UNIVERSIDADE FEDERAL DO CARIRI

CENTRO DE CIÊNCIAS E TECNOLOGIA

UNIDADE CURRICULAR: ESTATÍSTICA

Disciplina: Probabilidade e Estatística **Docente:** Rosilda Benício

(r*osilda.benicio@ufca.edu.br*)

UNIDADE III: NOÇÕES DE INFERÊNCIA

Sumário

UNIDADE III – NOÇÕES DE INFERÊNCIA

Nas Unidades anteriores, vimos como construir modelos probabilísticos para descrever alguns fenômenos. Nesta parte, iremos estudar um ramo muito importante da Estatística conhecido como Inferência Estatística, ou seja, como fazer afirmações sobre características de uma população, baseando-se em resultados de uma amostra.

O **objetivo da Inferência Estatística** é produzir afirmações sobre uma dada característica da população, na qual estamos interessados, a partir de informações colhidas de uma parte dessa população (amostra). Esta característica pode ser representada por uma variável aleatória. Se tivéssemos informação completa sobre a função de probabilidade, no caso discreto, ou sobre a função densidade de probabilidade, no caso contínuo, da variável em questão, não teríamos necessidade de colher uma amostra. Toda afirmação desejada seria obtida através da distribuição da variável, usando-se as propriedades estudadas anteriormente. Mas isso raramente acontece. Ou não temos qualquer informação a respeito da variável, ou ela é apenas parcial.

Conceitos básicos de Inferência e Amostragem

População: é o conjunto formado por indivíduos ou objetos que tem pelo menos uma variável comum e observável. Por exemplo, população dos alunos do primeiro período de uma faculdade; população de alturas em *cm* das pessoas de determinado bairro; população de peças fabricadas numa linha de produção, e assim por diante.

Definiremos como tamanho de uma população finita o número de elementos que a compõem. Usaremos N para designar esse número.

Amostra: fixada uma população, qualquer subconjunto formado exclusivamente por seus elementos é denominado *amostra* desta população. Usaremos n para indicar o número de elementos da amostra, o seu tamanho.

Amostragem: é o processo de seleção de uma amostra, que possibilita o estudo das características da população.

Parâmetro: é a medida usada para descrever uma característica numérica populacional. Genericamente representaremos por θ. A média (μ), a variância (σ^2), a proporção (p) e o coeficiente de correlação (ρ) são alguns exemplos de parâmetros populacionais.

Estimador: também denominado *estatística* de um parâmetro populacional; é uma característica numérica determinada na amostra, uma função de seus elementos. Genericamente representaremos por $\hat{\theta}$. A média amostral ($\bar{x}\ ou\ \hat{\mu}$), a variância amostral ($s^2 ou\ \hat{\sigma}^2$), a proporção amostral ($\hat{p}$) e o coeficiente de correlação amostral ($r\ ou\ \hat{\rho}$) são exemplos de estimadores.

Estimativa: é o valor numérico determinado pelo estimador, genericamente representaremos por $\hat{\theta}_0$.

Erro amostral: é o erro que ocorre justamente pelo uso da amostra.

Logo, o erro amostral, que designaremos por $\varepsilon = \hat{\theta} - \theta$.

O valor de $\hat{\theta}$ varia em cada uma das N^n amostras de tamanho n, tiradas da população, como segue:

amostra 1 → $\hat{\theta}_1$; amostra 2 → $\hat{\theta}_2$; ... ; amostra $k \rightarrow \hat{\theta}_k$.

Logo, $\hat{\theta}$ é uma variável aleatória, e como tal, podemos determinar a $E(\hat{\theta}),\ Var(\hat{\theta})$, isto é, a esperança matemática de $\hat{\theta}$ e sua variância.

Podemos desmembrar o erro amostral em duas partes:

$$\varepsilon = \hat{\theta} - \theta = [\hat{\theta} - E(\hat{\theta})] + [E(\hat{\theta}) - \theta],$$

em que $[\hat{\theta} - E(\hat{\theta})]$ é a parte casual e $[E(\hat{\theta}) - \theta]$ é o viés ou desvio.

O viés pode aparecer na seleção da amostra, na coleta dos dados ou na estimação dos parâmetros.

Viés de seleção: A amostragem pode ser probabilística e não probabilística. Amostragem probabilística é o processo de seleção de uma amostra no qual cada unidade amostral da população tem probabilidade diferente de zero e conhecida de pertencer a amostra.

Na amostragem não probabilística, a probabilidade de seleção é desconhecida para alguns ou todos os elementos da população, podendo alguns destes elementos ter probabilidade nula de pertencer à amostra, como em amostras intencionais, a esmo ou de voluntários.

O melhor modo de evitar o viés de seleção é o uso do sorteio, seja ele manual ou por meio de uma tabela de números aleatórios, ou então pela geração de números aleatórios por computador.

A amostragem probabilística é isenta de viés de seleção.

Viés na coleta de dados: Esse tipo de vício pode ocorrer principalmente quando se substitui a unidade de amostragem ou quando há falta de respostas.

Viés de estimação: Esse tipo de vício pode ser controlado fazendo-se amostragens probabilísticas.

Censo ou amostragem?

A palavra censo refere-se à pesquisa de todos os elementos da população. Geralmente, realizamos um censo quando: a população é pequena; as variáveis são fáceis de serem medidas ou observadas; ou, necessitamos de resultados exatos.

Grande parte das pesquisas científicas é feita por amostragem. A amostragem é particularmente interessante quando: a população é grande ou infinita; as observações ou mensurações tem alto custo; as medidas exigem testes destrutivos; ou, há necessidade de rapidez.

Em geral, o uso de amostragem leva à redução de custos e tempo. Mas a amostragem precisa ser feita com critérios, pois pretendemos ter amostras que permitam, a partir de uma análise estatística apropriada, obter conclusões satisfatórias.

Tipos de amostragem

Amostragem não probabilística - o pesquisador não sabe qual é a probabilidade de que um elemento da população tem de pertencer à amostra. Portanto, os resultados da amostra não podem ser estatisticamente generalizados para a população, porque não se pode estimar o erro amostral. Destacam-se os seguintes tipos de amostragem não probabilística: a esmo, intencional (por julgamento), por cotas, bola de neve

Amostragem probabilística, também é chamada de aleatória ou casual. A sua importância decorre do fato de que apenas os resultados provenientes de uma amostra probabilística podem ser generalizados estatisticamente para a população da pesquisa. Será de nosso interesse estudar este tipo de amostragem, como segue.

- Amostragem Casual Simples

Precisa-se ter uma lista completa dos elementos da população, isso por vezes é uma dificuldade. No entanto, as técnicas estatísticas, em geral, supõem esse tipo de amostragem. Além disso, deve haver homogeneidade na população em relação a variável de interesse.

Consideremos uma população $X_1, X_2, \ldots, X_N$ com elemento genérico X_j, com $1 \leq j \leq N$ e a amostra x_1, $x_2, \ldots, x_n$ com elemento genérico x_i, $1 \leq i \leq n$.

Definição: uma amostra se diz casual simples quando $P(X_j = x_i) = 1/N$, quaisquer que sejam $i = 1, 2, \ldots, n$ e $j = 1, 2, \ldots, N$.

Exemplo: Consideremos a população formada pelos números 1,2,3,4,5,6,7,8 e 9. A média da população é $\mu = 5$. Retiramos dessa população amostras de tamanho 3.

a) Com reposição:

- amostra com os menores valores: 1,1,1 → $\bar{x} = 1$ → $\varepsilon = 1 - 5 = -4$ (lembrando: $\varepsilon = \bar{x} - \mu$);
- amostra com os maiores valores: 9,9,9 → $\bar{x} = 9$ → $\varepsilon = 9 - 5 = 4$, portanto $|\varepsilon| = |\bar{x} - \mu| \leq 4$.

b) Sem reposição:

- amostra com os menores valores: 1,2,3 → $\bar{x} = 2$ → $\varepsilon = 2 - 5 = -3$;
- amostra com os maiores valores: 7,8,9 → $\bar{x} = 8$ → $\varepsilon = 8 - 5 = 3$, portanto $|\varepsilon| \leq 3$.

Concluímos que o erro amostral é menor quando se usa amostragem sem reposição. A variabilidade é reduzida, uma vez que os elementos extremos não poderão ser retirados mais do que uma vez.

Exemplo: Pretende-se realizar uma pesquisa de opinião sobre a qualidade de um curso universitário, que tem cerca de 1000 alunos, perguntando aspectos relativos ao encadeamento das disciplinas no currículo. A amostragem casual simples é apropriada para selecionar os respondentes?

A amostragem casual simples pode ser utilizada quando houver homogeneidade na população em relação à variável de interesse. No presente caso o interesse é na opinião dos alunos sobre o currículo. É razoável imaginar que um aluno do oitavo período tenha um conhecimento diferente do currículo, se comparado a um aluno do segundo período, acarretando em diferentes opiniões, que deveriam ser registradas. Se for utilizada uma amostragem casual simples, totalmente por acaso, apenas alunos dos períodos mais adiantados, ou dos períodos iniciais, comprometendo o resultado da pesquisa. Assim, como não há homogeneidade na população acerca da variável de interesse, a amostragem casual simples não é apropriada para este caso.

➢ Amostragem Sistemática

Quando a lista completa dos elementos da população for muito grande a utilização da amostragem aleatória simples pode ser um processo moroso. Utiliza-se então uma variação, a amostragem sistemática, que também supõe que a população é homogênea em relação à variável de interesse, mas que consiste em retirar elementos da população a intervalos regulares, até compor o total da amostra.

Consideremos uma população de tamanho N e dela tiramos uma amostra de tamanho n.

Definimos $s = \left[\frac{N}{n}\right]$: fator de sistematização.

Sorteamos (isso precisa ser feito para garantir que todos os elementos da população terão chance de pertencer à amostra) um número entre 1 e s. Seja m esse número:

- o primeiro elemento da amostra é o de número m;
- o segundo elemento da amostra é o de número $s + m$;
- o terceiro elemento da amostra é o de número $2s + m$;
- ...
- o n-ésimo elemento da amostra é o de número $(n-1)s + m$.

Para este tipo de amostragem é necessário que a população esteja ordenada, por exemplo, em nomes de uma lista telefônica ou em números das casas de uma rua.

Obs.: se s for fracionário, deve-se aumentar n até tornar o resultado inteiro; se N for um número primo, excluem-se *por sorteio* alguns elementos da população para tornar s inteiro.

Exemplo: De uma população de $N = 1000$ elementos ordenados, retirar uma amostra sistemática de tamanho 100.

$s = \left[\frac{1000}{100}\right]$=10

Seja $1 \leq m \leq 10$. Suponhamos que $m = 7$. Logo, temos:

1º elemento da amostra ....7°
2º elemento da amostra ....17°
3º elemento da amostra ....27°
...
100º elemento da amostra ......997°

➢ Amostragem por Estratificação

É bastante comum que a população alvo de uma pesquisa seja heterogênea em relação à variável de interesse, não podendo ser aplicada imediatamente a amostragem casual simples (acs). Assim, determina-se estratos segundo alguma característica relevante, e a seguir, nos estratos, pode-se aplicar a acs ou a amostragem sistemática. A definição dos estratos requer um conhecimento prévio dos elementos da população à cerca das variáveis de interesse.

Vamos considerar o mesmo exemplo abordado anteriormente, devemos usar uma variável "critério" para separar a população em estratos.

No exemplo, o critério de estratificação será:
E_1: grupo formado pelos três menores valores. $E_1 = 1, 2, 3$;
E_2: grupo formado pelos três valores centrais. $E_2 = 4, 5, 6$;
E_3: grupo formado pelos três maiores valores. $E_3 = 7, 8, 9$.

Selecionemos um elemento de cada estrato para formarmos as amostras de tamanho 3:

- amostra com menores elementos: 1,4,7 → $\bar{x} = 4$ → $\varepsilon = -1$;
- amostra com maiores elementos: 3,6,9 → $\bar{x} = 6$ → $\varepsilon = 1$, portanto $|\varepsilon| \leq 1$.

A amostragem por estratificação tem as seguintes características:

- dentro de cada estrato há uma grande homogeneidade, ou então uma pequena variabilidade;
- entre os estratos há uma grande heterogeneidade, ou então uma grande variabilidade.

Exemplos:

1. Se para estudar a dureza de certo aço temos corpos de prova de dois fornecedores, então a população dos corpos de prova pode ser dividida em dois estratos.

