

**UNIVERSIDADE FEDERAL DO PARANÁ  
SETOR DE CIÊNCIAS EXATAS  
DEPARTAMENTO DE ESTATÍSTICA**

**APOSTILA DE  
ESTATÍSTICA APLICADA A  
ENGENHARIA**

**SONIA ISOLDI MARTY GAMA MÜLLER  
2008**

## **APRESENTAÇÃO**

**A Estatística é uma ferramenta imprescindível a qualquer pesquisador ou pessoa que necessite tomar decisões. O seu estudo não representa uma tarefa muito fácil, principalmente no início, quando são apresentados muitos conceitos novos que exigem um tipo especial de raciocínio.**

**Uma boa base teórica é importante e necessária para que o estudo da Estatística seja prazeroso e não muito sofrido.**

**Com o intuito de facilitar a apresentação e aprendizado da Estatística desenvolveu-se este material para servir de apoio didático à disciplina de Estatística II, apresentada aos alunos dos cursos de Engenharia.**

**Sonia Isoldi Marty Gama Müller**

## ÍNDICE

<b>I- INTRODUÇÃO</b> .....	<b>4</b>
<b>II- ESTATÍSTICA DESCRITIVA</b> .....	<b>5</b>
<b>III - PROBABILIDADE</b> .....	<b>12</b>
<b>IV- AMOSTRAGEM</b> .....	<b>35</b>
<b>V- ESTIMAÇÃO DE PARÂMETROS</b> .....	<b>39</b>
<b>VI-TESTES DE HIPÓTESES</b> .....	<b>42</b>
<b>VI-TESTES DE HIPÓTESES</b> .....	<b>42</b>
<b>VII- ANÁLISE DA VARIÂNCIA (COMPARAÇÃO DE VÁRIAS MÉDIAS)</b> .....	<b>51</b>
<b>VIII- REGRESSÃO E CORRELAÇÃO LINEAR SIMPLES</b> .....	<b>54</b>
<b>BIBLIOGRAFIA</b> .....	<b>58</b>
<b>TABELAS</b> .....	<b>59</b>

## I- INTRODUÇÃO

### 1.1 DEFINIÇÕES .

#### 1.1.1. ESTATÍSTICA:

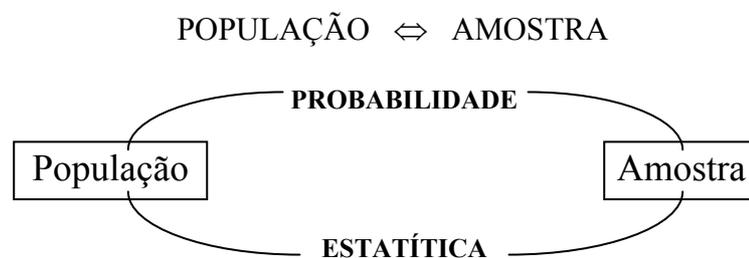
A Estatística refere-se às técnicas pelas quais os dados são "coletados", "organizados", "apresentados" e "analisados".

Pode-se dividir a ciência Estatística em dois grupos de estudo:

1. Estatística Descritiva: refere-se as técnicas de sintetização, organização e descrição de dados.
2. Estatística Inferencial: compreende as técnicas por meio das quais são tomadas decisões sobre a população baseadas na observações de amostras.

População é o conjunto "Universo" dos dados sobre os quais se quer estudar.

Amostra é um subconjunto da população que contenha todas suas propriedades.



#### 1.1.2 VARIÁVEL

Variável é uma característica da população à ser estudada.

Tipos de Variáveis: qualitativas e quantitativas  
discretas e contínuas

**1.1.2.1. QUALITATIVAS** : quando resultar de uma classificação por tipo ou atributo.

Ex: Pop.: canetas fabricadas  
Var.: cor (azul, vermelha, preta, etc)

**1.1.2.2. QUANTITATIVAS**: quando seus valores forem expressos em quantidades.

Podem-se dividir em dois tipos variáveis quantitativas discretas e variáveis quantitativas contínuas.

**1.1.2.2.1. VARIÁVEIS DISCRETAS**: quando seus valores forem expressos por números inteiros.

Ex.: Pop.: pessoas atendidas em um caixa de banco  
Var.: número de pessoas por sexo

**1.1.2.2.2. VARIÁVEIS CONTÍNUAS**: quando seus valores forem expressos em intervalos

Ex.: Pop.: salários de empregados de uma empresa  
Var.: valores (US\$)

## II- ESTATÍSTICA DESCRITIVA

A Estatística Descritiva tem como objetivo a organização e descrição de dados experimentais.

EXEMPLO: Tempo de atendimento (min.) aos clientes por um vendedor de uma loja de materiais de construção.

Dados brutos:	3,5	1,9	2,1	1,6	3,1
	1,0	1,4	1,8	1,2	1,3
	0,8	1,1	0,5	2,5	1,3
	0,7	1,7	1,4	1,3	1,6
Rol:	0,5	0,7	0,8	1,0	1,1
	1,2	1,3	1,3	1,3	1,4
	1,4	1,6	1,6	1,7	1,8
	1,9	2,1	2,5	3,1	3,5

### 2.1 TABELAS:

#### 2.1.1 AMPLITUDE TOTAL:

$$AT = \text{Máx.} - \text{Mín.}$$

No exemplo:  $AT=3,5 - 0,5 = 3,0$

#### 2.1.2 NÚMERO DE CLASSES:

Fórmula de Sturges para o cálculo do número de classes:

$$k = 1 + (3,3 \times \log n)$$

onde:

k = número de classes

n = número de dados disponíveis

No exemplo:

$$k = 1 + 3,3 \log 20 = 5,29 \quad (5 \text{ ou } 6 \text{ classes})$$

#### 2.1.3 AMPLITUDE DE CLASSE:

$$AC = \frac{AT}{k}$$

### 2.1.4 PONTO MÉDIO DE CLASSE

$$x_i = \frac{LI_i + LS_i}{2}$$

onde:

$x_i$  = ponto médio da classe i

$LI_i$  = limite inferior da classe i

$LS_i$  = limite superior da classe i

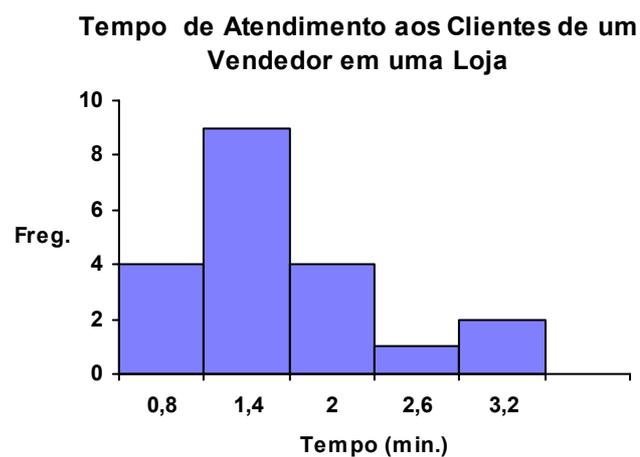
#### TEMPO DE ATENDIMENTO AOS CLIENTES DE UM VENDEDOR EM UMA LOJA

CLASSE (Tempo)	FREQUÊNCIA ABSOLUTA $f_i$	FREQUÊNCIA RELATIVA $F_r$	FREQUÊNCIA ACUMULADA $F_a$	PONTO MÉDIO $x_i$
0,5  —— 1,1	4	0,20 (20%)	4	0,8
1,1  —— 1,7	9	0,45 (45%)	13	1,4
1,7  —— 2,3	4	0,20 (20%)	17	2,0
2,3  —— 2,9	1	0,05 (5%)	18	2,6
2,9  —— 3,5	2	0,10 (20%)	20	3,2
TOTAL	20	1,00 (100%)	-	-

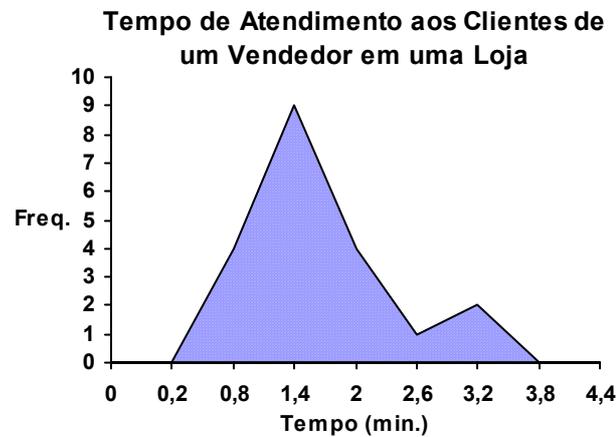
FONTE: Lojas Balasol.

## 2.2 GRÁFICOS:

### 2.2.1 HISTOGRAMA:



## 2.2.2 POLÍGONO DE FREQUÊNCIA



## 2.3 MEDIDAS DE TENDÊNCIA CENTRAL:

### 2.3.1 MÉDIA ARITMÉTICA:

	POPULAÇÃO	AMOSTRA
DADOS ISOLADOS:	$\mu = \frac{\sum_{i=1}^N x_i}{N}$	$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$
DADOS AGRUPADOS:	$\mu = \frac{\sum_{i=1}^k f_i x_i}{\sum_{i=1}^k f_i}$	$\bar{x} = \frac{\sum_{i=1}^k f_i x_i}{\sum_{i=1}^k f_i}$

onde:  $x$  = valores da variável  
 $n$  = número de elementos da amostra  
 $N$  = número de elementos da população  
 $f_i$  = frequência da classe  $i$   
 $x_i$  = ponto médio da classe  $i$   
 $k$  = número de classes

#### EXEMPLO 1:

Para dados isolados:  $\bar{x} = 1,59$

Para dados agrupados:  $\bar{x} = 1,61$

### 2.3.2 MEDIANA ( $\tilde{x}$ ):

Divide o conjunto “ordenado” de valores em 2 partes iguais, não considerando o valor numérico, mas somente a posição.

EXEMPLO 1: Seja o conjunto  $A = \{7 \ 8 \ 10 \ 5 \ 3 \ 2 \ 6\}$

Rol: 2 3 5 **6** 7 8 10  
então:  $\tilde{x} = 6$

EXEMPLO 2: Seja o conjunto  $B = \{10 \ 11 \ 12 \ 13 \ 15 \ 9 \ 200 \ 1000\}$

Rol: 9 10 11 **12 13** 15 200 1000

então:  $\tilde{x} = \frac{12 + 13}{2} = 12,5$

### 2.3.3 MODA ( $\hat{x}$ ):

Moda(s) de um conjunto de valores é o valor(es) de frequência máxima.

EXEMPLO 1: Seja o conjunto  $B = \{13 \ 10 \ 11 \ 12 \ 13 \ 15 \ 9 \ 200 \ 1000\}$   
 $\hat{x} = 13$

Distribuição Normal :  $\bar{x} = \tilde{x} = \hat{x}$

### 2.3.4 FRACTIS: Quartis, Decis e Percentis.

**2.3.4.1 QUARTIS:**  $Q_1$ ,  $Q_2$  e  $Q_3$ . Dividem os valores ordenados em quatro subconjuntos com iguais números de elementos.

**2.3.4.2 DECIS:**  $D_1, D_2, \dots, D_9$ . Dividem os valores ordenados e dez subconjuntos com iguais números de elementos.

**2.3.4.3 PERCENTIS:**  $P_1, P_2, \dots, P_{99}$ . Dividem os valores ordenados e cem subconjuntos com iguais números de elementos.

$Q_2 = D_5 = P_{50} = \tilde{x}$  (mediana)

## 2.4 MEDIDAS DE DISPERSÃO:

Dados os conjuntos:

$A = \{2, 3, 5, 6, 4\}$        $\tilde{x}_A = 4$        $\bar{x}_A = 4$

$B = \{1, 4, 0, 5, 10\}$        $\tilde{x}_B = 4$        $\bar{x}_B = 4$

### 2.4.1 AMPLITUDE TOTAL:

$$AT = \text{Máx} - \text{Mín}$$

$$AT_A = 6 - 2 = 4$$

### 2.4.2 VARIÂNCIA:

Média dos quadrados das diferenças dos valores em relação a sua média aritmética.

	POPULAÇÃO	AMOSTRA
DADOS ISOLADOS:	$\sigma^2 = \frac{\sum_{i=1}^N (x_i - \mu)^2}{N}$	$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}$
DADOS AGRUPADOS:	$\sigma^2 = \frac{\sum_{i=1}^k f_i \cdot (x_i - \mu)^2}{\left(\sum_{i=1}^k f_i\right)}$	$s^2 = \frac{\sum_{i=1}^k f_i \cdot (x_i - \bar{x})^2}{\left(\sum_{i=1}^k f_i\right) - 1}$

### 2.4.3 DESVIO PADRÃO:

Como a variância envolve o quadrado dos desvios, e é dada em número de unidades elevadas ao quadrado; o desvio padrão torna a unidade da variável igual a da média.

POPULAÇÃO	AMOSTRA
$\sigma = \sqrt{\sigma^2}$	$s = \sqrt{s^2}$

EXEMPLO 1:

Para dados isolados:  $s = 0,7553$

### 2.4.4 COEFICIENTE DE VARIAÇÃO:

Caracteriza a dispersão dos dados em termos relativos ao seu valor médio. É uma medida adimensional e é usada na comparação entre distribuições de dados de unidades diferentes.

$$cv = \frac{s}{\bar{X}} \times 100$$

EXEMPLO 1:

Para dados isolados:  $cv = 46,30\%$

Para dados agrupados:  $cv = 40,60\%$

### 2.5 EXERCÍCIOS:

- 2.5.1 Os dados a seguir referem-se ao diâmetro, em polegadas, de uma amostra de 60 rolamentos de esferas produzidos por uma companhia.

