UFRRJ
UNIVERSIDADE FEDERAL RURAL DO RIO DE JANEIRO
INSTITUTO TRÊS RIOS
DEPARTAMENTO DE CIÊNCIAS ECONÔMICAS E EXATAS

ESTATÍSTICA I

NOTAS DE AULA

Prof. Dr. Diógenes Ferreira Filho

TRÊS RIOS - RJ
Agosto de 2019

Sumário

I - Revisão de Estatística Descritiva 4

1 Tabelas 4

1.1 Distribuição de frequências para dados qualitativos 4

1.2 Distribuição de frequências para dados quantitativos 5

2 Gráficos 6

2.1 Gráfico de colunas (ou gráfico de barras verticais)..6

2.2 Histograma 6

3 Medidas de posição...7

3.1 Média aritmética 8

3.2 Mediana 8

3.3 Moda 10

3.4 Quartis...11

4 Medidas de dispersão 12

4.1 Variância...12

4.2 Desvio Padrão...14

5 Boxplot 15

II - Probabilidade 18

1 - Probabilidade 18

1.1 Espaço amostral, eventos 18

1.2 Técnicas de Contagem...26

1.3 Probabilidade condicional, Teorema de Bayes e independência de eventos 33

2. Variáveis aleatórias 42

2.1 Conceito. Valor esperado e variância de uma variável aleatória 42

2.2 Covariância e Correlação 57

2.3 Variáveis Aleatórias Múltiplas 60

3.1 Distribuição de Bernoulli 64

3.2 Distribuição Binomial 65

3.3 Distribuição de Poisson 68

3.4 Distribuição Geométrica 71

3.5 Distribuição Hipergeométrica..73

3.6 Distribuição exponencial 75

3.7 Distribuição Normal...77

III - Inferência Estatística 87

1 Introdução à inferência estatística. 87

1.1 Conceitos básicos. Amostra e população. 87

1.2 Amostragem aleatória simples: obtenção de uma amostra aleatória 88

1.3 Conceito de Distribuições amostrais 89

1.4 Distribuição amostral da média 90

2 Estimação 96

2.1 Conceitos básicos. Estimadores não viciados 96

2.2 Intervalo de confiança para média de uma população Normal com variância populacional conhecida. 99

2.3 Determinação do tamanho de uma amostra 104

2.4 Intervalo de confiança para a média de uma população Normal com variância populacional desconhecida 105

Apêndice 112

Apêndice 1. Tabela da distribuição Normal Padrão...112

Apêndice 2. Tabela (Bilateral) da distribuição t de Student 113

Bibliografia 114

Prefácio

Este material foi preparado com a intenção de cobrir o programa da disciplina Estatística I da Universidade Federal Rural do Rio de Janeiro. Ele é composto por notas de aula elaboradas à partir de livros e apostilas constantes na Bibliografia e não substitui a leitura dos mesmos.

O material não está livre de erros e/ou imperfeições e toda e qualquer contribuição será bem-vinda.

I - Revisão de Estatística Descritiva

1 Tabelas

1.1 Distribuição de frequências para dados qualitativos

Quando observamos dados qualitativos, classificamos cada observação em determinada categoria. Depois, contamos o número de observações em cada categoria. A idéia é resumir as informações na forma de uma tabela que mostre essas contagens (frequências) por categoria. Temos, então uma *tabela de distribuição de frequências*.

Exemplo

Considere os dados brutos do grau de instrução de 36 indivíduos:

Fundamental	Médio	Médio	Superior	Médio	Médio
Médio	Médio	Fundamental	Fundamental	Superior	Fundamental
Superior	Fundamental	Médio	Médio	Médio	Médio
Fundamental	Médio	Superior	Fundamental	Superior	Médio
Médio	Fundamental	Médio	Médio	Médio	Fundamental
Fundamental	Fundamental	Superior	Fundamental	Médio	Médio

Na Tabela 1.1 é apresentada a distribuição de frequências da variável grau de instrução cujos os dados foram apresentados acima.

Tabela 1.1. Distribuição de frequências de 36 indivíduos segundo o grau de instrução

Grau de instrução	**Frequência**
Fundamental	12
Médio	18
Superior	6
Total	36

Observando-se os resultados da segunda coluna da Tabela 1.1, vê-se que dos 36 indivíduos, 12 têm ensino fundamental, 18 o ensino médio e 6 possuem curso superior.

As tabelas de distribuição de frequência podem conter, ainda, além das frequências, as proporções e porcentagens, como pode-se observar na Tabela 1.2. As proporções (ou frequências relativas) são calculadas dividindo-se cada frequência pelo total, enquanto as porcentagens (ou frequências percentuais) são calculadas multiplicando-se cada proporção por 100. Assim, por exemplo para o ensino fundamental, obtemos a proporção fazendo $12/36 = 0{,}3333$, e a porcentagem fazendo $0{,}3333 \times 100 = 33{,}33\ \%$.

Tabela 1.2. Distribuição de frequências de 36 indivíduos segundo o grau de instrução

Grau de instrução	Frequência	Proporção	Porcentagem
Fundamental	12	0,3333	33,33
Médio	18	0,5000	50,00
Superior	6	0,1667	16,67
Total	36	1,0000	100,00

1.2 Distribuição de frequências para dados quantitativos

Tabela para dados contínuos

Costuma-se descrever as variáveis quantitativas contínuas através de **tabelas de classes de frequências** ou **tabelas de intervalos**.

Exemplo

Considere os salários (x sal. mín.) de 36 indivíduos:

4,00	4,56	5,25	5,73	6,26	6,66
6,86	7,39	7,59	7,44	8,12	8,46
8,74	8,95	9,13	9,35	9,77	9,80
10,53	10,76	11,06	11,59	12,00	12,79
13,23	13,60	13,85	14,69	14,71	15,99
16,22	16,61	17,26	18,75	19,40	23,30

Devemos agrupar os dados em faixas de salário, como podemos observar na Tabela 1.3.

Tabela 1.3. Distribuição de frequências dos salários de 36 indivíduos

Classes de salários	Frequência	Proporção	Porcentagem
$4{,}00 \vdash 8{,}00$	10	0,2778	27,78
$8{,}00 \vdash 12{,}00$	12	0,3333	33,33
$12{,}00 \vdash 16{,}00$	8	0,2222	22,22
$16{,}00 \vdash 20{,}00$	5	0,1389	13,89
$20{,}00 \vdash 24{,}00$	1	0,0278	2,78
Total	36	1,000	100,0

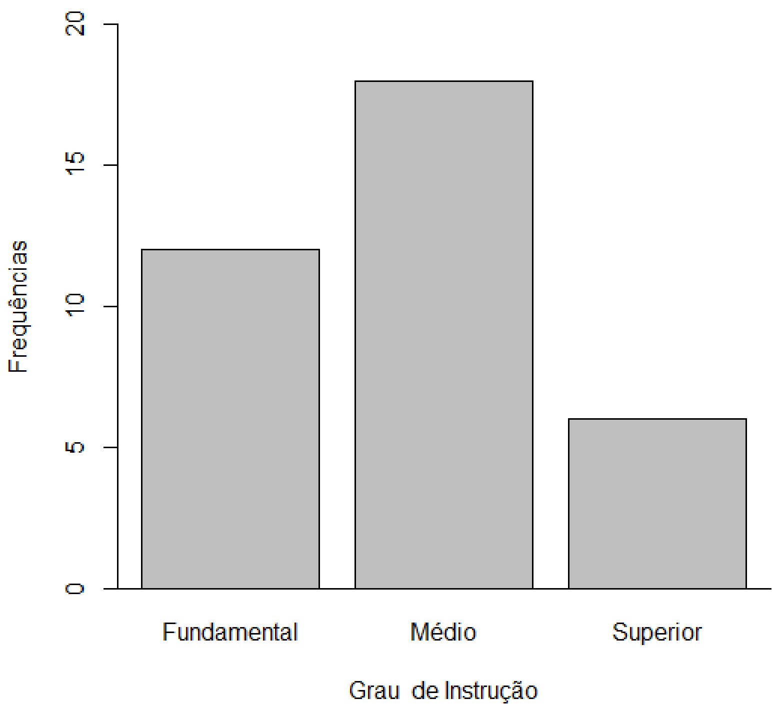
2 Gráficos

2.1 Gráfico de colunas (ou gráfico de barras verticais)

Exemplo

Considerando a tabela de distribuição de frequências para a variável *grau de instrução* de 36 indivíduos (Tabela 1.1), podemos representar essa distribuição de frequências por meio de um gráfico de colunas (Figura 2.1).

Figura 2.1. Gráfico de colunas para a variável *grau de instrução*



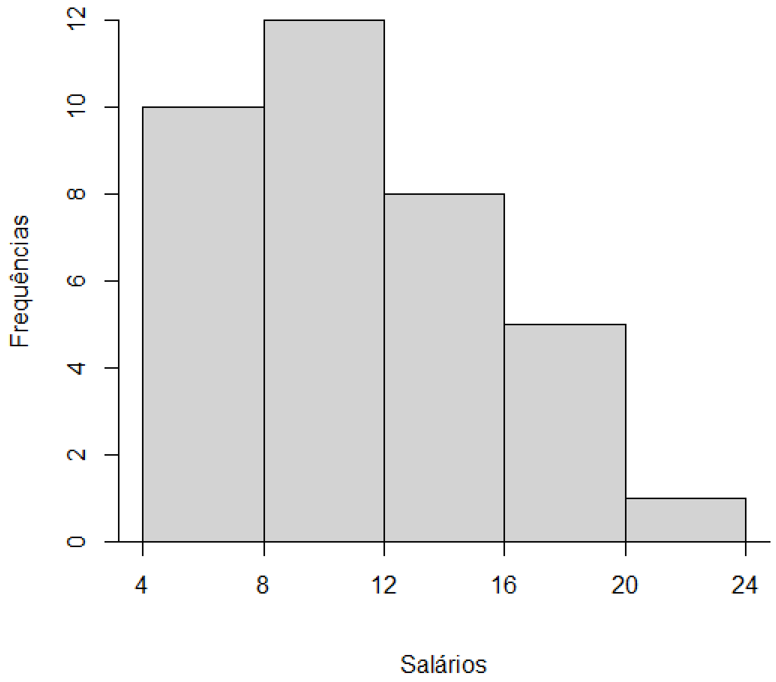
2.2 Histograma

O histograma é utilizado para apresentar graficamente dados *quantitativos contínuos* organizados em uma tabela de distribuição de frequências.

Exemplo

Considerando a tabela de distribuição de frequências para a variável *salário* (Tabela 1.3), podemos representar essa distribuição de frequências por meio de um histograma (Figura 2.2).

Figura 2.2. Histograma dos salários de 36 indivíduos



3 Medidas de posição

Vimos que o resumo de dados por meio de tabelas de frequências fornece muito mais informações sobre o comportamento de uma variável do que a própria tabela original de dados (dados brutos). Muitas vezes, queremos resumir ainda mais estes dados, apresentando um ou alguns valores que sejam representativos de todos os dados. Quando usamos um só valor, obtemos uma redução drástica dos dados. Usualmente, emprega-se uma das seguintes medidas de posição (ou localização) central: média, mediana ou moda.

3.1 Média aritmética

A média aritmética, ou simplesmente média do conjunto de dados, é obtida somando-se todos os dados e dividindo-se o resultado da soma pelo número deles.

$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n} = \frac{1}{n}\sum_{i=1}^{n} x_i.$$

Podemos, por simplicidade, escrever

$$\bar{x} = \frac{\sum x_i}{n}$$

que lê-se "x barra" é igual ao somatório de x, dividido por n. Quando omitirmos o índice significa que ele está variando de 1 à n, ou seja, $\sum x_i = \sum_{i=1}^{n} x_i$.

Exemplo

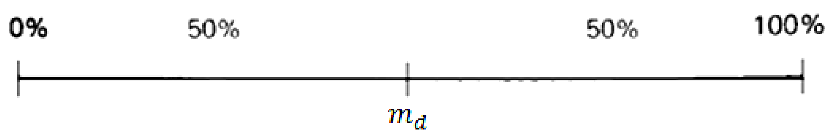
Se as cinco observações de uma variável forem 3, 4, 7, 8 e 8, então:

$$\bar{x} = \frac{1}{5}\sum_{i=1}^{5} x_i = \frac{3+4+7+8+8}{5} = \frac{30}{5} = 6.$$

3.2 Mediana

A mediana (m_d) é o valor que ocupa a posição central do conjunto de dados ordenados. A mediana divide o conjunto de dados em duas partes:

- uma com números menores ou iguais à mediana;
- outra com números maiores ou iguais à mediana.



Número ímpar de dados $(\boldsymbol{n}\ \boldsymbol{\acute{\imath}mpar})$

Quando o número de dados é **ímpar**, existe um único valor na posição central. Esse valor é a mediana.

Exemplo

O conjunto de dados

$$\{3, \mathbf{5}, 9\}$$

tem mediana igual a 5 $(m_d = 5)$, pois 5 é o valor central do conjunto de dados ordenados.

Número par de dados $(\boldsymbol{n}\ \boldsymbol{par})$

Quando o número de dados é **par**, existem dois valores na posição central. A mediana é a média desses dois valores.

Exemplo

O conjunto de dados

$$\{3, \mathbf{5}, \mathbf{7}, 9\}$$

tem mediana 6, pois 6 é a média de 5 e 7, que estão na posição central do conjunto de dados ordenados $\left(m_d = \frac{5+7}{2} = 6\right)$.

Fórmula da Mediana

Podemos utilizar a seguinte fórmula para determinar a mediana:

$$m_d = \begin{cases} x_{\left(\frac{n+1}{2}\right)} & \text{se } n \text{ for ímpar} \\ \dfrac{x_{\left(\frac{n}{2}\right)} + x_{\left(\frac{n}{2}+1\right)}}{2} & \text{se } n \text{ for par} \end{cases}$$

Exemplo

Para o conjunto de dados ordenados $\{3,\ 5,\ 9\}$, temos $x_1 = 3,\ \ x_2 = 5$ e $x_3 = 9$. Como $n = 3$ é ímpar, então:

$$m_d = x_{\left(\frac{n+1}{2}\right)} = x_{\left(\frac{3+1}{2}\right)} = x_{\left(\frac{4}{2}\right)} = x_{(2)} = 5.$$

Para o conjunto de dados ordenados $\{3, 5, 7, 9\}$, temos $x_1 = 3,\ \ x_2 = 5,\ \ x_3 = 7$ e $x_4 = 9$. Como $n = 4$ é par, então:

$$m_d = \frac{x_{\left(\frac{n}{2}\right)} + x_{\left(\frac{n}{2}+1\right)}}{2} = \frac{x_{\left(\frac{4}{2}\right)} + x_{\left(\frac{4}{2}+1\right)}}{2} = \frac{x_{(2)} + x_{(3)}}{2} = \frac{5+7}{2} = \frac{12}{2} = 6.$$

3.3 Moda

Moda é a realização do conjunto de dados que ocorre com maior frequência.

Exemplo

A moda do conjunto de dados:

$$\{0, 0, 2, 5, 3, \mathbf{7}, 4, \mathbf{7}, 8, \mathbf{7}, 9, 6\}$$

é o número 7, porque este é o valor que ocorre o maior número de vezes.

Um conjunto de dados pode não ter moda ou ter duas ou mais modas. Assim, o conjunto de dados:

$$\{0, 2, 4, 6, 8, 10\}$$

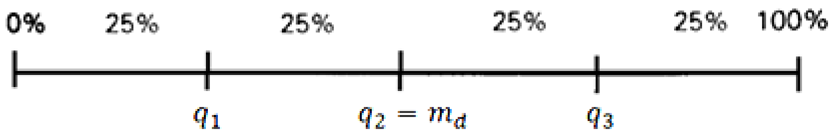
não tem moda, enquanto o conjunto de dados:

$$\{1, \mathbf{2}, \mathbf{2}, 3, \mathbf{4}, \mathbf{4}, 5, 6, 7\}$$

tem duas modas: 2 e 4.

3.4 Quartis

Os quartis (q_1, q_2 e q_3) são medidas que dividem o conjunto de dados em *quatro* partes iguais (ou aproximadamente iguais):



O primeiro quartil (q_1) é um número tal que 25% dos dados são menores ou iguais a ele e 75% dos dados são maiores ou iguais a ele. O segundo quartil $(q_2 = m_d)$ é um número tal que 50% dos dados são menores ou iguais a ele e 50% dos dados são maiores ou iguais a ele. O terceiro quartil (q_3) é um número tal que 75% dos dados são menores ou iguais a ele e 25% dos dados são maiores ou iguais a ele.

Exemplo

Obtendo os quartis de um conjunto com um número ímpar de dados.

Considere o conjunto de dados:

$$\{3, 4, 1, 5, 7, 9, 2, 10, 6\}.$$

Temos que $n = 9$ é **ímpar**. Então, a mediana é o valor central dos dados ordenados, ou seja, $m_d = 5$:

$$1, 2, 3, 4, \mathbf{5}, 6, 7, 9, 10.$$

Para obter o primeiro quartil, q_1, separe os dados **desde o menor dado até a mediana**. O primeiro quartil é a mediana do novo conjunto de dados, ou seja, $q_1 = 3$:

$$1, 2, \mathbf{3}, 4, 5.$$

Para obter o terceiro quartil, q_3, separe os dados **desde a mediana até o maior dado**. O terceiro quartil é a mediana do novo conjunto de dados, ou seja, $q_3 = 7$:

$$5, 6, \mathbf{7}, 9, 10.$$

Exemplo

Obtendo os quartis de um conjunto com um número par de dados.

Considere o conjunto de dados:

$$\{11, 3, 4, 1, 5, 7, 9, 2, 10, 6\}.$$

Temos que $n = 10$ é **par**. Então, a mediana é a média dos dois valores centrais dos dados ordenados, ou seja, $m_d = \frac{5+6}{2} = 5{,}5$:

$$1, 2, 3, 4, 5, (\mathbf{5.5}), 6, 7, 9, 10{,}11.$$

Para obter o primeiro quartil, q_1, separe os dados **abaixo da mediana**. O primeiro quartil é a mediana do novo conjunto de dados, ou seja, $q_1 = 3$:

$$1, 2, \mathbf{3}, 4, 5.$$

Para obter o terceiro quartil, q_3, separe os dados **acima da mediana**. O terceiro quartil é a mediana do novo conjunto de dados, ou seja, $q_3 = 9$:

$$6, 7, \mathbf{9}, 10{,}11.$$

As maneiras de se calcular os quartis apresentadas acima não são as únicas. Na realidade existem várias maneiras diferentes de se calcular quartis. Podemos ver, a seguir, que o R pode obter os quartis de formas diferentes.

4 Medidas de dispersão

As medidas de tendência central resumem a informação contida em um conjunto de dados, mas não contam toda a história. Por causa da variabilidade, a média, a mediana e a moda não são suficientes para descrever um conjunto de dados: informam apenas a tendência central, ou seja, onde está o centro, mas nada dizem sobre a variabilidade.

4.1 Variância

Considere as notas de cinco alunos em uma prova, apresentadas na Tabela 4.1.

Tabela 4.1. Desvios e quadrados dos desvios das notas de cinco alunos

Nota (x_i)	Desvio $(x_i - \bar{x})$	Quadrados dos desvios $(x_i - \bar{x})^2$
3	$(3-5) = -2$	$(3-5)^2 = (-2)^2 = 4$
4	$(4-5) = -1$	$(4-5)^2 = (-1)^2 = 1$
5	$(5-5) = 0$	$(5-5)^2 = (0)^2 = 0$
6	$(6-5) = 1$	$(6-5)^2 = (1)^2 = 1$
7	$(7-5) = 2$	$(7-5)^2 = (2)^2 = 4$
Total	$\sum(x_i - \bar{x}) = 0$	$\sum(x_i - \bar{x})^2 = 10$

A idéia central desta medida de dispersão é obter uma medida que represente a variabilidade ao redor da média, por isso utiliza-se os desvios ao redor da média $(x_i - \bar{x})$. Contudo o somatório dos desvios sempre será nulo. Uma forma de evitar que a soma dos desvios se anule é elevando cada desvio ao quadrado, ou seja, fazendo $(x_i - \bar{x})^2$. A partir dos quadrados dos desvios obtemos à variância, que é a medida de variabilidade mais utilizada. A variância pode ser entendida como se fosse praticamente a "média" dos quadrados de desvios em relação à média. Numa amostra de tamanho n, este valor (n) deveria ser usado como divisor desta soma de quadrados de desvios. No entanto, devido a motivos associados a propriedades dos estimadores, o divisor da variância amostral é dado por $n-1$ em lugar de n na expressão do estimador da variância. Assim, **se os dados são provenientes de uma amostra,** a **variância amostral** será denotada por S^2 e será calculada da seguinte maneira:

$$S^2 = \frac{\sum(x_i - \bar{x})^2}{n-1}.$$

Para as notas dos alunos, a variância será:

$$s^2 = \frac{\sum(x_i - \bar{x})^2}{n-1} = \frac{(3-5)^2 + (4-5)^2 + \cdots + (7-5)^2}{5-1} = \frac{10}{4} = 2{,}5.$$

Fórmula alternativa da variância

Outra forma de se calcular a variância é pela fórmula:

$$S^2 = \frac{1}{n-1}\left[\sum x_i^2 - \frac{(\sum x_i)^2}{n}\right].$$

Esta fórmula é resultado de manipulações algébricas da fórmula anterior. Assim, independente de qual fórmula utilizarmos o resultado da variância será o mesmo. Considerando as notas de cinco alunos, são apresentados na Tabela 4.2 os somatórios utilizados no cálculo da variância.

Tabela 4.2. Tabela de cálculos auxiliares para obtenção de $\sum x_i$ e $\sum x_i^2$

Aluno	**Nota** $(\boldsymbol{x_i})$	$\boldsymbol{x_i^2}$
1	3	$3^2 = 9$
2	4	$4^2 = 16$
3	5	$5^2 = 25$
4	6	$6^2 = 36$
5	7	$7^2 = 49$
Total	$\sum x_i = 25$	$\sum x_i^2 = 135$

Assim, obtidos $\sum x_i = 25$ e $\sum x_i^2 = 135$, basta substituí-los na fórmula da variância:

$$s^2 = \frac{1}{n-1}\left[\sum x_i^2 - \frac{(\sum x_i)^2}{n}\right] = \frac{1}{5-1}\left[135 - \frac{(25)^2}{5}\right] = \frac{1}{4}\left[135 - \frac{625}{5}\right]$$
$$= \frac{1}{4}[135 - 125] = \frac{1}{4}[10] = 2{,}5.$$

4.2 Desvio Padrão

No cálculo da variância, devido ao fato de se elevar os desvios ao quadrado, a unidade de medida da variância também fica elevada ao quadrado, gerando escalas sem sentido prático. Assim, se a unidade de medida dos dados seja metros (m), a unidade de medida da variância será m^2, se a unidade de medida dos dados for kg, a unidade de medida da variância será kg^2, etc.

Uma forma de se obter uma medida de dispersão com a mesma unidade de medida dos dados observados é, simplesmente, extrair a raiz quadrada da variância, obtendo-se o desvio padrão. O desvio padrão será denotado por S e será dado por:

$$S = \sqrt{S^2}.$$

Exemplo

Para dados de pesos (em quilogramas) de 86 indivíduos, obteve-se uma variância $s^2 = 322{,}35\ kg^2$. Assim, o desvio padrão é:

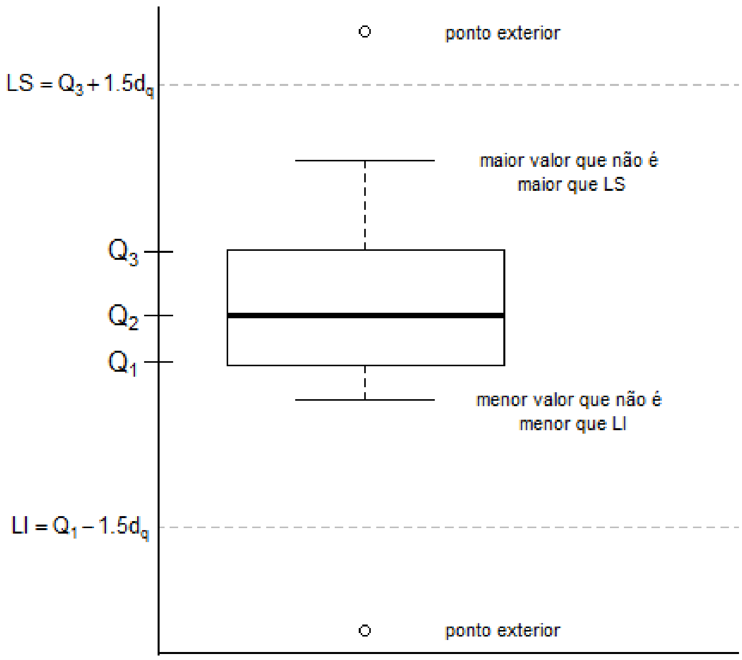
$$s = \sqrt{s^2} = \sqrt{322{,}35} = 17{,}95\ kg.$$

5 Boxplot

As medidas que vimos anteriormente: mínimo, 1º quartil, mediana, 3º quartil e máximo - permitem traçar o diagrama de caixa (*boxplot*), que ajuda a entender a informação contida em um conjunto de dados.

O boxplot (Figura 5.1) dá uma idéia da posição, dispersão, assimetria, caudas e dados discrepantes. A posição central é dada pela mediana e a dispersão por d_q. As posições relativas de Q_1, Q_2 e Q_3 dão uma noção da assimetria da distribuição.