2. Dada a população de 50.000 operários da construção civil, formar uma amostra de 5% de operários para estimar seu salário médio.

Usando a variável critério “cargo” para estratificar essa população, e considerando amostras de 5% de cada estrato obtido, chegamos ao seguinte quadro:

Cargos	**População**	**Amostra**
Chefes de seção	5.000	250
Operários especializados	15.000	750
Operários não especializados	30.000	1.500
Total	50.000	2.500

Neste exemplo fizemos uma partilha proporcional para cada cargo.

- Amostragem estratificada proporcional: a proporcionalidade do tamanho de cada estrato da população é mantida na amostra. Por exemplo, se um estrato abrange 20% da população, ele também deve abranger 20% da amostra.
- Amostragem estratificada uniforme: selecionamos o mesmo número de elementos em cada estrato. É o processo usual quando se deseja comparar os diversos estratos.

➢ Amostragem por Conglomerados

Teoricamente a amostragem estratificada proporcional apresenta os melhores resultados possíveis. A grande dificuldade em implementá-la deve-se ao grau de conhecimento necessário sobre a população, que geralmente não existe ou é impraticável de ser obtido. Uma forma alternativa de amostragem consiste no uso de conglomerados.

Os conglomerados também são grupos mutuamente exclusivos de elementos da população, mas são definidos de forma mais arbitrária do que os estratos: é bastante comum definir os conglomerados geograficamente. Por exemplo, os bairros de uma cidade, que constituiriam conglomerados de domicílios.

O procedimento para a amostragem por conglomerados: divide-se a população em conglomerados; sorteiam-se, aleatoriamente, os conglomerados; pesquisam-se todos os elementos dos conglomerados sorteados, ou sorteiam-se elementos deles. Observa-se que a unidade de sorteio muda, passamos de elemento para grupo. E, todos os elementos dos conglomerados podem ser observados.

Consideremos conglomerados os grupos de elementos com as seguintes características:

- dentro de cada conglomerado há uma grande heterogeneidade, ou então uma grande variabilidade;
- entre os conglomerados há uma pequena variabilidade, ou então uma grande homogeneidade.

Por exemplo, se estivermos interessados no salário médio dos operários da construção civil, como no exemplo anterior, podemos selecionar uma construtora e, dentro dela, estudar os salários.

Exemplo: Pesquisa Nacional por Amostra de Domicílios (PNAD) do IBGE. Coleta informações demográficas e sócio-econômicas sobre a população brasileira. Utiliza amostragem por conglomerados.

Primeiro estágio: amostras de municípios (conglomerados) para cada uma das regiões geográficas do Brasil;

Segundo estágio: setores censitários sorteados em cada município (conglomerado sorteado);
Terceiro estágio: domicílios sorteados em cada setor censitário.

A utilização de amostragem por conglomerados permite uma redução substancial nos custos de obtenção da amostra, sem comprometer em demasia a precisão, e em alguns casos é a única alternativa possível.

Distribuição amostral dos estimadores

O problema da Inferência Estatística é fazer uma afirmação sobre parâmetros da população através da amostra. Digamos que nossa afirmação deva ser feita sobre um parâmetro θ da população (média, variância ou qualquer outra medida). Decidimos que usaremos uma amostra casual simples (elementos são retirados ao acaso da população, assim todo elemento da população tem igual probabilidade de ser escolhido para a amostra), com reposição, de n elementos sorteados dessa população. Também decidimos que a nossa decisão será baseada na estatística T, que será uma função da amostra $(X_1, X_2, \ldots, X_n)$ ou seja, $T = f(X_1, X_2, \ldots, X_n)$. Colhida uma amostra, teremos observado um particular valor de T, digamos t_0, e baseados nesse valor é que faremos a afirmação sobre θ, o parâmetro populacional.

A validade da nossa resposta seria melhor compreendida se soubéssemos o que acontece com a estatística T, quando retiramos todas as amostras de uma população conhecida segundo o plano amostral adotado. Isto é, qual a distribuição de T quando $(X_1, X_2, \ldots, X_n)$ assume todos os valores possíveis. Esta distribuição é chamada de distribuição amostral da estatística T e desempenha papel fundamental na teoria de Inferência Estatística.

Exemplo: Numa urna tem 5 tiras de papel, numeradas 1, 3, 5, 5 e 7. Uma tira de papel é sorteada e recolocada na urna; então uma segunda tira é sorteada. Seja X_1 e X_2, respectivamente, o primeiro e o segundo número sorteados. Selecionamos todas as amostras de tamanho 2, com reposição. A distribuição conjunta da variável bidimensional (X_1, X_2) é dada na Tabela 1.

Tabela 1 - Distribuição das probabilidades das possíveis amostras de tamanho 2 que podem ser selecionadas com reposição da população {1,3,5,5,7}.

X_2 \ X_1	1	3	5	7	**Total**
1	1/25	1/25	2/25	1/25	1/5
3	1/25	1/25	2/25	1/25	1/5
5	2/25	2/25	4/25	2/25	2/5
7	1/25	1/25	2/25	1/25	1/5
Total	1/5	1/5	2/5	1/5	1

Vejamos qual a distribuição amostral da estatística $\bar{X} = \frac{X_1+X_2}{2}$. Esta distribuição é obtida através da Tabela 1. Por exemplo, quando a amostra selecionada é o par (1,1), corresponderá à média 1; então, temos $P(\bar{X} = 1) = 1/25$. Obteremos média igual a 3 quando ocorrer o evento $A = \{(1,5), (3,3), (5,1)\}$, logo, $P(\bar{X} = 3) = P(A) = 2/25 + 1/25 + 2/25 = 1/5$.

Procedendo de modo análogo, obtemos a distribuição amostral da estatística $\bar{X}$ (Tabela 2).

Tabela 2 - Distribuição amostral da estatística $\bar{X}$.

$\bar{x}$	1	2	3	4	5	6	7	**Total**
$P(\bar{X} = \bar{x})$	1/25	2/25	5/25	6/25	6/25	4/25	1/25	1

- **Distribuição amostral da média**

Vamos estudar agora a distribuição amostral da estatística $\bar{X}$, a média da amostra. Consideremos uma população identificada pela variável X, cujos parâmetros populacionais são: a média $\mu = E(X)$ e a variância $\sigma^2 = Var(X)$ são supostamente conhecidos. Vamos retirar todas as possíveis amostras casuais simples de tamanho n dessa população, e para cada uma calcular a média $\bar{X}$. Em seguida, vamos construir a distribuição amostral e estudar suas propriedades. Ilustremos com os dados do Ex. 1:

Exemplo: A população {1,3,5,5,7} tem média μ = 4,2 e σ^2 = 4,16.

$E(X_1) = E(X_2) = 1.\frac{1}{5} + 3.\frac{1}{5} + 5.\frac{2}{5} + 7.\frac{1}{5} = \frac{1}{5} + \frac{3}{5} + \frac{10}{5} + \frac{7}{5} = \frac{21}{5} = 4{,}2.$

$Var(X_1) = Var(X_2) = 4{,}16.$

E a distribuição amostral de $\bar{X}$ para $n = 2$ está na Tabela 2. Baseando-nos naqueles dados, podemos verificar que

$E(\bar{X}) = \frac{1}{25} + \frac{4}{25} + \frac{15}{25} + \frac{24}{25} + \frac{30}{25} + \frac{24}{25} + \frac{7}{25} = \frac{105}{25} = 4{,}2.$

$Var(\bar{X}) = 2{,}08.$

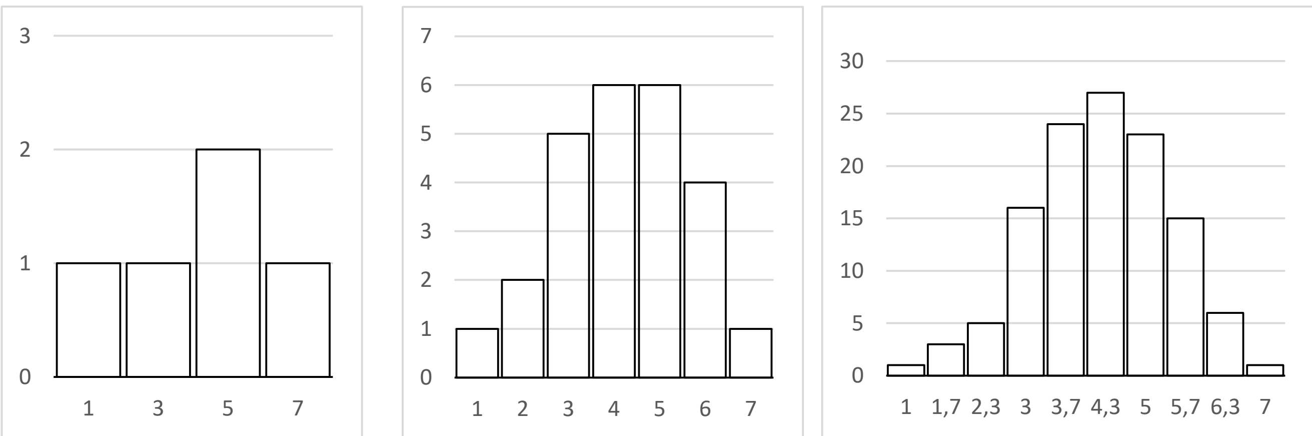
Observe que a média das médias amostrais é igual a média populacional. E a variância de $\bar{X}$ é igual a $Var(X)/2$, quando $n = 2$. Isso sempre acontece.

Teorema 1: Seja X uma v.a. com média μ e variância σ^2, e seja $(X_1, X_2, \ldots, X_n)$ uma amostra casual simples.

Então, se $\bar{X} = \frac{X_1 + X_2 + \cdots + X_n}{n}$, temos $E(\bar{X}) = \mu$ e $Var(\bar{X}) = \sigma^2/n$.

Já determinamos a média e a variância da distribuição de $\bar{X}$. Para derivar as demais propriedades de $\bar{X}$, bastaria agora determinar qual a forma (modelo probabilístico) da curva referente à distribuição (histograma) de $\bar{X}$. A derivação dessa propriedade exige recursos matemáticos que estão fora dos objetivos dessa disciplina. Assim, iremos mostrar empiricamente o que acontece com a distribuição de $\bar{X}$.

Exemplo: (continuação): Na população {1,3,5,5,7}, construímos os “histogramas” das distribuições de $\bar{X}$ para n = 1, 2 e 3.



$Var(\bar{X}) = 4{,}16$ (a) $Var(\bar{X}) = 2{,}08$ (b) $Var(\bar{X}) = 1{,}39$ (c)

(a) Distribuição de $\bar{X}$ para $n = 1$; a distribuição coincide com a distribuição de X.
(b) Distribuição de $\bar{X}$ para $n = 2$ (Tabela 2)
(c) Distribuição de $\bar{X}$ para $n = 3$.

Observe que, n vai aumentando, o histograma vai ficando mais serrilhado e tende a concentrar-se cada vez mais em torno de $E(\bar{X})$. Os casos extremos passam a ter pouca probabilidade de ocorrência. Quando n for suficientemente grande, o histograma alisado aproxima-se da distribuição normal. Este resultado fundamental na teoria de Inferência Estatística é conhecido como ***Teorema do Limite Central***.

Teorema 2: Para amostras casuais simples $(X_1, X_2, \ldots, X_n)$ retiradas de uma população com média μ e variância σ^2, a distribuição amostral da média $\bar{X}$ aproxima-se de uma distribuição normal com média μ e variância σ^2/n, quando n tende ao infinito.

Ou seja, a distribuição da variável $\bar{X}$ por amostragem casual simples será sempre normal da população X e variância n vezes menor. Isso significa que, quanto maior o tamanho da amostra, menor será a variância de $\bar{X}$, ou o estimador será mais preciso à medida que o tamanho da amostra aumentar.

Exemplo: Considere uma máquina empacotadora regulada para encher pacotes segundo uma distribuição normal com média 500 g e desvio padrão de 10 g, isto é $X \sim N(500{,}100)$. Colhendo uma amostra de 100 pacotes e pesando-os, sabemos pelo Teorema do Limite Central que $\bar{X}$ terá distribuição normal, com média 500 e variância 100/100 = 1 g^2. Assim, se a máquina estiver regulada, a probabilidade de encontrarmos em média 100 pacotes diferindo do peso de 500 g com menos de 2 g, será

$P(|\bar{X} - 500| < 2) = P(498 < \bar{X} < 502) = P(-2 < Z < 2) = 2P(0 < Z < 2) \approx 95\%$.