0,738 0,729 0,743 0,740 0,741 0,735 0,731 0,726 0,737 0,736  
 0,728 0,737 0,736 0,735 0,724 0,733 0,742 0,736 0,739 0,735  
 0,745 0,736 0,742 0,740 0,728 0,738 0,725 0,733 0,734 0,732  
 0,733 0,730 0,732 0,730 0,739 0,734 0,738 0,739 0,727 0,735  
 0,735 0,732 0,735 0,727 0,734 0,732 0,736 0,741 0,736 0,744  
 0,732 0,737 0,731 0,746 0,735 0,735 0,729 0,734 0,730 0,740

Baseado na distribuição de dados construa:

- Tabela da distribuição de frequência, utilizando a fórmula de Sturges.
- Polígono de frequência
- Histograma

- 2.5.2 Utilizando a tabela que você construiu no problema 2.5.1. determine:

- Média Aritmética para dados isolados
- Desvio padrão para dados isolados
- Mediana para dados isolados
- Moda para dados isolados
- Coefficiente de Variação

- 2.5.3 Uma amostra dos salários mensais em reais de 50 operários da Construção Civil de uma certa Empresa são apresentados a seguir.

415 424 477 454 397 424 549 441 513 425  
 391 450 524 410 413 543 560 469 585 556  
 449 442 424 447 527 457 544 420 465 514  
 473 398 389 340 401 391 382 397 437 383  
 433 524 497 513 429 389 440 427 491 414

Baseado na distribuição de dados construa:

- Tabela da distribuição de frequência, utilizando a fórmula de Sturges.
- Polígono de frequência
- Histograma

- 2.5.4 Utilizando a tabela que você construiu no problema 2.5.3. determine:

- Média Aritmética para dados isolados
- Desvio padrão para dados isolados
- Mediana para dados isolados
- Moda para dados isolados
- Coefficiente de Variação

- 2.5.5 Para uma amostra da vida útil de ferramentas de corte em um processo industrial apresentadas abaixo, determine:

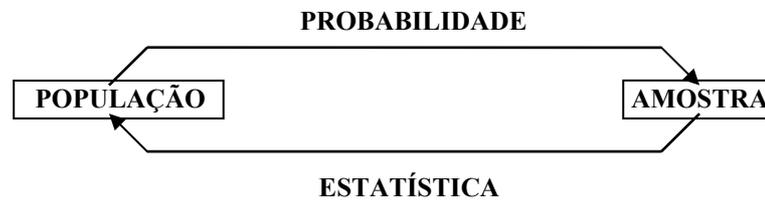
120	88	77	122	89	69	71	99	57	107
100	92	102	77	105	101	66	92	95	127
102	63	81	95	88	64	69	122	97	81
125	53	62	95	53	86	85	89	103	102

- a) Média Aritmética para dados isolados
- b) Desvio padrão para dados isolados
- c) Mediana para dados isolados
- d) Moda para dados isolados
- e) Coeficiente de Variação

### III - PROBABILIDADE

#### 3.1. DEFINIÇÕES BÁSICAS:

##### 3.1.1- INTRODUÇÃO:



Universo :  $\Omega$  ou  $U$

Vazio:  $\emptyset$

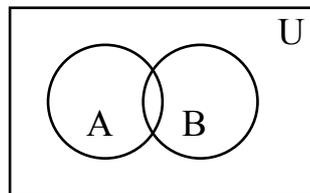
União:  $A \cup B$

Intersecção:  $A \cap B$

Complemento:  $A'$  ou  $\bar{A}$  ou  $A^c$

Diferença:  $A - B$  (A mas não B)

DIAGRAMA DE VENN:



Teorema 1: Lei Comutativa.

$$A \cup B = B \cup A \text{ e}$$

$$A \cap B = B \cap A$$

Teorema 2: Lei Associativa.

$$A \cup (B \cup C) = (A \cup B) \cup C \text{ e}$$

$$A \cap (B \cap C) = (A \cap B) \cap C$$

Teorema 3: Lei Distributiva.

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C) \text{ e}$$

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

Teorema 4:

$$(A^c)^c = A$$

Teorema 5:

$$A \cap U =$$

$$A \cup U =$$

$$A \cap \emptyset =$$

$$A \cup \emptyset =$$

Teorema 6:

$$A \cap A' = \emptyset$$

$$A \cup A' = U$$

$$A \cap A = A$$

$$A \cup A = A$$

Teorema 7 : Leis de De Morgan

$$(A \cup B)' = A' \cap B' \quad e$$

$$(A \cap B)' = A' \cup B'$$

Teorema 8:

$$A - B = A \cap B'$$

Teorema 9:

$$A = (A \cap B) \cup (A \cap B')$$

Teorema 10:

$$(A \cup B) = A \cup (A^c \cap B)$$

Teorema 11:

Se  $A \subset B$ , então  $A \cap B = A$  e  $A \cup B = B$

### 3.1.2- AMOSTRAS ORDENADAS E NÃO-ORDENADAS:

#### 3.1.2.1 PRINCÍPIO FUNDAMENTAL DA CONTAGEM:

Se um acontecimento pode ocorrer por várias etapas sucessivas e independentes de tal modo que:

$m_1$  é o nº de possibilidades da 1ª etapa

$m_2$  é o nº de possibilidades da 2ª etapa

:

$m_k$  é o nº de possibilidades da kª etapa

Então o número total de possibilidades do acontecimento ocorrer é  $m_1 \times m_2 \times \dots \times m_k$ .

Exemplo:

Número possíveis de placas de automóveis:

26 letras  $\rightarrow m_1 = 26^3$

10 algarismos  $\rightarrow m_2 = 10^4$

Total:  $\rightarrow m_1 \times m_2 = 175.760.000$

#### 3.1.2.2 AMOSTRAS ORDENADAS:

Suponha ter os conjuntos A e B. Se A tem  $m$  elementos distintos  $(a_1, a_2, \dots, a_m)$  e B tem  $p$  elementos distintos  $(b_1, b_2, \dots, b_p)$ , então o número de pares  $(a_i, b_j)$ , com  $i=1, 2, \dots, m$  e  $j=1, 2, \dots, p$ ; que podem ser formados, tomando-se um ponto de A e um ponto de B é:  $m \cdot p$  (pelo Princípio Fundamental da Contagem).

Suponha, ainda que ter  $n$  conjuntos  $A_1, A_2, \dots, A_n$  cada um tendo  $m_1, m_2, \dots, m_n$  elementos distintos, respectivamente. Então, o número de  $n$ -uplas  $(x_1, x_2, \dots, x_n)$  que podem ser formadas com um elemento  $x_i$  de cada  $A_i$  é  $m_1 \times m_2 \times \dots \times m_n$  (pelo Princípio Fundamental da Contagem).

Quando cada conjunto  $A_i$  é o mesmo conjunto  $A$  com  $N$  elementos distintos, tem-se  $N^n$   $n$ -uplas.

#### -AMOSTRAGEM COM REPOSIÇÃO:

Exemplo 1: Suponha que uma caixa tenha  $N$  bolas numeradas de 1 a  $N$ . Extrair uma bola e recolocar. Quantas  $n$ -uplas podem ser formadas com os  $n$  números obtidos nas extrações?

R:  $N^n$

Exemplo 2: Suponha que a caixa tenha 3 bolas, represente as possíveis  $n$ -uplas resultantes de  $n=2$  extrações com reposição.

R: São  $3^2=9$  possíveis resultados.

(1,1) (1,2) (1,3) (2,1) (2,2) (2,3) (3,1) (3,2) (3,3)

#### -AMOSTRAGEM SEM REPOSIÇÃO:

Exemplo 1: Suponha que uma caixa tenha  $N$  bolas numeradas de 1 a  $N$ . Extrair uma bola e não recoloca-la de volta na caixa. Quantas  $n$ -uplas podem ser formadas ?

R:  $A_{N,n} = \frac{N!}{(N-n)!}$

Exemplo 2: No caso da caixa conter 3 bolas, represente as possíveis  $n$ -uplas resultantes de  $n=2$  extrações.

R: São  $A_{3,2}=6$  possíveis resultados

(1,2) (1,3) (2,1) (2,3) (3,1) (3,2)

### 1.3.3 AMOSTRAS NÃO ORDENADAS:

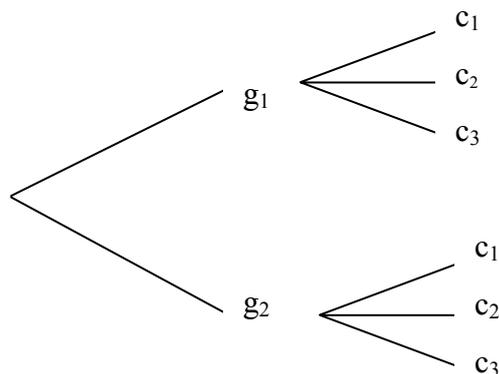
O número de amostras distintas de tamanho  $n$  que podem ser extraídas, sem reposição e sem considerar a ordem que eles aparecem, de um conjunto de  $N$  objetos distintos é denominado de Combinação, denotado por  $C_{N,n}$  e dado pela fórmula:

$$C_{N,n} = \binom{N}{n} = \frac{N!}{n!(N-n)!}$$

#### -DIAGRAMA DE ÁRVORE

É a representação esquemática do experimento de se combinar um elemento do conjunto  $A$  com um elemento do conjunto  $B$ . Também pode-se combinar com elementos de um terceiro conjunto, porém com 4 ou mais conjunto este procedimento não é recomendado.

Exemplo: Suponha que se queira combinar 2 gravatas ( $g_1, g_2$ ) com 3 camisas ( $c_1, c_2, c_3$ ).



### RESUMINDO:

#### Permutações:

Sem repetição: o número de maneiras de dispor  $N$  objetos diferentes é dado por:

$$P_N = N!$$

Com repetição: o número de maneiras de dispor de  $N$  objetos dos quais  $N_1$  são iguais,  $N_2$  são iguais, ..., é dado por:

$$P_N = \frac{N!}{N_1! \times N_2! \times \dots}$$

#### Arranjos:

Se tivermos  $N$  objetos diferentes e desejamos escolher  $n$  desses objetos ( $n \leq N$ ) e permutar os  $n$  escolhidos é dado por:

$$A_{N,n} = \frac{N!}{(N-n)!}$$

#### Combinações:

Se tivermos  $N$  objetos diferentes e queremos o número de maneiras de se obter  $n$  dentre esses  $N$ , sem considerar a ordem teremos:

$$C_{N,n} = \binom{N}{n} = \frac{N!}{n!(N-n)!}$$

### 3.1.3- ESPAÇO AMOSTRAL:

Experimento Aleatório:

Quando na vida real se realiza uma experiência (experimento) cujo resultado não pode ser previsto com certeza.

Espaço Amostral:

É o conjunto de todos os resultados possíveis de um experimento aleatório, denotado por  $\Omega$  ou  $\mathcal{S}$ .

Evento Aleatório:

Qualquer subconjunto de interesse do espaço amostral  $\Omega$ .

Evento Simples ou Elementar: único ponto amostral  $\{a\}$

Evento Certo:  $\Omega$

Eventos mutuamente exclusivos ou disjuntos:  $A \cap B = \emptyset$

### 3.2 DEFINIÇÕES DE PROBABILIDADE:

Quando se deseja associar números aos eventos, de tal forma que estes venham traduzir as possibilidades dos eventos ocorrerem, atribui-se à ocorrência dos eventos uma medida denominada de Probabilidade do Evento.

#### 3.2.1 DEFINIÇÃO CLÁSSICA:

Se existem **a** resultados possíveis favoráveis à ocorrência de um evento A e **b** resultados possíveis não favoráveis à ocorrência de A, e sendo todos os resultados igualmente prováveis e mutuamente exclusivos, então a probabilidade do evento A ocorrer é:

$$P(A) = \frac{a}{a + b}$$

ou

$$P(A) = \frac{\#A}{\#\Omega}$$

#### 3.2.2 DEFINIÇÃO AXIOMÁTICA:

Para todo  $A \in \mathcal{A}$  que associe um número real  $P(A)$ , chamado de Probabilidade de A, de modo que os axiomas a seguir sejam satisfeitos:

$$\text{Axioma 1: } P(A) \geq 0 \quad \forall A \in \mathcal{A}$$

$$\text{Axioma 2: } P(\Omega) = 1$$

Axioma 3: Sejam  $A_1, A_2, \dots$ ; uma sequência (finita ou infinita) de eventos mutuamente exclusivos onde:

$A_i \cap A_j = \emptyset$  com  $i \neq j$  então:

$$P\left(\bigcup_i A_i\right) = \sum_{i=1}^{n(\infty)} P(A_i)$$

### 3.3 PROPRIEDADES DE PROBABILIDADE:

#### 3.3.1- PROPRIEDADE 1:

$$P(A^c) = 1 - P(A)$$

Demo:  $A \cap A^c = \emptyset$  e  $\Omega = A \cup A^c$

$$1 = P(\Omega) = P(A) + P(A^c)$$

$$P(A^c) = 1 - P(A) \quad // \text{ c.q.d}$$

#### 3.3.2- PROPRIEDADE 2:

$$P(\emptyset) = 0$$

Demo:  $A = (A \cup \emptyset)$  ou  $P(\emptyset) = 1 - P(\Omega)$   
 $P(A) = P(A \cup \emptyset)$   $P(\emptyset) = \emptyset$  // c.q.d

$$P(A) = P(A) + P(\emptyset)$$

$$P(A) - P(A) = P(\emptyset)$$

$$0 = P(\emptyset) \quad // \text{ c.q.d}$$

#### 3.3.3- PROPRIEDADE 3:

Se A e B são dois eventos quaisquer, não necessariamente mutuamente exclusivos, então:

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

Demo:  $B = (A \cap B) \cup (A^c \cap B)$  e  $(A \cup B) = A \cup (A^c \cap B)$

$$P(B) = P(A \cap B) + P(A^c \cap B) \quad \text{e} \quad P(A \cup B) = P(A) + P(A^c \cap B)$$

$$P(A^c \cap B) = P(B) - P(A \cap B)$$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \quad // \text{ c.q.d}$$

### 3.4 PROB. CONDICIONAL E INDEPENDÊNCIA DE EVENTOS:

#### 3.4.1- PROBABILIDADE CONDICIONAL:

Seja  $(\Omega, \mathcal{A}, P)$  um espaço amostral. Se  $A \in \mathcal{A}$  e  $B \in \mathcal{A}$  e  $P(B) > 0$ , a probabilidade condicional de A dado que B ocorreu é definida por:

$$P(A / B) = \frac{P(A \cap B)}{P(B)} \quad A \in \mathcal{A}$$

Obs: Se  $P(B) = 0$  então  $P(A/B) = P(A)$  (Barry James)  
 não é definido ( Mood, Graybill & Goes)

