, connect/disconnected, enabled/disabled) in the tail decay. At the present time, there is no distribution with this high degree of flexibility. This thesis completes the description of the discrete heavy-tail distribution introduced in [24] by getting their moments and variance derived from a rigorous generating functions analysis. It validates the model's heavy-tail nature through mean excess functions and some related plots such as the Quantile-Quantile or Probability-Probability plot. Also, the model's stability and their match with the Pareto distribution are investigated. This work concludes with a discussion about the initial conditions influence in the model's tail decay.

URN MODELING FOR HEAVY-TAILED PHENOMENA

Oscar Rodríguez Morales, M.Sc.
Instituto Tecnológico y de Estudios Superiores de Monterrey, 2003

Durante la década pasada quedo de manifiesto que el uso de procesos Poisson para el modelado del tráfico presente en redes de Telecomunicaciones, subestima ciertas métricas de desempeño importantes tales como el bloqueo o el retardo en cola entre otras. Investigadores de todo el mundo coinciden con la presencia de un comportamiento de cola pesada en casi todas las métricas de interes en el tráfico de datos; tales como los arrivos de conexión, el tamaño de archivos, el tiempo de servicio demandado a la unidad de procesamiento central (CPU) por parte de procesos UNIX, etc. Como resultado, durante los próximos años las distribuciones de cola pesada jugarán un papel principal en el modelado y desarrollo de sistemas de telecomunicaciones.

Debido a la naturaleza del tráfico de datos, los investigadores demandan una distribucion discreta de cola pesada perfectamente bien descrita que les permita reflejar el impacto de los dos estados presentes en todos los sistemas digitales (encendido/apagado, exitoso/fallido, conectado/desconectado, habilitado/deshabilitado) en el decaimiento de la cola. A la fecha, no existe una distribución con este alto grado de flexibilidad. Esta tésis completa la descripción de la distribución de cola pesada discreta introducida en [24] mediante la obtención de sus momentos y varianza a partir de un análisis rigoroso de sus funciones generatrices. Valida la naturaleza de cola pesada del modelo a través de funciones de exceso medio y algunos gráficos relacionados, tales como los graficos de Quantile-Quantile y Probabilidad-Probabilidad. También la estabilidad del modelo y su correspondencia con la distribución Pareto es investigada. Este trabajo concluye con una discusión acerca de la influencia de las condiciones iniciales en el decaimiento de la cola del modelo.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

In the telephone era, teletraffic engineering was concerned with the statistical behavior on the call level only; a telephone call required a fixed amount of bandwidth, and consequently, knowledge of the number and duration of calls was sufficient to determine the needed resources. In contrast, multimedia traffic is characterized by a high variability in its bandwidth needs. Data communications between computer terminals usually result in short periods of high activity, followed by long periods of silence. So, in the multimedia era, we are interested not only in the number and duration of calls, but also in the statistical properties of the information flow during the call, in order to make efficient use of the resources while guaranteeing a high quality of service.

A related observation with the traffic behavior in computer networks is that file sizes in some systems have been shown to be well described, using distributions that are *heavy-tailed* (distributions whose tails follow a power law), meaning that file sizes also often span many orders of magnitude [4].

Heavy-tailed distributions behave quite differently from the distributions more commonly used to describe characteristics of computing systems, such as the normal distribution and the exponential distribution, which have tails that decline exponentially (or faster). In contrast, because their tails decline relatively slowly, the probability of very large observations occurring when sampling random variables that follow heavy-tailed distributions is non-negligible. One of the main characteristics of these kinds of distributions is their infinite variance, which reflects the extremely high variability that they capture.

As a result, designers of computing and telecommunications systems are increasingly interested in employing heavy-tailed distributions to generate workloads for use in simulation.

1.1 Objective

The objective of this thesis is to complete the description of the discrete heavy-tail distribution, presented in [24] through an exhaustive generating-function analysis, and to validate their matching with the Pareto distribution through mean-excess functions and certain related plots, such as the Quantile-Quantile or Probability-Probability plots. As a final contribution, some insight about their stability is given in order to provide the basis for their application in the modelling of the aggregate traffic at the input of telecommunications systems when the discrete heavy-tail distribution is used to represent the traffic generated by independent sources.

1.2 Justification

When modeling network traffic, packet and connection arrivals are often assumed to be Poisson processes since researchers expected a similar behavior such as in telephone networks, and mainly because such processes have attractive theoretical properties [12]. On the other hand, a number of exhaustive studies have shown evidence indicating that some aspects of computing and telecommunications systems can show heavy-tailed distributions. Measurements of computer-network traffic have shown that autocorrelations are often related to heavy tails; this is the phenomenon of *self-similarity* [13],[19]. Measurements of the file sizes in the Web [4], and I/O patterns [22], have shown evidence that file sizes can show heavy-tailed distributions. In addition, the CPU time demands of UNIX processes have also been shown to follow heavy-tailed distributions [15], [18].

Due to these observations, heavy-tailed distributions are increasingly used to represent workload characteristics of computing systems, and researchers interested in simulating such systems are beginning to use heavy-tailed inputs to simulations.

Unfortunately there is not a discrete distribution that enclosed the heavy-tail behavior completely, but there have been empirical work loads [8] that fit the measurements obtained from real environments that can be only applied in a specific context. On the other hand, the Zipf distribution (considered as the *discrete Pareto distribution*) has overestimated the Pareto Behavior [24].

Under this perspective, the complete characterization of the discrete heavy-tail distribution presented in [24] is of great importance since the behavior of vital parameters in telecommunications systems could be truly represented and handled in simulations carrying in consequence a best prediction of their impact in the global performance.

1.3 Contributions

In this thesis, I have completed the mathematical description of the discrete heavy-tailed distribution presented in [24] in order to establish the necessary tools for their practical use in analysis and simulations of telecommunications systems or of any discrete heavy-tailed phenomena. Since this distribution is based on the Urn theory, throughout this work we refer to it as the Urn Model distribution.

1.4 Organization

The organization of the present work is as follows. Chapter 2 includes a brief summary about transformation methods for probability distributions as well as the mathematical description of heavy-tail distributions and their importance for telecommunications. Chapter 3 presents a complete mathematical description of the Urn Model distribution. Their heavy-tail nature is validated through excess functions while their moments are derived from generating functions. Later, I investigate the stability of the Urn Model distribution, and towards the end of the chapter an analysis about the behavior of the tail is conducted. In Chapter 4, the parameters' influence on the behavior of the Urn Model distribution is investigated through graphics which are discussed in order to point to a range of interest. Finally, Chapter 5 contains the conclusions of this work, and the opportunities for further research are commented on.

Chapter 2

Heavy-Tailed Distributions and Telecommunications

The purpose of this chapter is to present the fundamental aspects of heavy-tailed distributions and their relationship with telecommunications. For a more detailed description of these topics, see references [23] and [6].

2.1 Pareto Distribution

The simplest heavy-tailed distribution is the *Pareto* distribution, which is power law over its entire range. The classical Pareto distribution, with shape parameter β and location parameter α, has the following cumulative distribution function

$$F(x) = P[X \leq x] = 1 - \left(\frac{\alpha}{x}\right)^{\beta}, \qquad \alpha, \beta \geq 0, \quad x \geq \alpha, \tag{2.1}$$

with the corresponding probability density function:

$$f(x) = \beta \alpha^{\beta} x^{-\beta-1}. \tag{2.2}$$

If $\beta \leq 2$, then the distribution has infinite variance, and if $\beta \leq 1$, then it has infinite mean.

The Pareto distribution (also referred to as the power-law distribution, the double-exponential distribution and the hyperbolic distribution) has been used to model distributions of incomes exceeding a minimum value and sizes of asteroids, islands, cities and extinction events [17], [20]. Leland and Ott also found that a Pareto distribution with 1.05< β <1.25 is a good model for the amount or CPU time consumed by an arbitrary process [18].

2.2 Heavy-Tailed Distributions

In communications, heavy-tailed distributions have been used to model telephone call holding times [9] and frame sizes for variable-bit-rate video [13]. The discrete Pareto (Zipf) distribution [11],

$$P[x = n] = \frac{1}{(n+1)(n+2)}, \quad \text{for } n \geq 0, \tag{2.3}$$

arises in connection with platoon lengths for cars at different speeds traveling on an infinite road with no passing, a model suggestively analogous to computer-network traffic.

We define a distribution as heavy-tailed if

$$P[X \geq x] \sim cx^{-\beta}, \quad \text{as } \ x \to \infty, \quad \beta \geq 0. \tag{2.4}$$

By this, we mean that for some β and some constant c, the ratio $P[X \geq x]/(cx^{-\beta})$ tends to 1 as $x \to \infty$. This definition includes the Pareto and Weibull distributions.

2.2.1 Excess Functions

A more strict definition of *heavy-tailed* defines a distribution as heavy-tailed if the *mean excess function* $e(a)$ of the random variable X is an increasing function of a [23], where

$$e(a) = E[X - a|X > a]. \tag{2.5}$$

Using this second definition of heavy-tailed, consider a random variable X that represents a waiting time. For waiting times with a light-tailed distribution, such as the uniform distribution, the mean excess function is a decreasing function of a. For such a light-tailed distribution, the longer you have waited, the sooner you are likely to be done. For waiting times with a medium-tailed distribution, such as the (memoryless) exponential distribution, the expected future waiting time is independent of the waiting time so far. In contrast, for waiting times with a heavy-tailed distribution, the longer you have waited, the longer your expected future waiting time is. In order to validate the last statement, Equation (2.6) gives the mean excess function for the Pareto distribution with $\beta > 1$ (that is, with finite mean) [23]. As it can be seen, this equation is a linear function of a, since

$$e(a) = \frac{a}{(\beta - 1)}. \tag{2.6}$$

In order to observe the influence of the parameter β in the Pareto's excess function, in Figure 2.1 we can see the plot of Equation (2.6) for certain values of β. From the picture, it is easy to observe that the heavy-tail behavior of the Pareto distribution is reduced as β is growing, due to the fact that the distribution represents a less disperse sample.

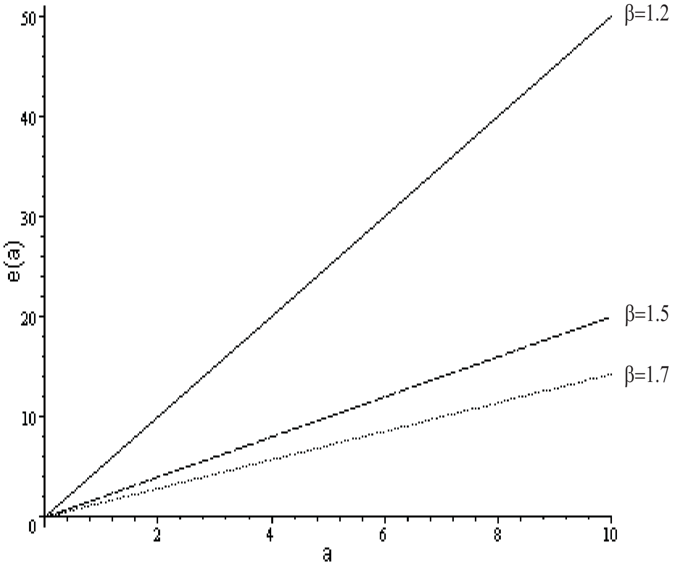

Figure 2.1: Plots of Equation (2.6) for different values of β.

Figure 2.2 shows the shapes of certain mean excess functions. This graphic was taken from [23] and is reproduced as a brief survey.

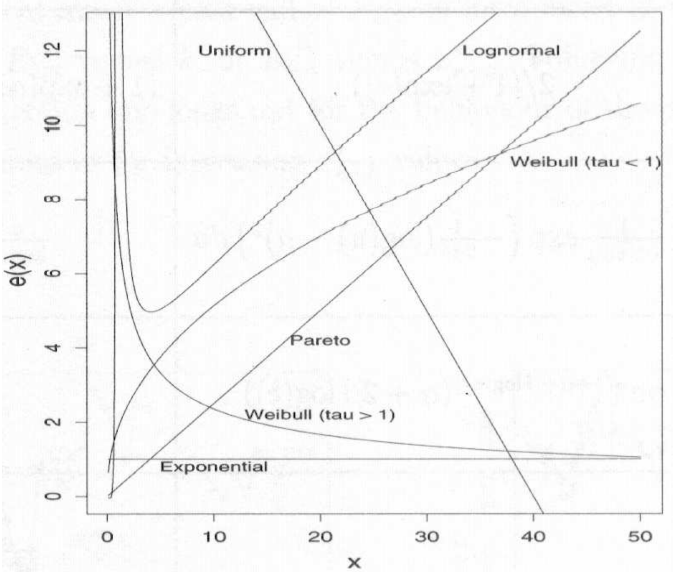

Figure 2.2: Shapes of certain mean-excess functions.

Equation (2.5) defines the mean-excess function in a general way. In order to derive the calculation of $e(a)$ for a specific distribution model $1 - F$ of positive random variables, it can be done using the following formula

$$e(a) = \frac{\int_a^{x+} (1 - F(u))du}{1 - F(a)}. \tag{2.7}$$

It is easy to see that Equation (2.7) can be expressed as

$$e(a) = \sum_{x=a}^{\infty} \frac{P[X > x]}{P[X > a]}, \tag{2.8}$$

which defines the mean excess function for a discrete random variable X. Equation (2.8) will be used to establish the mean excess function of the Urn Model.

Another important parameter in tail estimation when using moment estimators, is the *quadratic mean excess function* $s(a)$, defined as

$$s(a) = E[(X - a)^2 \,|X > a], \tag{2.9}$$

in addition $s(a)$ plays an important role in fixing a premium along a variance or standard-deviation principle.

The quadratic mean excess function $s(a)$ is only defined when X possesses a finite variance, in which case

$$s(a) = 2\frac{\int_a^{x+} (u - a)(1 - F(u))du}{1 - F(a)}. \tag{2.10}$$

In order to establish the quadratic mean excess function for a discrete random variable X, Equation (2.10) can be written as

$$s(a) = 2\frac{\sum_{x=a}^{\infty} (x - a)P[X > x]}{P[X > a]}, \tag{2.11}$$

Equation (2.11) will be used to establish the quadratic mean excess function of the Urn Model.

For the Pareto distribution with $\beta > 2$ (that is, with finite variance), the quadratic mean excess function is defined as

$$s(a) = \frac{2a^2}{(\beta - 2)(\beta - 1)}. \tag{2.12}$$

In order to observe the influence of parameter β in the Pareto's square excess function, in Figure 2.3 we can see the plot of Equation (2.12) for certain values of β. From the figure, we can observe that Pareto's square excess function reduces its slope as β grows as a consequence of a finite variance.

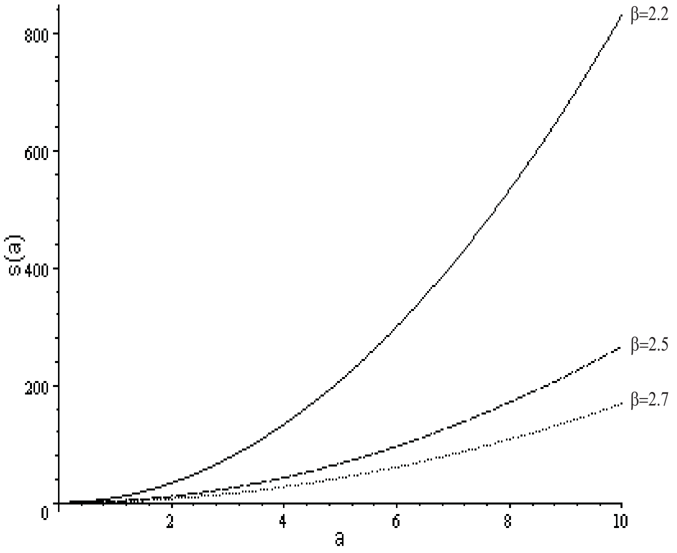

Figure 2.3: Plots of Equation (2.12) for different values of β.

