

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n} = \frac{\sum_{i=1}^n x_i}{n}$$

ESTATÍSTICA APLICADA À ADMINISTRAÇÃO

$$P\left(\bar{X} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} < \bar{X} < \bar{X} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right) =$$

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

Rudinei Luiz Bogo

$1 - \alpha$

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

SUMÁRIO

APRESENTAÇÃO	3
1. ESTATÍSTICA DESCRITIVA.....	3
1.1 População e amostra	5
1.2 Amostras aleatórias e não-aleatórias.....	5
1.3 Tipos de dados.....	6
1.4 Representação dos dados	7
1.4.1 Rol ou listas	8
1.4.2 Tabelas.....	9
1.4.3 Gráficos.....	11
1.5 Medidas de posição e dispersão.....	12
2. PROBABILIDADES	20
2.1 Experimentos aleatórios e não-aleatórios.....	20
2.2 Definição de probabilidade	21
2.2.1 Espaço Amostral	22
2.2.2 Evento	23
2.2.3 Definição de probabilidade	23
2.2.4 Probabilidade condicional	27
2.2.5. Eventos independentes.....	28
2.3 Distribuições discretas e contínuas	29
2.3.1 Distribuições discretas: Bernoulli.....	31
2.3.2 Distribuições discretas: Binomial.....	33
2.3.3 Distribuições contínuas: normal	35
3. ESTATÍSTICA INFERENCIAL.....	41
3.1 Distribuição amostral.....	42
3.1.1 Distribuição amostral da média para populações normais	45
3.2 Teorema Central do Limite.....	47
3.3 Intervalos de confiança.....	47
3.3.1 Dimensionamento de amostras.....	50
3.4 Testes de hipóteses.....	51
3.4.1 Testes de hipóteses para uma média.....	54
REFERÊNCIAS	58

APRESENTAÇÃO

Caro estudante, este material objetiva apresentar ideias básicas sobre a estatística e suas aplicações na administração, de modo a ampliar o conhecimento sobre o universo da estatística. Assim, pretende-se abordar os assuntos de maneira mais aplicada e sem envolver deduções e/ou demonstrações.

Bons estudos!

1. ESTATÍSTICA DESCRITIVA

Em algum momento no trabalho, o gestor público se depara com o problema de analisar e entender um conjunto de dados relevantes ao objetivo de uma simples análise ou para tomada de decisão. Será necessário transformar tais dados em informações relevantes e confiáveis para julgar a adequação, em alguma decisão. Para realizar um estudo estatístico, normalmente, existem várias etapas a serem percorridas, conhecidas como fases do método estatístico.

As fases do método estatístico são:

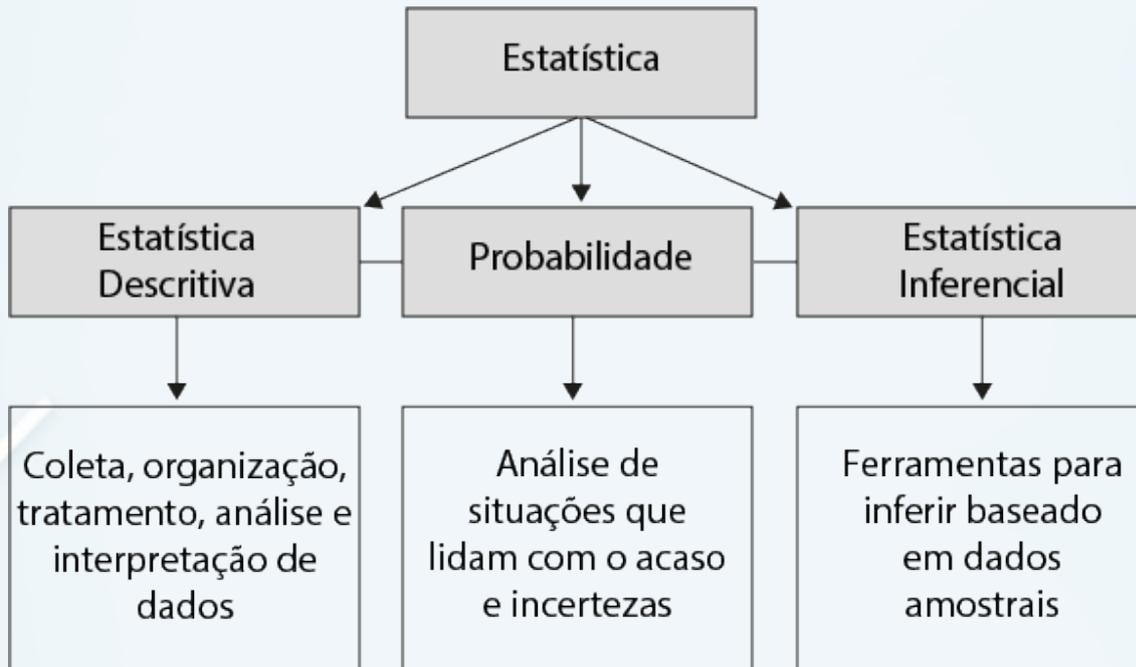
- definição do problema;
- planejamento do processo de resolução;
- obtenção dos dados;
- organização dos dados;
- apresentação dos dados;
- análise estatística e interpretação dos resultados.

Para obter êxito na resolução do problema analisado, é importante compreender todos os aspectos que envolvem o universo chamado estatística. Inicialmente, há que entender o que é estatística. A estatística é uma ciência que fornece métodos para a coleta, organização, descrição,

análise e interpretação de dados quantitativos ou qualitativos, para a utilização na tomada de decisões acertadas. A origem da palavra estatística vem do latim *statisticum collegium* (conselho de Estado) e do italiano *statista* (estadista ou político).

A estatística se divide em três subáreas, conforme o fluxograma a seguir:

Figura 1 – Fluxograma das subáreas da Estatística



Fonte: Elaborado pelo autor, 2022.

As fases do método estatístico estão inseridas nessas três subáreas. A Estatística descritiva, tratada neste capítulo, é a parte que se preocupa com a coleta, organização e tratamentos dos dados, mas sem tirar conclusões mais genéricas. A probabilidade é um número que varia de 0 (0%) a 1 (100%) e mede a chance de ocorrência de um determinado resultado. Já a inferência estatística é a parte da estatística que trabalha com análise e interpretação de dados. Nela, os resultados são obtidos utilizando uma amostra da população, evidenciando conclusões para o comportamento da população com precisão e a probabilidade confiável.

Em todas essas fases que compõem uma análise estatística de um determinado problema, é necessário ter o domínio da utilização de *softwares* estatísticos como o Excel, por exemplo.

Neste capítulo apresenta-se os elementos que compõem a estatística descritiva e os conceitos inerentes à estatística, importantes para a compreensão do processo como um todo.

1.1 População e amostra

Para a compreensão do que é uma população e o que é uma amostra, faz-se uma analogia com a teoria dos conjuntos. A população compreende todos os elementos de um determinado conjunto finito, por outro lado, a amostra se refere apenas a um subconjunto da população. Veja-se alguns exemplos:

- todos os habitantes de uma mesma cidade que contribuem com o pagamento do IPTU são considerados a população;
- optando-se por trabalhar apenas uma parte dessa população que contribui com o IPTU, que apresente as mesmas características da população, tem-se uma amostra.

Nem sempre é possível trabalhar com a população devido à dificuldade de dimensioná-la ou por conter um número muito grande de elementos. Assim, surgem algumas maneiras de se produzir uma determinada amostra, os chamados métodos de amostragem, a seguir.

1.2 Amostras aleatórias e não-aleatórias

Nas amostras aleatórias ou probabilísticas cada elemento da população tem a mesma probabilidade de ser selecionado para compor a amostra por meio de mecanismos de seleção. Por outro lado, no processo de amostragem não-aleatória, a seleção depende do julgamento do

pesquisador. Há uma escolha deliberada para compor a amostra utilizando mecanismos subjetivos.

Em geral, os métodos de amostragem aleatória apresentam resultados que representam melhor a população toda. Em compensação, os métodos de amostragem não-aleatórias são mais fáceis de serem utilizados. Para aprender um pouco mais sobre os diferentes tipos de amostragem, acesse:



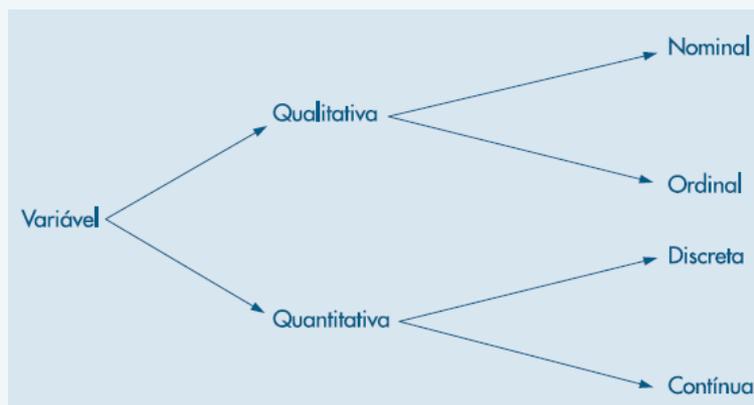
Tipos de Amostragem

1.3 Tipos de dados

Quando se realiza uma análise estatística, independentemente de utilizar uma população ou amostra, a classificação do tipo de dado (variável) analisado é muito importante. Eis diferentes tipos de variáveis e alguns exemplos.

As variáveis são classificadas de acordo com a figura 2:

Figura 2 – Tipos de variáveis



Fonte: BUSSAB e MORETTIN, 2017, p. 23.

- **Variável qualitativa:** representa uma qualidade (atributo) e não apresenta estrutura numérica. Exemplos: cor dos olhos, classe social, estado civil, nome da empresa etc. Este tipo de variável ainda é classificado em:

- **nominal:** quando não há ordenação ou hierarquia. Exemplos: sexo dos funcionários de uma prefeitura, estado civil, região de procedência etc.
- **ordinal:** quando há ordenação ou hierarquia. Exemplos: grau de instrução (fundamental, médio ou superior), cargos em uma organização pública (diretor, gerente, supervisor) etc.
- **Variável quantitativa:** é representada por meio de números provenientes de uma contagem ou mensuração. São de dois tipos:
 - **discreta:** observações que assumem valores fixos e inteiros. Exemplos: número de pessoas que pagam os impostos em dia, número de peças defeituosas etc.
 - **contínua:** observações que assumem quaisquer valores em um intervalo, os números reais, por exemplo. Exemplos: peso, altura, pressão arterial etc.

Compreendida a ideia de variáveis, é interessante, agora, analisar as diferentes possibilidades de se representar graficamente os dados coletados levando em consideração o tipo de variável que se está observando.

1.4 Representação dos dados

Feita a coleta, os dados originais (amostra) ainda não estão prontos para análise, por não estarem numericamente organizados. Por essa razão, são chamados de dados brutos, sem nenhum tipo de tratamento.

Veja-se o exemplo a seguir: foram coletados dados correspondentes ao percentual dos trabalhadores que contribuem com o Instituto Nacional de Seguridade Social (INSS) em 20 cidades de uma determinada região do Brasil no ano de 2008. Os dados são apresentados na tabela 1:

Tabela 1 – Percentual dos trabalhadores que contribuem com o INSS em 20 cidades de uma determinada região do Brasil no ano de 2008

45	51	50	58
50	44	46	57
42	41	60	58
41	50	54	60
52	46	52	51

Fonte: TAVARES, 2011, p. 50.

Os dados estão organizados na forma em que foram coletados, sem nenhum tratamento. Por essa razão, são chamados de dados brutos. Cabe ressaltar, também, que não estão agrupados.

A partir da coleta, existem basicamente três maneiras de representá-los: rol ou listas; tabelas e gráficos. Veja-se cada uma delas, na sequência.

1.4.1 Rol ou listas

Rol ou lista significa a simples organização dos dados brutos em ordem crescente ou decrescente. Tal organização aumenta muito o nível de informação dos dados. Considerando o exemplo anterior, a Tabela 2 apresenta os dados na forma de rol.

Tabela 2 – Percentual ordenado dos trabalhadores que contribuem com o INSS em 20 cidades de uma determinada região do Brasil no ano de 2008

41	46	51	57
41	46	51	58
42	50	52	58
44	50	52	60
45	50	54	60

Fonte: TAVARES, 2011, p. 50.