### 3.4.2- INDEPENDÊNCIA DE EVENTOS:

Seja  $(\Omega, \mathcal{A}, P)$  um espaço de probabilidade. Os eventos aleatórios  $A$  e  $B$  são independentes se :

$$P(A \cap B) = P(A) \cdot P(B)$$

Isto implica que:

$$P(B/A) = P(B) \quad \text{ou} \quad P(A/B) = P(A)$$

Assim:

$$P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C)$$

Generalizando temos que:

$$P(A_i \cap A_j) = P(A_i) \cdot P(A_j) \quad \text{são independentes 2 a 2, } \forall i \in I, i \neq j$$

Se os eventos  $A_i, i \in I$ , são independentes, então os eventos  $B_i, i \in I$ , também são independentes, onde cada  $B_i$ , é igual a  $A_i$ , ou  $A_i^c$ .

$$P\left(\bigcap_{i=1}^n B_i\right) = \prod_{i=1}^n P(B_i)$$

## 3.5 PROBABILIDADE TOTAL E FÓRMULA DE BAYES:

### 3.5.1 PROBABILIDADE TOTAL:

Seja  $A_1, A_2, \dots$  uma sequência de eventos aleatórios que forma uma partição de  $\Omega$ , ou seja  $A_i$  são mutuamente exclusivos e sua união é  $\Omega$  e seja  $B$  um evento qualquer assim:

$$\begin{aligned} B &= \Omega \cap B = (A_1 \cup A_2 \cup \dots \cup A_n) \cap B \\ B &= (A_1 \cap B) \cup (A_2 \cap B) \cup \dots \cup (A_n \cap B) \end{aligned}$$

onde  $(A_i \cap B)$  são mut. excl.

$$\begin{aligned} P(B) &= P(A_1 \cap B) + P(A_2 \cap B) + \dots + P(A_n \cap B) \\ P(B) &= P(A_1) \cdot P(B/A_1) + P(A_2) \cdot P(B/A_2) + \dots + P(A_n) \cdot P(B/A_n) \end{aligned}$$

$$P(B) = \sum_i P(A_i) P(B/A_i) \quad \forall B \in \mathcal{A}$$

### 3.5.2- FÓRMULA DE BAYES:

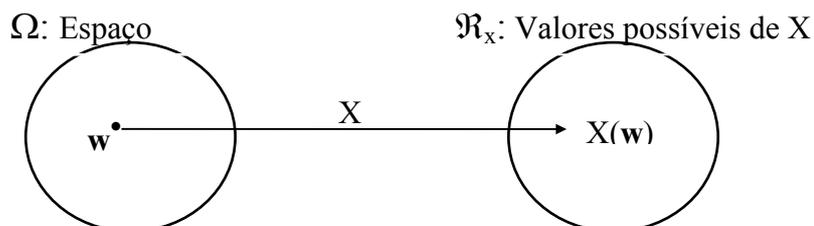
$$P(A_k / B) = \frac{P(A_k \cap B)}{P(B)} = \frac{P(A_k) \cdot P(B/A_k)}{\sum_{i=1}^n P(A_i) \cdot P(B/A_i)}$$

### 3.6-VARIÁVEIS ALEATÓRIAS UNIDIMENSIONAIS:

#### 3.6.1 CONCEITO DE VARIÁVEIS ALEATÓRIAS:

Informalmente, uma variável aleatória é um característico numérico do resultado de um experimento aleatório.

Definição: Seja  $\varepsilon$  um experimento e  $\Omega$  um espaço amostral associado ao experimento. Uma função  $X$ , que associe a cada elemento  $w$  pertencente ao  $\Omega$  um número real é denominada de variável aleatória.



Ex: 1. Dado o experimento:

$\varepsilon$  = lançamento de 2 moedas

$X$  = número de caras obtidos

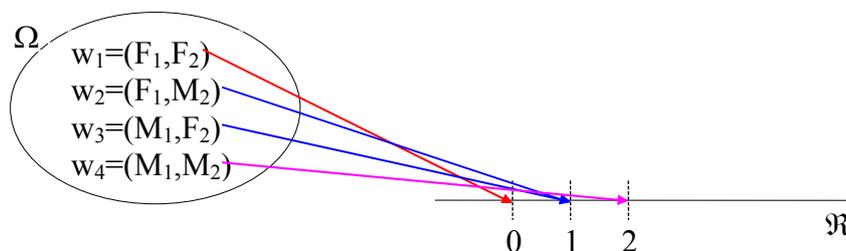
Então:

$$\Omega = \{(ca,ca), (ca,co), (co,ca), (co,co)\}$$

$$X(ca,ca) = 2 \quad X(ca,co) = X(co,ca) = 1 \quad X(co,co) = 0$$

2. Seja uma família com 2 crianças, v.a. número de meninos na família

$$\Omega = \{(F_1,F_2); (F_1,M_2); (M_1,F_2); (M_1,M_2)\}$$



a função  $X(w) \Rightarrow$  V.A. contradomínio da v.a.  $\{0,1,2\}$

#### 3.6.2- VARIÁVEIS ALEATÓRIAS DISCRETAS.

Def: Sendo  $X$  um v.a.. Se o número de valores possíveis de  $X$  for finito ou infinito numerável, denominamos  $X$  de v.a discreta.

#### 3.6.3- VARIÁVEIS ALEATÓRIAS CONTÍNUAS:

Def: Sendo  $X$  uma v.a. Suponha que  $\mathcal{R}_X$ , o contradomínio de  $X$  seja um intervalo ou uma coleção de intervalos, isto é, a v.a. toma um número infinito não numeráveis de valores.

### 3.6.4.- DISTRIBUIÇÃO DE PROBABILIDADE.

Seja  $X$  uma v.a. que assume os valores  $x_1, x_2, \dots, x_n$  com probabilidade  $p_1, p_2, \dots, p_n$  associadas a cada elemento de  $X$ , sendo  $p_1 + p_2 + \dots + p_n = 1$  diz-se que está definida um Distribuição de Probabilidade.

- v.a. discretas  $\Rightarrow$  Função de Probabilidade

- v.a. contínuas  $\Rightarrow$  Função Densidade de Probabilidade.

#### 3.6.4.1- FUNÇÃO DE PROBABILIDADE.

Seja  $X$  um v.a. discreta. A cada possível resultado  $x_i$  associaremos uma probabilidade  $p(x_i)$ , então  $p(x_i)$  é uma Função de Probabilidade se satisfizer as seguintes condições:

$$a) \sum_{i=1}^{\infty} p(x_i) = 1$$

$$b) p(x_i) \geq 0$$

Ex: Ao lançar um dado e seja  $X$  os valores observados:

Então:

$X$	1	2	3	4	5	6
$P(x_i)$	1/6	1/6	1/6	1/6	1/6	1/6

$$a) \sum_{i=1}^6 p(x_i) = 6 \times 1/6 = 1$$

$$b) p(x_i) = 1/6 > 0$$

#### 3.6.4.1- FUNÇÃO DENSIDADE DE PROBABILIDADE.

Se  $X$  é uma v.a contínua com Função Densidade de Probabilidade  $f_X(x)$  com domínio em  $\mathfrak{R}$  ou no conjunto  $A \subset \mathfrak{R}$ , tem-se:

$$a) f_X(x) \geq 0 \quad \forall x \in \mathfrak{R}$$

$$b) \int_{-\infty}^{+\infty} f_X(x) dx = 1$$

$$c) P(X=a) = P(X=b) = 0$$

$$d) P(a \leq X \leq b) = \int_a^b f_X(x) dx$$

Ex1: Seja uma v.a.contínua definida pela f.d.p:

$$f(x) = \begin{cases} 0 & \text{para } x < 0 \\ kx & \text{para } 0 \leq x \leq 2 \\ 0 & \text{para } x > 2 \end{cases}$$

Para que seja satisfeita a propriedade:

$$\int_{-\infty}^{+\infty} f_X(x) = 1 \Rightarrow \text{a área do triângulo compreendido entre 0 e 2 deve ser unitário}$$

então:

$$\frac{bxh}{2} = 1 \Rightarrow \frac{2 \times 2k}{2} = 1 \quad \text{resulta que } k=1/2$$

### 3.6.5- FUNÇÃO DISTRIBUIÇÃO.

DEFINIÇÃO:

Seja  $X$  uma v.a. Defini-se a função  $F_X$  como a Função Distribuição (Acumulada) da v. a.  $X$  como  $F_X(x) = P(X \leq x)$ ,  $x \in \mathcal{R}$ .

Ex: Dois dados são lançados. A v. a  $X$  é definida como a soma dos pontos obtidos:

	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

$X=x_i$	2	3	4	5	6	7	8	9	10	11	12
$p(x_i)$	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36
$F(x_i)$	1/36	3/36	6/36	10/36	15/36	21/36	26/36	30/36	33/36	35/36	1

### 3.6.6- ESPERANÇA DE UMA VARIÁVEL ALEATÓRIA:

#### 3.6.6.1- DEFINIÇÃO:

Esperança ou Expectância de uma v. a. é um valor médio dos possíveis valores de  $X$ , ponderada conforme sua distribuição, i.e.; é uma média ponderada onde os pesos são as probabilidades  $p(x_i)$ . É ainda, o centro de gravidade da unidade de massa que é determinada pela função densidade de  $X$ . Assim  $E(X)$  é uma medida de localização ou centro de v.a.

**3.6.6.2-CASO DISCRETO:**

$$E(X) = \sum_{i=1}^{\infty} x_i p(x_i)$$

**3.6.6.3- CASO CONTÍNUO:**

$$E(X) = \int_{-\infty}^{+\infty} x f(x) dx$$

**3.6.6.4- PROPRIEDADES:**

1.  $E(X=c) = c$

Demo:

$$E(X) = \int_{-\infty}^{+\infty} x \cdot f(x) dx = \int_{-\infty}^{+\infty} c \cdot f(x) dx = c \cdot 1 = c$$

2.  $E(cX) = cE(X)$

Demo:

$$E(cX) = \int_{-\infty}^{\infty} c \cdot x \cdot f(x) dx = c \cdot \int_{-\infty}^{\infty} x \cdot f(x) dx = cE(X)$$

3.  $E(aX + b) = a \cdot E(X) + b$

Demo : decorrente das demonstrações acima.

**3.6.7- VARIÂNCIA DE UMA VARIÁVEL ALEATÓRIA:****3.6.7.1-DEFINIÇÃO:**

Se X é uma v.a., definimos a variância de X como a dispersão da densidade de X em relação ao seu valor de localização central de densidade( $E(X)$ ) e é dada por:

$$\begin{aligned} V(X) &= E[X-E(X)]^2 \\ &= E(X-\mu)^2 = E\{X^2 - 2XE(X) + [E(X)]^2\} = \\ &= E(X^2) - 2E(X)E(X) + [E(X)]^2 = \\ &= E(X^2) - [E(X)]^2 \quad // \quad (E(X) \text{ é um constante}) \end{aligned}$$

Obs:  $E(X^2) = \sum_{i=1}^{\infty} x_i^2 p(x_i)$  para uma v.a. discreta

$E(X^2) = \int_{-\infty}^{+\infty} x^2 f(x) dx$  para uma v.a. contínua

### 3.6.7.2- PROPRIEDADES:

$$1. V(X + c) = V(X)$$

Demo:

$$\begin{aligned} V(X + c) &= E[(X+c) - E(X+c)]^2 = E[(X+c) - E(X)-c]^2 \\ &= E[X - E(X)]^2 = V(X) \end{aligned}$$

$$2. V(cX) = c^2 \cdot V(X)$$

Demo:

$$\begin{aligned} V(cX) &= E(cX)^2 - [E(cX)]^2 = c^2 E(X^2) - c^2 [E(X)]^2 \\ &= c^2 \{E(X^2) - [E(X)]^2\} = c^2 V(X) \end{aligned}$$

EXEMPLO 1: Seja uma distribuição de probabilidade dada por:

X	-1	0	1
P(X)	1/4	1/2	1/4

Determine a  $E(X)$  e  $V(X)$ .

R:  $E(X) = 0$  e  $V(X) = 1/2$

EXEMPLO 2: Seja a f.d.p dada abaixo, calcule a  $E(X)$  e  $V(X)$ :

$$f(x) = \begin{cases} 2x & 0 \leq x \leq 1 \\ 0 & c/c \end{cases}$$

R:  $E(X) = 2/3$  e  $V(X) = 1/18$

### 3.7. PRINCIPAIS DISTRIBUIÇÕES DE PROBABILIDADE.

#### 3.7.1. DISTRIBUIÇÕES DISCRETAS;

##### 3.7.1.1. UNIFORME DISCRETA

Uma v.a. tem distribuição uniforme discreta quando sua função de probabilidade for dada por:

$$p(x) = \begin{cases} \frac{1}{N} & x = 1, 2, \dots, N \\ 0 & \text{c/c} \end{cases}$$

PROPRIEDADES:

$$E(X) = \frac{1+N}{2} \qquad V(X) = \frac{N^2 - 1}{12}$$

EXEMPLO:

Seja  $\varepsilon$  lançar um dado, então:

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

$$p(x_i) = 1/6$$

$$E(X) = 3,5$$

$$V(X) = 2,92$$

##### 3.7.1.2. BERNOULLI:

Uma v.a.  $X$  tem distr. Bernoulli se sua f.p. for dada por:

$$p(x) = \begin{cases} p^x (1-p)^{1-x} & x = 0, 1 \\ 0 & \text{c/c} \end{cases}$$

PROPRIEDADES:

$$E(X) = p$$

$$V(X) = p.q \quad \text{onde } q = 1-p$$

PROCESSO DE BERNOULLI:

É o processo de amostragem no qual :

1. Em cada tentativa existem 2 resultados possíveis mutuamente exclusivos (sucesso e fracasso).
2. As séries de tentativas são independentes.
3. A probabilidade de sucesso ( $p$ ) permanece constante de tentativa para tentativa ou seja o processo é estacionário.

**3.7.1.3. BINOMIAL:**

Uma v.a. possui distribuição binomial se sua f.p. for dada por:

$$p(x) = \left\{ \binom{n}{x} \cdot p^x \cdot q^{n-x} \quad x = 0, 1, \dots, n \right\}$$

$$C_{n \ x} = \binom{n}{x} = \frac{n!}{x!(n-x)!}$$

$$x! = x \cdot (x-1) \cdot (x-2) \dots 1$$

A distribuição binomial é utilizada para determinar a probabilidade de obter um dado número de sucessos em um processo de Bernoulli.

