



NAB

Núcleo de Estudos em
Biomassa e Gerenciamento de Água

Manual de Estatística Básica Aplicada a Química

**-Núcleo de Estudos em Biomassa e
Gerenciamento de Água (NAB)-**

Setembro/ 2009
Versão 01

Elaborado	Aprovado	Data
Gerente da Qualidade	Coordenador do NAB	16/09/2009

Sumário

SUMÁRIO	2
1. INTRODUÇÃO.....	5
2. CONCEITOS FUNDAMENTAIS E DEFINIÇÕES	6
3. ALGARISMOS SIGNIFICATIVOS.....	10
3.1 REGRAS DE CONTAGEM DO NÚMERO DE ALGARISMOS SIGNIFICATIVOS DE UM RESULTADO	10
3.2 ARREDONDAMENTO E SUAS REGRAS	11
3.2.1 REGRAS DE ARREDONDAMENTO.....	11
3.3 OPERAÇÕES COM ALGARISMOS SIGNIFICATIVOS.....	12
3.3.1 MULTIPLICAÇÃO E DIVISÃO.....	12
3.3.2 ADIÇÃO E SUBTRAÇÃO	12
3.4 NOTAÇÃO CIENTÍFICA	13
4. MEDIDAS DE POSIÇÃO.....	14
4.1 MÉDIA ARITMÉTICA	14
4.2 MEDIANA	16
4.3 MODA.....	17
4.4 RELAÇÃO ENTRE A MÉDIA, A MEDIANA E A MODA	17
5. MEDIDAS DE DISPERSÃO	19
5.1 DESVIO PADRÃO	19
5.2 VARIÂNCIA.....	20
5.3 DESVIO MÉDIO (DM).....	21
5.4 COEFICIENTE DE VARIAÇÃO	22
5.5 AMPLITUDE.....	23
5.6 ERRO PADRÃO DA MÉDIA.....	23
6. DISTRIBUIÇÃO DE PROBABILIDADES	26
6.1 CONCEITOS BÁSICOS.....	26
6.2 DISTRIBUIÇÃO NORMAL OU GAUSSIANA	27
6.3 DISTRIBUIÇÃO BINOMIAL	34
6.4 DISTRIBUIÇÃO UNIFORME	34
6.5 DISTRIBUIÇÃO EXPONENCIAL	36
6.6 DISTRIBUIÇÃO RETANGULAR.....	38
6.7 DISTRIBUIÇÃO TRIANGULAR	38

6.8	DISTRIBUIÇÃO T DE STUDENT	38
6.9	DISTRIBUIÇÃO F DE SNEDECOR OU FISHER	40
7.	TESTES DE NORMALIDADE	41
7.1	TESTE DE KOLMOGOROV - SMIRNOV	41
7.2	TESTE DE SHAPIRO WILK.....	43
7.3	TESTE DE ANDERSON – DARLING.....	48
8.	TESTES DE HOMOGENEIDADE DE VARIÂNCIAS.....	51
8.1.	TESTE DE COCRHAN	51
8.2.	TESTE DO F MÁXIMO DE HARTLEY	52
8.3.	TESTE DE BARTLETT.....	53
9.	ANÁLISE DE VARIÂNCIA (ANOVA).....	55
9.1.	TESTE F (MESMO NÚMERO DE REPETIÇÕES).....	56
9.2.	TESTE DE TUKEY (COMPARAÇÃO DE MÉDIAS)	60
9.3.	ANÁLISE DE VARIÂNCIA (NÚMEROS DE REPETIÇÕES DIFERENTES)	60
10.	TESTE DE HIPÓTESES	64
10.1.	COMPARAÇÃO DE MÉDIAS, VARIÂNCIA CONHECIDA	66
10.2	COMPARAÇÃO DE MÉDIAS, VARIÂNCIA DESCONHECIDA	71
10.3	TESTE DE HIPÓTESES PARA A MÉDIA (DESVIO PADRÃO DESCONHECIDO)	72
10.4	TESTE DE HIPÓTESES PARA COMPARAÇÃO DE MÉDIAS (INDEPENDENTES)	73
11.	TEORIA DOS ERROS.....	84
11.1	ERROS ALEATÓRIOS	85
11.2	ERROS SISTEMÁTICOS	86
11.3	ERROS GROSSEIROS	87
11.4	OUTROS TIPOS DE ERROS	88
11.5	PREVENÇÃO DE ERROS SISTEMÁTICOS	91
11.6.	PRINCIPAIS FONTES DE ERRO NA MEDIÇÃO	93
11.7	FLUTUAÇÕES NAS MEDIDAS	93
12.	OUTLIERS.....	94
12.1	IDENTIFICAÇÃO DE OULIERS.....	94
12. 2	DEFINIÇÃO DE OUTLIERS	95
13.	MÉTODOS DE IDENTIFICAÇÃO DE OUTLIERS.....	96
13.1	GRÁFICO DE BOX-PLOT.....	96

13.2 TESTE DE DIXON.....	97
13.3 TESTE DE GRUBBS.....	100
13.4 COCHRAN	106
13.5 Z-SCORE.....	108
13.6 O CRITÉRIO DE CHAUVENET	108
14. REFERÊNCIAS BIBLIOGRÁFICAS	111
15. SUMÁRIO DE REVISÕES	112

1. Introdução

Este manual tem por objetivo orientar e proporcionar aos colaboradores do NAB um aprendizado dinâmico e simplificado no estudo de Estatística.

A Estatística é uma ferramenta indispensável, sendo aplicável em qualquer área de conhecimento, principalmente na área de química, sendo uma ciência que visa fornecer subsídios ao analista para coletar, organizar, descrever, resumir, analisar e interpretar dados. Esses dados são utilizados na tomada de decisões.

Normalmente, no trabalho estatístico, o pesquisador se vê obrigado a lidar com uma grande quantidade de valores numéricos resultantes de uma estimação ou da realização de um experimento.

Estes valores numéricos são chamados dados estatísticos. A estatística ensina métodos racionais para a obtenção de informações a respeito de um fenômeno coletivo, além de obter conclusões válidas para o fenômeno e também permitir decisões, através de dados estatísticos observados.

Desta forma, a estatística pode ser dividida em três áreas principais:

Estatística Descritiva: É a etapa inicial da análise utilizada pra descrever os dados. A disponibilidade de uma grande quantidade de dados e de métodos computacionais muitos eficientes revigorou esta área da Estatística.

Inferência Estatística: Procura com base nos dados amostrais tirar conclusões sobre a população. O termo população é usado em estatística como sinônimo de conjunto.

Probabilidade: As teorias das probabilidades nos permitem descrever os fenômenos aleatórios, ou seja, aqueles em que está presente a incerteza.

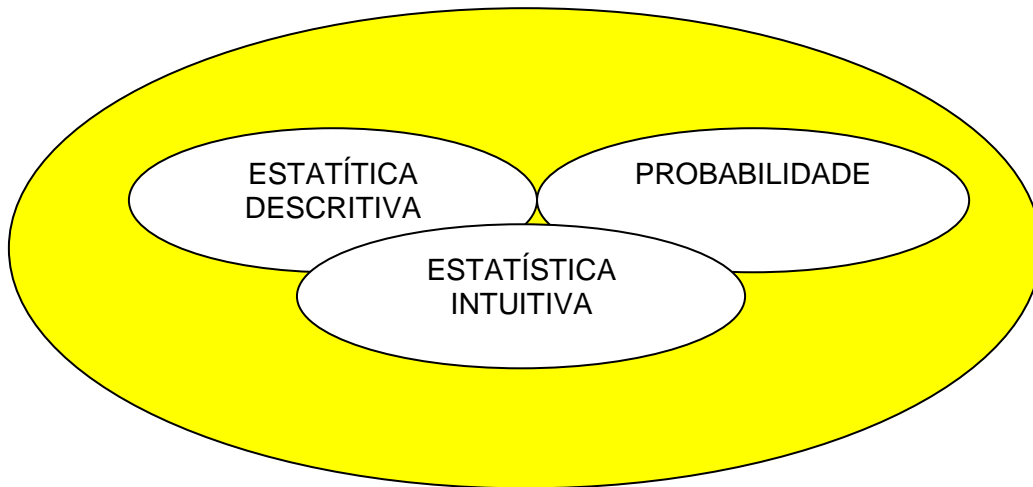
Dois importantes conceitos devem ser apresentados aqui: a estatística indutiva ou inferencial que busca obter conclusões para toda população p , a partir dos resultados encontrados na amostra retirada deste mesmo universo; ou seja, ela generaliza o resultado da amostra para o universo da população.

Mas, ao efetuarmos esta generalização, podemos notar que um erro, também chamado de desvio ou incerteza poderá ocorrer. Assim, como descrito anteriormente, o ramo da estatística que estuda essa situação é a probabilidade.

Conforme Costa Neto (1992):

“A probabilidade é um número associado a um evento, destinado a medir sua possibilidade de ocorrência”.

Os ramos da estatística interagem da seguinte forma:



Amostragem aleatória é uma forma de se obter a amostra no qual todos os elementos da população têm igual probabilidade de ocorrer.

O desenvolvimento e o aperfeiçoamento de técnicas estatísticas de obtenção e de análise de informações permitem o controle e o estudo adequado de fenômenos, fatos, eventos e ocorrências, em diversas áreas de conhecimentos. A estatística tem por objetivo fornecer métodos e técnicas para se lidar, racionalmente, com situações sujeitas a incertezas.

2. Conceitos Fundamentais e Definições

α : *É a probabilidade de erro na estimação por intervalo. No caso de teste de hipóteses, é a probabilidade de Rejeitar H_0 , quando ela é verdadeira. Neste caso, seria a probabilidade de afirmar que as atitudes em relação à Matemática interferem na formação das atitudes em relação à Estatística, quando na realidade não existe nenhuma relação.*

Amostra: Parcela representativa da população que é examinada com o objetivo de se tirar conclusões sobre a população em questão.

Amostragem: Uma técnica especial para recolher as amostras, que garante, tanto possível, o acaso na escolha, de modo a garantir à amostra, o caráter de representatividade.

Atributo: Quando os dados estatísticos apresentam um caráter qualitativo, o levantamento e os estudos necessários ao tratamento desses dados são designados genericamente de estatística de atributos.

Dado estatístico: É o valor numérico e será considerado como o responsável sobre o qual será aplicado o método estatístico.

Dados brutos: Dados originais que ainda não foram numericamente organizados após a coleta.

Distribuição de frequência: É um agrupamento de dados em classes, exibindo o número ou percentagem de observações em cada classe, pode ser apresentada sob forma gráfica ou tabular.

Cada um dos grupos ou intervalos da distribuição intervalar, formados a partir do agrupamento ou conjunto destes dados, é chamado de **Classe**.

Distribuição de frequência sem intervalos de classes: Tabela onde os valores da variável aparecem individualmente.

Ex:

Nº de aparelhos com defeitos	Nº de meses
0	06
1	11
2	09
3	08
4	08
5	05
6	01
Total	48

Distribuição de frequência com intervalos de classes: Tabela onde os valores da variável não aparecem individualmente, mas agrupados em classes.

Ex:

Notas	Nº de alunos
0 --- 20	020
20 --- 40	065
40 --- 60	230
60 --- 80	160
80 --- 100	025
Total	500

Estatística: Um conjunto de técnicas e métodos de pesquisa que entre outros tópicos envolve o planejamento dos experimentos a serem realizados, a coleta dos dados, a inferência, o processamento, a análise e a disseminação das informações. O desenvolvimento e o aperfeiçoamento de técnicas estatísticas de obtenção e análise de informações, permitem o controle e o estudo de fenômenos, fatos, eventos e ocorrências em diversas áreas de conhecimento. Ela fornece métodos e técnicas para lidarmos, racionalmente, com situações sujeitas a incertezas.

Erro: é a diferença entre a quantidade *medida* e o valor verdadeiro.

Estimativa: Valor aproximado do parâmetro e é calculado com o uso da amostra

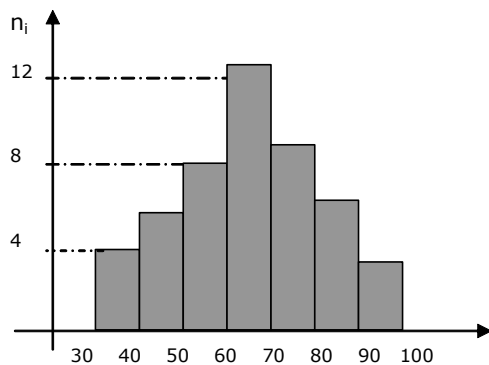
Exatidão: É a característica de um instrumento de medida que exprime o afastamento entre a medida nele observada e o valor de referência aceito como verdadeiro.

Frequência relativa de uma observação da série é o número de repetições dessa observação dividida pelo tamanho da amostra.

Frequência de uma observação da série é o número de repetições dessa observação.

Uma distribuição de frequência pode ser representada graficamente pelo histograma.

Histograma: É a representação gráfica da distribuição de frequências de uma massa de medições, normalmente em um gráfico de barras verticais. É composto por retângulos justapostos em que a base de cada um deles corresponde a um intervalo de classes e a sua altura à respectiva frequência.



Graus de liberdade: É um conceito ligado ao número de dados disponíveis (livres) para o cálculo da estatística. Por exemplo, ao estimarmos a média populacional, com a média amostral perdemos um grau de liberdade, assim a estatística t-student terá $n-1$ graus de liberdade. No caso da Tabela de ANOVA, os graus de liberdade do grupo são iguais ao número de grupos menos 1, os graus de liberdade total serão iguais a $(n-1)$ e os graus de liberdade do resíduo, a diferença entre esses dois.

Média populacional (μ): É o valor que representa um conjunto de valores da população. Definida como a soma de todos os valores da população dividido pelo número de observações. Por exemplo: renda *per capita* de um país, esperança de vida, renda familiar média, pontuação média na escala de atitudes em relação à Estatística, etc.

(1- α) Nível de confiança: É a probabilidade de que o intervalo de confiança contenha o verdadeiro valor do parâmetro.

Nível de significância α (alpha): É definida como a probabilidade de cometer o erro de tipo I, ou seja, rejeitar a hipótese nula (H_0), quando ela é verdadeira. Por exemplo:

H_0 : As atitudes em relação à Matemática **não interferem** na formação das atitudes em relação à Estatística ($\rho_{xy} = 0$)

H_1 : As atitudes em relação à Matemática **interferem** na formação das atitudes em relação à Estatística. ($\rho_{xy} \neq 0$)

População: Conjunto total de elementos portadores de pelo menos uma característica em comum.

Parâmetro: É uma medida usada para descrever, de forma resumida, uma característica da população, Por exemplo, a média populacional (μ), a proporção populacional (π), a variância populacional (σ), o coeficiente de correlação (ρ), etc. Os parâmetros, via de regra, são valores desconhecidos e desejamos estimar ou testar, a partir dos dados de uma amostra.

p-valor ou p-value: É a probabilidade de cometer o erro de tipo I (rejeitar H_0 quando ela é verdadeira), com os dados de uma amostra específica. Este valor é dado pelo pacote estatístico, assim o comparamos com o nível de significância escolhido e tomamos a decisão. Se o p-valor for menor que o nível de significância escolhido, rejeitamos H_0 , caso contrário, aceitamos H_0 .

Variável: Convencionalmente, o conjunto de resultados possíveis de um fenômeno.

As variáveis são classificadas em qualitativas e quantitativas.

Variáveis Qualitativas: Aquelas para as quais uma medição numérica não é possível (também chamadas de categóricas).

As variáveis **qualitativas** são classificadas em: **nominal** (quando seus valores são expressos por atributos, como por exemplo, sexo, cor dos olhos, tipo de pele) e **ordinal** (quando, apesar de não ser medida, a variável segue uma ordem, como por exemplo, patente, escolaridade, hierarquia no serviço público civil).

Variáveis Quantitativas: Aquelas para as quais é possível realizar uma medição numérica, tratam-se, portanto da estatística de variável.

As variáveis quantitativas são classificadas em **discretas** (resultam de um conjunto enumerável de valores, ou seja, resulta normalmente de contagens, aceitando apenas números inteiros, como por exemplo, número de filhos de um casal, quantidade de carros em um estacionamento) e **contínuas** (são aquelas originadas de medições, deste modo, podem assumir qualquer valor real entre dois extremos, como por exemplo, teor de enxofre em óleo diesel, temperatura, teor de óleo em uma determinada biomassa).

Tabela primitiva ou dados brutos: É uma tabela ou relação de elementos que não foram numericamente organizados, sendo difícil formar uma idéia exata do comportamento do grupo como um todo, a partir de dados não ordenados. Ex: 45, 41, 42, 41, 42, 43, 44, 41, 50, 46, 50, 46, 54, 52, 58, 57, 58, 60, 51.

Rol: Tabela obtida após a ordenação dos dados (crescente ou decrescente). Ex: 41, 41, 42, 42, 43, 44, 45, 46, 50, 50, 51, 52, 57, 58, 58, 60, 60.

Valor verdadeiro: É o valor exato da medida de uma grandeza obtido quando nenhum tipo de erro incide na medição. Na prática é impossível eliminar todos os erros e obter um valor aceito como verdadeiro. Utiliza-se uma medida de uma amostra de um determinado número de medidas técnicas, usando o mesmo material e mantendo-se as mesmas condições ambientais, usando então este valor como verdadeiro.

3. Algarismos Significativos

Ao realizarmos uma medida com um instrumento de medição, estamos sujeitos a incertezas do valor obtido. Essa medida é apresentada pelos números de algarismos significativos, que são formados pelos algarismos que temos certeza (valor lido no instrumento de medição ou equipamento) mais um algarismo duvidoso (“chutado”), que compreende um valor entre dois que temos certeza (provenientes da menor subdivisão do aparelho de leitura do instrumento).

3.1 Regras de Contagem do Número de Algarismos Significativos de um Resultado

A contagem dos algarismos significativos faz-se da esquerda para a direita, começando pelo primeiro diferente de zero.

- Qualquer algarismo diferente de zero é significativo,
- Zeros entre algarismos diferentes de zero são significativos,
- Para números superiores a 1, os zeros à direita da vírgula contam como algarismos significativos,
- As potências na base 10 não são significativas,
- Zeros a esquerda de um número superior a 1 não são significativos.

Exemplos

134g → 3 a.s

3005m → 4 a.s.

0,000456g → 3 a.s.

34,000g → 5 a.s

$5 \times 10^2 \rightarrow 1$ a.s

$5,0 \times 10^2 \rightarrow 2$ a.s

Nota: a. s. →algarismos significativos

3.2 Arredondamento e suas regras

É comum a redução do número de algarismos significativos de certo valor numérico. Essa operação é chamada de “arredondamento” e tem como objetivo, manter a precisão numérica sobre controle.

3.2.1 Regras de Arredondamento

Escolhida a casa decimal até onde se quer fazer a aproximação:

1. Despreze o algarismo seguinte se for inferior a 5:

Ex: 1,56849 = 1,568

2. Acrescente uma unidade à casa decimal, se o algarismo for superior a 5.

EX: 2,5698 = 2,57

3. Se o algarismo seguinte à casa escolhida for igual a 5, tem duas situações:

- Se o número da casa principal que pretende arredondar é par, permanece como está:

EX: 1,85 = 1,8

- Se o número da casa decimal que pretende arredondar é ímpar, acrescentar uma unidade:

EX: 2,735 = 2,74

3.3 Operações com Algarismos Significativos

3.3.1 Multiplicação e Divisão

Manter o resultado com uma quantidade de algarismos idêntica à da grandeza com menor número de dígitos significativos.

EX: $2,3 \times 3,1416 \times 245 = 1,8 \times 10^3$

O produto dos três números deu como resultado, em notação científica, $1,7702916 \times 10^3$. Todavia, o resultado final deve ser emitido com apenas dois algarismos significativos (o valor numérico 2,3 possui o menor número de algarismos significativos).

O número 1,7702916 foi arredondado para 1,8 porque seu terceiro dígito (7) é maior do que 5.

3.3.2 Adição e Subtração

1. Vamos supor que você queira fazer a seguinte adição:

$$250,657 + 0,0648 + 53,6 =$$

Verificar qual parcela apresenta o menor número de casas decimais, no caso 53,6 que apresenta apenas uma casa decimal. Esta parcela será mantida e as demais serão aproximadas para uma casa decimal.

São feitas as seguintes aproximações:

$$250,657 \approx 250,6$$

$$0,0648 \approx 0,1$$

O resultado final é o seguinte, após a adição dos números aproximados:

$$250,6 + 0,1 + 53,6 = 304,3 \text{ cm}$$

2. Quando se tratar da soma de potências de 10, expressar o resultado na maior potência. Verificar qual desses números tem o algarismo duvidoso de maior ordem. O algarismo duvidoso do resultado da adição e/ou subtração estará nessa mesma ordem.

Ex: (a) $2,247 \times 10^3 + 3,25 \times 10^2 = (2,247 + 0,325) \times 10^3 = 2,572 \times 10^3$

Neste exemplo, os algarismos duvidosos em cada uma das parcelas pertencem à mesma ordem, à dos milésimos.

$$(b) 3,18 \times 10^4 + 2,14 \times 10^2 = (3,18 + 0,0214) \times 10^4 = 3,20 \times 10^4$$

Observe que os algarismos duvidosos em 3,18 e 0,0214 pertencem a ordens distintas, respectivamente centésimos e décimos de milésimos. Neste caso, o resultado da soma será significativo até a ordem dos centésimos apenas:

$$3,18 + 0,0214 \rightarrow 3,18 + 0,02 = 3,20$$

$$(c) 2550,0 + 0,75 = 2550,8$$

Aqui o número 0,75 foi arredondado para 0,8. Observe que os algarismos duvidosos em 2550,0 e 0,75 também pertencem a ordens distintas, décimos e centésimos, respectivamente. O resultado da soma será significativo até a ordem dos décimos.

3.4 Notação Científica

É um procedimento matemático de escrever medidas, tais como, área, volume, temperatura, etc. que possibilita trabalhar com números muito grandes ou muito pequenos.

Regra Prática

- Para números **maiores** que 1, a vírgula é deslocada para a esquerda até atingirmos o primeiro algarismo do número, sendo que o número de casas decimais deslocadas para a esquerda corresponderá ao expoente positivo da potência de 10.

$$\text{EX: } 2000 = 2 \times 10^3 \\ 762500 = 7,625 \times 10^5$$

- Para números **menores** que 1, a vírgula é deslocada para a direita até atingirmos o primeiro algarismo diferente do zero, sendo que o número de casas decimais deslocadas para a direita corresponderá ao expoente negativo da potência de 10.

$$\text{EX: } 0,0008 = 8 \times 10^{-4} \\ 3,45 = 3,45 \times 10^{-7}$$

4. Medidas de Posição

As medidas de posição, ou medidas estatísticas de tendência central são valores calculados para um conjunto de dados numéricos e utilizados de alguma forma, para descrever estes dados, e geralmente localizadas em torno do centro deste conjunto de dados, estando eles ordenados. As mais importantes medidas de tendência central são: a média aritmética, a mediana e a moda.

4.1 Média Aritmética

É definida como sendo o quociente da soma dos valores de um grupo de dados numéricos pelo número de valores.

Simbologia da média aritmética populacional: μ

Simbologia da média aritmética amostral: \bar{X}
- Para dados não agrupados

Para a **população**, a média é definida como:

$$\mu = \frac{\sum X}{N}$$

Onde N é o nº de observações, ou nº de valores do conjunto de dados e $\sum X$ pode ser definido como o **total** dos valores dos dados.

Para a **amostra**, a média é definida como:

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

Onde:

\bar{X} = média aritmética;

X_i = os valores da variável;

n = o número de valores.

Exemplo: Se X: 2, 0, 5, 3; então:

$$\bar{X} = \frac{2+0+5+3}{4} = 2,5$$

- Para dados agrupados sem intervalos de classe (média aritmética ponderada)

Se os dados estão apresentados na forma de uma variável discreta, faremos a média aritmética ponderada considerando as frequências simples de f_i como sendo as ponderações dos elementos x_i correspondentes:

Assim a fórmula para o cálculo da média é:

$$\bar{X} = \frac{\sum x_i f_i}{\sum f_i}$$

Exemplo: Considerando a distribuição:

x_i	f_i
2	1
4	3
5	2

$$\bar{X} = \frac{\sum x_i f_i}{\sum f_i} = \frac{2.1 + 4.3 + 5.2}{1 + 3 + 2} = \frac{2 + 12 + 10}{6} = 4$$

- Para dados agrupados com intervalos de classe

Quando os dados estão agrupados, por convenção, as frequências se distribuem uniformemente ao longo da classe e que, portanto, o ponto médio da classe é o valor representativo do conjunto.

Neste caso, a média será calculada fazendo a média aritmética ponderada considerando as frequências simples de f_i como sendo as ponderações dos elementos \bar{X}_i correspondentes, onde \bar{X}_i é o ponto médio do intervalo. Assim, a fórmula para o cálculo da média é:

$$\bar{X} = \frac{\sum \bar{x}_i f_i}{\sum f_i}$$

Exemplo: Considere a distribuição:

Classe	f_i	x_i
[180, 200[4	190
[200, 220[18	210
[220, 240[10	230
[240, 260[5	250
[260, 280[3	270

$$\bar{X} = \frac{\sum x_i f_i}{\sum f_i} = \frac{190.4 + 210.18 + 230.10 + 250.5 + 270.3}{4 + 18 + 10 + 5 + 3} = \frac{8900}{40} = 222,50$$

4.2 Mediana

A mediana de um conjunto ordenado de valores é o valor do meio deste conjunto, ou o valor médio dos dois valores centrais. Ela divide o grupo ordenado de valores em 2 partes iguais (50% acima e 50% abaixo da mediana).

Se o número de itens é par, a mediana será a média dos 2 valores do meio. Se o número de itens for ímpar, a mediana será o valor do meio.

Simbologia da mediana populacional: μ_d

Podemos determinar a **posição da mediana** pela fórmula: $(N + 1) / 2$

Exemplos

1) **Dada a série:** 37, 28, 40, 41, 45, 37, 37, 41, 44.

Colocando os dados em **rol** temos: 28, 37, 37, 37, 40, 41, 41, 44, 45.

A distribuição tem um número ímpar (9) de dados. Há quatro valores à esquerda de 40 e quatro valores à direita de 40. Dizemos que o valor central dessa distribuição, 40, é a mediana.

$$M_d = 40$$

2) **Dada a série:** 25, 27, 28, 30, 32, 34, 38, 40.

Tem um número par (8) de elementos, não existe um valor central, mas dois valores centrais:

25, 27, 28, 30, 32, 34, 38, 40.

Neste caso, mediana será a média aritmética dos valores centrais:

$$M_d = \frac{30 + 32}{2} = 31 \Rightarrow M_d = 31$$

4.3 Moda

A moda de uma série de valores é o valor de maior frequência absoluta, ou seja, o valor que aparece o maior número de vezes na distribuição.

Podemos ter conjuntos de dados com uma moda (unimodal), 2 modas (bimodal) e várias modas (multimodais). Pode não haver moda no conjunto de dados, quando não há valores repetidos no grupo. Nesse caso, dizemos que o conjunto de dados é **amodal**.

