

# 10

## UNIDADE

# Testes de Hipóteses



Nesta Unidade você conhecerá e aplicará os conceitos de Testes de Hipóteses, especialmente para média de uma variável quantitativa, proporção de um dos valores de uma variável quantitativa, e associação entre duas variáveis qualitativas. Você aprenderá também qual é a importância de tais conceitos para o dia-a-dia do administrador.



Caro estudante, você viu anteriormente que uma determinada população pode ser descrita através de um modelo, que apresenta características e parâmetros. Muitas vezes esses parâmetros são desconhecidos e há interesse em estimá-los para obter um melhor conhecimento sobre a população: retira-se então uma amostra aleatória da população e através das técnicas de Estimação de Parâmetros (Unidade 9) procura-se obter uma estimativa de algum parâmetro de interesse, e associamos uma probabilidade de que a estimativa esteja correta. Nesta última e importantíssima Unidade veremos que a Estimação de Parâmetros é uma subdivisão da Inferência Estatística (que consiste em fazer afirmações probabilísticas sobre o modelo da população a partir de uma amostra aleatória desta população), a outra grande subdivisão constitui os **Testes de Hipóteses**. Vamos saber mais!

**Testes paramétricos** – testes de hipóteses sobre parâmetros do modelo da variável sob análise Fonte: elaborado pelo autor.

**Testes não paramétricos** – testes de hipóteses sobre outros aspectos do modelo da variável sob análise ou alternativas aos testes paramétricos quando as condições para uso destes não forem satisfeitas. Fonte: elaborado pelo autor.

**C**ontrariamente à Estimação de Parâmetros os Testes de Hipóteses permitem fazer inferências sobre outras características do modelo da população além dos parâmetros, como, por exemplo, a forma do modelo da população. Quando os Testes são feitos sobre os parâmetros da população são chamados de Testes Paramétricos, e quando são feitos sobre outras características são chamados de Testes Não Paramétricos. Não obstante vamos nos restringir aos Testes Paramétricos: de uma média de uma variável quantitativa e de uma proporção de um dos valores de uma variável qualitativa.

Vimos que uma determinada população pode ser descrita através de um modelo

### *Tô a fim de saber*

Na realidade a denominação correta deveria ser “Testes dependentes de distribuição de referência” (porque para fazer inferências sobre os parâmetros devemos supor que o modelo probabilístico populacional é normal, por exemplo, ou que a distribuição amostral do parâmetro pode ser aproximada por uma normal). Já os não paramétricos deveriam ser “Testes livres de distribuição” (porque os Testes Não Paramétricos não exigem que os dados tenham uma aderência a certo modelo).

A você estudante interessado em Testes Não Paramétricos recomendamos a seguinte referência: SIEGEL, Sidney. **Estatística Não Paramétrica (para as Ciências do Comportamento)**. McGraw-Hill: São Paulo, 1975. É uma boa referência no assunto, em português.

probabilístico, que apresenta características e parâmetros. Muitas vezes estes parâmetros são desconhecidos e há interesse em estimá-los para obter um melhor conhecimento sobre a população: retira-se então uma amostra aleatória da população e através das técnicas de **Estimação de Parâmetros** (Unidade 9) procura-se obter uma estimativa de algum parâmetro de interesse, e associamos uma probabilidade de que a estimativa esteja correta. A Estimação de Parâmetros é uma subdivisão da Inferência Estatística (que consiste em fazer afirmações probabilísticas sobre o modelo probabilístico da população a partir de uma amostra aleatória desta população), a outra grande subdivisão constitui os **Testes de Hipóteses**.

Contrariamente à Estimação de Parâmetros os Testes de Hipóteses permitem fazer inferências sobre outras características do modelo probabilístico da população além dos parâmetros (como por exemplo a forma do modelo probabilístico da população). Quando os Testes são feitos sobre os parâmetros da população são chamados de Testes Paramétricos, e quando são feitos sobre outras características são chamados de **Testes Não Paramétricos**. Não obstante vamos nos restringir aos Testes Paramétricos.

Imagine-se que um determinado pesquisador está interessado em alguma característica de uma população. Devido a estudos prévios, ou simplesmente por bom senso (melhor ponto de partida para o estudo) ele estabelece que a característica terá um determinado comportamento. Formula então uma hipótese estatística sobre a característica da população, e essa hipótese é aceita como válida até prova estatística em contrário.

