

UNIVERSIDADE ESTADUAL PAULISTA

Instituto de Geociências e Ciências Exatas

Câmpus de Rio Claro

**A DISTRIBUIÇÃO DE LEI DE POTÊNCIA  
GRADUALMENTE TRUNCADA NA EDUCAÇÃO:  
EXAME DE VESTIBULAR DA UNESP**

Fábio Roberto Chavarette

Orientador: Prof.Dr. Hari Mohan Gupta

Dissertação de Mestrado elaborada junto ao  
Curso de Pós-Graduação em Física - Área de  
Concentração em Física Aplicada, para  
Obtenção do Título de Mestre em Física.

Rio Claro (SP)  
2002

Comissão Examinadora

Prof. Dr. Hari Moham Gupta - Orientador  
Instituição: IGCE/Rio Claro

Prof. Dr. José Roberto Campanha  
Instituição: IGCE/Rio Claro

Prof. Dr. Elbert Einstein Nehrer Macau  
Instituição: INPE/São José dos Campos

Fábio Roberto Chavarette

Rio Claro, 27 de Novembro de 2002.

Resultado: Aprovado com Louvor.

## AGRADECIMENTOS

Ao meu mestre e orientador, Professor Doutor **Hari Mohan Gupta** gostaria de agradecer por ter-me recebido como seu aluno, e pela generosidade e atenção que sempre dispensou à minha pessoa.

Ao Professor Doutor **José Roberto Campanha**, pela inestimável ajuda, atenção e apoio.

Aos Professores Doutores **José Silvio Govone** e **José Manoel Balthazar**, que me proporcionou a oportunidade de realizar este sonho.

Ao Instituto de Geociências e Ciências Exatas e ao Departamento de Física da Unesp Campus de Rio Claro.

A todo o pessoal do Instituto de Geociências e Ciências Exatas, pela forma carinhosa e gentil com que sempre me trataram.

A todos os meus amigos, e em especial a **Danilo, Edna, Érika** e **Gislaine**, pelo apoio e carinho.

Àqueles que, mesmo ausentes, sempre estiveram comigo.

Aos meus pais **Silvio** e **Jane**, minha irmã **Alessandra**, e ao meu cunhado **Junior**, presenças fundamentais em minha vida.

À minha sobrinha **Thalita**, a quem pertence o futuro.

A **Fulvia**, eterna presença.

## SUMÁRIO

Índice.....	i
Resumo.....	iii
Abstract.....	iv
Índice de Tabelas.....	v
Índice de Figuras.....	vii
I – Sistemas Complexos.....	01
II – Distribuições Estatísticas.....	08
III – Vestibular um Sistema Complexo.....	17
IV – Conclusões.....	47
V – Referencias Bibliográficas.....	55

## ÍNDICE

I - Sistemas Complexos.....	01
1.1. – Introdução.....	01
1.2. – A Complexidade da Natureza.....	02
1.3. – Características de Sistemas Complexos.....	04
1.3.1. – Distribuição de Lei de Potência.....	05
1.3.2. – Geometria Fractal.....	05
1.3.3. – Ruído 1 / f .....	06
1.3.4. – Lei de Zipf.....	07
II – Distribuições Estatísticas.....	08
2.1. – Distribuição Normal.....	08
2.2. – Distribuição Log-Normal.....	10
2.3. – Distribuição de Lei de Potência e Realimentação Positiva.....	11
2.4. – Distribuição de Lévy.....	12
2.5. – Distribuição Gradualmente Truncada.....	13
2.6. – Gráficos do Tipo Zipf (Zipf Plot).....	15
III – Vestibular um Sistema Complexo.....	17
3.1. – O Vestibular da Unesp.....	17
3.2. – Distribuição Estatística com Ênfase no Período.....	20
3.2.1. – Área de Ciências Exatas.....	20
3.2.2. – Área de Ciências Biológicas.....	23
3.2.3. – Área de Ciências Humanas.....	26
3.3. – Distribuição Estatística com Ênfase no Tipo de Instituição.....	29
3.3.1. – Área de Ciências Exatas.....	29
3.3.2. – Área de Ciências Biológicas.....	32
3.3.3. – Área de Ciências Humanas.....	35
3.4. – Distribuição Estatística com Ênfase na Renda Familiar.....	38
3.4.1. – Área de Ciências Exatas.....	38
3.4.2. – Área de Ciências Biológicas.....	41
3.4.3. – Área de Ciências Humanas.....	44
IV– Conclusões.....	47
4.1. - Alunos que cursaram o ensino médio no período diurno ou noturno.....	47

4.2. - Alunos que cursaram o ensino médio em escola privada ou escola pública.....	49
4.3. - Alunos que cursaram o ensino médio que possuem renda familiar até 10 salários mínimos e acima de 10 salários mínimos.....	50
4.4. – Conclusão Geral.....	52
V – Referencias Bibliográficas.....	55
VI – Apêndice.....	60
5.1. – Artigo Aceito para Publicação.....	60

## RESUMO

Nesse trabalho analisamos a distribuição estatística das notas dos candidatos do vestibular da UNESP (Universidade Estadual Paulista) nos anos de 1998, 1999 e 2000, levando em conta o fato do aluno ter cursado o ensino médio no período diurno ou noturno, ter cursado em escola pública ou privada e possuir renda familiar até 10 salários mínimos ou superior a 10 salários mínimos. O vestibular é dividido em três áreas: Ciências Exatas, Ciências Biológicas e Ciências Humanas. Para cada área o candidato realiza três avaliações: conhecimentos específicos, conhecimentos gerais e língua portuguesa. O enfoque desse estudo é a avaliação de conhecimentos específicos das três áreas. Comparamos as distribuições das notas dos alunos que cursaram o ensino médio no período diurno e noturno, em escola pública e privada e com renda até 10 salários mínimos e superior a 10 salários mínimos, observamos a presença da distribuição de lei de potência gradualmente truncada e da distribuição normal como observado em vários sistemas naturais complexos. Finalmente discutimos e justificamos os resultados obtidos dessas distribuições.

**Palavras-chave:** Distribuição de Lei de Potência, Distribuição Gradualmente Truncada, Distribuição Normal, Vestibular, Ensino Médio.

## A B S T R A C T

We studied the statistical distribution of student's performance, which is measured through their marks, in university entrance examination (Vestibular) of UNESP (Universidade Estadual Paulista) with respect to (i) period of study –day vs. night period (ii) teaching conditions – private vs. public school (iii) economical conditions – high vs. low family income. We observed long ubiquitous power law tails in Physical and Biological sciences in all cases. The mean value increases with better study conditions followed by better teaching and economical conditions. In humanities, the distribution is close to normal distribution with very small tail. This indicates that these power law tails in science subjects are due to the nature of the subject itself. Further better study, teaching and economical conditions are more important for physical and biological science in comparison to humanities at this level of study. We explain these statistical distributions through Gradually Truncated Power Law Distributions. We discuss the possible reason for this peculiar behaviour.

**Keywords:** Power-Law Distributions, Gradually Truncated Distributions, Normal Distribution, University Entrance Examination, Average Education.

## LISTA DE TABELAS

- Tabela 1:** Tabela de parâmetros para o melhor ajuste com os dados empíricos na Distribuição de Lei de Potência Gradualmente Truncada, para os dados dos alunos que cursaram o ensino médio no período diurno e noturno para os anos de 1998, 1999 e 2000 na área de ciências exatas..... 22
- Tabela 2:** Tabela de parâmetros para o melhor ajuste com os dados empíricos na Distribuição de Lei de Potência Gradualmente Truncada, para os dados dos alunos que cursaram o ensino médio no período diurno e noturno para os anos de 1998, 1999 e 2000 na área de ciências biológicas..... 25
- Tabela 3:** Tabela de parâmetros para o melhor ajuste com os dados empíricos na Distribuição Normal, para os dados dos alunos que cursaram o ensino no período diurno e noturno para os anos de 1998, 1999 e 2000 na área de ciências humanas..... 28
- Tabela 4:** Tabela de parâmetros para o melhor ajuste com os dados empíricos na Distribuição de Lei de Potência Gradualmente Truncada para os dados dos alunos que cursaram o ensino médio na escola particular e publica para os anos de 1998, 1999 e 2000 na área de ciências exatas..... 31
- Tabela 5:** Tabela de parâmetros para o melhor ajuste com os dados empíricos na Distribuição de Lei de Potência Gradualmente Truncada para os dados dos alunos que cursaram o ensino médio na escola particular e publica para os anos de 1998, 1999 e 2000 na área de ciências biológicas..... 34
- Tabela 6:** Tabela de parâmetros para o melhor ajuste com os dados empíricos na Distribuição Normal, para os dados dos alunos que cursaram o ensino médio na escola particular e publica para os anos de 1998, 1999 e 2000 na área de ciências humanas..... 37

<b>Tabela 7:</b> Tabela de parâmetros para o melhor ajuste com os dados empíricos na Distribuição de Lei de Potência Gradualmente Truncada para os dados dos alunos que cursaram o ensino médio com renda acima de 10 salários mínimos e até 10 salários mínimos para os anos de 1998, 1999 e 2000 na área de ciências exatas.....	40
<b>Tabela 8:</b> Tabela de parâmetros para o melhor ajuste com os dados empíricos na Distribuição de Lei de Potência Gradualmente Truncada para os dados dos alunos que cursaram o ensino médio com renda acima de 10 salários mínimos e até 10 salários mínimos para os anos de 1998, 1999 e 2000 na área de ciências biológicas.....	43
<b>Tabela 9:</b> Tabela de parâmetros para o melhor ajuste com os dados empíricos na Distribuição Normal para os dados dos alunos que cursaram o ensino médio com renda acima de 10 salários mínimos e até 10 salários mínimos para os anos de 1998, 1999 e 2000 na área de ciências humanas.....	46

## LISTA DE FIGURAS

- Figura 1:** Números de candidatos que obtiveram nota  $x$  [ $N(x)$  versus notas obtidas ( $x$ )] para a avaliação de conhecimentos específicos da área de ciências exatas do ano de 1998 para alunos do período diurno ( ) e noturno (o) no ensino médio..... 21
- Figura 2:** Números de candidatos que obtiveram nota  $x$  [ $N(x)$  versus notas obtidas ( $x$ )] para a avaliação de conhecimentos específicos da área de ciências exatas do ano de 1999 para alunos do período diurno ( ) e noturno (o) no ensino médio..... 21
- Figura 3:** Números de candidatos que obtiveram nota  $x$  [ $N(x)$  versus notas obtidas ( $x$ )] para a avaliação de conhecimentos específicos da área de ciências exatas do ano de 2000 para alunos do período diurno ( ) e noturno (o) no ensino médio..... 22
- Figura 4:** Números de candidatos que obtiveram nota  $x$  [ $N(x)$  versus notas obtidas ( $x$ )] para a avaliação de conhecimentos específicos da área de ciências biológicas do ano de 1998 para alunos do período diurno ( ) e noturno (o) no ensino médio..... 24
- Figura 5:** Números de candidatos que obtiveram nota  $x$  [ $N(x)$  versus notas obtidas ( $x$ )] para a avaliação de conhecimentos específicos da área de ciências biológicas do ano de 1999 para alunos do período diurno ( ) e noturno (o) no ensino médio..... 24
- Figura 6:** Números de candidatos que obtiveram nota  $x$  [ $N(x)$  versus notas obtidas ( $x$ )] para a avaliação de conhecimentos específicos da área de ciências biológicas do ano de 2000 para alunos do período diurno ( ) e noturno (o) no ensino médio..... 25
- Figura 7:** Números de candidatos que obtiveram nota  $x$  [ $N(x)$  versus notas obtidas ( $x$ )] para a avaliação de conhecimentos específicos da área de ciências humanas do ano de 1998 para alunos do período diurno ( ) e noturno (o) no ensino médio..... 27
- Figura 8:** Números de candidatos que obtiveram nota  $x$  [ $N(x)$  versus notas obtidas ( $x$ )] para a avaliação de conhecimentos específicos da área de

ciências humanas do ano de 1999 para alunos do período diurno (†) e noturno (o) no ensino médio.....	27
<b>Figura 9:</b> Números de candidatos que obtiveram nota x [N(x) versus notas obtidas (x)] para a avaliação de conhecimentos específicos da área de ciências humanas do ano de 2000 para alunos do período diurno (†) e noturno (o) no ensino médio.....	28
<b>Figura 10:</b> Números de candidatos que obtiveram nota x [N(x) versus notas obtidas (x)] para a avaliação de conhecimentos específicos da área de ciências exatas do ano de 1998 para alunos da escola particular (†) e publica (o) no ensino médio.....	29
<b>Figura 11:</b> Números de candidatos que obtiveram nota x [N(x) versus notas obtidas (x)] para a avaliação de conhecimentos específicos da área de ciências exatas do ano de 1999 para alunos da escola particular (†) e publica (o) no ensino médio.....	30
<b>Figura 12:</b> Números de candidatos que obtiveram nota x [N(x) versus notas obtidas (x)] para a avaliação de conhecimentos específicos da área de ciências exatas do ano de 2000 para alunos da escola particular (†) e publica (o) no ensino médio.....	30
<b>Figura 13:</b> Números de candidatos que obtiveram nota x [N(x) versus notas obtidas (x)] para a avaliação de conhecimentos específicos da área de ciências biológicas do ano de 1998 para alunos da escola particular (†) e publica (o) no ensino médio.....	33
<b>Figura 14:</b> Números de candidatos que obtiveram nota x [N(x) versus notas obtidas (x)] para a avaliação de conhecimentos específicos da área de ciências biológicas do ano de 1999 para alunos da escola particular (†) e publica (o) no ensino médio.....	33
<b>Figura 15:</b> Números de candidatos que obtiveram nota x [N(x) versus notas obtidas (x)] para a avaliação de conhecimentos específicos da área de ciências biológicas do ano de 2000 para alunos da escola particular (†) e publica (o) no ensino médio.....	34
<b>Figura 16:</b> Números de candidatos que obtiveram nota x [N(x) versus notas obtidas (x)] para a avaliação de conhecimentos específicos da área	

de ciências humanas do ano de 1998 para alunos da escola particular (†) e publica (o) no ensino médio.....	36
<b>Figura 17:</b> Números de candidatos que obtiveram nota x [N(x) versus notas obtidas (x)] para a avaliação de conhecimentos específicos da área de ciências humanas do ano de 1999 para alunos da escola particular (†) e publica (o) no ensino médio.....	36
<b>Figura 18:</b> Números de candidatos que obtiveram nota x [N(x) versus notas obtidas (x)] para a avaliação de conhecimentos específicos da área de ciências humanas do ano de 2000 para alunos da escola particular (†) e publica (o) no ensino médio.....	37
<b>Figura 19:</b> Números de candidatos que obtiveram nota x [N(x) versus notas obtidas (x)] para a avaliação de conhecimentos específicos da área de ciências exatas do ano de 1998 para alunos com renda acima de 10 salários mínimos (†) e abaixo de 10 salários mínimos (o) no ensino médio.....	38
<b>Figura 20:</b> Números de candidatos que obtiveram nota x [N(x) versus notas obtidas (x)] para a avaliação de conhecimentos específicos da área de ciências exatas do ano de 1999 para alunos com renda acima de 10 salários mínimos (†) e abaixo de 10 salários mínimos (o) no ensino médio.....	39
<b>Figura 21:</b> Números de candidatos que obtiveram nota x [N(x) versus notas obtidas (x)] para a avaliação de conhecimentos específicos da área de ciências exatas do ano de 2000 para alunos com renda acima de 10 salários mínimos (†) e abaixo de 10 salários mínimos (o) no ensino médio.....	39
<b>Figura 22:</b> Números de candidatos que obtiveram nota x [N(x) versus notas obtidas (x)] para a avaliação de conhecimentos específicos da área de ciências biológicas do ano de 1998 para alunos com renda acima de 10 salários mínimos (†) e abaixo de 10 salários mínimos (o) no ensino médio.....	42
<b>Figura 23:</b> Números de candidatos que obtiveram nota x [N(x) versus notas obtidas (x)] para a avaliação de conhecimentos específicos da área de	

ciências biológicas do ano de 1999 para alunos com renda acima de 10 salários mínimos (†) e abaixo de 10 salários mínimos (o) no ensino médio.....	42
<b>Figura 24:</b> Números de candidatos que obtiveram nota x [N(x) versus notas obtidas (x)] para a avaliação de conhecimentos específicos da área de ciências biológicas do ano de 2000 para alunos com renda acima de 10 salários mínimos (†) e abaixo de 10 salários mínimos (o) no ensino médio.....	43
<b>Figura 25:</b> Números de candidatos que obtiveram nota x [N(x) versus notas obtidas (x)] para a avaliação de conhecimentos específicos da área de ciências humanas do ano de 1998 para alunos com renda acima de 10 salários mínimos (†) e abaixo de 10 salários mínimos (o) no ensino médio.....	45
<b>Figura 26:</b> Números de candidatos que obtiveram nota x [N(x) versus notas obtidas (x)] para a avaliação de conhecimentos específicos da área de ciências humanas do ano de 1999 para alunos com renda acima de 10 salários mínimos (†) e abaixo de 10 salários mínimos (o) no ensino médio.....	45
<b>Figura 27:</b> Números de candidatos que obtiveram nota x [N(x) versus notas obtidas (x)] para a avaliação de conhecimentos específicos da área de ciências humanas do ano de 2000 para alunos com renda acima de 10 salários mínimos (†) e abaixo de 10 salários mínimos (o) no ensino médio.....	46

# CAPÍTULO 1

## SISTEMAS COMPLEXOS

### 1.1 Introdução.

Neste trabalho, propomos tratar o sistema educacional como um sistema complexo. Vários conceitos de sistemas complexos como lei de potência, geometria fractal, ruído  $1/f$  e lei de zipf, vem sendo aplicados em sistemas físicos, biológicos e econômicos (Zipf, 1949; Mandelbrot, 1983; Bak, 1997).

O sistema educacional a ser tratado é o Exame de Vestibular da UNESP nos anos de 1998, 1999 e 2000, divididos nas áreas de ciências biológicas, ciências humanas e ciências exatas, e estudaremos as distribuições estatísticas obtidas a partir das notas dos candidatos.

Para executarmos tal tarefa, dividiremos o trabalho nas seguintes etapas. No capítulo 1, fazemos uma explanação sobre os objetivos do nosso trabalho e uma introdução sobre os principais conceitos e características dos sistemas complexos.

No segundo capítulo, apresentamos as técnicas estatísticas para o tratamento dos mesmos.

No terceiro capítulo, estudaremos o vestibular da Unesp como um sistema complexo, através das distribuições das notas dos alunos que cursaram o ensino médio nas seguintes categorias:

- alunos que cursaram o ensino médio nos períodos diurno ou noturno.
- alunos que cursaram o ensino médio em escolas públicas ou particulares.
- alunos cuja renda familiar seja menor que dez salários mínimos ou acima de dez salários mínimos.

No quarto capítulo, apresentaremos a discussão e conclusão dos nossos estudos.

Já no quinto capítulo, apresentaremos as referências bibliográficas utilizadas para este trabalho.

## **1.2 A Complexidade da Natureza.**

O Universo supostamente evoluiu de acordo com as leis da física. Os físicos através de observações e experimentos em laboratórios conseguiram encontrar leis gerais da natureza, bem como explicar a estrutura da matéria desde os quarks até grandes galáxias. Também através destas leis desenvolvem-se a tecnologia, com a produção de um grande número de aparelhos, dispositivos e materiais para o conforto do ser humano.