2.2.2 Statistics in Heavy-Tail Distributions

As discussed above, the distinguishing feature of heavy-tailed distributions is the presence of long-ranged, power-law tails, which might lead to the divergence of even the lowest order moments. Moreover, as we know, the most important parameters that summarize the behavior of a random variable are the expected values and the variance. From these facts, it is of vital importance to obtain the Urn Model moments in a suitable way that helps us make the mathematical complexity tractable. In this case, the most powerful mathematical tool is the generating-functions analysis. A brief survey of this topic is presented in the following paragraphs, which are compiled from [10].

Probability-Generating Function

Let us consider the sequence of real numbers $a_0, a_1, a_2, ...$; if

$$A(t) = a_0 + a_1 t + a_2 t^2 + \cdots , \tag{2.13}$$

converges in some interval $-t_0 < t < t_0$, then $A(t)$ is called the generating function of the sequence $\{a_j\}$.

The variable t itself has no significance. If the sequence $\{a_j\}$ is bounded, then a comparison with the geometric series shows that Equation (2.13) converges, at least for $|t| < 1$.

Now, let X be a discrete random variable assuming only the integral values $0, 1, 2, ...$ It will be convenient to have a notation both for the distribution of X and for its tails, and we shall write

$$P\{X = j\} = p_j, \qquad P\{X > j\} = q_j, \tag{2.14}$$

then

$$q_k = p_{k+1} + p_{k+2} + \cdots \qquad k \geq 0. \tag{2.15}$$

The generating functions of the sequences $\{p_j\}$ and $\{q_k\}$ are

$$\Phi_X(t) = p_0 + p_1 t + p_2 t^2 + p_3 t^3 + \cdots \tag{2.16}$$

$$Q_X(t) = q_0 + q_1 t + q_2 t^2 + q_3 t^3 + \cdots, \tag{2.17}$$

as $\Phi_X(1) = 1$, the series for $\Phi_X(t)$ converges absolutely, at least for $-1 \leq t \leq 1$. The coefficients of $Q_X(t)$ are less than unity, and so the series for $Q_X(t)$ converges, at least in the open interval $-1 < t < 1$.

Note that the coefficients of $\Phi_X(t)$ are the values of the probability-density function of the r.v. X, $f_X(x)$; evaluated in $x = 0, 1, 2, ...$ therefore, we can write the probability-generating function $\Phi_X(t)$ for a discrete random variable X [16] as follows

$$\Phi_X(t) = \sum_{x=0}^{\infty} P(X = x)t^x = \sum_{x=0}^{\infty} f_X(x)t^x, \qquad -1 \leq t \leq 1, \tag{2.18}$$

and it is called the probability-generating function due to the fact that values of pdf are obtained from

$$f_X(x) = \frac{1}{k!}\frac{d^x}{dt^x}\Phi_X(t) \mid_{t=0}, \tag{2.19}$$

note that Equation (2.18) can be viewed as the expected value of a function of x, t^x.

On the other hand, since the survival function defines the *"tail"* probabilities, we can also write the generating function for these probabilities based on the previous discussion [1], [10],

$$Q_X(t) = \sum_{x=0}^{\infty} P(X > x)t^x = \sum_{x=0}^{\infty} \overline{F}(x)t^x. \tag{2.20}$$

Note that $Q_X(t)$ is not a probability-generating function in a strict sense. Although the coefficients are probabilities, they do not in general constitute a probability distribution.

A useful result connecting $\Phi_X(t)$ and $Q_X(t)$ is that

$$(1-t)Q_X(t) = 1 - \Phi_X(t). \tag{2.21}$$

Proof. The coefficients of t^x in $(1-t)Q_X(t)$ equals $q_x - q_{x-1} = -p_x$ when $x \geq 1$, and equals $q_0 = p_1 + p_2 + p_3 + \cdots = 1 - p_0$ when $x = 0$. Therefore, $(1-t)Q_X(t) = 1 - \Phi_X(t)$ is asserted. As we will see, important results will be derived from it.

Simple formulas are available giving the mean and variance of the probability distribution $f_X(x)$ in terms of particular values of the generating functions and their derivatives. Thus, the mean is

$$E[X] = \sum_{x=0}^{\infty} xP(X = x) = \Phi'_X(t)|_{t=1}, \tag{2.22}$$

$$E[X] = \sum_{x=0}^{\infty} P(X > x) = Q_X(t)|_{t=1}, \tag{2.23}$$

where the prime in Equation (2.22) indicates differentiation. The validation from the last equations is derived from the relation given in (2.21) because

$$\begin{aligned} \Phi'_X(t)|_{t=1} &= [Q_X(t) - (1-t)Q'_X(t)]\,|_{t=1} \\ &= Q_X(t)|_{t=1}. \end{aligned} \tag{2.24}$$

On the other hand, it can be verified that

$$E[X(X-1)] = \sum_{x=0}^{\infty} x(x-1)P(X = x) = \Phi''_X(t)|_{t=1} = 2Q'_X(t)|_{t=1}; \tag{2.25}$$

hence the variance is

$$var[X] = \Phi''_X(1) + \Phi'_X(1) - \left[\Phi'_X(1)\right]^2, \tag{2.26}$$

$$var[X] = 2Q'_X(1) + Q_X(1) - \left[Q_X(1)\right]^2. \tag{2.27}$$

Similarly, we can obtain the rth factorial moment $\mu'_{[r]}$ about the origin as

$$\begin{aligned} E[X(X-1)\cdots(X-r+1)] &= \sum_{x=0}^{\infty} x(x-1)\cdots(x-r+1)P(X = x) \\ &= \Phi_X^{(r)}(1) \equiv rQ_X^{(r-1)}(1), \end{aligned} \tag{2.28}$$

i.e., by differentiating $\Phi_X(t)$ r times and putting $t = 1$.

In order to validate the results presented in this section for the given distribution under study and, mainly, to give some insight into its higher central moments, we need to investigate its moment-generating function. A review of this topic is presented in the following section.

Moment-Generating Function

The moment-generating function $M_X(t)$ for a discrete random variable X is defined by [16]

$$M_X(t) = \sum_{x=0}^{\infty} e^{xt} P(X = x). \tag{2.29}$$

In the last section, we defined the probability-generating function for such random variables as

$$\Phi_X(t) = \sum_{x=0}^{\infty} P(X = x) t^x, \qquad -1 \leq t \leq 1. \tag{2.30}$$

From Equations (2.29) and (2.30), it is clear that

$$M_X(t) = \Phi_X(e^t), \tag{2.31}$$

and Equation (2.31) allows us to determine the moment-generating function directly from the probability-generating function. On the other hand, it is a well know result that

$$EX^n = \frac{d^n}{dt^n} M_X(t) \mid_{t=0} . \tag{2.32}$$

With the previous background over generating functions, we can obtain the statistics of any heavy-tail distribution.

2.2.3 Their Importance for Telecommunications

Heavy-tailed distributions are an important probabilistic tool used to model the behavior of vital parameters in communications systems, such as the size of files in a web server and the transmission times and number of files being transmitted through a packet network. All of these elements, as well as some others have a deep impact in network performance, and the theory of heavy-tailed distributions plays a major role in the design and fine adjustment of telecommunications systems.

But how important are heavy-tailed distributions in telecommunications systems?. As we will see, this fact was a natural consequence of data traffic. This was pointed out by the scientific community during the last decade.

At the beginning of the packet-network era, network arrivals were modelled for analytic simplicity, as Poisson processes, mainly because there was not any previous experience on this kind of traffic. Subsequent traffic reports and studies showed that packet inter-arrivals were not exponentially distributed [8],[21]. What researches found was that user-initiated transmission control protocol (TCP) session arrivals, such as remote login and file transfer, were well modelled as Poisson processes with fixed hourly rates, but that other connection arrivals deviated considerably from Poisson. First efforts on this topic focused on the use of empirical workloads, such as the Teplib, to simulate packet inter-arrivals obtaining traffic where the inter-arrivals preserve burstinnes over many time scales; results agree with real environments far away from the expected results since of the point of view of exponential arrivals. Another related observation was that file transfer protocol (FTP) data-connection arrivals within FTP sessions came bunched into a “connection burst,” the largest of which were so large that they dominated the FTP data traffic.

Certainly, the main contribution of these first works was the observation that the arrival pattern of user-generated TELNET packets had an invariant distribution, independent of network details. Therefore, the natural question concerned the kind of distribution they had in from. Some insight came from of same work realized in [21]. Paxson and collaborators found that the distribution of the number of bytes in each burst had a very heavy upper tail and that a small fraction of the largest burst carried almost all of the FTP data-connection bytes. This implied that faithful modeling of FTP traffic should concentrate heavily on the characteristics of the largest burst. The last statement point towards a heavy-tail behavior of the network traffic, but the question in the air was left without an answer, although Paxson’s team finishes its work with a discussion of how its burstinnes results mesh with self similar models of network traffic.

At about the same time, the statistical analysis of Ethernet traffic measures collected at Bellcore between 1989 and 1992 is reported in [19]. From their study, they arrived at the conclusion that the Ethernet traffic is statistically self-similar. Moreover, they proposed a stochastic model for this self-similar behavior by means of a renewal-reward process through the aggregation of a sequence of independent, identically distributed (iid) random variables (r.v.), whose distinctive characteristic is their heavy-tail nature.

From the previous discussions, we can conclude that the fact that network traffic shows self-similarity means that it shows a noticeable burst at a wide range of time scales. A related observation is that the number of bytes in each burst could be described by using distributions that are heavy-tailed (distributions whose tails follow a power law) meaning that the number of bytes in each burst often span many orders of magnitude, also.

Information about other phenomena observing heavy-tailed behavior can be found in [7], [5] and [2]. These works show that the distribution of transmission times, the size of the files available on web servers, the number of files transmitted through the network, the average number of request Vs file size, the relative popularity of web pages, and certain others aspects of the WWW are heavy tailed distributed.

Of special interest is the fact that the results shown in [5] agree with the results shown in [19] concerning the "heaviness" of the tail presented in the distributions of WWW traffic. This is the main cause of the presence of long-range dependence (i.e., self-similar) observed in the WWW traffic. They showed that the tail of the distribution of the ON times (transmission times) is heavier than the distribution of the OFF times (silent times), meaning that the self-similarity of the WWW traffic will be governed by parameter α of the transmission times.

From previous discussion, it is clear, the importance in the modeling of Heavy Tailed distributions. Of special interest, is the availability of a discrete heavy tail distribution that enable to researches match the tail decay found in the characterization of actual telecommunications networks with a mathematical model. From this necessity and due to the mathematical implications in the study of heavy tailed distributions, the math background at the beginning of this section was presented.

Chapter 3

Model Description

This chapter introduces the discrete heavy-tail distribution proposed in [24]. The notation is defined, and the mathematical description of the model is completed. At the end of the chapter, a closed expression for the nth *factorial moment* is presented, and the stability of the urn model distribution is investigated.

3.1 Urn Model for Heavy-Tailed Phenomena

In this section, the Urn Model's functioning is explained, the main assumptions are discussed, and a first approximation to its heavy-tailed nature is given.

3.1.1 Probability-Mass Function

As established in chapter two, the holding times of actual network traffic are heavy-tailed distributed. This implies that the longer that a user has been connected, the longer the expected future connection time is. In [24] was proposed a discrete heavy-tail distribution that resembles this behavior for the holding times which can be see as an addictive-connection process.

From this point of view, the addictive-connection process is represented by a discrete-time urn process with "connect" and "disconnect" balls operating in such a form that at any observation time the connection state will be determined by the random selection of a ball. Any time a ball is selected, $m+1$ balls of the same class are replaced into the urn (m is known as the Polya constant).

In order to show that adequate selection of initial conditions allows for model on-off periods with the desired heavy-tailed characteristics, we consider a subscriber that, at a time t_0, is in on-state, and that at a time t_1 he will make the decision of quitting the connection with the following probability:

$$P(1) = P(X = 1) = \frac{d}{c + d}, \tag{3.1}$$

where c stands for the initial number of "connect" balls and d denotes the number of "disconnect" balls in the urn.

However, if the decision at time t_1 was to remain connected, at time t_2 of the second observation the probability of ending the connection will be $\frac{d}{c+d+m}$. Thus the probability of a connection ending at the second observation will be

$$p(2) = P(X = 2) = \left(\frac{c}{c + d}\right)\left(\frac{d}{c + d + m}\right). \tag{3.2}$$

This concept can readily be extended so that probability of ending a connection at the $x - th$ observation window becomes

$$P(X = x) = \left(\frac{c}{c + d}\right)\left(\frac{c + m}{c + d + m}\right)\left(\frac{c + 2m}{c + d + 2m}\right)...\left(\frac{c + (x - 2)m}{c + d + (x - 2)m}\right)\left(\frac{d}{c + d + (x - 1)m}\right). \tag{3.3}$$

After some algebraic reordering, Equation (3.3) can be expressed as

$$P(X = x) = \left(\frac{d}{m}\right)\frac{\left(x - 2 + \frac{c}{m}\right)\left(x - 3 + \frac{c}{m}\right)...\left(\frac{c}{m} + 1\right)\left(\frac{c}{m}\right)\Gamma\left(\frac{c}{m}\right)}{\left(x - 1 + \frac{c+d}{m}\right)\left(x - 2 + \frac{c+d}{m}\right)...\left(\frac{c+d}{m} + 1\right)\left(\frac{c+d}{m}\right)\Gamma\left(\frac{c}{m}\right)}. \tag{3.4}$$

It can be noted that for $m = 0$, Equation (3.3) is reduced to the geometric model. On the other hand, using the well known property of the gamma function -$\Gamma(n + c) = (n - 1 + c)\Gamma(n - 1 + c) = (n - 1 + c)(n - 2 + c)...(c)\Gamma(c)$- and after regrouping, we can write (3.4) as

$$P(X = x) = \left(\frac{d}{m}\right)\frac{\Gamma\left(x - 1 + \frac{c}{m}\right)}{\left(x - 1 + \frac{c+d}{m}\right)\left(x - 2 + \frac{c+d}{m}\right)...\left(\frac{c+d}{m} + 1\right)\left(\frac{c+d}{m}\right)\Gamma\left(\frac{c}{m}\right)}. \tag{3.5}$$

The last expression can be formulated as

$$P(X = x) = \frac{\Gamma\left(\frac{c+d}{m}\right)}{\Gamma\left(\frac{c}{m}\right)\Gamma\left(\frac{d}{m}\right)}\frac{\Gamma\left(x - 1 + \frac{c}{m}\right)\Gamma\left(\frac{d}{m} + 1\right)}{\Gamma\left(x + \frac{c+d}{m}\right)}. \tag{3.6}$$

It can be verified that for $x = 1$, expression (3.4) is reduced to (3.1), thus the support of (3.4) is $x = 1, 2, 3, 4....$

Recalling that $B(m, n) = \frac{\Gamma(m)\Gamma(n)}{\Gamma(m+n)}$ where $B(x, y) = \int_0^1 t^{x-1}(1-t)^{y-1} = \int_0^\infty \frac{t^{x-1}}{(1+t)^{x+y}} dt$ is the beta function which is known to have the following properties, [14]

$$B(x, y) = B(y, x), \tag{3.7}$$

$$\sum_{k=0}^{\infty} B(x + k, y) = B(x, y - 1). \tag{3.8}$$

We can rewrite the probability mass function (pmf) of the Urn Model as

$$p(x) = P(X = x) = \frac{1}{B\left(\frac{c}{m}, \frac{d}{m}\right)} B\left(x - 1 + \frac{c}{m}, \frac{d}{m} + 1\right), \qquad x = 1, 2, 3, ... \tag{3.9}$$

In order to verify that (3.9) is a valid *pmf*, we must to verify that $\sum_{x=1}^{\infty} p(x) = 1$. For such objective, if we define $j = x - 1$, which implies that $x = j + 1$, and if we use the relation given in (3.8) we can establish that

$$\begin{aligned}\sum_{x=1}^{\infty} \frac{B\left(x - 1 + \frac{c}{m}, \frac{d}{m} + 1\right)}{B\left(\frac{c}{m}, \frac{d}{m}\right)} &= \sum_{j=0}^{\infty} \frac{B\left(\frac{c}{m} + j, \frac{d}{m} + 1\right)}{B\left(\frac{c}{m}, \frac{d}{m}\right)}, \\ &= \frac{B\left(\frac{c}{m}, \frac{d}{m}\right)}{B\left(\frac{c}{m}, \frac{d}{m}\right)}, \\ &= 1, \end{aligned} \tag{3.10}$$

setting aside any doubt about the validity of (3.9). As a first snapshot, Figure 3.1 shows the plot of Equation (3.9) for two different values of the quotient $\frac{d}{m}$ denoted by α. As it can be seen, for the small value of α, the heavy-tail behavior is more stressed.