Com esse simples tratamento, verifica-se que o menor percentual foi 41% e o maior 60%, o que representa uma amplitude total de 19%. Outro fato observado é que nas cidades avaliadas, o valor 50 aparece com maior

frequência, é o que mais se repete. Diante disso, representa-se tais dados com base na frequência com que cada um ocorre.

1.4.2 Tabelas

O tipo de tabela mais comum para representar dados é a tabela de distribuição de frequências. Em todo caso é importante analisar o tipo de variável (quantitativa discreta ou contínua) para elaborar a tabela da maneira correta. No exemplo anterior, que considera o percentual de trabalhadores que contribuem ao INSS, trata-se de uma variável quantitativa contínua. Logo, deve-se distribuir os dados em classes (intervalos nos quais os valores da variável analisada são agrupados). A tabela 3 representa a distribuição de frequência em classes para o exemplo anterior.

Tabela 3 – Distribuição de Frequência em Classes do percentual de trabalhadores que contribuem ao INSS em 20 cidades de uma determinada região do Brasil no ano de 2008

Percentual (x_i)	Frequência (f_i)	Frequência Acumulada (f_{ac})	Frequência Relativa Acumulada (f_{rel})
38,5 – 43,5	3	3	0,15
43,5 – 48,5	4	7	0,35
48,5 – 53,5	7	14	0,7
53,5 – 58,5	4	18	0,9
58,5 – 63,5	2	20	1,0
Total	20		

Fonte: Elaborado pelo autor, 2022.

O passo-a-passo para produzir esta tabela e o que significa cada elemento nela contido é acessado no sítio



Passo-a-passo para montagem da Tabela de Distribuição de Frequência em Classes

Quando se trabalha com variáveis quantitativas discretas, também é possível construir a distribuição de frequência. Considere uma pesquisa durante 30 dias em um determinado mês com relação ao número de

reclamações (N.R.) no setor de tributos de uma prefeitura considerada como modelo de gestão em tributos. Os dados são apresentados pela tabela 4:

Tabela 4 - Dados referentes ao número de reclamações (N.R.) por dia no setor de tributos de uma prefeitura ao longo de 30 dias

DIA	N.R.								
1	0	7	1	13	0	19	1	25	0
2	2	8	2	14	0	20	0	26	3
3	1	9	2	15	1	21	0	27	4
4	5	10	3	16	2	22	2	28	0
5	3	11	0	17	3	23	0	29	2
6	2	12	3	18	5	24	4	30	1

Fonte: TAVARES, 2011, p. 59.

Organizando esses dados em uma lista (crescente) tem-se:

0 0 0 0 0 0 0 0 0 1 1 1 1 1 2 2 2 2 2 2 2 3 3 3 3 3 4 4 5 5

Como tais dados se referem ao número de reclamações (valores sempre inteiros), tem-se uma variável aleatória discreta. Nesse caso, as classes na tabela de distribuição de frequência são representadas pelos valores de ocorrência. O resultado será a tabela 5.

Tabela 5 - Distribuição de frequência de reclamações ocorridas diariamente durante certo mês

Número de Reclamações por dia (x_i)	Frequência (f_i)	Frequência Acumulada (f_{ac})	Frequência Relativa Acumulada (f_{rel})
0	9	9	0,3
1	5	14	0,47
2	7	21	0,70
3	5	26	0,86
4	2	28	0,93
5	2	30	1,00
Total	30		

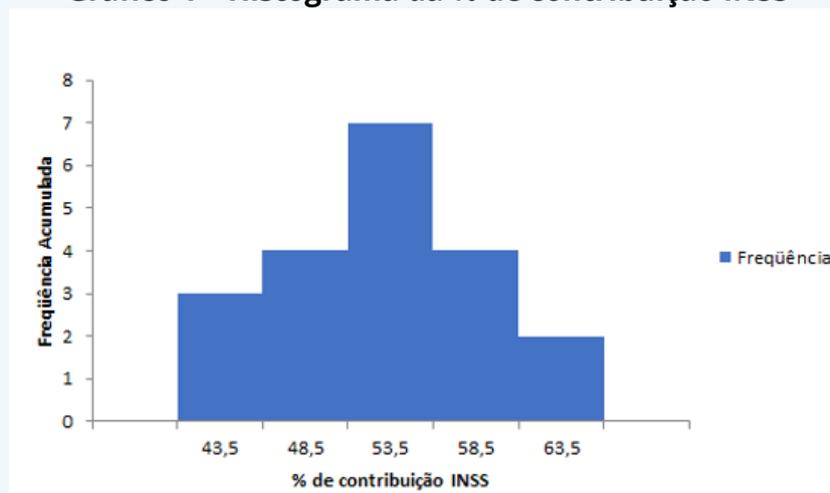
Fonte: Elaborado pelo autor, 2022.

Outra forma de apresentar um conjunto de dados é por meio de gráficos. Apresenta-se dois dos principais tipos de gráficos para representar variáveis quantitativas discretas e contínuas.

1.4.3 Gráficos

Existem diversos tipos de gráficos para representação de um conjunto de dados. Primeiramente, mostra-se como representar um conjunto de dados que utiliza uma variável quantitativa contínua. Tendo em conta o conjunto de dados que considera o percentual de trabalhadores que contribuem para o INSS em uma determinada cidade no ano de 2008, visualiza-se graficamente a distribuição dos dados utilizando o chamado histograma. O gráfico 1, apresenta o histograma destes dados.

Gráfico 1 - Histograma da % de contribuição INSS



Fonte: Elaborado pelo autor, 2022.

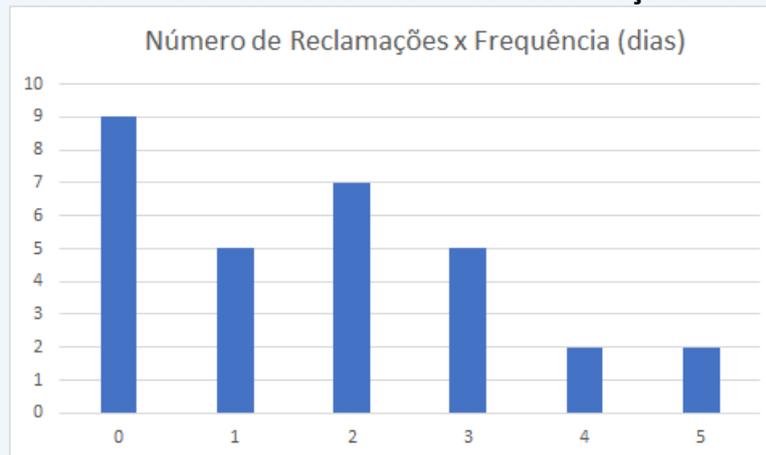
A principal característica do histograma é ser utilizado somente para variáveis quantitativas contínuas e as barras são adjacentes umas às outras. No eixo horizontal, tem-se as classes (intervalo) e no eixo vertical, a frequência acumulada. Para saber como elaborar um histograma utilizando a ferramenta Excel acesse:



Passo-a-passo para a construção do Histograma no Excel

Por outro lado, para representar um conjunto de dados que utiliza uma variável quantitativa discreta, utiliza-se o gráfico de barras. Tomando o conjunto de dados relacionado aos números de reclamações (N.R.), durante 30 dias de um certo mês, no setor de tributos de uma prefeitura, representa-se graficamente conforme o gráfico 2.

Gráfico 2 - Gráfico de barras nº de reclamações x dias



Fonte: Elaborado pelo autor, 2022.

Nota-se aqui que as barras estão no formato retangular e separadas, diferentemente do histograma. No eixo horizontal, tem-se o número de reclamações observadas em uma quantidade determinada de dias, representada pelo eixo vertical.

1.5 Medidas de posição e dispersão

Tendo feito a coleta, a organização e representação dos dados, caracteriza-se os dados utilizando medidas de posição e dispersão. As medidas de posição permitem determinar o valor central de um conjunto de dados, agrupados ou não agrupados. Já as medidas de dispersão permitem determinar a concentração dos demais valores em relação ao valor central dos dados.

Para facilitar o desenvolvimento dos estudos aqui, são enunciados dois exemplos como base e, a partir deles, será feita a aplicação das medidas de posição.

Exemplo base 1: A taxa de efetivação da cobrança de um determinado tributo que se apresentava atrasado em uma prefeitura após uma campanha realizada para que ele fosse saldado. Esses resultados são diários, conforme a tabela 6.

Tabela 6 – Taxa de efetivação de Cobrança

44	46	51	54	54	55	56	56	56
58	59	60	61	61	61	62	63	63

Fonte: TAVARES, 2011, p. 74.

Exemplo base 2: Em um projeto de aplicação de rede de esgoto de certa região de uma cidade, os engenheiros da Prefeitura Municipal tomaram uma amostra de 52 ruas, contando o número de casas por rua. Os dados referentes a uma pesquisa de mercado foram agrupados como mostra a tabela 7.

Tabela 7 – Distribuição de Frequências dos números de casas por rua de certa região de uma cidade

Número de casas por rua (x_i)	Frequência Absoluta (f_i)	Frequência Acumulada (f_{ac})	Frequência Relativa (f_{rel})
0 – 4	12	12	0,23
4 – 8	11	23	0,21
8 – 12	16	39	0,31
12 – 16	8	47	0,15
16 – 20	5	52	0,1
Total	52		

Fonte: Elaborado pelo autor, 2022.

As medidas de posição que tratadas a seguir são: média aritmética, mediana e moda.

- Média aritmética para dados não agrupados: a média aritmética de um conjunto de n valores x_1, x_2, \dots, x_n é definida por:

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n} = \frac{\sum_{i=1}^n x_i}{n}$$

Tomando o exemplo base 1, tem-se que

$$\bar{x} = \frac{44 + 46 + 51 + 54 + 54 + 55 + 56 + 56 + 56 + 58 + 59 + 60 + 61 + 61 + 61 + 62 + 63 + 63}{18}$$

resulta em,

$$\bar{x} = 56,67\%.$$

A taxa média de efetivação diária é de 56,67%.

- Média aritmética para dados agrupados: a média aritmética de um conjunto de n valores x_1, x_2, \dots, x_n que ocorrem com as frequências $f_i, i = 1, 2, \dots, k$, respectivamente, é definida por:

$$\bar{x} = \frac{x_1 f_1 + x_2 f_2 + x_3 f_3 + \dots + x_n f_n}{n} = \frac{\sum_{i=1}^n x_i f_i}{n}.$$

Observação: No caso em que os dados estão distribuídos em classes por intervalo, os valores x_1, x_2, \dots, x_n correspondem aos pontos médios das k classes.

Tomando o exemplo 2 como base, tem-se

$$\bar{x} = \frac{2 \cdot 12 + 6 \cdot 11 + 10 \cdot 16 + 14 \cdot 8 + 18 \cdot 5}{52} = 8,69$$

Logo, o número médio de casas na rua é de 8,69 casas. Importante destacar, em relação à notação matemática, quando calcula-se a média a partir de dados de uma população, que se indica a letra grega μ (mi) para designar a média populacional e para média amostral indica-se a notação \bar{x} . Contudo, a forma de cálculo é a mesma para as duas situações.

- Mediana para dados não agrupados: a mediana M_e de um conjunto ordenado de n valores ordenados $x_1, x_2, x_3, \dots, x_n$ é representada pelo valor central do conjunto para n ímpar e pela média aritmética dos dois valores centrais para n par.

Tomando o exemplo 1 por base, como o número de elementos é par, $n = 18$, a mediana será calculada pela média aritmética entre os elementos x_9 e x_{10} , isto é,

$$M_e = \frac{x_9 + x_{10}}{2} = \frac{56 + 58}{2} = \frac{114}{2} = 57.$$

- Mediana para dados agrupados: a mediana para dados agrupados em classes é definida por

$$Md = L_i + \left(\frac{P - f'_{ac}}{f_{M_e}} \right) \cdot h,$$

onde,

- L_i é o limite inferior da classe mediana;
- $P = n/2$ é a posição i da classe mediana;
- f'_{ac} é a frequência acumulada da classe anterior à classe mediana;
- f_{Md} é a frequência absoluta da classe mediana;
- e h é a amplitude da classe mediana.

Tomando o exemplo 2, tem-se que

- posição da mediana: $P = \frac{52}{2} = 26$. O valor 26 refere-se ao elemento x_{26} que pertence à classe 8 - 12, conforme distribuição de frequências.
- $L_i = 8$;
- $f'_{ac} = 23$;
- $f_{Md} = 16$;
- $h = 4$.

