

O PROBLEMA DISCRIMINATÓRIO NUMA  
FAMÍLIA DE DISTRIBUIÇÕES DISCRETAS

LISBETH KAISERLIAN CORDANI

DISSERTAÇÃO APRESENTADA AO  
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA  
DA  
UNIVERSIDADE DE SÃO PAULO  
PARA OBTENÇÃO DO GRAU DE MESTRE  
EM  
ESTATÍSTICA

ORIENTADOR:

PROF. DR. FLÁVIO WAGNER RODRIGUES

- SÃO PAULO, OUTUBRO DE 1976 -

"O ESSENCIAL É INVISÍVEL PARA OS OLHOS"

A.S.E.

Queremos agradecer a todos que, sendo, acima de tudo, nossos amigos, contribuíram direta ou indiretamente para a realização deste trabalho.

Em particular, destacamos:

- Dr. Flávio Wagner Rodrigues, nosso orientador;
- Dr. Carlos Alberto Barbosa Dantas, que possibilitou nossos primeiros contactos com o Departamento de Estatística do IME-USP.
- Dr. Rubens Murillo Marques que elaborou os programas de simulação, gentilmente executados através do Departamento de Estatística da Unicamp.
- Sr. João Baptista Esteves de Oliveira pela dedicação no trabalho de datilografia.

## Í N D I C E

CAP. 1 - INTRODUÇÃO. . . . .	1
1.1 - Considerações Gerais. . . . .	1
1.2 - Contágio. . . . .	3
1.3 - Distribuições Compostas e Generalizadas . . .	5
CAP. 2 - CARACTERIZAÇÃO DE ALGUMAS DISTRIBUIÇÕES DISCRETAS .	8
2.1 - Distribuição Binomial . . . . .	8
2.2 - Distribuição de Poisson . . . . .	14
2.3 - Distribuição Binomial Negativa. . . . .	23
CAP. 3 - MÉTODOS DE DISCRIMINAÇÃO. . . . .	42
3.1 - Série de Potências. . . . .	42
3.2 - Sistemas Definidos por Equações-diferença . .	45
CAP. 4 - ANÁLISE DOS RESULTADOS. . . . .	54
4.1 - Estimação . . . . .	54
4.2 - Análises. . . . .	62
APÊNDICE - PROGRAMAS EM "BASIC". . . . .	77
BIBLIOGRAFIA . . . . .	79

## CAPÍTULO 1

### INTRODUÇÃO

#### 1.1 - CONSIDERAÇÕES GERAIS

Desvendar todo o comportamento da Natureza através de modelos matemáticos seria, talvez, superestimar o poder dos métodos matemáticos ou adotar uma visão por demasiado simples dos fenômenos naturais. Entretanto, tradicionalmente, o pesquisador experimental tenta encontrar modelos que se adaptem a seus dados, com dois objetivos principais:

1. descobrir o mecanismo segundo o qual suas observações foram geradas;
2. usar o modelo para fazer inferências, previsões, correções, etc.

Há muitas situações nas quais não são válidas as su posições necessárias para uso da clássica análise de dados. Daí a vantagem de tentarmos descobrir um modelo subjacente às nossas observações para construir testes adequados, embo ra seja perfeitamente possível que tal modelo não esteja de acordo com nossa intuição ou experiência. Um exemplo de situação como a que acaba de ser descrita pode ser encontrada

na Física, onde algumas partículas se distribuem no espaço segundo a estatística de Bose-Einstein, enquanto que outras são distribuídas segundo a estatística de Fermi-Dirac, não existindo nenhuma explicação intuitiva para essa diferença de comportamento [12].

Nas ciências experimentais, como Biologia, Psicologia, Ecologia, etc., há uma grande classe de fenômenos que são medidos em escala discreta. São situações ligadas a contagens que poderiam ser, a priori, associadas a um dos clássicos modelos: Binomial, Poisson ou Binomial Negativa. No capítulo 2 definimos, caracterizamos (através de propriedades e formas limites) e apresentamos as formas mais usuais de ocorrência das variáveis aleatórias associadas a tais modelos. Para um dos itens dessa apresentação vamos necessitar das noções de distribuições compostas (misturas) e de distribuições generalizadas que achamos conveniente colocar neste capítulo introdutório, deixando para o capítulo 2 somente exemplos referentes às distribuições consideradas. Incluímos também aqui, neste capítulo, alguns comentários sobre distribuições de contágio, já que algumas distribuições descritas no capítulo seguinte são assim chamadas na literatura.

No capítulo 3 consideramos essas três distribuições como elementos de duas classes, sendo uma aquela definida por série de potências e outra por equações-diferença. Ainda no mesmo capítulo serão introduzidos critérios discriminatórios entre as referidas distribuições.

A fim de termos idéia acerca do comportamento do critério discriminatório proposto no Capítulo 3, utilizamos o método de simulação, gerando Distribuições Binomial, Poisson e Binomial Negativa, com parâmetros e tamanhos de amostra por nós fixados. Os resultados dessa simulação, assim como os problemas de estimação ligados ao critério mencionado,

são apresentados no Capítulo 4. Os programas utilizados na geração das distribuições foram construídos na linguagem "BASIC" e estão colocados no Apêndice alguns deles.

Antes de finalizarmos com aspectos sobre contágio e misturas conforme já mencionamos, queremos lembrar que a literatura é pródiga em transformações de variáveis aleatórias para posterior uso da clássica análise de dados. Grande parte dos pesquisadores está interessada em detectar o modelo correto para utilizar transformações que julgam adequadas para efetuar uma análise detalhada de seu fenômeno. Entretanto já podemos encontrar testes que podem ser aplicados diretamente aos dados: Hinz e Gurland [19] propõem uma estatística para o caso da Distribuição Binomial Negativa com o objetivo de testar uma hipótese linear geral utilizando os valores observados. Esta abordagem contribue para evitar os problemas que usualmente ocorrem em conexão com transformação de variáveis.

## 1.2 - CONTÁGIO

As situações descritas na literatura como sendo de contágio são associadas a ocorrências de eventos ligados a contagens em populações:

- i) não homogêneas, divididas em grupos, cada qual com sua própria susceptibilidade, onde a ocorrência de um evento favorável aumenta (ou diminui) a probabilidade da ocorrência de outro evento favorável, simplesmente pela maior informação fornecida pela amostragem;
- ii) homogêneas, no instante  $t = 0$ , mas o fato de aumentar (ou diminuir) a probabilidade de ocorrência de um evento favorável dado que ocorreu um evento favorável é inerente ao fenômeno observado.

Tanto i) como ii) são designadas como sendo de contágio, embora somente a segunda se refira a um contágio real, enquanto que podemos nos referir à primeira como sendo de contágio aparente. As vezes podemos ter um excelente ajuste de nossas observações a uma distribuição contagiosa embora não possamos inferir um fenômeno de contágio no mecanismo subjacente à distribuição observada. Como mencionamos no ítem i) o esquema de amostragem muitas vezes favorece um aparente contágio, como por exemplo no caso em que a distribuição de certos ovos de mosquito sobre plantação de soja é de contágio, se a unidade amostral for planta, mas poderá não sê-lo se a unidade amostral for vagem. Quando a Distribuição Binomial Negativa oferece melhor ajuste do que a Distribuição de Poisson (onde a ocorrência é aleatória), é impossível distinguir entre contágio real e aparente.

Embora a Distribuição Binomial Negativa ocorra na natureza de inúmeras maneiras, sendo muito utilizada pelos pesquisadores experimentais, temos mais três distribuições com dois parâmetros, com duas ou mais modas (a Binomial Negativa só tem uma), descritas por alguns autores, que construíram modelos matemáticos a partir de situações biológicas de seu interesse: Neyman tipo A, Thomas e Polya. A primeira delas assume a população inicial como sendo Poisson, mas a mobilidade desses indivíduos dentro da área onde o fenômeno é observado é feita a uma taxa muito lenta — apesar da característica multimodal parece que o domínio desta distribuição é mais restrito que o da Binomial Negativa. A segunda, Thomas, é descrita como tendo picos mais elevados do que a Neyman tipo A.

Neyman também derivou outras distribuições, chamadas de tipo B e C, que, segundo Feller,[11] são de natureza especializada e restrita aplicação.



Evans [10] concluiu que em situações associadas a contagens de insetos o ajuste melhor é através da Binomial Negativa, enquanto que para contagens de plantas, a Distribuição de Neyman tipo A se adapta melhor; sua justificativa se baseia na existência de competição maior no caso das plantas do que no de insetos. Entretanto Bliss [ 8 ] acha que as falhas de adaptação da Distribuição Binomial Negativa no caso de plantas se deva à ineficiência do estimador de  $k$  (um dos parâmetros da distribuição).

### 1.3 - DISTRIBUIÇÕES COMPOSTAS (MISTURAS) E GENERALIZADAS

Muitas vezes a situação analisada por um pesquisador fornece uma interpretação imediata para a noção de mistura. Entretanto, às vezes, temos que interpretá-la como um mecanismo para a construção de novas distribuições, sem uma imediata justificativa empírica.

DEFINIÇÃO - Seja  $X$  uma variável aleatória tal que a função distribuição seja

$$F(x|\theta_1, \dots, \theta_k),$$

onde  $\theta_1, \dots, \theta_k$  são parâmetros.

Se atribuirmos a alguns  $\theta$ 's (ou a todos) uma distribuição de probabilidade, então obtemos uma nova variável aleatória cuja função distribuição é dada por

$$E[F(x|\theta_1, \theta_2, \dots, \theta_k)]$$

sendo que este valor esperado é tomado com relação à distribuição conjunta dos  $\theta$ 's.

É conveniente usar a seguinte notação:

$$F_1 \underset{\theta}{\wedge} F_2 \sim F$$

onde

$F_1$  representa a distribuição original

$F_2$  representa a distribuição "misturante"

$\theta$  representa os parâmetros que variam e

$\sim$  significa equivalência de distribuições no sentido de Gurland [17], isto é se a distribuição de  $X_1$  é  $F_1(x|\alpha)$  e a de  $X_2$  é  $F_2(x|\beta)$  e se para cada  $\alpha$ , existe algum  $\beta$  (vice-versa), tal que

$$F_1(x|\alpha) = F_2(x|\beta), \quad \forall x,$$

então  $X_1$  e  $X_2$  são equivalentes e a notação é  $X_1 \sim X_2$ .

Como exemplo teríamos:

Poisson  $\wedge$  Gama  $\sim$  Binomial Negativa

DEFINIÇÃO - Sejam  $X_1$  e  $X_2$  variáveis aleatórias com funções geradoras de probabilidades  $g_1(z)$  e  $g_2(z)$  respectivamente. Seja  $X_1 \vee X_2$  a variável aleatória cuja função geradora de probabilidade é da forma  $g_1[g_2(z)]$ . Então  $X_1 \vee X_2$  é chamada variável  $X_1$  generalizada com respeito a  $X_2$ .

O seguinte teorema relaciona certa classe de distribuições compostas e generalizadas.

TEOREMA - Seja  $X_1$  uma variável aleatória cuja função geradora de probabilidade (fgp) é dada por

$$[h(z)]^\theta$$

onde  $\theta$  é um parâmetro dado.

Seja  $\theta = X_2$  uma variável aleatória com função distribuição  $F_2$  e função geradora de probabilidades  $g_2$ .

Então,  $\forall X_2$ ,

$$X_1 \wedge X_2 \sim X_2 \vee X_1$$

Exemplos:

Poisson A Poisson ~ Poisson V Poisson  
Pascal A Poisson ~ Poisson V Pascal  
Pascal A Gama ~ Gama V Pascal  
Poisson A Gama ~ Poisson V Logarítmica

É imediato observar, a partir do Teorema, que duas interpretações matemáticas diferentes podem "suportar" a mesma distribuição. Neste caso, a Distribuição Neyman Tipo A (ã qual nos referiremos no Capítulo 2) pode ser interpretada como Distribuição de Poisson Composta ou Distribuição de Poisson Generalizada.

## CAPÍTULO 2

### CARACTERIZAÇÃO DE ALGUMAS DISTRIBUIÇÕES DISCRETAS

Introduziremos neste capítulo as Distribuições Binomial, Poisson e Binomial Negativa, apresentando suas principais características, algumas das quais serão úteis ao desenvolvimento deste trabalho, enquanto que outras servirão como referência ao leitor. Estas distribuições são membros da classe das distribuições generalizadas em série de potências, mas este aspecto será abordado no próximo capítulo.

#### 2.1 - DISTRIBUIÇÃO BINOMIAL

A Distribuição Binomial foi derivada a partir do início do século XVIII, através de J. Bernoulli. Além de servir de modelo matemático a algumas situações práticas, sua importância também se deve ao fato de sua conexão com uma grande variedade de distribuições conhecidas.

##### 2.1.1 - DEFINIÇÃO

Diz-se que uma variável aleatória  $X$ , com valores inteiros não negativos tem Distribuição Binomial com parâmetros  $n$  e  $p$  se

$$(2.1) \quad P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k=0,1,2,\dots,n$$

onde  $n \geq 1$  e  $0 \leq p \leq 1$ . Usualmente coloca-se  $1-p = q$ .

Vamos supor que num experimento s3 podemos observar sucesso ou fracasso, associados, respectivamente, a ser observada ou n3o certa caracter3stica de uma popula33o. Se este experimento for realizado mediante  $n$  tentativas independentes, tal que permane3a constante a probabilidade  $p$  de sucesso, ent3o o n3mero de sucessos observados, para  $n$  fixado a priori, tem Distribui33o Binomial com par3metros  $n$  e  $p$ , como em (2.1).

Muitas situa33es pr3ticas podem ser adaptadas a este modelo, da3 o grande uso da Binomial. Al3m disso, o seu uso se torna ainda maior se consideradas as formas limites.

Algumas vezes nos referiremos 3 probabilidade definida em (2.1) com a notaa3o  $b(k;n,p)$ , ou  $B(n,p)$ .

### 2.1.2 - CARACTERIZA33O

A fun33o geradora de probabilidades 3 da forma:

$$(2.2) \quad P(s) = (q+ps)^n,$$

pois

$$P(s) = \sum_{k=0}^{\infty} \binom{n}{k} (ps)^k (1-p)^{n-k} = [(1-p)+ps]^n$$

Se, na express3o (2.2), colocamos  $s = e^t$  obteremos a fun33o geradora de momentos e para  $s = e^{it}$  obteremos a fun33o caracter3stica, ou seja

$$M_t = (q+pe^t)^n$$

$$\phi_t = (q + pe^{it})^n$$

Para a função geradora de cumulantes, basta considerar o logaritmo de  $\phi_t$ .

Usando as relações

$$E(X) = P'(1) \quad \text{e} \quad V(X) = P''(1) + P'(1) - P'^2(1)$$

podemos verificar que

$$E(X) = np \quad \text{e} \quad V(X) = np(1-p).$$

Para futura referência observamos que, na Distribuição Binomial temos  $V(X) \leq E(X)$ , com a igualdade ocorrendo apenas no caso trivial ( $X$ , constante)  $p = 0$ .

- Relações de recorrência entre momentos

Se considerarmos a função característica da variável aleatória  $(X-np)$ , temos

$$\phi_t = e^{-npe^{it}} (q+pe^{it})^n.$$

Fazendo-se  $it = \theta$  e diferenciando a expressão anterior com relação a  $\theta$ , obtém-se

$$(2.3) \quad \mu_r = npq \sum_{j=0}^{r-2} \binom{r-1}{j} \mu_j - p \sum_{j=0}^{r-2} \binom{r-1}{j} \mu_{j+1}$$

e

$$(2.4) \quad \mu_{r+1} = pq(nr\mu_{r-1} + \frac{\partial \mu_r}{\partial p}).$$

Obs: Qualquer momento inverso da distribuição binomial, isto é  $E(X^{-r})$  com  $r > 0$ , é infinito porque  $P(X=0) > 0$ .

- Reprodutividade

Se  $X_1$  e  $X_2$  são variáveis aleatórias independentes, com Distribuição Binomial com parâmetros  $(n_1, p)$  e  $(n_2, p)$ , respectivamente, então  $X_1 + X_2$  tem distribuição binomial com parâmetros  $(n_1 + n_2, p)$ . Em outras palavras, a Distribuição Binomial é reprodutiva com relação a  $n$ .

- Distribuição Condicionais

A distribuição condicional de  $X_1$  dada a soma  $X_1 + X_2$  é da forma

$$(2.5) \quad P(X_1=k | X_1+X_2=x) = \frac{\binom{n_1}{k} \binom{n_2}{x-k}}{\binom{n_1+n_2}{x}}$$

para  $\max(0, x-n_2) \leq k \leq \min(n_1, x)$ , expressão que representa a distribuição hipergeométrica.

Com relação à variável aleatória  $X_1 - X_2$  temos dois casos a considerar:

i)  $p = \frac{1}{2}$ ; então

$$P(X_1 - X_2 = k) = \binom{n_1+n_2}{n_2+k} / 2^{n_1+n_2} \quad (-n_2 \leq k \leq n_1)$$

o que é uma forma mais geral de (2.1) se substituímos  $k$  por  $(a+bk)$  com  $b \neq 0$  no termo direito de (2.1) [21].

ii)  $p \neq \frac{1}{2}$ ; neste caso a distribuição exata é bastante complicada e, na prática, utiliza-se a aproximação pela Distribuição Normal.

Se em vez de  $(n_1, p)$  e  $(n_2, p)$  serem respectivamente os parâmetros das variáveis aleatórias  $X_1$  e  $X_2$  tivermos  $(n_1, p_1)$  e  $(n_2, p_2)$ , então a distribuição condicional (2.5) não mais será hipergeométrica e em [18] foi obtido um resul

tado assintoticamente normal.

- Moda da distribuição

Se  $p = \frac{1}{2}$  a distribuição é simétrica.

De (2.1) observa-se que:

$$\frac{P(X=k+1)}{P(X=k)} = \frac{p}{q} \frac{(n-k)}{(k+1)}$$

e este quociente será maior do que a unidade para  
 $k < p(n+1) - 1$ .