3.1.2 Survival Function

Since our purpose is to study the Urn Model's tail behavior, we are mostly interested in the possibility of obtaining a subscriber connected still overshooting a given time observation x. This is, we are interested in the event

$$1 - F(x) = P[X > x] = \sum_{k=x+1}^{\infty} p(k), \tag{3.11}$$

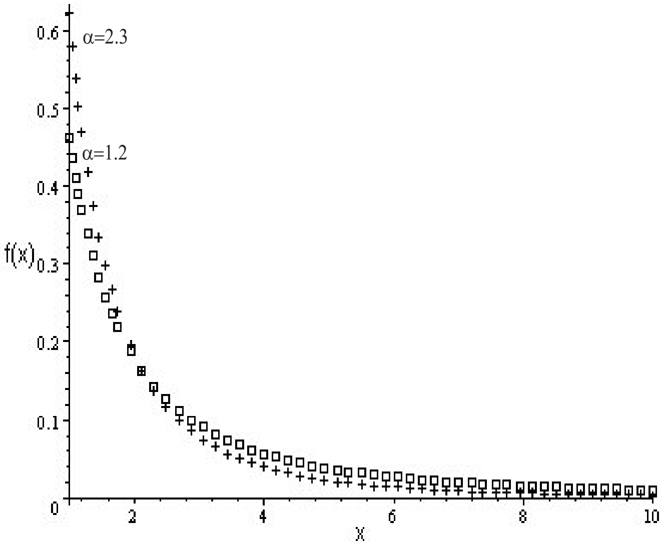

Figure 3.1: Urn Model probability-density function.

where $p(k)$ is calculated according to (3.9).

In biostatistical applications $1 - F(x)$, denoted by $\overline{F}$, is called the *survival function*. This is the notation that we will use throughout this work. Replacing Equation (3.9) in (3.11) we have

$$\overline{F}(x) = P[X > x] = \sum_{k=x+1}^{\infty} \frac{1}{B(\delta, \alpha)} B(k - 1 + \delta, \alpha + 1), \qquad x = 1, 2, 3, ... \tag{3.12}$$

where $\delta = \frac{c}{m} > 0$ and $\alpha = \frac{d}{m} > 0$ defining $i = k - x - 1$, we get

$$\overline{F}(x) = \frac{1}{B(\delta, \alpha)} \sum_{i=0}^{\infty} B(i + x + \delta, \alpha + 1), \qquad x = 1, 2, 3, ..., \tag{3.13}$$

and using (3.8) we obtain

$$\overline{F}(x) = \frac{1}{B(\delta, \alpha)} B(x + \delta, \alpha). \tag{3.14}$$

In Figure 3.2, we can see the plot of Equation (3.14). As we observe, the heavy-tail behavior is reduced for the larger value of α.

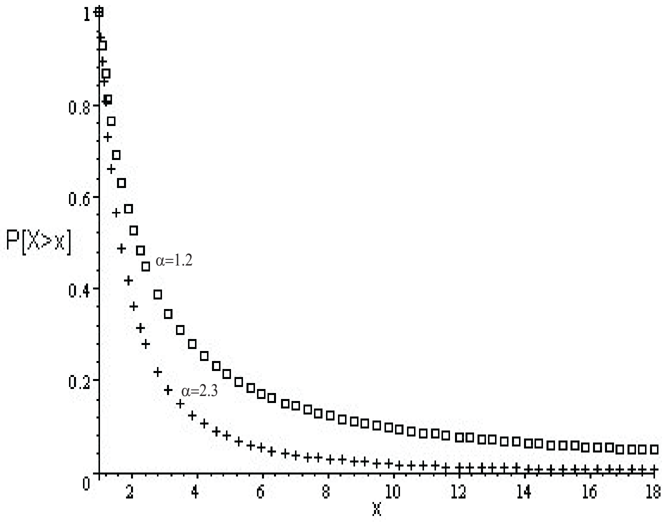

Figure 3.2: Urn Model survival function.

3.1.3 Distribution Function

From the definition $\overline{F}(x)$, the cumulative-distribution function of the Urn Model is given by

$$F(x) = 1 - \frac{B(x+\delta,\alpha)}{B(\delta,\alpha)}, \qquad x = 1, 2, 3, ... \tag{3.15}$$

In Figure 3.3, we can see the plot of Equation (3.15). As we observe, for the larger value of α, the *cdf* grows faster than for the smaller value.

3.1.4 Heavy-Tail Behavior of the Urn Model

In order to show that $p(x)$ exhibits a heavy-tailed behavior, we consider tail decay of the survivability function (3.14), which can be formulated as

$$\overline{F}(x) = \frac{1}{B(\delta,\alpha)} \frac{\Gamma(x+\delta)\Gamma(\alpha)}{\Gamma(x+\delta+\alpha)}. \tag{3.16}$$

The heavy-tail behavior of a random variable is characterized by the slow decay of the survival function for large values. That is, a random is heavy tailed if $P[x > \xi] \sim \xi^{-\alpha}$ for large values of ξ. Using property $\frac{\Gamma(i)}{\Gamma(i+k)} \approx i^{-k}$ [14], as $k \to \infty$ we have

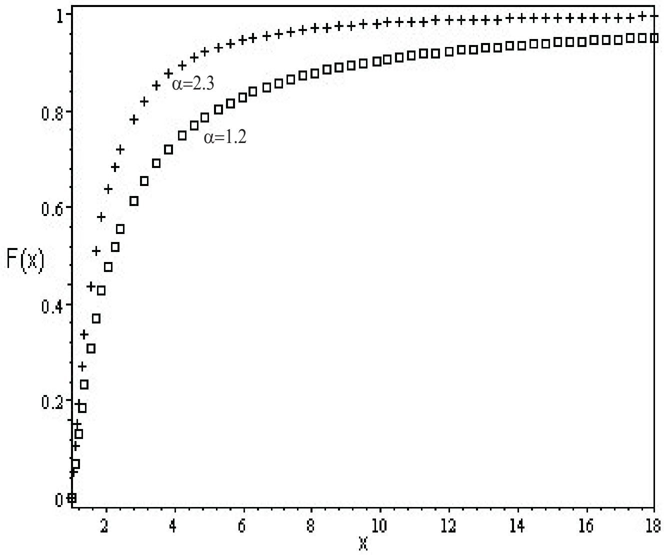

Figure 3.3: Urn Model cumulative-distribution function.

$$\overline{F}(x) = \frac{1}{B(\delta,\alpha)}\frac{\Gamma(x+\delta)\Gamma(\alpha)}{\Gamma(x+\delta+\alpha)} \approx \frac{\Gamma(\alpha)}{B(\delta,\alpha)}(x+\delta)^{-\alpha}, \tag{3.17}$$

which shows the heavy-tailed behavior of the distribution. Since δ and α were defined as $\delta = \frac{c}{m}$ and $\alpha = \frac{d}{m}$, it can be seen that the tail decay depends on the initial conditions of the urn experiment. In order to have an idea about the last approximation, in Figure 3.5, we can observe the plot of both equations. As we can see for large values of x, the tails show an identical behavior, confirming what was established: that the Urn Model presents heavy-tail behavior. In Chapter 4, Q-Q plots show close behavior between a Paretian distribution and results obtained by the proposed model.

3.2 Excess Functions

In this section, the Urn Model's excess functions are investigated and their similarity with the Pareto's excess functions is commented.

3.2.1 Mean-Excess Function

The excess function of a r.v. X has been defined as $e(\xi) = E(X - \xi | X > \xi)$ which for discrete r.v. can be written as $e(\xi) = \sum_{x=\xi}^{\infty} \frac{P[X>x]}{P[X>\xi]}$. Replacing Equation (3.14) into the last expression, we have

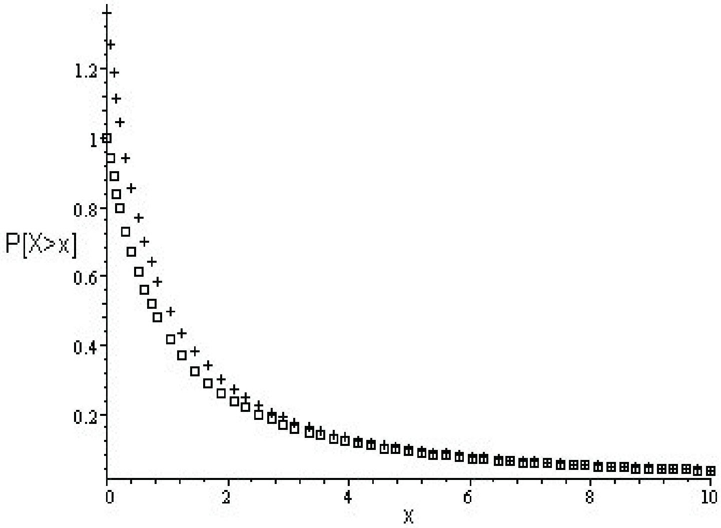

Figure 3.4: Comparison between $\overline{F}(x)$ (boxes) and their approximation (crosses).

$$e(\xi) = \sum_{x=\xi}^{\infty} \frac{B(\delta,\alpha)}{B(\delta+\xi,\alpha)} \frac{B(x+\delta,\alpha)}{B(\delta,\alpha)}. \tag{3.18}$$

Now take the dummy variable "y" defined as $y = x - \xi \Rightarrow x = y + \xi$. Using again the well-know reduction formula given in (3.8), the last equation takes the form

$$e(\xi) = \frac{1}{B(\delta+\xi,\alpha)} \sum_{y=0}^{\infty} B(y+\xi+\delta,\alpha) = \frac{B(\xi+\delta,\alpha-1)}{B(\xi+\delta,\alpha)}. \tag{3.19}$$

The previous expression can be reduced by using the relation between the Beta and Gamma functions, as shown bellow

$$e(\xi) = \frac{\Gamma(\delta+\xi+\alpha)}{\Gamma(\delta+\xi)\Gamma(\alpha)} \frac{\Gamma(\delta+\xi)\Gamma(\alpha-1)}{\Gamma(\delta+\xi+\alpha-1)} = \frac{\delta+\xi+\alpha-1}{\alpha-1}. \tag{3.20}$$

Finally, the mean-excess function of the Urn Model can be written as

$$e(\xi) = \frac{\delta+\xi}{\alpha-1} + 1. \tag{3.21}$$

In terms of the model's parameters, we have the following

$$e(\xi) = \frac{c + d + m(\xi - 1)}{d - m}. \tag{3.22}$$

In Chapter 2, it was said that the heavy-tail behavior of a r.v. X was demonstrated by the increasing behavior of its excess function. As a second validation of the heavy-tail nature of the Urn Model, in Figure 3.5 we can see the plot of Equation (3.21). As can be seen, the Urn Model really behaves as a heavy-tail distribution, and even more so if we compare the next figure with Figure 2.1. We can observe a high similarity between both Pareto and Urn Model mean-excess functions.

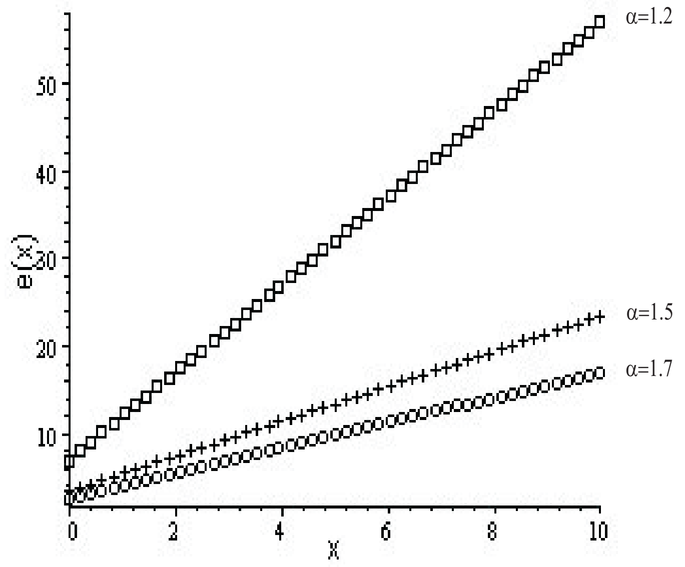

Figure 3.5: Urn Model mean-excess function.

3.2.2 Quadratic Mean-Excess Function

The Quadratic Mean-Excess function for a discrete r.v. X is defined as $s(\xi) = 2\sum_{x=\xi}^{\infty} \frac{(x-\xi)P[X>x]}{P[X>\xi]}$. Using (3.14), the Quadratic Mean-Excess function for the Urn Model is

$$s(\xi) = 2\sum_{x=\xi}^{\infty}(x-\xi)\frac{B(\delta,\alpha)}{B(\delta+\xi,\alpha)}\frac{B(x+\delta,\alpha)}{B(\delta,\alpha)}. \tag{3.23}$$

Setting aside the constant terms, we have

$$s(\xi) = \frac{2}{B(\delta+\xi,\alpha)}\sum_{x=\xi}^{\infty}(x-\xi)B(x+\delta,\alpha). \tag{3.24}$$

Now, making use of the dummy variable "y", defined as: $y = x - \xi \Rightarrow x = y + \xi$, last equation adopts the following form

$$s(\xi) = \frac{2}{B(\delta+\xi,\alpha)} \sum_{y=0}^{\infty} yB(y+\xi+\delta,\alpha), \tag{3.25}$$

in terms of Gamma functions

$$s(\xi) = 2\sum_{y=0}^{\infty} y \frac{\Gamma(\delta+\xi+\alpha)}{\Gamma(\delta+\xi)\Gamma(\alpha)} \frac{\Gamma(\alpha)\Gamma(y+\delta+\xi)}{\Gamma(y+\delta+\xi+\alpha)}, \tag{3.26}$$

simplifying and solving

$$s(\xi) = \frac{2}{(\alpha-1)(\alpha-2)} \frac{\Gamma(\delta+\xi+1)}{\Gamma(\delta+\xi)} \frac{\Gamma(\delta+\xi+\alpha)}{\Gamma(\delta+\xi+\alpha-1)}. \tag{3.27}$$

Finally the Urn Model quadratic mean-excess function is given by

$$s(\xi) = \frac{2(\delta+\xi)(\delta+\xi+\alpha-1)}{(\alpha-1)(\alpha-2)}. \tag{3.28}$$

In Figure 3.6, we can observe plots of (3.28) in order to have an idea about its behavior to different values of α. As it can be seen, the Urn Model quadratic mean-excess function behaves similarly to the Pareto quadratic mean-excess function.