Simbologia da moda populacional: μ_o

Exemplos: Verificar o valor da moda, para os seguintes conjuntos de dados:

- a) 12, 18, 20, 15, 12, 19, 15, 12. >>> $\mu_o = 12$.
- b) 15, 19, 21, 12, 15, 21, 17, 14. >>> $\mu_o = 15$ e $\mu_o = 21$.
- c) 12, 16, 13, 18, 20, 14, 25, 11 >>> amodal (não existe valor mais presente).
- d) 2, 0, 0, 5, 3 >>> $\mu_o = 0$.
- e) 1, 1, 2, 2, 2, 3, 3, 4, 5, 5, 6 >>> $\mu_o = 2$ e $\mu_o = 3$ (existem dois valores modais, a série é bimodal).

Para distribuições de frequências por intervalos não é possível determinarmos o valor exato da moda. Nesse caso, utilizamos a chamada moda bruta >>>ponto médio da classe de maior frequência.

4.4 Relação entre a média, a mediana e a moda

a) Principais características da média

A média reflete cada valor, possuindo propriedades matemáticas atraentes. É influenciada por valores extremos, portanto recomenda-se a eliminação de *outliers* (ou discrepantes ou aberrantes), antes de se efetuar um cálculo matemático. É recomendada sua utilização quando houver possibilidade de tratamento algébrico.

É a medida de tendência central de maior emprego, tendo grande aplicação nas distribuições simétricas.

b) Principais características da mediana

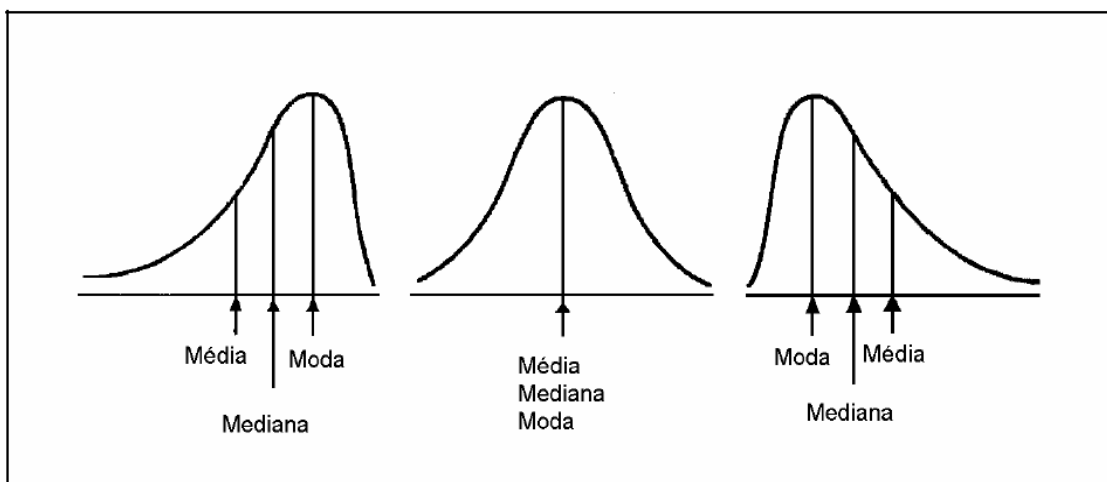
A mediana é menos sensível a valores extremos que a média, sendo utilizada para se obter o ponto que divide o conjunto em partes iguais. É uma medida separatriz, ou seja, são medidas que "separam" ou que dividem o conjunto em certo número de partes iguais. É influenciada pela posição ou localização dos valores e não pela magnitude deles.

c) Principais características da moda

A moda não é utilizada para análise matemática, devido ao fato de que pode não haver moda ou haver mais de uma moda para um dado conjunto de dados. É utilizada quando se deseja obter uma medida rápida e aproximada da posição. A medida da posição deve ser o mais típico da distribuição. É menos estável do que a média e apresenta grande instabilidade na amostra.

Em geral, a moda é menos empregada, sendo adequada para caracterizar situações onde estejam em causa os valores mais frequentes. Por exemplo, em estudos de mercado, o principal interesse pode ser pesquisar os produtos que mais se vendem.

A escolha é feita entre a média e a mediana. Em alguns casos, a mediana é mais resistente do que a média, principalmente nas séries assimétricas, por não ser influenciada pelos valores extremos. Já a média tem vantagens quando a curva de frequência é mais ou menos simétrica, transformando-a num bom estimador de parâmetros populacionais, através de dados amostrais.



Quando a distribuição é simétrica, sua média, moda e mediana coincidem.

Conforme, verificado na figura anterior, nas curvas de frequência, as diferenças entre os valores da média, mediana e moda são indicadores de forma da curva em termos de simetria, quando: média<mediana<moda ->>> assimétrica negativa,
 média=mediana=moda ->>> curva simétrica,
 média>mediana>moda ->>> assimétrica positiva.

5. Medidas de Dispersão

A função dessas medidas é avaliar quanto os valores observados numa distribuição de frequência ou de probabilidades estão dispersos, ou seja, o grau de afastamento ou de concentração entre os valores. As principais medidas de dispersão são:

- Desvio Padrão,
- Variância,
- Desvio Médio,
- Coeficiente de Variação.
- Amplitude ou *Range*,
- Erro padrão da Média.

5.1 Desvio Padrão

Esta medida de dispersão é a mais empregada, podendo ser considerada como uma medida de variabilidade dos dados de uma distribuição de frequências, isto é, mede a dispersão dos valores individuais em torno da média. As fórmulas utilizadas para o cálculo do desvio padrão são as seguintes:

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}} \quad (\text{para a população})$$

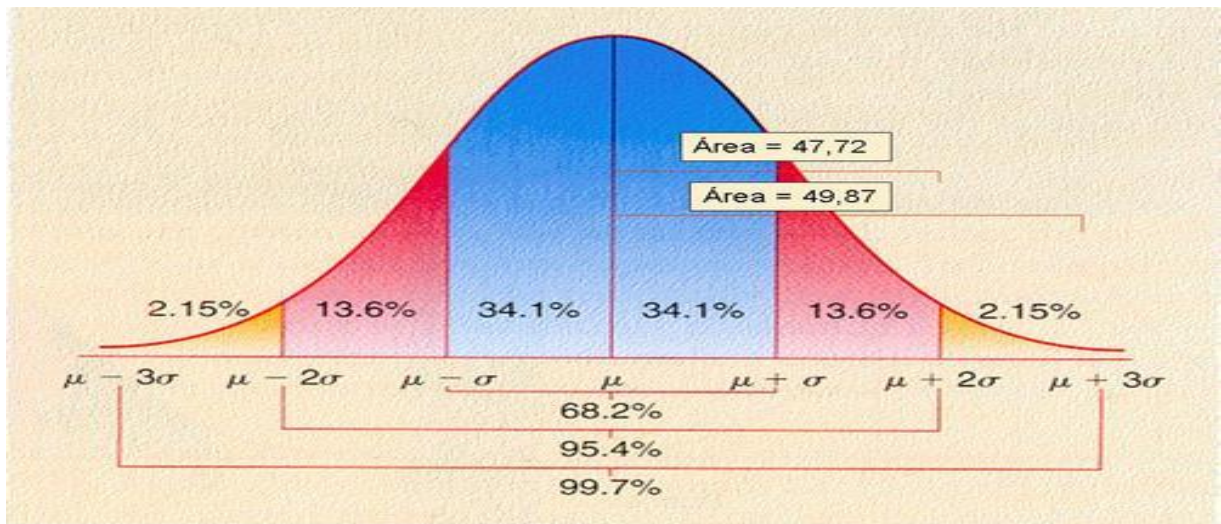
$$S = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}} \quad (\text{para a amostra})$$

Ela leva em consideração a totalidade dos valores da variável em estudo, sendo um conceito imprescindível para análises gráficas, determinação de confiabilidade e estudos de distribuições.

Interpretação do desvio padrão

Quanto *maior for* à amostra, maior a confiança que podemos depositar nesse processo de estimação, tanto para a média como para o desvio padrão. Assim, se aumentarmos o tamanho da amostra, não podemos dizer que o desvio padrão irá aumentar nem que irá diminuir. Mas apenas que o desvio padrão amostral irá estimar com maior confiança o desvio padrão populacional.

A interpretação do desvio padrão em determinada amostra vai avaliar o intervalo (2 desvios padrões, por exemplo) da média daquela amostra, o qual pode ser negativo ou positivo. Com esses dados, torna-se possível interpretar a probabilidade daquele evento acontecer ou não em uma determinada população. Numa distribuição normal, 1 desvio padrão indica que $\approx 68\%$ desses valores estão dentro deste intervalo; 2 desvios padrões correspondem a $\approx 95\%$ e 3 desvios padrões a $\approx 99,7\%$, respectivamente.



Por exemplo, assumindo que uma série estatística relativa a alguma medida de uma população e cujos valores tem média de 100 e desvio padrão de 10. De acordo com o que foi colocado anteriormente, podemos dizer que $\approx 68\%$ da amostra têm valores entre 90 ($100 - 10$) e 110 ($100 + 10$); da mesma forma, podemos dizer que $\approx 95\%$ da amostra têm valores que se situam entre 80 ($100 - 2 \times 10$) e 120 ($100 + 2 \times 10$) e finalmente, $\approx 99,7\%$, situa-se entre 70 ($100 - 3 \times 10$) e 130 ($100 + 3 \times 10$).

5.2 Variância

A variância é uma medida que expressa um desvio quadrático médio. A unidade da variância é, portanto o quadrado dos dados originais.

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

$$S^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1} \quad (\text{para amostras})$$

Nota

O denominador “n – 1” da variância é chamado de graus de liberdade. O princípio dos graus de liberdade é constantemente utilizado na estatística. Considerando um conjunto de “n” observações (dados) e fixando uma média para esse grupo, existe a liberdade de escolher os valores numéricos de n - 1 observações, o valor da última observação estará fixado para atender ao requisito de ser a soma dos desvios da média igual a zero. No caso específico do cálculo da variância, diz-se que os “n” graus de liberdade originalmente disponíveis no conjunto sofreram a redução de uma unidade porque uma estatística, a média já foi calculada dos dados do grupo e aplicada na determinação da variância.

Propriedades da variância

- Somando-se (ou subtraindo-se) a cada elemento de um conjunto de valores uma constante arbitrária, o desvio padrão não se altera.
- Multiplicando-se (ou dividindo-se) cada elemento de um conjunto de valores por um valor constante, desvio padrão fica multiplicado (ou dividido) pela constante.

Nota

Deseja-se comparar a variabilidade de duas ou mais distribuições, mesmo quando essas se referem a diferentes fenômenos e sejam expressas em unidades de medidas distintas, podemos utilizar o coeficiente de variação de Pearson ou CV (medida de dispersão relativa).

5.3 Desvio Médio (DM)

O desvio médio é a média dos desvios dos valores a contar da média, sendo uma medida de dispersão dos dados em relação à média de uma seqüência, “afastamento” em relação a essa média:

$$DM = \frac{\sum |X - \bar{X}|}{N}$$

Onde

DM = desvio médio

X = valor de cada resultado

N = tamanho da amostra

\bar{X} = média dos resultados.

O desvio médio é apenas uma medida do erro médio de cada ponto em relação à média. Ele está menos suscetível as não uniformidades da distribuição de frequências, enquanto o desvio padrão, por ser quadrático, leva mais em conta estas assimetrias.

5.4 Coeficiente de Variação

O coeficiente de variação mede a homogeneidade dos dados em conjunto, em relação à média, sua fórmula é expressa por:

$$CV = \frac{S}{\bar{X}} \times 100$$

O valor obtido geralmente é dado em porcentagem.

Dependendo do valor encontrado, o CV pode indicar como está a dispersão dos dados, ou seja,

$CV < 0,2 \rightarrow$ pouca dispersão nos dados,

$CV > 0,3 \rightarrow$ o conjunto de dados é considerado heterogêneo,

$CV > 0,3 \rightarrow$ o conjunto de dados é considerado homogêneo.

$CV > 0,5 \rightarrow$ a distribuição analisada tende a uma forma assimétrica ou não normal,

$CV > 1,0 \rightarrow$ dispersão bastante elevada nos dados

Deste modo, é possível comparar a dispersão entre duas variáveis, mesmo que tenham sido mensuradas em escalas de medidas diferentes e possuam médias diferentes. Por exemplo, através do coeficiente de variação, pode-se comparar diretamente, sem o recurso de transformações, a variabilidade existente em uma distribuição de alturas medidas em metros e outra em milímetros.

Nota

O CV é também chamado de coeficiente de variação de Pearson.

Por exemplo, tomemos os resultados das estaturas e dos pesos de um mesmo grupo de indivíduos: Estatura - 1,75 cm e 5,0 cm.

Pesos – 68 kg e 2,0 kg.

Qual das medidas (Estatura ou Peso) possui maior homogeneidade?

Após o cálculo da média e dos desvios padrões foram calculados os respectivos CV:

$$CV_{\text{estatura}} = (5 / 175) \times 100 = 2,85 \%$$

$$CV_{\text{peso}} = (2 / 68) \times 100 = 2,94 \%$$

Logo, nesse grupo de indivíduos, as estaturas apresentam menor grau de dispersão que os pesos.

5.5 Amplitude

A amplitude é uma medida de dispersão que pode ser definida como a diferença entre o valor maior e o valor menor de um grupo de observações. É a medida de dispersão mais simples, sendo a única medida que não tem média e é simbolizada por R (*range*) ou A (amplitude).

$$A = (X_{MAX} - X_{MIN})$$

Nota

Não é uma medida completa de variação porque seu cálculo utiliza apenas as variações extremas (dado o maior e menor do grupo), não sendo avaliados os valores intermediários.

O uso desta medida de dispersão é devido a estabelecer parâmetros de controle de um determinado processo de medição, ou seja, na construção de cartas de controle para controle de qualidade, onde são registrados os valores máximos de variação de um determinado produto ou processo.

5.6 Erro Padrão da Média

O erro padrão mede a discrepância entre a média calculada e a amostra verdadeira da população.

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{N}}$$

σ = desvio padrão

N = número total de amostras

Se $\sigma_{\bar{x}}$ é pequeno, há boa chance de a média de uma amostra estar próxima da média da população.

Se $\sigma_{\bar{x}}$ é grande, é mais provável que obtenhamos uma média amostral consideravelmente diferente da média da população.

Quanto maior o tamanho da amostra, N , menor será o erro padrão da média.

Por exemplo, o teste de QI é padronizado de modo a ter média populacional (μ) de 100 e um desvio padrão populacional (σ) de 15. Se considerarmos um tamanho amostral de 10, a média amostral estaria sujeita a um erro padrão de:

$$\sigma_{\bar{x}} = \frac{15}{\sqrt{10}} = 4,74$$

Enquanto, a população de escores de QI tem um desvio padrão de 15, a distribuição amostral da média da amostra para $N=10$ tem um erro padrão (desvio padrão teórico) de 4,74.

Como é possível estimarmos a verdadeira média populacional com base em uma única média amostral?

Com o auxílio do erro padrão da média, podemos achar o intervalo de valores das médias no qual, a nossa verdadeira média populacional provavelmente se situa. Podemos também estimar a probabilidade de nossa média populacional situar-se efetivamente dentro daquele intervalo de valores médios. Esse é o conceito de intervalo de confiança, descrito anteriormente.

Quando não se conhece o desvio padrão da população, utiliza-se o desvio padrão da amostra (**s**) ficando a fórmula:

$$s_{\bar{x}} = \frac{s}{\sqrt{N}}$$

Assim, o erro padrão não é mais do que o desvio padrão da distribuição das médias das amostras de uma população.

- Exemplos

Volume plasma (L)	Desvio padrão $(x_i - \bar{x})$	(Desvio) ² $(x_i - \bar{x})^2$
2,75	-0,2525	0,0638
2,86	-0,1425	0,00203
3,37	0,3675	0,1351
2,76	-0,2425	0,0588
2,62	-0,3825	0,1463
3,49	0,4875	0,2377
3,05	0,0475	0,0023
3,12	0,1175	0,0138
MÉDIA	3,0025	SOMA = 0,6780

O desvio padrão encontrado foi de 0,3112 L.

- Importância das medidas de dispersão:

Ex.: Calcular a média, mediana, moda e desvio padrão das amostras A e B.

Parâmetros	A	B
Média	12,5	12,3
Erro Padrão	0,5627	1,6197
Mediana	12,5	12
Moda	14	12
Desvio Padrão	1,7795	5,1218
Variância	3,1667	26,2333

2 – Observar as notas de quatro competidores em uma prova de manobras radicais com skates.

Competidor	A:	7,0	–	5,0	–	3,0
Competidor	B:	5,0	–	4,0	–	6,0
Competidor	C:	4,0	–	4,0	–	7,0

Calcular a média, variância e desvio padrão de cada competidor:

Competidor	Média	Variância	Desvio padrão
A	5,0	4	2
B	5,0	1	1
C	5,0	3	1,732

Podemos observar que a média dos três competidores são cinco, impossibilitando a análise sobre a regularidade dos competidores. Portanto, partindo-se da idéia que há necessidade de adotar uma medida que apresente a variação dessas notas, sem no entanto comprometer a análise.

O Competidor B, possui uma melhor regularidade nas notas, por possuir menor variabilidade.

6. Distribuição de Probabilidades

6.1 Conceitos Básicos

- Variável Aleatória

A observação dos resultados de um experimento aleatório é sempre um processo de medição onde uma série de mecanismos probabilísticos influencia os valores que são registrados. Nesse contexto, uma variável aleatória é aquilo que se obtém quando se associa a cada ponto do espaço amostral um valor numérico pertencente ao conjunto dos Números Reais.

- Função de Distribuição

A Função de Distribuição de uma variável aleatória é a relação matemática que fornece, para cada valor da variável, o somatório das probabilidades associadas a esse ponto e daquelas associadas a todos os valores abaixo dele. Num gráfico cartesiano onde o eixo horizontal expressa os valores de uma variável aleatória e o eixo vertical o valor da sua função de distribuição, a curva resultante sempre parte de uma altura mínima de zero e cresce sem parar até atingir a altura máxima de um. O aspecto e equação matemática específica de tal curva num caso em particular é uma consequência direta dos mecanismos probabilísticos subjacentes ao fenômeno em questão.

- Densidade de Probabilidade

A taxa de variação da função de distribuição em cada ponto da variável aleatória corresponde à probabilidade associada a tal ponto isoladamente, sendo esta a definição da Densidade de Probabilidade. Em outras palavras, trata-se de um reflexo preciso do conjunto de probabilidades associadas a todos os pontos do espaço amostral.

- Distribuição de Probabilidade

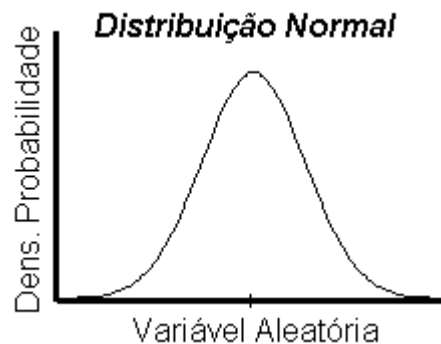
A densidade de probabilidade de uma variável é comumente chamada de sua Distribuição de Probabilidade. Como foi dito antes, trata-se de um modelo da

incerteza associada a um fenômeno ou a um conjunto de fenômenos, ou seja, de uma expressão do seu mecanismo probabilístico. A caracterização definitiva de uma distribuição de probabilidade é realizada a partir da equação ou conjunto de equações que a definem e que determinam a sua representação gráfica.

Dependendo da natureza do mecanismo probabilístico subjacente a um dado fenômeno, a sua correspondente distribuição de probabilidade (a densidade de probabilidade da variável aleatória concomitante) pode variar significativamente, sendo conhecidos e classificados diversos tipos, alguns dos quais são mostrados a seguir.

6.2 Distribuição Normal ou Gaussiana

É a mais importante e a mais comum das distribuições de probabilidade, estando associada a diversos tipos de fenômenos. Calcula-se por meio de uma função matemática baseada no número e (2.7182818284...). Apresenta o clássico formato de um "sino" simétrico, tal como se pode ver na representação ao lado.



Na distribuição gaussiana, os eventos mais extremos (mais perto dos níveis máximo ou mínimo) são menos prováveis do que os eventos intermediários (mais próximos ao ponto central da escala).

A vantagem da distribuição normal reside na facilidade de defini-la com apenas dois parâmetros, a média μ e o desvio padrão σ da distribuição.

Uma das características importantes é que a partir desses dois parâmetros será possível calcular, por exemplo, a percentagem de valores que deverão estar acima ou abaixo de um determinado valor da variável aleatória, ou entre esses dois valores definidos.

A variável aleatória X tem distribuição normal de probabilidade com parâmetros, média $-\infty < \mu < +\infty$ e variância $0 < \sigma^2 < +\infty$, se a função densidade $f(x)$ for:

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

Analisando a fórmula da função densidade $f(x)$, podemos verificar que para cada par de parâmetros μ e σ , há uma curva diferente de $f(x)$, ou que para qualquer outro par de parâmetros μ e σ , a curva $f(x)$ será diferente.

Portanto, não há apenas uma única distribuição normal, e sim uma família de distribuições normais representadas como $N(\mu, \sigma)$.

Propriedades da Distribuição Normal

A distribuição normal da variável aleatória X com média μ_x e variância σ_x^2 tem as seguintes propriedades:

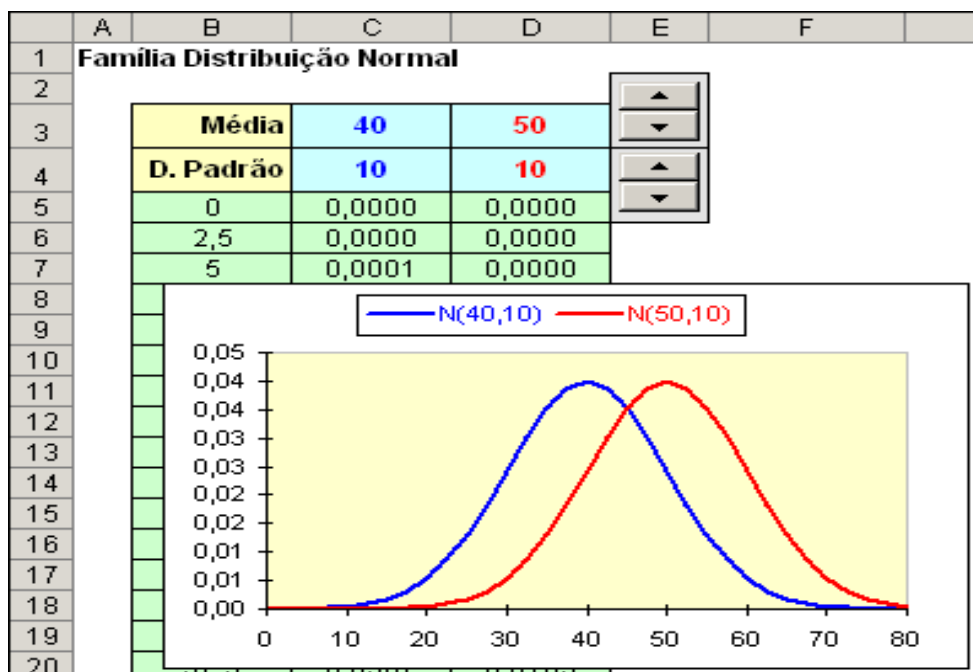
- $f(x)$ tende a zero, quando x tende a $\pm \infty$, ou $f(x) \rightarrow 0$ quando $x \rightarrow \pm \infty$,
- A área total sob a curva é 100 e cada metade da curva tem 50% da área total, pois a curva é simétrica ao redor da média μ_x ,

A probabilidade $P(a \leq X \leq b)$ é a área sob a curva no intervalo (a, b) .

Influência dos Parâmetros

O que ocorreria com a forma da distribuição normal $N(40, 10)$, se o valor da média for mudado de 40 para 50?

A forma da distribuição permaneceria a mesma, porém com a média deslocada de 40 para 50, sendo a nova distribuição normal definida como $N(50, 10)$.



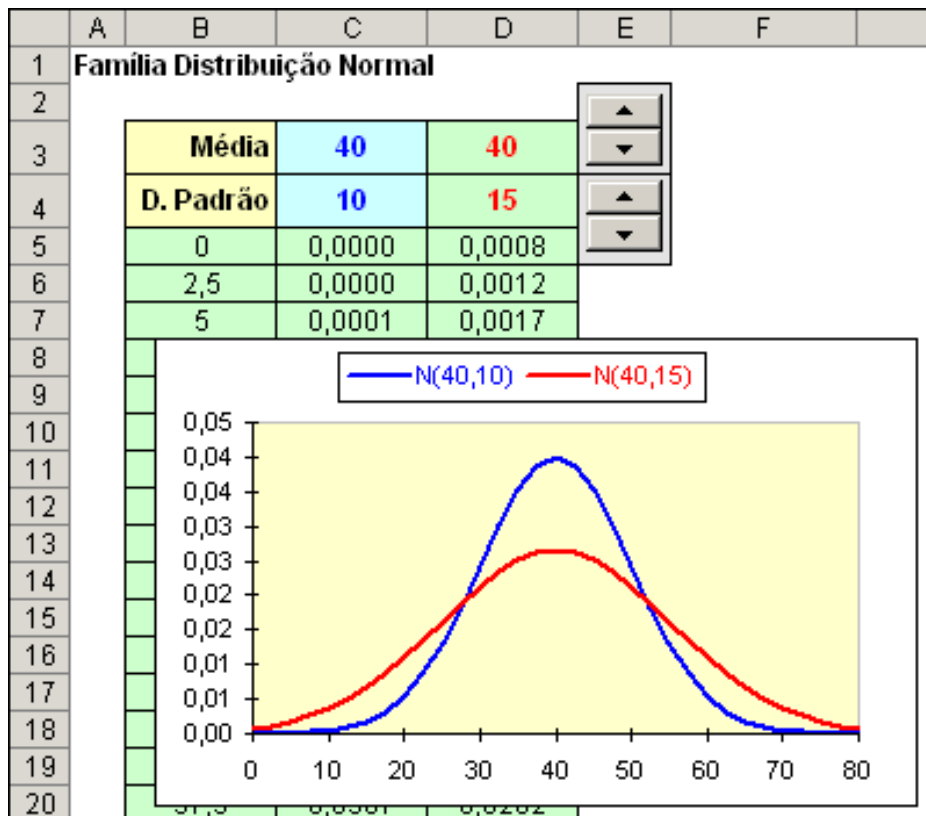
No intervalo C3:C4, é possível mudar os parâmetros da curva de cor azul registrando novos valores. Entretanto, os parâmetros da curva vermelha são mudados acionando

um dos dois controles giratórios por vez, células E3 e E4. Observe que ao variar o valor da média de 40 para 50, mantendo constante o desvio padrão $\sigma=10$, a distribuição normal mantém sua forma e se desloca para a direita. Da mesma maneira, para valores de média menores que 40, por exemplo $\mu=30$, a distribuição manterá sua forma e se deslocará para a esquerda de $\mu=40$.

Variando o Desvio Padrão

Para analisar o comportamento da forma da distribuição normal variando o desvio padrão, lembremos primeiro o significado do desvio padrão apresentado anteriormente. A regra prática estabelece que, em todas as distribuições, a porcentagem de valores contidos dentro de três desvios padrões ao redor da média será próxima de 100%. Como a distribuição normal é simétrica ao redor da média, seja qual for o valor do desvio padrão, praticamente 100% dos valores da variável aleatória estarão contidos dentro de três desvios padrões ao redor da média.