Então, à medida que  $k$  cresce,  $P(X=k)$  cresce e depois decresce, sendo que o máximo é atingido num inteiro  $k$ , satisfazendo

$$(n+1)p - 1 < k \leq (n+1)p.$$

Se  $(n+1)p$  for inteiro, então  $P(X=(n+1)p-1) = P(X=(n+1)p)$ .

### 2.1.3 - FORMAS LIMITES

Daremos a seguir algumas aproximações que se relacionam com a Distribuição Binomial. Os resultados serão enunciados da maneira usual, ou seja, como aparecem em textos elementares ou como utilizados na prática. Entretanto deve-se ter em mente que todos eles se referem a convergência fraca de seqüências de distribuições. Isto é válido para todos os itens relativos a formas limites que aparecem no capítulo.

#### i) Distribuição Normal

Se uma variável aleatória  $X$  tiver Distribuição Binomial com parâmetros  $n$  e  $p$ , já vimos que  $E(X) = np$  e  $V(X) = npq$ . Nessas condições, se  $n \rightarrow \infty$ , a variável aleatória

$$Z = \frac{X - np + \frac{1}{2}}{\sqrt{npq}}$$



tenderá à Distribuição Normal Padrão. Isto é equivalente a

$$\lim_{n \rightarrow \infty} P[Z \leq z] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-\frac{1}{2}t^2} dt.$$

Raff [32] mostrou que se  $np^{3/2} > 1,07$  o erro associado à aproximação à normal nunca excede 0,05, qualquer que seja  $z$ .

Em [2] e [5] há outras transformações de uso corrente em Análise de Variância.

ii) Distribuição Hipergeométrica

Uma variável aleatória  $X$  tem Distribuição Hipergeométrica quando

$$(2.6) \quad P(X=k) = \frac{\binom{Np}{k} \binom{Nq}{n-k}}{\binom{N}{n}}$$

onde

$N$ : tamanho da população

$n$ : tamanho da amostra

$p$ : probabilidade de sucesso.

A Distribuição Binomial é uma forma limite da Distribuição Hipergeométrica quando  $N \rightarrow \infty$ .

Podemos reescrever (2.6) do seguinte modo:

$$P(X=k) = \binom{n}{k} \frac{p(p - \frac{1}{N}) \dots (p - \frac{k-1}{N}) q(q - \frac{1}{N}) \dots (q - \frac{n-k-1}{N})}{(1 - \frac{1}{N}) \dots (1 - \frac{n-1}{N})}$$

É imediato verificar que

$$\lim_{n \rightarrow \infty} P(X=k) = \binom{n}{k} p^k q^{n-k}.$$

iii) Distribuição de Poisson

Em (2.2) se  $n \rightarrow \infty$  e  $p$  for suficientemente pequeno, isto é,  $p \rightarrow 0$ , tal que o produto  $np \rightarrow \lambda$  (constante) observe-se:

$$\lim_{\substack{n \rightarrow \infty \\ np = \lambda}} (q+ps)^n = \lim_{\substack{n \rightarrow \infty \\ np = \lambda}} \left( 1 + \frac{\lambda(s-1)}{n} \right)^n = e^{\lambda(s-1)}$$

Entretanto esta é justamente a função geradora de probabilidades de uma Distribuição de Poisson [ver pág. 18]. Como há correspondência bi-unívoca entre função distribuição e função geradora de probabilidade (se esta última existir), concluímos, usando o Teorema da Continuidade, que a Distribuição de Poisson é a forma limite da Distribuição Binomial nas condições mencionadas.

A função Distribuição Acumulada da Binomial pode ser escrita como

$$B(k;n,p) = \sum_{x=0}^k \binom{n}{x} p^x q^{n-x} = \sum_{x=0}^k b(x;n,p) \quad k=0,1,2,\dots$$

Pelo exposto acima, tal função será aproximada pela função Distribuição Acumulada da Poisson, isto é,

$$P(k;\lambda) = \sum_{x=0}^k \frac{e^{-\lambda} \lambda^x}{x!} = \sum_{x=0}^k p(x;\lambda) \quad k=0,1,2,\dots,$$

para  $\lambda = np$ .

Pode ser verificado [ 1 ] que  $P(k;\lambda) - B(k;n,p)$ , isto é, o erro associado à aproximação, é positivo se  $k < np - \frac{np}{n+1}$  e é negativo para  $np \leq k$ . Sob o ponto de vista prático, isto equivale a dizer que a Distribuição de Poisson superestima as probabilidades associadas às caudas. Esta informação é útil quando for necessária calcular a probabilidade as

sociada ao erro do tipo I num processo de decisão.

Além disso, pode ser verificado [13] que

$$(2.7) \quad \sum_{x=0}^{\infty} |b(x;n,p) - p(x;np)| \leq 2np^2$$

[isto implica em

$$\sum_{k=0}^{\infty} |b(k;n,p) - p(k;np)| \rightarrow 0 \text{ se } np = \lambda \text{ (constante) e } p \rightarrow 0].$$

Em [13] há referência a outros resultados ligados ao valor da soma em (2.7).

Esta aproximação pela Distribuição de Poisson pres<sup>u</sup> supõe tentativas independentes e probabilidade de sucesso constante - melhor dizendo, estas são as próprias suposições inerentes ao modelo binomial. Entretanto este resultado foi estendido para as seguintes situações:

- i) tentativas independentes com diferentes probabilidades de sucesso [25] e
- ii) algumas formas de dependência [34], com grande aplicação em pesquisa operacional.

#### 2.1.4 - MISTURAS

No capítulo anterior demos a noção de mistura de distribuições, que iremos agora exemplificar através da Distribuição Binomial.

- i) Binomial (n,P)  $\wedge$   $\frac{\Lambda}{p}$  Beta

Esta é a chamada Distribuição Polya - Eggenberger, e Ishii e Hayakawa [20] trataram dessa mistura considerando para P a Distribuição Beta com a justificativa de que a mesma é flexível no intervalo (0,1).

Então, se

$$P(X=k) = \binom{n}{p} P^k (1-P)^{n-k}, \quad k=0,1,\dots,n$$

e

$$f(p) = \frac{1}{B(a,b)} p^{a-1} (1-p)^{b-1}, \quad 0 \leq p \leq 1$$

tal processo acarretaria uma distribuição da forma

$$P(X=k) = \frac{B(a+k, n+b-k)}{(n+1)B(a,b)B(1+k, n+1-k)}, \quad k=0,1,\dots,n$$

onde

$$B(x,y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}, \quad x>0, \quad y>0.$$

Esta distribuição é considerada como sendo realmente de contágio, (ver comentário à página 3) e se adapta a modelos de doenças infecciosas, onde o aparecimento de um caso aumenta a probabilidade de infecção pela doença.

ii) Binomial  $(N,p)$   $\overset{\Lambda}{N/n}$  Poisson  $(\lambda)$

Esta distribuição é comumente chamada na literatura de Poisson - Binomial; [não deve ser confundida com a mistura Poisson  $(\lambda)$   $\overset{\Lambda}{\lambda/\phi}$  Binomial  $(n,p)$ ]. Certos modelos associados a contagem, notadamente em Biologia, em que as distribuições discretas clássicas não oferecem bons ajustes, são convenientemente adaptados a esta distribuição, apesar dos problemas relativos a cálculos por ela apresentados. Em [26] podem ser encontradas algumas aplicações.

Sua função geradora de probabilidade é da forma:

$$P(s) = e^{\lambda[(q+pt)^n - 1]}$$

e a distribuição é definida por

$$P(X=k) = \frac{e^{-\lambda} (p/q)^k}{k!} \sum_{j \geq \frac{k}{n}} (nj)^{(k)} [(\lambda q^n)^j / j!]$$

onde  $k=0,1,2,\dots$

No particular caso em que  $n=1$ , temos que

$$P(X=k) = \frac{e^{-\lambda p} (\lambda p)^k}{k!}, \quad k=0,1,2,\dots$$

expressão que corresponde à Distribuição de Poisson com parâmetro  $\lambda p$  (ver (2.8) ). Para este último caso, Fisz [ 15 ] usa o termo Distribuição Binomial Composta.

## 2.2 - DISTRIBUIÇÃO DE POISSON

Esta distribuição foi inicialmente abordada como forma limite da Distribuição Binomial. Entretanto descreve também eventos aleatórios e independentes no tempo. Como um campo de grande aplicação temos a ecologia: em muitas espécies de plantas e animais o número de descendentes em potencial por indivíduo é bem grande e a chance dos indivíduos sobreviverem é bem pequena; uma vez que os descendentes se desenvolvem independentemente uns dos outros e que o número esperado por família não flutua muito entre famílias, é razoável assumir que o número real de descendentes numa família siga o modelo de Poisson. Mesmo quando as premissas não estão satisfeitas, este procedimento é utilizado numa primeira aproximação, devido à sua grande simplicidade.

### 2.2.1 - DEFINIÇÃO

Diz-se que uma variável aleatória  $X$  que toma valores nos inteiros não negativos tem Distribuição de Poisson com parâmetro  $\lambda > 0$ , se

$$(2.8) \quad P(X=k) = \frac{e^{-\lambda} \lambda^k}{k!}, \quad k=0,1,2,\dots$$

### 2.2.2 - CARACTERIZAÇÃO

A função geradora de probabilidades é dada por

$$(2.9) \quad P(s) = \sum_{k=0}^{\infty} p_k s^k = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda s)^k}{k!}$$

$$\therefore P(s) = e^{\lambda(s-1)}$$

A função geradora de momentos e a função característica são dadas, respectivamente, pelas expressões:

$$M_t = E(e^{tX}) = e^{\lambda(e^t-1)}$$

$$Q_t = E(e^{itX}) = e^{\lambda(e^{it}-1)}$$

Além disso, a função geradora de cumulantes, é da forma:

$$\psi_t = \ln \phi_t = \lambda(e^{it}-1),$$

donde se conclui que todos os cumulantes são iguais a  $\lambda$ .

A partir da função geradora de probabilidades podemos calcular:

$$E(X) = P'(1) = \lambda$$

$$V(X) = P''(1) + P'(1) - P'^2(1) = \lambda.$$

Então, para a Distribuição de Poisson, média e variância coincidem, fato que sugere critérios para detectar ausência de aleatoriedade.

Considerando-se em (2.3) e (2.4) o limite para  $n \rightarrow \infty$ , vem:

$$(2.10) \quad \mu_r = \lambda \sum_{j=0}^{r-2} \binom{r-1}{j} \mu_j$$

$$(2.11) \quad \mu_{r+1} = r\lambda\mu_{r-1} + \frac{\lambda\partial\mu_r}{\partial\lambda}$$

- Reprodutividade

É fácil verificar, através da função geradora de probabilidades, que a Distribuição de Poisson é reprodutiva com relação ao parâmetro  $\lambda$ , isto é, se  $X_1$  e  $X_2$  forem duas variáveis aleatórias independentes, com Distribuição de Poisson, com parâmetros  $\lambda_1$  e  $\lambda_2$ , respectivamente, então  $X_1 + X_2$  terá também Distribuição de Poisson com parâmetro  $\lambda_1 + \lambda_2$ . A recíproca é verdadeira [21]. Este resultado pode ser estendido a um número finito arbitrário de variáveis aleatórias independentes.

- Distribuição Condicional

Se  $X_1$  e  $X_2$  forem variáveis aleatórias independentes, cada uma com Distribuição de Poisson com parâmetros  $\lambda_1$  e  $\lambda_2$ , respectivamente, vem:

$$P(X_1=x_1 | X_1+X_2=x) = \binom{x}{x_1} \left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^{x_1} \left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)^{x-x_1}$$

isto é, a distribuição condicional de  $X_1$  dada a soma  $X_1 + X_2$  é Binomial com parâmetros

$$x \text{ e } \frac{\lambda_1}{\lambda_1+\lambda_2}.$$

- Moda da Distribuição

De (2.8) observa-se que:

$$\frac{P(X=k+1)}{P(X=k)} = \frac{\lambda}{k+1}$$

Então  $P(X=k+1) > P(X=k)$  se  $k < \lambda - 1$ . Nessas condições,  $P(X=k)$  cresce com  $k$  até um valor máximo em  $k = [\lambda]$  e depois decresce quando  $k$  cresce.

### 2.2.3 - FORMAS LIMITES E APROXIMAÇÕES

i) Distribuição Binomial

Já tratada à página 14

ii) Distribuição Normal

Se X for uma variável aleatória com distribuição de Poisson com parâmetro  $\lambda$ , diz-se que a variável

$$Z = (X - \lambda + \frac{1}{2}) \lambda^{-1/2}$$

tem distribuição de Poisson padrão. Neste caso,

$$\lim_{\lambda \rightarrow \infty} P(Z \leq z) = (2\pi)^{-1/2} \int_{-\infty}^z e^{-u^2/2} du$$

Cabe aqui a observação final do item i) da secção 2.1.3.

iii) Distribuição Binomial Negativa

A convergência da Binomial Negativa para Poisson encontra-se à página 33.

iv) Relação da Distribuição de Poisson com  $\chi^2$  (central)

A função de Distribuição Acumulada de  $\chi^2$  pode ser usada como aproximação para a Distribuição de Poisson (e vice-versa). De fato, é possível mostrar que se a variável aleatória X tem distribuição de Poisson com parâmetro  $\lambda$ , podemos escrever:

$$P(X \leq k) \equiv P(\chi^2_{2(k+1)} > 2\lambda).$$

### 2.2.4 - MISTURAS

Considerando a noção de mistura vista à página 5 daremos a seguir, sem entrar em muitos detalhes, algumas distribuições decorrentes da atribuição de uma dada distribui-



ção ao parâmetro  $\lambda$  da Distribuição de Poisson.

i) Poisson ( $\lambda$ )  $\overset{\Lambda}{\lambda}$  Gama ( $\alpha, \beta$ )

A distribuição gerada é uma Binomial Negativa e a demonstração pode ser encontrada à página 16

ii) Poisson ( $\lambda$ )  $\overset{\Lambda}{\lambda}$  Retangular ( $a, b$ )

Apresentada como possível alternativa ao caso i), foi usada no estudo de repetidos acidentes ou de reincidência de determinada moléstia quando a susceptibilidade para acidentes, doenças, etc., varia de indivíduo para indivíduo [ 6 ].

Neste caso,

$$P(X=k) = [k!(b-a)]^{-1} \int_a^b \lambda^k e^{-\lambda} d\lambda, \quad k=0,1,2,\dots$$

iii) Poisson ( $\lambda$ )  $\overset{\Lambda}{\lambda}$  Normal Truncada ( $\xi, \sigma$ )

A distribuição decorrente desta mistura é também chamada Poisson-Normal ou Gauss-Poisson. O truncamento é usualmente no ponto 0, assegurando que  $\lambda$  tem probabilidade zero de ser negativo. Convém notar que  $\xi$  e  $\sigma$  são os parâmetros da Distribuição Normal antes do truncamento.

Tem-se, então:

$$P(X=r) = \frac{\exp[\frac{1}{2}\sigma^2 - \xi]}{r!} \cdot \frac{\int_0^{\infty} \lambda^r \exp[-\frac{1}{2}\left(\frac{\lambda - \xi + \sigma^2}{\sigma}\right)^2] d\lambda}{\int_0^{\infty} \exp[-\frac{1}{2}\left(\frac{\lambda - \xi}{\sigma}\right)^2] d\lambda}$$

iv) Poisson ( $\lambda$ )  $\overset{\Lambda}{\lambda}$  Log Normal ( $\xi, \sigma, a$ )

Esta distribuição é também denominada Lognormal Discreta. Neste caso os dados teriam que sofrer uma transformação logarítmica e mesmo assim, dado o fato de que a forma da distribuição apresenta uma integral não tabelada, não se vê

muita vantagem em usá-la.

v) Poisson ( $\lambda$ )  $\overset{\Lambda}{\lambda/\phi}$  Poisson ( $\theta$ )

É a chamada Neyman tipo A e tradicionalmente encarada como distribuição de contágio. Encontramos porém alguns autores que preferem usar o termo heterogeneidade porque preferem considerar o verdadeiro contágio em situações em que eventos em observação dependem do modelo de ocorrência de eventos anteriores. Inicialmente foi usada para descrever um modelo de distribuição de larvas numa região escolhida aleatoriamente numa área [27].

Para  $r > 0$ , temos:

$$P(X=r) = \sum_{j=1}^{\infty} e^{-\lambda} (\lambda^j / j!) e^{-j\phi} [(j\lambda)^r / r!]$$

Para  $r = 0$ ,

$$P(X=0) = \exp[-\theta(1-e^{-\phi})]$$

Esta distribuição de Neyman do Tipo A é baseada em um modelo preciso e se as suposições biológicas inerentes à distribuição não forem satisfeitas a dispersão da população não será adequadamente descrita por tal distribuição. Neste caso, aconselha-se adaptar a Distribuição Binomial Negativa, que será tratada a partir da secção 2.3.

#### 2.2.5 - PROCESSO DE POISSON

Este processo é utilizado para situações de desintegração de partículas, chegada de chamadas telefônicas, quebras de cromossomas sob determinadas condições, etc.

Um Processo de Poisson é aquele onde os tempos decorridos entre sucessivos eventos têm a mesma Distribuição Exponencial [isto é,  $f(x) = \lambda e^{-\lambda x}$ ,  $x > 0$ ] e são independentes; nessas condições o número de eventos num intervalo específico

tem Distribuição de Poisson. No caso das mencionadas chamadas telefônicas suponha que existe um grande número de linhas disponíveis e que o tempo de espera  $t$  (duração da chamada) para cada chamada tenha Distribuição Exponencial, e que as chegadas das chamadas sigam um Processo de Poisson com  $\lambda$  sendo o número médio de chamadas por unidade de tempo.

Então  $P_N$ , a probabilidade de exatamente  $N$  canais serem ocupados é da forma

$$P_N = \frac{e^{-\lambda} \lambda^N}{N!}, \quad N=0,1,2,\dots$$

Esta última equação segue da solução de equações diferenciais associadas ao processo.