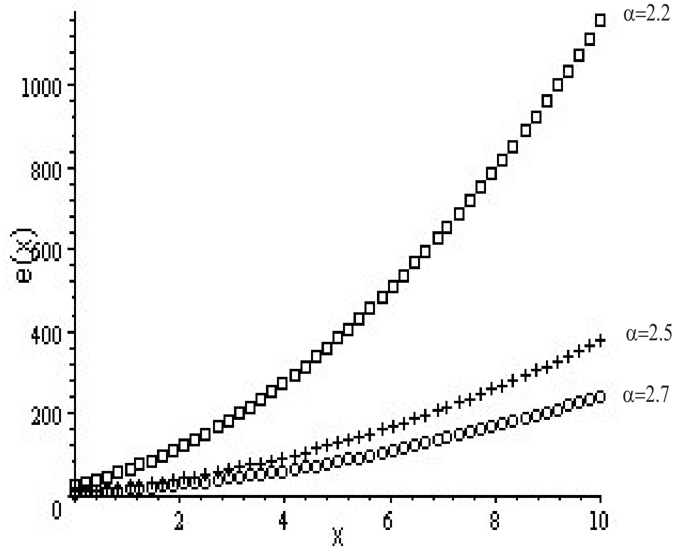

Figure 3.6: Urn Model quadratic mean-excess function.

3.3 Generating Functions

In this section, the Urn Model's generating functions are derived in order to investigate the moments and the variance of the model. The moments' existence and their influence in the sample space that the Urn Model could be represent are commented.

3.3.1 Probability-Generating Function

As was established in Chapter 2, the probability-generating function $\Phi_X(t)$ for a discrete random variable X is defined as

$$\Phi_X(t) = \sum_{x=0}^{\infty} P(X = x)t^x, \qquad -1 \leq t \leq 1, \tag{3.29}$$

so, in order to get the probability-generating function for the Urn Model, we need to substitute Equation (3.9) into (3.29), taking into consideration that the support of $p(x)$ given in (3.9) for the urn model is $x = 1, 2, 3, ..$; in such way we have

$$\Phi_X(t) = \sum_{x=1}^{\infty} \frac{1}{B\left(\delta, \alpha\right)} B\left(x - 1 + \delta, \alpha + 1\right) t^x. \tag{3.30}$$

The previous expression makes their manipulation difficult. Replacing beta functions with their equivalent in terms of gamma functions, and by moving the constant terms away, we have

$$\Phi_X(t) = \frac{\Gamma(\delta + \alpha)\Gamma(\alpha + 1)}{\Gamma(\delta)\Gamma(\alpha)} \sum_{x=1}^{\infty} \frac{\Gamma(x + \delta - 1)}{\Gamma(x + \delta + \alpha)} \cdot t^x. \tag{3.31}$$

By focusing ours efforts on the sum and by expanding certain terms, we have

$$\sum_{x=1}^{\infty} \frac{\Gamma(x + \delta - 1)}{\Gamma(x + \delta + \alpha)} \cdot t^x = \frac{\Gamma(\delta)t}{\Gamma(\delta + \alpha + 1)} + \frac{\Gamma(\delta + 1)t^2}{\Gamma(\delta + \alpha + 2)} + \frac{\Gamma(\delta + 2)t^3}{\Gamma(\delta + \alpha + 3)} + ... \tag{3.32}$$

It is easy to see that the previous expression can be written as

$$\sum_{x=1}^{\infty} \frac{\Gamma(x + \delta - 1)}{\Gamma(x + \delta + \alpha)} \cdot t^x = \left[\frac{\Gamma(\delta)t}{\Gamma(\delta + \alpha + 1)}\right] \left[1 + \frac{\delta t}{(\delta + \alpha + 1)} + \frac{\delta(\delta + 1)t^2}{(\delta + \alpha + 1)(\delta + \alpha + 2)} + ...\right]. \tag{3.33}$$

Based on the definition of the hypergeometric function [14], which is

$$hypergeom([n_1, n_2, ..., n_p], [d_1, d_2, ..., d_q], z) = \sum_{k=0}^{\infty} \frac{\prod_{i=1}^{p} \frac{\Gamma(n_i + k)}{\Gamma(n_i)} z^k}{\prod_{i=1}^{q} \frac{\Gamma(d_i + k)}{\Gamma(d_i)} k!}, \tag{3.34}$$

we can observe that in Equation (3.33) the right term is the extended form of the hypergeometric function of parameters $([1,\delta],[\delta+\alpha+1],t)$. Therefore, Equation (3.31) can be expressed as

$$\Phi_X(t) = \left[\frac{\Gamma(\delta+\alpha)\Gamma(\alpha+1)}{\Gamma(\delta)\Gamma(\alpha)}\right]\left[\frac{\Gamma(\delta)t}{\Gamma(\delta+\alpha+1)}\right] hypergeom\left([1,\delta],[\delta+\alpha+1],t\right). \tag{3.35}$$

For the sake of simplicity, Equation (3.35) can be reformulated as

$$\Phi_X(t) = \frac{\alpha\cdot t}{\delta+\alpha} hypergeom\left([1,\delta],[\delta+\alpha+1],t\right). \tag{3.36}$$

$Q_X(t)$ is another useful generating function that allows us to validate the results derived from $\Phi_X(t)$. On the other hand, as we will see, with $Q_X(t)$ we can establish another way to obtain the *tail* probabilities. From the explanation in chapter two, $Q_X(t)$ for the Urn Model comes from to replace Equation (3.14) into (2.20). That is,

$$Q_X(t) = \sum_{x=0}^{\infty}\frac{1}{B(\delta,\alpha)}B(x+\delta,\alpha)t^x. \tag{3.37}$$

Again, in order to facilitate their manipulation, we replace beta functions with their equivalent in terms of gamma functions, and by moving the constant terms away, we have

$$Q_X(t) = \frac{\Gamma(\delta+\alpha)}{\Gamma(\delta)}\sum_{x=0}^{\infty}\frac{\Gamma(x+\delta)}{\Gamma(x+\delta+\alpha)}\cdot t^x, \tag{3.38}$$

expanding certain terms and after regrouping, we have

$$Q_X(t) = \left[1+\frac{\delta\cdot t}{\delta+\alpha}+\frac{\delta(\delta+1)\cdot t^2}{(\delta+\alpha)(\delta+\alpha+1)}+\cdots+\right]. \tag{3.39}$$

Observe that Equation (3.39) is the extended form of the hypergeometric function of parameters $([1,\delta],[\delta+\alpha],t)$. Therefore, the Urn Model's tail probability-generating function is given by

$$Q_X(t) = hypergeom\left([1,\delta],[\delta+\alpha],t\right). \tag{3.40}$$

In Chapter 2, $E[X] = \Phi'_X(1)$ was established. Therefore, we need to find the derivative of Equation (3.36). To realize such an operation, we can use the chain rule. By the use of this mathematical tool, we can check that

$$\Phi'_X(t) = \frac{\alpha hypergeom([1,\delta],[\delta+\alpha+1],t)}{\delta+\alpha}+\frac{\alpha\delta t hypergeom([2,\delta+1],[\delta+\alpha+2],t)}{(\delta+\alpha)\,(\delta+\alpha+1)}. \tag{3.41}$$

In order to evaluate the previous equation in $t = 1$, we can make use of the following relation [14]

$$hypergeom([\tau,\beta],[\gamma],1) = \frac{\Gamma(\gamma)\Gamma(\gamma-\tau-\beta)}{\Gamma(\gamma-\tau)\Gamma(\gamma-\beta)}. \tag{3.42}$$

Thanks to reduction formula (3.42) we can write (3.41) as follows

$$\Phi'_X(1) = \left(\frac{\alpha}{\delta+\alpha}\right)\left(\frac{\Gamma(\delta+\alpha+1)\Gamma(\alpha)}{\Gamma(\delta+\alpha)\Gamma(\alpha+1)}\right) + \left(\frac{\alpha\delta}{(\delta+\alpha)(\delta+\alpha+1)}\right)\left(\frac{\Gamma(\delta+\alpha+2)\Gamma(\alpha-1)}{\Gamma(\delta+\alpha)\Gamma(\alpha+1)}\right), \tag{3.43}$$

and in consequence, after simplifying we have

$$\Phi'_X(1) = \left(\frac{\alpha}{\delta+\alpha}\right)\left(\frac{\delta+\alpha}{\alpha}\right) + \left(\frac{\alpha\delta}{(\delta+\alpha)(\delta+\alpha+1)}\right)\left(\frac{(\delta+\alpha)(\delta+\alpha+1)}{\alpha(\alpha-1)}\right). \tag{3.44}$$

A final simplification can be obtained from the previous formula, and in the end we get

$$E[X] = \Phi'_X(1) = \frac{\delta+\alpha-1}{\alpha-1}. \tag{3.45}$$

One way to validate the previous result is through $E[X] = Q_X(1)$. Therefore,

$$\begin{aligned} E[X] &= Q_X(1), \\ &= hypergeom\left([1,\delta],[\delta+\alpha],1\right), \\ &= \frac{\Gamma(\delta+\alpha)\Gamma(\alpha-1)}{\Gamma(\delta+\alpha-1)\Gamma(\alpha)}, \\ &= \frac{\delta+\alpha-1}{\alpha-1}. \end{aligned} \tag{3.46}$$

As the expected value of the Urn Model has been determined and validated, the next parameter of importance in the description of the behavior of the Urn Model is the second moment. The importance of $E[X^2]$ resides in the fact that it determines the variance and therefore gives some insight about the tail decay.

From Equation (2.25), we can establish that

$$\begin{aligned} E[X^2] &= \Phi''_X(1) + \Phi'_X(1), \\ &= 2Q'_X(1) + Q_X(1). \end{aligned} \tag{3.47}$$

In order to follow the same order in the presentation, the first expression in (3.47) will be determined and then will be validated through the tail probability-generating function. Since the derivations of the derivatives and simplifications could make it hard to follow the presentation, only final results will be presented. We have that the second derivative of $\Phi_X(t)$ given by

$$\begin{aligned}\Phi''_X(t) &= 2\frac{\alpha\delta hypergeom([2,\delta+1],[\delta+\alpha+2],t)}{(\delta+\alpha)(\delta+\alpha+1)}+\\ &\quad 2\frac{\alpha\delta(\delta+1)(t)hypergeom([3,\delta+2],[\delta+\alpha+3],t)}{(\delta+\alpha)\left(\delta+\alpha+1\right)\left(\delta+\alpha+2\right)}. \end{aligned}\tag{3.48}$$

Evaluating the last expression in $t=1$, we have

$$\begin{aligned}\Phi''_X(1) &= \frac{2\alpha\delta\Gamma(\delta+\alpha+2)\Gamma(\alpha-1)}{(\delta+\alpha)(\delta+\alpha+1)\Gamma(\delta+\alpha)\Gamma(\alpha+1)}+\\ &\quad \frac{2\alpha\delta(\delta+1)\Gamma(\delta+\alpha+3)\Gamma(\alpha-2)}{(\delta+\alpha)\left(\delta+\alpha+1\right)\left(\delta+\alpha+2\right)\Gamma(\delta+\alpha)\Gamma(\alpha+1)}.\\ &= \frac{2\delta}{\alpha-1}+\frac{2\delta(\delta+1)}{(\alpha-1)(\alpha-2)}, \end{aligned}\tag{3.49}$$

replacing the last result in (3.47), we find that $E[X^2]$ takes the following form

$$\begin{aligned}E[X^2] &= \frac{2\delta}{\alpha-1}+\frac{2\delta(\delta+1)}{(\alpha-1)(\alpha-2)}+\frac{\delta+\alpha-1}{\alpha-1},\\ &= \frac{2\delta(\delta+\alpha-1)}{(\alpha-1)(\alpha-2)}+\frac{\delta+\alpha-1}{\alpha-1}. \end{aligned}\tag{3.50}$$

As it can be seen $E[X^2]$ is only defined for $\alpha>2$.
On the other hand, the $Q'_X(t)$ is given by

$$Q'_X(t)=\frac{\delta}{\delta+\alpha}hypergeom\left([2,\delta+1],[\delta+\alpha+1],t\right). \tag{3.51}$$

Evaluating the previous expression in $t=1$ and replacing in the second expression of (3.47), we have

$$\begin{aligned}E[X^2] &= 2\left(\frac{\delta}{\delta+\alpha}\right)\left(\frac{\Gamma(\delta+\alpha+1)\Gamma(\alpha-2)}{\Gamma(\delta+\alpha-1)\Gamma(\alpha)}\right)+\frac{\delta+\alpha-1}{\alpha-1},\\ &= 2\frac{\delta(\delta+\alpha-1)}{(\alpha-1)(\alpha-2)}+\frac{\delta+\alpha-1}{\alpha-1}, \end{aligned}\tag{3.52}$$

from Equation (3.50) and (3.52), we can obtain $E[X^2]$ and validate Equation (3.47).

As we have seen, we get the same results from $\Phi_X(t)$ and $Q_X(t)$. Therefore, we will find the variance of X by replacing the appropriate results in Equation (2.26), so we have

$$Var[X] = \frac{2\delta(\delta+\alpha-1)}{(\alpha-1)(\alpha-2)} + \frac{\delta+\alpha-1}{\alpha-1} - \left[\frac{\delta+\alpha-1}{\alpha-1}\right]^2 . \tag{3.53}$$

In chapter two was also established a methodology to get the rth factorial moment μ'_r about the origin. It consists in finding the rth derivative of $\Phi_X(t)$ and evaluating it with $t = 1$. The importance of these factorial moments resides in the fact that the rth **central** moment can be derived from some combination of them reducing a lot of work, as will be shown when the moment-generating analysis is exposed. In order to facilitate the exposition, only final results will be shown.

Note that $\Phi'_X(1)$ and $\Phi''_X(1)$ already have been derived (Equations (3.45) and (3.49)), so let us obtain additional derivatives in order to find a relation for the rth factorial moment. That is,

$$\Phi^3_X(1) = 6\frac{\delta(\delta+1)(\delta+\alpha-1)}{(\alpha-1)(\alpha-1)(\alpha-3)}, \tag{3.54}$$

$$\Phi^4_X(1) = 24\frac{\delta(\delta+1)(\delta+2)(\delta+\alpha-1)}{(\alpha-1)(\alpha-1)(\alpha-3)(\alpha-4)}. \tag{3.55}$$

From the above results, it is easy to see

$$\mu'_r = \Phi^r_X(1) = r!\frac{(\delta+\alpha-1)\prod_{i=0}^{r-2}(\delta+i)}{\prod_{j=1}^{r}(\alpha-j)}. \tag{3.56}$$

From Equation (3.56), we can conclude that the Urn Model's rth factorial moment will exist only if $\alpha > r$.

Before continuing the next section, we have to validate what was established in Chapter 2 about Equations (3.36) and (3.40). Why were they called probability-generating functions?. As a reminder, this was due to the fact that $f_X(x) = \frac{1}{k!}\frac{d^x}{dt^x}\Phi_X(t) \mid_{t=0}$ and $\overline{F}(x) = \frac{1}{k!}\frac{d^x}{dt^x}Q_X(t) \mid_{t=0}$, respectively. That is, with these equations we are able to obtain the values of the Urn Model's pdf and survival function for the x th value, just by taking the xth derivative and realizing the operations indicated. In order to validate this fact, in Tables 3.1 and 3.2 we can observe the values obtained through both methods for each case. As it can be see, the results agree with theory.