As leis da física são simples e basicamente podem ser escritas como (i) Lei de Movimento de Newton (ii) Equações de Maxwell's para sistemas elétricos e magnéticos (iii) Equações da Relatividade de Einstein para altas velocidades e (iv) Mecânica Quântica para descrição de movimento e energia de partículas como elétrons,

etc. Estas leis poderiam ser escritas por equações em poucas páginas, mas, a matemática envolvida para resolver essas equações é extremamente difícil, principalmente envolvendo sistemas com dois ou mais objetos. Por exemplo, obter as soluções analíticas do movimento de dois planetas em presença de outros planetas e do sol, é extremamente difícil.

Em certos casos específicos, os físicos conseguiram explicar com muitos detalhes, o comportamento de um sistema envolvendo trilhões e trilhões de átomos, moléculas ou elétrons. Por exemplo, o comportamento de um cristal é bem entendido pelas leis básicas da física. Isso acontece porque um cristal pode ser dividido em células contendo muito poucos átomos. Todas as células são iguais e entendendo uma célula, conseguiria entender todo o cristal.

Mas nós não moramos em um mundo simples feito de cristal ou gás. A superfície da Terra possui rios, montanhas, mar, praias, materiais e pedras com estruturas complexas, cada um dos quais com suas características. Nós encontramos na natureza todos os tipos e tamanhos de montanhas, bem como rios, praias, etc. As maiorias das pedras que encontramos não são cristais e possuem estruturas variadas. A grande maioria dos sistemas naturais possuem várias interações ocorrendo ao mesmo tempo e na presença de muitos elementos. Tal fato define o que denominamos de sistema complexo.

Não é possível tratar tal tipo de sistema em detalhes através das leis básicas da física, mas utilizando modelos matemáticos e estatísticos poderemos fazer previsões sobre o comportamento dos mesmos, obtendo assim leis gerais estatísticas.

As técnicas matemáticas desenvolvidas no estudo de sistemas físicos complexos, também podem ser utilizadas em outras áreas, como biologia, educação, economia, sociologia, etc.

O comportamento complexo na natureza reflete a tendência dos grandes sistemas com muitos componentes a evoluir para estados “críticos”, fora do equilíbrio, onde distúrbios menores podem levar a eventos, chamamos de avalanches, de grandes

tamanhos. As maiorias das mudanças nesses sistemas ocorrem através de eventos catastróficos ao invés de uma forma gradual suave. As evoluções para estes estados “críticos” ocorrem sem um agente externo. O estado é estabelecido somente por causa das interações dinâmicas entre os elementos individuais do sistema. Vários mecanismos vêm sendo discutidos para estabelecer este estado. (Badii, 1997; Bak, et al, 1987; Bak, et al, 1988).

### **1.3 Características de Sistemas Complexos.**

A teoria da complexidade é essencialmente de natureza estatística e por isso não pode mostrar detalhes específicos. Por exemplo, podemos prever a probabilidade de acontecer um acidente ou roubo de carro para uma seguradora de carro, mas não podemos dizer que um carro em particular será roubado ou não. No estudo de um gás, não estamos interessados na velocidade e posição de cada molécula, mas sim na função que descreve a distribuição das velocidades das moléculas do gás. Da mesma forma, na educação não estamos interessados na nota individual dos alunos, mas sim nas distribuições estatísticas dessas notas, pois pode ser importante do ponto de vista do planejamento educacional. Mais ainda, mesmo conhecendo essas distribuições, é impossível para um educador prever os resultados de uma instituição ou organização educacional em particular, devido a vários fatores que envolvem o sistema.

Temos algumas observações empíricas gerais ubíquas, em que sistemas complexos de várias características não podem ser explicados por uma característica individual. Precisamos de uma teoria geral abstrata ou alguns mecanismos gerais que poderiam explicar estas observações em todas as áreas. Estas características dos sistemas complexos serão detalhadas a seguir.

### **1.3.1 Distribuição de Lei de Potência**

Devido a interação entre seus elementos, os sistemas complexos podem exibir um comportamento catastrófico, onde uma parte do sistema pode afetar outra, em um efeito dominó. Este mecanismo é o de realimentação positiva, em que quando acontece um evento, há um outro elemento que ajuda a aumentar a magnitude deste evento. Por exemplo, quando uma pessoa joga bem futebol, a sociedade em volta dela nota essa qualidade e fornece a essa pessoa melhores condições para ela praticar o futebol, fazendo com que se torne um melhor jogador de futebol. Este comportamento catastrófico poderia ser explicado pela distribuição de Lei de Potência. (Gutenberg, et al, 1949; Sepkoski, 1993; Tsallis, 1999; Mandelbrot, 1964)

### **1.3.2 Geometria Fractal**

Mandelbrot (Mandelbrot, 1983) inventou a palavra “fractal” para descrever certas estruturas geométricas que em geral tem dimensões que não são números inteiros.

Uma cadeia de montanhas inclui picos que podem variar de centímetros a quilômetros, do mesmo modo uma árvore possui ramos de galhos que variam de centímetros a metros, assim como as rugosidades da faixa litorânea de um país, são exemplos de estruturas fractais encontrados na natureza. Nós temos como caracterizar as propriedades geométricas dos fractais, mas o problema da sua origem dinâmica ainda persiste. De onde vem esta estrutura fractal? A resposta a esse problema ainda é desconhecida e não existe qualquer lei geral da física que possa explicar o surgimento da geometria fractal na natureza.

### 1.3.3 Ruído 1/f.

O ruído 1/f tem sido observado em sistemas tão diversos, como o fluxo de rio Nilo, luz de quasares e tráfego de auto-estradas. O geofísico britânico J.Hurst (Hurst, 1951) que dedicou parte de sua vida estudando o nível de água do rio Nilo, em várias escalas de tempo, (de minutos até anos) notou que esta serie temporal pode ser vista como uma superposição de todas as variações do nível do rio, isto é, como uma superposição de sinais periódicos de todas das frequências. Esta é uma maneira de dizer que há aspectos em todas as escalas de tempo.

Nos fenômenos da variação do preço da bolsa de valores, da temperatura média global, da luz emitida pelo quasares ou do tráfego de uma rodovia, podemos dizer que existem todas as escalas de tempo envolvidas. Quanto maior a frequência da variação dos itens descritos acima, maior a sua intensidade. Em geral podemos dizer que

$$I \propto \frac{1}{f^\alpha} \quad (1.1)$$

onde  $\alpha$  é um expoente com valor entre 0 e 2, e  $f$  é a frequência. Sistemas que são descritos por essa equação possuem o que denominamos ruído do tipo 1/f.

O ruído 1/f é diferente do ruído randômico branco, no qual não há correlações entre o valor do sinal de um momento para o próximo, o padrão do ruído branco não tem flutuações lentas. O ruído branco não tem qualquer relação com intensidade e frequência. Não existe uma compreensão geral sobre a sua origem, constituindo um dos problemas da física atual.

### 1.3.4 Lei de Zipf

Esta lei foi observada por Zipf (Zipf, 1949) na área das ciências humanas. Ele utilizou os dados de populações de várias cidades do mundo e traçou um gráfico do  $\log(\text{ranking})$  pelo  $\log(\text{tamanho})$ , onde observou que o resultado é uma linha reta com inclinação quase 1.

Ele também contou a frequência de uma dada palavra era usada em uma peça de literatura, e novamente verificou que o gráfico do  $\log(\text{frequência})$  pelo  $\log(\text{ranking})$  era uma linha reta, e tentou explicar esse resultado através do princípio do mínimo esforço.

## CAPÍTULO 2

### DISTRIBUIÇÕES ESTATÍSTICAS

Neste capítulo, discutimos as distribuições e técnicas estatísticas existentes para definir os sistemas complexos.

#### 2.1 Distribuição Normal

No século XVIII, a Distribuição Normal foi considerada simplesmente como uma aproximação da distribuição binomial (Moivre, 1733). No século XIX, esta distribuição foi reconhecida devido ao trabalho de Laplace (Laplace, 1781) e Gauss (Gauss, 1816), e hoje em dia é base para a maioria dos trabalhos estatísticos. Esta distribuição é dada por:

$$P(x) = \frac{1}{\sqrt{2\pi s}} \cdot \exp\left[-\frac{1}{2}\left(\frac{x - m}{s}\right)^2\right] \quad (2.1)$$

onde  $P(x)$  é a função densidade de probabilidade e representa a probabilidade de que a variável  $x$  seja encontrada entre os valores  $x$  e  $x+dx$ ,  $\mu$  é o valor médio e  $\sigma$  é o desvio padrão do sistema. Dada uma amostra de  $n$  valores da variável, as estimativas de  $\mu$  e  $\sigma$  são, respectivamente:

$$\mathbf{m} = \frac{1}{n} \sum_{i=1}^n x_i \quad (2.2)$$

$$\mathbf{s} = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \mathbf{m})^2} \quad (2.3)$$

Na forma padrão da Distribuição Normal, se considerarmos a variável  $U = \frac{x - \mathbf{m}}{\mathbf{s}}$  então  $P(U)$  é dado por:

$$P(U) = \frac{1}{\sqrt{2\mathbf{p}}} \cdot \exp\left(-\frac{1}{2}U^2\right) \quad (2.4)$$

que é a função densidade de probabilidade normal com média zero e variância um.

A probabilidade de  $\mu$  assumir um valor qualquer entre zero e  $\theta$  é:

$$\mathbf{f}(\mathbf{q}) = P[0 \leq U \leq \mathbf{q}] = \frac{1}{\sqrt{2\mathbf{p}}} \int_0^{\mathbf{q}} e^{-\frac{1}{2}x^2} dx \quad (2.5)$$

Fisicamente,  $\phi(\theta)$  resulta na probabilidade da variável  $U$  estar entre 0 e  $\theta$ .

Os valores das probabilidades estão tabelados na chamada tabela da distribuição normal.

Na maioria das bibliografias estatísticas (Ayres, 2000; Magalhães, 2001; Vieira, 1980), as tabelas são fornecidas somente para valores positivos da variável. Os valores negativos são encontrados pela simetria da distribuição em torno de zero.

Um dos teoremas mais importantes sobre as distribuições estatísticas é o teorema do limite central (Central Limit Theorem: CLT) (Feller, 1971). Pelo teorema, se  $x_1, x_2, x_3, \dots, x_n$  são valores independentes de uma variável  $x$ , com valor médio e desvio padrão finitos (condição necessária para todos os sistemas naturais), então a distribuição de  $S$ , onde

$$S = \sum_{i=1}^n x_i \quad (2.7)$$

tende à uma Distribuição Normal para grandes valores de  $n$ , isto é, as distribuições estatísticas convergem para a normal quando o número de elementos amostrais  $n$ , tende à infinito.

## 2.2 Distribuição Log-Normal

Em 1879, Galton (Galton, 1879) demonstrou que se  $x_1, x_2, x_3, \dots, x_n$  são variáveis aleatórias positivas independentes, então a distribuição do  $\log(T_n)$  onde  $T_n$  dado por

$$T_n = \prod_{j=1}^n x_j = x_1 x_2 x_3 \dots x_n \quad (2.8)$$

possui uma Distribuição Normal para  $n$  grande, pois

$$\log(T_n) = \log(x_1) + \log(x_2) + \log(x_3) + \dots + \log(x_n) \quad (2.9)$$

e se o teorema do limite central é válido para a variável aleatória independente  $\log(x_i)$ , a distribuição de  $\log(T_n)$  possuirá Distribuição Normal para valores de  $n$  muito grandes. Então, considerando  $x = T_n$ , a distribuição é dada por

$$P(\log x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2}\left(\frac{\log x - \mu}{\sigma}\right)^2\right] \quad (2.10)$$

onde  $\mu$  e  $\sigma$  são a média e o desvio padrão de  $\log(x)$ . Mc-Allister (Mc-Allister, 1879) obteve as expressões para a média e outros parâmetros de distribuição. Kapteyn e Van Uen (Kapteyn e Van Uen, 1916) apresentaram um método gráfico para obter os valores dos parâmetros da distribuição. Gibrat (Gibrat, 1931) mostrou que esta distribuição também podia ser aplicada em economia, no estudo do faturamento de empresas. A distribuição Log-Normal tem sido aplicada em várias áreas, como por exemplo na física, economia, medicina, agricultura, etc (Gibrat, 1931; Bliss, 1934; Gaddum, 1933; Hatch, 1929; Bain, 1964; Cochran, 1938; Williams, 1940).

### 2.3 Distribuição de Lei de Potência e Realimentação Positiva.

Pareto (Pareto, 1896) estudando a distribuição dos rendimentos pessoais propôs que

$$N=A x^{-\alpha} \quad (2.11)$$

onde  $N$  é o número de pessoas com rendimento maior ou igual a  $x$ ,  $A$  e  $\alpha$  são constantes da distribuição. Propôs ainda que esta lei é devido à **Realimentação Positiva**, isto é, pessoas com rendimento pessoal maiores conseguem em suas aplicações financeiras maiores taxas de retorno, conseguindo dessa forma um rendimento pessoal ainda maior.

Pareto sugeriu que a Distribuição de Lei de Potência pudesse ser utilizada na descrição de vários sistemas.

Em geral, a densidade da probabilidade desta distribuição (Mandelbrot, 1963) é dada por

$$P(x) = \frac{C_1 P(x_m)}{C_1 + (x - x_m)^{1+a}} \quad (2.12)$$

onde  $x_m$  é o valor de  $x$  para a probabilidade máxima e  $C_1$  é uma constante.

Esta distribuição foi observada em vários sistemas complexos (Tsallis, 1999; Bak, 1997; Schroeder, 1991), como por exemplo o estudo da energia liberada por um terremoto, de um tornado, da variação de preços de algodão, da extinção de espécies, da biodiversidade da Floresta Amazônica. Assim, quando o valor de  $x$  é grande em relação a  $x_m$ ,  $P(x)$  pode ser aproximada por

$$P(x) = \text{const} \frac{1}{x^{1+a}} \quad (2.13)$$

## 2.4 Distribuições de Lévy.

A função de distribuição simétrica de Levy (Levy, 1937; Frish, Shlesinger and Zaslavsky, 1994) é dada por:

$$L_a(x, \Delta t) = \frac{1}{\Gamma(\frac{1}{\alpha})} \int_0^{\infty} \exp(-g\Delta t q^\alpha) \cos(qx) dq \quad (2.14)$$

onde  $\alpha$  ( $0 < \alpha \leq 2$ ),  $\gamma$  é o fator de escala ( $\gamma > 0$ ) e  $\Delta t$  é o tempo entre observações sucessivas. Tal distribuição descreve por exemplo a variação de valor da ação de uma empresa ou conjunto de empresas na bolsa de valores ao longo do tempo.

Para grandes valores de  $x$ , a distribuição de Lévy se aproxima de uma Lei de Potência (Bergstrom, 1952).

## 2.5 Distribuição Gradualmente Truncada

A Distribuição de Lei de Potência possui desvio padrão infinito, mas geralmente os sistemas físicos reais possuem desvio padrão finito. Por isso ao aplicarmos a distribuição de Lévy ou Pareto, ou ainda Log-Normal, aos sistemas físicos reais, precisamos truncar as distribuições após um determinado valor, a fim de evitar um desvio padrão infinito. Os sistemas complexos possuem várias interações e grande número de componentes interagindo de maneira diferente, portanto esperamos que o truncamento desta **realimentação positiva** seja gradual após um passo crítico. Baseado neste aspecto, Gupta e Campanha (Gupta and Campanha, 1999) desenvolveram uma nova distribuição conhecida como Distribuição de Lévy Gradualmente Truncada, introduzindo um fator  $f(x)$  para essa capacidade física do componente ou sistema, dada por:

$$f(x) = \begin{cases} 1 & \text{se } |x| \leq x_c \\ \exp\left[-\left(\frac{x-x_c}{k}\right)^B\right] & \text{se } |x| > x_c \end{cases} \quad (2.15)$$

onde  $B = (2 - \alpha)$ ,  $x_c$  é o ponto crítico onde começa o truncamento gradual devido ao limite físico do sistema e  $k$  uma constante de truncamento gradual. Para valores de  $k$  menores, o truncamento da realimentação positiva será maior.

Esta distribuição tem desvio padrão finito, e no limite obtem-se uma Distribuição Normal como exigido pelo teorema do limite central (Gupta and Campanha, 2000). Esta distribuição também obedece a Lei de Potência em sua parte central e decaimento exponencial nos valores extremos de  $x$ . Podemos considerar esta distribuição

equivalente a distribuição de Tsallis (Tsallis, 1999), obtidas pela termodinâmica para sistemas complexos (Gupta and Campanha, 2002). A Distribuição Gradualmente Truncada torna mais simples a extração de informações úteis que descrevem o sistema real. Gupta e Campanha também desenvolveram o truncamento para a Distribuição Log-Normal (Gupta and Campanha, 2001).

Usando o fator de truncamento, temos as seguintes distribuições:

A) Distribuição de Lei de Potência Gradualmente Truncada:

$$P(x) = \frac{C_1 P(x_m)}{C_1 + (|x - x_m|)^{1+a}} * f(x) \quad (2.16)$$

B) Distribuição de Lévy Gradualmente Truncada:

$$P(x_\alpha) = C \cdot L_\alpha(x, \Delta t) \cdot f(x) \quad (2.17)$$

onde C é a constante de normalização de tal maneira que:

$$\int_{-a}^a P(x) dx = 1 \quad (2.18)$$

C) Distribuição Log-Normal Gradualmente Truncada

$$P(\ln x) = const. \left[ -\frac{1}{2} \left( \frac{\ln x - \mathbf{m}}{\mathbf{s}} \right)^2 \right] \cdot f(x) \quad (2.19)$$

onde  $const$  é obtido através da condição de normalização, mas é bem próximo de  $\frac{1}{\sqrt{2ps}}$ , e  $\sigma$  e  $\mu$  são os valores do desvio padrão e do valor médio de  $\ln(x)$ .  $f(x)$  é dado por:

$$f(x) = \begin{cases} 1 & \text{se } |x| \leq x_c \\ \exp\left[-\left(\frac{x-x_c}{k}\right)^B\right] & \text{se } |x| > x_c \end{cases} \quad (2.20)$$

## 2.6 Gráficos do Tipo Zipf (Zipf Plot).

Geralmente, temos muito poucos dados nos extremos das distribuições estatísticas dos sistemas reais que estudamos, daí, para grandes valores de  $x$ , qualquer distribuição estatística que se aplique pode parecer correta. Usamos então a técnica denominada de gráfico do tipo Zipf (Zipf, 1949). Esta técnica é importante por exemplo no faturamento das primeiras 1000 firmas ou citações, no caso de cientistas mais citados (Gupta and Campanha, 2001).

Suponhamos  $x_1, x_2 \dots x_N$  como o conjunto de  $N$  observações de um variável  $x$  as quais são ordenadas em ordem decrescente, isto é, o índice  $i$  nos dá o ranking de  $x_i$ , o gráfico do tipo Zipf é a curva  $\log(i)$  (abscissa) versus  $\log(x_i)$  (ordenada).