### 2.3 - DISTRIBUIÇÃO BINOMIAL NEGATIVA

Esta distribuição, muito utilizada em Ciências Biológicas, Psicologia, etc., é adequada a muitas situações, entre elas, quando

- não estão satisfeitos os requisitos de aleatoriedade para adaptação da Distribuição de Poisson;
- os modelos seriam mais convenientemente descritos como misturas de certas distribuições;
- o esquema de amostragem facilita a adaptação (amostragem inversa).

#### 2.3.1 - DEFINIÇÃO

Diz-se que uma variável aleatória  $X$ , que toma valores nos inteiros não negativos, tem Distribuição Binomial Negativa com parâmetros  $k$  e  $p$ , se

$$(2.12) \quad P(X=r) = \binom{r+k-1}{r} p^k q^r, \quad r=0,1,2,\dots$$

$k > 0$   
 $0 < p < 1$   
 $q = 1 - p$

Quando  $k$  for inteiro,  $X$  representa o número de falhas que antecedem o  $k$ -ésimo sucesso, numa sequência infinita de ensaios de Bernoulli, com probabilidade  $p$  de sucesso. Neste caso diz-se que a variável aleatória  $X$  tem Distribuição de Pascal. Em particular, se  $k = 1$ , a distribuição de  $X$  toma a forma

$$(2.13) \quad P(X=r) = pq^r, \quad r=0,1,2,\dots$$

que é a conhecida Distribuição Geométrica, interpretada como sendo o número de falhas que antecedem o 1º sucesso.

Se expandirmos a expressão  $(\frac{1}{p} - \frac{q}{p})^{-k}$ , o  $(r+1)$ -ésimo termo da expansão será da forma dada em (2.12). Entretanto, muitos autores como Fisher, R., Johnson, N., etc. preferem considerar tal termo usando a seguinte notação

$$(2.14) \quad P(X=r) = \binom{k+r-1}{r} \left(\frac{1}{Q}\right)^k \left(\frac{P}{Q}\right)^r,$$

proveniente da expansão de  $(Q-P)^{-k}$ . Através das expressões consideradas acima, parece natural o uso do nome de Binomial Negativa.

Apesar da forma (2.14) ser encontrada em muitos artigos, usaremos, neste texto, a expressão definida em (2.12).

### 2.3.2 - Caracterização

A função geradora de probabilidades pode ser determinada como segue:

$$P(s) = \sum_{r=0}^{\infty} (-1)^r \binom{-k}{r} p^k (qs)^r = p^k \sum_{r=0}^{\infty} \binom{-k}{r} (-qs)^r$$

Logo,

$$(2.15) \quad P(s) = \left(\frac{p}{1-qs}\right)^k$$

Obs.: a identidade usada acima envolvendo coeficientes binomiais é válida para  $k > 0$  [12].

Se fizermos  $s = e^t$  e  $s = e^{it}$  em (2.15), obteremos, respectivamente, as funções geradoras de momentos e característica, isto é,

$$M_t = \left( \frac{p}{1-qe^t} \right)^k$$

$$\phi_t = \left( \frac{p}{1-qe^{it}} \right)^k$$

Tomando-se o logaritmo de  $\phi_t$  teremos a função geradora de cumulantes. Utilizando as igualdades

$$E(X) = P'(1) \text{ e } V(X) = P''(1) + P'(1) + P'^2(1)$$

obtemos

$$E(X) = \frac{kq}{p} \text{ e } V(X) = \frac{qk}{p^2}.$$

Observa-se que  $V(X) > E(X)$ , característica das situações descritas na literatura como de contágio.

- Moda da Distribuição

De (2.12) podemos observar que:

$$\frac{P(X=r+1)}{P(X=r)} = \frac{k(1-p)+r(1-p)}{r+1}$$

Então

$$P(X=r+1) > P(X=r) \text{ se } r < \frac{k-1}{p} - k$$

e

$$P(X=r+1) < P(X=r) \text{ se } r > \frac{k-1}{p} - k$$

Nessas condições, temos:

i) se  $\frac{k-1}{p} < k$ , teremos a moda no valor  $r = 0$ ; observa-se que para  $0 < k \leq 1$ , qualquer que seja  $p$ , a moda é sempre zero; obviamente, considerando-se o caso degenerado onde  $p = 1$ , teremos, qualquer que seja  $k$ , a moda também no ponto zero;

ii) se  $r = \frac{k-1}{p} - k$ , teremos duas modas, uma em  $r$  e outra em  $r+1$ ;

iii) se  $\frac{k-1}{p} > k$ , haverá uma só moda cujo valor será  $[\frac{k-1}{p} - k]$ .

### 2.3.3 - FORMAS DE OCORRÊNCIA

i) Suponhamos um grande número de pessoas expostas ao mesmo risco com igual susceptibilidade. Se número de acidentes num dado intervalo de tempo seguir uma Distribuição de Poisson, a probabilidade de uma pessoa ter  $r$  acidentes é dada por

$$P(X=r) = \frac{e^{-\lambda} \lambda^r}{r!} \quad r=0,1,2,\dots$$

onde  $\lambda$  é o número médio de acidentes por pessoa por intervalo de tempo.

Se as pessoas não forem igualmente susceptíveis a acidentes  $\lambda$  será uma variável aleatória; suponhamos que sua densidade seja da forma

$$f(\lambda) = \begin{cases} \frac{a^k}{\Gamma(k)} \lambda^{k-1} e^{-a\lambda} & \text{se } \lambda > 0 \\ 0 & \lambda \leq 0 \end{cases}$$

onde  $k > 0$  e  $a > 0$ .

Neste caso temos uma mistura, de acordo com o descrito à página 5, com  $F_1$  sendo a Distribuição de Poisson com parâmetro  $\lambda$  e  $F_2$  sendo uma Distribuição Gama, com parâ-

metros  $\underline{a}$  e  $\underline{k}$ .

Então:

$$P(X=r) = \binom{r+k-1}{r} \left(\frac{a}{1+a}\right)^k \left(\frac{1}{1+a}\right)^r$$

que é a expressão da Distribuição Binomial Negativa.

Daremos a prova a seguir:

Seja  $(X, \lambda)$  uma variável aleatória bidimensional

$$P(X=r, \lambda) = f(\lambda)P(X=r|\lambda)$$

Como nosso interesse é a marginal  $P(X=r)$ , segue:

$$P(X=r) = \int_0^{\infty} P(X=r|\lambda) f(\lambda) d\lambda = \frac{a^k}{r! \Gamma(k)} \int_0^{\infty} \lambda^{r+k-1} e^{-\lambda(a+1)} d\lambda$$

Como

$$\frac{\Gamma(p)}{a^p} = \int_0^{\infty} y^{p-1} e^{-ay} dy,$$

segue

$$P(X=r) = k(k+1)\dots(k+r-2)(k+r-1) \cdot \frac{1}{r!} \left(\frac{a}{1+a}\right)^k \left(\frac{1}{1+a}\right)^r$$

Lembrando que, para  $k > 0$ , vale a expressão

$$\binom{k}{r} = \frac{k(k-1)(k-2)\dots(k-r+1)}{r!},$$

segue

$$P(X=r) = (-1)^r \binom{-k}{r} \left(\frac{a}{1+a}\right)^k \left(\frac{1}{1+a}\right)^r$$

e, de acordo com [12],

$$P(X=r) = \binom{r+k-1}{r} \left(\frac{a}{1+a}\right)^k \left(\frac{1}{1+a}\right)^r$$

sendo que  $p = \frac{a}{1+a}$  e  $q = \frac{1}{1+a}$ .

ii) Existem casos onde o tamanho da amostra não é fixado a priori. Suponhamos que uma proporção de indivíduos numa população possui certa característica. Se necessitamos de  $k$  indivíduos da população com a citada característica, a variável  $X$  que denota o número de falhas  $r$  até a obtenção dos  $k$  sucessos tem Distribuição Binomial Negativa. Este tipo de amostragem (conhecido como amostragem inversa) é usado comumente no caso de estimação do tamanho da população através de captura e recaptura de indivíduos.

Verifica-se a afirmação acima usando o seguinte argumento:

- a probabilidade de  $(k-1)$  sucessos nas primeiras  $(n-1)$  tentativas é da forma  $\binom{n-1}{k-1} p^{k-1} q^{n-k}$ ;
  - a probabilidade de sucesso na próxima tentativa é  $p$ ;
- então

$$(2.16) \quad P(X=r) = \binom{n-1}{k-1} p^k q^{n-k}.$$

Esta expressão é equivalente à forma (2.12).

iii) Suponhamos que colônias ou grupos de indivíduos se distribuem aleatoriamente numa área. Consideremos a variável aleatória  $X$  (número de colônias observadas numa certa área) como tendo Distribuição de Poisson

$$P(X=r) = \frac{e^{-\lambda} \lambda^r}{r!}, \quad r=0,1,2,\dots$$

Se a variável aleatória  $Z$  (número de indivíduos por colônia) possuir Distribuição Logarítmica então a variável



vel Y (número de indivíduos observados) terá Distribuição Binomial Negativa.

De fato,

$$P(X=r) = \frac{e^{-\lambda} \lambda^r}{r!}, \quad r=0,1,2,\dots$$

$$P(Y=s) = \sum_{r=0}^{\infty} P(X=r)P(Y=s|X=r)$$

Como Z tem Distribuição Logarítmica, então

$$P(Z=s) = \alpha \cdot \frac{x^s}{s}$$

o que é equivalente ao coeficiente de  $t^s$  em  $[-\alpha \ln(1-xt)]$ .

Generalizando, observar  $s$  indivíduos em  $r$  grupos será o coeficiente de  $t^s$  em  $[-\alpha \ln(1-xt)]^r$  onde  $\alpha = \frac{-1}{\ln(1-x)}$

$$\begin{aligned} \therefore P(Y=s) &= \sum_{r=0}^{\infty} \frac{e^{-\lambda} \lambda^r}{r!} \cdot \text{coef. de } t^s \text{ em } [-\alpha \ln(1-xt)]^r \\ &= \text{coef. de } t^s \text{ em } \sum_{r=0}^{\infty} \frac{e^{-\lambda} \lambda^r}{r!} (-\alpha \ln(1-xt))^r \\ &= \text{coef. de } t^s \text{ em } e^{-\lambda(1+\alpha \ln(1-xt))} \\ &= e^{-\lambda} \text{ coef. de } t^s \text{ em } (1-xt)^{-\lambda\alpha} \\ &= (1-x)^{\lambda\alpha} \cdot \text{coef. de } t^s \text{ em } (1-xt)^{-\lambda\alpha}. \end{aligned}$$

O  $(s+1)$ -ésimo termo da expansão de  $(1-xt)^{-\alpha\lambda}$  é,

$$\binom{\lambda\alpha+s-1}{\lambda\alpha-1} (xt)^s$$

$$\therefore P(Y=s) = (1-x)^{\lambda\alpha} \binom{\lambda\alpha+s-1}{\lambda\alpha-1} x^s$$

o que equivale a:

$$P(Y=s) = \binom{s+\lambda\alpha-1}{s} (1-x)^{\lambda\alpha} (x)^s$$

que é a expressão da Distribuição Binomial Negativa com

$$k = \lambda\alpha \quad e \quad p = 1-x.$$

iv) Em muitos modelos biológicos podemos notar como característica uma desusada dispersão da população e a Distribuição Binomial Negativa fornece uma boa adaptação. Às vezes na literatura estes casos são considerados como modelos de contágio. Por exemplo, é do interesse dos ecologistas estimar a abundância de uma dada espécie numa comunidade de plantas; o método consiste em dividir aleatoriamente a área em quadrados. Então, com repetidas amostras, podemos ter o número de quadrados com  $k$  plantas e a média desta distribuição nos dá uma estimativa da densidade da espécie. Em muitos casos, a distribuição de plantas por quadrado se adapta muito bem o modelo da Binomial Negativa.

v) Processo de Yule — Processo Puro de Nascimento

Seja uma população onde novos membros são gerados mas nenhum morre. Suponha que cada membro tenha probabilidade de  $\lambda h + o(h)$  de gerar outro (onde  $h$  é um intervalo de tempo no qual o indivíduo é gerado) e  $\lambda$  é o coeficiente de crescimento da população. Seja  $n$  o tamanho da população no instante  $t$ . A probabilidade  $P_n(t)$  de que o tamanho da população seja exatamente  $n$  é dada pela expressão

$$P_n(t) = \binom{n-1}{n-i} e^{-i\lambda t} (1-e^{-\lambda t})^{n-i}$$

que equivale ao modelo da Binomial Negativa descrito inicialmente com  $r = n - i$ ,  $k = i$  e  $p = e^{-\lambda t}$ .

vi) Urna de Polya

Consideremos uma urna com  $b$  bolas pretas e  $r$  brancas. Retira-se uma bola da urna e anota-se sua cor. A bola é recolocada e  $c$  bolas da mesma cor são adicionadas na urna. Faz-se uma retirada novamente. É preciso especificar algum critério de parada, como, por exemplo, parar até que  $n$  retiradas tenham sido feitas, ou até que  $k$  bolas pretas tenham aparecido.

Temos aqui um modelo bem simples de um fenômeno que poderíamos chamar de "contágio", onde cada ocorrência aumenta a probabilidade de futuras ocorrências.

Pode ser verificado que a probabilidade  $p_{k,n}$  de que foram necessárias  $n$  tentativas para o aparecimento de  $k$  bolas pretas, segue a expressão

$$p_{k,n} = \frac{\binom{-b/c}{k} \binom{-r/c}{n-k}}{\binom{-(b+r)/c}{n}}$$

Seja  $\frac{b}{b+r} = p$ ,  $\frac{r}{b+r} = q$  e  $\frac{c}{b+r} = \gamma$ .

Então

$$p_{k,n} = \frac{\binom{-p/\gamma}{k} \binom{-q/\gamma}{n-k}}{\binom{-1/\gamma}{n}}$$

Se  $n \rightarrow \infty$ ,  $p \rightarrow 0$ ,  $\gamma \rightarrow 0$ ,  $|np + \lambda|$  e  $n\gamma \rightarrow \frac{1}{\rho}$ , então para  $k$  fixado

$$p_{k,n} \rightarrow \binom{\lambda\rho + k - 1}{k} \left(\frac{\rho}{1+\rho}\right)^{\lambda\rho} \left(\frac{1}{1+\rho}\right)^k.$$

Observamos então que esta é a expressão da Distri-

ção Binomial Negativa, com parâmetros  $k$  e  $\frac{1}{1+\rho}$ .

vii) A distribuição de quebras (divisões) de cromossomas por célula, quando o tecido é submetido a irradiações segue uma Distribuição Binomial Negativa. Entretanto como tais quebras (divisões) só podem ocorrer naquelas células que estão num estágio particular do ciclo mitótico na época da irradiação, as não suscetíveis à quebra (divisão) confundem-se com aquelas suscetíveis, em que nenhuma quebra (divisão) ocorre. Dessa maneira, o número de indivíduos para os quais a variável tem valor zero não pode ser determinado. Deste modo deve-se adaptar a Distribuição Binomial Negativa truncada no ponto zero.

viii) Seja  $(y, z)$  os números de acidentes do mesmo indivíduo em dois períodos sucessivos de tempo. Sejam  $Y$  e  $Z$  variáveis aleatórias com Distribuição de Poisson, com o mesmo parâmetro  $\lambda$ . Se  $\lambda$  for uma variável aleatória com Distribuição Gama, isto é,

$$p(\lambda) = \left(\frac{k}{m}\right)^k \frac{1}{\Gamma(k)} e^{-\left(\frac{k}{m}\right)\lambda} \lambda^{k-1},$$

então

$$P(Z|y) = \frac{\Gamma(k+y+z)}{\Gamma(k+y)\Gamma(z+1)} \cdot \left(\frac{m+k}{2m+k}\right)^{k+y} \left(\frac{m}{2m+k}\right)^z,$$

ou seja

$$(2.17) \quad P(Z|y) = \binom{k+y+z-1}{z} \left(\frac{m+k}{2m+k}\right)^{k+y} \left(\frac{m}{2m+k}\right)^z$$

A demonstração deste resultado decorre imediatamente dos seguintes fatos:

$$i) \quad p(\lambda|y) = \frac{e^{-\lambda} \lambda^y p(\lambda)}{\int_0^{\infty} e^{-\lambda} \lambda^y p(\lambda) d\lambda} \quad e$$

$$ii) \quad P(Z|y) = \int_0^{\infty} \frac{e^{-\lambda} \lambda^z}{z!} p(\lambda|y) d\lambda.$$

Observa-se que (2.17) corresponde à forma da Distribuição Binomial Negativa, com parâmetros  $(k+y)$  e  $\frac{m+k}{2m+k}$ . O valor esperado da condicional é  $(k+y) \frac{m}{m+k}$  e a variância é  $(k+y) \frac{m}{m+k} \cdot (1 + \frac{m}{m+k})$ . Observa-se que a regressão de Z em Y é linear e Johnson [22] mostra que isso só ocorre quando  $\lambda$  tem Distribuição Pearson III.

#### 2.3.4 - FORMAS LIMITES

i) Se considerarmos  $q = \frac{\lambda}{k}$ ,  $p \rightarrow 1$  e  $q \rightarrow 0$  tal que  $\lambda$  seja fixo, é fácil verificar que

$$\lim_{k \rightarrow \infty} \binom{r+k-1}{r} p^k q^r = \frac{e^{-\lambda} \lambda^r}{r!},$$

isto é, a Distribuição Binomial Negativa, sob determinadas condições, converge fracamente para a Distribuição de Poisson. De fato,

$$\lim_{k \rightarrow \infty} \binom{r+k-1}{r} p^k q^r = \frac{\lambda^r}{r!} \lim_{k \rightarrow \infty} \frac{k(k+1) \dots (r+k-1)}{k^r} \cdot \left(1 - \frac{\lambda}{k}\right)^k = \frac{e^{-\lambda} \lambda^r}{r!}$$

Este resultado também seria imediatamente obtido se levássemos ao limite, para  $k \rightarrow \infty$ , a função geradora de probabilidades.

