NAVALE NAVTIS
BELLUM DOCERE

MINISTÉRIO DA MARINHA

ESCOLA DE GUERRA NAVAL

NOÇÕES FUNDAMENTAIS DA TEORIA DAS PROBABILIDADES

APLICAÇÕES

CMG (RRm) ADOLPHO GOMES BUSSE

NOÇOES FUNDAMENTAIS DA TEORIA DAS PROBABILIDADES

APLICAÇÕES

CMG (RRm) ADOLPHO GOMES BUSSE

1. INTRODUÇÃO

Sob a denominação genérica de Estatística eram designados antigamente todos aqueles dados de importância para a condução dos negócios do Estado, tais como, população, nascimentos, óbitos, impostos arrecadados, etc. Modernamente, a Estatística é entendida como o ramo de ciência (conjunto de conhecimentos coordenados e relativos a um determinado objeto ou aos fenômenos de uma ordem (ou classe) que, apoiada em métodos matemáticos e na teoria das probabilidades, trata do estudo de massas de dados numéricos de qualquer natureza. Nesse sentido, a Estatística penetrou nos mais variados ramos do conhecimento humano, tais como a Biologia, a Engenharia, a Economia, etc.

No estudo da Estatística, é comum, para fins didáticos, dividí-la em duas partes bem distintas: a Estatística Descritiva ou Metodológica e a Estatística Matemática ou Indutiva. A primeira compreende o tratamento de massas de dados numéricos sem qualquer generalização. A segunda trata da generalização (por meio de uma teoria matemática), de problemas de estimação, testes de hipótese e predição.

A Estatística Descritiva trata principalmente das seguintes etapas do procedimento estatístico: a produção ou coleta dos dados, a sua apuração e classificação, a sua apresentação sob forma tabular ou gráfica e o cálculo de medidas descritivas da massa de dados numéricos.

O estudo da Estatística Matemática pressupõe um conhecimento de matemáticas em nível de graduação e de teoria das probabilidades. Começaremos com esta última.

2. TEORIA DAS PROBABILIDADES

Segundo Fisz (Probability Theory and Mathematical Statistics), a teoria das probabilidades "é o ramo das matemáticas que trata da descoberta e da investigação das características regulares dos eventos aleatórios".

Essa definição não traz muita luz ao entendimento da teoria, pois nela há duas expressões (grifadas) que não foram explicadas previamente. Voltaremos a elas.

Segundo Parzen (Modern Probability Theory and its Applications), a teoria das probabilidades " é o ramo das matemáticas que trata do estudo dos métodos de análise que são comuns ao estudo dos fenômenos aleatórios, onde quer que eles ocorram". Nessa segunda definição, há também uma expressão (grifada), a qual não se ememprestou um sentido que torne fechada a definição. Por isto, inicialmente iremos introduzir as noções de fenômeno aleatório, experiência aleatória, evento aleatório e características regulares dos eventos aleatórios.

3. FENOMENO ALEATÓRIO. EXPERIÊNCIA ALEATÓRIA, EVENTO ALEATÓRIO E REGULARIDADE ESTATÍSTICA

Como já dissemos, a teoria das probabilidades penetrou nos mais variados ramos do conhecimento humano. Ocorre então a pergunta: qual a propriedade comum aos fenômenos que ocorrem nesses diferentes ramos do conhecimento humano, que torna a teoria das probabilidades apta á ter aplicação cada vez maior? Para responder a essa pergunta, introduzimos a noção de fenômeno aleatório: é um fenômeno empírico caracterizado pela propriedade de que, a sua observação debaixo de um dado conjunto de condições, não conduz sempre ao mesmo resultado (de modo que não há regularidade determinística) e sim a diferentes resultados, de modo que se observa uma regularidade estatística". Por exemplo, a produção de parafusos por um torno é um fenômeno aleatório, pois o parafuso produzido pode apresentar as seguintes características:ser bom ou defeituoso.

Introduzimos agora as noções de experiência aleatória, evento aleatório e regularidade estatística.

Suponha-se a experiência de lançamento de uma moeda dessa experiência participam o experimentador e a moeda, em local onde a experiencia vai ser realizada. Suponha-se que o experimentador crê nas leis do acaso (não determinista) e que a moeda seja honesta (tenha cara e coroa e seja homogénea e bem equilibrada). Suponha-se mais que as condições do local onde vai ser realizada a experiência estejam controladas(por exemplo: velocidade do vento, temperatura, pressão atmosférica, etc.). Antes da realização não se pode predizer qual o resultado que ocorrerá. A única coisa que podemos predizer é que poderá ocorrer cara ou coroa.

Uma vez realizada a experiência, ocorrerá um daqueles resultados, cara, por exemplo. Cara é um evento aleatório,isto é, um resultado eventual da experiência.

Uma experiência debaixo do conjunto de condições acima especificadas é dito ser uma experiência aleatória.

Vejamos agora a noção de regularidade estatística (dentro em breve daremos uma noção rigorosa de evento aleatório).

Considere-se novamente o lançamento de uma moeda e pensemos na realização dessa mesma experiência um grande número de vezes, sempre debaixo do mesmo (ou aproximadamente o mesmo)conjunto de condições. Chamemos de:

A - evento que consiste na ocorrência de cara.

$\nu(A)$ - número de vezes que ocorre cara, em n realizações da experiência considerada.

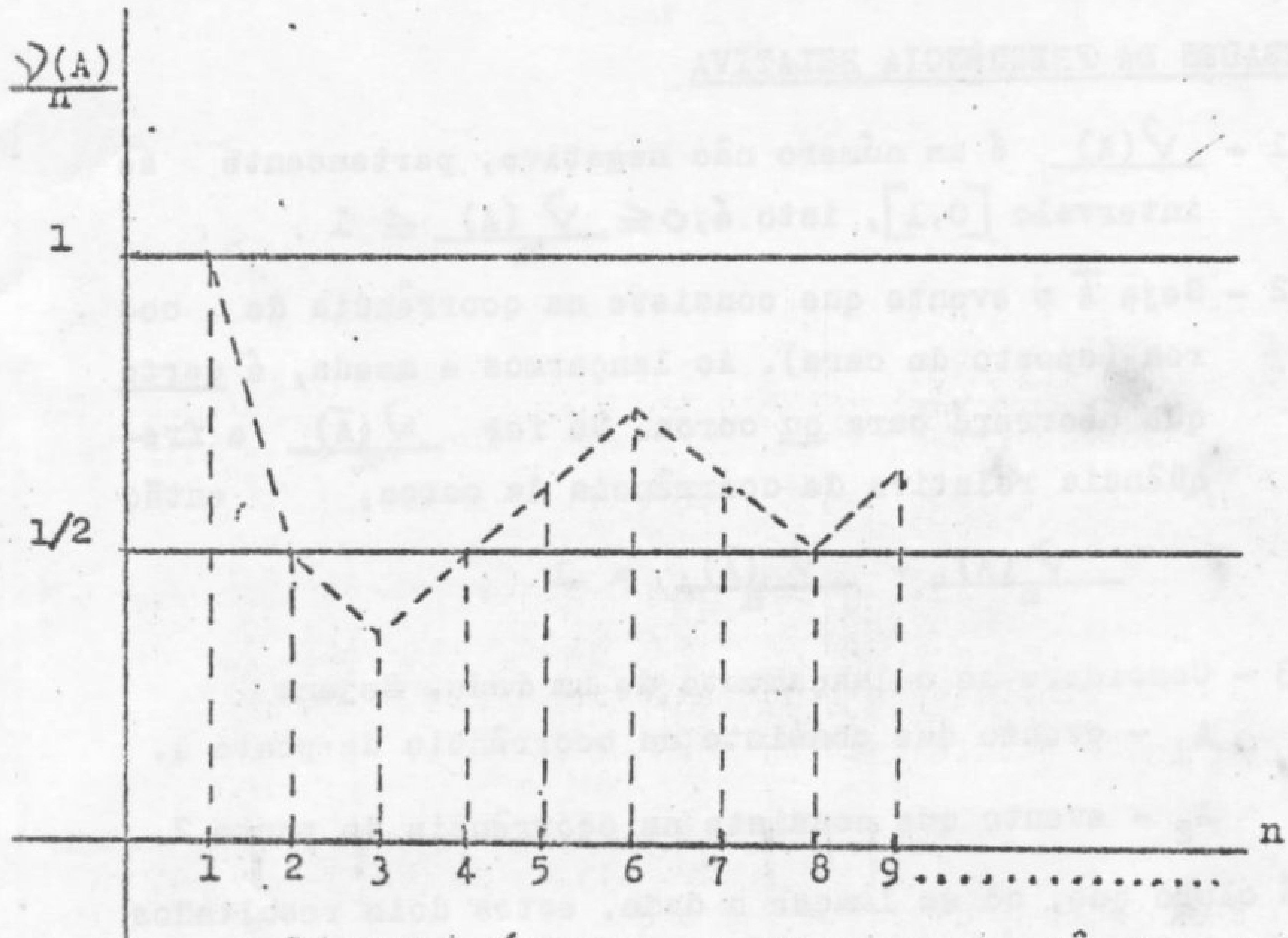
$\frac{\nu(A)}{n}$ - freqüência relativa de ocorrência de A.

Ao realizar a experiência e observar o resultado daremos o algarismo 1 para a ocorrência de cara e zero para a ocorrência de coroa, de modo que, sempre que ocorre cara, acumularemos o resultado.

Suponha-se que a experiência foi realizada um certo número de vezes (poderia prolongar-se indefinidamente) e construi-se o quadro seguinte:

n	Resultado	$\nu(A)$	$\frac{\nu(A)}{n}$
1	Cara	1	1
2	Coroa	1	1/2
3	Coroa	1	1/3
4	Cara	2	2/4
5	Cara	3	3/5
6	Cara	4	4/6
7	Coroa	4	4/7
8	Coroa	4	4/8
9	Cara	5	5/9
n	...	...	...

Se levarmos os resultados a um gráfico, com n nas abcissas e $\frac{\nu(A)}{n}$ nas ordenadas, obteremos o seguinte:



Se continuássemos a realizar a experiência indefinidamente, o que observaríamos seria que a freqüência relativa $\left(\frac{\nu(A)}{n}\right)$ se aproximaria cada vez mais de um valor estável (no caso 1/2), sem nunca contudo ser igual a 1/2. À proporção que <u>n</u> fosse crescendo indefinidamente, o afastamento absoluto de $\frac{\nu(A)}{n}$ em relação á 1/2, seria cada vez menor. Essse valor estável 1/2, é a probabilidade " a priori" de ocorrência de cara, na experiência considerada. Os fenômenos aleatórios apresentam essa característica regular, que se chama regularidade estatística. A observação nos mostra que, à proporção que cresce n, $\frac{\nu(A)}{n}$ se aproxima de $P[A]$, que é a probabilidade de ocorrência de A, ou seja:

$$\frac{\nu(A)}{n} \xrightarrow{P} P[A]$$

A notação $\xrightarrow{P}$ lê-se: "converge em probabilidade para".

4. PROPRIEDADES DA FREQUÊNCIA RELATIVA

1 - $\frac{\nu(A)}{n}$ é um número não negativo, pertencente ao intervalo $[0,1]$, isto é, $0 \leq \frac{\nu(A)}{n} \leq 1$

2 - Seja $\bar{A}$ o evento que consiste na ocorrência de coroa (oposto de cara). Ao lançarmos a moeda, é certo que ocorrerá cara ou coroa. Se for $\frac{\nu(\bar{A})}{n}$ a freqüência relativa de ocorrência de coroa, então

$$\frac{\nu(A)}{n} + \frac{\nu(\bar{A})}{n} = 1$$

3 - Considere-se o lançamento de um dado. Sejam:

A_1 - evento que consiste na ocorrência de ponto 1.

A_2 - evento que consiste na ocorrência de ponto 2.

É claro que, ao se lançar o dado, estes dois resultados não podem ocorrer simultaneamente. Diz-se que os 2 eventos são mutuamente exclusivos (ME).

Se perguntarmos: qual a probabilidade (em termos de freqüência relativa) da ocorrência do evento "A_1 ou A_2"?

A resposta será imediata: $\frac{2}{6}$.

Então, para eventos MF:

$$\frac{\nu(A_1 \text{ ou } A_2)}{n} = \frac{\nu(A_1)}{n} + \frac{\nu(A_2)}{n}$$

Essas propriedades serão usadas mais adiante, quando tratarmos do sistema de axiomas da teoria das probabilidades.

5. ESPAÇO AMOSTRA (OU CONJUNTO DE EVENTOS ELEMENTARES)

Chama-se Espaço Amostra de uma experiência aleatória, e designa-se por S, o conjunto de pontos representativos dos resultados dessa experiência. Por exemplo, no lançamento de uma moeda:

$$S = \{ \text{Cara, Coroa} \} = \{ C, c \}$$

no lançamento de um lado:

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

Para cada experiência, tem que ser construido o Espaço Amostra correspondente.

Nos dois exemplos acima, $\{C\}$ e $\{c\}$, $\{1\}$, $\{2\}$, $\{3\}$, $\{4\}$, $\{5\}$ e $\{6\}$ são chamados eventos elementares.

Pode-se reconhecer de imediato, particularmente no caso do dado, que ao se realizar a experiência, pode-se indagar a respeito da ocorrência de outros eventos, tais como, ocorrência de ponto par, ocorrência de ponto impar, ocorrência de ponto menor que 4, etc. Nota-se, então, que a partir dos eventos elementares de S, outros eventos podem ser gerados. Realmente, S gera todos os eventos possíveis, aos quais se pode atribuir probabilidades. Esses eventos são em número de 2^N, onde N é o número de eventos elementares de S (no caso do dado, N=6). Realmente, os eventos são os seguintes:

a) eventos de um só elemento : $\{1\}$, $\{2\}$, $\{3\}$, ... $\{6\}$ $C_6^1 =$ 6

b) eventos de dois elementos : $\{1,2\}$, ... $\{1,6\}$, $\{2,1\}$, .. $\{2,6\}$, .. $\{5,5\}$ $-C_6^2 =$ 15

c) eventos de três elementos : $\{1,2,3,\}$ .... $\{4,5,6\}$ $-C_6^3 =$ 20

d) eventos de quatro elementos : $\{1,2,3,4\}$, ... $\{3,4,5,6\}$ $-C_6^4 =$ 15

e) eventos de cinco elementos : $\{1,2,3,4,5\}$, ... $\{2,3,4,5,6\}$ $-C_6^5 =$ 6

62

A esses 62 eventos, devemos acrescentar o próprio S(evento de 6 elementos) e o evento que não contém elementos de S (conjunto vazio), o que totaliza 64 eventos ou 2^6 eventos.