Figura 5.1. Construção do *boxplot*



Boxplot comparativo

Outra utilidade do boxplot é na comparação de diferentes conjuntos de dados, como veremos no exemplo a seguir.

Exemplo

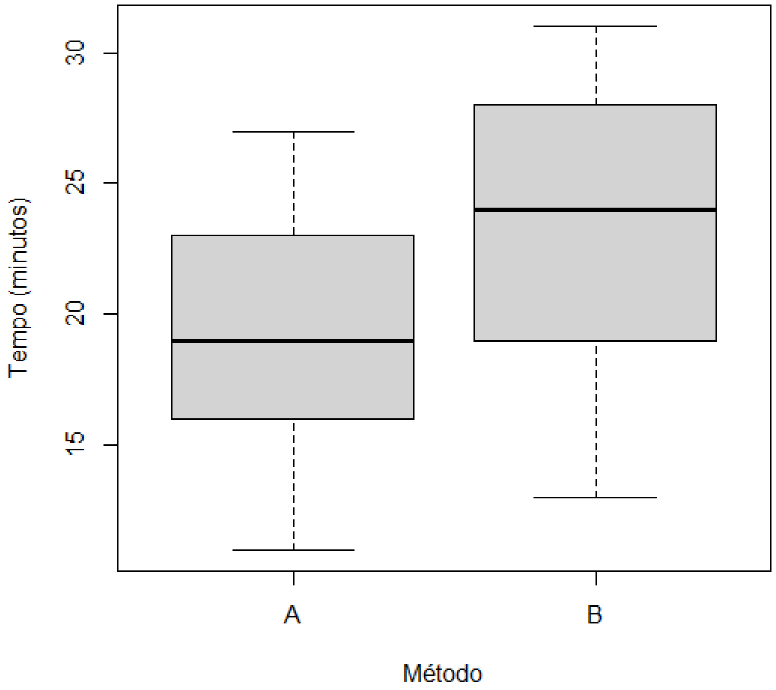
Foi feito um experimento para comparar dois programas de treinamento para a execução de um serviço especializado. Vinte homens foram selecionados para esse treinamento. Dez foram escolhidos ao acaso e treinados pelo método A. Outros dez foram treinados pelo método B. Concluído o período de treinamento, todos os homens executaram o serviço e foi medido o tempo de cada um. Os dados são apresentados na Tabela 5.1.

Tabela 5.1. Tempo (em minutos) despendido na execução do serviço, segundo o método de treinamento

Método A	Método B
15	23
20	31
11	13
23	19
16	23
21	17
18	28
16	26
27	25
24	28

A comparação do tempo de execução de serviço pelos métodos A e B é feita utilizando dois boxplots no mesmo gráfico, que chamaremos de *boxplot comparativo* dos métodos, como é apresentado na Figura 5.2.

Figura 5.2. Boxplot comparativo do tempo de execução em função do método de treinamento



Pode-se observar pelo boxplot comparativo (Figura 5.2) que o tempo de execução do serviço pelo método A é menor do que o tempo pelo método B.

II - Probabilidade

1 - Probabilidade

1.1 Espaço amostral, eventos

Experimento Aleatório

Um experimento aleatório é um experimento que:

a) pode ser repetido indefinidamente sob as mesmas condições;

b) não se conhece um particular valor do experimento "a priori", porém pode-se descrever todos os possíveis resultados - as possibilidades;

c) Quando o experimento for repetido um grande número de vezes surgirá uma regularidade, isto é, haverá uma estabilidade da fração $f = \frac{r}{n}$ (frequência relativa), em que n é o número de repetições e r o número de sucessos de um particular resultado estabelecido antes da realização.

Exemplos

- E_1: Retirar uma carta de um baralho com 52 cartas e observar seu "naipe".
- E_2: Jogar uma moeda 10 vezes e observar o número de caras obtidas.
- E_3: Jogar um dado e observar o número mostrado na face de cima.

Espaço amostral

Definição: Para cada experimento aleatório E, define-se Espaço Amostral Ω o conjunto de todos os possíveis resultados desse experimento.

Exemplos

a) Considere o experimento E: "jogar um dado e observar o nº da face de cima", então

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

b) Seja E: "jogar duas moedas e observar o resultado", então

$$\Omega = \{(c, c), (c, k), (k, c), (k, k)\}$$

em que $c =$ cara e $k =$ coroa.

Observe que sendo Ω um conjunto, poderá ser finito ou infinito. Também pode ser discreto ou contínuo (intervalos). A princípio consideraremos apenas conjuntos finitos.

Evento

Definição: Evento é um conjunto de resultados do experimento, em termos de conjuntos, é um subconjunto de Ω. Em particular, Ω e ∅ (conjunto vazio) são eventos, Ω é dito o evento certo e ∅ o evento impossível.

Exemplos

a) Seja o experimento E: "jogar 3 moedas e observar os resultados". Então

$$\Omega = \{(c, c, c), (c, c, k), (c, k, c), (k, c, c), (c, k, k), (k, c, k), (k, k, c), (k, k, k)\}$$

Seja o evento A: "ocorrer pelo menos 2 caras".

Então, $A = \{(c, c, c), (c, c, k), (c, k, c), (k, c, c)\}$

b) Seja o experimento E: "lançar um dado e observar o número de cima". Então

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

Seja o evento B: "ocorrer múltiplo de 2". Então

$$B = \{2, 4, 6\}$$

Usando as operações com conjuntos, podem-se formar novos eventos. Assim:

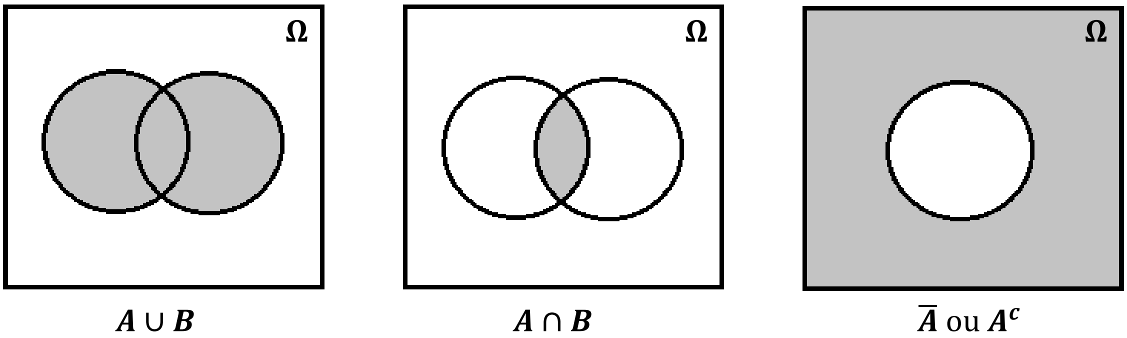
(i) $A \cup B \rightarrow$ é o evento que ocorre se A ou B ocorrem;

(ii) $A \cap B \rightarrow$ é o evento que ocorre se A e B ocorrem;

(iii) $\bar{A}$ ou $A^c \rightarrow$ é o evento que ocorre se A não ocorre ($\bar{A}$ é chamado complementar de A).

Podemos observar na Figura 1.1 os Diagramas de *Veen* dos conjuntos definidos acima.

Figura 1.1. Diagramas de *Veen* dos conjuntos definidos acima



Exemplo

Seja o experimento E: "lançar um dado e observar o número de cima". Então:

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

Seja o evento A: "ocorrer número ímpar". Então:

$$A = \{1, 3, 5\}$$

Seja o evento B: "ocorrer um número menor ou igual a 3". Então:

$$B = \{1, 2, 3\}$$

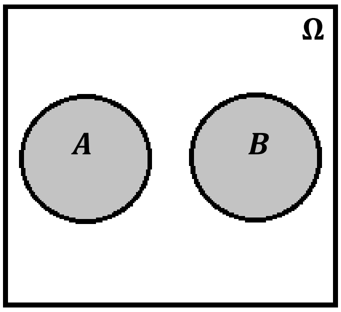
Assim:

- $A \cup B = \{1, 2, 3, 5\}$
- $A \cap B = \{1, 3\}$
- $\bar{A} = \{2, 4, 6\}$
- $\bar{B} = \{4, 5, 6\}$

Eventos Mutuamente Excludentes

Dois eventos A e B são denominados mutuamente excludentes, se eles não puderem ocorrer simultaneamente, isto é, $A \cap B = \emptyset$. Veja a Figura 1.2.

Figura 1.2. Eventos mutuamente excludentes



Exemplo

E: "jogar um dado e observar o resultado"

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

Sejam os eventos: A: "ocorrer nº par" e B: "ocorrer nº ímpar". Então:

$$A = \{2, 4, 6\} \text{ e } B = \{1, 3, 5\}, \text{ assim, } A \cap B = \emptyset.$$

A e B são mutuamente exclusivos, pois a ocorrência de um número par impede a ocorrência de um número ímpar e vice-versa.

Resultados adicionais envolvendo conjuntos

1. $(A^c)^c = A$
2. Regras distributivas:

   $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$ e $(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$
3. Leis de Morgan:

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

Definição (axiomática) de Probabilidade

Dado um experimento aleatório E e Ω o espaço amostral, probabilidade de um evento A, denotada por $P(A)$, é uma função definida em Ω que associa a cada evento um número real, satisfazendo os seguintes axiomas:

(i) $0 \leq P(A) \leq 1$;

(ii) $P(\Omega) = 1$;

(iii) Se A e B forem eventos mutuamente excludentes, $(A \cap B = \emptyset)$, então $P(A \cup B) = P(A) + P(B)$.

Principais Teoremas

1. Se $\emptyset$ é o conjunto vazio, então $P(\emptyset) = 0$.

Demonstração

Podemos escrever $\mathrm{A} = A \cup \emptyset$, logo $P(\mathrm{A}) = P(A \cup \emptyset)$. Como A e $\emptyset$ são mutuamente excludentes, decorre do axioma (iii) que

$$P(\mathrm{A}) = P(A \cup \emptyset) = P(A) + P(\emptyset).$$

Portanto $P(\emptyset) = 0$.

2. Se $\bar{A}$ é o complemento do evento A, então $P(\bar{A}) = 1 - P(A)$.

Demonstração

Podemos escrever $\Omega = A \cup \bar{A}$ e, pelos axiomas (ii) e (iii), obtemos

$$P(\Omega) = P(A \cup \bar{A}) = P(A) + P(\bar{A}) = 1$$

e, portanto,

$$P(\bar{A}) = 1 - P(A)$$

3. Se $A \subset B$, então $P(A) \leq P(B)$.

Demonstração

Podemos decompor B em dois eventos mutuamente excludentes, da seguinte forma: $B = A \cup (B \cap \bar{A})$. Consequentemente,

$$P(B) = P(A) + P(B \cap \bar{A}) \geq P(A)$$

Pois, pelo axioma (i) $P(B \cap \bar{A}) \geq 0$.

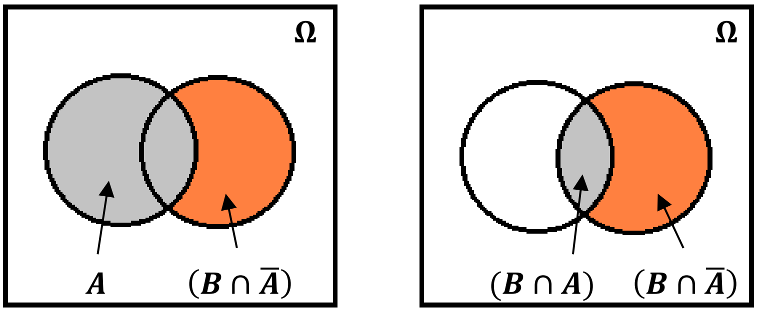
4. Teorema da soma: Se A e B são dois eventos quaisquer, então:

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

Demonstração

Primeiramente vamos decompor $A \cup B$ e B em eventos mutuamente excludentes:

$$A \cup B = A \cup (B \cap \bar{A}) \quad \text{e} \quad B = (B \cap A) \cup (B \cap \bar{A}).$$



Assim,

$$P(A \cup B) = P(A) + P(B \cap \bar{A})$$
$$P(B) = P(B \cap A) + P(B \cap \bar{A})$$

Subtraindo a segunda igualdade da primeira, obtemos

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

Portanto,

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

5. Se A, B e C são três eventos quaisquer, então:

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

Demonstração

Primeiramente vamos reescrever o evento $A \cup B \cup C$ como $(A \cup B) \cup C$. Assim,

$$P(A \cup B \cup C) = P[(A \cup B) \cup C]$$

e, pelo Teorema 4, temos:

$$
\begin{aligned}
P[(A \cup B) \cup C] &= P(A \cup B) + P(C) - P[(A \cup B) \cap C] \\
&= P(A) + P(B) - P(A \cap B) + P(C) - P[(A \cup B) \cap C] \\
&= P(A) + P(B) + P(C) - P(A \cap B) - P[(A \cup B) \cap C] \quad (*)
\end{aligned}
$$

$(*)$ Pela regra distributiva, temos

$$P[(A \cup B) \cap C] = P[(A \cap C) \cup (B \cap C)]$$

Aplicando o Teorema 4, temos:

$$
\begin{aligned}
&P[(A \cap C) \cup (B \cap C)] \\
&\quad = P(A \cap C) + P(B \cap C) - P[(A \cap C) \cap (B \cap C)] \\
&\quad = P(A \cap C) + P(B \cap C) - P[(A \cap B) \cap C] \\
&\quad = P(A \cap C) + P(B \cap C) - P(A \cap B \cap C)
\end{aligned}
$$

Portanto,

$$
\begin{aligned}
P[(A \cup B) \cup C] &= P(A) + P(B) + P(C) - P(A \cap B) - P[(A \cup B) \cap C] \\
&= P(A) + P(B) + P(C) - P(A \cap B) \\
&\quad -[P(A \cap C) + P(B \cap C) - P(A \cap B \cap C)] \\
&= P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) \\
&\quad +P(A \cap B \cap C)
\end{aligned}
$$

6. Uma coleção de eventos $E_1, E_2, \dots, E_k$ é dita mutuamente excludente se para todos os pares $E_i \cap E_j = \emptyset$. Para uma coleção de eventos mutuamente excludentes:

$$P(E_1 \cup E_2 \cup \dots \cup E_k) = P(E_1) + P(E_2) + \cdots + P(E_k).$$

Demonstração

Se E_1 e E_2 são mutuamente excludentes, então, pelo axioma (iii) temos:

$$P(E_1 \cup E_2) = P(E_1) + P(E_2).$$

Suponha, agora, que para $(k-1)$ eventos mutuamente excludentes seja verdadeira a igualdade:

$$P(E_1 \cup E_2 \cup \dots \cup E_{k-1}) = P(E_1) + P(E_2) + \cdots + P(E_{k-1}).$$

Então, para k eventos mutuamente excludentes, podemos escrever:

$$P(E_1 \cup E_2 \cup \ldots \cup E_{k-1} \cup E_k) = P[(E_1 \cup E_2 \cup \ldots \cup E_{k-1}) \cup E_k].$$

Como $E_1, E_2, \ldots, E_k$ são mutuamente excludentes (não tem elementos em comum), então os eventos $(E_1 \cup E_2 \cup \ldots \cup E_{k-1})$ e E_k também serão excludentes. Pelo axioma (iii), temos

$$P[(E_1 \cup E_2 \cup \ldots \cup E_{k-1}) \cup E_k] = P(E_1 \cup E_2 \cup \ldots \cup E_{k-1}) + P(E_k).$$

Como assumimos verdadeira a igualdade

$$P(E_1 \cup E_2 \cup \ldots \cup E_{k-1}) = P(E_1) + P(E_2) + \cdots + P(E_{k-1})$$

então:

$$P(E_1 \cup E_2 \cup \ldots \cup E_{k-1}) + P(E_k) = P(E_1) + P(E_2) + \cdots + P(E_{k-1}) + P(E_k)$$

e, portanto:

$$P(E_1 \cup E_2 \cup \ldots \cup E_{k-1} \cup E_k) = P(E_1) + P(E_2) + \cdots + P(E_{k-1}) + P(E_k)$$

para qualquer k inteiro, $K \geq 2$.

Espaços Amostrais Finitos Equiprováveis

Quando se associa a cada elemento do espaço amostral a mesma probabilidade, o espaço amostral chama-se equiprovável ou uniforme. Em particular se Ω contém n elementos, então, a probabilidade de cada ponto será $1/n$.

Por outro lado, se um evento A contém r pontos, então:

$$P(A) = r.\left(\frac{1}{n}\right) = \frac{r}{n}.$$

Este método de avaliar $P(A)$ é frequentemente enunciado da seguinte maneira:

$$P(A) = \frac{\text{nº de elementos do evento A}}{\text{nº de elementos do espaço amostral } \Omega}$$

ou, ainda,

$$\mathrm{P(A)} = \frac{\text{nº de casos favoráveis ao evento A}}{\text{nº total de casos possíveis}}.$$

Exemplo

Escolha aleatoriamente uma carta de um baralho com 52 cartas.

Sejam:

A: "a carta é de ouros"

B: "a carta é uma figura"

Calcular $P(A)$ e $P(B)$.

$$P(A) = \frac{\text{nº de cartas de ouros}}{\text{nº total de cartas}} = \frac{13}{52} = 0{,}25.$$

$$P(B) = \frac{\text{nº de figuras}}{\text{nº total de cartas}} = \frac{12}{52} = 0{,}2308.$$

Como se observa, o cálculo da probabilidade de um evento se reduz a um problema de contagem. Assim, a Análise Combinatória (Teoria da Contagem) tem fundamental importância para se contar o número de casos favoráveis e o total de casos.

1.2 Técnicas de Contagem

Regra da Multiplicação (para técnicas de contagem)

Considere uma operação que possa ser descrita como uma sequência de k etapas e:

- O número de maneiras de completar a etapa 1 for n_1 e;
- O número de maneiras de completar a etapa 2 for n_2 para cada maneira de completar a etapa 1 e;
- O número de maneiras de completar a etapa 3 for n_3 para cada maneira de completar a etapa 2 e assim por diante.

O número total de maneiras de completar a operação será

$$n_1 . n_2 \ldots n_k$$

Exemplo

Uma peça manufaturada deve passar por três estações de controle. Em cada estação, a peça é inspecionada para determinada característica e marcada adequadamente. Na primeira estação três classificações são possíveis, enquanto em cada uma das duas últimas estações quatro classificações são possíveis. Consequentemente, existem

$$3 \times 4 \times 4 = 48$$

maneiras pelas quais uma peça pode ser marcada.

Permutação

Uma permutação dos elementos é uma sequência ordenada dos elementos. O número de permutações de n elementos diferentes é dado por $n!$ (fatorial de n), sendo

$$n! = n\,.\,(n-1)\,.\,(n-2)\dots 2\,.\,1$$

Definimos $0! = 1$.

Exemplo

Considere o conjunto $\{A, B, C\}$, então:

$$ABC, \qquad ACB, \qquad BAC, \qquad BCA, \qquad CAB, \qquad CBA$$

são todas as permutações dos elementos desse conjunto. Note que o número de permutações é calculado por

$$P_3 = 3! = 3\,.\,2\,.\,1 = 6.$$

Permutações de subconjuntos (Arranjos)

O número de permutações de subconjuntos de r elementos selecionados de um conjunto de n elementos diferentes é:

$$A_{n,r} = n\,.\,(n-1)\,.\,(n-2)\dots(n-r+1) = \frac{n!}{(n-r)!}$$

Exemplo

Uma placa de circuito impresso tem 8 localizações diferentes em que um componente pode ser colocado. Se 4 componentes diferentes forem colocados na placa, quantos projetos diferentes são possíveis?

Solução

Cada projeto consiste em selecionar uma localização das 8 localizações disponíveis para o primeiro componente, uma localização das 7 restantes para o segundo componente, uma localização das 6 restantes para o terceiro componente e uma localização das 5 restantes para o quarto componente. Portanto,

$$A_{8,4} = 8\,.\,7\,.\,6\,.\,5 = \frac{8!}{4!} = 1680$$

projetos diferentes são possíveis.

Exemplo

Qual a probabilidade de que em um grupo de n pessoas $(n \leq 365)$ haja pelo menos uma coincidência de aniversários?

Solução

Considere o espaço amostral do experimento, ou seja, as possíveis datas de aniversários de n pessoas:

$$\begin{array}{cccccc} & \text{Pessoa 1} & \text{Pessoa 2} & \cdots & \text{Pessoa } n & \\ \Omega = & \left\{ \begin{array}{c} 1 \\ 2 \\ \vdots \\ 365 \end{array} \right. & \begin{array}{c} 1 \\ 2 \\ \vdots \\ 365 \end{array} & \begin{array}{c} \cdots \\ \cdots \\ \ddots \\ \cdots \end{array} & \left. \begin{array}{c} 1 \\ 2 \\ \vdots \\ 365 \end{array} \right\} & \end{array}$$

Observe que: (número de elementos de Ω) $= 365^n$.

O evento de interesse é A: "existe pelo menos uma coincidência de aniversários".

$$P(A) = P(1\ coincid\hat{e}ncia) + P(2\ coincid\hat{e}ncias) + \cdots + P(n\ coincid\hat{e}ncias)$$

É mais fácil calcular a probabilidade pelo complementar:

$$P(A) = 1 - P(\bar{A})$$

em que $\bar{A}$: "não existe nenhuma coincidência de datas de aniversário entre as n pessoas". Para obter o número de elementos em $\bar{A}$ observe que:

Pessoa 1	Pessoa 2	Pessoa 3	$\cdots$	Pessoa n
1	1	1	$\cdots$	1
2	2	2	$\cdots$	2
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
365	365 menos 1 data já ocupada pela pessoa 1, ou seja, 364	365 menos 2 datas já ocupadas pelas pessoas 1 e 2, ou seja, 363	$\cdots$	365 menos $(n-1)$ datas já ocupadas pelas $(n-1)$ pessoas anteriores

Assim, (número de elementos de $\bar{A}$) $= 365\,.\,364\,.\,363 \ldots \big(365-(n-1)\big)$

$$= 365\,.\,364\,.\,363 \ldots (365-n+1)$$

$$= \frac{365!}{(365-n)!}$$

$$= A_{365,n}$$

Portanto:

$$P(A) = 1 - P(\bar{A})$$

$$= 1 - \frac{A_{365,n}}{365^n}$$

Observação: Para valores grandes de n o cálculo acima se torna computacionalmente inviável ($A_{365,n}$ e 365^n geram valores extremamente grandes que calculadoras e alguns computadores não conseguem processar). Para contornar esse problema podemos reescrever o resultado como:

$$P(A) = 1 - \frac{365}{365}\,.\frac{364}{365}\,.\frac{363}{365}\ldots\frac{(365-n+1)}{365}$$

$$= 1 - \frac{364}{365}.\frac{363}{365}\ldots\frac{(365-n+1)}{365}$$

Exemplo

Qual a probabilidade de que em um grupo de $n = 10$ pessoas haja pelo menos uma coincidência de aniversários?

Solução

Como o valor de n é pequeno podemos usar o primeiro resultado encontrado no exemplo anterior:

$$P(A) = 1 - P(\bar{A})$$

$$= 1 - \frac{A_{365,10}}{365^{10}}$$

$$= 1 - 0{,}8831$$

$$= 0{,}1169$$

Permutações de Objetos Similares

O número de permutações de $n = n_1 + n_2 + \cdots + n_r$ objetos dos quais n_1 são de um tipo, n_2 são de um segundo tipo, ..., e n_r são de um r-ésimo tipo é

$$\frac{n!}{n_1!\,.\,n_2!\,.\,n_3!\,\ldots n_r!}$$

Exemplo

Um item é codificado pela impressão de 4 linhas espessas, 3 linhas médias e 2 linhas finas. Se cada ordenação das 9 linhas representa um código diferente, quantos códigos diferentes podem ser gerados pelo uso desse esquema?

Solução

O número de códigos possíveis é

$$\frac{9!}{4!\,.\,3!\,.\,2!} = 1260.$$

Combinação

Outro problema de interesse é o número de subconjuntos de r elementos que podem ser selecionados a partir de um conjunto de n elementos. Aqui a ordem não é importante. Esses subconjuntos são chamados de combinações n elementos tomados r à r.

Definição

O número de **combinações**, subconjuntos de tamanho r, que podem ser selecionados a partir de um conjunto de n elementos, é denotado como $\binom{n}{r}$ ou $C_{n,r}$ e é calculado por:

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

Exemplo

Um componente pode ser colocado em 8 localizações diferentes em uma placa de circuito impresso. Se 5 componentes idênticos forem colocados na placa, quantos projetos diferentes são possíveis?

Solução

O número de projetos possíveis é

$$\binom{8}{5} = \frac{8!}{5!\,(8-5)!} = \frac{8!}{5!\,.\,3!} = 56$$

Exemplo

Num lote de 12 peças, 4 são defeituosas. Duas peças são retiradas aleatoriamente (sem reposição). Determine:

a) A probabilidade de ambas serem defeituosas;

b) A probabilidade de ambas serem boas;

c) A probabilidade de ao menos uma ser defeituosa.