Em vez de considerar  $\lambda = qk$  como acaba de ser descrito, Fisz [15] determina o limite para  $\lambda = \frac{qk}{p}$  que é justamen-

te o valor esperado da Binomial Negativa. Ainda neste caso, temos:

$$\lim_{k \rightarrow \infty} \frac{\lambda^r}{r!} \frac{k(k+1) \dots (r+k-1)}{(\lambda+k)^r} \cdot \left( \frac{1}{\frac{\lambda}{k} + 1} \right)^k = \frac{e^{-\lambda} \lambda^r}{r!}$$

Entretanto, Bartko [ 3 ] faz uma análise da eficiência das duas aproximações, baseada no critério do erro máximo definido por

$$M_i(j, k, p) = \text{Max} |BN_i(j, k, p) - BN(j, k, p)|,$$

onde  $BN_i$  é a  $i$ -ésima aproximação da Binomial Negativa acumulada. Suas conclusões são baseadas em tabelas apresentadas para cada caso, e observa-se a melhor precisão quando da adaptação de  $\lambda = \frac{qk}{p}$ . Neste último caso os erros tendem a decrescer, para um dado  $k$ , quando  $p$  está se aproximando da unidade e também para um dado  $p$  quando  $k$  cresce. Quando da primeira aproximação,  $\lambda = qk$ , para um  $p$  fixado os erros tendem a crescer, à medida que  $k$  cresce; aqui a aproximação é melhor quando  $rq$  e  $\frac{rq}{p}$  diferem pouco. É interessante observar que para um  $k$  fixado tal diferença é mínima quando  $p$  se aproxima da unidade, entretanto para  $p$  fixado tal diferença aumenta com  $r$ , daí a situação descrita.

ii) A expressão em (2.16) poderia ser interpretada como distribuição do número de tentativas  $n = r+k$  para a obtenção de  $k$  sucessos. Neste caso teríamos

$$P(X=n) = \binom{n-1}{k-1} p^k q^{n-k}, \quad n=k, k+1, \dots$$

Vamos supor que cada tentativa leva um certo tempo  $\Delta t$  e que  $p$ , a probabilidade de sucesso, é proporcional a  $\Delta t$ , ou seja,  $p = \lambda \Delta t$ . Nessas condições, o tempo  $t = n \Delta t$  para obtermos  $k$  sucessos tem a seguinte distribuição

$$f(t)\Delta t = \binom{\frac{t}{\Delta t} - 1}{k-1} (\lambda\Delta t)^k (1-\lambda\Delta t)^{\frac{t}{\Delta t} - k} =$$
$$= \frac{(\frac{t}{\Delta t} - 1)!}{(k-1)! (\frac{t}{\Delta t} - k)!} \cdot (\lambda\Delta t)^k (1-\lambda\Delta t)^{\frac{t}{\Delta t} - k}$$

$$f(t) = (t-\Delta t)\dots(t-(k-1)\Delta t) \cdot \lambda^k \cdot (1-\lambda\Delta t)^{\frac{t}{\Delta t} - k} \cdot \frac{1}{(k-1)!}$$

Se  $\Delta t \rightarrow 0$ , então

$$f(t) = \frac{t^{k-1} \cdot \lambda^k e^{-\lambda t}}{\Gamma(k)}$$

Esta é a densidade da variável aleatória com Distribuição Gama, com parâmetros  $\lambda$  e  $k$ . Podemos interpretar  $f(t)$  como sendo o tempo de espera para ocorrência de um evento raro.

iii) A aproximação à Distribuição Normal, na sua conotação mais comum, seria da forma:

$$BN(j, k, p) \approx \Phi(x) = \int_{-\infty}^x \phi(t) dt$$

onde  $\phi(t)$  é a densidade normal padrão e

$$x = (j + \frac{1}{2} - \frac{kq}{p}) / \sqrt{\frac{kq}{p^2}}$$

Através das tabelas apresentadas em [ 3 ] observamos que os erros são menores para  $p \leq 0,5$ . Para um  $p$  fixado, os erros tendem a decrescer quando  $r$  cresce.

Em [ 2 ] e [ 5 ] há referência a outras transforma-

ções para uso em Análise de Variância.

iv) Existem espécies de animais na natureza que são tão raras que a chance de inclusão numa amostra é pequena - então as observações podem ser enquadradas na forma limite da Binomial Negativa, quando  $k \rightarrow 0$ , que é Distribuição Logarítmica.

Como já vimos,

$$P(X=r) = \binom{r+k-1}{r} p^k q^r = \frac{(r+k-1)!}{r!(k-1)!} p^k q^r.$$

Fazendo  $k \rightarrow 0$  e substituindo a constante  $\frac{1}{(k-1)!}$  por  $\alpha$ , segue

$$P(X=r) = \frac{\alpha}{r} q^r.$$

Esta é a expressão da Distribuição Logarítmica desenvolvida por Fisher [14] e que já utilizamos à página 29. Esta expressão só é válida para  $r=1,2,\dots$ , ou seja, se considerarmos a Binomial Negativa truncada no valor  $r=0$ .

### 2.3.5 - USO DE TABELAS

Quando estamos tentando adaptar uma Distribuição Binomial Negativa aos nossos dados faz-se necessário calcular as probabilidades inerentes ao modelo, depois que os parâmetros  $p$  e  $k$  forem convenientemente estimados (\*)

Voltemos à expressão (2.12)

$$P(X=r) = \binom{k+r-1}{r} p^k q^r \quad \begin{array}{l} 0 < p < 1 \\ 0 < k < \infty \\ r=0,1,2,\dots \end{array}$$

Para facilidade de notação, façamos

$$P(X=r) = y(r,p,k)$$

---

(\*) - o assunto estimação de parâmetros não será abordado neste trabalho.



e

$$Y(r,p,k) = \sum_{x=0}^r y(x,p,k).$$

Patil [30] sugere o uso da tabela da Distribuição Binomial para o caso da Distribuição de Pascal ( $k$  inteiro), pois:

$Y(r,p,k)$  = probabilidade de serem necessárias até  $(r+k)$  tentativas independentes para a obtenção de  $k$  sucessos, com probabilidade  $p$  de sucesso por tentativa

= probabilidade de no mínimo  $k$  sucessos em  $(r+k)$  tentativas independentes com probabilidade de sucesso  $p$  em cada tentativa

$$= \sum_{x=k}^{r+k} \binom{r+k}{x} p^x (1-p)^{r+k-x}$$

$$(2.18) \quad Y(r,p,k) = 1 - B(k-1; r+k, p)$$

Baseado nisso, Binns [7] propôs uma aproximação à Normal da forma:

$$ABN(r;k,p) = \Phi \left[ \frac{rp - kq + 0,5}{\sqrt{(r+k)pq}} \right]$$

onde  $\Phi$  é a área acumulada da Distribuição Normal Padrão até o ponto  $r$ . Verifica-se, então, através do mesmo erro máximo proposto por Bartko [3], que essa aproximação é mais eficiente do que a aproximação direta à Normal que já mencionamos à página 35.

De (2.18) concluímos que as aproximações e formas limites da Distribuição Binomial vistas em 2.1.3 podem ser aplicadas à correspondente Distribuição Binomial Negativa.

Se  $k > 0$  (não necessariamente inteiro), Patil [30] sugere usar a tabela da função Beta Incompleta, da seguinte forma:

$$Y(r;p,k) = I_p(k,r+1) \quad k > 0,$$

onde

$$I_p(m,n) = \frac{1}{B(m,n)} \int_0^p u^{m-1} (1-u)^{n-1} du.$$

Isto se deve ao seguinte fato:

$$\frac{d}{dp}[I_p(k,r+1)] = \frac{k}{p^2} y(r,p,k+1)$$

$$\frac{d}{dp}[Y(r;p,k)] = \frac{k}{p^2} y(r,p,k+1).$$

Ainda para  $k > 0$ , Guenther [16] se refere a outra a aproximação: considerando que

$$E(X) = \frac{kq}{p} \quad e \quad V(X) = \frac{kq}{p^2},$$

se fizermos  $\alpha = kq$  e  $\beta = 1/p$ , teremos  $E(X) = \alpha\beta$  e  $V(X) = \alpha\beta^2$ ; en tretanto  $\alpha\beta$  e  $\alpha\beta^2$  são respectivamente a média e variância de uma Distribuição Gama cuja densidade é dada por

$$\frac{1}{\beta^\alpha \Gamma(\alpha)} z^{\alpha-1} e^{-z/\beta}, \quad z > 0;$$

a sugestão é que se aproxime então a probabilidade acumulada da Distribuição Binomial Negativa pela Distribuição Gama com parâmetros  $\alpha = k(1-p)$  e  $\beta = 1/p$ ; isto levaria a

$$Y(r,k,p) \approx P(Z < r+1/2)$$

onde  $1/2$  é a usual correção de continuidade; se considerar-

mos  $z/\beta = pz = t/2$ , vem

$$\begin{aligned} P(Z < r+1/2) &= P(T < (2r+1)p) \\ &= P(N, X) \\ &= \int_0^X D(N, t) dt \end{aligned}$$

onde

$$D(N, t) = \frac{t^{N-1} e^{-t/2}}{2^N \Gamma(N)}, \quad N = \alpha, \quad X = (2r+1)p$$

$$\therefore Y(r; k, p) \approx P(N, X).$$

Binns [ 7 ] considera esta aproximação superior à sua em eficiência e Guenther [ 16 ] recomenda usá-la somente para  $r > 1$ .

Embora haja outras aproximações mais eficientes, não serão tratadas aqui, quer pela falta de tabelas, quer pela sofisticação das transformações usadas.

### 2.3.5 - MISTURAS

Como foi feito com as duas Distribuições anteriores, mencionaremos algumas misturas envolvendo a Distribuição Binomial Negativa.

i) Binomial Negativa  $(k, p)$   $\overset{\Delta}{p}$  Beta

Neste caso a função densidade de  $P$  é da forma

$$[B(\alpha, \beta)] p^{\alpha-1} (1-p)^{\beta-1}$$

e a distribuição resultante é a chamada Distribuição Hiper-

geométrica Generalizada [21], cuja expressão é obtida se considerarmos a possibilidade dos parâmetros da Distribuição Hipergeométrica não serem necessariamente positivos. [De fato, se dois deles são negativos e o terceiro for positivo, ainda teremos valores positivos para as probabilidades calculadas através da expressão em (2.5)].

ii) Binomial Negativa (K,p)  $\overset{\Lambda}{K/N}$  Poisson ( $\lambda$ )

A distribuição resultante é a chamada Distribuição Poisson-Pascal (a mais usada depois da Poisson-Binomial, página 16), que inclui as Distribuições Neyman Tipo A e Binomial Negativa como particulares limites.

A função geradora de probabilidades da Distribuição Poisson-Pascal é da forma

$$(2.19) \quad P(s) = \exp\left[\lambda\left\{\left(\frac{p}{1-qs}\right)^k - 1\right\}\right]$$

e a distribuição correspondente é definida por

$$P(X=0) = e^{-\lambda}$$

$$P(X=n) = e^{-\lambda} q^n \sum_{j \leq n/N} \binom{n-1}{Nj-1} \left\{ \left[ \lambda \left( \frac{p}{q} \right)^{-N} \right]^j / j! \right\}, \quad n \geq N$$

se considerarmos a Distribuição Binomial Negativa como em (2.16). Sob esta forma, a Distribuição Poisson-Pascal é conhecida como Distribuição Polya-Aeppli Generalizada e, no caso de  $N=1$ , temos a Distribuição Polya-Aeppli. Esta última corresponde a um modelo em que os indivíduos ocorrem em grupo, onde o número de grupos tem Distribuição de Poisson enquanto que o número de indivíduos por grupo tem Distribuição Geométrica.

Katti [23] descreve uma situação prática em que se adapta a distribuição (2.19) e, além disso, aborda as seguintes situações limites:

- i) se  $k \rightarrow \infty$ ,  $q \rightarrow 0$ , tal que  $qk \rightarrow \lambda_1$ ,  
a Distribuição Poisson-Pascal torna-se Neyman-Tipo A;
- ii) se  $k \rightarrow 0$ ,  $\lambda \rightarrow \infty$ , tal que  $\lambda k \rightarrow k_1$ , temos  
a Distribuição Binomial Negativa;
- iii) se  $q \rightarrow 0$ ,  $\lambda \rightarrow \infty$ , tal que  $\lambda k q \rightarrow \lambda_1$ , temos  
a Distribuição de Poisson;
- iv) se  $k \rightarrow 1$ , (2.19) é a função geradora de probabilidades da Distribuição Polya-Aeppli.

## CAPÍTULO 3

### MÉTODOS DE DISCRIMINAÇÃO

As distribuições associadas a variáveis aleatórias que tomam valores num látice<sup>(\*)</sup> de amplitude  $d$  (como as analisadas no capítulo anterior), podem ser descritas de várias formas diferentes. Inicialmente vamos considerá-las como um caso particular das distribuições em série de potências e, em seguida, como distribuições que satisfazem um sistema de equações-diferença. Através desse sistema obteremos critérios de discriminação entre as distribuições.

#### 3.1 - SÉRIE DE POTÊNCIAS

Uma variável aleatória  $X$  tem distribuição em série de potências (dsp) se

$$(3.1) \quad P(X=r) = \frac{a_r \theta^r}{h(\theta)}, \quad r = 0, 1, 2, \dots$$

onde

$$a_r \in \mathbb{R},$$

$$|\theta| < 1,$$

$$h(\theta) = \sum_{r=0}^{\infty} a_r \theta^r:$$

---

(\*) - Diz-se que os valores de uma variável aleatória  $X$  pertencem a um látice de amplitude  $d$ , quando  $\exists d \in \mathbb{N}^+$  tal que  $X/d$  toma valores em todos os inteiros não negativos. Para o caso das distribuições descritas no capítulo anterior,  $d=1$ .

Se a variável aleatória só possui valores num subconjunto não vazio dos inteiros não negativos, temos, de acordo com Patil [31], uma distribuição generalizada em série de potências (dgsp). Neste caso:

$$(3.2) \quad P(X=r) = \frac{a_r \theta^r}{g(\theta)}, \quad r \in T$$

onde

$$g(\theta) = \sum_{r \in T} a_r \theta^r, \quad \begin{array}{l} a_r > 0 \\ \theta \geq 0 \end{array}$$

Uma dgsp torna-se dsp quando T for o conjunto dos inteiros não negativos.

Para os momentos de uma dgsp valem as seguintes propriedades:

$$\mu = EX = \theta \frac{d}{d\theta} \log g(\theta)$$

$$\sigma^2 = VX = \mu + \theta^2 \frac{d^2}{d\theta^2} \log g(\theta) = \theta \frac{d\mu}{d\theta}.$$

Mostraremos em seguida como as Distribuições Binomial, Poisson e Binomial Negativa podem ser colocadas nesse contexto, e os resultados poderão ser comparados com os que foram obtidos no capítulo anterior.

a) Binomial (n,p)

Basta tomar  $g(\theta) = (1+\theta)^n$  onde  $\theta = \frac{p}{1-p}$ ; observa-se que

$$g(\theta) = \sum_{r \in T} \binom{n}{r} \theta^r, \quad T = \{0, 1, \dots, n\}$$

Então

$$P(X=r) = \frac{\binom{n}{r} \theta^r}{(1+\theta)^n}, \quad r=0, 1, \dots, n$$

$$\mu = EX = \theta \frac{\partial}{\partial \theta} n \log(1+\theta) = np$$

$$VX = \theta \frac{\partial}{\partial \theta} \frac{n\theta}{1+\theta} = np(1-p).$$

b) Poisson ( $\lambda$ )

Neste caso

$$h(\theta) = e^\theta = \sum_{r=0}^{\infty} \frac{\theta^r}{r!}, \text{ onde } \theta = \lambda.$$

Assim

$$P(X=r) = \frac{\theta^r (r!)^{-1}}{e^\theta}$$

$$\mu = EX = \theta \frac{\partial}{\partial \theta} \log h(\theta) = \theta$$

$$VX = \theta \frac{\partial \mu}{\partial \theta} = \theta.$$

c) Binomial Negativa ( $k;p$ )

Escolhe-se  $h(\theta) = (1-\theta)^{-k}$ ,  $k=0$   $\theta=(1-p)$ .

Com a notação de Feller [12], temos:

$$h(\theta) = \sum_{r=0}^{\infty} \binom{r+k-1}{r} \theta^r.$$

A distribuição pode ser escrita como:

$$P(X=r) = \frac{\binom{k+r-1}{r} \theta^r}{(1-\theta)^{-k}}, \quad r=0,1,2,\dots$$



Ainda

$$E(X) = \theta \cdot \frac{d}{d\theta} - k \log(1-\theta) = +\theta k \cdot \frac{1}{1-\theta} = \frac{qk}{p}$$

$$V(X) = \theta \cdot \frac{d}{d\theta} \cdot \frac{\theta k}{1-\theta} = \theta k \frac{(1-\theta)+\theta}{(1-\theta)^2} = \frac{qk}{p^2}$$

### 3.2 - SISTEMAS DEFINIDOS POR EQUAÇÕES-DIFERENÇA

Em analogia ao sistema de distribuições contínuas gerado por Pearson, baseado numa equação diferencial, pode-se obter um sistema de distribuições discretas baseado na seguinte equação-diferença:

$$(3.3) \quad \Delta f_{r-1} = \frac{(a-r)f_{r-1}}{b_0 + b_1 r + b_2 r(r-1)},$$

onde  $f_r = P(X=r)$  e  $\Delta f_{r-1} = f_r - f_{r-1}$ .

Esta família de distribuições é tão extensa quanto aquela determinada pela equação diferencial. No caso em que  $b_2 = 0$  temos as distribuições generalizadas em série de potências, definidas em 3.1.