Como vemos, a partir de S podemos gerar novos eventos e os 64 que foram gerados correspondem a todos os eventos possíveis que podem ser formados a partir de S. Chega-se assim a uma noção rigorosa de evento: evento é qualquer subconjunto de S.

6. DEFINIÇÕES E OPERAÇÃO SOBRE EVENTOS E PROPRIEDADES (a seqüencia a seguir encontra-se Fisz, op. cit)

DEFINIÇÃO 1

O conjunto, que chamaremos de Z, formado de todos os subconjuntos gerados por S , é chamado Corpo de Borel de Eventos (essa definição será ampliada adiante).

DEFINIÇÃO 2

O evento que contém todos os elementos de S é chamado evento certo e representa-se pela letra S.

DEFINIÇÃO 3

O evento que não contém elementos de S é chamado evento impossível e representa-se pelo símbolo ϕ.

DEFINIÇÃO 4

Sejam A_1 e A_2 eventos quaisquer (subconjuntos de S). Diz-se que A_1 está contido em A_2 e representa-se pela notação $A_1 \subset A_2$, se todo elemento de A_1 é também elemento de A_2. Isto significa que a ocorrência de A_1 implica na ocorrência de A_2, sem que o oposto se verifique,

Corolário - se A_1 está contido em A_2 e se A_2 está contido em A_1, diz-se que nos dois eventos são equivalentes e representa-se por $A_1 = A_2$.

Propriedade 1 - O conjunto Z contém como um de seus elementos o evento certo.

Propriedade 2 - O conjunto z contém como um de seus elementos o evento impossível.

DEFINIÇÃO 5

Diz-se que dois eventos A_1 e A_2 são mutuamente exclusivos (ME), se eles não contém em comum, elementos de S.

Exemplo: par e impar, no lançamento de um dado, são eventos ME.

DEFINIÇÃO 6

Sejam A_1, A_2, ..., An, ..., eventos quaisquer. O evento A que contém aqueles e somente aqueles elementos que pertencem a pelo menos um dos eventos A_1, A_2, ...,An,...., é chamado UNIÃO ou ALTERNATIVA dos eventos considerados e representa-se por:

$$A = A_1 \cup A_2 \ldots \quad \cup A_n \cup \ldots\ldots\ldots = \bigcup_{n=1}^{\infty} A_n$$

O símbolo $\cup$, da operação de união, deve ser lido ou; em outras palavras, o evento A ocorre se somente ocorrer pelo menos um dos eventos considerados.

No caso de 2 eventos A_1 e A_2, para exemplificar, há três alternativas, que são as seguintes: ocorrer A_1 sem ocorrer A_2, ou ocorrer A_2 sem ocorrer A_1, ou ocorrerem A_1 e A_2.

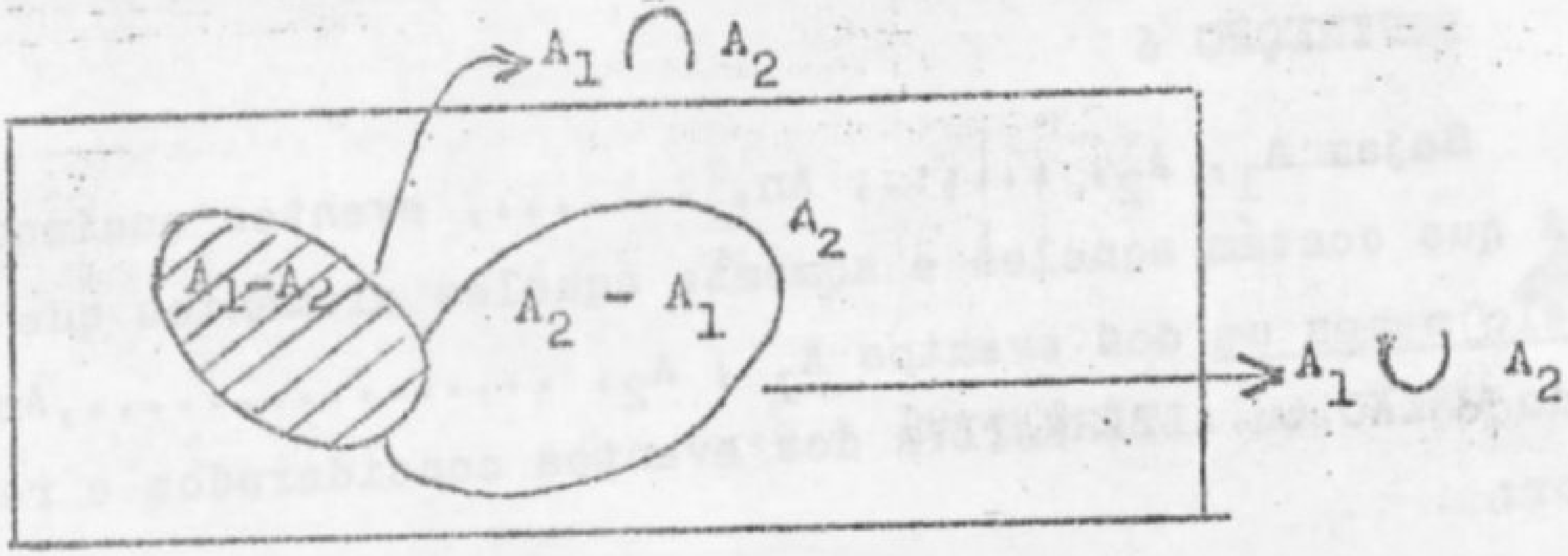
Propriedade 3 - Se os eventos A_1, A_2, ..., An ......, pertencem a Z, a sua alternativa também pertence a Z (diz-se que Z é fechado para a operação de União).

DEFINIÇÃO 7 (Operação de Diferença)

Sejam A_1 e A_2 eventos quaisquer. O evento A, que consiste na ocorrência de A_1, sem a ocorrência de A_2, é chamado DIFERENÇA entre os dois eventos e representa-se por $A = A_1 - A_2$.

Corolário - Seja A um evento qualquer. O evento $\overline{A}$, que consiste na ocorrência de S, sem a ocorrência de A, é chamado complementar de A em relação a S e representa-se por $\overline{A} = S - A$.

Na figura a seguir, representa-se a união e a diferença entre os eventos A_1 e A_2.



Propriedade 4 - Se A_1 e A_2 pertencem a Z, a sua diferença também pertence a Z (diz-se que Z é fechado para a operação da Diferença).

DEFINIÇÃO 8

Sejam A_1 A_2, ... , An, ... , eventos quaisquer. O evento A, que consiste na ocorrência de todos os eventos elementares que são comuns aos eventos A_1, A_2, ..., é chamado interseção dos eventos considerados e representa-se por:

$$A = A_1 \cap A_2 \cap \ldots\ldots\ldots \cap An \cap \ \ldots = \bigcap_{n=1}^{\infty} A_i$$

O símbolo $\cap$, de interseção, lê-se e.

Na figura anterior está representado o evento $A_1 \cap A_2$.

Propriedade 5 - Se os eventos A_1 , A_2 , ..., An, ..., pertencem a Z, a sua interseção também pertence a Z (diz-se que Z é fechado para a operação de Interseção).

DEFINIÇÃO 9

O conjunto Z, formado de todos os subconjuntos de S, com as propriedades 1 a 5, é chamado Corpo de Borel de Eventos e seus elementos são chamados eventos aleatórios.

7. SISTEMA DE AXIOMAS DA TEORIA DAS PROBABILIDADES

Axioma 1 - A todo evento A, pode-se associar um número real não negativo $P[A]$, chamado probabilidade de ocorrência de A, que satisfaz à dupla desigualdade : $0 \leq P[A] \leq 1$

Axioma 2 - A probabilidade do evento certo é igual a 1, isto é, $P[S] = 1$ (a recíproca não é verdadeira, isto é, o evento que tem probabilidade de 1 não é necessariamente o evento certo).

Axioma 3 - Sejam A_1 , A_2 ......, An........., eventos mutuamente exclusivos. Então :

$$P\left[\bigcup_{n=1}^{\infty} A_n\right] = \sum_{n=1}^{\infty} P\left[A_n\right],$$

isto é, a probabilidade da união de eventos mutuamente exclusivos é igual à soma das probabilidades desses eventos.

8. PRINCIPAIS TEOREMAS DA TEORIA DAS PROBABILIDADES

Teorema 1. A probabilidade do evento impossível é igual a zero, isto é, $P[\phi] = 0$

A demonstração é imediata, pois $S = S \cup \emptyset$; como S e $\emptyset$ são

ME, tem-se $P[S] = P[S] + P[\emptyset] \therefore P[\emptyset] = 0$

<u>Teorema 2</u>. Se $A_1 \subset A_2$, então $P[A_1] \leq P[A_2]$

Se $A_1 \subset A_2$, então:

$$A_2 = A_1 \cup (A_2 \cap \bar{A}_1) \ ;$$

ora, os eventos A_1 e A_2 são ME, logo,

$P[A_2] = P[A_1] + P[A_2 \cap \bar{A}_1]$ e como

$P[A_2 \cap A_1] \geq 0$, vem que $P[A_2] \geq P[A_1]$

<u>Teorema 3</u>. Sejam A_1 e A_2 eventos quaisquer.

Então $P[A_2] = P[A_1 \cap A_2] + P[A_2 \cap \bar{A}_1]$

A demonstração é imediata, pois $A_2 = (A_2 \cap \bar{A}_1) \cup (A_1 \cap A_2)$.

Como o segundo membro é a união de 2 eventos ME, segue-se de imediato a demonstração.

<u>Teorema 4</u>. Sejam A_1 e A_2 eventos quaisquer. Então:

$$P[A_1 \cup A_2] = P[A_1] + P[A_2] - P[A_1 \cap A_2]$$

<u>Demonstração</u>

$$A_1 \cup A_2 = A_1 \cup (A_2 \cap \bar{A}_1) \qquad (1)$$

$$A_2 = (A_1 \cap A_2) \cup (A_2 \cap \bar{A}_1) \qquad (2)$$

Combinando as 2 ultimas igualdades, demonstra-se o teorema. O teorema extende-se a um número qualquer de eventos. Se forem 3 eventos, por exemplo, A_1, A_2 e A_3, temos:

$$P[A_1 \cup A_2 \cup A_3] = P[A_1] + P[A_2] + P[A_3] - P[A_1 \cap A_2] - P[A_1 \cap P_3] - P[A_2 \cap A_3] + P[A_1 \cap A_2 \cap A_3],$$

sendo a fórmula geral obtida por extensão das fórmulas de 2 e 3 eventos.

Teorema 5 . Sejam A_1 e A_2 eventos quaisquer. Então:

$P\left[A_1 \cap A_2\right] = P\left[A_1\right] P\left[A_2/A_1\right] = P\left[A_2\right] \; P\left[A_1/A_2\right]$. Os eventos $A_1/A2$ e A_2/A_1 são chamados eventos condicionais e devem ser lidos respectivamente, "A_1 na hipótese de ter ocorrido A_2" e "A2 na hipótese de ter ocorrido A_1". Assim, P A_2/A_1 é a probabilidade de ocorrência do evento condicional A_2/A_1.

$P[A2] = 1/6$

$P[A1] = 3/6$ $\quad P[A2/A1] = \frac{1}{3}$

Teorema 6 . Se $\bar{A}$ é o evento oposto de A, então:

$$P\left[\bar{A}\right] = 1 - P\left[A\right]$$

Demonstração:

$$A \cup \bar{A} = S$$

$P\left[A \cup \bar{A}\right] = P\left[S\right] = 1$; mas, $P\left[A \cup \bar{A}\right] =$

$= P\left[A\right] + P\left[\bar{A}\right]$, o que demonstra o teorema.

APLICAÇÃO

Um exemplo simples de urna e bolas.

Extração de uma amostra de tamanho $n = 2$, com reposição ou sem reposição das bolas.

Considere-se uma urna com 3 bolas brancas e 5 bolas pretas.

Extraem-se sucessivamente dessa urna, com (sem) reposição, 2 bolas.

Calcular as probabilidades dos seguintes eventos:

a) ambas as bolas serem brancas;

b) a 2ª bola da amostra ser branca;

c) pelo menos uma bola ser branca;

SOLUÇÃO

Consideremos primeiro a extração com reposição.