Se na distribuição normal $N(40, 10)$, o desvio padrão for mudado para 15, a forma da curva $N(40, 15)$ será mais aberta que a anterior $N(40, 10)$, diminuindo sua altura para manter a área de 100%. Da mesma maneira, pode-se verificar que para desvios padrão menores que 10, a distribuição diminui sua base e aumenta sua altura.



Embora, não haja apenas uma única distribuição normal e sim uma família de distribuições normais representadas com $N(\mu, \sigma)$, elas mantêm algumas propriedades em comum, como, por exemplo, a porcentagem de resultados ao redor da média.

Um desvio padrão ao redor da média

Qual a percentagem de valores dentro de um desvio padrão ao redor da média = 200 da variável aleatória X com distribuição normal $N(200, 10)$?

Nesse caso, um valor de X incluído no intervalo de um desvio padrão ao redor da média deve estar no intervalo (190, 210).

A probabilidade de um valor da variável aleatória X estar dentro de um desvio padrão ao redor da média é $P(190 \leq X \leq 210) = 0,6827$ ou 68,27%.

Esse resultado mostra que 68,27% dos valores da variável X com distribuição normal se distribuem no intervalo de um desvio padrão ao redor da média, ou seja, que a probabilidade de um resultado do experimento com distribuição $N(200, 10)$ pertencer ao intervalo (190, 210) é 68,27%.

Dois desvios padrão ao redor da média

Qual a percentagem de valores da variável X com distribuição normal $N(200, 10)$ localizados dentro de dois desvios padrão ao redor da média?

Um valor de X incluído no intervalo de dois desvios padrão ao redor da média deve estar no intervalo (180, 220).

A probabilidade de um valor da variável aleatória X estar dentro de dois desvios padrão ao redor da média é $P(180 \leq X \leq 220) = 0,9545$ ou 95,45%.

Esse resultado mostra que 95,45% dos valores da variável X com distribuição normal se distribuem no intervalo de dois desvios padrão ao redor da média, ou seja, que a probabilidade de um resultado do experimento com distribuição $N(200, 10)$ pertencer ao intervalo (180, 220) é 95,45%.

Três desvios padrão ao redor da média

Qual a percentagem de valores da variável X com distribuição normal $N(200, 10)$ localizados dentro de três desvios padrão ao redor da média?

Um valor de X incluído no intervalo de dois desvios padrão ao redor da média deve estar no intervalo (170, 230).

A probabilidade de um valor da variável aleatória X estar dentro de três desvios padrão ao redor da média é $P(170 \leq X \leq 230) = 0,9973$ ou 99,73%.

Esse resultado mostra que 99,73% dos valores da variável X com distribuição normal se distribuem no intervalo de três desvios padrão ao redor da média, ou seja, que a probabilidade de um resultado do experimento com distribuição $N(200, 10)$ pertencer ao intervalo (170, 230) é 99,73%.

Modelo da Distribuição Normal

O gráfico da distribuição normal $f(x)$ construído na planilha **Modelo Distribuição Normal** inclui a representação dos intervalos dos três tipos de desvios padrão.

Com os controles giratórios, são informados os valores da média na célula C3 e o desvio padrão na célula C4 e, ao mesmo tempo, a planilha recalcula a tabela do intervalo B6:C31 de acordo com o tipo de probabilidade escolhida na caixa de grupo:

- $f(x)$, fornecerá o valor da função $f(x)$ para o valor x ,
- *Probabilidade acumulada até x* , fornecerá a probabilidade acumulada de menos infinito até x .

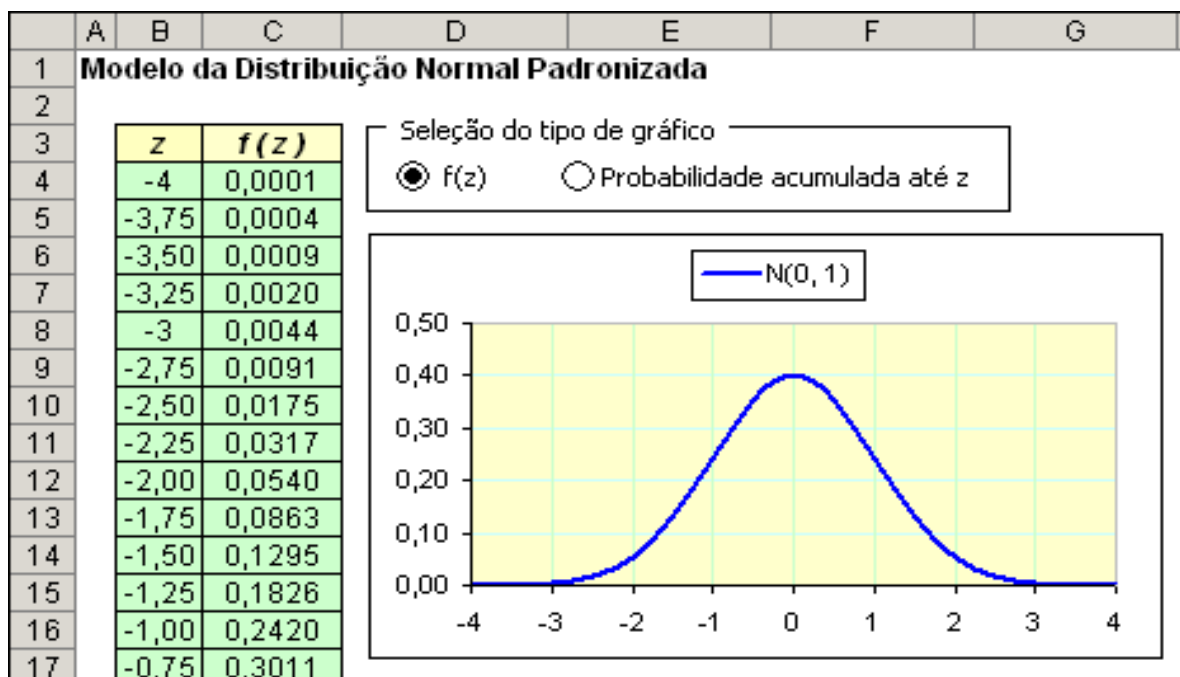
A variável aleatória, *desvio padrão normalizado Z* de uma *distribuição normal padronizada* é definida pela expressão:

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

A nova variável Z realiza cálculos de probabilidades com uma única curva de distribuição denominada distribuição normal padronizada $N(0, 1)$.

Depois da transformação da variável, a função densidade $f(x)$ passa a ser:

$$f(Z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$



Propriedades da Distribuição Normal Padronizada

A distribuição normal da variável aleatória Z com média μ_Z e variância σ_Z^2 têm as seguintes propriedades:

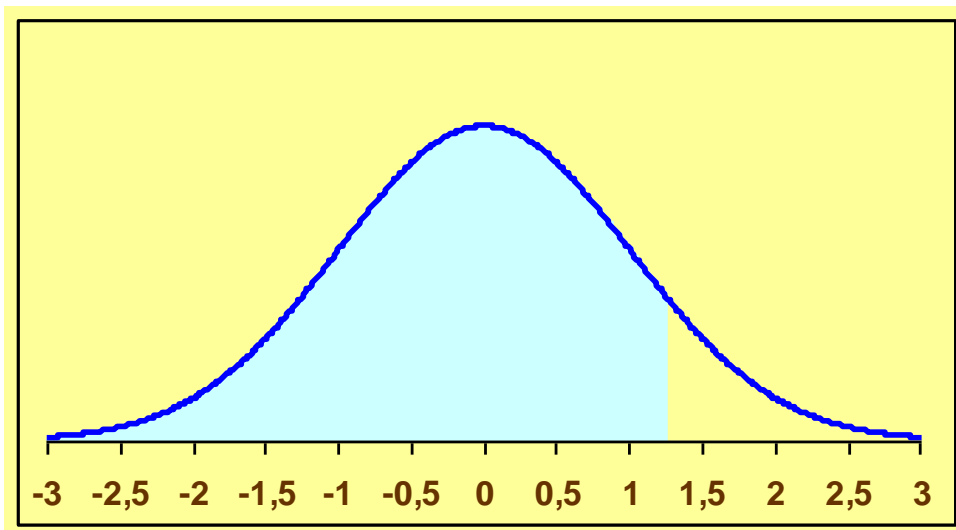
- $\mu_Z = 0$ e $\sigma_Z^2 = 1$
- $f(Z) \rightarrow 0$ quando $Z \rightarrow \pm\infty$.
- A curva é simétrica ao redor da média $\mu_Z = 0$. A área total sob a curva é 100%, e cada metade da curva tem 50% da área total.
- A probabilidade $P(Z_1 \geq Z \leq Z_2)$ é a área sob a curva no intervalo (Z_1, Z_2) .

Ex.

Qual a probabilidade de um valor de X ser menor ou igual a 52,4 sabendo que a variável tem distribuição normal (40,10)?

Solução: O objetivo é determinar a probabilidade $P(X \leq 52,4)$. Para utilizar a distribuição normal padronizada, $Z = 1,24$; resultado obtido através da seguinte fórmula:

$$Z = \frac{X - \mu}{\sigma} = \frac{52,4 - 40}{10} = +1,24$$



O valor da probabilidade $P(Z \leq 1,24)$ pode ser obtido com a função DIST.NORMP do Excel para a distribuição normal padronizada e também com a *Tabela Z* a seguir (0,8925 ou 89,25%).

Tabela Z

Z	0,00	0,01	0,02	0,03	0,04	0,05
0,00	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199
0,10	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596
0,20	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987
0,30	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368
0,40	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736
0,50	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088
0,60	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422
0,70	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734
0,80	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023
0,90	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289
1,00	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531
1,10	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749
1,20	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944
1,30	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115

A curva da distribuição Z tem a mesma forma que a curva da distribuição X, com a mudança dos valores do eixo X para o eixo Z.

Como há diversos cálculos de probabilidade, a tabela seguinte apresenta o resumo dos procedimentos de cálculo de probabilidade utilizando a tabela da distribuição normal padronizada Z.

Tipo de cálculo	Resultado da Tabela Z
$P(Z \leq a); a > 0$	Direto da tabela
$P(Z \geq a); a > 0$	$1 - P(Z \leq a)$
$P(Z \leq a); a < 0$	$1 - P(Z \leq a)$
$P(Z \geq a); a < 0$	$P(Z \leq a)$

Ex.:

Se a variável X tem distribuição normal com parâmetros $N(40,10)$. Qual a probabilidade $P(X \leq 50)$?

Solução: O valor equivalente de $X = 50$ na distribuição normal padronizada é $Z = +1$, resultado obtido com a seguinte fórmula:

$$Z = \frac{X - \mu}{\sigma} = \frac{50 - 40}{10} = +1$$

Da tabela Z, se obtém $P(X \leq 50) = P(Z \leq 1) = 0,8413$ ou 84,13%.

Qual a probabilidade de $P(X \leq 35)$, obtém-se $Z = -0,50$. Entretanto, não temos na tabela valores negativos, então, devemos calcular seu complemento, ou seja, $1 - P(Z \leq |a|) = 1 - P(Z \leq +0,50) = 1 - 0,6915 = 0,3085$ ou 30,85 %.

Com a variável X do exemplo anterior, qual a probabilidade $P(25 \leq X \leq 60)$?

Os valores de Z para $X = 25$ e $X = 60$ são $Z_1 = -1,50$ e $Z_2 = +2,0$. Como a probabilidade $P(-1,5 \leq Z \leq +2)$ não é obtida na tabela Z, ela será o resultado da diferença $P(Z \leq +2) - P(Z \leq -1,5)$, procedimento que deve ser realizado por partes:

→ $P(Z \leq +2) = 0,9772$, obtida diretamente da tabela de Z.

→ $P(-1,5 \leq Z \leq +2,0) = 1 - P(Z \leq +1,5) = 1 - 0,9332 = 0,0668$

Então: $P(-1,5 \leq Z \leq +2) = 0,9772 - 0,0668 = 0,9104$ ou 91,04%.

6.3 Distribuição Binomial

É uma distribuição de probabilidade para fenômenos discretos, sendo calculada a partir do binômio de Newton. Sua representação é a de uma série de barras separadas, uma para cada evento, havendo uma barra para cada resultado possível. O seu aspecto geral é o de uma série de colunas que, juntas, formam uma curva semelhante à da gaussiana.

6.4 Distribuição Uniforme

É uma distribuição onde todos os intervalos iguais da variável aleatória têm a mesma probabilidade, sendo a probabilidade de um intervalo diretamente proporcional ao seu comprimento. A sua aparência é a de um simples retângulo. A distribuição uniforme expressa um mecanismo probabilístico ao qual é atribuído um estado de incerteza plena.



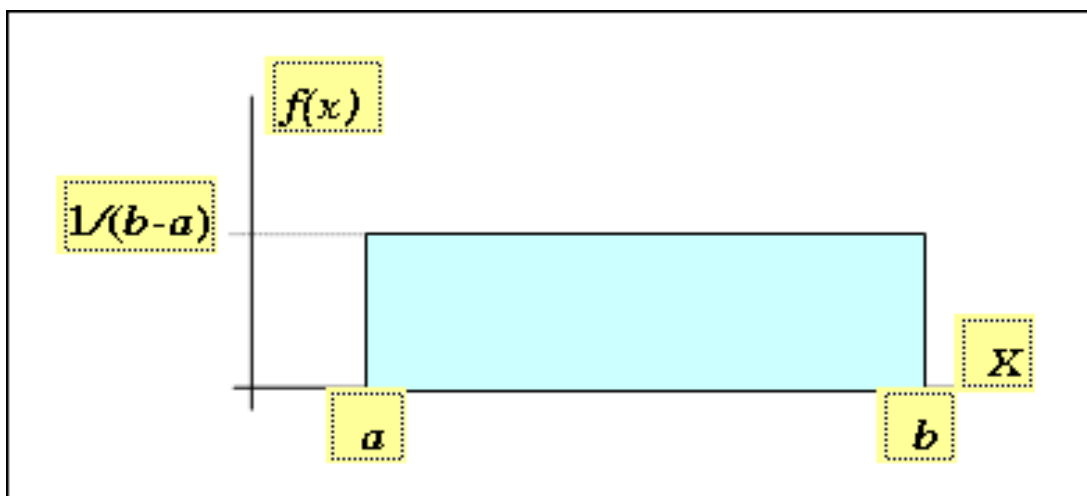
Uma distribuição de variável aleatória contínua é a *distribuição uniforme* cuja função densidade de probabilidade é constante dentro de um intervalo de valores da variável aleatória X.

Cada um dos possíveis valores que X com distribuição uniforme podem assumir tem a mesma probabilidade de ocorrer.

A variável aleatória X tem distribuição uniforme de probabilidades no intervalo (a,b), se a função densidade f(x) for:

$$f(x) = \frac{1}{b-a}$$

Com as seguintes condições $b \geq a$ e $a \leq x \leq b$.



A média μ_x e a variância σ_x^2 da variável aleatória X com distribuição uniforme de probabilidades (a,b) são:

$$\text{Média} = \frac{a+b}{2}$$

$$\text{Variância} = \frac{(b-a)^2}{12}$$

Ex.: Calcular a média e o desvio padrão da variável aleatória X com distribuição uniforme no intervalo (100,200).

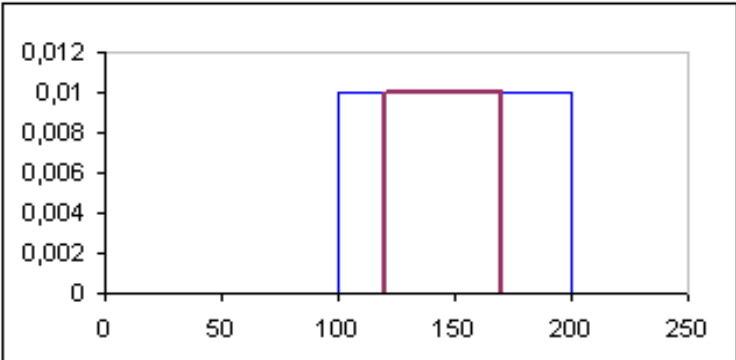
Solução:

A média obtida pela fórmula $\text{Média} = \frac{a+b}{2} = \frac{100+200}{2} = 150$.

A variância obtida pela fórmula $\text{Variância} = \frac{(b-a)^2}{12} = \frac{(200-100)^2}{12} = 833,33$.

O desvio padrão obtido pela fórmula $Desviopadrão = \sqrt{833,33} = 28,87$

A figura abaixo resume os resultados obtidos.

	A	B	C	D	E	F
1	Modelo da Distribuição Uniforme					
2						
3	Dados			Resultados		
4		Mínimo	100		Média	150,00
5		Máximo	200		Variância	833,33
6					Desvio Padrão	28,87
7	Cálculo de Probabilidades					
8		c	120			
9		d	170			
10				P(120 < X < 170)		50,00%
11						
12						
13						
14						
15						
16						
17						
18						
19						
20						
21						

6.5 Distribuição Exponencial

É uma distribuição onde os valores mais próximos de um dos extremos da variável aleatória apresentam probabilidade máxima e os mais próximos do outro extremo apresentam probabilidade que tende a zero. Está associada a um mecanismo probabilístico que favorece os eventos de intensidade mais alta ou mais baixa.



Na distribuição exponencial, a variável aleatória é definida como o tempo entre duas ocorrências, sendo a média de tempo entre ocorrências de $1/\lambda$.

Por exemplo, se a média de atendimentos no caixa bancário é de $l = 6/\text{min}$, então o tempo médio entre atendimentos é $1/l = 1/6$ de minuto, ou seja, 10 segundos.

Seja X é uma variável aleatória continua no intervalo $0 \leq x < \infty$, onde a sua função densidade de probabilidade é dada por:

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

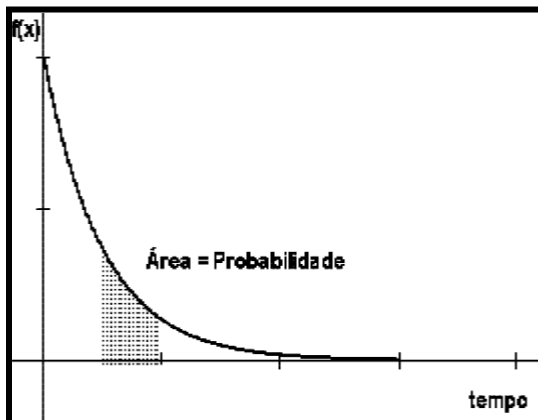
Ou seja,

$$f(x) = \begin{cases} 0 & x < 0 \\ \lambda e^{-\lambda x} & 0 \leq x < \infty \end{cases}$$

Onde λ é alguma constante positiva.

Notação: $X \sim \text{Exp}(\lambda)$

Graficamente, temos:



Aplicação

A densidade de distribuição exponencial tem importante aplicação em tempos de vida de determinados tipos de equipamento, ou como intervalo de tempo entre eventos que ocorrem em seqüência aleatória, como por exemplo, chamadas telefônicas em um centro telefônico, demandas de serviço de computador, chegadas sucessivas de raios cósmicos, teoria das filas e linha de espera para medir o tempo de ocorrido entre chegadas de clientes em locais de prestação de serviços, como caixas eletrônicos, chegadas de pacientes em um pronto-socorro, buscas em um site de internet.

6.6 Distribuição Retangular

Esta distribuição é mais conservadora, na qual a probabilidade de que o valor da componente esteja dentro de um intervalo de a_- até a_+ , para todos os fins práticos, é igual a 1, e 0, de que esteja fora deste intervalo. Neste caso, o componente-padrão associada será:

$$u = \frac{a_+ - a_-}{\sqrt{12}}$$

Se os parâmetros a_+ e a_- forem simétricos, sua diferença (amplitude do intervalo) pode ser designada por $2a$, implicando em:

$$u = \frac{a}{\sqrt{3}}$$

Essa distribuição deve ser utilizada sempre que inexistir qualquer informação de comportamento da componente de incerteza, mas apenas sua faixa de variação.

6.7 Distribuição Triangular

Normalmente, associada a distribuições de incerteza que possuem maior probabilidade de ocorrência nas imediações do valor central. Por exemplo, a incerteza devida à resolução de instrumentos analógicos segue esta distribuição, uma vez que a sua probabilidade de ocorrência é a de maior amplitude nas imediações da região central entre duas marcas consecutivas da escala de medição (para escalas lineares). Se $2a$ corresponde à faixa de valores atribuíveis à distribuição triangular de incerteza ($a_+ - a_-$, quando a_+ e a_- são simétricos), a incerteza padrão é:

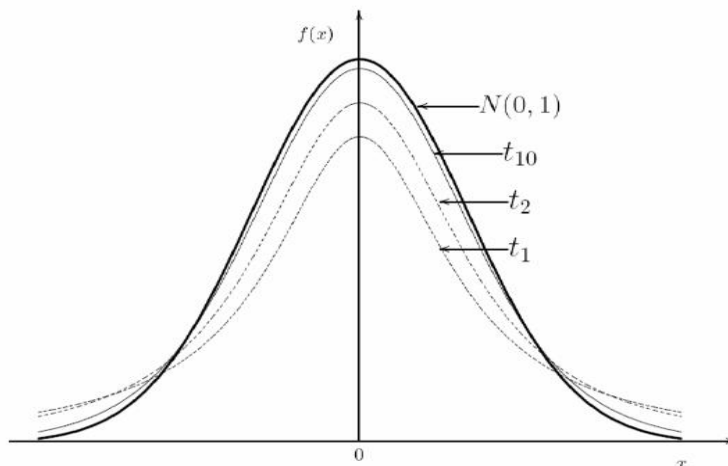
$$u = \frac{a}{\sqrt{6}}$$

6.8 Distribuição t de Student

Graus de liberdade

É um conceito ligado ao número de dados disponíveis (livres) para o cálculo da estatística. Por exemplo, ao estimarmos a média da populacional, com a média amostral, perdemos um grau de liberdade, assim a estatística t de student terá $n - 1$ graus de liberdade. No caso da tabela de Análise de Variância (ANOVA), os graus de liberdade do grupo serão iguais ao número de grupos menos 1, o grau de liberdade total será $a - 1$ e os graus de liberdade do resíduo, a diferença entre esses dois.

Esta distribuição é uma família de distribuição indexada por um parâmetro, que representa o número de graus de liberdade (g.l.). Quando X tem uma distribuição t com n graus de liberdade, temos que $X \sim t_n$.



A distribuição t de Student é utilizada quando se deseja determinar a média da população (que segue a distribuição normal) a partir de uma amostra. Neste problema, não se sabe qual é a média ou o desvio padrão da população, mas ela deve seguir a uma distribuição normal.

Quanto menor o número de gl, mais aplainada (platocurtica) é a forma da distribuição, resultando em maior área nas caudas da distribuição. Pode-se observar ainda que a distribuição t é muito semelhante à curva normal. À medida que aumentam os gl, a distribuição t aproxima-se da distribuição normal padronizada (média = 0, desvio-padrão = 1). A curva normal padronizada é um caso particular da distribuição t quando gl tende ao infinito. Para os propósitos práticos, os valores da distribuição t aproximam-se dos valores da distribuição normal padronizada relativamente depressa, tal que quando $gl = 30$ esses valores são quase idênticos.

Propriedades da distribuição t de student

Essa distribuição varia de acordo com a dimensão da amostra que vai determinar o número de graus de liberdade.

- A curva da distribuição t de Student tem a mesma forma em sino da distribuição normal, mas reflete a maior variabilidade (com curvas mais alargadas) que é de esperar em amostras pequenas.
- A distribuição t de Student tem valor médio zero (tal como a distribuição normal padronizada).

- O desvio padrão da distribuição de t de Student varia de acordo com o tamanho da amostra e é maior do que 1 (o que não acontece com a distribuição normal padrão, onde $\sigma = 1$).
- Quanto maior a dimensão da amostra, mais essa distribuição se aproxima da distribuição normal.

6.9 Distribuição F de Snedecor ou Fisher

Essa distribuição é utilizada em um dos testes mais importantes da estatística, que é a análise de variância. Suponhamos duas amostras independentes retiradas de populações normais forneçam variâncias amostrais e que desejamos conhecer a distribuição amostral do quociente s_1^2/s_2^2 . Isso será possível através da distribuição F, conhecendo-se os graus de liberdade, δ_1 e δ_2 , cuja distribuição é:

$$F = \frac{s_1^2}{s_2^2}$$

Onde $s_1 > s_2$

O valor obtido para F, chamado de F calculado, é comparado com o F, chamado F crítico, que se obtém em valores tabelados.

A consulta a essas tabelas é simples, através de graus de liberdade das amostras correspondentes ao numerador (ν_1 na horizontal) e ao denominador (ν_2 na vertical), cujos índices são previamente estabelecidos.

Quando $F_{\text{calculado}} < F_{\text{crítico}}$ é sempre calculado é sempre maior que a unidade, aceita-se a igualdade da variâncias; caso o contrário não se aceita a igualdade entre elas.

O emprego da distribuição F é amplo. É bastante empregado na comparação de médias e outras situações comparativas.

Os valores de estatística F apresentam-se tabelados, onde a primeira linha mostra os graus de liberdade do numerador e a segunda coluna os graus de liberdade do denominador. No cruzamento de uma linha com uma coluna aparecem os valores tabelados.

Ex. Propõem-se um novo método par a determinação de nitrato em águas pluviais, supondo-se que seja melhor que o atualmente empregado. Verificar se a suposição tem fundamento, ao nível de 95% de confiança.

Método em uso: $s = 25$ mg/kg, para $n = 8$ (considerar como índice 1)

Método proposto: $s = 13$ mg/kg, para $n = 10$ (considerar como índice 2)

$$F_{calc} = \frac{25^2}{13^2} = \frac{625}{169} = 3,70$$

F crítico = 3,29 ($\nu_1 = 7$ e $\nu_2 = 9$)

Comparando: $F_{calc} > F_{crítico}$, há evidência de supremacia do método anterior, pois as variâncias são consideradas diferentes.

7. Testes de Normalidade

Significância

Quando avaliamos os resultados de uma experiência, os descrevemos em função da sua significância. A significância refere o quão seguro podemos estar de que os nossos resultados são realmente diferentes. O nível de significância é representado pela letra α . Os níveis de significância adotados são: 0,05, 0,01 e 0,001 (5%, 1% e 0,1%, respectivamente);

São utilizados três testes estatísticos para analisar se o conjunto de observações segue ou não uma distribuição normal:

- Teste de Kolmogorov Smirnov, quando o conjunto de observações é grande ($N > 50$).
- Teste de Shapiro-Wilk, quando o conjunto de observações é pequeno ($N < 50$);
- Anderson Darling.

7.1 Teste de Kolmogorov - Smirnov

O teste de Kolmogorov-Smirnov baseia-se na máxima diferença entre a distribuição acumulada da amostra e distribuição acumulada esperada. Se o valor calculado de D é estatisticamente significativo (para $p = 0,05$), rejeita-se a hipótese que a distribuição estudada é normal, ou seja, para a distribuição ser considerada **normal**, o valor de p **deve ser maior que 0,05**. Quando a média e o desvio padrão da distribuição esperada (hipotética) não é conhecido, os valores de probabilidades de Lilliefors devem ser utilizados.

