Pedro Alberto Morettin

INTRODUÇÃO À ESTATÍSTICA

PREFÁCIO

O objetivo destas notas é fornecer subsídios para um primeiro curso de Estatística. Não há requisito prévio, exceto um curso de álgebra elementar. Somente em raras ocasiões fazemos uso de noções de limite e integral.

São óbvias as dificuldades encontradas ao se escrever um livro em nível elementar. Em primeiro lugar, há de se decidir "quão elementar" deve ser o livro. Depois, há o problema do que deve ser incluído. Quanto ao primeiro aspecto, devo frisar que estas notas não constituem um conjunto de métodos matemáticos da Estatística, mas sim uma abordagem inicial, na maior parte das vezes heurística, dos princípios da Estatística. Dada a exigüidade de tempo, não incluímos tópicos que pretendíamos abordar: alguns testes de hipóteses usuais para distribuições normais e uma introdução ao enfoque bayesiano.

Ênfase foi dada à distribuição binomial, para introduzir as idéias básicas de estimação e testes de hipóteses. Alguns tópicos mais difíceis ou especializados estão no apêndice.

Finalmente, gostaria de agradecer à Comissão Orga-

nizadora do 10º Colóquio Brasileiro de Matemática, por ter nos convidado para ministrar este curso, a Clóvis de A. Peres, por ter lido os originais e apresentado várias sugestões e ao Sr. João Baptista Esteves de Oliveira, pelo excelente trabalho de datilografia.

São Paulo, maio de 1975

Pedro A. Morettin

ÍNDICE

- o -

1. PRELIMINARES SOBRE PROBABILIDADES

Neste capítulo vamos introduzir alguns conceitos básicos de probabilidades de maneira simples e utilizando vários exemplos. Vamos nos restringir ao caso de espaços amostrais discretos, pois pretendemos introduzir as idéias principais de estimação e testes de hipóteses para parâmetros de uma distribuição discreta de probabilidades.

1.1 - MODELO PROBABILÍSTICO

Todo experimento ou fenômeno que envolva um elemento casual ficará especificado no momento que estabelecemos:

(i) um *espaço amostral*, Ω, que consiste, no caso discreto, da enumeração (finita ou infinita) de todos os resultados possíveis do experimento em questão: $\Omega=\{\omega_1,\omega_2,\ldots\}$. Os elementos ω são os *pontos amostrais*;

(ii) uma *probabilidade*, $P(\omega)$, para cada ponto amostral, de tal sorte que seja possível encontrar a probabilidade $P(A)$ de qualquer subconjunto A de Ω, isto é, a probabilidade do que chamaremos um *evento*.

Exemplo 1.1 - Uma urna contém duas bolas brancas (B) e três bolas vermelhas (V). Retirando-se uma bola ao acaso da urna, a probabilidade de se obter uma bola branca será $P(B) = \frac{2}{5}$ e a probabilidade de se obter uma bola vermelha será $P(V) = \frac{3}{5}$. Aqui, $\Omega = \{B, V\}$.

<u>Exemplo 1.2</u> - Suponha que lançamos uma moeda duas vezes. Se C indica cara e R indica coroa, então um espaço amostral será $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$, onde $\omega_1 = (C,C)$, $\omega_2 = (C,R)$, $\omega_3 = (R,C)$, $\omega_4 = (R,R)$. É razoável supor que cada ponto ω tem probabilidade $\frac{1}{4}$, se a moeda é perfeitamente simétrica e homogênea.

Se A é o evento que consiste na obtenção de faces iguais nos dois lançamentos, então

$$P(A) = P\{\omega_1, \omega_4\} = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}.$$

De um modo geral, se A é um evento, então

$$P(A) = \sum P(\omega_j), \tag{1.1}$$

onde a soma é estendida a todos os $\omega_j \in A$.

Vale, aqui, uma observação sobre espaços amostrais. Para um mesmo experimento podemos ter vários espaços amostrais, dependendo do objetivo do problema que se quer estudar. Por exemplo, suponha que lancemos uma moeda 5 vezes. Se estamos interessados apenas na seqüência de caras e coroas obtida, um espaço amostral é $\Omega_1 = \{(x_1, \ldots, x_5): x_i = 0 \text{ ou } x_i = 1, i = 1, \ldots, 5\}$. Mas se estamos interessados no número de caras obtidas, então um espaço amostral mais conveniente é $\Omega_2 = \{0, 1, 2, 3, 4, 5\}$.

1.2 - <u>ALGUMAS PROPRIEDADES</u>

Suponha que lancemos uma moeda; costumamos dizer que a probabilidade de cara ou coroa é $\frac{1}{2}$. Podemos dar a seguinte explicação em termos de freqüência relativa. Considere 10 lançamentos da moeda; se obtemos 6 caras, então $\frac{6}{10} = 0,6$ é a freqüência relativa de ocorrência de caras nestes 10 lançamentos. Suponha que aumentemos progressivamente o número de lançamentos, e para cada um deles calculemos a freqüência relativa. Esta variará de prova para prova mas a medida que aumentamos o número de lançamentos ela tenderá a se estabilizar ao redor de um número, que será a probabilidade de "cara"; este número será $\frac{1}{2}$ se a moeda for "honesta".

Esta interpretação freqüencial poderá ser útil ao leitor, se quizer verificar heuristicamente as propriedades que passamos a considerar, em especial a regra de adição. Em particular, como toda freqüência relativa é um número entre 0 e 1, vemos que $0 \leq P(A) \leq 1$, para qualquer evento A. Para efeito de completividade, será útil considerarmos o espaço todo Ω e o conjunto vazio, $\emptyset$, como eventos. O primeiro é denominado *evento certo* e $P(\Omega)=1$, enquanto que $\emptyset$ é o *evento impossível* e $P(\emptyset)=0$.

<u>Exemplo 1.3</u> - Suponha que o seguinte quadro represente uma possível divisão dos alunos matriculados em dado instituto de matemática, num dado ano. (Vide quadro na pág. seguinte).

Vamos indicar por P o evento que ocorre quando, escolhendo-se ao acaso um aluno do instituto, ele for um estudante de matemática pura. A, E, C, H e M têm significados a

	HOMENS	MULHERES	TOTAIS
MATEMÁTICA PURA . .	70	40	110
MATEMÁTICA APLICADA	15	15	30
ESTATÍSTICA.	10	20	30
COMPUTAÇÃO.	20	10	30
TOTAIS.	115	85	200

Tabela 1.1

nálogos. Desta maneira vemos que $P(E) = \frac{30}{200}$, ao passo que $P(H) = \frac{115}{200}$.

Dado os eventos A e H podemos considerar dois novos eventos:

$A \cup H$, chamado a *reunião* de A e H, que ocorre quando pelo menos um dos eventos ocorre;

$A \cap H$, chamado a *intersecção* de A e H, que ocorre quando A e H ocorrem simultaneamente.

É fácil ver que $P(A \cap H) = \frac{15}{200}$, pois o aluno escolhido terá que ser ao mesmo tempo, matriculado no curso de matemática aplicada e ser homem.

Vemos que $P(A) = \frac{30}{200}$ e $P(H) = \frac{115}{200}$;suponha que nosso cálculo para $P(A \cup H)$ fosse

$$P(A \cup H) = P(A) + P(H) = \frac{30}{200} + \frac{115}{200} = \frac{145}{200}.$$

Se assim o fizéssemos, estaríamos contando duas vezes os alunos que são homens *e* estão matriculados no curso de matemática aplicada, como está destacado na tabela 1.1. Portanto, a resposta correta é

$$P(A \cup H) = P(A) + P(H) - P(A \cap H)$$

$$= \frac{30}{200} + \frac{115}{200} - \frac{15}{200} = \frac{130}{200}.$$

No entanto, considerando-se os eventos A e C, vemos que $P(A) = \frac{30}{200}$, $P(C) = \frac{30}{200}$ e $P(A \cup C) = \frac{60}{200} = P(A) + P(C)$. Neste caso, os eventos A e C são disjuntos ou *mutuamente exclusivos*, pois se A ocorre então C não ocorre e vice-versa. Aqui, $A \cap C = \emptyset$ e $P(A \cap C) = 0$.

Portanto, se M e N são dois eventos quaisquer temos a chamada *regra da adição de probabilidades*

$$P(M \cup N) = P(M) + P(N) - P(M \cap N), \tag{1.2}$$

que se reduz a

$$P(M \cup N) = P(M) + P(N), \tag{1.3}$$

se M e N são eventos mutuamente exclusivos.

Suponha, agora, que estamos somente interessados em saber se um estudante escolhido ao acaso está matriculado como aluno de matemática pura, aplicada, estatística ou computação, não interessando saber se é homem ou mulher. Um espaço amostral é $\Omega = \{P, A, E, C\}$. Os eventos A e $B = \{P, E, C\}$ são chamados *complementares* e são tais que $A \cup B = \Omega$ e $A \cap B = \emptyset$. Vemos que $P(A) = \frac{30}{200}$, enquanto $P(B) = \frac{110}{200} + \frac{30}{200} + \frac{30}{200} = \frac{170}{200}$, isto é,

$$P(A) + P(B) = 1. \tag{1.4}$$

Em geral, vamos indicar por A^c o complementar de um evento A e teremos, então,

$$P(A^c) = 1 - P(A). \tag{1.5}$$

1.3 - PROBABILIDADE CONDICIONAL E INDEPENDÊNCIA

Voltemos ao quadro do exemplo 1.3. Dado que um estudante, escolhido ao acaso, esteja matriculado no curso de Estatística, a probabilidade dele ser mulher é $\frac{10}{30} = \frac{1}{3}$. Isto porque do total de 30 alunos que estudam estatística, 10 são mulheres. Escrevemos

$$P(\text{mulher} \mid \text{estatística}) = \frac{1}{3}.$$

Para dois eventos quaisquer A e B, sendo $P(B) > 0$, definimos a *probabilidade condicional de* A, *dado* B, $P(A|B)$, como sendo

$$P(A|B) = \frac{P(A \cap B)}{P(B)}. \tag{1.6}$$

Para o exemplo mencionado, se B e A indicam, respectivamente, os eventos "aluno matriculado em Estatística" e aluno é mulher", então $P(A \cap B) = 10/200$, $P(B) = 30/200$ e, portanto,

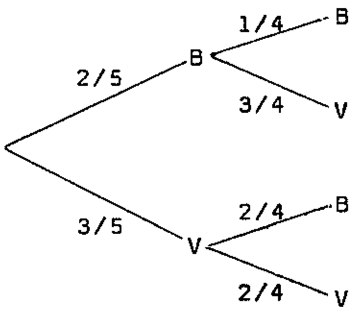
$$P(A|B) = \frac{10/200}{30/200} = \frac{1}{3},$$

como havíamos obtido.

Da relação (1.6) obtemos a chamada *regra do produto de probabilidades*,

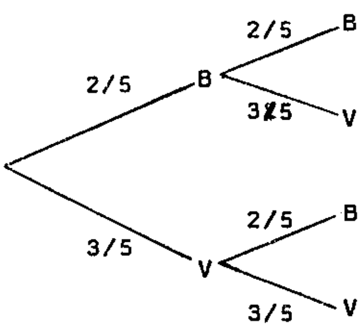
$$P(A \cap B) = P(B) \cdot P(A|B). \tag{1.7}$$

Exemplo 1.4 - Retomemos a urna do exemplo 1.1 mas suponha agora que extraímos duas bolas ao acaso, *sem reposição*. Isto significa que escolhemos a primeira bola, verificamos a sua cor e *não* a devolvemos à urna, misturamos as bolas restantes e retiramos a segunda bola. O diagrama em árvore abaixo ilustra as possibilidades. Em cada "galho" da árvore estão indicadas as probabilidades de ocorrência, sendo que, para as segundas bolas temos probabilidades condicionais. A probabilidade do resultado conjunto é, então, dada por (1.7).



RESULTADOS	PROBABILIDADES
BB	$\frac{2}{5} \times \frac{1}{4} = \frac{2}{20}$
BV	$\frac{2}{5} \times \frac{3}{4} = \frac{6}{20}$
VB	$\frac{3}{5} \times \frac{2}{4} = \frac{6}{20}$
VV	$\frac{3}{5} \times \frac{2}{4} = \frac{6}{20}$
SOMA	1

<u>Exemplo 1.5</u> - Imagine, agora, que as duas extrações são feitas da urna, mas a primeira bola é *reposta* na urna antes da extração da segunda bola. Nestas condições, as extrações são *independentes*, no sentido que o resultado de cada extração não tem influência no resultado da outra. Obtemos a situação a seguir.



RESULTADOS	PROBABILIDADES
BB	$\frac{2}{5} \times \frac{2}{5} = \frac{4}{25}$
BV	$\frac{2}{5} \times \frac{3}{5} \quad \frac{6}{25}$
VB	$\frac{3}{5} \times \frac{2}{5} \quad \frac{6}{25}$
VV	$\frac{3}{5} \times \frac{3}{5} \quad \frac{9}{25}$
SOMA	1

Observe que, aqui,

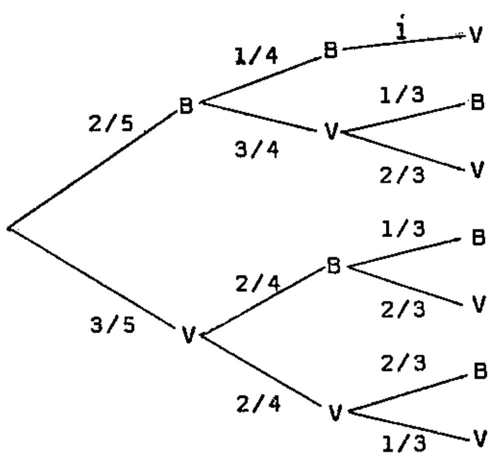
$$P(\text{branca na } 2^{\underline{a}} \mid \text{branca na } 1^{\underline{a}}) = \frac{2}{5} = P(\text{branca na } 2^{\underline{a}}),$$

Ou seja, se dois eventos A e B são *independentes*, $P(A|B) = P(A)$ e $P(B|A) = P(B)$. Se assim é, a relação (1.7) fica para A e B independentes:

$$P(A \cap B) = P(A) \cdot P(B). \qquad (1.8)$$

A fórmula (1.8) pode ser tomada como definição de independência: A e B são independentes se e somente se (1.8) é válida.

Exemplo 1.6 - Considere, ainda, a mesma urna dos exemplos anteriores, mas vamos fazer três extrações *sem reposição*. Obtemos o esquema abaixo.



RESULTADOS	PROBABILIDADES
BBV	$\frac{2}{5} \times \frac{1}{4} \times 1 = \frac{2}{20} = \frac{6}{60}$
BVB	$\frac{2}{5} \times \frac{3}{4} \times \frac{1}{3} = \frac{6}{60}$
BVV	$\frac{2}{5} \times \frac{3}{4} \times \frac{2}{3} = \frac{12}{60}$
VBB	$\frac{3}{5} \times \frac{2}{4} \times \frac{1}{3} = \frac{6}{60}$
VBV	$\frac{3}{5} \times \frac{2}{4} \times \frac{2}{3} = \frac{12}{60}$
VVB	$\frac{3}{5} \times \frac{2}{4} \times \frac{2}{3} = \frac{12}{60}$
VVV	$\frac{3}{5} \times \frac{2}{4} \times \frac{1}{3} = \frac{6}{60}$
SOMA	$60/60 = 1$

Observe que $P(B|B) = \frac{1}{4}$, ao passo que $P(V|B \text{ e } B) = 1$; daí

$$P(B \text{ e } B \text{ e } V) = P(B) \cdot P(B|B) \cdot P(V|B \text{ e } B) = \frac{2}{5} \times \frac{1}{4} \times 1 = \frac{1}{10}.$$

De modo geral, dados 3 eventos M, N e R, temos que

$$P(M \cap N \cap R) = P(M) \cdot P(N|M) \cdot P(R|M \cap N) \qquad (1.9)$$

PROBLEMAS PARA O CAPÍTULO 1

1.1 - Duas moedas são lançadas. Dê dois possíveis espaços amostrais para este experimento. Represente um destes como o *produto cartesiano* de dois outros espaços amostrais.

1.2 - Uma moeda e um dado são lançados. Dê um espaço amostral do experimento e depois represente-o como produto cartesiano dos dois espaços amostrais correspondente aos experimentos considerados individualmente.

1.3 - No exercício 1.1, liste os eventos:

a) pelo menos uma cara;

b) duas caras;

c) o complementar de b).

1.4 - Considere o lançamento de dois dados. Considere os eventos A: soma dos números obtidos igual a 9 e B: número no primeiro dado maior ou igual a 4. Enumere os elementos de A e B. Obtenha $A \cup B$, $A \cap B$ e A^c.

1.5 - Obtenha as probabilidades dos eventos que aparecem nos problemas 1.3 e 1.4.

1.6 - Considere uma urna contendo 3 bolas pretas e 5 bolas vermelhas. Retire duas bolas da urna, sem reposição. Obtenha os resultados possíveis e as respectivas probabilidades. Mesmo problema, para extrações com reposição.

1.7 - No problema anterior, calcule as probabilidades dos eventos:

a) bola preta na primeira e segunda extrações;

b) bola preta na segunda extração;

c) bola branca na primeira extração.

1.8 - Se A e B são independentes, mostre que também são independentes: A e B^c, A^c e B, A^c e B^c.

1.9 - As probabilidades que dois eventos independentes ocorram são p e q, respectivamente. Qual a probabilidade:

a) que nenhum destes eventos ocorra?

b) pelo menos um destes eventos ocorra?

1.10- Na tabela abaixo, os números que aparecem são probabilidades relacionadas com a ocorrência de A, B, A e B, etc. Assim, P(A) = 0,1, enquanto que $P(A \cap B) = 0{,}04$. Verifique se A e B são independentes.

	B	B^c	
A	0,04	0,06	0,10
A^c	0,08	0,82	0,90
	0,12	0,86	1,00

1.11- Suponha uma população de N elementos $a_1, a_2, \ldots, a_N$. Qualquer arranjo *ordenado* $a_{i_1}, a_{i_2}, \ldots, a_{i_n}$ de n símbolos é chamado uma *amostra ordenada de tamanho* n extraída da população. Considere o símbolo $(N)_n$ como significando $N(N-1)\ldots(N-n+1)$. Suponha $n<N$. Mostre que existem N^n amostras *com reposição* (um mesmo elemento pode ser retirado mais de uma vez) e $(N)_n$ amostras *sem reposição* (um elemento quando escolhido é removido da população, não havendo pois repetição na amostra).

1.12- Uma amostra de tamanho n extraída de uma população com N elementos diz-se *casual* (ou *aleatória*) se todas as possíveis amostras têm a mesma probabilidade de serem escolhidas; esta probabilidade será $1/N_n$ se a amostra que a probabilidade de cara ou coroa é $\frac{1}{2}$. Podemos dar a seguinte explicação em termos de freqüência relativa. Considere 10 lançamentos da moeda; se obtemos 6 caras, então $\frac{6}{10} = 0{,}6$ é a freqüência relativa de ocorrência de caras nestes 10 lançamentos. Suponha que aumentemos progressivamente o número de lançamentos, e para cada um deles calculemos a freqüência relativa. Esta variará de prova para prova mas a medida que aumentamos o número de lançamentos ela tenderá a se estabilizar ao redor de um número, que será a probabilidade de "cara"; este número será $\frac{1}{2}$ se a moeda for "honesta".

Esta interpretação freqüencial poderá ser útil ao leitor, se quizer verificar heuristicamente as propriedades que passamos a considerar, em especial a regra de adição. Em particular, como toda freqüência relativa é um número entre 0 e 1, vemos que $0 \leq P(A) \leq 1$, para qualquer evento A. Para efeito de completividade, será útil considerarmos o espaço todo Ω e o conjunto vazio, $\emptyset$, como eventos. O primeiro é denominado *evento certo* e $P(\Omega)=1$, enquanto que $\emptyset$ é o *evento impossível* e $P(\emptyset)=0$.

Exemplo 1.3 - Suponha que o seguinte quadro represente uma possível divisão dos alunos matriculados em dado instituto de matemática, num dado ano. (Vide quadro na pág. seguinte).

Vamos indicar por P o evento que ocorre quando, escolhendo-se ao acaso um aluno do instituto, ele for um estudante de matemática pura. A, E, C, H e M têm significados a

	HOMENS	MULHERES	TOTAIS
MATEMÁTICA PURA . .	70	40	110
MATEMÁTICA APLICADA	15	15	30
ESTATÍSTICA.	10	20	30
COMPUTAÇÃO.	20	10	30
TOTAIS.	115	85	200

Tabela 1.1

nálogos. Desta maneira vemos que $P(E) = \frac{30}{200}$, ao passo que $P(H) = \frac{115}{200}$.

Dado os eventos A e H podemos considerar dois novos eventos:

$A \cup H$, chamado a *reunião* de A e H, que ocorre quando pelo menos um dos eventos ocorre;

$A \cap H$, chamado a *intersecção* de A e H, que ocorre quando A e H ocorrem simultaneamente.

É fácil ver que $P(A \cap H) = \frac{15}{200}$, pois o aluno escolhido terá que ser ao mesmo tempo, matriculado no curso de matemática aplicada e ser homem.

Vemos que $P(A) = \frac{30}{200}$ e $P(H) = \frac{115}{200}$; suponha que nosso cálculo para $P(A \cup H)$ fosse

$$P(A \cup H) = P(A) + P(H) = \frac{30}{200} + \frac{115}{200} = \frac{145}{200}.$$

Se assim o fizéssemos, estaríamos contando duas vezes os alunos que são homens *e* estão matriculados no curso de matemática aplicada, como está destacado na tabela 1.1. Portanto, a resposta correta é

$$P(A \cup H) = P(A) + P(H) - P(A \cap H)$$

$$= \frac{30}{200} + \frac{115}{200} - \frac{15}{200} = \frac{130}{200}.$$

No entanto, considerando-se os eventos A e C, vemos que $P(A) = \frac{30}{200}$, $P(C) = \frac{30}{200}$ e $P(A \cup C) = \frac{60}{200} = P(A) + P(C)$. Neste caso, os eventos A e C são disjuntos ou *mutuamente exclusivos*, pois se A ocorre então C não ocorre e vice-versa. Aqui, $A \cap C = \emptyset$ e $P(A \cap C) = 0$.

Portanto, se M e N são dois eventos quaisquer temos a chamada *regra da adição de probabilidades*

$$P(M \cup N) = P(M) + P(N) - P(M \cap N), \tag{1.2}$$

que se reduz a

$$P(M \cup N) = P(M) + P(N), \tag{1.3}$$

se M e N são eventos mutuamente exclusivos.

Suponha, agora, que estamos somente interessados em saber se um estudante escolhido ao acaso está matriculado como aluno de matemática pura, aplicada, estatística ou computação, não interessando saber se é homem ou mulher. Um espaço amostral é $\Omega = \{P, A, E, C\}$. Os eventos A e $B = \{P, E, C\}$ são chamados *complementares* e são tais que $A \cup B = \Omega$ e $A \cap B = \emptyset$. Vemos que $P(A) = \frac{30}{200}$, enquanto $P(B) = \frac{110}{200} + \frac{30}{200} + \frac{30}{200} = \frac{170}{200}$, isto é,

$$P(A) + P(B) = 1. \tag{1.4}$$

Em geral, vamos indicar por A^c o complementar de um evento A e teremos, então,

$$P(A^c) = 1 - P(A). \tag{1.5}$$

1.3 - PROBABILIDADE CONDICIONAL E INDEPENDÊNCIA

Voltemos ao quadro do exemplo 1.3. Dado que um estudante, escolhido ao acaso, esteja matriculado no curso de Estatística, a probabilidade dele ser mulher é $\frac{10}{30} = \frac{1}{3}$. Isto porque do total de 30 alunos que estudam estatística, 10 são mulheres. Escrevemos

$$P(\text{mulher} \mid \text{estatística}) = \frac{1}{3}.$$

Para dois eventos quaisquer A e B, sendo $P(B) > 0$, definimos a *probabilidade condicional de* A, *dado* B, $P(A|B)$, como sendo

$$P(A|B) = \frac{P(A \cap B)}{P(B)}. \tag{1.6}$$

Para o exemplo mencionado, se B e A indicam, respectivamente, os eventos "aluno matriculado em Estatística" e aluno é mulher", então $P(A \cap B) = 10/200$, $P(B) = 30/200$ e, portanto,

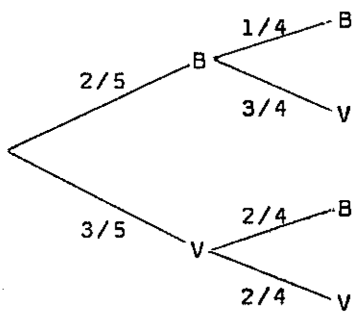
$$P(A|B) = \frac{10/200}{30/200} = \frac{1}{3},$$

como havíamos obtido.

Da relação (1.6) obtemos a chamada ***regra do produto de probabilidades***,

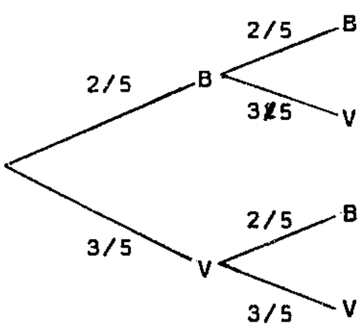
$$P(A \cap B) = P(B) \cdot P(A|B). \tag{1.7}$$

Exemplo 1.4 - Retomemos a urna do exemplo 1.1 mas suponha agora que extraímos duas bolas ao acaso, ***sem reposição***. Isto significa que escolhemos a primeira bola, verificamos a sua cor e *não* a devolvemos à urna, misturamos as bolas restantes e retiramos a segunda bola. O diagrama em árvore abaixo ilustra as possibilidades. Em cada "galho" da árvore estão indicadas as probabilidades de ocorrência, sendo que, para as segundas bolas temos probabilidades condicionais. A probabilidade do resultado conjunto é, então, dada por (1.7).



RESULTADOS	PROBABILIDADES
BB	$\frac{2}{5} \times \frac{1}{4} = \frac{2}{20}$
BV	$\frac{2}{5} \times \frac{3}{4} = \frac{6}{20}$
VB	$\frac{3}{5} \times \frac{2}{4} = \frac{6}{20}$
VV	$\frac{3}{5} \times \frac{2}{4} = \frac{6}{20}$
SOMA	1

<u>Exemplo 1.5</u> - Imagine, agora, que as duas extrações são feitas da urna, mas a primeira bola é *reposta* na urna antes da extração da segunda bola. Nestas condições, as extrações são *independentes*, no sentido que o resultado de cada extração não tem influência no resultado da outra. Obtemos a situação a seguir.



RESULTADOS	PROBABILIDADES
BB	$\frac{2}{5} \times \frac{2}{5} = \frac{4}{25}$
BV	$\frac{2}{5} \times \frac{3}{5} \quad \frac{6}{25}$
VB	$\frac{3}{5} \times \frac{2}{5} \quad \frac{6}{25}$
VV	$\frac{3}{5} \times \frac{3}{5} \quad \frac{9}{25}$
SOMA	1

Observe que, aqui,

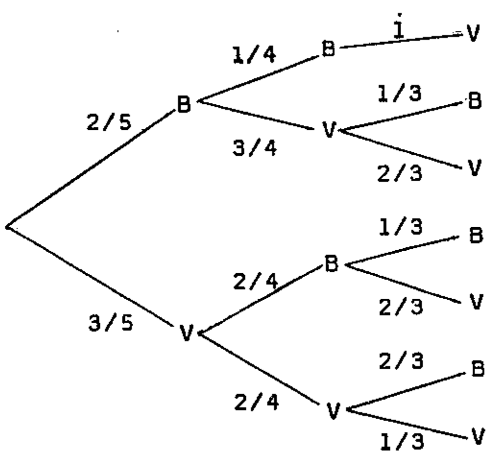
$$P(\text{branca na } 2^{a} \mid \text{branca na } 1^{a}) = \frac{2}{5} = P(\text{branca na } 2^{a}),$$

Ou seja, se dois eventos A e B são *independentes*, $P(A|B) = P(A)$ e $P(B|A) = P(B)$. Se assim é, a relação (1.7) fica para A e B independentes:

$$P(A \cap B) = P(A) \cdot P(B). \tag{1.8}$$

A fórmula (1.8) pode ser tomada como definição de independência: A e B são independentes se e somente se (1.8) é válida.

Exemplo 1.6 - Considere, ainda, a mesma urna dos exemplos anteriores, mas vamos fazer três extrações *sem reposição*. Obtemos o esquema abaixo.



RESULTADOS	PROBABILIDADES
BBV	$\frac{2}{5} \times \frac{1}{4} \times 1 = \frac{2}{20} = \frac{6}{60}$
BVB	$\frac{2}{5} \times \frac{3}{4} \times \frac{1}{3} = \frac{6}{60}$
BVV	$\frac{2}{5} \times \frac{3}{4} \times \frac{2}{3} = \frac{12}{60}$
VBB	$\frac{3}{5} \times \frac{2}{4} \times \frac{1}{3} = \frac{6}{60}$
VBV	$\frac{3}{5} \times \frac{2}{4} \times \frac{2}{3} = \frac{12}{60}$
VVB	$\frac{3}{5} \times \frac{2}{4} \times \frac{2}{3} = \frac{12}{60}$
VVV	$\frac{3}{5} \times \frac{2}{4} \times \frac{1}{3} = \frac{6}{60}$
SOMA	$60/60 = 1$

Observe que $P(B|B) = \frac{1}{4}$, ao passo que $P(V|B \text{ e } B) = 1$; daí

$$P(B \text{ e } B \text{ e } V) = P(B) \cdot P(B|B) \cdot P(V|B \text{ e } B) = \frac{2}{5} \times \frac{1}{4} \times 1 = \frac{1}{10}.$$

De modo geral, dados 3 eventos M, N e R, temos que

$$P(M \cap N \cap R) = P(M) \cdot P(N|M) \cdot P(R|M \cap N) \qquad (1.9)$$

PROBLEMAS PARA O CAPÍTULO 1

1.1 - Duas moedas são lançadas. Dê dois possíveis espaços amostrais para este experimento. Represente um destes como o *produto cartesiano* de dois outros espaços amostrais.

1.2 - Uma moeda e um dado são lançados. Dê um espaço amostral do experimento e depois represente-o como produto cartesiano dos dois espaços amostrais correspondente aos experimentos considerados individualmente.

1.3 - No exercício 1.1, liste os eventos:
   a) pelo menos uma cara;
   b) duas caras;
   c) o complementar de b).

1.4 - Considere o lançamento de dois dados. Considere os eventos A: soma dos números obtidos igual a 9 e B: número no primeiro dado maior ou igual a 4. Enumere os elementos de A e B. Obtenha $A \cup B$, $A \cap B$ e A^c.

1.5 - Obtenha as probabilidades dos eventos que aparecem nos problemas 1.3 e 1.4.

1.6 - Considere uma urna contendo 3 bolas pretas e 5 bolas vermelhas. Retire duas bolas da urna, sem reposição. Obtenha os resultados possíveis e as respectivas probabilidades. Mesmo problema, para extrações com reposição.

1.7 - No problema anterior, calcule as probabilidades dos eventos:

a) bola preta na primeira e segunda extrações;

b) bola preta na segunda extração;

c) bola branca na primeira extração.

1.8 - Se A e B são independentes, mostre que também são independentes: A e B^c, A^c e B, A^c e B^c.

1.9 - As probabilidades que dois eventos independentes ocorram são p e q, respectivamente. Qual a probabilidade:

a) que nenhum destes eventos ocorra?

b) pelo menos um destes eventos ocorra?

1.10- Na tabela abaixo, os números que aparecem são probabilidades relacionadas com a ocorrência de A, B, A e B, etc. Assim, P(A) = 0,1, enquanto que $P(A \cap B) = 0,04$. Verifique se A e B são independentes.

	B	B^c	
A	0,04	0,06	0,10
A^c	0,08	0,82	0,90
	0,12	0,86	1,00

1.11- Suponha uma população de N elementos $a_1, a_2, \ldots, a_N$. Qualquer arranjo *ordenado* $a_{i_1}, a_{i_2}, \ldots, a_{i_n}$ de n símbolos é chamado uma *amostra ordenada de tamanho* n extraída da população. Considere o símbolo $(N)_n$ como significando $N(N-1)\ldots(N-n+1)$. Suponha n<N. Mostre que existem N^n amostras *com reposição* (um mesmo elemento pode ser retirado mais de uma vez) e $(N)_n$ amostras *sem reposição* (um elemento quando escolhido é removido da população, não havendo pois repetição na amostra).

1.12- Uma amostra de tamanho n extraída de uma população com N elementos diz-se *casual* (ou *aleatória*) se todas as possíveis amostras têm a mesma probabilidade de serem escolhidas; esta probabilidade será $1/N_n$ se a amostra for com reposição e $1/(N)_n$ se for sem reposição.

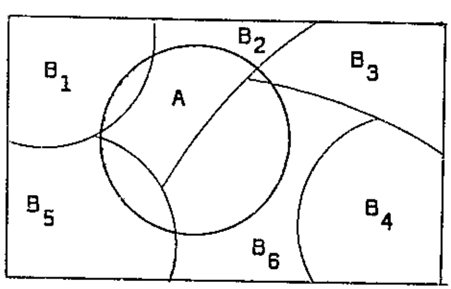
Uma amostra casual de tamanho n com reposição é extraída de uma população com N elementos. Encontre a probabilidade de não haver repetição na amostra.

1.13- Considere os aniversários de n pessoas como constituindo uma amostra de tamanho n da população dos 365 dias do ano. Aplique o resultado do problema anterior para encontrar a probabilidade que todos os n aniversários caiam em dias diferentes.

1.14- Denote por $\binom{N}{n} = \frac{(N)_n}{n!} = \frac{N!}{n!(N-n)!}$. Considere a situação do problema 1.11, mas na qual *não levamos em conta* a ordem do conjunto $a_{i_1}, a_{i_2}, \ldots, a_{i_n}$. Mostre que existem $\binom{N}{n}$ amostras sem reposição.

1.15- Suponha que $\{B_1, B_2, \ldots, B_m\}$ é um conjunto de eventos mutuamente exclusivos e exaustivos, ou seja, quaisquer dois eventos B_i e B_j não podem ocorrer ao mesmo tempo e um deles deve ocorrer. Simbolicamente, $B_i \cap B_j = \emptyset$, $i \neq j$ e $B_1 \cup B_2 \cup \ldots \cup B_m = \Omega$. Seja A um evento qualquer. Se A ocorre, então A deve ocorrer juntamente com um dos B_i. (Ver figura). Prove que

$$P(A) = \sum_{i=1}^{m} P(A|B_i) \cdot P(B_i).$$



Faça um diagrama em árvore.

1.16- Utilize o resultado anterior para resolver este problema. Uma companhia produz rádios em 3 fábricas A, B e C. A fábrica A produz 40% dos rádios, enquanto B e C produzem 30% cada uma. As probabilidades que um rádio produzido por estas fábricas não funcione são 0,01, 0,04 e 0,03, respectivamente. Escolhido um rádio da produção conjunta das três fábricas, qual a probabilidade do mesmo não funcionar?

1.17- Considere a situação do problema anterior, mas suponha agora que um rádio é escolhido ao acaso e ele é defeituoso. Determinar qual a probabilidade de ter sido fabricado por A. No caso geral referido em 1.15, queremos obter $P(B_j|A)$ para $j=1,\ldots,m$. Verificar que se chega a

$$P(B_j|A) = \frac{P(A|B_j)P(B_j)}{P(A)} = \frac{P(A|B_j)P(B_j)}{\sum_{i=1}^{m} P(A|B_i)P(B_i)}$$

Este resultado é denominado *fórmula de Bayes* e é útil em muitas aplicações.

1.18- O colégio A tem 40% de rapazes, o colégio B tem 30% e o colégio C tem 10%. De um colégio selecionado ao acaso são escolhidos 8 alunos, com reposição. Se obtemos (R para rapaz e M para moça) RRRMMMMM, qual a probabilidade de ter sido selecionado o colégio C?

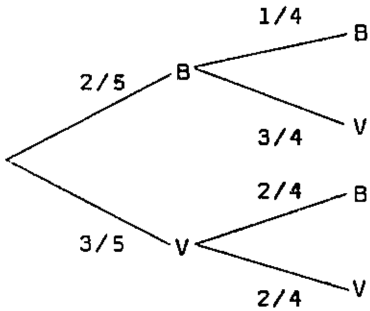
- o -

2. VARIÁVEIS ALEATÓRIAS

Em todas as situações de interesse prático que consideramos queremos estudar o comportamento de uma ou mais variáveis. Assim, podemos estar interessados em estudar a distribuição das alturas e pesos de pessoas de uma dada população. Escolhida uma pessoa ao acaso desta população, obtemos o seu peso e a sua altura e estamos, pois, considerando duas variáveis. Em muitas situações, como a descrita, a associação de números a resultados de experimentos é natural. Em outras, a associação pode ser arbitrária. Por exemplo, considere o caso de um questionário, em que uma pessoa é indagada a respeito de uma certa proposição e as respostas possíveis são "SIM" ou "NÃO". Podemos definir uma variável que tome dois valores, 1 ou 0, por exemplo, correspondentes às respostas "SIM" ou "NÃO, respectivamente. Teremos ocasião de estudar com pormenores este tipo de variável.

2.1 - O CONCEITO DE VARIÁVEL ALEATÓRIA

Exemplo 2.1 - Consideremos a situação do exemplo 1.4, em que consideramos duas extrações, sem reposição, de uma urna contendo 2 bolas brancas e 3 bolas vermelhas. Definamos a variá vel X igual número de bolas vermelhas obtidas nas duas extrações. Obtemos o seguinte esquema.



RESULTADOS	PROBABILIDADES	X
BB	1/10	0
BV	3/10	1
VB	3/10	1
VV	3/10	2

Vemos, pois, que a cada resultado do experimento está associado um valor da variável aleatória (v.a.) X; estes valores são 0, 1 e 2.

Dizemos que X=0, com probabilidade 1/10, pois X = 0 se e somente se ocorre o resultado BB; X=1 com probabilidade de $\frac{3}{10} + \frac{3}{10} = \frac{6}{10}$, pois X=1 se e somente se ocorrem os resultados BV ou VB, que são mutuamente exclusivos; finalmente, X=2 com probabilidade $\frac{3}{10}$, pois X=2 se e somente se ocorrem duas bolas vermelhas nas duas extrações. A notação que usaremos será P(X=0), P(X=1), etc. Temos, pois que

$$P(X=0) = P(BB) = \frac{1}{10},$$

$$P(X=1) = P(BV \text{ ou } VB) = \frac{6}{10},$$

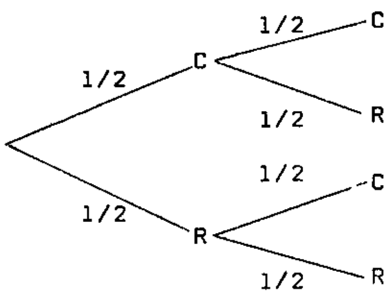
$$P(X=2) = P(VV) = \frac{3}{10}.$$

No quadro a seguir, esquematizamos a *distribuição de probabilidades da v.a.* X, que consiste nos possíveis valores de X, com as respectivas probabilidades.

Usaremos a notação $P(X=x_i) = p_i$, isto é, designare mos por letras minúsculas x_i os valores da v.a. X e por p_i as respectivas probabilidades, $i=1, 2, 3, \ldots$.

x_i	0	1	2
p_i	1/10	6/10	3/10

<u>Exemplo 2.2</u> - Retomemos o exemplo 1.2, em que consideramos o lançamento de uma moeda duas vezes. Definamos a v.a. Y = = número de "caras" obtidas nos dois lançamentos.