X = número de sucessos

n = número de tentativas

p = probabilidade de sucessos em cada tentativa.

PROPRIEDADES:

$$E(X) = np$$

$$V(X) = npq$$

EXEMPLO 1: Uma urna contém 6 bolas brancas e 4 pretas. Calcular a probabilidade de ao retirar com reposição 3 bolas, 2 sejam brancas.

EXEMPLO 2: Lançando 8 moedas, qual a chance de obter:

- a) 3 caras.
- b) Nenhuma cara.
- c) Pelo menos 1 cara.
- d) no mínimo 2 caras.
- e) no máximo 6 caras.

EXEMPLO 3: Sabe-se que 5% dos parafusos fabricados por certa indústria são defeituosos. Em um lote de 10 parafusos, calcular a probabilidade de:

- a) exatamente 2 serem defeituosos;
- b) menos de 2 serem defeituosos;
- c) três ou mais serem defeituosos.

Qual a média e o desvio padrão do número de parafusos defeituosos?

### 3.7.1.4. POISSON:

Uma v.a.  $X$  tem distr. Poisson se sua f.p. for dada por:

$$p(x) = \begin{cases} \frac{\lambda^x \cdot e^{-\lambda}}{x!} & x = 0, 1, 2, \dots \\ 0 & \text{c/c} \end{cases}$$

A distr. de Poisson pode ser usada para determinar a probabilidade de um dado número de sucessos quando os eventos ocorrem em um “continuum” de tempo ou espaço.

É similar ao processo de Bernoulli, exceto que os eventos ocorrem em um “continuum” ao invés de ocorrerem em tentativas fixadas, tal como o processo de Bernoulli os eventos são independentes e o processo é estacionário.

$\lambda$  = número médio de sucessos para uma específica dimensão de tempo e espaço.

$X$  = número de sucessos desejados.

PROPRIEDADE:

$$E(X) = \lambda$$

$$V(X) = \lambda$$

Obs: Quando o número de observações ou experimentos em um processo de Bernoulli for muito grande a distr. de Poisson é apropriada como uma aprox. das distr. Binomiais quando:

$$n \geq 30$$

$$np < 5$$

$$\lambda = np$$

EXEMPLO 1: Um técnico recebe em média de 5 chamadas por dia. Qual a probabilidade que, em uma dia selecionada aleatoriamente, sejam recebidas exatamente 3 chamadas?

EXEMPLO 2: Em média, 12 pessoas por hora são atendidas em um banco. Qual a probabilidade que 3 ou mais pessoas sejam atendidas durante um período de 10 minutos?

### 3.7.2 DISTRIBUIÇÕES CONTINUAS DE PROBABILIDADE:

#### 3.7.2.1. UNIFORME OU RETÂNGULAR:

Uma v.a.  $X$  é uniformemente distribuída em  $1 \leq x \leq b$  se sua f.d.p. for:

$$f(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{c/c} \end{cases}$$

PROPRIEDADES:

$$E(X) = \frac{a+b}{2}$$

$$V(X) = \frac{(b-a)^2}{12}$$

#### 3.7.2.2. DISTRIBUIÇÃO NORMAL (GAUSS):

##### 3.7.2.2.1. DEFINIÇÃO

Uma v.a.  $X \sim N(\mu, \sigma^2)$  se sua f.d.p. for dada por:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad -\infty < x < \infty, \quad -\infty < \mu < \infty \quad \text{e} \quad \sigma > 0$$

##### 3.7.2.2.2. PROPRIEDADES:

1.  $f_X(x) > 0$  ,  $x \in \mathfrak{R}$
2.  $f_X(x)$  é crescente para  $x \in (-\infty, \mu)$  e decrescente para  $x \in (\mu, \infty)$ .
3. Ponto de máximo da função em  $x = \mu$ . Então  $\mu$  é também a moda da distribuição.
4.  $f_X(x)$  é simétrica em relação a  $\mu$ .
8. Valor esperado :  $\mu$
9. Variância =  $\sigma^2$
10. A área da curva correspondente entre:
  - $(\mu - \sigma)$  e  $(\mu + \sigma) = 68,27\%$
  - $(\mu - 2\sigma)$  e  $(\mu + 2\sigma) = 95,45\%$
  - $(\mu - 3\sigma)$  e  $(\mu + 3\sigma) = 99,73\%$

### 3.7.2.2.3. IMPORTÂNCIA:

1. Poder de modelamento. Medidas produzidas em diversos processos aleatórios seguem a distr. normal.
2. Capacidade de aproximação de outras distr. como Binomial e Poisson.
3. As distr. de estatísticas da amostra freqüentemente seguem a distr. normal independente da distr. da população.

EXEMPLO 1: Construa uma distribuição normal com  $\mu = 20$  e  $\sigma = 2$  e determine a probabilidade de se encontrar valores entre:

- a) 18 e 22
- b) 20 e 24

### 3.7.2.2.4. DISTRIBUIÇÃO NORMAL REDUZIDA:

Quando  $\mu = 0$  e  $\sigma^2 = 1$  (caso particular) (chamada "standard", normalizada, padrão)

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

EXEMPLO 1: Determine a área limitada pela curva normal em cada um dos casos:

1. $0 \leq Z \leq 1,2$	R: 0,3849
2. $-0,68 \leq Z \leq 0$	0,2517
3. $-0,46 \leq Z \leq 2,21$	0,6637
4. $Z \leq -0,6$	0,2743
5. $Z \geq 0,62$	0,2676
6. $0,18 \leq Z \leq 0,26$	0,0312
7. $-0,95 \leq Z \leq -0,41$	0,1698
8. $Z < -1,51$ e $Z > 1,51$	0,1310
9. $Z > -0,5$	0,6915

EXEMPLO 2: Sendo os QI's Feminino e Masculino com média igual a 100 e desvio padrão 5 e 10 respectivamente. Calcular as probabilidades de encontrarmos QI's acima de 110 para ambos os sexos.

EXEMPLO 3: Com os dados do exercício anterior calcular as probabilidades de encontrarmos QI's abaixo de 85.

### 3.7.2.2.5. APROXIMAÇÕES PELA NORMAL:

#### 1. BINOMIAL:

quando  $n \geq 30$   
 $np \geq 5$   
 então:  $\mu = np$   
 $\sigma^2 = npq$

## 2. POISSON:

quando  $\lambda \geq 10$ então:  $\mu = \lambda$  $\sigma^2 = \lambda$ 

EXEMPLO 1: Uma moeda não viciada é lançada 500 vezes. Determinar a probabilidade do número de caras não diferir de 250 em:

a) mais de 10

b) mais de 30

EXEMPLO 2: Um dado é lançado 120 vezes. Determinar a probabilidade de aparecer a face 4:

a) 18 vezes ou mais

b) 14 vezes ou menos

EXEMPLO 3: Sabe-se que os pedidos de serviços chegam aleatoriamente e como um processo estacionário numa média de 5 por hora. Qual a probabilidade de que sejam recebidos mais de 50 pedidos em um período de 8 horas?

## 3.7.2.3. EXPONENCIAL:

Uma v.a.  $X$  tem distr. exponencial com parâmetro  $\lambda > 0$ , se sua f.d.p. for dada por:

$$f(x) = \lambda e^{-\lambda x} I_{(0, \infty)}(x) \quad , \quad \lambda > 0$$

PROPRIEDADES:

$$E(X) = \frac{1}{\lambda} \quad V(X) = \frac{1}{\lambda^2} \quad M_X(t) = \frac{\lambda}{\lambda - t}, t < \lambda$$

$$F(x) = P(X \leq x) = \int_0^x \lambda e^{-\lambda t} dt = 1 - e^{-\lambda x}, x \geq 0$$

Assim:

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

Obs: Se os eventos, ou sucessos, ocorrem em um contexto de um processo de Poisson, então o comprimento do tempo ou espaço entre 2 eventos sucessivos segue uma distribuição de probabilidade exponencial. Uma vez que tempo ou espaço são um “continuum”, a distr. será contínua.

EXEMPLO 1: Em média, um navio atraca em certo porto a cada 2 dias. Qual a prob. de que, a partir da partida de um navio, se passem 4 dias antes da chegada do próximo navio?

R: média por 2 dias = 1

$\lambda = \text{média por dia} = 1/2$

$P(X > 4) = e^{-\lambda x} = e^{-4 \cdot 1/2} = 13,53\%$

EXEMPLO 2: Um departamento de conserto de máquinas recebe, em média, 5 chamadas por hora. Iniciando em um ponto do tempo aleatoriamente escolhido, qual a prob de que a primeira chamada chegue dentro de  $1/2$  hora ?

R: média por hora = 5

$\lambda = \text{média por hora} = 5$

$P(X \leq 1/2) = 1 - e^{-\lambda x} = 1 - e^{-5 \cdot 1/2} = 1 - 0.08208 = 91,792\%$

### 3.8. TEOREMA DO LIMITE CENTRAL

Seja  $X_1, X_2, \dots, X_n$  v.a.independente identicamente distribuídas (iid); com a mesma  $\mu$  e  $\sigma^2$  e seja  $S_n = X_1 + X_2 + \dots + X_n$  a soma de v.a. iid:

$$\frac{S_n - E(S_n)}{\sqrt{V(S_n)}} \xrightarrow{x \rightarrow \infty} Z_n \sim N(0,1)$$

$$\frac{S_n - n\mu}{\sigma\sqrt{n}} \xrightarrow{x \rightarrow \infty} Z_n \sim N(0,1) \quad \text{ou seja}$$

pois  $E(S_n) = E(X_1 + X_2 + \dots + X_n) = \mu_1 + \mu_2 + \dots + \mu_n = n\mu$

$$V(S_n) = \sigma_1^2 + \sigma_2^2 + \dots + \sigma_n^2 = n\sigma^2$$

Uma dedução feita através do Teorema do Limite Central é que uma distribuição amostral de médias tende uma distr. normal quando  $n$  é suficientemente grande ( $n \geq 30$ ).

$$\frac{\bar{X}_n - E(\bar{X}_n)}{\sqrt{V(\bar{X}_n)}} \sim N\left(\mu_{\bar{X}} = \mu, \sigma_{\bar{X}}^2 = \frac{\sigma^2}{n}\right)$$

onde:

$$\mu_{\bar{X}} = E(\bar{X}) = E\left[\frac{X_1 + X_2 + \dots + X_n}{n}\right] = \frac{1}{n} [E(X_1) + E(X_2) + \dots + E(X_n)] = \frac{1}{n} n\mu = \mu$$

$$\sigma_{\bar{X}}^2 = V(\bar{X}) = V\left[\frac{X_1 + X_2 + \dots + X_n}{n}\right] = \frac{1}{n^2} [V(X_1) + V(X_2) + \dots + V(X_n)] = \frac{1}{n^2} n\sigma^2 = \frac{\sigma^2}{n}$$

#### NOTA HISTÓRICA:

A distribuição normal é chamada historicamente de lei dos erros. Foi usada por Gauss para modelar erros em observações astronômicas. Gauss derivou a distribuição normal, não como limite de somas de variáveis aleatórias independentes, mas a partir de certas hipóteses entre elas a de considerar a média aritmética das observações. Hoje em dia o Teorema do Limite Central dá apoio ao uso da normal como distribuição de erros, pois em muitas situações reais é possível interpretar o erro de uma observação como resultante de muitos erros pequenos e independentes. Pode-se interpretar também que uma observação é gerada da soma de muitos efeitos pequenos e independentes.

### 3.9 EXERCÍCIOS:

3.9.1 Seja um Conjunto Universo dado por  $U = \{0,1,2,3,4,5\}$  e seja os seguintes subconjuntos de U:

$$X = \{1, 2, 4\}$$

$$Y = \{0, 3, 4, 5\}$$

$$Z = \{0, 5\}$$

Encontre :

a)  $X \cap Y$

b)  $X \cup Y$

c)  $(X \cup Y) \cap Z$

d)  $Y' \cup Z'$

e)  $X - Y$

f)  $(Y \cap Z)'$

g)  $(X \cap Y) \cup (Y \cap Z)$

h)  $(X' \cup Y')' \cap (Y' \cup Z)'$

3.9.2 Suponha que se tenha 6 bolas de diferentes cores. De quantas maneiras diferentes elas podem aparecer ao serem colocadas em fila?

3.9.3 De quantas maneiras 8 pessoas podem sentar em 3 lugares diferentes?

3.9.4 Quantas diferentes saladas de frutas podem ser feitas com maçãs, laranjas, tangerinas e bananas.

3.9.5 De quantas maneiras diferentes podemos dispor as letras a,b,c e d.

3.9.6 Com as letras da palavra DADDY podemos ter quantas permutações com reposição?

3.9.7 Qual o número de maneiras de dispor 3 objetos diferentes tomados 2 a 2:

a) Considerando a ordem dos objetos?

b) Não considerando a ordem dos objetos?

3.9.8 Suponha que você retirou uma bola aleatoriamente de uma urna que contém 7 bolas vermelhas, 6 brancas, 5 azuis e 4 amarelas. Qual é a probabilidade que a bola retirada:

- a) seja vermelha
- b) não seja branca
- c) seja branca ou azul
- d) não seja vermelha e nem branca
- e) seja vermelha ou azul ou amarela

3.9.9 Joga-se um dado duas vezes. Encontre a probabilidade de se obter:

- a) uma face 5 ou 6 na primeira jogada e uma face 2 ou 3 na segunda jogada.
- b) um total de 5 ou 6 se somarmos o resultado obtido nas duas fases.

3.9.10 Qual a probabilidade de retirarmos 2 valetes de um baralho de 52 cartas, se:

- a) a primeira carta é recolocada no baralho após ser retirada.
- b) a primeira carta não é recolocada no baralho após ser retirada.

3.9.11 Uma urna contém 7 bolas vermelhas, 4 brancas e 8 azuis. Se 3 bolas forem retiradas aleatoriamente da urna sem reposição, determine a probabilidade de que:

- a) 2 sejam vermelhas e 1 azul.
- b) a última seja azul
- c) a primeira seja vermelha, a segunda branca e a última seja azul
- d) uma seja vermelha, outra branca e outra seja azul.

3.9.12 Os pneus de certa marca apresentam um certo defeito com probabilidade de 0,2. Se 3 pneus forem escolhidos aleatoriamente, qual a probabilidade de todos os 3 apresentarem esse defeito?