Este teste baseia-se no cálculo da razão entre a diferença máxima absoluta entre as frequências observadas e esperadas e o número de observações. Esta razão (D) é comparada com o valor da tabela de Kolmogorov – Smirnov. Assim:

$$D = \frac{d_{MAX}}{N}$$

Onde:

d_{MAX} = máxima diferença absoluta entre as frequências observadas e esperadas,

N = número de observações

Segundo Sachs, quando o número de observações for maior que 35, o valor de D tabelado pode ser calculado através das relações apresentadas no quadro 1.

Quadro 1 – Tabela de D

VALOR DE D	NÍVEL DE SIGNIFICÂNCIA
$1,07 \sqrt{n}$	0,20
$1,14 \sqrt{n}$	0,15
$1,22 \sqrt{n}$	0,10
$1,36 \sqrt{n}$	0,05
$1,63 \sqrt{n}$	0,01

Exemplo:

Realizaram-se oito titulações, tendo obtido os resultados:

25.13, 25.02, 25.11, 25.07, 25.03, 24.97, 25.14 e 25.09

Estes resultados provêm de uma população com distribuição normal de parâmetros valor médio =25 e desvio padrão=0.05?

Solução:

A standarização dos dados $x-25/0.05$ conduz aos valores:

2.6, 0.4, 2.2, 1.4, 0.6, -0.6, 2.8 e 1.8

Estes valores depois de ordenados, calculada a frequência cumulativa e a função de distribuição para uma normal reduzida, correspondem à seguinte tabela:

$(x - 25)/0,05$	Frequencia cumulativa	Função de distribuição	Diferença Valor Absoluto
-0,6	$1/8 = 0,125$	0,2743	0,1493
0,4	$2/8 = 0,250$	0,6554	0,4054
0,6	$3/8 = 0,375$	0,7257	0,3507
1,4	$4/8 = 0,500$	0,9192	0,4192 MAX
1,8	$5/8 = 0,625$	0,9641	0,3391
2,2	$6/8 = 0,750$	0,8961	0,2361
2,6	$7/8 = 0,875$	0,9953	0,1200
2,8	1	0,9974	0,0026

Comparando o valor observado = 0,419, com o valor crítico tabelado para 8 observações = 0,288 → rejeitamos a hipótese nula.

7.2 Teste de Shapiro Wilk

O teste de *Shapiro-Wilk* tem sido o mais utilizado para testar a normalidade. Se o valor calculado de *W* é estatisticamente significativo (para $p = 0,05$), rejeita-se a hipótese que a distribuição estudada é normal, ou seja, para a distribuição ser considerada normal, o valor de p deve ser maior que 0,05.

Passos para o teste de *Shapiro-Wilk*:

1º - Formulação da Hipótese

H_0 : A amostra provém de uma população normal

H_1 : A amostra não provém de uma população normal

2º - Estabelecer a Estatística

Teste de Shapiro-Wilk

3º - Nível de Significância do teste (α)

4º - Calcular a estatística de teste

→ Ordenar as N observações da amostra:

$$x_1 \leq x_2 \leq x_3 \leq \dots \leq x_N$$

→ Calcular:

$$\sum_{i=1}^N (x_i - \bar{X})^2 = \sum_{i=1}^N x_i^2 - \frac{\left(\sum_{i=1}^N x_i\right)^2}{N} = \sum_{i=1}^n x_i^2 - N * \bar{x}^{-2}$$

→ Calcular b:

$$b = \sum_{i=1}^{N/2} a_{N-i+1} * (x_{N-i+1} - x_i)$$

Se N é ímpar, desprezar a observação mediana.

Coeficientes do Teste de Shapiro Wilk

*Anexo 3: Coeficientes α_{N-i+1} para o teste de normalidade W de SHAPIRO – WILK
(Para N = 2(1)50)*

$i \setminus N$	2	3	4	5	6	7	8	9	10
1	0.7071	0.7071	0.6872	0.6646	0.6431	0.6233	0.6052	0.5888	0.5739
2		0.0000	0.1677	0.2413	0.2806	0.3031	0.3164	0.3244	0.3291
3				0.0000	0.0875	0.1401	0.1743	0.1976	0.2141
4						0.0000	0.0561	0.0947	0.1224
5								0.0000	0.0399

$i \setminus N$	11	12	13	14	15	16	17	18	19	20
1	0.5601	0.5475	0.5359	0.5251	0.5150	0.5056	0.4968	0.4886	0.4808	0.4734
2	0.3315	0.3325	0.3325	0.3318	0.3306	0.3290	0.3273	0.3253	0.3232	0.3211
3	0.2260	0.2347	0.2412	0.2460	0.2495	0.2521	0.2540	0.2553	0.2561	0.2565
4	0.1429	0.1586	0.1707	0.1802	0.1878	0.1939	0.1988	0.2027	0.2059	0.2085
5	0.0695	0.0922	0.1099	0.1240	0.1353	0.1447	0.1524	0.1587	0.1641	0.1686
6	0.0000	0.0303	0.0539	0.0727	0.0880	0.1005	0.1109	0.1197	0.1271	0.1334
7			0.0000	0.0240	0.0433	0.0593	0.0725	0.0837	0.0932	0.1013
8					0.0000	0.0196	0.0359	0.0496	0.0612	0.0711
9							0.0000	0.0163	0.0303	0.0422
10									0.0000	0.0140

→ Calcular a estatística de teste W:

$$W = \frac{b^2}{\sum (x_i - \bar{x})^2}$$

5º - Tomar a Decisão: Rejeitar H_0 ao nível de significância α se:

$$W_{calc} < W_{\alpha}$$

Teste de Shapiro-Wilk – Valores Críticos

Anexo 4: Valores críticos da estatística W de SHAPIRO-WILK

		Nível de significância α								
		0.01	0.02	0.05	0.10	0.50	0.90	0.95	0.98	0.99
Tamanho da Amostra, N	3	0.753	0.756	0.767	0.789	0.959	0.998	0.999	1.000	1.000
	4	0.687	0.707	0.748	0.792	0.935	0.987	0.992	0.996	0.997
	5	0.686	0.715	0.762	0.806	0.927	0.979	0.986	0.991	0.993
	6	0.713	0.743	0.788	0.826	0.927	0.974	0.981	0.986	0.989
	7	0.730	0.760	0.803	0.838	0.928	0.972	0.979	0.985	0.988
	8	0.749	0.778	0.818	0.851	0.932	0.972	0.978	0.984	0.987
	9	0.764	0.791	0.829	0.859	0.935	0.972	0.978	0.984	0.986
	10	0.781	0.806	0.842	0.869	0.938	0.972	0.978	0.983	0.986
	11	0.792	0.817	0.850	0.876	0.940	0.973	0.979	0.984	0.986
	12	0.805	0.828	0.859	0.883	0.943	0.973	0.979	0.984	0.986
	13	0.814	0.837	0.866	0.889	0.945	0.974	0.979	0.984	0.986
	14	0.825	0.846	0.874	0.895	0.947	0.975	0.980	0.984	0.986
	15	0.835	0.855	0.881	0.901	0.950	0.975	0.980	0.984	0.987
	16	0.844	0.863	0.887	0.906	0.952	0.976	0.981	0.985	0.987
	17	0.851	0.869	0.892	0.910	0.954	0.977	0.981	0.985	0.987
	18	0.858	0.874	0.897	0.914	0.956	0.978	0.982	0.986	0.988
	19	0.863	0.879	0.901	0.917	0.957	0.978	0.982	0.986	0.988
	20	0.868	0.884	0.905	0.920	0.959	0.979	0.983	0.986	0.988
	21	0.873	0.888	0.908	0.923	0.960	0.980	0.983	0.987	0.989
	22	0.878	0.892	0.911	0.926	0.961	0.980	0.984	0.987	0.989
	23	0.881	0.895	0.914	0.928	0.962	0.981	0.984	0.987	0.989
	24	0.884	0.898	0.916	0.930	0.963	0.981	0.984	0.987	0.989
	25	0.888	0.901	0.918	0.931	0.964	0.981	0.985	0.988	0.989
	26	0.891	0.904	0.920	0.933	0.965	0.982	0.985	0.988	0.989
	27	0.894	0.906	0.923	0.935	0.965	0.982	0.985	0.988	0.990
	28	0.896	0.908	0.924	0.936	0.966	0.982	0.985	0.988	0.990
	29	0.898	0.910	0.926	0.937	0.966	0.982	0.985	0.988	0.990
	30	0.900	0.912	0.927	0.939	0.967	0.983	0.985	0.988	0.990

Exemplo:

Considere a seguinte amostra,

8 12 10 24 12 10 16 19 9 10. Teste se a amostra provém de uma população normal.

1º - Formulação da Hipótese

H_0 : A amostra provém de uma população Normal

H_1 : A amostra não provém de uma população Normal

2º - Estabelecer a Estatística

Teste de Shapiro-Wilk

3º - Nível de Significância do teste $\alpha = 0,05$

4º - Calcular a estatística de teste

1- Ordenar em ordem crescente, as N observações da amostra: 8 9 10 10 10 12 12 16 19 24

2 – Calcular:

$$\sum_{i=1}^N (x_i - \bar{X})^2 = 236$$

3 – Calcular b:

$$b = \sum_{i=1}^{N/2} a_{N-i+1} * (x_{N-i+1} - x_i)$$



i	N-i+1	a_{N-i+1}	x_{N-i+1}	x_i	$a_{N-i+1}(x_{N-i+1}-x_i)$
1	10	0.5739	24	8	9.1824
2	9	0.3291	19	9	3.2910
3	8	0.2141	16	10	1.2846
4	7	0.1224	12	10	0.2448
5	6	0.0399	12	10	0.0798
					b=14.0826

Anexo 3: Coeficientes α_{N-i+1} para o teste de normalidade W de SHAPIRO – WILK
 (Para $N = 2(1)50$)

$i \setminus N$	2	3	4	5	6	7	8	9	10
1	0.7071	0.7071	0.6872	0.6646	0.6431	0.6233	0.6052	0.5888	0.5739
2		0.0000	0.1677	0.2413	0.2806	0.3031	0.3164	0.3244	0.3291
3				0.0000	0.0875	0.1401	0.1743	0.1976	0.2141
4						0.0000	0.0561	0.0947	0.1224
5								0.0000	0.0399

$i \setminus N$	11	12	13	14	15	16	17	18	19	20
1	0.5601	0.5475	0.5359	0.5251	0.5150	0.5056	0.4968	0.4886	0.4808	0.4734
2	0.3315	0.3325	0.3325	0.3318	0.3306	0.3290	0.3273	0.3253	0.3232	0.3211
3	0.2260	0.2347	0.2412	0.2460	0.2495	0.2521	0.2540	0.2553	0.2561	0.2565
4	0.1429	0.1586	0.1707	0.1802	0.1878	0.1939	0.1988	0.2027	0.2059	0.2085
5	0.0695	0.0922	0.1099	0.1240	0.1353	0.1447	0.1524	0.1587	0.1641	0.1686
6	0.0000	0.0303	0.0539	0.0727	0.0880	0.1005	0.1109	0.1197	0.1271	0.1334
7			0.0000	0.0240	0.0433	0.0593	0.0725	0.0837	0.0932	0.1013
8					0.0000	0.0196	0.0359	0.0496	0.0612	0.0711
9							0.0000	0.0163	0.0303	0.0422
10									0.0000	0.0140

4 – Calcular W :

$$W = \frac{b^2}{\sum (x_i - \bar{x})^2} = \frac{(14,0826)^2}{236} = 0,840$$

5º - Tomar a Decisão

$$W_{calc} = 0,840 < W_{(0,05;10)} = 0,842$$

Deve rejeitar a hipótese de normalidade da amostra.

$$p\text{-value} = \Pr(W > W_{calc}) = 0,0443 < \alpha = 0,05$$

Rejeitar H_0

6º - Conclusão

Pode-se afirmar com nível de significância de 5% que a amostra não provém de uma população normal ($p\text{-value} = 0,0443$).

ANEXO 4: Valores críticos da estatística W de SHAPIRO-WILK

		Nível de significância α								
		0.01	0.02	0.05	0.10	0.50	0.90	0.95	0.98	0.99
Tamanho da Amostra, N	3	0.753	0.756	0.767	0.789	0.959	0.998	0.999	1.000	1.000
	4	0.687	0.707	0.748	0.792	0.935	0.987	0.992	0.996	0.997
	5	0.686	0.715	0.762	0.806	0.927	0.979	0.986	0.991	0.993
	6	0.713	0.743	0.788	0.826	0.927	0.974	0.981	0.986	0.989
	7	0.730	0.760	0.803	0.838	0.928	0.972	0.979	0.985	0.988
	8	0.749	0.778	0.818	0.851	0.932	0.972	0.978	0.984	0.987
	9	0.764	0.791	0.829	0.859	0.935	0.972	0.978	0.984	0.986
	10	0.781	0.806	0.842	0.869	0.938	0.972	0.978	0.983	0.986
	11	0.792	0.817	0.850	0.876	0.940	0.973	0.978	0.984	0.986
	12	0.805	0.828	0.859	0.883	0.943	0.973	0.979	0.984	0.986
	13	0.814	0.837	0.866	0.889	0.945	0.974	0.979	0.984	0.986
	14	0.825	0.846	0.874	0.895	0.947	0.975	0.980	0.984	0.986
	15	0.835	0.855	0.881	0.901	0.950	0.975	0.980	0.984	0.987
	16	0.844	0.863	0.887	0.906	0.952	0.976	0.981	0.985	0.987
	17	0.851	0.869	0.892	0.910	0.954	0.977	0.981	0.985	0.987
	18	0.858	0.874	0.897	0.914	0.956	0.978	0.982	0.986	0.988
	19	0.863	0.879	0.901	0.917	0.957	0.978	0.982	0.986	0.988
	20	0.868	0.884	0.905	0.920	0.959	0.979	0.983	0.986	0.988
	21	0.873	0.888	0.908	0.923	0.960	0.980	0.983	0.987	0.989
	22	0.878	0.892	0.911	0.926	0.961	0.980	0.984	0.987	0.989
	23	0.881	0.895	0.914	0.928	0.962	0.981	0.984	0.987	0.989
	24	0.884	0.898	0.916	0.930	0.963	0.981	0.984	0.987	0.989
	25	0.888	0.901	0.918	0.931	0.964	0.981	0.985	0.988	0.989
	26	0.891	0.904	0.920	0.933	0.965	0.982	0.985	0.988	0.989
	27	0.894	0.906	0.923	0.935	0.965	0.982	0.985	0.988	0.990
	28	0.896	0.908	0.924	0.936	0.966	0.982	0.985	0.988	0.990
	29	0.898	0.910	0.926	0.937	0.966	0.982	0.985	0.988	0.990
	30	0.900	0.912	0.927	0.938	0.967	0.983	0.985	0.988	0.990

7.3 Teste de Anderson – Darling

O teste de *Anderson-Darling* é um dos testes estatísticos mais poderosos para detectar se há normalidade dos dados de uma amostra. É uma generalização do teste de Kolmogorov- Smirnov (KS), porém dando mais peso às caudas. Ele é também uma alternativa ao teste Qui-Quadrado. Pode ser utilizado com tamanhos de amostra pequenos ($n < 25$). Com amostras muito grandes, a suposição de normalidade poderá ser rejeitada apenas com pequenas imperfeições, mas dados industriais com tamanho de amostra de no mínimo 200 geralmente passam no teste de Anderson-Darling.

O teste avalia, em geral, se a amostra vem de uma distribuição qualquer especificada. A fórmula para o teste estatístico A para avaliar se os dados $\{Y_1 < \dots < Y_N\}$ (onde os dados devem ser ordenados) vem de uma distribuição com função distribuição acumulada F que é dada por:

$$A^2 = -N - S$$

O teste estatístico pode então ser comparado contra os valores críticos da distribuição teórica (dependendo de qual F é usada) para determinar o valor-p. O teste de *Anderson-Darling* para normalidade é um teste de distância ou de distribuição empírica

da função. É baseado no conceito de que quando uma distribuição de hipótese subjacente é dada, os dados podem ser transformados em uma distribuição uniforme. Os dados da amostra transformada podem então ter sua uniformidade testada com um teste de distância.

Em relação ao poder, foi verificado que A^2 é um dos melhores testes estatísticos de distribuições empíricas para detectar a normalidade dos dados. A única estatística que se aproxima é a estatística W^2 de teste de Cramér von-Mises.

O teste de *Anderson-Darling*, em caso de teste para a distribuição Normal da variável X segue seis passos:

- 1) Os dados X_i , para $i = 1, \dots, n$, que irão ser testados, devem ser ordenados do menor ao maior.
- 2) A média e o desvio padrão são calculados da amostra de X .
- 3) Os valores X_i são padronizados da seguinte forma:

$$Y_i = \frac{(X_i - \bar{X})}{S}$$

- 4) Com uma função distribuição acumulada de uma normal padronizada Φ , A^2

$$A^2 = -n - \frac{1}{n} \sum_{i=1}^n (2i-1) (\ln \Phi(Y_i) + \ln(1 - \Phi(Y_{n+1-i})))$$

- 5) A^{*2} , um ajustamento de aproximação para o tamanho de amostra, é calculado utilizando-se

$$A^{*2} = A^2 \left(1 + \frac{0,75}{n} + \frac{2,25}{n^2} \right)$$

- 6) Se A^{*2} for maior que 0,752, a hipótese de não normalidade é rejeitada com um nível de significância (α) de 5%.

Algumas observações, entretanto, devem ser levadas em consideração:

- 1) Se $s = 0$ ou se alguma $P_i = (0 \text{ ou } 1)$, então A^2 não poderá ser calculado e portanto é considerado indefinido.

2) O procedimento descrito acima serve apenas para o teste da distribuição normal. Qualquer outra distribuição teórica poderá também ser assumida utilizando-se a sua função distribuição acumulada. Cada distribuição teórica tem seus próprios valores críticos.

3) A hipótese nula é de que os dados amostrais vêm de uma população com distribuição especificada que se deseja testar (no exemplo acima, $N(0, 1)$).

Para melhor ilustração do uso do teste de Anderson-Darling, vamos ver o seguinte exemplo:

Exemplo: um pesquisador estava interessado em saber se o índice de flúor no dentes de crianças com até três anos de idade em Porto Alegre segue uma distribuição normal. Para isso, obteve a seguinte amostra de 12 crianças em diferentes clínicas dentárias:

13,4; 15,01; 12,48; 13,53; 14,21; 12,93; 8,21; 12,43; 20,81; 2,23; 4,43; 5,28.

A amostra apresenta uma média 11,24 e um desvio padrão de 5,26. Os dados foram ordenados e padronizados para, por fim, o $\Phi(Y)$ ser calculado, como mostra a tabela abaixo:

x	i	Y	$\Phi(Y)$
2,23	1	-1,71	0,04
4,34	2	-1,31	0,09
5,28	3	-1,13	0,13
8,21	4	-0,58	0,28
12,43	5	0,23	0,59
12,48	6	0,24	0,59
12,93	7	0,32	0,63
13,40	8	0,41	0,64
13,53	9	0,44	0,67
14,21	10	0,55	0,71
15,01	11	0,72	0,76
20,81	12	1,82	0,97

Com estas informações, chegou-se em $A^2 = 0,5409$ e por fim em $A^{*2} = 0,5831$, ambos calculados com as fórmulas vistas anteriormente.

Como $0,5831 < 0,752$, a hipótese de normalidade não pode ser rejeitada com nível de significância de 5% e, portanto, não há indícios de que índice de flúor no dentes de crianças com até três anos de idade em Porto Alegre não siga uma distribuição Normal.

O valor de p (p-value) do Teste de *Anderson-Darling* (AD) é $> 0,05$, portanto, não rejeita-se a hipótese H_0 , ou seja, os erros têm uma distribuição normal.

8. Testes de Homogeneidade de variâncias

Homogeneidade das variâncias.

O bom desempenho dos testes paramétricos exige que as variâncias nele envolvidas sejam homogêneas. Isso não implica, porém, que elas devam ser idênticas, porque nada é exatamente igual em estatística, havendo sempre uma faixa de tolerância em torno de qualquer suposta igualdade. O que os testes exigem é que elas não sejam discrepantes a ponto de ultrapassarem determinados limites de tolerância. Para entender as razões dessa exigência, basta imaginar o que aconteceria se alguém tentasse comparar a variação do crescimento de melancias com a variação do crescimento de jabuticabas. As variâncias nesse caso seriam heterogêneas, ou seja, tão diferentes que tornariam impossível qualquer comparação direta.

Todavia, mesmo nesse caso, aparentemente absurdo, a comparação estatística não é de todo impossível, bastando para isso que se encontre um denominador comum, capaz de permitir o confronto entre esses dois tipos de crescimento tão diferentes. Por exemplo, a variável adequada nesse caso, poderia perfeitamente ser algo como a taxa de crescimento de cada fruto em relação ao seu próprio peso, ou ao seu próprio volume, taxa essa considerada a intervalos regulares, ao longo do período de tempo estabelecido para as observações.

Resumindo, embora as variações (ou variâncias), nos dois tipos de crescimento considerados em nossa hipótese, pudessem ser heterogêneas e desproporcionais, as variações das taxas de crescimento relativo poderiam ser homogêneas, e portanto compatíveis e passíveis de comparação.

8.1. Teste de Cochran

Um teste muito simples e de fácil execução, para verificar a homogeneidade das variâncias, é o teste de Cochran que consiste em calcular todas as variâncias envolvidas no experimento e dividir a maior delas pela soma de todas. O valor resultante da divisão é então comparado com os valores críticos de uma tabela estatística apropriada, que leva em conta o número de variâncias envolvidas (k) e o número de graus de liberdade (gl) utilizado nos cálculos, número esse que evidentemente deve ser o mesmo para todas, pois a tabela é construída dessa forma.

Talvez a única dificuldade na execução desse teste seja decidir quais variâncias testar. A experiência de muitos anos acabou me ensinando que as variâncias que melhor se prestam a essa finalidade são as que se referem à interação maior envolvida no plano geral do experimento (binárias ou ternárias, conforme o experimento fatorial tenha dois ou três fatores de variação). Este teste está descrito no item 13.4.

Interpretação do resultado do teste de Cochran.

O teste de Cochran é um teste curioso, porque nele é de interesse que o valor calculado seja menor do que o valor crítico da tabela, e não maior como ocorre na maioria dos testes estatísticos, pois é exatamente isso que indica que as variâncias são homogêneas. De fato, se o valor calculado fosse maior, o resultado seria significativo, o que negaria a hipótese de igualdade (ou de homogeneidade) das variâncias envolvidas no experimento.

Dessa forma, já sabemos se a distribuição dos erros de nosso experimento é ou não normal e se as variâncias são ou não homogêneas. Se a distribuição for normal e as variâncias homogêneas, estamos autorizados a usar os testes paramétricos.

Caso as variâncias não sejam homogêneas, ou a distribuição não seja normal, só há duas alternativas, ou tentamos uma transformação de dados ou testes paramétricos, ou testes não-paramétricos.

Antes disso, porém, sobrou ainda um último detalhe, que até agora não foi resolvido, o problema da independência dos erros.

O ajuste dos erros à distribuição normal é uma pressuposição fundamental, pois sobre esta se baseiam a maioria dos testes clássicos para comparação de populações. A homogeneidade das variâncias deve ser a última a ser testada, não porque é menos importante, mas porque a eficiência de alguns testes para verificar a homogeneidade, depende da distribuição normal dos erros.

8.2. Teste do F máximo de Hartley

Dentre os testes utilizados para testar a homogeneidade das variâncias está o teste do F máximo de *Hartley* e o teste de *Bartlett*, ambos utilizados para testar a homogeneidade das variâncias de mais de duas populações ($t > 2$).

O teste do F máximo é utilizado para verificar a homogeneidade das variâncias de mais de duas populações ($t > 2$). Pela simplicidade de cálculo é um teste muito utilizado, mas deve ser evitado por apresentar baixa eficiência. Os passos para o teste são: hipóteses, significância, estatística apropriada e conclusão.

1o Passo: *Hipóteses*

H_0 : As variâncias entre as populações são homogêneas, ou seja, $\sigma^2 = \sigma_1^2 = \dots = \sigma_t^2$.

H_1 : As variâncias entre as populações são heterogêneas.

2o Passo: *Escolha da significância α .*

3o Passo: *Estatística apropriada*

A estatística apropriada é definida pelo quociente entre a maior e a menor variância calculada entre os tratamentos (t). A expressão é apresentada a seguir:

$$F_{Máx} = \frac{S^2_{Max}}{S^2_{Min}}$$

Onde: S^2_{Max} é a maior variância entre os tratamentos e S^2_{Min} é a menor variância entre os tratamentos.

4o Passo: Conclusão

Quando o valor calculado da estatística $F_{Máx}$, for maior que o valor tabelado F_t para t tratamentos ; $\nu = r-1$ graus de liberdade (Anexo 5) e a significância α , rejeita-se a hipótese H_0 e conclui-se que as variâncias entre as populações são heterogêneas. Mas, se o valor de $F_{Máx}$ calculado for menor que o valor tabelado F_t , aceita-se H_0 e conclui-se que as variâncias entre as populações são homogêneas.

8.3. Teste de Bartlett

O teste é utilizado para comparar as variâncias de mais de duas populações, apresentando uma maior eficiência no julgamento das hipóteses quando comparado ao teste de $F_{Máx}$ de Hartley. Este teste deve ser aplicado quando as populações cujas homogeneidades serão testadas apresentarem distribuição normal. Os passos para o teste são:

1o Passo: Hipóteses

H_0 : As variâncias entre as populações são homogêneas, ou seja, $\sigma_1^2 = \sigma_2^2 = \dots = \sigma_t^2$.
 H_1 : As variâncias entre as populações são heterogêneas.

2o Passo: Escolha da significância α .

3o Passo: Estatística apropriada.

$$\chi_c^2 = \left[\sum_{i=1}^t (n_i - 1) \right] * \ln s^{-2} - \sum_{i=1}^t (n_i - 1) \ln s_i^2$$

$$s^{-2} = \frac{\sum_{i=1}^t (n_i - 1) * s_i^2}{\sum_{i=1}^t (n_i - 1)} ; n_i = r_i$$

Onde:

n_i : número de repetições do i -ésimo tratamento (r_i);

t : número de tratamentos; ln: logaritmo natural;

s^{-2} : variância média dos tratamentos;

s_i^2 : variância do i -ésimo tratamento.

Quando o valor calculado de x_c^2 estiver próximo do valor tabelado é indispensável à utilização do fator de correção dado por:

$$C = 1 + \frac{1}{3(-1)} \left[\sum_{i=1}^t \frac{1}{n_i - 1} - \frac{1}{\sum_{i=1}^t (n_i - 1)} \right]$$

Onde:

n_i = número de repetições do i -ésimo tratamento (r_i);

t = número de tratamentos;

Assim, o valor de x_{aj}^2 ajustado passa a ser:

$$x_{aj}^2 = \frac{x_c^2}{c}$$

4o Passo: *Região crítica*

Verificar a região crítica em função da significância escolhida.