Ou seja, dificilmente 100 pacotes estarão, em média, fora do intervalo]498; 502[.

Corolário 1: Se $(X_1, X_2, \ldots, X_n)$ é amostra casual simples da população X com média μ e variância σ^2, e $\bar{X} = (X_1 + X_2 + \cdots + X_n)/n$, então,

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

Corolário 2: A distribuição de $(\bar{X} - \mu)$ aproxima-se de uma distribuição normal com média 0 e variância σ^2/n, isto é, $(\bar{X} - \mu) \sim N(0, \sigma^2/n)$. Pois, o Teorema do Limite Central afirma que $\bar{X}$ aproxima-se de uma normal quando n tende para infinito.

Há uma observação importante a ser feita: se a população for finita e de tamanho N conhecido, e se a amostra de tamanho n dela retirada for sem reposição, então $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$, o fator de correção $\left(\frac{N-n}{N-1}\right)$ deve ser usado. Porém, se $n/N \leq 0{,}05$, o fator de correção deve ser ignorado.

Exemplo: Temos uma população de 5000 alunos de uma faculdade. Sabemos que a altura média dos alunos é 175 cm e o desvio padrão, 5 cm. Retiramos uma amostra sem reposição de tamanho $n = 100$.

$X \sim N(175{,}25)$. Então $\mu_{\bar{X}} = E(\bar{X}) = 175$ e $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}} = \frac{5}{\sqrt{100}}\sqrt{\frac{5000-100}{5000-1}} = 0{,}495 \simeq 0{,}50$.

$\frac{n}{N} = \frac{100}{5000} = 0{,}02 < 0{,}05$; Ignorando o fator de correção, temos $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{5}{\sqrt{100}} = 0{,}50$

Logo, a média das médias amostrais é $175\ cm$ e o desvio padrão da média amostral é $0{,}50\ cm$.

Exemplo: A idade média de 250 empregados de uma empresa é de 39,5 anos, com desvio padrão de 9,3 anos. Se uma amostra aleatória de 35 empregados for tomada, qual a probabilidade de que a idade média na amostra seja inferior a 43 anos?

$\frac{n}{N} = \frac{35}{250} = 0{,}14 > 0{,}05$, assim, o fator de correção deve ser usado.

Portanto, $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}} = \frac{9{,}3}{\sqrt{35}}\sqrt{\frac{250-35}{250-1}} = 1{,}57 * 0{,}93 = 1{,}46$

$P(\bar{X} < 43) = P\left(\frac{\bar{X}-\mu}{\sigma_{\bar{X}}} < \frac{43-39{,}5}{1{,}46}\right) = P(z < 2{,}40) \simeq 99{,}2\%$.

- **Distribuição amostral da Proporção**

Vamos considerar uma população em que a proporção de elementos portadores de uma certa característica é p. Assim, a população pode ser considerada como a variável X, tal que

$$X = \begin{cases} 0, & se\ o\ indivíduo\ não\ é\ portador\ da\ característica \\ 1, & se\ o\ indivíduo\ é\ portador\ da\ característica \end{cases}$$

Com $P(X = 1) = p$, $P(X = 0) = q$, $p + q = 1$. E sabemos que $\mu = E(X) = p$ e $\sigma^2 = Var(X) = pq = p(1-p)$.

Retirando uma amostra casual simples dessa população, com reposição, e indicando por S_n o total de indivíduos portadores da característica na amostra, teremos que $S_n \sim B(n, p)$. Logo $E(S_n) = np$ e $Var(S_n) = npq$.

Definimos como $\hat{p}$ a proporção de indivíduos portadores da característica na amostra, isto é,

$$\hat{p} = \frac{S_n}{n}.$$

Calculando esperança e variância de $\hat{p}$, temos: $E(\hat{p}) = p$ e $Var(\hat{p}) = pq/n$.

Quando n tende a infinito, podemos considerar a distribuição amostral de $\hat{p}$ do seguinte modo: $\hat{p} \sim N(p, pq/n)$.

E, além disso, $\frac{\hat{p}-p}{\sqrt{pq/n}} \sim N(0{,}1)$.

Exemplo: Em uma população, a proporção de pessoas favoráveis a uma determinada lei é de 40%. Retiramos uma amostra de 300 pessoas dessa população, qual a probabilidade de que a proporção de pessoas favoráveis à lei, seja de pelo menos 50%?

Temos $n = 300$ e $p = 0{,}4$; $q = 1 - p = 0{,}6$. $\sigma_{\hat{p}} = \sqrt{pq/n} = \sqrt{0{,}4 * 0{,}6/300} = 0{,}0283$.

$$\begin{aligned} P(\hat{p} \geq 0{,}50) &= P\left(\frac{\hat{p}-p}{\sqrt{pq/n}} \geq \frac{0{,}50-0{,}40}{\sqrt{0{,}40 * 0{,}60/300}}\right) = P(Z \geq 3{,}54) = 0{,}5 - P(0 \leq Z \leq 3{,}54) \\ &= 0{,}5 - 0{,}49980 = 0{,}00020 \end{aligned}$$

Tamanho de uma amostra

Muitas vezes é importante sabermos qual deverá ser o tamanho de uma amostra, para que a probabilidade de que um determinado parâmetro ou estimador esteja dentro dos limites seja um valor dado, ou então, queremos delimitar o erro da amostragem dentro de um risco determinado.

	v.a. quantitativa	v.a. qualitativa
População finita	$n = \frac{N * (z_{\alpha/2})^2 * \sigma^2}{(N-1) * e^2 + (z_{\alpha/2})^2 * \sigma^2}$	$n = \frac{N * pq * (z_{\alpha/2})^2}{(N-1) * e^2 + pq * (z_{\alpha/2})^2}$
População infinita	$n = \left(\frac{z_{\alpha/2} * \sigma}{e}\right)^2$	$n = \left(\frac{z_{\alpha/2} * \sqrt{pq}}{e}\right)^2$

Em que: $z_{\alpha/2}$ é valor crítico para o grau de confiança desejado; σ: desvio padrão populacional conhecido; e: erro padrão /erro máximo tolerado; N: tamanho da população finita; p: proporção de resultados favoráveis; q: proporção de resultados desfavoráveis.

Obs.: para população finita, aplica-se o fator de correção $(N-n)/(N-1)$, às fórmulas de população infinita.

Quando não se conhece o desvio padrão ou as frequências populacionais da variável, e não se dispuser de dados semelhantes na literatura, deve-se realizar um pré-teste com uma amostra piloto. E para o cálculo do tamanho da amostra, a distribuição t-*Student* deverá ser usada ao invés da normal, considerando o desvio padrão amostral. $t_{\alpha/2}$ é o valor crítico para o grau de confiança desejado, com $(m-1)$ graus de liberdade, onde m é o tamanho da amostra piloto.

Exemplo: Seja X uma v.a. com distribuição $N(1200, 840)$. Qual deverá ser o tamanho de uma amostra de tal forma que $P(1196 \leq \bar{X} \leq 1204) = 0{,}90$?

Temos: $\mu = 1200;\ \sigma^2 = 840\ e\ \mu_{\bar{X}} = 1200.$

Assim,$\sigma_{\bar{X}} = \sqrt{\frac{\sigma^2}{n}} = \sqrt{\frac{840}{n}} = \frac{28{,}98}{\sqrt{n}}$. Portanto, $Z = \frac{\bar{x}-\mu}{\sigma_{\bar{X}}}$.

Uma vez que queremos $P(1196 \leq \bar{X} \leq 1204) = 0{,}90$, temos que $Z_{(1-\alpha)/2} = Z_{0{,}45} = 1{,}64$.

Logo, $-1{,}64 = \frac{1196-1200}{\frac{28{,}98}{\sqrt{n}}}$ ou $1{,}64 = \frac{1204-1200}{\frac{28{,}98}{\sqrt{n}}}$.

De modo que $\sqrt{n} = \frac{1{,}64*28{,}98}{4} = 11{,}88 = 141{,}13 \quad \rightarrow \quad n \simeq 142.$

Ou ainda,

$n = \left(\frac{z_{\alpha/2}*\sigma}{e}\right)^2 = \left(\frac{1{,}64*28{,}98}{4}\right)^2 = 141{,}18 \rightarrow \quad n \simeq 142.$

Intervalos de confiança para média e proporção

A necessidade de construir intervalos de confiança surge do fato de que estimar pontualmente especifica um único valor para o estimador. Este procedimento não permite julgar qual a possível magnitude do erro que estamos cometendo. Daí surge a ideia de construir os intervalos (limites) de confiança, que são baseados na distribuição amostral do estimador pontual e com certa probabilidade incluem o verdadeiro valor do parâmetro.

Logo, a estimação por intervalo consiste na fixação de dois valores, tais que $(1-\alpha)$ seja a probabilidade de que o intervalo, por eles determinado, contenha o verdadeiro valor do parâmetro.

α: nível de incerteza ou grau de desconfiança.

$1-\alpha$: coeficiente de confiança ou nível de confiabilidade.

Logo, a partir da informação da amostra, devemos calcular os limites de um intervalo, valores críticos, que em $(1-\alpha)\%$ dos casos inclua o valor do parâmetro a estimar e em $\alpha\%$ dos casos não inclua o valor do parâmetro.

- **Intervalo de confiança para a média (variância conhecida)**

Consideramos uma população normal com média desconhecida que desejamos estimar e com σ^2 conhecida, $X \sim N(?\,, \sigma^2)$. Procedemos da seguinte forma para construir o intervalo de confiança:

1. Retiramos uma amostra casual simples de n elementos.
2. Calculamos a média da amostra $\bar{X}$.
3. Calculamos o desvio padrão da média amostral: $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$.
4. Fixamos o nível de significância α, e com ele determinamos Z_α, tal que $P(|z| > z_\alpha) = \alpha$, ou seja: $P(z > z_\alpha) = \alpha/2$ e $P(z < -z_\alpha) = \alpha/2$. Logo, devemos ter: $P(|z| < z_\alpha) = 1-\alpha$.

Como $Z = \frac{\bar{X}-\mu_{\bar{X}}}{\sigma_{\bar{X}}} \;\; \rightarrow \;\; P\left(\left|\frac{\bar{X}-\mu}{\sigma_{\bar{X}}}\right| < z_\alpha\right)$

$$P(|\bar{X}-\mu| < z_\alpha * \sigma_{\bar{X}}) = 1-\alpha \quad \rightarrow \quad P(-z_\alpha * \sigma_{\bar{X}} < \bar{X}-\mu < z_\alpha * \sigma_{\bar{X}}) = 1-\alpha$$

Portanto,

$$P(\bar{X} - z_\alpha * \sigma_{\bar{X}} < \mu < \bar{X} + z_\alpha * \sigma_{\bar{X}}) = 1-\alpha$$

é a fórmula do IC para a média de populações normais com variâncias conhecidas.

Por exemplo, $P(\bar{X} - 1{,}96\sigma_{\bar{X}} < \mu < \bar{X} + 1{,}96\sigma_{\bar{X}}) = 0{,}95$

deve ser interpretada do seguinte modo: construídos todos os intervalos da forma $\bar{X} \pm 1{,}96\sigma_{\bar{X}}$, 95% deles conterão o parâmetro μ. Note que este intervalo pode ou não conter o parâmetro μ, mas pelo exposto temos 95% de confiança de que contenha.

Exemplo: Uma máquina enche garrafas de álcool em gel com uma variância igual a 100 ml^2. Ela estava regulada para enchê-las com 500 ml, em média. Agora ela se desregulou, e queremos saber qual a nova média μ. Uma amostra de 25 garrafas apresentou uma média igual a 492 ml. Vamos construir um intervalo de 95% de confiança para μ. Pelo corolário 1, teremos

$IC(\mu; 95\%) = 492 \pm 1{,}96 * 2 = (488{,}08, 495{,}92)$ uma vez que $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{10}{5} = 2$ ml.

Observação: Quando queremos estimar a média de uma população normal com variância desconhecida, consideramos dois procedimentos:

- Se $n \leq 30$, então usa-se a distribuição *t* de *Student*, com o estimador da variância, s^2;
- Se $n > 30$, então usa-se a distribuição normal com o estimador s^2 de σ^2, onde: $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$.