Para testar a hipótese é coletada uma amostra aleatória representativa da população, sendo calculadas as estatísticas necessárias para o teste. Naturalmente, devido ao fato de ser utilizada uma amostra aleatória, haverá diferenças entre o que se esperava, sob a condição da hipótese verdadeira, e o que realmente foi obtido na amostra. A questão a ser respondida é: as diferenças são significativas o bastante para que a hipótese estatística estabelecida seja rejeitada? Esta não é uma pergunta simples de responder: dependerá do que está sob teste (que parâmetro, por exemplo), da confiabilidade desejada para o resultado, entre outros. Basicamente, porém, será necessário comparar as diferenças com uma referência, a distribuição amostral de um parâmetro, por exemplo, que supõe que a hipótese sob teste é verdadeira: a comparação costuma ser feita através de uma estatística de teste que envolve os valores da amostra e os valores sob teste.

A tomada de decisão é feita da seguinte forma:

- se a diferença entre o que foi observado na amostra e o que era esperado (sob a condição da hipótese verdadeira) não for **significativa** a hipótese será **aceita**; e
- se a diferença entre o que foi observado na amostra e o que era esperado (sob a condição da hipótese verdadeira) for significativa a hipótese será **rejeitada**.

O valor a partir do qual a diferença será considerada significativa será determinado pelo Nível de Significância do teste. O Nível de Significância geralmente é fixado pelo pesquisador, muitas vezes de forma arbitrária, e também será a probabilidade de erro do Teste de Hipóteses: a probabilidade de cometer um erro no teste, rejeitando uma hipótese válida. Como a decisão do teste é tomada a partir dos dados de uma amostra aleatória da população há sempre a probabilidade de estar cometendo um erro, mas com a utilização de métodos estatísticos é possível calcular o valor desta probabilidade.

O Nível de Significância é uma probabilidade, portanto é, um número real que varia de 0 a 1 (0 a 100%), e como é a probabilidade de se cometer um erro no teste é interessante que seja o mais próximo possível de zero: valores típicos são 5%, 10%, 1% e até menores dependendo do problema sob análise. Contudo, não é possível usar um Nível de Significância igual a zero porque devido ao uso de uma amostra aleatória sempre haverá chance de erro, a não ser que a amostra fosse do tamanho da população. O complementar do Nível de Significância é chamado de **Nível de Confiança**, pois ele indica a confiabilidade do resultado obtido, a probabilidade de que a decisão tomada esteja correta

Usando outros métodos (empíricos) não há como ter ideia da chance de erro (pode ser um erro de 0% ou de 5.000%...).

**Nível de Significância**  
– probabilidade arbitrada pelo pesquisador, valor máximo de erro admissível para rejeitar a hipótese nula sendo ela verdadeira, espera-se que seja um valor baixo, de no máximo 10%.  
Fonte: Barbetta, Reis e Bornia (2008); Moore, McCabe, Duckworth e Sclove (2006).

Você deve estar lembrado desses dois conceitos de Estimção de Parâmetros: Nível de Confiança era a probabilidade de que o Intervalo de Confiança contivesse o valor real do parâmetro, e Nível de Significância, complementar daquele, era a probabilidade de que o Intervalo não contivesse o parâmetro, em suma a probabilidade da Estimção estar correta ou não, respectivamente.

## Tipos de Hipóteses

Para realizar um Teste de Hipóteses é necessário definir (enunciar) duas Hipóteses Estatísticas complementares (que abrangem todos os resultados possíveis): a chamada **Hipótese Nula** (denotada por  $H_0$ ) e a **Hipótese Alternativa** (denotada por  $H_1$  ou  $H_a$ ). Enunciar as hipóteses é o primeiro e possivelmente mais importante passo de um Teste de Hipóteses, pois todo o procedimento dependerá dele.

A Hipótese Nula ( $H_0$ ) é a hipótese estatística aceita como verdadeira até prova estatística em contrário: pode ser o ponto de partida mais adequado para o estudo, ou exatamente o contrário do que o pesquisador quer provar (ou o contrário daquilo que o preocupa).

A Hipótese Alternativa ( $H_1$ ), que será uma hipótese complementar de  $H_0$ , fornecerá uma alternativa à hipótese nula: muitas vezes é justamente o que o pesquisador quer provar (ou o que o preocupa).