5o Passo: *Conclusão*

Quando o valor calculado de x_c^2 ou x_{aj}^2 for maior que o valor tabelado, x_t^2 para $\nu = t-1$ à significância α , a hipótese H_0 é rejeitada e conclui-se que as variâncias são heterogêneas. Porém, se o valor calculado x_c^2 ou x_{aj}^2 for menor que o valor tabelado x_t^2 , aceita-se H_0 e conclui-se que as variâncias são homogêneas.

9. Análise de Variância (ANOVA)

- Conceituação da ANOVA

O objetivo da análise de variância é avaliar se as diferenças observadas entre as médias das amostras são estatisticamente significantes. Uma variação de médias das amostras pode ser consequência da variação amostral ou é uma boa evidência entre as médias das populações.

A variabilidade total ou fontes de variabilidade das amostras pode ser dividida em 2 partes. A primeira parte desta variabilidade é proveniente das populações serem diferentes, denominada variabilidade *entre*, enquanto a segunda parte é causada pelas diferenças dentro de cada amostra, denominada variabilidade *dentro*.

Quanto maior for à variabilidade *entre*, mais forte é a evidência de que as médias das populações são diferentes. Quanto maior for à variabilidade *dentro*, maior será a dificuldade para concluir se as médias das populações são diferentes.

- Premissas da ANOVA

As populações têm a mesma variância.
As amostras são retiradas de populações com distribuição normal.
As amostras são aleatórias e independentes.

- Teste de Hipótese da ANOVA

A classificação dos testes de análise da variância é de acordo com o número de fatores de interesse ou que influem na variável dependente. Por exemplo, na verificação da eficiência do crescimento de uma lavoura tratada com quatro tipos de fertilizantes, cada um dos fertilizantes é um fator. Da mesma maneira, na comparação do consumo de carros abastecidos com o mesmo tipo de combustível, porém de três marcas diferentes, cada marca de combustível é um fator.

- Introdução

Às vezes é preciso comparar médias de duas populações. Por exemplo, para verificar se pessoas com diferentes níveis de renda (alto, médio e baixo) têm, em média, o mesmo peso corporal, é preciso comparar as médias das três populações.

Outras vezes, é preciso comparar várias situações experimentais. Por exemplo, se um pesquisador separa, ao acaso, um conjunto de quatro pacientes e administra uma droga diferente a cada grupo, terá que comparar as médias dessas quatro "populações".

Para essa análise é utilizado a ANOVA, que é um método que tem como objetivo testar a igualdade de três ou mais médias populacionais, baseado na análise de variâncias amostrais.

Primeiramente, vamos levantar as hipóteses da ANOVA:

→ H_0 : Não existem diferenças entre as médias.

→ H_1 : Pelo menos uma das médias testadas é diferente das demais.

Para testar essa hipótese, temos que obedecer às seguintes premissas:

- as populações têm distribuições normais (Teste de Normalidade),
- as populações têm a mesma variância (Teste de homogeneidade das variâncias),

9.1. Teste F (mesmo número de repetições)

Ex.1:

Imagine que quatro amostras casuais simples, todas com quatro elementos mas cada uma proveniente de uma população, conduziram os dados apresentados na tabela abaixo. As médias dessas amostras estão na última linha dessa tabela. Será que as diferenças de médias das amostras são suficientemente grandes para que se possa afirmar que as médias das populações são diferentes? Vamos aplicar um teste estatístico.

Tabela 1 – Resultados das amostras

AMOSTRAS				
	A	B	C	D
	11	8	5	4
	8	5	7	4
	5	2	3	2
	8	5	3	0
	8	5	7	0
SOMA	40	25	25	10
MÉDIA	8	5	5	2

- Análise de Variância para Experimentos ao Acaso

Se a variável em estudo tem uma distribuição normal, para comparar mais de duas médias, aplica-se o Teste F. Primeiro é preciso estudar as causas de variação. Por que os dados variam? Uma explicação é o fato das amostras provirem de populações diferentes. Outra explicação é o acaso, porque mesmos dados provenientes da mesma população variam.

A tabela 2 apresenta os dados de **k tratamentos**, cada um com **r repetições** (no exemplo, denominam-se repetições os elementos de uma mesma amostra). A soma de r repetições de um mesmo tratamento constitui o total desse tratamento. O total geral é dado pela soma dos k totais de tratamentos.

Tabela 2 – k tratamentos

	Tratamentos					Total
	1	3	3	...	k	
	x_{11}	x_{21}	x_{31}		x_{k1}	
	x_{12}	x_{22}	x_{32}		x_{k2}	
	.	.	.			
	.	.	.			
	x_{1r}	x_{2r}	x_{3r}		x_{kr}	
Total	T_1	T_2	T_3		T_K	$\sum T = \sum x$
Nº de repetições	r	R	r		r	$n = kr$
Média	\bar{x}_1	\bar{x}_2	\bar{x}_3		\bar{x}_k	

Para fazer a análise de variância é preciso calcular as seguintes quantidades:

→ **os graus de liberdade:**

- A) De tratamentos: $K - 1$;
- B) Do total: $n - 1$
- C) Do resíduo: $(n - 1) - (K - 1) = n - k$

→ **o valor C:** dado pelo total geral elevado ao quadrado e dividido pelo número de dados. O valor C é chamado de “correção” e é obtido pela seguinte fórmula:

$$C = \frac{(\sum x)^2}{n}$$

→ **a soma de quadrados total:**

$$SQT = \sum x^2 - C$$

→a soma de quadrados de tratamentos:

$$SQT_r = \frac{\sum T^2}{r} - C$$

→a soma dos quadrados de resíduo:

$$SQR = SQT - SQT_r$$

→o quadrado médio de tratamentos:

$$QMT_r = \frac{SQT_r}{K - 1}$$

→o quadrado médio do resíduo:

$$QMR = \frac{SQR}{n - K}$$

→o valor de F:

$$F = \frac{QMT_r}{QMR}$$

Em seguida, é preciso comparar o valor calculado de F com o valor dado em uma tabela estatística, ao nível de significância estabelecido e com (K-1) graus de liberdade no numerador e (n - k) no denominador. Toda vez que o valor calculado de F for igual ou maior do que o da tabela conclui-se, ao nível de significância estabelecido, que as médias de tratamentos são estatisticamente diferentes, ou seja, não são iguais.

Para fazer a análise da variância com os dados apresentados na tabela 1, é preciso calcular:

→os graus de liberdade:

de tratamentos: $K - 1 = 4 - 1 = 3$

do total: $n - 1 = 20 - 1 = 19$

de resíduos: $n - K = 20 - 4 = 16$

→o valor C:

$$C = \frac{(11+8+\dots+0)^2}{20} = \frac{100^2}{20} = 500$$

→a soma dos quadrados total:

$$SQT = (11^2 + 8^2 + \dots + 0^2) - 500 = 658 - 500 = 158$$

→a soma dos quadrados de tratamentos:

$$SQT_r = \frac{(40)^2 + (25)^2 + (25)^2 + (10)^2}{5} - 500 = 590 - 500 = 90$$

→soma dos quadrados de resíduo:

$$SQR = 158 - 90 = 68$$

→o quadrado médio de tratamentos:

$$QMTr = \frac{90}{3} = 30$$

→o quadrado médio do resíduo:

$$QMR = \frac{68}{16} = 4,25$$

→o valor de F:

$$F = \frac{30}{4,25} = 7,06$$

As quantidades calculadas são apresentadas na tabela 3.

Tabela 3 – ANOVA dos dados da tabela 1

Causas da Variação	G. L.	SQ	QM	F
Tratamentos	3	90	30	7,06
Resíduo	16	68	4,25	
Total	19	158		

Ao nível de significância de 5%, o valor de F na tabela, com 3 e 16 graus de liberdade, é 3,24. Como o valor obtido (F calculado = 7,06) é maior do que F tabelado = 3,24, conclui-se que as médias não são iguais, ao nível de significância de 5%.

9.2. Teste de *Tukey* (Comparação de Médias)

Uma ANOVA permite estabelecer se as médias das populações em estudo são, ou não são, estatisticamente iguais. No entanto, esse tipo de análise não permite detectar quais são as médias estatisticamente diferentes das demais. Por exemplo, a ANOVA apresentada na tabela 3 mostrou que as médias das populações não são iguais, mas não permite concluir qual é, ou quais são, as médias diferentes das demais.

O teste de Tukey permite estabelecer a diferença mínima significativa, ou seja, a menor diferença (d.m.s.) de médias de amostras que deve ser tomada como estatisticamente significativa, em determinado nível. Essa diferença (d.m.s.) é dada por:

$$d.m.s. = q \sqrt{\frac{QMR}{r}}$$

Onde q é um valor fornecido em tabela, QMR é o quadrado médio do resíduo da análise de variância e r é o número de repetições em cada tratamento

A ANOVA apresentada na tabela 3 mostra um valor significativo ao nível de 5 %. Então as médias A,B, C e D não são estatisticamente iguais. Mas qual é, ou quais são, as médias diferentes entre si?

A pergunta pode ser respondida com aplicação do teste de *Tukey*. Ao nível de significância de 5%, o valor de q para comparar 4 tratamentos (A, B, C e D), com 16 graus de liberdade no resíduo é de 4,05.

$$d.m.s. = 4,05 \sqrt{\frac{4,25}{5}} = 3,73$$

De acordo com o teste de *Tukey*, duas médias são estatisticamente diferentes toda vez que o valor absoluto da diferença entre elas for igual ou superior ao valor da d.m.s..No caso, do exemplo, o valor de d.m.s. é 3,73 (calculado) e os valores absolutos das diferenças entre as médias estão apresentados na tabela 4. É fácil ver que diferença entre as médias A e D (6) é maior do que a d.m.s. (3,73). Então, ao nível de 5 %, a média de A, é significativamente maior do que a média de D.

Tabela 4 – Teste de *Tukey*

Pares de médias	Valor absoluto da diferença
A e B	 8 – 5 = 3
A e C	 8 – 5 = 3
A e D	 8 – 2 = 6
B e C	 5 – 5 = 0
B e D	 5 – 2 = 3
C e D	 5 – 2 = 3

9.3. Análise de Variância (Números de repetições diferentes)

Muitas vezes, o pesquisador dispõe de diversas amostras, cada uma proveniente de uma população, mas essas amostras não têm todas o mesmo tamanho.

Ex.:

Tabela 5 – Resultados das amostras

AMOSTRAS			
	A	B	C
	15	23	19
	10	16	15
	13	19	21
	18	18	14
	15		16
	13		
SOMA	14	19	17
MÉDIA	84	76	85

Para fazer a análise da variância com os dados apresentados na tabela 5, é preciso calcular:

→os graus de liberdade:

de tratamentos: $K - 1 = 3 - 1 = 2$

do total: $n - 1 = 15 - 1 = 14$

de resíduos: $n - K = 14 - 2 = 12$

→o valor C:

$$C = \frac{(245)^2}{15} = 4001,67$$

→a soma dos quadrados total:

$$SQT = (15)^2 + (10)^2 + \dots + (16)^2 - 4001,67 = 159,33$$

→a soma dos quadrados de tratamentos:

$$SQT_r = \frac{(84)^2}{6} + \frac{(76)^2}{4} + \frac{(85)^2}{5} - 4001,67 = 63,33$$

→soma dos quadrados de resíduo:

$$SQR = 159,33 - 63,33 = 96$$

→o quadrado médio de tratamentos:

$$QMTr = \frac{63,33}{2} = 31,67$$

→o quadrado médio do resíduo:

$$QMR = \frac{96}{12} = 8$$

→o valor de F:

$$F = \frac{31,67}{8} = 3,96$$

As quantidades calculadas são apresentadas na tabela 6.

Tabela 6 – ANOVA dos dados da tabela 5

Causas da Variação	G. L.	SQ	QM	F
Tratamentos	2	63,33	31,67	3,96
Resíduo	12	96	8	
Total	14	159,33		

Ao nível de significância de 5%, o valor de F na tabela, com 2 e 12 graus de liberdade, é 3,88. Como o valor obtido (F calculado = 3,96) é maior do que F tabelado = 3,88; conclui-se que as médias não são iguais, ao nível de significância de 5%.

Para comparar as médias de tratamentos duas a duas, aplica-se o teste de *Tukey* que neste caso, é aproximado, porque os tratamentos têm um número diferente de repetições. A diferença mínima significativa (d.m.s.) é dada pela formula:

$$d.m.s. = q \sqrt{\frac{1}{r_1} + \frac{1}{r_2} * \frac{QMR}{2}}$$

Onde r_i é o número de repetições do i-ésimo tratamento e r_j é o número de repetições do j-ésimo tratamento.

No caso do exemplo, para comparar a média de A com a média de B, têm-se:

$$d.m.s. = 3,77 \sqrt{\frac{1}{6} + \frac{1}{4} * \frac{8}{2}} = 4,87$$

Para comparar a média de A com a média de C, têm-se:

$$d.m.s. = 3,77 \sqrt{\frac{1}{6} + \frac{1}{5} * \frac{8}{2}} = 4,57$$

Para comparar a média de B com a média de C, têm-se:

$$d.m.s. = 3,77 \sqrt{\frac{1}{4} + \frac{1}{5} * \frac{8}{2}} = 5,06$$

Tabela 7 – Teste de Tukey

Pares de médias	Valor absoluto da diferença
A e B	 14 – 19 = 5
A e C	 14 – 17 = 3
B e C	 19– 17 = 2

Os valores obtidos da diferenças entre as médias estão na tabela 7 (3,77). Como o valor absoluto da diferença entre A e B é maior do que a respectiva d.m.s., conclui-se que, em média, A é diferente de B, ao nível de significância de 5 %.

10. Teste de Hipóteses

O teste de hipóteses é uma das aplicações em estatística mais utilizada. Via de regra, a hipótese nula é feita com base no comportamento passado do produto/processo/serviço, enquanto a alternativa é formulada em função de alterações/inovações recentes. Além disso, permite confirmar a eficácia das medidas de melhoria adotadas, no ambiente atual de melhoria contínua.

Ao testar a hipótese, toma-se uma amostra aleatória do sistema em estudo e se calcula o parâmetro desejado. Conforme o valor do parâmetro, a hipótese nula será aceita ou rejeitada, a partir de procedimentos estatísticos.

- Passos para realizar um Teste de Hipóteses:

Passo 1 : Definição da Hipótese

O primeiro passo é o estabelecimento das hipóteses: hipótese nula e hipótese alternativa.

Hipótese Nula (Ho): É um valor suposto para um parâmetro. Se os resultados da amostra não forem muito diferentes de Ho, ela não poderá ser rejeitada.

Hipótese Alternativa (H₁) : É uma hipótese que contraria a hipótese nula, complementar de Ho. Essa hipótese somente será aceita se os resultados forem muito diferentes de Ho.

Passo 2: Calcular a estatística do Teste

É o valor calculado a partir da amostra, que será usado na tomada de decisão. Uma maneira de tomar uma decisão é comparar o valor tabelado com a estatística do teste.

Para o caso de testes de médias, a estatística do teste é a variável padronizada Z:

$$Z_{cal} = \frac{(\bar{X} - \mu)}{(\sigma/\sqrt{n})}$$

Passo 3: Região Crítica

O valor da estatística do teste, no caso, o valor Z, é calculado supondo que a hipótese nula (H_0) é verdadeira. No entanto, o valor calculado pode estar associado a uma probabilidade de ocorrência muito baixa. Nesse caso, a hipótese nula deve ser rejeitada e aceitamos a hipótese alternativa.

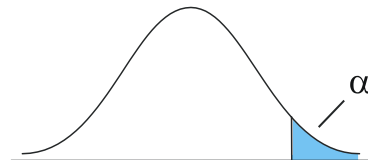
A região crítica é a região onde H_0 é rejeitada. A área da região crítica é igual ao nível de significância (α), que estabelece a probabilidade de rejeitar H_0 quando ela é verdadeira.

Por exemplo, se utilizarmos o nível de significância de 5%, a probabilidade de rejeitar H_0 quando ela é verdadeira é igual a 5%. Na prática, os valores usuais de alfa (α) são 0,01 ou 0,05 ou 0,10.

Unilateral à esquerda:

$H_0: \mu = 50$

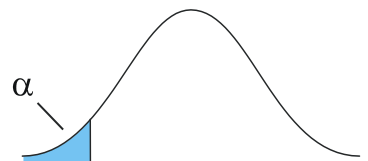
$H_1: \mu > 50$



Unilateral à direita:

$H_0: \mu = 50$

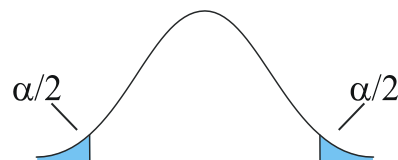
$H_1: \mu < 50$



Bilateral:

$H_0: \mu = 50$

$H_1: \mu \neq 50$



Passo 4. Regra de Decisão:

Se o valor da estatística do teste cair na região crítica, rejeita-se H_0 . Ao rejeitar a hipótese nula (H_0), existe uma forte evidência de sua falsidade.

Ao contrário, quando aceitamos, dizemos que não houve evidência amostral significativa no sentido de permitir a rejeição de H_0 .



Passo 5: Conclusão

Aceitar H_0 , implica que a hipótese nula não pode ser rejeitada.

Rejeitar H_0 implica que temos evidências estatísticas para rejeitá-la com um risco conhecido : α .

1. Comparação de médias, variância conhecida
2. Comparação de médias, variância desconhecida
3. Comparação de pares de observações
4. Comparação de variâncias

10.1. Comparação de médias, variância conhecida

Suponha que X é uma variável aleatória com média μ desconhecida e variância conhecida. E queremos testar a hipótese de que a média é igual a certo valor especificado μ_0 . O teste de hipótese pode ser formulado como segue:

$$H_0 : \mu = \mu_0$$

$$H_1 : \mu \neq \mu_0$$

Para testar a hipótese, toma-se uma amostra aleatória de n observações e calcula-se a estatística:

$$Z_o = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$$

Note que o teste é feito usando-se σ / \sqrt{n} no denominador, uma vez que esse é o desvio padrão da média.

A hipótese H_0 é rejeitada se $|Z_0| > |Z_{\alpha/2}|$ onde $\pm Z_{\alpha/2}$ é um valor limite da distribuição normal reduzida tal que a probabilidade de se obter valores extremos a $\pm Z_{\alpha/2}$ é α .

A probabilidade do valor de Z_0 acontecer, segundo a hipótese nula é menor do que α , logo a hipótese nula H_0 é rejeitada.

Se \bar{X} resultar próximo de μ_0 , $|Z_0| \leq |Z_{\alpha/2}|$ a hipótese H_0 é aceita;

Se \bar{X} resultar longe de μ_0 , $|Z_0| > |Z_{\alpha/2}|$ a hipótese H_0 é rejeitada

- Teste de Hipótese para a média

EX 1.:

A resistência à tração do aço inoxidável produzido numa usina permanecia estável, com uma resistência média de 72 kg/mm² e um desvio padrão de 2,0 kg/mm². Recentemente, a máquina foi ajustada. A fim de determinar o efeito do ajuste, 10 amostras foram testadas.

76,2 78,3 76,4 74,7 72,6 78,4 75,7 70,2 73,3 74,2

Presuma que o desvio padrão seja o mesmo que antes do ajuste. Podemos concluir que o ajuste mudou a resistência à tração de aço? (Adote um nível de significância de 5%)

Passo 1 : Definição da Hipótese

$H_0: \mu = 72 \text{ kg/mm}^2$

$H_1: \mu \neq 72 \text{ kg/mm}^2$

$s = 2 \text{ kg/mm}^2$

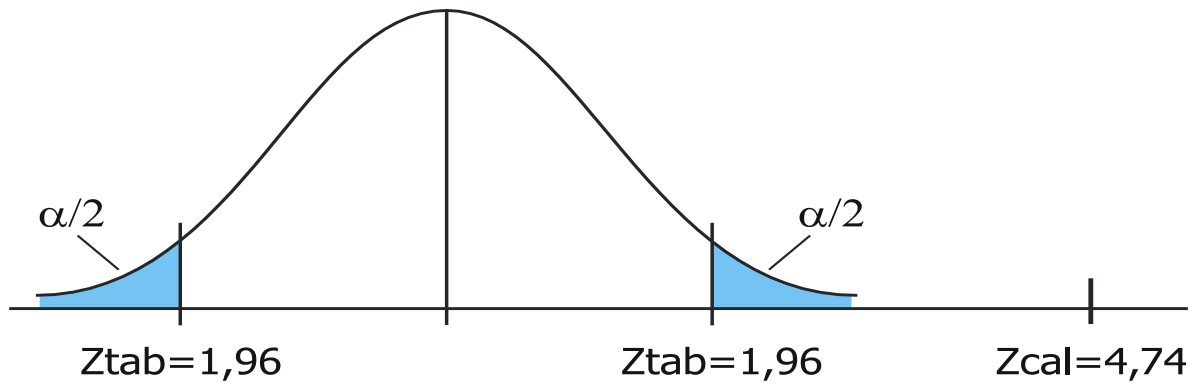
Passo 2: Calcular a estatística do Teste

Sendo $\bar{X} = 75,0$ e $s = 2 \text{ kg/mm}^2$, temos:

$$Z_{cal} = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} = \frac{75 - 72}{2/\sqrt{10}} = \frac{3}{0,6325} = 4,74$$

Isso significa que a média da amostra retirada aleatoriamente da produção está a 4,74 desvios-padrão da média alegada em H_0 que é 72.

Passo 3: Região Crítica



Passo 4: Regra de Decisão

Como o valor crítico para 5% é 1,96 desvios (Z tabelado), estamos na região de rejeição de H_0 .

Passo 5: Conclusão

H_0 é rejeitada e concluímos que a resistência à tração do aço mudou.

EX 2.:

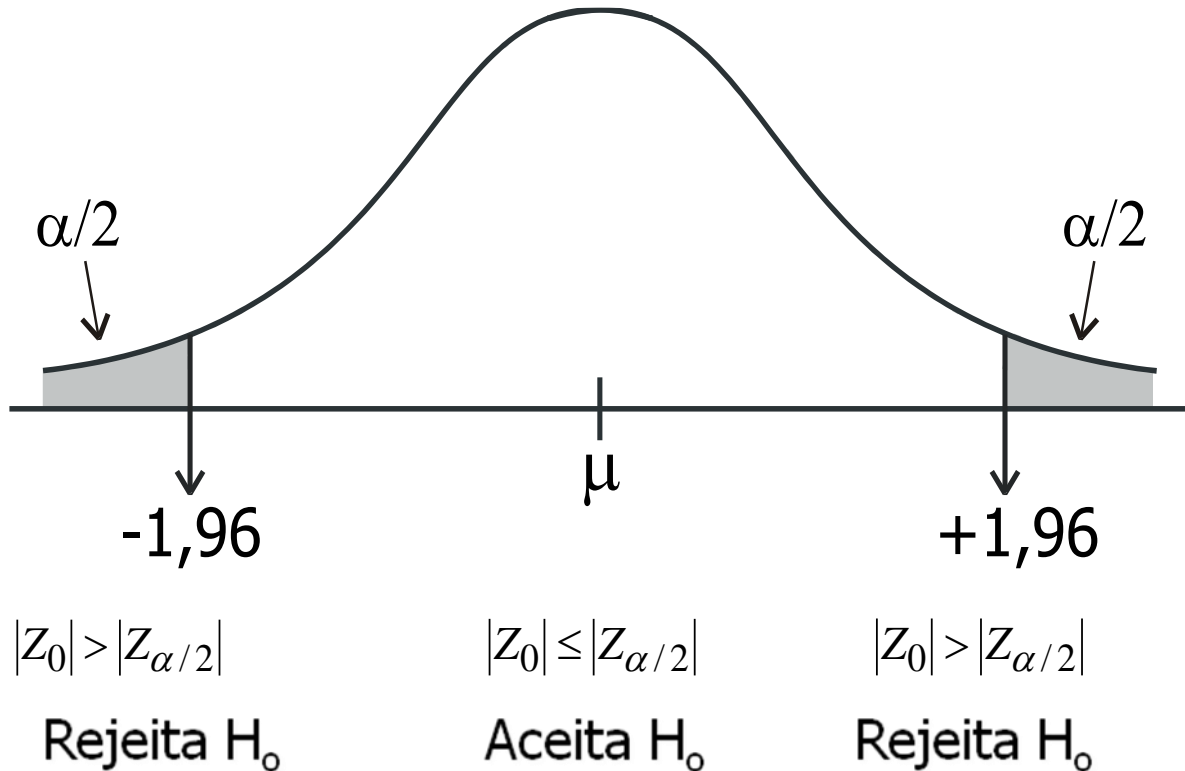
Um processo deveria produzir bancadas com 0,85 m de altura. O engenheiro desconfia que as bancadas que estão sendo produzidas são diferentes que o especificado. Uma amostra de 8 valores foi coletada e indicou $\bar{X} = 0,87$. Sabendo que o desvio padrão é $\sigma = 0,010$, teste a hipótese do engenheiro usando um nível de significância $\alpha = 0,05$.

Solução:

$$H_0 : \mu = 0,85$$

$$H_1 : \mu \neq 0,85$$

$$Z_0 = \frac{0,87 - 0,85}{0,010 / \sqrt{8}} = 5,66$$



Em alguns casos, o objetivo pode ser rejeitar H_0 somente se a verdadeira *média* for maior que μ_0 . Assim, a hipótese alternativa unilateral será $H_1 : \mu > \mu_0$, e a hipótese nula será rejeitada somente se $Z_o > Z_\alpha$.

Se o objetivo for rejeitar H_0 somente quando a verdadeira *média* for menor que μ_0 , a hipótese alternativa será $H_1 : \mu < \mu_0$ e a hipótese nula será rejeitada somente se $Z_o < -Z_\alpha$ ou $|Z_o| > |Z_\alpha|$.

Quando há duas populações com médias desconhecidas, μ_0 e μ_1 digamos e variâncias conhecidas, σ_1^2 e σ_2^2 , o teste para verificar a hipótese que as médias sejam iguais é o seguinte:

$$H_0 : \mu_1 = \mu_2$$

$$H_1 : \mu_1 \neq \mu_2$$

Nesse caso, a partir de uma amostra aleatória de n_1 observações da população 1 e n_2 observações da população 2, calculam-se:

$$Z_o = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

E H_0 é rejeitada se $|Z_0| > |Z_{\alpha/2}|$.

No caso da alternativa unilateral $H_1 : \mu_1 > \mu_2$, a hipótese nula H_0 será rejeitada quando $Z_0 > Z_\alpha$.

E se a alternativa unilateral for $H_1 : \mu_1 < \mu_2$, a hipótese nula H_0 será rejeitada quando resultar $Z_0 < -Z_\alpha$ ou $|Z_0| > |Z_\alpha|$.