Assim,

$$Md = L_i + \left(\frac{P - f'_{ac}}{f_{M_e}} \right) \cdot h = 8 + \left(\frac{26 - 23}{16} \right) \cdot 4 = 8,75.$$

- Moda para dados não agrupados: a moda M_o de um conjunto de n valores $x_1, x_2, x_3, \dots, x_n$ é o número desse conjunto que tiver a maior repetição. Se o conjunto não tiver valores repetidos não existirá moda (amodal) e se dois valores estiverem igualmente repetidos, tem-se então duas modas e o conjunto será dito bimodal. A moda é o valor ao qual está associada a frequência mais alta.

Observando os dados do exemplo base 1, os valores que mais se repetem são 56 e 61. Portanto, o conjunto de dados é bimodal.

- Moda para dados agrupados: a moda para dados agrupados em classes é definida por

$$M_o = L_i + \left(\frac{\Delta_1}{\Delta_1 + \Delta_2} \right) \cdot h,$$

onde,

- L_i é o limite inferior da classe modal;
- Δ_1 é a diferença entre a frequência da classe modal e a imediatamente anterior;
- Δ_2 é a diferença entre a frequência da classe modal e a imediatamente posterior
- e h é a amplitude da classe modal.

Utilizando o exemplo base 2, tem-se que

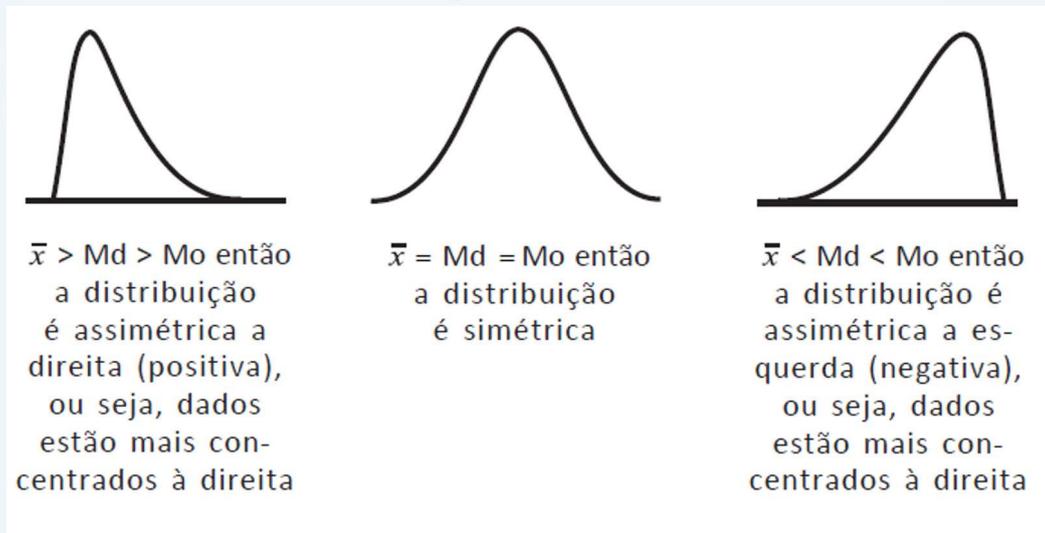
- a classe modal é classe que apresenta a maior frequência. Sendo assim, é a terceira classe 8 - 12, cujo valor é 16;
- $L_i = 8$;
- $\Delta_1 = 16 - 11 = 5$;
- $\Delta_2 = 16 - 8 = 8$;
- e $h = 4$.

Assim, a moda para esse conjunto de dados é

$$Mo = L_i + \left(\frac{\Delta_1}{\Delta_1 + \Delta_2} \right) \cdot h = 8 + \left(\frac{5}{5 + 8} \right) \cdot 4 = 9,53.$$

Resumindo, em um conjunto de dados, a média, a mediana e a moda não apresentam o mesmo valor. A mediana divide o conjunto de dados em dois subconjuntos de mesmo tamanho e não é influenciada pelos extremos. Utiliza-se estas três informações para ter um indicativo da natureza da distribuição dos dados, seguindo as regras definidas no gráfico 3.

Gráfico 3 – Distribuições assimétricas e simétricas



Fonte: TAVARES, 2011, p. 82.

As medidas de dispersão têm por objetivo avaliar a concentração dos dados em torno da média. As medidas de dispersão aqui apresentadas são: variância, desvio padrão e coeficiente de variação.

- Variância: se os dados são provenientes de uma população, a variância é definida pela expressão

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

Onde N é o tamanho da população e μ a respectiva média.

Para uma amostra da população, a variância é calculada por

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

Onde n é o tamanho da amostra. A variância mede o grau de dispersão dos dados em relação à média. A sua unidade de medida é o quadrado da unidade de medida das observações. O valor $n - 1$ é denominado grau de liberdade.

- Desvio-Padrão: se os dados são provenientes de uma população, a variância é definida pela expressão

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

Onde N é o tamanho da população e μ a respectiva média. Para uma amostra da população, a variância é calculada por

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

Onde n é o tamanho da amostra. A variância mede o grau de dispersão dos dados em relação à média. O desvio-padrão também mede o grau de dispersão dos pontos em relação à média. Porém, a unidade de medida é a mesma dos dados observados, facilitando o entendimento. Além disso, o desvio-padrão estabelece um intervalo de confiança indicando o quão é confiável esse valor (média), isto é,

$$(\bar{x} - s, \bar{x} + s)$$

Exemplo: considerando as séries de dados apresentadas pelos gastos com transportes em relação ao total gasto em várias prefeituras, conforme descrição a seguir, suponha que você precise efetuar uma estimativa com base nesses dados. Sobre qual série é mais fácil fazer estimativas precisas? Por quê?

Série A: {3,96; 3,17; 3,55; 3,61; 4,11; 4,57; 4,97; 5,91; 5,99; 5,74}

Série B: {1,46; 2,09; 3,04; 5,12; 7,80; 8,25; 9,95; 15,24; 17,40; 21,74}

Série A:

- Média:

$$\bar{x} = \frac{\sum x_i}{n} = \frac{3,96 + 3,17 + \dots + 5,74}{10} = 4,558$$

- Variância:

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$
$$= \frac{(3,96 - 4,558)^2 + (3,17 - 4,558)^2 + \dots + (5,74 - 4,558)^2}{10 - 1}$$
$$= 1,0939$$

- Desvio-padrão:

$$s = \sqrt{s^2} = \sqrt{1,0939} = 1,0459$$

Série B:

- Média:

$$\bar{x} = \frac{\sum x_i}{n} = \frac{1,46 + 2,09 + \dots + 21,74}{10} = 9,206$$

- Variância:

$$\begin{aligned}s^2 &= \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \\ &= \frac{(1,46 - 9,206)^2 + (2,09 - 9,206)^2 + \dots + (21,74 - 9,206)^2}{10 - 1} \\ &= 47,748\end{aligned}$$

- Desvio-padrão:

$$s = \sqrt{s^2} = \sqrt{47,748} = 6,91$$

Observe que na série A é mais fácil fazer estimativas precisas, pois ela apresenta menor dispersão dos dados em relação à média aritmética.

2. PROBABILIDADES

A origem da teoria das probabilidades está relacionada aos jogos de azar, desde o século XVII, devido a necessidade de um método racional para calcular os riscos envolvidos em jogos de cartas, roletas, entre outros.

A partir de então, governos e empresas utilizam probabilidades em processos de decisões, ajudando a desenvolver estratégias. Na área de Gestão, é uma ferramenta para tomada de decisões e análise de riscos.

Para compreender os principais conceitos de probabilidade, é necessário entender o que são experimentos aleatórios e não-aleatórios.

2.1 Experimentos aleatórios e não-aleatórios

Um experimento é dito não-aleatório quando o resultado é previsível se repetido sob condições específicas. Exemplos: aceleração da gravidade na ausência do ar, temperatura que a água entra em ebulição, entre outros. É possível responder a essas duas perguntas antes mesmo da realização dos experimentos.

Um experimento é dito aleatório quando não é possível prever o resultado, mesmo quando realizado sob as mesmas condições. Exemplos: reação de um contribuinte quando ele é atendido, lançamento de um dado, uma carta retirada do baralho, entre outros. Em todas estas situações, não há como prever o resultado antecipadamente.

É justamente nos experimentos aleatórios que a Teoria das Probabilidades auxilia na análise e na previsão de um resultado futuro.

Para saber mais sobre a história da teoria das probabilidades, acesse:



História das Probabilidades

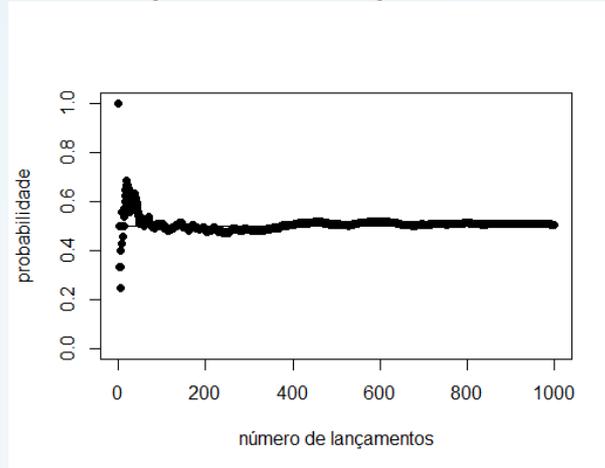
2.2 Definição de probabilidade

Para calcular uma probabilidade, é necessário ter um experimento aleatório que gere um resultado incerto ou casual. O experimento aleatório deve ter as seguintes características:

- deve ser repetido indefinidamente sob as mesmas condições (n vezes);
- não se conhece o resultado *a priori*, mas pode-se elencar todos os resultados possíveis (espaço amostral);
- a frequência relativa $f = \frac{r}{n}$, que corresponde ao número de vezes que um resultado favorável ocorreu (r) dividido pelo número total de repetições (n), deve tender a algum valor entre 0 (zero) e 1 (um).

Como exemplo, considera-se o experimento de jogar uma moeda inúmeras vezes e observar o número de caras obtidos. O gráfico 4 ilustra os resultados obtidos em 1000 lançamentos. Observa-se que a frequência relativa fica próxima de 0,5 após 1000 jogadas, diferentemente do que ocorre no início.

Gráfico 4 – Simulação de 1000 lançamentos de uma moeda



Fonte: Elaborado pelo autor, 2022.

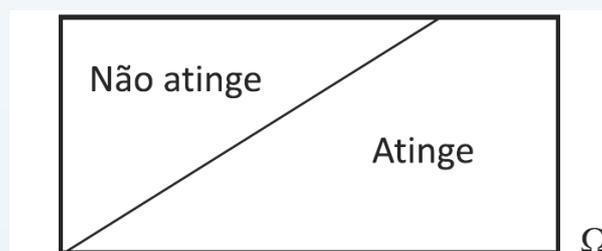
Além do conceito de experimento aleatório, para calcular probabilidade, é necessário o entendimento de espaço amostral e evento.

2.2.1 Espaço Amostral

Considere a seguinte situação: funcionários públicos de uma prefeitura são selecionados aleatoriamente e classificados em atinge ou não atinge uma determinada meta de produtividade. Há, apenas, dois resultados possíveis. O conjunto de todos os resultados possíveis é definido como espaço amostral e denotado por S ou Ω (ômega). Nesse caso, tem-se $\Omega = \{\text{atinge}; \text{não atinge}\}$.

Representa-se o espaço amostral de um experimento aleatório utilizando o Diagrama de Venn, estudado em Matemática para Administradores. A figura 2 utiliza o espaço amostral do exemplo citado no início desta seção.

Figura 3 – Espaço amostral



Fonte: TAVARES, 2011, p. 105.

2.2.2 Evento

É qualquer subconjunto do espaço amostral Ω associado a um experimento aleatório. Então, no exemplo dos funcionários públicos em que o espaço amostral é $\Omega = \{\text{atinge; não atinge}\}$, um evento representa um resultado particular de Ω .

Geralmente, as probabilidades são calculadas com base em um evento de um espaço amostral associado a um experimento aleatório.

2.2.3 Definição de probabilidade

Seja um determinado experimento aleatório e Ω um espaço amostral finito, cujos resultados de Ω sejam igualmente prováveis. Considere um evento A de Ω , então a probabilidade de ocorrência do evento A , representada por $P(A)$, é o número real definido por:

$$P(A) = \frac{n_A}{n_\Omega}$$

Onde n_A é o número de resultados favoráveis do evento A , e n_Ω é o número de resultados possíveis do espaço amostral do experimento realizado. O espaço amostral Ω é igualmente provável quando todos os seus elementos têm a mesma probabilidade de ocorrência.