Chamemos:

A_1 – evento que consiste na ocorrência da bola branca na 1ª extração;

A_2 – idem, bola branca na 2ª extração;

$\bar{A}_1$ – evento que consiste na ocorrencia de bola preta na 1ª extração;

$\bar{A}_2$ – idem, bola preta na 2ª extração ;

Então:

a) $P\left[A_1 \cap A_2\right] = P\left[A_1\right] P\left[A_2/A_1\right] = \frac{3}{8} \times \frac{3}{8} = \frac{9}{64}$

b) $P\left[A_2\right] = P\left[(A_1 \cap A_2) \cup (\bar{A}_1 \cap A_2)\right] = P\left[A_1 \cap A_2\right] + P\left[\bar{A}_1 \cap A_2\right] =$

$= \frac{5}{8} \times \frac{3}{8} + \frac{3}{8} \times \frac{3}{8} = \frac{3}{8}$

c) $P\left[A_1 \cup A_2\right] = P\left[A_1\right] + P\left[A_2\right] - P\left[A_1 \cap A_2\right] = \frac{3}{8} + \frac{3}{8} - \frac{9}{64} = \frac{39}{64}$

Consideremos agora a extração sem reposição:

a) $P\left[A_1 \cap A_2\right] = P\left[A_1\right] P\left[A_2/A_1\right] = \frac{3}{8} \times \frac{2}{7} = \frac{6}{56}$

b) $P\left[A_2\right] = P\left[(A_1 \cap A_2) \cup (\bar{A}_1 \cap A_2)\right] = P\left[A_1\right] P\left[A_2/A_1\right] + P\left[\bar{A}_1\right] \times$ $P\left[A_2/\bar{A}_1\right] = \frac{3}{8} \times \frac{2}{7} + \frac{5}{8} \times \frac{3}{7} = \frac{3}{8}$! (Compare com o resultado na amostra com reposição).

c) $P\left[A_1 \cup A_2\right] = P\left[A_1\right] + P\left[A_2\right] - P\left[A_1 \cap A_2\right] = \frac{3}{8} + \frac{3}{8} - \frac{6}{56} =$ $= \frac{36}{56}$. Este último resultado poderia ser obtido do seguin te modo:

$$P\left[A_1 \cup A_2\right] = 1 - P\left[\bar{A}_1 \cap \bar{A}_2\right] = 1 - \frac{5}{8} \times \frac{4}{7} = 1 - \frac{20}{56} = \frac{36}{56}$$

9. NOÇÃO DE VARIÁVEL ALEATÓRIA

Podemos associar a qualquer resultado de uma experiência aleatória (evento elementar), um número ou um conjunto de números do eixo dos reais. Assim, à ocorrência de cara, no lançamento de moeda, podemos associar o valor 1, e à ocorrência de coroa, o valor 0. Da mesma forma, à ocorrência de parafuso defeituoso produzido por um torno, podemos associar o valor 1 e parafuso bom o valor 0.

Pode-se introduzir a noção de variável aleatória através a noção de imagem inversa.

Definição 1

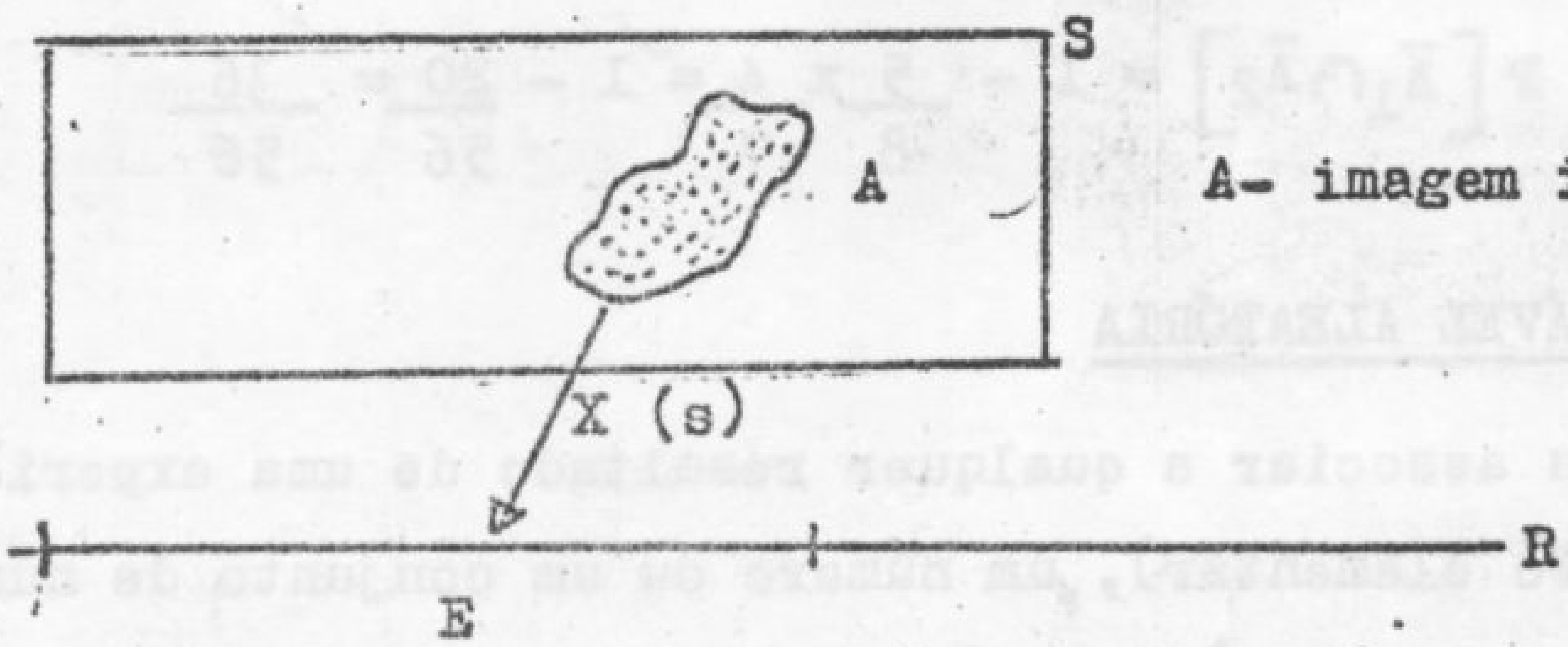
Seja X(s) uma função real unívoca definida sobre o conjunto s. O conjunto A de todos os eventos elementares aos quais a função X(s) associa valores num dado conjunto F dos números reais, é chamado imagem inversa de E.

É claro que a imagem inversa do conjunto R de todos os números reais é o próprio conjunto S (Espaço Amostra).

Definição 2

Uma função real unívoca X (s) definida sobre o conjunto dos eventos elementares S é chmada uma variável aleatória se a imagem inversa de todo intervalo I do eixo dos reais da forma $(-\infty, x)$ é um evento aleatório.

A figura abaixo mostra de maneira simplificada como se passa do Espaço Amostra S para o eixo dos reais.



A- imagem inversa de E

Então, a probabilidade de ocorrência de A, $P[A]$ é através da correspondência, a mesma que $X(s) \in E$ ou seja, $P[A] = P[X(s) \in E]$

Daqui por diante usaremos somente X para anotação da variável aleatória X. Se o evento aleatório A é a imagem inversa de um ponto x, então $P[A] = P[X = x]$

No exemplo simples do lançamento da moeda, chamando A evento cara, $P[A] = P[X = 1] = \frac{1}{2}$

10. FUNÇÃO DE DISTRIBUIÇÃO

Considere-se o lançamento de um dado. A todo evento elementar de S, associamos um número real, no caso 1, 2, 3, 4, 5 e 6. Nesse exemplo, a variável aleatória X assume os valores $x_i = i$ $(i = 1, 2, 3, \ldots\ldots\ldots, 6)$, com probabilidade,

$$P[X = x_i] = \frac{1}{6}$$

probabilidade de

$$X \leq 1, \quad P[X \leq 1] = \frac{1}{6}$$

$$P[X \leq 2] = P[X \leq 1] + P[X = 2] = \frac{1}{6} + \frac{1}{6} = \frac{2}{6},$$

$$P[X \leq 3] = P[X \leq 2] + P[X = 3] = \frac{3}{6}, \text{ etc },$$

$$P\left[X \leq 6\right] = P\left[X \leq 5\right] + P\left[X = 6\right] = \frac{6}{6} = 1$$

e resultados acima podem ser resumidos na seguinte fórmula:

$$P\left[X \quad x\right] = \sum_{i=1}^{x} P\left[X = i\right]$$

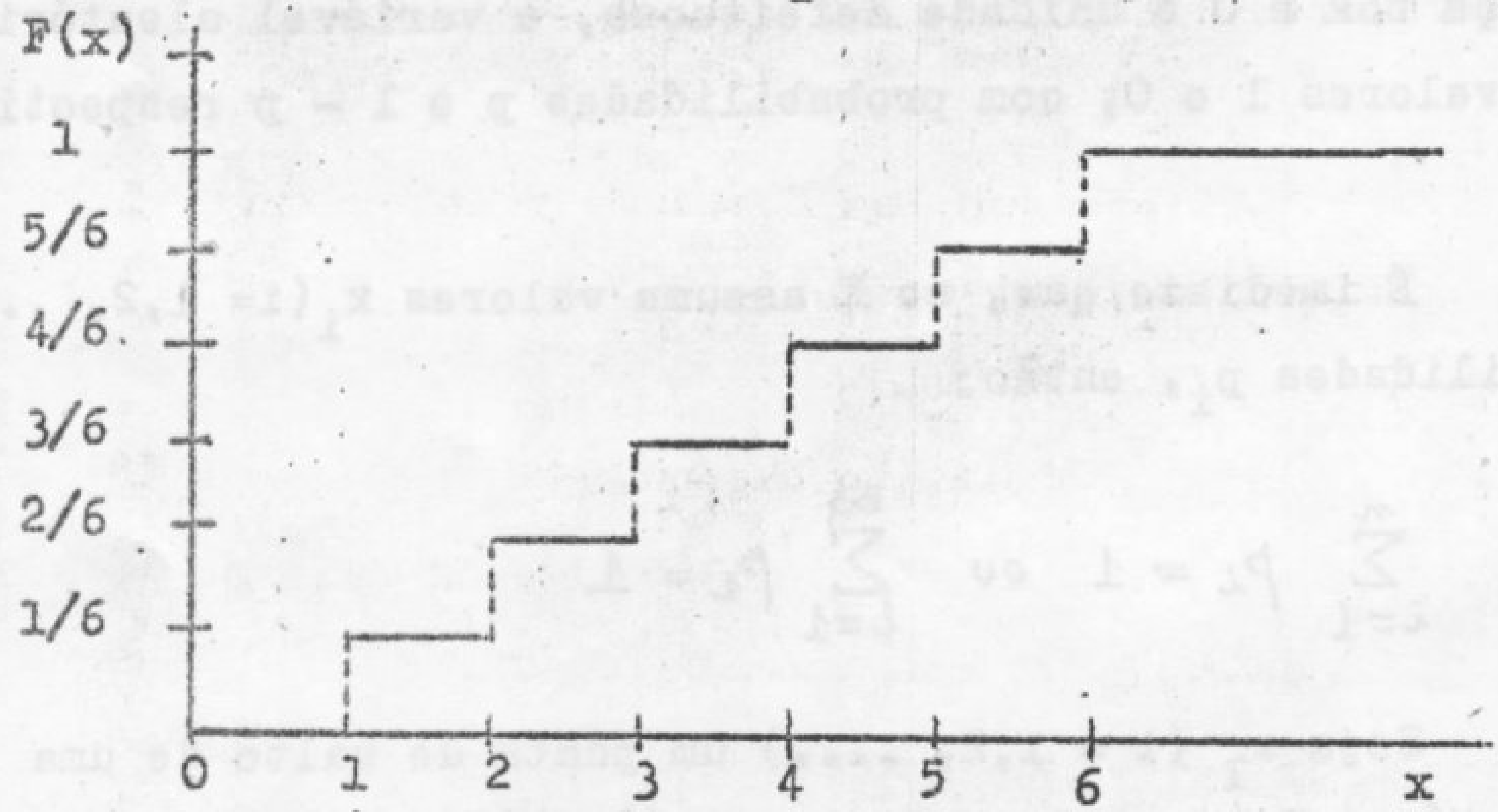
A função $F(x) = P\left[X \leq x\right]$ é chamada função de distribuição e variáveis aleatória X.

Da definição de F (x) segue-se imediatamente que:

$(-\infty) = 0$ e $F(+\infty) = 1$

Valem também as seguintes propriedades:

1a.) Se $x_1 < x_2$, $F(x_1) \leq F(x_2)$, isto é, a F(x) é uma função não decrescente. No exemplo citado, o gráfico da F(x) tem uma forma de escada, mostrada na figura abaixo:



2a.) A F(x) é contínua pelo menos à direita, isto é, $F(x) = F(x + 0)$

11. TIPOS DE VARIÁVEIS ALEATÓRIAS

Em teoria das probabilidades estudam-se dois tipos de variações aleatórias, que são a do tipo discreto e a do tipo contínuo(há ainda um terceiro tipo, misto, que não estudaremos).

Definição 1

Uma variável aleatória X é dita ser do tipo discreto, se ela assume, com probabilidade 1, valores pertencentes a um conjunto E, que é no máximo infinito enumerável, e todo valor em E tem probabilidade positiva. Esses valores são chamados pontos de salto e suas probabilidades, saltos.

Exemplo:

Um torno produz unidades boas e defeituosas. O Espaço Amostra é constituido de 2 eventos elementares, unidades boas e defeituosas. Seja p a probabilidade de extrair uma unidade boa e 1- p a de extrair uma unidade defeituosa ($0<p<1$). Atribuindo o valor 1 à peça boa e 0 à unidade defeituosa, a variável aleatória X assume os valores 1 e 0, com probabilidades p e 1 - p respectivamente.

É imediato que, se X assume valores x_i (i= 1,2, ...) com probabilidades p_i, então:

$$\sum_{i=1}^{n} p_i = 1 \quad \text{ou} \quad \sum_{i=1}^{\infty} p_i = 1.$$

Seja x_1 (i = 1,2, ...) um ponto de salto de uma variável aleatória X do tipo discreto. A probabilidade de que a variável aleatória X assume o valor x_1 é chamada função de probabilidade de X e escreve-se

$$P\left[X = x_i\right] = p_i$$

A função de distribuição de x é:

$$F(x) = \sum_{x_i \leq x} P \; X = x_i$$

Definição 2

Uma variável aleatória X é dita ser do tipo contínuo se existir uma função real não negativa f (x) tal que para todo número real x, se verifica a seguinte igualdade, ligando F(x) a f(x):

$$F(x) = \int_{-\infty}^{x} f(y)\, dy$$

A função f(x) é chamada função de densidade de probabilidade de X (ou função de densidade de X).