Então  $x_i$  pode ser estimado por:

$$\int_{x_i}^a N(x) d(x) = N \int_{x_i}^a P(x) d(x) = i \quad (2.21)$$

Isso especifica que temos  $i$  números dentro do conjunto de  $N$  observações com valor igual ou maior que  $x_i$ . Pela dependência de  $x_i$  sobre  $i$  no gráfico de Zipf, nós podemos verificar se a distribuição está de acordo com a distribuição assumida para  $P(x)$ .

## CAPÍTULO 3

### **VESTIBULAR, UM SISTEMA COMPLEXO.**

Neste capítulo, descreveremos o processo do Vestibular da Universidade Estadual Paulista “Julio de Mesquita Filho” (UNESP) e abordaremos o Vestibular como um sistema complexo, encontrando distribuições estatísticas para vários fatores estudados.

#### **3.1 O Vestibular da Unesp.**

Desde a sua criação, a UNESP vem se consolidando como uma das principais universidades brasileiras. Durante esse tempo, um dos principais destaques é o processo de seleção dos alunos, considerado um dos melhores do País. No último exame realizado (2002), 85.762 candidatos concorrem a 5.695 vagas, onde esses candidatos provem de várias classes econômicas, culturais e com variadas tradições familiares.

Um dos principais aspectos desse exame é sua realização em uma única fase, em três dias, geralmente realizado no mês de dezembro pela fundação VUNESP.

Os candidatos do Vestibular da UNESP são divididos em três áreas: Ciências Exatas, Ciências Biológicas e Ciências Humanas. Em cada área o candidato é submetido a três avaliações: Conhecimentos Gerais, Conhecimentos Específicos e Língua Portuguesa.

A Avaliação de Conhecimento Gerais, consta de 84 questões sob forma de teste de múltipla escolha, igualmente distribuídas pelas disciplinas de Matemática, Biologia, Geografia, Física, Química e Língua Estrangeira (Inglês ou Francês, conforme opção).

A Avaliação de Conhecimentos Específicos, consta de 25 questões, sob forma discursiva. As disciplinas que compõem esta avaliação variam conforme a área de opção, que podem ser:

- Área de Ciências Biológicas: Biologia (10 questões); Química (6 questões); Física (5 questões) e Matemática (4 questões).
- Área de Ciências Exatas: Matemática (10 questões); Física (9 questões) e Química (6 questões).
- Área de Ciências Humanas: História (10 questões); Geografia (9 questões) e Língua Portuguesa (6 questões).

A Avaliação de Língua Portuguesa, consta de 10 questões e uma redação, sob a forma discursiva, idêntica para todos os candidatos.

O vestibular é um processo muito competitivo, o qual reúne candidatos de todo País, de classes sociais diversas, com renda familiar e nível de conhecimento variado, reduzindo assim a capacidade de prever quais candidatos serão aprovados, sendo, por conseguinte, um processo extremamente difícil.

Recentemente Gupta, Campanha e Prado (Gupta, Campanha and Prado, 2000) estudaram a distribuição estatística das notas dos alunos que prestaram o vestibular da Unesp e observaram a presença da lei de potência para os alunos da área de ciências exatas e biológicas; já para os alunos da área de ciências humanas a distribuição que melhor se aproximou foi a Distribuição Normal. Neste estudo, foram utilizados métodos da física para sistemas complexos, o que tornou possível analisar quantitativamente alguns aspectos do sistema educacional. Lembramos que este estudo é o primeiro com este tipo de análise do sistema educacional utilizando essa técnica.

No presente trabalho decidimos não considerar as notas da avaliação de conhecimentos gerais, devido a esta prova ter sua resposta em forma de múltipla escolha, o que significa que a probabilidade do candidato acertar uma resposta é de 20%, considerando que não sabe a resposta correta. Isto cria um grande ruído na distribuição, acarretando possivelmente erro em um julgamento, devido à não possuir uma nota negativa para resposta errada. Decidimos também não trabalhar com a Avaliação de Língua Portuguesa, devido ao estudo anterior (Gupta, Campanha and Prado, 2000) ter apresentado uma Distribuição Normal.

Escolhemos então trabalhar com as notas da avaliação de conhecimentos específicos devido ao fato destas terem apresentado realimentação positiva nesta avaliação (Gupta, Campanha and Prado, 2000).

Devido a esta escolha, neste capítulo abordamos as distribuições das notas dos candidatos da Avaliação de Conhecimentos Específicos na área de Ciências Exatas, Ciências Biológicas e Ciências Humanas, divididos da seguinte forma:

1. alunos que cursaram o ensino médio no período diurno ou noturno;
2. alunos que cursaram o ensino médio em escola particular ou pública;
3. alunos com renda familiar superior a dez salários mínimos ou inferior a dez salários mínimos.

Para realizar as análises, utilizamos o software Mathematica® 4.1, com a biblioteca **Statistics NonlinearFit**, onde neste acionamos a função **NonlinearRegress** que trabalha com o método **Levenberg Marquardt** de ajuste de curvas não lineares (Press, 1992). Esse método é um dos mais utilizados para estimativas de parâmetros na simulação de sistemas não lineares, o qual realiza sucessivas interações dos parâmetros calculados com os dados experimentais.

### **3.2 Distribuição Estatística com Ênfase no Período.**

Neste tópico, abordamos apenas as notas dos candidatos da avaliação de Conhecimentos Específicos de cada área, que foram divididos em dois grupos: alunos que cursaram o ensino médio no período diurno ou noturno.

#### **3.2.1 Área de Ciências Exatas**

Como mostram as Figuras 1, 2 e 3 comparamos as notas da avaliação de conhecimentos específicos da área de ciências exatas do período diurno e noturno dos anos de 1998, 1999 e 2000, onde obtivemos a Distribuição de Lei de Potência Gradualmente Truncada, utilizando os parâmetros  $1+\alpha$ ,  $x_c$ ,  $x_m$ , e  $k$  como descrito na Tabela 1, sendo que na curva teórica  $x_c$  e  $x_m$  são dados empíricos e  $1+\alpha$  e  $k$  são os melhores valores que se ajustam às curvas.

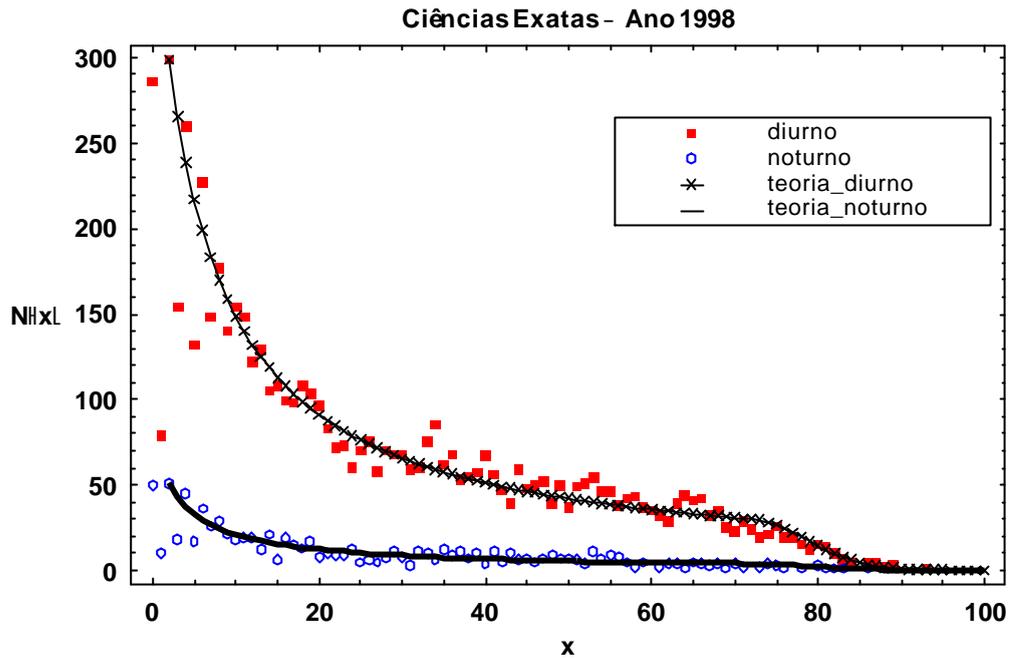


Figura 1: Números de candidatos que obtiveram nota  $x$  [ $N(x)$  versus notas obtidas ( $x$ )] para a avaliação de conhecimentos específicos da área de ciências exatas do ano de 1998 para alunos do período diurno (■) e noturno (○) no ensino médio.

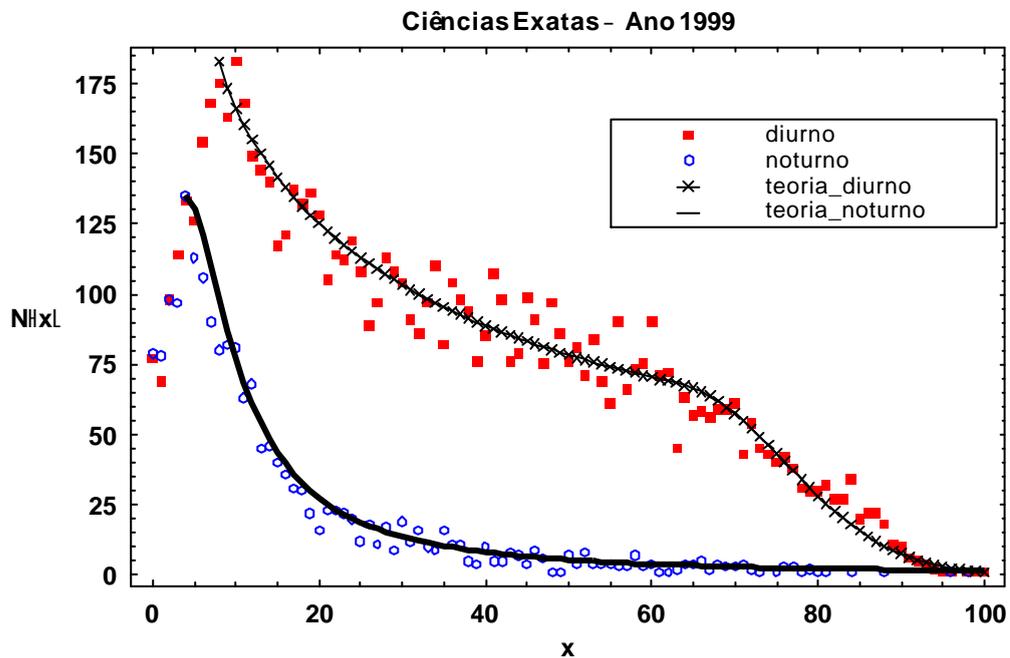


Figura 2: Números de candidatos que obtiveram nota  $x$  [ $N(x)$  versus notas obtidas ( $x$ )] para a avaliação de conhecimentos específicos da área de ciências exatas do ano de 1999 para alunos do período diurno (■) e noturno (○) no ensino médio.

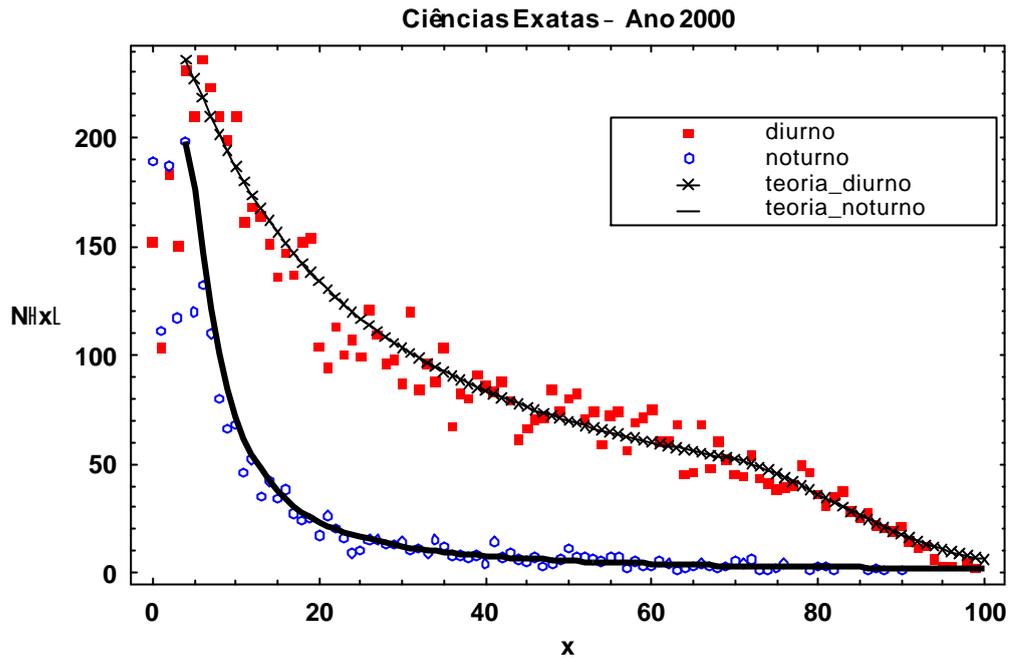


Figura 3: Números de candidatos que obtiveram nota  $x$  [ $N(x)$  versus notas obtidas ( $x$ )] para a avaliação de conhecimentos específicos da área de ciências exatas do ano de 2000 para alunos do período diurno (■) e noturno (○) no ensino médio.

Tabela 1: Tabela de parâmetros para o melhor ajuste com os dados empíricos na Distribuição de Lei de Potência Gradualmente Truncada, para os dados dos alunos que cursaram o ensino médio no período diurno e noturno para os anos de 1998, 1999 e 2000 na área de ciências exatas.

Área de Ciências Exatas						
Ano	Ano de 1998		Ano de 1999		Ano de 2000	
Período	Diurno	Noturno	Diurno	Noturno	Diurno	Noturno
Nº de alunos	6.018	819	7.710	1.863	8.262	2.148
$1+a$	1,00	0,98	0,85	1,69	1,07	1,47
$x_m$	2	2	8	4	4	4
$x_c$	73	74	64	-	70	-
$k$	100	200	600	-	400	-
$C_1$	7,96	5,40	17,99	28,05	26,12	8,02
$-x$	25,59	22,98	33,94	14,68	31,59	13,16
$s$	21,95	20,67	23,21	15,83	24,39	16,36

Podemos verificar na Tabela 1, que para o período diurno do ano 1998,  $1+\alpha$  é aproximadamente 1.0 e para o período noturno é 0.9 caracterizando assim a

realimentação positiva, em que o ponto crítico das notas dos candidatos do período diurno é 73 e para o noturno é 74, causando assim um truncamento até o seu limite físico que é a nota 100.

Observamos também que para o período diurno do ano 1999,  $1+\alpha$  é aproximadamente 0.8 e para o período noturno é 1.7 caracterizando assim a realimentação positiva somente para o período diurno, em que o ponto crítico das notas dos candidatos é 64, caracterizando assim um truncamento até o seu limite físico que é a nota 100. Já no período noturno, notamos que após a nota 20 ha um decaimento para perto de 0, descaracterizando assim a Distribuição de Lei de Potência Gradualmente Truncada.

Finalmente para o período diurno do ano 2000,  $1+\alpha$  é aproximadamente 1.1 e para o período noturno é 1.5 caracterizando assim a realimentação positiva somente para o período diurno, em que o ponto crítico das notas dos candidatos é 70, caracterizando assim um truncamento até o seu limite físico que é a nota 100. Já no período noturno, notamos que após a nota 20 a um decaimento para perto de 0, descaracterizando assim a Distribuição de Lei de Potência Gradualmente Truncada.

Notamos que o ajustamento da Distribuição de Lei de Potência Gradualmente Truncada com os dados empíricos é muito boa para o período diurno dos anos de 1998, 1999 e 2000, já para o período noturno a Distribuição de Lei de Potência Gradualmente Truncada descreve os dados somente no ano de 1998, e para os anos de 1999 e 2000 a que melhor descreve os dados é a Distribuição de Lei de Potência.

### **3.2.2 Área de Ciências Biológicas**

Como mostram as Figuras 4, 5 e 6 comparamos as notas da avaliação de conhecimentos específicos da área de ciências biológicas do período diurno e noturno dos anos de 1998, 1999 e 2000, onde obtivemos a Distribuição de Lei de Potência Gradualmente Truncada, utilizando os parâmetros  $1+\alpha$ ,  $x_c$ ,  $x_m$ , e  $k$  como descrito na

Tabela 2, sendo que na curva teórica  $x_c$  e  $x_m$  são dados empíricos e  $1+\alpha$  e  $k$  são os melhores valores que se ajustam às curvas.

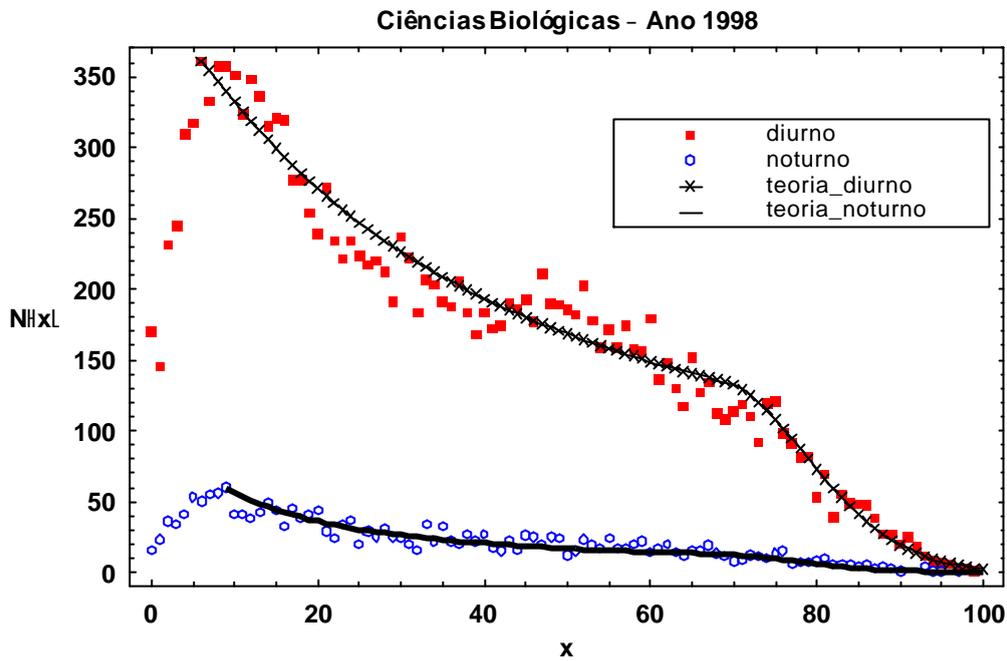


Figura 4: Números de candidatos que obtiveram nota  $x$  [ $N(x)$  versus notas obtidas ( $x$ )] para a avaliação de conhecimentos específicos da área de ciências biológicas do ano de 1998 para alunos do período diurno (■) e noturno (●) no ensino médio.