Table 3.1: Numerical comparison for values of $f_X(x)$.

x	$\frac{1}{k!}\Phi_X^x(0)$	$f_X(x)$
1	.592592593	.592592592
2	.176176176	.176176176
3	.078717015	.078717015
4	.042811008	.042811008
5	.026197781	.026197781
6	.017351777	.017351777
7	.012166189	.012166189
8	.008905148	.008905148
9	.00674128	.00674128
10	.005243218	.005243218

Table 3.2: Numerical comparison for values of $\overline{F}(x)$.

x	$\frac{1}{k!}Q_X^x(0)$	$\overline{F}(x)$
1	.4074074074	.4074074074
2	.2312312312	.2312312312
3	.1525142163	.1525142163
4	.1097032082	.1097032082
5	.08350542718	.08350542718
6	.06615365009	.06615365009
7	.05398746155	.05398746155
8	.04508231326	.04508231326
9	.03834103278	.03834103278
10	.03309781463	.03309781463

3.3.2 Moment-Generating Function

In Chapter 2, the moment-generating function $M_X(t)$ for a discrete random-variable X was defined as

$$M_X(t) = \sum_{x=0}^{\infty} e^{xt} P(X = x). \tag{3.57}$$

On the other hand, it is a well know result that

$$EX^n = \frac{d^n}{dt^n} M_X(t) \mid_{t=0} . \tag{3.58}$$

Therefore, in order to get the Urn Model´s central moments, we need first to find its moment-generating function. Substituting Equation (3.9) into (3.57), we have

$$M_X(t) = \sum_{x=1}^{\infty} e^{xt} \frac{1}{B(\delta, \alpha)} B(x - 1 + \delta, \alpha + 1), \tag{3.59}$$

the previous expression makes their manipulation difficult. Replacing beta functions with their equivalent in terms of gamma functions, and moving the constant terms away, we have

$$M_X(t) = \frac{\Gamma(\delta + \alpha)\Gamma(\alpha + 1)}{\Gamma(\delta)\Gamma(\alpha)} \sum_{x=1}^{\infty} \frac{\Gamma(x + \delta - 1)}{\Gamma(x + \delta + \alpha)} \cdot e^{xt}. \tag{3.60}$$

By focusing ours efforts on the sum, and by expanding certain terms, we have

$$\sum_{x=1}^{\infty} \frac{\Gamma(x + \delta - 1)}{\Gamma(x + \delta + \alpha)} \cdot e^{xt} = \frac{\Gamma(\delta) e^t}{\Gamma(\delta + \alpha + 1)} + \frac{\Gamma(\delta + 1) e^{2t}}{\Gamma(\delta + \alpha + 2)} + \frac{\Gamma(\delta + 2) e^{3t}}{\Gamma(\delta + \alpha + 3)} + ... \tag{3.61}$$

It is easy to see that the previous expression can be written as

$$\sum_{x=1}^{\infty} \frac{\Gamma(x + \delta - 1)}{\Gamma(x + \delta + \alpha)} \cdot e^{xt} = \left[\frac{\Gamma(\delta) e^t}{\Gamma(\delta + \alpha + 1)} \right] \left[1 + \frac{\delta e^t}{(\delta + \alpha + 1)} + \frac{\delta(\delta + 1) e^{2t}}{(\delta + \alpha + 1)(\delta + \alpha + 2)} + ... \right]. \tag{3.62}$$

We can observe that in Equation (3.62) the term of the right, is the extended form of the hypergeometric function of parameters $([1, \delta], [\delta + \alpha + 1], e^t)$. Therefore, Equation (3.60) can be expressed as

$$M_X(t) = \left[\frac{\Gamma(\delta + \alpha)\Gamma(\alpha + 1)}{\Gamma(\delta)\Gamma(\alpha)} \right] \left[\frac{\Gamma(\delta) e^t}{\Gamma(\delta + \alpha + 1)} \right] hypergeom \left([1, \delta], [\delta + \alpha + 1], e^t \right), \tag{3.63}$$

for the sake of simplicity, Equation (3.63) can be reformulated as

$$M_X(t) = \frac{\alpha}{\delta + \alpha} e^t hypergeom \left([1, \delta], [\delta + \alpha + 1], e^t \right). \tag{3.64}$$

As a validation of Equation (3.64), can be verified that $M_X(0) = 1$ because

$$\begin{aligned} M_X(0) &= \left(\frac{\alpha}{\delta+\alpha}\right) hypergeom\left([1,\delta],[\delta+\alpha+1],1\right), \\ &= \left(\frac{\alpha}{\delta+\alpha}\right)\left(\frac{\Gamma(\delta+\alpha+1)\Gamma(\alpha)}{\Gamma(\delta+\alpha)\Gamma(\alpha+1)}\right), \\ &= 1. \end{aligned} \tag{3.65}$$

With the Urn Model moment-generating function, we are able to research its central moments. For this purpose, let us obtain the first derivative of Equation (3.64). That is

$$\begin{aligned} M'_X(t) &= \frac{\alpha e^t}{\delta+\alpha}\left[\frac{\delta e^t}{\delta+\alpha+1}+\frac{2\delta(\delta+1)e^{2t}}{(\delta+\alpha+1)(\delta+\alpha+2)}+...\right]+ \\ & \quad hypergeom\left([1,\delta],[\delta+\alpha+1],e^t\right)\frac{\alpha e^t}{\delta+\alpha}, \end{aligned} \tag{3.66}$$

from the definition of the hypergeometric function and after certain algebraic manipulations, it can be verified that the term between brackets can be reduced to $\frac{\delta e^t hypergeom([2,\delta+1],[\delta+\alpha+2],e^t)}{\delta+\alpha+1}$. Therefore, the first derivative of Equation (3.64) can be written as

$$\begin{aligned} M'_X(t) &= \frac{\alpha\delta e^{2t} hypergeom\left([2,\delta+1],[\delta+\alpha+2],e^t\right)}{(\delta+\alpha)(\delta+\alpha+1)}+ \\ & \quad \frac{\alpha e^t hypergeom\left([1,\delta],[\delta+\alpha+1],e^t\right)}{\delta+\alpha}, \end{aligned} \tag{3.67}$$

according to (3.58), we have $E[X]=M'_X(0)$. For this reason, when evaluating Equation (3.67) for $t=0$ and then simplifying, we find that the first moment of the Urn Model is given by

$$E[X]=\frac{\delta+\alpha-1}{\alpha-1}. \tag{3.68}$$

We note that Equation (3.68) agrees with (3.45). Moreover, it is equivalent to $e(0)$. Therefore, as a result, we can establish

$$E[X]=e(0)=\frac{\delta}{\alpha-1}+1. \tag{3.69}$$

In order to obtain an expression for the Urn Model's second moment, we need to find the derivative of Equation (3.67). Again, we will obtain the derivatives of the series' terms and, after regrouping, we will try to find a closed expression in terms of hypergeometric functions. Note that the second term in (3.67) is already derivative because it is the generating-moment function, while for the first term we have

$$\begin{aligned} M''_X(t) &= \frac{\alpha\delta e^{2t}}{(\delta+\alpha)(\delta+\alpha+1)}\left[\frac{2(\delta+1)e^t}{(\delta+\alpha+2)} + \frac{6(\delta+1)(\delta+2)e^{2t}}{(\delta+\alpha+2)(\delta+\alpha+3)} + ...\right] + \\ & \quad hypergeom([2,\delta+1],[\delta+\alpha+2],e^t)\frac{2\alpha\delta e^{2t}}{(\delta+\alpha)(\delta+\alpha+1)} + \\ & \quad M'_X(t), \end{aligned} \tag{3.70}$$

when replacing $M'_X(t)$ by its extended form (3.67) and regrouping, we have

$$\begin{aligned} M''_X(t) &= \frac{2\alpha\delta(\delta+1)e^{3t}}{(\delta+\alpha)(\delta+\alpha+1)(\delta+\alpha+2)}\left[1 + \frac{3(\delta+2)e^t}{(\delta+\alpha+3)} + ...\right] + \\ & \quad \frac{3\alpha\delta e^{2t}}{(\delta+\alpha)(\delta+\alpha+1)} hypergeom([2,\delta+1],[\delta+\alpha+2],e^t) + \\ & \quad \frac{\alpha e^t}{\delta+\alpha} hypergeom\left([1,\delta],[\delta+\alpha+1],e^t\right). \end{aligned} \tag{3.71}$$

From the definition of the hypergeometric function and after certain algebraic manipulations, it can be verified that the term between brackets is the extended form of $hypergeom([3,\delta+2],[\delta+\alpha+3],e^t)$. Therefore, the second derivative of Equation (3.64) can be written as

$$\begin{aligned} M''_X(t) &= \frac{2\alpha\delta(\delta+1)e^{3t}}{(\delta+\alpha)(\delta+\alpha+1)(\delta+\alpha+2)} hypergeom\left([3,\delta+2],[\delta+\alpha+3],e^t\right) + \\ & \quad \frac{3\alpha\delta e^{2t}}{(\delta+\alpha)(\delta+\alpha+1)} hypergeom\left([2,\delta+1],[\delta+\alpha+2],e^t\right) + \\ & \quad \frac{\alpha e^t}{\delta+\alpha} hypergeom\left([1,\delta],[\delta+\alpha+1],e^t\right), \end{aligned} \tag{3.72}$$

according to (3.58) we have $E[X^2] = M''_X(0)$. For this reason, when evaluating Equation (3.72) for $t=0$ and then simplifying, we find that the second moment of the Urn Model is given by

$$E[X^2] = \frac{\alpha^2 - 3\alpha + 2 + 3\delta\alpha - 4\delta + 2\delta^2}{(\alpha-1)(\alpha-2)}. \tag{3.73}$$

After certain algebraic work, it can be verified that Equation (3.73) can be written as

$$E[X^2] = \frac{2\delta(\delta+\alpha-1)}{(\alpha-1)(\alpha-2)} + \frac{\delta+\alpha-1}{(\alpha-1)}. \tag{3.74}$$

Before proceeding to find the third moment of the Urn Model, we note that Equation (3.74) agrees with (3.50), denoting an absolute convergence in the analysis of the transformation methods to the Urn Model.

Continuing with our analysis, and in order to find the third moment of the Urn Model, we need to derive Equation (3.72). Since the methodology has been already shown and since the intermediate steps are irrelevant, we will show only the final expression for the third derivative of the moment-generating function, which has the form

$$\begin{aligned} M_X'''(t) &= \frac{6\alpha\delta\,(\delta+1)\,(\delta+2)\,e^{4\,t}}{(\delta+\alpha)\,(\delta+\alpha+1)\,(\delta+\alpha+2)\,(\delta+\alpha+3)} hypergeom([4,\delta+3],[\delta+\alpha+4],e^t)+ \\ &\quad \frac{12\alpha\delta\,(\delta+1)\,e^{3\,t}}{(\delta+\alpha)\,(\delta+\alpha+1)\,(\delta+\alpha+2)} hypergeom([3,\delta+2],[\delta+\alpha+3],e^t)+ \\ &\quad \frac{7\alpha\delta e^{2\,t}}{(\delta+\alpha)\,(\delta+\alpha+1)} hypergeom([2,\delta+1],[\delta+\alpha+2],e^t)+ \\ &\quad \frac{\alpha e^t}{\delta+\alpha} hypergeom([1,\delta],[\delta+\alpha+1],e^t). \end{aligned} \tag{3.75}$$

Because we know that $E[X^3] = M_X'''(0)$, when evaluating Equation (3.75) for $t = 0$ and then simplifying, we find that the third moment of the Urn Model is given by

$$E[X^3] = \frac{\alpha^3 + 7\alpha^2\delta - 6\alpha^2 + 11\alpha - 23\delta\alpha + 12\delta^2\alpha - 18\delta^2 - 6 + 6\delta^3 + 18\delta}{(\alpha-1)(\alpha-2)(\alpha-3)}. \tag{3.76}$$

After certain algebraic work, it can be verified that Equation (3.76) can be written as

$$E[X^3] = \frac{6\delta(\delta+1)(\delta+\alpha-1)}{(\alpha-1)(\alpha-2)(\alpha-3)} + 3\left[\frac{2\delta(\delta+\alpha-1)}{(\alpha-1)(\alpha-2)} + \frac{(\delta+\alpha-1)}{(\alpha-1)}\right] - \frac{2(\delta+\alpha-1)}{(\alpha-1)}. \tag{3.77}$$

As can be seen, a lot of effort needs to be made in finding the central moments from the moment-generating function $M_X(t)$. We can get at least for the first three central moments, a significant reduction of work, if we follow the next procedure. Based on the definition of $\Phi_X(t)$, we have

$$\begin{aligned} \frac{d}{dt}\Phi_X(t)|_{t=1} &= \sum_{x=0}^{\infty} f_X(x) x t^{x-1}|_{t=1}, \\ &= \sum_{x=0}^{\infty} x f_X(x), \\ &= E[X], \\ &= \mu_1', \end{aligned} \tag{3.78}$$

where μ_k' is the k th factorial moment. In the same way,

$$\begin{aligned}\frac{d^2}{dt^2}\Phi_X(t)|_{t=1} &= \sum_{x=0}^{\infty} f_X(x)x(x-1)t^{x-2}|_{t=1},\\ &= \sum_{x=0}^{\infty} x^2 f_X(x) - \sum_{x=0}^{\infty} x f_X(x),\\ &= E[X^2]-E[X]. \qquad (3.79)\end{aligned}$$

From the previous expression, we can establish

$$E[X^2]=\mu_2'+\mu_1', \qquad (3.80)$$

while the third moment is found as

$$\begin{aligned}\frac{d^3}{dt^3}\Phi_X(t)|_{t=1} &= \sum_{x=0}^{\infty} f_X(x)x(x-1)(x-2)t^{x-3}|_{t=1},\\ &= \sum_{x=0}^{\infty} x^3 f_X(x) - 3\sum_{x=0}^{\infty} x^2 f_X(x) + 2\sum_{x=0}^{\infty} x f_X(x),\\ &= E[X^3]-3E[X^2]+2E[X]. \qquad (3.81)\end{aligned}$$

Setting $E[X^2]$ and $E[X]$ in term of factorial moments and after sorting out, we can verify that

$$E[X^3]=\mu_3'+3\mu_2'+\mu_1'. \qquad (3.82)$$

Obtain larger central moments represents the same difficulty following any methodology. So, the election of which methodology to follow is letting to the reader. On the other hand, as in any study, the moments of interest are the first three. The exposition is focused on these. As a validation of the given relations, note that Equations (3.68), (3.74) and (3.77) agrees with (3.78), (3.80) and (3.82), respectively.

3.3.3 Stability of the Urn Model

A probability density is called *stable* if it is invariant under convolution [25]; i.e., if there are constants $a>0$ and b, such that

$$\begin{aligned}p(u) &= f(x)*g(x),\\ &= p(a_1x+b_1)*p(a_2x+b_2),\\ &= \sum_{-\infty}^{\infty} p(a_1(u-x)+b_1)p(a_2x+b_2),\\ &= p(au+b), \qquad (3.83)\end{aligned}$$

for all real constants $a_1 > 0$, b_1, $a_2 > 0$, b_2. In (3.83) u represents the sum of two independent and identically distributed (iid) r.v. and $*$ denotes the convolution operation.