A função de densidade de X satisfaz à relação:

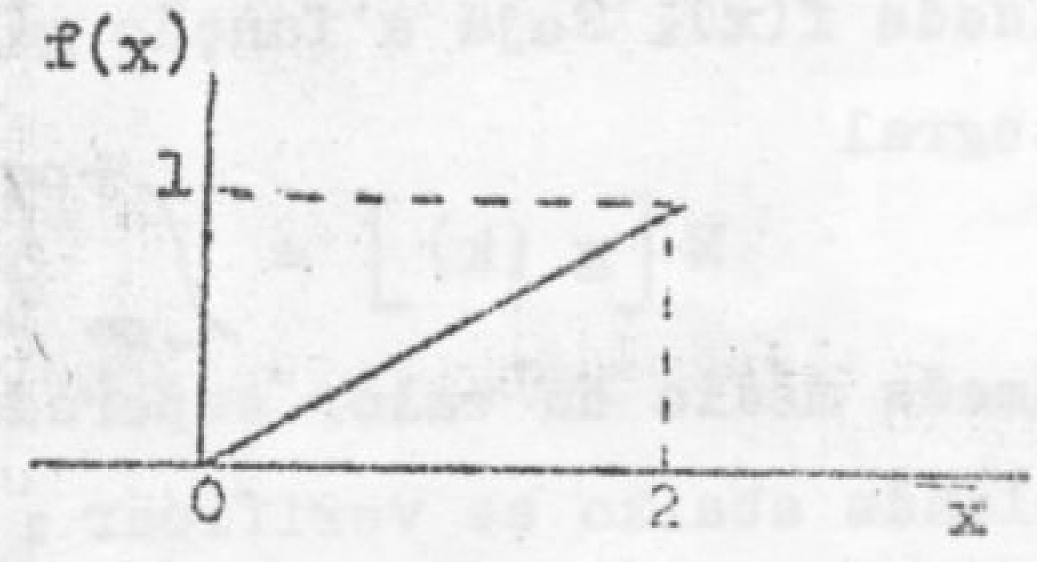
$$F(+\infty) = \int_{-\infty}^{+\infty} f(Y)\, dy = 1$$

Além disto, para todo real a e b, tal que $a < b$,

$$P\left[a \leq X \leq b\right] = F(b) - F(a) = \int_a^b f(x)\, dz$$

isto é, a probabilidade de um intervalo é medida pela integral de f (x) entre os extremos do intervalo. Por outro lado, se f(x) é continua em algum ponto x, tem-se F'(x) = f (x), isto é, a função de densidade é a derivada da função de distribuição.

Exemplo: $f(x) = \begin{cases} 0, & x < 0 \\ \frac{1}{2} x, & 0 < x < 2 \\ 0, & x > 2 \end{cases}$



Portanto, a F(x) será:

$$F(x) = \begin{cases} 0, \text{para } x \leq 0 \\ \frac{1}{4} x^2, \text{ para } 0 < x \leq 2 \\ 1, \text{ para } x > 2 \end{cases}$$

12. PARÂMETROS DA DISTRIBUIÇÃO DE UMA VARIÁVEL ALEATÓRIA

A toda distribuição de uma variável aleatória X estão associados números que chamamos os parâmetros dessa distribuição.

Esses parâmetros são chamados momentos, Funções desses momentos são também parâmetros, como veremos.

Definição

Seja X uma variável aleatória do tipo discreto com pontos de descontinuidade x_k e saltos (probabilidades dos pontos de descontinuidade) p_k.

A série

$$E\left[g(x)\right] = \sum_k g(x_k)\, p_k$$

é chamada média (ou valor esperado) da variável aleatória g(x), se se verifica a seguinte desigualdade:

$$\sum_k \left|g(x_k)\right| p_k < \infty$$

Definição

Seja X uma variável do tipo contínuo, com função de densidade f(x). Seja a função g(x) integrável no sentido de Riemann. A integral

$$E\left[g(x)\right] = \int_{-\infty}^{+\infty} g(x) f(x)\, dx$$

é chamada média ou valor esperado da variável aleatória g(x) se a desigualdade abaixo se verificar:

$$\int_{-\infty}^{+\infty} |g(x)|\, f(x)\, dx < \infty$$

O símbolo E é chamado Operador Média, isto é, operando sobre g(x), ele realiza a integral de g(x), relativamente a f(x), no intervalo $(-\infty, +\infty)$ (ou a soma, se a variável for do tipo discreto). O operador média é um operador linear, isto é, tem as seguintes propriedades:

a) $E\left[aX\right] = aE\left[X\right]$

b) $E\left[X+Y\right] = E\left[X\right] + E\left[Y\right]$

c) $E\left[aX+b\right] = aE\left[X\right] + E\left[b\right]$

Exemplo:

A variável aleatória X assume valores $x_1 = -1$, com probabilidade $p = 0{,}4$ e $x = +1$, com probabilidade $p = 0{,}6$. A média de $g(x) = X$ será:

$$E\left[X\right] = \sum_{k=1}^{2} x_k p_k = -1 \times 0{,}4 + 1 \times 0{,}6 = 0{,}2$$

(note que a média está compreendida entre os valores extremos que X assume).

Definição

A média (ou valor esperado) da função $g(X) = x^j$, isto é, $E\left[X^j\right] = \alpha_j$, é chamado momento de origem zero, de ordem j, da variação aleatória X.

A formula que dá o momento de origem zero, de ordem J, é:

$$\alpha_j = \sum_k x_k^j p_k$$

(note que, se j = 1, cae-se no caso da média da variável aleatória X)

Se a variável aleatória X for do tipo continuo , então $E\left[X^j\right] = \alpha_j$ é dado pela formula:

$$\alpha_j = E\left[X^j\right] = \int_{-\infty}^{+\infty} x^j f(x)\, dx$$

Momento de ordem j, de origem <u>c</u>, de uma variável aleatória X, é definido pela formula:

$$E\left[(X - c)^j\right] = \sum_k (x_k - c)^j\, p(x_k) ;$$

se a variável aleatória for do tipo contínuo, a formula de definição é:

$$E\left[(X - c)^j\right] = \int_{-\infty}^{+\infty} (x - c)^j\, f(x)\, dx$$

De importância extraordinária em Estatística é o 2º momento de origem $c=\alpha_1=\mu$, isto é, o 2º momento tomado em relação à média μ, que se chama variância. A sua importância decorre de uma propriedade deduzida da prófia definição. A formula de definição da variância é :

$$V^2[X]=\sigma^2=\sum_k (x_k-\mu)^2 p_k, \text{se a variável aleatória}$$

é discreta ou

$$V^2[X]=\sigma^2=\int_{-\infty}^{+\infty}(x-\mu)^2 f(x)dx$$

Demonstremos a seguinte propriedade: a soma dos quadrados dos desvios em relação à media é mínima; de fato,

$$\frac{dE[X-c]^2}{dc}=\sum_k 2(x_k-c)[-1]p_k ;$$

igualado a zero para achar a condição de extremo, tem-se:

$$c=\sum_k x_k p_k$$

que é a própria definição de média, o que demonstra que a soma dos quadrados dos desvios é mínima, se ela for tomada em relação à media e não em relação a qualquer outro valor de X.

Geometricamente, a média é uma medida de localização e a variância, uma medida de dispersão. Na prática, faz-se uso do desvio padrão, que é a raiz quadrada positiva da variância, isto é, $\sigma=+\sigma^2$.

Para se ter idéiaintuitiva da variância, podemos citar 2 exemplos:

1º) Uma turma de alunos tem 30 alunos e todos tem a mesma altura. Nesse caso, a média é igual à altura comum dos alunos e a variância é nula, pois não há desvio entre a altura dos alunos e a sua média e, portanto, a soma dos quadrados dos desvios é nula.

2º) Uma outra turma tem 30 alunos, o mais baixo tem 1,50m e o mais alto, 2,10m. Nesse caso, a média seria igual à soma das alturas de todos os alunos dividida por 30 e seria grande a variância, porque o intervalo entre a menor e a maior altura é grande.

13. COEFICIENTES DE CORRELAÇÃO

Sejam X e Y duas variáveis aleatórias. Se os desvios padrões dessas 2 variáveis, σ_1 e σ_2, existirem e forem diferentes de zero, defini-se o coeficiente de correlação ρ entre essas duas variáveis, que é um parametro que dá a relação de dependência estatística entre elas. O coeficiente de correlação é dado pela formula :

$$\rho = \frac{E\left[(X-\mu_1)\ (Y-\mu_2)\right]}{\sigma_1\ \sigma_2}$$

onde μ_1 e μ_2 são as médias de x e y respectivamente.

Demonstra-se que o coeficiente de correlação é um número compreendido no intervalo

$$\left[-1; +1\right]$$

Se as duas variáveis aleatórias são independentes, $\rho = 0$; se $\rho = 0$, diz-se que as variáveis aleatórias X e Y são não correlacionadas. Demonstra-se, por outro lado, que se X e Y estão ligados por uma relação da forma $Y = aX + b$, então $\rho^2 = 1$.

14. PRINCIPAIS DISTRIBUIÇÕES DE PROBABILIDADES DO TIPO DISCRETO

14.1 - Distribuição de Bernoulli

É a distribuição de uma variável aleatória X, que assume os valores 1 e 0, respectivamente com probabilidades p e q, cuja função de probabilidade é dada pela formula :

$$f(x) = p^x \quad (1-p)^{1-x}$$

ou seja

$$f(1) = p \text{ e } f(0) = 1 - p = q$$

A média e a variância da variável aleatória da distribuição de Bernoulli são respectivamente:

$$E[X] = \mu = \sum_{x=0}^{1} x p^x (1-p)^{1-x} = 0 + p = p$$

$$V^2[X] = \sigma^2 = \sum_{x=0}^{1} (x-p)^2 \, p^x (1-p)^{1-x} = \sum (x^2 - 2px + p^2) \, p^x (1-p)^{1-x}$$

$$= \sum_{x=0}^{1} x^2 p^x (1-p)^{1-x} - 2p \sum_{x=0}^{1} x p^x (1-p)^{1-x} +$$

$$+ p^2 \sum_{x=0}^{1} p^x (1-p)^{1-x} = p - 2p^2 + p^2 =$$

$$= p - p^2 = p(1-p).$$

14.2 - Distribuição Binomial

É a distribuição de uma variável aleatória X do tipo discreto, definida para x = 0, 1, 2, ..., n, cuja função de probabilidade é dada pela fórmula:

$$f(x) = P[X = x] = C_n^x \, p^x \, (1-p)^{n-x} \qquad (1)$$

Antes de achar os principais parâmetros dessa distribuição - média e variância - vamos estudar como se chegou a essa fórmula.

Para chegar ao estabelecimento da formula (1), imaginemos o seguinte problema, que envolve urnas e bolas: uma urna contém bolas brancas e bolas pretas, sendo p a proporção de bolas brancas na urna. Dessa urna são extraídas, com reposição, n bolas. É claro que, se a extração é aleatória, na amostra de tamanho n podem ocorrer 0, 1, 2, ..., n bolas brancas. Esses valores são aqueles que a variável aleatória X assume. O que se deseja é estabelecer uma formula que dê a probabilidade de ocorrerem $\underline{x}$ bolas brancas, dentre as n extraidas. Seja A_i o evento que consiste na extração de uma bola branca, na 1ª extração, e $\bar{A}_i$ o evento que consiste em extrair bola preta. Então $P[A_i] = p$ e $P[\bar{A}_i] = 1 - p = q$

Considere-se uma seqüência particular

$$\underbrace{A_1\ A_2\ A_3\ \ldots\ldots\ldots\ldots A_x}_{x}\ \underbrace{\bar{A}_1\ \bar{A}_2\ \bar{A}_3\ldots\ldots\ldots\ldots\bar{A}_{n-x}}_{n-x}$$

A probabilidade dessa seqüência é $p^x(1-p)^{n-x}$. Como as bolas brancas podem permutar de posição com as bolas pretas, a seqüência acima admite

$$\frac{n!}{x!\,(n-x)!}$$

variantes; logo, a probabilidade de que, dentre as n bolas, ocorram exatamente x brancas, é dada por:

$$f(x)=\frac{n!}{x!\,(x-x)!}\ p^x(1-p)^{n-x} = Cn^x\ p^x(1-p)^{n-x}$$

Conhecendo os valores que a variável aleatória assume e a sua função de probabilidade, estamos em condições de achar a média e a variância da distribuição binomial

$$\mu=E[X]=\sum_{x=0}^{n} x\,C_n^x\,p^x(1-p)^{n-x}=$$
$$=\sum_{x=1}^{n} x\frac{n!}{x!\,(n-x)!}p^x(1-p)^{n-x}=$$
$$=\sum_{x=1}^{n}\frac{n!}{(x-1)!\,(n-x)!}p^x(1-p)^{n-x}=$$
$$=\mu p\sum_{x=1}^{n}\frac{(n-1)!}{(x-1)!\,(n-x)!}p^{x-1}(1-p)^{n-x}=$$
$$=np\sum_{y=0}^{n-1}\binom{n-1}{y}p^y(1-p)^{n-1-y}=np,$$

que é média. A variância é calculada por meio de relação entre os momentos centrais e os momentos de origem zero:

$$\mu_2=\sigma^2=\alpha_2-\mu^2=\alpha_2-(np)^2.$$

Precisamos achar

$$\alpha_2=\sum_{x=0}^{n}x^2C_n^x p^x(1-p)^{n-x}=\sum_{x=0}^{n}[x(x-1)+x]\,C_n^x p^x(1-p)^{n-x}$$

Desenvolvendo esse somatório, encontramos:

$$\alpha_2 = n^2p^2 - np^2 + np \text{, e então}$$

$$\mu_2 = \alpha_2 - n^2p^2 = np - np^2 = np(1-p) =$$
$$= npq.$$

A distribuição binomial é tabelada, para valores de n e p, obtendo-se a propabilidade do evento X = x, para valores de x de 0 até n.