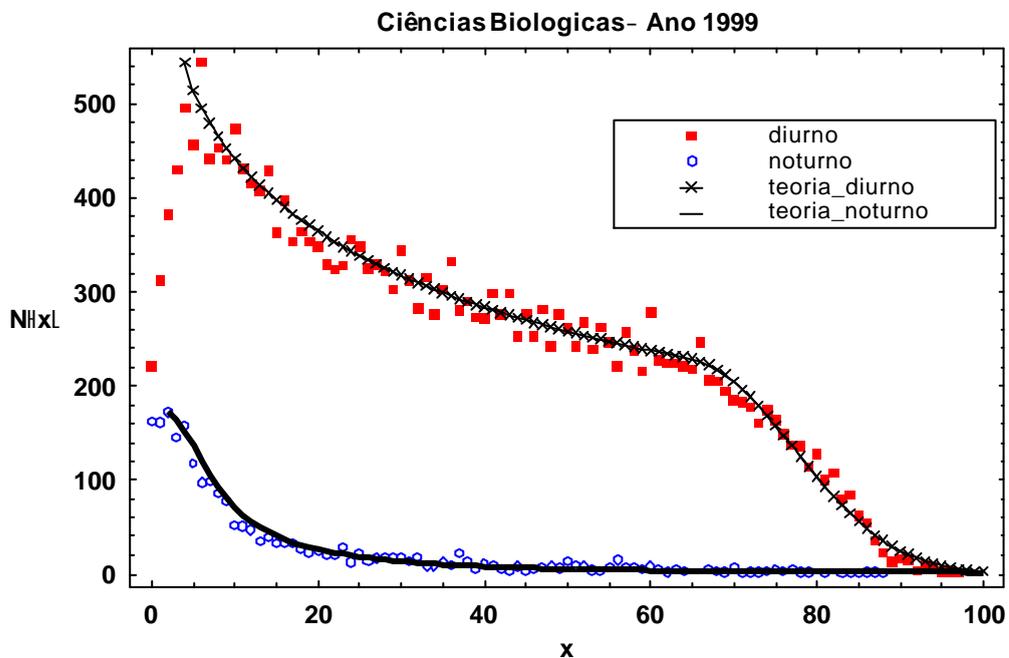


Figura 5: Números de candidatos que obtiveram nota  $x$  [ $N(x)$  versus notas obtidas ( $x$ )] para a avaliação de conhecimentos específicos da área de ciências biológicas do ano de 1999 para alunos do período diurno (■) e noturno (●) no ensino médio.

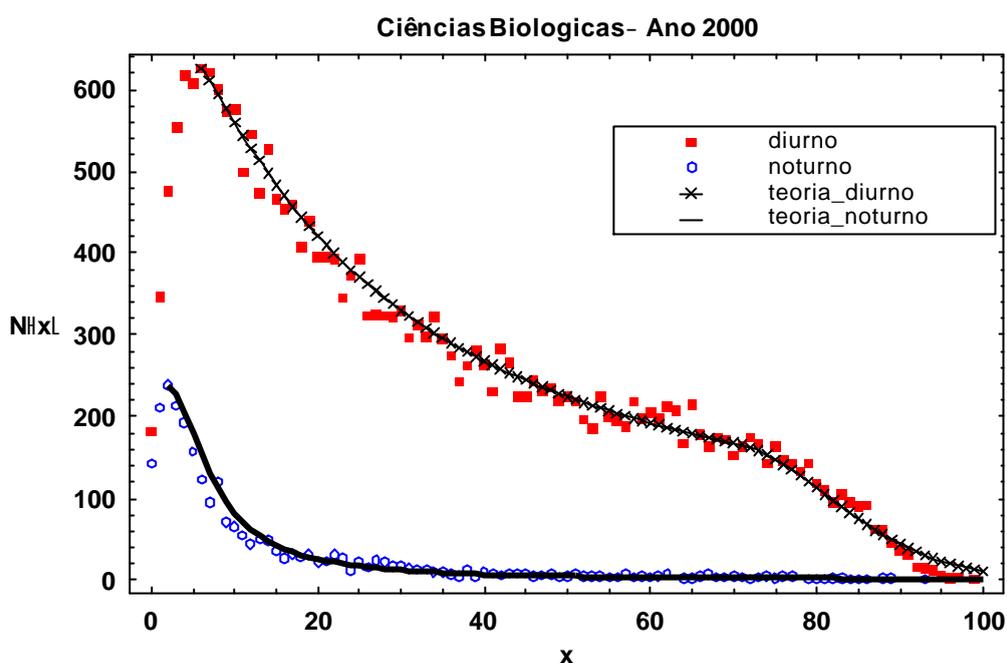


Figura 6: Números de candidatos que obtiveram nota  $x$  [ $N(x)$  versus notas obtidas ( $x$ )] para a avaliação de conhecimentos específicos da área de ciências biológicas do ano de 2000 para alunos do período diurno (■) e noturno (○) no ensino médio.

Tabela 2: Tabela de parâmetros para o melhor ajuste com os dados empíricos na Distribuição de Lei de Potência Gradualmente Truncada, para os dados dos alunos que cursaram o ensino médio no período diurno e noturno para os anos de 1998, 1999 e 2000 na área de ciências biológicas.

Área de Ciências Biológicas						
Ano	Ano de 1998		Ano de 1999		Ano de 2000	
Período	Diurno	Noturno	Diurno	Noturno	Diurno	Noturno
Nº de alunos	16.704	2.088	24.179	2.191	25.416	2.487
$1+\alpha$	1,07	1,01	0,76	1,69	1,13	1,80
$x_m$	6	9	4	2	6	2
$x_C$	70	70	65	-	72	-
$K$	200	200	700	-	220	-
$C_1$	51,72	16,98	17,06	23,98	41,28	22,38
$\bar{x}$	34,03	31,61	34,63	14,39	32,27	12,57
$s$	23,43	23,01	23,43	17,23	23,96	15,97

Podemos verificar na Tabela 2, que para o período diurno,  $1+\alpha$  do ano 1998 é aproximadamente 1.0 e para o período noturno é 1.0 caracterizando assim a

realimentação positiva, em que o ponto crítico das notas dos candidatos do período diurno é 70 e para o noturno é 70, causando assim um truncamento até o seu limite físico que é a nota 100.

Observamos também que para o período diurno,  $1+\alpha$  do ano de 1999 é aproximadamente 0.7 e para o período noturno é 1.7 caracterizando assim a realimentação positiva somente para o período diurno, em que o ponto crítico das notas dos candidatos é 65 caracterizando assim um truncamento até o seu limite físico que é a nota 100. Já no período noturno, notamos que após a nota 20 a um decaimento para perto de 0, descaracterizando assim a Distribuição de Lei de Potência Gradualmente Truncada.

Finalmente para o período diurno do ano 2000,  $1+\alpha$  é aproximadamente 1.1 e para o período noturno é 1.8 caracterizando assim a realimentação positiva somente para o período diurno, em que o ponto crítico das notas dos candidatos é 72, caracterizando assim um truncamento até o seu limite físico que é a nota 100. Já no período noturno, notamos que após a nota 20 a um decaimento para perto de 0, descaracterizando assim a Distribuição de Lei de Potência Gradualmente Truncada.

Notamos que o ajustamento da Distribuição de Lei de Potência Gradualmente Truncada com os dados empíricos é muito boa para o período diurno dos anos de 1998, 1999 e 2000, já para o período noturno a Distribuição de Lei de Potência Gradualmente Truncada descreve os dados somente no ano de 1998 e para os anos de 1999 e 2000 a que melhor descreve os dados é a Distribuição de Lei de Potência.

### **3.2.3 Área de Ciências Humanas**

Como mostram as Figuras 7, 8 e 9 comparamos as notas da avaliação de conhecimentos específicos da área de ciências humanas do período diurno e noturno dos anos de 1998, 1999 e 2000, onde obtivemos a Distribuição Normal, utilizando os

parâmetros  $\bar{x}$  e  $s$  como descrito na Tabela 3, sendo que na curva teórica são os melhores valores que se ajustam aos dados.

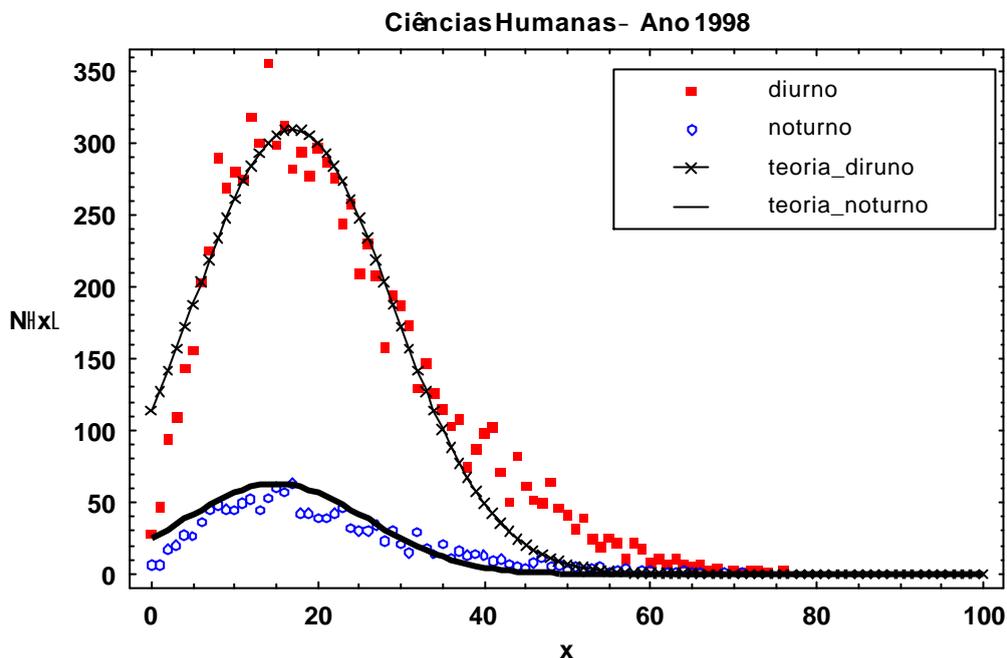


Figura 7: Números de candidatos que obtiveram nota  $x$  [ $N(x)$  versus notas obtidas ( $x$ )] para a avaliação de conhecimentos específicos da área de ciências humanas do ano de 1998 para alunos do período diurno (■) e noturno (○) no ensino médio.

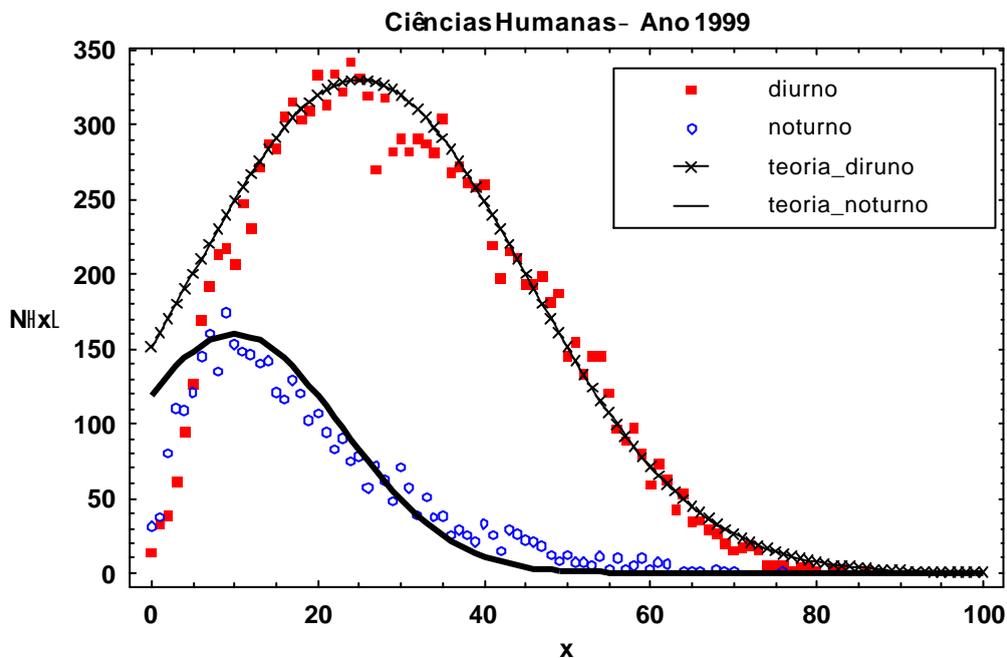


Figura 8: Números de candidatos que obtiveram nota  $x$  [ $N(x)$  versus notas obtidas ( $x$ )] para a avaliação de conhecimentos específicos da área de ciências humanas do ano de 1999 para alunos do período diurno (■) e noturno (○) no ensino médio.

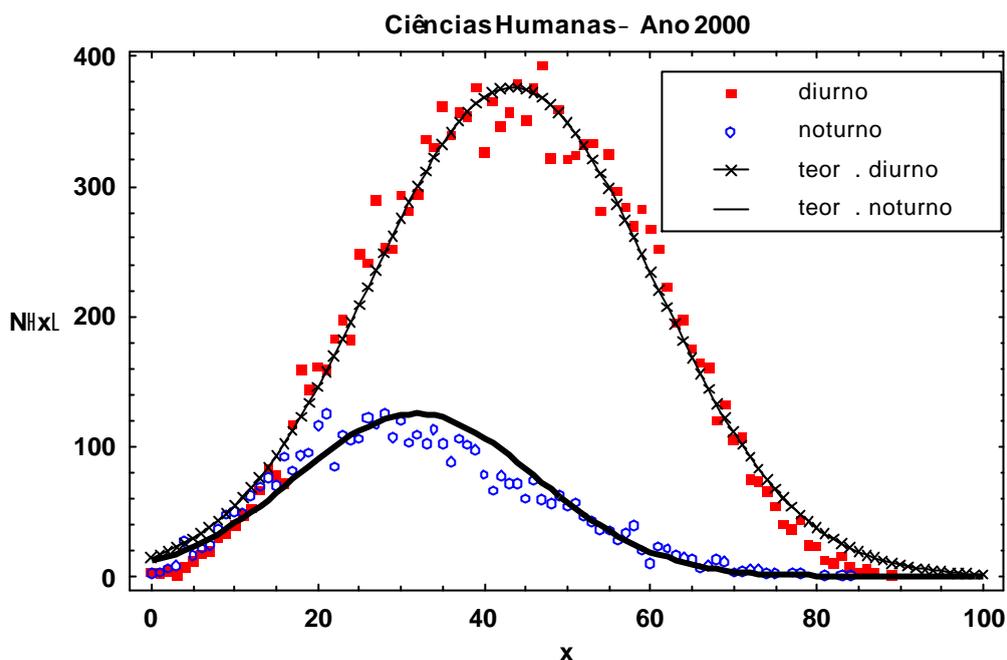


Figura 9: Números de candidatos que obtiveram nota  $x$  [ $N(x)$  versus notas obtidas ( $x$ )] para a avaliação de conhecimentos específicos da área de ciências humanas do ano de 2000 para alunos do período diurno (■) e noturno (●) no ensino médio.

Tabela 3: Tabela de parâmetros para o melhor ajuste com os dados empíricos na Distribuição Normal, para os dados dos alunos que cursaram o ensino no período diurno e noturno para os anos de 1998, 1999 e 2000 na área de ciências humanas.

Área de Ciências Humanas						
Ano	Ano de 1998		Ano de 1999		Ano de 2000	
Período	Diurno	Noturno	Diurno	Noturno	Diurno	Noturno
Nº de alunos	9.161	1.417	13.652	3.881	15.669	4.313
$m$	17	15	25	9	43,44	31,79
$-x$	21,74	20,37	30,14	18,24	43,44	31,79
$s$	12	10	20	10	17	14,55
$s$	12,88	12,41	15,57	12,60	15,48	14,55

Notamos que o ajustamento da Distribuição de Normal com os dados empíricos é muito boa para o período diurno e noturno dos anos de 1998, 1999 e 2000.

### 3.3 Distribuição Estatística com Ênfase no Tipo de Instituição.

Neste tópico, abordamos apenas as notas dos candidatos da avaliação de conhecimentos específicos de cada área, que foram divididos em dois grupos: alunos que cursaram o ensino médio em escola pública ou em escola particular.

#### 3.3.1 Área de Ciências Exatas

Como mostram as Figuras 10, 11 e 12 comparamos as notas da avaliação de conhecimentos específicos da área de ciências exatas da escola particular e pública dos anos de 1998, 1999 e 2000, onde obtivemos a Distribuição de Lei de Potência Gradualmente Truncada, utilizando os parâmetros  $1+\alpha$ ,  $x_c$ ,  $x_m$ , e  $k$  como descrito na Tabela 4, sendo que na curva teórica  $x_c$  e  $x_m$  são dados empíricos e  $1+\alpha$  e  $k$  são os melhores valores que se ajustam às curvas.

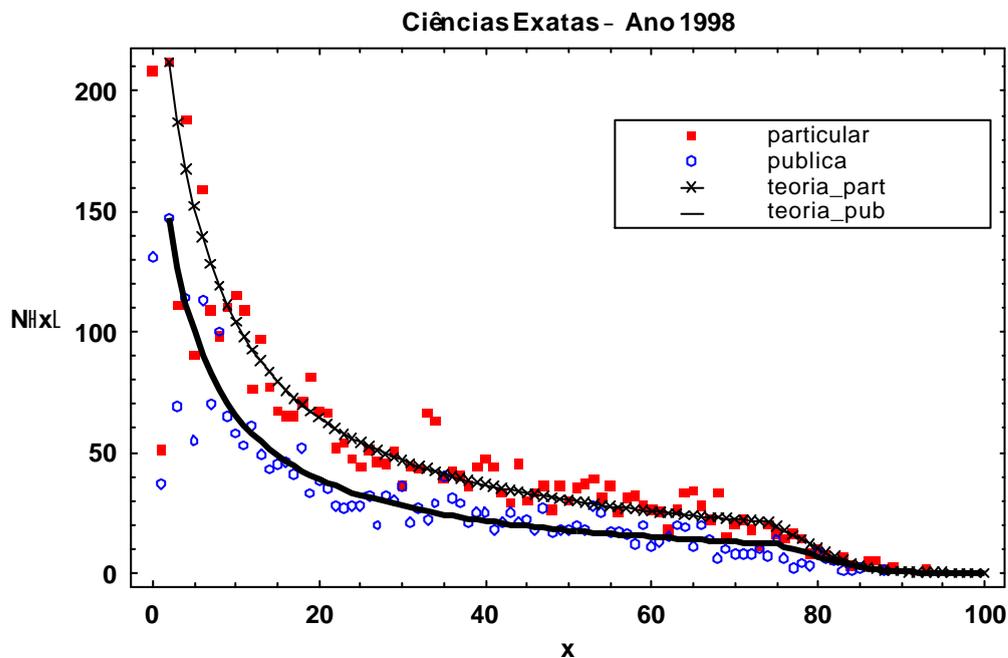


Figura 10: Números de candidatos que obtiveram nota  $x$  [ $N(x)$  versus notas obtidas ( $x$ )] para a avaliação de conhecimentos específicos da área de ciências exatas do ano de 1998 para alunos da escola particular (■) e pública (○) no ensino médio.