3.9.13 Dados 2 eventos mutuamente exclusivos, A e B, sendo  $P(A)$  e  $P(B) \neq 0$ , serão A e B independentes? Porquê?

3.9.14 A probabilidade de um item defeituoso num processo de fabricação é de 10%. Qual a probabilidade de que dois itens aleatoriamente selecionados:

- a) os dois apresentem defeitos
- b) os dois não apresentem defeitos
- c) um dos dois apresente defeito
- d) somente o primeiro apresente defeito

3.9.15 Durante um período particular 80% das ações emitidas por uma indústria tiveram elevações no mercado. Se um investigador escolhe aleatoriamente 4 ações determine a probabilidade de que:

- a) todas tiveram suas cotações aumentadas.
- b) Somente um delas teve sua cotação aumentada.

- 3.9.16 As probabilidades de um marido, sua esposa e um filho estarem vivos daqui a 30 anos são, respectivamente; 0,3; 0,6; e 0,9. Determine a probabilidade de que, daqui a 30 anos nenhum esteja vivo?
- 3.9.17 A probabilidade de um componente apresentar o defeito tipo 1 é de 3% e do tipo 2 é de 6%, e a probabilidade do componente apresentar ambos é de 2%. Encontre:
- a) a probabilidade de um componente apresentar o defeito tipo 1 ou o defeito tipo 2.
  - b) a probabilidade não apresentar nenhum dos dois defeitos.
- 3.9.18 Uma empresa produz circuitos integrados em três fábricas A, B e C. A fábrica A produz 40% dos circuitos, enquanto que as outras produzem 30% cada uma. As probabilidades de que um circuito integrado produzido por estas fábricas não funcione são 1%, 4% e 3%, respectivamente. Responda:
- a) Escolhido um circuito qual a probabilidade do circuito não funcionar?
  - b) Escolhido um circuito e verificou-se ser defeituoso, qual a probabilidade dele ter vindo da fábrica A?
- 3.9.19 Supondo que 20% dos funcionários de uma empresa são mulheres. Numa amostra de 100 funcionários, qual a probabilidade de obtermos:
- a) exatamente 20 mulheres?
  - b) no mínimo 5 mulheres?
  - c) pelo menos 4 mulheres?
- 3.9.20 Uma partida de certo componente consiste de uma caixa com 50 deles, sendo 4 fora da especificação. Retirando-se ao acaso 5 componentes de uma partida, qual a probabilidade de que todas estejam dentro da especificação?
- 3.9.21 A máquina M produz esferas para rolamentos. Se o diâmetro das esferas puder ser considerado um variável aleatória normalmente distribuída, com média 5mm e desvio padrão de 0,05mm, quantas terão diâmetro superior a 5,07 se 200 esferas forem selecionadas? Se o controle de qualidade rejeitar os itens que se afastarem mais do que 0,1mm da média, quantas esferas serão rejeitadas?
- 3.9.22 Uma certa empresa produz canaletas com comprimento médio de 5,1m e desvio padrão de 8cm a um custo unitário de R\$ 6,00. As peças com comprimento entre 4,95m e 5,25m são vendidas a R\$ 10,00; as produzidas com menos de 4,95m são refugadas a R\$ 2,00; as de mais de 5,25m são encurtadas, a um custo de R\$1,00. Qual o lucro médio esperado por unidade produzida?
- 3.9.23 As vendas de determinado produto têm distribuição normal, com média 500 e desvio padrão de 50 unidades. Se a empresa decide fabricar 600 unidades no mês em estudo, qual a probabilidade de que não possa atender a todos os pedidos desse mês, por estar com a produção esgotada?

- 3.9.24 Suponha que a vida de dois aparelhos elétricos A e B tenham distribuições  $N(46,9)$  e  $N(46,36)$ , respectivamente. Se o aparelho é para ser usado por um período de no mínimo de 45 horas, qual aparelho deve ser preferido?
- 3.9.25 Uma máquina produz recipientes cujos diâmetros são normalmente distribuídos com média 0,498 e desvio padrão de 0,02. Se as especificações exigem recipientes com diâmetro igual a 0,500 polegadas mais ou menos 0,04 polegadas, que fração dessa produção será inaceitável?
- 3.9.26 Sabe-se que 30% de todas as chamadas destinadas a uma mesa telefônica são chamadas DDD. Se 1200 chamadas chegarem a essa mesa, qual é a probabilidade de pelo menos 50 serem DDD?
- 3.9.27 Numa central telefônica, o número médio de chamadas é de 8 por minuto. Determinar qual a probabilidade de que num minuto se tenha:
- 10 ou mais chamadas.
  - Menos de 9 chamadas.
  - Entre 7 (inclusive) e 9 (exclusive) chamadas.

## IV- AMOSTRAGEM

### 4.1-TIPOS DE AMOSTRAGEM:

**4.1.1- AMOSTRAGEM PROBABILÍSTICA:** onde todos os elementos tem probabilidade conhecida e diferente, isto somente será possível se a população for finita e totalmente acessível.

1. Amostragem Aleatória Simples: todos os elementos tem igual probabilidade de pertencer a amostra.(Ex: números aleatórios, sorteio).
2. Amostragem Aleatória Sistemática: quando os elementos se apresentam ordenados e a retirada dos elementos da amostra é feita periodicamente (Ex: linha de produção).
3. Amostragem Aleatória Estratificada: a população se divide em sub-populações ou estratos, e a variável de interesse possui comportamento homogêneo dentro de cada estrato. Esta amostragem consiste em especificar quantos elementos serão retirados de cada estrato para constituir a amostra.
4. Amostragem Aleatória Agregada: a população é subdividida em pequenos grupos, chamados de conglomerados ou agregados, sorteia-se um número suficiente de agregados, cujos elementos constituirão a amostra (Ex: bairros, turmas).

**4.1.2- AMOSTRAGEM NÃO-PROBABILÍSTICA:** usada quando for impossível se obter amostras probabilísticas, como seria o desejável (Ex: retirar 100 parafusos de uma caixa que contém 10.000 ).

### 4.2- DISTRIBUIÇÕES AMOSTRAIS:

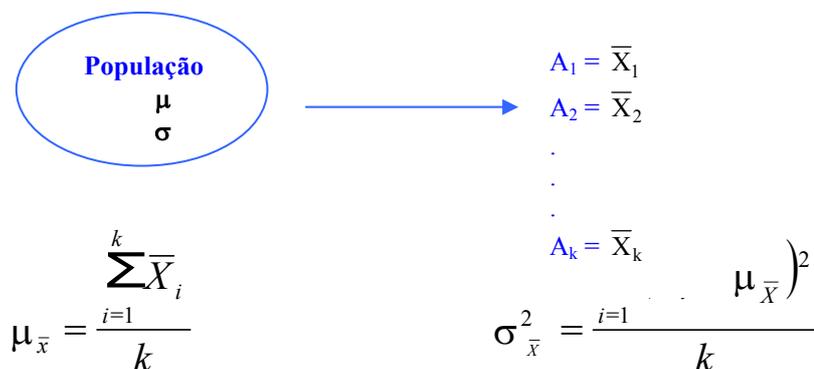
São Aquelas que consideram todas as amostras “possíveis” que possam ser retiradas de uma população.

Para cada amostra pode-se calcular um grandeza estatística, como média aritmética, desvio padrão, proporção, etc.; que varia de amostra para amostra. Assim, calculando-se a média e a variância da grandeza obtêm-se as distribuições amostrais da grandeza, isto é , se a grandeza adotada for a média teremos uma distribuição amostral de médias. De modo análogo teríamos para a variância, proporção , etc.

#### 4.2.1- DISTRIBUIÇÃO AMOSTRAL DE MÉDIAS:

Vimos anteriormente utilizando o T.L.C. que uma distribuição de médias é aproximadamente Normal padronizada

Admita-se todas as amostras possíveis de tamanho  $n$  retiradas de uma população de tamanho  $N$ , então poderemos determinar a média e a variância.



Assim:

$$\mu_{\bar{x}} = E(\bar{X}) = E\left(\frac{X_1 + X_2 + \dots + X_n}{n}\right) = \frac{1}{n} E(X_1 + X_2 + \dots + X_n) = \frac{1}{n} (\mu + \mu + \dots + \mu) = \frac{1}{n} n\mu = \mu$$

$$\sigma_{\bar{x}}^2 = Var(\bar{X}) = Var\left(\frac{X_1 + X_2 + \dots + X_n}{n}\right) = \frac{1}{n^2} Var(X_1 + X_2 + \dots + X_n) = \frac{1}{n^2} n\sigma^2 = \frac{\sigma^2}{n}$$

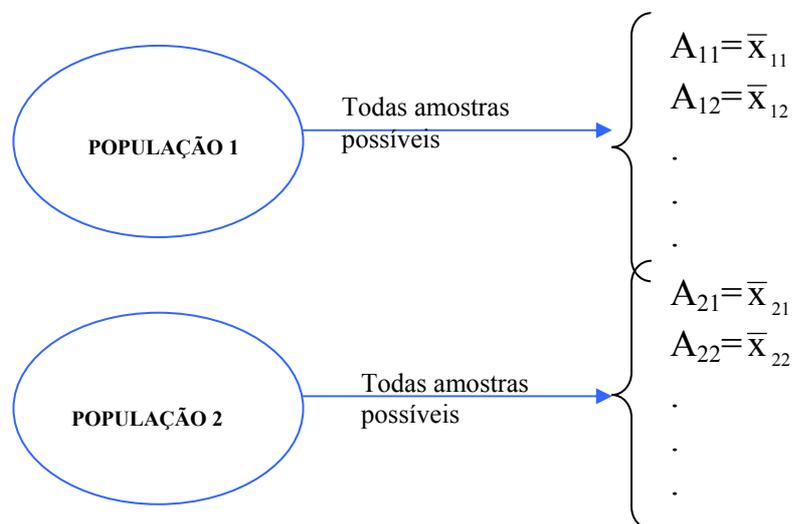
Amostragem	Média	Desvio Padrão
Com Reposição (pop. Infinita)	$\mu_{\bar{x}} = \mu$	$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$
Sem Reposição (pop. Finita)	$\mu_{\bar{x}} = \mu$	$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}$

Obs:

1. Para grandes valores de  $n$  ( $n \geq 30$ ), a distribuição amostral das médias é aproximadamente Normal com média  $\mu_{\bar{x}}$  e Variância  $\sigma_{\bar{x}}^2$ , independentemente da população (desde que o tamanho da população seja, no mínimo, o dobro do tamanho da amostra).
2. No caso da população ser normalmente distribuída, a distribuição amostral das médias também será, mesmo para valores pequenos de  $n$ .
3. A variável reduzida ou padronizada  $Z$  para a distribuição amostral de médias será:

$$Z = \frac{\bar{X} - \mu_{\bar{x}}}{\sigma_{\bar{x}}}$$

#### 4.2.2- DISTRIBUIÇÕES AMOSTRAIS DAS DIFERENÇAS E SOMAS:



A Distribuição Amostral de Diferenças de Médias é obtida através das diferenças entre  $(\bar{X}_{11} - \bar{X}_{21}), (\bar{X}_{12} - \bar{X}_{22}),$  etc.

Analogamente, para a Distr. Amostral de Somas de Médias e para Distr. Amostrais de Diferenças ou Somas de Proporções, ou qualquer outra estatística.

##### 4.2.2.1- DISTRIBUIÇÃO AMOSTRAL DE DIFERENÇAS (SOMAS) DE MÉDIAS:

$$\mu_{\bar{X}_1 \pm \bar{X}_2} = \mu_{\bar{X}_1} \pm \mu_{\bar{X}_2}$$

$$\sigma_{\bar{X}_1 \pm \bar{X}_2}^2 = \sigma_{\bar{X}_1}^2 + \sigma_{\bar{X}_2}^2$$

desde que sejam independentes

Amostragem	Média	Desvio Padrão
Com Reposição (pop. Infinita)	$\mu_{\bar{X}_1 \pm \bar{X}_2} = \mu_{\bar{X}_1} \pm \mu_{\bar{X}_2} = \mu_1 \pm \mu_2$	$\sigma_{\bar{X}_1 \pm \bar{X}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$
Sem Reposição (pop. Finita)	$\mu_{\bar{X}_1 \pm \bar{X}_2} = \mu_{\bar{X}_1} \pm \mu_{\bar{X}_2} = \mu_1 \pm \mu_2$	$\sigma_{\bar{X}_1 \pm \bar{X}_2} = \sqrt{\frac{\sigma_1^2}{n_1} \left( \frac{N_1 - n_1}{N_1 - 1} \right) + \frac{\sigma_2^2}{n_2} \left( \frac{N_2 - n_2}{N_2 - 1} \right)}$

A variável reduzida Z é dada por:

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - \mu_{\bar{X}_1 - \bar{X}_2}}{\sigma_{\bar{X}_1 - \bar{X}_2}}$$

### 4.2.3- DISTRIBUIÇÃO AMOSTRAL DE PROPORÇÕES:

Seja  $p$  a probabilidade de sucesso de um evento e  $q$  o insucesso. Consideremos todas as amostras possíveis de tamanho  $n$ , obtidas com e sem reposição, e para cada uma vamos calcular a proporção  $P$  de sucessos. Obtemos assim a Distribuição Amostral de Proporções com os parâmetros:  $\mu_p$  e  $\sigma_p$ .

Amostragem	Média	Desvio Padrão
Com Reposição (pop. Infinita)	$\mu_p = p$	$\sigma_p = \sqrt{\frac{pq}{n}}$
Sem Reposição (pop. Finita)	$\mu_p = p$	$\sigma_p = \sqrt{\frac{pq}{n}} \sqrt{\frac{N-n}{N-1}}$

Obs:

1. Para grandes valores de  $n$  a distr. É aproximadamente normal.
2. A população é binomial.
3. A variável padronizada  $Z$  será:

$$Z = \frac{P - \mu_p}{\sigma_p}$$

## V- ESTIMAÇÃO DE PARÂMETROS

### 5.1. INTRODUÇÃO:

Seja  $X_1, X_2, \dots, X_n$  uma amostra aleatória com função (densidade) de probabilidade conhecida, seja ainda  $\theta$  um vetor dos parâmetros desta variável aleatória. Assim  $\theta = \{\theta_1, \theta_2, \dots, \theta_k\}$  os  $k$  parâmetros que chamamos de espaço de parâmetros denotado por  $\Theta$ . Então o objetivo da inferência estatística é encontrar funções das observações  $X_1, X_2, \dots, X_n$  para usar como estimador de  $\theta_j$  onde  $j=1,2,\dots,k$ .