RESULTADOS	PROBABILIDADES	Y
CC	1/4	2
CR	1/4	1
RC	1/4	1
RR	1/4	0

Temos, então:

$$P(Y=0) = P(RR) = \frac{1}{4} ,$$

$$P(Y=1) = P(CR \text{ ou } RC) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2} ,$$

$$P(Y=2) = P(CC) = \frac{1}{4}.$$

A distribuição de probabilidades de Y está indicada abaixo.

y_i	0	1	2
p_i	1/4	1/2	1/4

2.2 - <u>VALOR ESPERADO DE UMA VARIÁVEL ALEATÓRIA</u>

Se $x_1, x_2, \ldots$ são os possíveis valores da v.a. X e

$p_1, p_2, \ldots$ são as respectivas probabilidades, então o *valor esperado* (ou *esperança* ou *média*) de X é definido como

$$E(X) = \sum_{i=1}^{\infty} x_i \cdot p_i. \qquad (2.1)$$

Exemplo 2.3 - Para a distribuição da v.a. X do exemplo 2.1, temos que

$$E(X) = 0 \times \frac{1}{10} + 1 \times \frac{6}{10} + 2 \times \frac{3}{10} = \frac{12}{10} = 1,2.$$

Para a distribuição da v.a. Y do exemplo 2.2,

$$E(Y) = 0 \times \frac{1}{2} + 1 \times \frac{1}{2} + 2 \times \frac{1}{4} = 1.$$

Exemplo 2.4 - Consideremos o lançamento de um dado e seja X a v.a. que representa o número obtido na face voltada para cima. Então, X toma os valores 1, 2, 3, 4, 5, 6 com probabilidades todas iguais a 1/6. Calculemos $E(X)$ e $E(X^2)$.

$$E(X) = 1 \times \frac{1}{6} + 2 \times \frac{1}{6} + \ldots + 6 \times \frac{1}{6} = (1+2+\ldots+6) \times \frac{1}{6} = \frac{21}{6} = 3,5.$$

Os valores da v.a. X^2 são 1, 4, 9, 16, 25, 36 e as probabilidades respectivas são iguais a 1/6. De fato,

$$P(X^2=9) = P(X=3) = \frac{1}{6}, \text{ etc.}$$

Então,

$$E(X^2) = (1+4+9+16+25+36) \times \frac{1}{6} = \frac{91}{6}.$$

Exemplo 2.5 - Considere a v.a. V com a distribuição dada pela tabela abaixo:

v	-2	-1	0	1	2
p	1/5	1/5	1/5	1/5	1/5

Calculemos $E(V)$, $E(V^2)$ e $E(2V)$.

$$E(V) = (-2) \times \frac{1}{5} + (-1) \times \frac{1}{5} + 0 \times \frac{1}{5} + 1 \times \frac{1}{5} + 2 \times \frac{1}{5} = 0$$

Obtemos, depois, as distribuições de $2V$ e V^2.

2v	-4	-2	0	2	4
p	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$

v^2	0	1	4
p	$\frac{1}{5}$	$\frac{2}{5}$	$\frac{2}{5}$

Por exemplo,

$$P(2V=4) = P(V=2) = \frac{1}{5} \quad \text{e}$$

$$P(V^2=1) = P(V=-1 \text{ ou } V=1) = \frac{1}{5} + \frac{1}{5} = \frac{2}{5}.$$

Então,

$$E(2V) = (-4) \times \frac{1}{5} + \ldots + 4 \times \frac{1}{5} = 0,$$

enquanto que

$$E(V^2) = 0 \times \frac{1}{5} + 1 \times \frac{2}{5} + 4 \times \frac{2}{5} = \frac{10}{5} = 2.$$

Observe as distribuições de V e $2V$; dizemos que elas são simétricas ao redor de 0; a média será, obviamente, o ponto de simetria. (Veja Fig. 2.1).

Observe, também, que $E(V^2) = 2 \neq [E(V)]^2 = 0$. No entanto, é fácil ver que

$$E(cX) = cE(X), \tag{2.2}$$

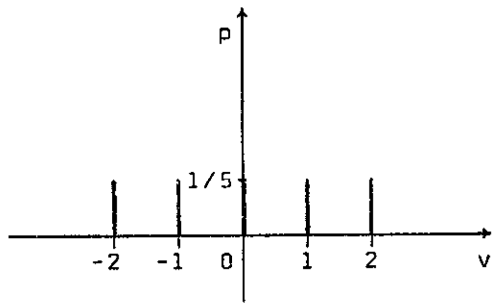


Figura 2.1

se c é uma constante, o que decorre imediatamente da definição (2.1).

Observando o gráfico da figura 2.1, vemos que E(V) é uma medida de posição central da distribuição de V.

A v.a. 2V pode ser representada pelo gráfico da figura 2.2.

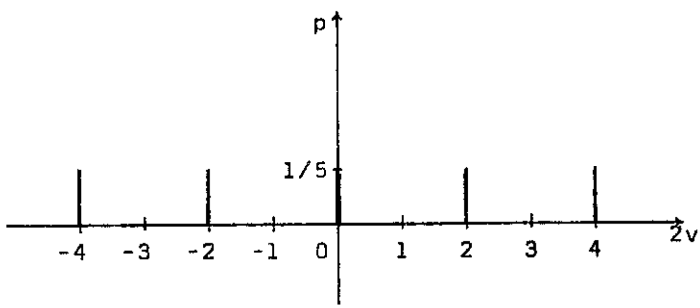


Figura 2.2

As variáveis alaatórias V e 2V têm a mesma média, zero, mas notamos que 2V é "mais espalhada" ao redor de 0 do que V. Uma medida da dispersão de uma v.a. ao redor de sua média é dada pela *variância* da v.a.

Se X é uma v.a. com média E(X), então a variância de X é definida por

$$\text{Var}(X) = \sum_{i=1}^{\infty} (x_i - E(X))^2 \cdot p_i \tag{2.3}$$

e é fácil ver que

$$\text{Var}(X) = E(X^2) - [E(X)]^2. \tag{2.4}$$

<u>Exemplo 2.6</u> - Para a v.a. V do exemplo anterior, $E(V) = 0$ e $E(V^2) = 2$, logo $\text{Var}(V) = 2$.

Vimos que $E(2V) = 0$. Vejamos $E[(2V)^2]$; a distribuição de $(2V)^2$ é aquela abaixo:

$(2v)^2$	0	4	16
p	$\frac{1}{5}$	$\frac{2}{5}$	$\frac{2}{5}$

Logo,

$$E[(2V)^2] = 0 \times \frac{1}{5} + 4 \times \frac{2}{5} + 16 \times \frac{2}{5} = \frac{40}{5} = 8,$$

do que segue que $\text{Var}(2V) = 8$, usando (2.4).

Notamos que

$$\text{Var}(2V) = 8 = 2^2 \cdot \text{Var}(V) = 4 \times 2.$$

De modo geral, utilizando (2.3), vemos que

$$\text{Var}(cX) = c^2 \cdot \text{Var}(X) \tag{2.5}$$

se c é uma constante.

<u>Exemplo 2.7</u> - Considere a v.a. V do exemplo 2.5 e seja $W = V + 6$.

Então, a distribuição de W é dada por:

w	4	5	6	7	8
p	1/5	1/5	1/5	1/5	1/5

Segue-se que

$$E(W) = (4+5+6+7+8) \times \frac{1}{5} = \frac{30}{5} = 6 = E(V) + 6$$

e

$$E(W^2) = (16+25+36+49+64) \times \frac{1}{5} = \frac{190}{5} = 38,$$

do que decorre que

$$Var(W) = 38-6^2 = 38-36 = 2 = Var(V).$$

Genericamente, se c é uma constante,

$$E(X+c) = E(X)+c \tag{2.6}$$

e

$$Var(X+c) = Var(X). \tag{2.7}$$

A relação (2.6) expressa o fato intuitivo que, ao transladarmos todos os valores de X por uma constante c, a média será transladada pela mesma constante, ao passo que (2.7) expressa o fato que a variabilidade não muda, pois a posição relativa dos novos valores em relação à nova média é a mesma que antes.

As fórmulas (2.2), (2.5), (2.6) e (2.7) podem ser combinadas para se obter:

$$E(aX+b) = aE(X) + b, \tag{2.8}$$

$$Var(aX + b) = a^2 \cdot Var(X), \tag{2.9}$$

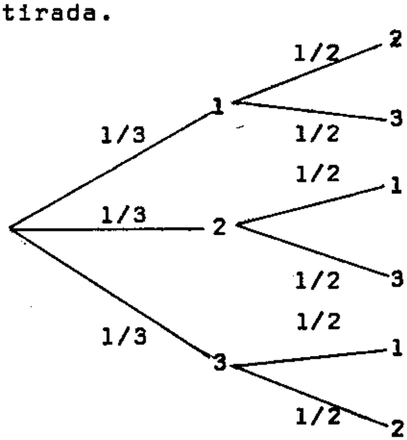
sendo a e b constantes.

O *desvio padrão* de X é definido como a raiz quadrada positiva da Var(X) e será indicado por $\sigma(X)$ ou σ_X; a variância de X também é denotada por $\sigma^2(X)$ ou σ^2_X. A relação (2.9) fica, para $\sigma(X)$:

$$\sigma(aX+b) = |a| \cdot \sigma(X). \tag{2.10}$$

2.3 - MAIS DE UMA VARIÁVEL ALEATÓRIA

Exemplo 2.8 - Suponha que temos uma urna contendo 3 bolas numeradas 1, 2 e 3. Retiramos duas delas, sem reposição; seja X o número da primeira bola e Y o número da segunda bola retirada.



RESULTADOS	PROBABILIDADES	X	Y
1 2	1/6	1	2
1 3	1/6	1	3
2 1	1/6	2	1
2 3	1/6	2	3
3 1	1/6	3	1
3 2	1/6	3	2

Temos, então, duas variáveis aleatórias, X e Y, cu jas distribuições de probabilidades são:

x	1	2	3
p	1/3	1/3	1/3

y	1	2	3
p	1/3	1/3	1/3

O quadro seguinte representa a chamada *distribuição conjunta de* X e Y, onde P(X=x, Y=y) denota a probabilidade do evento {X=x e Y=y} = $(X=x)\cap(Y=y)$.

(X,Y)	P(X=x,Y=y)
(1,2)	1/6
(1,3)	1/6
(2,1)	1/6
(2,3)	1/6
(3,1)	1/6
(3,2)	1/6

Lembremos como P(X=x,Y=y) é calculada; de acordo com a fórmula (1.7),

$$P(X=x,Y=y) = P(X=x)\cdot P(Y=y|X=x). \qquad (2.11)$$

Por exemplo,

$$P(X=2,Y=3) = P(X=2)P(Y=3|X=2) = \frac{1}{3} \times \frac{1}{2} = \frac{1}{6}.$$

Uma maneira mais cômoda e usual de se representar a distribuição conjunta de X e Y é através de uma tabela de dupla entrada, como a seguir:

Y \ X	1	2	3	P(Y=y)
1	0	1/6	1/6	1/3
2	1/6	0	1/6	1/3
3	1/6	1/6	0	1/3
P(X=x)	1/3	1/3	1/3	1

A vantagem desta tabela é que, tendo-se a distribuição conjunta de X e Y podemos obter as distribuições de X e Y (a recíproca não é verdadeira, em geral;por que?) que são chamadas *distribuições marginais*. Assim, a primeira e a última coluna da tabela dão a distribuição de Y, ao passo que a primeira e a última linha dão a distribuição de X.

Dada a distribuição de X e Y acima, podemos considerar, por exemplo, a v.a. X+Y, ou a v.a. XY. A soma X+Y é definida de modo natural: a cada resultado do experimento ela associa a soma dos valores de X e Y.

Podemos, pois, obter a tabela seguinte:

(X,Y)	X+Y	XY	PROBABILIDADES
(1,2)	3	2	1/6
(1,3)	4	3	1/6
(2,1)	3	2	1/6
(2,3)	5	6	1/6
(3,1)	4	3	1/6
(3,2)	5	6	1/6

As distribuições de X+Y e XY serão:

x+y	3	4	5
p	1/3	1/3	1/3

xy	2	3	6
p	1/3	1/3	1/3

Calculemos E(X), E(Y), E(X+Y) e E(XY). Temos:

$$E(X) = \frac{1}{3} + \frac{2}{3} + \frac{3}{3} = 2 = E(Y)$$

$$E(X+Y) = \frac{3}{3} + \frac{4}{3} + \frac{5}{3} = 4 = E(X) + E(Y)$$

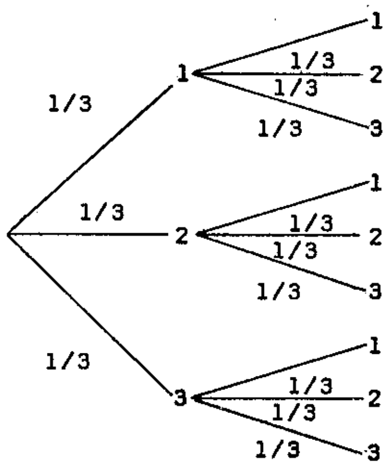
$$E(XY) = \frac{2}{3} + \frac{3}{3} + \frac{6}{3} = \frac{11}{3} \neq (X) \cdot E(Y) = 4.$$

De modo geral, é verdade que

$$E(X+Y) = E(X) + E(Y), \tag{2.12}$$

para quaisquer v.a. X e Y, mas não é verdade que $E(XY) = E(X) \cdot E(Y)$; isto só acontece num caso particular que veremos a seguir.

<u>**Exemplo 2.9**</u> - Considere a situação do exemplo 2.8, mas onde as extrações são feitas com reposição, isto é, a primeira bola é reposta antes de se retirar a segunda bola. Nestas condições, a $P(Y=y|X=x) = P(Y=y)$, isto é, o valor assumido por X não tem influência nenhuma no valor que Y vai assumir. A situação é descrita a seguir.



RESULTADOS	PROBABILIDADES	X	Y
1 1	1/9	1	1
1 2	1/9	1	2
1 3	1/9	1	3
2 1	1/9	2	1
2 2	1/9	2	2
2 3	1/9	2	3
3 1	1/9	3	1
3 2	1/9	3	2
3 3	1/9	3	3

A distribuição conjunta de X e Y é dada abaixo, juntamente com as marginais de X e de Y.

Y \ X	1	2	3	P(Y=y)
1	1/9	1/9	1/9	1/3
2	1/9	1/9	1/9	1/3
3	1/9	1/9	1/9	1/3
P(X=x)	1/3	1/3	1/3	1

Observe que, embora as distribuições marginais de X e Y sejam as mesmas que no exemplo anterior, a distribuição conjunta é diferente. A relação (2.11) ficará, neste caso,

$$P(X=x, Y=y) = P(X=x) \cdot P(Y=y), \qquad (2.13)$$

para todos os possíveis pares (x,y), $x=1,2,3$; $y=1,2,3$.

Dizemos que X e Y *são independentes*.

Ainda temos, aqui, E(X) = E(Y) = 2 e (2.12) continua válida. A distribuição de XY será:

xy	1	2	3	4	6	9
p	1/9	2/9	2/9	1/9	2/9	1/9

de modo que $E(XY) = \frac{1}{9} + \frac{4}{9} + \frac{6}{9} + \frac{4}{9} + \frac{12}{9} + \frac{9}{9} = \frac{36}{9} = 4$, que é igual a $E(X) \cdot E(Y)$. Então, obtemos o fato importante:

"Se X e Y são v.a. independentes,

$$E(XY) = E(X) \cdot E(Y)". \tag{2.14}$$

Uma observação importante é a seguinte: para que X e Y sejam independentes é necessário que (2.13) esteja verificada para *todos os pares* (x,y). Basta que ela não se verifique para *um* par para que X e Y não sejam independentes. Neste caso diremos que X e Y são *dependentes*. A pergunta que pode surgir é: pode (2.14) valer para v.a. dependentes? Ver exercícios para a resposta.

<u>**Exemplo 2.10**</u> - Vamos calcular Var(X), Var(Y) e Var(X+Y) para o exemplo 2.9.

Vimos que E(X) = E(Y) = 2, também

$$E(X^2) = E(Y^2) = 1 \times \frac{1}{3} + 4 \times \frac{1}{3} + 9 \times \frac{1}{3} = \frac{14}{3}.$$

Portanto,

$$\mathrm{Var}(X) = \mathrm{Var}(Y) = \frac{14}{3} - 2^2 = \frac{2}{3}$$

A distribuição de X+Y é:

x+y	2	3	4	5	6
p	1/9	2/9	3/9	2/9	1/9

logo $E(X+Y) = 4$ e

$$E[(X+Y)^2] = 4 \times \frac{1}{9} + 9 \times \frac{2}{9} + 16 \times \frac{3}{9} + 25 \times \frac{2}{9} + 36 \times \frac{1}{9} = \frac{156}{9}$$

de modo que

$$Var(X+Y) = \frac{156}{9} - 16 = \frac{12}{9} = \frac{4}{3} = Var(X) + Var(Y).$$

Este resultado também é geral para v.a. independentes:

"Se X e Y são independentes, então

$$Var(X+Y) = Var(X) + Var(Y)." \qquad (2.15)$$

Esta relação pode não ser verdadeira se X e Y são dependentes. Para um exemplo, ver exercícios.

2.4 - REPRESENTAÇÕES GRÁFICAS PARA VARIÁVEIS

No exemplo 2.5 vimos um tipo de representação gráfica para uma v.a., V, naquele caso. Outro tipo de gráfico que pode ser usado é o histograma.

Exemplo 2.11 - Consideremos o lançamento de 3 moedas e sejam as v.a. X e Y definidas como segue: X é o número de "caras" que aparecem e Y o número de "seqüências", onde uma se qüência é um conjunto (uma ou mais) de letras iguais.

RESULTADO	X	Y
C C C	3	1
C C R	2	2
C R C	2	3
C R R	1	2
R C C	2	2
R C R	1	3
R R C	1	2
R R R	0	1

A distribuição conjunta de X e Y é obtida facilmente,bem como as marginais.

Y \ X	0	1	2	3	P(Y=y)
1	1/8	0	0	1/8	2/8
2	0	2/8	2/8	0	4/8
3	0	1/8	1/8	0	2/8
P(X=x)	1/8	3/8	3/8	1/8	1

Na figura 2.3 temos representações gráficas para as distribuições de X e Y na forma de gráfico em barra e histogramas.

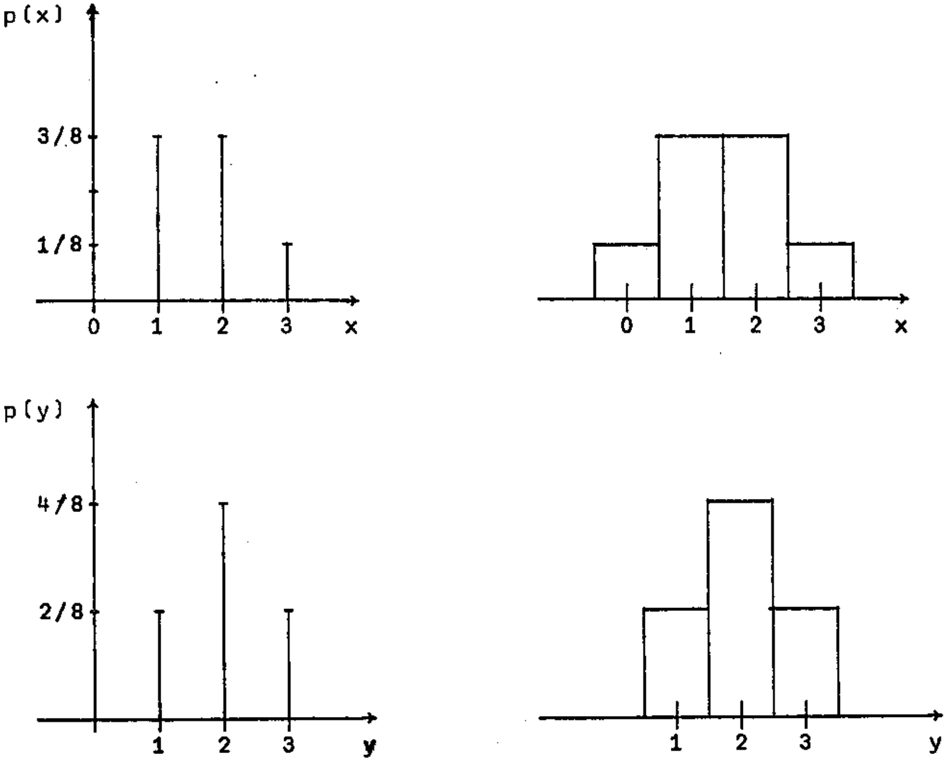


Figura 2.3

Podemos obter, também, um gráfico para a distribuição conjunta de X e Y, como na figura 2.4.

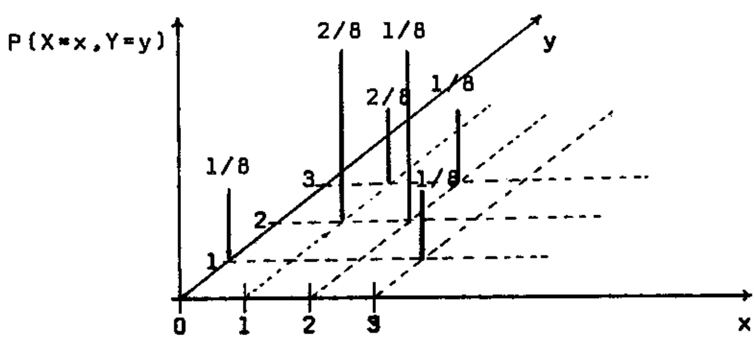


Figura 2.4

Um outro tipo de gráfico bastante útil na análise de dependência entre v.a. é o gráfico de dispersão, que consiste em colocar, num gráfico cartesiano, os possíveis pontos (x,y). Para o caso em questão obtemos a figura 2.5, onde os círculos duplos indicam os pontos com **probabilidade** 2/8.

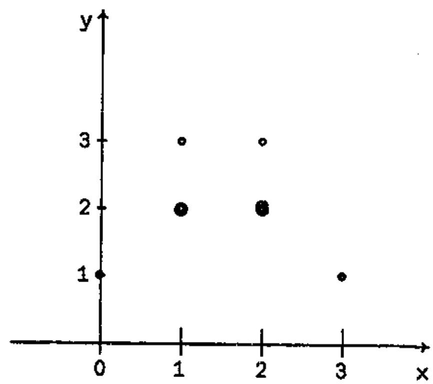


Figura 2.5

Este tipo de gráfico pode nos mostrar, por exemplo: que à medida que x cresce, y também cresce; ou a medida que x cresce, y decresce, etc.

PROBLEMAS PARA O CAPÍTULO 2

2.1 - Considere uma urna contendo 3 bolas vermelhas e 5 bolas pretas. Retire três bolas, sem reposição e defina a v.a. X igual ao número de bolas pretas. Obtenha a distribuição de X, 3X e X^2.

2.2 - Repita o problema anterior, mas agora considerando extrações com reposição.

2.3 - Considere o lançamento de três moedas. Se ocorre o evento CCC dizemos que temos *uma seqüência,* ao passo que se ocorre CRC temos *três seqüências*. Defina as v.a.: X igual ao número de caras obtidas e Y igual ao número de seqüências, isto para cada resultado possível. Assim, X(CRR) = 1 e Y(CRR) = 2. Obtenha as distribuições de X, Y, X+Y, |X-Y|. Faça o gráfico de dispersão das v.a. X e X+Y.

2.4 - Obtenha: E(X) e Var(3X), para o problema 2.1; E(X), E(Y), E(X+Y), Var(X+Y) para o problema 2.3.

2.5 - Suponha que a v.a. V tem a distribuição seguinte:

v	0	1
p	q	1-q

Obtenha E(V) e Var(V).

2.6 - Lance uma moeda e defina a v.a. V=0 se ocorre cara e V=1 se ocorre coroa e seja q = P(cara). Lance a moeda, agora, n vezes e para cada lançamento defina a v.a. V_i como acima, i=1,...,n. Seja $W = V_1+...+V_n$.

a) qual o significado de W?

b) obtenha a distribuição de $W = V_1+V_2$;

c) obtenha E(W) e Var(W).

2.7 - Utilizando a fórmula (1.3) prove que a variância de X pode ser obtida pela fórmula (2.4).

2.8 - Verifique se as v.a. X e Y definidas no problema 2.2 são independentes.

2.9 - No problema 2.2, calcule E(XY), E(X/Y) e E(X+Y).

2.10- Verifique, para o problema 2.2, se E(XY) = E(X)·E(Y). Utilizando a resposta do problema 2.8, o que você pode concluir? Verifique se $E(X/Y) = \frac{E(X)}{E(Y)}$.

2.11- Calcule Var(X+Y), Var(X) e Var(Y) para as v.a. do problema 2.2. É verdade que Var(X+Y) = Var(X) + Var(Y)?

2.12- Se X é uma v.a. e x é um número real qualquer denominamos *função de distribuição acumulada* de X à função

$$F(x) = P(X \leq x).$$

Obtenha F(x) para a v.a. X com distribuição dada por

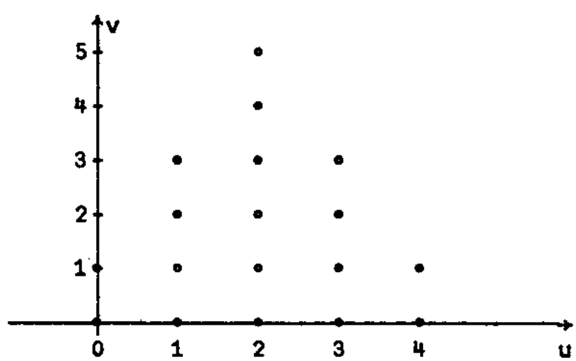
x	-2	-1	0	1	2
p	1/5	1/5	1/5	1/5	1/5

Faça o gráfico de F(x) e verifique que é uma "função em escada", não decrescente e tal que

$$F(x) = 0,\ x<0 \text{ e } F(x) = 1,\ x \geq 2.$$

2.13- Para a função F(x) do problema anterior, obtenha F(-5), F(-2), F(0,5) e F(556). Qual o valor de F(1)-F(0)? e de F(-1)-F(-2)?

2.14- O diagrama de dispersão entre as variáveis U e V é dado abaixo. Obter a distribuição conjunta e as distribuições marginais de U e V; somente pelo exame do gráfico, obter E(U).



2.15- Seja X com distribuição dada abaixo.

x	0	1	2
p	1/2	1/4	1/4

Calcule $E(X)$. Considere a v.a. $(X-a)^2$ e calcule $E(X-a)^2$ para $a = 0,\ 1/4,\ 1/2,\ 3/4,\ 1$. Obtenha o gráfico de $E(X-a)^2 = g(a)$. Para qual valor de a, $g(a)$ é mínimo?

2.16- Obtenha formalmente o resultado de 2.15, utilizando o fato que $X-a = X-E(X)+E(X)-a$.

2.17- Se $E(X) = \mu_X$, $E(Y) = \mu_Y$, a *covariância* entre X e Y é definida como

$$\mathrm{Cov}(X,Y) = E[(X-\mu_X)(Y-\mu_Y)].$$

Mostre que:

a) $\mathrm{Cov}(X,Y) = \mathrm{Cov}(Y,X)$;

b) $\mathrm{Cov}(X,X) = \mathrm{Var}(X)$;

c) $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\,\mathrm{Cov}(X,Y)$;

d) $\mathrm{Cov}(X,Y) = E(XY) - E(X)\cdot E(Y)$;

e) Se X e Y são independentes, $\mathrm{cov}(X,Y) = 0$.

- o -

3 - ALGUNS MODELOS PROBABILÍSTICOS

Vamos apresentar, neste capítulo, alguns modelos que se adequam a uma série de problemas práticos. Estes modelos envolvem certos *parâmetros*, cujo conhecimento é indispensável quando queremos calcular certas probabilidades. Na maioria dos problemas reais, estes parâmetros são desconhecidos e há necessidade de fazer algum tipo de inferência sobre eles; este assunto será abordado em capítulos seguintes.

3.1 - A DISTRIBUIÇÃO BINOMIAL

Muitos experimentos são tais que os resultados possíveis apresentam ou não uma determinada característica.

(1) Uma moeda é lançada; o resultado ou é "cara" ou não é (coroa);

(2) Um dado é lançado; ou ocorre face 5 ou não (ocorrendo então, uma das faces 1, 2, 3, 4 ou 6);

(3) Uma peça é escolhida ao acaso de um lote contendo 500 peças; esta peça é defeituosa ou não defeituosa;

(4) Uma pessoa escolhida ao acaso dentre 1.000 pessoas é ou não do sexo masculino;

(5) Um livro é escolhido ao acaso de uma biblioteca; este li-

vro é ou não um livro de Matemática.

Em todos estes casos, estamos interessados na ocorrência de um *sucesso* (ocorrência de cara, face 5, peça defeituosa, etc.) ou *fracasso* (ocorrência de coroa, face diferente de 5, peça boa, etc.). Esta terminologia será usada freqüentemente.

Podemos, para cada experimento acima, definir uma v.a. X, que tem apenas dois valores: o valor 1, digamos, se ocorre sucesso e o valor 0 se ocorre fracasso. Se indicamos por p = P(sucesso), então a distribuição de probabilidades de X é dada por

x	1	0
P(X=x)	p	1-p

Os experimentos (1)-(5) são chamados *ensaios de Bernoulli* e a v.a. X tem uma *distribuição de Bernoulli*.

Imagine, agora, que repetimos um ensaio de Bernoulli n vezes, ou, como se diz também, obtemos uma *amostra de tamanho* n de uma distribuição de Bernoulli. Suponha, ainda, que as repetições sejam *independentes*, isto é, o resultado de um ensaio não tem influência nenhuma no resultado de outro qualquer ensaio. Uma amostra particular será constituída de uma seqüência de 0 e 1, tal como (0, 1, 1, 0, 1); aqui, n=5. A probabilidade de uma tal amostra será

$$(1-p)p\cdot p\cdot(1-p)p = p^3(1-p)^2.$$

O número de sucessos nesta amostra é igual a 3,sendo 2 o número de fracassos.

Podemos, considerar, pois as situações:

(1') Uma moeda é lançada três vezes; qual a probabilidade de se obter exatamente duas caras?

(2') Um dado é lançado cinco vezes; qual a probabilidade de se obter face 5 no máximo 3 vezes?

(3') 10 peças são extraídas ao acaso, com reposição, de um lote contendo 500 peças idênticas; qual a probabilidade de que todas sejam defeituosas, sabendo-se que 10% das peças são defeituosas?

(4') Qual a probabilidade que, dentre 5 pessoas escolhidas ao acaso dentre 1.000 pessoas, duas sejam do sexo masculino?

(5') Escolhemos ao acaso 3 livros de uma biblioteca contendo 10.000 volumes; qual a probabilidade de que 2 sejam de Matemática, sabendo-se que a biblioteca possui 1.000 livros de Matemática?

Observe que nos exemplos (4') e (5'), o fato de estarmos extraindo os objetos de um conjunto muito grande, implica que podemos supor que as extrações sejam praticamente independentes.

Vamos resolver o problema (1'), supondo-se que a moeda é "honesta", isto é, P(sucesso) = P(cara) = $\frac{1}{2}$; indiquemos sucesso (cara) por S e fracasso (coroa) por F. Então, estamos interessados na probabilidade do evento

$$A = \{SSF,\ SFS,\ FSS\},$$

ou, em termos da notação anterior, na probabilidade de obter as amostras (1, 1, 0), (1, 0, 1) ou (0, 1, 1). É claro que P(A) = P(SSF) + P(SSF) + P(SFS) + P(FSS) e, devido à independência,

$$P(SSF) = \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{8} = P(SFS) = P(FSS).$$

Portanto, P(A) = 3/8.

Se a moeda não tiver $P(S) = \frac{1}{2}$ e esta for igual a p, $0<p<1$, então, sendo $q = 1-p = P(F)$,

$$P(SSF) = p \cdot p \cdot q = p^2 \cdot q,$$

$$P(SFS) = p \cdot q \cdot p = p^2 \cdot q,$$

$$P(FSS) = q \cdot p \cdot p = p^2 \cdot q, \text{ de modo que}$$

$$P(A) = 3p^2 \cdot q.$$

Uma característica importante dos experimentos que estamos considerando é que estamos interessados apenas no *número total* de sucessos, e não na ordem em que eles ocorrem. Podemos construir a tabela que segue, para n=3 lançamentos da moeda. (Vide tabela na página seguinte).

Vamos designar por $b(k;n,p)$ a probabilidade de se obter k sucessos em n provas de Bernoulli, sendo p a probabilidade de sucesso em cada uma delas. No exemplo dado,

$$b(2;3,\tfrac{1}{2}) = \frac{3}{8},$$

NÚMERO DE SUCESSOS	PROBABILIDADE	p=1/2
0	q^3	1/8
1	$3pq^2$	3/8
2	$3p^2q$	3/8
3	p^3	1/8

enquanto que

$$b(2;3,p) = 3p^2q.$$

Note que $q^3+3pq^2+3p^2q+p^3 = (q+p)^3 = 1$, pois $p+q = 1$.

Em termos de v.a., se S_n indica o número total de sucessos em n provas de Bernoulli, então $P(S_n=k) = b(k;n,p)$. É claro que os possíveis valores de S_n são $0,1,2,\ldots,n$ e os pares $(k;b(k;n,p))$ constituem *a chamada distribuição binomial*. Para o exemplo (1'), com $p=\frac{1}{2}$, obtemos o gráfico da figura 3.1.

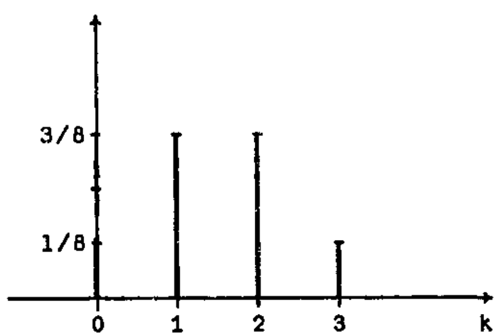


Figura 3.1

Em uma seqüência de n provas, uma seqüência qualquer com k sucessos e n-k fracassos, ou, o que é o mesmo, uma seqüência de k uns e n-k zeros, terá probabilidade

$$p^k \cdot q^{n-k}. \tag{3.1}$$

Esta função, para esta seqüência de resultados, encarada como função de p, é chamada *função de verossimilhança*.

Ainda no exemplo (1'),

$$P(S_3=1) = b(1;3,p) = 3pq^2 = \binom{3}{1}pq^2.$$

Para n=3, podemos escrever, pois,

$$P(S_3=k) = b(k;3,p) = \binom{3}{k}p^k q^{3-k}, \tag{3.2}$$

para k=0, 1, 2, 3.

Para se obter uma fórmula geral, basta saber quantas seqüências com k sucessos e n-k fracassos podemos formar, cada uma tendo probabilidade dada por (3.1). É fácil ver que existem $\frac{n!}{k!(n-k)!} = \binom{n}{k}$ tais seqüências, de modo que

$$b(k;n,p) = \binom{n}{k}p^k q^{n-k}, \tag{3.3}$$

k=0, 1, 2, ..., n.

Vamos considerar, agora, o problema (3'). Aqui, temos n=10 provas de Bernoulli, cada uma com P(sucesso) igual a P(peça defeituosa) = p = 0,1. Queremos

$$P(S_{10}=10) = b(10;10,\tfrac{1}{10}).$$

Utilizando (3.3),

$$b(10;10,\tfrac{1}{10}) = \binom{10}{10}\cdot\left(\frac{1}{10}\right)^{10}\cdot\left(\frac{9}{10}\right)^{0} = \left(\frac{1}{10}\right)^{10}$$

A variável aleatória S_n pode ser pensada como a soma de n variáveis aleatórias, definidas do seguinte modo:

$$X_1 = \begin{cases} 1, & \text{se na primeira prova ocorre sucesso} \\ 0, & \text{se na primeira prova ocorre fracasso} \end{cases}$$

$$X_2 = \begin{cases} 1, & \text{se na segunda prova ocorre sucesso} \\ 0, & \text{se na segunda prova ocorre fracasso} \end{cases}$$

$$\vdots$$

$$X_i = \begin{cases} 1, & \text{se na i-ésima prova ocorre sucesso} \\ 0, & \text{se na i-ésima prova ocorre fracasso} \end{cases}$$

$$\vdots$$

$$X_n = \begin{cases} 1, & \text{se na n-ésima prova ocorre sucesso} \\ 0, & \text{se na n-ésima prova ocorre fracasso} \end{cases}$$

Portanto, $S_n = X_1+X_2+\ldots+X_n$. Observe que todas as variáveis X_1, X_2, ..., X_n têm os mesmos valores (0 ou 1), com as mesmas probabilidades, isto é, $P(X_i=1) = p$, $P(X_i=0)=q=1-p$. Devido à independência das provas, as v.a. $X_1,\ldots,X_n$ são independentes. Dizemos que elas são *independentes e identicamente distribuídas* (i.i.d.).

É imediato que

$$E(X_i) = 1 \times p + 0 \times q = p, \qquad (3.4)$$

ao passo que

$$E(X_i^2) = 1^2 \times p + 0^2 \times q = p,$$

do que decorre

$$Var(X_i) = p - p^2 = p(1-p) = pq. \qquad (3.5)$$

Daí, pela generalização da propriedade (2.12),

$$E(S_n) = np, \qquad (3.6)$$

enquanto que, pela independência e usando uma generalização imediata de (2.15),

$$Var(S_n) = npq. \qquad (3.7)$$

Para o exemplo (3'),

$$E(S_{10}) = 10 \times \frac{1}{10} = 1, \quad Var(S_{10}) = 10 \times \frac{1}{10} \times \frac{9}{10} = \frac{9}{10}.$$

Enquanto foi relativamente fácil obter $b(2;3,\frac{1}{2})$ para o exemplo (1'), não seria imediato obter, por exemplo,

$$b(20;17;0,9) = \binom{20}{17}(0,9)^{17}\cdot(0,1)^3.$$

Felizmente, existem tabelas que dão as probabilidades (3.3) para um grande número de valores de n e p. Utilizando a tabela 1 do Apêndice 1, vemos que a probabilidade acima é dada por 0,19.

Para n grande (maiores que 10, se p não está muito próximo de 0 ou 1) veremos uma maneira de aproximar $b(k;n,p)$.

Se n grande e p está próximo de 0, podemos usar a aproximação

$$b(k;n,p) \cong \frac{e^{-np}(np)^k}{k!}, \qquad (3.8)$$

e uma tabela da função do segundo membro de (3.8) encontra-se no Apêndice 1 (tabela 4).

Por exemplo,

$$b(2;1.000;0,0001) = \frac{e^{-0,1}\cdot(0,1)^2}{2!} = 0,0045.$$

3.2 - A DISTRIBUIÇÃO HIPERGEOMÉTRICA

Este modelo é adequado quando consideramos extrações casuais feitas de uma população, considerada dividida em dois atributos, e estas extrações são feitas *sem reposição*. Para ilustrar, considere uma população de N objetos, r dos quais são defeituosos e N-r são não defeituosos. Um grupo de n elementos é escolhido ao acaso. Vamos calcular a probabilidade de que este grupo contenha k defeituosos. Observe que $0 \le k \le \min(r,n)$. É fácil ver que esta probabilidade é dada por

$$p_k = \frac{\binom{r}{k}\binom{N-r}{n-k}}{\binom{N}{n}} . \qquad (3.9)$$

Os pares (k,p_k) constituem a *distribuição hipergeométrica*.

Se definimos a v.a. X igual ao número de defeituosos na amostra, então $P(X=k)$ é dada por (3.9).