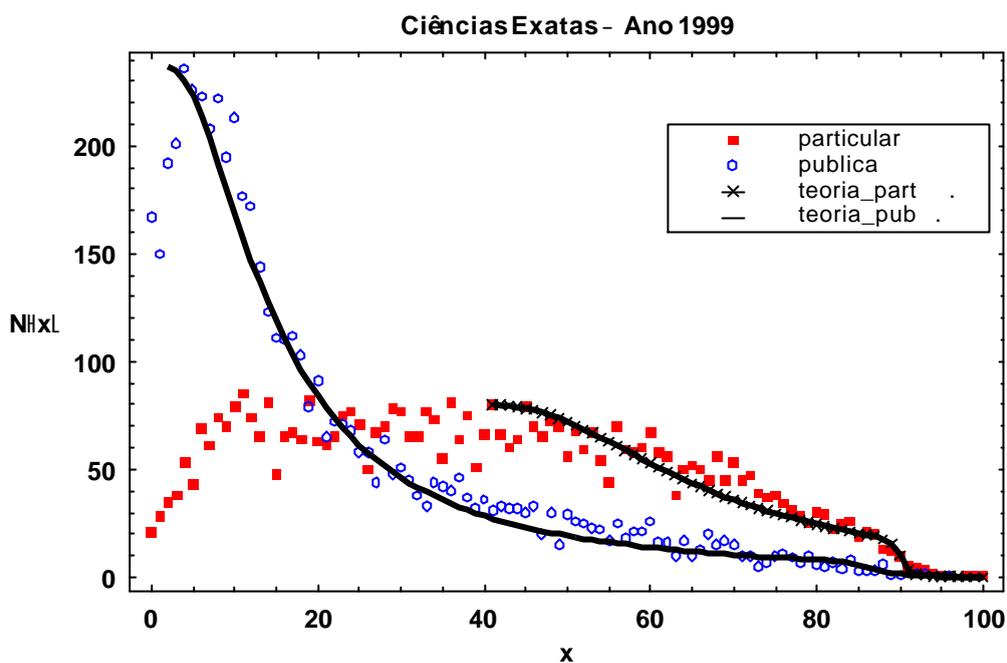


Figura 11: Números de candidatos que obtiveram nota  $x$  [ $N(x)$  versus notas obtidas ( $x$ )] para a avaliação de conhecimentos específicos da área de ciências exatas do ano de 1999 para alunos da escola particular (■) e pública (○) no ensino médio.

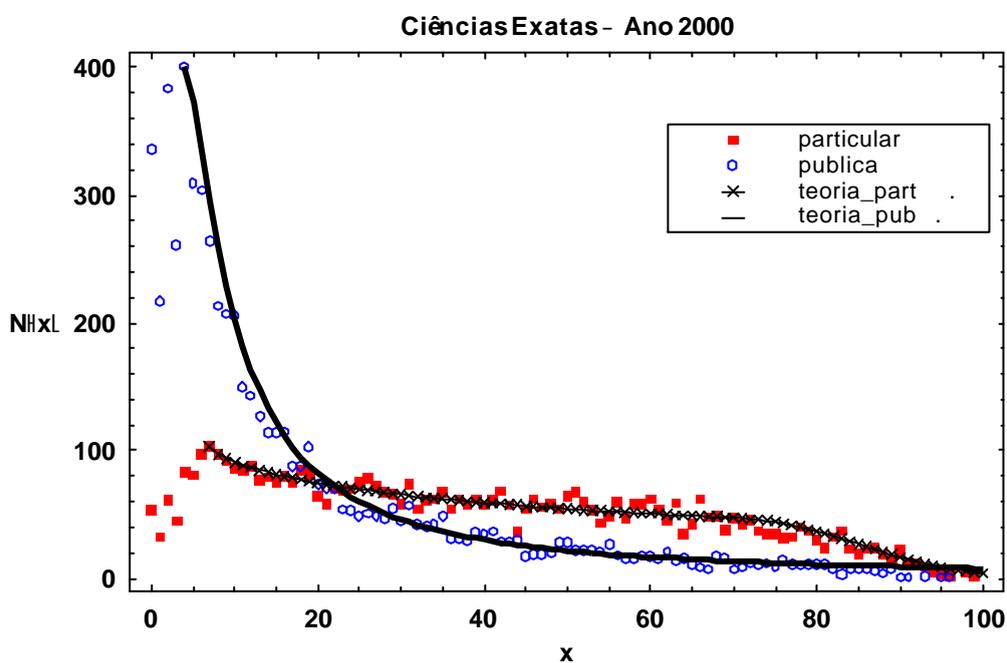


Figura 12: Números de candidatos que obtiveram nota  $x$  [ $N(x)$  versus notas obtidas ( $x$ )] para a avaliação de conhecimentos específicos da área de ciências exatas do ano de 2000 para alunos da escola particular (■) e pública (○) no ensino médio.

Tabela 4: Tabela de parâmetros para o melhor ajuste com os dados empíricos na Distribuição de Lei de Potência Gradualmente Truncada para os dados dos alunos que cursaram o ensino médio na escola particular e pública para os anos de 1998, 1999 e 2000 na área de ciências exatas.

<b>Área de Ciências Exatas</b>						
<b>Ano</b>	<b>Ano de 1998</b>		<b>Ano de 1999</b>		<b>Ano de 2000</b>	
<b>Escola</b>	<b>Particular</b>	<b>Pública</b>	<b>Particular</b>	<b>Pública</b>	<b>Particular</b>	<b>Pública</b>
<b>Nº de alunos</b>	<b>4.281</b>	<b>2.598</b>	<b>4.989</b>	<b>5.236</b>	<b>5.170</b>	<b>5.922</b>
<b>1+<math>\alpha</math></b>	<b>0,98</b>	<b>0,98</b>	<b>2,47</b>	<b>1,86</b>	<b>0,71</b>	<b>1,41</b>
<b><math>x_m</math></b>	<b>2</b>	<b>2</b>	<b>41</b>	<b>4</b>	<b>7</b>	<b>4</b>
<b><math>x_C</math></b>	<b>70</b>	<b>72</b>	<b>90</b>	<b>81</b>	<b>70</b>	<b>-</b>
<b>k</b>	<b>100</b>	<b>100</b>	<b>1,2</b>	<b>6,72</b>	<b>1000</b>	<b>-</b>
<b><math>C_1</math></b>	<b>7,49</b>	<b>6,26</b>	<b>6289,26</b>	<b>120,31</b>	<b>16,85</b>	<b>13,08</b>
<b><math>\bar{x}</math></b>	<b>25,93</b>	<b>24,41</b>	<b>39,72</b>	<b>19,89</b>	<b>37,95</b>	<b>17,50</b>
<b>s</b>	<b>22,17</b>	<b>21,43</b>	<b>23,03</b>	<b>19,07</b>	<b>24,72</b>	<b>19,22</b>

Podemos verificar na Tabela 4, que para os alunos que cursaram em escola particular no ano 1998,  $1+\alpha$  é aproximadamente 0.9 e para os alunos que cursaram a escola pública é 0.9 caracterizando assim a realimentação positiva, em que o ponto crítico das notas dos candidatos da escola particular é 70 e para a escola pública é 72, causando assim um truncamento até o seu limite físico que é a nota 100.

Observamos também que para os alunos que cursaram em escola particular no ano 1999,  $1+\alpha$  é aproximadamente 2.4 e para os alunos que cursaram a escola pública é 1.8 caracterizando assim a realimentação positiva, em que o ponto crítico das notas dos candidatos da escola particular é 90 e para a escola pública é 81, causando assim um truncamento até o seu limite físico que é a nota 100, mostrando a Distribuição de Lei de Potência Gradualmente Truncada

Finalmente para os alunos que cursaram em escola particular no ano 2000,  $1+\alpha$  é aproximadamente 0.7 e para os alunos que cursaram a escola pública é 1.4 caracterizando assim a realimentação positiva somente para os alunos que cursaram o ensino médio em escola particular, em que o ponto crítico é 70 causando assim um

truncamento até o seu limite físico que é a nota 100. Já para os alunos que cursaram em escola pública, notamos que após a nota 20 a um decaimento para perto de 0, descaracterizando assim a Distribuição de Lei de Potência Gradualmente Truncada.

Notamos que o ajustamento da Distribuição de Lei de Potência Gradualmente Truncada com os dados empíricos é muito boa para a escola particular nos anos de 1998 e 2000, já para o ano de 1999 a que melhor descreve os dados é a Distribuição de Lei de Potência. No caso da escola pública a Distribuição de Lei de Potência Gradualmente Truncada descreve os dados dos anos de 1998 e 1999, já para o ano 2000 a que melhor descreve os dados é a Distribuição de Lei de Potência.

### **3.3.2 Área de Ciências Biológicas**

Como mostram as Figuras 13, 14 e 15 comparamos as notas da avaliação de conhecimentos específicos da área de ciências biológicas dos alunos da escola particular e pública dos anos de 1998, 1999 e 2000, onde obtivemos a Distribuição de Lei de Potência Gradualmente Truncada, utilizando os parâmetros  $1+\alpha$ ,  $x_C$ ,  $x_m$ , e  $k$  como descrito na Tabela 5, sendo que na curva teórica  $x_C$  e  $x_m$  são dados empíricos e  $1+\alpha$  e  $k$  são os melhores valores que se ajustam às curvas.

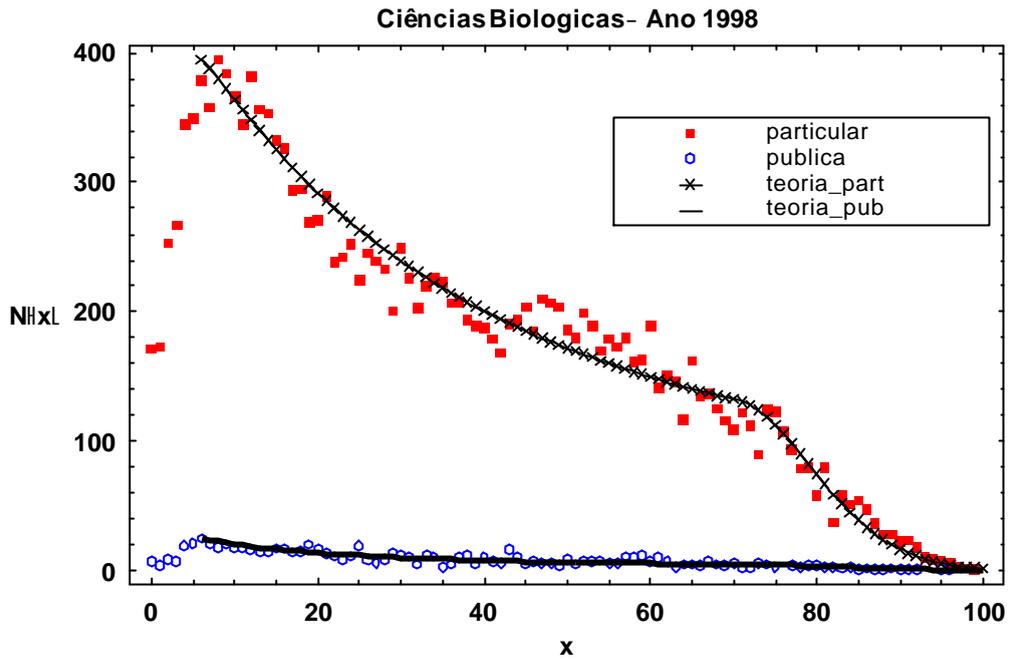


Figura 13: Números de candidatos que obtiveram nota  $x$  [ $N(x)$  versus notas obtidas ( $x$ )] para a avaliação de conhecimentos específicos da área de ciências biológicas do ano de 1998 para alunos da escola particular (■) e pública (○) no ensino médio.

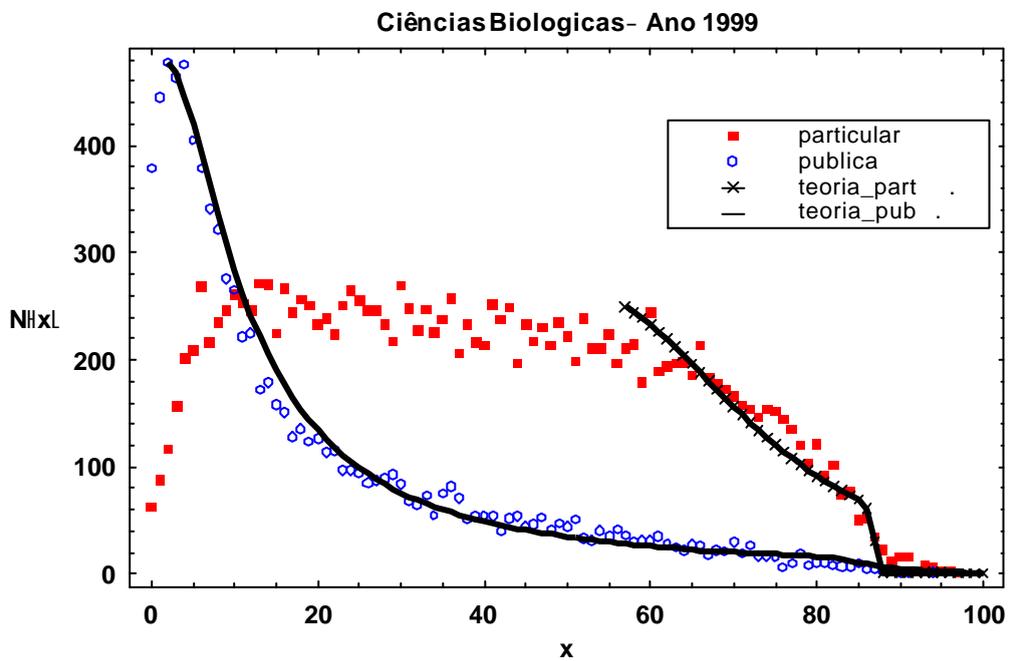


Figura 14: Números de candidatos que obtiveram nota  $x$  [ $N(x)$  versus notas obtidas ( $x$ )] para a avaliação de conhecimentos específicos da área de ciências biológicas do ano de 1999 para alunos da escola particular (■) e pública (○) no ensino médio.

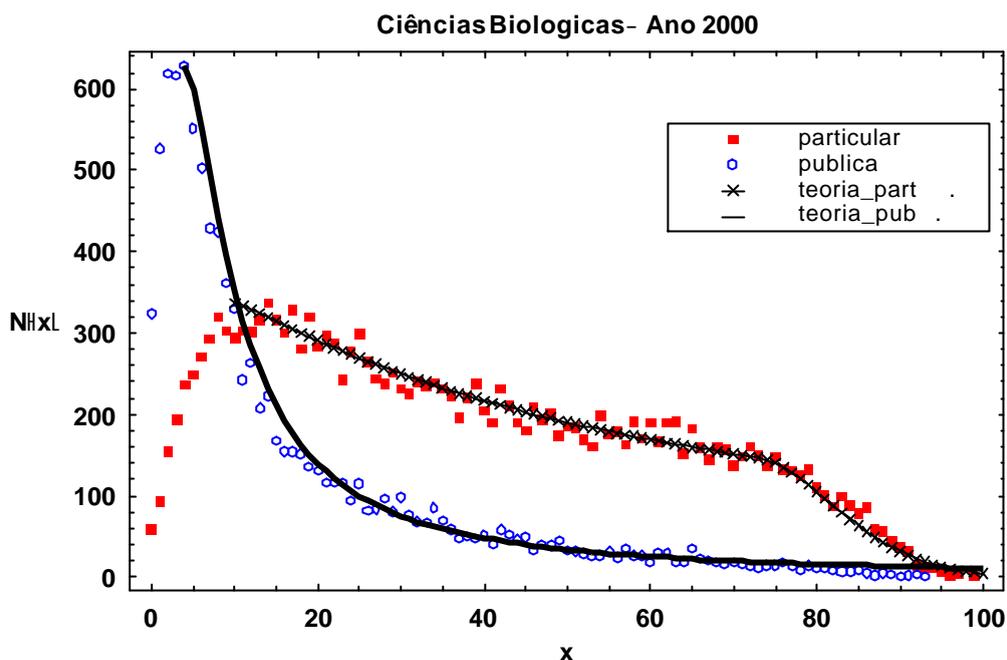


Figura 15: Números de candidatos que obtiveram nota  $x$  [ $N(x)$  versus notas obtidas ( $x$ )] para a avaliação de conhecimentos específicos da área de ciências biológicas do ano de 2000 para alunos da escola particular (◻) e pública (◉) no ensino médio.

Tabela 5: Tabela de parâmetros para o melhor ajuste com os dados empíricos na Distribuição de Lei de Potência Gradualmente Truncada para os dados dos alunos que cursaram o ensino médio na escola particular e pública para os anos de 1998, 1999 e 2000 na área de ciências biológicas.

Área de Ciências Biológicas						
Ano	Ano de 1998		Ano de 1999		Ano de 2000	
Escola	Particular	Pública	Particular	Pública	Particular	Pública
Nº de alunos	17.716	796	17.538	8.746	18.011	9.902
$1+\alpha$	1,13	1,05	2,92	1,63	1,14	1,52
$x_m$	6	6	45	2	10	4
$x_c$	72	75	87	82	74	91
$k$	115	200	0,5	19,63	140	13,79
$C_1$	56,90	19,90	16645,70	44,52	88,30	19,81
$-x$	33,66	32,05	39,37	18,76	37,60	16,60
$s$	23,36	23,01	22,88	19,28	23,89	17,97

Podemos verificar na Tabela 5, que para os alunos que cursaram escola particular no ano 1998,  $1+\alpha$  é aproximadamente 1.1 e para os alunos que cursaram a escola pública é 1.0 caracterizando assim a realimentação positiva, em que o ponto crítico das

notas dos candidatos da escola particular é 72 e para a escola pública é 75, causando assim um truncamento até o seu limite físico que é a nota 100.

Observamos também que para os alunos que cursaram escola particular do ano 1999,  $1+\alpha$  é aproximadamente 2.9 e para os alunos que cursaram a escola pública é 1.6 caracterizando assim a realimentação positiva, em que o ponto crítico das notas dos candidatos da escola particular é 87 e para a escola pública é 82, causando assim um truncamento até o seu limite físico que é a nota 100.

Finalmente para os alunos que cursaram escola particular no ano 2000,  $1+\alpha$  é aproximadamente 1.1 e para os alunos que cursaram a escola pública é 1.5 caracterizando assim a realimentação positiva, em que o ponto crítico das notas dos candidatos da escola particular é 74 e para a escola pública é 91, causando assim um truncamento até o seu limite físico que é a nota 100.

Notamos que o ajustamento da Distribuição de Lei de Potência Gradualmente Truncada com os dados empíricos é muito boa para a escola particular nos anos de 1998 e 2000, já para o ano de 1999 a que melhor descreve os dados é a Distribuição de Lei de Potência. No caso da escola pública a Distribuição de Lei de Potência Gradualmente Truncada descreve os dados dos anos de 1998, 1999 e 2000.

### **3.3.3 Área de Ciências Humanas**

Como mostram as Figuras 16, 17 e 18 comparamos as notas da avaliação de conhecimentos específicos da área de ciências humanas da escola particular e pública dos anos de 1998, 1999 e 2000, onde obtivemos a Distribuição Normal, utilizando os parâmetros  $\bar{x}$  e  $s$  como descrito na Tabela 6, sendo que na curva teórica são os melhores valores que se ajustam aos dados.