**ESTIMADOR:** é um estatística (função conhecida de v.a. observáveis que também é um v.a.) cujos valores são usados para estimar alguma função do parâmetro  $\theta$ . Ex: para estimar  $\mu$  (média populacional) o estimador mais adequado é  $\bar{X}$  (média aritmética da amostra).

**ESTIMATIVA:** é o valor numérico obtido para o estimador numa certa amostra.

**TIPOS :**

**PONTUAL:** a estimativa é representado por um único valor.

**POR INTERVALO:** a estimativa é representada por um intervalo.

### 5.2 . ESTIMAÇÃO PONTUAL:

Melhor estimador para a  $\mu$  (média populacional) é  $\bar{X}$  a média aritmética amostral dada por:

$$\hat{\mu} = \bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$

Melhor estimador para a  $\sigma^2$  (variância populacional) e  $s^2$  a variância amostral dada por:

$$\hat{\sigma}^2 = s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}$$

### 5.3. ESTIMAÇÃO POR INTERVALO:

Estimação por intervalo consiste na construção de um intervalo em torno da estimativa pontual, de modo que esse tenha uma probabilidade conhecida de conter o verdadeiro valor do parâmetro.

### 5.3.1. INTERVALO DE CONFIANÇA PARA MÉDIA POPULACIONAL:

1. Quando  $n \geq 30$  ou  $\sigma$  for conhecido:

$$P(\bar{x} - z_{\alpha/2} \cdot \sigma_{\bar{X}} \leq \mu \leq \bar{x} + z_{\alpha/2} \cdot \sigma_{\bar{X}}) = 1 - \alpha$$

2. Quando  $n < 30$ ,  $\sigma$  desconhecido e população normalmente distribuída:

$$P(\bar{x} - t_{n-1, \alpha/2} \cdot \hat{\sigma}_{\bar{X}} \leq \mu \leq \bar{x} + t_{n-1, \alpha/2} \cdot \hat{\sigma}_{\bar{X}}) = 1 - \alpha$$

#### Observação:

Podemos determinar o tamanho de amostra isolando o valor de  $n$  na precisão da estimativa (semi-amplitude) que no caso da média populacional é dada por:

$$e_0 = z_{\alpha/2} \sigma_{\bar{x}}$$

### 5.3.2. INTERVALO DE CONFIANÇA PARA PROPORÇÃO POPULACIONAL:

$$P(\hat{p} - z_{\alpha/2} \sigma_p \leq p \leq \hat{p} + z_{\alpha/2} \sigma_p) = 1 - \alpha$$

onde  $\hat{p}$  é o estimador de  $p$ , que pode ser dado por:

$$\hat{p} = \frac{f}{n}$$

sendo:  $e_0 = z_{\alpha/2} \hat{\sigma}_p$

### 5.4 EXERCÍCIOS:

5.4.1 Uma amostra de 100 esferas apresenta diâmetro médio de 2,09cm e desvio padrão de 0,11cm. Estime o diâmetro médio populacional, com 95% de confiança.

5.4.2 De 500 lâmpadas fabricadas por uma companhia retira-se uma amostra de 40 lâmpadas, e obtém-se a vida média de 800 horas e o desvio padrão de 100 horas. Qual o intervalo de confiança para a média populacional, utilizando um coeficiente de confiança de 99%.

5.4.3 Foram realizadas 16 determinações de densidade ( $\text{g/cm}^3$ ) de certo metal, obtendo-se os resultados:

19,7	19,8	19,9	19,9	19,7	19,6	19,5	19,6
19,9	19,6	19,9	19,5	19,8	19,6	19,8	19,7

Determine o intervalo de confiança de 99% para a média populacional.

- 5.4.4 Sabe-se por pesquisas já realizadas que o desvio padrão das tensões limites de tração de barras de aço é  $15 \text{ kgf/mm}^2$ , e que uma amostra de 26 barras foram ensaiadas apresentando tração média igual a  $70 \text{ kgf/mm}^2$ . Estimar a verdadeira tensão limite de tração utilizando 99% de confiança.
- 5.4.5 Determine uma estimativa pontual para a média populacional do problema anterior.
- 5.4.6 Uma amostra de 625 donas-de-casa revela que 70% delas preferem a marca X de detergente. Construir um intervalo de confiança de 99% para  $p =$  proporção das donas-de-casa que preferem o detergente X.
- 5.4.7 Antes de uma eleição em que existiam 2 candidatos  $A$  e  $B$ , foi feita uma pesquisa com 400 eleitores escolhidos ao acaso, e verificou-se que 208 deles pretendiam votar no candidato  $A$ . Construa um intervalo de confiança, com 95% para a porcentagem de eleitores favoráveis ao candidato  $A$  na época das eleições.
- 5.4.8 Qual deve ser o tamanho de uma amostra cujo desvio padrão é 10 para que a diferença da média amostral para a média populacional, em valor absoluto seja menor que 1, com coeficiente de confiança igual a 95%.

## VI-TESTES DE HIPÓTESES

### 6.1 INTRODUÇÃO:

**6.1.1 HIPÓTESES:** são suposições que fazemos para testar a fixação de decisões, que poderão ser verdadeiras ou não.

### 6.1.2 HIPÓTESES ESTATÍSTICA:

Hipótese Nula ( $H_0$ ): a ser validada pelo teste.

Hipótese Alternativa ( $H_1$  ou  $H_a$ ): complementar a  $H_0$ .

Assim, o teste poderá aceitar ou rejeitar a hipótese nula, sendo que no último caso implicaria na aceitação da hipótese alternativa.

### 6.1.3 RISCOS DE TOMADAS DE DECISÕES:

DECISÃO	REALIDADE	
	$H_0$ Verdadeira	$H_0$ Falsa
Aceita $H_0$	Decisão Correta (1 - $\alpha$ )	Erro Tipo II $\beta$
Rejeita $H_0$	Erro Tipo I $\alpha$	Decisão Correta (1 - $\beta$ )

Ex1: Decisão de um professor:

DECISÃO	REALIDADE	
	$H_0$ Verdadeira <b>Estudou</b>	$H_0$ Falsa <b>Não Estudou</b>
Aceita $H_0$ <b>Aprova Aluno</b>	Decisão Correta (1 - $\alpha$ )	Erro Tipo II $\beta$
Rejeita $H_0$ <b>Reprova Aluno</b>	Erro Tipo I $\alpha$	Decisão Correta (1 - $\beta$ )

Ex 2: Decisão de um médico:

DECISÃO	REALIDADE	
	$H_0$ Verdadeira <b>Precisa Operar</b>	$H_0$ Falsa <b>Não Precisa Operar</b>
Aceita $H_0$ <b>Opera</b>	Decisão Correta (1 - $\alpha$ )	Erro Tipo II $\beta$
Rejeita $H_0$ <b>Não Opera</b>	Erro Tipo I $\alpha$	Decisão Correta (1 - $\beta$ )

Ex 3: Decisão do julgamento de um réu:

DECISÃO	REALIDADE	
	H <sub>0</sub> Verdadeira <b>Inocente</b>	H <sub>0</sub> Falsa <b>Culpado</b>
Aceita H <sub>0</sub> <b>Não Prende o Réu</b>	Decisão Correta (1 - $\alpha$ )	Erro Tipo II $\beta$
Rejeita H <sub>0</sub> <b>Prende o Réu</b>	Erro Tipo I $\alpha$	Decisão Correta (1 - $\beta$ )

Ex4: Decisão em Controle de Qualidade:

DECISÃO	REALIDADE	
	H <sub>0</sub> Verdadeira <b>Lote Bom</b>	H <sub>0</sub> Falsa <b>Lote Ruim</b>
Aceita H <sub>0</sub> <b>Aceita o Lote</b>	Decisão Correta (1 - $\alpha$ )	Erro Tipo II $\beta$ (Risco do Consumidor)
Rejeita H <sub>0</sub> <b>Rejeita o Lote</b>	Erro Tipo I $\alpha$ (Risco do Produtor)	Decisão Correta (1 - $\beta$ )

#### 6.1.3.1 ERRO TIPO I ( $\alpha$ )-

$\alpha$  = P(rejeitar H<sub>0</sub> / H<sub>0</sub> verdadeira) chamado de **nível de significância**

#### 6.1.3.2 ERRO TIPO II ( $\beta$ )-

$\beta$  = P(aceitar H<sub>0</sub> / H<sub>0</sub> falsa)

#### 6.1.4 REGIÕES DE DECISÃO:

Região de Aceitação (R.A.): é a região da curva, delimitada por valores tabelados a um determinado nível de significância ( $\alpha$ ), que contém os valores para os quais aceitamos a hipótese nula.

Região Crítica ou de Rejeição (RC ou RR): é a região da curva, delimitada por valores tabelados a um determinado nível de significância ( $\alpha$ ), que contém os valores para os quais rejeitamos a hipótese nula.

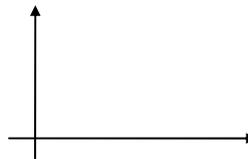
## 6.1.5 CLASSES DE TESTES:

### 6.1.5.1 UNILATERAIS:

#### 6.1.5.1.1 UNILATERAL À DIREITA

$$H_0: \theta = \theta_0 \quad (\theta \leq \theta_0)$$

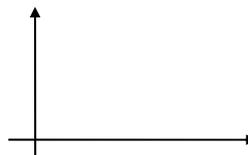
$$H_1: \theta > \theta_0$$



#### 6.1.5.1.2 UNILATERAL À ESQUERDA:

$$H_0: \theta = \theta_0 \quad (\theta \geq \theta_0)$$

$$H_1: \theta < \theta_0$$



### 6.1.5.2 BILATERAIS:

$$H_0: \theta = \theta_0$$

$$H_1: \theta \neq \theta_0$$



## 6.2 TESTES PARAMÉTRICOS:

### 6.2.1 TESTES DA MÉDIA POPULACIONAL:

#### 6.2.1.1 TESTE PARA UMA MÉDIA POPULACIONAL

Estatística do Teste:

1) Quando  $n \geq 30$  ou  $\sigma$  conhecido

$$Z_c = \frac{\bar{X} - \mu_0}{\sigma_{\bar{x}}}$$

onde:  $\mu_0$  = valor de  $\mu$  proposto pelo teste

$\sigma_{\bar{x}}$  = desvio padrão da distribuição amostral de médias

Obs: Caso o desvio padrão seja desconhecido usar  $S_{\bar{x}}$  como estimador pontual de  $\sigma_{\bar{x}}$ .

2) Quando  $n < 30$ ,  $\sigma$  desconhecido e a população normal:

$$t_c = \frac{\bar{X} - \mu_0}{S_{\bar{x}}}$$

CONCLUSÃO: quando o valor calculado (  $Z_c$  ou  $t_c$  ) cair na região de aceitação, deve-se aceitar  $H_0$ , caso contrário deve-se rejeitar  $H_0$  e toma-se  $H_1$  como verdadeira.

### 6.2.1.2 TESTE PARA DUAS MÉDIAS POPULACIONAIS:

As hipóteses serão enunciadas da seguinte maneira:

$$H_0 = \mu_1 - \mu_2 = \Delta \Rightarrow \text{quando } \Delta = 0 \Rightarrow \mu_1 = \mu_2$$

$$H_1 = \mu_1 - \mu_2 \neq \Delta \Rightarrow \text{quando } \Delta = 0 \Rightarrow \mu_1 \neq \mu_2$$

$$\text{ou } \mu_1 - \mu_2 > \Delta \Rightarrow \text{quando } \Delta = 0 \Rightarrow \mu_1 > \mu_2$$

$$\text{ou } \mu_1 - \mu_2 < \Delta \Rightarrow \text{quando } \Delta = 0 \Rightarrow \mu_1 < \mu_2$$

#### 6.2.1.2.1 DADOS EMPARELHADOS:

Quando duas amostras estão correlacionadas segundo algum critério (por exemplo o caso antes e depois).

Estatística do Teste:

$$t_c = \frac{\bar{d} - \Delta}{s_d / \sqrt{n}}$$

onde:

$\bar{d}$  = média da amostra das diferenças

$\Delta$  = valor testado da média das diferenças nas populações

$s_d$  = desvio padrão da amostra das diferenças

$n$  = tamanho da amostra das diferenças (iguais para as 2 amostras)

#### 6.2.1.2.2 DADOS NÃO EMPARELHADOS:

Neste caso as amostras podem ter tamanhos diferentes ( $n_1$  e  $n_2$ ).

PRIMEIRO CASO: As duas variâncias são conhecidas.

Estatística do Teste:

$$Z_c = \frac{(\bar{x}_1 - \bar{x}_2) - \Delta}{\sqrt{\sigma_{\bar{x}_1}^2 + \sigma_{\bar{x}_2}^2}}$$

SEGUNDO CASO: As duas variâncias não são conhecidas, mas podemos admitir que as variâncias sejam iguais ( $\sigma_1^2 = \sigma_2^2 = \sigma^2$ ).

Estatística do Teste:

$$t_{n_1+n_2-2} = \frac{(\bar{x}_1 - \bar{x}_2) - \Delta}{\sqrt{s_p^2 [(1/n_1) + (1/n_2)]}}$$

$$v = gl = n_1 + n_2 - 2$$

onde:

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

TERCEIRO CASO: As duas variâncias não são conhecidas e diferentes ( $\sigma_1^2 \neq \sigma_2^2$ ).