Exemplo: Considerando o lançamento de um dado honesto (cada face possui a mesma chance de ocorrência), seja o evento A que considera a ocorrência dos números 3 ou 4. Qual é a probabilidade associada ao evento A ?

O espaço amostral é dado por $\Omega = \{1,2,3,4,5,6\}$ e o evento A é o subconjunto de Ω dado por $A = \{3,4\}$. Logo, $n_A = 2$ e $n_\Omega = 6$. Assim, a probabilidade de cair o número 3 ou 4 num único lançamento desse dado é igual a

$$P(A) = \frac{n_A}{n_\Omega} = \frac{2}{6} = \frac{1}{3}$$

A probabilidade é um número real entre 0 (zero) e 1 (um). Ao se multiplicar o resultado por 100, tem-se o valor correspondente em porcentagem.

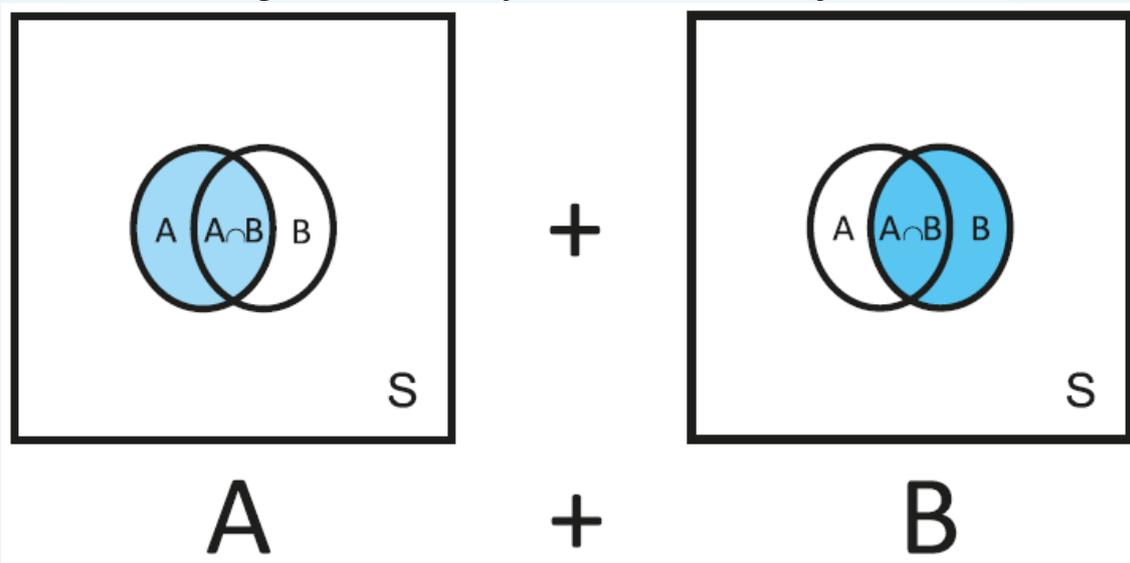
A probabilidade é definida, ainda, como uma função que atribui um número real aos eventos do espaço amostral Ω (se A é um evento de Ω , então $P(A)$ é a probabilidade de A), que satisfaz:

- $P(\emptyset) = 0$ (probabilidade do conjunto vazio é igual zero);
- $P(\Omega) = 1$ (probabilidade do espaço amostral é igual a um);
- $0 \leq P(A) \leq 1$ (probabilidade de um evento qualquer de Ω está entre 0 (zero) e 1 (um)).

Existem, ainda, algumas propriedades envolvendo operações entre eventos que auxiliam a calcular probabilidades. São elas:

- União entre dois eventos: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Figura 4 - Intersecção entre dois subconjuntos

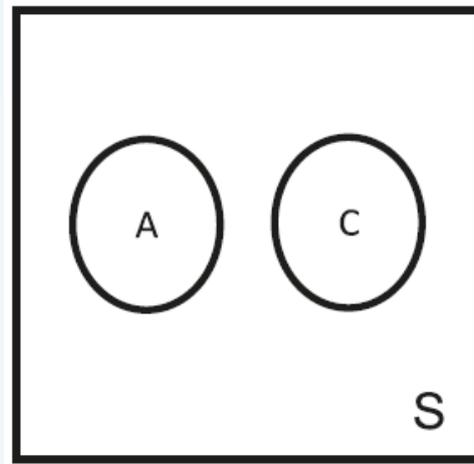


Fonte: TAVARES, 2011, p. 111.

Na figura 3 quando se soma $P(A) + P(B)$, soma-se duas vezes a intersecção ($A \cap B$).

- Eventos mutuamente exclusivos: $P(A \cup B) = P(A) + P(B)$

Figura 5 - Conjuntos disjuntos

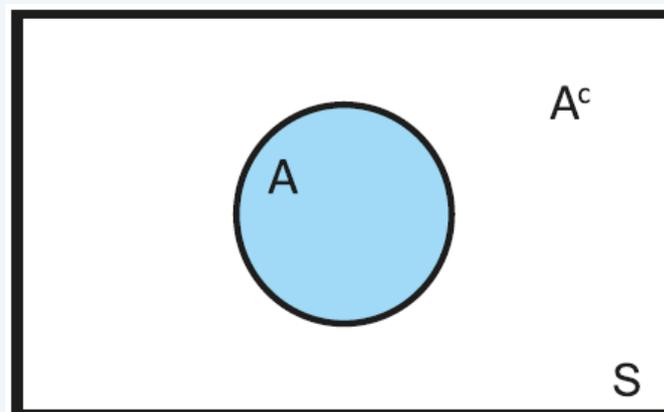


Fonte: TAVARES, 2011, p. 111.

Nesse caso, os eventos A e B não podem acontecer simultaneamente, o que significa $A \cap B = \emptyset$ (intersecção é igual ao conjunto vazio).

- Evento complementar: $P(A) = 1 - P(A^c)$

Figura 6 - Evento complementar



Fonte: TAVARES, 2011, p. 112.

A^c é o evento em que A não ocorre (complementar de A).

Exemplo: Os dados da tabela 8 referem-se ao total de contribuintes, de três classes sociais, que pagaram e não pagaram em dia um determinado tributo.

Tabela 8 - Dados sobre total de contribuintes de um determinado tributo

	MÉDIA-BAIXA	ALTA	TOTAL
Pagaram (P)	39.577	8.672	48.249
Não Pagaram (NP)	46.304	7.297	53.601
Total	85.881	15.969	101.850

Fonte: TAVARES, 2011, p. 113.

Selecionado um contribuinte aleatoriamente, responda:

(a) Há probabilidade de ele ser da classe média-baixa?

Todos os contribuintes da amostra representam o espaço amostral $\Omega = \{101.850 \text{ contribuintes}\}$. $MB = \text{contribuintes da classe média-baixa}$. Logo, a probabilidade $P(MB)$ é igual a

$$P(MB) = \frac{\text{n}^\circ \text{ de contribuintes da classe média - baixa}}{\text{n}^\circ \text{ total de contribuintes}} = \frac{85.881}{101.850} = 0,843.$$

(b) Há probabilidade de ele pagar em dia o tributo ou ser da classe média-baixa?

Considerando, $P = \text{pagar em dia}$ e $MB = \text{classe média-baixa}$, há que calcular

$$P(P \cup MB) = P(P) + P(MB) - P(P \cap MB).$$

Os dois eventos não são disjuntos ou mutuamente exclusivos, visto que existem contribuintes que pagam em dia e são classe-média. Logo,

$$P(P \cup MB) = \frac{48.249}{101.850} + \frac{85.881}{101.850} - \frac{39.577}{101.850}$$

$$P(P \cup MB) = 0,473 + 0,843 - 0,388$$

$$P(P \cup MB) = 0,928.$$

Portanto, a probabilidade de um contribuinte selecionado aleatoriamente pagar em dia o tributo ou ser da classe média-baixa é igual a 0,928.

2.2.4 Probabilidade condicional

Suponha, considerando o exemplo anterior, que haja uma informação parcial sobre o contribuinte selecionado, por exemplo, ele não paga o tributo em dia. Com base nessa informação, qual a probabilidade de que ele seja da classe alta?

Essa probabilidade é chamada de probabilidade condicional, e denotada por

$$P(MB|P).$$

Leia-se: “probabilidade do evento *MB* ocorrer sabendo que o evento *P* já ocorreu”.

Nesse caso, tem-se

$$P(MB|P) = \frac{\text{n}^\circ \text{ de contribuintes que são da classe média – baixa e pagam em dia}}{\text{n}^\circ \text{ total de contribuintes que pagam em dia}}$$

$$P(MB|P) = \frac{39.577}{48.249} = 0,820.$$

Houve uma redução no espaço amostral, devido à informação parcial de que o contribuinte selecionado pagava em dia. Sendo assim,

$$\begin{aligned} P(MB|P) &= \frac{\frac{\text{n}^\circ \text{ de contribuintes que são da classe média – baixa e pagam em dia}}{\text{n}^\circ \text{ total de contribuintes}}}{\frac{\text{n}^\circ \text{ de contribuintes que pagam em dia}}{\text{n}^\circ \text{ total de contribuintes}}} \\ &= \frac{P(MB \cap P)}{P(P)} \end{aligned}$$

De uma forma geral, considerando dois eventos A e B de um experimento aleatório qualquer, diz-se que a probabilidade de A ocorrer dado que o evento B ocorreu é definida como

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

2.2.5. Eventos independentes

Dois eventos A e B são ditos independentes, quando um evento não afeta a probabilidade do outro. Então

$$P(A|B) = P(A) \text{ ou } P(B|A) = P(A).$$

Logo, pela fórmula da probabilidade condicional, tem-se

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

Esta nova fórmula é conhecida como regra do produto. Contudo, se os eventos A e B forem independentes, escreve-se

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

Para compreender melhor essa discussão, analisa-se o seguinte experimento: lançar uma moeda (não viciada) 3 (três) vezes. Considere, agora, os eventos

$$A = \{\text{no primeiro lançamento da moeda sai cara}\}$$

$$B = \{\text{no segundo lançamento da moeda sai cara}\}$$

Verificar que os eventos A e B são independentes nesse experimento.
Fazendo C = cara e K = coroa, tem-se

$$\Omega = \{CCC, CCK, CKC, CKK, KCC, KCK, KKC, KKK\}$$

$$A = \{CCC, CCK, CKC, CKK\}$$

$$B = \{CCC, CCK, KCC, KCK\}$$

$$P(A \cap B) = \frac{2}{8} = \frac{1}{4}$$

$$P(A) = \frac{4}{8} = \frac{1}{2}$$

$$P(B) = \frac{4}{8} = \frac{1}{2}$$

Logo, é fácil ver que $P(A \cap B) = P(A)P(B)$. Ou ainda, considerando a regra da probabilidade condicional

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{1/4}{1/2} = \frac{1}{2} = P(A)$$

Portanto, os eventos A e B são independentes.

Vistos e entendidos os conceitos relacionados a probabilidades até aqui, apresenta-se a seguir, as chamadas distribuições de probabilidades. Essas distribuições são de extrema utilidade na estatística inferencial, além de auxiliar no próprio cálculo de probabilidades.

2.3 Distribuições discretas e contínuas

Para entender as distribuições de probabilidade, é necessário compreender o conceito de variável aleatória, que pode ser discreta ou contínua. Uma variável aleatória é definida como uma função que associa valores reais aos eventos de um espaço amostral.

Um exemplo de variável aleatória discreta é $X = \text{quantidade de ações que tiveram queda em um determinado dia}$. Nesse caso, X é uma variável aleatória que assume os valores 0, 1, 2, 3, 4, ...

Por outro lado, um exemplo de variável aleatória contínua é $Y = \text{tempo gasto por um funcionário público para atender um contribuinte}$. Y pode assumir um número infinito de valores (tempo é contínuo).