Tabelas completas da distribuição binomial foram publicadas pelo National Bureau of Standars e pela Universidade de Harvard e são de extrema utilidade em Controle de Qualidade.

A função de distribuição é dada por:

$$F(x) = P[X \leq x] = \sum_{x=0}^{x} C_n^x p^x (1-p)^{n-x}$$

Exemplos:

1º) Um dado é lançado 5 vezes. Qual é a probabilidade de se obter no máximo duas vezes a face 1?

$$P[X=2] = C_5^2 \left(\tfrac{1}{6}\right) 2 \left(\tfrac{5}{6}\right) 3 = 0{,}16$$

2º) No exemplo 1, qual a probabilidade de se obter no máximo duas vezes a face 1?

$$P[X \leq 2] = P[X=0] + P[X=1] + P[X=2] =$$
$$= \sum_{x=0}^{2} C_5^x \left(\frac{1}{6}\right)^x \left(\frac{5}{6}\right)^{5-x} = 0{,}40 + 0{,}40 + 0{,}16 = 0{,}96$$

3º) Uma indústria montadora compra grandes partidas de fusíveis e de cada partida, extrai uma amostra de tamanho n = 200 para submeter a inspeção de qualidade (fusível bom-fusível defeituoso). O comprador sabe que o produtor dos fusíveis, por contrato, deve fabricar com qualidade média de 0,02, isto é, a qualidade do processo de fabricação é tal que a probabilidade de ocorrer um fusível defeituoso é de 0,02.

O critério de aceitação do comprador é aceitar a partida, se no amostra ocorrerem no máximo 5 fusíveis defeituosos. Qual é essa probabilidade?

$$P\left[X \leq 5\right] = \sum_{x=0}^{5} C_{200}^{x} \quad 0{,}02^{x} \quad 0{,}98^{200-x}$$

= 0,788 (esse valor foi tirado diretamente das tabelas de Distribuição Binomial do N.B.S., entrando-se com n = 200 e p = 0,02). Seria praticamente impossível obter tal resultado com máquinas comuns de calcular ou a mão. Veremos mais adiante que a distribuição de Poisson pode ser usada com bastante precisão, como aproximação da distribuição binomial.

14.3 - Distribuição de Poisson

É a distribuição de uma variável aleatória X do tipo discreto, definida para x = 0, 1, 2, ...e cuja função de probabilidade é dada pela fórmula:

$$f(x) = P\left[X = x\right] = \frac{e^{-\lambda}\lambda^{x}}{x!}$$

onde λ é o parâmetro da distribuição.

A média da distribuição Poisson é dada por

$$\mu = \sum_{x=0}^{\infty} x \frac{e^{-\lambda}\lambda^{x}}{x!} = e^{-\lambda} \sum_{x=0}^{\infty} x \frac{\lambda^{x}}{x!}$$

isto é, média igual ao parâmetro da distribuição.

A variância $\sigma^2 = \alpha_2 - \mu = \alpha_2 - \lambda^2$

Precisamos achar:

$$\sigma_2 = \sum_{x=0}^{\infty} x^2 \frac{e^{-\lambda}\lambda^x}{x!} = \lambda + \lambda^2 \text{; então:}$$

$$\sigma_2 = \lambda + \lambda^2 - \lambda^2 = \lambda.$$

Este resultado é notável e altamente informativo, pois nos diz que na distribuição Poisson a média e a variância são iguais ao parâmetro λ.

A distribuição Poisson é obtida por um processo de passagem ao limite da distribuição binomial, fazendo-se n tender para infinito, ao mesmo tempo que se faz p tender para zero, de modo que o produto np se mantenha constante, isto é,

$$\lim_{\substack{n \to \infty \\ p \to 0 \\ np = \lambda = \text{constante}}} C_n^x \cdot p^x (1-p)^{n-x} = \frac{e^{-np}\, np^x}{x!}$$

A distribuição Poisson é usada na prática como uma aproximação bastante razoável da distribuição bimonial, quando $n \geq 50$ e $p \leq 0,10$. Vejamos alguns exemplos de aplicação .

1º) No exemplo nº 3 da Seção 14.2 qual a probabilidade de que a amostra contenha no máximo 5 fusíveis defeituosos?

$$P[X \leq 5] = \sum_{x=0}^{5} \frac{e^{-\lambda}\lambda^x}{x!} \text{; com}$$

$$\lambda = np = 200 \times 0{,}02 = 4$$

É claro que pode aplicar-se a distribuição Poisson em lugar da binomial pois :

$$n = 200\ (>50) \text{ e } p = 0{,}02\ (<0{,}10)$$

Então temos que

$$P[X \leq 5] = \sum_{x=0}^{5} \frac{e^{-4}\ 4^{x}}{x!} =$$

$$= e^{-4}\left[1+4+\frac{4^2}{2}+\frac{4^3}{6}+\frac{4^4}{24}+\frac{4}{120}\right] =$$

$$= 0,785$$

Comparando este resultado aproximado, com o resultado exato fornecido pela distribuição binomial (0,788), verifica-se que a aproximação pela Poisson é mais que satisfatória e de cálculo imediato.

2º) Uma folha de papel produzida por uma indústria de papel é submetida a inspeção e verifica-se que ela possue pequenas imperfeições, na forma de pequenas descontinuidades que lhe dão uma aparência de pequenos pontos sem côr. A folha é dividida em 400 retangulos iguais. A contagem das descontinuidades mostra que há retângulos em que não ocorrem imperfeições, retângulos em que só ocorre uma imperfeição, retângulos em que ocorre 2 imperfeições, etc. O investigador suspeita que a distribuição das imperfeições segue a lei de probabilidade de Poisson e decidi-se a investigar o fenômeno e constroi a tabela seguinte, que dá o número de retângulos em que ocorrem zero imperfeições, uma imperfeição, duas imperfeições, etc., e para testar se a distribuição é Poisson usa o resultado teórico de que a média e a variância devem ser iguais.

Nº de imperfeições (x)	0	1	2	3	4	5	6	7	8	9	10	$\sum$
Nº de retangulos (Freq. Observada)	103	143	98	42	8	4	2	0	0	0	0	400
Freq. Teórica	106	141	93	41	14	4	1	0	0	0	0	400

Para testar se a média da distribuição observada, $\overline{x}$ (veremos esta fórmula na Seção ....) é igual à variância da mesma distribuição, s^2 (veremos ésta fórmula também na seção.......), o investigador calcula $\overline{x}$ e s^2 pelas fórmulas:

$$\overline{x} = \frac{\sum_{i=1}^{10} x_i f_i}{n} =$$

$$= \frac{0\times105+1\times143+2\times98+3\times42+4\times8+5\times4+6\times2}{400} =$$

$$= 1,32 \quad \text{e}$$

$$s^2 = \frac{\sum_{i=1}^{10} (x_i - \overline{x})^2 f_i}{n} = 1,2901$$

Embora os 2 resultados não sejam exatamente iguais, são muito proximos. O investigador, no próximo passo, testa se p é realmente pequeno. Como $\lambda = np = 1,32$, tem-se que:

$$p = \frac{1,32}{400} = 0,0033$$

que é um valor relamente pequeno, O investigador tem elementos de convicção bastante fortes para suspeitar que o fenômeno sob análise segue a lei de Poisson e passa à penúltima etapa, que é a de calcular as freqüências teóricas. Isto ele faz usando a fórmula:

$$n \frac{e^{-1,32} \; 1,32^x}{x!}$$ e substituindo

$x = 0, 1, 2, 3, 4, 5,$ e 6,

obtendo os resultados mostrados na tabela, que são bastante aproximados dos resultados observados.

A última etapa consistiria em testar a hipótese de que o fenômeno segue a lei de Poisson, contra a hipótese alternativa de que o fenômeno não segue a mencionada lei. Isto será visto mais adiante.

14.4 - <u>Distribuição hipergeométrica</u>

É a distribuição de uma variável aleatória X do tipo discreto, definida para x = 0, 1, 2, ..., n e cuja função de probabilidade é dada pela fórmula:

$$f(x) = P\left[X = x\right] = \frac{C_M^x \, C_{N-M}^{n-x}}{C_N^n}$$

Vamos ver como se chegou a esta fórmula. - Imaginemos uma urna com N bolas, das quais M são brancas e as restantes N - M são pretas. A proporção de bolas brancas na urna é

$$p = \frac{M}{N} .$$

Dessa urna são extraidas n, bolas, sem reposição. Se a extração é aleatória, na amostra de tamanho n podem figurar 0, 1, 2, ... n, bolas brancas, que são os valores que a variável aleatória X assume. O que se deseja é estabelecer uma formula que dê a probabilidade de ocorrerem x bolas brancas, dentre as n extraidas. Seja A_i o evento que consiste em que na 1ª extração ocorra bola branca e $\overline{A}_i$ o evento oposto. É claro que, não havendo reposição, a constituição da urna se altera e a probabilidade de extrair bola branca também, em cada extração.

O Espaço Amostra da experiência é constituido de n bolas, das quais uma delas é a seguinte:

$$\underbrace{A_1 \, A_2 \, A_3 \ldots\ldots\ldots\ldots\ldots\ldots A_x}_{\text{x bolas brancas}} \; \underbrace{\overline{A}_1 \, \overline{A}_2 \, \overline{A}_3 \ldots\ldots\ldots \overline{A}_{n-x}}_{\text{n-x bolas pretas}}$$

A probabilidade dessa configuração é: $\dfrac{A_M^x \, A_{N-M}^{n-x}}{A_N^n}$

Como existem $\dfrac{n!}{x!\,(n-x)!}$ configurações no espaço amostra, que pertencem ao evento considerado, tem-se que

$$f(x) = P\left[X = x\right] = \frac{n!}{x!\,(n-x)!} \; \frac{C_M^x \, x! \, C_{N-M}^{n-x} \, (n-x)!}{C_N^n \, n!} = \frac{C_M^x \; C_{N-M}^{n-x}}{C_N^n}$$

como queríamos demonstrar.

A média e a variância da distribuição hipergeométrica são dadas respectivamente por:

$$E[X] = np \quad e \quad V^2[X] = \frac{N-n}{N-1}\, npq$$

Vê-se que a média da hipergeométrica é igual a média da binomial e a variância da hipergeométrica é igual à da binomial, multiplicada por um multiplicador finito. Vê-se que fazendo N tender para infinito, a variância da hipergeométrica tende para a variância da binominal. De fato, se o tamanho da urna crescer indefinidamente sendo finito o tamanho da amostra, a fração de amostragem $\frac{n}{N} \to 0$

A distribuição hipergeométrica é a distribuição principal do Controle de Qualidade por atributos.

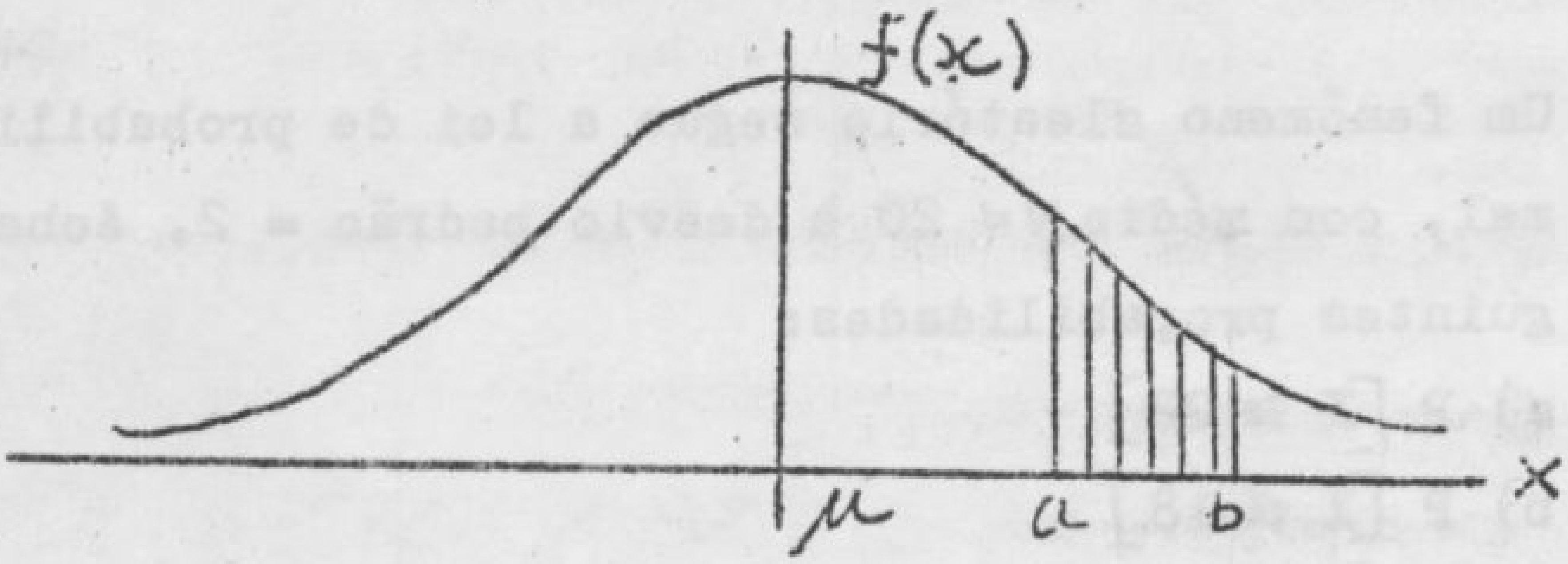
15. PRINCIPAIS DISTRIBUIÇÕES DE PROBABILIDADE DO TIPO CONTÍNUO

15.1 - Distribuição Normal

É a distribuição de uma variável aleatória X do tip contínuo, definida no intervalo ($-\infty$, $+\infty$), e cuja função de densidade de probabilidade é dada por

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

onde μ e σ são parâmetros, que representam respectivamente a média e o desvio padrão da distribuição. É fácil ver que f(x) é simétrica em relação a μ, pelo gráfico abaixo:



Chama-se atenção do aluno que a ordenada f(x) não é uma probabilidade e sim uma densidade de probabilidade.