Estatística do Teste:

$$t_v = \frac{(\bar{x}_1 - \bar{x}_2) - \Delta}{\sqrt{(s_1^2 / n_1) + (s_2^2 / n_2)}}$$

onde os graus de liberdade são dados por:

$$v = \frac{(w_1 + w_2)^2}{[w_1^2 / (n_1 + 1)] + [w_2^2 / (n_2 + 1)]} - 2$$

$w_1$  e  $w_2$  são calculados por:

$$w_1 = \frac{s_1^2}{n_1} \quad e \quad w_2 = \frac{s_2^2}{n_2}$$

## 6.2.2 TESTES DA VARIÂNCIA POPULACIONAL:

### 6.2.2.1 TESTE DE UMA VARIÂNCIA POPULACIONAL:

As hipóteses a serem testadas serão:

$$H_0 : \sigma^2 = \sigma_0^2$$

$$H_1 : \sigma^2 \neq \sigma_0^2$$

$$\sigma^2 > \sigma_0^2$$

$$\sigma^2 < \sigma_0^2$$

Estatística do Teste:

$$\chi_{n-1}^2 = \frac{(n-1)S^2}{\sigma^2}$$

### 6.2.2.2 TESTE DE COMPARAÇÃO DE DUAS VARIÂNCIAS:

As hipóteses a serem testadas serão:

$$H_0 : \sigma_1^2 = \sigma_2^2$$

$$H_1 : \sigma_1^2 \neq \sigma_2^2$$

$$\sigma_1^2 > \sigma_2^2$$

$$\sigma_1^2 < \sigma_2^2$$

Estatística do Teste:

$$F_{v_N, v_D} = \frac{\max(s_1^2, s_2^2)}{\min(s_1^2, s_2^2)}$$

### 6.3 TESTES NÃO PARAMÉTRICOS:

#### 6.3.1 TESTE DE ADERÊNCIA:

Objetivo: Verificar a boa ou má aderência dos dados de uma amostra a um modelo proposto.

Utiliza-se o teste  $\chi^2$  de Karl Pearson:

Evento	Frequência Observada na Amostra	Frequência Esperada por	Alguns
Modelo			
A <sub>1</sub>	O <sub>1</sub>		E <sub>1</sub>
A <sub>2</sub>	O <sub>2</sub>		E <sub>2</sub>
·	·		·
·	·		·
·	·		·
A <sub>n</sub>	O <sub>n</sub>		E <sub>n</sub>

As hipóteses a serem testadas serão:

H<sub>0</sub>: Não há diferença entre as frequências observadas e as esperadas.

H<sub>1</sub>: Há diferença entre as frequências observadas e as esperadas.

Estatística do Teste:

$$\chi_v^2 = \sum_{i=1}^n \frac{(O_i - E_i)^2}{E_i}$$

Obs:

1.  $v = n-1 \Rightarrow$  se as freq. esperadas foram calculadas sem recorrer a estimação de algum parâmetro populacional

$v = n - 1 - m \Rightarrow$  se as freq. esperadas foram calculadas a partir de  $m$  parâmetros populacionais.

2. Se  $E_i < 5$  deve-se agrupar as classes adjacentes.

### 6.3.2 TABELAS DE CONTINGÊNCIA - TESTE DE INDEPENDÊNCIA:

Utilizamos este teste quando existem duas ou mais variáveis de interesse e desejamos verificar se existe associação entre elas.

As hipóteses a serem testadas serão:

$H_0$ : As variáveis são independentes..

$H_1$ : As variáveis não são independentes ou seja elas apresentam algum grau de associação entre si.

Estatística do Teste:

$$\chi_v^2 = \sum_{j=1}^r \sum_{i=1}^s \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \quad \text{onde: } r = n^\circ \text{ de linhas}$$

$$s = n^\circ \text{ de colunas}$$

e

$$v = (r - 1).(s - 1)$$

$$E_{ij} = \frac{\text{Total coluna} \times \text{Total linha}}{\text{Total Geral}}$$

### 6.4 EXERCÍCIO:

6.4.1 Sabe-se que o consumo mensal *per capita* de um determinado produto tem distribuição normal com desvio padrão de 2kg. A diretoria de uma firma que fabrica esse produto resolveu que retiraria o produto da linha de produção se a média de consumo *per capita* fosse menor que 8kg. Caso contrário, continuaria a fabricá-lo. Foi realizada uma pesquisa de mercado com 25 indivíduos dados a seguir:

9.8 8.5 7.0 10.4 8.9 7.9 5.9 7.7 8.8 7.7 7.9 5.1 9.9 7.6 8.4  
3.7 6.8 7.5 8.5 5.1 7.6 5.0 6.1 6.1 10.2

Verificar, ao nível de 5% de significância, que posição deve tomar a diretoria.

6.4.2 Uma amostra de 5 cabos de aço foi ensaiada, antes e após sofrer um tratamento para aumentar a sua resistência. Os resultados são apresentados a seguir:

Cabos	1	2	3	4	5
Antes	50	54	51	50	55
Depois	60	61	57	54	59

Verifique se o tratamento foi eficiente, utilizando um nível de 5% de significância.

6.4.3 Dois candidatos a um emprego, A e B, foram submetidos a um conjunto de oito questões, sendo anotados os tempos que cada um gastou na solução. Podemos, ao nível de 5% de significância, concluir que B seja mais rápido que A, em termos do tempo médio gasto para resolver questões do tipo das formuladas?

Questão	1 <sup>a</sup>	2 <sup>a</sup>	3 <sup>a</sup>	4 <sup>a</sup>	5 <sup>a</sup>	6 <sup>a</sup>	7 <sup>a</sup>	8 <sup>a</sup>
Tempo de A	11	8	15	2	7	18	9	10
Tempo de B	5	7	13	6	4	10	3	12

6.4.4 Foram ensaiadas lâmpadas das marcas A e B. Verificou-se que os tempos de vida (em horas) foram:

A	1500	1450	1480	1520	1510
B	1000	1300	1180	1250	

Podemos concluir, ao nível de significância de 1%, que o tempo médio de vida da marca A supera o de B em mais de 300h?

6.4.5 A fim de comparar duas marcas de cimento, A e B, fizemos experiência com quatro corpos de prova da marca A e cinco da marca B, obtendo-se as seguintes resistências à ruptura:

Marca A	184	190	185	186	
Marca B	189	188	183	186	184

Verifique se as resistências médias das duas marcas diferem si ao nível de 5%.

6.4.6 Dois fertilizantes A e B, para a produção de certa variedade de tomate vão ser comparados, em termos do peso médio de produção. As produções, em kg, de 10 pés de tomate sob o fertilizante A e 12 sob o B foram:

A	1,6	1,7	1,8	1,4	1,5	1,9	2,3	2,1	1,9	1,7		
B	2,0	2,1	1,5	1,9	1,9	2,3	1,8	1,9	2,1	2,4	2,5	2,7

Qual a conclusão ao nível de 1% de significância

6.4.7 As freq. observadas de 120 jogadas de um dado apresentam-se na tabela abaixo. Teste a hipótese de que o dado é honesto, utilizando um nível de significância de 5%.

Face	1	2	3	4	5	6
Freq.Observada	25	17	15	23	24	16

6.4.8 A indústria K.B.S. usa oito máquinas para a produção de 9 mil unidades/dia. Uma amostra retirada em determinado tempo apresentada abaixo permite supor que as máquinas são igualmente produtivas?

Máquina	1	2	3	4	5	6	7	8
Produção	6	9	8	5	4	6	10	7

6.4.9 Verificar se existe associação entre gênero (masculino e feminino) e tabagismo (fumante e não-fumante) utilizando um nível de significância de 1%, numa certa população, onde se observou uma amostra aleatória de 300 pessoas adultas. Os dados são apresentados a seguir.

	Masculino	Feminino	Total
Fumante	92	38	130
Não Fumante	108	62	170
Total	200	100	300

## VII- ANÁLISE DA VARIÂNCIA (COMPARAÇÃO DE VÁRIAS MÉDIAS)

A Análise da Variância é um método suficientemente poderoso para identificar diferenças entre as médias populacionais devidas a várias causas, atuando simultaneamente sobre os elementos da população.

### 7.1. HIPÓTESES.

$$H_0: \mu_1 = \mu_2 = \dots = \mu_k$$

$H_1$ : pelo menos uma das médias é diferente

Ainda se impusermos algumas condições:

- as k populações tem a mesma variância (Homoscedasticidade)
- as k populações sejam normalmente distribuídas.

teremos ainda um **modelo robusto** (mesmo quando as hipóteses básicas não forem válidas, o modelo ainda leva a resultado com razoável aproximação)

### 7.2 ANÁLISE DA VARIÂNCIA COM UM CRITÉRIO DE CLASSIFICAÇÃO.

#### 7.2.1 AMOSTRAS DE TAMANHOS IGUAIS.

Fonte de Variação	Soma dos Quadrados	Graus de Liberdade	Quadrados Médios	$F_c$	$F_\alpha$
Entre as Amostras	$SQE = \sum_{i=1}^k \frac{T_i^2}{n} - \frac{T^2}{nk}$	k-1	$s_E^2 = \frac{SQE}{k-1}$	$F_c = \frac{s_E^2}{s_R^2}$	$F_{k-1, k(n-1)}$
Residual	$SQR = Q - \sum_{i=1}^k \frac{T_i^2}{n}$	k(n-1)	$s_R^2 = \frac{SQR}{k(n-1)}$		
Total	$SQT = Q - \frac{T^2}{n.k}$	nk-1			

Onde:

$$T_i = \sum_{j=1}^n x_{ij} = \text{soma dos valores da i-ésima amostra}$$

$$Q_i = \sum_{j=1}^n x_{ij}^2 = \text{soma dos quadrados dos valores da i-ésima amostra}$$

$$T = \sum_{i=1}^k T_i = \sum_{i=1}^k \sum_{j=1}^n x_{ij} = \text{soma de todos os valores}$$

$$Q = \sum_{i=1}^k Q_i = \sum_{i=1}^k \sum_{j=1}^n x_{ij}^2 = \text{soma dos quadrados de todos os valores}$$

n = tamanho da amostra

k = número de critérios

### 7.2.2 AMOSTRAS DE TAMANHOS DIFERENTES.

Fonte de Variação	Soma dos Quadrados	Graus de Liberdade	Quadrados Médios	$F_c$	$F_\alpha$
Entre as Amostras	$SQE = \sum_{i=1}^k \frac{T_i^2}{n_i} - \frac{T^2}{\sum_{i=1}^k n_i}$	$k-1$	$s_E^2 = \frac{SQE}{k-1}$	$F_c = \frac{S_e^2}{S_R^2}$	$F_{k-1, \sum_{i=1}^k n_i - k}$
Residual	$SQR = Q - \sum_{i=1}^k \frac{T_i^2}{n_i}$	$\sum_{i=1}^k n_i - k$	$s_R^2 = \frac{SQR}{\sum_{i=1}^k n_i - k}$		
Total	$SQT = Q - \frac{T^2}{\sum_{i=1}^k n_i}$	$\sum_{i=1}^k n_i - 1$			

Onde:

$$T_i = \sum_{j=1}^n x_{ij} = \text{soma dos valores da } i\text{-ésima amostra}$$

$$Q_i = \sum_{j=1}^n x_{ij}^2 = \text{soma dos quadrados dos valores da } i\text{-ésima amostra}$$

$$T = \sum_{i=1}^k T_i = \sum_{i=1}^k \sum_{j=1}^n x_{ij} = \text{soma de todos os valores}$$

$$Q = \sum_{i=1}^k Q_i = \sum_{i=1}^k \sum_{j=1}^n x_{ij}^2 = \text{soma dos quadrados de todos os valores}$$

$n$  = tamanho da amostra e  $k$  = número de critérios

### 7.3 EXERCÍCIOS:

7.3.1 Quatro pneus de cada uma das marcas A, B e C foram testados quanto a durabilidade. O resultados obtidos foram:

Marca	Durabilidade (meses)			
A	34	38	31	35
B	32	34	31	29
C	30	25	28	23

Ao nível de significância de 1%, há evidência de que os pneus tenham diferentes durabilidades médias?

7.3.2 Foram testados três tipos de lâmpadas elétricas e o tempo de vida (em horas) são dados a seguir.

Lâmpada A	1245	1354	1367	1289	
Lâmpada B	1235	1300	1230	1189	1250
Lâmpada C	1345	1450	1320		

Podemos identificar, ao nível de 5% de significância, a existência de diferença entre as médias das populações das quais provieram essas amostras?

## VIII- REGRESSÃO E CORRELAÇÃO LINEAR SIMPLES

### 8.1 INTRODUÇÃO:

1. Relacionamento entre variáveis :  
- requer conhecimento

$$Y = f(X) + \varepsilon$$

Ex: 1.

Y = Produção agrícola  
X = Fertilizante

Y(v.aleatória) em função de X(v.determinística), onde Y é a variável explicada por X .