Quando as hipóteses são formuladas sobre os parâmetros do modelo probabilístico da população o Teste de Hipóteses é chamado de Paramétrico. Quando as hipóteses são formuladas sobre outras características do modelo o Teste é chamado de Não Paramétrico.

A decisão do teste consiste em aceitar ou rejeitar a Hipótese Nula ( $H_0$ ): aceitar-se ou não a hipótese até então considerada verdadeira.

É importante ter a noção exata do que significa aceitar ou rejeitar a Hipótese Nula ( $H_0$ ). A decisão é tomada sobre esta hipótese e não sobre a Hipótese Alternativa porque é a Hipótese Nula é que considerada verdadeira (até prova em contrário). Quando se aceita a Hipótese Nula significa que não há provas suficientes para rejeitá-la. Já quando a decisão é por rejeitar a Hipótese Nula há evidências suficientes de que as diferenças obtidas (entre o que era esperado e o que foi observado na amostra) não ocorreram por acaso. Usando uma analogia com o direito dos EUA, aceitar  $H_0$  seria comparável a um veredito de não culpado “*not guilty*”, ou seja, não há provas suficientes para condenar o réu. Por outro lado rejeitar  $H_0$  seria comparável a um veredito de culpado “*guilty*”, ou seja, as provas reunidas são suficientes para condenar o réu. O Nível de Significância será a probabilidade assumida de **Rejeitar  $H_0$  sendo  $H_0$  verdadeira**.

## Tipos de Testes Paramétricos

A formulação das hipóteses é o ponto inicial do problema, e deve depender única e exclusivamente das conclusões que se pretende obter com o teste. A formulação da hipótese alternativa determinará o tipo de teste: se Unilateral ou Bilateral.

Se a formulação da hipótese alternativa indicar que o parâmetro é maior ou menor do que o valor de teste (valor considerado verdadeiro até prova em contrário) o teste será **Unilateral**: somente há interesse se as diferenças entre os dados da amostra e o valor de teste forem em uma determinada direção. Se a formulação da hipótese alternativa indicar que o parâmetro é diferente do valor de teste o teste será **Bilateral**: há interesse nas diferenças em qualquer direção. As hipóteses então seriam:

### Testes Unilaterais

$H_0$ : parâmetro = valor de teste.

$H_1$ : parâmetro < valor de teste.

$H_0$ : parâmetro = valor de teste.

$H_1$ : parâmetro > valor de teste.

### Testes Bilaterais

$H_0$ : parâmetro = valor de teste.

$H_1$ : parâmetro  $\neq$  valor de teste.

A escolha do tipo de teste dependerá das condições do problema sob estudo, sejam as três situações abaixo:

- Um novo protocolo de atendimento foi implementado no Banco RMG, visando reduzir o tempo que as pessoas passam na fila do caixa. O protocolo será considerado satisfatório se a média do tempo de fila for menor do que 30 minutos. Um teste **Unilateral** seria o adequado.
- Cerca de 2.000 formulários de pedidos de compra estão sendo analisados. Os clientes podem ficar insatisfeitos se houver erros nos formulários. Neste caso, admite-se que a proporção máxima de formulários com erros seja de 5%.

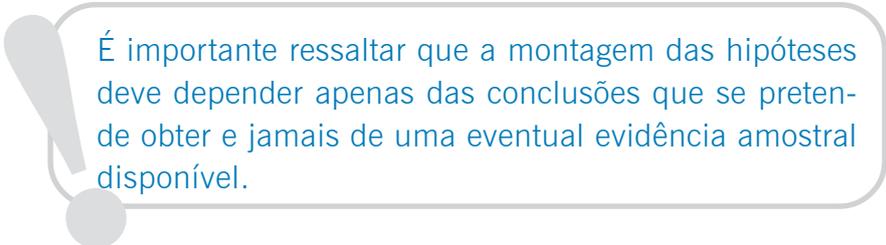
**Teste unilateral** – teste no qual a região de rejeição da hipótese nula está concentrada em apenas um dos lados da distribuição amostral da variável de teste. Fonte: Barbetta, Reis e Bornia (2008).

**Teste bilateral** – teste no qual a região de rejeição da hipótese nula está dividida em duas partes, em cada um dos lados da distribuição amostral da variável de teste. Fonte: Barbetta, Reis e Bornia (2008).

Ou seja, um valor maior do que 5% causaria problemas. Um teste **Unilateral** seria o adequado.