Solução

a) A: "ambas são defeituosas".

$$A \text{ contém } \binom{4}{2} = \frac{4!}{2!(4-2)!} = \frac{4.3.2!}{2\,.2!} = 6 \text{ elementos.}$$

$$\Omega \text{ contém } \binom{12}{2} = \frac{12!}{2!(12-2)!} = \frac{12.11.10!}{2\,.10!} = 66 \text{ elementos.}$$

Logo, $P(A) = \frac{\text{nº de casos favoráveis}}{\text{nº total de casos}} = \frac{6}{66} = 0{,}0909.$

b) B: "ambas são boas".

$$B \text{ contém } \binom{8}{2} = \frac{8!}{2!(8-2)!} = \frac{8.7.6!}{2\,.6!} = 28 \text{ elementos.}$$

Logo, $P(B) = \frac{\text{nº de casos favoráveis}}{\text{nº total de casos}} = \frac{28}{66} = 0{,}4242.$

c) C: "ao menos uma é defeituosa".

1ª forma (usando o evento complementar)

Observe que C (uma é defeituosa ou as duas são defeituosas) é o complemento de B (as duas são boas), ou seja, $C = \bar{B}$.

Possibilidades:

$$\left\{\underbrace{\text{(duas boas)}}_{\text{B}}, \underbrace{\text{(uma defeituosa e uma boa)}, \text{(duas defeituosas)}}_{\text{C}}\right\}$$

Logo, $C = \bar{B}$.

Portanto:

$$P(C) = 1 - P(B) = 1 - 0{,}4242 = 0{,}5758.$$

2ª forma (usando a regra da multiplicação para técnicas de contagem)

O evento C pode ser reescrito como a união de dois eventos mutuamente excludentes:

$$C = C_1 \cup A$$

em que:

C_1: "apenas uma peça é defeituosa" e A: "as duas peças são defeituosas".

Já vimos que $P(A) = 0{,}0909$, resta então calcular $P(C_1)$. Pela regra da multiplicação, temos:

- nº de elementos de $C_1 = 4\ (\text{defeituosas}) \times\ 8\ (\text{boas}) = 32$;
- nº de elementos de $\Omega = 66$.

Portanto,

$$P(C_1) = \frac{32}{66} = 0{,}4848$$

Como C_1 e A são mutuamente excludentes, então:

$$P(C) = P(C_1) + P(A) = 0{,}4848 + 0{,}0909 = \mathbf{0,5757} \quad \text{(arredondamento)}$$

1.3 Probabilidade condicional, Teorema de Bayes e independência de eventos

Probabilidade Condicional

Exemplo

Seja E: "lançar um dado", e o evento A: "sair o nº 3". Então $\Omega = \{1, 2, 3, 4, 5, 6\}$, $A = \{3\}$ e, portanto:

$$P(A) = \frac{1}{6}.$$

Considere agora o evento B: "sair um nº ímpar", ou seja, $B = \{1, 3, 5\}$. Pode-se estar interessado em avaliar a probabilidade do evento A condicionada à ocorrência do evento B. Em símbolos, designa-se por $P(A|B)$; lê-se "probabilidade do evento A condicionada à ocorrência de B", ou ainda, "probabilidade de A dado B".

Dada a informação da ocorrência do evento B (ocorreu um nº ímpar), temos a redução do espaço amostral $\Omega = \{1, 2, 3, 4, 5, 6\}$ para $\Omega^* = \{1, 3, 5\}$ (pois já sabemos que o nº que ocorreu é ímpar), e é nesse espaço amostral reduzido que se avalia a probabilidade do evento.

- $A = \{3\} \quad \Rightarrow \quad \text{nº de elementos de } A = 1;$
- $\Omega^* = \{1, 3, 5\} \quad \Rightarrow \quad \text{nº de elementos de } \Omega^* = 3;$

Portanto:

$$P(A|B) = \frac{1}{3} = 0{,}3333.$$

Podemos, também, calcular a probabilidade condicional de A dado B sem ter que considerar o espaço amostral reduzido.

Definição: Dados dois eventos A e B, denota-se $P(A|B)$ a probabilidade condicional do evento A, quando B tiver ocorrido, por:

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

com $P(B) \neq 0$, pois B já ocorreu.

Exemplo

Seja E: "lançar um dado", então $\Omega = \{1,2,3,4,5,6\}$. Sejam:

- A: "sair o nº 3" $= \{3\} \;\; \Rightarrow \;\; P(A) = \frac{1}{6}$
- B: "sair um nº ímpar" $= \{1, 3, 5\} \;\; \Rightarrow \;\; P(B) = \frac{3}{6}$.

Então, $A \cap B = \{3\} \;\; \Rightarrow \;\; P(A \cap B) = \frac{1}{6}$.

Assim, podemos calcular $P(A|B)$ como:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{1/6}{3/6} = \frac{1}{3} = 0{,}3333.$$

Podemos, também, calcular a probabilidade de B dado A como:

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

Teorema do Produto

A partir da definição de probabilidade condicional pode-se enunciar o teorema do produto:

"A probabilidade da ocorrência simultânea de dois eventos, A e B, do mesmo espaço amostral, é igual ao produto da probabilidade de um deles pela probabilidade condicional do outro, dado o primeiro."

Assim:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad \Rightarrow \quad P(A \cap B) = P(B).P(A|B)$$

ou

$$P(B|A) = \frac{P(A \cap B)}{P(A)} \quad \Rightarrow \quad P(A \cap B) = P(A).P(B|A)$$

Usamos o Teorema do Produto quando já conhecemos a probabilidade condicional (quando não precisamos calculá-la usando fórmula) e queremos encontrar a probabilidade da interseção de A e B, como será visto no exemplo a seguir.

Exemplo

Em um lote de 12 peças, 4 são defeituosas. Duas são retiradas uma após a outra sem reposição. Qual a probabilidade de que ambas sejam boas?

D_1	D_2	D_3	D_4	B_1	B_2	B_3	B_4	B_5	B_6	B_7	B_8

Solução

- A: "a primeira peça é boa"
- B: "a segunda peça é boa"

$$P(A \cap B) = P(A).P(B|A) = \frac{8}{12}.\frac{7}{11} = \frac{56}{132} = 0{,}4242.$$

Regra Geral do Produto

Seja $A_1, A_2, \ldots, A_n$ uma sequência de eventos de um espaço amostral Ω. Então

$$P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1).P(A_2|A_1) \cdots P(A_n|A_1 \cap A_2 \cap \cdots \cap A_{n-1})$$

Exemplo

Suponhamos uma extração de 3 cartas sem reposição, de um baralho, onde estamos interessados no evento "nenhuma copas".

Seja C = {carta de copas}. Logo $P(C) = 13/52$ e seu complementar será $P(\overline{C}) = 39/52$. Então:

$$P(\bar{C}_1 \cap \bar{C}_2 \cap \bar{C}_3) = P(\bar{C}_1).P(\bar{C}_2|\bar{C}_1).P(\bar{C}_3|\bar{C}_1 \cap \bar{C}_2) = \frac{39}{52}.\frac{38}{51}.\frac{37}{50} = 0{,}4135$$

Independência Estatística

Um evento A é considerado independente de um outro evento B se a probabilidade de A é igual à probabilidade condicional de A dado B, isto é, se

$$P(A) = P(A|B)$$

ou seja, A é independente de B se o fato de B ter ocorrido não afeta em nada a probabilidade da ocorrência de A.

É evidente que, se A é independente de B, B é independente de A; assim:

$$P(B) = P(B|A).$$

Considerando o teorema do produto, pode-se afirmar que: se A e B são independentes, então:

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

Exemplo

Em uma caixa de 10 peças, 4 são defeituosas. São retiradas duas peças, uma após a outra, com reposição. Calcular a probabilidade de ambas serem boas.

D_1	D_2	D_3	D_4	B_1	B_2	B_3	B_4	B_5	B_6

Solução

Sejam os eventos:

A: "a primeira peça é boa",

B: "a segunda peça é boa".

Note que A e B são independentes, pois, $P(B) = P(B|A)$. Ou seja, o fato de a 1ª peça ter sido boa não alterou a probabilidade de a 2ª peça ser boa (pois a 1ª peça foi colocada de volta na caixa). Logo:

$$P(A \cap B) = P(A).P(B) = \frac{6}{10}.\frac{6}{10} = \frac{36}{100} = 0{,}36.$$

Regra Geral da Multiplicação (Eventos Independentes)

Seja $A_1, A_2, \ldots, A_n$ uma sequência de eventos independentes de um espaço amostral Ω. Então

$$P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1).P(A_2) \cdots P(A_n)$$

Exemplo

As probabilidades de 3 jogadores marcarem um penalty são respectivamente $2/3$, $4/5$ e $7/10$. Se cada um "cobrar" uma única vez, qual a probabilidade de todos acertarem?

Solução

Os eventos são A_1: {o 1º jogador acertar}, A_2: {o 2º jogador acertar} e A_3: {o 3º jogador acertar}. Observe que os eventos são independentes (a ocorrência de um evento não altera a probabilidade de ocorrência do outro), então:

$$P(A_1 \cap A_2 \cap A_3) = P(A_1).P(A_2).P(A_3) = \frac{2}{3}.\frac{4}{5}.\frac{7}{10} = 0{,}3733$$

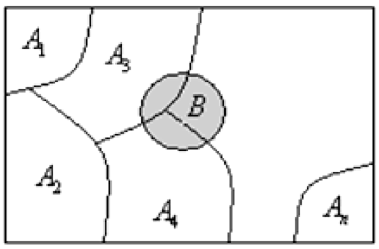
Teorema da Probabilidade Total

Sejam A_1, A_2, ..., A_n, n eventos mutuamente exclusivos tais que

$$A_1 \cup A_2 \cup \ldots \cup A_n = \Omega.$$

Sejam $P(A_i)$ as probabilidades conhecidas dos vários eventos, e B um evento qualquer de Ω tal que são conhecidas todas as probabilidades condicionais $P(B|A_i)$. Então

$$P(B) = P(A_1).P(B|A_1) + P(A_2).P(B|A_2) + \cdots + P(A_n).P(B|A_n).$$



Exemplo

Considere um lote contendo 100 peças, das quais 20 são defeituosas e 80 são não-defeituosas. São extraídas duas peças, sem reposição. Definindo-se:

A: {a primeira peça extraída é defeituosa},

B: {a segunda peça extraída é defeituosa},

determine $P(B)$.

Solução

Na primeira extração pode ocorrer apenas um dos seguintes eventos:

A_1: {a primeira peça extraída é defeituosa}, ou

A_2: {a primeira peça extraída é não-defeituosa},

então, a probabilidade de a segunda peça extraída ser defeituosa é calculada por:

$$\begin{aligned} P(B) &= P(A_1).P(B|A_1) + P(A_2).P(B|A_2) \\ &= \frac{20}{100}.\frac{19}{99} + \frac{80}{100}.\frac{20}{99} \\ &= 0{,}2 \end{aligned}$$

Teorema de Bayes

Sejam A_1, A_2, ..., A_n, n eventos mutuamente exclusivos tais que

$$A_1 \cup A_2 \cup \ldots \cup A_n = \Omega.$$

Sejam $P(A_i)$ as probabilidades conhecidas dos vários eventos, e B um evento qualquer de Ω tal que são conhecidas todas as probabilidades condicionais $P(B|A_i)$. Então, para cada índice "i", tem-se:

$$P(A_i|B) = \frac{P(A_i).P(B|A_i)}{P(A_1).P(B|A_1) + P(A_2).P(B|A_2) + \cdots + P(A_n).P(B|A_n)}$$

Exemplo

Considere três urnas, cada uma contendo bolas pretas, brancas e vermelhas. Admita a seguinte configuração:

Urnas / Cores	u_1	u_2	u_3
Pretas	3	4	2
Brancas	1	3	3
Vermelhas	5	2	3

Escolheu-se uma urna ao acaso e dela extraiu-se uma bola ao acaso, verificando-se que a bola é branca. Qual a probabilidade da bola ter vindo da urna 2?

Solução

- Como temos três urnas então a probabilidade de se escolher uma urna ao acaso é de $1/3$.

- Dado que a urna escolhida foi u_1, então a probabilidade de se escolher uma bola branca é de $1/9$.

Urnas / Cores	$P(u_1) = 1/3$	u_2	u_3
Pretas	$P(p\|u_1) = 3/9$	4	2
Brancas	$\boldsymbol{P(b\|u_1) = 1/9}$	3	3
Vermelhas	$P(v\|u_1) = 5/9$	2	3

- Dado que a urna escolhida foi u_2, então a probabilidade de se escolher uma bola branca é de $3/9$.

Urnas / Cores	u_1	$P(u_2) = 1/3$	u_3
Pretas	3	$P(p\|u_2) = 4/9$	2
Brancas	1	$\boldsymbol{P(b\|u_2) = 3/9}$	3
Vermelhas	5	$P(v\|u_2) = 2/9$	3

- Dado que a urna escolhida foi u_3, então a probabilidade de se escolher uma bola branca é de $3/8$.

Urnas / Cores	u_1	u_2	$P(u_3) = 1/3$
Pretas	3	4	$P(p\|u_3) = 2/8$
Brancas	1	3	$\boldsymbol{P(b\|u_3) = 3/8}$
Vermelhas	5	2	$P(v\|u_3) = 3/8$

Assim:

$$P(u_1) = \tfrac{1}{3}, \quad P(u_2) = \tfrac{1}{3}, \quad P(u_3) = \tfrac{1}{3}.$$

$$P(b|u_1) = \tfrac{1}{9}, \quad P(b|u_2) = \tfrac{3}{9}, \quad P(b|u_3) = \tfrac{3}{8}$$

deseja-se calcular $P(u_2|b)$.

Aplicando-se o teorema de Bayes, tem-se:

$$P(A_i|B) = \frac{P(A_i).P(B|A_i)}{P(A_1).P(B|A_1) + P(A_2).P(B|A_2) + \cdots + P(A_n).P(B|A_n)}$$

$$P(u_2|br) = \frac{P(u_2).P(br|u_2)}{P(u_1).P(br|u_1) + P(u_2).P(br|u_2) + P(u_3).P(br|u_3)}$$

$$= \frac{\frac{1}{3}\cdot\frac{1}{3}}{\frac{1}{3}\cdot\frac{1}{9} + \frac{1}{3}\cdot\frac{1}{3} + \frac{1}{3}\cdot\frac{3}{8}} = 0{,}4068.$$

Observe que a probabilidade a priori de u_2 era de $1/3$. Dada a informação que saiu uma bola branca, a probabilidade a posteriori de u_2 será $0{,}4068$.

Exemplo

Uma doença ataca 3% de uma população. Um determinado teste rápido de sangue consegue identificar corretamente 98% das pessoas que possuem a doença. Contudo, o teste identifica como positivo para a doença 8% das pessoas que na realidade não possuem a doença (falsos positivos). Qual é a probabilidade de um indivíduo que foi classificado como positivo no teste ter efetivamente a doença?

Solução

Sejam:

- I: Infectado (o indivíduo possui a doença);
- S: Saudável (i indivíduo não possui a doença);
- CP: Classificado como positivo no teste;
- CN: Classificado como negativo no teste.

Desejamos encontrar a probabilidade de um indivíduo que, foi classificado como positivo no teste, realmente ter a doença, ou seja, dado que o indivíduo foi classificado como positivo, qual a probabilidade dele realmente ter a doença? Podemos escrever isso como:

$$P(I|CP) = ?$$

Pelo Teorema de Bayes temos que:

$$P(A_i|B) = \frac{P(A_i).P(B|A_i)}{P(A_1).P(B|A_1) + P(A_2).P(B|A_2)}$$

$$P(I|CP) = \frac{P(I).P(CP|I)}{P(I).P(CP|I) + P(S).P(CP|S)}$$

O enunciado nos dá as seguintes informações:

- $P(I) = 0{,}03$ (a doença ataca 3% da população)
- $P(S) = 0{,}97$ (logo, a doença não ataca 97% da população)
- $P(CP|I) = 0{,}98$ (o teste identifica como infectados 98% das pessoas doentes)
- $P(CP|S) = 0{,}08$ (o teste identifica como infectados 8% das pessoas saudáveis)

Portanto:

$$P(I|CP) = \frac{P(I).P(CP|I)}{P(I).P(CP|I) + P(S).P(CP|S)}$$

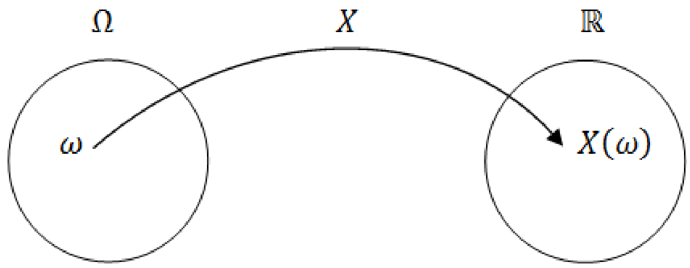
$$= \frac{0{,}03\,.\,0{,}98}{0{,}03\,.\,0{,}98 + 0{,}97\,.\,0{,}08}$$

$$= 0{,}2747.$$

2. Variáveis aleatórias

2.1 Conceito. Valor esperado e variância de uma variável aleatória

Definição: Sejam E um experimento e Ω o espaço amostral associado ao experimento. Uma função X, que associe a cada elemento $\omega \in \Omega$ um número real $X(\omega)$ é denominada variável aleatória. Veja a ilustração.



Exemplo

Sejam: E: "lançamento de duas moedas";

X: "nº de caras obtidas nas duas moedas".

Então: $\Omega = \{(c,c),(c,k),(k,c),(k,k)\}$, em que: $c =$ cara e $k =$ coroa.

- $X = 0 \rightarrow$ corresponde ao elemento: (k,k)
- $X = 1 \rightarrow$ corresponde aos elementos: (c,k) e (k,c)
- $X = 2 \rightarrow$ corresponde ao elemento: (c,c)

Observações

1. "Variável Aleatória" é uma função cujo domínio é Ω e o contradomínio é $\mathbb{R}$;
2. Nas aplicações é conveniente trabalhar com números e não com eventos, daí o uso da variável aleatória;
3. Se Ω é numérico, então $X(\omega) = \omega$;

4. Uma variável aleatória X será discreta se o número de valores possíveis de X (seu contradomínio) for finito ou infinito numerável. Caso seu contradomínio seja um intervalo ou uma coleção de intervalos, ela será uma variável aleatória contínua.

Função de Probabilidades

Seja X uma variável aleatória discreta. Portanto, o contradomínio de X será formado no máximo por um número infinito numerável de valores $x_1, x_2, \ldots$ A cada possível resultado x_i associaremos um número $p(x_i) = P(X = x_i)$, denominado probabilidade de x_i. Os números $p(x_i)$, $i = 1, 2, \ldots$ devem satisfazer às seguintes condições:

i. $p(x_i) \geq 0$ para todo i,

ii. $\sum_{i=1}^{n} p(x_i) = 1$.

A função p, definida acima, é denominada **Função de Probabilidades** da variável aleatória X. A probabilidade de um determinado valor da variável aleatória é igual a probabilidade do evento associado ao valor da variável.

Exemplo

Sejam: E: "lançamento de duas moedas";

X: "nº de caras obtidas nas duas moedas".

Então:

- $P(X = 0) = P[(k, k)] = \frac{1}{4}$;
- $P(X = 1) = P[(c, k), (k, c)] = \frac{2}{4} = \frac{1}{2}$;
- $P(X = 2) = P[(c, c)] = \frac{1}{4}$.

A coleção de pares $[x_i, p(x_i)], i = 1, 2, \ldots$, é algumas vezes denominada **Distribuição de Probabilidades** de X.

Exemplo

Sejam: E: "lançamento de duas moedas";

X: "nº de caras obtidas".

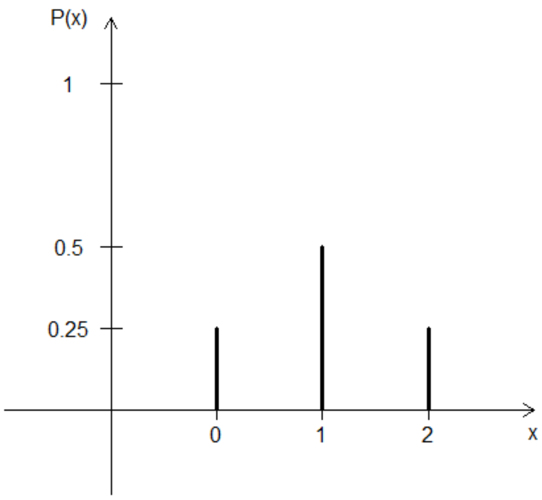
Eis as várias expressões para $P(x)$:

1. Tabela

Tabela 2.1. Distribuição de probabilidades da variável aleatória X: "nº de caras obtidas"

$\boldsymbol{x}$	$\boldsymbol{P(x)}$
0	$1/4$
1	$1/2$
2	$1/4$

2. Gráfico



3. Fórmula

$$P(x) = \frac{1}{4}\binom{2}{x}, \qquad \text{para} \qquad x = 0, 1, 2.$$

Verificando a fórmula:

- $P(X = 0) = \frac{1}{4}\binom{2}{0} = \frac{1}{4}\frac{2!}{0!(2-0)!} = \frac{1}{4} \cdot \frac{2!}{1!\,.\,2!} = \frac{1}{4} \cdot 1 = \frac{1}{4}$
- $P(X = 1) = \frac{1}{4}\binom{2}{1} = \frac{1}{4}\frac{2!}{1!(2-1)!} = \frac{1}{4} \cdot \frac{2!}{1!\,.\,1!} = \frac{1}{4} \cdot 2 = \frac{1}{2}$
- $P(X = 2) = \frac{1}{4}\binom{2}{2} = \frac{1}{4}\frac{2!}{2!(2-2)!} = \frac{1}{4} \cdot \frac{2!}{2!\,.\,0!} = \frac{1}{4} \cdot 1 = \frac{1}{4}$

Note que

$$\sum_{i=1}^{3} P(x_i) = P(0) + P(1) + P(2) = \frac{1}{4} + \frac{1}{2} + \frac{1}{4} = 1.$$

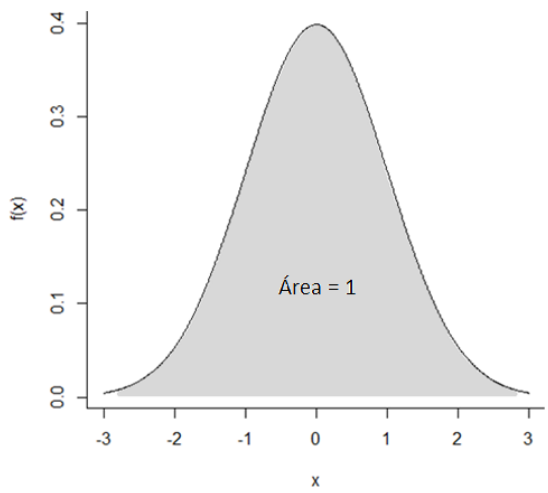
Função Densidade de Probabilidade

Definição: Uma variável aleatória contínua X é contínua em $\mathbb{R}$, se existir uma função $f(x)$, tal que:

i) $f(x) \geq 0, \quad \forall\, x \in \mathbb{R}$;

ii) Área entre o gráfico da função f e o eixo x é igual a 1, ou seja,

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



A função $f(x)$ é chamada **função densidade de probabilidade** (f.d.p.).

Exemplo

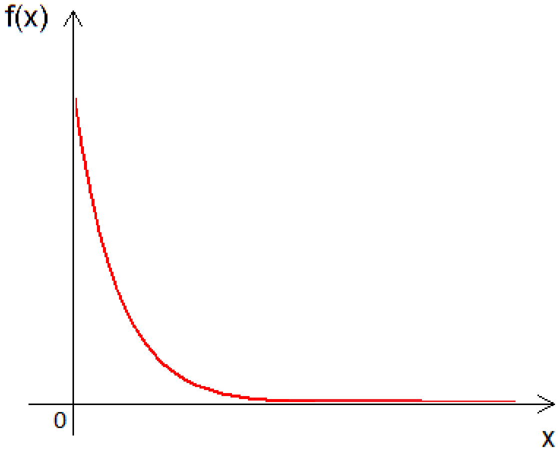
Seja X uma variável aleatória com função de densidade de probabilidade dada por:

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

Verifique se $f(x)$ é uma fdp de X.

Solução

Note que $f(x) \geq 0$ para todo x, como pode-se observar pelo gráfico de f;



Além disso,

$$\int_0^\infty \lambda e^{-\lambda x}\, dx = -e^{-\lambda x}\Big|_0^\infty$$

$$= \lim_{x\to\infty}\left(-e^{-\lambda x}\right) \; -(-e^{-0})$$

$$= \lim_{x\to\infty}\left(-\frac{1}{e^{\lambda x}}\right) + e^0$$

$$= 0 + 1$$

$$= 1$$

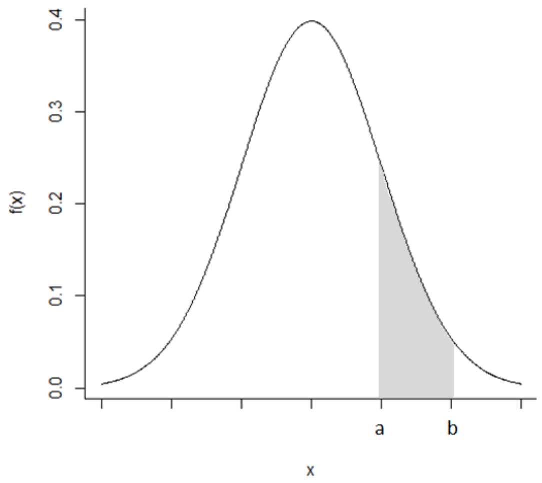
ou seja, $f(x)$ é realmente uma *fdp* da variável X.