<u>Exemplo 3.1</u> - Em problemas de controle de qualidade, lotes com N elementos são examinados. O número de elementos com defeito, r, é desconhecido. Colhemos uma amostra de n elementos e determinamos k. Somente para ilustrar, suponha que num lote de N=100 peças, r=10 são defeituosos. Escolhendo-se n=5 peças ao acaso, a probabilidade de não se obter peças defeituosas é

$$p_0 = \frac{\binom{10}{0}\binom{90}{5}}{\binom{100}{5}} = \frac{\binom{90}{5}}{\binom{100}{5}},$$

enquanto que a probabilidade de se obter pelo menos uma defeituosa é

$$p_1+p_2+\dots+p_5 = 1-p_0 = 1-\frac{\binom{90}{5}}{\binom{100}{5}}$$

Pode ser demonstrado que a v.a. X definida acima tem média e variância dadas por

$$E(X) = np, \qquad (3.10)$$

$$Var(X) = np(1-p)\cdot\frac{N-n}{N-1}, \qquad (3.11)$$

onde $p=\frac{r}{N}$ é a probabilidade de se obter uma peça defeituosa em uma única extração. Se N é grande quando comparado com n, então extrações com ou sem reposição serão praticamente equivalentes, de modo que as probabilidades dadas por (3.9) serão quase iguais as dadas pela fórmula (3.3), isto é, $p_k \approx b(k;n,p)$. Do mesmo modo, os resultados (3.10) e (3.11) serão praticamente iguais aos valores correspondentes da distribuição binomial(*).

3.3 - A DISTRIBUIÇÃO NORMAL

Vamos considerar a distribuição binomial para $p = \frac{1}{2}$ e n variável, digamos n=5, 10 e 20. Obtemos os histogramas da figura 3.2.

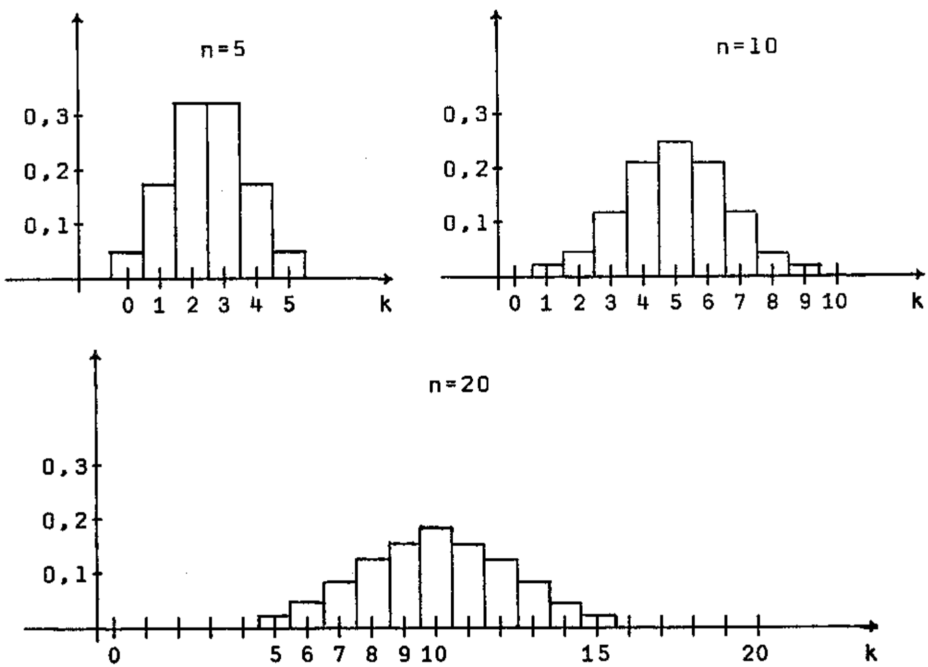


Figura 3.2

(*) Formalmente, considere limites para $N \to \infty$, com $\frac{r}{N} = p$ fixo.

Dado que p=1/2, as distribuições serão simétricas ao redor das respectivas médias, $\frac{5}{2}$, 5 e 10.

Vejamos, agora, os histogramas para o caso $p = \frac{1}{4}$ e os mesmos valores de n acima, isto é, 5, 10 e 20 (Fig. 3.3).

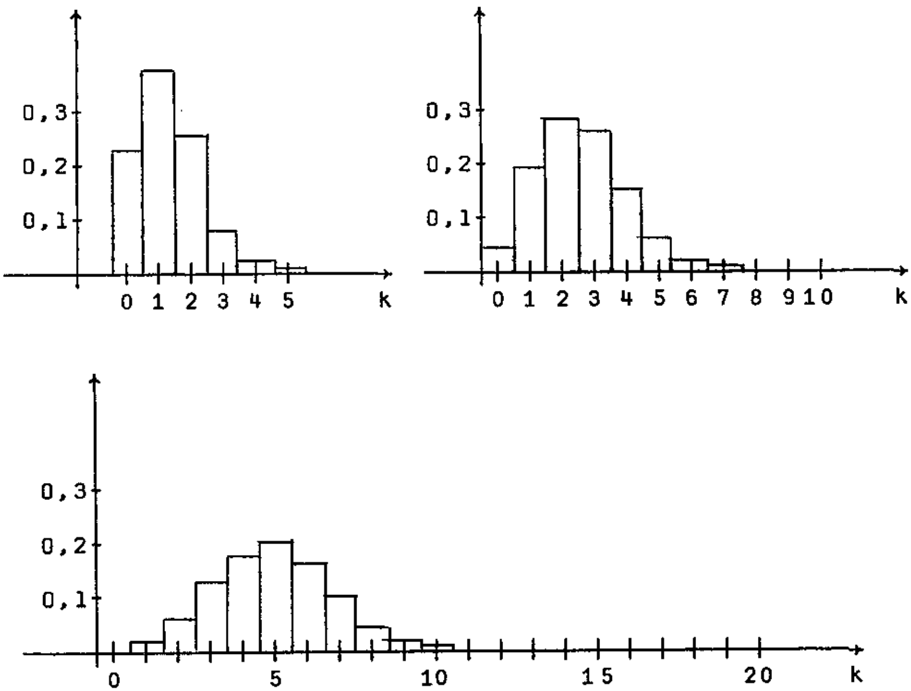


Figura 3.3

As médias agora são 1,25; 2,5 e 5 e as distribuições não são simétricas ao redor destas médias.

Observando as figuras podemos observar os seguintes fatos:

a) para $p=\frac{1}{2}$, por exemplo, à medida que n cresce, a distribuição move-se para a direita, torna-se mais acha tada e mais espalhada;

b) para $p=\frac{1}{4}$, quando n cresce, a distribuição vai "perdendo a sua assimetria" (observe o histograma para n=20)

Vimos que, quando n é grande, é difícil calcular as probabilidades $b(k;n,p)$. A idéia é obtermos aproximações para estas probabilidades. Observando os histogramas acima, vemos que uma sugestão é aproximar a área do histograma pela área sob uma curva contínua. Tal curva é chamada *curva normal*. E aproximaremos probabilidades relativas a uma binomial por probabilidades relativas a uma *distribuição normal*. (Veja figura 3.4, para $n=10$, $p=\frac{1}{2}$).

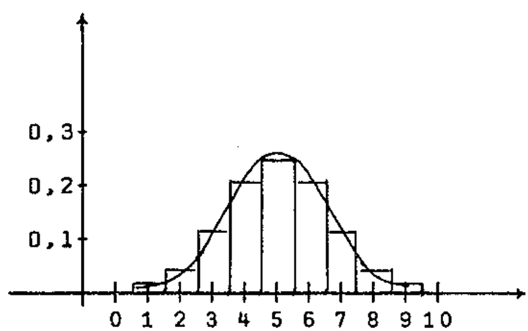


Figura 3.4

A distribuição normal difere das distribuições apresentadas até agora no sentido que é uma distribuição variável contínua, ou seja, através desta distribuição associamos probabilidades a intervalos de números reais. Por exemplo, se escolhemos ao acaso um indivíduo de uma população e medimos sua altura, podemos supor que ao resultado do experimento associamos uma v.a. X que assume valores em um certo intervalo de números reais, [0,3], por exemplo, tomando o metro como unidade. É claro que esta é uma aproximação conveniente da realidade, pois a medida da altura do indivíduo dependerá da precisão do instrumento de medida. Assim, se dizemos que a altura encontrada é 1,75 m, na realidade esta será um valor qualquer entre 1,745 m e 1,755 m.

Na seção seguinte daremos algumas idéias sobre distribuições de variáveis contínuas; para uma tal v.a. X a probabilidade de X estar num certo intervalo, [a,b] digamos, é calculada como sendo a área, entre a e b, sob uma curva, que é chamada função de densidade de X. Tal função será sempre não negativa e a área total sob a curva será igual à unidade. Compare com um histograma, cuja área total também é unitária. (Ver figura 3.5).

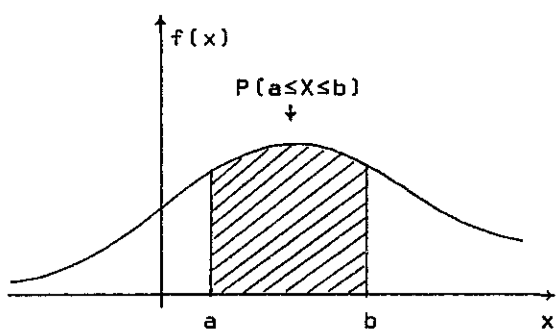


Figura 3.5

A distribuição normal é caracterizada por uma função densidade, e o gráfico desta função tem a forma vista na figura 3.4.

Na realidade, não temos *uma* distribuição normal, mas uma *família* de distribuições normais, que são caracterizadas por dois parâmetros, que chamaremos μ e σ, $-\infty < \mu < \infty$ e $\sigma>0$. Variando μ e σ teremos os diversos membros da família.

Nos gráficos da figura 3.6 temos três exemplos:

a) curva normal com $\mu=0$ e $\sigma=1$;

b) curva normal com $\mu=0$ e $\sigma=2$;

c) curva normal com $\mu=2$ e $\sigma=1$.

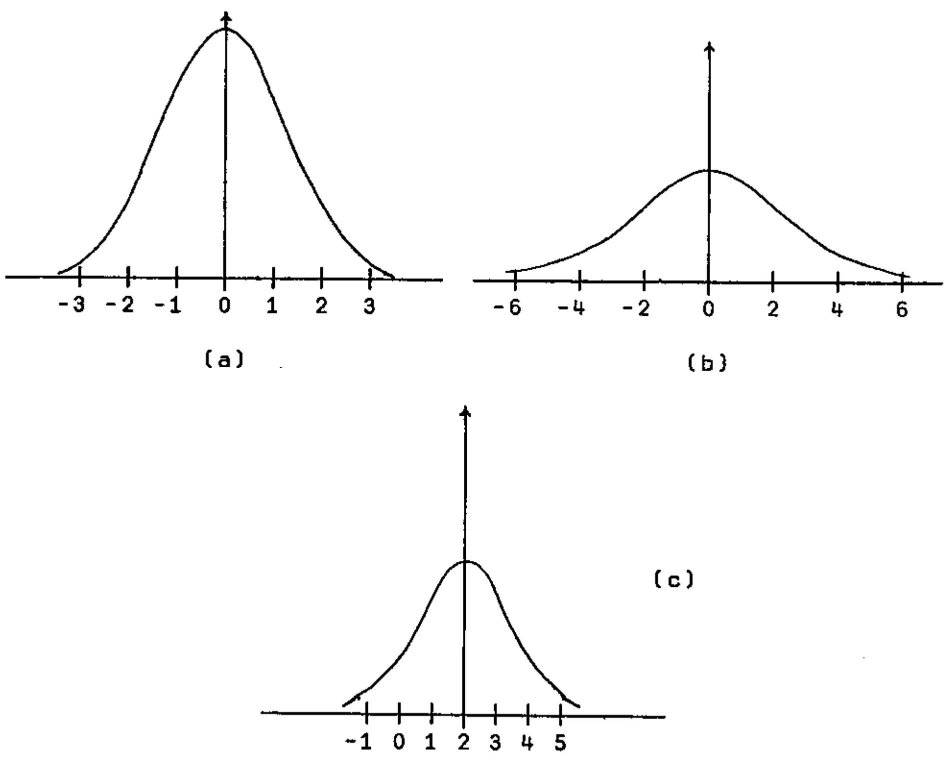


Figura 3.6

Algumas observações sobre as curvas:

a) Para um mesmo μ ($\mu=0$), a curva é mais achatada e mais espalhada para um σ maior;

b) as curvas são simétricas em relação ao ponto μ;

c) praticamente toda a área está concentrada entre os pontos $\mu-3\sigma$ e $\mu+3\sigma$.

O que pode ser demonstrado é que os parâmetros μ e σ correspondem à média e ao desvio padrão da distribuição; portanto σ^2 é a variância da v.a. em questão.

Para calcularmos probabilidades sob uma curva normal é necessário, antes, sabermos calcular probabilidades sob uma particular normal, chamada *normal reduzida* ou *normal padrão*, que é caracterizadas pelos valores $\mu=0$ e $\sigma=1$. Dizemos, também, que temos uma "normal 0,1" e escrevemos: $N(0,1)$. Ver figura 3.6 (a). A tabela 2 do Apêndice 1 dá as probabilidades sob uma curva normal padrão, que nada mais

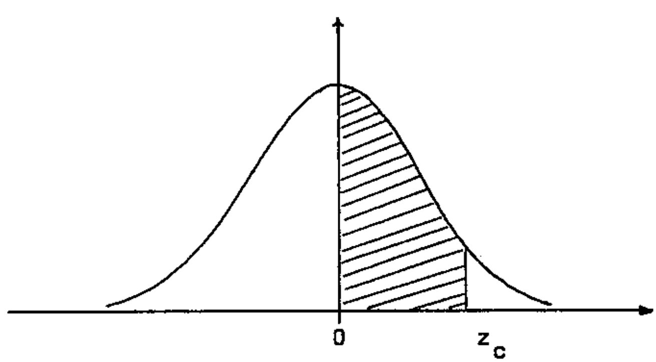


Figura 3.7

são do que as correspondentes áreas sob a curva. A tabela em questão fornece a probabilidade de que a variável Z, normal padrão, esteja entre 0 e um valor, z_c: $P(0 \leq Z \leq z_c)$. (Fig.3.7).

Assim, se $z_c = 1{,}73$, $P(0 \leq Z \leq 1{,}73) = 0{,}4582$.

Observe que (figura 3.8):

1) $P(-1{,}73 \leq Z \leq 0) = P(0 \leq Z \leq 1{,}73) = 0{,}4582$, devido à simetria da curva.

2) $P(Z \geq 1{,}73) = 0{,}5 - P(0 \leq Z \leq 1{,}73) = 0{,}5 - 0{,}4582 = 0{,}0418$, pois $P(Z \geq 0) = 0{,}5 = P(Z \leq 0)$.

3) $P(Z < -1{,}73) = P(Z \geq 1{,}73) = 0{,}0418$.

4) $P(0{,}47 \leq Z \leq 1{,}73) = P(0 \leq Z \leq 1{,}73) - P(0 \leq Z \leq 0{,}47) =$

$= 0{,}4582 - 0{,}1808 = 0{,}2774$.

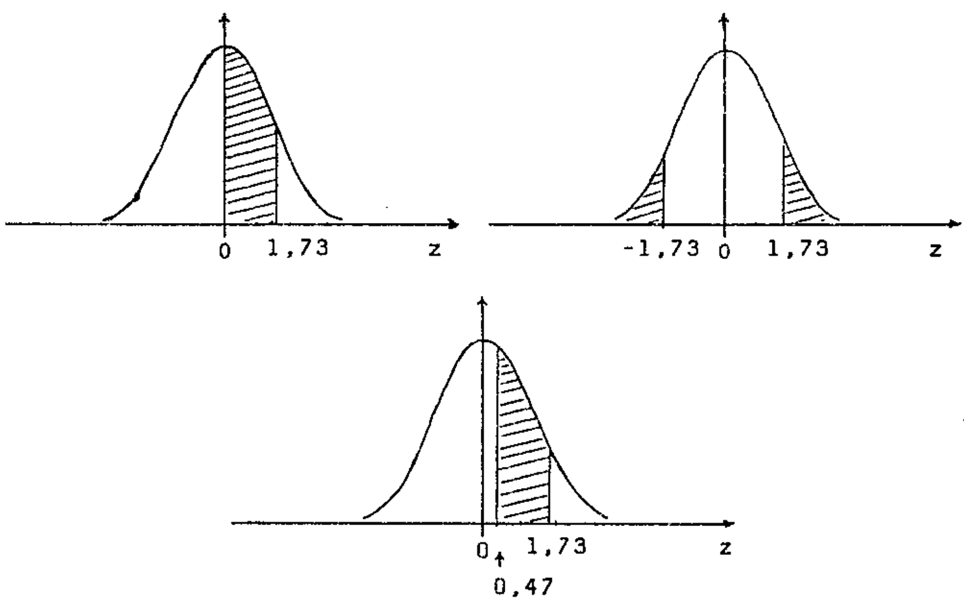


Figura 3.8

Suponha, agora, que X seja uma v.a. com distribuição normal de média μ e variância σ^2, que indicaremos $N(\mu,\sigma^2)$. Então, a v.a. Z, tal que

$$Z = \frac{X-\mu}{\sigma} \tag{3.12}$$

terá distribuição normal com média 0 e variância 1, como é fácil ver (O fato que Z é normal não será provado aqui).

De fato,

$$E(Z) = E\left(\frac{X-\mu}{\sigma}\right) = \frac{1}{\sigma}E(X-\mu) = \frac{1}{\sigma}E(X) - \frac{1}{\sigma}E(\mu) = \frac{\mu}{\sigma} - \frac{\mu}{\sigma} = 0,$$

$$Var(Z) = Var\left(\frac{X-\mu}{\sigma}\right) = \frac{1}{\sigma^2}Var(X-\mu) = \frac{1}{\sigma^2}Var(X) = \frac{\sigma^2}{\sigma^2} = 1$$

Por exemplo, calculemos $P(2 \leq X \leq 5)$ se X é $N(3,16)$, isto é, $\mu=3$, $\sigma^2=16$. Utilizando (3.12),

$$P(2 \leq X \leq 5) = P\left(\frac{2-\mu}{\sigma} \leq \frac{X-\mu}{\sigma} \leq \frac{5-\mu}{\sigma}\right) = P\left(\frac{2-3}{4} \leq Z \leq \frac{5-3}{4}\right) = P\left(-\frac{1}{4} \leq Z \leq \frac{1}{2}\right).$$

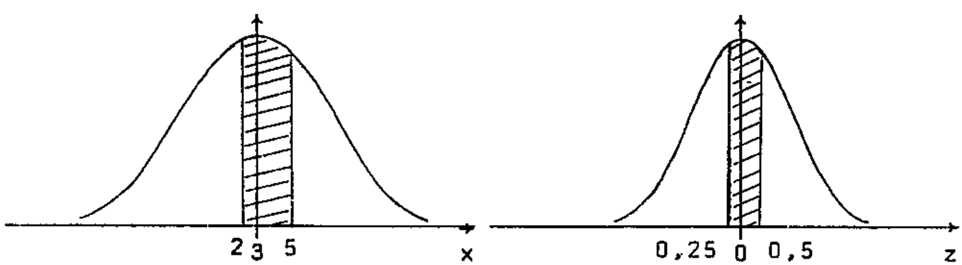


Figura 3.9

Portanto, a probabilidade que X esteja entre 2 e 5 é igual à probabilidade que Z esteja entre -0,25 e 0,5, e u tilizando a tabela 2, vemos que

$$P(-0,25 \leq Z \leq 0,5) = 0,0987 + 0,1915 = 0,2902$$

ou seja,

$$P(2 \leq X \leq 5) = 0,2902.$$

Vejamos, agora, como podemos aproximar probabilida des binomiais através da normal.

Suponha que a v.a. Y tem distribuição binomial com $n=10$ e $p=\frac{1}{2}$ e queremos calcular $P(Y \geq 7)$. Vemos que (fig.3.10)

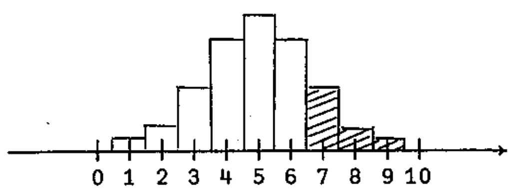


Figura 3.10

$P(Y=7)$ = área do retângulo de base unitária e altu ra igual à $P(Y=7)$, e assim por diante, logo $P(Y \geq 7)$ é a soma das áreas dos retângulos hachurados na figura 3.10. A idéia é aproximar tal área pela área, sob a curva normal, *à direi ta de 6,5*. Qual curva normal? Aquela de média

$\mu = np = 10 \times \frac{1}{2} = 5$ e variância $\sigma^2 = np(1-p) = 10 \times \frac{1}{2} \times \frac{1}{2} = 2,5$.

Ver figura 3.11.

Suponha, agora, que X seja uma v.a. com distribuição normal de média μ e variância σ^2, que indicaremos $N(\mu,\sigma^2)$. Então, a v.a. Z, tal que

$$Z = \frac{X-\mu}{\sigma} \tag{3.12}$$

terá distribuição normal com média 0 e variância 1, como é fácil ver (O fato que Z é normal não será provado aqui).

De fato,

$$E(Z) = E\left(\frac{X-\mu}{\sigma}\right) = \frac{1}{\sigma}E(X-\mu) = \frac{1}{\sigma}E(X) - \frac{1}{\sigma}E(\mu) = \frac{\mu}{\sigma} - \frac{\mu}{\sigma} = 0,$$

$$Var(Z) = Var\left(\frac{X-\mu}{\sigma}\right) = \frac{1}{\sigma^2}Var(X-\mu) = \frac{1}{\sigma^2}Var(X) = \frac{\sigma^2}{\sigma^2} = 1$$

Por exemplo, calculemos $P(2 \leq X \leq 5)$ se X é $N(3,16)$, isto é, $\mu=3$, $\sigma^2=16$. Utilizando (3.12),

$$P(2 \leq X \leq 5) = P\left(\frac{2-\mu}{\sigma} \leq \frac{X-\mu}{\sigma} \leq \frac{5-\mu}{\sigma}\right) = P\left(\frac{2-3}{4} \leq Z \leq \frac{5-3}{4}\right) = P\left(-\frac{1}{4} \leq Z \leq \frac{1}{2}\right).$$

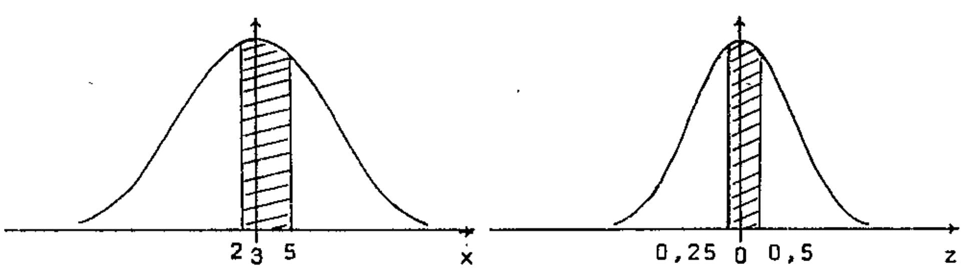


Figura 3.9

Portanto, a probabilidade que X esteja entre 2 e 5 é igual à probabilidade que Z esteja entre -0,25 e 0,5, e u tilizando a tabela 2, vemos que

$$P(-0,25 \leq Z \leq 0,5) = 0,0987 + 0,1915 = 0,2902$$

ou seja,

$$P(2 \leq X \leq 5) = 0,2902.$$

Vejamos, agora, como podemos aproximar probabilidades binomiais através da normal.

Suponha que a v.a. Y tem distribuição binomial com $n=10$ e $p=\frac{1}{2}$ e queremos calcular $P(Y \geq 7)$. Vemos que (fig.3.10)

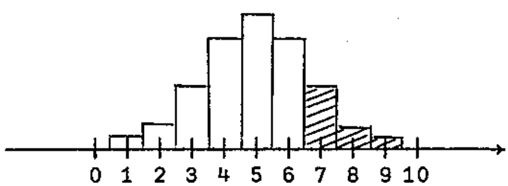


Figura 3.10

$P(Y=7)$ = área do retângulo de base unitária e altura igual à $P(Y=7)$, e assim por diante, logo $P(Y \geq 7)$ é a soma das áreas dos retângulos hachurados na figura 3.10. A idéia é aproximar tal área pela área, sob a curva normal, *à direita de 6,5*. Qual curva normal? Aquela de média

$\mu = np = 10 \times \frac{1}{2} = 5$ e variância $\sigma^2 = np(1-p) = 10 \times \frac{1}{2} \times \frac{1}{2} = 2,5$.
Ver figura 3.11.

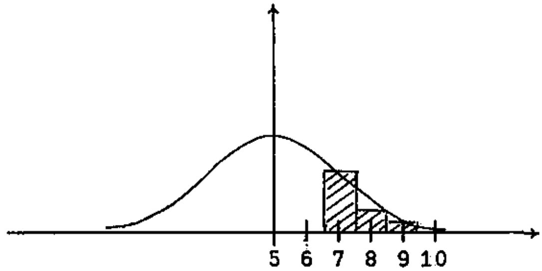


Figura 3.11

Chamando X tal v.a. normal,

$$P(Y\geq 7) \simeq P(X\geq 6,5) = P\left(\frac{X-\mu}{\sigma}\geq\frac{6,5-\mu}{\sigma}\right) = P\left(\frac{X-5}{\sqrt{2,5}}\geq\frac{6,5-5}{\sqrt{2,5}}\right) =$$

$$= P\left(Z \geq \frac{1,5}{1,58}\right) = P(Z \geq 0,94) = 0,1736.$$

A verdadeira probabilidade é 0,172.

Vamos calcular $P(3<Y\leq 6) = P(Y=4) + P(Y=5) + P(Y=6)$.

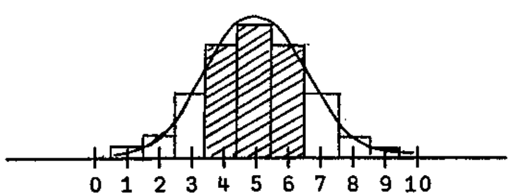


Figura 3.12

Vemos, através da figura 3.12, que a aproximação a ser feita deve ser

$$P(3<Y\leq 6) \cong P(3,5\leq X\leq 6,5) = P\left(\frac{3,5-5}{1,58}\leq Z\leq\frac{6,5-5}{1,58}\right) =$$

$$= P(-0,94 \leq Z \leq 0,94) = 0,6528,$$

ao passo que a probabilidade verdadeira é 0,656.

A justificativa formal de tal aproximação é dada pelo chamado Teorema de De Moivre-Laplace, que é caso particular do chamado Teorema do Limite Central. Ver Apêndice 2.

3.4 - VARIÁVEIS ALEATÓRIAS CONTÍNUAS

Vimos até agora, com exceção da normal, variáveis aleatórias que têm distribuições discretas, ou seja, os valores da v.a. pertencem a um conjunto enumerável (finito ou infinito) de números.

Quando isto não acontece, necessitamos de variáveis que podem tomar valores em um contínuo de pontos, de tal sorte que a "massa" em um particular ponto seja 0. Só terá sentido, para uma variável aleatória contínua X, falar na probabilidade de X pertencer a um certo intervalo [a,b]. Para uma v.a. discreta Y falamos na probabilidade de Y assumir um valor y_j, isto é, $P(Y=y_j)$.

A maneira de caracterizar uma v.a. contínua X é através de uma função $f(\cdot)$, que seja não negativa e de tal modo que a área total sob a curva representativa de $f(\cdot)$ seja unitária. Deste modo, definimos (figura 3.13)

$$P(a \leq X \leq b) = \text{área sob } f(\cdot) \text{ desde a até b} \qquad (3.13)$$

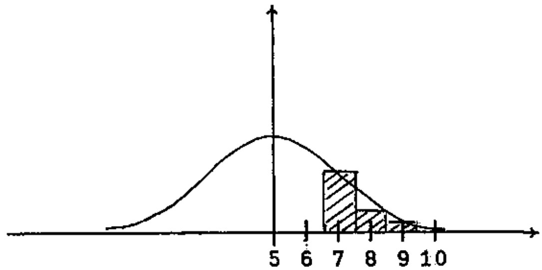


Figura 3.11

Chamando X tal v.a. normal,

$$P(Y\geq 7) \simeq P(X\geq 6,5) = P\left(\frac{X-\mu}{\sigma}\geq\frac{6,5-\mu}{\sigma}\right) = P\left(\frac{X-5}{\sqrt{2,5}}\geq\frac{6,5-5}{\sqrt{2,5}}\right) =$$

$$= P\left(Z \geq \frac{1,5}{1,58}\right) \simeq P(Z \geq 0,94) = 0,1736.$$

A verdadeira probabilidade é 0,172.

Vamos calcular $P(3<Y\leq 6) = P(Y=4) + P(Y=5) + P(Y=6)$.

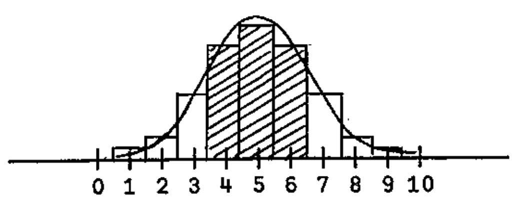


Figura 3.12

Vemos, através da figura 3.12, que a aproximação a ser feita deve ser

$$P(3<Y\leq 6) \simeq P(3,5\leq X\leq 6,5) = P\left(\frac{3,5-5}{1,58}\leq Z\leq\frac{6,5-5}{1,58}\right) =$$

$$= P(-0,94 \leq Z \leq 0,94) = 0,6528,$$

ao passo que a probabilidade verdadeira é 0,656.

A justificativa formal de tal aproximação é dada pelo chamado Teorema de De Moivre-Laplace, que é caso particular do chamado Teorema do Limite Central. Ver Apêndice 2.

3.4 - VARIÁVEIS ALEATÓRIAS CONTÍNUAS

Vimos até agora, com exceção da normal, variáveis aleatórias que têm distribuições discretas, ou seja, os valores da v.a. pertencem a um conjunto enumerável (finito ou infinito) de números.

Quando isto não acontece, necessitamos de variáveis que podem tomar valores em um contínuo de pontos, de tal sorte que a "massa" em um particular ponto seja 0. Só terá sentido, para uma variável aleatória contínua X, falar na probabilidade de X pertencer a um certo intervalo [a,b]. Para uma v.a. discreta Y falamos na probabilidade de Y assumir um valor y_j, isto é, $P(Y=y_j)$.

A maneira de caracterizar uma v.a. contínua X é através de uma função $f(\cdot)$, que seja não negativa e de tal modo que a área total sob a curva representativa de $f(\cdot)$ seja unitária. Deste modo, definimos (figura 3.13)

$$P(a \leq X \leq b) = \text{área sob } f(\cdot) \text{ desde a até b} \qquad (3.13)$$

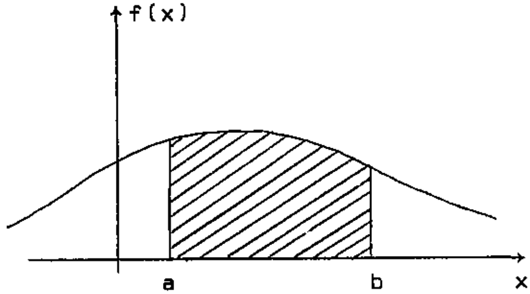


Figura 3.13

A função $f(\cdot)$ é chamada *função densidade* de X.

Exemplo 3.2 - Se $f(x) = 2x$, para $0 \leq x \leq 1$ e zero fora deste intervalo, vemos que $f(x) \geq 0$, $\forall x$ e a área sob $f(\cdot)$ é unitária. (Figura 3.14).

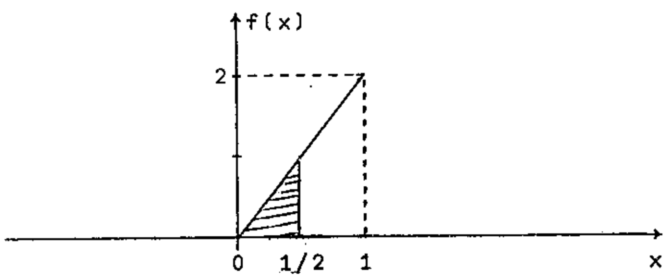


Figura 3.14

Se queremos calcular $P(0 \leq X \leq \frac{1}{2})$, vemos que esta probabilidade é igual a área do triângulo de base $\frac{1}{2}$ e altura 1 hachurado na figura 3.14, logo a probabilidade em questão é $\frac{1/2 \times 1}{2} = 1/4$.

Exemplo 3.3 - A chamada *distribuição uniforme* sobre o intervalo $[0,1]$ é aquela cuja função densidade é dada por:

$$f(x) = \begin{cases} 1, & x \in [0,1] \\ 0, & x \notin [0,1]. \end{cases} \qquad (3.14)$$

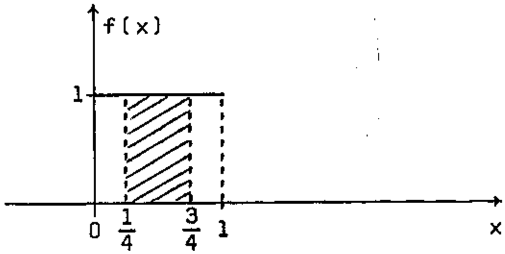


Figura 3.15

Observe que $P\left(\frac{1}{4} \leq X \leq \frac{3}{4}\right) = \frac{1}{2}$, que é a área do retângulo de lados 1 e 3/4 - 1/4 = 2/4 = 1/2 da figura 3.15.

Formalmente,

$$P(a \leq X \leq b) = \int_a^b f(x)dx. \qquad (3.15)$$

Se queremos calcular a média de uma v.a. X com função densidade f(x), podemos pensar da seguinte maneira. Primeiramente, observe que f(x) não é uma probabilidade,mas se h é pequeno, (figura 3.16).

$$P(x \leq X \leq x+h) \simeq h \cdot f(x). \qquad (3.16)$$

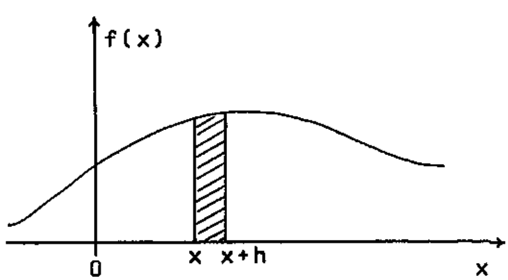


Figura 3.16

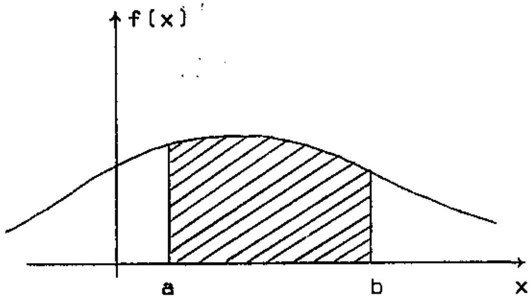


Figura 3.13

A função $f(\cdot)$ é chamada *função densidade* de X.

Exemplo 3.2 - Se $f(x) = 2x$, para $0 \le x \le 1$ e zero fora deste intervalo, vemos que $f(x) \ge 0$, $\forall x$ e a área sob $f(\cdot)$ é unitária. (Figura 3.14).

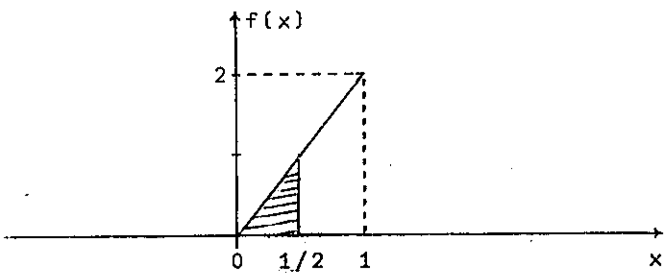


Figura 3.14

Se queremos calcular $P(0 \le X \le \frac{1}{2})$, vemos que esta probabilidade é igual a área do triângulo de base $\frac{1}{2}$ e altura 1 hachurado na figura 3.14, logo a probabilidade em questão é $\frac{1/2 \times 1}{2} = 1/4$.

Exemplo 3.3 - A chamada *distribuição uniforme* sobre o intervalo $[0,1]$ é aquela cuja função densidade é dada por:

$$f(x) = \begin{cases} 1, & x \in [0,1] \\ 0, & x \notin [0,1]. \end{cases} \tag{3.14}$$

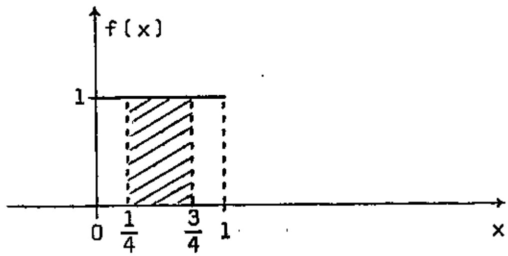


Figura 3.15

Observe que $P\left(\frac{1}{4} \le X \le \frac{3}{4}\right) = \frac{1}{2}$, que é a área do retângulo de lados 1 e 3/4 - 1/4 = 2/4 = 1/2 da figura 3.15.

Formalmente,

$$P(a \le X \le b) = \int_a^b f(x)dx. \tag{3.15}$$

Se queremos calcular a média de uma v.a. X com função densidade f(x), podemos pensar da seguinte maneira. Primeiramente, observe que f(x) não é uma probabilidade, mas se h é pequeno, (figura 3.16).

$$P(x \le X \le x+h) \approx h \cdot f(x). \tag{3.16}$$

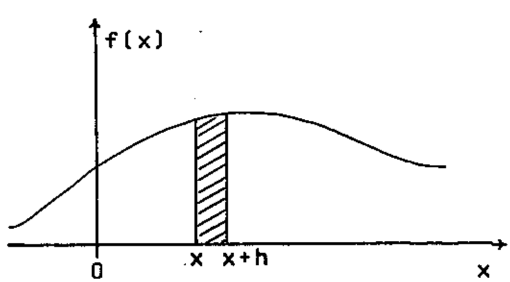


Figura 3.16

Portanto, dividamos o intervalo $[a,b]$ onde $f(x) > 0$ em n partes de amplitudes iguais a $h = \frac{b-a}{n}$ (figura 3.17).

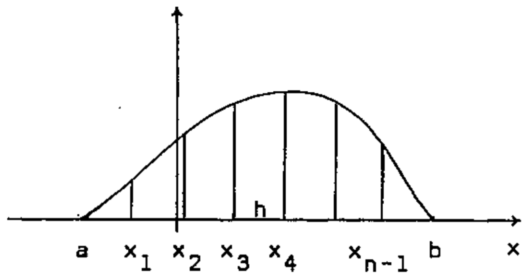


Figura 3.17

Defina a v.a. Y que assume os valores $x_1, x_2, \ldots$. Por (3.16), temos, por exemplo,

$$P(Y=x_3) \approx P(x_3 \leq Y \leq x_3+h) \approx f(x_3)\cdot h.$$

Desta maneira,

$$E(Y) = \sum_{i=1}^{n} x_i \; f(x_i)h. \qquad (3.17)$$

Defina, então,

$$E(X) = \lim_{n\to\infty} E(Y). \qquad (3.18)$$

Para o exemplo 3.2, cada intervalo terá amplitude $1/n$ e $x_i = i/n$, $i=0,\ldots,n-1$, portanto,

$$E(Y) = \sum_{i=0}^{n-1} \frac{i}{n}.\frac{2i}{n}.\frac{1}{n} = \frac{2}{n^3}\sum_{i=0}^{n-1} i^2 = \frac{2}{6}(1-\frac{1}{n})(2-\frac{1}{n}) \to \frac{4}{6},$$

quando $n\to\infty$.

É claro que, da definição de integral, de (3.17) vem

$$E(X) = \int_a^b x\ f(x)dx. \qquad (3.19)$$

<u>Exemplo 3.4</u> - A *função de densidade normal* é dada por

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\ e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \quad -\infty<x<\infty, \qquad (3.20)$$

com $-\infty<\mu<\infty$ e $\sigma>0$.