Um exemplo mais claro sobre variável aleatória. Considere o experimento em que duas moedas são lançadas. Nesse caso, o espaço amostral é dado por

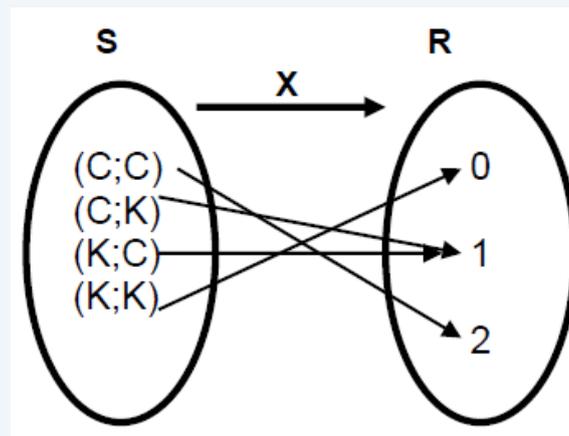
$$S = \{(CC), (CK), (KC), (KK)\},$$

em que C = cara e K = coroa. Considere, agora, a variável aleatória

$X = \text{número de caras (C) obtidas no lançamento de duas moedas}$.

Assim, define-se X como uma função que associa cada elemento do espaço amostral à quantidade de caras obtidas. A figura 7 ilustra os possíveis resultados para X .

Figura 7 - Resultados da v.a. X



Fonte: Elaborado pelo autor, 2022.

A cada resultado que X assume é possível atribuir uma probabilidade, definindo assim, uma distribuição de probabilidades. Para o exemplo anterior, os resultados são mostrados na tabela 9.

Tabela 9 – Distribuição de probabilidades da v.a. X

X	$x_1 = 0$	$x_2 = 1$	$x_3 = 2$
$P(X = x_i)$	1/4	1/2	1/4

Fonte: Elaborado pelo autor, 2022.

Se a variável X em questão for discreta, a distribuição é caracterizada por uma função de probabilidade ($P(X = x)$) e que satisfaz

$$\sum_{i=1}^n P(X = x_i) = 1,$$

isto é, o somatório de todas as probabilidades associadas aos valores X é igual a 1 (um). Existem várias distribuições de probabilidades para caracterizar uma variável aleatória discreta. No entanto, nesse curso, apresenta-se apenas as distribuições de Bernoulli e Binomial.

2.3.1 Distribuições discretas: Bernoulli

São experimentos no qual somente dois possíveis resultados ocorrem, sucesso e fracasso, por exemplo. Também são conhecidos por experimentos de Bernoulli (Jacques Bernoulli, 1654 – 1705, matemático suíço, foi o primeiro a descrever esses experimentos, no século XVII).

Alguns exemplos:

- uma venda é efetuada ou não em uma ligação de *call center*;
- um contribuinte é adimplente ou inadimplente;
- uma guia recolhida tem o preenchimento correto ou não;
- em uma linha de produção, um item é classificado defeituoso ou não.

Tais situações correspondem à Distribuição de Bernoulli, cujos resultados possíveis de X são:

- i. $X=1$ se o resultado for um sucesso;
- ii. $X=0$ se o resultado for um fracasso.

Se p a probabilidade de ocorrer um sucesso, a probabilidade de ocorrer um fracasso será $q = 1 - p$ e a função de probabilidade da Distribuição de Bernoulli será

$$P(X = x) = \begin{cases} p, & \text{para } x = 1 \\ q = 1 - p, & \text{para } x = 0 \end{cases}$$

ou ainda,

$$P(X = x) = p^x q^{1-x}.$$

Observação: O sucesso deve ser relacionado com o objetivo do exercício ou problema a ser resolvido, sendo que, pode não ser algo bom.

Uma distribuição de probabilidade é caracterizada pelas estimativas da média e do desvio padrão. No caso da distribuição de Bernoulli, são mostradas que a média e a variância são dadas por:

- $E(X) = p$ (também conhecida como esperança de X e corresponde à probabilidade de sucesso);
- $V(X) = pq$ (corresponde à probabilidade de sucesso p vezes à probabilidade de fracasso q)

Por fim, se X é uma variável aleatória que segue uma distribuição de Bernoulli, denota-se por $X \sim Ber(p)$.

2.3.2 Distribuições discretas: Binomial

A origem de uma distribuição de probabilidade Binomial vem dos experimentos de Bernoulli. Seja um experimento dentro das seguintes condições:

- são realizadas n tentativas independentes;
- cada tentativa é uma prova de Bernoulli em que somente dois resultados são possíveis (sucesso ou fracasso)
- as probabilidades p de sucesso e $q = 1 - p$ de fracasso se mantêm constantes em todas as provas.

A função de probabilidade da distribuição Binomial é definida por

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

onde p é probabilidade de sucesso; q é a probabilidade de fracasso;

$\binom{n}{x} = \frac{n!}{x!(n-x)!}$ é a combinação de n valores tomados x a x .

Simbolicamente, $X \sim B(n, p)$.

A média (esperança) e a variância da distribuição Binomial são dadas por:

- $E(X) = np$;
- $V(X) = npq$.

Exemplo: em uma determinada repartição pública, 10% das guias preenchidas estão incorretas. Essas guias correspondem a uma liberação na qual 10 (dez) guias estão preenchidas conjuntamente. Considere que cada guia tem a mesma probabilidade de ser preenchida corretamente (como se houvesse repetição no experimento de retirar guias).

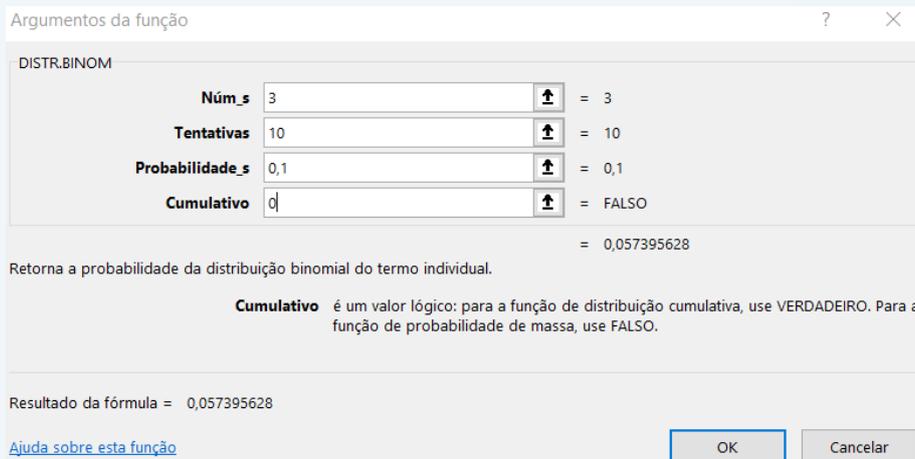
- (a) Qual a probabilidade de haver exatamente 3 (três guias) incorretas nas 10 (dez) guias para liberação?

O sucesso é a ocorrência de guias preenchidas incorretamente, isto é, $p = 0,1$ e $n = 10$. Logo,

$$P(X = 3) = \binom{10}{3} 0,1^3 0,9^{10-3} = 0,0574$$

Tal valor é obtido, facilmente, utilizando o Excel, conforme figura 8.

Figura 8 – Cálculo da Binomial utilizando o Excel



Argumentos da função

DISTR.BINOM

Núm_s	3	= 3
Tentativas	10	= 10
Probabilidade_s	0,1	= 0,1
Cumulativo	0	= FALSO

= 0,057395628

Retorna a probabilidade da distribuição binomial do termo individual.

Cumulativo é um valor lógico: para a função de distribuição cumulativa, use VERDADEIRO. Para a função de probabilidade de massa, use FALSO.

Resultado da fórmula = 0,057395628

[Ajuda sobre esta função](#) OK Cancelar

Fonte: Elaborado pelo autor, 2022.

Para acessar esta tela, basta seguir os passos:
abrir Excel -> Fórmulas -> Estatística -> Dist. BINOMIAL.

(b) Qual a probabilidade de haver duas ou mais guias incorretas nas dez guias para liberação?

$$\begin{aligned} P(X \geq 2) &= P(X = 2) + P(X = 3) + \dots + P(X = 10) \\ &= 1 - [P(X = 0) + P(X = 1)] \\ &= 0,2639. \end{aligned}$$

Se quiser saber mais sobre distribuições de probabilidade de variáveis aleatórias discretas, basta acessar:



Algumas distribuições de probabilidades discretas

A seguir, discorre-se sobre algumas das principais distribuições de probabilidades envolvendo as variáveis aleatórias contínuas, essencial para o estudo em diversas áreas de conhecimento.

2.3.3 Distribuições contínuas: normal

Inúmeras variáveis contínuas que descrevem fenômenos naturais e sociais apresentam distribuições de probabilidades próximas da distribuição normal. Aliada ao Teorema Central do Limite (TCL), apresentado mais adiante, a distribuição normal aparece em qualquer estudo que realizado mediante a distribuição amostral das médias.

Diferentemente das variáveis discretas, em que a função de probabilidade é definida como $P(X = x)$, aqui há uma função contínua conhecida como função densidade de probabilidade. No caso da distribuição normal, tem-se

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

em que, μ e σ são a média e o desvio padrão, respectivamente; $\pi = 3,14165 \dots$; e $e = 2,718 \dots$, conhecido como número de Euler. Simbolicamente, $X \sim N(\mu, \sigma^2)$, onde σ^2 representa a variância.

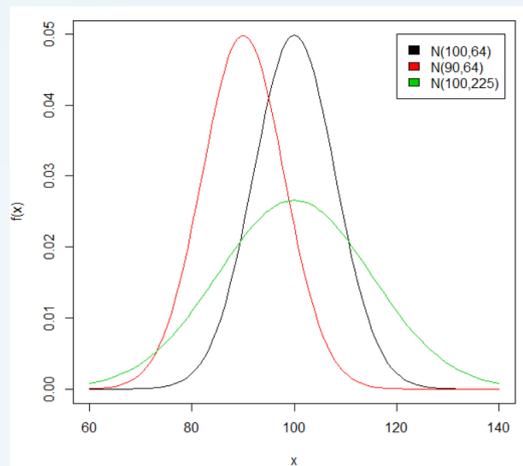
A origem da distribuição normal está relacionada com a distribuição Binomial. Caso queira saber mais, acesse:



Origem da Distribuição Normal

O gráfico da distribuição normal, tem características muito importantes.

Gráfico 5 - Exemplos de curvas normais

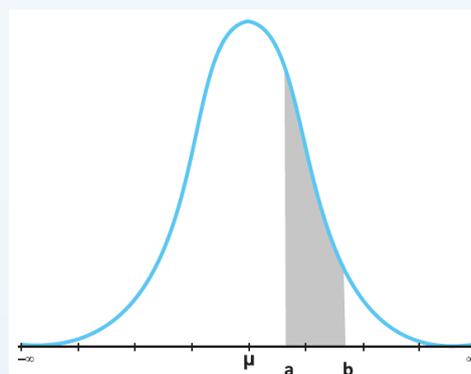


Fonte: Elaborado pelo autor, 2022.

- É simétrica em relação a $x = \mu$ (divide os dados em duas partes iguais);
- tem formato de sino;
- a média, a mediana e a moda são iguais;
- a área compreendida entre a curva e o eixo x é igual a 1 (conceito de probabilidades);
- $E(X) = \mu$ e $V(X) = \sigma^2$.

O cálculo das probabilidades envolvendo uma variável aleatória contínua é equivalente ao cálculo da área sob a curva entre os valores $x = a$ e $x = b$, onde $a < b$, e representa a probabilidade da variável X assumir um valor entre a e b (região escura), conforme o gráfico 6.

Gráfico 6 - Área entre $x=a$ e $x=b$



Fonte: TAVARES, 2011, p. 142.

As probabilidades da distribuição normal são calculadas via cálculo de uma integral, conceito visto em cursos de Cálculo. Na prática, é necessário dispor de tabelas com os resultados das probabilidades desejadas, porém, ter-se-ia que desenvolver inúmeras tabelas diferentes para cada combinação particular de μ e σ , o que é praticamente impossível.