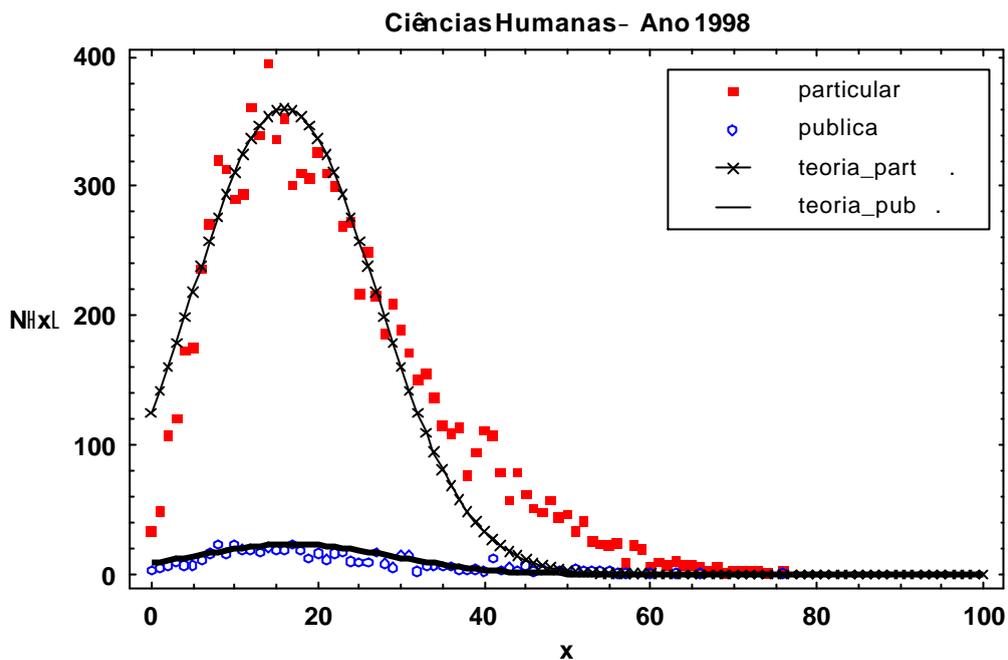


Figura 16: Números de candidatos que obtiveram nota  $x$  [ $N(x)$  versus notas obtidas ( $x$ )] para a avaliação de conhecimentos específicos da área de ciências humanas do ano de 1998 para alunos da escola particular (■) e pública (○) no ensino médio.

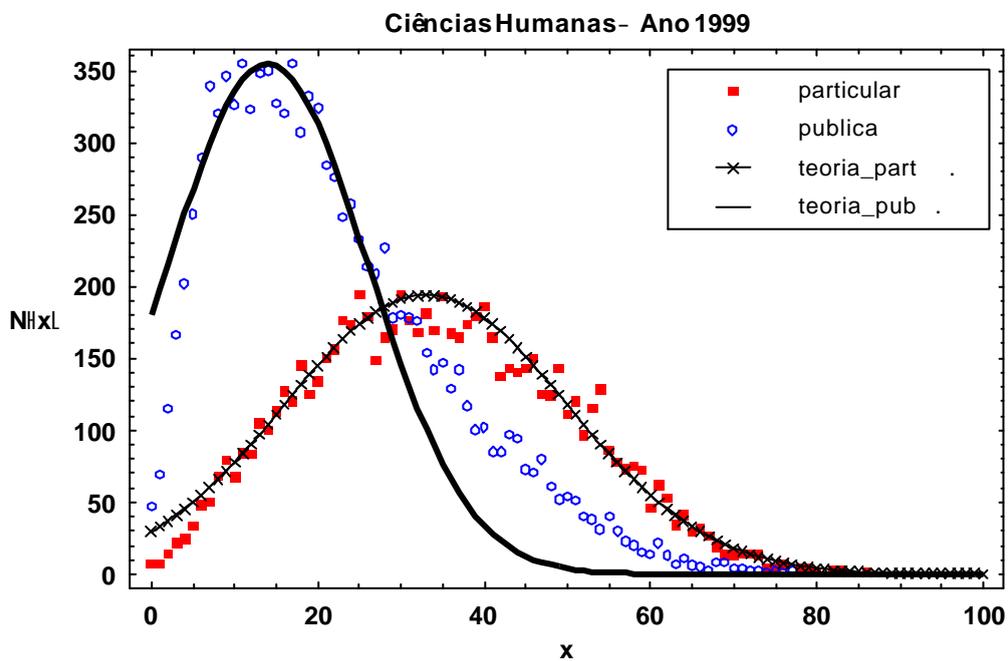


Figura 17: Números de candidatos que obtiveram nota  $x$  [ $N(x)$  versus notas obtidas ( $x$ )] para a avaliação de conhecimentos específicos da área de ciências humanas do ano de 1999 para alunos da escola particular (■) e pública (○) no ensino médio.

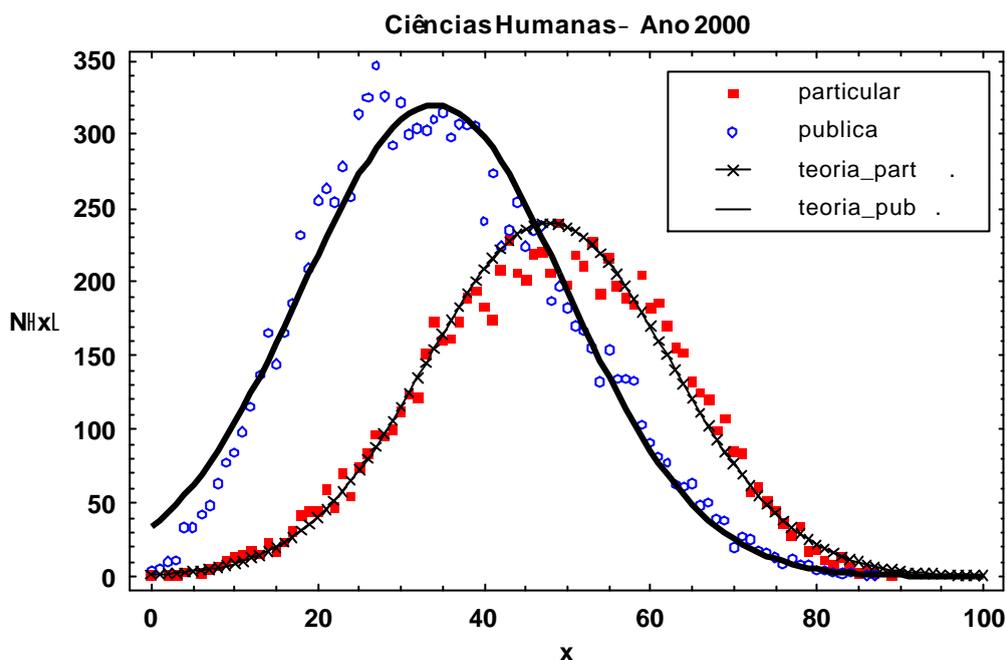


Figura 18: Números de candidatos que obtiveram nota  $x$  [ $N(x)$  versus notas obtidas ( $x$ )] para a avaliação de conhecimentos específicos da área de ciências humanas do ano de 2000 para alunos da escola particular (■) e pública (○) no ensino médio.

Tabela 6: Tabela de parâmetros para o melhor ajuste com os dados empíricos na Distribuição Normal, para os dados dos alunos que cursaram o ensino médio na escola particular e pública para os anos de 1998, 1999 e 2000 na área de ciências humanas.

Área de Ciências Humanas						
Ano	Ano de 1998		Ano de 1999		Ano de 2000	
Escola	Particular	Pública	Particular	Pública	Particular	Pública
Nº de alunos	9.972	537	7.675	10.660	8.658	12.400
$m$	16	17	33	14	47,69	34
$-x$	21,37	20,69	34,40	21,66	47,69	35,22
$S$	11	13,03	17	12	14,69	16
$s$	12,78	13,03	15,43	13,67	14,69	14,94

Notamos que o ajustamento da Distribuição de Normal com os dados empíricos é muito boa para o período diurno e noturno dos anos de 1998, 1999 e 2000.

### 3.4 Distribuição Estatística com Ênfase na Renda Família.

Neste tópico, abordamos apenas as notas dos candidatos da avaliação de conhecimentos específicos de cada área, que foram divididos em dois grupos: alunos que cursaram o ensino médio com renda familiar até 10 salários mínimos e com renda acima de 10 salários mínimos.

#### 3.4.1 Área de Ciências Exatas

Como mostram as Figuras 19, 20 e 21, comparamos as notas da avaliação de conhecimentos específicos da área de ciências exatas para os alunos que cursaram o ensino médio com renda familiar até 10 salários mínimos e com renda acima de 10 salários mínimos dos anos de 1998, 1999 e 2000, onde obtivemos a Distribuição de Lei de Potência Gradualmente Truncada, utilizando os parâmetros  $1+\alpha$ ,  $x_C$ ,  $x_m$ , e  $k$  como descrito na Tabela 7, sendo que na curva teórica  $x_C$  e  $x_m$  são dados empíricos e  $1+\alpha$  e  $k$  são os melhores valores que se ajustam às curvas.

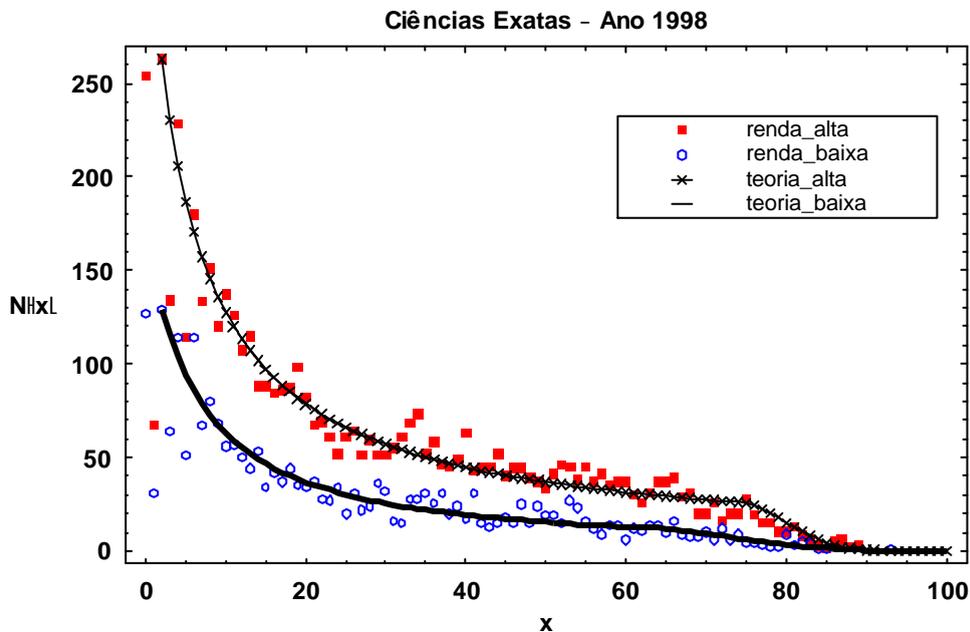


Figura 19: Números de candidatos que obtiveram nota  $x$  [ $N(x)$  versus notas obtidas ( $x$ )] para a avaliação de conhecimentos específicos da área de ciências exatas do ano de 1998 para alunos com renda acima de 10 salários mínimos (■) e abaixo de 10 salários mínimos (○) no ensino médio.

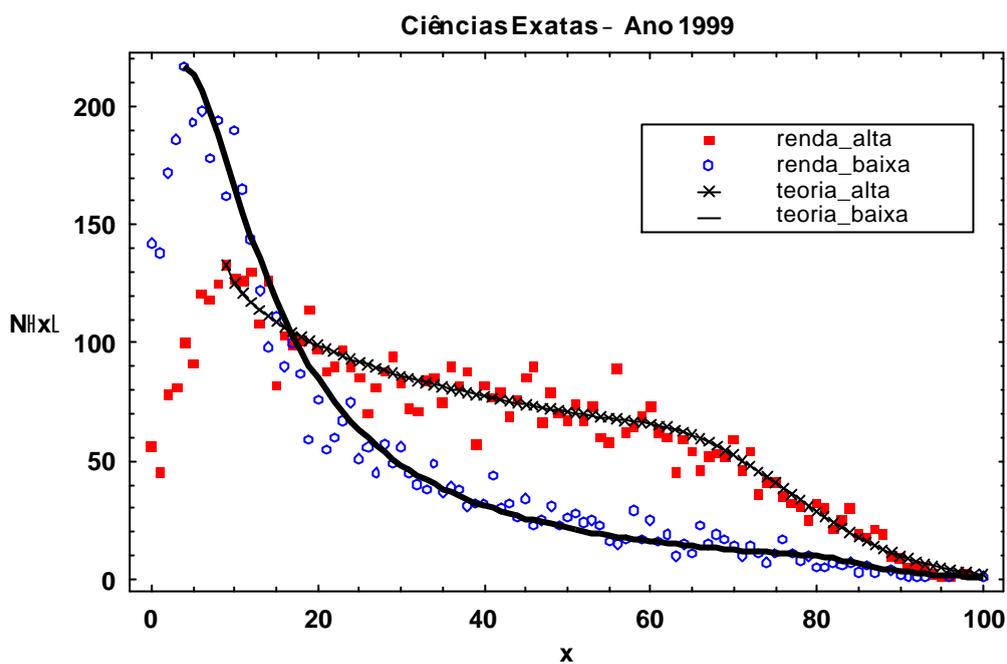


Figura 20: Números de candidatos que obtiveram nota  $x$  [ $N(x)$  versus notas obtidas ( $x$ )] para a avaliação de conhecimentos específicos da área de ciências exatas do ano de 1999 para alunos com renda acima de 10 salários mínimos (■) e abaixo de 10 salários mínimos (○) no ensino médio.

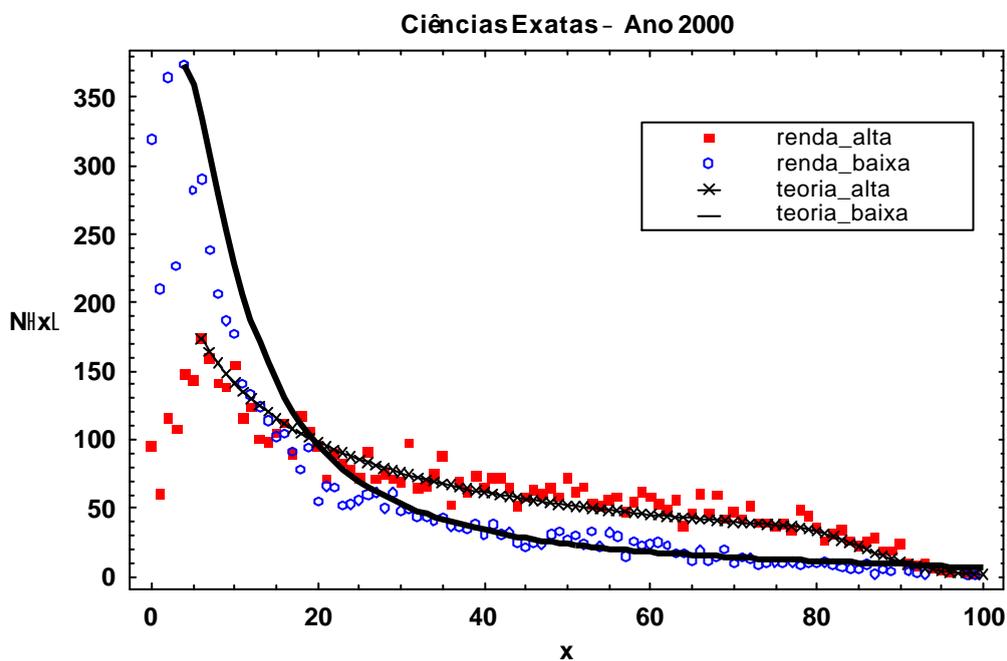


Figura 21: Números de candidatos que obtiveram nota  $x$  [ $N(x)$  versus notas obtidas ( $x$ )] para a avaliação de conhecimentos específicos da área de ciências exatas do ano de 2000 para alunos com renda acima de 10 salários mínimos (■) e abaixo de 10 salários mínimos (○) no ensino médio.

Tabela 7: Tabela de parâmetros para o melhor ajuste com os dados empíricos na Distribuição de Lei de Potência Gradualmente Truncada para os dados dos alunos que cursaram o ensino médio com renda acima de 10 salários mínimos e até 10 salários mínimos para os anos de 1998, 1999 e 2000 na área de ciências exatas.

<b>Área de Ciências Exatas</b>						
<b>Ano</b>	<b>Ano de 1998</b>		<b>Ano de 1999</b>		<b>Ano de 2000</b>	
<b>Renda</b>	<b>até 10</b>	<b>Acima 10</b>	<b>até 10</b>	<b>Acima 10</b>	<b>até 10</b>	<b>Acima 10</b>
	<b>S.M.</b>	<b>S.M.</b>	<b>S.M.</b>	<b>S.M.</b>	<b>S.M.</b>	<b>S.M.</b>
<b>Nº de alunos</b>	<b>2.387</b>	<b>5.247</b>	<b>4.798</b>	<b>6.398</b>	<b>5.768</b>	<b>6.364</b>
<b>1+<math>\alpha</math></b>	<b>1,075</b>	<b>0,97</b>	<b>1,65</b>	<b>0,71</b>	<b>1,33</b>	<b>0,96</b>
<b><math>x_m</math></b>	<b>2</b>	<b>2</b>	<b>4</b>	<b>9</b>	<b>4</b>	<b>6</b>
<b><math>x_C</math></b>	<b>65</b>	<b>75</b>	<b>81</b>	<b>60</b>	<b>87</b>	<b>77</b>
<b>k</b>	<b>200</b>	<b>80</b>	<b>27,22</b>	<b>1500</b>	<b>227</b>	<b>200</b>
<b><math>C_1</math></b>	<b>8,90</b>	<b>7,13</b>	<b>63,75</b>	<b>16,23</b>	<b>24,62</b>	<b>16,53</b>
<b><math>\bar{x}</math></b>	<b>25,89</b>	<b>23,62</b>	<b>35,63</b>	<b>21,11</b>	<b>34,18</b>	<b>18,89</b>
<b>s</b>	<b>20,96</b>	<b>22,14</b>	<b>19,96</b>	<b>23,40</b>	<b>20,11</b>	<b>25,02</b>

Podemos verificar na Tabela 7, que para os alunos que possuem renda familiar até 10 salários mínimos do ano 1998,  $1+\alpha$  é aproximadamente 1.0 e para os alunos que possuem renda familiar acima de 10 salários mínimos é 0.9 caracterizando assim a realimentação positiva, em que o ponto crítico das notas dos candidatos com renda até 10 salários mínimos é 65 e para a renda acima de 10 salários mínimos é 75, causando assim um truncamento até o seu limite físico que é a nota 100.

Observamos também que para os alunos que possuem renda familiar até 10 salários do ano 1999, mínimos  $1+\alpha$  é aproximadamente 0.7 e para os alunos que possuem renda familiar acima de 10 salários mínimos é 1.3 caracterizando assim a realimentação positiva, em que o ponto crítico das notas dos candidatos com renda até 10 salários mínimos é 60 e para a renda acima de 10 salários mínimos é 76, causando assim um truncamento até o seu limite físico que é a nota 100.

Finalmente para os alunos que possuem renda familiar até 10 salários mínimos do ano 2000,  $1+\alpha$  é aproximadamente 1.3 e para os alunos que possuem renda familiar acima de 10 salários mínimos é 0.9 caracterizando assim a realimentação positiva, em que o ponto crítico das notas dos candidatos com renda até 10 salários mínimos é 87 e para a renda acima de 10 salários mínimos é 77, causando assim um truncamento até o seu limite físico que é a nota 100.