For the sake of simplicity, in our case, let us choose $a_1 = a_2 = 1$, $b_1 = b_2 = 0$. Moreover, since the Urn Model's *pmf* is only supported for $x = 1, 2, 3, 4...$ and the convolution's arguments can not be negative, we have that the convolution of two random variables distributed according the Urn Model is obtained from

$$f_U(u) = \sum_{x=1}^{U} \frac{B\left(U - x + \delta - 1, \alpha + 1\right) B\left(x + \delta - 1, \alpha + 1\right)}{B\left(\delta, \alpha\right)^2}, \tag{3.84}$$

When replacing the Beta functions with their equivalent in terms of gamma functions, and after making certain algebraic simplifications, Equation (3.84) can be written as

$$f_U(u) = \left[\frac{\alpha\Gamma(\delta + \alpha)}{\Gamma(\delta)}\right]^2 \sum_{x=1}^{U} \frac{\Gamma\left(U - x + \delta - 1\right)\Gamma\left(x + \delta - 1\right)}{\Gamma\left(U - x + \delta + \alpha\right)\Gamma\left(x + \delta + \alpha\right)}. \tag{3.85}$$

As a first step to find a closed expression for the last sum, let us develop certain terms in order to clarify the term's sequence, which is given below:

$$\begin{aligned} f_U(u) \quad &= \quad \left[\frac{\alpha\Gamma(\delta + \alpha)}{\Gamma(\delta)}\right]^2 \frac{\Gamma(U + \delta - 2)\Gamma(\delta)}{\Gamma(U + \delta + \alpha - 1)\Gamma(\delta + \alpha + 1)} + \frac{\Gamma(U + \delta - 3)\Gamma(\delta + 1)}{\Gamma(U + \delta + \alpha - 2)\Gamma(\delta + \alpha + 2)} + \cdots \\ &\quad + \frac{\Gamma(\delta)\Gamma(\delta + U - 2)}{\Gamma(\delta + \alpha + 1)\Gamma(\delta + \alpha + U - 1)} + \frac{\Gamma(\delta - 1)\Gamma(\delta + U - 1)}{\Gamma(\delta + \alpha)\Gamma(\delta + \alpha + U)}, \end{aligned} \tag{3.86}$$

From our experience handling hypergeometric functions, it is easy to observe that the sum sequence can be expressed as

$$\frac{\Gamma(U + \delta - 2)\Gamma(\delta) hypergeom([1, 2 - U - \delta - \alpha, \delta], [\delta + \alpha + 1, 3 - U - \delta], 1)}{\Gamma(U + \delta + \alpha - 1)\Gamma(\delta + \alpha + 1)}, \tag{3.87}$$

the inconvenient resides in the fact that the hypergeometric function is an infinite series, so it is necessary to eliminate the terms beyond $v = U + 1$, as can be proof, these remanent terms can be enclosure by

$$\frac{\Gamma(\delta - 2)\Gamma(\delta + U) hypergeom([1, 2 - \delta - \alpha, U + \delta], [U + 1 + \delta + \alpha, 3 - \delta], 1)}{\Gamma(\delta + \alpha - 1)\Gamma(\delta + \alpha + U + 1)}. \tag{3.88}$$

From the previous discussion, we can establish that the convolution of two random variables distributed according the Urn Model is given by

$$\begin{aligned} f_U(u) \;=\; & \frac{K_0\Gamma(U+\delta-2)\Gamma(\delta) hypergeom([1,2-U-\delta-\alpha,\delta],[\delta+\alpha+1,3-U-\delta],1)}{\Gamma(U+\delta+\alpha-1)\Gamma(\delta+\alpha+1)} - \\ & \frac{K_0\Gamma(\delta-2)\Gamma(\delta+U) hypergeom([1,2-\delta-\alpha,U+\delta],[U+1+\delta+\alpha,3-\delta],1)}{\Gamma(\delta+\alpha-1)\Gamma(\delta+\alpha+U+1)}, \end{aligned} \tag{3.89}$$

where $K_0 = \left[\frac{\alpha\Gamma(\delta+\alpha)}{\Gamma(\delta)}\right]^2$, and $u = 2, 3, 4, 5, ...$

There is not simplification for Equation (3.89) in order that it takes the form of a ratio of Beta functions and the Urn Model stability can not be determined. On the order hand, Equation (3.83) becomes particularly simple in Fourier space, where the convolution $p(u) = f(x) * g(x)$ reduces to a product of the Fourier transforms. Following this analysis it is clear that $p(u)$ can be obtained taking the inverse transform of Fourier of this product. This is

$$f_U(u) = F^{-1}\left[\left(\Phi_X(e^{jw})\right)^2\right], \tag{3.90}$$

under the assumption that both r.v. are *iid* according to the Urn Model. The operations indicated in (3.90) take the form

$$f_U(u) = \frac{1}{2\pi}\int_0^{2\pi}\left[\frac{\alpha e^{j\omega}}{\delta+\alpha}\right]^2 \left[hypergeom\left([1,\delta],[\delta+\alpha+1],e^{j\omega}\right)\right]^2 e^{-j\omega x}d\omega. \tag{3.91}$$

From Fourier theory, $f_U(u)$ exist only if the argument of Equation (3.91) is infinitively summable, this is if $[hypergeom\,([1,\delta],[\delta+\alpha+1],e^{j\omega})]^2$ converges. Simplification formulas are not available and we conclude that the stability of the Urn Model can not be determined by traditional methods. Fortunately, there are other ways to investigate the Urn Model domain of attraction problem as we will in the next section.

3.3.4 Domain of attraction for extremes

In practical engineering work, the order statistic $T = x_{(i)}$ is one of the simplest and most useful, because it allows the decision maker to focus on a specific region of the distribution. In particular, the extreme $x_{(1)}$ or $x_{(n)}$ is important because it is often a required design input.

The *exact* sampling pdf of the i*th* order statistic is known [3]

$$f_i(x;\theta,n) = \frac{n!}{(i-1)!(n-i)!}[F(x;\theta)]^{i-1}[1-F(x;\theta)]^{n-i}f(x;\theta), \tag{3.92}$$

where f and F are the pdf and cdf of the measurement variable X, respectively. Moreover, $f(x;\theta)$ simply represent the probability model of the *rv* X but indexed by a parameter θ, which often is a vector of two or more parameters.
For the special case of $i = n$, the pdf of the largest observation $x_{(n)}$ in a sample of size n is

$$f_n(x;\theta,n) = n[F(x;\theta)]^{n-1} f(x;\theta), \tag{3.93}$$

with cdf

$$F_n(x;\theta,n) = [F(x;\theta)]^n, \tag{3.94}$$

Of special interest for us, is the fact, that if the initial distribution $f(x;\theta)$ has an unbounded upper tail, but not all of its moments are finite, then the Frechet distribution arises as the limiting form of the distribution given in (3.93). In this context the Frechet distribution is termed a *type II extreme value* distribution of maxima [3].
A continuous random variable X has a Frechet distribution if its pdf has the form

$$f(x;\sigma,\lambda) = \frac{\lambda}{\sigma}\left(\frac{\sigma}{x}\right)^{\lambda+1} e^{-\left(\frac{\sigma}{x}\right)^{\lambda}}; \qquad x \geq 0; \quad \sigma,\lambda>0. \tag{3.95}$$

A Frechet variable X, as defined by (3.95), has the cdf

$$F(x;\sigma,\lambda) = e^{-\left(\frac{\sigma}{x}\right)^{\lambda}}. \tag{3.96}$$

This model has scale structure, with σ a scale parameter and λ a shape parameter. The expected value of a *rv* X distributed according to the Frechet model is defined as

$$E[X] = \sigma\Gamma\left(1 - \frac{1}{\lambda}\right). \tag{3.97}$$

The Frechet distribution features a reproductive property for its maximum extreme. That is, the distribution of $X_{(n)}$ is again Frechet, with the same shape parameter but with the scale parameter increased to $\sigma n^{1/\lambda}$ [3]. Thus, the pdf of $X_{(n)}$ has the same shape as that of X but rescaled as given above.

Back to the Urn Model, we known that it has an unbounded upper tail and that only the moments of order $n < \alpha$ exits. On the other hand, the Urn Model has not a well defined scale structure but along this work we have seen that its α parameter determine the tail's shape while that its δ parameter influence the tail's size; this is its scale. For this reason we can expect that for a given Urn Model distribution its Frechet representation can be obtained just exchanging parameters of shape and scale. This is

$$\frac{B(x+\delta-1,\alpha+1)}{B(\delta,\alpha)} \approx \frac{\alpha}{\delta}\left(\frac{\delta}{x}\right)^{\alpha+1} e^{-\left(\frac{\delta}{x}\right)^{\alpha}}; \tag{3.98}$$

In order to observe how last approximation works, in Figure 3.7 we can see the plot of a Urn Model distribution with $\alpha = 1.2$, $\delta = 1.2$ and its Frechet representation. Despite, the Frechet representation of a Urn model distribution is defined for $x \geq 0$ its tail decay is very close to the Urn Model's tail decay that the engender.

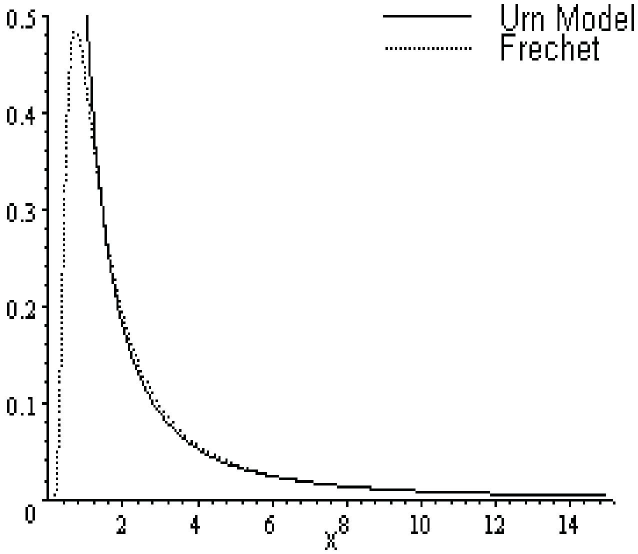

Figure 3.7: Urn Model distribution and its Frechet representation.

Figure 3.7 gives a good idea about the approximation between an Urn Model and its Frechet equivalent representation but is desired that this approximation could be quantified. One way to do it, is quantifying the difference between the expected value of both distributions. This can be done through the next error function

$$error = \sqrt{\left(\frac{E[X_{UM}] - E[X_F]}{E[X_F]}\right)^2} 100\%, \tag{3.99}$$

where $E[X_{UM}]$ represents the Mean Value of a *rv* X distributed according the Urn Model and $E[X_F]$ represents the Mean Value of a *rv* X distributed according the Frechet distribution. After replacing (3.45) and (3.97) into (3.99) and simplifying, last expression take the form

$$error = \sqrt{\left(\frac{\delta + \alpha - 1}{(\alpha - 1)(\delta)\Gamma(1 - \frac{1}{\alpha})} - 1\right)^2} 100\%, \tag{3.100}$$

Fixing the shape parameter α and varying the scale parameter δ along the range of interest $1 < \delta < 2$ we can obtain a plot about the error incurred when we represent an Urn Model through the Frechet model by exchanging their parameters of shape and scale. As a survey, a family of error plots is show in Figure 3.8 for different values of α. As we can see, when α grows, the error becomes bigger. Another relative observation, is that for the example given in Figure 3.7 ($\delta = 1.2$), the error between mean values is about 4.8%.

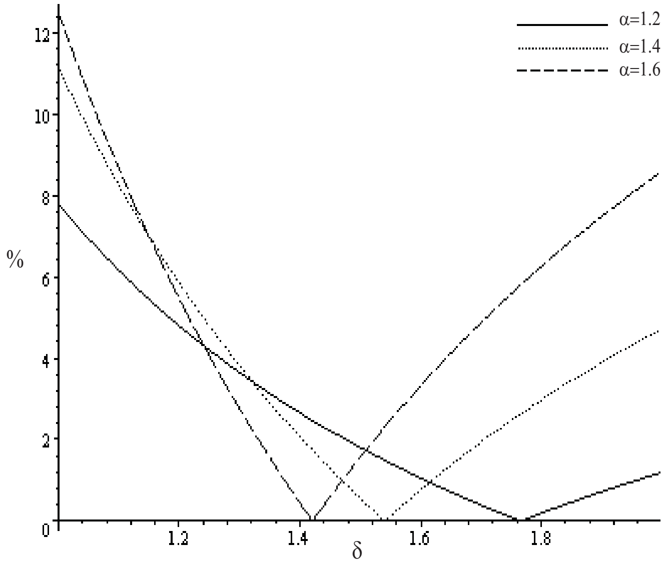

Figure 3.8: Family of error Plots as defined by (3.100) for different values of α.

Back to the problem of extremes, from exposed at the beginning of section 3.3.4, we have that the pdf of the largest observation $x_{(n)}$ in a sample of *rv* X distributed according the Urn Model of size n is given by

$$f_n(x;\delta,\alpha,n) = n\left[1 - \frac{B(x+\delta,\alpha)}{B(\delta,\alpha)}\right]^{n-1}\frac{B(x+\delta-1,\alpha+1)}{B(\delta,\alpha)}, \tag{3.101}$$

From the Frechet's reproductive property for its maximum extreme, we know that the pdf of the largest observation $x_{(n)}$ in a sample of *rv* X distributed according the Frechet model of size n is given by

$$f_n(x;\sigma,\lambda,n) = \frac{\lambda}{\sigma n^{1/\lambda}}\left(\frac{\sigma n^{1/\lambda}}{x}\right)^{\lambda+1} e^{-\left(\frac{\sigma n^{1/\lambda}}{x}\right)^{\lambda}}; \qquad x \geq 0; \quad \sigma,\lambda{>}0. \tag{3.102}$$

Now, since the Urn Model has an unbounded upper tail and not all its moments are finite, we can expect that equation (3.101) converges to the Frechet probability model. The question that arise, is to which specific member of the Frechet family (3.101) tends. Since we already have seen, that from a given Urn Model we can obtain its equivalent Frechet representation with the same shape and scale parameters; we can establish the next limiting density for equation (3.101) as n grows

$$f_n(x;\delta,\alpha,n) = \frac{\alpha}{\delta n^{1/\alpha}}\left(\frac{\delta n^{1/\alpha}}{x}\right)^{\alpha+1} e^{-\left(\frac{\delta n^{1/\alpha}}{x}\right)^{\alpha}}; \qquad x \geq 0; \quad \delta,\alpha>0. \tag{3.103}$$

In Figure 3.9 we can see the plots of (3.101) and (3.103) for $\alpha = 1.2$, $\delta = 1.2$ and $n = 20$. Despite n is small, the approximation is very acceptable.

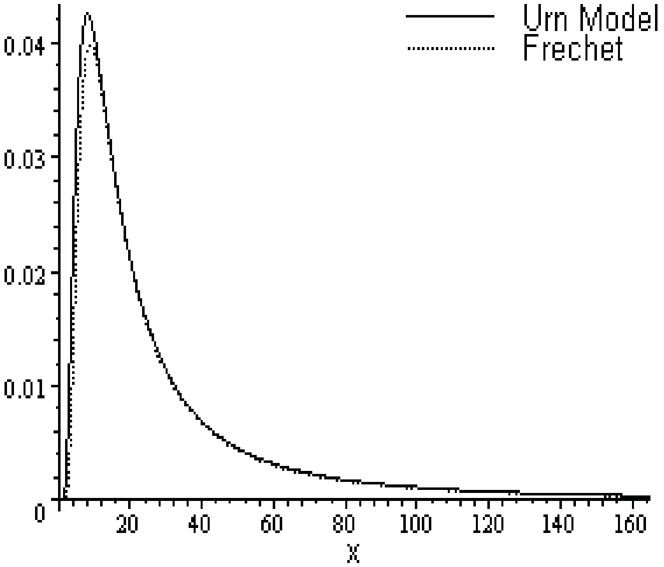

Figure 3.9: Plot of (3.101) and its approximation (3.103) for $n = 20$.

As a second observation, in Figure 3.10 we can see the plots of (3.101) and (3.103) for $\alpha = 1.2$, $\delta = 1.2$ but now with $n = 40$. As we can see the densities are almost identical even when n is not really big.

3.3.5 Urn Model - Pareto Match

In order to facilitate the use of the Urn Model in the modeling of heavy-tailed phenomena, we deduce a methodology to obtain the Urn Model's parameters from a given Pareto distribution. That is, if we already have at hand a well defined Pareto distribution, and if it is in our interest to dispose of an Urn Model equivalent representation, we can proceed in the following way.