A seguir mostraremos que as Distribuições Binomial, Poisson e Binomial Negativa, pertencem à classe de distribuições definidas por (3.3).

i) Binomial (n;p)

$$\begin{aligned} \Delta f_{r-1} &= \binom{n}{r} p^r q^{n-r} - \binom{n}{r-1} p^{r-1} q^{n-r+1} = \\ &= \frac{n! p^{r-1} q^{n-r+1}}{(r-1)!(n-r+1)!} \left[ \frac{(n-r+1)p-qr}{q \cdot r} \right] = \left[ \frac{(n+1)p-r}{rq} \right] f_{r-1}. \end{aligned}$$

$$\text{Então: } a = (n+1)p$$

$$b_1 = q = (1-p)$$

$$b_0 = b_2 = 0.$$

ii) Poisson ( $\lambda$ )

$$\Delta f_{r-1} = \frac{e^{-\lambda} \lambda^r}{r!} - \frac{e^{-\lambda} \lambda^{r-1}}{(r-1)!} = \frac{e^{-\lambda} \lambda^{r-1}}{(r-1)!} \left[ \frac{r-1}{r} \right] \dots \begin{matrix} a = \lambda \\ b_0 = b_2 = 0 \\ b_1 = 1 \end{matrix}$$

iii) Binomial Negativa ( $k, p$ )

$$\begin{aligned} \Delta f_{r-1} &= \binom{k+r-1}{r} p^k q^r - \binom{k+r-2}{r-1} p^k q^{r-1} = \\ &= \frac{(k+r-2)! p^k q^{r-1}}{(r-1)! (k-1)!} \left[ \frac{(k+r-1)q-r}{r} \right] = \frac{\frac{(k-1)q-r}{1-q}}{\frac{r}{1-q}}. \end{aligned}$$

Temos:

$$a = \frac{(k-1)q}{1-q}$$

$$b_0 = b_2 = 0$$

$$b_1 = \frac{1}{1-q}.$$

Observamos então que a equação diferença inicialmente definida em (3.3) reduz-se, no caso das três distribuições mencionadas, à seguinte expressão:

$$(3.4) \quad \Delta f_{r-1} = \frac{(a-r)f_{r-1}}{b_1 r}$$

ou, equivalentemente, a

$$(3.5) \quad \frac{f_{r+1}}{f_r} = \frac{[(a+b_1-1)/b_1] + [(b_1-1)/b_1]r}{r+1}$$

Foi a partir de (3.5) que Katz [24] desenvolveu um método discriminatório para as três distribuições consideradas, isto é, Binomial, Poisson e Binomial Negativa.

Simplificando a notação, segue:

$$(3.6) \quad \frac{f_{r+1}}{f_r} = \frac{\alpha + \beta r}{r+1}$$

Para obter relações entre  $\alpha$ ,  $\beta$  e os momentos das distribuições pertencentes à classe, basta reescrever (3.6) da seguinte forma:

$$(3.7) \quad \sum (r+1)^{p+1} f(r+1) = \sum (r+1)^p (\alpha + \beta r) f(r).$$

Se  $p = 0$  na expressão (3.7) verifica-se que

$$EX = \mu = \alpha + \beta \mu$$

$$(3.8) \quad \mu = \frac{\alpha}{1 - \beta}$$

Se  $p = 1$  na mesma expressão, vem

$$EX^2 = V(X) + \mu^2 = \sum (\alpha r + \alpha + \beta r^2 + \beta r) \cdot f(r).$$

Levando em conta (3.8) segue:

$$(3.9) \quad V(X) = \sigma^2 = \frac{\alpha}{(1-\beta)^2}$$

As expressões

$$(3.10) \quad \alpha = \frac{\mu^2}{\sigma^2}$$

$$(3.11) \quad \beta = 1 - \frac{\mu}{\sigma^2}$$

decorrem de (3.8) e (3.9).

Como  $\mu > 0$ , exceto no caso em que  $P(X=0) = 1$ , observa-se que os parâmetros da classe em que estamos trabalhando são tais que  $\alpha > 0$  e  $\beta < 1$ . Além disso, como as distribuições da referida classe estão bem determinadas através de seus dois primeiros momentos, podemos expressar  $\mu_3$  e  $\mu_4$  (que são, respectivamente os 3º e 4º momentos centrais) em função de  $\mu$  e  $\sigma^2$ , bastando fazer  $p = 2$  e  $p = 3$ , respectivamente, em (3.7). Deste modo, segue

$$\mu_3 = \sigma^2(2c-1)$$

$$\mu_4 = \sigma^2(6c^2-6c+1)+3\sigma^4$$

onde  $c = \frac{\sigma^2}{\mu}$ .

Mostraremos a partir de (3.6), que a escolha dos parâmetros segundo o que foi descrito nos itens i, ii e iii determina as Distribuições Binomial, Poisson e Binomial Negativa de maneira única.

$$\frac{f(r)}{f(r-1)} = \frac{\alpha + \beta(r-1)}{r}$$

(3.12)  $\therefore f(r) = \frac{\alpha + \beta(r-1)}{r} \cdot f(r-1)$

a) Se  $\beta = 0$  então (3.12) se torna

$$f(r) = \frac{\alpha}{r} \cdot f(r-1)$$

$$f(r) = \frac{\alpha^r}{r!} \cdot f(0).$$

Para que  $\sum_r f(r) = 1$  toma-se  $f(0) = e^{-\alpha}$

$$\therefore f(r) = \frac{e^{-\alpha} \alpha^r}{r!}.$$

Esta última expressão é justamente a Distribuição de Poisson, conforme já visto em (2.8).

b) Se  $\beta \neq 0$ , reescrevemos (3.12) da seguinte forma:

$$(3.13) \quad f(r) = \frac{\beta}{r} \left( \frac{\alpha}{\beta} + r - 1 \right) \cdot f(r-1)$$

Como

$$\Gamma \left( \frac{\alpha}{\beta} + r \right) = \left( \frac{\alpha}{\beta} + r - 1 \right) \left( \frac{\alpha}{\beta} + r - 2 \right) \dots \left( \frac{\alpha}{\beta} \right) \Gamma \left( \frac{\alpha}{\beta} \right),$$

vem:

$$f(r) = \frac{\beta^r}{r!} \cdot \frac{\Gamma \left( \frac{\alpha}{\beta} + r \right)}{\Gamma \left( \frac{\alpha}{\beta} \right)} \cdot f(0).$$

Toma-se  $f(0) = (1-\beta)^{\alpha/\beta}$  e temos:

$$(3.14) \quad f(r) = \frac{\Gamma \left( \frac{\alpha}{\beta} + r \right)}{\Gamma \left( \frac{\alpha}{\beta} \right)} \cdot \frac{\beta^r}{r!} \cdot (1-\beta)^{\alpha/\beta}$$

Dentro deste item b), podemos considerar:

b<sub>1</sub>)  $\beta$  positivo, isto é,  $0 < \beta < 1$ .

Seja  $\frac{\alpha}{\beta} = k$  e  $\beta = 1-p$ . Então

$$f(r) = \frac{(1-p)^r}{r!} \cdot \frac{(k+r-1)!}{(k-1)!} \cdot p^k$$

$$f(r) = \binom{k+r-1}{r} \cdot p^k (1-p)^r,$$

ou seja, obtivemos a expressão da Distribuição Binomial Negativa como definida em (2.12).

b<sub>2</sub>)  $\beta < 0$ .

Façamos  $\frac{\alpha}{\beta} = -v$  e  $\beta = \frac{p}{p-1}$ . Então

$$f(r) = \left(\frac{p}{p-1}\right)^r \cdot \frac{(-v+r-1)!}{(-v-1)!} \cdot \left(1 - \frac{p}{p-1}\right)^{-v} \cdot \frac{1}{r!}.$$

Usando a identidade:

$$\binom{v}{r} = (-1)^r \cdot \binom{-v+r-1}{r}, \quad v > 0$$

segue

$$f(r) = \binom{v}{r} \cdot p^r \cdot (1-p)^{v-r}$$

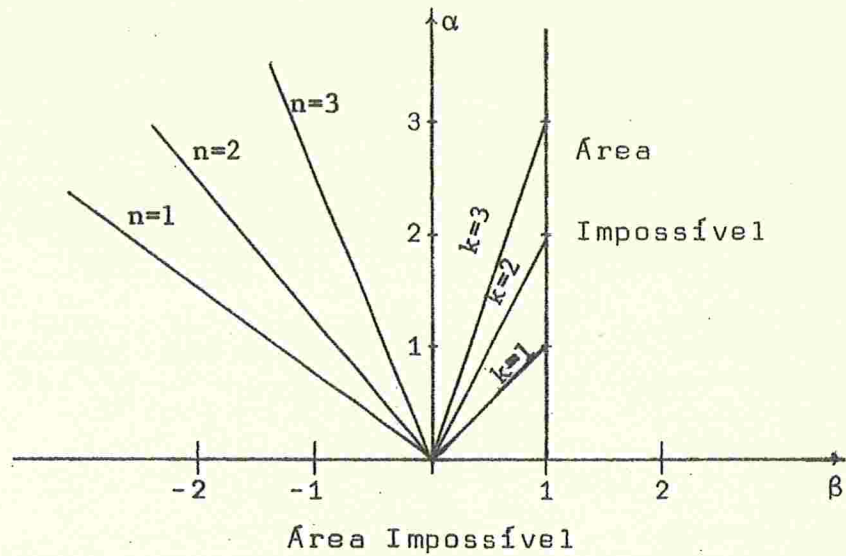
Esta é a expressão da Distribuição Binomial já descrita em (2.1).

Identificando os parâmetros  $\alpha$  e  $\beta$ , podemos resumir os resultados como na Tabela 3.1.

TABELA 3.1

DISTRIBUIÇÃO	$\alpha$	$\beta$	a	b <sub>1</sub>
Binomial	$\frac{np}{q}$	$-\frac{p}{q}$	$(n+1)p$	q
Poisson	$\lambda$	0	$\lambda$	1
Binomial Negativa	kq	q	$\frac{(k-1)q}{1-q}$	$\frac{1}{1-q}$

Além disso, usando as informações acima, podemos relacionar  $\alpha$  e  $\beta$  através do seguinte gráfico:



Ord [29] utilizou a mesma equação-diferença em (3.5) para propor um critério discriminatório para todas as distribuições de séries hipergeométricas que em particular se aplica às três distribuições em questão. Esse critério u tiliza-se dos quocientes:

$$(3.15) \quad I = \frac{\mu_2}{\mu_1^2} \quad e \quad S = \frac{\mu_3}{\mu_2}$$

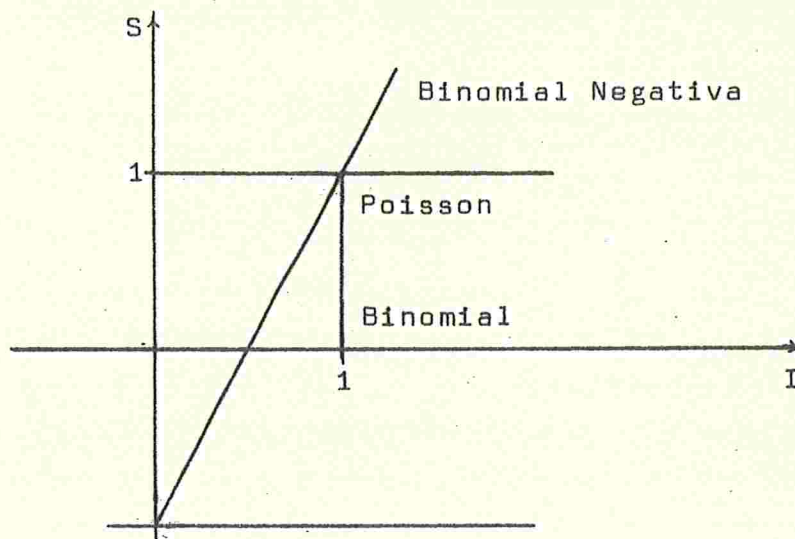
onde  $\mu_1'$  = 1º momento;  $\mu_2$  = 2º momento central e  $\mu_3$  = 3º momento central.

Nesse caso temos:

DISTRIBUIÇÃO	I	S
Binomial	q	2I-1
Poisson	1	1
Binomial Negativa	1/p	2I-1

Observa-se que para a Distribuição de Poisson  $I=S=1$ ; para a Binomial  $0 < I < 1$  e  $-1 < S < 1$ , enquanto que para a Binomial Negativa  $I > 1$  e  $S > 1$ .

O plano  $(I,S)$  pode ser descrito do seguinte modo:



Por outro lado nota-se que para essas três distribuições valem as relações  $I = \eta + 1$  e  $S = 2\eta + 1$ . Portanto, do ponto de vista dos parâmetros populacionais, os métodos são equivalentes.

Este mesmo autor [28] sugere um outro critério para discriminação entre essas distribuições, que consiste em colocar

$$(3.16) \quad u_r = \frac{rf_r}{f_{r-1}}$$

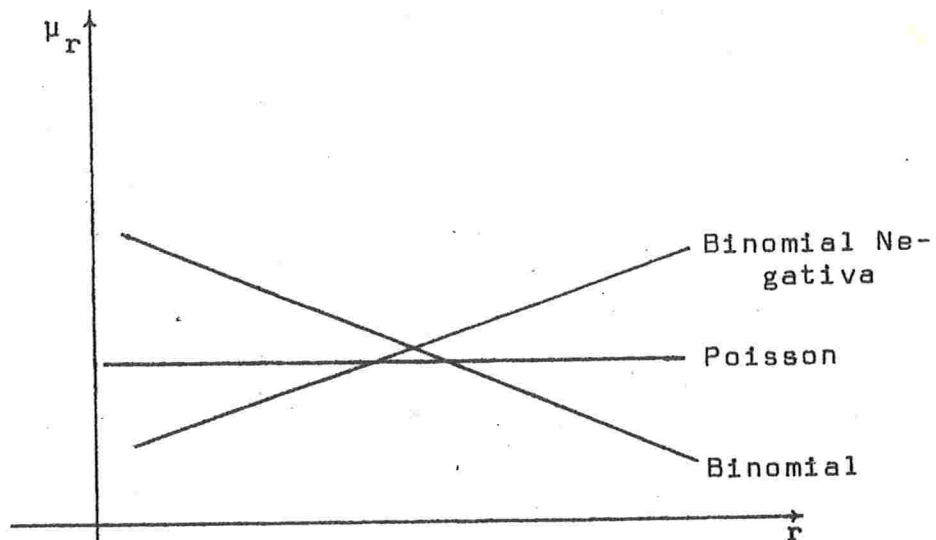
contra  $r$  num gráfico de coordenadas cartesianas. De (3.12) observa-se que  $u_r = (\alpha - \beta) + \beta r$ . Utilizando a notação do autor,  $u_r = c_0 + c_1 r$ , temos  $c_0 = \alpha - \beta$  e  $c_1 = \beta$ , cujos valores podem ser obtidos a partir da Tabela 3.1, e estão relacionadas na Tabela 3.2:



TABELA 3.2

DISTRIBUIÇÃO	$c_0$	$c_1$
Binomial	$(n+1)p/q$	$-p/q$
Poisson	$\lambda$	0
Binomial Negativa	$(k-1)q$	$q$

Podemos esquematizar graficamente a tabela anterior do seguinte modo:



O autor sugere o uso desta técnica somente para

$$f_{r-1} \geq 5.$$

NOTA - No próximo capítulo usaremos  $N$  para denotar um dos parâmetros da Distribuição Binomial em lugar de  $n$ , o qual reservaremos para tamanho da amostra.

## CAPÍTULO 4

### ANÁLISE DOS RESULTADOS

#### 4.1 - ESTIMAÇÃO

Nesta secção discutiremos o problema de estimação dos parâmetros introduzidos no capítulo anterior com base numa amostra aleatória de tamanho  $n$  da distribuição (3.14).

O logaritmo da função de verossimilhança tem a forma

$$L(\alpha, \beta) = \sum x_i \ln \beta + n \frac{\alpha}{\beta} \ln(1-\beta) + \\ + \sum \ln \Gamma\left(\frac{\alpha}{\beta} + x_i\right) - n \ln \Gamma\left(\frac{\alpha}{\beta}\right) - \sum \ln(x_i!)$$

de onde segue que

$$\frac{\partial h}{\partial \alpha} = \frac{n}{\beta} \ln(1-\beta) + \sum F\left(\frac{\alpha}{\beta} - x_i - 1\right) \frac{1}{\beta} - n F\left(\frac{\alpha}{\beta} - 1\right) \cdot \frac{1}{\beta}$$

e

$$\frac{\partial h}{\partial \beta} = \frac{\sum x_i}{\beta} + n \alpha \left(-\frac{1}{\beta^2} \ln(1-\beta) + \frac{1}{1-\beta} - \frac{1}{\beta}\right) + \\ + \sum F\left(\frac{\alpha}{\beta} + x_i - 1\right) \left(-\frac{\alpha}{\beta^2}\right) - n F\left(\frac{\alpha}{\beta} - 1\right) \left(-\frac{\alpha}{\beta^2}\right)$$

Observação:  $F$  é a função digama, definida por

$$F(\alpha) = \frac{\partial}{\partial \alpha} \ln \Gamma(\alpha)$$

Igualando a zero as duas últimas equações e somando membro a membro, vem

$$\begin{aligned} & \frac{n}{\beta} \ln(1-\beta) + \sum [F(\frac{\alpha}{\beta} + x_i - 1)] (\frac{1}{\beta} - \frac{\alpha}{\beta^2}) - n[F(\frac{\alpha}{\beta} - 1)] \cdot \\ & \cdot (\frac{1}{\beta} - \frac{\alpha}{\beta^2}) + \frac{\sum x_i}{\beta} + n\alpha (-\frac{1}{\beta^2} \ln(1-\beta) + \frac{1}{\beta(1-\beta)}) = 0. \end{aligned}$$

Simplificando, vem

$$(4.1) \quad \frac{\hat{\alpha}}{1-\beta} = \frac{\sum x_i}{n} = \bar{x} \text{ (média amostral).}$$

Substituindo  $\alpha$ , obtido em (3.15) na expressão  $\frac{dL}{d\alpha}$ , segue:

$$(4.2) \quad \frac{n}{\beta} \ln(1-\beta) + \sum [F(\frac{\bar{x}(1-\beta)}{\beta} + x_i - 1)] \frac{1}{\beta} - n[F(\frac{\bar{x}(1-\beta)}{\beta} - 1)] \frac{1}{\beta} = 0.$$

A solução desta equação nos fornecerá o estimador de  $\beta$ .