Cálculo de probabilidades para uma variável aleatória contínua

Para qualquer $a < b$ pertencente ao contradomínio de X, temos:

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

Graficamente, a probabilidade acima corresponde à área limitada pela função $f(x)$, eixo x e pelas retas $x = a$ e $x = b$:



Da relação entre a probabilidade e a área sob a função, a inclusão ou não dos extremos a e b não afetará os resultados. Assim, será admitido que

$$P(a \leq X \leq b) = P(a \leq X < b) = P(a < X \leq b) = P(a < X < b).$$

Exemplo

Considere a variável aleatória contínua X cuja fdp é dada por:

$$f(x) = \begin{cases} 2x, & \text{se } \ 0 < x < 1, \\ 0, & \text{se } \ x \leq 0 \text{ ou } x \geq 1 \end{cases}$$

Evidentemente, $f(x) \geq 0$. Além disso,

$$\int_{-\infty}^{+\infty} f(x)dx = \int_{-\infty}^{0} 0\,dx + \int_{0}^{1} 2x\,dx + \int_{1}^{+\infty} 0\,dx = \int_{0}^{1} 2x\,dx = x^2\Big|_0^1 = 1^2 - 0^2 = 1.$$

Logo, $f(x)$ é realmente uma fdp.

Para exemplificar o cálculo de uma probabilidade, vamos calcular, por exemplo, $P(0 < X < 1/2)$:

$$P\left(0 < X < \frac{1}{2}\right) = \int_{0}^{\frac{1}{2}} 2x\,dx = x^2\Big|_0^{\frac{1}{2}} = \left(\frac{1}{2}\right)^2 - 0^2 = \frac{1}{4}.$$

Função de Distribuição Acumulada

Seja X uma variável aleatória discreta. Define-se função de distribuição acumulada (ou função de repartição) da variável aleatória X, no ponto x, como sendo a probabilidade de que X assuma um valor menor ou igual a x, isto é:

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

Propriedades

1. $F(x) = \sum_{x_i \leq x} P(x_i)$
2. $F(-\infty) = 0$
3. $F(+\infty) = 1$
4. $P(a < X \leq b) = F(b) - F(a)$
5. $P(a \leq X \leq b) = F(b) - F(a) + P(X = a)$
6. $P(a < X < b) = F(b) - F(a) - P(X = b)$
7. $F(x)$ é conínua à direita $\rightarrow \lim_{x \to x_0} F(x) = F(x_0)$
8. $F(x)$ é descontínua à esquerda, nos pontos em que a probabilidade é diferente de zero.
9. A função é não decrescente, isto é, $F(b) \geq F(a)$, para $b > a$.

Exemplo

Sejam: E: "lançamento de duas moedas";

X: "nº de caras obtidas".

Vimos que a distribuição de probabilidades de X é dada por:

Recordando a Tabela 2.1

x	$P(x)$
0	$1/4$
1	$1/2$
2	$1/4$

Então:

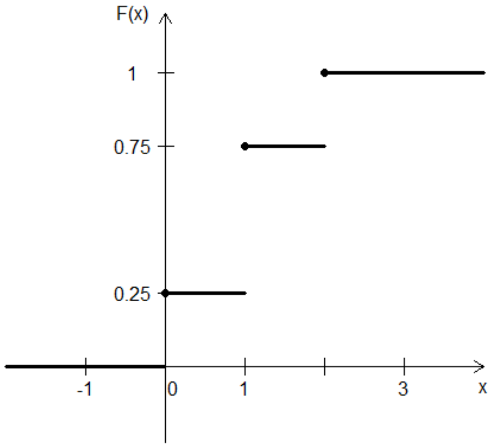
$F(x) = 0$ se $x < 0$ $\quad (F(x) = 0)$

$F(x) = \frac{1}{4}$ se $0 \leq x < 1$ $\quad \left(F(x) = \frac{1}{4} = 0{,}25\right)$

$F(x) = \frac{3}{4}$ se $1 \leq x < 2$ $\quad \left(F(x) = \frac{1}{4} + \frac{1}{2} = \frac{3}{4} = 0{,}75\right)$

$F(x) = 1$ se $x \geq 2$ $\quad \left(F(x) = \frac{1}{4} + \frac{1}{2} + \frac{1}{4} = 1\right)$

Eis o gráfico de $F(x)$:



Se X é uma variável aleatória contínua com função densidade de probabilidade f(y) define-se a sua função de distribuição acumulada $F(x)$ como:

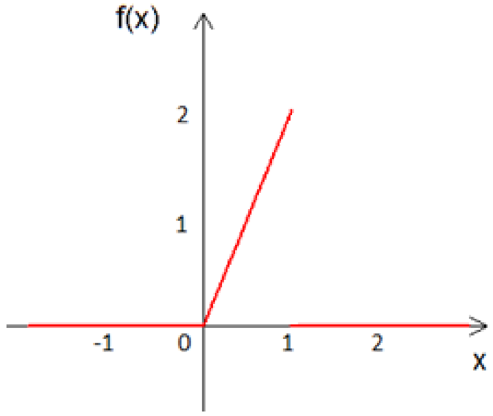
$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)dt$$

Exemplo

Considere a variável aleatória contínua X cuja fdp é dada por:

$$f(x) = \begin{cases} 0, & se \;\; x \leq 0, \\ 2x, & se \;\; 0 < x < 1, \\ 0, & se \;\; x \geq 1. \end{cases}$$

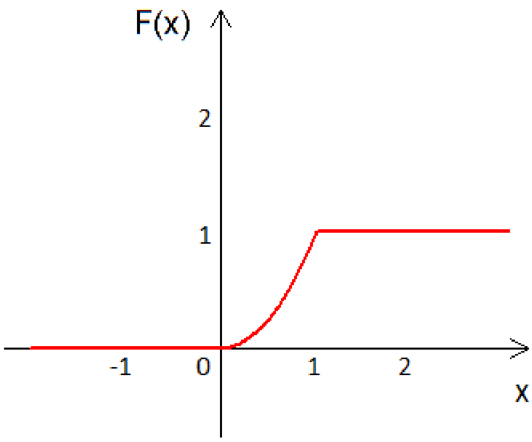
Eis o gráfico de $f(x)$:



Portanto, a função de distribuição é dada por:

$$F(x) = \begin{cases} 0 & \text{se} \quad x \leq 0, \\ \displaystyle\int_0^x 2t\, dt = x^2 & \text{se} \quad 0 < x \leq 1, \\ 1 & \text{se} \quad x > 1. \end{cases}$$

Eis o gráfico de $F(x)$:



Teorema

Seja F a função de distribuição acumulada de uma variável aleatória contínua, com fdp f. Então,

$$f(x) = \frac{d}{dx} F(x),$$

para todo x no qual F seja derivável.

Demonstração

$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\,dt$. Aplicando-se o Teorema Fundamental do Cálculo, obteremos $F'(x) = f(x)$.

Exemplo

Seja X uma variável aleatória contínua com função de distribuição acumulada F dada por

$$F(x) = \begin{cases} 0, & \text{se } x \leq 0 \\ 1 - e^{-x}, & \text{se } x > 0 \end{cases}$$

Neste caso, $F'(x) = e^{-x}$ para $x > 0$, e, por isso, a fdp será dada por

$$f(x) = \begin{cases} 0, & \text{se } x \leq 0 \\ e^{-x}, & \text{se } x > 0 \end{cases}$$

Se a e b forem dois números reais quaisquer, tem-se que:

$$P(a < X < b) = \int_{a}^{b} f(x)dx = F(b) - F(a)$$

Média ou Esperança Matemática

Define-se Esperança Matemática ou Média de uma variável aleatória discreta como:

$$E(X) = \mu = \sum_{i=1}^{n} x_i P(x_i).$$

Define-se Esperança Matemática ou Média de uma variável aleatória contínua como:

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

Exemplo

Seja o experimento E: "lançamento de duas moedas", e seja a variável aleatória X: "número de caras". Então:

$$\Omega = \{(c,c),(c,k),(k,c),(k,k)\}.$$

A distribuição de probabilidades de X é dada por:

Tabela 2.2. Distribuição de probabilidades da variável aleatória X: "nº de caras"

$\boldsymbol{x}$	$\boldsymbol{P(x)}$
0	1/4
1	1/2
2	1/4
Total	1

Então,

$$\begin{aligned} E(X) &= \sum_{i=1}^{n} x_i P(x_i) \\ &= 0\,.\frac{1}{4} + 1\,.\frac{1}{2} + 2\,.\frac{1}{4} \\ &= 1. \end{aligned}$$

Exemplo

Seja X uma variável aleatória com função de probabilidade dada por:

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

Onde n é um número inteiro positivo, $0 \leq p \leq 1$, e para cada par fixo n e p a função de probabilidade soma 1. O valor esperado desta variável aleatória é dado por:

$$\begin{aligned} E(X) &= \sum_{x=0}^{n} x \binom{n}{x} p^x (1-p)^{n-x} \\ &= \sum_{x=1}^{n} x \binom{n}{x} p^x (1-p)^{n-x} \qquad \text{(em } x = 0 \text{ o termo é 0)} \end{aligned}$$

Utilizando a identidade $x\binom{n}{x} = n\binom{n-1}{x-1}$, temos

$$E(X) = \sum_{x=1}^{n} n \binom{n-1}{x-1} p^x (1-p)^{n-x}$$

$$= \sum_{y=0}^{n-1} n \binom{n-1}{y} p^{y+1} (1-p)^{n-(y+1)} \qquad \text{(substituindo } y = x - 1\text{)}$$

$$= \sum_{y=0}^{n-1} n \binom{n-1}{y} p^{y}.p.(1-p)^{(n-1)-y}$$

$$= np \underbrace{\sum_{y=0}^{n-1} \binom{n-1}{y} p^{y} (1-p)^{(n-1)-y}}_{1}$$

$$= np$$

uma vez que a última soma deve ser 1, sendo a soma de todos os valores possíveis de uma função de probabilidade.

Exemplo

Seja X uma variável aleatória com função de densidade de probabilidade dada por:

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

A Esperança Matemática da variável aleatória X é dada por:

$$E(X) = \int_0^{\infty} x\, \lambda e^{-\lambda x}\, dx$$

Resolveremos usando integração por partes:

$$\int u\, dv = uv - \int v\, du$$

então $u = x$ e $dv = \lambda e^{-\lambda x} dx$ teremos $du = dx$ e $v = -e^{-\lambda x}$. Logo

$$E(X) = \int_0^{\infty} x\, \lambda e^{-\lambda x}\, dx = -xe^{-\lambda x} \Big|_0^{\infty} - \int_0^{\infty} -e^{-\lambda x}\, dx$$

$$= -\frac{x}{e^{\lambda x}} \Big|_0^{\infty} + \int_0^{\infty} e^{-\lambda x}\, dx$$

$$= -\frac{x}{e^{\lambda x}} \Big|_0^{\infty} + \frac{-e^{-\lambda x}}{\lambda} \Big|_0^{\infty}$$

$$= -\frac{x}{e^{\lambda x}} \Big|_0^{\infty} + \frac{-\left(\frac{1}{e^{\lambda x}}\right)}{\lambda} \Big|_0^{\infty}$$

$$= \ [-0-(-0)] + \left[-0-\left(-\frac{1}{\lambda}\right)\right]$$

$$= \frac{1}{\lambda}.$$

Propriedades da Média

a. $E(k) = k,$ sendo k uma constante.

Demonstração

$$E(k) = \sum_{i=1}^{n} k \, . \, p(x_i)$$

$$= k \, . \sum_{i=1}^{n} p(x_i)$$

$$= k \, . \, 1$$

$$= k$$

b. $E(kX) = k.E(X)$, sendo k uma constante.

Demonstração

$$E(kX) = \sum_{i=1}^{n} kx_i \, . \, p(x_i)$$

$$= k \, . \sum_{i=1}^{n} x_i \, p(x_i)$$

$$= k.E(X)$$

c. $E(aX \pm b) = aE(X) \pm b$, sendo a e b constantes.

Demonstração

$$E(aX + b) = \sum_{i=1}^{n} (ax_i + b).p(x_i)$$

$$= \sum_{i=1}^{n} \big(ax_i \, p(x_i) + b \, p(x_i)\big)$$

$$= \sum_{i=1}^{n} ax_i \, p(x_i) \ + \sum_{i=1}^{n} b \, p(x_i)$$

$$= a \, . \sum_{i=1}^{n} x_i \, p(x_i) \ + b \, . \sum_{i=1}^{n} p(x_i)$$

$$= aE(X) + b\,.\,1$$
$$= aE(X) + b$$

Definição

Para cada número inteiro n, o n-ésimo momento da variável aleatória X, chamado μ_n, é:

$$\mu_n = E(X^n).$$

O n-ésimo momento central de X, $\mu_n' = E(X-\mu)^n$.

Variância

Define-se variância de uma variável aleatória X como sendo o segundo momento central:

$$Var(X) = E(X-\mu)^2 = E[X - E(X)]^2$$

Podemos reescrever a fórmula da variância como:

$$\begin{aligned} V(X) &= E\big(X - E(X)\big)^2 \\ &= E\big(X^2 - 2\,X\,E(X) + E^2(X)\big) \\ &= E(X^2) - 2E(X)E(X) + E^2(X) \\ &= E(X^2) - 2E^2(X) + E^2(X) \\ &= E(X^2) - E^2(X) \end{aligned}$$

ou seja,

$$V(X) = E(X^2) - E^2(X)$$

A raiz quadrada positiva de $V(X)$ é o desvio padrão de X, ou seja,

$$dp(X) = \sqrt{V(X)}.$$

A variância da uma medida do grau de dispersão de uma distribuição ao redor de sua média. O ***desvio padrão*** é mais fácil de ser interpretado, no sentido de que a unidade de medida no ***desvio padrão*** é a mesma que para a variável original X. A unidade de mediada na variância é o quadrado da unidade original.

Se X é uma variável aleatória discreta, então:

$$\sigma^2 = \sum (x_i - \mu)^2 P(x_i).$$

Exemplo

Seja o experimento E: "lançamento de duas moedas", e seja a variável aleatória X: "número de caras". Vimos a distribuição de probabilidades de X na Tabela 2.2.

Recordando a Tabela 2.2

x	$P(x)$
0	1/4
1	1/2
2	1/4
Total	1

Já vimos, também, que $E(X) = 1$. Logo:

$$\begin{aligned} Var(X) &= \sum (x_i - \mu)^2 P(x_i) \\ &= (0-1)^2 . \frac{1}{4} + (1-1)^2 . \frac{1}{2} + (2-1)^2 . \frac{1}{4} \\ &= 0{,}5. \end{aligned}$$

Usando a fórmula alternativa da variância:

$$\begin{aligned} E(X) &= \sum_{i=1}^{n} x_i P(x_i) \\ &= 0 . \frac{1}{4} + 1 . \frac{1}{2} + 2 . \frac{1}{4} = 1{,}0. \end{aligned}$$

e

$$\begin{aligned} E(X^2) &= \sum_{i=1}^{n} x_i^2 \, P(x_i) \\ &= 0^2 . \frac{1}{4} + 1^2 . \frac{1}{2} + 2^2 . \frac{1}{4} = 1{,}5. \end{aligned}$$

Portanto:

$$\begin{aligned} V(X) &= E(X^2) - [E(X)]^2 \\ &= 1{,}5 - 1{,}0^2 \\ &= 1{,}5 - 1{,}0 \\ &= 0{,}5. \end{aligned}$$

O desvio padrão da variável aleatória X é obtido por:

$$dp(X) = \sqrt{V(X)} = \sqrt{0{,}5} \approx 0{,}7071.$$

2.2 Covariância e Correlação

Anteriormente discutimos a ausência ou presença de uma relação entre duas variáveis aleatórias, independência ou não independência. Mas se houve uma relação, esta poderá ser forte ou fraca. Uma das medidas para medir este grau de dependência é a *covariância*, como já definimos anteriormente. Uma outra medida, mais utilizada, e de melhor interpretação, pode ser obtida através do *Coeficiente de Correlação*.

Definição

A covariância mede o grau de dependência entre duas variáveis aleatórias X e Y.

$$\begin{aligned} COV(X,Y) &= E\{[X - E(X)].[Y - E(Y)]\} \\ &= E\{[XY - XE(Y) - YE(X) + E(X)E(Y)]\} \\ &= E(XY) - E(X)E(Y) - E(Y)E(X) + E(X)E(Y) \\ &= E(XY) - E(X)E(Y) \end{aligned}$$

Se X e Y forem independentes temos $E(XY) = E(X)E(Y)$ então $COV(X,Y) = 0$.

Exemplo

Considere a distribuição conjunta de X e Y dada na tabela a seguir:

$Y \backslash X$	**0**	**1**	**2**	$p(y)$
1	$3/20$	$3/20$	$2/20$	$8/20$
2	$1/20$	$1/20$	$2/20$	$4/20$
3	$4/20$	$1/20$	$3/20$	$8/20$
$p(x)$	$8/20$	$5/20$	$7/20$	$1{,}00$

Temos que:

$$E(X) = 0 \times \frac{8}{20} + 1 \times \frac{5}{20} + 2 \times \frac{7}{20} = 0{,}95$$

$$E(Y) = 1 \times \frac{8}{20} + 2 \times \frac{4}{20} + 3 \times \frac{8}{20} = 2{,}00$$

$$\begin{aligned} E(XY) = {} & (0 \times 1) \times \frac{3}{20} + (0 \times 2) \times \frac{1}{20} + (0 \times 3) \times \frac{4}{40} + \\ & +(1 \times 1) \times \frac{3}{20} + (1 \times 2) \times \frac{1}{20} + (1 \times 3) \times \frac{1}{20} + \\ & +(2 \times 1) \times \frac{2}{20} + (2 \times 2) \times \frac{2}{20} + (2 \times 3) \times \frac{3}{20} \\ & = 1{,}90 \end{aligned}$$

Obtemos, então:

$$\begin{aligned} COV(X,Y) &= E(XY) - E(X)E(Y) \\ &= 1{,}90 - (0{,}95)(2{,}00) \\ &= 0. \end{aligned}$$

Portanto, as variáveis aleatórias X e Y desse exemplo são independentes.

Propriedades da Variância

a. $V(k) = 0$;

Demonstração

$$\begin{aligned} V(k) &= E[k - E(k)]^2 \\ &= E[k - k]^2 \\ &= E[0]^2 \\ &= E[0] \\ &= 0 \end{aligned}$$

b. $V(kX) = k^2V(X)$, sendo k uma constante;

Demonstração

$$\begin{aligned} V(kX) &= E[kX - E(kX)]^2 \\ &= E[kX - kE(X)]^2 \\ &= E\{k\,.[X - E(X)]\}^2 \\ &= E\{k^2.[X - E(X)]^2\} \\ &= k^2\,.E[X - E(X)]^2 \\ &= k^2V(X) \end{aligned}$$

c. $V(X \pm Y) = V(X) + V(Y) \pm 2COV(X,Y)$;

Demonstração

$$\begin{aligned} V(X+Y) &= E\{[(X+Y) - E(X+Y)]^2\} \\ &= E\left\{\left[\left(X - E(X)\right) + \left(Y - E(Y)\right)\right]^2\right\} \\ &= E\left\{\left(X - E(X)\right)^2 + 2\left(X - E(X)\right)\left(Y - E(Y)\right) + \left(Y - E(Y)\right)^2\right\} \\ &= E\left\{\left(X - E(X)\right)^2 + 2[XY - XE(Y) - YE(X) + E(X)E(Y)] \right. \\ &\qquad \left. + \left(Y - E(Y)\right)^2\right\} \\ &= E\left(X - E(X)\right)^2 + 2E[XY - XE(Y) - YE(X) + E(X)E(Y)] \\ &\qquad + E\left(Y - E(Y)\right)^2 \\ &= V(X) + 2[E(XY) - E(X)E(Y) - E(X)E(Y) + E(X)E(Y)] + V(Y) \\ &= V(X) + V(Y) + 2[E(XY) - E(X)E(Y)] \\ &= V(X) + V(Y) + 2COV(X,Y) \end{aligned}$$

Analogamente, mostra-se que $V(X - Y) = V(X) + V(Y) - 2COV(X,Y)$

Sabemos que se X e Y são **independentes** então $COV(X,Y) = 0$. Logo,

$$V(X \pm Y) = V(X) + V(Y).$$

d. $V(aX \pm bY) = a^2\,V(X) + b^2\,V(Y) \pm 2\,ab\,COV(X,Y)$ em que a e b são constantes.

Demonstração: fica como exercício.

Definição

A correlação (*Coeficiente de Correlação*) de X e Y é o número definido por

$$\rho_{XY} = \frac{COV(X,Y)}{\sigma_X \sigma_Y}$$

A correlação é sempre entre -1 e 1, com os valores -1 e 1 indicando uma perfeita relação linear entre X e Y, isto é, $-1 \leq \rho_{XY} \leq 1$.

a) Quando $\rho_{XY} > 0$ existe uma correlação positiva entre as duas variáveis;

b) Quando $\rho_{XY} < 0$ existe uma correlação negativa entre as duas variáveis;

c) Quando $\rho_{XY} = 0$ não existe uma correlação entre as duas variáveis.

Exemplo

Duas variáveis aleatórias X e Y resultaram $COV(X,Y) = 0{,}25$, $Var(X) = 0{,}75$ e $Var(Y) = 0{,}25$, logo:

$$\begin{aligned} \rho(X,Y) &= \frac{COV(X,Y)}{\sigma(X)\sigma(Y)} \\ &= \frac{0{,}25}{\sqrt{0{,}75}.\sqrt{0{,}25}} \\ &= 0{,}58. \end{aligned}$$

2.3 Variáveis Aleatórias Múltiplas

Em muitas situações estamos interessados em observar duas ou mais características simultaneamente. Por exemplo, podemos estudar a estatura X e o peso Y de alguma pessoa escolhida, que forneceria o resultado (x, y). Denominaremos (X, Y) uma variável aleatória bidimensional (algumas vezes chamada de vetor aleatório).

Definição

a. Uma variável aleatória bidimensional (X, Y) será discreta se os valores possíveis de (X, Y) forem finitos ou infinitos numeráveis.

b. Uma variável aleatória bidimensional (X, Y) será contínua se puder tomar todos os valores em algum conjunto não-numerável do plano euclidiano.

Definição

a. Seja (X, Y) uma variável aleatória discreta bidimensional. A cada resultado possível (x_i, y_i) associaremos um número $p(x_i, y_i)$ representando $P(X = x_i, Y = y_i)$ e satisfazendo às seguintes condições:

1. $p(x_i, y_i) \geq 0$ para todo (x, y);

2. $\sum_{j=1}^{\infty}\sum_{i=1}^{\infty} p(x_i y_i) = 1.$

b. Se (X, Y) for uma variável aleatória contínua bidimensional. A função densidade de probabilidade conjunta f é uma função que satisfaz às seguintes condições:

1. $f(x, y) \geq 0$ para todo $(x, y) \in \mathbb{R}$;

2. $\iint_{\Re} f(x, y)\, dx\, dy = 1.$

Exemplo

Duas linhas de produção fabricam um certo tipo de peça. Suponha que a capacidade (em qualquer dia) seja 5 peças na linha I e 3 peças na linha II. Admita que o número de peças realmente produzidas em qualquer linha seja uma variável aleatória e que (X, Y) represente a variável aleatória bidimensional que fornece o número de peças produzidas pelas linhas I e II, respectivamente. A tabela a seguir dá a distribuição de probabilidades conjunta de (X, Y).

Y \ X	**0**	**1**	**2**	**3**	**4**	**5**
0	0	0,01	0,03	0,05	0,07	0,09
1	0,01	0,02	0,04	0,05	0,06	0,08
2	0,01	0,03	0,05	0,05	0,05	0,06
3	0,01	0,02	0,04	0,06	0,06	0,05

Cada célula no interior da tabela representa $p(x_i, y_i)$.

Considere o evento $A =$ "Mais peças são produzidas pela linha I que pela linha II". Determine $P(A)$.

Solução

Vamos considerar apenas o s pares (x, y) tais que $x > y$ e somar estas probabilidades.

Y \ X	0	1	2	3	4	5
0		0,01	0,03	0,05	0,07	0,09
1			0,04	0,05	0,06	0,08
2				0,05	0,05	0,06
3					0,06	0,05

Portanto,

$$P(B) = 0{,}01 + 0{,}03 + 0{,}05 + 0{,}07 + 0{,}09 + 0{,}04 + \cdots + 0{,}06 + 0{,}05 = 0{,}75.$$

Exemplo

Suponha que a va contínua bidimensional (X, Y) tenha fdp conjunta dada por

$$f(x, y) = x^2 + \frac{xy}{3}, \quad 0 \leq x \leq 1, \quad 0 \leq y \leq 2$$

e $f(x, y) = 0$ para quaisquer outros valores. Verifique se f é realmente uma fdp conjunta de (X, Y).

Solução

Observa-se facilmente que $f(x, y) \geq 0$ para todo (x, y). Além disso,

$$\begin{aligned}
\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x, y)\ dx\ dy &= \int_0^2\int_0^1 \left(x^2 + \frac{xy}{3}\right)\ dx\ dy \\
&= \int_0^2 \left(\frac{x^3}{3} + \frac{x^2y}{6}\right)\Bigg|_0^1\ dy \\
&= \int_0^2 \left(\frac{1}{3} + \frac{y}{6}\right)\ dy \\
&= \left(\frac{y}{3} + \frac{y^2}{12}\right)\Bigg|_0^2 \\
&= \frac{2}{3} + \frac{4}{12} \\
&= 1
\end{aligned}$$

Portanto, $f(x, y)$ é uma fdp conjunta de (X, Y).