Como a amostra é grande, $s^2 \cong \sigma^2$ e $\sigma_{\bar{X}} \cong \frac{s}{\sqrt{n}}$.

Portanto

$$P(\bar{X} - z_\alpha * \sigma_{\bar{X}} < \mu < \bar{X} + z_\alpha * \sigma_{\bar{X}}) = 1-\alpha$$

é a formulação já conhecida para o IC para a média.

Exemplo: Testes de compressão foram aplicados na marca A de cimento, para avaliar sua resistência em concretos. Foram produzidos 13 corpos de prova. A empresa afirma que o processo tem variabilidade amostral $s^2 = 21{,}76\ MPa^2$. Construa um intervalo de confiança com nível de significância $\alpha = 0{,}05$ para a resistência à compressão média. Sabendo-se que foi observada uma média amostral igual a 33,76 Mpa.

$$P\left(\bar{X} - t_{(n-1);\frac{\alpha}{2}} * \frac{s}{\sqrt{n}} \leq \mu \leq \bar{X} + t_{(n-1);\frac{\alpha}{2}} * \frac{s}{\sqrt{n}}\right) = 1 - \alpha$$

$$P\left(33{,}76 - 2{,}179 * \frac{4{,}665}{\sqrt{13}} \leq \mu \leq 33{,}76 + 2{,}179 * \frac{4{,}665}{\sqrt{13}}\right) = 0{,}95 \dashrightarrow (30{,}94; 36{,}58)$$

- **Intervalo de confiança para a proporção**

Seja o seguinte exemplo:

Uma amostra de $n = 500$ pessoas de uma cidade é escolhida, e a cada pessoa da amostra é feita uma pergunta a respeito de um problema municipal, para o qual é apresentada uma solução pela prefeitura. A resposta à pergunta poderá ser SIM (favorável a solução) ou NÃO (contrária a solução). Deseja-se estimar a proporção (p) de pessoas na cidade favoráveis à solução apresentada.

Vamos obter um intervalo de confiança para p. Sabemos que se X = número de sucessos nas n perguntas, então X tem uma distribuição aproximadamente normal, com média $\mu = p$ e variância $\sigma^2 = pq$, com $q = 1 - p$.

Portanto, com probabilidade 0,95, temos que $-1{,}96\sqrt{\frac{\hat{p}\hat{q}}{n}} \leq \hat{p} - p \leq 1{,}96\sqrt{\frac{\hat{p}\hat{q}}{n}}$

Assim, $P(\hat{p} - 1{,}96\sigma_{\hat{p}} < p < \hat{p} + 1{,}96\sigma_{\hat{p}}) = 0{,}95$.

Sorteada uma amostra e encontrado a proporção $\hat{p}$, podemos construir o intervalo $\hat{p} \pm 1{,}96\sigma_{\hat{p}}$.

Exemplo: Numa pesquisa de mercado, $n = 400$ pessoas foram entrevistadas sobre determinado produto, e 60% destas pessoas preferiram a marca A. Aqui $\hat{p} = 0{,}6$, e um intervalo de confiança para p, com confiança de 95%, será

$0{,}6 \pm 1{,}96\sqrt{\frac{(0{,}6)(0{,}4)}{400}} = 0{,}6 \pm 0{,}049$, ou seja, $IC(p{:}\,95\%) = (0{,}551; 0{,}649)$.

Testes de Hipóteses

Em Estatística, uma **hipótese** é uma alegação, ou afirmação, sobre uma propriedade de uma população.

As afirmações seguintes são exemplos de hipóteses a serem testadas pelos processos a serem apresentados:

- Pesquisadores médicos afirmam que a temperatura média do corpo humano não é igual a 98,6º F.

- A percentagem de motoristas hospitalizados em consequência de acidentes é menor no caso de carros equipados com *airbag* do que no caso de carros sem esse equipamento.

- Quando se utiliza equipamento novo na fabricação de altímetros para aviões, a variação nos erros é reduzida, de modo que os resultados apresentados são mais consistentes.

Antes de iniciar um estudo deste item, deve-se ter em mente esta diretriz geral para o raciocínio estatístico: *analisar uma amostra para distinguir entre resultados que podem ocorrer facilmente e os que dificilmente ocorrem.*

Podemos explicar a ocorrência de resultados altamente improváveis dizendo que ou ocorreu efetivamente um evento raro, ou as coisas não são como supúnhamos.

Formulamos duas hipóteses básicas:

H_0: hipótese nula ou da existência.

H_1: hipótese alternativa.

Podemos apresentar as hipóteses genéricas que englobam a maioria dos casos:

1. $\begin{cases} H_0: \theta = \theta_0 \\ H_1: \theta \neq \theta_0 \end{cases}$ para testes bilaterais
2. $\begin{cases} H_0: \theta \leq \theta_0 \\ H_1: \theta > \theta_0 \end{cases}$ para testes unilaterais à direita
3. $\begin{cases} H_0: \theta \geq \theta_0 \\ H_1: \theta < \theta_0 \end{cases}$ para testes unilaterais à esquerda

Em geral, admite-se a hipótese nula como igualdade. No entanto, o imprescindível é que as duas hipóteses se complementem.

Há dois tipos possíveis de erro ao testar uma hipótese estatística. Rejeitar uma hipótese quando ela é, de fato, verdadeira, ou aceitar uma hipótese quando ela é, de fato, falsa. Define-se:

Erro tipo I: Rejeitar H_0 quando H_0 for verdadeira.

Erro tipo II: Não rejeitar H_0 quando H_0 for falsa.

	Não rejeitar H_0	Rejeitar H_0
H_0 verdadeira	Decisão correta	Erro do tipo I
H_0 falsa	Erro do tipo II	Decisão correta

E temos: $P(Erro\ tipo\ I) = \alpha$ e $P(Erro\ tipo\ II) = \beta$.

Onde, a situação ideal é aquela em que ambas as probabilidades, α e β, são próximas de zero. No entanto, é fácil ver que a medida que diminuímos α, β aumenta.

A probabilidade α do erro tipo I é denominada <u>nível de significância</u> do teste.

Há uma sequência de passos que podem ser usados sistematicamente para qualquer teste de hipóteses, como segue:

1º Passo: Fixe qual a hipótese H_0 a ser testada e qual a hipótese alternativa H_1.

2º Passo: Use a teoria estatística e as informações disponíveis para decidir qual estatística (estimador) será usada para julgar a hipótese H_0.

3º Passo: Fixe a probabilidade α de cometer um erro tipo I, e use este valor para construir a região crítica (RC).

4º Passo: Use as informações fornecidas pela amostra para encontrar o valor da estatística que definirá a decisão.

5º Passo: Se o valor da estatística observado na amostra não pertencer a região crítica, não rejeite H_0; caso contrário, rejeite H_0.

p-valor: também denominado nível descritivo do teste, é a probabilidade de que a estatística de teste tenha valor extremo em relação ao valor observado, quando H_0 é verdadeira. Também podemos interpretá-lo como o menor valor do nível de significância para o qual rejeitamos a hipótese (se $\alpha > p.valor$).

$$p.valor = \begin{cases} 2[1 - \Phi(|z_0|)], teste\ bilateral\ (H_1: \theta \neq \theta_0) \\ 1 - \Phi(z_0), teste\ unilateral\ \text{à}\ direita\ (H_1: \theta > \theta_0) \\ \Phi(z_0), teste\ unilateral\ \text{à}\ esquerda\ \ (H_1: \theta < \theta_0) \end{cases}$$

Em que $\Phi(z) = P(Z \leq z)$ é a distribuição cumulativa N(0,1).

- **Teste para a média de uma população**

Vejamos agora uma aplicação dos cinco passos, definidos anteriormente, para testar a hipótese de que a média de uma população μ é igual a um número fixado μ_0, supondo-se a variância σ^2 da população conhecida.

Exemplo: Uma máquina automática de encher garrafas de álcool em gel, enche-as segundo uma distribuição normal, com média μ e variância 100 ml^2. O valor de μ pode ser fixado num mostrador situado numa posição um pouco inacessível dessa máquina. Desejamos, a cada meia hora, colher uma amostra de 16 garrafas e verificar se o processo de enchimento está sob controle, isto é, se μ = 500 ml ou não. Se uma dessas amostras apresentasse uma média $\bar{x}$ = 492 ml, você pararia ou não de encher as garrafas para verificar se o mostrador está na posição correta?

Este é um exemplo típico de teste de hipóteses.

Indiquemos por X a quantidade de álcool em cada garrafa; então, $X\sim N(\mu, 100)$. E as hipóteses que nos interessa são:

$$\begin{cases} H_0: \mu = 500ml \\ H_1: \mu \neq 500\ ml \end{cases}$$

pois a máquina pode desregular para mais ou para menos. O problema fornece informações sobre a alternativa, que poderia ter uma das três formas abaixo:

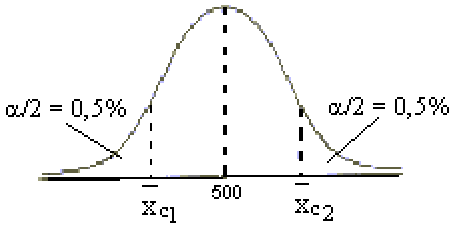
(i) $H_1: \mu \neq 500\ ml$ (teste bilateral)
(ii) $H_1: \mu > 500\ ml$ (teste unilateral à direita)
(iii) $H_1: \mu < 500\ ml$ (teste unilateral à esquerda)

Pela afirmação do problema, σ^2 = 100 será sempre a mesma; assim, qualquer que seja a média μ, a média $\bar{X}$ de 16 garrafas terá distribuição

$\bar{X}\sim N(\mu, 100/16)$, isto é, $\bar{X}\sim N(\mu; 6{,}25)$.

Em particular, se H_0 é verdadeira, $\bar{X}\sim N(500; 6{,}25)$.

Vamos fixar $\alpha = 1\%$; pela hipótese alternativa, vemos que a hipótese H_0 deve ser rejeitada quando $\bar{X}$ for muito pequeno ou muito grande (teste bilateral). Assim, nossa região crítica será como a figura abaixo.



Da tabela da curva normal, obtemos que

$Z_1 = -2{,}58 = \frac{\bar{x}_{c1}-500}{2{,}5}\ \therefore\ \bar{x}_{c1} = 493{,}55$ e $Z_1 = 2{,}58 = \frac{\bar{x}_{c2}-500}{2{,}5}\ \therefore\ \bar{x}_{c2} = 506{,}45$.

Logo, $RC = \{\bar{x} \in \mathbb{R} | \bar{x} < 493{,}55\ ou\ \bar{x} > 506{,}45\}$.

A informação pertinente da amostra é a sua média, que neste caso particular é $\bar{x}_0 = 492\ ml$.

Como $\bar{x}_0 \in RC$, a nossa conclusão será de rejeitar H_0, ou seja, é possível que a máquina esteja desregulada. Destacando que esse é o resultado particular de uma amostra observada. Em outras amostras, outros resultados para a média são esperados. O aconselhável é olhar o mostrador.

- **Teste para Proporção**

Construção do teste para proporções:

Temos uma população, e temos uma hipótese sobre a proporção p de indivíduos portadores de uma certa característica. Esta hipótese afirma que essa proporção é igual a um certo número p_0. Então, $H_0: p = p_0$.

O problema fornece informações sobre a alternativa, que pode ter uma das três formas abaixo:

(i) $H_1: p \neq p_0$ (teste bilateral)
(ii) $H_1: p > p_0$ (teste unilateral à direita)
(iii) $H_1: p < p_0$ (teste unilateral à esquerda)

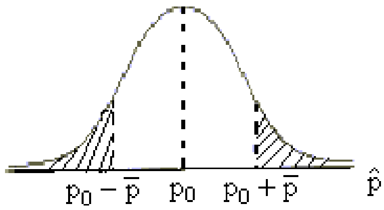
Como vimos, a estatística p, a proporção da amostra, tem uma distribuição aproximadamente normal, isto é:

$\hat{p} \sim N(p, pq/n)$.

Fixado um valor α, devemos construir a região crítica para p na suposição de que os parâmetros definidos em H_0 sejam verdadeiros.

$\hat{p} \sim N(p_0, p_0 q_0/n)$,

e, consequentemente teremos a região crítica da figura abaixo (no caso, supondo $H_1: p \neq p_0$), onde $\bar{p} = z_\alpha \sqrt{p_0 q_0/n}$.



Os últimos passos irão depender da amostra, e o procedimento está descrito no exemplo seguinte.