Para uma v.a. com densidade (3.20) temos:

a) $E(X) = \mu$ e $Var(X) = \sigma^2$;

b) $\mu\pm\sigma$ são dois pontos de inflexão da curva;

c) $\frac{1}{\sigma\sqrt{2\pi}}$ é o ponto máximo de $f(x)$;

d) $\lim_{x\to\pm\infty} f(x) = 0$.

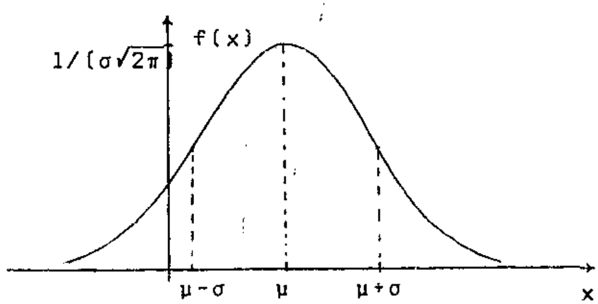


Figura 3.18

Temos, na realidade, uma família $f(x;\mu,\sigma)$ de densidades normais, variando μ e σ.

<u>Exempo 3.5</u> - Dizemos que X tem *distribuição exponencial* se

Portanto, dividamos o intervalo $[a,b]$ onde $f(x) > 0$ em n partes de amplitudes iguais a $h = \frac{b-a}{n}$ (figura 3.17).

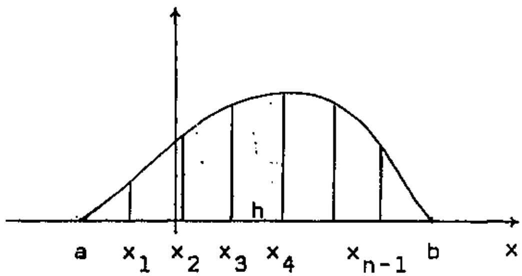


Figura 3.17

Defina a v.a. Y que assume os valores $x_1, x_2, \ldots$. Por (3.16), temos, por exemplo,

$$P(Y=x_3) \approx P(x_3 \leq Y \leq x_3+h) \approx f(x_3)\cdot h.$$

Desta maneira,

$$E(Y) = \sum_{i=1}^{n} x_i \; f(x_i)h. \qquad (3.17)$$

Defina, então,

$$E(X) = \lim_{n\to\infty} E(Y). \qquad (3.18)$$

Para o exemplo 3.2, cada intervalo terá amplitude $1/n$ e $x_i = i/n$, $i=0,\ldots,n-1$, portanto,

$$E(Y) = \sum_{i=0}^{n-1} \frac{i}{n}\cdot\frac{2i}{n}\cdot\frac{1}{n} = \frac{2}{n^3}\sum_{i=0}^{n-1} i^2 = \frac{2}{6}(1-\frac{1}{n})(2-\frac{1}{n}) \to \frac{4}{6},$$

quando $n\to\infty$.

É claro que, da definição de integral, de (3.17) vem

$$E(X) = \int_a^b x\ f(x)dx. \tag{3.19}$$

Exemplo 3.4 - A *função de densidade normal* é dada por

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\ e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2},\quad -\infty<x<\infty, \tag{3.20}$$

com $-\infty<\mu<\infty$ e $\sigma>0$.

Para uma v.a. com densidade (3.20) temos:

a) $E(X) = \mu$ e $Var(X) = \sigma^2$;

b) $\mu\pm\sigma$ são dois pontos de inflexão da curva;

c) $\frac{1}{\sigma\sqrt{2\pi}}$ é o ponto máximo de $f(x)$;

d) $\lim_{x\to\pm\infty} f(x) = 0$.

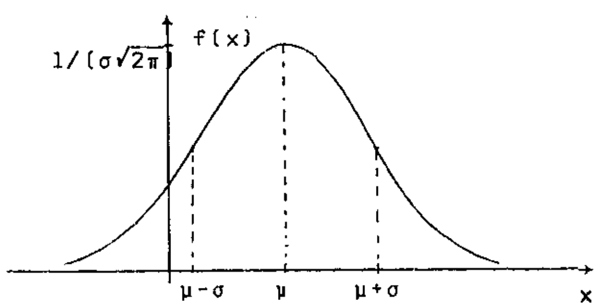


Figura 3.18

Temos, na realidade, uma família $f(x;\mu,\sigma)$ de densidades normais, variando μ e σ.

Exempo 3.5 - Dizemos que X tem *distribuição exponencial* se a densidade de X é dada por

$$f(x) = \begin{cases} \alpha \; e^{-\alpha x}, & x \geq 0 \\ 0 & , \; x < 0 \end{cases} \qquad (3.21)$$

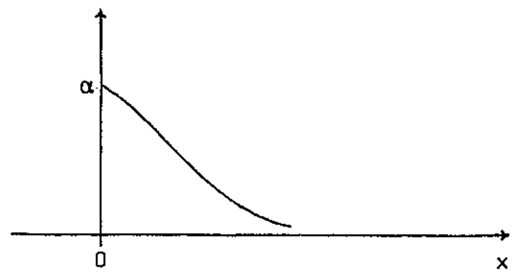


Figura 3.19

É fácil ver que $E(X) = \frac{1}{\alpha}$ e $Var(X) = \frac{1}{\alpha^2}$. (Ver exercícios).

PROBLEMAS PARA O CAPÍTULO 3

3.1 - Obtenha $b(k;n,p)$ para:

a) $n=10$, $p=0,4$, $k=3$;

b) $n=10$, $p=0,4$, $k=9$;

c) $n=15$, $p=0,5$, $k=10$.

3.2 - Se X é uma v.a. binomial com parâmetro $n=5$ e $p=\frac{1}{2}$, obtenha:

a) $P(X=0 \text{ ou } X=1)$

b) $P(X \geq 2)$

c) $P(X \leq 4)$

d) $P(2 < X \leq 5)$.

3.3 - Num teste tipo certo-errado, com 50 questões, qual é a probabilidade que um aluno acerte 80% das questões, supondo que ele as responde ao acaso?

3.4 - Mesmo problema, com 5 alternativas para cada questão.

3.5 - Numa partida de 20 rádios foram escolhidas, ao acaso, 5 unidades para serem inspecionadas. Sabe-se que há 20% de rádios defeituosos. Calcular as probabilidades abaixo, nos dois caos:

a) extração com reposição;

b) extração sem reposição:

1) probabilidade que a amostra seja formada só por rádios com defeito;

2) probabilidade que haja apenas um com defeito;

3) probabilidade que todos sejam bons.

3.6 - Em um experimento binomial com 3 provas, a probabilidade de exatamente 2 sucessos é 12 vezes a probabilidade de 3 sucessos. Encontre p.

3.7 - Em uma pesquisa de opinião pública, 3 dentre 4 pessoas entrevistadas são favoráveis a uma certa proposição. Em uma amostra de 10 pessoas entrevistadas, qual a probabilidade que pelo menos 3 sejam favoráveis?

3.8 - Prove que

$$b(k+1;n,p) = \frac{(n-k)p}{(k+1)(1-p)}\, b(k;n,p).$$

3.9 - Y é uma v.a. binomial com n=10 e p=0,4. Determine a aproximação normal para:

a) $P(3<Y<8)$

b) $P(Y\geq 7)$

c) $P(Y<5)$.

3.10- Se uma moeda é lançada 500 vezes, e P(cara) = 0,6, obter a probabilidade de ocorrerem 285 caras.

3.11- Vimos que, para n grande e p pequeno, as probabilidades binomiais b(k;n,p) podem ser aproximadas pelas pro babilidades de Poisson

$$p_k = \frac{e^{-np}\cdot(np)^k}{k!} .$$

Suponha que um processo de fabricação produza itens dos quais 2 em 1.000 são defeituosos. Em um lote de 500 itens, qual a probabilidade que nenhum dos itens seja defeituoso?

3.12- Seja X com distribuição normal com média $\mu=10$ e variância $\sigma^2=4$. Calcular:

a) $P(8<X<10)$

b) $P(9\leq X\leq 12)$

c) $P(X>10)$

d) $P(X<8$ ou $X>11)$.

3.13- As vendas de um determinado produto têm distribuição aproximadamente normal com média 500 e desvio padrão 50. Se a empresa decide fabricar 600 unidades no mês em estudo, qual é a probabilidade que não possa atender a todos os pedidos deste mês, por estar com a produção esgotada?

3.14- Encontre a média e a variância das v.a. cujas densidades são:

a) $f(x) = 6x(1-x)$, $0<x<1$.

b) $f(x) = e^{-x}$, $x>0$.

c) $f(x) = 2x^{-3}$, $x>1$.

3.15- Obtenha E(X) e Var(x) para X com distribuição normal dada por (3.2) do texto.

3.16- Idem, para X com distribuição exponencial dada por (3.21) do texto.

- o -

4 - INTRODUÇÃO À INFERÊNCIA ESTATÍSTICA

"O cidadão comum pensa em "estatísticas" como sendo colunas de números em páginas de esportes ou em seções econômicas de jornais, ilustradas com gráficos em zig-zag, pilhas de moedas e linhas de pessoas. Mas o estatístico de hojenão compila tabelas de dados e os ilustra graficamente. Ele trabalha em tarefas científicas e profissionais que são interessantes e complicadas. Seu trabalho é o de ajudar a planejar experimentos, interpretar dados obtidos de observações e apresentar os resultados de maneira a facilitar a tomada de decisões razoáveis."

4.1 - PROBABILIDADE E ESTATÍSTICA

O trecho acima, extraído do folheto "Carreiras em Estatística", publicado pelo Comitê de Presidentes das Sociedades Estatísticas Americanas, pode tentar descrever o campo de ação da Estatística.

O objetivo da Teoria das Probabilidades é constituir um modelo matemático para uma situação física e, a partir deste modelo, *deduzir* propriedades da situação.

A matéria prima da Estatística é um conjunto de *dados*. A grosso modo, podemos dizer que a Estatística se preocupa em coletar, analisar e fazer inferências a partir de dados.

Portanto, a Inferência Estatística, que se baseia na Teoria das Probabilidades, é uma parte apenas do campo de ação da Estatística. O objetivo da Inferência Estatística é o de *inferir* propriedades de um agregado maior (a *população*) a partir de um conjunto menor (a *amostra*).

A tomada de decisões "razoáveis", mencionada acima, será feita, pois, a partir de dados, que fornecem informação a respeito de uma dada característica da população, na qual estamos interessados. Esta característica pode ser representada por uma variável aleatória, e os dados são, então, valores desta variável. Se tivéssemos informação completa a respeito da distribuição de X, isto é, se conhecessemos os pares de valores $(x_i, P(X=x_i))$, bem como os parâmetros da distribuição (n e p, no caso da binomial, por exemplo), então não haveria necessidade de colher uma amostra de X. Normalmente, nossa informação a respeito de X é parcial, ou mesmo nada conhecemos. Podemos, por exemplo, saber que a distribuição de X é binomial, mas desconhecemos os parâmetros que a caracterizam, ou podemos ter uma idéia da média e da variância de uma distribuição e não sabemos qual a sua forma.

4.2 - MODELOS ESTATÍSTICOS

Vamos apresentar alguns exemplos simples, que nos darão uma idéia do tipo de problemas com os quais um estatístico pode lidar.

Exemplo 4.1 - No exemplo (1') do capítulo 3, tinhamos um experimento envolvendo uma distribuição binomial, na qual supomos que $p=\frac{1}{2}$. Conhecendo-se n, podemos calcular $b(k;n,p)$, para qualquer $k=0,1,\ldots,n$. Suponha, agora, que lançamos uma moeda $n = 50$ vezes e obtemos 40 "caras" (sucessos). Com esta informação (número de sucessos) queremos *estimar* o valor de $p = P$ (sucesso). Temos, aqui, um problema estatístico. Qual seria um bom *estimador* de p? Por outro lado, observando-se o número de caras obtidas, 40, podemos suspeitar da "honestidade" da moeda usada, isto é, podemos querer *testar a hipótese* que $p = \frac{1}{2}$. Com os dados obtidos, podemos ser levados a aceitar ou rejeitar a hipótese formulada. Temos, então, dois problemas básicos da Estatística: o primeiro diz-se um *problema de estimação*, enquanto que o segundo é um *problema de teste de hipóteses*. Os dois capítulos seguintes tratarão com pormenores estes problemas para o caso específico do parâmetro p de uma binomial.

Exemplo 4.2 - Consideremos, novamente, o problema da inspeção de qualidade do capítulo anterior. Lotes de N objetos são sujeitos ao controle de qualidade, mas suponha que o nú mero de peças defeituosas, r, seja desconhecido. Uma amostra de tamanho n é escolhida ao acaso e o número de peças defeituosas, k, é determinado. O modelo para a situação será

$$P(X=k) = \frac{\binom{r}{k}\binom{N-r}{n-k}}{\binom{N}{n}}, \qquad (4.1)$$

onde X = número de defeituosos na amostra. A diferença entre esta situação e a descrita no capítulo 3, é que agora *r é desconhecido.*

Suponha que um comprador tem que tomar uma decisão sobre um lote com N peças. Ele pode, como no exemplo 4.1, estar interessado em estimar r, baseado no valor observado de X. Ou então, pode tomar a decisão de comprar o lote, se k é menor ou igual a um determinado valor e devolver o lote se k é maior que este valor.

Exemplo 4.3 - Suponha que temos uma esfera de aço e medimos o seu diâmetro, obtendo-se o valor d. Se D representa o verdadeiro valor do diâmetro, então podemos estabelecer o modelo

$$d = D + e, \qquad (4.2)$$

onde e representa o erro de mensuração. Se efetuarmos várias medidas, d_1, d_2, ..., d_m do diâmetro, podemos escrever, do mesmo modo,

$$d_i = D + e_i \ , \ i=1,2,\ldots,m. \qquad (4.3)$$

Aqui, d_i é uma v.a., porque e_i será uma v.a. O problema que se apresenta é o de estimar D, ou testar hipóteses a respeito de D, baseando-nos nas m observações d_1, d_2, ..., d_m. Normalmente, várias suposições são feitas a respeito da distribuição das variáveis aleatórias e_i. Ver capítulo 8 para a solução deste problema.

Exemplo 4.4 - Um dado é lançado 300 vezes, com os seguintes resultados:

OCORRÊNCIA	1	2	3	4	5	6
FREQÜÊNCIA	43	49	56	45	66	41

Com estes dados, queremos saber se o dado é "honesto", isto é, a probabilidade de ocorrência de qualquer face seja 1/6. Isto é, queremos testar a hipótese que $p_1=p_2=\ldots=p_6=1/6$, onde $p_i=P(\text{face } i)$, $i=1,\ldots,6$. Este é um tipo de teste bastante comum na prática, chamado *teste de aderência* ou *ajustamento*.

Exemplo 4.5 - Podemos estar interessados em mais de uma categoria em um teste de aderência.

Considere a tabela abaixo, em que 1.000 indivíduos foram classificados de acordo com o sexo e de acordo com o fato de serem ou não daltônicos:

	HOMENS	MULHERES
NORMAIS	442	514
DALTÔNICOS	38	6

Temos, pois, duas categorias (ou atributos). De acordo com um modelo genético, estes números deveriam ser consistentes com

$p/2$	$p^2/2 + pq$
$q/2$	$q^2/2$

onde $q = 1-p$ é a proporção de gens que causam daltonismo na população.

O objetivo, pois, é testar se os dados obtidos são consistentes com o modelo genético proposto. Ver problema 4.5.

Outras vezes, podemos querer testar a independência entre dois atributos, dada uma tabela do tipo acima.

Tais tabelas são chamadas *tabelas de contingência*.

Nos exemplos 4.4 e 4.5, temos casos em que o experimento tem mais de dois resultados possíveis. A distribuição apropriada para descrever tais experimentos é chamada *distribuição multinomial* (ver exercícios), que tem papel relevante na obtenção dos testes mencionados nos dois exemplos anteriores (os chamados *testes do qui-quadrado*).

PROBLEMAS PARA O CAPÍTULO 4

4.1 - Suponha que temos n provas de Bernoulli com

$$P(\text{sucesso}) = p$$

e seja X igual ao número de sucessos obtidos. Qual é um estimador razoável de p?

4.2 - Obtenha $E(\hat{p})$ e $Var(\hat{p})$, onde $\hat{p}$ é o estimador que você encontrou em 4.1.

4.3 - Suponha que capturamos, marcamos e depois soltamos r animais de uma população com N animais, sendo N desconhecido. Depois de algum tempo, capturamos n animais, dos quais constatamos que k são marcados. Descreva um modelo para a situação em dois casos:

a) supondo que n é suficientemente pequeno, comparado com N, de modo que a amostragem pode ser considerada com reposição;

b) amostragem sem reposição.

No caso b, qual um limite inferior para o valor de N?

4.4 - Suponha que quando realizamos um experimento, um e somente um dos eventos $A_1,\ldots,A_k$ ocorre. Ou seja, $A_1,\ldots,A_k$ constituem uma *partição* do espaço amostral Ω. Seja $p_i = P(A_i)$ e suponha que p_i permaneça constante durante as repetições do experimento. Defina as v.a. $X_1,\ldots,X_k$ como segue:

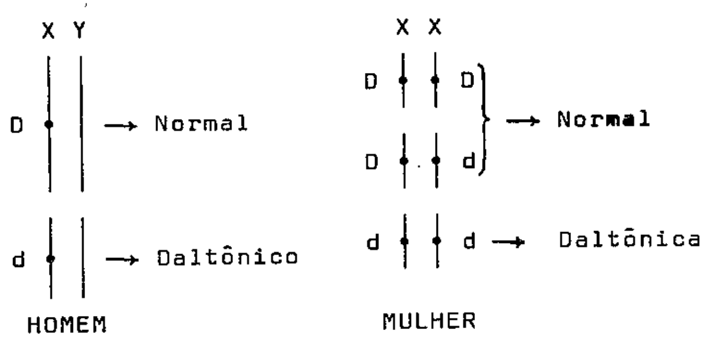
X_i = número de vezes que A_i ocorre dentre as n repetições do experimento, $i=1,\ldots,k$.

Prove que

$$P(X_1=n_1,\ldots,X_k=n_k) = \frac{n!}{n_1!n_2!\ldots n_k!}\, p_1^{n_1}p_2^{n_2}\ldots p_k^{n_k} ,$$

onde $\sum_{i=1}^{k} n_i = n$. Esta é chamada *distribuição multinomial* de probabilidades.

4.5 - Justifique as probabilidades teóricas dadas no exemplo 4.5 tendo em vista o seguinte. O daltonismo é uma característica genética ligada ao sexo. Cada pessoa apresenta um par de cromossomas sexuais, sendo que no homem este par é denominado XY e, na mulher, XX. O daltonismo é caracterizado por um par de gens, que chamaremos D e d, sendo que, no homem, D ou d só podem estar localizados no cromossoma X, ao passo que na mulher podem estar localizados nos dois cromossomas X.



A situação é a esquematizada acima e, no exemplo 4.5, $q = P(d)$, $0<q<1$, isto é, d é o gen que ocasiona o daltonismo. Vemos que se, na mulher, aparece o par DD ou Dd, a pessoa é normal, ao passo que se aparece dd, a pessoa é daltônica. Suponha que a probabilidade de homem ou mulher na população seja 1/2. Em genética, q é chamado *freqüência gênica* do gen para daltonismo.

- o -

5 - ESTIMAÇÃO: PRIMEIRAS IDÉIAS

Neste capítulo vamos considerar tão somente o problema de estimar o parâmetro desconhecido, p, de uma distribuição binomial. Este caso particular servirá para introduzir alguns conceitos importantes.

5.1 - UM ESTIMADOR POR PONTO DE p

Vamos considerar o seguinte

Exemplo 5.1 - Uma amostra de $n = 500$ pessoas de uma cidade é escolhida e a cada pessoa da amostra é feita uma pergunta a respeito de um dado problema municipal, para o qual foi apresentada uma solução pela Prefeitura. A resposta à pergunta poderá ser "SIM" (favorável à solução) ou "NÃO" (contrário à solução). Deseja-se estimar a proporção de pessoas na cidade favorável à solução apresentada.

Como no capítulo 3, definimos as v.a. $X_1, X_2, \ldots, X_n$ onde

$$X_i = \begin{cases} 1, \text{ se a i-ésima pessoa na amostra responde "SIM"}, \\ 0, \text{ se a i-ésima pessoa na amostra responde "NÃO"}, \end{cases}$$

e seja $p = P(\text{sucesso})$, onde aqui "sucesso" significa que a pessoa responde SIM à questão formulada.

Logo, se $S_n = \sum_{i=1}^{n} X_i$, S_n tem distribuição binomial com parâmetros n e p e o problema consiste em estimar p, a probabilidade de uma pessoa escolhida ao acaso da população, responder "SIM". É claro que S_n representa o número de pessoas da amostra que responderam "SIM", logo um possível *estimador* de p é

$$\hat{p} = \frac{S_n}{n} = \frac{\sum_{i=1}^{n} X_i}{n}. \qquad (5.1)$$

Então, se $S_n = k$, isto é, observamos o valor k da v.a. S_n, obtemos $\hat{p} = k/n$ como uma *estimativa* de p. Embora não vamos enfatizar esta diferença, observe que $\hat{p}$ dado por (5.1) é uma v.a., ao passo que k/n é um número, isto é, um valor da v.a. Por exemplo, se 300 pessoas responderam "SIM" à pergunta, então $\hat{p} = \frac{300}{500} = 0{,}6$ e dizemos que 0,6 é uma estimativa da proporção de pessoas na população favorável à iniciativa da Prefeitura.

Vejamos algumas propriedades de $\hat{p}$. Como $E(S_n) = np$ e $Var(S_n) = npq$, $q = 1-p$, segue-se que

$$E(\hat{p}) = p, \qquad (5.2)$$

e

$$Var(\hat{p}) = \frac{npq}{n^2} = \frac{pq}{n}. \qquad (5.3)$$

Para n fixo, a função $\frac{pq}{n} = \frac{p(1-p)}{n}$ atinge seu máximo para $p = 1/2$, logo

$$\mathrm{Var}(\hat{p}) \leq \frac{1}{4n}. \tag{5.4}$$

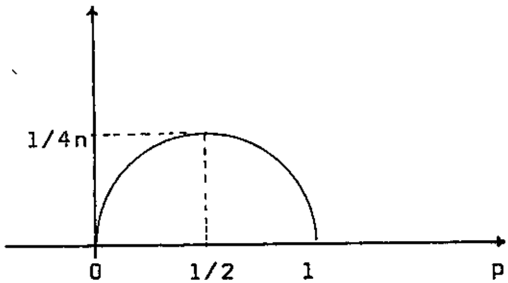


Figura 5.1

No exemplo dado, $\mathrm{var}(\hat{p}) \leq \frac{1}{2.000} = 0{,}0005$; é claro que $\mathrm{Var}(\hat{p})$ tende a decrescer quando n aumenta. Em particular, devido a (5.3), $\mathrm{Var}(\hat{p}) \to 0$ quando $n \to \infty$. Esta propriedade, juntamente com (5.2) nos diz que $\hat{p}$ é um *estimador consistente* de p.

O estimador $\hat{p}$ é também chamado *não viciado*, pelo fato de (5.2) ser válida; ela nos diz que, em média, $\hat{p}$ "acerta p", ou seja, se escolhermos várias amostras, e para cada uma calculamos $\hat{p}$, a média destes valores estará próxima de p. Esta qualidade de um estimador é desejável, mas não é suficiente, como veremos depois. Pode acontecer que para uma particular amostra $\hat{p}$ esteja bem distante de p.

O estimador dado por (5.1) é chamada *estimador por ponto*, pois obtemos um único número que escolhemos como representante do verdadeiro valor de p, que é desconhecido.

Exemplo 5.2 - Somente a título de ilustração, vejamos a distribuição de $\hat{p}$ num caso particular, no qual supomos p conhecido e igual a 1/2. Considere n = 3. Portanto, S_3 assume os valores 0, 1, e 3 com probabilidades 1/8, 3/8, 3/8, e 1/8, respectivamente (ver capítulo 3).

Os valores possíveis de $\hat{p} = \frac{S_n}{n} = \frac{S_3}{3}$ são 0, 1/3, 2/3 e 1, com probabilidades 1/8, 3/8, 3/8 e 1/8, respectivamente. De fato,

$$P(\hat{p}=2/3) = P(S_3/3=2/3) = P(S_3=2) = 3/8, \text{ etc.}$$

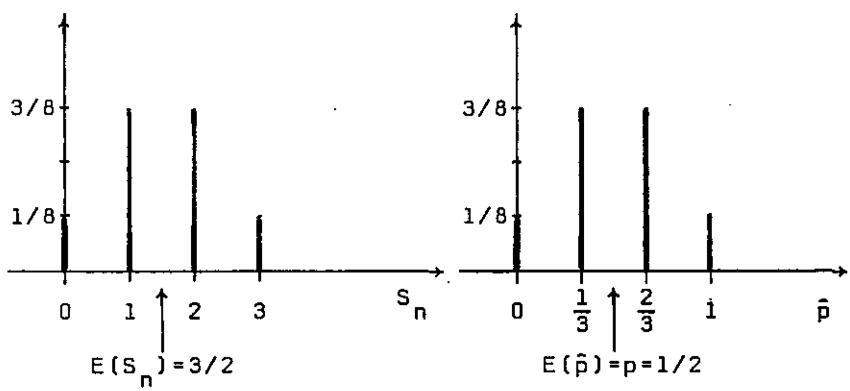


Figura 5.2

Temos que

$$E(S_3) = 3 \times \frac{1}{2} = 1,5$$

e

$$E(\hat{p}) = 0 \times \frac{1}{8} + \frac{1}{3} \times \frac{3}{8} + \frac{2}{3} \times \frac{3}{8} + 1 \times \frac{1}{8} = \frac{1}{2} = p.$$

Ver figura 5.2.

5.2 - A LEI DOS GRANDES NÚMEROS

No capítulo 1 mencionamos que uma maneira de olhar para a probabilidade de um evento A de maneira intuitiva é repetir o experimento que dá origem a A um certo número n de vezes e contar quantas vezes, k, o evento A ocorre nestas n repetições. Para n grande, $\frac{k}{n} \simeq P(A)$.

Podemos formalizar esta idéia intuitiva para o caso em que temos n provas de Bernoulli com probabilidade p de sucesso. O fato de, para n grande, $\frac{k}{n}$ estar suficientemente próximo de p é posto nos seguintes termos: para qualquer número $\varepsilon>0$,

$$P\left\{\left|\frac{k}{n}-p\right|\geq\varepsilon\right\} \leq \frac{p(1-p)}{n\cdot\varepsilon^2}, \qquad (5.5)$$

ou, de modo equivalente,

$$P\left\{\left|\frac{k}{n}-p\right|<\varepsilon\right\} \geq 1-\frac{p(1-p)}{n\,\varepsilon^2} \qquad (5.6)$$

O resultado (5.5) (ou (5.6)) é chamado *lei dos grandes números* para provas de Bernoulli. A prova deste resultado, embora simples, será omitida. (Ver Apêndice 2).

A relação (5.6) implica que, para todo número $\varepsilon>0$, a $P\left\{\left|\frac{k}{n}-p\right|<\varepsilon\right\} \longrightarrow 1$, quando $n \to \infty$.

Podemos formular a seguinte questão: quantas repetições do experimento, n, devemos realizar a fim de que k/n difira de p de menos de ε, com probabilidade maior ou igual a γ? Isto é, qual deve ser o valor de n para que tenhamos

$$P\left\{\left|\frac{k}{n}-p\right|<\varepsilon\right\} \geq \gamma ? \qquad (5.7)$$

Portanto,

$$1-\frac{p(1-p)}{n\varepsilon^2}=\gamma \Longrightarrow n=\frac{p(1-p)}{\delta\cdot\varepsilon^2},$$

onde $\delta = 1-\gamma$. Como não conhecemos p, usamos o fato que $p(1-p) \leq 1/4$, logo, basta tomar o número de provas n tal que

$$n \leq \frac{1}{4\delta\varepsilon^2}. \qquad (5.8)$$

Exemplo 5.3 - Quantos ensaios com probabilidade p de sucesso devemos realizar se queremos $\varepsilon = 0{,}01$ e $\gamma = 0{,}95$? Portanto, queremos n tal que

$$P\left\{\left|\frac{k}{n}-p\right|<0{,}01\right\} \geq 0{,}95.$$

$\delta = 1-\gamma = 0{,}05$ portanto basta tomar $n = \dfrac{1}{4\cdot(0{,}05)(0{,}01)^2} =$
$= 50.000$.

Para alguns valores de ε os respectivos valores de n estão dados na tabela abaixo, para $\gamma = 0{,}95$.

ε	0,01	0,05	0,1	0,2
n	50.000	2.000	500	125

Estes valores de n são desnecessariamente grandes. Utilizando a aproximação normal, sabemos que

$$P\{-2\sqrt{npq} \leq k-np \leq 2\sqrt{npq}\} \approx 0{,}95, \qquad (5.9)$$

lembrando que o número esperado de sucessos é np e o desvio padrão é $\sqrt{npq}$.

Portanto,

$$P\{|k/n-p|\leq 2\sqrt{(pq)/n}\} \approx 0{,}95$$

e chamando $\varepsilon = 2\sqrt{(pq)/n}$, vemos que $n = \frac{4pq}{\varepsilon^2}$. Novamente, como $pq \leq 1/4$, basta tomar $n = 1/\varepsilon^2$, no máximo. Obtemos a nova tabela abaixo.

ε	0,01	0,05	0,1	0,2
n	10.000	400	100	25

<u>Exemplo 5.4</u> - A proporção de fumantes de uma dada população é p, desconhecida. Queremos determinar p com um erro de, no máximo, 0,05. Qual deveria ser o tamanho da amostra n, a ser escolhida com reposição, se $\gamma = 0{,}95$? Devemos ter

$$P\left\{\left|\frac{k}{n} - p\right| \leq 0{,}05\right\} = P\{|S_n - np| \leq 0{,}05n\} \geq \gamma.$$

onde S_n = número de fumantes na amostra. Portanto, devemos ter $n = 400$.

5.3 - <u>ESTIMAÇÃO DE p POR INTERVALO</u>

Vamos considerar, agora, o caso em que queremos construir um *intervalo de confiança* para p, isto é, um intervalo $[\underline{p},\bar{p}]$ que contenha p com uma dada probabilidade, chamada *coeficiente de confiança*. Fixado este coeficiente, γ, devemos ter que a $P\{\underline{p} \le p \le \bar{p}\} = \gamma$.

Sabemos que $\frac{S_n - np}{\sqrt{npq}}$ tem distribuição aproximadamente normal padrão. Chamando Y este quociente, temos que

$$Y = \frac{S_n/n - p}{\sqrt{(pq)/n}} = \frac{\hat{p} - p}{\sqrt{(pq)/n}} \sim N(0,1) \qquad (5.10)$$

Assim, se $\gamma = 0,95$, temos, consultando a tabela 2

$$P\{-1,96 \le Y \le 1,96\} = 0,95,$$

ou seja,

$$P\{-1,96 \le \frac{\hat{p} - p}{\sqrt{(pq)/n}} \le 1,96\} = 0,95.$$

Portanto, com probabilidade 0,95 temos que

$$-1,96\sqrt{(pq)/n} \le \hat{p} - p \le 1,96\sqrt{(pq)/n},$$

do que segue

$$\hat{p} - 1,96\sqrt{(pq)/n} \le p \le \hat{p} + 1,96\sqrt{(pq)/n}. \qquad (5.11)$$

Novamente, como não conhecemos p, usamos o fato que $pq \le 1/4$, logo $\sqrt{(pq)/n} \le 1/\sqrt{4n}$, obtendo-se

$$\hat{p} - 1,96/\sqrt{4n} \le p \le \hat{p} + 1,96/\sqrt{4n}. \qquad (5.12)$$

Dizemos que $[\hat{p} - 1,96/\sqrt{4n}, \hat{p} + 1,96/\sqrt{4n}]$ é um intervalo de confiança para p com coeficiente de confiança de 95%.

Exemplo 5.5 - Numa pesquisa de mercado, $n = 400$ pessoas foram entrevistadas sobre um determinado produto e 60% destas pessoas preferiram a marca A. Aqui, $\hat{p} = 0,6$, e um intervalo de confiança para p, com $\gamma = 0,95$ será

$$0,6 \pm (1,96) \cdot \frac{1}{\sqrt{1.600}} = 0,6 \pm 0,049,$$

ou seja, $[\underline{p}, \bar{p}] = [0,551; 0,649]$.

Para um coeficiente de confiança qualquer γ, temos que (5.12) fica

$$\hat{p} - y_0/\sqrt{4n} \leq p \leq \hat{p} + y_0/\sqrt{4n}, \tag{5.13}$$

onde y_0 é obtido da tabela 2 tal que $P\{-y_0 \leq Y \leq y_0\} = \gamma$. Este intervalo é chamado pelos estatísticos de *conservativo*, pois se p não for igual a 1/2 e estiver próximo de 0 ou de 1, então (5.13) fornece um intervalo de amplitude desnecessariamente grande, pois substituimos pq pelo valor máximo 1/4. A memos que $\hat{p} = 1/2$, podemos usar o intervalo do exemplo a seguir.

Exemplo 5.6 - ("Enfoque otimista"). Suponha que em $n = 400$ provas obtemos $k = 80$ sucessos. Vamos obter um intervalo de confiança para p com $\gamma = 0,90$. Para tanto, usamos $\hat{p}\hat{q}$ como estimador de pq e então (5.11) fica (com y_0 no lugar de 1,96)

$$\hat{p} - y_0\sqrt{(\hat{p}\hat{q})/n} \leq p \leq \hat{p} + y_0\sqrt{(\hat{p}\hat{q})/n}. \tag{5.14}$$

No nosso caso, $\hat{p} = k/n = 80/400 = 0{,}2$ e $\hat{q} = 1 - k/n = 0{,}8$.

Portanto, o intervalo será

$$0{,}2 \pm (1{,}645)\sqrt{((0{,}2)(0{,}8))/400} = 0{,}2 \pm 0{,}033,$$

isto é,

$$[\underline{p},\bar{p}] = [0{,}167;\ 0{,}233].$$

Usando (5.13) obtemos

$$0{,}2 \pm (1{,}645)/40 = 0{,}2 \pm 0{,}041.$$

ou seja, $$[\underline{p},\bar{p}] = [0{,}159;\ 0{,}241].$$

Observe que o primeiro intervalo tem amplitude menor que o segundo.

Outra observação importante é a seguinte. Usando (5.13), para um γ fixo, os intervalos que podemos obter em sucessivas amostras terão todos a *mesma* amplitude, dada por $2y_0/\sqrt{4n}$. Por outro lado, usando-se (5.14), a amplitude do intervalo será $2y_0\sqrt{(\hat{p}\hat{q})/n}$, que é *variável* de amostra para amostra, pois $\hat{p}$ (e conseqüentemente $\hat{q}$) variará de amostra para amostra.

Em qualquer caso, a interpretação do que seja um intervalo de confiança para p com coeficiente de confiança γ é a seguinte: construindo 100 intervalos, correspondentes a 100 amostras de tamanho n, $100\gamma\%$ deles conterão o valor p.

Graficamente, teríamos a situação da figura 5.3, supondo intervalos dados por (5.13), isto é, de amplitudes constantes.

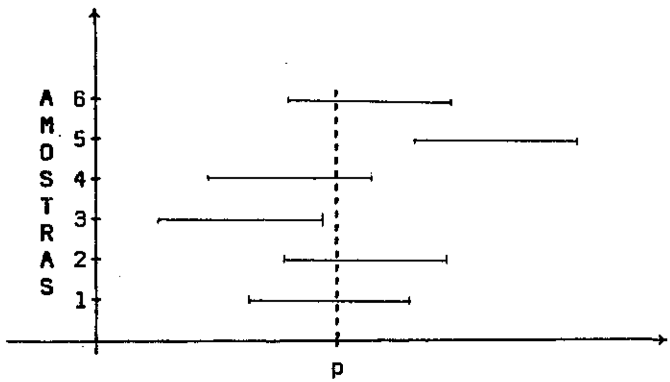


Figura 5.3

PROBLEMAS PARA O CAPÍTULO 5

5.1 - Obtenha a distribuição de $\hat{p} = S_n/n$, quando $p = 0,2$ e $n = 5$. Obtenha $E(\hat{p})$ e $Var(\hat{p})$.

5.2 - Deseja-se usar $\hat{p}$ como estimador de p, com probabilidade de 0,975, ou mais, que $\hat{p}$ difira 0,05 de p. Quão grande deve ser n?

5.3 - Encontre o máximo para $Var(\hat{p})$ quando n = 10, 25, 100 e 400.

5.4 - Encontre os intervalos de confiança para p se $k/n=0,3$ com coeficiente de confiança 95%. Utilize os dois enfoques, conservativo e otimista.

5.5 - Utilizando a relação (5.13) do texto, encontre um intervalo de confiança para p, sabendo-se que $n = 100$, $\hat{p} = 0,6$ e que a amplitude do intervalo deve ser igual a $d = 0,096$.

5.6 - Dê uma interpretação gráfica análoga à da figura 5.3, para intervalos dados por (5.14).

5.7 - Uma amostra casual de 625 donas de casa revela que 70 por cento delas preferem a marca X de detergente. Construir um intervalo de confiança para p, a proporção da população de donas de casa que preferem a marca X, com coeficiente de confiança de 95%.

5.8 - Suponha que X seja uma v.a. com média μ e variância σ^2, finitas. Se c é um número positivo, então

$$P(|X-\mu|>c\sigma) \leq 1/c^2. \qquad (*)$$

Utilizando esta desigualdade, prove a Lei dos Grandes Números (5.5). A desigualdade (*) é chamada *Desigualdade de Chebyshev*.

5.9 - Suponha que X tenha distribuição

x	0	1	2	3
p	1/8	1/4	1/2	1/8

Obtenha $P(|X-\mu|>3\sigma)$ e compare com o limite superior dada pela desigualdade (*) de (5.8).

- o -

6 - TESTES DE HIPÓTESES: PRIMEIRAS IDÉIAS

Como no capítulo anterior, vamos considerar aqui somente o caso de testes de hipóteses a respeito do parâmetro p de uma distribuição binomial. Em particular, veremos a conexão entre intervalos de confiança e testes de hipóteses.

6.1 - HIPÓTESE ESTATÍSTICA

Consideremos uma *população* cujos elementos podem ser classificados segundo dois atributos somente, que chamaremos "sucesso" e "fracasso". Por exemplo, podemos ter N objetos, dos quais alguns são defeituosos e os restantes não defeituosos. Se designarmos por p a proporção de sucessos na população, então o objetivo é fazer algum tipo de inferência sobre p. Vimos, no capítulo 5, como estimar p. Neste capítulo, estamos interessados em testar alguma hipótese sobre p. Entenderemos por *hipótese estatística* a qualquer afirmação que se faça sobre o parâmetro desconhecido p. Baseados em uma *amostra* da população vamos estabelecer uma *regra de decisão*, segundo a qual *rejeitaremos* ou *aceitaremos* a hipótese proposta. Uma tal regra de decisão é chamada *teste*.

Acima usamos vagamente os termos população e amostra, se bem que no capítulo 1 (ver exercícios) introduzimos já algumas noções sobre amostras.

Entenderemos por população a qualquer conjunto de observações, enquanto que uma amostra é qualquer subconjunto próprio da população. Por exemplo:

1) todos os possíveis lançamentos de uma moeda constituem uma população, enquanto que cinco lançamentos desta moeda constituem uma amostra;
2) as alturas de todos os estudantes de uma escola constituem uma população; as alturas dos estudantes de uma particular classe da escola constituem uma amostra;
3) os artigos produzidos por uma certa máquina em um dado dia constituem uma população; se escolhermos dez desses artigos para inspeção, teremos uma amostra de tamanho dez.