Tabela 1: Teste de Médias, Variância Conhecida

Hipótese	Estatística	Critério para rejeitar H_0
$H_0 : \mu = \mu_0$ $H_1 : \mu \neq \mu_0$	$Z_0 = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$	$ Z_0 > Z_{\alpha/2} $
$H_0 : \mu = \mu_0$ $H_1 : \mu > \mu_0$		$Z_0 > Z_\alpha$
$H_0 : \mu = \mu_0$ $H_1 : \mu < \mu_0$		$Z_0 < -Z_\alpha$ ou $ Z_0 > Z_\alpha $
$H_0 : \mu_1 = \mu_2$ $H_1 : \mu_1 \neq \mu_2$	$Z_0 = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$	$ Z_0 > Z_{\alpha/2} $
$H_0 : \mu_1 = \mu_2$ $H_1 : \mu_1 > \mu_2$		$Z_0 > Z_\alpha$
$H_0 : \mu_1 = \mu_2$ $H_1 : \mu_1 < \mu_2$		$Z_0 < -Z_\alpha$ ou $ Z_0 > Z_\alpha $

10.2 Comparação de médias, variância desconhecida

Suponha que X é uma variável aleatória normal com média μ e variância σ^2 desconhecidas. Para testar a hipótese de que a média é igual a um valor especificado μ_0 , formulamos:

$$H_0 : \mu = \mu_0$$

$$H_1 : \mu \neq \mu_0$$

Esse problema é idêntico aquele da seção anterior, exceto que agora a variância é desconhecida. Como a variância é desconhecida, é necessário fazer a suposição adicional de que a variável tenha distribuição normal.

Essa suposição é necessária para poder desenvolver a estatística do teste, contudo, os resultados ainda serão válidos se o afastamento da normalidade não for forte.

Como σ^2 não é conhecido, usa-se a distribuição de student para construir a estatística do teste:

$$t_o = \frac{\bar{X} - \mu_0}{S / \sqrt{n}}$$

E a hipótese nula $H_0 : \mu = \mu_0$ é rejeitada se $|t_o| > t_{\alpha/2, n-1}$, onde $t_{\alpha/2}$ é um valor limite da distribuição de Student tal que a probabilidade de se obter valores externos a $t_{\alpha/2}$ é α . A Tabela 2 mostra os testes apropriados para os casos de hipóteses unilaterais.

Tabela 2 - Teste de Médias, Variância desconhecida

Hipótese	Estatística	Critério para rejeitar H_0
$H_0 : \mu = \mu_0$ $H_1 : \mu \neq \mu_0$	$t_o = \frac{\bar{X} - \mu_0}{S / \sqrt{n}}$	$ t_o > t_{\alpha/2, n-1}$
$H_0 : \mu = \mu_0$ $H_1 : \mu > \mu_0$		$t_o > t_{\alpha, n-1}$
$H_0 : \mu = \mu_0$ $H_1 : \mu < \mu_0$		$t_o < -t_{\alpha, n-1}$ ou $ t_o > t_{\alpha, n-1}$
$H_0 : \mu_1 = \mu_2$ $H_1 : \mu_1 \neq \mu_2$	$t_o = \frac{\bar{x}_1 - \bar{x}_2}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$ $v = n_1 + n_2 - 2$	$ t_o > t_{\alpha/2, v}$
$H_0 : \mu_1 = \mu_2$ $H_1 : \mu_1 > \mu_2$	$t_o = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$	$t_o > t_{\alpha, v}$
$H_0 : \mu_1 = \mu_2$ $H_1 : \mu_1 < \mu_2$	$v = \frac{[(S_1^2/n_1) + (S_2^2/n_2)]^2}{\frac{(S_1^2/n_1)^2}{n_1+1} + \frac{(S_2^2/n_2)^2}{n_2+1}} - 2$	$t_o < -t_{\alpha, v}$ ou $ t_o > t_{\alpha, v}$

10.3 Teste de Hipóteses para a média (desvio padrão desconhecido)

Um trecho de uma rodoviária estadual, quando é utilizado o radar, é verificado em média 7 infrações diárias por excesso de velocidade. O chefe de polícia acredita que este número pode ter aumentado. Para verificar isso, o radar foi mantido por 10 dias consecutivos. Os resultados foram:

8, 9, 5, 7, 8, 12, 6, 9, 6, 10

Os dados trazem evidência de aumento nas infrações?

Passo 1 : Definição da Hipótese

$H_0: m = 7$

$H_1: m > 7$

Passo 2: Calcular a estatística do Teste

Temos $\bar{X} = 8$.

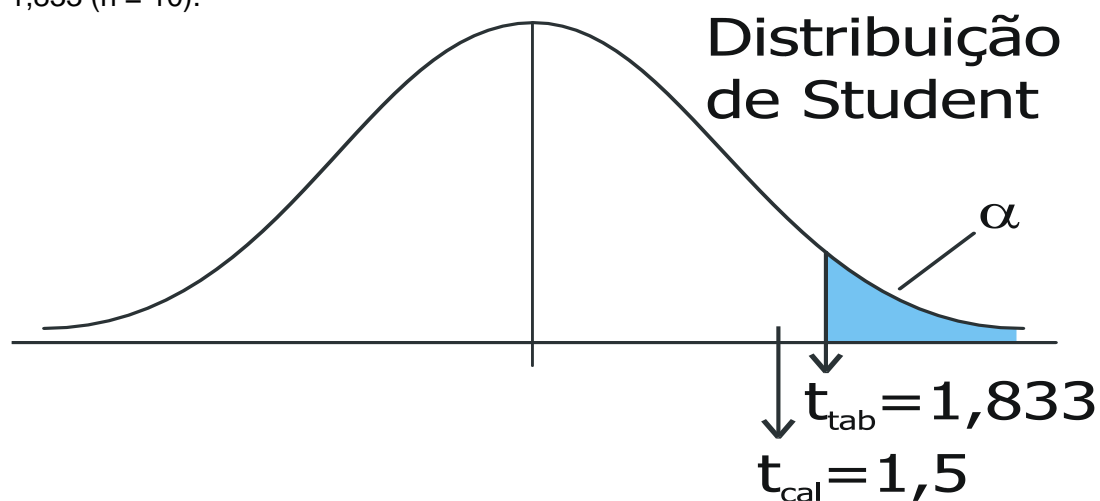
Não conhecendo σ , estimamos por S (desvio-padrão da amostra), logo, $S = 2,11$ (Desvio-padrão foi estimado a partir dos 10 resultados obtidos acima, ou seja, do número de amostras coletadas), deve-se usar a estatística t-student.

$$t_{cal} = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} = \frac{8 - 7}{2,10/\sqrt{10}} = \frac{1}{0,666} = 1,5$$

Isso significa que a média da amostra retirada aleatoriamente da produção está a 1,5 desvios-padrão da média alegada em H_0 que é 7.

Passo 3: Região Crítica

O valor tabelado de t depende do nível de significância (5%) e dos graus de liberdade, que são função do tamanho da amostra: $GL = n - 1 = 9$. Nesse exemplo, t tabelado = 1,833 ($n = 10$).



Passo 4: Regra de Decisão

O valor calculado de t está dentro da região de aceitação de H_0 .

Passo 5: Conclusão

Como aceitamos H_0 , a conclusão é que não houve um aumento significativo no número de infrações. Veja que, apesar de 8 ser maior que 7, a diferença não foi significativa para concluir que o número de infrações aumentou. É como se não houvesse provas suficientes para condenar o réu. E além disso, o t calculado se encontra na região de aceitação de H_0 .

Outro exemplo: Um empresário desconfia que o tempo médio de espera para atendimento de seus clientes é superior a 20 minutos. Para testar essa hipótese ele entrevistou 20 pessoas e questionou quanto tempo demorou para ser atendido. O resultado dessa pesquisa aparece a seguir:

22	20	21	23	22	20	23	22	20	24
21	20	21	24	22	22	23	22	20	24

$$H_0 : \mu = 20 \text{ min}$$

$$H_1 : \mu > 20 \text{ min}$$

$$\bar{X} = 21,8 \text{ min}$$

$$S = 1,40 \text{ min}$$

$$t_o = \frac{\bar{X} - \mu_o}{S / \sqrt{n}} = \frac{21,8 - 20}{1,40 / \sqrt{20}} = 5,75 \quad |t_o| = 5,75 > t_{0,05,19} = 1,729$$

Rejeita-se H_0 . (Há diferenças significativas entre os resultados, ou seja, o tempo médio de atendimento aumentou.).

10.4 Teste de Hipóteses para comparação de médias (independentes)

Existem situações que queremos comparar duas amostras independentes, por exemplo, queremos verificar se existe diferença significativa entre dois lotes em relação à média de uma característica de qualidade importante.

Neste caso, temos duas amostras e utilizaremos a diferença entre as médias amostrais. Se esta diferença for significativa, dizemos que as populações possuem médias diferentes quanto à característica utilizada.

Passo 1 : Definição da Hipótese

Quando há duas populações normais com médias e variâncias desconhecidas, as hipóteses para testar se as médias são iguais são as seguintes:

$$H_o : \mu_1 = \mu_2$$

$$H_1 : \mu_1 \neq \mu_2$$

Passo 2: Calcular a estatística do Teste

O procedimento do teste irá depender de que $\sigma_1^2 = \sigma_2^2$. Se essa suposição for razoável, então calcula-se a variância combinada.

E a seguir calcula-se a estatística do teste:

$$t_{cal} = \frac{\bar{x}_1 - \bar{x}_2}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \qquad S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

Passo 3: Região Crítica

Similar aos demais testes.

Passo 4: Regra de Decisão

Comparar o valor da estatística do teste t_{cal} com o valor tabelado t_{tab} com $n_1 + n_2 - 2$ graus de liberdade.

Ho será rejeitada se $|t_0| > t_{\alpha/2, n_1 + n_2 - 2}$.

EXEMPLO

- Teste de Hipóteses para comparação de médias (Independentes).

Um engenheiro desconfia que a qualidade de um material pode depender da matéria-prima utilizada. Há dois fornecedores de matéria-prima sendo usados. Testes com 10 observações de cada fornecedor indicaram,

$$\bar{X}_1 = 39 \qquad S_1 = 7 \qquad \bar{X}_2 = 43 \qquad S_2 = 9$$

$$H_o : \mu_1 = \mu_2$$

$$H_1 : \mu_1 \neq \mu_2$$

$$S_p^2 = \frac{(9) \times 7^2 + (9) \times 9^2}{10 + 10 - 2} = 65 \Rightarrow S_p = 8,06$$

$$t_{cal} = \frac{39 - 43}{8,06 \sqrt{\frac{1}{10} + \frac{1}{10}}} = -1,11 \quad t_{cal} = -1,11 < t_{tab} = 2,101$$

⇒ Aceito H_0

Ou seja, a matéria-prima tem a mesma qualidade, não tem diferenças significativas.

Se houver evidências de que $\sigma_1^2 \neq \sigma_2^2$, então a estatística a ser usada é:

$$t_0 = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

e o número de graus de liberdade para t é calculado da forma aproximada:

$$v = \frac{[(S_1^2/n_1) + (S_2^2/n_2)]^2}{\frac{(S_1^2/n_1)^2}{n_1 + 1} + \frac{(S_2^2/n_2)^2}{n_2 + 1}} - 2$$

H_0 será rejeitada se $|t_0| > t_{\alpha/2, v}$. Os testes unilaterais correspondentes aparecem na Tabela 3.

Tabela 3: Teste de Médias, Variância desconhecida

Hipótese	Estatística	Critério para rejeitar H_0
$H_0 : \mu = \mu_0$ $H_1 : \mu \neq \mu_0$	$t_o = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$	$ t_o > t_{\alpha/2, n-1}$
$H_0 : \mu = \mu_0$ $H_1 : \mu > \mu_0$		$t_o > t_{\alpha, n-1}$
$H_0 : \mu = \mu_0$ $H_1 : \mu < \mu_0$		$t_o < -t_{\alpha, n-1}$ ou $ t_o > t_{\alpha, n-1}$
$H_0 : \mu_1 = \mu_2$ $H_1 : \mu_1 \neq \mu_2$	$t_o = \frac{\bar{x}_1 - \bar{x}_2}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$ $v = n_1 + n_2 - 2$	$ t_o > t_{\alpha/2, v}$
$H_0 : \mu_1 = \mu_2$ $H_1 : \mu_1 > \mu_2$	$t_o = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$	$t_o > t_{\alpha, v}$
$H_0 : \mu_1 = \mu_2$ $H_1 : \mu_1 < \mu_2$	$v = \frac{[(S_1^2/n_1) + (S_2^2/n_2)]^2}{\frac{(S_1^2/n_1)^2}{n_1+1} + \frac{(S_2^2/n_2)^2}{n_2+1}} - 2$	$t_o < -t_{\alpha, v}$ ou $ t_o > t_{\alpha, v}$

A forma habitual de apresentar os resultados de uma ANOVA num trabalho científico consiste em apresentar características amostrais de cada grupo (médias e desvios padrões) e depois indicar o valor observado da estatística de Teste F e o valor de *p-value* da mesma.

Exemplos de testes de diferenças entre médias para grupos relacionados ou independentes.

EXEMPLO 1

Dez cobaias foram submetidas ao tratamento de engorda com certa ração. Os pesos em gramas, antes e após o teste são dados a seguir (supõe-se que provenham de distribuições normais). A 1% de significância, podemos concluir

que o uso da ração contribuiu para o aumento do peso médio dos animais?

Cobaia	1	2	3	4	5	6	7	8	9	10
Antes	635	704	662	560	603	745	698	575	633	669
Depois	640	712	681	558	610	740	707	585	635	682

Resolução:

Trata-se de uma situação em que queremos comparar as médias de duas distribuições normais, supondo que se trata da mesma população, mas em dois momentos diferentes, antes e após um tratamento de engorda. Há interesse em verificar se a dieta contribuiu para o peso médio dos animais, ou seja, queremos verificar se a média de peso antes do tratamento é menor do que a média de peso após o tratamento (se a dieta fez efeito e os animais estarão em média mais pesados ao final do tratamento). Reparem que é exigido que se tome uma decisão, o que configura um problema de Teste de Hipóteses.

Iremos então aplicar um teste de diferenças entre médias populacionais, para dados **pareados** (mesma população, antes e depois).

1) *Enunciar as hipóteses*

De acordo com o que foi dito acima, queremos verificar se a média antes é menor do que a média depois; o melhor ponto de partida, que servirá para a definição da hipótese H_0 , é que a dieta não faz efeito, ou seja, as médias antes e após o tratamento são iguais (costumamos colocar em H_0 o contrário do que queremos provar), ou seja, a diferença entre diferença entre as médias deve ser suposta igual a zero, teremos então:

$$H_0: \mu_d = 0$$

$$\mu_d = \mu_{antes} - \mu_{depois}$$

$$H_1: \mu_d < 0$$

2) *Estabelecer o nível de significância ou nível de confiança.*

Conforme foi estabelecido no enunciado do problema:

$$\alpha = 0,01, \text{ ou seja, } 1 - \alpha = 0,99 .$$

3) *Identificar a variável de teste.*

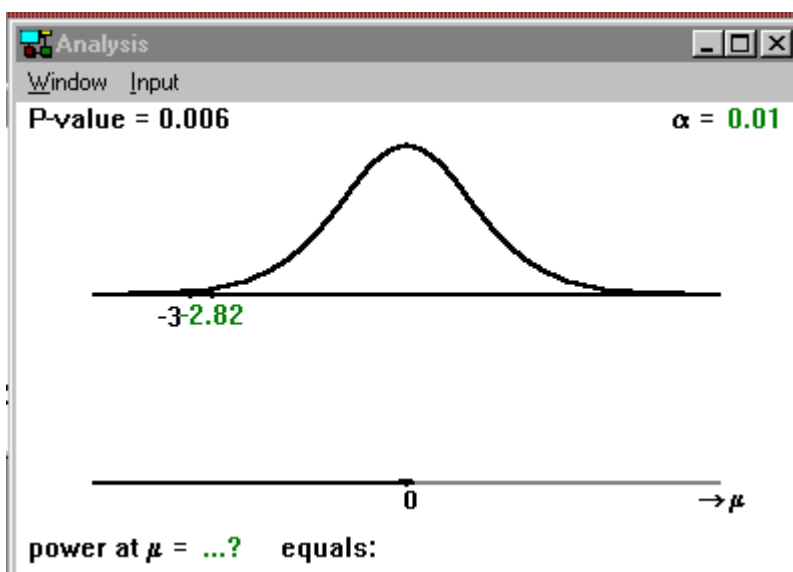
No presente problema, temos uma amostra de apenas 10 elementos. Como a amostra tem menos de 30 elementos, a variável de teste que será utilizada será a variável t_{n-1} da distribuição t de Student.

4) Definir a região de aceitação de H_0 , de acordo com o tipo de teste e variável.

Trata-se de um teste unilateral à esquerda (com 1% de significância), e a variável de teste é t_{n-1} (a amostra tem 10 elementos), então o valor crítico (obtido da tabela da distribuição t de Student) será:

$$t_{n-1, crítico} = t_{10-1; 0,01} = t_{9; 0,01} = -t_{9; 0,99} = 2,82$$

Observe a região de aceitação de H_0 na figura abaixo:



Para valores maiores de -2,82; aceitaremos H_0 (ou seja, a dieta não faz efeito, a diferença entre as médias é nula). Se t_{n-1} for menor do que -2,82 rejeitaremos H_0 (a média depois, aumentou demais em relação à média antes da dieta para que a diferença seja devida apenas ao acaso. Claro que há uma chance de 1% de que venhamos a rejeitar H_0 , sendo ela verdadeira.

5) Através dos valores das amostras antes e depois, calcular a diferença d_i entre cada par de valores, onde

$$d_i = X_{antes} - X_{depois}$$

Para o conjunto sob análise teremos:

Cobaia	1	2	3	4	5	6	7	8	9	10
Antes	635	704	662	560	603	745	698	575	633	669
Depois	640	712	681	558	610	740	707	585	635	682
d_i	-5	-8	-19	2	-7	5	-9	-10	-2	-13
d_i^2	25	64	361	4	49	25	81	100	4	169

6) e 7) Calcular a diferença média e o desvio padrão da diferença média.

Para o presente problema:

$$\bar{d} = \frac{\sum d_i}{n} = \frac{-66}{10} = -6,6 \text{ gramas}$$

$$sd = \sqrt{\frac{\sum d_i^2 - [(\sum d_i)^2 / n]}{n-1}} \quad sd = \sqrt{\frac{882 - [(-66)^2 / 10]}{10-1}} = 7,04 \text{ gramas}$$

8) Calcular o valor da variável de teste.

Neste problema é a variável t_{n-1} :

$$t_{n-1} = \frac{\bar{d}}{(s_d / \sqrt{n})} \quad t_{10-1} = t_9 = \frac{-6,6}{(7,04 / \sqrt{10})} = -2,96$$

9) Decidir pela aceitação ou rejeição de H_0 .

Conforme foi visto anteriormente, se o valor da variável de teste fosse menor do que -2,82 ; a hipótese H_0 seria rejeitada:

$$t_{n-1} = t_9 = -2,96 < t_{n-1,crítico} = -t_{9;0,01} = -2,82$$

Assim, rejeitamos H_0 a 1% de significância.

10) Interpretar a decisão dentro do contexto do problema.

Assim, concluímos com 99% de confiança (ou uma chance de erro de 1%) que a dieta contribuiu para o aumento do peso médio dos animais.

EXEMPLO 2

Deseja-se saber se 2 máquinas de empacotar café estão fornecendo o mesmo peso médio em kg. Extraem-se duas amostras, uma de cada máquina (supondo que os pesos das amostras sigam uma distribuição normal):

Máquina Nova - 36 amostras, média = 0,81 kg, variância = 0,00020 kg².

Máquina Velha - 39 amostras, média = 0,78 kg, variância = 0,00135 kg².

Qual é a sua conclusão a 5% de significância?

Resolução:

Trata-se de uma situação em que queremos comparar as médias de duas distribuições normais, supondo que se trata de duas populações distintas, podemos supor que as amostras são independentes: estão sendo avaliadas as produções de 2 máquinas diferentes. Há interesse em verificar simplesmente se há diferença entre as médias das duas populações: em qualquer direção. Reparem que é exigido que se tome uma decisão, o que configura um problema de teste de hipóteses.

Iremos então aplicar um teste de diferenças entre médias populacionais, para dado não pareados (populações distintas).

1) Enunciar as hipóteses

De acordo com o que foi dito acima, queremos verificar se as médias diferem entre si; o melhor ponto de partida, que servirá para a definição da hipótese H_0 , será considerar que não há diferença entre as médias, ou seja, a média da máquina velha seria igual à da máquina nova (costumamos colocar em H_0 , o contrário do que queremos provar), teremos então:

$$H_0: \mu_1 = \mu_2$$

$$H_1: \mu_1 \neq \mu_2$$

Onde $\mu_1 = \mu_{\text{nova}}$ e $\mu_2 = \mu_{\text{velha}}$

2) Estabelecer o nível de significância ou nível de confiança.

Conforme foi estabelecido no enunciado do problema:

$$\alpha = 0,05 \qquad 1 - \alpha = 0,95 \qquad \alpha/2 = 0,025$$

3) Identificar a variável de teste.

Neste ponto do roteiro é necessário ter muito cuidado. Há 3 variáveis de teste possíveis, dependendo das condições do problema, mais especificamente das variâncias das duas populações.

Se as variâncias de ambas as populações forem conhecidas (caso muito raro na prática) deverá ser usada a variável Z da distribuição normal padrão.

Se as variâncias populacionais forem desconhecidas, deverá ser usada a variável t da distribuição t de Student, mas a determinação dos graus de liberdade desta variável dependerá do fato de as variáveis serem iguais (quando a variável terá $n_1 + n_2 - 2$ graus de liberdade, onde n_1 e n_2 são os tamanhos das amostras), ou diferentes. Uma pergunta bastante razoável que surge é: "Se eu não conheço as variâncias populacionais, como poderei saber se são iguais ou diferentes?"

Para resolver este problema é necessário usar o teste F de diferença entre variâncias populacionais (ou teste de razão entre variâncias, já que a variável do teste é um quociente entre as 2 variâncias amostrais). Em suma, quando fazemos um teste de diferença entre médias de duas populações distintas, devemos inicialmente fazer um teste para verificar se suas variâncias, caso desconhecidas, são iguais ou diferentes.

No presente problema, as variâncias populacionais são desconhecidas, portanto devemos realizar o teste F para obter a variável de teste com o número correto de graus de liberdade.

Teste F

$$H_0 : \sigma^2_1 = \sigma^2_2$$

$$H_1 : \sigma^2_1 \neq \sigma^2_2$$

$$\text{Onde: } \sigma^2_1 = \sigma^2_{nova}$$

$$\sigma^2_2 = \sigma^2_{velha}$$

Queremos apenas verificar se há diferença entre as variâncias, por isso o teste será sempre bilateral.

Nível de significância: como a tabela da distribuição F apresenta valores apenas para 5% de significância (teste bilateral), este será o nível adotado em todos os testes F (se for possível obter outras tabelas, níveis diferentes poderão ser utilizados).

Encontrar a maior variância amostral, que será chamada de s_A^2 (e por conseguinte n_A) e a menor que será chamada s_B^2 (e por conseguinte n_B). Neste problema teremos:

$$s_A^2 = s^2_{velha} = 0,00135$$

$$n_A = 39$$

$$s_B^2 = s^2_{nova} = 0,00020$$

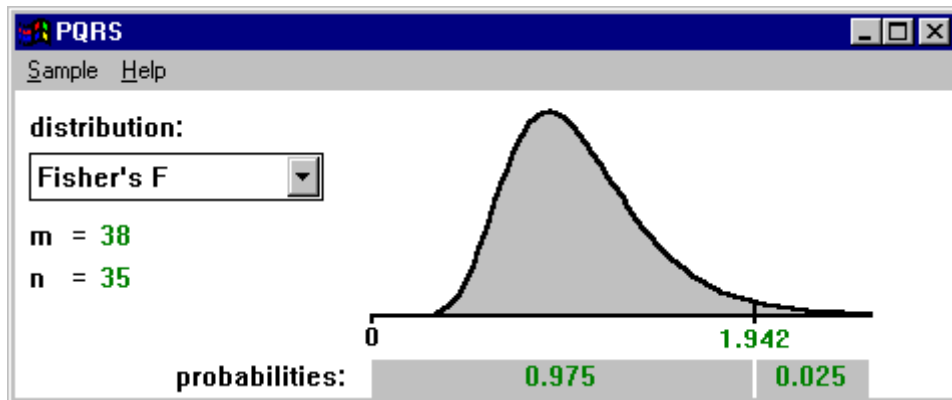
$$n_B = 36$$

Então a variável de teste do teste F será:

$$F_{n_A-1, n_B-1} = \frac{s_A^2}{s_B^2} = F_{38,35} = \frac{0,00135}{0,00020} = 6,75$$

Esta variável segue uma distribuição amostral chamada de distribuição F de Fisher (ou de Snedecor), trata-se de uma distribuição assimétrica, que somente pode assumir

valores positivos, e que possui graus de liberdade associados ao numerador e ao denominador de um quociente. Veja a figura abaixo, de uma distribuição F com 38 e 35 graus de liberdade:



Esta variável será comparada com um valor crítico, que será $F_{38,35;0,025}$: bastaria procurá-lo em uma tabela adequada (como a tabela, sendo 38 o número de graus de liberdade do numerador da estatística e 35 o número de graus de liberdade do denominador da estatística).

Usando a tabela citada acima, porém, só há os valores $F_{30,30;0,025} = 2,07$ e $F_{40,40;0,025} = 1,88$. O valor procurado $F_{38,35;0,025}$ devem estar entre estes 2 valores (na realidade ele vale 1,942). Como o valor da variável $F_{38,35}(6,75)$ é 3 vezes maior do que ambos os valores (2,07 e 1,88), podemos tranqüilamente rejeitar a hipótese de que as variâncias populacionais, desconhecidas, são iguais, com uma chance de erro de 5%. Sendo assim, a nossa variável t de Student terá ν graus de liberdade.

Então, como as variâncias populacionais são desconhecidas e supostas diferentes (de acordo com o teste F) a variável t terá ν graus de liberdade, que serão calculados a seguir:

$$\nu = \left[\frac{(\varpi_1 + \varpi_2)^2}{\frac{\varpi_1^2}{n_1 + 1} + \frac{\varpi_2^2}{n_2 + 1}} \right] - 2$$

$$\varpi_1 = \frac{s_1^2}{n_1} = \frac{0,00020}{36} = 5,56 * 10^{-6} \quad (\text{máquina nova})$$

$$\varpi_2 = \frac{s_2^2}{n_2} = \frac{0,00135}{39} = 3,46 * 10^{-5} \quad (\text{máquina velha})$$

$$v = \left[\frac{(5,56 * 10^{-6} + 3,46 * 10^{-5})^2}{\frac{(5,56 * 10^{-6})^2}{36 + 1} + \frac{(3,46 * 10^{-5})^2}{39 + 1}} \right] - 2 \cong 50,41 \cong 50$$

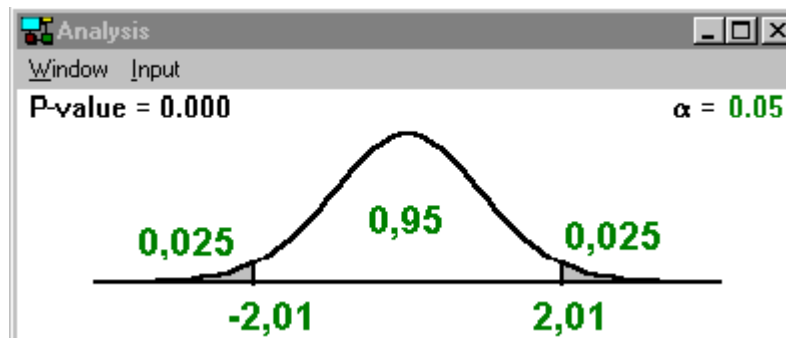
Então a variável de teste t de Student terá 50 graus de liberdade.