Exemplo: Em uma eleição municipal, o candidato a prefeito X afirma que 60% dos eleitores votam nele. O candidato Y deseja contestar essa afirmação. Para isso, realiza uma pesquisa com 200 eleitores. Qual deve ser o procedimento adotado para julgar a veracidade da afirmação do candidato X? No passo 4 abaixo, daremos o resultado da amostra, pois é importante ficar claro que o resultado da amostra não deve influenciar a escolha da hipótese alternativa.

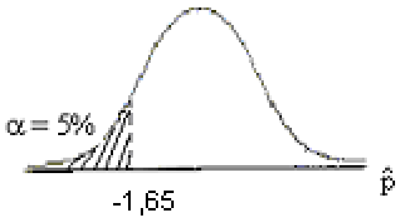
Vamos colocar à prova a afirmação do candidato X, isto é, $H_0: p = 0{,}60$.

Sabemos que, se esta hipótese não for verdadeira, espera-se uma proporção menor, nunca maior. O candidato sempre divulgaria o máximo possível. Isto nos leva à hipótese alternativa $H_1: p < 0{,}60$.

A estatística a ser usada é p, a proporção de 200 eleitores que foram entrevistados, e da teoria temos $\hat{p} \sim N(p, p(1-p)/n)$.

Fixaremos α = 5%, e sob a suposição de que H_0 seja verdadeira, a distribuição de p será:

$\hat{p} \sim N(0{,}60; 0{,}24/200)$, o que irá fornecer a região crítica (de acordo com a figura a seguir): $RC = \{p \in \mathbb{R} | \hat{p} \leq -1{,}65\}$.



Pois devemos achar o valor $\hat{p}_c$, tal que $P(\hat{p} \leq \hat{p}_c) = 0{,}05$ e como p é aproximadamente normal, temos $P\left(z \leq (\hat{p}_c - 0{,}60)/\sqrt{0{,}24/200}\right) = 0{,}05$. E assim, $\hat{p}_c = -1{,}65$, correspondendo a região crítica acima.

Admitamos que, do trabalho de campo, entrevistando os 200 eleitores sorteados aleatoriamente, obtivemos 104 respostas afirmativas. Isto equivale a um valor observado da proporção de $\hat{p} = \frac{104}{200} = 0{,}52$, o que implica $z = \frac{0{,}52-0{,}60}{\sqrt{0{,}24/200}} = -2{,}31$.

Do resultado acima, vemos que $-2{,}31 \in RC$; portanto, somos levados a rejeitar H_0. Isto é, há evidências de que o candidato X não tem a intenção de votos de 60% e sim, inferior a este número.

Exemplo: No passado, 15% das solicitações postais feitas por uma instituição de caridade resultaram em contribuição financeira. Uma nova carta de solicitação foi redigida. Esta nova carta elevou a taxa de contribuição. A nova carta é enviada a uma amostra de 200 pessoas e 45 responderam com uma contribuição.

- Ao nível de significância de 0,05 pode-se concluir que a nova carta é mais efetiva?

- Estabeleça a Hipótese nula e a Hipótese alternativa:
  $\begin{cases} H_0: p \leq 0{,}15 \\ H_1: p > 0{,}15 \end{cases}$
- Estabeleça a regra de decisão:

H_0 é rejeitada se o $z_{obs} > 1{,}645$.

- Calcule o valor da estatística de teste (valor do z_{obs}): $z = \frac{\frac{45}{200}-0{,}15}{\sqrt{(0{,}15)(0{,}18)/200}} = 2{,}97$.

- Qual é a decisão sobre a hipótese nula? Interprete os resultados.

Desde que o $z_{obs} = 2{,}97 > z_c(1{,}645)$, H_0 é rejeitada. A nova carta é mais efetiva.

Poder do teste: Seja T um teste estatístico, com região crítica RC, para avaliarmos as hipóteses a respeito do parâmetro θ. A função poder do teste é a probabilidade de rejeitarmos H_0 dado o valor de θ. Neste caso, temos que

$\pi(\theta) = P[rejeitar\ H_0|\theta] = P[T \in RC|\theta]$, para todo valor de θ.

O poder do Teste tem como objetivo conhecer o quanto o teste de hipóteses controla um erro do tipo II, ou qual a probabilidade de rejeitar a hipótese nula se realmente falsa. O poder de um teste é afetado por três fatores:

- Tamanho da amostra: mantendo todos os outros parâmetros iguais, quanto maior o tamanho da amostra, maior o poder do teste.
- Nível de significância: quanto maior o nível de significância, maior o poder do teste. Aumentando-o, reduz-se a região de aceitação, resultando em maior chance de rejeitar a hipótese nula.
- O verdadeiro valor do parâmetro a ser testado: quanto maior a diferença entre o "verdadeiro" valor do parâmetro e o valor especificado pela hipótese nula, maior o poder do teste.

Consideremos a estatística $Z = \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}$ e o teste a hipótese $\begin{cases} H_0: \mu = \mu_0 \\ H_1: \mu \neq \mu_0 \end{cases}$.

O erro do tipo II é o erro cometido ao aceitar H_0, quando esta é falsa.

$$P(Erro\ tipo\ II) = P(aceitar\ H_0 | H_1 \text{ é } verdadeira\ ) = \beta$$

Para que isto seja possível, suponha que a hipótese nula é falsa e que o verdadeiro valor da média é $\mu = \mu_0 + \delta$, então a estatística do teste é

$$Z_0 = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} = \frac{\bar{X} - (\mu_0 + \delta)}{\sigma/\sqrt{n}} + \frac{\delta}{\sigma/\sqrt{n}}$$

Portanto, a distribuição de Z_0 quando $\mu = \mu_0 + \delta$ é, $Z_0 \sim N\left(\frac{\delta}{\sigma/\sqrt{n}}, 1\right)$.

E, com isso para um teste bilateral, temos que β, é a probabilidade de que Z_0 esteja entre $-z_{\alpha/2}$ e $z_{\alpha/2}$, dado que H_1 verdadeira. Essa probabilidade é calculada da seguinte maneira:

$$\beta = \Phi\left(z_{\alpha/2} - \frac{\delta}{\sigma/\sqrt{n}}\right) - \Phi\left(-z_{\alpha/2} - \frac{\delta}{\sigma/\sqrt{n}}\right)$$

em que Φ é a função acumulada da distribuição normal padrão.

Para os testes unilaterais à direita e à esquerda, temos que as probabilidades de erro do tipo II são dadas, respectivamente por:

$\Phi\left(z_\alpha - \frac{\delta}{\sigma/\sqrt{n}}\right)$ e $1 - \Phi\left(-z_\alpha - \frac{\delta}{\sigma/\sqrt{n}}\right)$.

O poder do teste é calculado como sendo $P = 1 - \beta$.

Exemplo: Num certo banco de dados, o tempo para realização das buscas é aproximadamente normal com média 53s e desvio padrão 14s. Modificou-se o sistema para reduzir o tempo. Foram contados os tempos para 30 buscas. Admita que as 30 observações possam ser consideradas uma amostra aleatória e que não houve alteração na variância. Use $\alpha = 1\%$. Calcule o poder do teste se a verdadeira média do tempo fosse: a) 40s; b) 45s; c) 52s.

Temos $\begin{cases} H_0: \mu = 53\ s \\ H_1: \mu < 53\ s \end{cases}$, $n = 30$, $\alpha = 1\%$, $\sigma = 14\ s$, $z_\alpha = -2{,}33$,

$$\bar{x}_c = \mu + z_\alpha * \frac{\sigma}{\sqrt{n}} = 53 - 2{,}33 * \frac{14}{\sqrt{30}} = 47{,}04$$

a) $\mu = 40$; $z_\beta = \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}} = \frac{47{,}04-40}{\frac{14}{\sqrt{30}}} = 2{,}75$

$$\beta = P(z_{obs} < 2{,}75) = 1 - \Phi(2{,}75) = 1 - 0{,}99702 = 0{,}00298$$

$$\mathrm{P} = 1 - \beta = 0{,}99702$$

Logo, 0,99702 é a probabilidade de rejeitar H_0, quando esta hipótese é falsa.

b) $\mu = 45$; $z_\beta = \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}} = \frac{47{,}04-45}{\frac{14}{\sqrt{30}}} = 0{,}80$

$$\beta = P(z_{obs} < 0{,}80) = 1 - \Phi(0{,}80) = 1 - 0{,}78814 = 0{,}21186$$

$$\mathrm{P} = 1 - \beta = 0{,}78814$$

c) $\mu = 52$; $z_\beta = \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}} = \frac{47{,}04-52}{\frac{14}{\sqrt{30}}} = -1{,}94$

$$\beta = P(z_{obs} < -1{,}94) = \Phi(1{,}94) = 0{,}97381$$

$$\mathrm{P} = 1 - \beta = 0{,}02619$$

Exemplo: Um certo tipo de pneu dura, em média, 50000 Km. O fabricante investiu em uma nova composição de borracha para pneus, objetivando aumentar sua durabilidade. Vinte pneus, fabricados com esta nova composição, apresentaram desvio padrão de 4000 Km. Use $\alpha = 1\%$. Calcule o poder do teste se a verdadeira média de durabilidade dos pneus fosse de: a) 55000 Km; b) 51000 Km.

Temos $\begin{cases} H_0: \mu = 50000\ Km \\ H_1: \mu < 50000\ Km \end{cases}$, $n = 20$, $s = 4000\ Km$, $\alpha = 1\%$, $t_{1\%;19} = 2{,}539$,

$$\bar{x}_c = \mu + t_\alpha * \frac{s}{\sqrt{n}} = 50000 + 2{,}539 * \frac{4000}{\sqrt{20}} = 52270{,}95$$

a) $\mu = 55000\ Km$; $t_\beta = \frac{\bar{x}-\mu}{\frac{s}{\sqrt{n}}} = \frac{52270{,}95-55000}{\frac{4000}{\sqrt{20}}} = -3{,}051$

$$\beta = P(t_{obs} < -3{,}051) = 0{,}00329$$
$$\mathrm{P} = 1 - \beta = 0{,}99671$$

b) $\mu = 51000\ Km$; $t_\beta = \frac{\bar{x}-\mu}{\frac{s}{\sqrt{n}}} = \frac{52270{,}95-51000}{\frac{4000}{\sqrt{20}}} = 1{,}42$

$$\beta = P(t_{obs} < 1{,}42) = 0{,}91410$$
$$\mathrm{P} = 1 - \beta = 0{,}08590$$

Testes Paramétricos: *t* e ANOVA

Às vezes, é preciso comparar dois ou mais tratamentos, como, por exemplo dois processos de têmpera na produção de aço, três tipos de cimento-e-cola para fixar azulejos, etc. Para realizar as comparações, podemos planejar experimentos com amostras submetidas a cada tratamento. Em cada ensaio, observamos uma resposta adequada. Por exemplo, no caso do aço, a resposta pode ser a resistência mecânica; no cimento-e-cola, o grau de aderência, e assim por diante.

Na prática, porém, muitas vezes não temos total liberdade de escolher o projeto de experimento mais adequado. Ou ainda, quando queremos comparar o efeito de certo evento, em que é natural observar as unidades experimentais antes e depois do evento, resultando em dados pareados.

Teste *t*

O chamado teste t é apropriado para comparar dois conjuntos de dados quantitativos, em termos de seus valores médios. Mais especificamente,

$$\begin{cases} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 \neq \mu_2 \end{cases}$$

Teste *t* para observações pareadas

Para realizar o teste t emparelhado, primeiro devemos estabelecer uma das hipóteses:

$$\begin{cases} H_0: \mu_d = 0 \\ H_1: \mu_d \neq 0 \end{cases} \qquad \begin{cases} H_0: \mu_d = 0 \\ H_1: \mu_d > 0 \end{cases} \qquad \begin{cases} H_0: \mu_d = 0 \\ H_1: \mu_d < 0 \end{cases}$$

Em que $\mu_d = \mu_2 - \mu_1$.