O problema é resolvido por meio de uma transformação de variáveis que conduz à chamada distribuição normal padronizada ou distribuição normal reduzida. Fazendo a transformação de variável

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

tem-se a função densidade de probabilidade dada por

$$f(z) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{1}{2}(z)^2}, -\infty \leq z \leq +\infty.$$

A média (esperança) e variância de Z são definidas por

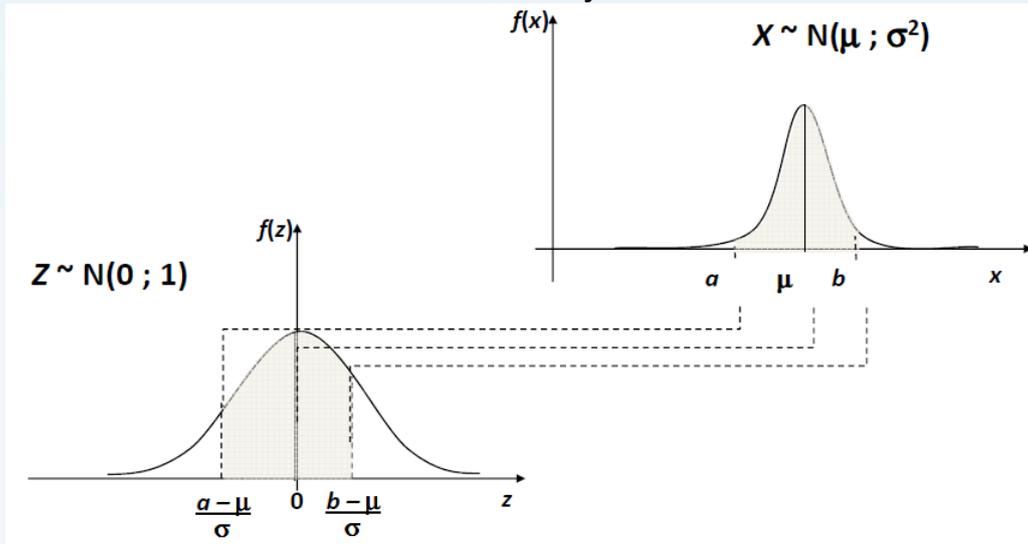
$$E(Z) = E\left(\frac{X - \mu}{\sigma}\right) = 0, \text{ pois } E(X) = \mu,$$

e

$$V(Z) = V\left(\frac{X - \mu}{\sigma}\right) = 1, \text{ pois } V(X) = \sigma^2.$$

Portanto, tem-se uma nova distribuição, cuja média é sempre igual a 0 (zero) e a variância igual a 1 (um). O gráfico 7 ilustra a transformação de variáveis realizada.

Gráfico 7 - Transformação de variável



Fonte: Elaborado pelo autor, 2022.

Agora, é viável estabelecer uma única tabela para obter qualquer probabilidade de uma variável aleatória com distribuição normal. Para acessar a tabela, acesse:



Tabela da Distribuição Normal Padronizada Z

Nesta tabela são fornecidas as áreas ou probabilidades de valores que estão entre 0 (zero) e o valor de z. Para ilustrar e entender a utilização da distribuição normal padronizada, utiliza-se o exemplo: considere que a arrecadação de um tributo, em uma pequena cidade, segue uma distribuição normal com valor médio arrecadado de R\$ 60.000,00 e desvio padrão de R\$10.000,00. Pergunta-se:

- (a) Qual a probabilidade de uma arrecadação ser maior do que R\$75.000,00?

Tem-se que $\mu = 60000$ e $\sigma^2 = 10000^2$, e daí, $X \sim N(60000, 10000^2)$. Quer-se obter $P(X > 75000)$. Padronizando a variável, tem-se

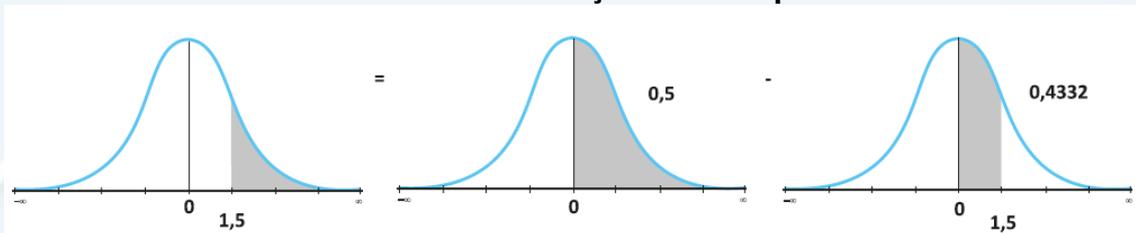
$$P(X > 75000) = P\left(\frac{X - \mu}{\sigma} > \frac{75000 - \mu}{\sigma}\right).$$

Substituindo, $\frac{x-u}{\sigma}$ por Z , μ por 60000 e σ por 10000, tem-se

$$P(Z > 1,50).$$

Olhando na Tabela de Z, e buscando o valor de $z = 1,50$ (1,5 na primeira coluna e o zero na primeira linha), encontra-se o valor de 0,4332. Esse valor ainda não é o desejado, pois corresponde à probabilidade de z estar entre 0 (zero) e 1,5 (quer-se $P(Z > 1,5)$). Conforme o gráfico 8, precisa-se obter área escura da parte (a), que corresponde à área da parte (b) menos a área da parte (c).

Gráfico 8 - Resolução do exemplo



Fonte: TAVARES, 2011, p. 145.

Então,

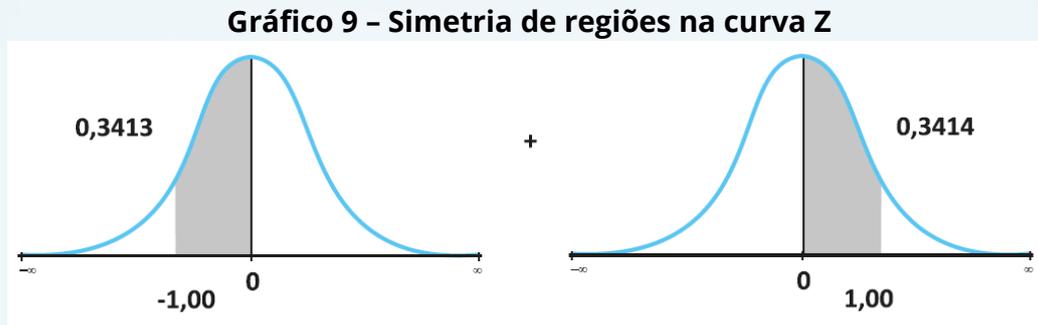
$$\begin{aligned} P(X > 75000) &= P(Z > 1,5) = 0,5 - P(0 < z < 1,50) \\ &= 0,5 - 0,4332 \\ &= 0,0668. \end{aligned}$$

Portanto, a probabilidade de obter uma arrecadação superior a R\$75.000,00 é de 0,0668 ou 6,68%.

(b) Qual a probabilidade de a arrecadação estar entre R\$50.000,00 e R\$70.000,00?

$$\begin{aligned} P(50000 < X < 70000) &= P\left(\frac{50000 - \mu}{\sigma} < \frac{X - \mu}{\sigma} < \frac{70000 - \mu}{\sigma}\right) \\ &= P\left(\frac{50000 - 60000}{10000} < z < \frac{70000 - 60000}{10000}\right) \\ &= P(-1,00 < z < 1,00) \end{aligned}$$

O valor de $z = 1,00$ na tabela é igual a 0,3413. Logo, a área entre 0 (zero) e $z = 1,00$ é igual 0,3413. No gráfico 9, a área entre $z = -1,00$ e 0 (zero) também é 0,3413. Isso se deve ao fato de que a curva é simétrica em $z = 0$.



Fonte: TAVARES, 2011, p. 146.

Portanto,

$$\begin{aligned}
 P(50000 < X < 70000) &= P(-1,00 < z < 1,00) \\
 &= 0,3413 + 0,3413 \\
 &= 0,6826.
 \end{aligned}$$

Existem, ainda, diversas outras distribuições de probabilidades envolvendo variáveis aleatórias contínuas. Destaque para a distribuição Qui-quadrado, t de *Student* e F de *Snedecor*. Para conhecer estas distribuições, acesse:



Distribuições Contínuas de Probabilidade

O assunto a seguir permite realizar a última etapa de um processo estatístico de análise de dados. Corresponde justamente à etapa de interpretação dos resultados e tomada de decisões. É nela que se infere, a partir de uma amostra, conclusões sobre o comportamento da população toda.

3. ESTATÍSTICA INFERENCIAL

Um dos principais objetivos da estatística inferencial é estimar os valores de parâmetros populacionais desconhecidos utilizando, evidentemente, dados a partir de uma única amostra.

Exemplo. Em um estudo antropométrico em nível nacional, uma amostra de 5000 adultos é selecionada dentre os adultos brasileiros e uma das variáveis de estudo é a altura.

Neste exemplo, a população é o conjunto de todos os brasileiros adultos e o interesse está na altura dos brasileiros. Então, para cada sujeito da população associa-se um número correspondente à altura. Se determinado sujeito é sorteado para entrar na amostra, o que interessa é a respectiva altura, o que define uma variável aleatória $X = \text{altura do adulto brasileiro}$. Como X é uma variável aleatória contínua, a ela está associada uma função densidade de probabilidade e da literatura, sabe-se que essa densidade é normal. Assim, a população é representada por v.a. $X \sim N(\mu, \sigma^2)$. Conhecidos a média e o desvio padrão ter-se informações completas sobre a população.

Pergunta: como determinar os valores de μ e σ ? A primeira maneira, e mais complexa, é medindo as alturas de todos os brasileiros adultos. Porém, seria um procedimento caro e demorado. Uma solução, então, é considerar a amostra, com reposição, obtida com 5000 adultos e determinar a altura de cada brasileiro contido na amostra. A partir da amostra, calcula-se a altura média \bar{x} e o seu desvio padrão s . Nesse caso, estas estimativas são chamadas estimativas pontuais, de modo que a média \bar{x} é utilizada para estimar a média populacional μ . O grande problema é que estimativas pontuais não dão uma ideia sobre confiança e sobre as margens de erro consideradas no problema. Não se sabe, também, se a amostra selecionada tem algum tipo de vício.

Entretanto, estima-se os valores populacionais utilizando os chamados intervalos de confiança, que dão o intervalo dentro do qual se espera que esteja o parâmetro populacional analisado, com a probabilidade ou nível de confiança.

Os intervalos de confiança fornecem uma informação mais precisa acerca do valor exato do parâmetro populacional. No entanto, para estimar parâmetros populacionais por meio de dados amostrais, é necessário o conhecimento da distribuição amostral da estatística que usada como estimador.

3.1 Distribuição amostral

Para compreender melhor o conceito de distribuição amostral, considera-se o exemplo de uma população cujos elementos são {1,5,6}. Assim, para esta população, tem-se

$$E(X) = \mu = \frac{1 + 5 + 6}{3} = 4$$

$$V(X) = \sigma^2 = \frac{1}{3}[(1 - 4)^2 + (5 - 4)^2 + (6 - 4)^2] = \frac{14}{3} \approx 4,66667.$$

Dessa população extrai-se todas as amostras aleatórias simples de tamanho 2 e determina-se a probabilidade de cada amostra e a média amostral. Em suma, analisa-se, inicialmente, a distribuição amostral das médias amostrais.

Quadro 1 - Todas as amostras de tamanho $n = 2$ da população $\{1,5,6\}$

Amostra	Probabilidade	Média Amostral \bar{x}
{1,1}	1/9	$(1+1)/2 = 1$
{1,5}	1/9	$(1+5)/2 = 3$
{1,6}	1/9	$(1+6)/2 = 3,5$
{5,1}	1/9	$(5+1)/2 = 3$
{5,5}	1/9	$(5+5)/2 = 5$
{5,6}	1/9	$(5+6)/2 = 5,5$
{6,1}	1/9	$(6+1)/2 = 3,5$
{6,5}	1/9	$(6+5)/2 = 5,5$
{6,6}	1/9	$(6+6)/2 = 6$

Fonte: Elaborado pelo autor, 2022.

Construindo a função de distribuição de probabilidade para a média amostral \bar{X} , tem-se:

Tabela 10 - Distribuição de probabilidade de \bar{X}

\bar{X}	1	3	3,5	5	5,5	6
$P(\bar{X} = \bar{x})$	1/9	2/9	2/9	1/9	2/9	1/9

Fonte: Elaborado pelo autor, 2022.

A variável de interesse é a média amostral \bar{X} . Daí, segue que

$$E(\bar{X}) = 1 \cdot \frac{1}{9} + 3 \cdot \frac{2}{9} + 3,5 \cdot \frac{2}{9} + 5 \cdot \frac{1}{9} + 5,5 \cdot \frac{2}{9} + 6 \cdot \frac{1}{9} = 4 = \mu$$

e

$$V(\bar{X}) = \frac{1}{9}(1 - 4)^2 + \frac{2}{9}(3 - 4)^2 + \frac{2}{9}(3,5 - 4)^2 + \frac{1}{9}(5 - 4)^2 + \frac{2}{9}(5,5 - 4)^2 + \frac{1}{9}(6 - 4)^2$$

$$V(\bar{X}) = 2,333333 = \frac{4,66667}{2} = \frac{\sigma^2}{n}.$$

Estes resultados mostram que, independentemente da natureza dos dados, a média de todas as amostras de uma população é igual à média populacional, isto é, $E(\bar{X}) = \mu$. Já a variância de todas as amostras desta população é igual à variância populacional dividida pelo tamanho de cada amostra considerada, $V(\bar{X}) = \frac{\sigma^2}{n}$.