4) Definir a região de aceitação de H_0 , de acordo com o tipo de teste e variável.

Trata-se de um teste bilateral (com 5% de significância), e a variável de teste é t com 50 graus de liberdade, então o valor crítico (obtido da tabela da distribuição t de Student) será:

$$|t_{v,critico}| = |t_{50;0,975}| = 2,01$$

Observe a região de aceitação de H_0 na figura abaixo:



Para valores de t_{50} menores do que -2,01 ou maiores do que 2,01, ou seja, valores em módulo maiores do que 2,01, rejeitaremos H_0 , ou seja, há diferença entre as médias de peso dos pacotes das 2 máquinas (claro que há 5% de chance de que venhamos a rejeitar H_0 sendo ela verdadeira).

5) Calcular o desvio padrão das diferenças.

Como as duas variâncias são desconhecidas, e o teste F mostrou que são diferentes, iremos utilizar a terceira expressão para calcular o desvio padrão das diferenças que está no roteiro.

$$sd = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = \sqrt{\frac{0,00020}{36} + \frac{0,00135}{39}} = 6,34 * 10^{-3}$$

6) Calcular a variável de teste.

Novamente, como as duas variâncias são desconhecidas, e o teste F mostrou que são diferentes, iremos utilizar a terceira expressão para calcular o valor da variável de teste que está no roteiro.

$$t_v = \frac{\bar{x}_1 - \bar{x}_2}{s_d} = t_{50} = \frac{0,81 - 0,78}{6,34 * 10^{-3}} = 4,73 \rightarrow |t_{50}| = 4,73$$

7) Decidir pela aceitação ou rejeição de H_0 .

Conforme foi visto anteriormente se o módulo da variável de teste fosse maior do que 2,01 a hipótese H_0 seria rejeitada:

$$|t_v| = |t_{50}| = 4,73 > |t_{v,critico}| = |t_{50;0,975}| = 2,01$$

Assim, rejeitamos H_0 a 5% de significância.

8) Interpretar a decisão dentro do contexto do problema.

Assim, concluímos com 95% de confiança (ou uma chance de erro de 5%) que há diferença entre os pesos médios dos pacotes fornecidos pelas duas máquinas. Seria recomendável descobrir qual das duas está com problemas para efetuar as correções necessárias.

11. Teoria dos Erros

Algumas grandezas possuem seus valores reais conhecidos e outras não. Quando conhecemos o valor real de uma grandeza e experimentalmente encontramos um resultado diferente, dizemos que o valor obtido está afetado de um erro.

Erro é a diferença entre um valor obtido ao se medir uma grandeza e o valor real ou correto da mesma. Matematicamente, o erro é a diferença entre o valor medido e o valor verdadeiro.

É válido lembrar que o valor real ou exato da maioria das grandezas físicas nem sempre é conhecido. O que se sabe, na maioria dos casos, é o valor mais provável desta grandeza, determinado através de experimentos com incerteza. Neste caso, ao efetuarmos uma medida desta grandeza e compararmos com este valor, falamos em desvios e não erros.

Desvio é a diferença entre um valor obtido ao se medir uma grandeza e um valor adotado que mais se aproxima do valor real. Na prática, se trabalha na maioria das vezes com desvios e não erros.

Podem ser, por exemplo, comprimentos, intervalos de tempo, voltagem entre dois pontos, carga elétrica transportada, intensidade luminosa, e muitas outras. Para se caracterizar o sistema de freios de um automóvel, por exemplo, realiza-se uma medida da distância percorrida após o acionamento dos freios, quando o carro se movia a uma determinada velocidade. Ao se realizar uma medida, há sempre **fontes de erro** que a afetam. As **fontes de erro** fazem com que toda medida realizada, por mais cuidadosa que seja, esteja afetada por um **erro experimental**.

O erro é inerente ao próprio processo de medida, isto é, nunca será completamente eliminado, mas poderão ser minimizados, quando puder eliminar o máximo possível as fontes de erros acima citadas. Portanto, ao realizar medidas, é necessário avaliar quantitativamente os erros cometidos.

Quando realizamos medidas experimentais obtemos uma série de valores que em geral não são idênticos. Nosso objetivo é saber qual deve ser o valor mais provável da grandeza medida, qual a diferença entre este valor e cada valor medido em particular.

Por mais cuidadosa que seja uma medição e por mais preciso que seja o instrumento, não é possível realizar uma medida direta perfeita. Ou seja, sempre existe uma incerteza ao se comparar uma quantidade de uma dada grandeza física ou química com sua unidade. Segundo sua natureza, os erros são geralmente classificados em três categorias: aleatórios, sistemáticos e grosseiros.

11.1 Erros Aleatórios

Eles aparecem devido ao fato de que as condições em que é realizado um experimento não são exatamente as mesmas, a cada vez que se repete o mesmo. Suponhamos que se queira estimar o tempo de queda de um corpo que se encontra numa altura **h**. Ao se repetir o processo, se o corpo estiver ligeiramente acima ou abaixo da posição da situação anterior, haverá uma incerteza na altura **h**, que produzirá uma incerteza no tempo de queda. Observa-se neste caso, que uma das variáveis do experimento não está controlada, produzindo flutuações em torno de um valor chamado mais provável. O erro aleatório é inerente a todo processo de medida e deve ser tratado convenientemente. Neste tipo de erro existe uma probabilidade de se errar para mais ou para menos, e o procedimento natural que se usa para tratá-los é a análise estatística.

Praticamente todo o trabalho experimental, ainda que muito dele automatizado, está sujeito a pequenas variações, pois mesmo a instrumentação tem limites quanto ao número de dígitos significativos do valor que pretende quantificar.

Algumas fontes típicas de erros aleatórios:

— flutuações no suprimento de energia, ruído na rede de transmissão e conversão de sinal, mudanças nas condições ambientais. Não podem ser completamente eliminados (estão sempre presentes), mas são geralmente de pequena magnitude.

— método de observação: erros devidos ao julgamento feito pelo observador ao fazer uma leitura abaixo da menor divisão de uma escala, como por exemplo, medir o comprimento de uma folha de papel com uma régua cuja menor divisão é 1 mm com precisão na medida de 0,5 mm;

— flutuações ambientais: mudanças não previsíveis na temperatura, voltagem da linha, correntes de ar, vibrações (por exemplo causadas por passagem de pessoas perto do aparato experimental ou veículos nas vizinhanças).

— erro de paralaxe na leitura de uma escala.

— imperícia do operador.

Os erros aleatórios afetam a precisão e a reprodutibilidade dos dados. Eles podem ser minimizados, por exemplo, pela perícia do operador, mas jamais eliminados por completo.

11.2 Erros Sistemáticos

Os erros sistemáticos são entre outros, normalmente decorrentes de má condução da experiência, má calibração dos instrumentos, descuidos de planejamento. Em qualquer dos casos resultam na distorção da medição, alterando todos os resultados causando um desvio acentuado do valor correto. Suponhamos que uma dada experiência é conduzida em diversos dias e o resultado depende das condições atmosféricas, como por exemplo a umidade. Não sendo esta constante, é natural que se este fator não estiver controlado, contribua para a distorção dos resultados. Outro exemplo, pode ser considerado pensando numa análise onde uma dada mistura é feita num recipiente, sendo a experiência replicada para uma mesma amostra, a não limpeza do recipiente pode contribuir para a alteração da concentração dos reagentes, provocando sucessivas distorções nos resultados.

Os erros sistemáticos (exatidão) são causados por fontes identificáveis, e, em princípio, podem ser eliminados ou compensados. Ocorrem sempre num mesmo sentido.

Estes fazem com que as medidas feitas estejam consistentemente acima ou abaixo do valor real, prejudicando a exatidão da medida. Erros sistemáticos podem ser devido a vários fatores, tais como:

- instrumento de medida; Ex: intervalos de tempo feitos com um relógio que atrasa.
- Método de observação; Ex: medir o instante da ocorrência de um trovão pelo ruído liberado.

- Efeitos ambientais; Ex: medida de um dado comprimento que dependa da temperatura ambiente.
- Simplificações do modelo teórico utilizado; Ex: não incluir o efeito da resistência do ar numa medida da aceleração da gravidade baseada na medida do tempo de queda de um objeto a partir de uma dada altura.
- Uma balança não zerada na ausência de massa sobre seu prato, por aproximações teóricas incorretas.
- O operador que sempre superestima ou subestima os valores de medidas.
- Não calibração do instrumento de medição, interferências, brancos de reagentes não corrigidos, variação instrumental.

Uma das principais tarefas do operador que realiza as medidas é identificar e eliminar o maior número possível de fontes de erro sistemático.

Nota

O erro aleatório não pode ser compensado, mas pode ser reduzido aumentando-se o número de observações. Apesar de frequentemente citado, o desvio padrão da média não é o erro aleatório da média. Representa sim, uma medida da incerteza da média devido aos efeitos aleatórios. O erro sistemático, em geral, não pode ser eliminado, mas pode eventualmente ser reduzido ou, caso seja identificado, deve ser corrigido.

11.3 Erros Grosseiros

Os erros grosseiros podem ser provocados por falhas ocasionais e/ou anormais dos instrumentos, do observador ou de outros parâmetros intervenientes. São normalmente detectáveis, quer por produzirem medições substancialmente fora do esperado, ou por identificação do agente causador do erro. Como consequência, exige a repetição da experiência.

São causados por eventos não-aleatórios como mau funcionamento (instalação inadequada), problema de calibração, desgaste ou corrosão de sensores, depósito de sólidos, falta de qualificação do operador, erro de leituras, erro na transcrição de resultados, contaminação da amostra, desvios no uso do procedimento operacional, reagentes fora da validade ou de especificação, instabilidade do equipamento, erros de cálculo, etc. Têm magnitude e sinal, que podem ser desconhecidos. Se uma medição for repetida com o mesmo instrumento, sob condições idênticas, a contribuição do erro sistemático para o valor medido será a mesma. Com bons procedimentos de instalação e manutenção é possível assegurar que não haverá erros grosseiros nos valores medidos, pelo menos por algum tempo.

Eles devem ser detectados e removidos usando técnicas de filtragem ou controle estatístico. São imprevisíveis, variam de sinal e magnitude.

11.4 Outros Tipos de Erros

- Erros acidentais:

Esses erros são difíceis de eliminar e ocorrem por várias razões. O mais comum é o “chute” do experimentador quando este é obrigado a fazer uma leitura a “olho nu”. Para minimizar este erro só mesmo a experiência do experimentador: aqui a concentração e perícia do operador do aparelho são fundamentais. Podemos representar os erros acidentais da seguinte forma:

→ **Erro absoluto (EA):** é o módulo da diferença entre o valor encontrado e o valor tabelado.

$$EA = |valor_{tabelado} - valor_{encontrado}|$$

Ex. 1: Num experimento onde se quer encontrar o valor da aceleração da gravidade terrestre, o aluno encontra o valor $9,0 \text{ m/s}^2$. Mas o valor tabelado para a gravidade é $9,8 \text{ m/s}^2$. Então o erro absoluto é:

$$EA = |9,8 - 9,0| = 0,8 \text{ m/s}^2$$

→ **Erro relativo percentual (ER):** É o quociente entre o valor do erro absoluto e o valor tabelado multiplicado por 100%. Expresso em porcentagem.

$$ER\% = \frac{(EA)}{valor_{tabelado}} * 100$$

EX. 2: Usando os valores do exemplo 1, temos:

$$ER = \frac{(0,8)}{9,8} * 100 = 8,2\%$$

Nota

• Como não se conhece o valor verdadeiro, o erro relativo é estimado pela razão do desvio padrão amostral e pelo valor médio:

$$ER\% = \frac{S}{X} * 100$$

Geralmente, o erro relativo aceitável é menor que 10%.

Quando se tratar de uma população, utilizar o desvio padrão populacional ou o desvio padrão da média.

→ **Erro Instrumental (ou Erro de Escala)**

A construção de uma escala de um equipamento de medida implica na escolha de subdivisões, em partes iguais, da unidade padrão. No entanto, pode ocorrer que a grandeza a ser medida não corresponda a um número inteiro das subdivisões existentes no aparelho. Deparamo-nos, desta forma, com o problema de estimar a fração da subdivisão considerada. Ao estimarmos esta fração, introduzimos, o erro instrumental que indica o grau de precisão do aparelho utilizado na medida. Este tipo de erro representa a limitação do instrumento. A estimativa do erro instrumental depende da habilidade do experimentador. Geralmente, adota-se o erro instrumental como sendo *igual à metade da menor divisão de escala do instrumento (precisão do aparelho)*. Assim, por exemplo, se a menor divisão de uma régua for o milímetro (mm), então o erro instrumental será de 0,5mm.

→ **Erros Estatísticos**

- A repetição de uma experiência em idênticas condições não fornece resultados idênticos.
- Flutuações estatísticas nos resultados: erros aleatórios ou estatísticos.
- Os erros estatísticos indicam a precisão das medidas.

Erros estatísticos (ou aleatórios) resultam de variações aleatórias nas medições, provenientes de fatores que não podem ser controlados ou que, por algum motivo, não foram controlados. Por exemplo, na medição de massa com balança, correntes de ar ou vibrações (fatores aleatórios) podem introduzir erros estatísticos na medição. São aqueles cometidos na definição e aplicação da técnica estatística.

Nota:

A distinção entre erro sistemático e estatístico é um tanto arbitrária. A distinção entre os dois tipos de erros depende do universo de medidas considerado.

Exemplos:

→ Normalmente o erro de calibração de um instrumento é considerado sistemático. Entretanto, quando são consideradas medições com instrumentos semelhantes, mas de diferentes marcas e procedências, espera-se que os erros de calibração sejam estatísticos.

→ Em um determinado equipamento, flutuações aleatórias da temperatura podem resultar em erro estatístico. Entretanto, pode existir correlação entre a flutuação da temperatura e os resultados da própria experiência, levando a um erro sistemático.

→ **Erro Tolerável**

É necessário utilizar um critério para a eliminação de dados que se apresentem com um erro muito grande.

$$E_{Tolerável} = 3 * \sigma$$

O critério para se descartar um dado é quando seu erro ($X_i - \mu$) for maior do que o erro tolerável.

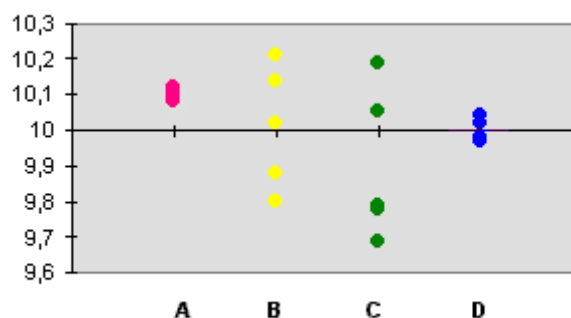
Exemplo

Tomemos como exemplo uma experiência realizada por quatro alunos numa análise com 10.00 ml de 0.1 M de hidróxido de sódio titulado com 0.1M de ácido clorídrico. Cada estudante realizou cinco replicadas tendo-se obtido os resultados:

Alunos	A	B	C	D
	10,08	9,88	10,19	10,04
	10,11	10,14	9,79	9,98
	10,09	10,02	9,69	10,02
	10,10	9,80	10,05	9,97
	10,12	10,21	9,78	10,04
Média	10,10	10,01	9,90	10,01
D. Padrão	0,0158	0,1718	21,0500	0,0332

O valor correto seria de 10 ml, o que nos permite com este conjunto de dados fazer uma distinção entre precisão e exatidão, caracterizado assim o tipo de erro inerente.

Para os estudantes A e D, verificamos que os dados são precisos, com uma dispersão pequena, erros aleatórios reduzidos. No entanto, enquanto que para o aluno D eles são também exatos, o mesmo não se pode afirmar para A. Neste conjunto de dados, existe um nítido desvio relativamente ao valor correto 10, ou seja, um erro sistemático evidente. Para os alunos B e C, os erros aleatórios são de uma ordem de grandeza superior aos de A e D, portanto menos precisos, mas para o aluno B são exatos enquanto que para C é inexato, o desvio em relação ao valor 10 é de 0.01 enquanto que para C este valor é 0.1. O quadro seguinte dá-nos essas mesmas indicações graficamente.



Podemos desta forma, classificar os dados relativos a cada aluno da seguinte forma:

Tipo de Erro		
Aluno	Aleatório	Sistemático
A	Precisos	Inexatos
B	Imprecisos	Exatos
C	<IMPRECISOS< TD>	Inexatos
D	Precisos	Exatos

Como conclusão, podemos afirmar que os dados obtidos pelo aluno C são inaceitáveis, e os de D os mais aceitáveis. A classificação de A e B em termos de aceitabilidade é no entanto mais delicada.

São devidos a causas diversas e incoerentes, bem como a causas temporais que variam durante observações sucessivas e que escapam a uma análise em função de sua imprevisibilidade. Podem ter várias origens, entre elas:

- Os instrumentos de medida;
- Pequenas variações das condições ambientais;
- Fatores relacionados com o próprio observador sujeito à flutuações, em particular visão e audição.

Os erros sistemáticos podem ser corrigidos e aproximados os dados ao valor correto, se por exemplo, tiverem origem numa errada calibração dos instrumentos. No entanto, ao contrário do exemplo apresentado, na prática normalmente os valores exatos não são conhecidos, não havendo portanto possibilidade de identificá-los. Nestes casos, evitar erros sistemáticos só é possível eliminando à partida a possibilidade da sua ocorrência, com uma cuidadosa preparação e análise dos métodos utilizados.

De uma forma geral, os erros aleatórios são mais difíceis de controlar, mas uma técnica cuidadosa pode minimizá-los. A repetição da experiência, o número de replicadas pode ajudar a quantificar e avaliar a sua amplitude.

11.5 Prevenção de erros sistemáticos

Podemos realizar os seguintes questionamentos em relação aos erros sistemáticos:

- Quais as causas mais freqüentes?
- Como evitá-los?

Como já foi descrito, a sua consequência é o desvio do valor médio em relação aos valores corretos. Ao contrário dos erros aleatórios, o incremento no número de replicadas não possibilita a sua identificação, uma vez que eles se repercutem sistematicamente. A não ser que o valor correto seja conhecido à partida, não é pois

possível determiná-los apenas por observação dos dados, levando portanto a conclusões errôneas.

Nos últimos anos, tem sido objeto de interesse o nível de metais de transição em amostras biológicas tais como o soro sanguíneo. Obtiveram-se deste modo muitas determinações dos níveis, por exemplo, de cromo no soro. Diferentes resultados apontaram para concentrações na ordem de 1 a 200 ng/ml. Os valores mais baixos foram os mais recentemente obtidos, tornando-se cada vez mais plausível que os anteriores elevados valores foram devidos em parte à contaminação das amostras pelas seringas de aço inoxidável, tampas dos tubos, etc. A determinação da presença de cromo, por espectrometria de absorção atômica, é relativamente segura e embora os erros aleatórios fossem minimizados, os erros sistemáticos foram completamente descuidados, porque a possibilidade de contaminação pelos materiais não foi considerada. Erros nas metodologias são as causas mais frequentes de erros sistemáticos, tais como, lavagem incompleta do precipitado na análise gravimétrica.

Analistas experientes sabem que os monocromadores dos espectrômetros se desajustam gradualmente, levando a erros de vários nanômetros bastante comuns. Até simples instrumentos como materiais de vidro volumétricos, relógios, medidores de pH, podem apresentar erros substanciais. Mesmo os aparelhos mais sofisticados hoje em dia, tidos em conta como perfeitamente infalíveis, não estão em absoluto isentos da possibilidade de erros.

No entanto os erros sistemáticos podem ser ocasionados, não só por erros instrumentais, como também por erros humanos. Os problemas de visão como astigmatismo e daltonismo podem ocasionar erros de leitura. Até mesmo a anotação de valores, onde são freqüentes os arredondamentos dos últimos dígitos a 0 e 5, a preferência de dígitos ímpares a pares, ou vice versa.

Os processos de diminuição dos erros sistemáticos devem ter em linha de conta três principais atitudes:

1. Identificação à partida das possibilidades de ocorrência de erros sistemáticos na metodologia .
2. Um cuidadoso delineamento da experiência em cada passo.
3. Uso de materiais e métodos padronizados
4. Ensaio colaborativos.

A prevalência de erros sistemáticos em trabalhos analíticos está bem patente nos resultados dos ensaios colaborativos. Seria de esperar que o mesmo tipo de análise em laboratórios diferentes conduzisse a resultados similares. Na prática, verifica-se que muitas vezes, esta suposição ocorre mais como exceção do que regra. As diferenças de resultados ultrapassam largamente o que seria de esperar na base de erros aleatórios. A gravidade destas situações, com sérias implicações para os analistas, levou ao desenvolvimento de metodologias em ensaios colaborativos.

O ato de medir é, em essência, um ato de comparar, e essa comparação envolve erros de diversas origens (dos instrumentos, do operador, dos processos de medida etc.). Vamos estudar os erros e suas conseqüências, de modo a expressar os resultados de dados experimentais em termos que sejam compreensíveis a outras pessoas.

Quando se pretende medir o valor de uma grandeza, pode-se realizar apenas uma ou várias medidas repetidas, dependendo das condições experimentais particulares ou ainda da postura adotada frente ao experimento. Em cada caso, deve-se extrair do processo de medida um valor adotado como melhor na representação da grandeza e, ainda um limite de erro dentro do qual deve estar compreendido o valor real.

11.6. Principais fontes de erro na medição

— **VARIAÇÃO DE TEMPERATURA:** A temperatura padrão de referência é de 20°C para todos os países industrializados (normas NBR 6165 do INMETRO e PB 18/56 da ABNT). Se a temperatura muda, a peça se expande ou contrai, afetando o resultado da medição. Quando não é possível controlar a temperatura devem ser feitos cálculos para compensar o erro.

— **FORÇA DE MEDIÇÃO:** Os instrumentos simples de medida envolvem o contato entre o instrumento e a peça, sendo que a força que promove este contato deve ser tal que **não** promova a deformação da peça ou do instrumento.

— **PARALAXE:** Quando os traços de uma escala principal e outra secundária, estiverem em planos diferentes, dependendo da direção de observação, podem-se obter valores de leitura diferentes, que implicam em erro.

Assim, a observação da leitura de um instrumento deve ser feita sempre no melhor posicionamento perpendicular da vista.

— **ESTADO DE CONSERVAÇÃO DO INSTRUMENTO:** Folgas provocadas pelo desgaste no instrumento poderão acarretar erros de consideração. Um programa de verificação e calibração periódica são a garantia de uma medição confiável.

11.7 Flutuações nas medidas

Ao se realizar várias medidas experimentais, de uma grandeza física ou química, temos como objetivo alcançar o seu “valor verdadeiro” ou “valor real”. Mas, atingir este objetivo é praticamente impossível. Pode-se chegar, após uma série de medidas, a um valor que mais se aproxima do valor verdadeiro, ou seja, ao valor mais provável de uma grandeza medida.

O valor verdadeiro seria aquele obtido teoricamente por meio de algum modelo “exato” (que incluísse todos os efeitos físicos) ou então aquele obtido por meio de uma medida experimental “perfeita”. Ambos os casos são situações ideais não alcançadas na prática. Se conhecermos, o valor real da grandeza e o compararmos com o valor medido pode-se definir o que denominamos “Erro”. Erro é a diferença entre o valor medido e o verdadeiro valor da grandeza.

“Erro = valor medido – valor verdadeiro”.

As flutuações que acompanham todas as medidas são as causas que limitam o objetivo de se atingir o valor verdadeiro da grandeza. E estas flutuações ou erros são de origem sistemática e de origem acidentais ou aleatórias.

12. Outliers

Todo o investigador já deparou com um conjunto de dados em que algumas observações se afastam demasiadamente das restantes, parecendo que foram geradas por um mecanismo diferente. As observações que apresentam um grande afastamento das restantes ou são inconsistentes com elas são habitualmente designadas por outliers ou valores aberrantes.

12.1 Identificação de outliers

Um outlier é caracterizado pela sua relação com as restantes observações que fazem parte da amostra. O seu distanciamento em relação a essas observações é fundamental para se fazer a sua caracterização. Estas observações são também designadas por observações "anormais", contaminantes, estranhas, extremas ou aberrantes (Definição proposta por Maria Manuela Caria Figueira).

Antes de decidir o que deverá ser feito às observações *outliers* é conveniente ter conhecimento das causas que levam ao seu aparecimento. Em muitos casos, as razões da sua existência determinam as formas como devem ser tratadas. Assim, as principais causas que levam ao aparecimento de outliers são: erros de medição, erros de execução e variabilidade inerente dos elementos da população.

O estudo de outliers, independentemente da(s) sua(s) causa(s), pode (m) ser (em) realizado (s) em várias fases. A fase inicial é a da identificação das observações que são potencialmente aberrantes. A identificação de *outliers* consiste na detecção, com métodos subjetivos, das observações surpreendentes. A identificação é feita, geralmente, por análise gráfica ou, no caso de o número de dados ser pequeno, por observação direta dos mesmos. São assim identificadas as observações que têm fortes possibilidades de virem a ser designadas por *outliers*.

Na segunda fase, tem-se como objetivo a eliminação da subjetividade inerente à fase anterior. Pretende-se saber se as observações identificadas como *outliers* potenciais o são, efetivamente. São efetuados testes à ou às observações denominadas como "preocupantes". Devem ser escolhidos os testes mais adequados para a situação em estudo. Estes dependem do tipo de *outlier* em causa, do seu número, da sua origem, do conhecimento da distribuição subjacente à população de origem das observações, etc.

As observações suspeitas são testadas quanto à sua discordância. Se for, aceite a hipótese de algumas observações serem *outliers*, elas podem ser designadas como discordantes. Uma observação diz-se discordante se puder considerar-se inconsistente com os valores restantes depois da aplicação de um critério estatístico objetivo. Muitas vezes, o termo discordante é usado como sinónimo de *outlier*.

Na terceira e última fase é necessário decidir o que fazer com as observações discordantes. A maneira mais simples de lidar com essas observações é eliminá-las. Como já foi dito, esta abordagem, apesar de muito utilizada, não é aconselhada. Ela só se justifica no caso de os *outliers* serem devidos a erros cuja correção é inviável. Caso contrário, as observações consideradas como *outliers* devem ser tratadas cuidadosamente, pois contêm informação relevante sobre características subjacentes aos dados e poderão ser decisivas no conhecimento da população à qual pertence à amostra em estudo.

A forma alternativa à eliminação é a acomodação dos outliers (accommodation of outliers). Tenta-se "viver" com os outliers. A acomodação passa pela inclusão na análise de todas as observações, incluindo os possíveis outliers. Independentemente de existirem ou não outliers, opta-se por construir proteção contra eles. Para tal, são efetuadas modificações no modelo básico e/ou nos métodos de análise. Às observações outliers é atribuído um peso reduzido. Ao serem menosprezadas, estas observações não influenciam demasiadamente o valor das estimativas dos parâmetros. Esta abordagem passa pelo menosprezo das observações aberrantes que poderiam, eventualmente, influenciar demasiadamente os resultados.