Para achar a probabilidade de um intervalo [a, b,] tem-se que integrar a função f(x) entre esses 2 limites, isto é,

$$P\left[a \leq x \leq b\right] = \int_a^b f(x)\,dx =$$

$$= \int_a^b \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}}\, dx$$

Não é possível resolver esta integral, sendo necessário recorrer a métodos de integração numérica. A integral de f(x) é tabelada para

$$\mu = 0 \quad e \quad \sigma = 1$$

recorrendo-se a transformação

$Z = \dfrac{x-\mu}{\sigma}$, em que a variável aleatória Z assim definida tem média 0 e variância 1. Substituindo-se, tem-se:

$$P\left[a \leq x \leq b\right] = P\left[a-\mu \leq x-\mu \leq b-\mu\right] =$$

$$= P\left[\frac{a-\mu}{\sigma} \leq \frac{x-\mu}{\sigma} \leq \frac{b-\mu}{\sigma}\right] = P\left[z_1 \leq z \leq z_2\right] \int_{z_1}^{z_2} \frac{1}{\sqrt{2\pi}}\, e^{-\frac{z^2}{2}}\, dz$$

que é tabelada.

Exemplos: Um fenômeno aleatório segue a lei de probabilidade normal, com média $\mu = 20$ e desvio padrão = 2. Achar as seguintes propabilidades:

a) $P[X \geq 22]$

b) $P[X \leq 18]$

c) $P[16 \leq X \leq 24]$

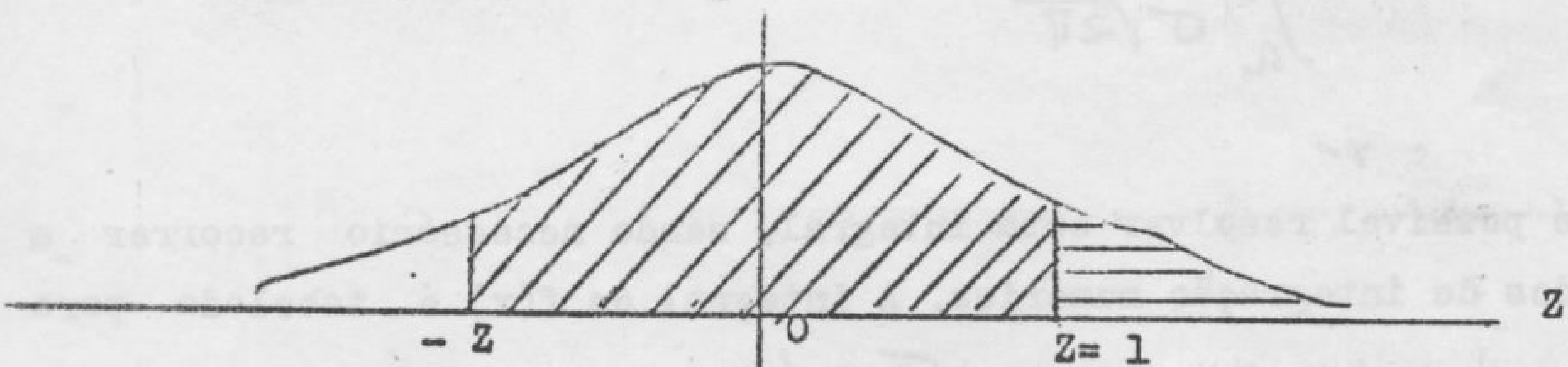
d) $P[22 \leq X \leq 24]$

Soluções: a) $P[X \geq 22] = P[X - 20 \geq 22 - 20] = P\left[\frac{X - 20}{2} \geq \frac{2}{2}\right] = P[z \geq 1]$

Procurando na tabela de distribuição, encontramos 0,15865.

Notar que a tabela da probabilidade,

$P[-z_1 \leq z \leq z_1]$ ou seja



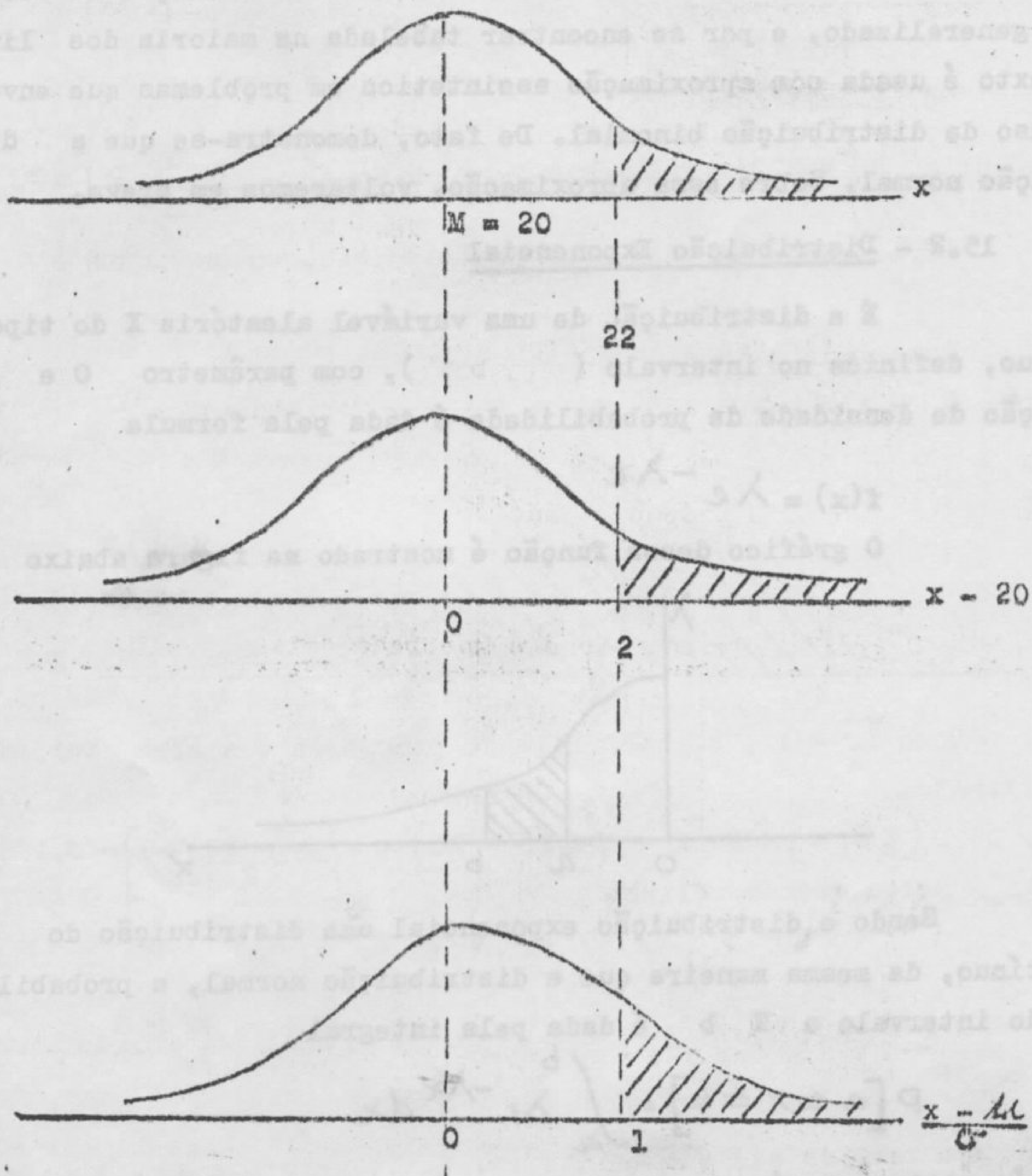
Como $z_1 = 1$, a probabilidade procurada é a área tarjada em horizontal

b) $P[X \leq 18] = P[X - 20 \leq 18 - 20] = P\left[\frac{x - 20}{2} \leq -1\right] = 0,15865$

pelos mesmos motivos

Há tabelas que dão a integral desde zero até a abcissa procurada e outras que dão a referida integral desde $-\infty$ até a abcissa mencionada.

Nesse exemplo, como em problemas reais, é sempre interessante desenhar a figura da distribuição normal, assinalando-se o intervalo, cuja probabilidade se pr cura. Assim, para achar $P[X \geq 22]$ faz-se as seguintes figuras:



A área tarjada, nas 3 figuras, é a probabilidade procurada. Na última figura, a bcissa

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

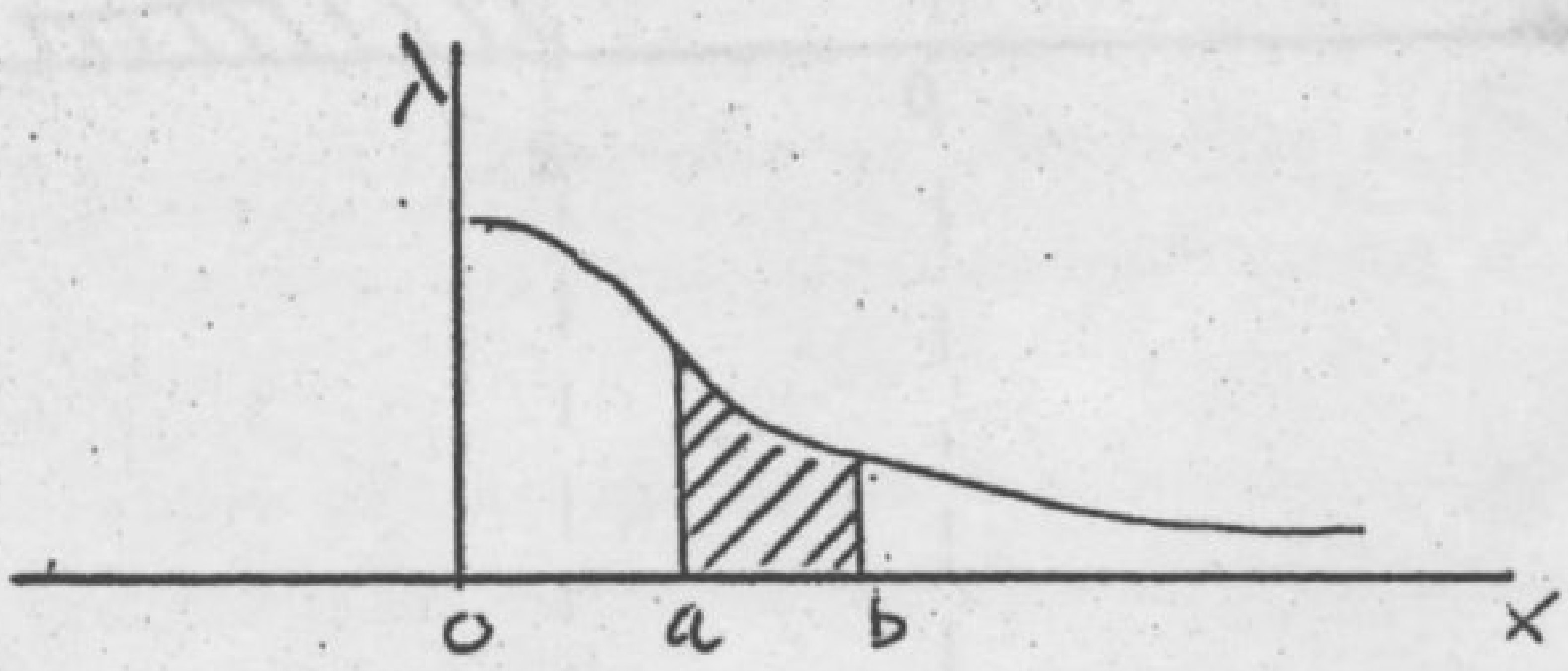
é a abcissa de entrada na tabela da distirbuição normal, mostrada no Anexo 1. A distribuição normal, por ser uma distribuição de uso muito generalizado, e por se encontrar tabelada na maioria dos livros-texto é usada com aproximação assintetica em problemas que envolvem o uso da distribuição binomial. De fato, demonstra-se que a distribuição normal. Sobre essa aproximação, voltaremos em breve.

15.2 - Distribuição Exponencial

É a distribuição de uma variável aleatória X do tipo contínuo, definida no intervalo (, b), com parâmetro 0 e cuja função de densidade de probabilidade é dada pela formula

$$f(x) = \lambda e^{-\lambda x}$$

O gráfico dessa função é mostrado na figura abaixo



Sendo a distribuição exponencial uma distribuição do tipo contínuo, da mesma maneira que a distribuição normal, a probabilidade do intervalo a X b é dada pela integral

$$P\left[a \leq x \leq b\right] = \int_a^b \lambda e^{-\lambda x}\, dx$$

que é medida pela área tarjada.

A distribuição exponencial aparece em muitos problemas reais, em que se aplica a teoria das probabilidades. Em pesquisa Operacional, é a distribuição comumente usada para exprimir o tempo de atendimento em postos de serviço.

Os parâmetros dessa distribuição são:

média $\mu = E[X] = \int_0^\infty x \lambda e^{-x\lambda} dx = \lambda \int_0^\infty x e^{-\lambda x} dx = \frac{1}{\lambda}$

Variância $\sigma^2 = V^2[X] = \int_0^\infty (x-\mu)^2 \lambda e^{-\lambda x} dx = \frac{1}{\lambda^2}$

16. CORRELAÇÃO E REGRESSÃO

As revistas de economia, os jornais especializados em ecomia e muitos estudos de viabilidade costumam dizer que existe uma intima relação entre o Produto Interno Bruto (PIB) e o consumo de energia elétrica. Isto nada mais é do que o resultado da observação ao longo do tempo, do comportamento conjunto dessas duas variáveis. De fato, se um período bastante longo, observamos os pares (PIB , Consumo de energia elétrica), veremos para os mesmos que os valores do PIB e do consumo de energia elétrica crescem no mesmo sentido.