Somente consideraremos *amostras aleatórias*, tal como foi definida no capítulo 1 (problema 1.13). Note que no exemplo 2 acima não temos uma amostra aleatória.

Escolhida uma amostra de tamanho n, definimos a v. a. X = número de sucessos na amostra e consideramos

$$\hat{p} = \frac{X}{n}, \qquad (6.1)$$

que será a proporção de sucessos na amostra.

Basicamente, nossa regra de decisão utilizará $\hat{p}$ para decidir sobre a rejeição ou não rejeição da hipótese estipulada para p.

6.2 - TESTE DE UMA HIPÓTESE PARA p

Exemplo 6.1 - Um professor aplica um teste envolvendo 10 questões do tipo certo-errado. Ele quer testar a hipótese "o estudante está adivinhando."

Vamos designar por p a probabilidade do estudante responder corretamente a uma questão. Então, a hipótese a ser testada pode ser expressa na forma $p = \frac{1}{2}$. É comum designar-se a hipótese formulada por H. Então queremos por à prova

$$H: p = \frac{1}{2}. \qquad (6.2)$$

O teste de H será baseado no "número de sucessos nas n repetições independentes do experimento". Aqui, temos $n = 10$ repetições do experimento e independência significa que acertar uma questão não tem influência na resposta dada a uma outra questão qualquer. O número de sucessos será o número de acertos. É claro que, se H for verdadeira, o núme ro de sucessos deverá estar próximo de $np = 10 \times \frac{1}{2} = 5$.

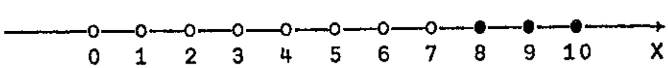
Suponha que o professor adote a seguinte *regra de decisão*:

"Se oito ou mais resposta estão corretas, o estudante não está adivinhando, enquanto que se menos do que oito questões estão corretas, o estudante está adivinhando."

Vamos indicar por X o número de respostas certas e por S o espaço amostral dos valores de X. Neste caso,

$$S = \{0,\ 1,\ 2,\ 3,\ \ldots,\ 9,\ 10\}.$$

A hipótese H será *rejeitada* se observamos os valores $X = 8$, $X = 9$ ou $X = 10$. Estes valores constituem a chamada *região crítica* ou *região de rejeição* do teste, que será indicada por S_0: $S_0 = \{8,\ 9,\ 10\}$. A região de *aceitação* (ou de *não rejeição*) do teste será S_1: $S - S_1 = \{0,\ 1,\ 2,\ \ldots,\ 6,\ 7\}$. Um gráfico ilustrativo é o que segue, onde indicamos por pontos cheios os valores de X que pertencem à região crítica.



O professor sabe, no entanto, que é possível que um estudante esteja adivinhando e ainda assim ele acerta 8 ou mais questões. Isto é, H é verdadeira, mas será rejeitada. O professor quer, certamente, que a probabilidade deste evento seja pequena. É fácil calculá-la:

$$P(X=8 \text{ ou } X=9 \text{ ou } X=10 \mid p=\tfrac{1}{2}) =$$

$$= \binom{10}{8}\left(\frac{1}{2}\right)^{10} + \binom{10}{9}\left(\frac{1}{2}\right)^{10} + \binom{10}{10}\left(\frac{1}{2}\right)^{10} =$$

$$= \frac{56}{1.024} = \frac{7}{128} \simeq 0{,}054.$$

O significado desta probabilidade é o seguinte: se o teste pudesse ser aplicado 128 vezes, o professor esperaria rejeitar H ("o aluno está advinhando"), 7 vezes. A probabilidade $\frac{7}{128}$ é chamada *nível de significância do teste* e será indicada por α. Temos, acima, um teste do tipo *unilateral* (à direita, no caso) porque só valores "de um lado do espaço amostral" foram usados como pontos críticos.

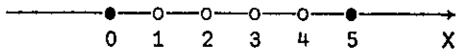
Podemos ter testes *bi-laterais*, como ilustramos a seguir.

Exemplo 6.2 - Dois quadros de futebol estão aguardando o juiz lançar uma moeda para ver quem dará a saída. Um dos capitães dá ao juiz uma moeda. Como o árbitro tem conhecimentos de testes de hipóteses e quer saber se a moeda é honesta, ele vai lançá-la 5 vezes; ele decide que a moeda é considerada viciada se sair 5 caras ou 5 coroas. Nos outros casos a moeda será considerada "honesta".

Se $p = P(\text{cara})$, queremos testar

$$H: p = 1/2.$$

Aqui, $S = \{0,1,2,3,4,5\}$, $S_0 = \{0,5\}$ e $S_1 = \{1,2,3,4\}$, sendo, novamente, X = número de caras nos 5 lançamentos.



O nível de significância do teste é

$$\alpha = P(X=0 \text{ ou } X=5 \mid p=\tfrac{1}{2}) = \left(\frac{1}{2}\right)^5 + \left(\frac{1}{2}\right)^5 = \frac{1}{16}.$$

Exemplo 6.3 - Um praticante de tiro ao alvo vai comprar um lote muito grande de munição e o vendedor garante que a porcentagem de projéteis em bom estado é 90%. No entanto, o comprador decide fazer uma experiência para testar a veracidade da afirmação do vendedor. Ele escolheu 10 projéteis e vai verificar quantos são bons. Ele decide *não* comprar o lote se X = número de sucessos na amostra (número de bons) é muito pequeno. A hipótese a ser testada é

$$H: p = 0,9,$$

onde p = proporção de bons projéteis no lote. Suponha que ele decide manter α abaixo de 2,5%, isto é, $\alpha < 0,025$. Então, para obter os valores críticos (que levarão à rejeição de H) ele calcula algumas probabilidades:

Dado que $p = 0,9$, temos (ver tabelas):

$P(X=0) = P(X=1) = P(X=2) = P(X=3) = P(X=4) \approx 0$

$P(X=5) = 0,001$

$P(X=6) = 0,011$

$P(X=7) = 0,057.$

Portanto,

$$P(X=0 \text{ ou } X=1 \text{ ou } \dots \text{ ou } X=5 \mid p=0{,}9) = 0{,}001,$$

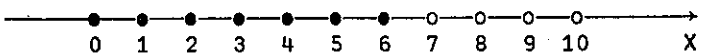
$$P(X=0 \text{ ou } X=1 \text{ ou } \dots \text{ ou } X=6 \mid p=0{,}9) = 0{,}012,$$

$$P(X=0 \text{ ou } X=1 \text{ ou } \dots \text{ ou } X=7 \mid p=0{,}9) = 0{,}069.$$

Portanto, a região crítica escolhida é

$$S_0 = \{0,1,2,3,4,5,6\}$$

pois incluindo-se o ponto 7 obtemos $\alpha > 0{,}025$.



Se, por exemplo, o comprador encontrou 8 projéteis bons, ele não rejeitará H.

6.3 - HIPÓTESE ALTERNATIVA - UM OUTRO TIPO DE ERRO

Exemplo 6.4 - Retomemos o exemplo 6.1, no qual estamos testando a hipótese H: $p = \frac{1}{2}$. Um valor de p maior do que $\frac{1}{2}$ indicará que o estudante não está simplesmente adivinhando. Vamos fixar α aproximadamente igual a 0,05. A hipótese H acima é comumente chamada *hipótese nula*. Vimos na seção anterior que $P(X=8 \text{ ou } X=9 \text{ ou } X=10 \mid p=\frac{1}{2}) \simeq 0{,}054$, enquanto que $P(X=9 \text{ ou } X=10 \mid p=\frac{1}{2}) = \frac{1}{1.024} \simeq 0{,}01$, logo decidimos usar a região crítica $S_0 = \{8,9,10\}$. A probabilidade, α, de rejeitar uma hipótese verdadeira, é também chamada *probabilidade do erro de tipo I*.

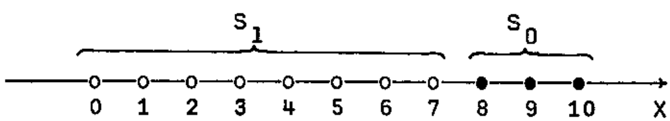
Suponha, agora, que o aluno acertou apenas 6 questões. Então, não há razão para rejeitar H e diríamos que o aluno está adivinhando. Mas, é possível que o aluno *não* esteja adivinhando (isto é, $p>\frac{1}{2}$) e no entanto acertou apenas 6 questões. Portanto, há um outro erro que está envolvido neste processo decisório: *aceitar uma hipótese H, sendo ela falsa*. Suponha que na realidade, $p=0,8$. Podemos pensar o problema em termos de testar

$$H: p = 0,5$$

contra a *hipótese alternativa*

$$K: p = 0,8 \qquad (6.3)$$

Vamos calcular a probabilidade, β, de aceitar H, quando K é verdadeira, para $\alpha = \frac{7}{128} = 0,054$, ou seja, a região crítica acima, $S_0 = \{8,9,10\}$.



Temos (ver tabela 1, $n = 10$, $p = 0,8$)

$$\beta = P(X=0 \text{ ou } X=1 \text{ ou } \ldots \text{ ou } X=7 \mid p=0,8) =$$

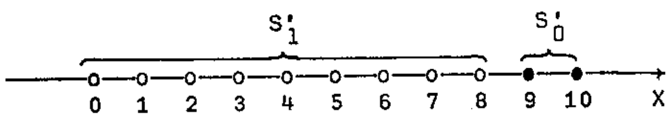
$$= 0 + 0 + 0 + 0,001 + 0,006 + 0,026 + 0,088 + 0,201 = 0,322.$$

Este erro, aceitar uma hipótese falsa, é denominado *erro do tipo II* e β é a probabilidade do erro de tipo II.

Podemos descrever os erros que cometemos e as decisões que podemos tomar através do quadro seguinte:

DECISÃO	p = 0,5	p = 0,8
ACEITAR H	Decisão correta	erro de tipo I, probabilidade α
ACEITAR K	erro de tipo II, probabilidade β	Decisão correta

Vamos ver o que acontece quando tentamos diminuir um dos erros, digamos α. Suponha que tomamos por região crítica o conjunto $S_0' = \{9,10\}$. Então, vimos que $\alpha' = 0{,}01 < 0{,}054$.



O valor correspondente da probabilidade do erro de tipo II será

$$\beta' = P(X=0 \text{ ou } X=1 \text{ ou} \ldots \text{ou } X=8 \mid p=0{,}8) = 0{,}322+0{,}302 = 0{,}624.$$

Portanto, diminuindo α, β aumenta. Se a região crítica é $S_0'' = \{7,8,9,10\}$, $\alpha'' = 0{,}17$ e $\beta = 0{,}121$ (calcule!). Podemos, pois, obter o quadro seguinte:

REGIÃO CRÍTICA	α	β
{7,8,9,10}	0,17	0,121
{8,9,10}	0,054	0,322
{9,10}	0,01	0,624

A situação que ocorre é que, para n fixado a priori, não é possível tomar α e β a nossa vontade; gostaríamos que α e β fossem pequenos, mas devido ao fato acima é costume fixar um valor para α; valores comumente usados para α são 0,01 e 0,05.

Para efeito de ilustração, consideremos o caso em que $\alpha = 0{,}054$ e $\beta = 0{,}322$ (figuras 6.1 e 6.2).

região hachurada = α

$n = 10$

$p = 1/2$

$np = 5 =$ média

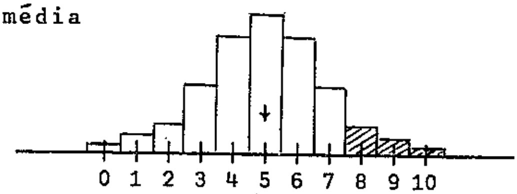


Fig. 6.1

região hachurada = β

$n = 10$

$p = 0{,}8$

$np =$ média $= 8$

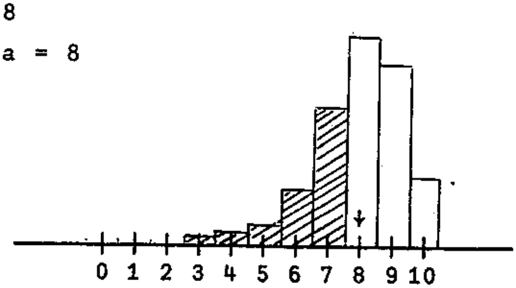


Fig. 6.2

Utilizando "aproximações contínuas" para os gráficos das figuras 6.1 e 6.2 obtemos a figura 6.3, onde estão destacados α e β.

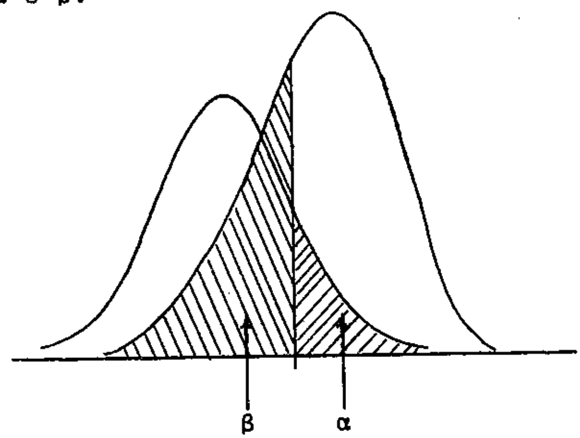


Figura 6.3

A justificativa de fixar α é dada pelo fato que escolhemos a hipótese nula que fornece o erro de primeira espécie mais grave. Para exemplificar, suponhamos que uma vacina contra uma doença vai ser testada em um grupo de pessoas, enquanto que um grupo de controle recebe apenas soro. Após algum tempo verificamos quais pessoas adquiriram a doença (afetados) e quais não adquiriram (não afetados). Obtemos uma tabela como aquela que segue.

	AFETADOS	NÃO AFETADOS
receberam vacina	n_{11}	n_{12}
receberam soro	n_{21}	n_{22}

Assim, n_{12} pessoas foram vacinadas e não ficaram doentes, enquanto que n_{21} pessoas receberam apenas soro e ficaram doentes, etc.

Suponha que queremos escolher uma das seguintes hipóteses como hipótese nula.

H_1: a vacina tem efeito positivo

H_2: a vacina é inóqua.

Se $H = H_1$, o erro de tipo I consiste em rejeitar H_1 sendo ela verdadeira, isto é, a vacina é eficiente, mas a consideramos inóqua.

Se $H = H_2$, o erro de tipo I consiste em **considerar** a vacina eficiente, quando ela é inóqua. Tomamos H_2 como hipótese nula, pois o erro de I decorrente nos parece o mais grave.

<u>Exemplo 6.5</u> - Um fabricante de dois produtos similares, A e B, tem motivos para acreditar que pelo menos 70% dos consumidores preferem A. Uma pesquisa de mercado revelou que, em uma amostra de 625 pessoas, 65% dos elementos preferiam A, e o restante, B.

Se indicarmos por p a proporção (desconhecida) dos consumidores da população que preferem A, podemos pensar no problema como um teste da forma

$$H: p \geq 0,7 \qquad (6.4)$$

contra a alternativa

$$K: p < 0,7 \qquad (6.5)$$

Se X designa o número de consumidores na amostra que preferem A, então rejeitaremos H para "valores pequenos" de X.

Em termos de $\hat{p} = \frac{X}{n}$, podemos dizer que rejeitaremos H se o valor observado de $\hat{p}$ for menor que um certo $\hat{p}_c$, tal que

$$P(\hat{p} < \hat{p}_c \mid H \text{ verdadeira}) = \alpha, \qquad (6.6)$$

α sendo o nível de significância do teste.

A probabilidade $P(\hat{p} < \hat{p}_c \mid H)$ envolve o conhecimento da *distribuição de $\hat{p}$*. No exemplo 5.2, do capítulo anterior, vimos a distribuição de $\hat{p}$ para o caso particular de $n = 3$ e $p = 1/2$. Aqui, não seria prático obter a distribuição de $\hat{p}$, quando H verdadeira, pois n é grande e p pode assumir qualquer valor maior ou igual a 0,7.

Felizmente, é possível obter um teste aproximado, utilizando-se uma aproximação para a distribuição de $\hat{p}$, quando n é grande.

Não teremos ocasião de tratar deste problema nestas notas, mas o exemplo serve para ilustrar a importância de obtermos o que chamamos a *distribuição amostral de um estimador*, no caso $\hat{p}$. Desta maneira podemos obter o valor crítico, $\hat{p}_c$, e conseqüentemente, a região crítica do teste. Veremos algumas noções sobre este assunto no próximo capítulo.

Os exemplos vistos até agora sugerem os seguintes tipos de testes que podemos encontrar:

$$\text{Exemplo 6.1: } \begin{cases} H: \ p = 1/2 \\ K: \ p > 1/2 \end{cases}$$

$$\text{Exemplo 6.2: } \begin{cases} H: \ p = 1/2 \\ K: \ p \neq 1/2 \end{cases}$$

$$\text{Exemplo 6.3: } \begin{cases} H: \ p = 0{,}9 \\ K: \ p < 0{,}9 \end{cases}$$

$$\text{Exemplo 6.5: } \begin{cases} H: \ p \geq 0{,}7 \\ K: \ p < 0{,}7 \end{cases}$$

De um modo bastante geral temos as 3 situações seguintes:

$$1) \begin{cases} H: \ p \leq p_0 \\ K: \ p > p_0 \end{cases} \tag{6.7}$$

$$2) \begin{cases} H: \ p \geq p_0 \\ K: \ p < p_0 \end{cases} \tag{6.8}$$

$$3) \begin{cases} H: \ p = p_0 \\ K: \ p \neq p_0 \end{cases} \tag{6.9}$$

6.4 - RELAÇÃO ENTRE INTERVALO DE CONFIANÇA E TESTE DE HIPÓTESES

Podemos testar hipóteses referentes ao parâmetro p de uma binomial utilizando determinados gráficos já prontos, para diversos valores de p, diversos valores de n e diversos níveis de significância. Vejamos como tais gráficos são construídos.

Para fixar idéias, seja $n = 10$ e suponha que queremos testar $H: p = p_0$ contra $K: p \neq p_0$, de modo que temos um teste bilateral. Vamos fixar os níveis das regiões críticas laterais, α_1 e α_2, menores ou iguais a 0,025 (2,5%). Vamos tomar, para valores de p_0, 0,1; 0,2; 0,3; ...; 0,9; 1.

O gráfico da figura 6.4 ilustra, para cada p_0, os valores críticos como pontos cheios e os restantes como círculos abertos.

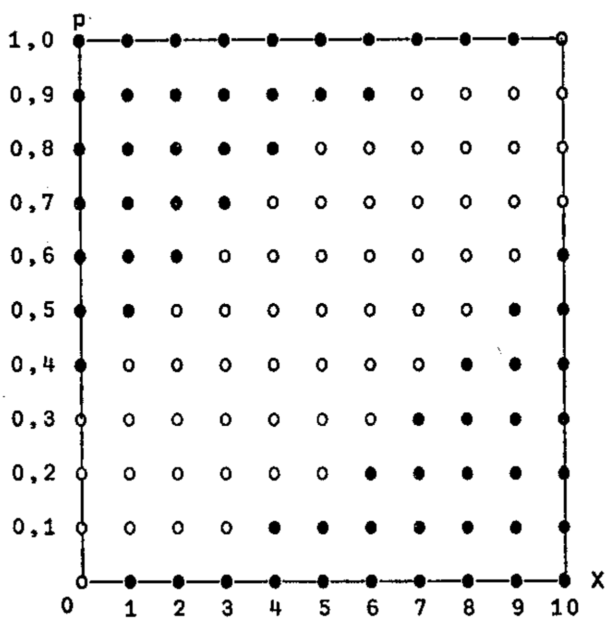


Figura 6.4

Por exemplo, para $p_0 = 0,4$,

$$P(X=0) = 0,006, \text{ menor do que } 0,025,$$

enquanto que

$$P(X=0 \text{ ou } X=1) = 0,006 + 0,0403 = 0,0409,$$

que é maior que 0,025. Portanto, a região crítica à esquerda é $\{0\}$.

Analogamente,

$$P(X=8 \text{ ou } X=9 \text{ ou } X=10) = 0,0106+0,0016+0,0001 = 0,0123,$$

menor do que 0,025, mas

$$P(X=7 \text{ ou } X=8 \text{ ou } X=9 \text{ ou } X=10) = 0,0425+0,0123 = 0,0548.$$

que é maior do que 0,025, logo a região crítica à direita é $\{8, 9, 10\}$.

De modo análogo obtemos as demais regiões críticas.

Se obtemos $X=2$ sucessos, basta olhar a coluna correspondente a $X=2$ e a linha correspondente a $p = 0,4$; encontramos um círculo aberto, portanto aceitamos H.

Mas se queremos testar $H: p = 0,25$ versus $K: p = 0,25$, não podemos usar este quadro.

Teríamos que acrescentar novos valores de p; mas podemos ter uma idéia do que vai acontecer; para um dado valor de X, olhando a coluna correspondente, de baixo para cima, notaremos a seqüência: pontos cheios — cículos — pontos cheios, de modo que podemos saber, aproximadamente, para quais valores de p rejeitaremos H e para quais valores de p aceitaremos H, para aquele dado X.

Para n = 10 elementos amostrais obtemos o gráfico da figura 6.5, com nível de significância menor que 5%, substituindo X por $X/n = \hat{p}$.

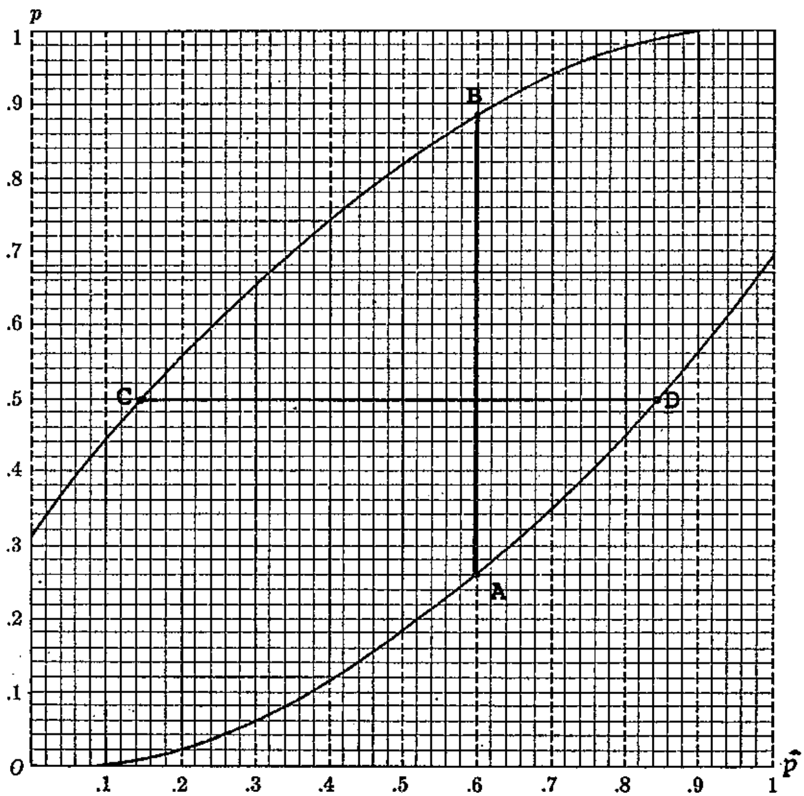


Figura 6.5

Por exemplo, se X/n = 0,6, a reta vertical passando por 0,6 intercepta as duas curvas nos pontos A e B; os pontos do segmento AB indicam aceitação da hipótese $p = p_0$, enquanto que pontos abaixo de A e acima de B indicam rejeição de $p = p_0$. Os pontos limites são 0,26 e 0,88, de modo que rejeitamos $H: p = p_0$ contra $K: p \neq p_0$ para $p < 0{,}26$ e $p > 0{,}88$ e aceitamos H para $0{,}26 \leq p \leq 0{,}88$.

Por outro lado, se queremos testar $H: p = 0{,}5$ contra $K: p \neq 0{,}5$, considere a reta horizontal passando por 0,5; esta intercepta as duas curvas nos pontos C e D; estes correspondem os valores 0 e 0,1, à esquerda, e 0,9 e 1, à direita, de modo que rejeitamos H para $X/n = \hat{p}$ pertencente a $\{0;\ 0{,}1;\ 0{,}9;\ 1\}$.

A região entre as curvas é chamada *região de confiança* e as curvas são as *curvas limites* da região.

No Apêndice 1 apresentamos regiões de confiança para alguns valores de n e para $\alpha < 5\%$ e $\alpha < 10\%$.

Voltemos ao caso $n = 10$ acima e $\alpha < 0{,}05$. Suponha que obtemos $\hat{p} = 0{,}6$. Então, o intervalo encontrado acima, $[0{,}26;\ 0{,}88]$ será um *intervalo de confiança* para p, com *coeficiente confiança* de pelo menos 0,95. Este intervalo corresponde à região de aceitação da hipótese, ao nível de 0,05. Isto é, se obtemos $\hat{p} = 0{,}6$, nós não rejeitaremos a hipótese $H: p = p_0$, para $0{,}26 \leq p \leq 0{,}88$.

De modo geral, a região de aceitação de um teste do tipo acima, de nível α, corresponde a um intervalo de confiança para p, com coeficiente de confiança $\gamma = 1 - \alpha$.

6.5 - UMA APLICAÇÃO: TESTE DO SINAL

Exemplo 6.6 - Dois fertilizantes A e B são comparados escolhendo-se 10 pares de canteiros de testes semelhantes e aplicando-se A a um canteiro de cada par e B ao outro canteiro. As produções obtidas estão abaixo. A questão de interesse é: podemos concluir que a produção decorrente do uso de A é maior do que a produção decorrente do uso de B? (estas produções serão indicadas também por A e B).

A	20	19	23	17	18	20	24	19	21	19
B	18	16	21	20	17	18	20	18	23	20
SINAL DE A-B	+	+	+	-	+	+	+	+	-	-

Vemos que a diferença A-B é positiva para 7 dentre 10 pares de canteiros; ou seja, em 7 pares de canteiros, a produção decorrente da aplicação de A é maior do que a produção decorrente do uso de B.

Se não há diferença entre as produções de A e B podemos pensar no problema como um experimento binomial com $n = 10$ e $p = 1/2$, onde $p = P(A-B > 0)$ (justifique a independência necessária ao modelo binomial).

Se X indica o número de sinais "+", temos que

$$\hat{p} = \frac{X}{n} = \frac{7}{10} = 0{,}7.$$

Podemos, pois, testar

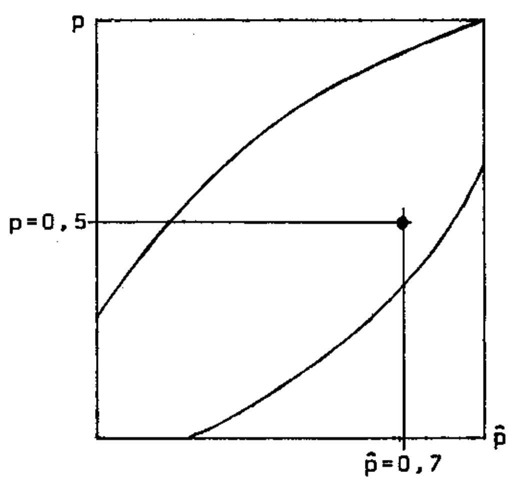
$$H: p = 0,5$$

contra

$$K: p > 0,5.$$

Estamos utilizando uma região crítica à direita porque temos alguma razão para esperar mais sinais "+" do que sinais "-".

Suponha $\alpha = 0,025$ para podermos utilizar o gráfico do Apêndice 1. A reta horizontal por $p = 0,5$ intercepta a reta vertical por $\hat{p} = 0,7$ dentro da região de confiança, portanto não rejeitamos H.



Portanto, podemos concluir que rejeitamos a idéia de uma diferença substancial na produção a favor do fertilizante A. Se não tivéssemos, baseados nos dados, razão para esperar maior número de "+" do que "-" poderíamos utilizar um teste bi-lateral, isto é, considerar

$$H: p = 0{,}5$$

contra

$$K: p \neq 0{,}5.$$

Ainda assim não rejeitaríamos H ao nível de 5%.(Verifique).

Observe que, para o caso de região crítica à direita ($\alpha<2{,}5\%$) os valores críticos são X=9 ou X=10, ao passo que para região crítica bi-lateral ($\alpha < 5\%$) os valores críticos são X=0 ou X=1 ou X=9 ou X=10.

PROBLEMAS PARA O CAPÍTULO 6

6.1 - Temos que tomar uma decisão sobre se uma moeda é viciada ou não baseados em 6 lançamentos da moeda. A moeda será considerada viciada se o resultado for:

a) 0 ou 6 caras;
b) 0 caras;
c) 6 caras.

Construa os testes correspondentes, especificando: as hipóteses H e K, a região crítica, a região de aceitação, o gráfico da região crítica e o valor de α.

6.2 - Para $\alpha = 0{,}02$, encontre o valor de β no problema anterior.

6.3 - Utilizando a figura 6,4 resolva:

a) encontrar os valores para os quais $H: p = 0{,}8$ é rejeitada, usando um teste bi-lateral;
b) mesma questão de (a), usando um teste uni-lateral;

c) a amostra de 10 elementos contém 4 sucessos; quais hipóteses não são rejeitadas, num teste bi-lateral?

d) em uma amostra de tamanho 10, X=6; quais hipóteses são rejeitadas? Use: teste uni-lateral à esquerda; teste uni-lateral à direita; teste bi-lateral.

6.4 - Em uma população temos $p = 0,45$. Uma amostra com $X \leq 1$ ou $X \geq 9$ supostamente contradiz $p = 0,45$. Qual a probabilidade de rejeitar $p = 0,45$?

6.5 - Queremos testar H: $p = 1$. Se $n = 10$ e $X = 9$, a hipótese é rejeitada? Se H: $p = 0$, $n = 10$ e $X=0$, a hipótese é rejeitada?

6.6 - Utilizando a figura 6.5, $n = 10$, resolva:

a) qual a região crítica de um teste bi-lateral para H: $p = 0,42$? e para H: $p = 0,76$?

b) para um teste uni-lateral à direita, quais valores de p são rejeitados se $\hat{p} = 0,3$? e se $\hat{p} = 0,8$?

6.7 - Um estudo é feito para determinar a proporção de famílias em uma comunidade que têm telefone. Uma amostra de 10 famílias é escolhida ao acaso e 8 têm telefone. Se $\alpha \leq 0,05$, qual a conclusão pode ser tirada sobre a população?

6.8 - Um médico examina aumentos na pressão sanguínea usan- uma certa droga. Ele mede a pressão de 20 homens an-

HOMENS	1	2	3	4	5	6	7	8	9	10
ANTES	121	130	127	125	140	121	132	135	124	136
DEPOIS	123	127	132	132	135	122	138	176	127	138
HOMENS	11	12	13	14	15	16	17	18	19	20
ANTES	127	118	131	120	128	123	128	133	141	124
DEPOIS	135	120	129	123	127	128	122	134	142	126

tes e depois de administrar a droga. A tabela na página anterior fornece os resultados. Existe razão suficiente para concluir que a pressão sanguínea em homens aumenta com o uso da droga? $\alpha \leq 5\%$.

6.9 - Encontre o número de sucessos, em uma amostra de 10, para o qual $p = 0,2$ e $p = 0,7$ são ambos rejeitados usando um teste bi-lateral. Tome $\alpha \leq 5\%$ e $\alpha \leq 10\%$.

6.10- Suponha que estamos testando $H: p = 0,5$ contra a alternativa $K: p \neq 0,5$. Suponha que para uma amostra de tamanho $n = 10$, decidimos pela região crítica

$$S_0 = \{0,1,2,8,9,10\}.$$

Determine o nível de significância, α. Denomina-se *poder* do teste à probabilidade de rejeitar a hipótese H quando ela é falsa. Assim, para $p = 0,6$, o poder é

$$P\{X=0 \text{ ou } X=1 \text{ ou } X=2 \text{ ou } X=8 \text{ ou } X=9 \text{ ou } X=10 \mid p=0,6\}.$$

Calcule o poder para $p = 0,2$, $p = 0,4$, $p = 0,6$, $p = 0,8$. Faça um gráfico da função obtida (chamada *função poder* do teste). Qual o poder do teste para $p = 0,5$?

6.11- Para testar se um dado é honesto, ele é lançado 1.000 vezes e é considerado viciado se números pares ocorrerem mais do que 510 ou menos do que 490 vezes. Qual o nível de significância do teste? Se a probabilidade de ocorrer número par é 0,51, qual o poder do teste?

- o -

7. DISTRIBUIÇÕES AMOSTRAIS

Frisamos, no capítulo anterior, a importância de se obter aquilo que denominamos distribuição amostral de um estimador, pois dela necessitamos para efetuar testes de hi hipóteses. Neste capítulo iremos introduzir algumas noções sobre este assunto. Em particular, estudaremos com algum pormenor a distribuição amostral da média amostral.

7.1 - ESTATÍSTICAS

Já temos idéia do que seja uma amostra aleatória. Suponha que X represente uma característica de interesse de uma dada população. Por exemplo, podemos estar interessados na distribuição das medidas das alturas dos indivíduos de uma dada cidade. Dizemos que a distribuição de X é a distribuição da população. Normalmente não temos conhecimento desta distribuição. Para termos uma idéia da forma desta distribuição, bem como estimarmos alguns parâmetros da mesma, extraímos uma amostra aleatória de X.

Suponha que os valores da variável X nos elementos da população sejam X_1, X_2, ..., X_N. Escolhendo-se uma amostra de tamanho n, podemos obter valores repetidos; suponha que obtemos n_1 valores iguais a X_1, n_2 valores iguais a X_2, ..., n_N valores iguais a X_N, onde $n_1+n_2+...+n_N = n$.

Obtemos a chamada *distribuição de freqüências* dos valores observados:

FREQÜÊNCIA	n_1	n_2	...	n_i	...	n_N
VALOR DE X	X_1	X_2	...	X_i	...	X_N

Definimos,então, a *média amostral*

$$\bar{X} = \frac{\sum_{i=1}^{N} n_i X_i}{\sum_{i=1}^{N} n_i} = \frac{1}{n} \sum_{i=1}^{N} n_i x_i. \qquad (7.1)$$

Observe a correspondência entre a distribuição de freqüências acima e a distribuição de probabilidades da v. a. X,

X	X_1	X_2	...	X_i	...	X_N
PROB.	p_1	p_2	...	p_i	...	p_N

bem como a analogia de $\bar{X}$ com o valor esperado de X,

$$E(X) = \sum_{i=1}^{N} X_i p_i \qquad (7.2)$$

Podemos, também, definir a *variância amostral*

$$S^2 = \frac{\sum_{i=1}^{N} (X_i-\bar{X})^2 \cdot n_i}{n-1}, \qquad (7.3)$$

que corresponde à variância de X,

$$Var(X) = \sum_{i=1}^{N} (X_i - E(X))^2 \cdot p_i. \qquad (7.4)$$

Dizemos que $\bar{X}$ e S^2 são *estatísticas*, isto é, funções dos valores amostrais. Os valores correspondentes da população são chamados *parâmetros*, como E(X) e Var(X).

Vamos indicar uma amostra de tamanho n por $(X_1, \ldots, X_n)$; como $X_1, X_2, \ldots, X_n$ são observações da mesma variável X, dizemos que $X_1, \ldots, X_n$ têm a mesma distribuição que X; e são, além disso, independentes, devido ao fato de termos uma amostra aleatória (com reposição). Uma estatística, T, será uma função de $X_1, \ldots, X_n$: $T = f(X_1, \ldots, X_n)$. Considere $\bar{X}$, por exemplo. Para diversas amostras escolhidas obteremos diversos valores de $\bar{X}$; esta será também uma variável aleatória. O mesmo acontece com S^2. Outras estatísticas que podem ser consideradas são:

$W = X_n - X_1$: amplitude total da amostra

$Y_1 = \max(X_1, \ldots, X_n)$: o maior valor da amostra

$Y_n = \min(X_1, \ldots, X_n)$: o menor valor da amostra.

Y_1 e Y_n são duas *estatísticas de ordem*; suponha que ordenemos a amostra obtida:

$$\begin{aligned} Y_1 &= \min(X_1, \ldots, X_n) \\ Y_2 &= \text{segunda maior em ordem crescente} \\ &\vdots \\ Y_n &= \max(X_1, \ldots, X_n). \end{aligned}$$

Então, $Y_1 \leq Y_2 \leq \ldots \leq Y_n$ são as estatísticas de ordem da amostra $(X_1,\ldots,X_n)$. Uma particular e importante estatística de ordem é a *mediana*, denotada por $Y_{1/2}$. Ordenados os valores amostrais, $Y_{1/2}$ é o valor central se n é ímpar e é a média aritmética dos dois valores centrais se n é par.

<u>Exemplo 7.1</u> - Os valores obtidos em uma amostra $(X_1,\ldots,X_5)$ foram:

$x_1=7$, $x_2=5$, $x_3=8$, $x_4=12$, $x_5=5$ (lembre-se que os valores de uma v.a. são denotados por letras minúsculas).

Então:

$$\bar{x} = \frac{1}{5}(7 + 5 + 8 + 12 + 5) = 7{,}4,$$

$$s^2 = \frac{1}{4}(0{,}4^2+2{,}4^2+0{,}6^2+4{,}6^2+2{,}4^2) = 8{,}3.$$

As estatísticas de ordem são:

$$Y_1 = Y_2 = 5;\ Y_3 = Y_{1/2} = 7;\ Y_4 = 8;\ Y_5 = 12.$$

Se $Y_1=5$, $Y_2=8$, $Y_3=12$, $Y_4=13$; $Y_5=15$ e $Y_6=16$, a mediana seria

$$Y_{1/2} = \frac{12+13}{2} = 12{,}5.$$

Como toda estatística é uma v.a. tem sentido em considerar a *distribuição desta estatística*, bem como considerar a média ou a variância dela.

Consideremos, em especial, $\bar{X}$. Suponha que a "população", isto é, a v.a. X, tem média $E(X) = \mu$ e variância

$Var(X) = \sigma^2$. Quando consideramos $E(\bar{X})$ e $Var(\bar{X})$, será que há alguma relação entre estes valores e μ, σ^2? Isto é o que veremos na próxima secção.

7.2 - DISTRIBUIÇÕES AMOSTRAIS

Denominaremos *distribuição amostral* à distribuição de uma estatística determinada a partir de uma amostra. Para efeito de ilustração, consideremos uma população de tamanho $N = 3$ e a v.a. X tendo os valores 1, 2 e 3 sobre os elementos desta população.

A média da população é

$$\mu = E(X) = \frac{1+2+3}{3} = 2$$

e a variância da população é

$$\sigma^2 = E(X-\mu)^2 = \frac{\sum(X_i-\mu)^2}{3} = \frac{2}{3} = 0,667.$$

X	$X - \mu$	$(X-\mu)^2$
1	-1	1
2	0	0
3	1	1
6	0	2

$\mu = 2$

$\sigma^2 = 2/3 = 0,667$

$\sigma = 0,818.$

Retiramos todas as possíveis amostras (com reposição) de tamanho $n = 2$. O número de amostras será $N^n = 3^2 = 9$ e cada amostra tem a mesma probabilidade, $\frac{1}{9}$, de ser selecionada. Obtemos o quadro seguinte.

AMOSTRA	VALORES AMOSTRAIS	PROBABILIDADE	$\bar{X}$	S^2	S
1	(1,1)	1/9	1,0	0	0
2	(1,2)	1/9	1,5	0,5	0,71
3	(1,3)	1/9	2,0	2,0	1,41
4	(2,1)	1/9	1,5	0,5	0,71
5	(2,2)	1/9	2,0	0	0
6	(2,3)	1/9	2,5	0,5	0,71
7	(3,1)	1/9	2,0	2,0	1,41
8	(3,2)	1/9	2,5	0,5	0,71
9	(3,3)	1/9	3,0	0	0
			18,0	6,0	5,66

Por exemplo, para a amostra (1,3), $\bar{x} = \frac{1+3}{2} = 2$, enquanto $s^2 = \frac{(1-2)^2+(3-2)^2}{1} = \frac{1+1}{1} = 2$ $(n-1 = 1)$.