Notamos que o ajustamento da Distribuição de Lei de Potência Gradualmente Truncada com os dados empíricos é muito boa para os alunos com renda até dez salários mínimos e acima de dez salários mínimos para os anos de 1998, 1999 e 2000.

### **3.4.2 Área de Ciências Biológicas**

Como mostram as Figuras 22, 23 e 24 comparamos as notas da avaliação de conhecimentos específicos da área de ciências biológicas para os alunos que cursaram o ensino médio com renda familiar até 10 salários mínimos e com renda acima de 10 salários mínimos dos anos de 1998, 1999 e 2000, onde obtivemos a Distribuição de Lei de Potência Gradualmente Truncada, utilizando os parâmetros  $1+\alpha$ ,  $x_c$ ,  $x_m$ , e  $k$  como descrito na Tabela 8, sendo que na curva teórica  $x_c$  e  $x_m$  são dados empíricos e  $1+\alpha$  e  $k$  são os melhores valores que se ajustam às curvas.

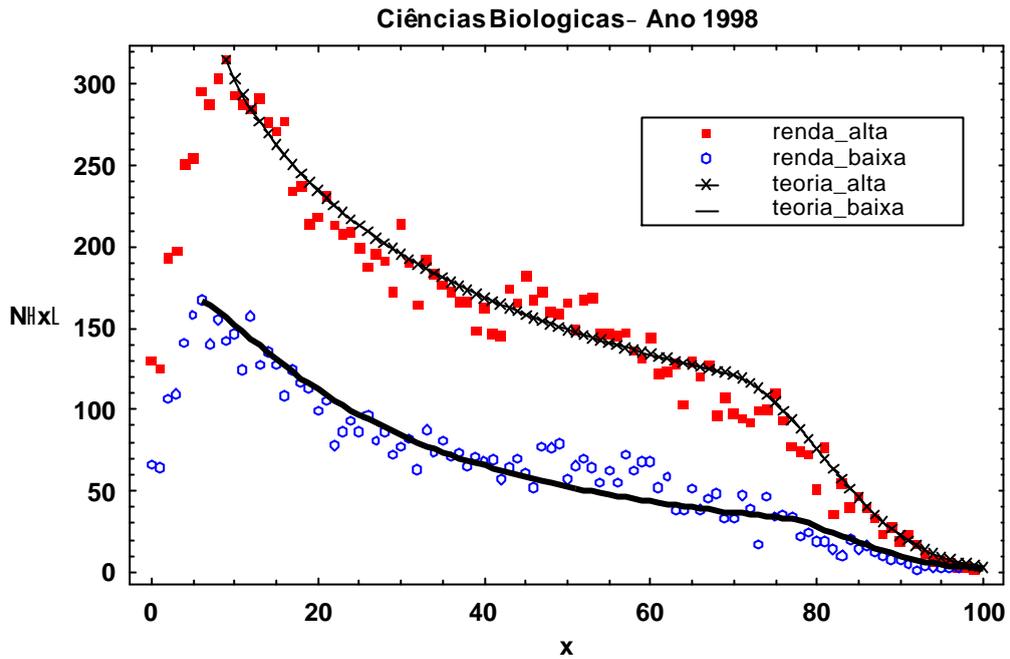


Figura 22: Números de candidatos que obtiveram nota  $x$  [ $N(x)$  versus notas obtidas ( $x$ )] para a avaliação de conhecimentos específicos da área de ciências biológicas do ano de 1998 para alunos com renda acima de 10 salários mínimos (■) e abaixo de 10 salários mínimos (○) no ensino médio.

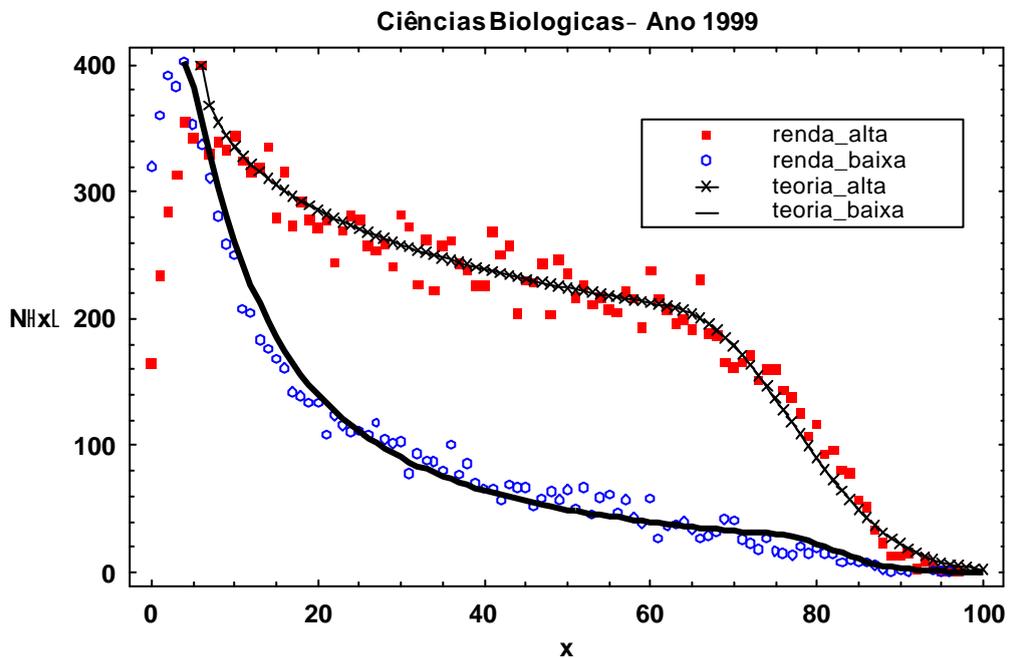


Figura 23: Números de candidatos que obtiveram nota  $x$  [ $N(x)$  versus notas obtidas ( $x$ )] para a avaliação de conhecimentos específicos da área de ciências biológicas do ano de 1999 para alunos com renda acima de 10 salários mínimos (■) e abaixo de 10 salários mínimos (○) no ensino médio.

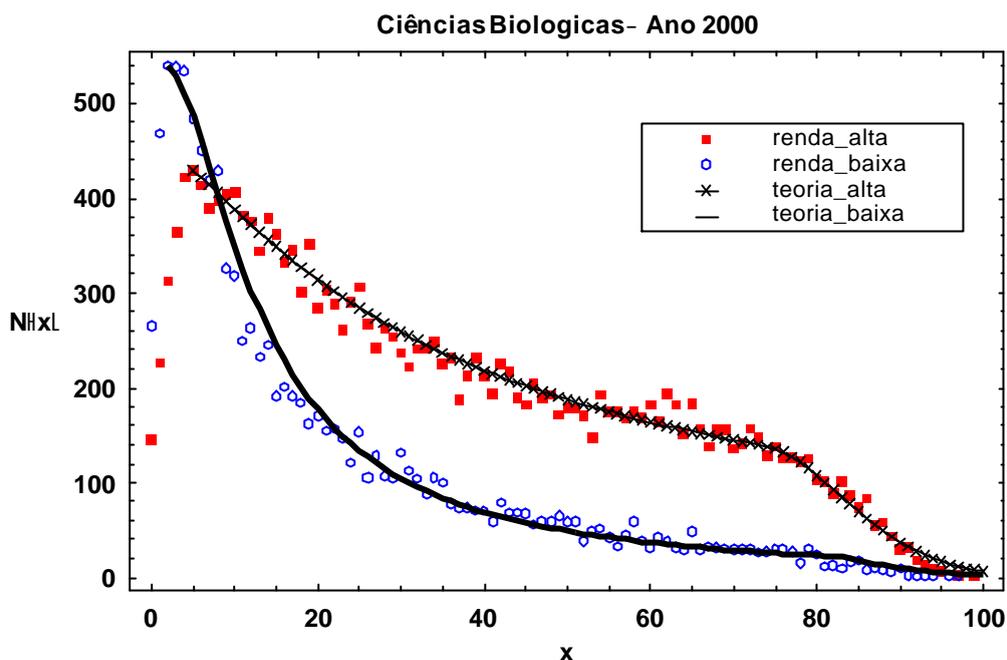


Figura 24: Números de candidatos que obtiveram nota  $x$  [ $N(x)$  versus notas obtidas ( $x$ )] para a avaliação de conhecimentos específicos da área de ciências biológicas do ano de 2000 para alunos com renda acima de 10 salários mínimos (■) e abaixo de 10 salários mínimos (○) no ensino médio.

Tabela 8: Tabela de parâmetros para o melhor ajuste com os dados empíricos na Distribuição de Lei de Potência Gradualmente Truncada para os dados dos alunos que cursaram o ensino médio com renda acima de 10 salários mínimos e até 10 salários mínimos para os anos de 1998, 1999 e 2000 na área de ciências biológicas.

Área de Ciências Biológicas						
Ano	Ano de 1998		Ano de 1999		Ano de 2000	
Renda	até 10 S.M.	Acima 10 S.M.	até 10 S.M.	Acima 10 S.M.	até 10 S.M.	Acima 10 S.M.
Nº de alunos	6.424	14.564	9.007	19.981	10.849	19.727
$1+a$	1,30	0,89	1,26	0,58	1,62	1,14
$x_m$	6	9	4	6	2	5
$x_C$	77	70	77	62	83	75
$k$	87,14	400	200	1600	26,68	150
$C_1$	63,45	25,02	17,82	11,75	54,11	60,28
$-x$	34,54	31,94	36,42	22,30	34,44	20,63
$s$	22,96	23,47	20,73	23,65	20,27	24,36

Podemos verificar na Tabela 8, que para os alunos que possuem renda familiar até 10 salários mínimos do ano 1998,  $1+\alpha$  é aproximadamente 1.3 e para os alunos que possuem renda familiar acima de 10 salários mínimos é 0.9 caracterizando assim a realimentação positiva, em que o ponto crítico das notas dos candidatos com renda até 10 salários mínimos é 77 e para a renda acima de 10 salários mínimos é 70, causando assim um truncamento até o seu limite físico que é a nota 100.

Observamos também que para os alunos que possuem renda familiar até 10 salários mínimos do ano 1999,  $1+\alpha$  é aproximadamente 1.2 e para os alunos que possuem renda familiar acima de 10 salários mínimos é 0.6 caracterizando assim a realimentação positiva, em que o ponto crítico das notas dos candidatos com renda até 10 salários mínimos é 77 e para a renda acima de 10 salários mínimos é 62, causando assim um truncamento até o seu limite físico que é a nota 100.

Finalmente para os alunos que possuem renda familiar até 10 salários mínimos do ano 2000,  $1+\alpha$  é aproximadamente 1.6 e para os alunos que possuem renda familiar acima de 10 salários mínimos é 1.1 caracterizando assim a realimentação positiva, em que o ponto crítico das notas dos candidatos com renda até 10 salários mínimos é 83 e para a renda acima de 10 salários mínimos é 75, causando assim um truncamento até o seu limite físico que é a nota 100.

Notamos que o ajustamento da Distribuição de Lei de Potência Gradualmente Truncada com os dados empíricos é muito boa para os alunos com renda até dez salários mínimos e acima de dez salários mínimos para os anos de 1998, 1999 e 2000.

### **3.4.3 Área de Ciências Humanas**

Como mostram as Figuras 25, 26 e 27 comparamos as notas da avaliação de conhecimentos específicos da área de ciências humanas para os alunos que cursaram o ensino médio com renda familiar até 10 salários mínimos ou com renda acima de 10 salários mínimos dos anos de 1998, 1999 e 2000, onde obtivemos a Distribuição

Normal, utilizando os parâmetros  $\bar{x}$  e  $s$  como descrito na Tabela 9, sendo que na curva teórica são os melhores valores que se ajustam aos dados.

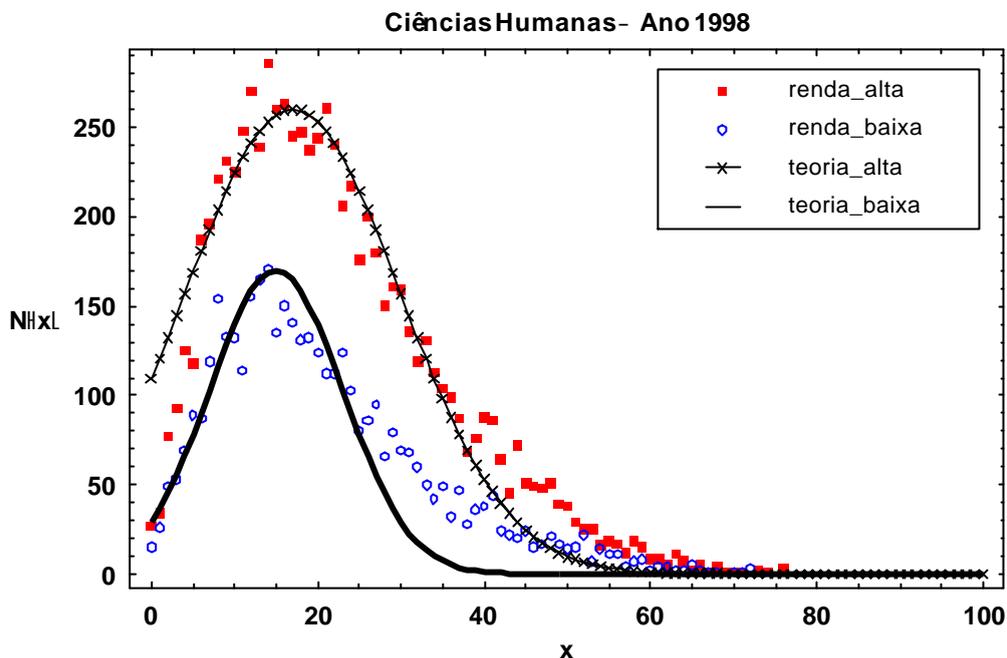


Figura 25: Números de candidatos que obtiveram nota  $x$  [ $N(x)$  versus notas obtidas ( $x$ )] para a avaliação de conhecimentos específicos da área de ciências humanas do ano de 1998 para alunos com renda acima de 10 salários mínimos (■) e abaixo de 10 salários mínimos (○) no ensino médio.

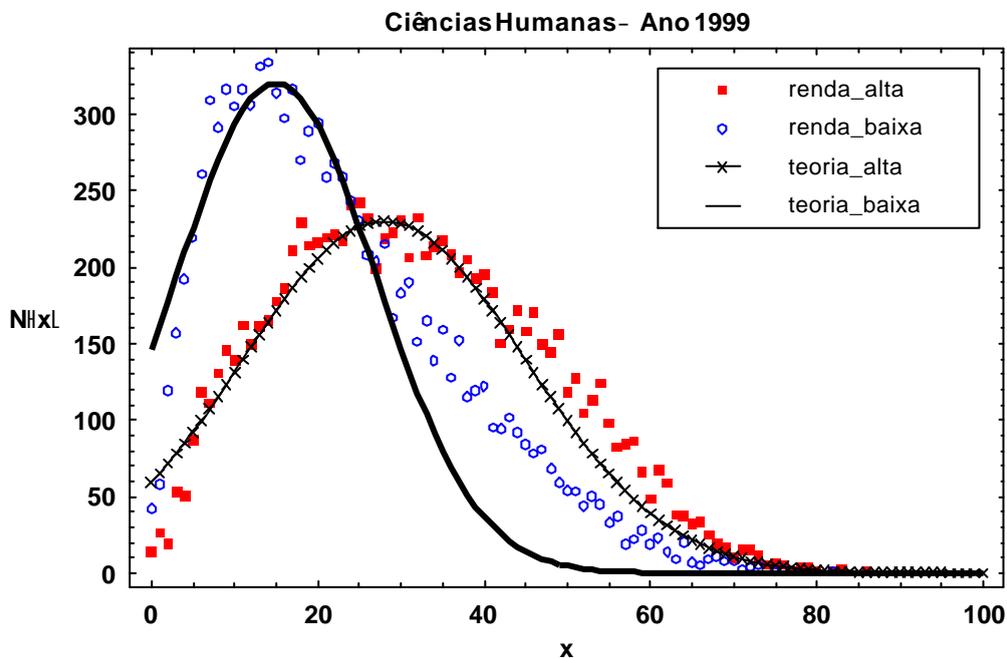


Figura 26: Números de candidatos que obtiveram nota  $x$  [ $N(x)$  versus notas obtidas ( $x$ )] para a avaliação de conhecimentos específicos da área de ciências humanas do ano de 1999 para alunos com renda acima de 10 salários mínimos (■) e abaixo de 10 salários mínimos (○) no ensino médio.

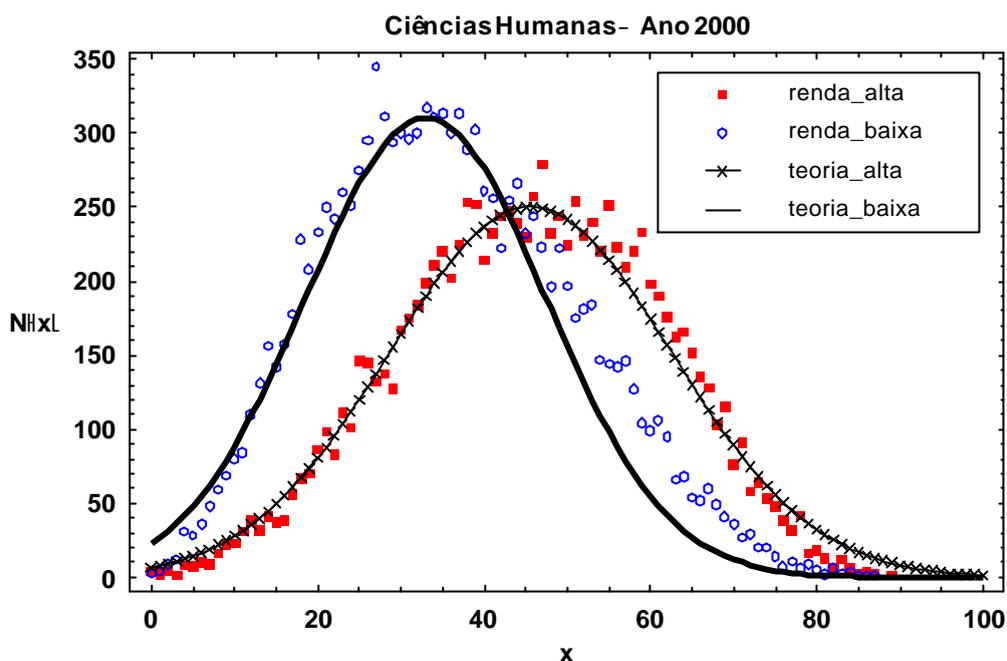


Figura 27: Números de candidatos que obtiveram nota  $x$  [ $N(x)$  versus notas obtidas ( $x$ )] para a avaliação de conhecimentos específicos da área de ciências humanas do ano de 2000 para alunos com renda acima de 10 salários mínimos (■) e abaixo de 10 salários mínimos (○) no ensino médio.

Tabela 9: Tabela de parâmetros para o melhor ajuste com os dados empíricos na Distribuição Normal para os dados dos alunos que cursaram o ensino médio com renda acima de 10 salários mínimos e até 10 salários mínimos para os anos de 1998, 1999 e 2000 na área de ciências humanas.