Distribuições de probabilidade Marginal e Condicional

Exemplo

Considere o exemplo anterior em que $X =$ "capacidade de produção da linha I" e $Y =$ "capacidade de produção da linha II". Considere a tabela a seguir:

Y \ X	**0**	**1**	**2**	**3**	**4**	**5**	$q(y_j)$
0	0	0,01	0,03	0,05	0,07	0,09	0,25
1	0,01	0,02	0,04	0,05	0,06	0,08	0,26
2	0,01	0,03	0,05	0,05	0,05	0,06	0,25
3	0,01	0,02	0,04	0,06	0,06	0,05	0,24
$p(x_i)$	0,03	0,08	0,16	0,21	0,24	0,28	1

Cada célula no interior da tabela representa $p(x_i, y_i)$. A primeira e última linha dão a distribuição de X, isto é, $\left(x, p(x)\right)$, enquanto a primeira e última coluna dão a distribuição de Y, isto é, $\left(y, p(y)\right)$. Essas distribuições são chamadas distribuições marginais de X e Y.

Se quisermos, por exemplo, determinar a probabilidade condicional de $X = 2$ dado $Y = 2$, ou seja, $P(X = 2|Y = 2)$, pela definição de probabilidade condicional temos que:

$$P(X = 2|Y = 2) = \frac{P(X = 2 \cap Y = 2)}{P(Y = 2)} = \frac{P(X = 2\,, Y = 2)}{P(Y = 2)} = \frac{0{,}05}{0{,}25} = 0{,}20.$$

3. Principais modelos de distribuição de probabilidade

Veremos, a seguir, alguns dos principais modelos de distribuição de probabilidade discretos e contínuos.

3.1 Distribuição de Bernoulli

Nos experimentos de Bernoulli o espaço amostral é composto por apenas dois resultados possíveis: "sucesso" (resultado de interesse) ou "fracasso" (resultado pelo qual não estamos interessados).

Exemplos

(i) E_1: "Lançar uma moeda": pode sair cara ou coroa;

(ii) E_2: "Plantar uma semente": pode germinar ou não.

Seja X a variável aleatória: sucesso ou fracasso. Seja p a probabilidade de ocorrer sucesso e $1-p$ a probabilidade de fracasso. Podemos ver, pela Tabela 3.1, as possibilidades, as probabilidades e os valores da variável aleatória de Bernoulli. Note que nos valores da variável codificamos sucesso como $x_i = 1$ e fracasso como $x_i = 0$.

Tabela 3.1. Resultados possíveis, probabilidades e valores (x_i) da variável aleatória X de Bernoulli

Resultados Possíveis	**Probabilidades**	$\boldsymbol{x_i}$
Sucesso	p	1
Fracasso	$1-p$	0

A distribuição de probabilidade de X com distribuição de Bernoulli, com parâmetro p é apresentada na Tabela 3.2.

Tabela 3.2. Distribuição de probabilidade da variável aleatória X de Bernoulli

$\boldsymbol{x_i}$	$\boldsymbol{P(X = x_i)}$
0	$1-p$
1	p
Total	1

Pode-se calcular a média desta distribuição:

$$\mu = E(X) = \sum_{i=1}^{n} x_i P(xi)$$
$$= 0.(1-p) + 1.p$$
$$= p.$$

Pode-se calcular a variância desta distribuição:

$$Var(X) = \sum_{i=1}^{n} [x_i - \mu]^2 P(x_i)$$
$$= (0-\mu)^2.(1-p) + (1-\mu)^2.p$$
$$= (0-p)^2.(1-p) + (1-p)^2.p$$
$$= p^2(1-p) + p(1-p)^2$$
$$= p^2(1-p) + p(1-2p+p^2)$$
$$= p^2 - p^3 + p - 2p^2 + p^3$$
$$= p - p^2$$
$$= p(1-p).$$

Portanto:

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

Notação: $X \sim Bernoulli(p)$.

Função de probabilidades

A função de probabilidades de uma distribuição de Bernoulli é dada por:

$$P(X = x_i) = p^{x_i}(1-p)^{1-x_i}, \quad x_i = 0, 1.$$

3.2 Distribuição Binomial

Trata-se de uma distribuição de probabilidade adequada aos experimentos que apresentam apenas dois resultados (sucesso ou fracasso). Este modelo fundamenta-se nas seguintes hipóteses:

1. São realizados n ensaios **independentes** (o resultado de um experimento não é afetado pelo resultado dos outros) de Bernoulli;

2. A probabilidade de sucesso em cada ensaio é p e de fracasso é $q = 1 - p$;
3. O número observado de sucessos é um número inteiro entre 0 e n.

Diz-se que a variável X: "número de sucessos em n ensaios" tem **distribuição binomial** com parâmetros n e p, onde n é o número de ensaios e p é a probabilidade de sucesso em cada ensaio.

Função de Probabilidades

A função de probabilidades de uma variável X com distribuição Binomial, $X \sim Bin(n, p)$, é dada por:

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

ou

$$P(X = x) = \binom{n}{x} p^x q^{n-x}, \qquad x = 0, 1, \dots, n$$

em que $\binom{n}{x} = \frac{n!}{x!(n-x)!}$.

Notação: $X \sim Bin(n, p)$.

Exemplo

Suponha que em uma maternidade irão nascer três bebês. Vamos estudar as possibilidades do sexo dos bebês nos três nascimentos. Considerando a variável X: "número de meninos em três nascimentos", e considerando que a probabilidade de nascer menino é $p = 0{,}5$ (então $q = 0{,}5$, pois $q = 1 - p$), temos:

- $P(X = 0) = \binom{3}{0} 0{,}5^0 . 0{,}5^{3-0} = 1 \,.\, 1 \,.\, 0{,}5^3 = 0{,}125$
- $P(X = 1) = \binom{3}{1} 0{,}5^1 . 0{,}5^{3-1} = 3 \,.\, 0{,}5 \,.\, 0{,}5^2 = 0{,}375$
- $P(X = 2) = \binom{3}{2} 0{,}5^2 . 0{,}5^{3-2} = 3 \,.\, 0{,}5^2 \,.\, 0{,}5 = 0{,}375$
- $P(X = 3) = \binom{3}{3} 0{,}5^3 . 0{,}5^{3-3} = 1 \,.\, 0{,}5^3 \,.\, 1 = 0{,}125$

Comandos no Software R para calcular as probabilidades da distribuição Binomial:

```
dbinom(0, 3, 0.5) #x=0, n=3, p=0.5
dbinom(1, 3, 0.5) #x=1, n=3, p=0.5
dbinom(2, 3, 0.5) #x=2, n=3, p=0.5
dbinom(3, 3, 0.5) #x=3, n=3, p=0.5
```

Exemplo

Considere a variável X: "número de meninos em três nascimentos", do exemplo anterior, porém, considere agora que a probabilidade de nascer menino é $p = 0{,}6$ (então $q = 0{,}4$, pois $q = 1 - p$). Assim, temos:

- $P(X = 0) = \binom{3}{0} 0{,}6^0 . 0{,}4^{3-0} = 1 \, . \, 1 \, . \, 0{,}4^3 = 0{,}064$
- $P(X = 1) = \binom{3}{1} 0{,}6^1 . 0{,}4^{3-1} = 3 \, . \, 0{,}6 \, . \, 0{,}4^2 = 0{,}288$
- $P(X = 2) = \binom{3}{2} 0{,}6^2 . 0{,}4^{3-2} = 3 \, . \, 0{,}6^2 \, . \, 0{,}4 = 0{,}432$
- $P(X = 3) = \binom{3}{3} 0{,}6^3 . 0{,}4^{3-3} = 1 \, . \, 0{,}6^3 \, . \, 1 = 0{,}216$

Comandos no Software R para calcular as probabilidades da distribuição Binomial:

```
dbinom(0, 3, 0.6) #x=0, n=3, p=0.6
dbinom(1, 3, 0.6) #x=1, n=3, p=0.6
dbinom(2, 3, 0.6) #x=2, n=3, p=0.6
dbinom(3, 3, 0.6) #x=3, n=3, p=0.6
```

Média e variância na Distribuição Binomial

A média μ de uma variável $X \sim Bin(n, p)$ é dada por:

$$\mu = np$$

e a variância σ^2 de uma variável $X \sim Bin(n, p)$ é dada por:

$$\sigma^2 = npq.$$

Exemplo

A probabilidade de nascer um menino é $p = 0{,}5$ (ignorando nascimento de gêmeos e nascimentos múltiplos). Calcule a média e a variância do número de meninos em 1.000 nascimentos.

Solução

A média é

$$\mu = np = 1000\ .\ 0{,}5 = 500 \text{ meninos},$$

e a variância é

$$\sigma^2 = npq = 1000\ .\ 0{,}5\ .\ 0{,}5 = 250.$$

3.3 Distribuição de Poisson

A distribuição de Poisson é largamente empregada quando se deseja contar o número de ocorrências de um evento de interesse, por unidade de tempo, comprimento, área ou volume.

Exemplos

a) Número de insetos de uma espécie coletados por armadilha por dia;
b) Número de chamadas recebidas por um telefone durante cinco minutos;
c) Número de falhas de um computador num dia de operação;
d) Número de bactérias por *ml* de urina;
e) Número de pacientes que chegam a um pronto atendimento de uma pequena cidade durante a madrugada.

Note que os possíveis valores que as variáveis descritas podem assumir são: 0, 1, . . . ,. O comportamento dessas variáveis pode ser descrito pela chamada distribuição de Poisson.

Função de Probabilidades

A função de probabilidades de uma variável X com distribuição Poisson é dada por:

$$P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}, \quad x = 0, 1, \ldots$$

em que $e = 2{,}7182\ldots$ é o número de Euler, e λ é o número médio de ocorrências do evento de interesse por unidade de tempo, distância ou área.

Notação: $X \sim Pois(\lambda)$

Média e variância na distribuição de Poisson

A esperança e a variância de uma variável aleatória $X \sim Pois(\lambda)$ são dadas, respectivamente, por:

$$\mu = \lambda \qquad \text{e} \qquad \sigma^2 = \lambda.$$

Exemplo

Uma telefonista recebe, em média, 2 chamadas por hora. Calcular a probabilidade de se receber:

- **a)** nenhuma chamada em 1 hora;
- **b)** 1 chamada em 1 hora;
- **c)** 2 chamadas em 1 hora;
- **d)** no máximo 2 chamadas em 1 hora.
- **e)** 3 chamadas em 2 horas;

Solução

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

a) $P(X = 0) = \frac{e^{-2}\, 2^0}{0!} = 0{,}1353$

b) $P(X = 1) = \frac{e^{-2}\, 2^1}{1!} = 0{,}2707$

c) $P(X=2)=\frac{e^{-2}\,2^2}{2!}=0{,}2707$

d) $P(X\leq 2)=P(X=0)+P(X=1)+P(X=2)$

$$=0{,}1353+0{,}2707+0{,}2707=0{,}6767$$

e) Como $\lambda=2$ chamadas em 1 hora, então, em 2 horas, teremos $\lambda^*=4$ chamadas. Logo:

$$P(X=3)=\frac{e^{-(\lambda^*)}(\lambda^*)^x}{x!}=\frac{e^{-4}4^3}{3!}=0{,}1954$$

Comandos no Software R para calcular as probabilidades da distribuição de Poisson:

```
dpois(0, 2) #x=0, lambda=2
dpois(1, 2) #x=1, lambda=2
dpois(2, 2) #x=2, lambda=2
sum(dpois(0:2, 2)) #x=0:2, lambda=2
dpois(3, 4) #x=3, lambda=4
```

Exemplo

Um pesquisador está interessado no número de ovos depositados por uma espécie de pássaro. Na primavera, ele procura e acha 80 ninhos. O número médio de ovos por ninho foi 3,8 e a variância foi 3,1. Porque a variância é aproximadamente igual á média, ele acha que pode ser razoável descrever o número de ovos por ninho como tendo uma distribuição Poisson com média $\lambda=3{,}8$.

a) Se esta realmente representa a distribuição populacional, qual seria a probabilidade de encontrar um ninho com nenhum ovo?

b) Qual seria a probabilidade de encontrar um ninho com no máximo 4 ovos?

c) Qual seria a probabilidade de encontrar um ninho com mais de 4 ovos?

Solução

a) $P(X=0)=\frac{e^{-3{,}8}\,3{,}8^0}{0!}=0{,}0224$

b) $P(X \leq 4) = P(0) + P(1) + P(2) + P(3) + P(4)$

$$= \frac{e^{-3,8}\,3{,}8^0}{0!} + \frac{e^{-3,8}\,3{,}8^1}{1!} + \frac{e^{-3,8}\,3{,}8^2}{2!} + \frac{e^{-3,8}\,3{,}8^3}{3!} + \frac{e^{-3,8}\,3{,}8^4}{4!}$$

$$= 0{,}0224 + 0{,}0850 + 0{,}1615 + 0{,}2046 + 0{,}1944$$

$$= 0{,}6679$$

c) $P(X > 4) = 1 - P(X \leq 4) = 1 - 0{,}6679 = 0{,}3321$

Comandos no Software R para calcular as probabilidades da distribuição de Poisson:

```
dpois(0, 3.8) #x=0, lambda=3.8
sum(dpois(0:4, 3.8)) #x=0:4, lambda=3.8
1-sum(dpois(0:4, 3.8)) #x=5:Inf, lambda=3.8
```

3.4 Distribuição Geométrica

Considere uma sequência de ensaios de Bernoulli independentes (com probabilidade de sucesso p, $0 < p < 1$). Defina X como o número de fracassos anteriores ao primeiro sucesso ou, em outras palavras, o tempo de espera (em termos de ensaios anteriores) para o primeiro sucesso. A variável X segue o modelo Geométrico com parâmetro p.

Função de Probabilidades

A função de probabilidade do modelo Geométrico é dada por:

$$P(X = x) = (1-p)^x . p, \quad x = 0, 1, 2, \ldots$$

Notação: $X \sim Geo(p)$.

Média e variância da Distribuição Geométrica

$$E(X) = \frac{1-p}{p} \quad \text{e} \quad V(X) = \frac{1-p}{p^2}.$$

Exemplo

Uma linha de fabricação de um equipamento de precisão é interrompida na primeira ocorrência de um defeito. A partir da manutenção, o equipamento tem probabilidade de 0,01 de apresentar defeito em um dia qualquer. Seja X a variável que conta o número de dias que antecedem a interrupção. Admitindo que o desempenho, nos dias sucessivos, sejam independentes, temos que $X \sim Geo(p = 0{,}01)$. Dessa forma,

$$P(X = x) = 0{,}99^x . 0{,}01, \quad x = 0, 1, 2, \ldots$$

Por exemplo, para uma interrupção no sexto dia temos:

$$P(X = 5) = 0{,}99^5 . 0{,}01 = 0{,}0095.$$

Observação: Para que a interrupção ocorra no sexto dia então o defeito deve ocorrer no sexto dia, assim, o número de dias que antecedem ao defeito é $x = 5$ dias.

Considerando ainda, o mesmo exemplo, qual seria o intervalo de tempo ideal para uma manutenção preventiva, se desejamos uma probabilidade de, pelo menos, 0,90 de que o defeito não ocorrerá?

Para isso precisamos saber quantos dias são necessários para acumular uma probabilidade de ocorrer defeito próxima de 0,10 (que é análogo a acumular uma probabilidade de não ocorrer defeito próxima de 0,90).

Tabela 3.3. Probabilidades antes de ocorrer defeito

Tempo de espera (dias) para ocorrer defeito	Probabilidade	Prob. acumulada de **Ocorrer defeito**	Prob. acumulada de **Não ocorrer defeito**
$\boldsymbol{x}$	$\boldsymbol{P(X = x)}$	$\boldsymbol{P(X \leq x)}$	$\boldsymbol{1 - P(X \leq x)}$
0	0,0100	0,0100	0,9900
1	0,0099	0,0199	0,9801
2	0,0098	0,0297	0,9703
3	0,0097	0,0394	0,9606
4	0,0096	0,0490	0,9510
5	0,0095	0,0585	0,9415
6	0,0094	0,0679	0,9321
7	0,0093	0,0773	0,9228
8	0,0092	0,0865	0,9135
9	0,0091	**0,0956**	**0,9044**
10	0,0090	0,1047	0,8953

Observe que obtivemos $P(X \leq 9) = 0{,}0956$, ou seja, a probabilidade (acumulada) do tempo de espera antes de ocorrer defeito está próxima de $0{,}10$. Portanto, a manutenção preventiva deve ser feita à partir de 9 dias de operação.

3.5 Distribuição Hipergeométrica

Considere um conjunto com n objetos dos quais m são do tipo I e $n - m$ são do tipo II. Uma amostra é escolhida, ao acaso e sem reposição, com tamanho r $(r < n)$ e definimos X como o número de objetos com a característica I, na amostra. Nesse caso, diremos que a variável aleatória X segue o modelo Hipergeométrico de parâmetros m, n e r.

Função de Probabilidades

A função de probabilidade do modelo Hipergeométrico é dada por:

$$P(X = x) = \frac{\binom{m}{x}\binom{n-m}{r-x}}{\binom{n}{r}},$$

sendo x inteiro e tal que $\max\{0, r - (n - m)\} \leq x \leq \min\{r, m\}$. Estes limites para x garantem que situações absurdas sejam evitadas.

Notação: $X \sim Hgeo(m, n, r)$.

Exemplo

Considere que, num lote de 20 peças, existam 4 defeituosas. Selecionando-se 5 peças, sem reposição, qual seria a probabilidade de 2 defeituosas serem escolhidas?

Solução

A variável aleatória que indica o número de peças defeituosas na amostra selecionada segue o modelo Hipergeométrico. Denominando por X essa variável, temos:

$$n = 20, \quad m = 4, \quad r = 5, \quad x = 2, \quad \text{e} \quad P(X = x) = \frac{\binom{m}{x}\binom{n-m}{r-x}}{\binom{n}{r}},$$

então:

$$P(X = 2) = \frac{\binom{4}{2}\binom{20-4}{5-2}}{\binom{20}{5}} = \frac{\binom{4}{2}\binom{16}{3}}{\binom{20}{5}} = 0{,}2167.$$

Exemplo

Pequenos motores elétricos são expedidos em lotes de 50 unidades. Antes que uma remessa seja aprovada, um inspetor escolhe 5 desses motores e os inspeciona. Se nenhum dos motores inspecionados for defeituoso o lote é aprovado. Se um ou mais forem verificados defeituosos, todos os motores da remessa são inspecionados. Suponha que existam, de fato, três motores defeituosos no lote. Qual é a probabilidade de que a inspeção de 100% do lote seja necessária?

Solução

A inspeção de 100% do lote será necessária se, e somente se, $X \geq 1$. Logo,

$$P(X \geq 1) = 1 - P(X = 0)$$

Temos, que:

$$n = 50, \quad m = 3, \quad r = 5 \quad x = 0 \quad \text{e} \quad P(X = x) = \frac{\binom{m}{x}\binom{n-m}{r-x}}{\binom{n}{r}}$$

Portanto:

$$\begin{aligned} P(X \geq 1) &= 1 - P(X = 0) \\ &= 1 - \frac{\binom{3}{0}\binom{50-3}{5-0}}{\binom{50}{5}} \\ &= 1 - \frac{\binom{3}{0}\binom{47}{5}}{\binom{50}{5}} \\ &= 1 - 0{,}7240 \\ &= 0{,}2760. \end{aligned}$$

3.6 Distribuição exponencial

A variável aleatória X, que mede a "distância" entre contagens sucessivas de um processo de Poisson, com média $\lambda > 0$, é uma variável aleatória com distribuição **Exponencial** com parâmetro λ. **Notação:** $X \sim Exp(\lambda)$.

Função densidade de probabilidade

$$f(x) = \lambda e^{-\lambda x}, \qquad \text{para } 0 \leq x < \infty.$$

Cálculo de probabilidades

$$P(a \leq X \leq b) = \int_a^b \lambda e^{-\lambda x}\, dx.$$

Função de distribuição

$$F(x) = P(X \leq x) = \begin{cases} 1 - e^{-\lambda x}, & \text{se } x \geq 0 \\ 0, & \text{se } x < 0 \end{cases}$$

Demonstração

$$\int_{-\infty}^{x} \lambda e^{-\lambda t}\, dt = \int_0^x \lambda e^{-\lambda t}\, dt = \left(-e^{-\lambda t}\right)\Big|_0^x = \left(-e^{-\lambda x}\right) - \left(-e^{-\lambda(0)}\right) = -e^{-\lambda x} + e^0$$
$$= -e^{-\lambda x} + 1 = \mathbf{1 - e^{-\lambda x}}$$

Média e variância

Se $X \sim Exp(\lambda)$, então:

$$\mu = \frac{1}{\lambda} \quad \text{e} \quad \sigma^2 = \frac{1}{\lambda^2}$$

Exemplo

Uma média de cinco chamadas por hora são recebidas em um departamento de manutenção. Considerando que o tempo (X) entre as chamada é uma variável aleatória com distribuição Exponencial com $\lambda = 5$ e, começando a observação em qualquer ponto no tempo, determine a probabilidade de que a primeira chamada de serviço chegue dentro de meia hora.

Solução

A probabilidade de ocorrer uma chamada dentro de meia hora (0,5 horas) é determinada por:

$$\begin{aligned} P(X \leq 0{,}5) &= 1 - e^{-5(0{,}5)} \\ &= 1 - e^{-2{,}5} \\ &= 1 - 0{,}0821 \\ &= \mathbf{0{,}9179}. \end{aligned}$$

Exemplo

Em uma grande rede de computadores, as conexões dos usuários do sistema podem ser modeladas como um processo de Poisson, com média de 25 conexões por hora. Qual é a probabilidade de :

a) o tempo até a próxima conexão estar entre 2 e 3 minutos?

b) não haver conexão nos próximos 6 minutos, ou, analogamente, o tempo até a próxima conexão ser maior que 6 minutos?

Solução

a) $\lambda = 25$ conexões por hora, então, convertendo minutos para horas temos:

- 2 minutos $= \frac{2}{60} = 0{,}033$ horas;
- 3 minutos $= \frac{3}{60} = 0{,}05$ horas.

Logo,

$$\begin{aligned} P(0{,}033 < X < 0{,}05) &= \int_{0{,}033}^{0{,}05} 25e^{-25x}\,dx \\ &= \left(-e^{-25x}\right)\Big|_{0{,}033}^{0{,}05} \\ &= \left(-e^{-25(0{,}05)}\right) - \left(-e^{-25(0{,}033)}\right) \\ &= \mathbf{0{,}1517} \end{aligned}$$

b) $\lambda = 25$ conexões por hora, então, convertendo minutos para horas temos:

- 6 minutos $= \frac{6}{60} = 0{,}1$ horas.

Logo,

$$
\begin{aligned}
P(X > 0{,}1) &= 1 - P(X \leq 0{,}1) \\
&= 1 - \left[1 - e^{-25(0{,}1)}\right] \\
&= e^{-25(0{,}1)} \\
&= \mathbf{0{,}0821}
\end{aligned}
$$

3.7 Distribuição Normal

Dentre todas as distribuições de probabilidades, sejam discretas ou contínuas, a mais estudada e mais utilizada é a distribuição Normal. As principais razões que fazem a distribuição Normal o modelo mais importante na estatística são:

1) Muitas variáveis biométricas tendem a ter distribuição Normal. Isto ocorre principalmente quando a variável é influenciada por um grande número de fatores que atuam de modo independente e aditivo;

2) A distribuição das médias amostrais de uma variável qualquer tendem a ter distribuição Normal, mesmo que a variável em si não tenha distribuição Normal;

3) Muitos testes e modelos estatísticos têm como pressuposição a normalidade dos dados isto é, que os dados possuem distribuição Normal.

A distribuição Normal é também conhecida como distribuição Gaussiana em homenagem a Karl F. Gauss (1777-1855), brilhante matemático e físico alemão, que a desenvolveu no início do século XIX.

Função Densidade da Normal

A função densidade de probabilidade de uma variável aleatória contínua X, é dada por:

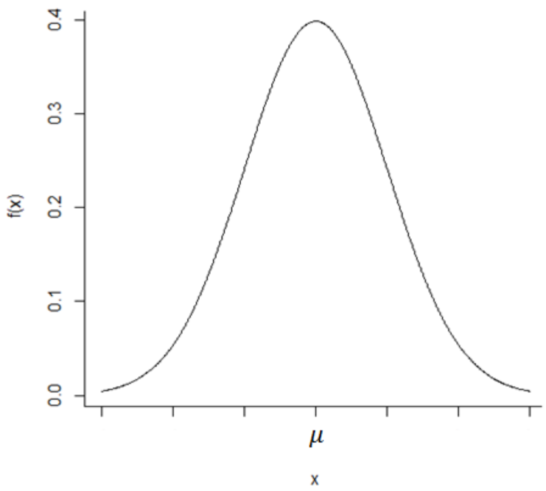
$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \qquad \text{para } -\infty < x < \infty.$$

em que:

- μ é a posição central da distribuição (média);
- σ^2 é a dispersão da distribuição (variância);
- x são os valores que a variável aleatória em estudo X assume.

Notação: $X \sim N(\mu, \sigma^2)$.