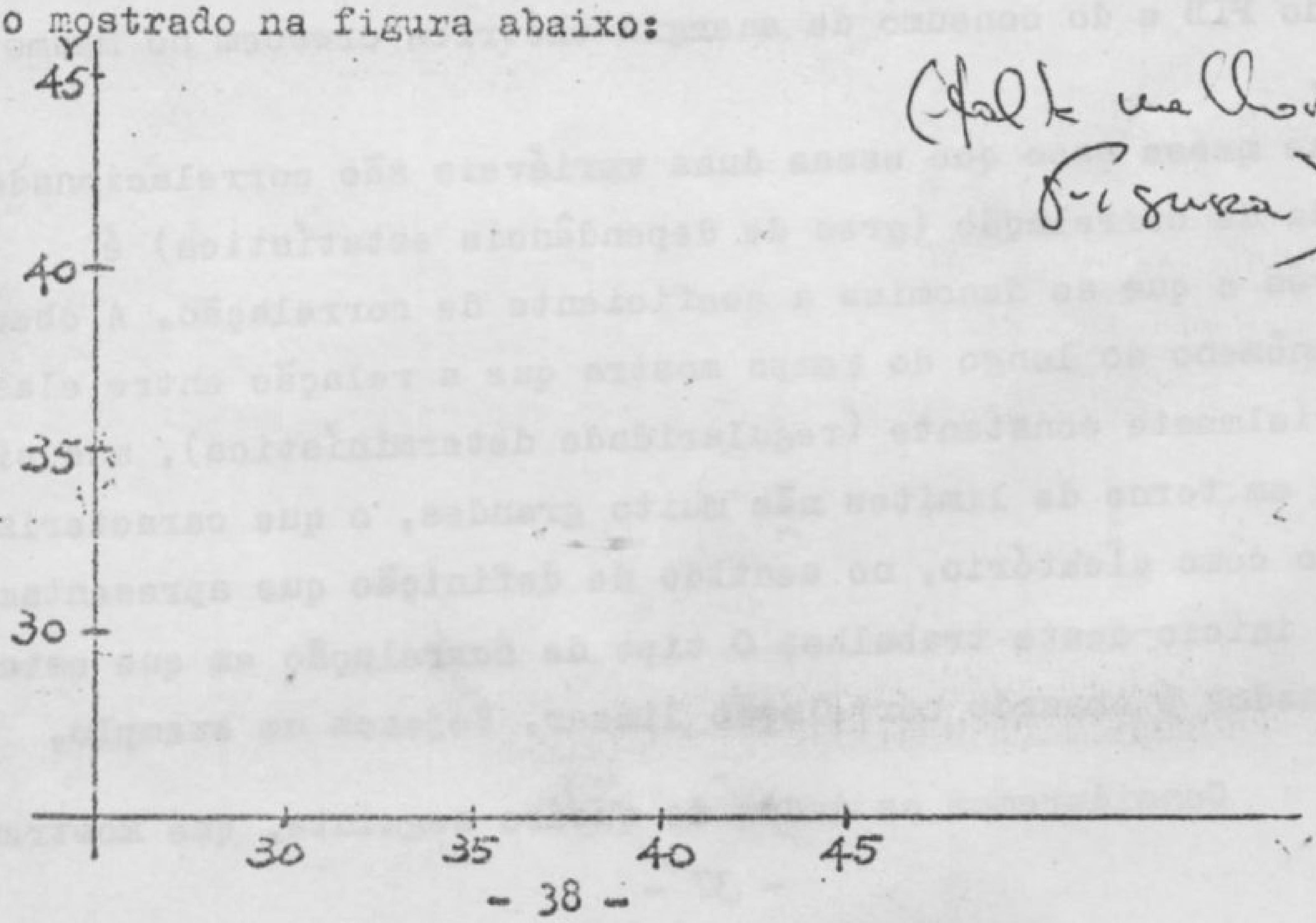
Diz-se nesse caso que essas duas variáveis são correlacionadas e a medida de correlação (grau de dependência estatística) é obtida através o que se denomina a coeficiente de correlação. A observação do fenômeno ao longo do tempo mostra que a relação entre elas não é invarialmente constante (regularidade determinística), mas sim que varia em torno de limites não muito grandes, o que caracteriza o fenômeno como aleatório, no sentido de definição que apresentamos logo no inicio deste trabalho. O tipo de correlação em que estamos interessados é chamado correlação linear. Vejamos um exemplo.

Consideremos os dados do quadro seguinte, que mostram as notas de 30 alunos, aos quais foram aplicados 2 testes, um de linguagem e um de ciências. As notas de linguagem estão na coluna X e as de ciências, na coluna Y, para uma nota máxima 50.

X	Y	X	Y	X	Y
34	37	28	30	39	36
37	37	30	34	33	29
36	34	32	30	30	29
32	34	41	37	33	40
32	33	38	40	43	42
36	40	36	42	31	29
35	39	37	40	38	40
34	37	33	36	34	31
29	36	32	31	36	38
35	35	33	31	34	32

Com os dados do quadro, foi construido o diagrama de dispersão mostrado na figura abaixo:

(Falta na xerox a figura)



A inspeção do diagrama de dispersão mostra que há uma tendência de os menores valores de X estarem associados a menores valores de Y e vice-versa. Além disto, a tendência geral do diagrama de dispersão é uma linha reta. Para variáveis como as citadas, é desejável medir-se o grau com que elas estão lineramente relacionadas. Essa medida, como já foi dito, é o coeficiente de correlação, que nos casos práticos em que se tem os pares.
(Xi, Yi , i = 1, 2, ...n, é dado pela fórmula.

$$r = \frac{\sum_{i=1}^{n} (x_i - \bar{X})(Y_i - \bar{Y})}{n s_x s_y}$$

onde $\bar{X}$, $\bar{Y}$, s_x e s_y são respectivamente, a medida de X, a média de Y, o desvio-padrão de X e o desvio-padrão de Y. Uma formula mais simples de operar que a precedente, é a seguinte:

$$r = \frac{n \sum X_i Y_i - (\sum X_i)(\sum Y_i)}{\sqrt{[n \sum X_i^2 - (\sum X_i)^2][n \sum Y_i^2 - (\sum Y_i)^2]}}$$

A aplicação desta formula ao exemplo do quadro anterior, leva ao valor de r= 0,66

Para aplicação desta formula, é útil e operacionalmente desejável construir um quadro da forma do quadro abaixo:

X_i	Y_i	$X_i Y_i$	X_i^2	Y_i^2
X_1	Y_1	X_1Y_1	X_1^2	Y_1^2
X_2	Y_2	X_2Y_2	X_2^2	Y_2^2
X_3	Y_3	X_3Y_3	X_3^2	Y_3^2
X_n	Y_n	X_nY_n	X_n^2	Y_n^2
$\sum X_i$	$\sum Y_i$	$\sum X_iY_i$	$\sum X_i^2$	$\sum Y_i^2$

Nos problemas de correlação (simples ou múltipla) procura-se sempre medir o grau de dependência estatística entre duas ou mais variáveis. No caso anterior, medimo-lo entre 2 variáveis X e Y.

Uma outra formula que não é operacional, mas serve para mostrar os limites de variação de r, é a seguinte:

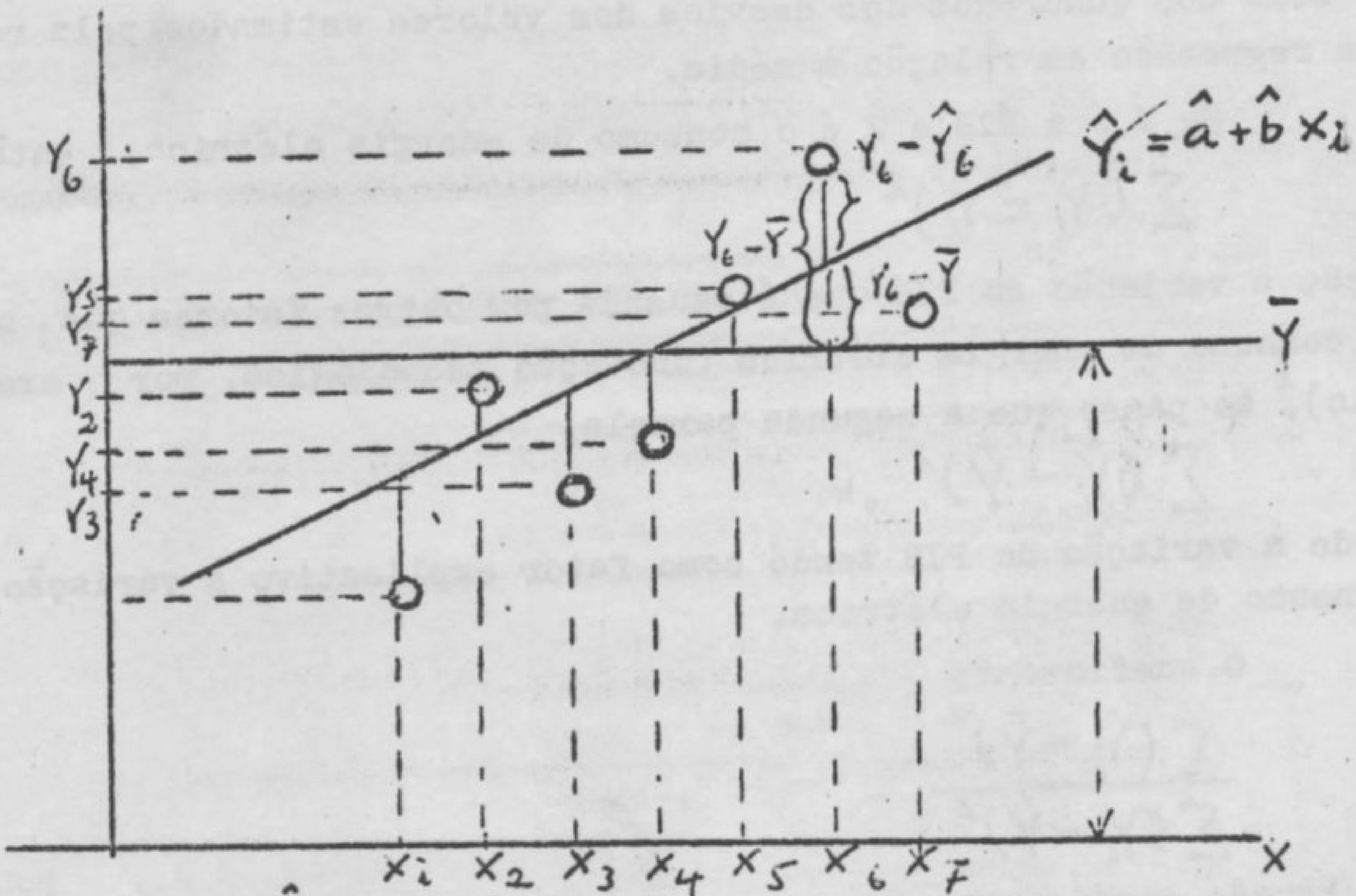
$$r=\sqrt{1-\frac{\sum_{i=1}^{n}(Y_i-\hat{Y}_i)^2}{\sum_{i=1}^{n}(Y_i-\overline{Y})^2}}$$

onde $\overline{Y}$ é a média de Y e $\hat{Y}_i$ são os valores estimados de Y_i, por regressão linear, como veremos na seção seguinte.

A formula mostra que $r \in [-1, +1]$ e será máximo em valor absoluto, se $\sum(Y_i-\hat{Y}_i)^2$ for zero.

A interpretação de r é simples. Se ele é zero, não existe correlação entre X e Y e o conhecimento de X em nada nos ajuda na predição de Y. Se ele é + 1 ou - 1, isto significa que os pontos (X_i , Y_i) ficam exatamente sobre uma linha reta ascendente ou descendente, no diagrama de dispersão, nesses casos, pode-se fazer predições em Y, conhecendo-se a evolução de X no tempo. Para valores intermediários de r, a interpretação já não é tão simples. Seria errôneo pensar que um coeficiente de correlação de 0,80 é duas vezes tão bom quanto um coeficiente de correlação de 0,40.

Procuraremos interpretar tais valores intermediários. Suponha-se que o diagrama de dispersão de um conjunto de pares (X_i , Y_i) tenha a apresentação seguinte:



A reta $\hat{Y}_i = \hat{a} + \hat{b}X_i$ é chamada reta de regressão e a maneira como estimá-la, veremos na seção seguinte. Por ora, o que interessa, é observar que entre os valores dados por $\hat{Y}_i = \hat{a} + \hat{b}X_i$ e os valores Y_i existe uma diferença $\hat{Y}_i - Y_i$ e também que existe uma diferença entre os valores Y_i e $\bar{Y}$, dada por $Y_i - \bar{Y}$. Na expressão do coeficiente de correlação aparecem as somas dos quadrados

$$\sum (Y_i - \hat{Y}_i)^2 \quad e \quad \sum (Y_i - \bar{Y})^2$$

Por meio de manipulação algébrica, demonstra-se que:

$$\sum (Y_i - \bar{Y})^2 = \sum (Y_i - \hat{Y}_i)^2 + \sum (\hat{Y}_i - \bar{Y})^2$$

isto é, a soma dos quadrados desvios das observações em relação à média é igual à soma dos quadrados de duas parcelas: a primeira, a soma dos quadrados dos desvios dos valores observados em relação aos valores estimados pela reta de regressão e a segunda, a soma dos quadrados dos desvios dos valores estimados pela reta de regressão em relação à média.

Se Y é o PIB e X é o consumo de energia elétrica, então

$$\sum (Y_i - \hat{Y}_i)^2$$

mede a variação do PIB que é causada por outros fatores que não o consumo de energia elétrica (inovação tecnológica, por exemplo), ao passo que a segunda parcela,

$$\sum (\hat{Y}_i - \bar{Y})^2 ,$$

mede a variação do PIB tendo como fator explicativo a variação do consumo de energia elétrica.