A *distribuição amostral de* $\bar{X}$ será, então:

$\bar{x}$	1,0	1,5	2,0	2,5	3,0
p	1/9	2/9	3/9	2/9	1/9

Portanto,

$$E(\bar{X}) = 1 \times \frac{1}{9} + (1,5) \times \frac{2}{9} + 2 \times \frac{3}{9} + (2,5) \times \frac{2}{9} + 3 \times \frac{1}{9} = \frac{18}{9} = 2 = \mu,$$

enquanto que

$$E(\bar{X}^2) = 1 \times \frac{1}{9} + (2,25) \times \frac{2}{9} + 4 \times \frac{3}{9} + (6,25) \times \frac{2}{9} + 9 \times \frac{1}{9} = 4,333,$$

do que segue

$$Var(\bar{X}) = 4,333 - 2^2 = 0,333.$$

É fácil ver que $Var(\bar{X}) = \frac{Var(X)}{n} = \frac{0,667}{2}$.

Então:

$$E(\bar{X}) = 2 = \mu$$

$$Var(\bar{X}) = 0,333 = \frac{\sigma^2}{n} \qquad (7.5)$$

Da mesma forma obtemos a distribuição amostral de S^2,

s^2	0	0,5	2,0
p	3/9	4/9	2/9

e a distribuição amostral de S (o desvio padrão amostral),

s	0	0,71	1,41
p	3/9	4/9	2/9

É fácil ver que

$$E(S^2) = \frac{6,0}{9} = \frac{2}{3} = 0,667 = \sigma^2,$$

mas

$$E(S) = \frac{5,66}{9} = 0,629 \neq \sigma = 0,818.$$

Vamos comparar as distribuições de X e de $\bar{X}$; observe que ambas possuem a mesma média, 2, mas as variâncias são diferentes: $Var(X) = \sigma^2 = 0{,}667$, enquanto que $Var(\bar{X}) = \frac{\sigma^2}{n} =$ $= 0{,}333$ (figura 7.1).

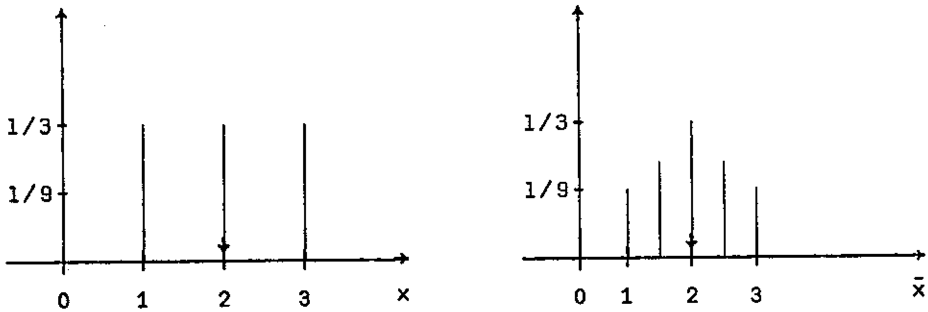


Figura 7.1

Quanto maior for o tamanho da amostra, n, menor será a variância de $\bar{X}$, devido ao fato que $Var(\bar{X}) = \frac{Var(X)}{n}$. Isto indica que a distribuição de $\bar{X}$ será mais concentrada ao redor de sua média, $E(\bar{X})$. A figura 7.2 ilustra este fato, para o caso em que X tem distribuição normal $N(\mu, \sigma^2)$.

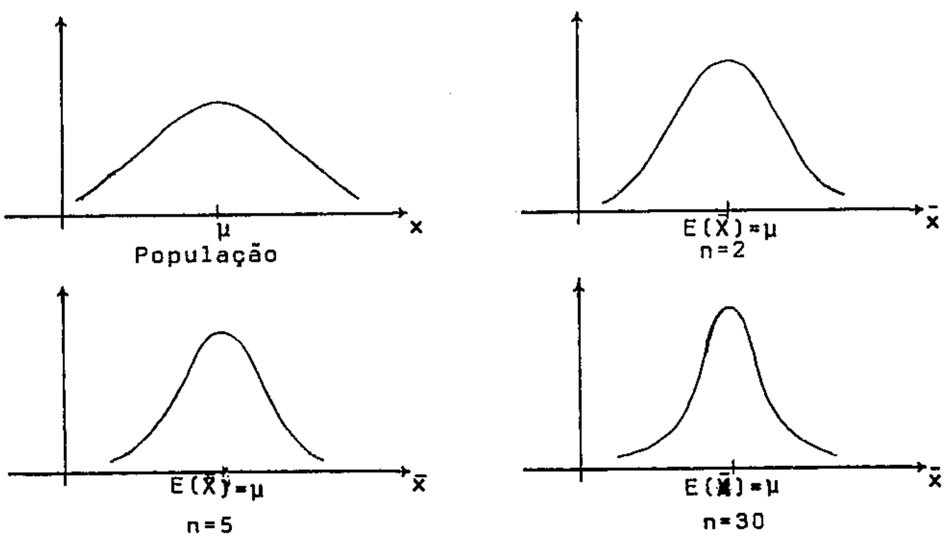


Figura 7.2

Outra observação que podemos fazer do que foi exposto é que $E(\bar{X}) = \mu$ e $E(S^2) = \sigma^2$. Estas relações são válidas em geral e não só para o exemplo visto. Mas $E(S) \neq \sigma$. Dizemos que $\bar{X}$ e S^2 são estimadores *não viciados* de μ e σ^2, respectivamente, ao passo que S é um estimador *viciado* de σ. Voltaremos a este assunto no capítulo seguinte.

Para *demonstrar* que $E(\bar{X}) = \mu$ e $Var(\bar{X}) = \dfrac{\sigma^2}{n}$ é relativamente fácil, bastando utilizar as propriedades da média e da variância; a demonstração de que $E(S^2) = \sigma^2$ é um pouco mais trabalhosa. Ver exercícios.

Para as figuras 7.2, no caso em que X é normal, percebe-se claramente que a distribuição de $\bar{X}$ também é normal; este é um resultado importante e decorre da seguinte *propriedade reprodutiva da normal*: se X_1 é uma v.a. $N(\mu_1, \sigma_1^2)$ e X_2 é uma v.a. $N(\mu_2, \sigma_2^2)$, com X_1 e X_2 *independentes*, então $X_1 + X_2$ será uma v.a. $N(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)$. Esta propriedade estende-se a um número finito de v.a. normais independentes, bem como para uma combinação linear de v.a. normais independentes. Em particular, como $\bar{X} = \frac{1}{n} X_1 + \ldots + \frac{1}{n} X_n$, se $X_1, X_2, \ldots, X_n$ são normais, segue-se que $\bar{X}$ será normal. Um fato importante é o seguinte. Mesmo que a população não seja normal, à medida que n cresce, a distribuição de $\bar{X}$ vai-se aproximando de uma normal.

No exemplo dado, a população é uniforme sobre os inteiros 1, 2 e 3 e mesmo para $n = 2$ já vemos que a distribuição de $\bar{X}$ tem uma forma que lembra a normal (figura 7.3).

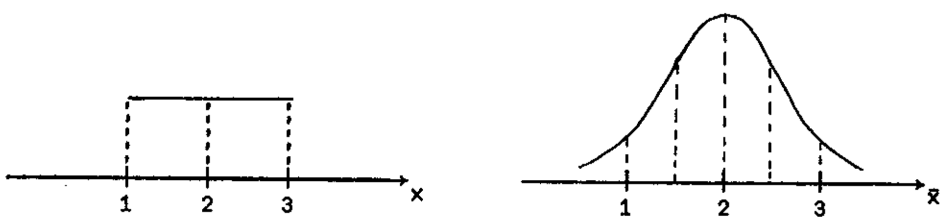


Figura 7.3

As figuras 7.4 dão uma idéia do que acontece quando n cresce, partindo de uma população que tem distribuição do tipo acima, isto é, "uniforme" num certo intervalo.

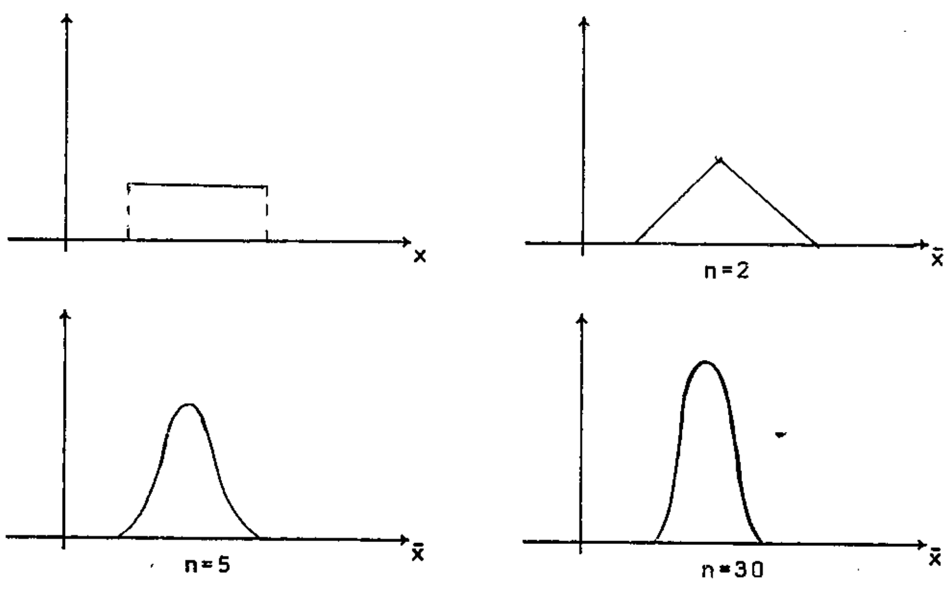


Figura 7.4

A justificativa teórica do que ocorre é baseada no que é chamado o *Teorema do Limite Central*. Quando aproximamos uma distribuição binomial por uma normal, estamos utilizando um caso especial deste teorema. Este teorema afirma o seguinte: se $X_1, X_2, \dots, X_n$ são v.a. independentes, com uma distribuição comum, de média μ e variância σ^2, então a v. a.

$$S_n = X_1 + X_2 + \dots + X_n \tag{7.7}$$

é tal que, para cada valor fixo z e para $n \to \infty$,

$$P\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \geq z\right) \longrightarrow P(Z \geq z), \tag{7.8}$$

onde Z é a v.a. N(0,1). Ver Apêndice 2.

Como $\bar{X} = S_n/n$ vem

$$\frac{S_n - n\mu}{\sigma\sqrt{n}} = \frac{(S_n/n) - \mu}{\sigma/\sqrt{n}} = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}},$$

portanto, na prática, para n grande, $\dfrac{\bar{X} - \mu}{\sigma/\sqrt{n}}$ tem distribuição aproximada normal padrão.

Devido a sua importância, vamos repetir o que foi dito a respeito da distribuição amostral de $\bar{X}$.

<u>Fato</u>: Seja X uma v.a. ("população") com média $E(X) = \mu$ e variância $Var(X) = \sigma^2$. Seja $\bar{X}$ a média amostral de uma amostra aleatória de tamanho n. Então:

(i) $E(\bar{X}) = \mu$.

(ii) $Var(\bar{X}) = \sigma^2/n$.

(iii) Para n grande, $\dfrac{\bar{X} - \mu}{\sigma/\sqrt{n}}$ tem, aproximadamente, distribuição N(0,1).

PROBLEMAS PARA O CAPÍTULO 7

7.1 - Para o exemplo discutido na seção 7.2, obtenha a distribuição do mínimo $Y_1 = \min(X_1, X_2)$ e do máximo $Y_2 = \max(X_1, X_2)$. Calcule $E(Y_1)$ e $E(Y_2)$.

7.2 - Prove que $Var(\bar{X}) = \sigma^2/n$ se a variância de X é σ^2.

7.3 - Prove que $E(S^2) = \sigma^2$. Para isso escreva

$$(n-1)S^2 = \sum_{i=1}^{n} (X_i - \bar{X})^2$$

e então, $X_i - \bar{X} = (X_i - \mu) - (\bar{X} - \mu)$.

7.4 - Considere uma população de tamanho $N = 4$ e a v.a. X tendo os valores 2, 3, 5, 6 sobre os elementos desta população.

a) Calcule E(X) e Var(X).

b) Considere amostras de tamanho $n = 2$, com reposição, da população. Obtenha a distribuição amostral de $\bar{X}$, S^2 e S.

7.5 - *Correção para populações finitas.* Vimos que $Var(\bar{X}) = \sigma^2/n$, se extraímos amostras com reposição da população. Na prática, utilizamos amostragem sem reposição. Se a população é infinita, a fórmula acima ainda é válida, mas se a população é finita, então pode-se demonstrar que

$$Var(\bar{X}) = \frac{\sigma^2}{n} \cdot \frac{N-n}{N-1}, \qquad (*)$$

sendo N o tamanho da população. Se $n \ll N$, $\frac{N-n}{N-1} \approx 1$.

Considere uma população de tamanho $N = 4$ e os valores $X_1 = 0$, $X_2 = 3$, $X_3 = 2$, $X_4 = 3$. Retire amostras de tamanho

$n = 2$, *sem reposição*. Obtenha a distribuição amostral de $\bar{X}$ e calcule $E(\bar{X})$ e $Var(\bar{X})$. Verifique que $E(\bar{X}) = E(X) = \mu$ e $Var(\bar{X})$ é dada por (*).

7.6 - Dada uma população normal com média $\mu = 600$ e desvio padrão $\sigma = 16$, se uma amostra de tamanho $n = 64$ é escolhida, com reposição, desta população, calcular:

a) $P(\bar{X} > 602)$ b) $P(\bar{X} > 610)$

c) $P(602 \leq \bar{X} \leq 610)$ d) $P(|\bar{X}| > 610)$.

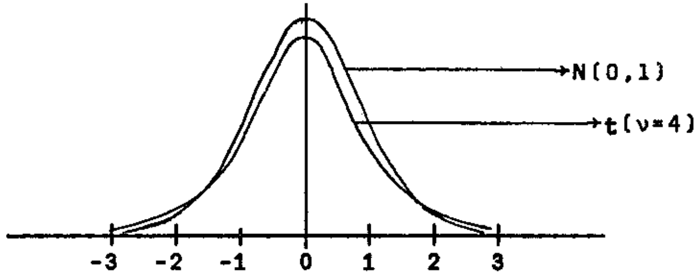
7.7 - Vimos que, se a população é normal com média μ e variância σ^2, então $Z = (\bar{X}-\mu)/(\sigma/\sqrt{n})$ tem distribuição: $N(0,1)$. Suponha que σ seja desconhecido e usamos:

$$S = \frac{\Sigma(X_i - \bar{X})^2}{n-1}$$

no seu lugar. Sabemos que S é um estimador viciado de σ e para valores pequenos de n, o quociente

$$\frac{\bar{X} - \mu}{S/\sqrt{n}}$$

não tem distribuição normal; para n grande a distribuição deste quociente é aproximadamente normal. De qualquer modo, para qualquer valor de n, a distribuição de $(\bar{X}-\mu)/(S/\sqrt{n})$ é conhecida como *distribuição t de Student* e se parece bastante com a $N(0,1)$, sendo que para n pequeno, ela apresenta uma maior dispersão que a $N(0,1)$.



A distribuição de $t = (\bar{X}-\mu)/(S/\sqrt{n})$ depende do valor de n e para $n > 30$ a distribuição de t é praticamente normal padrão. A tabela 3 do Apêndice 1 fornece probabilidades $P(t > t_0)$, para valores de $\nu = n-1$, que é chamado *número de graus de liberdade*.

Por exemplo, se $n = 25$, $\nu = n-1 = 24$ e $P(t > 1,711) = 0,05$.

Calcular:

a) $n = 24$, $P(t > 1,319)$

b) $n = 25$, $P(t < -2,060)$

c) $n = 30$, $P(|t| > 1,96)$

d) ν de tal sorte que $P(t > 2,479) = 0,01$.

7.8 - Suponha que de uma população normal com média $\mu=100$ e variância desconhecida, uma amostra de tamanho $n = 16$ forneceu $S^2 = 2,25$. Obter:

a) $P(\bar{X} > 105)$

b) $P(\bar{X} < 96)$

c) $P(|\bar{X}| > 103)$

- o -

8 - IDÉIAS GERAIS SOBRE INFERÊNCIA ESTATÍSTICA

Para introduzir os conceitos de estimação e testes de hipóteses utilizamos apenas a distribuição binomial: obtivemos estimadores para p e testamos hipóteses relativas a p. Vamos, agora, ver algumas noções gerais, aplicáveis a um grande número de situações. Trataremos, em particular, do chamado princípio de máxima verossimilhança.

8.1 - PROPRIEDADES DOS ESTIMADORES

Vimos que a Inferência Estatística tem por objetivo fazer generalizações sobre uma população com base em dados de uma amostra.

Consideremos uma amostra $(X_1, X_2, \ldots, X_n)$ de uma v.a. X, que descreve uma característica de interesse da população.

Sabemos que um *estimador* é qualquer função das observações $X_1, \ldots, X_n$. Uma estimativa é um valor particular do estimador, para uma amostra particular.

O problema da estimação é, então, com base em uma amostra, determinar uma função $T = f(X_1, \ldots, X_n)$ que seja "próxima" do parâmetro populacional de interesse.

Podemos ter vários estimadores de um mesmo parâmetro θ. Necessitamos, então, ter critérios que nos permitam escolher o "melhor" estimador de θ. Por exemplo, a média amostral $\bar{X}$ e a mediana amostral $Y_{1/2}$ são dois estimadores da média populacional $\mu = E(X)$.

Seja

$$e = T - \theta \tag{8.1}$$

o erro amostral. Ao valor

$$E(e^2) = E(T-\theta)^2 \tag{8.2}$$

denominamos *erro quadrático médio* ou *risco* do estimador T, e será indicado por $R(T,\theta)$. Poderíamos, então, tentar determinar T que minimizasse o risco (8.2), para todo θ. Isto não é possível, em geral.

Se temos dois estimadores T_1 e T_2 do parâmetro θ, se $R(T_1,\theta) < R(T_2,\theta)$, para todo θ, então T_1 seria preferível a T_2. Ver figura 8.1.

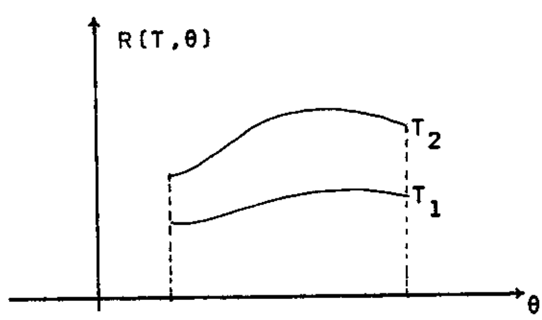


Figura 8.1

Mas a situação pode não ser essa, como indica a figura 8.2.

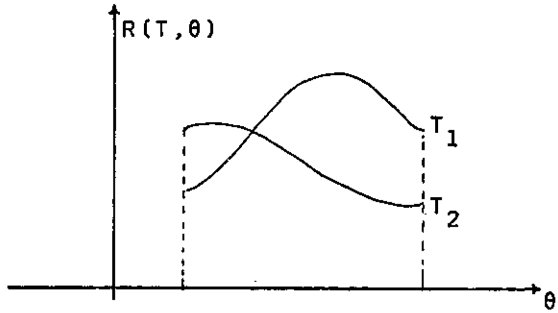


Figura 8.2

Para alguns valores de θ, $R(T_1,\theta) < R(T_2,\theta)$, acontecendo o oposto para outros valores de θ.

O que se faz, então, é tentar restringir a classe de estimadores a usar.

Uma possibilidade é considerar a classe dos *estimadores não viciados* de θ, isto é, se T é um membro desta classe,

$$E(T) = \theta, \tag{8.3}$$

para todo θ. Se esta condição está verificada, então

$$R(T,\theta) = E[T-E(T)]^2, \tag{8.4}$$

que é a variância de T. Escolher, pois, um estimador na classe dos estimadores não viciados do parâmetro θ, que minimiza o risco, significa escolher um estimador que minimiza a variância (8.4). Se existe tal estimador, dizemos que ele é um *estimador não viciado de variância mínima* de θ. Para algumas situações é possível encontrar o estimador não viciado de variância mínima de um parâmetro θ. Mas para isso é necessário recorrer a conceitos que estão fora do alcance destas notas.

<u>Exemplo 8.1</u> - $\bar{X}$ é um estimador não viciado da média $\mu = E(X)$ de uma v.a. X. De fato,

$$E(\bar{X}) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i).$$

Como cada X_i tem a mesma distribuição que X, $E(X_i) = \mu$, logo

$$E(\bar{X}) = \frac{1}{n}\sum_{i=1}^{n}\mu = \frac{n\mu}{\mu} = \mu.$$

Do mesmo modo, num experimento binomial, se X = número de sucessos, $P(\text{sucesso}) = p$, $\hat{p} = X/n$ é um estimador não viciado de p, como já vimos.

Vejamos, agora, um outro critério. Considere, por exemplo, $\bar{X}$ calculado para diversos tamanhos de amostras; obtemos, na realidade, uma seqüência de estimadores $\{\bar{X}_n, n=1, 2, 3, \ldots\}$. Suponha que a v.a. X é normal $N(\mu, \sigma^2)$. A distribuição de $\bar{X}_n$ é $N(\mu, \frac{\sigma^2}{n})$. À medida que n cresce, a distribuição de $\bar{X}_n$ torna-se mais concentrada ao redor de μ. Ver figura 8.3.

Dizemos que $\{\bar{X}_n\}$ é uma seqüência *consistente* de estimadores de μ. Formalmente, uma seqüência $\{T_n\}$ de estimadores de um parâmetro θ é consistente se, para todo $\varepsilon > 0$,

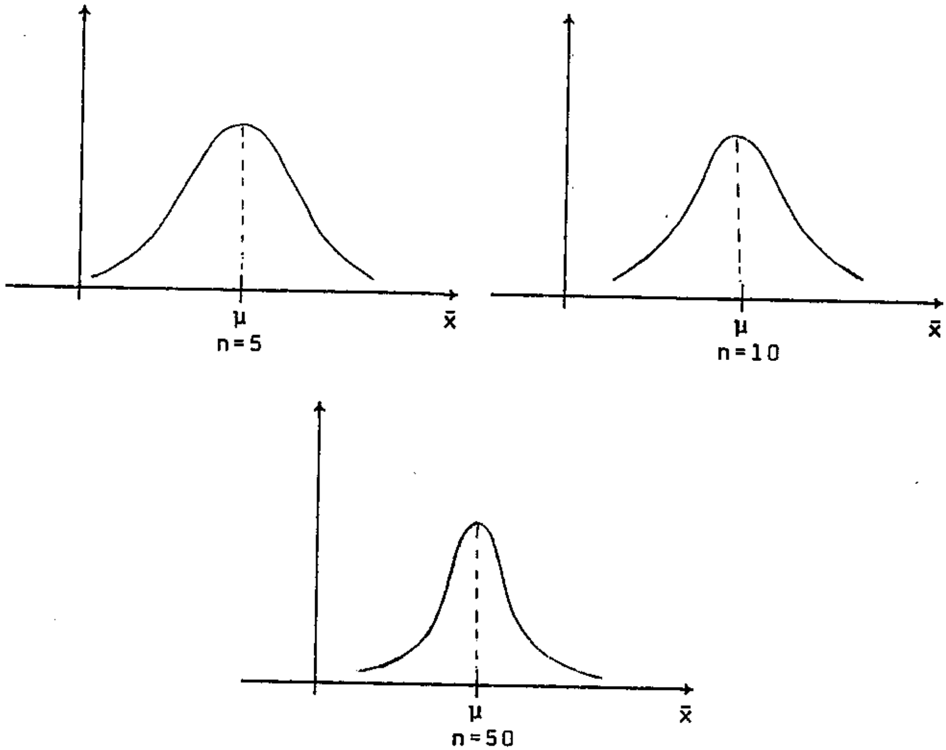


Figura 8.3

$$P\{|T_n - \theta| > \varepsilon\} \longrightarrow 0, \; n \to \infty. \tag{8.5}$$

É fácil ver que esta condição está satisfeita para $\{\bar{X}_n\}$, utilizando-se a *desigualdade de Chebyshev*, que afirma: se X é uma v.a. com $E(X) = \mu$ e $Var(X) = \sigma^2$, então

$$P\{|X - \mu| > \varepsilon\} \leq \frac{\sigma^2}{\varepsilon^2}, \tag{8.6}$$

para todo número $\varepsilon > 0$. Ver Apêndice 2.

Usando (8.6), vemos que

$$P\{|\bar{X}_n-\mu|>\varepsilon\} \leq \frac{Var(\bar{X}_n)}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2} \longrightarrow 0,\ n\to\infty, \tag{8.7}$$

o que mostra que $\{\bar{X}_n\}$ é consistente.

8.2 - ESTIMADORES DE MÍNIMOS QUADRADOS E ESTIMADORES DE MÁXIMA VEROSSIMILHANÇA

Retomemos o exemplo 4.3, em que tínhamos o modelo

$$d_i = D+e_i,\ i=1,\ldots,n. \tag{8.8}$$

Queremos estimar D com base nas n observações $d_1,\ldots,d_n$. O *método dos mínimos quadrados* consiste em escolher o estimador $\hat{D}$, de D, que minimiza a soma dos quadrados dos erros e_i. Como

$$e_i = d_i-D,\ i=1,\ldots,n, \tag{8.9}$$

vemos que

$$\sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (d_i-D)^2, \tag{8.10}$$

que é uma função de D, g(D) digamos. Se queremos determinar o mínimo desta função, temos

$$g'(D) = \sum_{i=1}^{n} (d_i-\hat{D})(-2) = 0, \tag{8.11}$$

do que decorre

$$\hat{D} = \frac{\sum_{i=1}^{n} d_i}{n} , \tag{8.12}$$

isto é, o chamado *estimador de mínimos quadrados* de D é a média amostral dos d_i.

Se queremos obter um intervalo de confiança para D ou testar hipóteses a respeito de D é necessário fazer algumas hipóteses a respeito da distribuição dos erros e_i. Normalmente, as suposições que se fazem são:

(i) e_i é uma v.a. com média 0, ou é simétrica ao redor de 0;

(ii) e_i é uma v.a. normal;

(iii) a variância de e_i é constante.

A situação geral é aquela em que queremos estimar um parâmetro vetorial $\underset{\sim}{\theta} = (\theta_1,\ldots,\theta_r)$ e as observações satisfazem ao modelo

$$X_i = g_i(\underset{\sim}{\theta}) + e_i , \tag{8.13}$$

$i=1,\ldots,n$. Escolhemos o estimar $\hat{\theta}$ que minimiza

$$\sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (X_i - g_i(\underset{\sim}{\theta}))^2. \tag{8.14}$$

Outro método de estimação que é bastante utilizado é o chamado *método de máxima verossimilhança*.

O princípio de máxima verossimilhança afirma o seguinte: devemos escolher aquele valor de θ que maximiza a probabilidade de obter a amostra observada, na ordem particular na qual os itens da amostra apareceram.

Exemplo 8.2 - Suponha que temos n provas de Bernoulli com P(sucesso) = p e X = número de sucessos. Devemos tomar como estimador aquele valor de p que torna a amostra observada a mais provável de ocorrer.

Suponha que $n = 3$ e obtemos 2 sucessos e 1 fracasso. A *função de verossimilhança* é

$$L(p) = P(2 \text{ sucessos e } 1 \text{ fracasso}) = p^2(1-p) \qquad (8.15)$$

Maximizando esta função, obtemos:

$L'(p) = 2p(1-p) - p^2 = 0 \Longrightarrow p(2-3p) = 0$, do que seguem $p = 0$ e $p = 2/3$. É fácil ver que o ponto de máximo é $\hat{p} = 2/3$ que é o estimador de máxima verossimilhança de p.

De modo geral, o estimador de máxima verossimilhança do parâmetro p de uma binomial, com X sucessos em n prova é

$$\hat{p} = \frac{X}{n} . \qquad (8.16)$$

Observe que este é o estimador que temos usado constantemente. Para se chegar ao resultado (8.16) observamos que a função de verossimilhança neste caso é

$$L(p) = p^x(1-p)^{n-x}, \qquad (8.17)$$

e que o máximo desta função ocorre no mesmo ponto que

$$\log L(p).$$

Portanto, sendo

$$\ell(p) = \log L(p) = x \log p + (n-x) \log(1-p),$$

temos

$$\ell'(p) = \frac{x}{p} - \frac{n-x}{1-p} = 0 \Longrightarrow \hat{p} = \frac{x}{n}.$$

<u>Exemplo 8.3 -</u> Em muitos problemas biológicos, tem importância o problema da estimação do tamanho de populações animais ou vegetais. Implicações práticas referem-se ao controle de insetos e manutenção de suprimentos de alimentos.

Suponha que capturamos, marcamos e depois soltamos r animais de uma população total de N animais, sendo N desconhecido. Quando os animais marcados tiverem se dispersado entre os não marcados, nós capturamos n animais (isto é, colhemos uma amostra de tamanho n da população de N animais) dos quais k são marcados. Vamos supor, primeiramente, que n é suficientemente pequeno, comparado com N, para que possamos ignorar complicações com a amostragem sem reposição.

A função de verossimilhança será

$$L(N) = \binom{n}{k}\left(\frac{r}{N}\right)^k\left(1 - \frac{r}{N}\right)^{n-k}, \tag{8.18}$$

e tomando logaritmo,

$$\ell(N) = \text{constante} + (n-k)\log(N-r) - n \log N.$$

Se bem que N seja discreto, será conveniente tratá-lo como se fosse contínuo, de modo que

$$\frac{\partial \ell}{\partial N} = \frac{N-k}{N-r} - \frac{n}{N}. \tag{8.19}$$

Donde, o estimador de máxima verossimilhança de N será

$$\hat{N} = \frac{rn}{k}. \tag{8.20}$$

Se capturamos, por exemplo, 100 animais e os marcamos e após algum tempo efetuamos nova captura de 100 animais e vemos que 10 dentre eles são marcados, então uma estimativa de N será $\hat{N} = \frac{100 \times 100}{10} = 1.000$ animais.

Utilizando a distribuição hipergeométrica, no caso em que a suposição acima feita, sobre a relação entre n e N não está satisfeita, temos que a função de verossimilhança será

$$L(N) = \frac{\binom{r}{k}\binom{N-r}{n-k}}{\binom{N}{n}} \tag{8.21}$$

Como $N-r \geq n-k$, vemos que podemos ter, de início, que

$$N \geq r + n - k.$$

No exemplo numérico acima, $N \geq 190$. Por exemplo, podemos testar a hipótese

$$H: N \leq 190$$

contra (8.22)

$$K: N > 190.$$

É fácil ver que rejeitamos H, bastando calcular a probabilidade de se obter 10 marcados dentre os 100 escolhidos, para N=190. Esta probabilidade será bem pequena. Para obter o estimador de máxima verossimilhança, neste caso, o trabalho é mais complicado. Pode-se demonstrar que o estimador é dado por

$$\hat{N} = \left[\frac{rn}{k}\right], \tag{8.23}$$

onde [y] indica o maior inteiro contido em y. No exemplo acima, $\hat{N} = \left[\frac{100\times 100}{10}\right] = 1.000$, que é o mesmo estimador obtido antes. (Ver problemas).

8.3 - O PROBLEMA GERAL DO TESTE DE HIPÓTESES

No capítulo 6 introduzimos as primeiras idéias sobre teste de hipóteses e tratamos extensivamente do caso em que queremos por à prova afirmações sobre o parâmetro p de uma distribuição binomial.

Vamos considerar os seguintes exemplos.

<u>Exemplo 8.4</u> - Há uma certa variabilidade no tempo necessário para que uma máquina complete dada tarefa. Deseja-se testar a validade da afirmação:

"A máquina leva, em média, 20 minutos para efetuar a tarefa".

Exemplo 8.5 - Um gerente de vendas deseja testar dois esquemas de promoção de vendas, I e II digamos. Especificamente, ele quer testar a hipótese:

"Ambos os esquemas são semelhantes, no sentido que as quantidades do produto vendidas, decorrentes do uso de I e II, são as mesmas."

Cada afirmação acima é uma *hipótese estatística*. Para testar cada uma delas, teremos que usar algum processo ou regra de decisão que envolva:

a) escolher uma amostra aleatória;

b) encontrar uma estatística conveniente;

c) usar a distribuição amostral da estatística para tomar uma decisão.

Estes três passos constituem um teste da hipótese em questão. De maneira um pouco mais formal diremos que uma hipótese estatística é uma afirmação sobre o valor de um parâmetro desconhecido da distribuição de uma variável aleatória, que representa uma característica de interesse de uma população.

Denotando por H cada afirmação dos exemplos citados, poderemos escrever, em uma notação concisa,

$$H: \mu = 20 \quad e \quad H: \mu_I = \mu_{II}, \tag{8.24}$$

para os exemplos 8.4 e 8.5, respectivamente. No primeiro caso, o tempo necessário para completar a tarefa é uma v.a. cuja distribuição é desconhecida, totalmente ou parcialmente, e μ representa a média desta distribuição. Interpretação semelhante para o segundo caso.

Chamaremos H de *hipótese nula*; para cada hipótese nula formulada teremos uma *hipótese alternativa* e que pode tomar formas diferentes, dependendo do problema. Denotando-se por K uma hipótese alternativa, podemos ter para os exemplos dados:

$$\text{Exemplo 8.4: } \begin{cases} H: \mu = 20 \\ K: \mu \neq 20 \end{cases} \tag{8.25}$$

$$\text{Exemplo 8.5: } \begin{cases} H: \mu_I = \mu_{II} \\ K: \mu_I > \mu_{II} \end{cases} \tag{8.26}$$

Vamos analisar detidamente o exemplo 8.4. Aqui, o parâmetro que está sendo testado é uma média desconhecida.

O conjunto de todos os valores possíveis de μ é chamado o *espaço do parâmetro* e é indicado pela letra Θ. Podemos supor aqui que

$$\Theta = \{\mu;\ \mu \geq 0\}, \tag{8.27}$$

já que a variável em questão é tempo.

Sob a hipótese nula, $\mu = 20$ e sob a hipótese alternativa, $\mu \neq 20$. Estas duas hipóteses, H e K, definem dois subconjuntos de Θ:

$$\Theta_0 = \{\mu: \mu = 20\} = \{20\} \tag{8.28}$$

e

$$\Theta_1 = \Theta - \Theta_0 = \{\mu: \mu \neq 20\}. \tag{8.29}$$

Em termos dos conjuntos acima, a formulação (8.25) é equivalente a

$$\begin{aligned} H&: \mu \in \Theta_0 \\ K&: \mu \in \Theta_1 \end{aligned} \tag{8.30}$$

Esta formulação de um problema de teste de hipóteses é a mais geral possível e sempre pode ser feita.

Dizemos que uma hipótese H (ou K) é simples se Θ_0 (ou Θ_1) reduz-se a um ponto, isto é, se especifacamos completamente o valor do parâmetro. No exemplo 8.4, H: $\mu = 20$ é uma hipótese simples(*). Uma hipótese H (ou K) diz-se *composta* se Θ_0 (ou Θ_1) possui mais que um ponto. A hipótese K: $\mu \neq 20$ do mesmo exemplo é composta.

A situação mais fácil de ser estudada é quando temos uma hipótese simples contra uma alternativa simples:

$$\begin{aligned} H&: \theta = \theta_0 \\ K&: \theta = \theta_1 \end{aligned} \tag{8.31}$$

No exemplo 8.4, poderíamos estar interessados em testar a hipótese que a média é 20, contra a alternativa que

---

(*) Estamos supondo aqui que μ é o único parâmetro desconhecido.

a média é 23.

Lembremos que ao testar uma hipótese podemos tomar uma decisão errada em duas situações:

(i) rejeitando-se H, quando ela é verdadeira: *erro de tipo I*;

(ii) aceitando-se H, quando ela é falsa: *erro de tipo II*.

Designamos por α e β, respectivamente, as probabilidades de cada tipo de erro.

Para a situação (8.31), o quadro abaixo resume as possibilidades que podem ocorrer.

DECISÃO	"ESTADO DA NATUREZA"	
	θ_0	θ_1
ACEITAR H	decisão correta	decisão incorreta: ERRO DE TIPO I
ACEITAR K	decisão incorreta: ERRO DE TIPO II	decisão correta

Por "estado da natureza" entendemos o verdadeiro valor do parâmetro θ.

Vimos que o primeiro passo para testar H contra K é dispor de observações $(X_1,\ldots,X_n)$ da população. Construímos, então, uma *função de decisão*, que é uma função definida no espaço de todas as amostras de tamanho n (o *espaço a-*

mostral) e com valores no que chamaremos *espaço das ações*. Designando-se estes espaços por S e $\mathcal{A}$, respectivamente, e por d uma tal função de decisão, então d: $S \to \mathcal{A}$ é uma função tal que a todo ponto $(X_1, \ldots, X_n) \in S$ corresponda uma ação $a \in \mathcal{A}$: $a = d(X_1, \ldots, X_n)$.

No caso de um teste de hipóteses, o espaço $\mathcal{A}$ consiste apenas de duas ações, a_0 e a_1, digamos, com:

a_0: ação de rejeitar H(e aceitar K);

a_1: ação de rejeitar K(e aceitar H).

Portanto, $\mathcal{A} = \{a_0, a_1\}$. O conjunto S ficará, automáticamente, particionado em dois subconjuntos, S_0 e S_1, disjuntos e complementares, tais que:

S_0 = conjunto de todos os pontos $(X_1, \ldots, X_n)$ tais que somos levados a tomar a ação a_0: rejeitar H;

S_1 = conjunto de todos os pontos $(X_1, \ldots, X_n)$ tais que somos levados a tomar a ação a_1: rejeitar K;

O conjunto S_0 é a *região crítica* do teste, ou o conjunto (ou região) de rejeição da hipótese nula H.

Usualmente, o conjunto S_0 é determinado em termos da estatística usada no teste. Por exemplo, suponha que vamos basear nossa decisão em aceitar ou rejeitar H: $\mu = 20$, na estatística $\bar{X}$, isto é, na média amostral. Então, a região crítica poderá ser da forma $S_0 = \{(x_1, \ldots, x_n): \bar{x} \geq c\}$, onde c é uma constante a ser determinada convenientemente. Neste caso, as ações a_0 e a_1 seriam:

a_0: rejeitar H quando $\bar{x} \geq c$;

a_1: aceitar H (rejeitar K) quando $\bar{x} < c$.

Lembremos que, sendo impossível minimizar α e β simultaneamente, fixamos um valor para α e tentamos encontrar, dentre todos os testes com nível de significância α, aquele que minimiza β.

Qualquer procedimento que utilizemos nos levará a encontrar uma estatística T, que é uma função das observações, $T(X_1,\ldots,X_n)$, que será utilizada no teste. Normalmente, teremos que determinar um *valor crítico*, T_c, da estatística, que nos permite escolher entre H e K. Este valor crítico delimitará a região crítica, que será da forma

$$\{(x_1,\ldots,x_n): T(x_1,\ldots,x_n) \geq T_c\} \qquad (8.32)$$

ou

$$\{(x_1,\ldots,x_n): T(x_1,\ldots,x_n) < T_c\} \qquad (8.33)$$

<u>Exemplo 8.6</u> - Suponha que queremos testar

$$H: \mu = 50$$

contra a alternativa

$$K: \mu < 50,$$

onde μ é a média de uma normal $N(\mu,900)$. Suponha que a regra de decisão seja a seguinte:

"Extraída uma amostra de tamanho $n = 36$, rejeite H se $\bar{X} < \bar{X}_c$, onde $\bar{X}_c$ é o valor crítico de $\bar{X}$, determinado de tal sorte $\alpha = 0{,}025$".