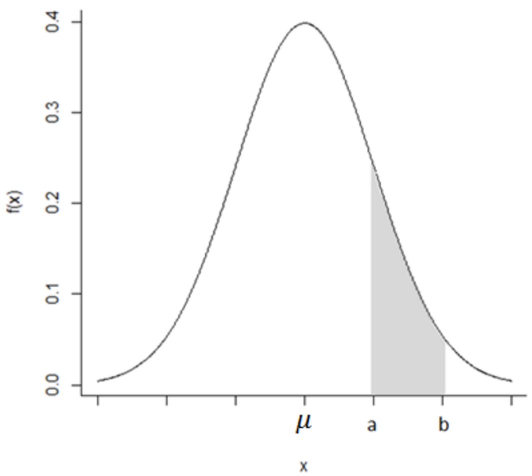
O gráfico da função densidade da distribuição Normal é apresentado a seguir:



Para se calcular a probabilidade da variável aleatória X assumir valores entre a e b basta calcular a área compreendida entre estes intervalos usando a fórmula:

$$P(a \leq X \leq b) = \int_a^b \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

cuja área é mostrada na figura a seguir.



Propriedades da Distribuição Normal

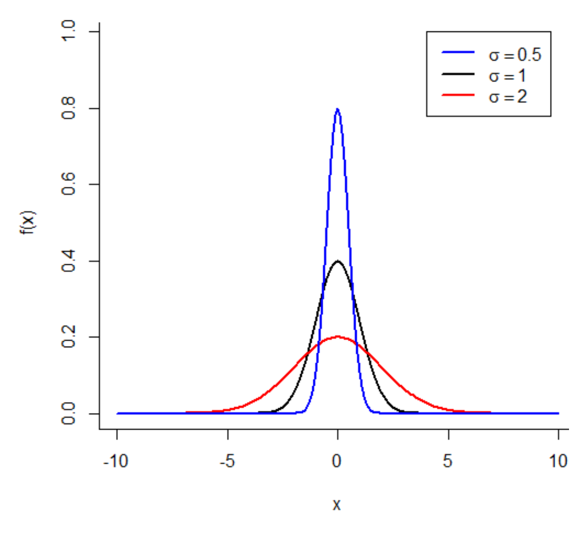
As principais características dessa função são:

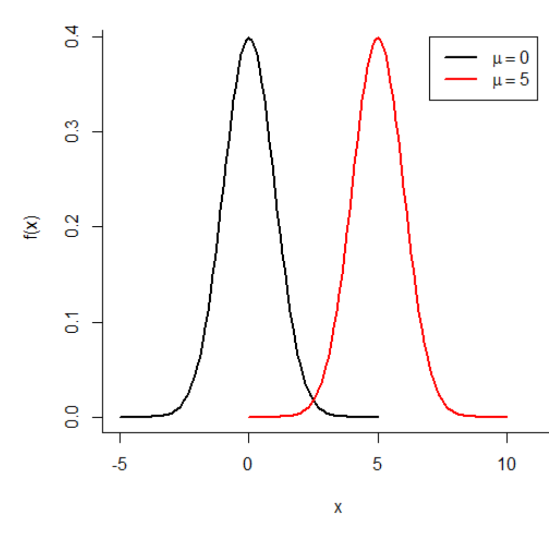
1) A função gera um gráfico em forma de sino, sendo unimodal e simétrica;

2) é definida por dois parâmetros: a média (μ) e o desvio padrão (σ), sendo que a média controla a localização do centro da distribuição (é o ponto de simetria), já o desvio padrão controla a dispersão da curva ao redor da média;

3) O ponto de máximo de $f(x)$ ocorre no ponto de abscissa $x = \mu$;

4) Não possui limite inferior ou superior;





O cálculo direto de probabilidades envolvendo a distribuição normal exige recursos do cálculo avançado e, mesmo assim, dada a forma da função densidade, não é um processo muito elementar. Por isso, elas foram tabeladas, permitindo-nos obter diretamente o valor da probabilidade desejada. Devido as dificuldades de cálculo e da dificuldade em se construir tabelas da função dependendo de dois parâmetros (μ e σ^2), recorre-se a uma mudança de variável, transformando a variável aleatória X na variável aleatória Z. Essa nova variável chama-se **variável normal padronizada**, ou reduzida.

Distribuição Normal Padrão

Denomina-se distribuição normal padrão, a distribuição normal de média zero e variância 1, ou seja:

$$Z \sim N(\mu = 0,\ \sigma^2 = 1),\ \text{ ou simplesmente, } Z \sim N(0{,}1).$$

As probabilidades associadas a distribuição normal reduzida são facilmente obtidas em tabelas.

Uso da Tabela da Distribuição Normal Padrão

Exemplo

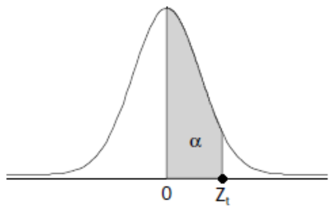
Seja $Z \sim N(0,1)$. Usando a tabela da distribuição normal padrão, calcular

$$P(0 < Z < 1{,}57).$$

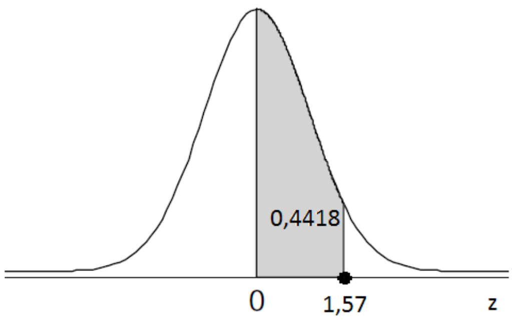
Solução

Tabela da distribuição Normal Padrão

Probabilidades (α) da distribuição Normal Padrão $N(0,1)$ para valores do quantil Z_t padronizado de acordo com o seguinte evento: $P(0 < Z < Z_t) = \alpha$.



Z_t	0	1	2	3	4	5	6	7	8	9
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0753
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
⋮										
1,1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1,2	0,3849	0,3869	0,3888	0,3907	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
1,3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
1,4	0,4192	0,4207	0,4222	0,4236	0,4251	0,4265	0,4279	0,4292	0,4306	0,4319
1,5	0,4332	0,4345	0,4357	0,4370	0,4382	0,4394	0,4406	0,4418	0,4429	0,4441
1,6	0,4452	0,4463	0,4474	0,4484	0,4495	0,4505	0,4515	0,4525	0,4535	0,4545



Portanto,

$$P(0 < Z < 1{,}57) = 0{,}4418.$$

Exercício

Seja $Z \sim N(0,1)$. Usando a tabela da distribuição normal padrão, calcular:

a) $P(Z > 0)$

b) $P(Z < 0)$

c) $P(0 < Z < 1{,}08)$

d) $P(-1{,}89 < Z < 0)$

e) $P(-1{,}23 < Z < 1{,}05)$

f) $P(Z < 1)$

g) $P(Z > 1)$

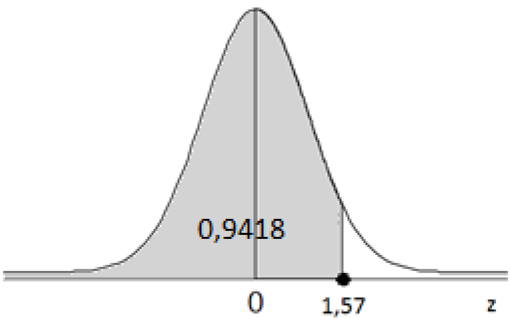
h) $P(0{,}5 < Z < 1)$

Distribuição Normal usando o software R

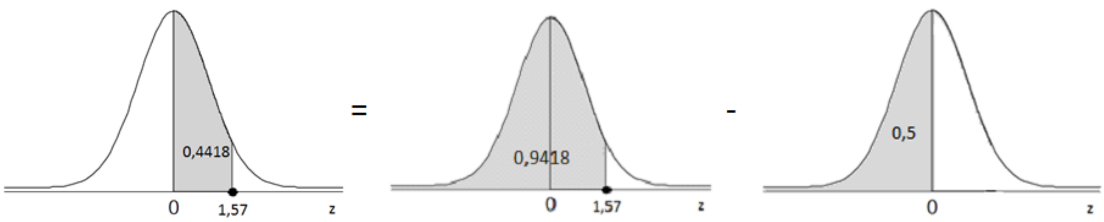
Na tabela que utilizamos acima as probabilidades são calculadas de 0 até Z_t, ou seja, $P(0 < Z < Z_t)$. Porém, no software R as probabilidades são calculadas de $-\infty$ até Z_t, ou seja, $P(-\infty < Z < Z_t)$. Assim, quando usamos o comando:

```
pnorm(1.57, mean=0, sd=1)
```

o R calcula $P(-\infty < Z < 1{,}57)$:



Logo, para obter $P(0 < Z < 1{,}57)$, precisamos subtrair $P(-\infty < Z < 0)$ de $P(-\infty < Z < 1{,}57)$:



$P(0 < Z < 1{,}57) = P(-\infty < Z < 1{,}57) - P(-\infty < Z < 0)$

`= pnorm(1.57,mean=0,sd=1) - pnorm(0,mean=0,sd=1)`

Comandos no Software R para calcular as probabilidades da distribuição Normal:

```
pnorm(1.57, mean=0, sd=1)                               #P(-Inf < Z <1.57)
pnorm(1.57, mean=0, sd=1) - pnorm(0,mean=0,sd=1)    #P(0 < Z <1.57)
```

Padronização de uma variável X com distribuição Normal

Os problemas da vida real, entretanto, não se apresentam na forma reduzida da variável Z, ao contrário, são formulados em termos da variável normal original X , com média μ e desvio-padrão σ. É preciso então, antes de passarmos à sua resolução, padronizar ou reduzir a variável aleatória normal X, transformando-a na variável aleatória Z:

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

Exemplo 1

Seja $X \sim N(4, 1)$. Determine $P(X \leq 4)$.

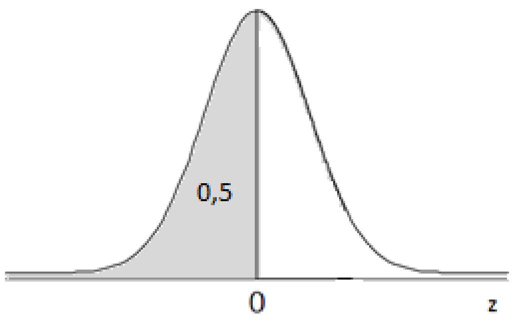
Solução

Como $X \sim N(4,\ 1)$, então:

$$\mu = 4 \quad \text{e} \quad \sigma = \sqrt{\sigma^2} = \sqrt{1} = 1$$

Logo:

$$P(X \leq 4) = P\left(\frac{X - \mu}{\sigma} \leq \frac{4 - \mu}{\sigma}\right) = P\left(Z \leq \frac{4 - \mu}{\sigma}\right) = P\left(Z \leq \frac{4 - 4}{1}\right)$$
$$= P(Z \leq 0) = 0{,}5$$

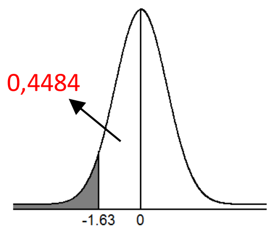


Comandos no Software R para calcular as probabilidades da distribuição Normal:

```
pnorm(4, mean=4, sd=1)  #Usando a variável X~N(4, 1)
```

Exemplo 2

A duração da gravidez humana, da concepção ao parto, varia segundo uma distribuição aproximadamente normal com média 266 dias e desvio padrão de 16 dias. Qual a probabilidade de uma gravidez durar menos de 240 dias?



$$
\begin{aligned}
P(X < 240) &= P\left(\frac{X-\mu}{\sigma} < \frac{240-\mu}{\sigma}\right)\\
&= P\left(Z < \frac{240-\mu}{\sigma}\right)\\
&= P\left(Z < \frac{240-266}{16}\right)\\
&= P(Z < -1{,}63)\\
&= 0{,}5 - 0{,}4484\\
&= \mathbf{0{,}0516}
\end{aligned}
$$

Exercício 1

Seja $X \sim N(4, 1)$. Determine:

a) $P(4 < X < 5)$

b) $P(2 < X < 5)$

c) $P(5 < X < 7)$

d) $P(X \leq 1)$

e) $P(0 \leq X \leq 2)$

Exercício 2

Seja $X \sim N(3, 16)$, ou seja, a variável X tem distribuição Normal com média $\mu = 3$ e variância $\sigma^2 = 16$. Faça o gráfico da distribuição e determine $P(3 < X < 8)$.

Exercício 3

A estatura média dos alunos da UFRRJ é $\mu = 1{,}75m$ e desvio padrão $\sigma = 0{,}15m$. Assumindo-se que a variável estatura (X) seja normalmente distribuída, calcule a probabilidade de um aluno aleatoriamente selecionado ter estatura entre $1{,}70m$ e $1{,}80m$.

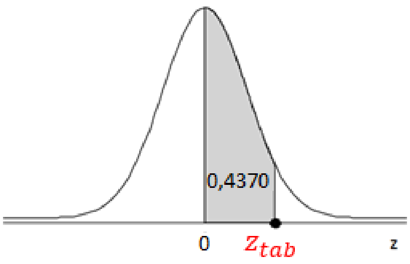
Fixada a probabilidade determinar o quantil (z_{tab}) da distribuição Normal

Exemplo

Sabendo-se que $Z \sim N(0,1)$ e usando a tabela da Distribuição Normal Padrão, obter z_{tab} tal que $P(0 < Z < z_{tab}) = 0{,}4370$.

Solução

Foi dado $P(0 < Z < z_{tab}) = 0{,}4370$, ou seja, temos a probabilidade e queremos encontrar o valor de z_{tab} tal que a área do gráfico, compreendida entre 0 e z_{tab}, é igual a $0{,}4370$. Graficamente, temos:



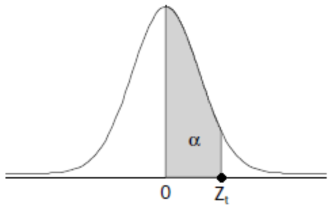
Observe que este é exatamente o tipo de probabilidade encontrada na tabela da Distribuição Normal Padrão, ou seja, $P(0 < Z < Z_{tab})$. Então, devemos localizar a probabilidade $0{,}4370$ no interior da tabela da Distribuição Normal Padrão:

1,1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1,2	0,3849	0,3869	0,3888	0,3907	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
1,3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
1,4	0,4192	0,4207	0,4222	0,4236	0,4251	0,4265	0,4279	0,4292	0,4306	0,4319
1,5	0,4332	0,4345	0,4357	$\boxed{0{,}4370}$	0,4382	0,4394	0,4406	0,4418	0,4429	0,4441
1,6	0,4452	0,4463	0,4474	0,4484	0,4495	0,4505	0,4515	0,4525	0,4535	0,4545

e depois, obtemos, nas bordas da tabela, o valor de z_{tab} correspondente a esta probabilidade:

Tabela da distribuição Normal Padrão

Probabilidades (α) da distribuição Normal Padrão $N(0,1)$ para valores do quantil Z_t padronizado de acordo com o seguinte evento: $P(0 < Z < Z_t) = \alpha$.

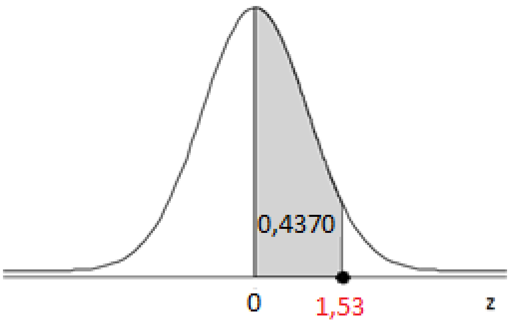


Z_t	0	1	2	3	4	5	6	7	8	9
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0753
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0,4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
0,5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0,6	0,2257	0,2291	0,2324	0,2357	0,2389	0,2422	0,2454	0,2486	0,2517	0,2549
0,7	0,2580	0,2611	0,2642	0,2673	0,2704	0,2734	0,2764	0,2794	0,2823	0,2852
0,8	0,2881	0,2910	0,2939	0,2967	0,2995	0,3023	0,3051	0,3078	0,3106	0,3133
0,9	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315	0,3340	0,3365	0,3389
1,0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621
1,1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1,2	0,3849	0,3869	0,3888	0,3907	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
1,3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
1,4	0,4192	0,4207	0,4222	0,4236	0,4251	0,4265	0,4279	0,4292	0,4306	0,4319
1,5	0,4332	0,4345	0,4357	0,4370	0,4382	0,4394	0,4406	0,4418	0,4429	0,4441
1,6	0,4452	0,4463	0,4474	0,4484	0,4495	0,4505	0,4515	0,4525	0,4535	0,4545

Portanto, o valor de z tal que $P(0 < Z < z_{tab}) = 0{,}4370$ é $z_{tab} = 1{,}53$, ou seja:

$$P(0 < Z < 1{,}53) = 0{,}4370.$$

Graficamente, temos:

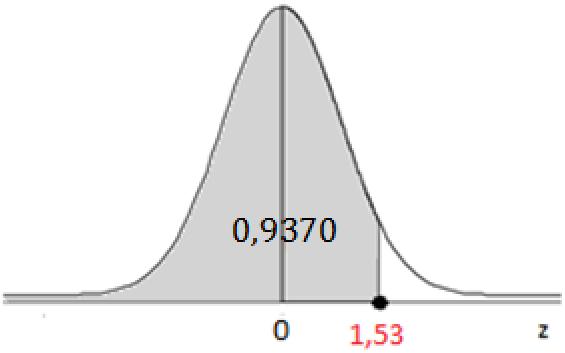


Comandos no Software R para calcular os quantis da distribuição Normal:

```
qnorm(0.9370, mean=0, sd=1)
```

Observação: Note que usamos 0,9370 ao invés de 0,4370 porquê o R considera a probabilidade de $-\infty$ até z_{tab} (e não de 0 até z_{tab}), ou seja, no R usamos:

$$P(Z < z_{tab}) = P(Z < 0) + P(0 < Z < z_{tab}) = 0{,}5 + 0{,}4370 = 0{,}9370$$



Exercício 1

Sabendo-se que $Z \sim N(0, 1)$ e usando a tabela da Distribuição Normal Padrão, obter z tal que:

a) $P(0 < Z < z_{tab}) = 0{,}475$

b) $P(z_{tab} < Z < 0) = 0{,}3531$

c) $P(z_{tab} < Z < 0) = 0{,}4949$

d) $P(-z_{tab} < Z < z_{tab}) = 0{,}95$

e) $P(-z_{tab} < Z < z_{tab}) = 0{,}97$

f) $P(Z < z_{tab}) = 0{,}8212$

g) $P(Z < z_{tab}) = 0{,}3023$

h) $P(Z > z_{tab}) = 0{,}9525$

i) $P(Z > z_{tab}) = 0{,}0749$

j) $P(Z < z_{tab}) = 0{,}5$

Observação: Se acaso a tabela não contiver a probabilidade desejada, utilize a probabilidade mais próxima.

III - Inferência Estatística

1 Introdução à inferência estatística.

1.1 Conceitos básicos. Amostra e população.

Até o momento vimos a *Estatística Descritiva*, que mostra como relatar os dados que temos em mãos. A interpretação do material coletado é feita por meio de tabelas e gráficos e da apresentação de estatísticas como média e desvio-padrão. Por exemplo, se você medir o peso e a altura de 100 estudantes da UFRRJ, saberá apresentar e resumir os dados, ou seja, descrever os resultados que você encontrou nesse grupo de estudantes.

É possível generalizar os resultados obtidos à partir das observações feitas nesses estudantes (uma amostra) para todos os estudantes da UFRRJ (a população). Mas para isso é preciso usar um conjunto de técnicas de Estatística que permitem, com base em uma amostra, fazer *inferência* para a população de onde foi retirada.

População e Amostra

População: é o conjunto de unidades sobre o qual desejamos informação.

Amostra: é qualquer subconjunto de unidades retiradas da população para obter a informação desejada.

A chave para o bom entendimento da Estatística é saber distinguir entre os dados observados (amostra) e a vasta quantidade de dados que poderiam ter sido observados (população). O uso de amostras permite obter respostas para a questão estudada, com *margens de erro* conhecidas.

O termo *população* não se restringe, porém a um conjunto de pessoas, referindo-se, sim, a qualquer conjunto "grande" de unidades que têm algo em comum, como, por exemplo:

- radiografias feitas pelos alunos de uma faculdade em determinado curso;
- prontuários de pacientes atendidos pelo SUS durante todo o ano;

- auditorias das contas hospitalares de uma maternidade;
- certidões de óbito registradas numa cidade em determinado período.

Parâmetro, estimador e estimativa

Parâmetro: é um valor em geral desconhecido (e, portanto, que precisa ser estimado) que representa determinada característica da população.

São exemplos de parâmetros a média populacional μ, a variância populacional σ^2 e o desvio padrão populacional σ.

Estimador: é uma função dos elementos da amostra. É usado para estimar o parâmetro correspondente, da população de onde a amostra foi retirada.

Assim, por exemplo, a média amostral $\bar{X}$ é um estimador da média populacional μ, a variância amostral S^2 é um estimador da variância populacional σ^2 e o desvio padrão amostral S é um estimador do desvio padrão populacional σ.

Estimativa: O valor numérico de um estimador é conhecido como estimativa. Assim $\bar{x} = 17{,}8$ é uma estimativa da média populacional μ.

Algumas razões para se tomar uma amostra ao invés de usar a população toda são as seguintes:

- custo alto para obter informação da população toda;
- tempo muito longo para obter informação da população toda;
- algumas vezes é impossível, por exemplo, estudo de poluição atmosférica;
- algumas vezes é logicamente impossível, por exemplo, em ensaios destrutivos.

1.2 Amostragem aleatória simples: obtenção de uma amostra aleatória

É compreensível que o estudo de todos os elementos da população possibilita preciso conhecimento das variáveis que estão sendo pesquisadas; todavia, nem sempre é possível obter as informações de todos os elementos da população. Torna-se claro que a representatividade da amostra dependerá de seu tamanho (quanto maior, melhor) e de outras considerações de ordem metodológica. Isto é, o pesquisador procurará acercar-se de cuidados, visando a obtenção de uma amostra que de fato represente "o melhor possível" toda a população.

A amostragem aleatória simples é o processo mais elementar e frequentemente utilizado. Atribui-se a cada elemento da população um número distinto de 1 a N (tamanho da população). Efetuam-se sucessivos sorteios até completar-se o tamanho da amostra: n.

Exemplo

Um dentista quer obter uma amostra de 2% dos quinhentos pacientes de sua clínica para entrevistá-los sobre a qualidade de atendimento da secretária. Para obter uma amostra aleatória de 2% dos quinhentos pacientes, é preciso sortear 10 pacientes.

Isso pode ser feito de maneira mais antiga e mais conhecida (e também mais trabalhosa): atribuem-se números de 1 a 500 a cada um dos pacientes e escrevem-se esses números em pedaços de papel. Coloca-se todos os pedaços de papel em uma caixa, misturando-os bem, e retira-se um papel. O procedimento é repetido até serem retirados 10 papéis referentes a 10 pacientes.

O procedimento pode ser feito mais facilmente utilizando-se um gerador de números aleatórios. Na calculadora isso pode ser feito utilizando a tecla #Ran (em laranja em cima da tecla do ponto "."). Para sortear 10 números entre 1 e 500 digitamos:

500 Shift #Ran =

Para sortear outro número basta teclar " = " novamente. Considera-se apenas a parte inteira do número. Se um número repetir ignore e sorteie outro.

1.3 Conceito de Distribuições amostrais

No início do curso vimos medidas que caracterizam uma amostra, como, por exemplo, a média, a variância, o desvio padrão, ... Nas seções anteriores vimos os principais modelos de distribuição de probabilidade, particularmente o modelo da distribuição Normal.

Neste capítulo juntam-se os modelos de distribuição de probabilidade e as medidas descritivas obtendo-se as distribuições amostrais dos principais estimadores. O conceito de distribuição amostral é fundamental na inferência estatística.

Considere todas as possíveis amostras de tamanho n que podem ser extraídas de determinada população. Se para cada uma delas se calcular um valor do estimador (por exemplo da média amostral $\bar{X}$), tem-se uma distribuição amostral desse estimador. Como o estimador é uma variável aleatória, pode-se determinar suas características, isto é, encontrar sua média, variância, desvio-padrão...

1.4 Distribuição amostral da média

A distribuição amostral da média refere-se à distribuição gerada pelas estimativas da média, obtidas a partir de todas as amostras, de tamanho n, retiradas de uma população de referência.

Exemplo

Considere uma população fictícia, de tamanho $N = 3$, $X = \{1, 2, 3\}$, cuja média é $\mu = 2$ e a variância é $\sigma^2 = 2/3$. Obter a distribuição amostral da média $\bar{X}$ para amostras de tamanho $n = 2$ retiradas com reposição dessa população.