Considerando a mesma situação, porém, estudando a distribuição amostral variância s^2 .

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

onde n é o tamanho da amostra.

Considera-se, também, o estimador $\hat{\sigma}^2$ definido por

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

onde N é o tamanho da amostra.

Seguindo o mesmo procedimento anterior, chega-se aos seguintes valores:

$$E(s^2) = 4,666667 = Var(X)$$

e

$$E(\hat{\sigma}^2) = 2,333333 \neq Var(X).$$

Em média, s^2 é igual à variância populacional, o que não ocorre com $\hat{\sigma}^2$. Isso justifica o fato de dividir, no cálculo da variância amostral s^2 , por $n - 1$ e não simplesmente por n . Ainda, os estimadores \bar{x} e s^2 são chamados de estimadores não-viesado ou não viciado. O desvio padrão da distribuição amostral de qualquer estatística é usualmente chamado de erro padrão. Então, o erro padrão da média amostral é definido por

$$EP(\bar{X}) = \frac{\sigma}{\sqrt{n}}$$

3.1.1 Distribuição amostral da média para populações normais

Quando há uma amostra aleatória simples de tamanho n de uma população normal, $X \sim N(\mu, \sigma^2)$, então a distribuição amostral da média amostral \bar{X} também será normal com média μ e variância σ^2/n ,

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

Afirma-se que, se X segue uma distribuição normal, então \bar{X} também segue uma distribuição normal. O exemplo a seguir ajuda a compreender melhor essa discussão.

Exemplo: A capacidade máxima de um elevador é de 500kg. Se a distribuição dos pesos dos usuários é $N(70,100)$, qual é a probabilidade de que 7 pessoas ultrapassem este limite? E de 6 pessoas?

É uma amostra aleatória simples com $n = 7$ pessoas. Se o peso máximo é 500, para que 7 pessoas ultrapassem o limite de segurança tem-se

$$\sum_{i=1}^7 X_i > 500 \Rightarrow \frac{1}{7} \sum_{i=1}^7 X_i > \frac{500}{7} \Rightarrow \bar{X} > 71,729.$$

Logo, como $X \sim N(70, 100)$ então

$$\bar{X} \sim N\left(70, \frac{100}{7}\right).$$

Daí,

$$P(\bar{X} > 71,729) = P\left(\frac{\bar{X} - 70}{\sqrt{\frac{100}{7}}} > \frac{71,729 - 70}{\sqrt{\frac{100}{7}}}\right) = P(Z > 0,46).$$

Aplicando a tabela da distribuição normal padronizada, obtém-se

$$P(\bar{X} > 71,729) = P(Z > 0,46) = 0,5 - 0,17724 = 0,32276.$$

De forma análoga, com 6 pessoas obtém-se

$$P\left(\bar{X} > \frac{500}{7}\right) = P(z > 3,27) = 0,5 - 0,49946 = 0,00054.$$

Concluindo, existe uma probabilidade alta (32%) de que 7 pessoas ultrapassem o limite de segurança. Já com 6 pessoas, essa probabilidade (0,054%) é bastante pequena. Assim, o número máximo de pessoas no elevador deve ser de 6 pessoas.

Na próxima seção, ilustra-se o mais importante resultado da estatística inferencial. O Teorema Central do Limite (TCL) é utilizado, principalmente, quando não se sabe nada a respeito da distribuição da população X .

3.2 Teorema Central do Limite

Considere uma amostra aleatória simples de uma população X qualquer tal que $E(X) = \mu$ e $V(x) = \sigma^2$. Então, a distribuição de \bar{X} converge para a distribuição normal com média μ e variância $\frac{\sigma^2}{n}$ quando n é número muito grande. Equivalentemente,

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

A interpretação prática do Teorema Central do Limite é a seguinte: para amostras grandes (geralmente $n > 30$) de qualquer população, aproxima-se a distribuição amostral de X por uma distribuição normal com a mesma média populacional e variância igual à variância populacional dividida pelo tamanho da amostra.

Para visualizar na prática a utilização do TCL, acesse:



Exemplo de aplicação: Teorema de Central do Limite

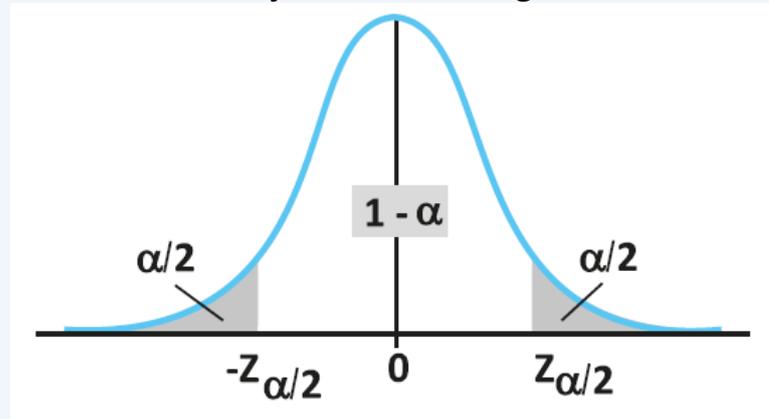
3.3 Intervalos de confiança

A estimação por intervalos permite obter um intervalo no qual o verdadeiro parâmetro populacional está contido.

Tomando informações a partir de uma amostra de uma população normal, calcula-se os limites de um intervalo, valores críticos, que em $(1 - \alpha)\%$ dos casos incluem o valor do parâmetro populacional (μ por exemplo) a estimar e em $\alpha\%$ dos casos não incluem o valor do parâmetro. O α é chamado de nível de significância e dá a medida da incerteza dessa inferência. Geralmente, assume valores entre 1 e 10%.

O nível de confiança $1 - \alpha$ é a probabilidade de o intervalo de confiança conter o parâmetro estimado. Em termos da variável normal padronizada Z , isso representa a área central sob a curva normal entre os pontos $-z$ e z , conforme o gráfico 10.

Gráfico 10 - Ilustração do nível de significância α



Fonte: TAVARES, 2011, p. 162.

A área à esquerda de $-z$ é igual a $\alpha/2$ e a área à direita de z é igual $\alpha/2$. A área total sob a curva normal é igual a 1 (um). Então, a probabilidade de Z estar entre os valores críticos de $-z_{\alpha/2}$ e $z_{\alpha/2}$ é igual a

$$P\left(-z_{\frac{\alpha}{2}} \leq Z < z_{\frac{\alpha}{2}}\right) = 1 - \alpha.$$

Isso vale para a distribuição normal padrão, em geral. Como $Z = \frac{\bar{X} - \mu}{\sqrt{n}/\sigma}$, tem-se

$$P\left(-z_{\frac{\alpha}{2}} \leq \frac{\bar{X} - \mu}{\sqrt{n}/\sigma} < z_{\frac{\alpha}{2}}\right) = 1 - \alpha,$$

que é equivalente a

$$P\left(\bar{X} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \leq \mu < \bar{X} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha.$$

Esta última expressão diz que a probabilidade da média populacional μ pertencer ao intervalo $\left[\bar{X} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}, \bar{X} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right]$ é igual a $1 - \alpha$. É claro, supõe-se

que a variância populacional aqui é conhecida, pois o interesse é a média populacional μ .

Caso a variância σ^2 não seja conhecida e/ou a população não siga uma distribuição, mas há uma amostra grande ($n > 30$) e uma população com média μ e variância σ^2 , o Teorema Central do Limite diz que a distribuição

$$\frac{\bar{X} - \mu}{S/\sqrt{n}} \approx N(0,1)$$

quando $n \rightarrow \infty$. Nesse caso, o intervalo de confiança aproximado de nível de confiança $1 - \alpha$ para μ é

$$\left[\bar{X} - z_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}, \bar{X} + z_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}} \right],$$

em que S é o desvio padrão amostral estimado para σ .

O valor de $e = z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$ é conhecido como margem de erro e representa o fato de que diferentes amostras fornecem diferentes valores de \bar{X} cuja média é igual a μ .

Exemplo: o Departamento de Recursos Humanos de uma prefeitura informa que o tempo de execução de tarefas que envolvem participação manual varia de tarefa para tarefa, mas que o desvio padrão permanece aproximadamente constante, em 3 minutos. Novas tarefas estão sendo implantadas na prefeitura. Uma amostra aleatória do tempo de execução de 50 das novas tarefas forneceu o valor médio de 15 minutos. Determine um intervalo de confiança de 95% para o tempo médio de execução de uma dessas novas tarefas.

Solução: Não se sabe se a população é normal. No entanto, há uma amostra grande ($n = 50 > 30$) e desvio padrão conhecido $\sigma = 3$ min. Precisa-se, agora, obter o valor de $z_{\frac{\alpha}{2}}$. Calculando,

$$1 - \alpha = 0,95 \Rightarrow \alpha = 0,05 \Rightarrow \frac{\alpha}{2} = 0,025.$$

Olhando na tabela de Z, encontra-se o valor de $z_{0,025} = 1,96$. Daí,

$$e = z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \Rightarrow e = 1,96 \frac{3}{\sqrt{50}} = 0,8315$$

$$P(\bar{X} - e \leq \mu \leq \bar{X} + e) = 0,95$$

$$P(15 - 0,8325 \leq \mu \leq 15 + 0,8315) = 0,95$$

$$P(14,168 \leq \mu \leq 15,831) = 0,95.$$

Interpretação do resultado: em cada 100 amostras retiradas contendo 50 pessoas cada, espera-se que, em 95 delas, a verdadeira média populacional μ esteja dentro do intervalo de 14,168 a 15,831.

3.3.1 Dimensionamento de amostras

É possível determinar o tamanho de uma amostra considerando um determinado nível de confiança e margem de erro máximo. Por exemplo, qual tamanho uma amostra deve ter para gerar um intervalo com 95% de confiança, com erro igual a 1,0, se o desvio padrão populacional é 5,0?

Utilizando a expressão para a margem de erro, isola-se n , isto é,

$$e = z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \Rightarrow n = \frac{z_{\frac{\alpha}{2}}^2 \sigma^2}{e^2}.$$

Daí,

$$n = \frac{z_{\frac{\alpha}{2}}^2 \sigma^2}{e^2} = \frac{1,96^2 5^2}{1^2} = 96,04 \cong 97.$$

3.4 Testes de hipóteses

Na etapa de análise de decisão estatística, os testes de hipóteses permitem avaliar se duas populações são, de fato, iguais ou diferentes em relação a um determinado parâmetro. Para isso utiliza-se amostras dessas populações e a tomada de decisão de um gestor público deve ser baseada na análise de dados a partir de um teste de hipótese.

Em um teste de hipótese, decide-se se uma determinada hipótese é verdadeira ou falsa, se ela deve ser aceita ou rejeitada, considerando uma determinada amostra. A seguir, elenca-se os passos que a serem seguidos para um teste qualquer:

1. definir a hipótese de igualdade (H_0) e a hipótese alternativa (H_1) para tentar rejeitar (H_0);
2. definir o nível de confiança $1 - \alpha$;
3. definir os limites da região de rejeição e de aceitação;
4. calcular a estatística da distribuição escolhida a partir dos valores amostrais obtidos e tomar a decisão;
5. se o valor da estatística da distribuição calculado estiver na região de rejeição, rejeita-se a hipótese nula. Caso contrário, aceita-se a hipótese nula ao nível de confiança especificado.

Para detalhar cada passo descrito anteriormente, considera-se o exemplo: querendo testar a hipótese de que o tempo médio de atendimento na retirada de uma guia, em uma prefeitura considerada referência, é igual a 50 segundos. Quais hipóteses formular? Primeiramente, há que estabelecer as hipóteses nula e alternativa. A hipótese nula é representada da seguinte maneira:

- $H_0: \mu = 50$ (hipótese de nulidade ou de igualdade)

Nesse caso, quer-se determinar se o tempo médio de atendimento de toda a população é igual a 50 segundos, tempo esse, obtido a partir de uma amostra. Esta hipótese será aceita ou rejeitada. Se for rejeitada, significa que a hipótese alternativa será aceita.