De entre estas duas abordagens, identificação e acomodação, a primeira parece-nos ser a de maior importância. Os métodos de acomodação requerem grande informação sobre o processo gerador dos outliers e são criados para serem imunes à presença desse tipo de observação. Desta forma, eles tendem a esconder ou menosprezar informação essencial contida nos dados. Pelo contrário, os métodos de identificação pretendem dar a conhecer essa informação e apresentar as características do conjunto de dados em análise.

Valores outliers (discrepantes ou aberrantes) são valores que "fogem" da distribuição dos dados.

12. 2 Definição de Outliers

São observações que apresentam um grande afastamento dos restantes dos resultados ou são inconsistentes. Estas observações são também designadas por observações "*anormais*", *contaminantes*, *estranhas*, *extremas* ou *aberrantes*.

A preocupação com observações outliers é antiga e data das primeiras tentativas de analisar um conjunto de dados. Inicialmente, pensava-se que a melhor forma de lidar com este tipo de observações seriam através da sua eliminação da análise. As opiniões não eram unânimes: uns defendiam a rejeição das observações "inconsistentes com as restantes", enquanto outros afirmavam que as observações nunca deveriam ser rejeitadas, simplesmente por parecerem inconsistentes com os restantes dados e que todas as observações deviam contribuir com igual peso para o resultado final.

Antes de decidir o que deverá ser feito às observações outliers é conveniente ter conhecimento das causas que levam ao seu aparecimento. Em muitos casos, as razões da sua existência determinam as formas como devem ser tratadas. Assim, as principais causas que levam ao aparecimento de outliers são:

- Erros de medição.

- Erros de execução.
- Variabilidade inerente dos elementos da população.

13. Métodos de identificação de *Outliers*

13.1 Gráfico de Box-Plot

O gráfico de Box mostra a mediana, o primeiro e o terceiro quartis de uma distribuição (pontos 50%, 25% e 75% na distribuição acumulada) e a noção de *outlier*: observação que se encontra a mais de um dado múltiplo (1.5 ou 3.0) do intervalo inter-quartilício, acima ou abaixo dos percentis 75% e 25%, respectivamente.

Este gráfico é construído da seguinte forma:

- Calcula-se a mediana, o quartil inferior (Q1) e o quartil superior (Q3);
- Subtrai-se o quartil superior do quartil inferior = (L)
- Os valores que estiverem no intervalo de $Q_3+1,5L$ e Q_3+3L e no intervalo $Q_1-1,5L$ e Q_1-3L , serão considerados *outliers* podendo, portanto ser aceitos na população com alguma suspeita;
- Os valores que forem maiores que Q_3+3L e menores que Q_1-3L devem ser considerados suspeitos de pertencer à população, devendo ser investigada a origem da dispersão. Estes pontos são chamados de extremos.

Exemplo:

280 283 287 288 288 289 289 290 290 290 292 293 293 293

– **Primeiro Quartil: Posição---** $(14+1)/4 = 3.75$

- 25% de 14 = 3,5 $Q_1 = 288$

– **Segundo Quartil (Mediana)~:** Posição = $14/2=7$

- $Q_2 = 289,5 = (289+290)/2$

– **Terceiro Quartil: Posição---** $3(14+1)/4=11,25$

- 75% de 14 = 10,5 $Q_3 = 292$

- Construção

– Calcular o Intervalo Interquartil:

- $[Q_1 - 1,5(Q_3 - Q_1) ; Q_3 + 1,5(Q_3 - Q_1)]$

– **Obs. fora desses intervalos, são consideradas discrepantes**

- Construir um retângulo com base inferior igual a Q_1 , e base superior igual a Q_3 .
- A partir do pto. médio da base superior, levantar haste indo até a maior observação, que esteja dentro do Intervalo Interquartis.
- A partir do pto. médio da base inferior, traçar haste indo até a menor observação, que esteja dentro do Intervalo Interquartis.
- Assinalar com um traço, dentro do retângulo, a mediana.
- Observações que estão fora do intervalo (Discrepante), assinalar com *, acima da haste superior, ou abaixo da inferior, conforme o caso.

13.2 Teste de Dixon

Esse teste é utilizado quando se deseja comparar valores individuais obtidos por um laboratório ou valores médios obtidos por vários laboratórios, ou ainda diferenças entre os dois resultados obtidos por vários laboratórios.

É um teste bilateral, isto é, é testado o valor mínimo e máximo.

Para um conjunto de resultados $Z(h)$, $h = 1, 2, \dots, H$, agrupados em ordem crescente, o teste de Dixon utiliza o seguinte critério:

3 a 7, 8 a 12, 13 a 40. (tabela amarela).

Tabela 4 – Valores críticos para o Teste de Dixon

Critério		H	5%	1%
$D = \frac{Z(2) - Z(1)}{Z(H) - Z(1)}$ e $= \frac{Z(H) - Z(H-1)}{Z(H) - Z(1)}$	3 a 7	3	0,970	0,994
		4	0,829	0,926
		5	0,710	0,821
		6	0,628	0,740
		7	0,569	0,680
<hr/>				
$D = \frac{Z(2) - Z(1)}{Z(H-1) - Z(1)}$ e $\frac{Z(H) - Z(H-1)}{Z(H) - Z(2)}$	8 a 12	8	0,608	0,717
		9	0,564	0,672
		10	0,530	0,635
		11	0,502	0,605
		12	0,479	0,579
<hr/>				
$D = \frac{Z(3) - Z(1)}{Z(H-2) - Z(1)}$ e $\frac{Z(H) - Z(H-2)}{Z(H) - Z(3)}$	>13	13	0,611	0,697
		14	0,586	0,670
		15	0,565	0,647
		16	0,546	0,627
		17	0,529	0,610
		18	0,514	0,594
		19	0,501	0,580
		20	0,489	0,567
		21	0,478	0,555
		22	0,468	0,544
		23	0,459	0,535
		24	0,451	0,526
		25	0,443	0,517
	26	0,436	0,510	
	27	0,429	0,502	
	28	0,423	0,495	
	29	0,417	0,489	
	30	0,412	0,483	

Os valores críticos deste estatístico encontram-se tabelados (tabela 4).

Se $D_{\text{calc}} < D_{\text{tab}} 5\%$ → Valor Aceito

Se $D_{\text{tab}} 1\% > D_{\text{calc}} > D_{\text{tab}} 5\%$ → Valor Suspeito

Se $D_{\text{calc}} > D_{\text{tab}} 1\%$ → Valor Disperso

Exemplo:

Dados:

Labi	n	média
1	3	0,708
2	3	0,680
3	3	0,667
4	3	0,660
5	3	0,690
6	3	0,733
7	3	0,703
8	3	0,677

Passo 1: Ordenar de forma crescente.

Z(1)	Z(2)	Z(3)	Z(4)	Z(5)	Z(6)	Z(7)	Z(8)
0,660	0,667	0,677	0,680	0,690	0,703	0,708	0,733

Passo 2 – Selecionar o menor e o maior valor
 Verificar se 0,660 é um valor disperso
 Verificar se 0,733 é um valor disperso

Passo 3 – Níveis de significância

Valores tabelados D_{tab} 1% ou D_{tab} 5%

D_{tab} 1% = 0,608

D_{tab} 5% = 0,717

Passo 4 – Cálculo da relação

$$\{Z(2)-Z(1)\}/\{Z(H-1)-Z(1)\} = (0,667-0,660)/(0,708-0,660) = 0,146$$

$$\{Z(H)-Z(H-1)\}/\{Z(H)-Z(2)\} = (0,733-0,708)/(0,733-0,667) = 0,379$$

Conclusão:

$D_{calc} < D_{tab}$ 5%, então tanto o menor valor (laboratório 4) quanto o maior valor (laboratório 6) são aceitos.

13.3 Teste de *Grubbs*

O teste de Grubbs é primeiramente realizado verificando a existência de um valor disperso em cada extremidade do conjunto. Se nesta primeira análise, um dos dois valores for considerado disperso, ele é rejeitado, retirado do conjunto e novo teste, verificando a existência de um valor disperso em cada extremidade do conjunto, é realizado e assim sucessivamente, caso contrário, se nesta primeira análise, ambos os valores forem aceitos como não dispersos, o teste é então realizado verificando-se a existência de dois valores dispersos em cada extremidade do conjunto. Se nesta segunda análise os dois resultados de uma das extremidades forem considerados como dispersos, eles devem ser rejeitados, retirados do conjunto e novo teste verificando a existência de dois valores dispersos em cada extremidade do conjunto é realizado e assim sucessivamente.

Assim como nos testes de Cochran e Dixon, não mais do que 15 % dos valores podem ser rejeitados. Se o número de valores rejeitados for maior do que 15%, então, o conjunto de dados que está sendo analisado é considerado não satisfatório e portanto, novo conjunto de valores, deverá ser obtido para análise.

Os valores críticos para Grubbs são tabelados. O teste de Grubbs substitui o teste de Dixon.

Observação de um valor disperso

Dado um conjunto de resultados g_i , para $i=1, 2, \dots, p$, dispostos em ordem crescente, então para determinar se um determinado valor é um disperso, usando o teste de Grubbs, devemos calcular as seguintes estatísticas:

Para testar o maior valor, utilizar $G_p = \frac{(g_p - \bar{g})}{S}$

Para testar o menor valor, utilizar $G_1 = \frac{(\bar{g} - g_1)}{S}$

Onde:

$\bar{g} = \sum \frac{g_i}{p}$ é a média aritmética do conjunto com todos os valores, e

$S = \sqrt{\frac{\sum_{i=1}^p (g_i - \bar{g})^2}{p-1}}$ é o desvio padrão do conjunto com todos os valores

Conclusão:

Um valor é considerado disperso e estranho quando os valores calculados forem maiores que os tabelados a 1% e 5%, respectivamente.

Observação de dois valores dispersos

Dado um conjunto de resultados g_i , para $i=1, 2, \dots, p$, dispostos em ordem crescente, então para determinar se dois (maiores ou menores) valores são considerados como dispersos, usando o teste de Grubbs, devemos calcular as seguintes estatísticas:

Para testar se dois maiores valores podem ser considerados como dispersos, calcular a relação entre as diferenças quadráticas (G):

$$G = \frac{(S_{p-1,p}^2)}{S_0^2}$$

Onde

$$S_0^2 = \sum_{i=1}^p (g_i - \bar{g})^2 \quad \text{é a diferença quadrática no conjunto com todos os valores.}$$

$$S_{p-1,p}^2 = \sum_{i=1}^{p-2} (g_i - \bar{g}_{p-1,p})^2 \quad \text{é a diferença quadrática no conjunto, sem os dois maiores valores.}$$

$$\bar{g}_{p-1,p} = \sum_{i=1}^{p-2} \frac{g_i}{(p-2)} \quad \text{é a média do conjunto, sem os dois maiores valores.}$$

Alternativamente, para os dois menores valores

$$G = \frac{S_{1,2}^2}{S_0^2}$$

Onde

$$S_0^2 = \sum_{i=1}^p (g_i - \bar{g})^2 \quad \text{é a diferença quadrática no conjunto com todos os valores}$$

$$S_{1,2}^2 = \sum_{i=3}^p (g_i - \bar{g}_{1,2})^2 \quad \text{é a diferença quadrática no conjunto, sem os dois menores valores}$$

$$\bar{g}_{1,2} = \sum_{i=3}^p \frac{g_i}{(p-2)} \quad \text{é a média do conjunto, sem dois menores valores}$$

Conclusão:

Dois valores: são considerados dispersos e suspeitos quando os valores calculados forem menores que os tabelados a 1% e 5%, respectivamente.

EX.:

Dados

Lab i	N	média
1	3	0,708
2	3	0,680
3	3	0,667
4	3	0,660
5	3	0,690
6	3	0,733
7	3	0,703
8	3	0,677

Passo 1 – Ordenar de forma crescente

g (1)	g(2)	g(3)	g(4)	g(5)	g(6)	g(7)	g(8)
0,660	0,667	0,677	0,680	0,690	0,703	0,708	0,733

Passo 2 – Níveis de significância

Valores tabelados: $G_{TAB 1\%}$ ou $G_{TAB 5\%}$

Para 8 laboratórios, testando um valor, temos:

$$G_{TAB 1\%} = 2,126$$

$$G_{TAB 5\%} = 2,274$$

Passo 3 – Selecionar o menor ou maior valor.

Verificar 0,660 é um disperso

Verificar 0,733 é um disperso

Passo 4 – Calcular a média e o desvio padrão do conjunto de dados

$$\bar{g} = \sum \frac{g^i}{p} = 0,68975$$

$$S = \sqrt{\frac{\sum_{i=1}^p (g^i - \bar{g})^2}{p-1}} = 0,0240022$$

Passo 5 – Calcular a estatística para o menor valor:

$$G_1 = \frac{\bar{g} - g_1}{S} = \frac{(0,68975 - 0,660)}{0,0240} = 1,2396$$

Conclusão

Como o valor calculado é menor do que os valores tabelados, a 1% como a 5%, ele não é considerado nem disperso, nem suspeito, respectivamente.

OBS.; Como tanto o menor valor quanto o maior valor foram aceitos, seguimos o teste de Grubbs, testando agora dois valores de cada extremidade.

Passo 6 – Níveis de significância

Valores tabelados

$G_{TAB1\%}$ ou $G_{TAB5\%}$

Para 8 laboratórios, testando dois valores, temos:

$G_{TAB1\%} = 0,1101$ ou $G_{TAB5\%} = 0,0563$

Passo 7 – Selecionar os dois menores valores e os dois maiores valores.

Verificar se 0,660 e 0,667 é disperso.

Verificar se 0,708 e 0,733 é disperso.

Passo 8 – Calcular a média do conjunto de dados, sem os dois valores a serem testados.

Portanto, sem os dois menores valores,teremos:

$$\bar{g}_{1,2} = \sum_{i=3}^p \frac{g_i}{(p-2)} = 0,6795$$

Passo 9 – Calcular a soma das diferenças quadráticas, dos dois conjuntos, ou seja, um conjunto com todos os valores e o outro conjunto sem os valores a serem testados.

Conjunto com todos os valores:

$$S_0^2 = \sum_{i=1}^p (g_i - \bar{g})^2 = 0,00404$$

Lab _i	g_i	$(g_i - \bar{g})$	$(g_i - \bar{g})^2$
1	0,708	0,01825	0,000333
2	0,680	-0,00975	0,000095
3	0,667	-0,02275	0,000518
4	0,660	-0,02975	0,000885
5	0,690	0,00025	0,0000000625
6	0,733	0,04325	0,001871
7	0,703	0,01325	0,000176
8	0,677	-0,01275	0,000163

$$\text{Média } \bar{g} = 0,68975 \quad S^2 = 0,00404$$

$$\text{Ou } S_0^2 = (\text{desvio padrão do conjunto})^2 \times (p-1) = (0,024022)^2 \times 7 = 0,00404$$

Conjunto sem os dois menores valores

$$S_{1,2}^2 = \sum_{i=3}^p (g_i - \bar{g}_{1,2})^2 = 0,002178$$

Passo 10 – Calcular o valor de Grubbs

$$G = \frac{S_{1,2}^2}{S_0^2}$$

$$G = \frac{0,002178}{0,004040} = 0,539$$

Conclusão

Como o valor calculado é menor do que os valores tabelados a 1% e 5%, os dois valores testados não são nem dispersos, nem suspeitos.

Passo 11 – Calcular a média do conjunto de dados, sem os dois maiores valores a serem testados.

Portando, sem os dois maiores valores, teremos:

$$\bar{g}_{p-1,p} = \sum_{i=1}^{p-2} \frac{g_i}{(p-2)} = 0,6795$$

Passo 12 – Calcular a soma das diferenças quadráticas, dos dois conjuntos, ou seja, um conjunto como todos os valores a serem testados.

Conjunto com todos os valores

$$S_0^2 = \sum_{i=1}^p (g_i - \bar{g})^2 = 0,00404$$

Conjunto sem os dois maiores valores

$$S_{p-2,p}^2 = \sum_{i=1}^{p-2} (g_i - \bar{g}_{p-1,p})^2 = 0,001206$$

Passo 13 – Calcular o valor de Grubbs

$$G = \frac{S_{p-1,p}^2}{S_0^2} = \frac{0,001206}{0,004040} = 0,29839$$

Conclusão

Como o valor calculado é maior do que os valores tabelados a 1% e 5%, os dois valores testados não são dispersos, nem suspeitos.

13.4 Cochran

É utilizado quando se deseja comparar variâncias, ou seja, verificar se a variância dos resultados obtidos por um laboratório é excessiva em relação à dos demais laboratórios.

É um teste unilateral, isto é, só verifica o maior valor.

Consiste em calcular todas as variâncias envolvidas no experimento e dividir a maior delas pela soma de todas. O valor resultante da divisão é então, comparado com os valores críticos de uma tabela estatística apropriada, que leva em conta o número de variâncias envolvidas e o número de graus de liberdade utilizado nos cálculos. Os valores críticos para o teste de Cochran são tabelados e o critério utilizado foi $C_{calculado} < C_{tabelado}$ (95%) (ISO 5725).

Para um conjunto de p laboratórios, com desvios padrão S_i ($i=1,2,\dots,p$) todos computados com o mesmo número de repetições n , o teste de Cochran é dado por:

$$C_{calculado} = \frac{S^2_{\max}}{\sum_{i=1}^p S^2_i}$$

Onde:

S^2 = estimativa de variância

S^2_{\max} = maior valor

S^2_{\max} = maior valor encontrado como estimativa da variância, no conjunto.

p = número de laboratórios

No caso de duas repetições, podem ser utilizadas as amplitudes R_i ao invés dos desvios padrão S_i , e o teste de Cochran é dado por:

$$C_{calculado} = \frac{R^2_{\max}}{\sum_{i=1}^p R^2_i}$$

Onde,

R^2 = amplitude quadrática

R^2_{\max} = maior valor encontrado como amplitude quadrática, no conjunto

p = número de laboratórios

Os valores críticos para o teste de Cochran são tabelados

Se $C_{\text{calc}} < C_{\text{tab}} 5\%$ → Valor Aceito

Se $C_{\text{tab}} 1\% > C_{\text{calc}} > C_{\text{tab}} 5\%$ → Valor Suspeito

Se $C_{\text{calc}} < C_{\text{tab}} 1\%$ → Valor Disperso

Exemplo:

Lab _i	n	S _i	(S _i) ²
1	3	0,005	0,000025
2	3	0,010	0,000100
3	3	0,021	0,000441
4	3	0,010	0,000100
5	3	0,019	0,000361
6	3	0,006	0,000036
7	3	0,012	0,000144
8	3	0,025	0,000625

Passo 1 – Calcular a soma das variâncias $\Sigma(S_i)^2$

$$\Sigma(S_i)^2 = 0,001832$$

Passo 2 – Relacionar a maior variância encontrada (0,000625) com a soma.

$$C_{\text{calculado}} = \frac{S^2_{\max}}{\sum_{i=1}^p S^2_i} = \frac{0,000625}{0,001832} = 0,341$$

Passo 3 – Comparar o valor calculado ($C_{calculado} = 0,341$) com o valor tabelado (C_{Tab} 1% ou C_{Tab} 5%)

Para $p=8$ e $n = 3$, temos:

$$C_{Tab} \text{ 1\%} = 0,516$$

$$C_{Tab} \text{ 5\%} = 0,615$$

Como $C_{calculado} < C_{Tab} \text{ 5\%}$, então o valor da maior variância, laboratório 8 é aceita.

13.5 Z-score

Calcular o z-escore, a partir dos valores de z padronizados dos dados.

Os valores os valores de z que estejam compreendidos entre 2 e 3 evidenciam que o resultado fornecido é **questionável**. Para os valores de z menores que 2, o resultado é considerado **satisfatório** e para z maiores que 3, resultado **insatisfatório**.

O z-escore foi calculado para cada um dos resultados obtidos pelos laboratórios participantes através da equação:

$$Z = \frac{(x - X)}{S}$$

Onde:

z = z-escore,

x = valor obtido pelo laboratório,

X = melhor estimativa do valor real,

S = desvio padrão.

O valor de z corresponde ao número de desvios padrão entre o resultado obtido pelo laboratório e o resultado médio estimado.

13.6 O critério de Chauvenet

O critério de Chauvenet estabelece que uma determinada leitura pode ser rejeitada se a probabilidade de obter um desvio particular em relação à média for menor que $\frac{1}{2} N$, onde N é o número máximo das medições realizadas.

Sejam n medições de uma grandeza, onde n é grande, de modo que os resultados seguem uma distribuição Gaussiana. Como já vimos, esta distribuição fornece a probabilidade de que um dado valor medido esteja desviado de certo valor da média esperada. Não se pode esperar probabilidade menor que $1/n$. Assim, se a probabilidade para o desvio de certo valor medido é menor que $1/n$, pode-se qualificá-lo para ser eliminado do conjunto de dados. Na realidade, testes mais restritivos são aplicados para eventualmente eliminar valores medidos. O critério de *Chauvenet* é um deles: especifica que um valor medido pode ser rejeitado se a probabilidade m de obter o desvio em relação à média é menor que $1/2n$. A tabela 1 lista valores da razão do desvio em relação ao desvio padrão para vários valores de n conforme este critério.

Tabela 1 – Critério de Chauvenet para rejeição de valor medido

Número de leituras, n	Razão entre o máximo desvio d_{\max} / σ
3	1,38
4	1,54
5	1,65
6	1,73
7	1,80
10	1,96
15	2,13
25	2,33
50	2,57
100	2,81
300	3,14

Para aplicar o critério de Chauvenet para eliminar valores duvidosos, em primeiro lugar calcula-se o desvio médio e o desvio padrão do conjunto de dados medidos. O desvio de cada um dos pontos é comparado com o desvio padrão, conforme os valores da Tabela 1, para assim se eliminar os pontos duvidosos. Para a apresentação dos dados finais, um novo valor médio e um novo desvio padrão são calculados, sem incluir os pontos eliminados. O critério poderia até ser repetido mais uma vez, para a eliminação de mais pontos, entretanto o procedimento não é recomendável, basta uma vez. Se diversos pontos extrapolarem o critério-limite estabelecido, é provável que o sistema de instrumentação seja inadequado, ou que o processo sendo medido seja extremamente variável.

Como já vimos, erros em medições podem ser caracterizados por uma função normal, e o desvio padrão σ pode ser reduzido aumentando-se o número de medições. Na maioria dos casos, o custo experimental coloca um limite sobre o número de medições. Deve-se considerar, também, que o erro sistemático presente na medição não é uma variável aleatória, e os procedimentos estatísticos podem ser efetivamente usados para “condicionar” dados experimentais com valores medidos erroneamente.

O critério de *Chauvenet*, então, fornece uma base consistente para tomar a tomada de decisão de excluir ou não um dado de um conjunto de valores medidos. Para sua aplicação, basta calcular a razão do desvio individual para o desvio padrão, ΔR , e comparar com um valor de referência, ΔR_0 . ΔR_0 depende do número de medidas, enquanto que ΔR é definida como: $\Delta R = (x_i - \bar{x}) / \sigma$. O valor medido é rejeitado quando $\Delta R > \Delta R_0$.

A seguir é apresentado um exemplo da aplicação do critério de *Chauvenet*. Considere os resultados na medição do comprimento de um corpo apresentados na tabela abaixo. Verifique se algum ponto pode ser rejeitado.

Leitura	Comprimento (cm)
1	49,36
2	50,16
3	48,98
4	49,24
5	49,26
6	50,56
7	49,58
8	49,89
9	49,33
10	49,39

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i = 49,53$$

$$\sigma = \left[\frac{1}{n-1} \sum (x_i - \bar{x})^2 \right]^{1/2} = 0,495$$

Leitura	$d = x_i - \bar{x}$	(d_i / σ)
1	-0,171	0,345
2	0,589	1,190
3	-0,551	1,113
4	-0,291	0,588
5	-0,271	0,547
6	1,029	2,079
7	-0,351	0,709
8	0,359	0,705
9	-0,201	0,405
10	-0,141	0,285

Verifica-se que o desvio da leitura número 6 dividido pelo desvio padrão é o único que ultrapassa o valor de 1,96 estabelecido como o máximo desvio para uma amostra de 10 medidas, como pode ser verificado na Tabela 1. Assim, este ponto pode ser rejeitado. Recalculando-se a média e o desvio padrão obtém-se:

$$\bar{x} = 42,49$$

$$\sigma = 0,359$$

14. Referências Bibliográficas

Andrews, D. F. e Pregibon, D.(1978); "Finding outliers that matter" J.R.Statist. Soc. B, **40**, 85-93.

Anscombe, F.J.(1960); "Rejection of outliers".Technometrics, **2**, 123-147.

Atkinson, A. C. (1981); "Two graphical displays for outlying and influential observations in regression". Biometrika, **68**, 13-20.

Barnett, V. e Lewis, T. (1978); "Outliers in statistical data", John Wiley & Sons, New York.

Barnett, V. e Lewis, T. (1994); "Outliers in statistical data", John Wiley & Sons, New York.

Beckman, R. J. e Cook, R. D.(1983);"Outliers".Technometrics, **25**,119-163.

Brant, R. (1990); "Comparing classical and resistant outlier rules". Journal of the American Statistical Association, **85**, 1083-1090.

Campbell, N. A.(1978); "The influence function as an aid in outlier detection in discriminant analysis".

<http://www.educacional.com.br/upload/blogSite/5053/5053974/16037/apostila1a.pdf>

<http://leq.ufpr.br/~shimakur/CE003/node30.html>

<http://www.inf.ufsc.br/~marcelo/testes2.html>

<http://www.ipv.pt/millennium/arq12.htm>

http://www.famat.ufu.br/espe/eea3/aulas_arquivos/denise/homogeneidade.pdf

Figueira, M.M.C, Identificação de Outliers, MILLENIUM nº12 – Outubro de 1998.

Morel P., Validação e Incerteza na Medição Analítica, Ministério da Saúde, ANVISA / GGLAS

http://www.anvisa.gov.br/reblas/cursos_gglas/validacao_incertezas_pierre_2.pdf

Andrade, E.A. e Robin J., Seminário - Mineração de Exceções

www.cin.ufpe.br/~compint/aulas-IAS/kdd-012/Outliers.ppt

Miler, J.C. e Miler, J.N. (1988), Statistics for Analytical Chemistry – second edition, John Wiley & Sons, New York, Chichester, Brisbane, Toronto.

P.R.G. Couto, “Estimativa da incerteza da massa específica da gasolina pelo ISO GUM 95 e Método de Monte Carlo e seu impacto na transferência de custódia”, *Dissertação de Mestrado*, Rio de Janeiro, 2006.

ISO 4787/1984 (E) Laboratory Glassware- Volumetric Glassware- Methods for use and testing of capacity.

B, Barros Neto, I.S. Scarmínio, R. E. BRUNS, “Como fazer experimentos: pesquisa e desenvolvimento na ciência e na indústria”, , Editora da Unicamp, Campinas, 2002.

15. Sumário de Revisões

<i>Edição</i>	<i>Item (s) Revisado (s)</i>	<i>Responsável pela Revisão</i>	<i>Data da Revisão</i>
<i>01</i>	<i>Emissão Inicial</i>	<i>Olivia Woyames Pinto</i>	<i>16/09/2009</i>