Para isso, primeiro, estabelece-se o nível de significância do teste e depois calcula:

a) A diferença entre as unidades de cada um dos n pares: $d = x_1 - x_2$.
b) A média das diferenças: $\bar{d} = \frac{\sum d}{n}$.
c) A variância das diferenças: $s^2 = \frac{\sum d^2 - \frac{(\sum d)^2}{n}}{n-1}$

d) O valor de t, $t = \frac{\bar{d}}{\sqrt{s^2/n}}$

que está associado a $n - 1$ graus de liberdade.

$$p - valor = \begin{cases} 2 * P(|t| > t_{obs}), teste\ bilateral\ (H_1: \mu_d \neq 0) \\ P(t > t_{obs}), teste\ unilateral\ \text{à}\ direita\ (H_1: \mu_d > 0) \\ P(t < t_{obs}), teste\ unilateral\ \text{à}\ esquerda\ \ (H_1: \mu_d < 0) \end{cases}$$

Exemplo: A fim de determinar a eficiência de um medicamento antitérmico, a temperatura corporal (em graus Celsius) de 10 indivíduos foi medida. Em seguida, foi administrado o medicamento e após uma hora a temperatura foi medida novamente. Os resultados podem ser encontrados na tabela abaixo.

	Temperatura	
Indivíduo	**Antes**	**Depois**
1	37,5	37,8
2	36,0	36,4
3	39,0	37,6
4	38,0	37,2
5	37,8	36,9
6	38,5	37,7
7	36,9	36,8
8	39,4	38,1
9	37,2	36,7
10	38,1	37,3

$$\begin{cases} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 > \mu_2 \end{cases}$$

Em termos da variável diferença D, as hipóteses são descritas como $\begin{cases} H_0: \mu_d = 0 \\ H_1: \mu_d > 0 \end{cases}$

Para fazer o teste, é preciso primeiro estabelecer o nível de significância. Seja $\alpha = 1\%$.

a) As diferenças ($d = x_1 - x_2$) entre os valores das temperaturas antes e depois de ser administrado o medicamento: -0,3; -0,4; 1,4; 0,8; 0,9; 0,8; 0,1; 1,3; 0,5; 0,8

b) A média das diferenças: $\bar{d} = \frac{5,9}{10} = 0,59$

c) A variância das diferenças: $s^2 = \frac{0,79+0,98+0,66+0,04+0,10+0,04+0,24+0,50+0,01+0,04}{9} = \frac{3,4}{9} = 0,38$

d) O valor de t é $t_0 = \frac{0,59}{\sqrt{0,38/10}} = 3,03$, que está associado a $n - 1 = 10 - 1 = 9$ graus de liberdade.

De acordo com a tabela, ao nível de significância de 1% e 9 graus de liberdade, temos $t = 2,82$.

Como o valor de t_0 calculado é maior do que o valor de t tabelado, rejeitamos a hipótese de que as médias de temperaturas são iguais. E conclui-se que o tratamento tem efeito significativo ao nível de 1%. Em termos práticos, podemos dizer que o medicamento é eficiente.

Teste t para observações independentes

Se a variável em análise tem distribuição normal, ou aproximadamente normal, aplica-se o teste t para comparar duas médias. Mas, primeiro é preciso estabelecer o nível de significância α.

Dados os dois grupos 1 e 2, calculam-se:

a) A média de cada grupo: $\bar{x}_1$ e $\bar{x}_2$

b) A variância de cada grupo: s_1^2 e s_2^2

c) Variância ponderada dada pela fórmula: $s^2 = \frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}$,

em que n_1e n_2 são os números de elementos dos grupos 1 e 2, respectivamente.

d) O valor de t é definido por: $t = \frac{\bar{x}_1-\bar{x}_2}{\sqrt{s^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$.

Feitos os cálculos, é preciso comparar o valor calculado de t com um valor tabelado de t, ao nível de significância estabelecido e $(n_1 + n_2 - 2)$ graus de liberdade.

Se o valor calculado de t (em valor absoluto) for igual ou maior ao valor tabelado, conclui-se que as médias não são iguais, ao nível de significância estabelecido.

Exemplo: Um produto fabricado por injeção de plástico é analisado em dois níveis de percentual de talco. Os dados seguintes apresentam resultados da dureza (HRc), segundo o percentual de talco utilizado.

Baixo	Alto
51,7	75,2
49,4	76,0
65,9	63,7
60,0	69,6
71,1	67,1
72,9	69,1
71,9	52,8
75,1	57,6

Os dados mostram evidência suficiente para afirmar que a dureza média do produto é diferente nos dois níveis de percentual de talco? Usando $\alpha = 5\%$.

$$\begin{cases} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 \neq \mu_2 \end{cases}$$

Para proceder ao teste t, é preciso primeiro, estabelecer o nível de significância. Seja $\alpha = 5\%$. E depois calcula:

a) Média de cada grupo:

$\bar{x}_1 = \frac{518}{8} = 64{,}75$ e $\bar{x}_2 = \frac{531{,}1}{8} = 66{,}39$

b) Variância de cada grupo:

$$s_1^2 = \frac{(51{,}7 - 64{,}75)^2 + (49{,}4 - 64{,}75)^2 + \cdots + (75{,}1 - 64{,}75)^2}{7} = \frac{694{,}8}{7} = 99{,}26$$

$$s_2^2 = \frac{(75{,}2 - 66{,}39)^2 + (76{,}0 - 66{,}39)^2 + \cdots + (57{,}6 - 66{,}39)^2}{7} = \frac{457{,}31}{7} = 65{,}33$$

c) A variância ponderada:

$$s^2 = \frac{(8-1) * 99{,}26 + (8-1) * 65{,}33}{8+8-2} = \frac{1152{,}11}{14} = 82{,}29$$

d) O valor de t:

$$t = \frac{64{,}75 - 66{,}39}{\sqrt{82{,}29 * \left(\frac{1}{8}+\frac{1}{8}\right)}} = -\frac{1{,}64}{4{,}54} = -0{,}36$$

que está associado a $n_1 + n_2 - 2 = 8 + 8 - 2 = 14$ graus de liberdade. Na tabela, ao nível de significância $\alpha = 5\%$ (teste bilateral) e com 14 graus de liberdade, $t = 2{,}145$. Como o valor calculado não pertence a região de rejeição, conclui-se que os dados não mostram evidência suficiente para afirmar que a dureza média do produto é diferente nos dois níveis de percentual de talco.

Teste F e Análise de Variância – ANOVA

- **Teste F para duas variâncias**

Suponha que queremos comparar se duas populações, supostamente com distribuições normais, têm a mesma variância. Formulamos as hipóteses por:

$$\begin{cases} H_0: \sigma_1^2 = \sigma_2^2 \\ H_1: \sigma_1^2 \neq \sigma_2^2 \end{cases}$$

em que:

σ_1^2: variância da população 1;

σ_2^2: variância da população 2.

A hipótese alternativa também pode ser $H_1: \sigma_1^2 > \sigma_2^2$, com as amostras da população 1 e da população 2, a estatística de teste é calculada por:

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

em que: s_1^2: variância da amostra de n_1 elementos; E, s_2^2: variância da amostra de n_2 elementos.

Considerando $s_1^2 > s_2^2$, ou seja, a maior variância deve ser colocada no numerador.

A distribuição de referência para este teste é a chamada distribuição F com $\mathrm{g.l.} = n_1 - 1$ no numerador e $\mathrm{g.l.} = n_2 - 1$ no denominador (para os níveis de significância de 10%; 5%; 2,5% e 1%). Assim, estabelecido o nível de significância α, podemos obter o valor f_c, que deixa área igual a $\alpha/2$ na cauda superior da distribuição (teste bilateral) ou, no caso de teste unilateral, área igual a α. A regra de decisão da abordagem clássica é dada por:

$$\text{Se } \mathrm{f} < \mathrm{f_c}, \text{ não rejeita } \mathrm{H_0};$$
$$\text{Se } \mathrm{f} \geq \mathrm{f_c}, \text{ rejeita } \mathrm{H_0}.$$

Exemplo: Em estudos anteriores verificou-se que não há evidências de que dois catalisadores A e B tenham efeitos médios diferentes no rendimento de certa reação química. Vamos verificar, agora, se eles produzem efeitos diferentes nas variâncias. As hipóteses podem ser:

$\mathrm{H_0}$: as variâncias do rendimento são iguais para os dois catalisadores;

$\mathrm{H_1}$: as variâncias do rendimento são diferentes para os dois catalisadores.

Dados os resultados dos experimentos:

Amostra 1: $\mathrm{n_1} = 10$; $\bar{x}_1 = 49{,}900$ e $\mathrm{s_1^2} = 35{,}656$;

Amostra 2: $\mathrm{n_2} = 10$; $\bar{x}_2 = 44{,}700$ e $\mathrm{s_2^2} = 42{,}233$.

No cálculo de f, a maior variância é colocada no numerador, assim:

$$f = \frac{\mathrm{s_2^2}}{\mathrm{s_1^2}} = \frac{42{,}233}{35{,}656} = 1{,}18$$

Para obter o valor crítico f_c ao nível de significância de 5%, devemos obter área igual a 2,5% na cauda superior (teste bilateral) da distribuição F, com $g.l. = 9$ no numerador e $g.l. = 9$ no denominador, o que acarreta $f_c = 4{,}03$.

Como $\mathrm{f} < \mathrm{f_c}$, o teste aceita $\mathrm{H_0}$.

- **Teste F e ANOVA**

Vimos como comparar médias de dois grupos de observações. Mas às vezes é preciso comparar médias de g ($g > 2$) grupos. Para isso, existe um método chamado Análise de Variância – ANOVA, onde aplica-se o teste F (desde que a variável em estudo tenha distribuição normal ou aproximadamente normal). Trata-se de um teste de hipótese nula de igualdade das médias dos grupos, contra a hipótese alternativa de pelo menos dois grupos diferirem em relação a suas médias.

A estatística de teste é construída como uma razão entre a variabilidade entre os grupos (numerador) e a variabilidade dentro dos grupos (denominador). A distribuição amostral dessa razão é chamada distribuição F de Snedecor com n_1 graus de liberdade no numerador e n_2 graus de liberdade no denominador. Para haver diferença significativa entre os grupos, a variabilidade entre eles tem que ser significativamente maior que a variabilidade dentro deles. A regra de decisão para rejeitar a hipótese de nulidade é análoga às dos testes de comparação de dois grupos. Procura-se em tabelas apropriadas da distribuição F os valores críticos que delimitam as regiões de rejeição ou não da hipótese nula.

- **Amostras independentes**

A análise estatística para comparação de g grupos independentes é tradicionalmente feita por uma análise de variância, acompanhada de um teste F, que, da mesma forma como o teste t, supõe:

1. As observações devem ser independentes;
2. As variâncias populacionais devem ser iguais nos grupos;
3. A distribuição das observações em cada grupo deve ser normal.

Formalmente, temos as seguintes hipóteses:

$$\begin{cases} H_0: \mu_1 = \mu_2 = \cdots = \mu_g \\ H_1: \mu_i \neq \mu_j, para\ algum\ i \neq j \end{cases}$$

onde μ_i representa o valor esperado da resposta sob o tratamento i ($i = 1, 2, \dots, g$).

Considerando n replicações sob cada tratamento (amostra de n elementos em cada grupo, totalizando $N = ng$ observações), podemos representar os dados pelo seguinte modelo estatístico:

$$Y_{ij} = \mu + \tau_i + \epsilon_{ij} \quad (i = 1, 2, \dots, g; j = 1, 2, \dots, n)$$

onde:

Y_{ij} é a variável aleatória associada à j-esima observação do i-esimo tratamento;

μ é a média global da resposta (independente do tratamento);

τ_ié o efeito do i-esimo tratamento;

ϵ_{ij} é o efeito aleatório ou erro experimental, o qual é suposto com distribuição aproximadamente normal, média zero e variância constante.