A região crítica será da forma

$$S_0 = \{(x_1,\ldots,x_{36}): \bar{x} < \bar{x}_c\}. \qquad (8.34)$$

Se a população é $N(\mu,900)$, então $\bar{X}$ tem distribuição (distribuição amostral de $\bar{X}$) normal, com média μ e variância $\frac{900}{n} = \frac{900}{36} = 25$. Logo,

$$\alpha = P(\bar{X} < \bar{X}_c \mid H \text{ verdadeira}) = P\left(\frac{\bar{X}-\mu}{\sigma_{\bar{X}}} < \frac{\bar{X}_c-\mu}{\sigma_{\bar{X}}} \middle| H\right),$$

$\sigma_{\bar{X}} = 5$ é o desvio padrão de $\bar{X}$. Portanto,

$$\alpha = \left(Y < \frac{\bar{X}_c-50}{5}\right), \qquad (8.35)$$

onde Y tem distribuição normal $N(0,1)$. Como $\alpha = 0{,}025$, a relação (8.35) implica (ver tabela 2 e figura 8.4)

$$\frac{\bar{X}_c-50}{5} = 1{,}96,$$

ou seja, $\bar{X}_c = 40{,}2$. Logo, as regras de decisão a_0 e a_1 serão:

a_0: rejeitar H se $\bar{X} < 40{,}2$;

a_1: rejeitar K se $\bar{X} \geq 40{,}2$.

8.4 - O TESTE DA RAZÃO DE VEROSSIMILHANÇA

Vamos considerar, nesta seção, apenas o caso em que

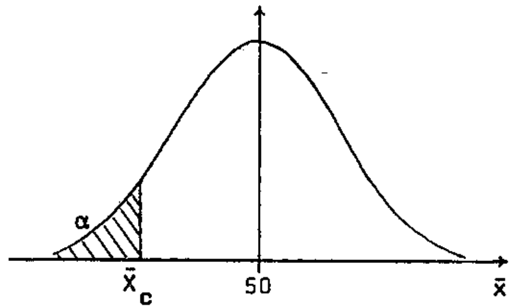


Figura 8.4

queremos testar uma hipótese simples contra uma alternativa simples, ou seja,

$$\begin{aligned} H&: \theta = \theta_0 \\ K&: \theta = \theta_1 \end{aligned} \tag{8.36}$$

Como vimos, nós controlamos o nível de significância, α, e tentamos minimizar a probabilidade do erro de tipo II, β. Se existe uma região crítica que tenha nível α e que minimiza β dentre todas as regiões críticas de nível no máximo α, esta será a *melhor região crítica de nível* α. Obtemos, então, um *melhor teste*.

Suponha que temos uma amostra $(X_1,\ldots,X_n)$ de uma v.a. X que tenha distribuição dependendo de um parâmetro θ que pode assumir somente os valores θ_0 e θ_1. Fixemos α, $0<\alpha<1$. Designemos por $L(\theta;x_1,\ldots,x_n)$ a função de verossimilhança correspondente à amostra $x_1,\ldots,x_n$, isto é, a probabilidade $P(X_1=x_1,\ldots,X_n=x_n)$, sendo que destacamos a depen-

$$\alpha = P(\text{rejeitar } H|\theta=\tfrac{1}{3}) = P\left(\sum_{i=1}^{n} X_i \geq C \,|\, \theta=\tfrac{1}{3}\right) = \sum_{i=C}^{5}\binom{5}{i}\left(\frac{1}{3}\right)^{i}\left(\frac{2}{3}\right)^{5-i}.$$

Se $\alpha=0{,}0453$, então $C=4$ e $S_0=\{(x_1,\ldots,x_5): \sum_{i=1}^{5} X_i \geq 4\}$ e se $\alpha=0{,}2099$, então $C=3$ e $S_0=\{(x_1,\ldots,x_5): \sum_{i=1}^{5} X_i \geq 3\}$, mas, se $\alpha=0{,}05$, não existe constante C e região crítica S_0 satisfazendo às condições estipuladas. Na prática, mudamos o valor de α para o qual o teste pode ser encontrado, ou então usamos um procedimento chamado *aleatorização*, que permite obter um teste de nível exatamente igual ao valor α fixado. Mas este assunto não será discutido aqui.

<u>Exemplo 8.8</u> - Suponha que queiramos testar

$$\begin{aligned} H&: \theta = \theta_0 \\ K&: \theta = \theta_1 \end{aligned} \qquad (8.43)$$

onde $\theta_1>\theta_0$, e θ é a média desconhecida de uma normal $N(\theta,\sigma_0^2)$, onde σ_0^2 é conhecido.

Aqui, para uma amostra de tamanho n, onde e^y será denotado por $\exp\{y\}$,

$$L(\theta;x_1,\ldots,x_n) = \prod_{i=1}^{n}\frac{1}{\sigma_0\sqrt{2\pi}}\exp\left\{-\frac{1}{2}\left(\frac{x_i-\theta}{\sigma_0}\right)^2\right\} =$$

$$= \left(\frac{1}{2\pi\sigma_0^2}\right)^{n/2}\exp\left\{-\frac{1}{2\sigma_0^2}\sum_{i=1}^{n}(x_i-\theta)^2\right\} \qquad (8.44)$$

Portanto,

$$\lambda = \frac{L_0}{L_1} = \frac{\exp\left\{-\frac{1}{2\sigma_0^2} \sum_{i=1}^{n} (x_i-\theta_0)^2\right\}}{\exp\left\{-\frac{1}{2\sigma_0^2} \sum_{i=1}^{n} (x_i-\theta_1)^2\right\}} = \exp\left\{-\frac{1}{2\sigma_0^2}\left[\sum(x_i-\theta_0)^2-\sum(x_i-\theta_1)^2\right]\right\}$$

Segue-se que $\lambda \leq k$ é equivalente a

$$\exp\left\{-\frac{1}{2\sigma_0^2}\left[2(\theta_1-\theta_0)\sum x_i+n(\theta_0^2-\theta_1^2)\right]\right\}\leq k,$$

ou

$$\exp\left\{\frac{\theta_0 - \theta_1}{\sigma_0^2} \sum x_i + \frac{n(\theta_1^2-\theta_0^2)}{2\sigma_0^2}\right\} \leq k,$$

ou ainda,

$$\frac{\theta_0 - \theta_1}{\sigma_0^2} \sum x_i + \frac{n(\theta_1^2 - \theta_0^2)}{2\sigma_0^2} \leq \log k. \qquad (8.45)$$

Como $\theta_0-\theta_1<0$, obtemos

$$\sum x_i \geq \frac{(\log k)\sigma_0^2}{\theta_0-\theta_1} + \frac{n\cdot(\theta_1+\theta_0)}{2},$$

isto é,

$$\sum_{i=1}^{n} x_i \geq C, \qquad -\infty<C<+\infty. \qquad (8.46)$$

Portanto, a região crítica é

$$S_0 = \{(x_1,\ldots,x_n): \sum_{i=1}^{n} x_i \geq C\}$$

é não viciado e de variância mínima?

8.4 - Suponha que obtemos uma amostra $(X_1, \ldots, X_n)$ da distribuição de uma v.a. X que tem densidade

$$f(x) = \alpha e^{-\alpha x}, \quad x \geq 0.$$

Obtenha o estimador de máxima verossimilhança de α.

8.5 - A v.a. Y é observada para cada um dos n níveis $X_1, \ldots, X_n$ de uma variável fixa X, estabelecendo-se o seguinte modelo:

$$Y_i = \alpha + \beta X_i + \varepsilon_i, \qquad i=1, \ldots, n$$

onde α e β são constantes a serem determinadas e ε_i são erros (variáveis aleatórias). Obtenha os estimadores de mínimos quadrados para α e β, isto é, aqueles valores $\hat{\alpha}$ e $\hat{\beta}$ que minimizam

$$\sum_{i=1}^{n} (Y_i - \alpha - \beta X_i)^2.$$

8.6 - Suponha que os valores observados de Y, para os valores $X_1=0$, $X_2=1$, $X_3=2$, $X_4=3$ e $X_5=4$ são, respectivamente, $Y_1=-3{,}5$, $Y_2=0{,}5$, $Y_3=1{,}2$, $Y_4=2{,}8$ e $Y_5=5{,}1$. Calcule os estimadores $\hat{\alpha}$ e $\hat{\beta}$ do problema 8.5.

8.7 - No exemplo 8.3, considere o caso em que temos a função de verossimilhança dada por (8.21), isto é, o modelo é a distribuição hipergeométrica. Para se obter o estimador de máxima verossimilhança de N, $\hat{N}$, considere a razão

$$\frac{L(N)}{L(N-1)} = \frac{(N-n)(N-r)}{(N-n-r+k)N}.$$

Veja o que acontece quando $\frac{L(N)}{L(N-1)} \gtrless 1$.

8.8 - Queremos testar H: $\theta=\theta_0$ contra a alternativa K: $\theta=\theta_1$, se X é uma v.a. discreta assumindo os valores 0,1,2, 3,4 e 5. Suponha que fazemos uma só observação, X_1. Na tabela abaixo damos as distribuições de X sob H e K.

x	0	1	2	3	4	5
$P(X=x\|\theta_0)$	0,01	0,03	0,04	0,04	0,38	0,50
$P(X=x\|\theta_1)$	0,04	0,05	0,010	0,010	0,40	0,31

Fixe o nível de significância $\alpha = 0,04$. Determine as possíveis regiões críticas e o erro de tipo II, β, associado a cada uma delas. Encontre o melhor teste de nível α.

8.9 - Para o exemplo 8.8, encontre o erro de tipo II, β.

8.10- Para o mesmo exemplo 8.8, determine o valor de n que nos dê valores fixados de α e β, precisamente, $\alpha=0,05$ e $\beta=0,10$.

8.11- No exemplo 8.8, vimos que o teste (8.46) é uniformemente melhor para testar H: $\theta=\theta_0$ contra K: $\theta>\theta_0$. Suponha, agora, que queremos testar H: $\theta=\theta_0$ contra K: $\theta\neq\theta_1$. Usando considerações de simetria, explique porque não esperamos obter um melhor teste, uniformemente, quando o valor alternativo de θ pode ser maior ou menor que θ_0.

8.12- Determinar o teste da razão de verossimilhança para testar

$$H: \lambda = \lambda_0$$

$$K: \lambda = \lambda_1 < \lambda_0$$

onde λ é a média desconhecida de uma distribuição de Poisson,

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

Aplique o resultado para $\lambda_0=5$, $\lambda_1=4,8$.

- o -

REFERÊNCIAS

[1] - Bailey, N.T.J.: ON ESTIMATING THE SIZE OF MOBILE POPULATIONS FROM RECAPTURE DATA. Biometrika, Vol. 38, 1951, pp. 293-306.

[2] - Blackwell D.: ESTATÍSTICA BÁSICA. Ed.McGraw-Hill do Brasil, 1973.

[3] - Feller, W.: AN INTRODUCTION TO PROBABILITY THEORY AND ITS APPLICATIONS. J.Wiley & Sons,3ª edição, 1968.

[4] - Fernandez, P.J.: INTRODUÇÃO À TEORIA DAS PROBABILIDADES. Livros Técnicos e Científicos S.A., 1973.

[5] - Lindgren, B.: ELEMENTS OF DECISION THEORY.The Mac Millan Company, 1971.

[6] - Meyer, P.L.: INTRODUCTORY PROBABILITY AND STATISTICAL APPLICATIONS. Addison Wesley, 1965.

[7] - Mosteller, Rourke & Thomas: PROBABILITY WITH STATISTICAL APPLICATIONS, 2ª edição, Addison Wesley,1970.

[8] - Noether,G.: INTRODUCTION TO STATISTICS. A FRESH APPROACH. Houghton Mifflin, 1971.

- o -

APÊNDICE 1

TABELA 1 - PROBABILIDADES BINOMIAIS

Exemplo: b(6;10,0,3) = 0,037 = b(4;10,0,7)

n=2								
p →	0,05	0,10	0,20	0,25	0,30	0,40	0,50	n=2
x=0	902	810	640	562	490	360	250	2
1	095	180	320	375	420	480	500	1
2	002	010	040	062	090	160	250	0=x
n=2	0,95	0,90	0,80	0,75	0,70	0,60	0,50	← p

n=3								
p →	0,05	0,10	0,20	0,25	0,30	0,40	0,50	n=3
x=0	857	729	512	422	343	216	125	3
1	135	243	384	422	441	432	375	2
2	007	027	096	141	189	288	375	1
3	0+	001	008	016	027	064	125	0=x
n=3	0,95	0,90	0,80	0,75	0,70	0,60	0,50	← p

n=4								
p →	0,05	0,10	0,20	0,25	0,30	0,40	0,50	n=4
x=0	815	656	410	316	240	130	062	4
1	171	292	410	422	412	346	250	3
2	014	049	154	211	265	346	375	2
3	0+	004	026	047	076	154	250	1
4	0+	0+	002	004	008	026	062	0=x
n=4	0,95	0,90	0,80	0,75	0,70	0,60	0,50	← p

n=29								
p →	0,05	0,10	0,20	0,25	0,30	0,40	0,50	n=29
x=0	226	047	002	0+	0+	0+	0+	29
1	345	152	011	002	0+	0+	0+	28
2	254	238	039	011	002	0+	0+	27
3	120	236	088	032	009	0+	0+	26
4	041	171	144	070	026	002	0+	25
5	011	095	179	116	055	006	0+	24
6	002	042	179	155	095	015	001	23
7	0+	015	147	170	133	034	003	22
8	0+	005	101	156	157	062	008	21
9	0+	001	059	121	157	096	019	20
10	0+	0+	030	081	135	128	037	19
11	0+	0+	013	047	100	147	064	18
12	0+	0+	005	023	064	147	097	17
13	0+	0+	002	010	036	128	126	16
14	0+	0+	0+	004	018	098	144	15
15	0+	0+	0+	001	008	065	144	14
16	0+	0+	0+	0+	003	038	126	13
17	0+	0+	0+	0+	001	019	097	12
18	0+	0+	0+	0+	0+	009	064	11
19	0+	0+	0+	0+	0+	003	037	10
20	0+	0+	0+	0+	0+	001	019	9
21	0+	0+	0+	0+	0+	0+	008	8
22	0+	0+	0+	0+	0+	0+	003	7
23	0+	0+	0+	0+	0+	0+	001	6
24	0+	0+	0+	0+	0+	0+	0+	5
25	0+	0+	0+	0+	0+	0+	0+	4
26	0+	0+	0+	0+	0+	0+	0+	3
27	0+	0+	0+	0+	0+	0+	0+	2
28	0+	0+	0+	0+	0+	0+	0+	1
29	0+	0+	0+	0+	0+	0+	0+	0=x
n=29	0,95	0,90	0,80	0,75	0,70	0,60	0,50	← p

n=30								
p →	0,05	0,10	0,20	0,25	0,30	0,40	0,50	n=30
x=0	215	042	001	0+	0+	0+	0+	30
1	339	141	009	002	0+	0+	0+	29
2	259	228	034	009	002	0+	0+	28
3	127	236	079	027	007	0+	0+	27
4	045	177	133	060	021	001	0+	26
5	012	102	172	105	046	004	0+	25
6	003	047	179	145	083	012	001	24
7	0+	018	154	166	122	026	002	23
8	0+	006	111	159	150	050	005	22
9	0+	002	068	130	157	082	013	21
10	0+	0+	035	091	142	115	028	20
11	0+	0+	016	055	110	140	051	19
12	0+	0+	006	029	075	147	081	18
13	0+	0+	002	013	044	136	112	17
14	0+	0+	001	005	023	110	135	16
15	0+	0+	0+	002	011	078	144	15
16	0+	0+	0+	001	004	049	135	14
17	0+	0+	0+	0+	001	027	112	13
18	0+	0+	0+	0+	0+	013	081	12
19	0+	0+	0+	0+	0+	005	051	11
20	0+	0+	0+	0+	0+	002	028	10
21	0+	0+	0+	0+	0+	001	013	9
22	0+	0+	0+	0+	0+	0+	005	8
23	0+	0+	0+	0+	0+	0+	002	7
24	0+	0+	0+	0+	0+	0+	001	6
25	0+	0+	0+	0+	0+	0+	0+	5
26	0+	0+	0+	0+	0+	0+	0+	4
27	0+	0+	0+	0+	0+	0+	0+	3
28	0+	0+	0+	0+	0+	0+	0+	2
29	0+	0+	0+	0+	0+	0+	0+	1
30	0+	0+	0+	0+	0+	0+	0+	0=x
n=30	0,95	0,90	0,80	0,75	0,70	0,60	0,50	← p

n=31								
p →	0,05	0,10	0,20	0,25	0,30	0,40	0,50	n=31
x=0	204	038	001	0+	0+	0+	0+	31
1	333	131	008	001	0+	0+	0+	30
2	263	219	029	007	001	0+	0+	29
3	134	235	070	022	006	0+	0+	28
4	049	183	122	052	017	001	0+	27
5	014	110	164	094	039	003	0+	26
6	003	053	178	135	072	009	0+	25
7	001	021	159	161	110	020	001	24
8	0+	007	119	161	142	041	004	23
9	0+	002	076	137	155	070	009	22
10	0+	0+	042	101	146	107	021	21
11	0+	0+	020	064	120	130	039	20
12	0+	0+	008	036	085	144	066	19
13	0+	0+	003	017	054	141	096	18
14	0+	0+	001	007	030	120	123	17
15	0+	0+	0+	003	014	091	140	16
16	0+	0+	0+	001	006	061	140	15
17	0+	0+	0+	0+	002	036	123	14
18	0+	0+	0+	0+	001	019	096	13
19	0+	0+	0+	0+	0+	008	066	12
20	0+	0+	0+	0+	0+	003	039	11
21	0+	0+	0+	0+	0+	001	021	10
22	0+	0+	0+	0+	0+	0+	009	9
23	0+	0+	0+	0+	0+	0+	004	8
24	0+	0+	0+	0+	0+	0+	001	7
25	0+	0+	0+	0+	0+	0+	0+	6
26	0+	0+	0+	0+	0+	0+	0+	5
27	0+	0+	0+	0+	0+	0+	0+	4
28	0+	0+	0+	0+	0+	0+	0+	3
29	0+	0+	0+	0+	0+	0+	0+	2
30	0+	0+	0+	0+	0+	0+	0+	1
31	0+	0+	0+	0+	0+	0+	0+	0=x
n=31	0,95	0,90	0,80	0,75	0,70	0,60	0,50	← p

Tabelas reproduzidas de "TABUAS DE ESTATÍSTICA E MATEMÁTICA" de W.O.Bussab e J.S.C.Pereira,sob permissão da Editora Brasiliense S.A. (Tabelas 1, 2, 3 e 4).

TABELA 1 - PROBABILIDADES BINOMIAIS (continuação)

p →	0,05	0,10	0,20	0,25	0,30	0,40	0,50	n=5
x=0	774	590	328	237	168	078	031	5
1	204	328	410	396	360	259	156	4
2	021	073	205	264	309	346	312	3
3	001	008	051	088	132	230	312	2
4	0+	0+	006	015	028	077	156	1
5	0+	0+	0+	001	002	010	031	0=x
n=5	0,95	0,90	0,80	0,75	0,70	0,60	0,50	← p

p →	0,05	0,10	0,20	0,25	0,30	0,40	0,50	n=6
x=0	735	531	262	178	118	047	016	6
1	232	354	393	356	303	187	094	5
2	031	098	246	297	324	311	234	4
3	002	015	082	132	185	276	312	3
4	0+	001	015	033	060	138	234	2
5	0+	0+	002	004	010	037	094	1
6	0+	0+	0+	0+	001	004	016	0=x
n=6	0,95	0,90	0,80	0,75	0,70	0,60	0,50	← p

p →	0,05	0,10	0,20	0,25	0,30	0,40	0,50	n=7
x=0	698	478	210	133	082	028	008	7
1	257	372	367	311	247	131	055	6
2	041	124	275	311	318	261	164	5
3	004	023	115	173	227	290	273	4
4	0+	003	029	058	097	194	273	3
5	0+	0+	004	012	025	077	164	2
6	0+	0+	0+	001	004	017	055	1
7	0+	0+	0+	0+	0+	002	008	0
n=7	0,95	0,90	0,80	0,75	0,70	0,60	0,50	← p

p →	0,05	0,10	0,20	0,25	0,30	0,40	0,50	n=28
x=0	238	052	002	0+	0+	0+	0+	28
1	350	163	014	003	[illegible]	0+	0+	27
2	249	244	046	013	003	[illegible]	0+	26
3	114	235	099	039	012	001	[illegible]	25
4	037	163	155	080	032	002	0+	24
5	009	087	186	128	085	008	0+	23
6	002	037	178	164	107	020	001	22
7	0+	013	140	172	145	043	004	21
8	0+	004	092	150	163	074	012	20
9	0+	001	051	111	156	110	026	19
10	0+	0+	024	071	126	140	049	18
11	0+	0+	010	038	088	152	080	17
12	0+	0+	004	018	054	144	113	16
13	0+	0+	001	007	026	118	139	15
14	0+	0+	0+	003	013	084	149	14
15	0+	0+	0+	001	005	053	139	13
16	0+	0+	0+	0+	002	028	113	12
17	0+	0+	0+	0+	001	013	080	11
18	0+	0+	0+	0+	0+	005	049	10
19	0+	0+	0+	0+	0+	002	026	9
20	0+	0+	0+	0+	0+	001	012	8
21	0+	0+	0+	0+	0+	0+	004	7
22	0+	0+	0+	0+	0+	0+	001	6
23	0+	0+	0+	0+	0+	0+	0+	5
24	0+	0+	0+	0+	0+	0+	0+	4
25	0+	0+	0+	0+	0+	0+	0+	3
26	0+	0+	0+	0+	0+	0+	0+	2
27	0+	0+	0+	0+	0+	0+	0+	1
28	0+	0+	0+	0+	0+	0+	0+	0=x
n=28	0,95	0,90	0,80	0,75	0,70	0,60	0,50	← p

p →	0,05	0,10	0,20	0,25	0,30	0,40	0,50	n=27
x=0	250	058	002	0+	0+	0+	0+	27
1	356	174	016	004	001	0+	0+	26
2	243	252	053	017	004	0+	0+	25
3	107	233	111	046	015	001	0+	24
4	034	156	166	092	039	004	0+	23
5	008	080	191	141	077	011	001	22
6	002	032	175	172	121	027	002	21
7	0+	011	131	172	155	053	007	20
8	0+	003	082	143	166	089	017	19
9	0+	001	043	101	150	125	035	18
10	0+	0+	019	060	116	150	063	17
11	0+	0+	008	031	077	154	097	16
12	0+	0+	003	014	044	137	130	15
13	0+	0+	001	005	022	105	149	14
14	0+	0+	0+	002	009	070	149	13
15	0+	0+	0+	001	003	041	130	12
16	0+	0+	0+	0+	001	020	097	11
17	0+	0+	0+	0+	0+	009	063	10
18	0+	0+	0+	0+	0+	003	035	9
19	0+	0+	0+	0+	0+	001	017	8
20	0+	0+	0+	0+	0+	0+	007	7
21	0+	0+	0+	0+	0+	0+	002	6
22	0+	0+	0+	0+	0+	0+	001	5
23	0+	0+	0+	0+	0+	0+	0+	4
24	0+	0+	0+	0+	0+	0+	0+	3
25	0+	0+	0+	0+	0+	0+	0+	2
26	0+	0+	0+	0+	0+	0+	0+	1
27	0+	0+	0+	0+	0+	0+	0+	0=x
n=27	0,95	0,90	0,80	0,75	0,70	0,60	0,50	← p

p →	0,05	0,10	0,20	0,25	0,30	0,40	0,50	n=26
x=0	264	065	003	001	0+	0+	0+	26
1	361	187	020	005	001	0+	0+	25
2	237	259	061	020	006	0+	0+	24
3	100	230	123	054	019	001	0+	23
4	030	147	176	104	047	005	0+	22
5	007	072	194	153	089	015	001	21
6	001	028	170	178	134	034	003	20
7	0+	009	121	170	164	066	010	19
8	0+	002	072	134	167	104	023	18
9	0+	001	036	090	143	139	047	17
10	0+	0+	015	051	104	157	079	16
11	0+	0+	006	025	065	152	115	15
12	0+	0+	002	010	035	127	144	14
13	0+	0+	0+	004	016	091	155	13
14	0+	0+	0+	001	006	056	144	12
15	0+	0+	0+	0+	002	030	115	11
16	0+	0+	0+	0+	001	014	079	10
17	0+	0+	0+	0+	0+	005	047	9
18	0+	0+	0+	0+	0+	002	023	8
19	0+	0+	0+	0+	0+	001	010	7
20	0+	0+	0+	0+	0+	0+	003	6
21	0+	0+	0+	0+	0+	0+	001	5
22	0+	0+	0+	0+	0+	0+	0+	4
23	0+	0+	0+	0+	0+	0+	0+	3
24	0+	0+	0+	0+	0+	0+	0+	2
25	0+	0+	0+	0+	0+	0+	0+	1
26	0+	0+	0+	0+	0+	0+	0	0=x
n=26	0,95	0,90	0,80	0,75	0,70	0,60	0,50	← p

TABELA 1 - PROBABILIDADES BINOMIAIS (continuação)

p →	0,05	0,10	0,20	0,25	0,30	0,40	0,50	n=8
x=0	663	430	168	100	058	017	004	8
1	279	383	336	267	198	090	031	7
2	051	149	294	311	296	209	109	6
3	005	033	147	208	254	279	219	5
4	0*	005	046	087	136	232	273	4
5	0*	0*	009	023	047	124	219	3
6	0*	0*	001	004	010	041	109	2
7	0*	0*	0*	0*	001	008	031	1
8	0*	0*	0*	0*	0*	001	004	0=x
n=8	0,95	0,90	0,80	0,75	0,70	0,60	0,50	← p

p →	0,05	0,10	0,20	0,25	0,30	0,40	0,50	n=9
x=0	630	387	134	075	040	010	002	9
1	299	387	302	225	156	060	018	8
2	063	172	302	300	267	161	070	7
3	008	045	176	234	267	251	164	6
4	001	007	066	117	172	251	246	5
5	0*	001	017	039	074	167	246	4
6	0*	0*	003	009	021	074	164	3
7	0*	0*	0*	001	004	021	070	2
8	0*	0*	0*	0*	0*	004	018	1
9	0*	0*	0*	0*	0*	0*	002	0=x
n=9	0,95	0,90	0,80	0,75	0,70	0,60	0,50	← p

p →	0,05	0,10	0,20	0,25	0,30	0,40	0,50	n=10
x=0	599	349	107	056	028	006	001	10
1	315	387	268	188	121	040	010	9
2	075	194	302	282	233	121	044	8
3	010	057	201	250	267	215	117	7
4	001	011	088	146	200	251	205	6
5	0*	001	026	058	103	201	246	5
6	0*	0*	006	016	037	111	205	4
7	0*	0*	001	003	009	042	117	3
8	0*	0*	0*	0*	001	011	044	2
9	0*	0*	0*	0*	0*	002	010	1
10	0*	0*	0*	0*	0*	0*	001	0=x
n=10	0,95	0,90	0,80	0,75	0,70	0,60	0,50	← p

p →	0,05	0,10	0,20	0,25	0,30	0,40	0,50	n=25
x=0	277	072	004	001	0*	0*	0*	25
1	365	199	024	006	001	0*	0*	24
2	231	266	071	025	007	0*	0*	23
3	093	226	136	064	024	002	0*	22
4	027	138	187	118	057	007	0*	21
5	006	065	196	165	103	020	002	20
6	001	024	163	183	147	044	005	19
7	0*	007	111	165	171	080	014	18
8	0*	002	062	124	165	120	032	17
9	0*	0*	029	078	134	151	061	16
10	0*	0*	012	042	092	161	097	15
11	0*	0*	004	019	054	147	133	14
12	0*	0*	001	007	027	114	155	13
13	0*	0*	0*	002	011	076	155	12
14	0*	0*	0*	001	004	043	133	11
15	0*	0*	0*	0*	001	021	097	10
16	0*	0*	0*	0*	0*	009	061	9
17	0*	0*	0*	0*	0*	003	032	8
18	0*	0*	0*	0*	0*	001	014	7
19	0*	0*	0*	0*	0*	0*	005	6
20	0*	0*	0*	0*	0*	0*	002	5
21	0*	0*	0*	0*	0*	0*	0*	4
22	0*	0*	0*	0*	0*	0*	0*	3
23	0*	0*	0*	0*	0*	0*	0*	2
24	0*	0*	0*	0*	0*	0*	0*	1
25	0*	0*	0*	0*	0*	0*	0*	0=x
n=25	0,95	0,90	0,80	0,75	0,70	0,60	0,50	← p

p →	0,05	0,10	0,20	0,25	0,30	0,40	0,50	n=24
x=0	292	080	005	001	0*	0*	0*	24
1	369	213	028	008	002	0*	0*	23
2	223	272	081	031	010	001	0*	22
3	086	221	149	075	031	003	0*	21
4	024	129	196	132	069	010	001	20
5	005	057	196	176	118	027	003	19
6	001	020	155	185	160	056	008	18
7	0*	006	100	159	176	096	021	17
8	0*	001	053	112	160	136	044	16
9	0*	0*	024	067	122	161	078	15
10	0*	0*	009	033	079	161	117	14
11	0*	0*	003	014	043	137	149	13
12	0*	0*	001	005	020	099	161	12
13	0*	0*	0*	002	008	061	149	11
14	0*	0*	0*	0*	003	032	117	10
15	0*	0*	0*	0*	001	014	078	9
16	0*	0*	0*	0*	0*	005	044	8
17	0*	0*	0*	0*	0*	002	021	7
18	0*	0*	0*	0*	0*	0*	008	6
19	0*	0*	0*	0*	0*	0*	003	5
20	0*	0*	0*	0*	0*	0*	001	4
21	0*	0*	0*	0*	0*	0*	0*	3
22	0*	0*	0*	0*	0*	0*	0*	2
23	0*	0*	0*	0*	0*	0*	0*	1
24	0*	0*	0*	0*	0*	0*	0*	0=x
n=24	0,95	0,90	0,80	0,75	0,70	0,60	0,50	← p

p →	0,05	0,10	0,20	0,25	0,30	0,40	0,50	n=23
x=0	307	089	006	001	0*	0*	0*	23
1	372	226	034	010	003	0*	0*	22
2	215	277	093	038	013	001	0*	21
3	079	215	163	068	038	004	0*	20
4	021	120	204	146	082	014	001	19
5	004	051	194	185	133	035	004	18
6	001	017	145	185	171	070	012	17
7	0*	005	088	150	178	113	029	16
8	0*	001	044	100	153	151	058	15
9	0*	0*	018	056	109	168	097	14
10	0*	0*	006	026	065	157	136	13
11	0*	0*	002	010	033	123	161	12
12	0*	0*	0*	003	014	062	161	11
13	0*	0*	0*	001	005	046	136	10
14	0*	0*	0*	0*	002	022	097	9
15	0*	0*	0*	0*	0*	009	058	8
16	0*	0*	0*	0*	0*	003	029	7
17	0*	0*	0*	0*	0*	001	012	6
18	0*	0*	0*	0*	0*	0*	004	5
19	0*	0*	0*	0*	0*	0*	001	4
20	0*	0*	0*	0*	0*	0*	0*	3
21	0*	0*	0*	0*	0*	0*	0*	2
22	0*	0*	0*	0*	0*	0*	0*	1
23	0*	0*	0*	0*	0*	0*	0*	0=x
n=23	0,95	0,90	0,80	0,75	0,70	0,60	0,50	← p

TABELA 1 - PROBABILIDADES BINOMIAIS (continuação)

p →	0,05	0,10	0,20	0,25	0,30	0,40	0,50	n=11
x=0	569	314	086	042	020	004	0*	11
1	329	384	236	155	093	027	005	10
2	087	213	295	258	200	089	027	9
3	014	071	221	258	257	177	081	8
4	001	016	111	172	220	236	161	7
5	0*	002	039	080	132	221	226	6
6	0*	0*	010	027	057	147	226	5
7	0*	0*	002	006	017	070	161	4
8	0*	0*	0*	001	004	023	081	3
9	0*	0*	0*	0*	001	005	027	2
10	0*	0*	0*	0*	0*	001	005	1
11	0*	0*	0*	0*	0*	0*	0*	0=x
n=11	0,95	0,90	0,80	0,75	0,70	0,40	0,50	← p

p →	0,05	0,10	0,20	0,25	0,30	0,40	0,50	n=12
x=0	540	282	069	032	014	002	0*	12
1	341	377	206	127	071	017	003	11
2	099	230	283	232	168	064	016	10
3	017	085	236	258	240	142	054	9
4	002	021	133	194	231	213	121	8
5	0*	004	053	103	158	227	193	7
6	0*	0*	016	040	079	177	226	6
7	0*	0*	003	011	029	101	193	5
8	0*	0*	001	002	008	042	121	4
9	0*	0*	0*	0*	001	012	054	3
10	0*	0*	0*	0*	0*	002	016	2
11	0*	0*	0*	0*	0*	0*	003	1
12	0*	0*	0*	0*	0*	0*	0*	0=x
n=12	0,95	0,90	0,80	0,75	0,70	0,60	0,50	← p

p →	0,05	0,10	0,20	0,25	0,30	0,40	0,50	n=13
x=0	513	254	055	024	010	001	0*	13
1	351	367	179	103	054	011	002	12
2	111	245	268	206	139	045	010	11
3	021	100	246	252	218	111	035	10
4	003	028	154	210	234	184	087	9
5	0*	006	069	126	180	221	157	8
6	0*	001	023	056	103	197	209	7
7	0*	0*	006	019	044	131	209	6
8	0*	0*	001	005	014	066	157	5
9	0*	0*	0*	001	003	024	087	4
10	0*	0*	0*	0*	001	006	035	3
11	0*	0*	0*	0*	0*	001	010	2
12	0*	0*	0*	0*	0*	0*	002	1
13	0*	0*	0*	0*	0*	0*	0*	0=x
n=13	0,95	0,90	0,80	0,75	0,70	0,60	0,50	← p

p →	0,05	0,10	0,20	0,25	0,30	0,40	0,50	n=22
x=0	324	098	007	002	0*	0*	0*	22
1	375	241	041	013	004	0*	0*	21
2	207	281	107	046	017	001	0*	20
3	073	208	178	102	047	008	0*	19
4	018	110	211	161	096	019	002	18
5	003	044	190	193	149	046	006	17
6	001	014	134	183	181	086	018	16
7	0*	004	077	139	177	131	041	15
8	0*	001	036	087	142	164	076	14
9	0*	0	014	045	095	170	119	13
10	0*	0*	005	020	053	148	154	12
11	0*	0*	001	007	025	107	168	11
12	0*	0*	0*	002	010	066	154	10
13	0*	0*	0*	001	003	034	119	9
14	0*	0*	0*	0*	001	014	076	8
15	0*	0*	0*	0*	0*	005	041	7
16	0*	0*	0*	0*	0*	001	018	6
17	0*	0*	0*	0*	0*	0*	006	5
18	0*	0*	0*	0*	0*	0*	002	4
19	0*	0*	0*	0*	0*	0*	0*	3
20	0*	0*	0*	0*	0*	0*	0*	2
21	0*	0*	0*	0*	0*	0*	0*	1
22	0*	0*	0*	0*	0*	0*	0*	0=x
n=22	0,95	0,90	0,80	0,75	0,70	0,60	0,50	← p

p →	0,05	0,10	0,20	0,25	0,30	0,40	0,50	n=21
x=0	341	109	009	002	001	0*	0*	21
1	376	255	048	017	005	0*	0*	20
2	198	284	121	055	022	002	0*	19
3	066	200	192	117	058	009	001	18
4	016	100	216	176	113	026	003	17
5	003	038	183	199	164	059	010	16
6	0*	011	122	177	188	105	026	15
7	0*	003	065	126	172	149	055	14
8	0*	001	029	074	129	174	097	13
9	0*	0*	010	036	080	168	140	12
10	0*	0*	003	014	041	134	168	11
11	0*	0*	001	005	018	089	168	10
12	0*	0*	0*	001	006	050	140	9
13	0*	0*	0*	0*	002	023	097	8
14	0*	0*	0*	0*	0*	009	055	7
15	0*	0*	0*	0*	0*	003	026	6
16	0*	0*	0*	0*	0*	001	010	5
17	0*	0*	0*	0*	0*	0*	003	4
18	0*	0*	0*	0*	0*	0*	001	3
19	0*	0*	0*	0*	0*	0*	0*	2
20	0*	0*	0*	0*	0*	0*	0*	1
21	0*	0*	0*	0*	0*	0*	0*	0=x
n=21	0,95	0,90	0,80	0,75	0,70	0,60	0,50	← p

p →	0,05	0,10	0,20	0,25	0,30	0,40	0,50	n=20
x=0	358	122	012	003	001	0*	0*	20
1	377	270	058	021	007	0*	0*	19
2	189	285	137	067	028	003	0*	18
3	060	190	205	134	072	012	001	17
4	013	090	218	190	130	035	005	16
5	002	032	175	202	179	075	015	15
6	0*	009	109	169	192	124	037	14
7	0*	002	055	112	164	166	074	13
8	0*	0*	022	061	114	180	120	12
9	0*	0*	007	027	065	160	160	11
10	0*	0*	002	010	031	117	176	10
11	0*	0*	0*	003	012	071	160	9
12	0*	0*	0*	001	004	035	120	8
13	0*	0*	0*	0*	001	015	074	7
14	0*	0*	0*	0*	0*	005	037	6
15	0*	0*	0*	0*	0*	001	015	5
16	0*	0*	0*	0*	0*	0*	005	4
17	0*	0*	0*	0*	0*	0*	001	3
18	0*	0*	0*	0*	0*	0*	0*	2
19	0*	0*	0*	0*	0*	0*	0*	1
20	0*	0*	0*	0*	0*	0*	0*	0=x
n=20	0,95	0,90	0,80	0,75	0,70	0,60	0,50	← p