Observando (4.1) e (4.2) verifica-se que é conveniente, em termos de estimação, fazer uso da seguinte reparametrização:

$$\xi = \frac{\alpha}{1-\beta}$$

e

$$\eta = \frac{\beta}{1-\beta}$$

Então, em correspondência às relações acima construímos, a partir da Tabela 3.1, a Tabela 4.1.

É fácil ver que

$$\eta \text{ Binomial} \leq \eta \text{ Poisson} \leq \eta \text{ Binomial Negativa}$$

TABELA 4.1

DISTRIBUIÇÃO	$\xi$	$\eta$
Binomial	$np$	$-p$
Poisson	$\lambda$	$0$
Binomial Negativa	$qk/p$	$q/p$

O estimador de máxima verossimilhança para  $\xi$  é

$$\hat{\xi} = \bar{X}$$

e para  $\eta$  é a solução da equação (4.3).

$$(4.3) \quad n \ln\left(\frac{1}{\eta+1}\right) + \sum_i [F\left(\frac{\bar{X}}{\eta} + x_i - 1\right)] - n[F\left(\frac{\bar{X}}{\eta} - 1\right)] = 0.$$

Esta equação pode ser resolvida por métodos de aproximação mas, dadas as dificuldades associadas à solução, recorreu-se a um estimador que tivesse ao menos uma das propriedades do estimador de máxima verossimilhança. Usando-se  $\hat{\eta}$  definido por

$$\hat{\eta} = \frac{S^2 - \bar{X}}{\bar{X}}$$

como estimador de  $\eta$ , podemos demonstrar, através dos teore-

mas de Khintchin e Slutsky [15], que  $\eta$  é um estimador consistente para  $\eta$ .

Considerando que  $\hat{\eta}$  só está definido para  $\bar{x} \neq 0$ , e que nosso processo de simulação (conforme descrito à página 62). abandonava a distribuição que estava sendo gerada caso

$$\sum_{i=1}^n x_i = 0,$$

achamos interessante obter expressões exatas para  $E(\hat{\eta})$ , através da distribuição de

$$\sum_{i=1}^n X_i,$$

truncada no ponto zero, a fim de compará-las com os resultados simulados. A partir das expressões:

$$s^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n} \quad \text{e} \quad \bar{X} = \frac{\sum_{i=1}^n X_i}{n}$$

obtidas de uma amostra casual de tamanho  $n$  de cada uma das três distribuições em estudo, vamos calcular o valor esperado de  $\hat{\eta}$ .

### 1. Distribuição de Poisson ( $\lambda$ )

Para cálculo da  $E(\hat{\eta})$  vamos nos utilizar do seguinte fato:

Se  $X_1, X_2, \dots, X_n$  for uma amostra casual de uma distribuição de Poisson, a distribuição condicional de  $X_j$ ,  $1 \leq j \leq n$ , dada

$$\sum_{i=1}^n X_i$$

é Binomial

$$B\left(\sum_{i=1}^n X_i; \frac{1}{n}\right).$$

Segue-se então que

$$(4.4) \quad E\left(X_j^2 \mid \sum_{i=1}^n X_i = \ell\right) = \frac{\ell(n-1) + \ell^2}{n^2}$$

Somando a expressão em (4.4) para todo  $j, 1 \leq j \leq n$ , e dividindo por  $\ell \neq 0$ , vem

$$(4.5) \quad E\left(\frac{\sum X_i^2}{\sum X_i} \mid \sum X_i = \ell\right) = \frac{(n-1) + \ell}{n}$$

Então temos que

$$E(\hat{\eta}) = E\left[\frac{n-1}{n} + \bar{X}\right] - E\bar{X} - 1$$

$$(4.6) \quad E(\hat{\eta}) = -\frac{1}{n}.$$

Observamos que o valor esperado em (4.6) é calculado em função da distribuição de  $\bar{X}$  truncada no ponto zero, dada a restrição que colocamos ao estabelecermos (4.5).

Se tivéssemos usado  $(n-1)$  no denominador de  $S^2$  ao invés de  $n$ , obteríamos um estimador não viciado para  $n$ , isto é,  $E(\hat{\eta}) = 0$ . De (4.6) vemos que  $E(\hat{\eta}) \rightarrow 0$  quando  $n \rightarrow \infty$ .

## 2. Distribuição Binomial $(N, p)$

Se considerarmos uma amostra aleatória de tamanho  $n$  desta distribuição, temos que, para  $1 \leq j \leq n$ , a distribuição condicional de  $X_j$  dada

$$\sum_{i=1}^n X_i$$

é hipergeométrica com parâmetros

$$(n, \sum_{i=1}^n X_i, N).$$

Dessa forma,

$$(4.7) \quad E(X_j^2 | \sum X_i = \ell) = \frac{(nN-\ell)(n-1)\ell}{n^2(Nn-1)} + \frac{\ell^2}{n^2}.$$

Somando (4.7) para todo  $j$ ,  $1 \leq j \leq n$ , e dividindo por  $\ell \neq 0$  temos:

$$(4.8) \quad E \left( \frac{\sum_{i=1}^n X_i^2}{\sum_{i=1}^n X_i} \mid \sum_{i=1}^n X_i = \ell \right) = \frac{(nN-\ell)(n-1)}{n(Nn-1)} + \frac{\ell}{n}.$$

Para esperança da variável aleatória

$$E \left( \frac{\sum_{i=1}^n X_i^2}{\sum_{i=1}^n X_i} \mid \sum X_i \right) = \phi_1 \left( \sum_{i=1}^n X_i \right)$$

temos que considerar a distribuição da variável aleatória

$$\sum_{i=1}^n X_i$$

como truncada no ponto zero, em virtude da restrição que impusemos ao escrevermos (4.8).

Assim sendo,

$$E \left( \sum_{i=1}^n X_i \right) = \frac{E[B(Nn, p)]}{1-(1-p)^{nN}} = \frac{nNp}{1-(1-p)^{nN}}.$$

Dessa forma temos:

$$E(\hat{\eta}) = E \left[ \frac{(nN - \sum_{i=1}^n X_i)(n-1)}{n(Nn-1)} + \bar{X} \right] = E \bar{X} - 1$$

$$\therefore (4.9) \quad E(\hat{\eta}) = \frac{1-N}{Nn-1} - \frac{(n-1)Np}{(Nn-1)[1-(1-p)^{nN}]}$$

De (4.9) observamos que para  $p = 1$ ,  $E(\hat{\eta}) = -1$ . Além disso,  $\lim_{n \rightarrow \infty} E(\hat{\eta}) = -p$ .

### 3. Distribuição Binomial Negativa ( $k;p$ )

Para  $1 \leq j \leq n$ , a distribuição condicional de  $X_j$  dada

$$\sum_{i=1}^n X_i = \ell$$

é da forma:

$$(4.10) \quad \binom{j+k-1}{j} \binom{\ell-j+k(n-1)-1}{\ell-j} / \binom{\ell+nk-1}{\ell}$$

Com um pouco de cálculo chega-se

$$(4.11) \quad E \left[ X_j \mid \sum_{i=1}^n X_i = \ell \right] = \frac{\ell[(k+1)\ell - (1-n)k]}{n(1+nk)}$$

Somando (4.11) para todo  $j$ ,  $1 \leq j \leq n$ , e dividindo por  $\ell \neq 0$ , vem

$$E \left[ \frac{\sum_{i=1}^n X_i^2}{\sum_{i=1}^n X_i} \mid \sum_{i=1}^n X_i = \ell \right] = \frac{(k+1)\ell - (1-n)k}{(1+nk)}$$



#### 4.2 - ANÁLISES

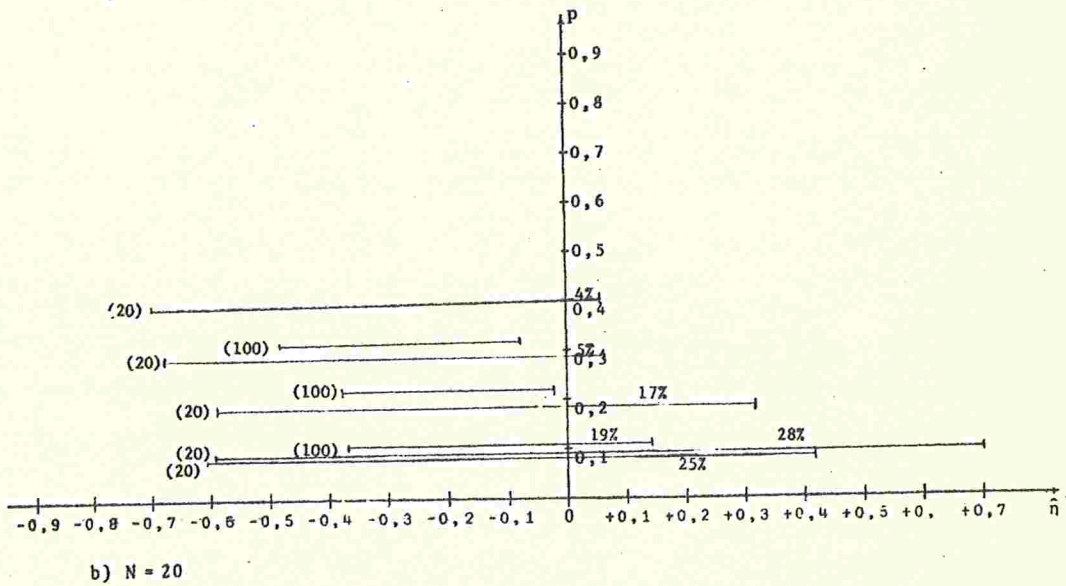
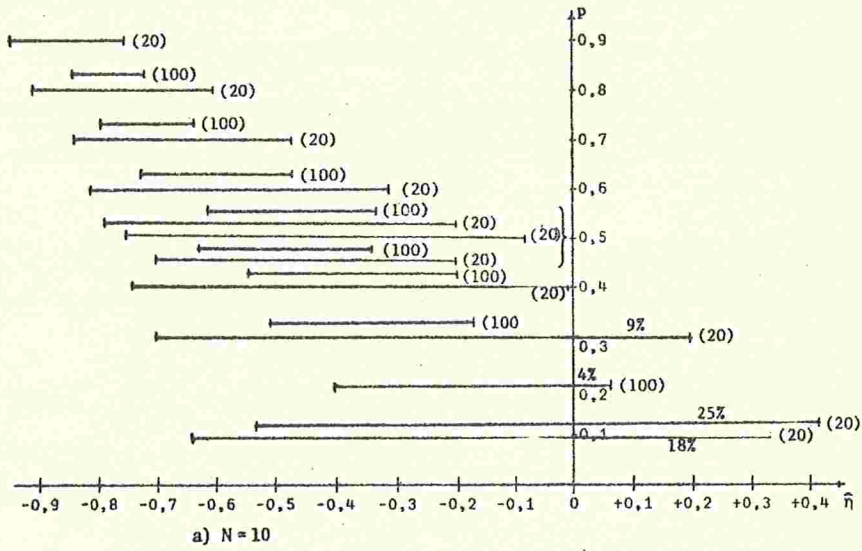
Como vimos anteriormente, Katz [24] propõe uma medida discriminatória entre algumas distribuições discretas, cuja expressão é  $\eta = \frac{\sigma^2 - \mu}{\mu}$  [lembramos que esta expressão é sempre definida para nós porque não estamos considerando distribuições degeneradas, onde a probabilidade de sucesso é nula]. Simulamos através do computador várias vezes cada uma dessas distribuições, especificando em cada caso os respectivos parâmetros. Isto foi feito para as Distribuições Binomial, Poisson e Binomial Negativa e passaremos a descrever os resultados obtidos em cada caso.

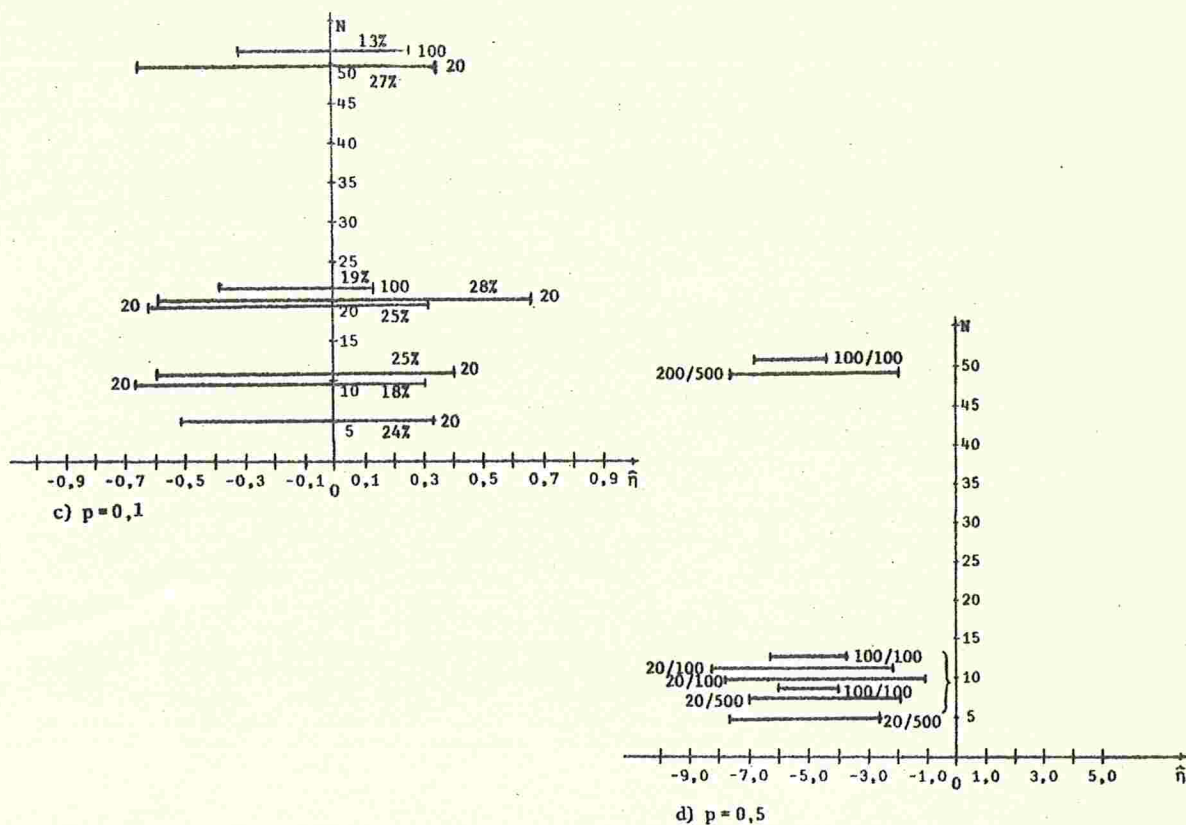
##### a) Distribuição Binomial B(N;p)

Foram selecionados 35 distribuições com os valores de p variando entre 0,1 e 0,9 e alguns valores de n (5; 10; 20; 50). Para cada caso foram consideradas amostras de tamanho n(20, 100 e excepcionalmente 10.000), colhidas 10, 100 ou 500 vezes. Vamos chamar de  $T_n$  o número de amostras de tamanho n para cada B(N;p). Em seguida para cada amostra estimamos o valor de  $\hat{\eta}$  através de  $\hat{\eta} = \frac{s^2 - \bar{x}}{\bar{x}}$ : desse modo obtivemos  $T_n$  valores de  $\hat{\eta}$ 's, para cada distribuição considerada.

Não conhecendo a distribuição de  $\hat{\eta}$ , tentamos fazer uma análise de seu comportamento através da construção de intervalos de tolerância não paramétricos. Estes intervalos nos fornecem dois valores limites, dentro dos quais esperamos que, com probabilidade  $\gamma$ , esteja  $100\pi\%$  da população.

Os gráficos abaixo, referem-se aos intervalos de tolerância construídos usando  $\gamma = \pi = 0,90$ , e podem nos dar uma idéia geral do comportamento dessa medida discriminatória.





Os números que aparecem à direita de alguns intervalos de tolerância correspondem à porcentagem de  $\hat{\eta}$ 's  $> 0$ . Exemplificando, nas 100 vezes que simulamos  $B(10;0,2)$  com amostras de tamanho 20, obtivemos 15% dos valores  $> 0$ , embora o intervalo de tolerância determinando fosse  $(-0,609091; 0,373404)$ .

Como os intervalos de tolerância podem ser influenciados por valores extremos, resolvemos calcular os percentis de ordem 90, a fim de dar maior informação sobre a distribuição de  $\hat{\eta}$ .

Para alguns casos apresentamos médias e variâncias dos  $T_n \hat{\eta}$ 's. Os resultados estão na Tabela 4.2.