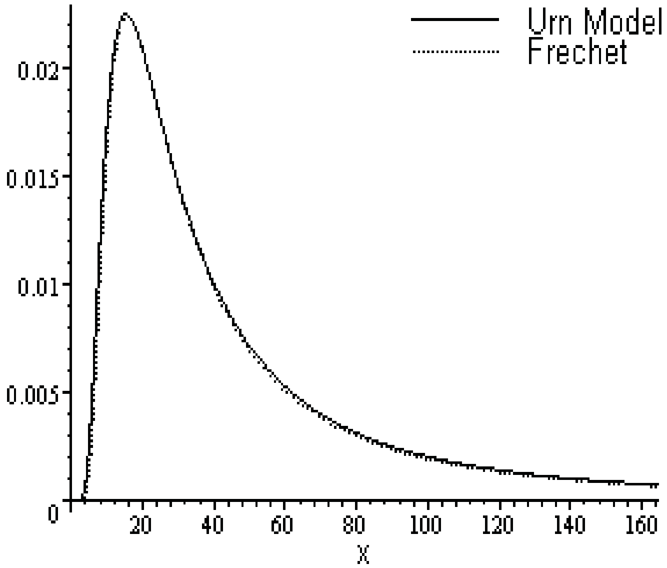

Figure 3.10: Plot of (3.101) and its approximation (3.103) for $n = 40$.

From a mathematical point of view, the Urn Model version from a Pareto distribution must have the same mean value, so we can establish

$$\frac{\zeta\beta}{\beta - 1} = \frac{\delta + \alpha - 1}{\alpha - 1}, \tag{3.104}$$

where the left term is the Pareto's mean value. Remember that ζ and β are the location and shape parameters, respectively, from a Pareto distribution, while that α is the Urn Model parameter, defined as $\alpha = d/m$.

On the other hand, in Section 3.1.4 it was indicated that the Urn Model's heavy-tail behavior depends predominantly on α; such a fact suggests us that we can assume that α can be of the same value as β. Under such assumptions, from Equation (3.104), we can write

$$\delta = \beta(\zeta - 1) + 1. \tag{3.105}$$

Even though the Urn Model's parameters have been defined, the definition of the Urn Model's initial conditions are still missing. From the two definitions given to α, namely

$$\alpha = \beta = \frac{d}{m}, \tag{3.106}$$

it is easy to see that d and m are given by the smallest rational number that gives the best approximation to β. The value of c can be derived form Equation (3.105), and by replacing

$\delta = \frac{c}{m}$ and $\beta = \frac{d}{m}$ and after making simplifications, we have

$$c = d(\zeta - 1) + m. \tag{3.107}$$

Under the last conditions in Figure 3.11, we can see the P-P plot for a Pareto distribution and its Urn Model equivalent representation for $\zeta = 5$ and $\beta = 1.25$. As can be seen, the Urn Model's representation sub-estimates the Pareto distribution that generates it at the beginning of the sample space but, towards the end, their tail decay is almost identical.

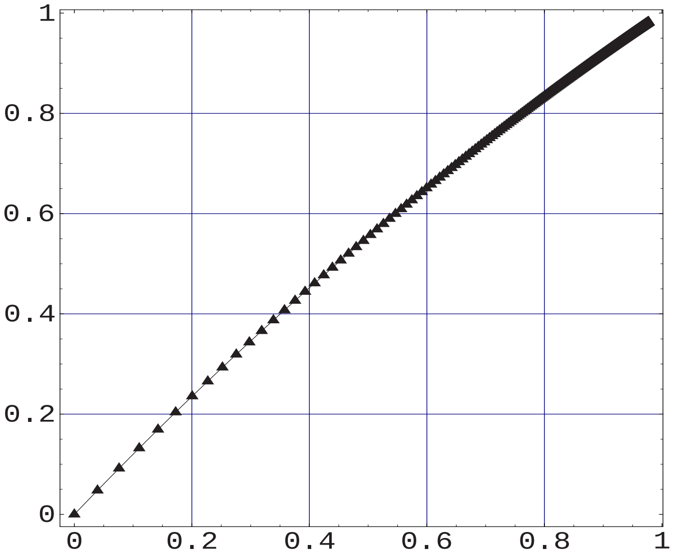

Figure 3.11: P-P plot, Pareto distribution versus its Urn Model representation.

Chapter 4

Urn Model Behavior

In this chapter, the Urn Model's heavy-tail behavior is validated from a practical point of view. We make an intensive use of graphical tools such as Q-Q and P-P plots. Since the probability of observing a connection ending at a given time observation x depends mainly on the "disconnect" balls in the model, graphics are obtained for different values of α, and ranges of interests are pointed. On the other hand, in order to show the versatility of the Urn Model, the match of the degenerated case ($m = 0$) with the geometric distribution is collaborated with the same methodology.

4.1 Heavy-Tail Behavior

In chapter two, it was said that the Pareto distribution was the simplest heavy-tailed distribution, so as a first snapshot of the heavy-tail nature of the Urn Model in Figure 4.1, we can see the Pareto and Urn Model pdf's plots in order to compare their shapes. As can be seen, both densities present a high similarity in their tail decay, even since low values of x.

As expected, their cdf's must present almost identical shapes, as we can see in Figure 4.2. As these graphics denote a similar behavior with the Pareto distribution, we can conclude that the Urn Model really possesses a heavy-tail nature.

As well the Urn Model distribution presents a high similarity with the Pareto distribution, we need to validate this similarity to a specific description of a heavy-tail distribution. One way to do this, is by obtaining the Quantile-Quantile plot of both distributions. The idea of quantile plots, (*QQ-plots for short*) has come forward from the observation that for important classes of distributions the quantiles $Q(p)$ are *linearly related* with the corresponding quantiles of a standard example from this class of distributions. A 45-degree reference line can also be mounted on the QQ-plot. If the quantiles of both distributions are similar, the points should fall approximately along this reference line. The greater the departure from this reference line, the greater the evidence for the conclusion that the two

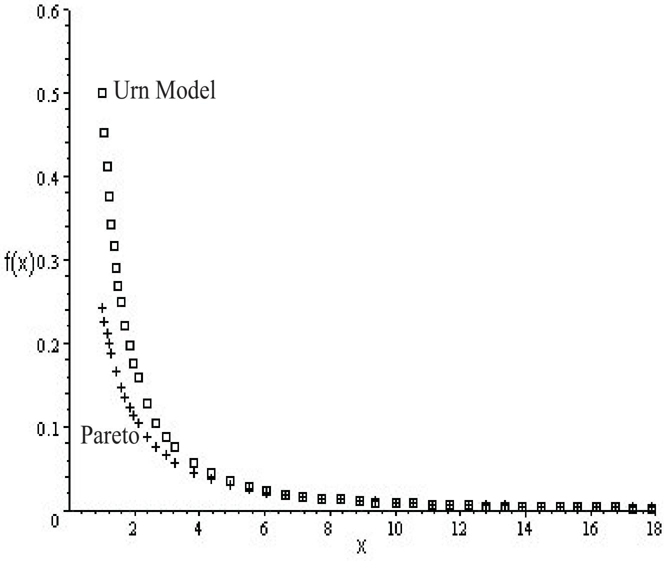

Figure 4.1: Comparison between the Pareto and Urn Model pdf's.

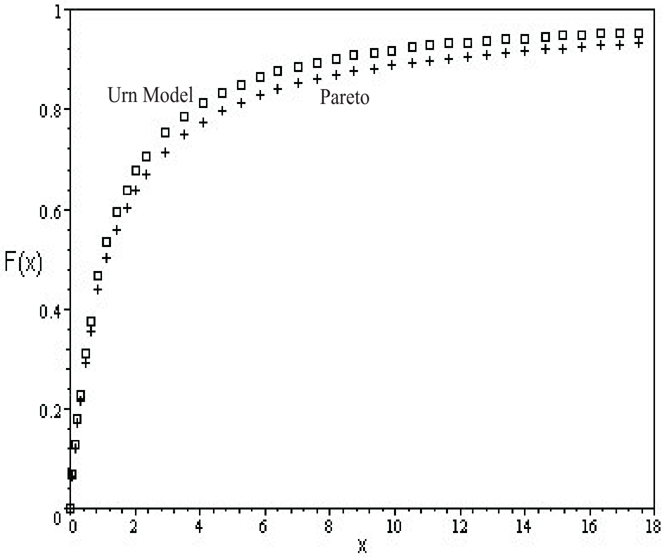

Figure 4.2: Comparison between the Pareto and Urn Model cdf's.

distributions present divergent behaviors. On the other hand, if both distributions differ only by a shift in location, the points should lie along a straight line that is displaced either up or down from the 45−degree reference line. Therefore, as linearity in a graph can be easily checked by the naked eye, we can compare the Urn Model's quantiles against the Pareto distribution quantiles, and validate the Urn Model heavy-tail behavior in a trustworthy way. For more details about this topic, see [23].

In Figure 4.3, we can see the comparison between the quantiles of the Urn Model and the Pareto distribution for the case of $\alpha > 1$. As we can observe, the Urn Model's heavy-tail behavior is stressed for $c > d$, while for $d > c$, the heavy-tail behavior is reduced as a result of the increased probabilities of a disconnection event.

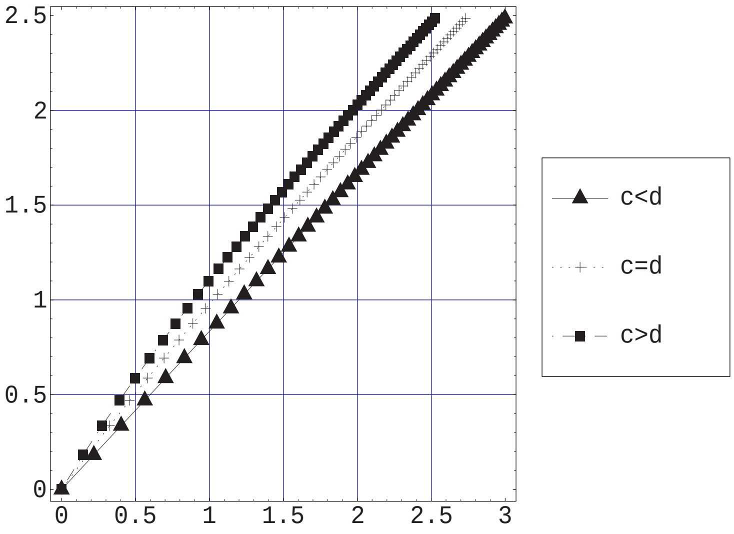

Figure 4.3: Q-Q plot of the Urn Model for $\alpha > 1$.

Another way to validate the similarity between the Urn Model and the Pareto distribution is by obtaining the Probability-Probability plot (*PP-plots for short*) for both distributions. The probability plot is a graphic technique for assessing whether or not a data set follows a given distribution. The data are plotted against a theoretical distribution in such a way that the points should form approximately a straight line. Departures from this straight line indicate departures from the specified distribution. In our case, as we already have a model at hand, this plot is formed by the cdf's values of each distribution obtained to the same value of x. Figure 4.4 shows the Urn Model P-P plot versus the Pareto distribution for the case $\alpha > 1$. As we can see, the results agree, since for $c > d$ the Urn Model's probabilities are lower than the Pareto ones as a result of a stressed heavy-tail behavior for this case. In the same way, for $d > c$, the Urn Model heavy-tail behavior is similar to Paretian as result of the increased probabilities of a disconnection event.

Another case of interest is for $\alpha < 1$, that is, for infinite media. What this means is that

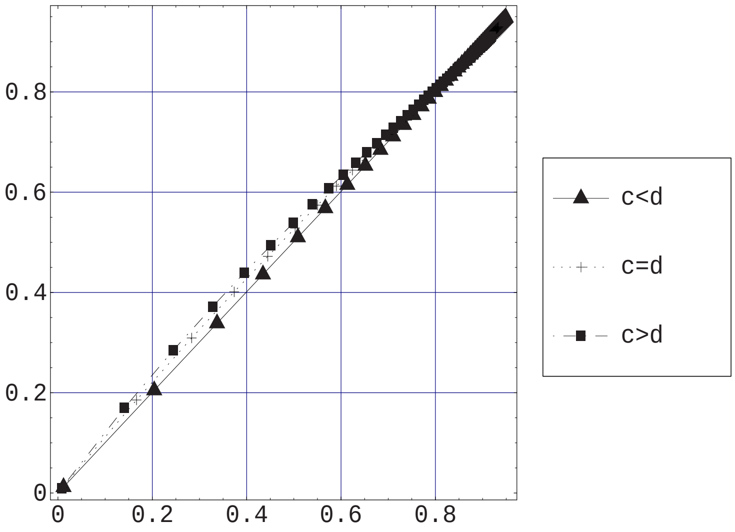

Figure 4.4: P-P plot of the Urn Model for $\alpha > 1$.

a highest level of variability is captured by the Urn Model distribution. Therefore, what we can expect is very heavy tail behavior, which implies a slope beyond the 45 degrees. This is what is exactly shown in Figure 4.5. Against we observe a stressed heavy-tail behavior for $c > d$, a match with the Paretian behavior for $c = d$ and a reduced heavy-tail behavior for $c < d$, despite $\alpha < 1$.

Figure 4.6 reaffirms the previous conclusions. That is, for $c > d$ and $c = d$, the Urn model really presents a heavy-tail behavior that falls inside the Paretian “family”, while in that for $c < d$, the Urn Model even presents a heavy-tail behavior with some discrepancies at the beginning of the distribution’s *low values*.

The third case of interest is for finite variance; that is, for $\alpha > 2$. If the variance is finite, the Urn Model captures a reduced variability which is far away from the heavy-tail behavior. But, in order to observe how the Urn model breaks the heavy-tail nature, in Figure 4.7 we can see the QQ-plots of interest. As we can observe, for $c > d, c = d$ and $c < d$, the Urn Model loses its heavy-tail behavior, and there is almost no difference in each case. This can be collaborated in Figure 4.8, where the PP-plots for $c > d, c = d$ and $c < d$ are shown. As we can see, there are not any visible differences among each case.

From previous figures, we can make the following conclusions:

- The Urn model presents a truly heavy-tail behavior only for $0 < \alpha < 2$.
- For practical proposes (*mean finite*), the range of interest can be limited to $1 < \alpha < 2$.

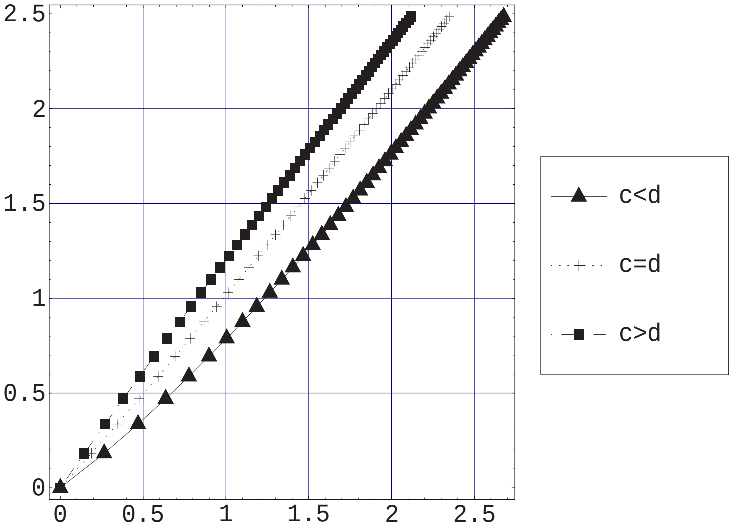

Figure 4.5: Q-Q plot of the Urn Model for $\alpha < 1$.

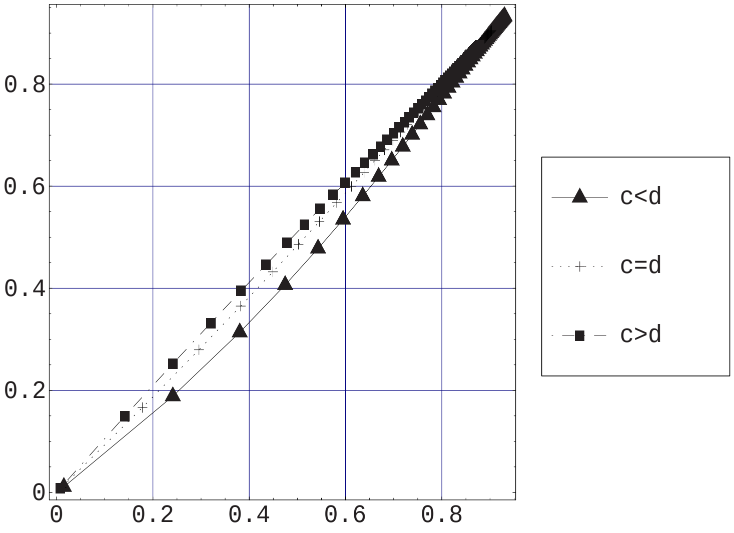

Figure 4.6: P-P plot of the Urn Model for $\alpha < 1$.