O coeficiente

$$\frac{\sum (\hat{Y}_i - \bar{Y})^2}{\sum (Y_i - \bar{Y})^2}$$

é chamado coeficiente de determinação, é igual ao quadrado do coeficiente de correlação e mede o quanto a variação de Y é explicada por X somente.

Realmente,

$\sum (\hat{Y}_i - \bar{Y})^2$, chama-se variância explicada e

$\sum (Y_i - \hat{Y}_i)^2$,

variância não explicada. A soma das duas variâncias, $\sum (Y_i - \bar{Y})^2$ é chamada variância total. Como a explicação de X sobre Y é dada pelo quadrado do coeficiente de correlação, isto significa que um coeficiente de correlação 0,80 é 4 vezes tão bom (no sentido de explicação) quanto um coeficiente de correlação de 0,40.

17 - CONTROLE ESTATÍSTICO DA QUALIDADE

17.1 - Utilização das Military Standard 105 e 414

A Military Standard 105 é o resultado de um estudo conduzido por um Grupo de Trabalho Americano-Britânico-Canadense (ABC) que procurou desenvolver um padrão comum para as 3 Nações. Foi liberada pelo Governo Americano em 1963. A Military Standard 414 é a equivalente da MILSTD 105 D, para inspeção por variáveis.

O ponto focal da MILSTD 105 D, bem como da MILSTD 414 é o Nível de Qualidade Aceitável (Acceptable Quality Level - AQL, que é a percentagem máxima de defeituosos que, para os propósitos de inspeção por amostragem, pode ser considerado satisfatório como média de um processo. Na aplicação dos MILSTD 105 D e 414 espera-se que em uma conferência de alto nível entre um fornecedor e um consumidor, o fornecedor será informado pelo consumidor o que este considera ser o Nível de Qualidade Aceitável para uma característica de um dado produto.

Na MILSTD 105 são previstos AQL de 0,01 até 10% e na 414, de 0,04 até 15%. Outro ponto importante nos MILSTD 105 D e 414 é o Nível de Inspeção (Inspection Level). O Nível de Inspeção determina uma relação entre o tamanho do lote e o tamanho da amostra. A decisão quanto ao Nível de Inspeção a ser usado varia de acordo com o produto a ser inspecionado. O Nível de Inspeção, quanto mais alto maior proteção oferece ao consumidor, diminuindo o risco de aceitar um lote defeituoso. Portanto, quanto maior o Nível de Inspeção, maior o tamanho da amostra, para um dado tamanho de lote.

17.2 - Inspeção por atributos

É a seguinte a rotina para emprego da Military Standard 105-D na inspeção por atributos:

a) Especificar o TAMANHO DO LOTE

É um número conhecido, pois é declarado na Nota Fiscal. O consumidor pode, através entendimentos com o fabricante, estabelecer os tamanhos de lotes desejados, para cada item adquirido.

De um modo geral, este número deve corresponder ao Lote Econômico.

b) Estabelecer, para o item a ser inspecionado, o Nível de Inspeção, que pode ser ATENUADO (NÍVEL I), NORMAL (NÍVEL II) ou RIGOROSO (NÍVEL III).

O estabelecimento deste nível não é uma decisão unilateral do consumidor e sim resultante de entendimentos entre as duas partes. De um modo geral a INSPEÇÃO deve ser iniciada em NÍVEL II.

c) Com o tamanho do Lote e com o NÍVEL DE INSPEÇÃO, determina-se através uma tabela de correspondência, a LETRA DE CÓDIGO DO TAMANHO DA AMOSTRA, que indica o número de unidades que devem ser extraídas do lote, para serem inspecionadas.

d) Especificar o NÍVEL DE QUALIDADE ACEITÁVEL (AQL)

Este Nível é aquele que o consumidor considera aceitável e deve ser informado ao fabricante. Caso este último concorde, fica então estabelecido como aceitável por ambas as partes.

e) Decidir quanto ao PLANO DE AMOSTRAGEM a ser empregado. A Mil. Std. 105-D permite o emprego de Plano de Amostragem simples, duplos ou múltiplos. A escolha do PLANO a ser empregado depende de conveniências administrativas e a decisão cabe ao Órgão responsável pelo Controle de Qualidade, dentro da estrutura administrativa.

f) Entrar na TABELA correspondente ao NÍVEL DE INSPEÇÃO e ao PLANO DE AMOSTRAGEM. Com a LETRA DE CÓDIGO DO TAMANHO DA AMOSTRA e o NÍVEL DE QUALIDADE ACEITÁVEL obtém-se o tamanho da amostra e os números de aceitação e rejeição correspondentes.

g) Aceitar ou rejeitar o lote, de acordo com os resultados da inspeção.

Segue-se um exemplo:

Recebido um lote de 10.000 unidades

NÍVEL DE INSPEÇÃO - NORMAL

NÍVEL DE QUALIDADE ACEITÁVEL - 4%

Determinar tamanho da amostra e os números de aceitação e rejeição.

Solução

Com o NÍVEL DE INSPEÇÃO NORMAL E TAMANHO DO LOTE 10000, a Mil. Std 105-D fornece a LETRA DE CÓDIGO DO TAMANHO DA AMOSTRA L.

Se for usado um PLANO DE AMOSTRAGEM SIMPLES, entra-se na Tabela correspondente a PLANO DE AMOSTRAGEM SIMPLES e NÍVEL DE INSPEÇÃO NORMAL, obtendo-se:

TAMANHO DA AMOSTRA: 200

NÚMERO DE ACEITAÇÃO: 14

NÚMERO DE REJEIÇÃO: 15

A técnica de inspeção então é a seguinte:

1º) Do lote de 10 000 unidades, são retiradas ao acaso 200, verificando-se as defeituosas. Se o número de defeituosas for menor ou igual que 14, o lote é aceito; se for maior ou igual que 15, o lote é rejeitado.

Se for decidido empregar um Plano de Amostragem Dupla, então entra-se na Tabela de NÍVEL DE INSPEÇÃO NORMAL, e PLANO DE AMOSTRAGEM DUPLA, com a mesma LETRA DE CÓDIGO DO TAMANHO DA AMOSTRA (L) e o mesmo NÍVEL DE QUALIDADE ACEITÁVEL (4%).

Obtém-se então 2 tamanhos de amostra e dois números de aceitação e rejeição que são os seguintes:

Tamanho da 1ª amostra	:	125
Tamanho da 2ª amostra	:	125
Número de aceitação da 1ª amostra	:	7
Número de aceitação da 2ª amostra	:	18
Número de rejeição da 1ª amostra	:	11
Número de rejeição da 2ª amostra	:	19

Então a técnica de inspeção é a seguinte:

1º) Extrai-se do lote de 10.000 unidades, uma amostra de 125 unidades; se o número de defeituosas for menor ou igual que 7, aceita-se o lote sem precisar extrair uma segunda amostra; se o número de defeituosas na 1ª amostra for maior ou igual que 11, rejeita-se o lote sem necessidade de extrair uma segunda amostra. Finalmente, caso o número de defeituosas na 1ª amostra seja maior que 7 e menor que 11, é então necessário extrair uma 2ª amostra.

Extraindo-se a 2ª amostra, também de 125 unidades, conta-se o número de defeituosas encontradas: se a soma de defeituosas encontradas na primeira amostra, mais as encontradas na 2ª amostra for menor ou igual que 18, aceita-se o lote; caso seja maior ou igual que 19, rejeita-se definitivamente o lote.

17.3 - <u>Inspeção por variáveis</u>

É a seguinte a rotina para o emprego da Military Standard 414 na inspeção por variáveis.

a) O mesmo que para a Mil Std 105-D (TAMANHO DO LOTE)

b) O mesmo que para a Mil Std 105-D (NÍVEL DE INSPEÇÃO)

A única diferença reside em que na Mil Std 414 existem 5

NÍVEIS DE INSPEÇÃO.

NÍVEL I

NÍVEL II

NÍVEL III - ATENUADO

NÍVEL IV - NORMAL

NÍVEL V - RIGOROSO

No Nível IV é designado como NORMAL e deve normalmente ser usado, a menos que seja especificado o contrário.

Os níveis I e II são empregados para aqueles casos em que o custo de inspeção por unidade é muito elevado, quando comparado ao custo unitário da unidade.

c) O mesmo que para a Mill Std 105-D (LETRA DE CÓDIGO DO TAMANHO DA AMOSTRA).

d) O mesmo que para a Mil Std 105-D (NÍVEL DE QUALIDADE ACEITÁVEL).

e) Identificar se na inspeção por amostragem a ser realizada a variabilidade é conhecida ou desconhecida. Considerar-se-á sempre que a variabilidade da característica do item pertencente ao lote é sempre desconhecido, que é o caso normal que se apresenta na prática. Mesmo no caso de ser conhecida a variabilidade, será ela tratada como desconhecida, pois a inspeção por amostragem servirá como verificação de variabilidade.

f) Idenficar se o tipo de inspeção refere-se a um único limite de tolerância ou a dois limites de tolerância.

g) Com os elementos dos itens procedentes, entrar na Tabela apropriada de Mil Std 414, correspondente ao NÍVEL DE INSPEÇÃO PARA PLANOS BASEADOS EM VARIABILIDADE DESCONHECIDA; MÉTODO DO DESVIO PADRÃO, UM LIMITE DE TOLERÂNCIA OU DOIS LIMITES, conforme o caso. Nessa tabela, entrar com a LETRA DE CÓDIGO DO TAMANHO DA AMOSTRA, obtendo o tamanho da amostra e o número K (se se tratarem de dois LIMITES DE TOLERÂNCIA). Estes números dão os critérios para aceitação ou rejeição do lote.

h) Extrair a amostra, cujo tamanho foi determinado em g e decidir quanto à aceitação ou rejeição do lote.

Os exemplos a seguir esclarecem o emprego das tabelas e como decidir quanto à aceitação ou rejeição do lote.

Este exemplo diz respeito ao exame de um lote de material em que só é especificado o limite de tolerância inferior, que chamaremos L. O exemplo é o seguinte:

L = 20 000

NÍVEL DE INSPEÇÃO : NORMAL

NÍVEL DE QUALIDADE ACEITÁVEL : 2,5%

TAMANHO DO LOTE : 250

Supõe-se desconhecida a variabilidade.

Com o NÍVEL DE INSPEÇÃO NORMAL (IV) e TAMANHO DO LOTE 250, determina-se a LETRA DE CÓDIGO DO TAMANHO DA AMOSTRA H.

Entrando na Tabela apropriada encontra-se para tamanho da amostra n= 20 e K = 1,51.

Extrai-se então uma amostra de tamanho 20 e encontram-se as seguintes medidas para a característica investigada, que chamaremos X:

22.030; 21.800; 20.980; 20.750; 21.480;
20.570; 21.110; 20.270; 18.970; 22.110;
22.740; 21.220; 21.300; 21.920; 21.050;
21.780; 21.800; 21.580; 20.990; 21.740;

Com as medidas acima calculam-se a média, e a variância e o desvio padrão pelas fórmulas abaixo:

$$X = \frac{X_i}{n} = 21\ 309{,}5$$

$$s^2 = \frac{X_i^2 - \frac{(X_i)^2}{n}}{} = 64\ 783$$

$$s = s^2 = 804{,}9$$

Em seguida calcula-se

$$Z_L = \frac{\overline{X} - L}{S} = 1{,}63$$

Se Z_L K, aceita-se o lote, caso contrário, rejeita-se.

Como Z_L K, então aceita-se o lote.

No exemplo mostrado, foi especificado o limite de tolerância inferior L. Caso fosse especificado o limite de tolerância superior U, as etapas seriam exatamente as mesmas com a diferença de que o valor de Z, Z_U, ao invés de ser calculado pela fórmula.

$Z_L = \frac{\overline{X} - L}{s}$, ou seria pela fórmula.

$$Z_U = \frac{U - \overline{X}}{s}$$

O critério de aceitação seria Z_U K e não Z_L K

Exemplo para o caso em que sejam especificados 2 limites de tolerância, um inferior L = 20.000 e outro superior.

U = 23 000

NÍVEL DE INSPEÇÃO : IV

NÍVEL DE QUALIDADE ACEITÁVEL : 2,5 %

TAMANHO DO LOTE : 250

Acha-se a LETRA DE CÓDIGO DO TAMANHO DA AMOSTRA H.

Entrando na Tabela apropriada (NÍVEL DE INSPEÇÃO NORMAL PLANOS BASEADOS EM VARIABILIDADE DESCONHECIDA) encontra-se n = 20 (TAMANHO DA AMOSTRA) e M = 6,17%

Extrai-se então uma amostra de tamanho n= 20 e acham-se os valores de característica investida X, que são os mesmos antes indicados. Calculam-se $\overline{X}$, s^2 e s.

A seguir, calculam-se $Z_L = \frac{\overline{X} - L}{s}$ e $Z_U = \frac{U - \overline{X}}{s}$.

Se os valores de Z_U ou Z_L ou ambos forem negativos, rejeita-se o lote; caso sejam positivos, é necessário uma condição a mais a ser satisfeita.

Para encontrar a condição, tem-se que determinar as percentagens defeituosas P_L e P_U, correspondente a Z_L e Z_U.

Calculando Z_L , encontra-se 1,63

Calculando Z_U , encontra-se

$$Z_U = \frac{23.000 - 21.309,5}{804,9} = 2,10$$

Os valores de P_L e P_U , correspondentes a Z_L e Z_U são obtidos na figura 12.3 (Chart for determining from Z: Standard Deviation Plans) constante do exemplar das " Military Standard " .

entrando-se com os valores de Z e de n.

Obtém-se:

$P_L = 4,75$

$P_U = 1,50$

Finalmente o critério de aceitabilidade indica que, se $P_L + P_U$ for menor ou igual que M, então aceita-se o lote.

No caso, $P_L + P_U = 6,25 \quad 6,17$

O lote é rejeitado.

Busse, Adolpho Gomes

Noções fundamentais da teoria das probabilidades aplicações

1-E-80

(272/86)

EGN 147