Área de Ciências Humanas						
Ano	Ano de 1998		Ano de 1999		Ano de 2000	
Renda	até 10 S.M.	Acima 10 S.M.	até 10 S.M.	Acima 10 S.M.	até 10 S.M.	Acima 10 S.M.
Nº de alunos	4.065	7.822	10.301	9.971	12.386	10.602
$m$	15	17	15	28	33	45,58
$-x$	20,59	21,93	22,57	31,55	35,98	45,58
$S$	8	12,90	12	17	14,5	17
$s$	12,69	12,90	14,21	15,77	15,11	15,44

Notamos que o ajustamento da Distribuição de Normal com os dados empíricos é muito boa para os alunos que cursaram o ensino médio com renda familiar até 10 salários mínimos e com renda acima de 10 salários mínimos dos anos de 1998, 1999 e 2000.

## CAPÍTULO 4

### CONCLUSÕES

Neste capítulo, apresentaremos as conclusões por fatores estudados e uma conclusão geral do estudo, em que nela apresentaremos algumas suposições e sugestões do sistema estudado.

#### **4.1 Alunos que cursaram o ensino médio no período diurno ou noturno.**

Verificamos que para todos os anos, em todas as áreas, os candidatos que cursaram o ensino médio no período diurno obtiveram um desempenho melhor no vestibular da UNESP em relação aos alunos que cursaram no período noturno.

Para verificarmos esse resultado, calculamos a percentagem das médias dos alunos contidos na tabela1, da seguinte forma:

$$\left[ \frac{\text{(média dos alunos do ensino médio diurno)}}{\text{(média dos alunos do ensino médio do noturno)}} - 1 \right] * 100 = \text{percentagem do aproveitamento. Exemplo: } \left[ \frac{25.5937}{22.9878} - 1 \right] * 100 = 11.33\%.$$

Em seguida, mostraremos esses resultados divididos por área:

- Para a área de ciências exatas no ano de 1998, a média das notas dos alunos do período diurno foi 11,33% melhor do que do período noturno. Em 1999, a média das notas dos alunos do período diurno foi 131,16% melhor do que do período noturno. No ano de 2000, a média das notas dos alunos do período diurno foi 140,00% melhor do que do período noturno.
- Para a área de ciências biológicas no ano de 1998, a média das notas dos alunos do período diurno foi 7,63% melhor do que do período noturno. Em 1999, a média das notas dos alunos do período diurno foi 140,53% melhor do que do período noturno. No ano de 2000, a média das notas dos alunos do período diurno foi 156,61% melhor do que do período noturno.
- Para a área de ciências humanas no ano de 1998, a média das notas dos alunos do período diurno foi 6,71% melhor do que do período noturno. Em 1999, a média das notas dos alunos do período diurno foi 65,23% melhor do que do período noturno. No ano de 2000, a média das notas dos alunos do período diurno foi 36,62% melhor do que do período noturno.

Verifica-se que os alunos que cursaram o ensino médio no período diurno obtiveram médias melhores que os do período noturno. Em relação às áreas nos 3 anos, as percentagens são as seguintes:

- 94,16% na área de ciências exatas
- 101,59% na área de ciências biológicas
- 36,19% na área de ciências humanas.

No caso das distribuições das notas dos alunos, verificamos que para as áreas de ciências biológicas e ciências exatas nos observamos que a distribuição que melhor se enquadra é a Distribuição de Lei de Potência Gradualmente Truncada em todos os casos, mas na área de ciências humanas a que melhor se aproximou foi a Distribuição Normal.

## **4.2 Alunos que cursaram o ensino médio em escola particular ou escola pública.**

Verificamos também que para todos os anos, em todas as áreas, os candidatos que cursaram o ensino médio em escola particular obtiveram um desempenho melhor no vestibular da UNESP em relação aos alunos que cursaram em escola pública.

Para melhor ilustrar essa conclusão, mostramos a seguir as percentagens obtidas do mesmo modo citado anteriormente em cada área nos 3 anos:

- Para a área de ciências exatas no ano de 1998, a média das notas dos alunos que cursaram o ensino médio em escola particular foi 6,19% melhor do que da escola pública. Em 1999, a média das notas dos alunos que cursaram o ensino médio em escola particular foi 99,71% melhor do que da escola pública. No ano de 2000, a média das notas dos alunos que cursaram o ensino médio em escola particular foi 116,89% melhor do que da escola pública.
- Para a área de ciências biológicas no ano de 1998, a média das notas dos alunos que cursaram o ensino médio em escola particular foi 5,01% melhor do que da escola pública. Em 1999, a média das notas dos alunos que cursaram o ensino médio em escola particular foi 109,77% melhor do que da escola pública. No ano de 2000, a média das notas dos alunos que

cursaram o ensino médio em escola particular foi 126,49% melhor do que da escola pública.

- Para a área de ciências humanas no ano de 1998, a média das notas dos alunos que cursaram o ensino médio em escola particular foi 3,25% melhor do que da escola pública. Em 1999, a média das notas dos alunos que cursaram o ensino médio em escola particular foi 58,84% melhor do que da escola pública. No ano de 2000, a média das notas dos alunos que cursaram o ensino médio em escola particular foi 35,63% melhor do que da escola pública.

Verificamos que os alunos que cursaram o ensino médio em escola particular obtiveram médias melhores que os alunos da escola pública. Pode-se verificar que o desempenho por área ficou no seguinte:

- 74,26% na área de ciências exatas
- 80,42% na área de ciências biológicas
- 32,57% na área de ciências humanas.

No caso da distribuição das notas dos alunos, verificamos que para as áreas de ciências biológicas e ciências exatas a distribuição que melhor se enquadra é a Distribuição de Lei de Potência Gradualmente Truncada em todos os casos, mas na área de ciências humanas a que melhor se aproximou foi a Distribuição Normal.

### **4.3 Alunos que cursaram o ensino médio que possuem renda familiar até 10 salários mínimos ou acima de 10 salários mínimos.**

Concluimos que para todos os anos, em todas as áreas, os candidatos que cursaram o ensino médio com renda familiar acima de 10 salários mínimos obtiveram um desempenho melhor no vestibular da UNESP em relação aos alunos que possuem renda familiar até 10 salários mínimos.

Para detalharmos isso, calculamos as percentagens do desempenho dos alunos por área nos 3 anos seguindo o mesmo procedimento anteriormente mostrado, como segue:

- Para a área de ciências exatas no ano de 1998, a média das notas dos alunos que cursaram o ensino médio com renda familiar acima de 10 salários mínimos foi 9,63% melhor do que os alunos com renda familiar até 10 salários mínimos. Em 1999, a média das notas dos alunos que cursaram o ensino médio com renda familiar acima de 10 salários mínimos foi 68,76% melhor do que os alunos com renda familiar até 10 salários mínimos. No ano de 2000, a média das notas dos alunos que cursaram o ensino médio com renda familiar acima de 10 salários mínimos foi 80,87% melhor do que os alunos com renda familiar até 10 salários mínimos.
- Para a área de ciências biológicas no ano de 1998, a média das notas dos alunos que cursaram o ensino médio com renda familiar acima de 10 salários mínimos foi 8,14% melhor do que os alunos com renda familiar até 10 salários mínimos. Em 1999, a média das notas dos alunos que cursaram o ensino médio com renda familiar acima de 10 salários mínimos foi 63,34% melhor do que os alunos com renda familiar até 10 salários mínimos. No ano de 2000, a média das notas dos alunos que cursaram o ensino médio com renda familiar acima de 10 salários mínimos foi 66,92% melhor do que os alunos com renda familiar até 10 salários mínimos.
- Para a área de ciências humanas no ano de 1998, a média das notas dos alunos que cursaram o ensino médio com renda familiar acima de 10 salários mínimos foi 6,53% melhor do que os alunos com renda familiar até 10 salários mínimos. Em 1999, a média das notas dos alunos que cursaram o ensino médio com renda familiar acima de 10 salários mínimos foi 39,78% melhor do que os alunos com renda familiar até 10 salários mínimos. No ano de 2000, a média das notas dos alunos que cursaram o ensino médio com renda familiar acima de 10 salários mínimos foi 26,67% melhor do que os alunos com renda familiar até 10 salários mínimos.

Verificamos que em média os alunos que cursaram o ensino médio com renda familiar acima de 10 salários mínimos obtiveram desempenho melhor que os alunos com renda familiar até 10 salários mínimos. Notamos que em relação às áreas dos 3 anos esses desempenhos ficam no seguinte:

- 53,09% na área de ciências exatas
- 46,13% na área de ciências biológicas
- 24,33% na área de ciências humanas.

No caso da distribuição das notas dos alunos, verificamos que para as áreas de ciências biológicas e ciências exatas nos observamos que a distribuição que melhor se enquadra é a Distribuição de Lei de Potência Gradualmente Truncada em todos os casos, mas na área de ciências humanas a que melhor se aproximou foi a Distribuição Normal.

#### **4.4 Conclusão Geral**

Verificamos que o sistema educacional é um sistema complexo, devido ao desempenho de um aluno no vestibular depender de vários fatores, como por exemplo à condição de ensino, a condição de tempo de estudo, a tradição da família, a renda familiar, a natureza das áreas de estudo.

Observamos que na área de Ciências Exatas e Ciências Biológicas, as distribuições das notas dos alunos se enquadram na Distribuição de Lei de Potência Gradualmente Truncada e a área de Ciências Humanas se enquadra na Distribuição Normal, independente dos fatores citados acima.

Concluimos que para os três fatores estudados, período em que aluno estudou, tipo de escola que frequentou e a renda familiar do aluno, o fator que maior se destacou foi o período, seguido pelo tipo de escola e logo depois a renda familiar do aluno. Para o

primeiro fator, os alunos que cursaram o ensino médio no período diurno tiveram um desempenho 77,31% melhor no vestibular da UNESP que os alunos que cursaram o ensino médio no período noturno.

Para o segundo fator, os alunos que cursaram o ensino médio em escola particular tiveram um desempenho 62,42% melhor no vestibular da UNESP que os alunos que cursaram o ensino médio em escola pública.

No terceiro fator estudado, os alunos que cursaram o ensino médio com renda familiar acima de 10 salários mínimos tiveram um desempenho 41,18% melhor no vestibular da UNESP que os alunos que cursaram o ensino médio com renda até 10 salários mínimos.

Todos esses resultados podem ser condensados através de algumas suposições, como:

- alunos que cursaram o ensino médio no período noturno tem que trabalhar durante o dia, isso diminui o tempo de estudo.
- melhores condições no ensino médio privado do que no público.
- melhores laboratórios de ensino no ensino privado.
- os alunos por terem melhores condições financeiras, possuem material didático melhor.
- os alunos com melhores condições financeiras, não precisam trabalhar e dedicam-se somente ao estudo.

Verificamos que o ensino privado demonstra uma superioridade em relação ao ensino público, numa tentativa de melhorar as condições do ensino médio público, sugerimos baseados nos resultados acima, o seguinte:

i) o governo deveria investir recursos na infra-estrutura da escola do ensino médio, pagando um melhor salário para seus professores atraindo assim um melhor profissional, principalmente para as áreas de ciências exatas e biológicas.

ii) o governo deveria ajudar os alunos de famílias de menor renda com uma bolsa para que o aluno com renda familiar baixa não tenha que trabalhar durante o dia e estudar no período noturno, particularmente o aluno com talento para as áreas de ciências exatas e biológicas.

iii) o governo deveria avaliar ainda mais (o ENEM já é um grande passo nessa direção) a qualidade do ensino nas escolas. Achamos que uma das principais avaliações do ensino de uma escola deveria ser feita através da porcentagem de alunos dessa escola que passam no vestibular, ao invés da nota que o aluno recebe do professor da própria escola.

iv) o governo deveria criar melhores condições para que os alunos permaneçam mais freqüentemente nas aulas das áreas de ciências exatas e biológicas.

v) o governo deveria montar laboratórios na área de ciências exatas e biológicas para que os alunos tenham uma noção real e não algo abstrato (Gennes, 1994).

Observamos neste estudo, que o fator econômico é importante, mas a qualidade do ensino e as condições de estudo são mais importantes.

Concluimos que para as áreas de Ciências Exatas e Ciências Biológicas isso ocorre devido à natureza das áreas mesmo, a Distribuição de Lei de Potência Gradualmente Truncada ocorre quando a distribuição possui uma realimentação positiva, enquanto que para a área de Ciências Humanas ocorre a Distribuição Normal devido às observações serem independentes, ou seja, os alunos conseguem estudar assuntos da área de ensino sem ter que estudar algum pré-requisito antes, então o conhecimento de cada assunto é independente. Já para as áreas de Ciências Exatas e Ciências Biológicas, como a matemática ou física os assuntos são interligados, um aluno necessita ser regular nos estudos, pois ele precisa aprender gradativamente, ou seja, acumular conhecimento de assunto para assunto; isso caracteriza a realimentação positiva e acreditamos que isso produz a Distribuição de Lei de Potência Gradualmente Truncada.

## REFERENCIAS BIBLIOGRAFICAS

AYRES, M.; AYRES Jr, M.; AYRES, D. L.; SANTOS, A.S. **Bioestat 2.0**: aplicações estatísticas nas áreas das ciências biológicas e médicas. Belém: Sociedade Civil Mamirauá; MCT – CNPq, 2000.

BADII, R; POLITI, A. **Complexity Hierarchical Structures and Scaling in Physics.** Cambridge: Cambridge University Press, 1997.

BAIN, A. **The growth of television ownership in the united kingdon since the war.** London: Cambridge University Press, 1964.

BAK, P. **How nature works:** The science of self-organized criticality. Oxford: Oxford University Press, 1997.

BAK, P.; TANG, C.; WIESENFELD, K. Self-organized criticality: an explanation of *1/f* noise, **Physical Review Letters**, New York, v.59, p.381, 1987.

BAK, P.; TANG, C.; WIESENFELD, K. Self-organized criticality, **Physical Review A**, New York, v.38, p.364, 1988.

BERGSTROM, H., On some expansions of stable distributions, **Arkiv fur Matematik** v.2, n° 18, p.375-378, 1952.

BLISS, C. I. The method of probits. **Science**, New York, v.79, p.38-39, 2002.

COCHRAN, W. G. Some difficulties in the statistical analysis of replicated experiments, **Empire Journal of Experimental Agriculture**, Washington, v.6, p.157-163, 1938.

FELLER, W. **An introduction to probability theory and its applications**. New York: Wiley, 1971.

FRISH, U.; SHLESINGER, M. F.; ZASLAVASKY, G. **Lévy flights and related phenomena in physics**, Berlin: Springer-Verlag, 1994.

GADDUM, J. H. **Reports on biological standards III**: methods of biological assay depending on a quantal response. London: Medical Research Council, 1993. (Special Report Series, 183).

GALTON, F. **The geometric mean in vital and social statistics**, London: Royal Society of London, v.29, p.365-367, 1879.

GAUSS, C. F. Bestimmung der genauigkeit der beobachtungen, **Zeitschrift für Astronomie**, Berlin, v.1, p.185-197, 1816.

GENNES, P.G.; BADOZ, J. **Soft Matter, Hard Science, and the Thrill of Discovery**. Paris: Librairie Plon, 1994.

GIBRAT, R. **Les inégalités économiques**. Paris: Libraire du Recueil Sirey, 1931.

GUPTA, H. M.; CAMPANHA, J. R. Gradually truncated log-normal distribution – size distribution of firms. arXiv.org e-Print archive, v.1, 30 nov. 2001. Disponível em:<**http://arxiv.org/pdf/cond-mat/0111579**>. Acesso em 14.ago.2002.

GUPTA, H. M.; CAMPANHA, J. R. The gradually truncated Lévy flight: stochastic process for complex systems, **Physica A**, New York, v.268, p.231-239, 1999.

GUPTA, H. M.; CAMPANHA, J. R. The gradually truncated Lévy flight: stochastic process for complex systems, **Physica A**, New York, v.275, p.531-543, 2000.

GUPTA, H. M.; CAMPANHA, J. R. Tsallis statistics and gradually truncated Lévy flight – distribution of an economical index, **Physica A**, New York, v.309, p.381-387, 2002.

GUPTA, H. M.; CAMPANHA, J. R.; PRADO, F. D. Power law distribution in education: University Entrance Examination. **International Journal of Modern Physics C**, New York, v.11, n.6, p.1273-1279, 2000.

GUTENBERG, B.; RICHTER, C. F. **Seismicity the earth**. Princeton: Princeton University Press, 1949.

HATCH, T.; CHOUTE, S. P. Statistical description of the size properties of non-uniform particulate substances. **Journal of the Franklin Institute**, Philadelphia, v.207, p.369-388, 1929.

HURST, H. E. Long-term storage of reservoirs. **Transactions of the American Society of Civil Engineers**, Reston, v.116, p.770-799, 1951.

KAPTEYN, J. C.; VAN UEN, M. J. **Skew frequency curves in biology and statistics**. Groningen: Astronomical Laboratory at Groningen, 1916.

LAPLACE, P. S. Mémoire sur les probabilités, **Histoire de l'Académie Royale de Sciences**, Paris, v.9, p.227-332, 1778(1781).

LÉVY, P. **Théorie de L'Addition des Variables Aléatoires**, Paris: Gauthier-Villars, 1937.

MAGALHÃES, M. V.; LIMA, A. C. P. **Nocões de probabilidade e estatística**. 3.ed. São Paulo: IME/USP, 2001.

MANDELBROT, B. **The fractal geometry of nature**. New York: Freeman, 1983.

MANDELBROT, B. The variation of certain speculative prices. **Journal of Business of the University of Chicago**, Chicago, v.36, p.394-419, 1963.

MANDELBROT, B. The variation of some other speculative prices. **Journal of Business of the University of Chicago**, Chicago, v.37, p.393, 1964.

Mc-ALLISTER, D. The law of the geometrical mean. **Royal Society of London**, London, v.29, p.369-375, 1879.

MOIVRE, A. de. **Approximatio ad summam terminorum binomii  $(a+b)^n$  in seriem expansi**. London: Supplement to Miscellanea Analytica, 1733.

PARETO, V. **Cours d'economic politique**. Lausanne: Rounge, 1896.

PRESS, W.H. e Cols., **Numerical Recipes – The Art of Scientific Computing**. Cambridge: Cambridge University Press, 1992.

SEPKOSKI Jr, J. J. Ten years in the library: new data confirm paleontological patterns. **Paleobiology**, Chicago, v.19, p.43-51, 1993.

SCHROEDER, M. **Fractals, chaos, power laws: minutes from an intinite paradise**. New York: W. H. Freeman and Company, 1991.

TSALLIS C. Nonextensive statistics: theoretical, experimental and computational evidences and connections. **Brazilian Journal of Physics**, São Paulo, v.29, p.1-35, 1999.

VIEIRA, S. **Introdução à bioestatística**. São Paulo: Campus - 3.ed., 1980.

WILLIAMS, C. B. A note on the statistical analysis of sentence length. **Biometrika**,  
Winnipeg, v.31, p.356-361, 1940.

ZIPF, G. K. **Human behavior and the principle of least effort**. Cambridge: Addison-  
Wesley, 1949.

## APÊNDICE

### **ARTIGO ACEITO PARA PUBLICAÇÃO.**

Gupta, H.M.; Campanha, J.R.; Chavarette, F.R. Power law distribution in Education: Effect of Economical, Teaching and study conditions in University Entrance Examination. **International Journal of Modern Physics – Computation**, New York, v. 14, n. 4, 2003.