- The Urn model's heavy-tail behavior can be handled from a *typical or Paretian behavior* until a *stressed or very heavy-tail* approximation for $c = d$ and $c > d$, respectively.

- The case of $d > c$ results in a light heavy-tail behavior, and it is not of great interest.

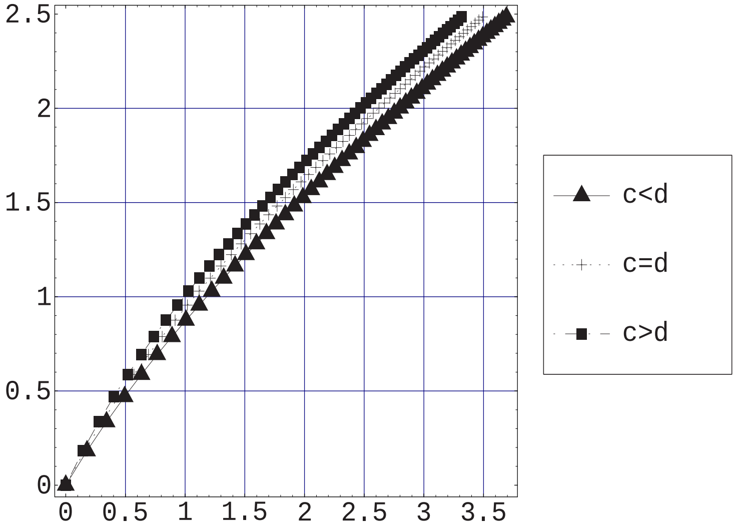

Figure 4.7: Q-Q plot of the Urn Model for $\alpha > 2$.

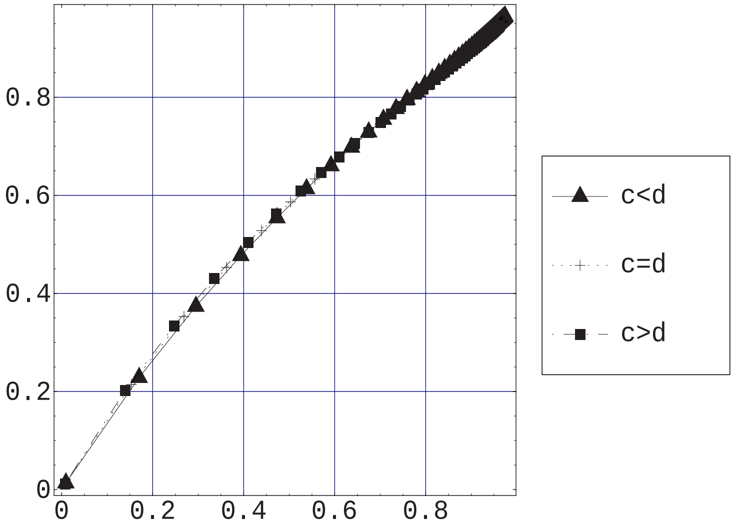

Figure 4.8: P-P plot of the Urn Model for $\alpha > 2$.

- Urn Model can be used to get the discrete version from a given Pareto distribution with identical characteristics.

4.2 Geometric Behavior

In chapter three, it was pointed out that for $m = 0$, the Urn Model reduces to the geometric model. Since the Urn Model's mathematical description is not supported for $m = 0$, what we can do is replace m for a value near 0 and observe its behavior. In Figure 4.9, we can see the Urn model and geometric cdf's, as we can observe; the Urn model really behaves as a Geometric distribution for $m \sim 0$. In order to appreciate the match between the degenerated case of the Urn Model with the geometric distribution, Figure 4.10 shows the P-P plot for both distributions. As we can observe, the match is perfect for the entire sample space.

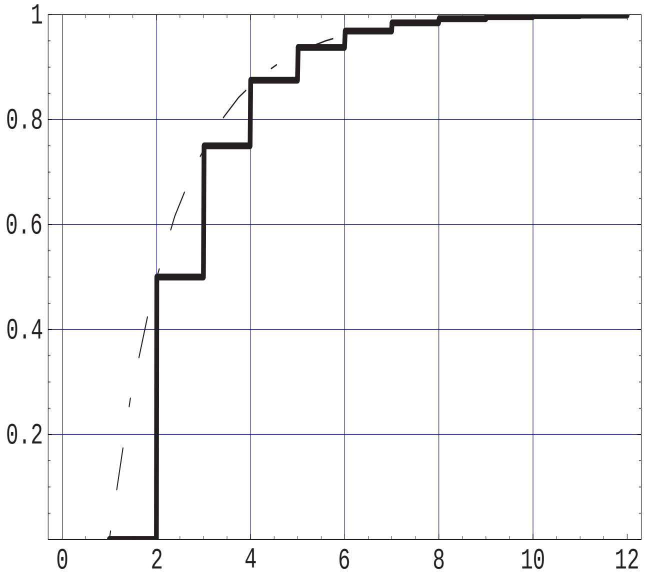

Figure 4.9: Urn Model and Geometric cdf's.

On the other hand, in order to appreciate how far away the Urn Model is from the Paretian behavior in the degenerated case, in Figure 4.11 we can see the P-P plots for the Pareto and geometric distributions. As can be easily noted, the Urn Model for $m \approx 0$ loses all of its heavy-tail nature and reduces to the geometric model.

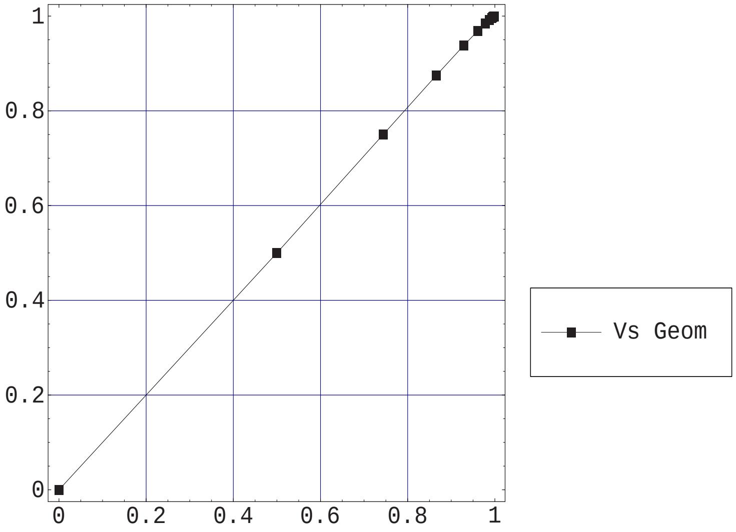

Figure 4.10: P-P plot of the Urn Model for $m \approx 0$.

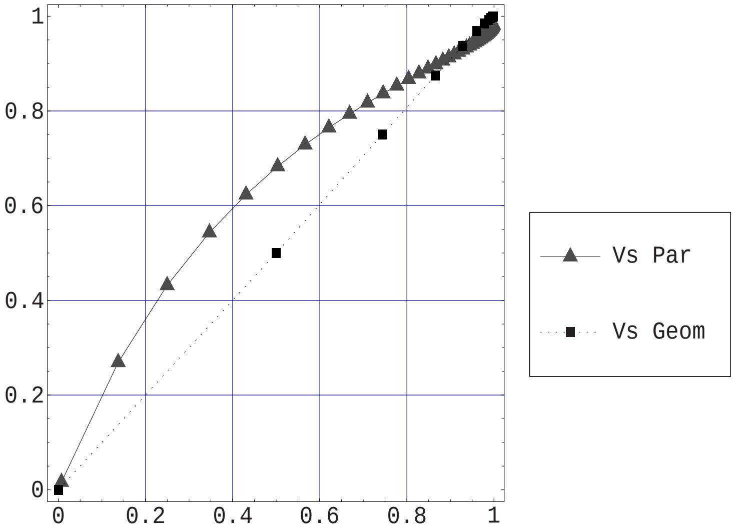

Figure 4.11: Urn Model P-P plots for $m \approx 0$.

Chapter 5

Conclusions

In this chapter, general conclusions of this work are presented. Also, certain projects for further research are suggested.

5.1 General Conclusions

This work has completed the description of the discrete heavy tail distribution presented in [24]. It has shown through mean excess functions, that the Urn Model really poses a heavy tail; which decay is predominantly dominated by the original *disconnect* condition. Moreover, by a graphical comparison with the Pareto's mean excess functions we found a high similarity between them. Of special attention is the fact that the Urn Model‘s parameter denoted as "α" plays a similar role that the Pareto's shape parameter denoted as "β". This is in the sense that, both parameters determinate the tail decay of their respective distribution. From the Urn Model's generating functions analysis realized in this work, were derived the most representative model's moments. Due to moments existence in specific values for "α"; the Urn Model heavy tail behavior has been identified in the interval $0 < \alpha < 2$. Last fact points towards an Urn Model's performance near to the Paretian behavior and is the motivation for the study of the match between both models presented at the end of Chapter 3. From this study, we conclude that is possible to find an Urn Model representation from a well defined Pareto Distribution at the hand, with the same mean value and identical tail decay.
In order to investigate the domain of attraction for the Urn Model maximum extreme X_n; a match between the Urn Model and the Frechet distribution was done. During this analysis, was stressed the shaping nature of the parameter α and the scaling properties of the δ parameter in the Urn Model. Thanks to these properties the Frechet distribution arise as the limiting distribution for the Urn Model maximum extreme X_n even for reduced samples.

Summarizing we conclude that

- The Urn Model really poses a heavy tail; which decay is predominantly dominated by the original *disconnect* condition.
- The Urn Model's parameter denoted as "α" plays a similar role that the Pareto's shape parameter denoted as "β".
- The Urn Model heavy tail behavior has been identified in the interval $0 < \alpha < 2$.
- It is possible to find an Urn Model representation from a given Pareto Distribution, with the same mean value and identical tail decay.
- The Urn Model can be an important tool in the design and simulation of internetworking devices, because it allows handling of the traffic's shape from light tailed (m=0) until heavy tailed ($0 < d < 2m$).
- From a given Urn Model is possible to obtain its Frechet equivalent representation with a small difference between its mean values.
- The Frechet distribution arise as the limiting distribution for the Urn Model maximum extreme X_n.
- Only the Urn Model mean value is useful in the establishment of the maximum value expected.

5.2 Future Research

There are certain research projects that can continue this work, due to stability implications. Among these are the following

- Investigate the Urn Model stability through alpha stable distributions theory.
- Pareto type workloads have been used in important works [6]. As a validation of the Urn Model utility, some of these works could be repeated with its respective Urn Model version of its workload. As well, this can be done with the finality to clarify the match between both distributions, the persistence of certain Paretian workloads properties can be investigated.
- Perform a statistical study about the buffer occupancy in a switch with heavy-tail input traffic generated through the Urn Model.

Bibliography

[1] Bailey, Norman T. J., *The elements of Stochastics Processes,* First Edition, John Wiley & Sons, Inc., 1964.

[2] Barford, P., Bestavros, A., Bradley, A. and Crovella, M.E., "Changes in Web Client Access Patterns: Characteristics and Caching Implications," *World Wide Web, Special Issue on Characterization and Performance Evaluation*, vol. 2, pp. 15-28, 1999.

[3] Bury, Karl, *Statistical Distributions in Engineering,* Cambridge University Press, 1999.

[4] Crovella, M.E. and Bestavros, A., "Self-Similarity in WWW Traffic: Evidence and Possible Causes," *IEEE/ACM Trans. on Net.*, vol. 5, No 6, pp. 835-846, Dec. 1997.

[5] Crovella, M.E., Bestavros, A. and Taqqu M., "Heavy Tailed Probability Distributions in the WWW," *A Practical Guide to Heavy Tails*, Birkhäuser, 1998.

[6] Crovella, M.E. and Lipsky, L., "Simulations with Heavy-Tailed workloads," *Self-Similar Network Traffic and Performance Evaluation, Edited by Kihong Park and Walter Willinger*, John Wiley & Sons, Inc., pp. 89-100, 2000.

[7] Cunha, C.R., Bestavros, A. and Crovella M.E., "Characteristics of WWW Client based Traces," *Technical Report TR-95-010, Boston University Computer Science Department*, June 1995.

[8] Danzig, P., Jamin, S., Cáceres, R., Mitzel, D. and Estrin, D., "An empirical workload model for driving wide-area TCP/IP network simulations," *Internetworking: Res., Experience*, vol. 3, no. 1, pp. 1-26, March 1992.

[9] Duffy, D., McIntosh, A., Rosenstein, M. and Willinger, W., "Statistical analysis of CCSN/SS7 traffic data from working CCS subnetworks," *IEEE Journal Select. Areas Commun.*, vol. 12, pp. 544-551, April 1994.

[10] Feller, William, *An Introduction to Probability Theory and Its Applications,* Third Edition, vol. 1, John Wiley & Sons, Inc., 1968.

[11] Feller, William, *An Introduction to Probability Theory and Its Applications,* Third Edition, vol. 2, John Wiley & Sons, Inc., 1968.

[12] Frost, V. and Melamed, B., "Traffic modeling for telecommunications networks," *IEEE Communications Magazine*, vol. 33, pp. 70-80, March 1994.

[13] Garrett, M. and Willinger, W., "Analysis, modeling, and generation of self-similar VBR video traffic," *ACM/SIGCOMM Proccedings*, September 1994.

[14] Gradshteyn, I.S. and Ryzhic, I.M., *Table of Integrals, Series and Products,* Fifth Edition, Academic Press, 1996.

[15] Harchol Balter, M. and Downey A., "Exploiting process lifetime distributions for dynamic load balancing," *ACM Trans. Comput. Syst.*, vol. 3, pp. 253-285, 1997.

[16] Hoel, P.G., Port, S.C. and Stone, C.J., *Introduction to Probability Theory,* First Edition, Houghton Mifflin Company, 1971.

[17] Kauffman, S., *The Origins of Order: Self-Organization and Selection in Evolution,*, London, England; Oxford University Press, 1993.

[18] Leland, W.E. and Ott, T.J., "Load-balancing heuristics and process behavior," *Proceedings of Performance and ACM Sigmetrics*, pp. 54-69, 1986.

[19] Leland, W.E., Taqqu, M.S., Willinger, W. and Wilson, D.V., "On the self-similar nature of Ethernet traffic (extended version)," *IEEE/ACM Transactions on Networking*, vol. 2, pp. 1-15, 1994.

[20] Mandelbrot, B., "New methods in statistical economics," *Political Economic Journal*, vol. 71, no. 5, pp. 421-440, Oct. 1963.

[21] Paxson, V. and Floyd, S., "Wide Area Traffic: The Failure of Poisson Modeling," *IEEE/ACM Transactions on Networking*, vol. 3, No 3, pp. 226 - 244, June 1995.

[22] Peterson, D.L., "Data center I/O patterns and power laws," *CMG Proceedings*, December 1996.

[23] Teugels, J.L., Beirlant, J. and Vynckier, P., *Practical Analysis of Extreme Values,* First Edition, Leuven University Press, 1996.

[24] Villareal Reyes, Salvador, "Discrete Heavy Tailed Distributions for Network Traffic Modeling," *M. Sc. in Electronic Systems Engineering (Telecommunications)ITESM-Campus Monterrey*, June 2001.

[25] Wolfgang, P. and Jörg, B., *Stochastic Processes: From Physics to Finance,* Springer-Verlag, 1999.