TABELA 4.2

B(N;p)	n	T <sub>n</sub>	90° η̄	E(η̄) <sup>(1)</sup>	√V(η̄) <sup>(2)</sup>	MÉDIA <sup>(3)</sup>	DESVIO PADRÃO <sup>(3)</sup>
(5;0,1)	20	500	0,2214	-0,1364	0,2563*	-0,1328	0,2639
(5;0,1)	10.000	10	-0,0855	-0,1001	0,0115	-0,1019	0,0124
(10;0,1)	20	100	0,1026	-0,1407	0,2708*	-0,1879	0,2340
(10;0,1)	10.000	10	-0,0862	-0,1001	0,0121	-0,0980	0,0092
(20;0,1)	20	100	0,3000	-0,1429	0,2778	-0,1694	0,2820
(20;0,1)	100	100	0,0444	-0,1086	0,1242		
(20;0,1)	20	100	0,2288	-0,1429			
(50;0,1)	20	100	0,2152	-0,1441	0,2819*	-0,1792	0,2720
(50;0,1)	100	100	0,0316	-0,1088	0,1261	-0,1142	0,1343
(50;0,016)	20	500	0,3500	-0,0643			
(10;0,2)	20	100	0,1000	-0,2362	0,2040*	-0,2633	0,2484
(10;0,2)	100	100	-0,0775	-0,2072	0,1081		
(20;0,2)	20	100	0,0951	-0,2381			
(20;0,2)	100	100	-0,0502	-0,2076	0,1106		
(10;0,3)	20	100	-0,0297	-0,3317			
(10;0,3)	100	100	-0,2065	-0,3063	0,0950		
(20;0,3)	20	100	-0,0524	-0,3333			
(20;0,3)	100	100	-0,1550	-0,3067	0,0970		
(10;0,4)	20	100	-0,1617	-0,4271			
(10;0,4)	100	100	-0,3304	-0,4054	0,0820		
(20;0,4)	20	100	-0,1933	-0,4286			
(5;0,5)	20	500	-0,3147	-0,5202			
(5;0,5)	20	500	-0,3276	-0,5202			
(10;0,5)	20	100	-0,3403	-0,5226	0,1541*	-0,5076	0,1615
(10;0,5)	100	100	-0,4220	-0,5045	0,0689	-0,5016	0,0565
(10;0,5)	20	500	-0,3307	-0,5226			
(50;0,5)	20	500	-0,3280	-0,5245			
(10;0,6)	20	100	-0,4374	-0,6181			
(10;0,6)	100	100	-0,5208	-0,6036	0,0559		
(10;0,7)	20	100	-0,6111	-0,7136			
(10;0,7)	100	100	-0,6500	-0,7027	0,0428		
(10;0,8)	20	100	-0,7088	-0,8090			
(10;0,8)	100	100	-0,7641	-0,8018	0,0297		
(5;0,9)	20	500	-0,8545	-0,9040	0,0412*	-0,9027	0,0394
(10;0,9)	20	500	-0,8909	-0,9045	0,0367*	-0,9036	0,0401

(1) Calculada segundo (4.9).

(2) Calculada segundo (4.14).

## COMENTÁRIOS

Como nesta distribuição  $\eta = -p$ , é de se supor que o critério discriminatório praticamente independa do parâmetro  $N$ , o que pode ser observado às páginas 63 e 64. Entretanto notamos uma pequena influência de  $N$  nos valores de  $E(\hat{\eta})$  na Tabela 4.2, isto é, à medida que  $N$  aumenta, a esperança decresce.

Ainda através da Tabela 4.2, concluímos que quando  $p = 0,1$ , obtem-se, quase sempre, pelo menos 10% de  $\hat{\eta}$ 's positivos, sendo que estes valores induziriam, pelo critério, a uma adaptação da Distribuição Binomial Negativa ou Poisson. Os casos mais críticos são aqueles provenientes de amostras de tamanho 20. Esta porcentagem não ocorreu somente nos dois casos em que as amostras eram de tamanho 10.000: parece razoável, dada a consistência de  $\hat{\eta}$ .

Para  $p = 0,2$ , só houve problemas com amostras de tamanho 20. Para os outros casos, a porcentagem de  $\hat{\eta}$ 's positivos é inferior ao apresentado no parágrafo anterior.

A situação tende a melhorar, à medida que aumentamos o valor de  $p$ : o percentil determinado aproxima-se de -1 quando  $p \rightarrow 1$ . Observa-se também que, para uma mesma distribuição, os valores de  $\hat{\eta}$  decrescem à medida que aumentamos o tamanho da amostra.

Em todos os casos, a mediana está à esquerda da média, o que indica uma assimetria da distribuição de  $\hat{\eta}$ .

Na mesma Tabela encontramos alguns valores exatos de  $E(\hat{\eta})$  e os valores do desvio padrão de  $\hat{\eta}$ , segundo (4.14). Para um mesmo valor de  $N$ , o valor esperado e o desvio padrão decrescem à medida que  $p$  cresce, enquanto que para um dado  $p$ , o desvio cresce com  $N$ . Além disso, para valores iguais de  $p$  o valor esperado decresce à medida que  $N$  cresce. Apesar da

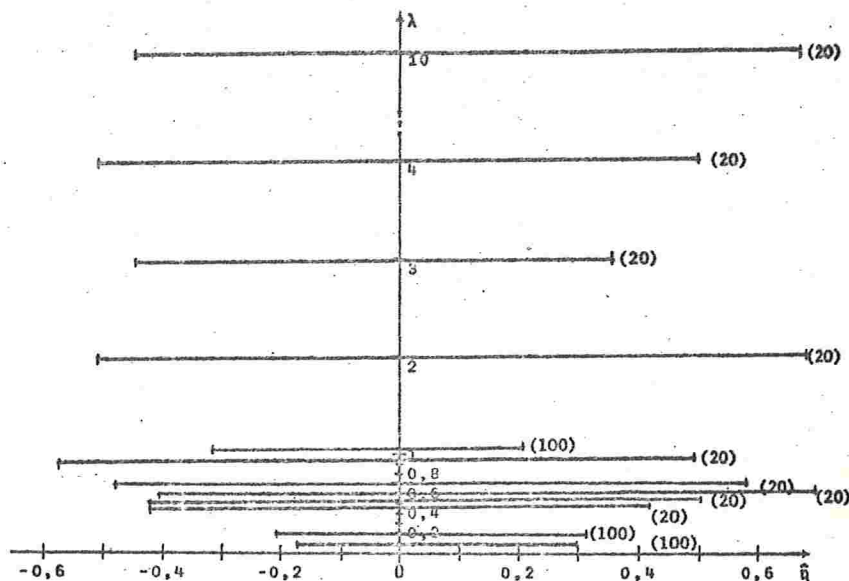
ordem de aproximação do resultado (4.14), calculamos também para alguns casos em que  $n = 20$  (assinalados com asterisco) porque tínhamos o desvio padrão amostral e gostaríamos de fazer comparações, ainda que pouco precisas. Sendo assim, apresentamos na Tabela 4.3, o desvio percentual em módulo ( $|P_1|$ ) de cada desvio padrão amostral com relação ao desvio padrão proposto por Katz. Fizemos o mesmo com relação às médias amostrais e  $E(\hat{n})$  (da Tabela 4.2), chamando  $|P_2|$  ao desvio correspondente.

TABELA 4.3

B(N;p)	n	$T_n$	$ P_1 $	$ P_2 $	B(N;p)	n	$T_n$	$ P_1 $	$ P_2 $
(5;0,1)	20	500	2,97	2,64	(50;0,1)	100	100	6,50	4,96
	10.000	10	7,83	1,80	(10;0,2)	20	100	21,76	11,47
(10;0,1)	20	100	13,59	33,55	(10;0,5)	20	100	4,80	2,87
	10.000	10	23,97	2,10		100	100	18,00	0,57
(20;0,1)	20	100	1,51	18,54	(5;0,9)	20	500	4,37	0,14
(50;0,1)	20	100	3,51	24,36	(10;0,9)	20	500	9,26	0,10

b) Distribuição de Poisson

Para este caso também construímos intervalos de tolerância para  $\hat{n}$ , com  $\gamma = \pi = 0,9$ , e os resultados podem ser resumizados no gráfico abaixo, onde  $(\lambda, n, T_n)$  é tal que  $\lambda$  é o



parâmetro da Distribuição de Poisson,  $n$  é o tamanho da amostra considerada e  $T_n$  é o número de amostras geradas.

Ainda aqui colocamos a porcentagem de valores de  $\hat{\eta}$  encontrados acima de zero.

Com o mesmo objetivo indicado para o caso da Distribuição Binomial, calculamos os percentis de ordem 5 e 95 ou seja, os valores que englobam os 90% centrais das observações (Tabela 4.4). Apresentamos também algumas médias e variâncias de  $\hat{\eta}$ , juntamente com a porcentagem dos valores de  $\hat{\eta}$ : a) maiores de -0,1; b) entre -0,2 e 0,2 exclusive.

TABELA 4.4

$\lambda$	$n$	$T_n$	90% CENTRAIS	$E(\hat{\eta})^{(1)}$	$\sqrt{V(\hat{\eta})}^{(2)}$	MÉDIA	DESVIO PADRÃO	a	b
0,1	100	100	-0,14 ; 0,2157	-0,01	0,1414	-0,0138	0,1315	72	92
0,2	100	100	-0,18 ; 0,25	-0,01	0,1414	-0,0137	0,1353	74	86
0,5	20	100	-0,4143; 0,4269	-0,05	0,3162	-0,0645	0,2387	59	54
0,5	20	500	-0,433 ; 0,443	-0,05	0,3162	-0,0549	0,2865	47	46
0,6	20	100	-0,425 ; 0,5071	-0,05	0,3162	-0,0528	0,3084	54	40
0,7	20	100	-0,4962; 0,4429	-0,05	0,3162	-0,0823	0,3236	46	40
0,8	20	10	-0,2277; 0,45	-0,05	0,3162	-	-	-	-
1,0	20	100	-0,4543; 0,3545	-0,05	0,3162	-0,0483	0,2596	61	51
1,0	100	100	-0,25 ; 0,1908	-0,01	0,1414	-0,0196	0,1319	74	85
2,0	20	100	-0,4402; 0,4160	-0,05	0,3162	-0,0538	0,2775	52	60
3,0	20	100	-0,4415; 0,3774	-0,05	0,3162	-0,0738	0,2540	53	46
4,0	20	100	-0,4143; 0,4366	-0,05	0,3162	-0,0102	0,3036	56	48
10,0	20	100	-0,4262; 0,5967	-0,05	0,3162	-0,0517	0,3018	45	44

(1) Calculada segundo (4.6).

(2) Calculada segundo (4.14).

### COMENTÁRIOS

Aqui temos  $\eta = 0$ , uma vez que  $\sigma^2 = \mu = \lambda$ , onde  $\lambda$  é o parâmetro da distribuição. Vemos que  $\eta$  independe de  $\lambda$  e supõe-se que  $\hat{\eta}$  também.

Apesar da grande maioria dos casos apresentar amostras de tamanho 20, o que justificaria talvez a grande flutuação dos valores  $\hat{\eta}$  em torno de zero, dada a ineficácia de  $\bar{X}$  e  $S^2$ , mesmo nos casos de amostras maiores (tamanho 100) no-

ta-se que os intervalos deixam muito a desejar quanto à precisão.

A partir da Tabela 4.4, notamos desvios padrão muito grandes em relação às médias, em todos os casos. As colunas a) e b) também apresentam resultados desfavoráveis, que possibilitariam confusão indevida com as Distribuições Binomial e Binomial Negativa. Sendo assim, pelo menos para pequenas amostras, achamos delicada a discriminação, dada a indesejável flutuação dos valores de  $\hat{\eta}$ .

A Tabela 4.5, similar àquela apresentada à página 67, oferece desvios bastante grandes tanto no que se refere a médias como a desvios padrão. Quanto a esses últimos, já seria de se esperar, dado que  $V(\hat{\eta})$ , segundo a expressão de Katz, é válida assintoticamente.

TABELA 4.5

$\lambda$	n	$T_n$	$ P_1 $	$ P_2 $
0,1	100	100	7,00	38,00
0,2	100	100	4,31	37,00
0,5	20	100	24,51	29,00
	20	500	9,39	9,80
0,6	20	100	2,47	5,60
0,7	20	100	2,34	64,60
1,0	20	100	17,90	3,40
	100	100	6,72	96,00
2,0	20	100	12,24	7,60
3,0	20	100	19,67	47,60
4,0	20	100	3,98	70,60
10,0	20	100	4,50	3,40

c) Distribuição Binomial Negativa

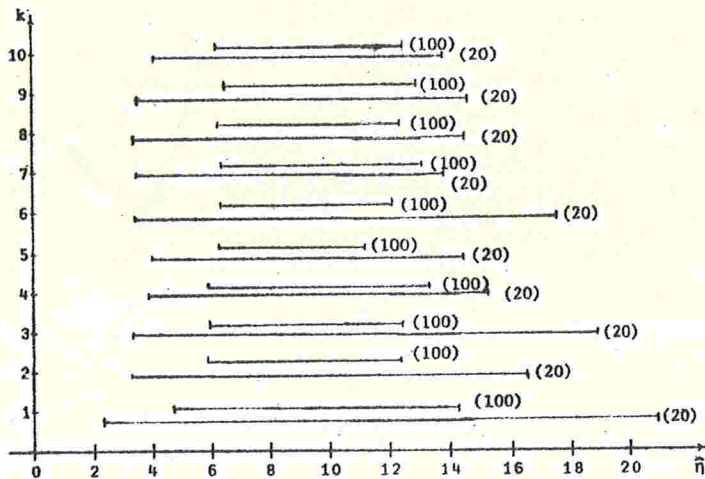
Foram escolhidas 176 combinações diferentes entre  $k$ ,  $p$  e  $n$  sendo que os dois primeiros são os parâmetros associados à Distribuição Binomial Negativa e  $n$  é o tamanho de



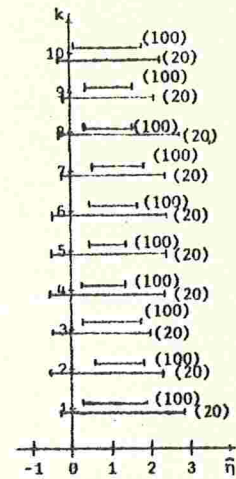
amostra para cada caso. Para cada combinação simulamos  $T_n$  distribuições do tipo Binomial Negativa e coletamos todos os valores de

$$\hat{\eta} = \frac{s^2 - \bar{x}}{\bar{x}}$$

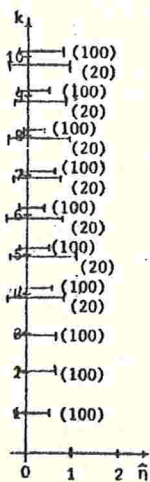
calculados. Em seguida, como nos casos anteriores, construímos intervalos de tolerância para  $\hat{\eta}$  com  $\pi = 0,90$ . A seguir apresentamos somente alguns destes intervalos, que cremos serem suficientes para termos uma idéia do comportamento do critério.



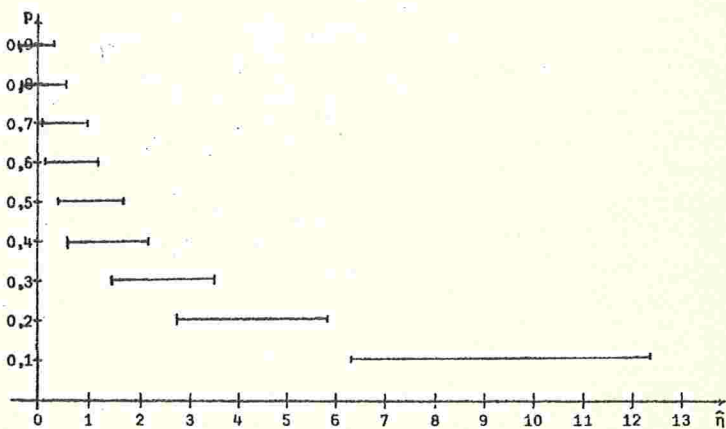
a)  $p=0,1$



b)  $p=0,5$



c)  $p=0,9$



d)  $k=6$   $n=100$

A Tabela 4.6, refere-se à determinação do percentil de ordem 10 (isto é, o valor que deixa 90% das observações à sua direita), para algumas das distribuições simuladas. Acrescentamos também alguns valores esperados e desvios padrão teóricos e, como os amostrais não foram calculados através da simulação, escolhemos algumas distribuições simuladas (que estão assinaladas com um número à direita da mesma tabela), e calculamos média e desvio padrão dos  $\hat{\eta}$ 's para cada caso. Os resultados estão na Tabela 4.7.

TABELA 4.6

k	p	n	$T_n$	$10^{\circ} \hat{\eta}$	$E(\hat{\eta})^*$	$\sqrt{V(\hat{\eta})}^{**}$	
1	0,1	20	100	3,6831	8,0476	2,2136	
	0,1	100		5,9871	8,8020		
	0,2	100		2,6641	3,9010		1,0954 (1)
	0,4	20		0,3000	1,2619		
	0,4	100		0,8265	1,4505		0,5362
1	0,7	100	100	0,0072	0,4003	0,2962	
	0,8	100		-0,0100	0,2252		0,2562 (2)
	0,9	100		-0,0900	0,0891		0,2250
4	0,1	20	100	4,5511	8,3827	1,6508	
	0,1	100		6,3267	8,8753		
4	0,2	100	100	3,0248	3,9376	0,8216	
	0,4	20		0,5222	1,3457		
	0,4	100		1,1059	1,4688		0,4070
	0,7	100		0,1522	0,4108		0,2292
	0,8	100		-0,0391	0,2344		0,1996
4	0,9	20	100	-0,4000	0,0425	0,1766	
	0,9	100		-0,0944	0,0972		
6	0,1	20	100	4,8933	8,4215	1,5759	
	0,1	100		6,9595	8,8835		0,7853 (3)
	0,2	100		3,0544	3,9417		
6	0,4	20	100	0,4630	1,3554	0,3900	
	0,4	100		0,9432	1,4709		
	0,7	100		0,1204	0,4119		0,2205
	0,8	100		0,0125	0,2354		0,1923 (4)
	0,9	20		-0,3500	0,0468		
6	0,9	100	100	-0,0999	0,0982	0,1703	
10	0,1	20	100	4,7860	8,4528	1,5133	
	0,1	100		6,9567	8,8901		
	0,2	100		2,9204	3,9451		0,7550
	0,4	20		0,6117	1,3632		
10	0,4	100	100	1,0371	1,4725	0,3758 (5)	
	0,7	100		0,0944	0,4128		0,2133
	0,8	100		0,0760	0,2362		0,1863
	0,9	20		-0,2881	0,0503		
	0,9	100		-0,0830	0,0989		0,1652

\* Calculada segundo (4.13).

\*\* Calculada segundo (4.14).

COMENTÁRIOS

Aqui  $\eta = \frac{1-p}{n}$  e independe de  $k$ , sendo provável que  $\hat{\eta}$  também independa desse parâmetro.