Y - variável explicada ou dependente de X  
X - variável explicativa ou independente

### 8.2. DIAGRAMA DE DISPERSÃO:

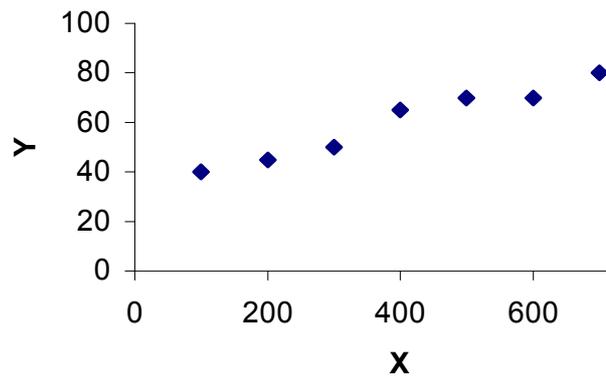
X (lb/acre)	Y (bushel/acre)
100	40
200	45
300	50
400	65
500	70
600	70
700	80

### 8.3. CORRELAÇÃO LINEAR:

$$r_{XY} = \frac{n \sum_{i=1}^n XY - \left( \sum_{i=1}^n X_i \right) \left( \sum_{i=1}^n Y_i \right)}{\sqrt{\left[ n \sum_{i=1}^n X_i^2 - \left( \sum_{i=1}^n X_i \right)^2 \right] \left[ n \sum_{i=1}^n Y_i^2 - \left( \sum_{i=1}^n Y_i \right)^2 \right]}}$$

$$-1 \leq r_{XY} \leq 1$$

## DIAGRAMA DE DISPERSÃO



No caso do exemplo 1:

$$r_{XY} = \frac{(19.000)}{\sqrt{(28 \times 10^8) \cdot (1.350)}} = 0,977$$

Estudaremos o relacionamento linear entre as variáveis, assim:

$$Y = \alpha + \beta X + \varepsilon$$

Suposições:

1. A relação de X e Y é linear e há efeito causal entre elas.
2. X é uma variável não estocástica e conhecida
3. Considerações a cerca do erro:

- 3.1.  $\varepsilon_i \sim N(0, \sigma^2)$

donde vemos que:

$$E(\varepsilon_i) = 0 \quad \forall i$$

$$Var(\varepsilon_i) = \sigma^2 \quad (\text{constante, por isto não é indexada})$$

Modelo Homoscedático

- 3.2. Não há correlação serial entre o erro aleatório, isto é, Os erros são independentes.

$$E[\varepsilon_i \varepsilon_j] = 0 \quad \forall i \neq j$$

#### 8.4. LEAST SQUARE SOLUTION (SOLUÇÃO DOS MÍNIMOS QUADRADOS) OU ORDINARY LEAST SQUARE (OLS) (MÍNIMOS QUADRADOS ORDINÁRIOS):

Seja a equação da reta:

$$Y = \alpha + \beta X$$

A idéia é estimar os parâmetros  $\alpha$  e  $\beta$  de tal maneira que a soma dos quadrados dos desvios seja mínima, isto é;

$$\text{minimizar } \sum_{i=1}^n (\varepsilon_i)^2 = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$$

substituindo  $\hat{Y}_i = a + bX$  e derivando em relação a  $a$  e  $b$ , e igualando as expressões a zero, temos:

$$\begin{cases} \sum Y_i = na + b \sum X_i \\ \sum X_i Y_i = a \sum X_i + b \sum X_i^2 \end{cases}$$

Resolvendo o sistema temos:

$$a = \frac{\Delta a}{\Delta} = \frac{\begin{vmatrix} \sum Y_i & \sum X_i \\ \sum X_i Y_i & \sum X_i^2 \end{vmatrix}}{\begin{vmatrix} n & \sum X_i \\ \sum X_i & \sum X_i^2 \end{vmatrix}}$$

$$b = \frac{\Delta b}{\Delta} = \frac{\begin{vmatrix} n & \sum Y_i \\ \sum X_i & \sum X_i Y_i \end{vmatrix}}{\begin{vmatrix} n & \sum X_i \\ \sum X_i & \sum X_i^2 \end{vmatrix}}$$

No caso do exemplo teremos:

$$a = 32,857143 \quad \text{e} \quad b = 0,0678571$$

Assim:

$$\hat{Y} = 32,857143 + 0,06785571 X$$

**8.5 EXERCÍCIO:**

8.5.1 A tabela abaixo apresenta os dados relacionados com o número de semanas de experiência de colocar fios em pequenos componentes eletrônicos bem como o número de tais componentes que foram rejeitados durante uma determinada semana, dados estes referentes a 12 trabalhadores aleatoriamente escolhidos.

Trabalhador amostrado	1	2	3	4	5	6	7	8	9	10	11	12
Semanas de experiência	7	9	6	14	8	12	10	4	2	11	1	8
Quantidade de rejeitados	26	20	28	16	23	18	24	26	38	22	32	25

- Verifique se há correlação entre os dados.
- Determine a equação de regressão linear.
- Estimar o número de componentes rejeitados para um empregado com 3 semanas de experiência.

**BIBLIOGRAFIA**

1. BUSSAB, W. O. & MORETTIN, P. A. Estatística Básica. Atual Editora, 1987.
2. COSTA NETO, P. L. de O. Estatística. Editora Edgard Blücher, 2002.
3. CHAVES NETO, A. Notas de aulas da disciplina Probabilidade e Estatística Aplicada. DEST/UFPR, 2000.
4. DEVORE, Jay L. Probabilidade e Estatística para Engenharia e Ciências. Editora Thomson, 2006
5. LIPSCHUTZ, Seymour. Probabilidade. Coleção Schaum, Editora McGraw-Hill, 1993.
6. MARQUES, Jair M. Notas de aula da disciplina Probabilidade e Estatística Aplicada. DEST/UFPR, 1994.
7. MEYER, Paul L. Probabilidade. Aplicações à Estatística. Livros Técnicos e Científicos Editora S.A, 1976
8. MENDENHALL, W. Probabilidade e Estatística. Editora Campus, Vol.1 e Vol.2, 1985.
9. MOOD A. M, GRAYBILL F., BOES, D. C. Introduction to the Theory of Statistics. Editora McGraw-Hill, 1974.
7. MORETTIN, Luiz Gonzaga. Estatística Básica : Inferência. Makron Books, Vol. 2, 2000.
8. SPIEGEL, M. L. Estatística. Coleção Schaum, Editora McGraw-Hill, 1972.
9. SPIEGEL, M. L. Probabilidade e Estatística. Coleção Schaum, Editora McGraw-Hill, 1978.





TABELA 2- Distribuição  $t$  de Student

<b>v</b> <b>Graus de</b> <b>Liberdade</b>	<b><math>\alpha</math> (nível de significância)</b>			
	<b>0,005</b> <b>(unilateral)</b> <b>0,01</b> <b>(bilateral)</b>	<b>0,01</b> <b>(unilateral)</b> <b>0,02</b> <b>(bilateral)</b>	<b>0,025</b> <b>(unilateral)</b> <b>0,05</b> <b>(bilateral)</b>	<b>0,05</b> <b>(unilateral)</b> <b>0,10</b> <b>(bilateral)</b>
1	63,657	31,821	12,706	6,314
2	9,925	6,965	4,303	2,920
3	5,841	4,541	3,182	2,353
4	4,604	3,747	2,776	2,132
5	4,032	3,365	2,571	2,015
6	3,707	3,143	2,447	1,943
7	3,500	2,998	2,365	1,895
8	3,355	2,896	2,306	1,860
9	3,250	2,821	2,262	1,833
10	3,169	2,764	2,228	1,812
11	3,106	2,718	2,201	1,796
12	3,054	2,681	2,179	1,782
13	3,012	2,650	2,160	1,771
14	2,977	2,625	2,145	1,761
15	2,947	2,602	2,132	1,753
16	2,921	2,584	2,120	1,746
17	2,898	2,567	2,110	1,740
18	2,878	2,552	2,101	1,734
19	2,861	2,540	2,093	1,729
20	2,845	2,528	2,086	1,725
21	2,831	2,518	2,080	1,721
22	2,819	2,508	2,074	1,717
23	2,807	2,500	2,069	1,714
24	2,797	2,492	2,064	1,711
25	2,787	2,485	2,060	1,708
26	2,779	2,479	2,056	1,706
27	2,771	2,473	2,052	1,703
28	2,763	2,467	2,048	1,701
29	2,756	2,462	2,045	1,699
$\infty$ (z)	<b>2,575</b>	<b>2,327</b>	<b>1,960</b>	<b>1,645</b>

TABELA 3- Distribuição  $\chi^2$ 

v (gl)	$\alpha$					
	0,995	0,975	0,05	0,025	0,01	0,005
1	3,9 E-5	1E-4	3,841	5,024	6,635	7,879
2	0,010	0,051	5,991	7,378	9,210	10,597
3	0,072	0,216	7,815	9,348	11,345	12,838
4	0,207	0,484	9,488	11,143	13,277	14,860
5	0,412	0,831	11,070	12,833	15,086	16,750
6	0,676	1,237	12,592	14,450	16,812	18,548
7	0,989	1,690	14,067	16,013	18,475	20,278
8	1,344	2,180	15,507	17,535	20,090	21,955
9	1,735	2,700	16,919	19,023	21,666	23,589
10	2,156	3,247	18,307	20,483	23,209	25,188
11	2,603	3,816	19,675	21,920	24,725	26,757
12	3,074	4,404	21,026	23,337	26,217	28,300
13	3,565	5,009	22,362	24,736	27,688	29,819
14	4,075	5,629	23,685	26,119	29,141	31,319
15	4,601	6,262	24,996	27,488	30,578	32,802
16	5,142	6,908	26,296	28,845	32,000	34,267
17	5,697	7,564	27,587	30,191	33,409	35,718
18	6,265	8,231	28,869	31,526	34,805	37,156
19	6,844	8,907	30,144	32,852	36,191	38,582
20	7,434	9,591	31,410	34,170	37,566	39,997
21	8,034	10,283	32,671	35,479	38,932	41,401
22	8,643	10,982	33,924	36,781	40,289	42,796
23	9,260	11,689	35,172	38,076	41,638	44,181
24	9,886	12,401	36,415	39,364	42,980	45,558
25	10,520	13,120	37,652	40,647	44,314	46,928
26	11,160	13,844	38,885	41,923	45,642	48,290
27	11,808	14,573	40,113	43,195	46,963	49,645
28	12,461	15,308	41,337	44,461	48,278	50,994
29	13,121	16,047	42,557	45,722	49,588	52,336
30	13,787	16,791	43,773	46,979	50,892	53,672

**TABELA 4- Distribuição  $F$  (Snedecor)  $\alpha = 0,01$**

$v_D$	$v_n$																	
	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	$\infty$
<b>1</b>	4052	4999	5403	5625	5764	5859	5928	5982	6022	6056	6106	6157	6209	6235	6261	6287	6313	6366
<b>2</b>	98,50	99,00	99,17	99,25	99,30	99,33	99,36	99,37	99,39	99,40	99,42	99,43	99,45	99,46	99,47	99,47	99,48	99,50
<b>3</b>	34,12	30,82	29,46	28,17	28,24	27,91	27,67	27,49	27,35	27,23	27,05	26,87	26,69	26,60	26,50	26,41	26,32	26,13
<b>4</b>	21,20	18,00	16,69	15,98	15,52	15,21	14,98	14,80	14,66	14,55	14,37	14,20	14,02	13,93	13,84	13,75	13,65	13,46
<b>5</b>	16,20	13,27	12,06	11,39	10,97	10,67	10,46	10,29	10,16	10,05	9,89	9,72	9,55	9,47	9,38	9,29	9,20	9,02
<b>6</b>	13,75	10,92	9,78	9,15	8,75	8,47	8,26	8,10	7,98	7,87	7,72	7,56	7,40	7,31	7,23	7,14	7,06	6,88
<b>7</b>	12,25	9,55	8,45	7,85	7,46	7,19	6,99	6,84	6,72	6,62	6,47	6,31	6,16	6,07	5,99	5,91	5,82	5,65
<b>8</b>	11,26	8,65	7,59	7,01	6,63	6,37	6,18	6,03	5,91	5,81	5,67	5,52	5,36	5,28	5,20	5,12	5,03	4,86
<b>9</b>	10,56	8,02	6,99	6,42	6,06	5,80	5,61	5,47	5,35	5,26	5,11	4,96	4,81	4,73	4,65	4,57	4,48	4,31
<b>10</b>	10,04	7,56	6,55	5,99	5,64	5,39	5,20	5,06	4,94	4,85	4,71	4,56	4,41	4,33	4,25	4,17	4,08	3,91
<b>12</b>	9,33	6,93	5,95	5,41	5,06	4,82	4,64	4,50	4,39	4,30	4,16	4,01	3,86	3,78	3,70	3,62	3,54	3,36
<b>15</b>	8,68	6,36	5,42	4,89	4,56	4,32	4,14	4,00	3,89	3,80	3,67	3,52	3,37	3,29	3,21	3,13	3,05	2,87
<b>20</b>	8,10	5,85	4,94	4,43	4,10	3,87	3,70	3,56	3,46	3,37	3,23	3,09	2,94	2,86	2,78	2,69	2,61	2,42
<b>24</b>	7,82	5,61	4,72	4,22	3,90	3,67	3,50	3,36	3,26	3,17	3,03	2,89	2,74	2,66	2,58	2,49	2,40	2,21
<b>30</b>	7,56	5,39	4,51	4,02	3,70	3,47	3,30	3,17	3,07	2,98	2,84	2,70	2,55	2,47	2,39	2,30	2,21	2,01
<b>40</b>	7,31	5,18	4,31	3,83	3,51	3,29	3,12	2,99	2,89	2,80	2,66	2,52	2,37	2,29	2,20	2,11	2,02	1,80
<b>60</b>	7,08	4,98	4,13	3,65	3,34	3,12	2,95	2,82	2,72	2,63	2,50	2,35	2,20	2,12	2,03	1,94	1,84	1,60
<b><math>\infty</math></b>	6,63	4,61	3,78	3,32	3,02	2,80	2,64	2,51	2,41	2,32	2,18	2,04	1,88	1,79	1,70	1,59	1,47	1,00

TABELA 5- Distribuição  $F$  (Snedecor)  $\alpha = 0,05$ 

$v_D$	$v_n$																	
	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	$\infty$
1	161,4	199,5	215,7	224,6	230,2	234,0	236,8	238,9	240,5	241,9	243,9	245,9	248,0	249,1	250,1	251,1	252,2	254,3
2	18,51	19,00	19,16	19,25	19,30	19,33	19,35	19,37	19,38	19,40	19,41	19,43	19,45	19,45	19,46	19,47	19,48	19,50
3	10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,74	8,70	8,66	8,64	8,62	8,59	8,57	8,53
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,91	5,86	5,80	5,77	5,75	5,72	5,69	5,63
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,68	4,62	4,56	4,53	4,50	4,46	4,43	4,36
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	4,00	3,94	3,87	3,84	3,81	3,77	3,74	3,67
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64	3,57	3,51	3,44	3,41	3,38	3,34	3,30	3,23
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35	3,28	3,22	3,15	3,12	3,08	3,04	3,01	2,93
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	3,07	3,01	2,94	2,90	2,86	2,83	2,79	2,71
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	2,91	2,85	2,77	2,74	2,70	2,66	2,62	2,54
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75	2,69	2,62	2,54	2,51	2,47	2,43	2,38	2,30
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54	2,48	2,40	2,33	2,29	2,25	2,20	2,16	2,07
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35	2,28	2,20	2,12	2,08	2,04	1,99	1,95	1,84
24	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36	2,30	2,25	2,18	2,11	2,03	1,98	1,94	1,89	1,84	1,73
30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21	2,16	2,00	2,04	1,93	1,89	1,84	1,79	1,74	1,62
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,08	2,00	1,92	1,84	1,79	1,74	1,69	1,64	1,51
60	4,00	3,15	2,76	2,53	2,37	2,25	2,17	2,10	2,04	1,99	1,92	1,81	1,75	1,70	1,65	1,59	1,53	1,29
$\infty$	3,84	3,00	2,60	2,37	2,21	2,10	2,01	1,94	1,88	1,83	1,75	1,67	1,57	1,52	1,45	1,39	1,32	1,00