Alguns exemplos de hipóteses alternativas mais utilizadas:

- $H_1: \mu > 50$ (teste unilateral ou unicaudal à direita. O tempo médio de resposta de retirada da guia é superior a 50 segundos.)
- $H_1: \mu < 50$ (teste unilateral ou unicaudal à esquerda. O tempo médio de resposta de retirada da guia é inferior a 50 segundos.)
- $H_1: \mu \neq 50$ (teste bilateral ou bicaudal. O tempo médio de resposta de retirada da guia é diferente de 50 segundos.)

A dúvida que surge aqui é: como definir qual hipótese alternativa utilizar? Isso depende de cada situação. Por exemplo: uma grande de remessa de produtos que chega a uma prefeitura é inspecionada. Constata-se que 8% desses produtos são defeituosos. O fornecedor garante que não haverá mais de 6% de produtos defeituosos em cada remessa. Nesse caso, há que responder se o fornecedor está dizendo a verdade. Logo, as hipóteses a formular são:

$$H_0: p = 0,06;$$

$$H_1: p > 0,06.$$

O sinal de igual da hipótese nula H_0 corresponde a um sinal de menor ou igual, pois o teste é unilateral à direita ($p > 0,06$).

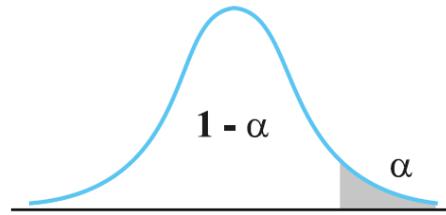
Definidas as hipóteses do teste, o próximo passo é definir o nível de significância α a ser utilizado. Geralmente, os valores mais utilizados são 0,01 e 0,05. O complementar do nível de significância é chamado de nível de confiança (ilustrado pela área clara do gráfico 11).

Gráfico 11 - Exemplificação da alteridade dos testes de hipóteses

Unilateral à direita:

$H_0: \mu = 50$

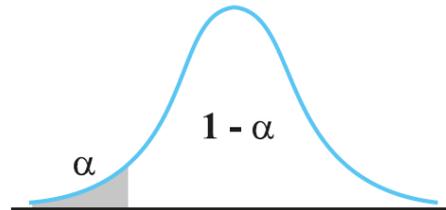
$H_1: \mu > 50$



Unilateral à esquerda:

$H_0: \mu = 50$

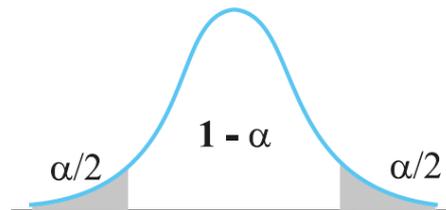
$H_1: \mu < 50$



Bilateral:

$H_0: \mu = 50$

$H_1: \mu \neq 50$



Fonte: TAVARES, 2011, p. 175.

Além disso, as áreas escuras dos gráficos correspondem à probabilidade de se cometer erro do tipo I (quando se rejeita a hipótese nula e essa hipótese é verdadeira). Determina-se o valor crítico, que separa a região de rejeição (área escura) da hipótese H_0 , da região de aceitação da hipótese H_0 .

A estatística a ser utilizada no teste é escolhida com base na distribuição amostral que os dados seguem. Nesse curso, utiliza-se a distribuição normal padronizada e a t de *Student*, quando a amostra for pequena ($n < 30$). A estrutura de cálculo da estimativa, de forma geral, é dada por

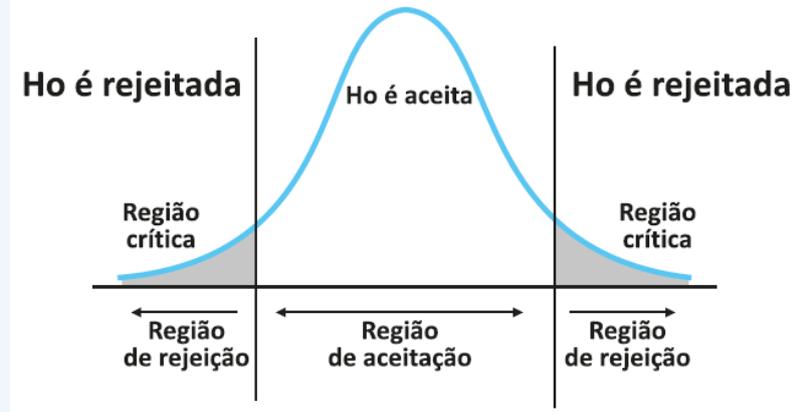
$$\text{Estatística da distribuição} = \frac{\text{estimativa} - \text{parâmetro}}{\text{erro padrão da estimativa}}$$

Um exemplo disso é dado pela distribuição normal padronizada Z , em que

$$Z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

Se o valor da estatística estiver na região crítica (rejeição), rejeita-se a hipótese H_0 , caso contrário, aceita-se H_0 . O gráfico 12 ilustra esse esquema, considerando um teste bilateral.

Gráfico 12 – Regiões de aceitação (RA) e rejeição (RR)



Fonte: TAVARES, 2011, p. 177.

Na seção a seguir, ilustra-se o teste de hipótese para uma média apenas. Existem diversos outros testes acessados em:



Teste de Hipótese

3.4.1 Testes de hipóteses para uma média

Para ilustrar a aplicação do teste de hipótese, supõe-se que a população segue uma distribuição normal com média μ e variância σ^2 . Nesta hipótese, há duas situações a considerar:

1. O desvio padrão σ é conhecido ou a amostra é grande ($n > 30$):

Quando o desvio padrão da população é conhecido ou a amostra é grande, a distribuição amostral a utilizar para a aplicação do teste é distribuição normal padronizada, cuja estatística do teste é dada por

$$Z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

onde:

\bar{x} é a média amostral;

μ é a média populacional;

σ é o desvio padrão populacional; e

n o tamanho da amostra.

Exemplo: um gestor público sabe que, para montar um determinado negócio em um bairro de Curitiba, é necessário que nele circulem, no mínimo, 1.500 pessoas por dia. Para o tipo de bairro em questão, supõe-se o desvio padrão como igual a 200 pessoas. Uma amostra aleatória formada por 12 observações revelou que passam pelo local escolhido 1.400 pessoas por dia, em média. O negócio pode ser montado ou não? Assuma $\alpha = 5\%$ e suponha população normalmente distribuída.

Solução: Primeiramente, extrai-se os dados principais do problema: $n = 12$; $\bar{x} = 1400$ e $\sigma = 200$, estabelecendo as hipóteses com base no enunciado do problema:

$$H_0: \mu = 1500;$$

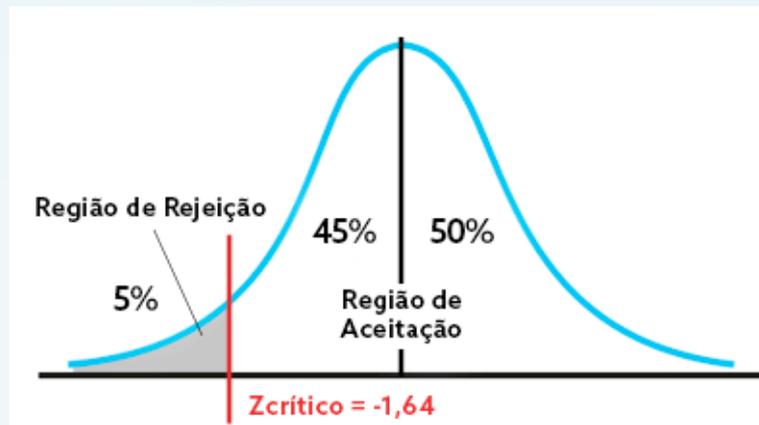
$$H_1: \mu < 1500.$$

Como a população é normal e o desvio padrão é conhecido, utiliza-se a estatística Z (distribuição normal padronizada). Assim,

$$Z_c = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{1400 - 1500}{\frac{200}{\sqrt{12}}} = -1,73$$

Considerando, agora, o valor de $\alpha = 0,05$, obtém-se o valor de Z tabelado, $Z_t = -1,64$. O teste é unilateral à esquerda, então o valor de Z_t foi obtido na tabela Z o que corresponde à área igual a 0,45. O valor mais próximo de 0,45, tabelado, é de 0,4495. O gráfico 13 ilustra mais claramente a obtenção de Z_t .

Gráfico 13 - Resolução do exemplo



Fonte: Elaborado pelo autor, 2022.

Observe que $Z_c = -1,73$ é menor do que o valor de Z tabelado, $Z_t = -1,64$. Isso significa que o valor de Z_t caiu na região de rejeição da hipótese nula ($H_0: \mu = 1500$). Assim, há fortes indícios para afirmar que transitam menos de 1500 pessoas por dia no bairro e, assim, não é viável abrir o negócio no local.

2. O desvio padrão σ não é conhecido e a amostra é pequena ($n \leq 30$):

Quando o desvio padrão da população não é conhecido e a amostra é pequena, a distribuição amostral a ser utilizada para a aplicação do teste é t de *Student*, cuja estatística do teste é dada por

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$

onde:

\bar{x} é a média amostral;

μ é a média populacional;

s é o desvio padrão amostral; e

n o tamanho da amostra.

Para conhecer a distribuição t de *Student*, foi deixado um link ao final da subseção 2.3.3. No entanto, quando se trabalha com amostras grandes, maiores que 30, a distribuição normal e a t de *Student* apresentam gráficos parecidos. A distribuição t de *Student* também tem uma tabela para obtenção das áreas (probabilidades, assim como a normal). Acesse:



Tabela t de Student

Exemplo: A Construtora Estrada Forte Ltda. alega ser capaz de produzir concreto com, no máximo, 15 kg de impurezas para cada tonelada fabricada. Mas, segundo a legislação municipal, caso essa quantidade seja maior do que 15 kg, a obra deve ser embargada pela prefeitura. Dezenove amostras de uma tonelada cada revelaram impurezas com média amostral igual a 23 kg e desvio padrão igual a 9 kg. Assumindo $\alpha = 5\%$ e população normalmente distribuída, a obra deve ser embargada ou não?

Solução: Tem-se $n = 19, \bar{x} = 23, s = 9$ e $\alpha = 0,05$. As hipóteses do problema são:

$$H_0: \mu = 15;$$

$$H_1: \mu > 15.$$

Caso a hipótese H_0 seja aceita, a obra não será embargada. Caso contrário, a prefeitura embarga a obra. Calcular a estatística considerando a distribuição t de *Student*:

$$t_c = \frac{\bar{x} - \sigma}{\frac{s}{\sqrt{n}}} = \frac{23 - 15}{\frac{9}{\sqrt{19}}} = 3,87$$

Supondo, agora, o valor de $\alpha = 0,05$, obtém-se o valor de t tabelado, $t_t = 1,734$. O teste é unilateral à direita, então o valor de t_t foi obtido procurando na tabela t considerando 18 graus de liberdade

($gl = n - 1 = 19 - 1 = 18$). Como o t calculado é maior que o t tabelado ($t_c > t_t$), o valor de t_t pertence a região de rejeição (R.R.) de H_0 , conforme o gráfico 14.

Gráfico 14 – Resolução do exemplo



Fonte: Elaborado pelo autor, 2022.

Assim, afirma-se, com 95% de certeza, que a alegação da construtora não é verdadeira, isto é, a capacidade de produção não é de 15kg, no máximo, de impurezas para cada tonelada fabricada. Portanto, a prefeitura deve embargar a obra.

REFERÊNCIAS

MORETTIN, Pedro Alberto. BUSSAB, Wilton de Oliveira. Estatística básica. 9 ed. São Paulo: Saraiva, 2017.

TAVARES, Marcelo. Estatística aplicada à administração. Florianópolis: Departamento de Ciências da Administração - UFSC, 2011.



Estatística Aplicada à Administração – Rudinei Luiz Bogo

UNIVERSIDADE ESTADUAL DO CENTRO-OESTE DO PARANÁ - UNICENTRO
NÚCLEO DE EDUCAÇÃO A DISTÂNCIA - NEAD
UNIVERSIDADE ABERTA DO BRASIL - UAB

Maria Aparecida Crissi Knuppel
Coordenador Geral UAB

Ademir Juracy Fanfa Ribas
Coordenador Geral Curso

Cleber Trindade Barbosa
Coordenador Geral NEAD

Fabíola de Medeiros
Apoio Pedagógico

Ruth Rieth Leonhadt
Revisão

Murilo Holubovski
Designer Gráfico

Robert Stump/Unsplash
Foto da Capa

05/2022