Considerando o modelo acima, o valor esperado da resposta no i-esimo tratamento é dado por $\mu_i = \mu + \tau_i$. E as hipóteses podem ser escritas por:

$$\begin{cases} H_0: \tau_1 = \tau_2 = \cdots = \tau_g = 0 \\ H_1: \tau_i \neq 0, para\ algum\ i = 1, 2, \dots, g \end{cases}$$

As observações, as somas e as médias por tratamento são representadas por:

Replicação	**Tratamento**			
	1	**2**	**...**	***g***
1	y_{11}	y_{21}	...	y_{g1}
2	y_{12}	y_{22}	...	y_{g2}

...	...	...	...	...	
n	y_{1n}	y_{2n}	...	y_{gn}	
Soma	y_1	y_2	...	y_g	$y_T = \sum_{i=1}^{g} y_i$
Média	$\bar{y}_1$	$\bar{y}_2$	...	$\bar{y}_g$	$\bar{y} = \frac{1}{g}\sum_{i} \bar{y}_i$

Considere a seguinte soma de quadrados:

$$SQ_{Tot} = \sum_{i=1}^{g}\sum_{j=1}^{n}\left(y_{ij} - \bar{y}\right)^2$$

Se $H_0$for verdadeira e, portanto, todas as observações provêm de uma mesma população, então SQ_{Tot} é o numerador do cálculo da variância, s^2, de todas as $N = ng$ observações. Pode-se mostrar que SQ_{Tot} (soma de quadrados total) é decomposta na soma de quadrados do erro dadas, respectivamente, por:

$$SQ_{Trat} = \sum_{i=1}^{g}\sum_{j=1}^{n}(\bar{y}_i - \bar{y})^2 = n\sum_{i=1}^{g}(\bar{y}_i - \bar{y})^2$$

$$SQ_{Erro} = \sum_{i=1}^{g}\sum_{j=1}^{n}\left(y_{ij} - \bar{y}_i\right)^2$$

Os g diferentes desvios de SQ_{Trat} são feitos em relação a única média amostral ($\bar{y}$). Por isso, dizemos que SQ_{Trat} tem $g - 1$ graus de liberdade. Enquanto os N desvios de SQ_{Erro} são feitos em relação a g média amostrais ($\bar{y}_i, i = 1, 2, \dots, g$) e, por isso, SQ_{Erro} tem $N - g$ graus de liberdade. A divisão das somas de quadrados pelos correspondentes graus de liberdade leva aos chamados quadrados médios. Assim:

$$QM_{Trat} = \frac{SQ_{Trat}}{gl_{Trat}} = \frac{n\sum_{i=1}^{g}(\bar{y}_i - \bar{y})^2}{g - 1}$$

$$QM_{Erro} = \frac{SQ_{Erro}}{gl_{Erro}} = \frac{\sum_{i=1}^{g}\sum_{j=1}^{n}\left(y_{ij} - \bar{y}_i\right)^2}{N - g}$$

onde $N = ng$.

Observe que QM_{Trat} é uma medida da variância entre as médias dos grupos, enquanto QM_{Erro} é uma medida da variância dentro dos grupos. Define-se a razão:

$$f = \frac{QM_{Trat}}{QM_{Erro}}$$

que pode ser interpretada como uma medida de discriminação entre os g grupos. Para testar a hipótese $H_0: \mu_1 = \mu_2 = \cdots = \mu_g$, usamos a distribuição F, com $gl = g - 1$ no numerador e $gl = N - g$ no denominador. Assim, estabelecido o nível de significância α, podemos obter tabelado o valor de f_c, que deixa área igual a α na cauda superior da distribuição. A regra de decisão é dada por:

se $f < f_c$, então aceita H_0;
se $f \geq f_c$, então rejeita H_0.

As somas de quadrados são mais facilmente calculadas, conforme as expressões apresentadas no quadro abaixo:

Fonte de variação	Somas de quadrados	gl	Quadrados médios	Razão f
Entre tratamentos	$SQ_{Trat} = \sum_{i=1}^{g} \frac{y_i^2}{n} - \frac{y_T^2}{N}$	$g-1$	$QM_{Trat} = \frac{SQ_{Trat}}{gl_{Trat}}$	$f = \frac{QM_{Trat}}{QM_{Erro}}$
Dentro de trat. (Erro)	$SQ_{Erro} = SQ_{Tot} - SQ_{Trat}$	$N-g$	$QM_{Erro} = \frac{SQ_{Erro}}{gl_{Erro}}$	
Total	$SQ_{Tot} = \sum_{i=1}^{g}\sum_{j=1}^{n} y_{ij}^2 - \frac{y_T^2}{N}$	$N-1$		

Exemplo: Considere o problema de comparar três tipos de rede de computadores, C1, C2 e C3, em termos do tempo médio de transmissão de pacotes de dados entre duas máquinas. Realizou-se um experimento com oito replicações com cada tipo de rede, aleatorizando a ordem dos 24 ensaios e mantendo fixos os demais fatores controláveis (ver resultados na tabela seguinte). Deseja-se testar as hipóteses:

H_0: os tempos esperados de transmissão são iguais para os três tipos de rede;

H_1: os tempos esperados de transmissão não são todos iguais (depende do tipo de rede).

Replicação	***Tipo de rede***		
	C1	**C2**	**C3**
1	7,2	7,8	6,3
2	9,3	8,2	6,0
3	8,7	7,1	5,3
4	8,9	8,6	5,1
5	7,6	8,7	6,2
6	7,2	8,2	5,2
7	8,8	7,1	7,2
8	8,0	7,8	6,8
Soma	65,7	63,5	48,1
Média	8,21	7,94	6,01

Soma global: $y_T = 177{,}3$

Soma de quadrados: $\sum_{i=1}^{g}\sum_{j=1}^{n} y_{ij}^2 = (7{,}2)^2 + (9{,}3)^2 + \cdots + (6{,}8)^2 = 1344{,}25$

$$SQ_{Trat} = \sum_{i=1}^{g} \frac{y_i^2}{n} - \frac{y_T^2}{N} = \frac{(65{,}7)^2 + (63{,}5)^2 + (48{,}1)^2}{8} - \frac{(177{,}3)^2}{24} = 22{,}99$$

$$SQ_{Tot} = \sum_{i=1}^{g}\sum_{j=1}^{n} y_{ij}^2 - \frac{y_T^2}{N} = 1344{,}25 - \frac{(177{,}3)^2}{24} = 34{,}45$$

$$SQ_{Erro} = SQ_{Tot} - SQ_{Trat} = 34{,}45 - 22{,}99 = 11{,}46$$

Resultando no quadro da ANOVA:

Fonte de variação	**SQ**	**gl**	**QM**	***f***
Entre grupos	22,99	2	11,50	21,07
Dentro dos grupos	11,46	21	0,55	

Total	34,45	23		

Adotando $\alpha = 0{,}05$, temos o valor crítico $f_c = 3{,}47$. Como o valor calculado ($f = 21{,}07$) é superior ao valor crítico, então o teste rejeita H_0, provando estatisticamente que há diferença entre os três tipos de rede, em termos do tempo médio de transmissão.

- **Estimação das médias**

As médias aritméticas de cada grupo servem como uma estimativa pontual da resposta esperada de cada tratamento. Podemos, também, usar a abordagem de intervalos de confiança para efetuarmos uma inferência mais completa. O erro padrão pode ser estimado através da variância conjunta das g amostras. Daí, o intervalo de confiança para o valor esperado da resposta sob o i-esimo tratamento é dado por:

$$IC(\mu_i, 1-\alpha) = \bar{y}_i \pm t_{1-\alpha} \sqrt{\frac{QM_{erro}}{n}}$$

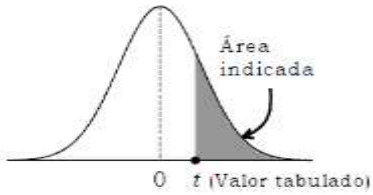
onde $t_{1-\alpha}$, é tabelado 4, em função do nível de confiança $(1-\alpha)$ estabelecido e dos graus de liberdade: $gl = g(n - 1)$.

Exemplo (continuação): Usando nível de confiança de 95% e $gl = N - g = 24 - 3 = 21$, temos o valor tabelado $t = 2{,}08$. Então, para a rede C1, temos:

$$IC(\mu_1, 95\%) = 8{,}21 \pm 2{,}08 \sqrt{\frac{0{,}55}{8}} = 8{,}21 \pm 0{,}55$$

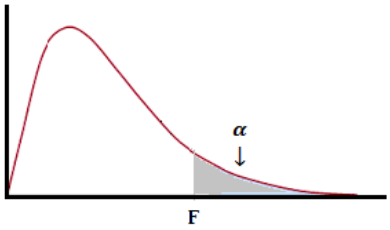
Construindo intervalos de confiança para os demais tipos de rede, sugere-se que as redes C1 e C2 podem ser consideradas com o mesmo tempo esperado de transmissão, pois seus intervalos de confiança se sobrepõem, indicando que a diferença entre as médias aritméticas pode ser meramente casual. Já a rede C3 parece ser a causa do teste F da ANOVA ter rejeitado H_0, pois apresenta um intervalo de confiança para a média não sobreposto com os outros intervalos de confiança, evidenciando produzir tempo esperado de transmissão menor.

Distribuição **t de *Student***



	Área da cauda superior								
gl	**0,2500**	**0,1000**	**0,0500**	**0,0250**	**0,0100**	**0,0050**	**0,0025**	**0,0010**	**0,0005**
1	1,000	3,078	6,314	12,710	31,820	63,660	127,300	318,300	636,600
2	0,816	1,886	2,920	4,303	6,965	9,925	14,090	22,330	31,600
3	0,765	1,638	2,353	3,182	4,541	5,841	7,453	10,210	12,920
4	0,741	1,533	2,132	2,776	3,747	4,604	5,598	7,173	8,610
5	0,727	1,476	2,015	2,571	3,365	4,032	4,773	5,894	6,869
6	0,718	1,440	1,943	2,447	3,143	3,707	4,317	5,208	5,959
7	0,711	1,415	1,895	2,365	2,998	3,499	4,029	4,785	5,408
8	0,706	1,397	1,860	2,306	2,896	3,355	3,833	4,501	5,041
9	0,703	1,383	1,833	2,262	2,821	3,250	3,690	4,297	4,781
10	0,700	1,372	1,812	2,228	2,764	3,169	3,581	4,144	4,587
11	0,697	1,363	1,796	2,201	2,718	3,106	3,497	4,025	4,437
12	0,695	1,356	1,782	2,179	2,681	3,055	3,428	3,930	4,318
13	0,694	1,350	1,771	2,160	2,650	3,012	3,372	3,852	4,221
14	0,692	1,345	1,761	2,145	2,624	2,977	3,326	3,787	4,140
15	0,691	1,341	1,753	2,131	2,602	2,947	3,286	3,733	4,073
16	0,690	1,337	1,746	2,120	2,583	2,921	3,252	3,686	4,015
17	0,689	1,333	1,740	2,110	2,567	2,898	3,222	3,646	3,965
18	0,688	1,330	1,734	2,101	2,552	2,878	3,197	3,610	3,922
19	0,688	1,328	1,729	2,093	2,539	2,861	3,174	3,579	3,883
20	0,687	1,325	1,725	2,086	2,528	2,845	3,153	3,552	3,850
21	0,686	1,323	1,721	2,080	2,518	2,831	3,135	3,527	3,819
22	0,686	1,321	1,717	2,074	2,508	2,819	3,119	3,505	3,792
23	0,685	1,319	1,714	2,069	2,500	2,807	3,104	3,485	3,768
24	0,685	1,318	1,711	2,064	2,492	2,797	3,091	3,467	3,745
25	0,684	1,316	1,708	2,060	2,485	2,787	3,078	3,450	3,725
26	0,684	1,315	1,706	2,056	2,479	2,779	3,067	3,435	3,707
27	0,684	1,314	1,703	2,052	2,473	2,771	3,057	3,421	3,689
28	0,683	1,313	1,701	2,048	2,467	2,763	3,047	3,408	3,674
29	0,683	1,311	1,699	2,045	2,462	2,756	3,038	3,396	3,660
30	0,683	1,310	1,697	2,042	2,457	2,750	3,030	3,385	3,646
35	0,682	1,306	1,690	2,030	2,438	2,724	2,996	3,340	3,591
40	0,681	1,303	1,684	2,021	2,423	2,704	2,971	3,307	3,551
45	0,680	1,301	1,679	2,014	2,412	2,690	2,952	3,281	3,520
50	0,679	1,299	1,676	2,009	2,403	2,678	2,937	3,261	3,496
z	0,674	1,282	1,645	1,960	2,326	2,576	2,807	3,090	3,291