A partir dos gráficos conclui-se que não há problemas discriminatórios para grandes valores de  $p$ . Além disso, percebemos claramente que a variância de  $\hat{\eta}$  é inversamente proporcional a  $p$  e independe de  $k$ . Infelizmente, quando o valor de  $p$  tende à unidade (variância mínima), é que ocorrem problemas de discriminação entre as Distribuições de Poisson e a Binomial Negativa, pois os respectivos valores de  $\hat{\eta}$  estão em torno do zero, sendo que nem sempre são válidas as condições de convergência analisadas no Capítulo 2. O último gráfico refere-se a um valor fixo de  $k$ , onde se percebe nitidamente a influência de  $p$  sobre o valor de  $\hat{\eta}$  e sobre sua variância.

Os cinco últimos percentis apresentados na Tabela 4.7, apresentam problemas discriminatórios, segundo o critério, uma vez que há valores de  $\hat{\eta}$  que poderiam ser interpretados como provenientes de uma Poisson ou Binomial.

TABELA 4.7

	MÉDIA	DESVIO PADRÃO	$ P_1 $	$ P_2 $
(1)	3,7660	0,8844	19,26	3,46
(2)	0,2110	0,1871	26,97	6,31
(3)	3,9565	0,7171	8,68	0,38
(4)	0,2254	0,1642	14,61	4,25
(5)	1,4627	0,3339	11,15	0,67

Após todas essas simulações não nos pareceu que o método proposto por Katz seja eficiente para pequenas amostras, a não ser em casos onde a probabilidade de sucesso é

muito grande na Distribuição Binomial ou muito pequena na Distribuição Binomial Negativa. Evidentemente, como já mencionamos, isto se deve à pouca eficiência do estimador de  $\eta$  para  $n$  pequeno.

Tentando nos colocar no papel do pesquisador em áreas aplicadas, simulamos algumas distribuições de cada um dos três tipos, e os resultados obtidos estão apresentados na Tabela 4.8.

Para alguns casos, comparamos, probabilisticamente, a distribuição gerada com aquela que seria sugerida pelo critério ao pesquisador, cabendo a este último analisar se as diferenças encontradas nas caudas ou no meio das distribuições são significantes, ao nível de sua análise. Passaremos em seguida às comparações:

- 1)  $B(10;0,1)$   
 $n = 100$   
 $\hat{\eta} = 0,1027$

O critério sugere o uso da Poisson ( $\lambda$ ) ou da Binomial Negativa ( $k;p$ ).

Para o caso da Poisson, o parâmetro estimado através da amostra gerada foi  $\hat{\lambda} = 1,1$  e então as comparações foram feitas entre Poisson (1,1) e  $B(10;0,1)$ . Encontramos 0,0632 como soma dos valores absolutos das diferenças entre as probabilidades das duas distribuições. As maiores diferenças individuais são encontradas nos valores: 0: -0,0158; 1: -0,0212; 3: 0,0164. O nível descritivo associado ao valor 4 da variável aleatória com distribuição  $B(10;0,1)$  é 0,0128 e o associado à variável aleatória com distribuição  $P(1,1)$  é 0,0257, isto é, o dobro. Comparando  $P(1)$  com  $B(10;0,1)$ , a soma dos valores absolutos das diferenças seria 0,0584 e o nível descritivo associado ao valor 4 na  $P(1)$  seria  $0,02 < 0,0257$ . Também nes

te caso as maiores diferenças individuais são encontradas nos valores 0 e 1 da variável aleatória.

Se adaptarmos o modelo da Distribuição Binomial Negativa, com  $\hat{k} = 11$  e  $\hat{p} = 0,91$  (valores obtidos da amostra gerada através do método dos momentos, mais rápido embora menos preciso do que o de máxima verossimilhança), obtemos 0,0817, como soma dos valores absolutos das diferenças entre  $B(10;0,1)$  e  $\text{Bineg}(11;0,91)$ , valor esse bem maior que os anteriores. Além disso, o nível descritivo associado ao ponto 4 vale 0,03 e a maior diferença individual se encontra no valor 1 da variável aleatória e é igual a 0,0366. Diferenças individuais sensíveis também são encontradas no meio da distribuição.

Se fôssemos usar o critério sugerido por Ord, conforme descrito à página 51, teríamos o seguinte resultado:

$$\left. \begin{array}{l} \hat{\mu}_3 = 1,384758 \\ \hat{\mu}_2 = 1,2019 \\ \hat{\mu}_1 = 1,09 \end{array} \right\} \begin{array}{l} \hat{I} = 1,1027 \\ \hat{S} = 1,1521407 \end{array}$$

Esses valores apresentam o mesmo problema discriminatório de  $\hat{\eta}$ , o que seria de se esperar, uma vez que, como vimos no capítulo anterior,  $S = 2\eta + 1$ . Evidentemente, como  $\hat{\eta}$  não é o estimador de máxima verossimilhança de  $\eta$ , não temos  $\hat{S} = 2\hat{\eta} + 1$ ; entretanto cremos que a diferença entre  $\hat{2\eta+1}$  e  $2\hat{\eta}+1$  não seja tão grande a ponto de melhorar o critério discriminatório.

2) Poisson (0,5)

$$\hat{n} = 20$$

$$\hat{\eta} = -0,2278$$

A partir desse valor de  $\hat{\eta}$  pensaríamos em adaptar u

ma Distribuição Binomial. No presente caso, através dos valores simulados, teríamos obtido  $\hat{p} = 0,23$  e  $\hat{n} = 2$ . A soma dos valores absolutos das diferenças entre  $B(2;0,23)$  e  $P(0,5)$  é 0,0875 e a maior diferença individual se encontra no valor 1 e vale 0,050935.

3) Poisson (6)

$$n = 100$$

$$\hat{\eta} = 1,877$$

Sabemos que este valor foi obtido da geração de uma Poisson com  $\lambda = 6$ , mas o resultado sugere a adaptação de uma Distribuição Binomial e, neste caso, os parâmetros estimados foram:  $\hat{p} = 0,19$  e  $\hat{n} = 31$ . Comparando as duas distribuições mencionadas teríamos 0,1028 como soma dos valores absolutos das diferenças entre probabilidades associadas ao mesmo valor da distribuição. A maior diferença individual encontra-se no valor 7 e vale 0,0179. O nível descritivo associado ao valor 12 vale 0,2 na Poisson e 0,0085 na Binomial; o associado ao valor 11 vale 0,0425 na Poisson e 0,0231 na Binomial.

4) Poisson (4)

$$n = 100$$

$$\hat{\eta} = 0,306$$

O resultado sugere a adaptação de uma Distribuição Binomial Negativa. Nesse caso, os parâmetros estimados através da amostra seriam:  $\hat{p} = 0,77$  e  $\hat{k} = 13,5$ . Como tivemos acesso a tabelas, com  $k = 13$  e  $k = 14$ , analisamos esses dois casos: a)  $\hat{p} = 0,77$  e  $\hat{k} = 13$ : a soma dos valores absolutos das diferenças entre as probabilidades dessa Binomial Negativa e uma Poisson com parâmetro  $\lambda = 4$  resultou 0,1457, ocorrendo diferenças maiores de  $\frac{1}{100}$  para os valores de 0 a 6; b)  $\hat{p} = 0,77$

e  $\hat{k} = 14$ : a referida soma vale, neste caso, 0,1235 e as diferenças maiores de 0,01 se encontram nos valores 3, 4, 5 e 8 da variável aleatória. Para a Poisson com  $\lambda = 4$  temos

$$P(X \geq 8) = 0,051.$$

No caso a) esta probabilidade vale 0,0675 e no b) vale 0,0868.

5) Binomial Negativa ( $k=1$ ;  $p = 0,7$ )

$$n = 100$$

$$\hat{\eta} = 0,07176$$

O pesquisador, mediante tal valor, estaria inclinado a adaptar uma Poisson. Através da amostra gerada teríamos  $\hat{\lambda} = 0,3$ . A soma dos valores absolutos das diferenças entre as probabilidades associadas à Binomial Negativa (1;0,7) e à Poisson (0,3) é 0,1061, sendo que a maior diferença ocorre à esquerda da distribuição, para o valor zero. O nível descritivo associado ao valor 2 é igual a 0,037 para Poisson e 0,09 para a Binomial Negativa.

TABELA 4.8

DISTRIBUIÇÃO	n	$\hat{\eta}$	DISTRIBUIÇÃO	n	$\hat{\eta}$
B(10;0,1)	20	-0,2	P(6)	20	0,633
	100	0,1027		100	0,1877
B(20;0,2)	20	-0,2	P(10)	100	-0,148
	100	-0,2826	Bineg(1;0,1)	20	4,1738
B(10;0,3)	100	-0,1849		100	8,7414
P(0,5)	20	-0,2278	Bineg(1;0,3)	100	2,8228
	10.000	-0,0068	Bineg(1;0,4)	100	0,9668
P(1)	20	0,43	Bineg(1;0,5)	10.000	0,9825
	100	-0,039		100	0,6427
	1.000	-0,0543	Bineg(1;0,7)	10	-0,2667
P(2)	20	0,2091	Bineg(1;0,7)	20	-0,1
	100	-0,183		100	0,0718
P(4)	20	-0,2714	Bineg(1;0,8)	100	0,3
	100	0,306	Bineg(1;0,9)	20	-0,1

## APÊNDICE

Seguem abaixo alguns dos programas em "BASIC" utilizados nas simulações deste trabalho.

### 1. Distribuição Binomial

```
00001 FILE #1,"BINOMI.LPT"
00002 SCRATCH #1
00005 DIM Y (1000)
00010 FOR N=10 TO 20 BY 10
00020 FOR P=.1 TO .9 BY .1
00030 FOR M=20 TO 100 BY 80
00040 U=100
00050 RANDOMIZE
00120 FOR J=1 TO U
00130 X1=X2=X3=M1=V=0
00150 DIM E(>J0)
00180 FOR I=1 TO M
00190 S=X=0
00200 G=RND
00210 IF G>P THEN 260
00220 S=S+1
00230 X=X+1
00240 IF X>=N THEN 290
00250 GO TO 230
00260 X=X+1
00270 IF X>=N THEN 290
00280 GO TO 240
00290 X1=X1+S
00300 X2=X2+S**2
00310 Y(S)=Y(S)+1
00330 NEXT I
00340 M1=X1/M
00350 V=(X2/M-M1**2)
00380 E=(V-M1)/M1
00410 E(J)=E
00420 NEXT J
00430 FOR I=1 TO U-1
00440 FOR J=I+1 TO U
00450 IF E(I)<=E(J) THEN 490
00460 C1=E(I)
00470 E(I)=E(J)
00480 E(J)=C1
00490 NEXT J
00500 NEXT I
00502 PRINT #1
00503 PRINT #1,"M=","M","P=","P","N=","N
00504 PRINT #1,
00510 PRINT #1,"VALORES DE ETA"
00511 FOR K1=1 TO U BY 5
00512 PRINT #1,E(K1),E(K1+1),E(K1+2),E(K1+3),E(K1+4)
00513 NEXT K1
00530 NEXT M
00540 NEXT P
00550 NEXT N
00600 END
```



## 2. Distribuição Binomial Negativa

```
R BASIC

READY, FOR HELP TYPE HELP
OLD BINEG.BAS

READY
LISTNH
10 FILE #1,"BINEG.LPT"
20 SCRATCH #1
30 DIM P(1000),E(1000),X(1000)
40 DIM S(1000)
50 PRINT #1,"SIMULACAO POR MONTE CARLO DA BINOMIAL NEGATIVA"
60 M=1
70 K=1
80 INPUT P
90 PRINT #1,
100 X=0
110 I=1
120 P(I)=P**K
130 S(I)=P(I)
140 P(I+1)=(1-P)*P(I)*(X+K)/(X+1)
150 S(I+1)=S(I)+P(I+1)
160 IF S(I+1)>=0.9999 THEN 210
170 X=X+1
180 N1=X
190 I=I+1
200 GO TO 140
210 V=I+1
220 RANDOMIZE
230 INPUT N
240 FOR U=1 TO M
250 MAT X=ZER(V)
260 DIM W(1000)
270 S1=S2=0
280 FOR W=1 TO N
290 G=RND
300 IF G<=0.9999 THEN 320
310 GO TO 290
320 J=1
330 IF G<S(J) THEN 360
340 J=J+1
350 GO TO 330
360 X(J)=X(J)+1
370 NEXT W
380 FOR C=1 TO V
390 S1=S1+X(C)*(C-1)
400 S2=S2+((C-1)**2)*X(C)
410 NEXT C
420 R1=S1/N
440 R2=(N*S2-S1**2)/(N**2)
450 E(U)=(R2-R1)/R1
460 NEXT U
470 PRINT #1,"K="K,"P="P,"N="N
480 PRINT #1,<PA>
490 PRINT #1,"VALORES DE ETA"
510 PRINT#1,E(1)
540 PRINT
550 MAT W=ZER(1000)
560 PRINT#1,"K", "PROB ", "PROB. AC. ", "PROB. SIM. ", "PROB. AC. SIM. "
570 FOR K1=1 TO V
580 W(K1)=W(K1-1)+X(K1)
590 PRINT#1,K1,P(K1),S(K1),X(K1)/N,W(K1)/N
600 NEXT K1
610 END
```

READY

MON

### BIBLIOGRAFIA

- [1] - Anderson, T.W e Samuels, S.M. - Some inequalities among Binomial and Poisson probabilities - Proceedings of the 5<sup>th</sup> Berkeley Symposium on Mathematical Statistics and Probability, 1, (1965); 1-12.
- [2] - Anscombe, F.J. - The transformation of Poisson, Binomial and Negative Binomial Data - Biometrika, 35, (1948); 246-254.
- [3] - Bartko, J.J. - Approximating the Negative Binomial, Biometrics, 8 (1966); 340-342.
- [4] - - The negative binomial distribution: a review of properties and applications. Virg. J. Sci., 12, (1961); 18-37.
- [5] - Bartlett, M.S. - The use of transformations. Biometrics, 3, (1947); 39-52.
- [6] - Bhattacharya, S.K. et al - On a discrete distribution with special reference to the theory of accident proneness. JASA, 60, (1965); 1060-1066.
- [7] - Binns, M.R. - Approximating the negative binomial via the positive binomial. Technometrics, 16, (1974); 323-324.
- [8] - Bliss, C.I. et al - Fitting the Negative Binomial Distribution to Biological Data. Biometrics, 9, (1953) - 176-200.

- [9] - Cramer, H. - Mathematical Methods of Statistics. Princeton University Press, (1957).
- [10] - Evans, D.A. - Experimental evidence concerning contagious distributions in ecology. Biometrika, 40, (1953); 186-211.
- [11] - Feller, W. - On a general class of contagious distributions. Ann. Math. Stat., 14, (1943); 389-400.
- [12] - Feller, W. - An introduction to probability theory and its applications. John Wiley and Sons, (1968).
- [13] - Fernandez, P. - Introdução à Teoria das Probabilidades. Monografias de Matemática nº 6, IMPA (1971).
- [14] - Fisher, R.A. et al - The relation between the number of species and the number of individuals in a random sample of animal population. J. An. Ecology, 12, (1943).
- [15] - Fisz, M. - Probability Theory and Mathematical Statistics. John Wiley and Sons, (1963).
- [16] - Guenther, W.C. - A simple approximation to the negative binomial (and regular binomial). Technometrics, 14, (1972); 385-389.
- [17] - Gurland, J. - Some interrelations among compound and generalized distributions. Biometrika, 44, (1957); 265-268.
- [18] - Hannan, J e Harkness, W. - Normal approximation to the distribution of two independent binomials, conditional on a fixed sum. Ann. Math. Stat., 34, (1963); 1593-1595.
- [19] - Hinz, P. e Gurland, J. - A method of analysing untransformed data from the negative binomial and other contagious distributions. Biometrika, 55, (1968); 163-170.

- [20] - Ishii, G. e Hayakawa, R. - On the compound binomial distribution. Annals of the Institute of Statistical Math., Tokyo, 12 (1960); 69-80.
- [21] - Johnson, N.L. - Discrete Distributions. Houghton Mifflin Company, (1969).
- [22] - Johnson, N.L. - Uniqueness of a result in the theory of accident proneness. Biometrika, 44, (1957).
- [23] - Katti, S.K. e Gurland, I. - The Poisson Pascal Distribution. Biometrics, 17, (1961); 527-538.
- [24] - Katz, L. - Unified treatment of a broad class of discrete probability distributions. Proceedings of the International Symposium - Montreal, (1963); 175-182.
- [25] - Koopman, B.D. - Necessary and sufficient conditions for Poisson's Distribution. Proc. Amer. Math. Soc., 1, (1950); 813-823.
- [26] - McGuire, J.U. et al - The distribution of the European corn-borer larvac (Hbn), in field Corn. Biometrics, 13, (1947); 65-78.
- [27] - Neyman, J. - On a new class of contagious distributions, applicable in entomology and bacteriology. Ann. Math. Stat., 10, (1939); 35-57.
- [28] - Ord, J.K. - Families of frequency distributions. Griffin, London, (1972).
- [29] - Ord, J.K. - On a system of discrete distributions. Biometrika, 54, (1967); 649-655.
- [30] - Patil, G.P. - On the evaluation of the negative binomial distribution with examples. Technometrics, 2, (1960); 501-505.
- [31] - - Certain properties of the generalized power series distributions - Annals of the Institute

- te Statistics Mathematics, Tokyo, 14, (1962); 179-182.
- [32] - Raff, M.S. - On approximating the point binomial. JASA, 51, (1956); 293-303.
- [33] - Sommerville, R. - Tables for obtaining non-parametric tolerance limits. Annals of Math. Stat., 29, (1958); 599-560.
- [34] - Walsh, J.E. - The Poisson Distribution as a limit for dependent binomial events with unequal probabilities. Operations Research, 3, (1955); 198-204.
- [35] - Anscombe, F.J. - Sampling theory of the negative binomial and logarithmic series distributions. Biometrika, 37, (1950); 358-381.
- [36] - Pearson, K. - Tables of the Incomplete Beta Functions. Cambridge University Press, London.
- [37] - Roming, H.G. - 50-100 Binomial Tables. John Wiley and Sons, (1953).