TABELA 1 - PROBABILIDADES BINOMIAIS (conclusão)

p →	0,05	0,10	0,20	0,25	0,30	0,40	0,50	n=14
x=0	488	229	044	018	007	001	0*	14
1	359	356	154	083	041	007	001	13
2	123	257	250	180	113	032	006	12
3	026	114	250	240	194	085	022	11
4	004	035	172	220	229	155	061	10
5	0*	008	086	147	196	207	122	9
6	0*	001	032	073	126	207	183	8
7	0*	0*	009	028	062	157	209	7
8	0*	0*	002	008	023	092	183	6
9	0*	0*	0*	002	007	041	122	5
10	0*	0*	0*	0*	001	014	061	4
11	0*	0*	0*	0*	0*	003	022	3
12	0*	0*	0*	0*	0*	001	006	2
13	0*	0*	0*	0*	0*	0*	001	1
14	0*	0*	0*	0*	0*	0*	0*	0=x
n=14	0,95	0,90	0,80	0,75	0,70	0,60	0,50	← p

p →	0,05	0,10	0,20	0,25	0,30	0,40	0,50	n=15
x=0	463	206	035	013	005	0*	0*	15
1	366	343	132	067	031	005	0*	14
2	135	267	231	156	092	022	003	13
3	031	129	250	225	170	063	014	12
4	005	043	188	225	219	127	042	11
5	001	010	103	165	206	186	092	10
6	0*	002	043	092	147	207	153	9
7	0*	0*	014	039	081	177	196	8
8	0*	0*	003	013	035	118	196	7
9	0*	0*	001	003	012	061	153	6
10	0*	0*	0*	001	003	024	092	5
11	0*	0*	0*	0*	001	007	042	4
12	0*	0*	0*	0*	0*	002	014	3
13	0*	0*	0*	0*	0*	0*	003	2
14	0*	0*	0*	0*	0*	0*	0*	1
15	0*	0*	0*	0*	0*	0*	0*	0=x
n=15	0,95	0,90	0,80	0,75	0,70	0,60	0,50	← p

p →	0,05	0,10	0,20	0,25	0,30	0,40	0,50	n=16
x=0	440	185	028	010	003	0*	0*	16
1	371	329	113	053	023	003	0*	15
2	146	275	211	134	073	015	002	14
3	036	142	246	208	146	047	009	13
4	006	051	200	225	204	101	028	12
5	001	014	120	180	210	162	067	11
6	0*	003	055	110	165	198	122	10
7	0*	0*	020	052	101	189	175	9
8	0*	0*	006	020	049	142	196	8
9	0*	0*	001	006	019	084	175	7
10	0*	0*	0*	001	006	039	122	6
11	0*	0*	0*	0*	001	014	067	5
12	0*	0*	0*	0*	0*	004	028	4
13	0*	0*	0*	0*	0*	001	009	3
14	0*	0*	0*	0*	0*	0*	002	2
15	0*	0*	0*	0*	0*	0*	0*	1
16	0*	0*	0*	0*	0*	0*	0*	0=x
n=16	0,95	0,90	0,80	0,75	0,70	0,60	0,50	← p

p →	0,05	0,10	0,20	0,25	0,30	0,40	0,50	n=19
x=0	377	135	014	004	001	0*	0*	19
1	377	285	068	027	009	001	0*	18
2	179	285	154	080	036	005	0*	17
3	053	180	218	152	087	017	002	16
4	011	080	218	202	149	047	007	15
5	002	027	164	202	192	093	022	14
6	0*	007	095	157	192	145	052	13
7	0*	001	044	097	153	180	096	12
8	0*	0*	017	049	098	180	144	11
9	0*	0*	005	020	051	146	176	10
10	0*	0*	001	007	022	098	176	9
11	0*	0*	0*	002	008	053	144	8
12	0*	0*	0*	0*	002	024	096	7
13	0*	0*	0*	0*	001	008	052	6
14	0*	0*	0*	0*	0*	002	022	5
15	0*	0*	0*	0*	0*	001	007	4
16	0*	0*	0*	0*	0*	0*	002	3
17	0*	0*	0*	0*	0*	0*	0*	2
18	0*	0*	0*	0*	0*	0*	0*	1
19	0*	0*	0*	0*	0*	0*	0*	0=x
n=19	0,95	0,90	0,80	0,75	0,70	0,60	0,50	← p

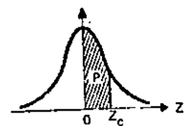
p →	0,05	0,10	0,20	0,75	0,30	0,40	0,50	n=18
x=0	397	150	018	006	002	0*	0*	18
1	376	300	081	034	013	001	0*	17
2	168	284	172	096	046	007	001	16
3	047	168	230	170	105	025	003	15
4	009	070	215	213	168	061	012	14
5	001	022	151	199	202	115	033	13
6	0*	005	082	144	187	166	071	12
7	0*	001	035	082	138	189	121	11
8	0*	0*	012	038	081	173	167	10
9	0*	0*	003	014	039	128	185	9
10	0	0*	001	004	015	077	167	8
11	0*	0*	0*	001	005	037	121	7
12	0*	0*	0*	0*	001	015	071	6
13	0*	0*	0*	0*	0*	004	033	5
14	0*	0*	0*	0*	0*	001	012	4
15	0*	0*	0*	0*	0*	0*	003	3
16	0*	0*	0*	0*	0*	0*	001	2
17	0*	0*	0*	0*	0*	0*	0*	1
18	0*	0*	0*	0*	0*	0*	0*	0=x
n=18	0,95	0,90	0,80	0,75	0,70	0,60	0,50	← p

p →	0,05	0,10	0,20	0,25	0,30	0,40	0,50	n=17
x=0	418	167	023	008	002	0*	0*	17
1	374	315	096	043	017	002	0*	16
2	158	280	191	114	058	010	001	15
3	041	156	239	189	125	034	005	14
4	008	060	209	221	187	080	018	13
5	001	017	136	191	208	138	047	12
6	0*	004	068	178	178	184	094	11
7	0*	001	027	067	120	193	148	10
8	0*	0*	008	028	064	161	185	9
9	0*	0*	002	009	028	107	185	8
10	0*	0*	0*	002	009	057	148	7
11	0*	0*	0*	001	003	024	094	6
12	0*	0*	0*	0*	001	008	047	5
13	0*	0*	0*	0*	0*	002	018	4
14	0*	0*	0*	0*	0*	0*	005	3
15	0*	0*	0*	0*	0*	0*	001	2
16	0*	0*	0*	0*	0*	0*	0*	1
17	0*	0*	0*	0*	0*	0*	0*	0=x
n=17	0,95	0,90	0,80	0,75	0,70	0,60	0,50	← p

TABELA 2 - ÁREAS SOB A CURVA NORMAL PADRÃO

Exemplo: P(Z>1,2) =

0,5-0,38493 = 0,11507



parte inteira e primeira decimal de Z_c	SEGUNDA DECIMAL DE Z_c										parte inteira e primeira decimal de Z_c
	0	1	2	3	4	5	6	7	8	9	
	p = 0,										
0,0	00000	00399	00798	01197	01595	01994	02392	02790	03188	03586	0,0
0,1	03983	04380	04776	05172	05567	05962	06356	06749	07142	07535	0,1
0,2	07926	08317	08706	09095	09483	09871	10257	10642	11026	11409	0,2
0,3	11791	12172	12552	12930	13307	13683	14058	14431	14803	15173	0,3
0,4	15542	15910	16276	16640	17003	17364	17724	18082	18439	18793	0,4
0,5	19146	19497	19847	20194	20540	20884	21226	21566	21904	22240	0,5
0,6	22575	22907	23237	23565	23891	24215	24537	24857	25175	25490	0,6
0,7	25804	26115	26424	26730	27035	27337	27637	27935	28230	28524	0,7
0,8	28814	29103	29389	29673	29955	30234	30511	30785	31057	31327	0,8
0,9	31594	31859	32121	32381	32639	32894	33147	33398	33646	33891	0,9
1,0	34134	34375	34614	34850	35083	35314	35543	35769	35993	36214	1,0
1,1	36433	36650	36864	37076	37286	37493	37698	37900	38100	38298	1,1
1,2	38493	38686	38877	39065	39251	39435	39617	39796	39973	40147	1,2
1,3	40320	40490	40658	40824	40988	41149	41309	41466	41621	41774	1,3
1,4	41924	42073	42220	42364	42507	42647	42786	42922	43056	43189	1,4
1,5	43319	43448	43574	43699	43822	43943	44062	44179	44295	44408	1,5
1,6	44520	44630	44738	44845	44950	45053	45154	45254	45352	45449	1,6
1,7	45543	45637	45728	45818	45907	45994	46080	46164	46246	46327	1,7
1,8	46407	46485	46562	46638	46712	46784	46856	46926	46995	47062	1,8
1,9	47128	47193	47257	47320	47381	47441	47500	47558	47615	47670	1,9
2,0	47725	47778	47831	47882	47932	47982	48030	48077	48124	48169	2,0
2,1	48214	48257	48300	48341	48382	48422	48461	48500	48537	48574	2,1
2,2	48610	48645	48679	48713	48745	48778	48809	48840	48870	48899	2,2
2,3	48928	48956	48983	49010	49036	49061	49086	49111	49134	49158	2,3
2,4	49180	49202	49224	49245	49266	49286	49305	49324	49343	49361	2,4
2,5	49379	49396	49413	49430	49446	49461	49477	49492	49506	49520	2,5
2,6	49534	49547	49560	49573	49585	49598	49609	49621	49632	49643	2,6
2,7	49653	49664	49674	49683	49693	49702	49711	49720	49728	49736	2,7
2,8	49744	49752	49760	49767	49774	49781	49788	49795	49801	49807	2,8
2,9	49813	49819	49825	49831	49836	49841	49846	49851	49856	49861	2,9
3,0	49865	49869	49874	49878	49882	49886	49889	49893	49897	49900	3,0
3,1	49903	49906	49910	49913	49916	49918	49921	49924	49926	49929	3,1
3,2	49931	49934	49936	49938	49940	49942	49944	49946	49948	49950	3,2
3,3	49952	49953	49955	49957	49958	49960	49961	49962	49964	49965	3,3
3,4	49966	49968	49969	49970	49971	49972	49973	49974	49975	49976	3,4
3,5	49977	49978	49978	49979	49980	49981	49981	49982	49983	49983	3,5
3,6	49984	49985	49985	49986	49986	49987	49987	49988	49988	49989	3,6
3,7	49989	49990	49990	49990	49991	49991	49992	49992	49992	49992	3,7
3,8	49993	49993	49993	49994	49994	49994	49994	49995	49995	49995	3,8
3,9	49995	49995	49996	49996	49996	49996	49998	49996	49997	49997	3,9
4,0	49997	49997	49997	49997	49997	49997	49998	49998	49998	49998	4,0
4,5	49999	50000	50000	50000	50000	50000	50000	50000	50000	50000	4,5

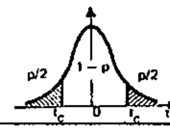
parte inteira e primeira decimal de Z_c	SEGUNDA E TERCEIRA DECIMAIS DE Z_c										parte inteira e primeira decimal de Z_c
	05	15	25	35	45	55	65	75	85	95	
	p = 0,										
0,0	00199	00598	00997	01396	01795	02193	02591	02989	03387	03784	0,0
0,1	04181	04578	04974	05369	05764	06159	06553	06946	07339	07730	0,1
0,2	08121	08512	08901	09290	09677	10064	10450	10834	11218	11600	0,2
0,3	11982	12362	12741	13119	13495	13871	14244	14617	14988	15358	0,3
0,4	15726	16093	16453	16822	17184	17545	17903	18261	18500	18970	0,4
0,5	19322	19672	20021	20368	20712	21055	21396	21735	22073	22408	0,5
0,6	22741	23072	23401	23729	24054	24377	24697	25016	25333	25647	0,6
0,7	25959	26270	26577	26883	27186	27488	27786	28083	28377	28669	0,7
0,8	28959	29246	29531	29814	30094	30372	30648	30921	31192	31461	0,8
0,9	31727	31990	32262	32511	32767	33021	33273	33522	33769	34013	0,9
	05	15	25	35	45	55	65	75	85	95	

TABELA 3 - ÁREAS SOB A DENSIDADE t

Exemplo: $\nu = 24$,

$P(t > 1{,}318) = 0{,}10$,

$P(t < -1{,}318) = 0{,}10$



ν	90%	80%	70%	60%	50%	40%	30%	20%	10%	5%	4%	2%	1%	0,2%	0,1%	ν
1	0,158	0,325	0,510	0,727	1,000	1,376	1,963	3,078	6,314	12,706	15,894	31,821	63,657	318,309	636,619	1
2	0,142	0,289	0,445	0,617	0,816	1,061	1,386	1,886	2,920	4,303	4,849	6,965	9,925	22,327	31,598	2
3	0,137	0,277	0,424	0,584	0,765	0,978	1,250	1,638	2,353	3,182	3,482	4,541	5,841	10,214	12,924	3
4	0,134	0,271	0,414	0,569	0,741	0,941	1,190	1,533	2,132	2,776	2,998	3,747	4,604	7,173	8,610	4
5	0,132	0,267	0,408	0,559	0,727	0,920	1,156	1,476	2,015	2,571	2,756	3,365	4,032	5,893	6,869	5
6	0,131	0,265	0,404	0,553	0,718	0,906	1,134	1,440	1,943	2,447	2,612	3,143	3,707	5,208	5,959	6
7	0,130	0,263	0,402	0,549	0,711	0,896	1,119	1,415	1,895	2,365	2,517	2,998	3,499	4,785	5,408	7
8	0,130	0,262	0,399	0,546	0,706	0,889	1,108	1,397	1,860	2,306	2,449	2,896	3,355	4,501	5,041	8
9	0,129	0,261	0,398	0,543	0,703	0,883	1,100	1,383	1,833	2,262	2,398	2,821	3,250	4,297	4,781	9
10	0,129	0,260	0,397	0,542	0,700	0,879	1,093	1,372	1,812	2,228	2,359	2,764	3,169	4,144	4,587	10
11	0,129	0,260	0,396	0,540	0,697	0,876	1,088	1,363	1,796	2,201	2,328	2,718	3,106	3,025	4,437	11
12	0,128	0,259	0,395	0,539	0,695	0,873	1,083	1,356	1,782	2,179	2,303	2,681	3,055	3,930	4,318	12
13	0,128	0,259	0,394	0,538	0,694	0,870	1,079	1,350	1,771	2,160	2,282	2,650	3,012	3,852	4,221	13
14	0,128	0,258	0,393	0,537	0,692	0,868	1,076	1,345	1,761	2,145	2,264	2,624	2,977	3,787	4,140	14
15	0,128	0,258	0,393	0,536	0,691	0,866	1,074	1,341	1,753	2,131	2,248	2,602	2,947	3,733	4,073	15
16	0,128	0,258	0,392	0,535	0,690	0,865	1,071	1,337	1,746	2,120	2,235	2,583	2,921	3,686	4,015	16
17	0,128	0,257	0,392	0,534	0,689	0,863	1,069	1,333	1,740	2,110	2,224	2,567	2,898	3,646	3,965	17
18	0,127	0,257	0,392	0,534	0,688	0,862	1,067	1,330	1,734	2,101	2,214	2,552	2,878	3,610	3,922	18
19	0,127	0,257	0,391	0,533	0,688	0,861	1,066	1,328	1,729	2,093	2,205	2,539	2,861	3,579	3,883	19
20	0,127	0,257	0,391	0,533	0,687	0,860	1,064	1,325	1,725	2,086	2,197	2,528	2,845	3,552	3,850	20
21	0,127	0,257	0,391	0,532	0,686	0,859	1,063	1,323	1,721	2,080	2,189	2,518	2,831	3,527	3,819	21
22	0,127	0,256	0,390	0,532	0,686	0,858	1,061	1,321	1,717	2,074	2,183	2,508	2,819	3,505	3,792	22
23	0,127	0,256	0,390	0,532	0,685	0,858	1,060	1,319	1,714	2,069	2,177	2,500	2,807	3,485	3,768	23
24	0,127	0,256	0,390	0,531	0,685	0,857	1,059	1,318	1,711	2,064	2,172	2,492	2,797	3,467	3,745	24
25	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,316	1,708	2,060	2,165	2,485	2,787	3,450	3,725	25
26	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,315	1,706	2,056	2,162	2,479	2,779	3,435	3,707	26
27	0,127	0,256	0,389	0,531	0,684	0,855	1,057	1,314	1,703	2,052	2,158	2,473	2,771	3,421	3,690	27
28	0,127	0,256	0,389	0,530	0,684	0,855	1,056	1,313	1,701	2,048	2,154	2,467	2,763	3,408	3,674	28
29	0,127	0,256	0,389	0,530	0,683	0,854	1,055	1,311	1,699	2,045	2,150	2,462	2,756	3,396	3,659	29
30	0,127	0,256	0,389	0,530	0,683	0,854	1,055	1,310	1,697	2,042	2,147	2,457	2,750	3,385	3,646	30
35	0,126	0,255	0,388	0,529	0,682	0,852	1,052	1,306	1,690	2,030	2,133	2,438	2,724	3,340	3,591	35
40	0,126	0,255	0,388	0,529	0,681	0,851	1,050	1,303	1,684	2,021	2,123	2,423	2,704	3,307	3,551	40
50	0,126	0,254	0,387	0,528	0,679	0,849	1,047	1,299	1,676	2,009	2,109	2,403	2,678	3,261	3,496	50
60	0,126	0,254	0,387	0,527	0,679	0,848	1,045	1,296	1,671	2,000	2,099	2,390	2,660	3,232	3,460	60
120	0,126	0,254	0,386	0,526	0,677	0,845	1,041	1,289	1,658	1,980	2,076	2,358	2,617	3,160	3,373	120
∞	0,126	0,253	0,385	0,524	0,674	0,842	1,036	1,282	1,645	1,960	2,054	2,326	2,576	3,090	3,291	∞
	$\rho =$ 90%	80%	70%	60%	50%	40%	30%	20%	10%	5%	4%	2%	1%	0,2%	0,1%	

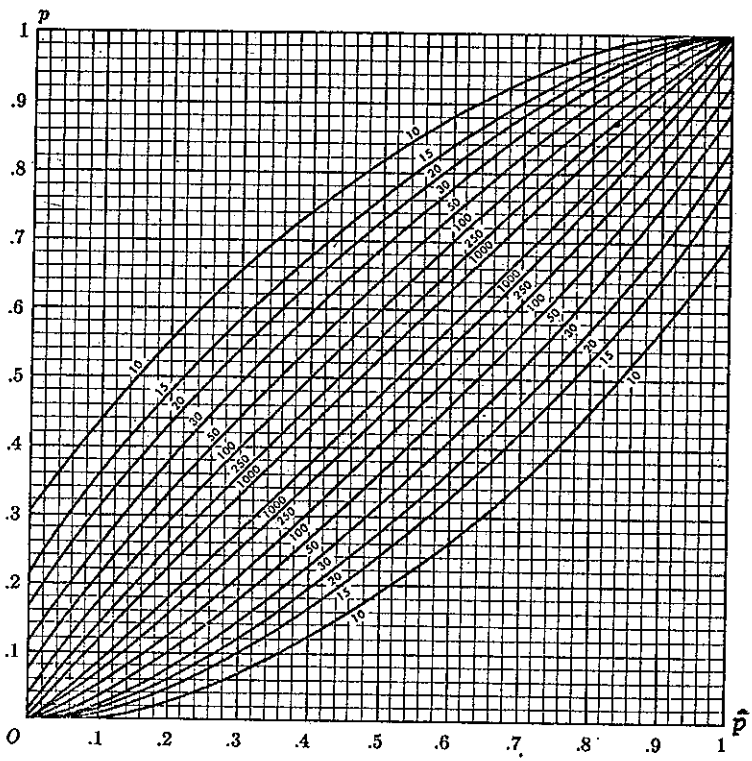
TABELA 4 - PROBABILIDADES DE POISSON: $P(X=r) = \frac{e^{-\alpha} \cdot \alpha^{r}}{r!}$

X \ α	0,001	0,005	0,010	0,015	0,020	0,025	0,030	0,035	0,040	0,045	0,050	0,055	0,060	0,065	0,070	0,075	α / X
0	9990	9950	9900	9851	9802	9753	9704	9656	9608	9560	9512	9465	9418	9371	9324	9277	0
1	0010	0050	0099	0148	0196	0244	0291	0338	0384	0430	0476	0521	0565	0609	0653	0696	1
2	0+	0+	0+	0001	0002	0003	0004	0006	0008	0010	0012	0014	0017	0020	0022	0026	2
3	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0001	0001	3
≥ 4	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	≥ 4

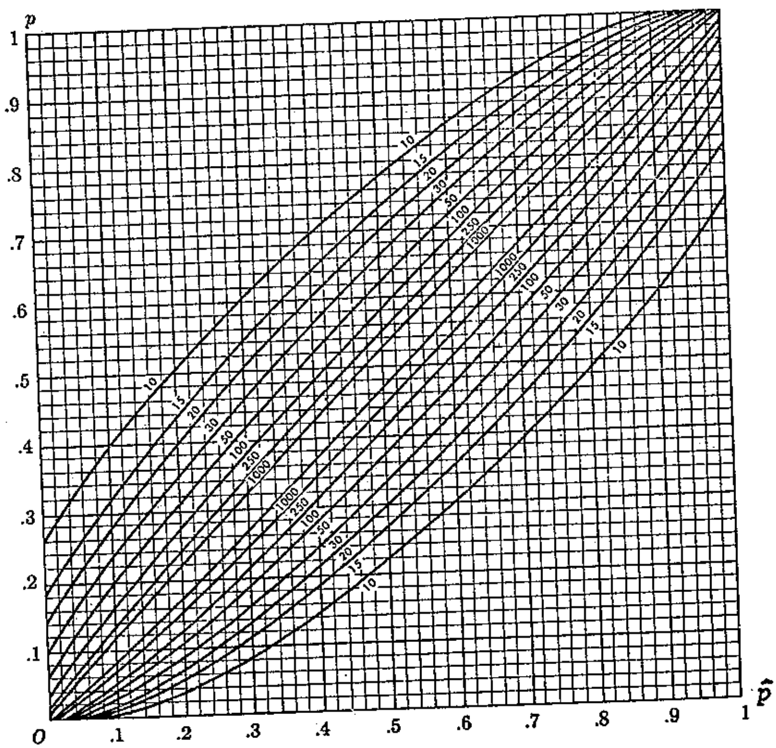
X \ α	0,080	0,085	0,090	0,095	0,100	0,200	0,300	0,400	0,500	0,600	0,700	0,800	0,900	1,000	1,200	1,400	α / X
0	9231	9185	,9139	9094	9048	8187	7408	6703	6065	5488	4966	4493	4066	3679	3012	2466	0
1	0739	0781	0823	0864	0905	1637	2222	2681	3033	3293	3476	3595	3659	3679	3614	3452	1
2	0030	0033	0037	0041	0045	0164	0333	0536	0758	0988	1217	1438	1647	1839	2169	2417	2
3	0001	0001	0001	0001	0002	0011	0033	0071	0126	0198	0284	0383	0494	0613	0867	1128	3
4	0+	0+	0+	0+	0+	0001	0003	0007	0016	0030	0050	0077	0111	0153	0260	0395	4
5	0+	0+	0+	0+	0+	0+	0+	0001	0002	0004.	0007	0012	0020	0031	0062	0111	5
6	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0001	0002	0003	0005	0012	0026	6
7	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0001	0002	0005	7
8	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0001	8
≥ 9	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	≥ 9

X \ α	1,600	1,800	2,000	2,500	3,000	3,500	4,000	4,500	5,000	5,500	6,000	6,500	7,000	8,000	9,000	10,00	α / X
0	2019	1653	1353	0821	0498	0302	0183	0111	0067	0041	0025	0015	0009	0003	0001	0+	0
1	3230	2975	2707	2052	1494	1057	0733	0500	0337	0225	0149	0098	0064	0027	0011	0005	1
2	2584	2678	2707	2565	2240	1850	1465	1125	0842	0618	0446	0318	0223	0107	0050	0023	2
3	1378	1607	1804	2138	2240	2158	1954	1687	1404	1133	0892	0688	0521	0286	0150	0076	3
4	0551	0723	0902	1336	1680	1888	1954	1898	1755	1551	1339	1118	0912	0573	0337	0189	4
5	0176,	0260	0361	0668	1008	1322	1563	1708	1755	1714	1606	1454	1277	0916	0607	0378	5
6	0047	0078	0120	0278	0504	0771	1042	1281	1462	1571	1606	1575	1490	1221	0911	0631	6
7	0011	0020	0034	0100	0216	0385	0595	0824	1044	1234	1377	1462	1490	1396	1171	0901	7
8	0002	0005	0009	0031	0081	0169	0298	0463	0653	0849	1033	1188	1304	1396	1318	1126	8
9	0+	0001	0002	0009	0027	0066	0132	0232	0363	0519	0688	0858	1014	1241	1318	1251	9
10	0+	0+	0+	0002	0008	0023	0053	0104	0181	0285	0413	0558	0710	0993	1186	1251	0
11	0+	0+	0+	0+	0002	0007	0019	0043	0082	0143	0225	0330	0452	0722	0970	1137	11
12	0+	0+	0+	0+	0+	0002	0006	0016	0034	0065	0113	0179	0264	0481	0728	0948	12
13	0+	0+	0+	0+	0+	0+	0002	0006	0013	0028	0052	0089	0142	0296	0504	0729	13
14	0+	0+	0+	0+	0+	0+	0+	0002	0005	0011	0022	0041	0071	0169	0324	0521	14
15	0+	0+	0+	0+	0+	0+	0+	0001	0002	0004	0009	0018	0033	0090	0194	0347	15
16	0+	0+	0+	0+	0+	0+	0+	0+	0+	0001	0003	0007	0014	0045	0109	0217	16
17	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0001	0003	0006	0021	0058	0128	17
18	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0001	0002	0009	0029	0071	18
19	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0001	0004	0014	0037	19
20	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0002	0006	0019	20
21	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0001	0003	0009	21
22	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0001	0004	22
23	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0002	23
24	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0001	24
≥25	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	≥25

REGIÕES DE CONFIANÇA PARA $p(\alpha \leq 0,05)$



REGIÕES DE CONFIANÇA PARA p($\alpha \leq 0,10$)



APÊNDICE 2

A.2.1 - DESIGUALDADE DE CHEBYSHEV

Seja X uma v.a. com $E(X) = \mu$ e $Var(X) = \sigma^2$ finita. Então, para todo $k > 0$, temos

$$P(|X-\mu| \geq k) \leq \frac{Var(X)}{k^2} \qquad (A.2.1)$$

DEMONSTRAÇÃO:

$$P(|X-\mu| \geq k) = \sum_{|x_j-\mu| \geq k} p(x_j),$$

onde $p(x_j) = P(X=x_j)$ e a soma é estendida a todos os x_j tais que $|x_j-\mu| \geq k$. Como $|x_j-\mu| \geq k$ se e somente se $(x_j-\mu)^2 \geq k^2$, temos que

$$\sum_{|x_j-\mu| \geq k} p(x_j) = \sum_{(x_j-\mu)^2 \geq k^2} p(x_j) \leq k^{-2} \sum_{(x_j-\mu)^2 \geq k^2} (x_j-\mu)^2 p(x_j)$$

$$\leq k^{-2} \sum_j (x_j-\mu)^2 p(x_j) = k^{-2}. \, Var(X),$$

o que prova (A.2.1).

Em particular, (A.2.1) é equivalente a

$$P(|X-\mu| \geq c\sigma) \leq \frac{1}{c^2} \qquad (A.2.2)$$

onde $c > 0$. Basta tomar $k = c\sigma$. Ou ainda,

$$P(|X-\mu|<c\sigma) \geq 1 - \frac{1}{c^2}. \qquad \text{(A.2.3)}$$

Para $c = 2$, (A.2.3) nos diz que

$$P(|X-\mu|<2\sigma) \geq \frac{3}{4},$$

ou seja,

$$P(\mu-2\sigma<X<\mu+2\sigma) \geq \frac{3}{4}. \qquad \text{(A.2.4)}$$

Isto é, para qualquer v.a. X, com variância finita, pelo menos 3/4 da massa está compreendida no intervalo $[\mu-2\sigma,\mu+2\sigma]$.

Se X é normal, com média μ e variância σ^2, sabemos que $P(\mu-2\sigma<X<\mu+2\sigma) \approx 0{,}95$, portanto o limite inferior dado por (A.2.4) é bastante impreciso. Todavia, como nada se pressupõe a respeito da distribuição de X, a não ser que tenha variância finita, a desigualdade (A.2.1) é bastante útil.

A.2.2 - LEI DOS GRANDES NÚMEROS

Consideremos n provas de Bernoulli com $p = P(\text{sucesso})$ e seja k o número de sucessos nas n provas. Então, a Lei dos Grandes Números afirma que, para todo $\varepsilon>0$,

$$P\left(\left|\frac{k}{n} - p\right| \geq \varepsilon\right) \leq \frac{p(1-p)}{n\varepsilon^2}. \qquad \text{(A.2.5)}$$

DEMONSTRAÇÃO: Utilizando (A.2.1) para a v.a. k/n,

$$P\left(\left|\frac{k}{n} - p\right| \geq \varepsilon\right) \leq \frac{Var(k/n)}{\varepsilon^2}.$$

Mas, sabemos que $Var(k/n) = \frac{p(1-p)}{n}$, do que decorre (A.2.5).

Vemos que (A.2.5) é equivalente a

$$P(\left|\frac{k}{n} - p\right| < \varepsilon) \geq 1 - \frac{p(1-p)}{n\varepsilon^2}. \qquad \text{(A.2.6)}$$

É claro que (A.2.5) também implica

$$\lim_{n\to\infty} P(\left|\frac{k}{n} - p\right| \geq \varepsilon) = 0 \ , \ \forall\varepsilon>0. \qquad \text{(A.2.7)}$$

A Lei dos Grandes Números afirma que, para n grande, a proporção de sucessos k/n está próxima de $p = P(\text{sucesso})$.

No caso geral, se $X_1, X_2, \ldots$ é uma seqüência de v. a. independentes com uma distribuição comum, e se $E(X_i) = \mu$ existe, então, para todo $\varepsilon>0$,

$$\lim_{n\to\infty} P(\left|\frac{X_1+\ldots+X_n}{n} - \mu\right| \geq \varepsilon) = 0. \qquad \text{(A.2.8)}$$

A.2.3 - TEOREMA DO LIMITE CENTRAL

No Capítulo 3 discutimos a aproximação de uma binomial por uma normal. A justificativa teórica é dada pelo chamado *Teorema de De Moivre-Laplace*, que afirma o seguinte.

Sejam $X_1, X_2, \ldots, X_n$ variáveis binomiais, isto é, X_n é o número de sucessos em n provas de Bernoulli, P(sucesso) = p. Seja:

$$Y_n = \frac{X_n - np}{\sqrt{npq}}.$$

Então, quando $n \to \infty$,

$$P(Y_n \geq z) \to P(Z \geq z),$$

onde Z é a v.a. N(0,1) e z é constante.

Ou seja, para n grande, a distribuição de Y_n pode ser aproximada por uma distribuição normal padrão.

Este teorema, para X_n binomial, é um caso especial de um resultado mais geral, chamado *Teorema do Limite Central*, que apresentamos na forma seguinte.

Seja $\{X_n\}$ uma seqüência de v.a. independentes e com a mesma distribuição, $E(X_n) = \mu$ e $Var(X_n) = \sigma^2$. Seja

$$Y_n = X_1 + X_2 + \ldots + X_n.$$

Então, para z fixo, quando $n \to \infty$,

$$P\left(\frac{Y_n - n\mu}{\sigma\sqrt{n}} \geq z\right) \to P(Z \geq z),$$

onde Z é a v.a. N(0,1).

O teorema pode ser generalizado, por exemplo, para o caso em que $\{X_n\}$ são independentes, mas não têm a mesma distribuição.

A.2.4 - HIPÓTESES COMPOSTAS

No Capítulo 8 vimos como testar uma hipótese simples contra uma alternativa simples. Vimos, também, para um caso especial, como testar uma hipótese composta contra uma alternativa composta. Para outros casos mais gerais temos que usar o *teste da razão de verossimilhança generalizado*, que passamos a discutir.

Suponhamos o teste

$$H:\ \theta \in \Theta_0$$
$$K:\ \theta \in \Theta_1,$$

como apresentado no Capítulo 8, $\Theta_0 \cup \Theta_1 = \Theta$. Seja $L(\theta; x_1, \ldots, x_n)$ a função de verossimilhança para uma amostra $(X_1, \ldots, X_n)$. A razão de verossimilhança generalizada é definida por

$$\lambda = \frac{\sup\limits_{\theta \in \Theta_0} L(\theta; x_1, \ldots, x_n)}{\sup\limits_{\theta \in \Theta} L(\theta; x_1, \ldots, x_n)} \tag{A.2.9}$$

O princípio do teste da razão de verossimilhança generalizado afirma que rejeitamos H se $\lambda \leq \lambda_0$, onde λ_0 é uma constante a ser determinada de modo que o teste tenha nível α, $0 \leq \lambda_0 \leq 1$.

Se, em (A.2.9) substituímos as observações $(x_1, \ldots, x_n)$ pelas v.a. correspondentes, escrevemos Λ para λ, pois a razão será uma v.a.

EXEMPLO - Vamos testar

$$H: \mu=\mu_0$$

$$K: \mu\neq\mu_0,$$

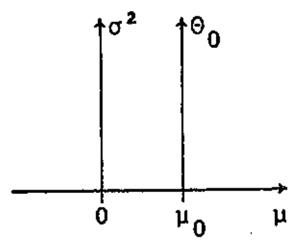
se $(X_1,\ldots,X_n)$ é uma amostra aleatória de uma distribuição normal $N(\mu,\sigma^2)$, onde μ e σ^2 são desconhecidos.

O espaço do parâmetro é

$$\Theta = \{(\mu,\sigma^2): -\infty<\mu<\infty\ ,\ \sigma^2>0\}$$

e $$\Theta_0 = \{(\mu_0,\sigma^2):\ 0<\sigma^2<\infty\}.$$

$$\Theta_1 = \{(\mu,\sigma^2):\ \mu\neq\mu_0\ ,\ \sigma^2>0\}.$$



A função de verossimilhança é

$$L(\mu,\sigma^2;x_1,\ldots,x_n) = \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right\}.$$

Vamos determinar sup L para $(\mu,\sigma^2)\in\Theta$. Tomando o log L e derivando em relação a μ e a σ^2 obtemos

$$\frac{1}{2\sigma^2}\sum(x_i-\mu) = 0$$

$$\frac{-n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum(x_i-\mu)^2 = 0,$$

do que decorrem

$$\hat{\mu} = \bar{x},$$
$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2. \tag{A.2.10}$$

Logo, as estatísticas que maximizam L são os estimadores de máxima verossimilhança de μ e σ^2. Portanto, chamando L^* o $\sup\limits_{(\mu,\sigma^2)\in\Theta} L$, temos que

$$L^*(x_1,\ldots,x_n) = \left(\frac{1}{2\pi\hat{\sigma}^2}\right)^{n/2} \exp\left(-\frac{1}{2\hat{\sigma}^2}\cdot n\hat{\sigma}^2\right) =$$
$$= \left(\frac{1}{2\pi\hat{\sigma}^2}\right)^{n/2} \exp\left(-\frac{n}{2}\right) \tag{A.2.11}$$

Agora, vamos determinar sup L para $(\mu,\sigma^2)\in\Theta_0$, isto é, $\mu = \mu_0$ e $\sigma^2 > 0$. Portanto, temos que maximizar

$$L(\mu_0,\sigma^2;x_1,\ldots,x_n) = \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} \exp\left(-\frac{1}{2\sigma^2} \sum (x_i-\mu_0)^2\right).$$

Derivando log L em relação a σ^2 obtemos

$$\overset{*}{\sigma}{}^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i-\mu_0)^2, \tag{A.2.12}$$

portanto, sendo $L_* = \sup\limits_{(\mu,\sigma^2)\in\Theta_0} L$, vem que

$$L_* = \left(\frac{1}{2\pi\overset{*}{\sigma}{}^2}\right)^{n/2} \exp\left(-\frac{n}{2}\right). \tag{A.2.13}$$

A razão de verossimilhança (A.2.9) fica

$$\lambda = \frac{L_*}{L^*} = \frac{\left(\frac{1}{2\pi\hat{\hat{\sigma}}^2}\right)^{n/2}}{\left(\frac{1}{2\pi\hat{\sigma}^2}\right)^{n/2}} = \left(\frac{\hat{\sigma}^2}{\hat{\hat{\sigma}}^2}\right)^{n/2}. \qquad \text{(A.2.14)}$$

Como $x_i-\mu_0 = x_i-\bar{x}+\bar{x}-\mu_0$, vemos facilmente que

$$(x_i-\mu_0)^2 = n\hat{\sigma}^2+n(\bar{x}-\mu_0)^2,$$

isto é,

$$\hat{\hat{\sigma}}^2 = \hat{\sigma}^2 + (\bar{x}-\mu_0)^2. \qquad \text{(A.2.15)}$$

Substituindo (A.2.15) em (A.2.14), obtemos:

$$\lambda = \left[\frac{1}{1+\left(\frac{\bar{x}-\mu_0}{\hat{\sigma}}\right)^2}\right]^{n/2}. \qquad \text{(A.2.16)}$$

Portanto,

$$\lambda \le \lambda_0 \iff \left|\frac{\bar{x} - \mu_0}{\sigma}\right| > k \qquad \text{(A.2.17)}$$

O que pode ser demonstrado é que a v.a.

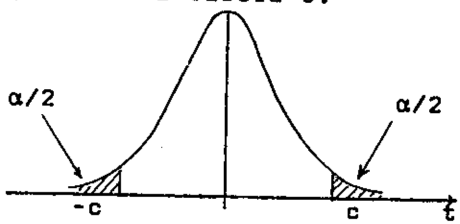
$$\frac{\bar{X}-\mu_0}{\hat{\sigma}/\sqrt{n-1}} \qquad \text{(A.2.18)}$$

tem uma distribuição t com n-1 graus de liberdade.

Como (A.2.17) é eqüivalente a

$$\left|\frac{\bar{x}-\mu_0}{\hat{\sigma}/\sqrt{n-1}}\right| > c, \qquad \text{(A.2.19)}$$

a constante c é determinada pela condição que $P(|t|>c) = \alpha =$ = nível do teste, usando a tabela 3.



Por exemplo, se queremos testar

$$H: \mu = 10$$

$$K: \mu \neq 10,$$

ao nível $\alpha = 0{,}05$, e obtemos $\bar{X} = 9{,}8$, $\hat{\sigma}^2 = 2{,}25$ a partir de u ma amostra de tamanho $n = 17$, então o valor de

$$\frac{\bar{x}-\mu_0}{\hat{\sigma}/\sqrt{n-1}} \text{ é } \frac{9{,}8 - 10}{(1{,}5)/4} = -0{,}33.$$

Como, para $\alpha = 0{,}05$, $c = 2{,}120$, vemos que a região de rejeição é $|t| > 2{,}120$. No caso, aceitamos H.

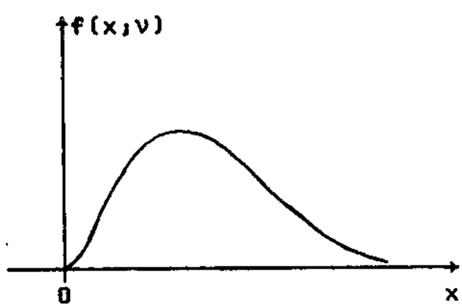
Nem sempre é possível obter de maneira fácil a distribuição da estatística a ser usada no teste. Sob determinadas condições, pode-se demonstrar que $-2 \log \Lambda$ tem uma distribuição limite, que é chamada *distribuição do qui-quadrado*. Esta distribuição, como a distribuição t, é caracterizada por um parâmetro, chamado também *número de graus de* *liberdade*. No caso em questão, a distribuição limite tem $\nu = \nu_1 - \nu_2$ graus de liberdade, onde ν_1 = número de parâmetros independentes em Θ e ν_2 = número de parâmetros independentes em Θ_0. No exemplo da normal acima, $\nu_1 = 2$ e $\nu_2 = 1$, e a distribuição limite é uma distribuição do qui-quadrado com 1 grau de liberdade, que denotamos χ_1^2. De modo geral, uma distribuição qui-quadrado com ν graus de liberdade tem função densidade

$$
\begin{aligned}
f(x;\nu) &= \frac{1}{\Gamma(\nu/2)}\left(\frac{1}{2}\right)^{\nu/2} \cdot x^{(\nu/2)-1} \exp\left(-\frac{x}{2}\right), \quad x>0 \\
&= 0 \qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad , \quad x<0
\end{aligned}
\qquad \text{(A.2.20)}
$$

Aqui, $\Gamma(\cdot)$ é a função gama, definida por

$$\Gamma(p) = \int_0^\infty x^{p-1}\, e^{-x} dx,$$

para $p > 0$ real. O gráfico de (A.2.20) tem a forma da figura abaixo, para $\nu > 2$.



Esta distribuição é tabelada para diferentes valores de ν, mas não apresentamos uma tal tabela no apêndice.

Quando falamos em distribuição limite, estamos considerando n grande. Portanto, a aproximação $-2 \log \Lambda \approx \chi^2_\nu$ só é razoável para amostras grandes. Não discutiremos, aqui, pormenores sobre esta aproximação, bem como a utilização da mesma em alguns testes (os chamados *testes do qui-quadrado*).

- o -