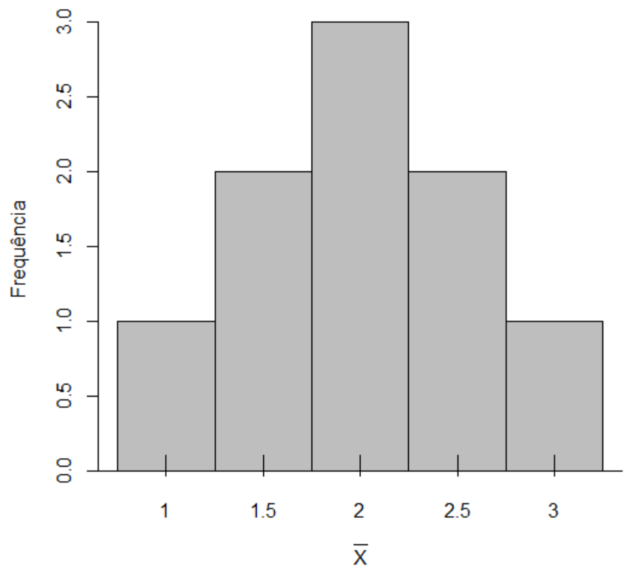
As amostras *com reposição* de tamanho $n = 2$ juntamente com a média amostral são apresentadas a seguir:

Amostras com reposição	$\bar{X}$
$(1, 1)$	1,0
$(1, 2)$	1,5
$(1, 3)$	2,0
$(2, 1)$	1,5
$(2, 2)$	2,0
$(2, 3)$	2,5
$(3, 1)$	2,0
$(3, 2)$	2,5
$(3, 3)$	3,0

Agrupando as médias comuns e computando suas frequências, obtêm-se os seguintes resultados.

$\bar{X}$	Frequência
1,0	1
1,5	2
2,0	3
2,5	2
3,0	1

O gráfico da distribuição amostral das médias da amostra é apresentado a seguir:



Calculando-se, agora a média e a variância de $\bar{X}$ para todas as 9 médias amostrais (população de médias amostrais) obtidas, têm-se:

$$\mu_{\bar{X}} = \frac{\sum \bar{X}_i}{9} = \frac{1 + 1{,}5 + 1{,}5 + \cdots + 3{,}0}{9} = 2$$

e

$$\sigma_{\bar{X}}^2 = \frac{1}{9}\left[\sum \bar{X}_i^2 - \frac{(\sum \bar{X}_i)^2}{9}\right] = \frac{1}{9}\left[(1^2 + 1{,}5^2 + \cdots 3^2) - \frac{18^2}{9}\right] = \frac{1}{3}$$

Teorema

Seja X uma variável aleatória com média μ e variância σ^2, e seja $(X_1, X_2, \dots, X_n)$ uma amostra aleatória de X. Seja $\bar{X} = (X_1 + X_2 + \cdots + X_n)/n$ a média amostral, então:

$$\mu_{\bar{X}} = E(\bar{X}) = \mu \qquad \text{e} \qquad \sigma_{\bar{X}}^2 = Var(\bar{X}) = \frac{\sigma^2}{n}.$$

Podemos verificar numericamente este resultado a partir do exemplo anterior:

- Médias:
  - $\mu = 2$ (média populacional)
  - $\mu_{\bar{X}} = 2$ (média das médias amostrais)

- Variâncias:
  - $\sigma^2 = \frac{2}{3}$ (variância populacional)
  - $\sigma^2_{\bar{X}} = \frac{1}{3}$ (variância das médias amostrais)

Temos que:

$$\frac{\sigma^2}{n} = \frac{\left(\frac{2}{3}\right)}{2} = \frac{1}{3} = \sigma^2_{\bar{X}}$$

ou seja, verificamos numericamente que $\mu_{\bar{X}} = \mu$ e $\sigma^2_{\bar{X}} = \frac{\sigma^2}{n}$.

Teorema

Seja $X \sim N(\mu, \sigma^2)$, ou seja, X é uma variável aleatória com distribuição Normal, com média μ e variância σ^2, então:

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right),$$

ou seja, a média amostral $(\bar{X})$ tem distribuição Normal com média μ e variância σ^2/n.

Corolário

Seja $X \sim N(\mu, \sigma^2)$, então:

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1).$$

Exemplo

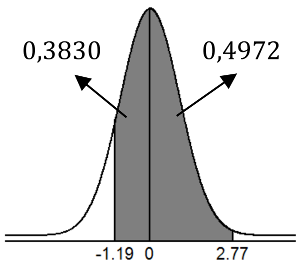
A duração da gravidez humana, da concepção ao parto, varia segundo uma distribuição aproximadamente normal com média $\mu = 266$ dias e desvio padrão de $\sigma = 16$ dias. Qual a probabilidade de em uma amostra de $n = 10$ mulheres grávidas a duração média da gravidez durar entre 260 e 280 dias?

Solução

$$P(260 < \bar{X} < 280) = P\left(\frac{260-\mu}{\sigma/\sqrt{n}} < \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} < \frac{280-\mu}{\sigma/\sqrt{n}}\right) = P\left(\frac{260-\mu}{\sigma/\sqrt{n}} < Z < \frac{280-\mu}{\sigma/\sqrt{n}}\right)$$

$$= P\left(\frac{260-266}{16/\sqrt{10}} < Z < \frac{280-266}{16/\sqrt{10}}\right) = P(-1{,}19 < Z < 2{,}77)$$

$$= 0{,}3830 + 0{,}4972 = \mathbf{0{,}8802}$$



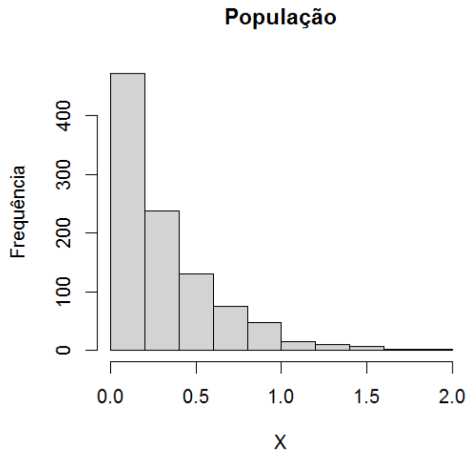
Teorema Central do Limite

Seja $(X_1, X_2, \ldots, X_n)$ uma amostra de variáveis aleatórias com *distribuição de probabilidades qualquer* (não precisa ser normal), com média μ e variância σ^2. A distribuição amostral da média $\bar{X}$ **aproxima-se, para n grande** $(n \geq 30)$, de uma distribuição normal, com média μ e variância σ^2/n, ou seja:

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right).$$

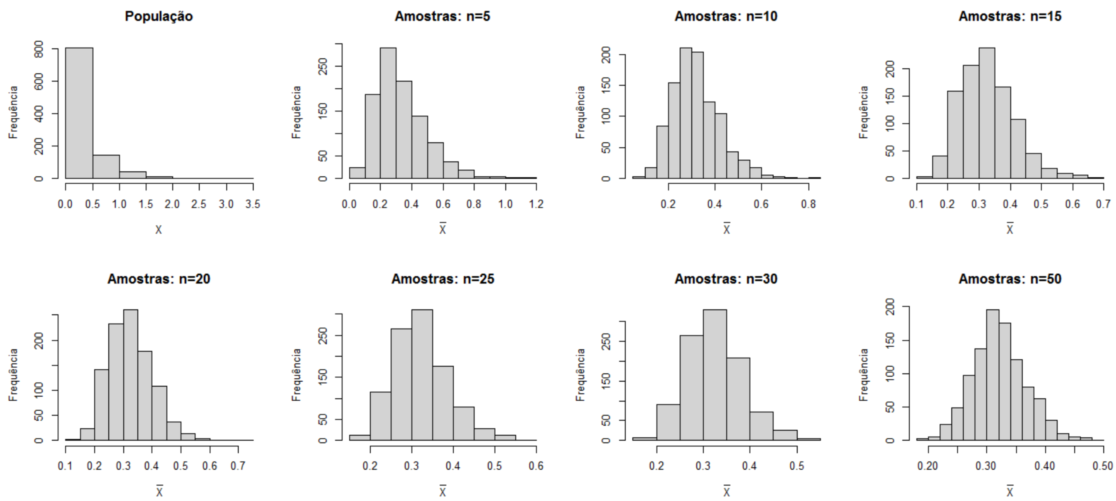
Exemplo

Para ilustrar o Teorema Central do Limite considere uma população com distribuição Exponencial, cujo histograma é apresentado a seguir:



- São extraídas, desta população, 1000 amostras aleatórias de tamanho $n = 5$, outras 1000 amostras aleatórias de tamanho $n = 10$, ..., $n = 15,\ 20,\ 30,\ 40$ e 50;
- Para cada uma das amostras foi calculada a média amostral $\bar{X}$;
- Foi feito um histograma para o conjunto de todas as médias obtidas de amostras de tamanho $n = 5$; também foi feito um histograma para o conjunto de todas as médias obtidas de amostras de tamanho $n = 10$, ..., também foi feito um histograma para o conjunto de todas as médias obtidas de amostras de tamanho $n = 50$.

São apresentados, a seguir, o histograma dos dados da população (X) e os histogramas das médias amostrais $(\bar{X})$ para as médias obtidas de amostras de tamanho $n = 5, 10, 15, 20, 30$ e 50:



Observe que, mesmo a população (X) não tendo distribuição Normal (histograma muito assimétrico), para amostras de tamanho $n = 30$ ou maior, a distribuição da média amostral $(\bar{X})$ tem uma distribuição aproximadamente Normal (histograma quase simétrico).

Corolário (do Teorema Central do Limite)

Seja $(X_1, X_2, \dots, X_n)$ uma amostra de variáveis aleatórias com *distribuição de probabilidades qualquer* (não precisa ser normal), com média μ e variância σ^2, e $\bar{X} = (X_1 + X_2 + \cdots + X_n)/n$, então, **para n grande** $(n \geq 30)$:

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1).$$

Utilizando o Teorema Central do Limite podemos calcular probabilidades relacionadas à média amostral independente de conhecermos a distribuição de probabilidades da população da qual a amostra é proveniente.

Exemplo

Suponha que numa população de camarões machos adultos a média dos comprimentos seja igual a $\mu = 27{,}3\ mm$ e o desvio padrão seja $\sigma = 7{,}8\ mm$. Qual a probabilidade de que numa amostra de $n = 35$ camarões, obtenha-se uma média menor que 22 milímetros (ou seja, $\bar{X} < 22\ mm$)?

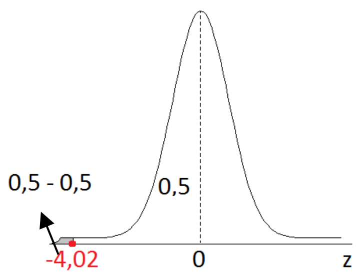
Solução

Deseja-se obter $P(\bar{X} < 22)$.

Note que nada foi dito sobre a distribuição de probabilidades dos comprimentos dos camarões, mas como $n > 30$ (n grande) podemos utilizar a distribuição Normal para calcular probabilidades sobre a média amostral $(\bar{X})$. Então:

$$P(\bar{X} < 22) = P\left(\frac{\bar{X} - \mu}{(\sigma/\sqrt{n}\,)} < \frac{22 - \mu}{(\sigma/\sqrt{n}\,)}\right) = P\left(Z < \frac{22 - \mu}{(\sigma/\sqrt{n}\,)}\right) = P\left(Z < \frac{22 - 27{,}3}{(7{,}8/\sqrt{35}\,)}\right)$$

$$= P(Z < -4{,}02) = 0{,}5 - 0{,}5 = 0$$



Observação: Na tabela da distribuição Normal Padrão, para $z = 3{,}9$ temos que $P(0 < Z < 3{,}9) = 0{,}5$, ou seja, $z = 3{,}9$ é um valor tão distante nesta distribuição que a área entre 0 e $3{,}9$ já ocupa toda a metade da área do gráfico. Então, para qualquer valor "z_{maior}" maior do que 3,9, temos $P(0 < Z < z_{maior}) = 0{,}5$.

2 Estimação

2.1 Conceitos básicos. Estimadores não viciados

Vimos que a Inferência Estatística tem por objetivo fazer generalizações sobre uma população, com base nos dados de uma amostra. Salientamos que dois problemas básicos nesse processo são:

a) Estimação de parâmetros;

b) Teste de hipóteses sobre parâmetros.

Lembremos que parâmetros são funções de valores populacionais, por exemplo a média populacional μ, enquanto estatísticas são funções de valores amostrais, por exemplo a média amostral $\bar{X}$.

Estimação Pontual

É usada quando a partir da amostra procura-se obter um único valor para "tentar adivinhar" o verdadeiro valor de certo parâmetro populacional, ou seja, obter estimativas a partir dos valores amostrais.

Principais estimadores:

- Média Amostral $(\bar{X})$;

- Proporção Amostral ($\hat{p}$);
- Variância Amostral (S^2).

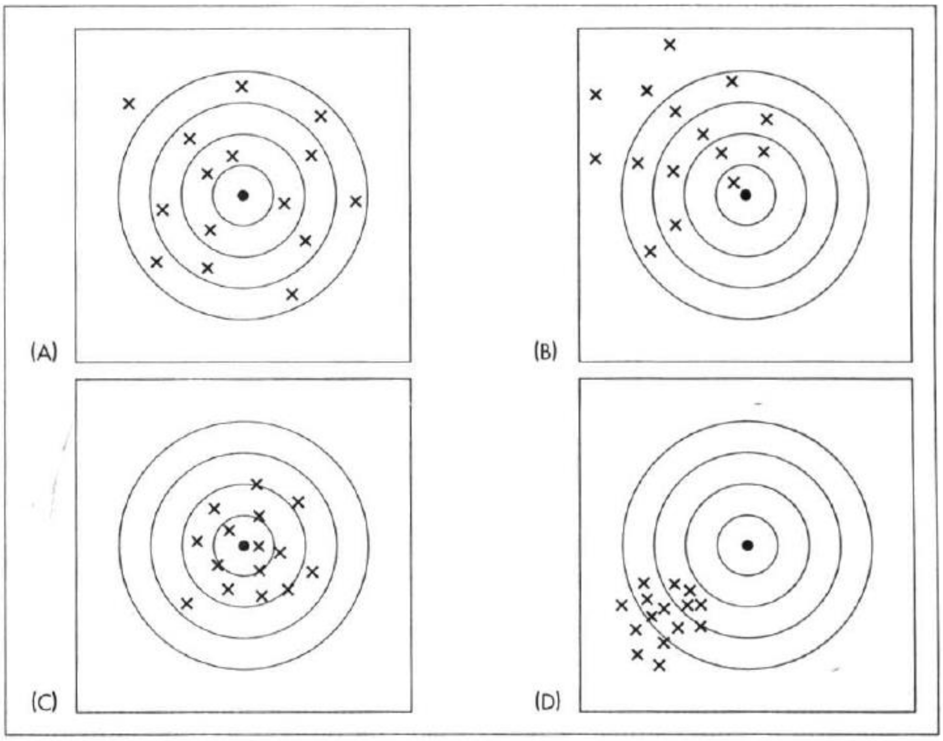
Estimadores não viciados

Definição

O estimador amostral $\hat{\theta}$ do parâmetro populacional θ é não viciado (ou não viesado) para o parâmetro θ se $E(\hat{\theta}) = \theta$.

Exemplo

Desejamos comprar um rifle e, após algumas seleções, restaram quatro alternativas, que chamaremos de rifles A, B, C e D. Foi feito um teste com cada rifle, que constituiu em fixá-lo num cavalete, mirar o centro do alvo e disparar 15 tiros. Os resultados estão ilustrados na figura abaixo:



Para analisar qual a melhor arma, podemos fixar critérios. Por exemplo, segundo o critério de "em média acertar o alvo" (não viciado), escolheríamos as armas A e C.

Segundo o critério de "não ser muito dispersivo" (variância pequena), a escolha recairia nas armas C e D. Note que a arma C é aquela que reúne as duas propriedades e, segundo esses critérios, seria a melhor arma.

Exemplo

Considere uma população com N elementos e a média populacional:

$$\mu = \frac{\sum X_i}{N}.$$

Sejam $X_1, X_2, \dots, X_n$ uma amostra aleatória de tamanho n desta população. A média amostral:

$$\bar{X} = \frac{\sum X_i}{n}$$

é um estimador **não viciado** da média populacional μ pois $E(\bar{X}) = \mu$.

Demonstração

Cada X_i tem média $E(X_i) = \mu$, pois $X_1, X_2, \dots, X_n$ são uma amostra aleatória de uma população com média μ. Então:

$$E(\bar{X}) = E\left(\frac{\sum X_i}{n}\right) = \frac{1}{n}E\left(\sum X_i\right) = \frac{1}{n}\sum [E(X_i)] = \frac{1}{n}\sum_{i=1}^{n}\mu = \frac{1}{n}\left(\underbrace{\mu + \mu + \cdots + \mu}_{n \text{ vezes}}\right)$$

$$= \frac{1}{n}.n\mu = \mu$$

Exemplo

Considere uma população com N elementos e a variância populacional

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{n}(X_i - \mu)^2.$$

Um possível estimador para σ^2, baseado numa amostra aleatória simples de tamanho n extraída dessa população, pode ser:

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2.$$

Pode-se mostrar, de modo semelhante ao do exemplo anterior (porém mais trabalhoso), que:

$$E(\hat{\sigma}^2) = \left(\frac{n-1}{n}\right)\sigma^2,$$

ou seja:

$$E(\hat{\sigma}^2) \neq \sigma^2.$$

Portanto, $\hat{\sigma}^2$ é um estimador **viciado** para σ^2.

Estimador não viciado para σ^2

Observe que, se multiplicarmos $\hat{\sigma}^2$ por $\left(\frac{n}{n-1}\right)$ teremos:

$$E\left(\frac{n}{n-1}\hat{\sigma}^2\right) = \frac{n}{n-1}E(\hat{\sigma}^2) = \frac{n}{n-1}\left(\frac{n-1}{n}\right)\sigma^2 = \sigma^2$$

Então, podemos obter um estimador não viciado para a variância populacional como:

$$S^2 = \frac{n}{n-1}.\hat{\sigma}^2 = \frac{n}{n-1}.\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2.$$

Observação

O motivo pelo qual dividimos por $(n-1)$ na variância amostral ao invés de dividir por n é para que o estimador da variância amostral seja não viciado, ou seja:

$$E(S^2) = E\left[\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2\right] = \sigma^2.$$

Assim, se retirarmos várias amostras da população e calcularmos a variância amostral (S^2) de cada uma das amostras, em média, estaremos "acertando" a verdadeira variância da população (σ^2).

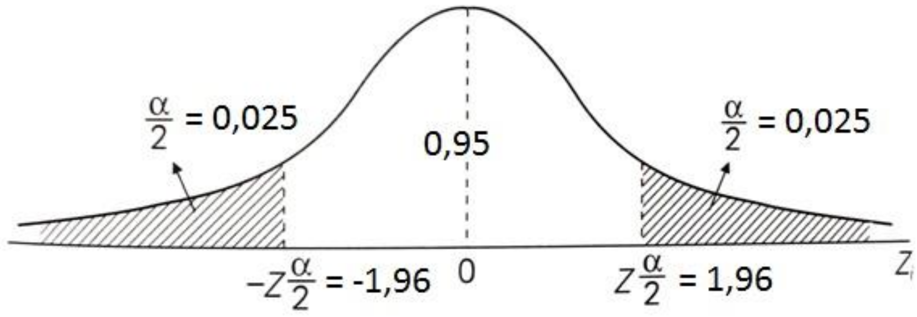
2.2 Intervalo de confiança para média de uma população Normal com variância populacional conhecida.

Os estimadores pontuais especificam um único valor para o parâmetro. Esse procedimento não permite julgar qual a possível magnitude do erro que se está cometendo. Assim, surge a idéia de construir intervalos de confiança, que serão baseados na distribuição amostral do estimador pontual.

Foi visto anteriormente que se $X \sim N(\mu, \sigma^2)$, então $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$. Assim, a variável padronizada para $\bar{X}$ será:

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}.$$

Foi visto também $P(-1{,}96 \leq Z \leq 1{,}96) = 0{,}95$.



Substituindo Z por $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$, temos:

$$P\left(-1{,}96 \leq \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \leq 1{,}96\right) = 0{,}95.$$

Resolvendo-se as inequações para μ, temos:

$$P\left(\bar{X} - 1{,}96.\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + 1{,}96.\frac{\sigma}{\sqrt{n}}\right) = 0{,}95.$$

Assim, o intervalo de confiança de 95% para a média populacional (μ) é dado por:

$$IC_{(0{,}95)}(\mu) = \left[\bar{x} - 1{,}96.\frac{\sigma}{\sqrt{n}} \;\; ; \;\; \bar{x} + 1{,}96.\frac{\sigma}{\sqrt{n}}\right].$$

Interpretação

Com 95% de confiança (ou probabilidade) o intervalo

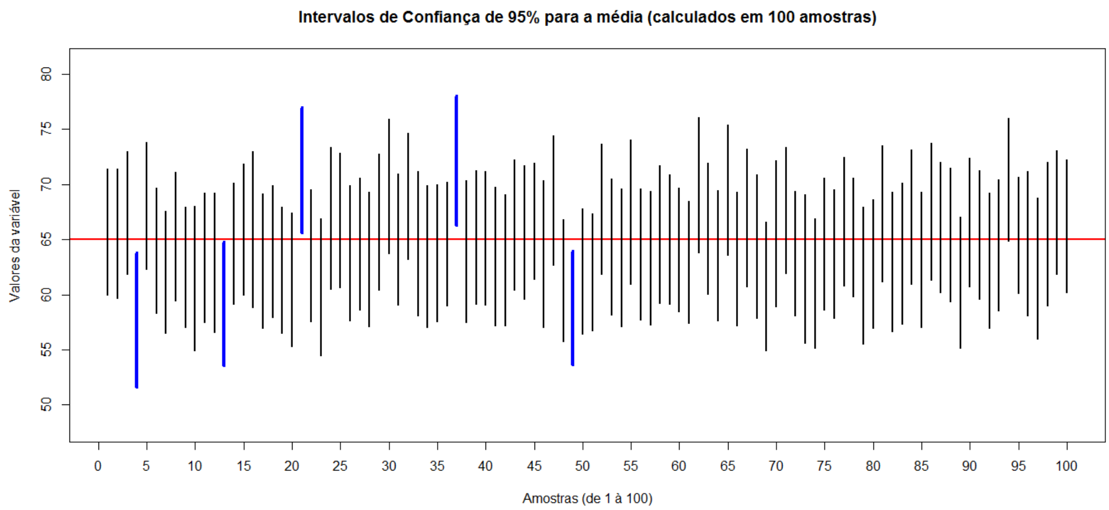
$$\left[\bar{x} - 1{,}96.\frac{\sigma}{\sqrt{n}} \;\; ; \;\; \bar{x} + 1{,}96.\frac{\sigma}{\sqrt{n}}\right]$$

contém o verdadeiro valor da média populacional μ.

Exemplo

Suponhamos uma população cujo verdadeiro valor da média é $\mu = 65$ (lembre-se que na prática μ é desconhecido). Se retirarmos uma amostra dessa população e, a partir dessa amostra, calcularmos o intervalo de confiança de 95% para a média populacional, este intervalo tem 95% de chances (probabilidade) de conter a verdadeira média populacional $\mu = 65$.

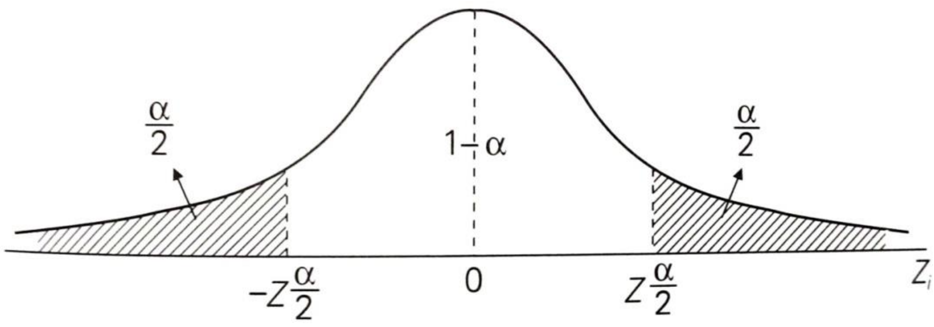
Dessa forma, se retirarmos 100 amostras desta população e, para cada amostra, obtermos o respectivo intervalo de confiança de 95% para a média, espera-se que aproximadamente 95 desses intervalos contenham a verdadeira média populacional $\mu = 65$. Observe na figura abaixo que para 100 amostras retiradas desta população, 95 dos 100 intervalos de confiança contém o valor da verdadeira média populacional $(\mu = 65)$, enquanto apenas 5 intervalos (destacados em azul) não contém a verdadeira média populacional.



Fixando-se um nível de confiança qualquer: $1 - \alpha$ temos:

$$P\left(-z_{\alpha/2} \leq Z \leq z_{\alpha/2}\right) = 1 - \alpha.$$

Graficamente, temos:



Substituindo Z por $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$, temos:

$$P\left(-z_{\alpha/2} \leq \frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \leq z_{\alpha/2}\right) = 1-\alpha.$$

Resolvendo-se as inequações para μ, temos:

$$P\left(\bar{X} - z_{\alpha/2}.\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + z_{\alpha/2}.\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha.$$

Assim, o intervalo de confiança de $100(1-\alpha)\%$ para a média populacional (μ) é dado por:

$$IC_{(1-\alpha)}(\mu) = \left[\bar{x} - z_{\alpha/2}.\frac{\sigma}{\sqrt{n}}\ ;\ \bar{x} + z_{\alpha/2}.\frac{\sigma}{\sqrt{n}}\right].$$

Exemplo

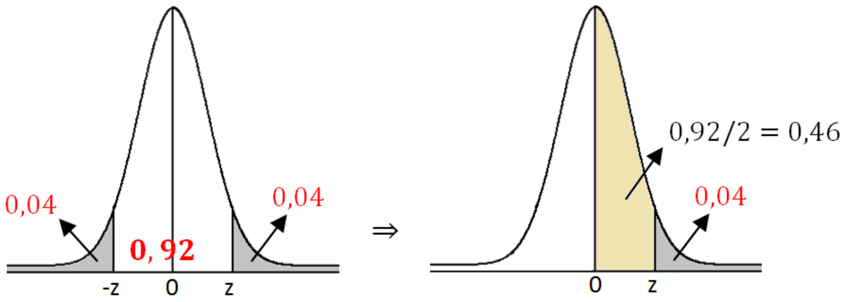
Sabe-se que o consumo mensal *per capita* de um determinado produto tem distribuição normal com desvio padrão de 2 kg. Foi realizada uma pesquisa de mercado, tomando-se uma amostra de 25 indivíduos e obteve-se um consumo médio *per capita* $\bar{x} = 7{,}2$ kg para esta amostra. Pede-se:

a) Obtenha o valor de z tal que $P(-z \leq Z \leq z) = 0{,}92$;

b) Estabeleça um intervalo de 92% de confiança para o consumo médio *per capita* deste produto. Interprete.