Distribuição **F de Snedecor**
($\alpha = 1\%$)



gl denom.	graus de liberdade no numerador									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	4052,20	4999,30	5403,50	5624,30	5764,00	5859,00	5928,30	5981,00	6022,40	6055,90
2	98,50	99,00	99,16	99,25	99,30	99,33	99,36	99,38	99,39	99,40
3	34,12	30,82	29,46	28,71	28,24	27,91	27,67	27,49	27,35	27,23
4	21,20	18,00	16,69	15,98	15,52	15,21	14,98	14,80	14,66	14,55
5	16,26	13,27	12,06	11,39	10,97	10,67	10,46	10,29	10,16	10,05
6	13,75	10,93	9,78	9,15	8,75	8,47	8,26	8,10	7,98	7,87
7	12,25	9,55	8,45	7,85	7,46	7,19	6,99	6,84	6,72	6,62
8	11,26	8,65	7,59	7,01	6,63	6,37	6,18	6,03	5,91	5,81
9	10,56	8,02	6,99	6,42	6,06	5,80	5,61	5,47	5,35	5,26
10	10,04	7,56	6,55	5,99	5,64	5,39	5,20	5,06	4,94	4,85
11	9,65	7,21	6,22	5,67	5,32	5,07	4,89	4,74	4,63	4,54
12	9,33	6,93	5,95	5,41	5,06	4,82	4,64	4,50	4,39	4,30
13	9,07	6,70	5,74	5,21	4,86	4,62	4,44	4,30	4,19	4,10
14	8,86	6,52	5,56	5,04	4,70	4,46	4,28	4,14	4,03	3,94
15	8,68	6,36	5,42	4,89	4,56	4,32	4,14	4,00	3,90	3,81
16	8,53	6,23	5,29	4,77	4,44	4,20	4,03	3,89	3,78	3,69
17	8,40	6,11	5,19	4,67	4,34	4,10	3,93	3,79	3,68	3,59
18	8,29	6,01	5,09	4,58	4,25	4,02	3,84	3,71	3,60	3,51
19	8,19	5,93	5,01	4,50	4,17	3,94	3,77	3,63	3,52	3,43
20	8,10	5,85	4,94	4,43	4,10	3,87	3,70	3,56	3,46	3,37
21	8,02	5,78	4,87	4,37	4,04	3,81	3,64	3,51	3,40	3,31
22	7,95	5,72	4,82	4,31	3,99	3,76	3,59	3,45	3,35	3,26
23	7,88	5,66	4,77	4,26	3,94	3,71	3,54	3,41	3,30	3,21
24	7,82	5,61	4,72	4,22	3,90	3,67	3,50	3,36	3,26	3,17
25	7,77	5,57	4,68	4,18	3,86	3,63	3,46	3,32	3,22	3,13
26	7,72	5,53	4,64	4,14	3,82	3,59	3,42	3,29	3,18	3,09
27	7,68	5,49	4,60	4,11	3,79	3,56	3,39	3,26	3,15	3,06
28	7,64	5,45	4,57	4,07	3,75	3,53	3,36	3,23	3,12	3,03
29	7,60	5,42	4,54	4,05	3,73	3,50	3,33	3,20	3,09	3,01
30	7,56	5,39	4,51	4,02	3,70	3,47	3,31	3,17	3,07	2,98
40	7,31	5,18	4,31	3,83	3,51	3,29	3,12	2,99	2,89	2,80
50	7,17	5,06	4,20	3,72	3,41	3,19	3,02	2,89	2,79	2,70
60	7,08	4,98	4,13	3,65	3,34	3,12	2,95	2,82	2,72	2,63
80	6,96	4,88	4,04	3,56	3,26	3,04	2,87	2,74	2,64	2,55
100	6,90	4,82	3,98	3,51	3,21	2,99	2,82	2,69	2,59	2,50

Distribuição **F de Snedecor** ($\alpha = 2,5\%$)										
gl	graus de liberdade no numerador									
denom.	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	647,80	799,50	864,20	899,60	921,80	937,10	948,20	956,60	963,30	968,60
2	38,51	39,00	39,17	39,25	39,30	39,33	39,36	39,37	39,39	39,40
3	17,44	16,04	15,44	15,10	14,89	14,74	14,62	14,54	14,47	14,42
4	12,22	10,65	9,98	9,60	9,36	9,20	9,07	8,98	8,91	8,84
5	10,01	8,43	7,76	7,39	7,15	6,98	6,85	6,76	6,68	6,62
6	8,81	7,26	6,60	6,23	5,99	5,82	5,70	5,60	5,52	5,46
7	8,07	6,54	5,89	5,52	5,29	5,12	5,00	4,90	4,82	4,76
8	7,57	6,06	5,42	5,05	4,82	4,65	4,53	4,43	4,36	4,30
9	7,21	5,72	5,08	4,72	4,48	4,32	4,20	4,10	4,03	3,96
10	6,94	5,46	4,83	4,47	4,24	4,07	3,95	3,86	3,78	3,72
11	6,72	5,26	4,63	4,28	4,04	3,88	3,76	3,66	3,59	3,53
12	6,55	5,10	4,47	4,12	3,89	3,73	3,61	3,51	3,44	3,37
13	6,41	4,97	4,35	4,00	3,77	3,60	3,48	3,39	3,31	3,25
14	6,30	4,86	4,24	3,89	3,66	3,50	3,38	3,29	3,21	3,15
15	6,20	4,77	4,15	3,80	3,58	3,42	3,29	3,20	3,12	3,06
16	6,12	4,69	4,08	3,73	3,50	3,34	3,22	3,13	3,05	2,99
17	6,04	4,62	4,01	3,67	3,44	3,28	3,16	3,06	2,99	2,92
18	5,98	4,56	3,95	3,61	3,38	3,22	3,10	3,01	2,93	2,87
19	5,92	4,51	3,90	3,56	3,33	3,17	3,05	2,96	2,88	2,82
20	5,87	4,46	3,86	3,52	3,29	3,13	3,01	2,91	2,84	2,77
21	5,83	4,42	3,82	3,48	3,25	3,09	2,97	2,87	2,80	2,74
22	5,79	4,38	3,78	3,44	3,22	3,06	2,93	2,84	2,76	2,70
23	5,75	4,35	3,75	3,41	3,18	3,02	2,90	2,81	2,73	2,67
24	5,72	4,32	3,72	3,38	3,16	3,00	2,87	2,78	2,70	2,64
25	5,69	4,29	3,69	3,35	3,13	2,97	2,85	2,75	2,68	2,61
26	5,66	4,27	3,67	3,33	3,11	2,95	2,82	2,73	2,65	2,59
27	5,63	4,24	3,65	3,31	3,08	2,92	2,80	2,71	2,63	2,57
28	5,61	4,22	3,63	3,29	3,06	2,90	2,78	2,69	2,61	2,55
29	5,59	4,20	3,61	3,27	3,04	2,88	2,76	2,67	2,59	2,53
30	5,57	4,18	3,59	3,25	3,03	2,87	2,75	2,65	2,58	2,51
40	5,42	4,05	3,46	3,13	2,90	2,74	2,62	2,53	2,45	2,39
50	5,34	3,98	3,39	3,05	2,83	2,67	2,55	2,46	2,38	2,32
60	5,29	3,93	3,34	3,01	2,79	2,63	2,51	2,41	2,33	2,27
80	5,22	3,86	3,28	2,95	2,73	2,57	2,45	2,36	2,28	2,21
100	5,18	3,83	3,25	2,92	2,70	2,54	2,42	2,32	2,24	2,18

Distribuição **F de Snedecor** ($\alpha = 5\%$)										
gl	graus de liberdade no numerador									
denom.	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	161,40	199,50	215,70	224,60	230,20	234,00	236,80	238,90	240,50	241,90
2	18,51	19,00	19,16	19,25	19,30	19,33	19,35	19,37	19,39	19,40
3	10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14
10	4,97	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98
11	4,84	3,98	3,59	3,36	3,20	3,10	3,01	2,95	2,90	2,85
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49
17	4,45	3,59	3,20	2,97	2,81	2,70	2,61	2,55	2,49	2,45
18	4,41	3,56	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35
21	4,33	3,47	3,07	2,84	2,69	2,57	2,49	2,42	2,37	2,32
22	4,30	3,44	3,05	2,82	2,66	2,55	2,46	2,40	2,34	2,30
23	4,28	3,42	3,03	2,80	2,64	2,53	2,44	2,38	2,32	2,28
24	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36	2,30	2,26
25	4,24	3,39	2,99	2,76	2,60	2,49	2,41	2,34	2,28	2,24
26	4,23	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,27	2,22
27	4,21	3,35	2,96	2,73	2,57	2,46	2,37	2,31	2,25	2,20
28	4,20	3,34	2,95	2,71	2,56	2,45	2,36	2,29	2,24	2,19
29	4,18	3,33	2,93	2,70	2,55	2,43	2,35	2,28	2,22	2,18
30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21	2,17
40	4,09	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,08
50	4,03	3,18	2,79	2,56	2,40	2,29	2,20	2,13	2,07	2,03
60	4,00	3,15	2,76	2,53	2,37	2,25	2,17	2,10	2,04	1,99
80	3,96	3,11	2,72	2,49	2,33	2,21	2,13	2,06	2,00	1,95
100	3,94	3,09	2,70	2,46	2,31	2,19	2,10	2,03	1,98	1,93

Distribuição **F de Snedecor** ($\alpha = 10\%$)										
gl denom.	graus de liberdade no numerador									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	39,86	49,50	53,59	55,83	57,24	58,20	58,91	59,44	59,86	60,20
2	8,53	9,00	9,16	9,24	9,29	9,33	9,35	9,37	9,38	9,39
3	5,54	5,46	5,39	5,34	5,31	5,29	5,27	5,25	5,24	5,23
4	4,55	4,33	4,19	4,11	4,05	4,01	3,98	3,96	3,94	3,92
5	4,06	3,78	3,62	3,52	3,45	3,41	3,37	3,34	3,32	3,30
6	3,78	3,46	3,29	3,18	3,11	3,06	3,01	2,98	2,96	2,94
7	3,59	3,26	3,07	2,96	2,88	2,83	2,79	2,75	2,73	2,70
8	3,46	3,11	2,92	2,81	2,73	2,67	2,62	2,59	2,56	2,54
9	3,36	3,01	2,81	2,69	2,61	2,55	2,51	2,47	2,44	2,42
10	3,29	2,92	2,73	2,61	2,52	2,46	2,41	2,38	2,35	2,32
11	3,23	2,86	2,66	2,54	2,45	2,39	2,34	2,30	2,27	2,25
12	3,18	2,81	2,61	2,48	2,39	2,33	2,28	2,25	2,21	2,19
13	3,14	2,76	2,56	2,43	2,35	2,28	2,23	2,20	2,16	2,14
14	3,10	2,73	2,52	2,40	2,31	2,24	2,19	2,15	2,12	2,10
15	3,07	2,70	2,49	2,36	2,27	2,21	2,16	2,12	2,09	2,06
16	3,05	2,67	2,46	2,33	2,24	2,18	2,13	2,09	2,06	2,03
17	3,03	2,65	2,44	2,31	2,22	2,15	2,10	2,06	2,03	2,00
18	3,01	2,62	2,42	2,29	2,20	2,13	2,08	2,04	2,01	1,98
19	2,99	2,61	2,40	2,27	2,18	2,11	2,06	2,02	1,98	1,96
20	2,98	2,59	2,38	2,25	2,16	2,09	2,04	2,00	1,97	1,94
21	2,96	2,58	2,37	2,23	2,14	2,08	2,02	1,98	1,95	1,92
22	2,95	2,56	2,35	2,22	2,13	2,06	2,01	1,97	1,93	1,90
23	2,94	2,55	2,34	2,21	2,12	2,05	2,00	1,95	1,92	1,89
24	2,93	2,54	2,33	2,20	2,10	2,04	1,98	1,94	1,91	1,88
25	2,92	2,53	2,32	2,18	2,09	2,02	1,97	1,93	1,90	1,87
26	2,91	2,52	2,31	2,17	2,08	2,01	1,96	1,92	1,88	1,86
27	2,90	2,51	2,30	2,17	2,07	2,01	1,95	1,91	1,87	1,85
28	2,89	2,50	2,29	2,16	2,06	2,00	1,94	1,90	1,87	1,84
29	2,89	2,50	2,28	2,15	2,06	1,99	1,94	1,89	1,86	1,83
30	2,88	2,49	2,28	2,14	2,05	1,98	1,93	1,88	1,85	1,82
40	2,84	2,44	2,23	2,09	2,00	1,93	1,87	1,83	1,79	1,76
50	2,81	2,41	2,20	2,06	1,97	1,90	1,84	1,80	1,76	1,73
60	2,79	2,39	2,18	2,04	1,95	1,88	1,82	1,78	1,74	1,71
80	2,77	2,37	2,15	2,02	1,92	1,85	1,79	1,75	1,71	1,68
100	2,76	2,36	2,14	2,00	1,91	1,83	1,78	1,73	1,70	1,66