Solução

a)



$\Rightarrow \boldsymbol{z = 1,75}$

Z_t	0	1	2	3	4	5	6	7	8	9
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0753
⋮										
1,6	0,4452	0,4463	0,4474	0,4484	0,4495	0,4505	0,4515	0,4525	0,4535	0,4545
1,7	0,4554	0,4564	0,4573	0,4582	0,4591	0,4599	0,4608	0,4616	0,4625	0,4633
1,8	0,4641	0,4649	0,4656	0,4664	0,4671	0,4678	0,4686	0,4693	0,4699	0,4706

b) Temos $(1-\alpha) = 0{,}92$, e vimos no item (a) que $z_{\alpha/2} = 1{,}75$. Substituindo em:

$$IC_{(1-\alpha)}(\mu) = \left[\bar{x} - z_{\alpha/2}.\frac{\sigma}{\sqrt{n}} \;\; ; \;\; \bar{x} + z_{\alpha/2}.\frac{\sigma}{\sqrt{n}}\right]$$

segue que:

$$\begin{aligned} IC_{0,92}(\mu) &= \left[\bar{x} - \mathbf{1,75}\frac{\sigma}{\sqrt{n}} \quad , \quad \bar{x} + \mathbf{1,75}\frac{\sigma}{\sqrt{n}}\right] \\ &= \left[7{,}2 - 1{,}75\frac{2}{\sqrt{25}} \quad , \quad 7{,}2 + 1{,}75\frac{2}{\sqrt{25}}\right] \\ &= [6{,}50 \quad , \quad 7{,}90] \end{aligned}$$

Interpretação: O intervalo que vai de 6,50 kg até 7,90 kg contém o consumo médio *per capita* deste produto (de toda a população) com 92% de confiança (ou probabilidade).

Comandos no Software R para calcular o Intervalo de Confiança:

```
require(asbio) #Precisa instalar o pacote asbio
ci.mu.z(conf=0.92, sigma=2, xbar=7.2, n=25, summarized=TRUE)
```

2.3 Determinação do tamanho de uma amostra

Podemos reescrever a fórmula do intervalo de confiança apresentado na seção anterior como:

$$IC_{(1-\alpha)}(\mu) = \bar{X} \pm z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

ou, simplesmente

$$IC_{(1-\alpha)}(\mu) = \bar{X} \pm e$$

em que $e = z_{\alpha/2} . \frac{\sigma}{\sqrt{n}}$ é o erro de estimação.

Surge, então, a seguinte questão: "qual deve ser o tamanho da amostra (n) para se ter determinada precisão na estimação da média populacional?", ou seja, qual deve ser o n para que se tenha, no máximo, um determinado erro e?

Isolando o n na equação $e = z_{\alpha/2} . \frac{\sigma}{\sqrt{n}}$, tem-se:

$$n = \left(\frac{\sigma . z_{\alpha/2}}{e}\right)^2 .$$

Exemplo

Sabe-se que o consumo mensal *per capita* de um determinado produto tem distribuição normal com desvio padrão de 2 kg. Deseja-se realizar uma pesquisa de mercado para estimar o consumo médio *per capita* deste produto. Qual deve ser o tamanho da amostra (n) para que se tenha um erro (e) de no máximo $0{,}5\ kg$, com um nível de confiança de 95%?

Solução

Sabe-se que $\sigma = 2$ e portanto $\sigma^2 = 4$. Além disso foi fixado um nível de confiança $1 - \alpha = 0{,}95$, e, portanto $z_{\alpha/2} = 1{,}96$. Como se deseja um erro de estimação (e) de no máximo $0{,}5\ kg$, segue que:

$$n = \frac{\sigma^2 z_{\alpha/2}^2}{e^2} = \frac{4 \times (1{,}96)^2}{(0{,}5)^2} \approx 62$$

ou seja, para que se tenha um erro de no máximo $0{,}5\ kg$ para mais ou para menos na estimação da média populacional é necessário que a amostra seja composta por 62 indivíduos.

2.4 Intervalo de confiança para a média de uma população Normal com variância populacional desconhecida

A pressuposição de normalidade para a média amostral $(\bar{X})$ é garantida para amostras grandes $(n \geq 30)$, sendo que para amostras pequenas $(n < 30)$ esta pressuposição é válida apenas se sua população for normalmente distribuída e σ for conhecido.

Nesta seção será considerado o caso em que a amostra é pequena $(n < 30)$ e a população é normalmente distribuída, porém σ é desconhecido. Neste caso, o desvio padrão populacional (σ) deve ser substituído pelo desvio padrão amostral (S) e a distribuição normal (Z) deve ser substituída pela distribuição t de *Student*.

Distribuição *t* de *Student*

Seja $X_1, X_2, \ldots, X_n$ uma amostra aleatória proveniente de uma distribuição normal, com média μ com variância σ^2 desconhecidas. A variável aleatória

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}}$$

tem distribuição t, com $n - 1$ graus de liberdade.

A função densidade de probabilidade da distribuição t de *Student* é

$$f(x) = \frac{\Gamma\left(\frac{\phi + 1}{2}\right)}{\sqrt{\pi\phi}\ \Gamma\left(\frac{\phi}{2}\right)} . \frac{1}{\left(1 + \frac{x^2}{\phi}\right)^{\frac{\phi+1}{2}}}, \quad -\infty < x < \infty,$$

em que $\phi > 0$ é o número de graus de liberdade.

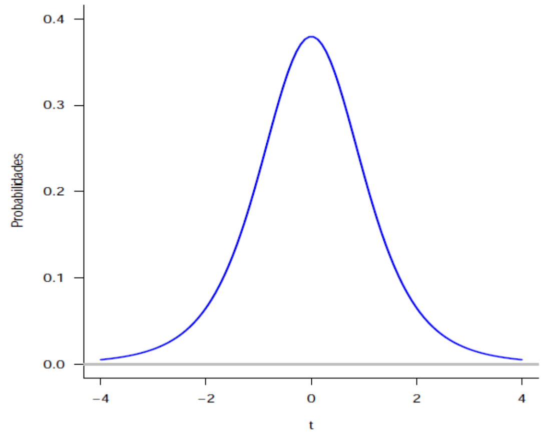


Figura. Gráfico da função densidade de probabilidade da distribuição t de *Student*

Quando (na prática) $n > 30$ a distribuição t tende para a normal padrão.

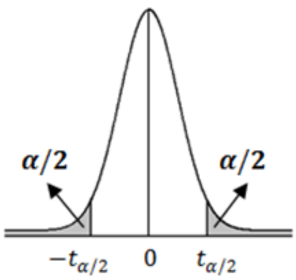
Tabela da distribuição *t* de *Student*

Neste material os quantis da distribuição t de *Student* são apresentados em uma tabela (bilateral) em função do número de graus de liberdade (ϕ) e da probabilidade (bilateral) α tal que $P(t \leq -t_{\alpha/2}) + P(t \geq t_{\alpha/2}) = \alpha$.

Apêndice 2. Tabela (Bilateral) da distribuição *t* de Student

Quantis da distribuição t de Student, com ϕ graus de liberdade, e probabilidades α, de acordo com o seguinte evento:

$$P(t \leq -t_{\alpha/2}) + P(t \geq t_{\alpha/2}) = \alpha$$

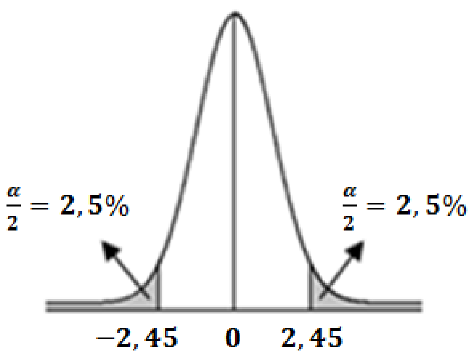


Graus de Liberdade	α					
	10%	5%	2%	1%	0,5%	0,1%
1	6,31	12,71	31,82	63,66	127,32	636,62
2	2,92	4,30	6,96	9,92	14,09	31,60
3	2,35	3,18	4,54	5,84	7,45	12,92
4	2,13	2,78	3,75	4,60	5,60	8,61
5	2,02	2,57	3,36	4,03	4,77	6,87
6	1,94	2,45	3,14	3,71	4,32	5,96
7	1,89	2,36	3,00	3,50	4,03	5,41

Por exemplo, para $\phi = 6$ graus de liberdade, se considerarmos $\alpha = 5\%$ de probabilidade, então o quantil $t_{\alpha/2} = t_{0,025} = 2,45$ é tal que:

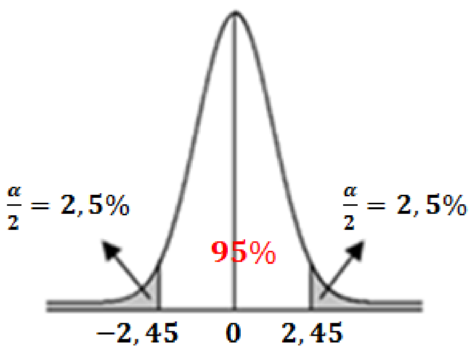
$$P(t \leq -2{,}45) + P(t \geq 2{,}45) = 5\%$$

Graus de Liberdade	α					
	10%	$\boxed{\textbf{5\%}}$	**2%**	**1%**	**0,5%**	**0,1%**
1	6,31	12,71	31,82	63,66	127,32	636,62
2	2,92	4,30	6,96	9,92	14,09	31,60
3	2,35	3,18	4,54	5,84	7,45	12,92
4	2,13	2,78	3,75	4,60	5,60	8,61
5	2,02	2,57	3,36	4,03	4,77	6,87
$\boxed{6}$	1,94	$\boxed{2{,}45}$	3,14	3,71	4,32	5,96
7	1,89	2,36	3,00	3,50	4,03	5,41



ou, ainda, $t_{0,025} = 2{,}45$ é o quantil da distribuição t de *Student* tal que:

$$P(-2{,}45 \leq t \leq 2{,}45) = 95\%$$



Intervalo de Confiança

O intervalo de confiança para a média (μ) de uma população com variância populacional (σ^2) desconhecida é dado por:

$$IC_{(1-\alpha)}(\mu) = \left[\bar{X} - t_{\alpha/2}.\frac{S}{\sqrt{n}} \quad , \quad \bar{X} + t_{\alpha/2}.\frac{S}{\sqrt{n}}\right],$$

em que $t_{\alpha/2}$ é o quantil da distribuição t de *Student* com $n-1$ graus de liberdade.

Exemplo (Levine, et al. 2012, pg. 271)

O Departamento de Transportes dos EUA exige que fabricantes de pneus forneçam informações sobre o desempenho dos pneus na banda lateral, de modo a que os potenciais consumidores possam ser mais bem informados ao tomar uma decisão de compra. Uma medida de grande importância do desempenho do pneu é o índice de desgaste da banda de rodagem, que indica a resistência do pneu em relação ao desgaste da banda de rodagem comparada a um pneu graduado com uma base de 100. Isso significa que um pneu com graduação de 200 deve durar duas vezes mais, em média, do que um pneu graduado com base de 100. Uma organização de defesa do consumidor deseja estimar o verdadeiro índice de desgaste da banda de rodagem de uma determinada marca de pneus, que declara "graduação 200" na banda lateral de seu pneu. Uma amostra aleatória de tamanho $n = 18$ indica uma média aritmética $\bar{x} = 195{,}3$ para o índice de desgaste de banda de rodagem, com desvio padrão da amostra $s = 21{,}4$.

Pressupondo que a população dos índices de desgaste de banda de rodagem seja distribuída conforme uma distribuição normal, construa um intervalo de confiança de 95% para a média do índice de desgaste das bandas de rodagem de todos os pneus produzidos por esse fabricante (de toda a população de pneus).

* LEVINE et. al. **Estatística Teoria e Aplicações: Usando o Microsoft Excel em Português**. 6 ed. Rio de Janeiro: LTC, 2012.

Solução

Temos: $n = 18$, $\bar{x} = 195{,}3$ e $s = 21{,}4$. Para construir o IC de 95% precisamos encontrar $t_{\alpha/2} = t_{0,025}$ com $n - 1 = 17$ graus de liberdade. Como a tabela é bilateral devemos procurar diretamente em $\alpha = 5\%$:

Graus de Liberdade	α					
	10%	$\boxed{5\%}$	2%	1%	0,5%	0,1%
1	6,31	12,71	31,82	63,66	127,32	636,62
2	2,92	4,30	6,96	9,92	14,09	31,60
3	2,35	3,18	4,54	5,84	7,45	12,92
⋮	⋮	⋮	⋮	⋮	⋮	⋮
16	1,75	2,12	2,58	2,92	3,25	4,01
$\boxed{17}$	1,74	$\boxed{2,11}$	2,57	2,90	3,22	3,97
18	1,73	2,10	2,55	2,88	3,20	3,92

Assim, $t_{0,025} = 2{,}11$. Portanto,

$$IC_{0,95}(\mu) = \left[\bar{X} - 2{,}11.\frac{S}{\sqrt{n}} \quad , \quad \bar{X} + 2{,}11.\frac{S}{\sqrt{n}}\ \right]$$

$$= \left[195{,}3 - 2{,}11.\frac{21{,}4}{\sqrt{18}} \quad , \quad 195{,}3 + 2{,}11.\frac{21{,}4}{\sqrt{18}}\ \right]$$

$$= [184{,}66 \quad , \quad 205{,}94\ ]$$

Interpretação: O intervalo que vai de 184,66 até 205,94 contém a "graduação" média de todos os pneus produzidos por esse fabricante (ou seja, de toda a população de pneus) com 95% de confiança.

Pergunta

Com base no resultado encontrado no exemplo anterior, você acredita que a organização de defesa do consumidor deveria acusar o fabricante de fabricar pneus que não atendem às informações de desempenho apresentadas na banda lateral do pneu?

Resposta

Como o intervalo de confiança de 95% é $[184{,}66 \quad , \quad 205{,}94\ ]$, então temos 95% de confiança de que este intervalo contém a verdadeira "graduação" média μ (de toda esta população de pneus), e, como a "graduação 200" está dentro deste intervalo então **não existe motivo** para suspeitar que a verdadeira "graduação" média, de toda esta população de pneus, seja diferente de 200. Assim, a organização de defesa do consumidor **<u>não deve</u>** acusar o fabricante de fabricar pneus que não atendem às informações de desempenho apresentadas na banda lateral do pneu.

Comandos no Software R para calcular o Intervalo de Confiança:

```
#Carregando o pacote BSDA
Library(BSDA) #Precisa instalar o pacote antes

#Intervalo de confiança:
tsum.test(mean.x=195.3, s.x=21.4, n.x=18, conf.level = 0.95,
          alternative="two.sided")

#Saída do R:

        One-sample t-Test

data:  Summarized x
t = 38.719, df = 17, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 184.658 205.942
sample estimates:
mean of x
    195.3                           IC de 95% para a média

Warning message:
In tsum.test(mean.x = 195.3, s.x = 21.4, n.x = 18, alternative =
"two.sided") :
  argument 'var.equal' ignored for one-sample test.
```

Obtenção do intervalo de confiança no R usando dados brutos

Considerando o exemplo anterior ("graduação de pneus"), suponha que ao invés das estatísticas descritivas (média, desvio padrão e tamanho da amostra) tivéssemos os dados brutos das graduações dos 18 pneus da amostra:

196.9	207.2	168.9	193.2	223.8	218.9
179.2	192.9	220.3	207.5	210.2	183.0
190.0	199.8	206.6	133.3	199.9	183.2

Neste caso, poderíamos obter o intervalo de confiança diretamente no R usando o comando `t.test()`.

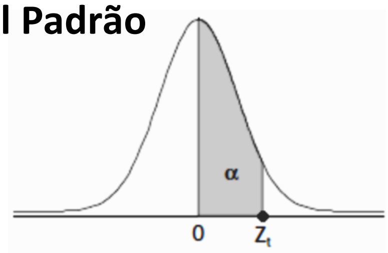
Comandos no Software R para calcular o Intervalo de Confiança:

```
#Entrando com os dados brutos (não foram apresentados no exemplo):
dados <- c(196.9, 179.2, 190.0, 207.2, 192.9, 199.8, 168.9, 220.3,
206.6, 193.2, 207.5, 133.3, 223.8, 210.2, 199.9, 218.9, 183.0,
183.2)

#Medidas descritivas (média, desvio padrão e tamanho da amostra):
mean(dados);sd(dados); length(dados)

#Saída do R (medidas descritivas):
[1] 195.2667
[1] 21.43998
[1] 18
#Observe que se a média e o desvio padrão forem arredondados para
uma casa decimal seus valores ficariam, respectivamente, 195.3 e
21.4, como no exemplo anterior.

#Intervalo de confiança:
t.test(dados, conf.level=0.95)

#Saída do R:
        One Sample t-test

data:  dados
t = 38.64, df = 17, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 184.6048 205.9285
sample estimates:
mean of x
 195.2667
```

IC de 95% para a média

Apêndice

Apêndice 1. Tabela da distribuição Normal Padrão



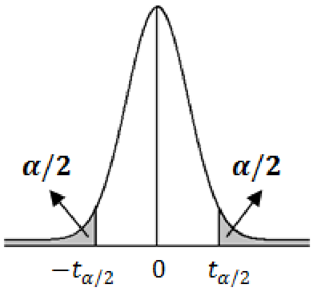
Probabilidades da distribuição Normal Padrão $N(0, 1)$ para valores do quantil Z_t padronizado de acordo com o seguinte evento: $P(0 < Z < Z_t) = \alpha$.

Z_t	0	1	2	3	4	5	6	7	8	9
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0753
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0,4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
0,5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0,6	0,2257	0,2291	0,2324	0,2357	0,2389	0,2422	0,2454	0,2486	0,2517	0,2549
0,7	0,2580	0,2611	0,2642	0,2673	0,2704	0,2734	0,2764	0,2794	0,2823	0,2852
0,8	0,2881	0,2910	0,2939	0,2967	0,2995	0,3023	0,3051	0,3078	0,3106	0,3133
0,9	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315	0,3340	0,3365	0,3389
1,0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621
1,1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1,2	0,3849	0,3869	0,3888	0,3907	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
1,3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
1,4	0,4192	0,4207	0,4222	0,4236	0,4251	0,4265	0,4279	0,4292	0,4306	0,4319
1,5	0,4332	0,4345	0,4357	0,4370	0,4382	0,4394	0,4406	0,4418	0,4429	0,4441
1,6	0,4452	0,4463	0,4474	0,4484	0,4495	0,4505	0,4515	0,4525	0,4535	0,4545
1,7	0,4554	0,4564	0,4573	0,4582	0,4591	0,4599	0,4608	0,4616	0,4625	0,4633
1,8	0,4641	0,4649	0,4656	0,4664	0,4671	0,4678	0,4686	0,4693	0,4699	0,4706
1,9	0,4713	0,4719	0,4726	0,4732	0,4738	0,4744	0,4750	0,4756	0,4761	0,4767
2,0	0,4772	0,4778	0,4783	0,4788	0,4793	0,4798	0,4803	0,4808	0,4812	0,4817
2,1	0,4821	0,4826	0,4830	0,4834	0,4838	0,4842	0,4846	0,4850	0,4854	0,4857
2,2	0,4861	0,4864	0,4868	0,4871	0,4875	0,4878	0,4881	0,4884	0,4887	0,4890
2,3	0,4893	0,4896	0,4898	0,4901	0,4904	0,4906	0,4909	0,4911	0,4913	0,4916
2,4	0,4918	0,4920	0,4922	0,4925	0,4927	0,4929	0,4931	0,4932	0,4934	0,4936
2,5	0,4938	0,4940	0,4941	0,4943	0,4945	0,4946	0,4948	0,4949	0,4951	0,4952
2,6	0,4953	0,4955	0,4956	0,4957	0,4959	0,4960	0,4961	0,4962	0,4963	0,4964
2,7	0,4965	0,4966	0,4967	0,4968	0,4969	0,4970	0,4971	0,4972	0,4973	0,4974
2,8	0,4974	0,4975	0,4976	0,4977	0,4977	0,4978	0,4979	0,4979	0,4980	0,4981
2,9	0,4981	0,4982	0,4982	0,4983	0,4984	0,4984	0,4985	0,4985	0,4986	0,4986
3,0	0,4987	0,4987	0,4987	0,4988	0,4988	0,4989	0,4989	0,4989	0,4990	0,4990
3,1	0,4990	0,4991	0,4991	0,4991	0,4992	0,4992	0,4992	0,4992	0,4993	0,4993
3,2	0,4993	0,4993	0,4994	0,4994	0,4994	0,4994	0,4994	0,4995	0,4995	0,4995
3,3	0,4995	0,4995	0,4995	0,4996	0,4996	0,4996	0,4996	0,4996	0,4996	0,4997
3,4	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4998
3,5	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998
3,6	0,4998	0,4998	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3,7	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3,8	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3,9	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000

Apêndice 2. Tabela (Bilateral) da distribuição t de Student

Quantis da distribuição t de Student, com ϕ graus de liberdade, e probabilidades α, de acordo com o seguinte evento:

$$P(t \leq -t_{\alpha/2}) + P(t \geq t_{\alpha/2}) = \alpha$$



Graus de Liberdade	α					
	10%	**5%**	**2%**	**1%**	**0,5%**	**0,1%**
1	6,31	12,71	31,82	63,66	127,32	636,62
2	2,92	4,30	6,96	9,92	14,09	31,60
3	2,35	3,18	4,54	5,84	7,45	12,92
4	2,13	2,78	3,75	4,60	5,60	8,61
5	2,02	2,57	3,36	4,03	4,77	6,87
6	1,94	2,45	3,14	3,71	4,32	5,96
7	1,89	2,36	3,00	3,50	4,03	5,41
8	1,86	2,31	2,90	3,36	3,83	5,04
9	1,83	2,26	2,82	3,25	3,69	4,78
10	1,81	2,23	2,76	3,17	3,58	4,59
11	1,80	2,20	2,72	3,11	3,50	4,44
12	1,78	2,18	2,68	3,05	3,43	4,32
13	1,77	2,16	2,65	3,01	3,37	4,22
14	1,76	2,14	2,62	2,98	3,33	4,14
15	1,75	2,13	2,60	2,95	3,29	4,07
16	1,75	2,12	2,58	2,92	3,25	4,01
17	1,74	2,11	2,57	2,90	3,22	3,97
18	1,73	2,10	2,55	2,88	3,20	3,92
19	1,73	2,09	2,54	2,86	3,17	3,88
20	1,72	2,09	2,53	2,85	3,15	3,85
21	1,72	2,08	2,52	2,83	3,14	3,82
22	1,72	2,07	2,51	2,82	3,12	3,79
23	1,71	2,07	2,50	2,81	3,10	3,77
24	1,71	2,06	2,49	2,80	3,09	3,75
25	1,71	2,06	2,49	2,79	3,08	3,73
26	1,71	2,06	2,48	2,78	3,07	3,71
27	1,70	2,05	2,47	2,77	3,06	3,69
28	1,70	2,05	2,47	2,76	3,05	3,67
29	1,70	2,05	2,46	2,76	3,04	3,66
30	1,70	2,04	2,46	2,75	3,03	3,65
40	1,68	2,02	2,42	2,70	2,97	3,55
60	1,67	2,00	2,39	2,66	2,91	3,46
120	1,66	1,98	2,36	2,62	2,86	3,37
∞	1,65	1,96	2,33	2,59	2,82	3,31

Bibliografia

BUSSAB, W. O., MORETTIN, P. A. **Estatística Básica**. 5. ed. São Paulo: Saraiva, 2002.

COSTA, S. C. **Estatística Aplicada à Veterinária**. Londrina: UEL, [ca. 2012]. (Apostila).

FERREIRA, D. F. **Estatística Básica**. Lavras: Editora UFLA, 2005.

FONSECA, J. S., MARTINS, G. A. **Curso de estatística**. 6. ed. São Paulo: Atlas, 1996.

LEVINE et. al. **Estatística Teoria e Aplicações: Usando o Microsoft Excel em Português**. 6 ed. Rio de Janeiro: LTC, 2012.

MONTGOMERY, D. C., RUNGER, G. C. **Estatística Aplicada e Probabilidade para Engenheiros**. 5. ed. Rio de Janeiro: LTC, 2012.

VIEIRA, S. **Introdução à Bioestatística**. 3. ed. rev. Rio de Janeiro: Campus, 1998.