- c) Uma peça automotiva precisa ter 100 mm de diâmetro, exatamente. Neste caso, a dimensão não pode ser maior ou menor do que 100 mm (em outras palavras não pode ser diferente de 100 mm), pois isso indicará que a peça não está de acordo com as especificações. Um teste **Bilateral** seria o adequado.

Após definir as hipóteses é coletada uma amostra aleatória da população para testá-las.



É importante ressaltar que a montagem das hipóteses deve depender apenas das conclusões que se pretende obter e jamais de uma eventual evidência amostral disponível.

A decisão de aceitar ou rejeitar  $H_0$  dependerá das **regiões de aceitação e rejeição de  $H_0$** , que por sua vez dependem dos seguintes fatores:

- do parâmetro sob teste (e da estatística ou variável de teste usada para testá-lo).
- do tipo de teste, se Unilateral ou Bilateral.
- do valor de teste (valor do parâmetro considerado verdadeiro até prova em contrário).
- do Nível de Significância ( $\sigma$ ) ou Nível de Confiança ( $1 - \sigma$ ) adotado.
- de um valor crítico da estatística ou variável de teste a partir do qual a hipótese será rejeitada, e esse valor dependerá por sua vez do Nível de Significância, do tipo de teste e da **Distribuição Amostral** do parâmetro.

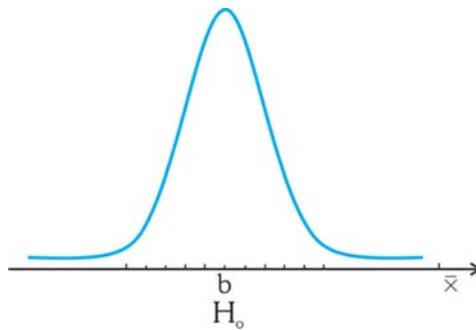
A **Região de Aceitação de  $H_0$**  será a faixa de valores da estatística (ou da variável de teste) associada ao parâmetro em que as diferenças entre o que foi observado na amostra e o que era esperado não são significativas.

A **Região de Rejeição de  $H_0$**  será a faixa de valores da estatística (ou da variável de teste) associada ao parâmetro em que as diferenças entre o que foi observado na amostra e o que era esperado **são significativas**.

Essa abordagem é chamada de abordagem clássica dos testes de hipóteses. Há também a do valor-p, muito usada por programas computacionais. Mas, neste texto, vamos usar apenas a clássica por considerá-la mais clara.

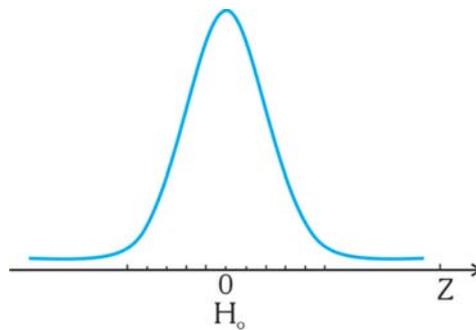
Para entender melhor os conceitos acima, observe a situação a seguir:

Há interesse em realizar um teste de hipóteses sobre o comprimento médio de uma das dimensões de uma peça mecânica. O valor nominal da média (aceito como verdadeiro até prova em contrário) é igual a “**b**” (valor genérico),  $H_0: \mu = \mathbf{b}$ . Supondo que a distribuição amostral do parâmetro (distribuição de  $\bar{x}$ ) seja normal, e será centrada em **b**: é possível fazer a conversão para a distribuição normal padrão (média zero e desvio padrão 1, variável Z) (Figuras 73 e 74).



$H_0: \mu = \mathbf{b}$

Figura 73: Hipótese nula: média populacional = b.  
Fonte: elaborada pelo autor.



$H_0: \mu = \mathbf{0}$

Figura 74: Hipótese nula normal padrão.  
Fonte: elaborada pelo autor.

O valor de **b** (média da dimensão e média de  $\bar{x}$ ) corresponde a zero, média da variável **Z**. Dependendo da formulação da Hipótese Alternativa haveria diferentes Regiões de Rejeição de **H<sub>0</sub>**.

Se a Hipótese Alternativa fosse **H<sub>1</sub>:  $\mu < b$**  (**H<sub>1</sub>:  $\mu < 0$** ), ou seja, se o teste fosse Unilateral à esquerda a Região de Rejeição de **H<sub>0</sub>** seria (Figuras 75 e 76):

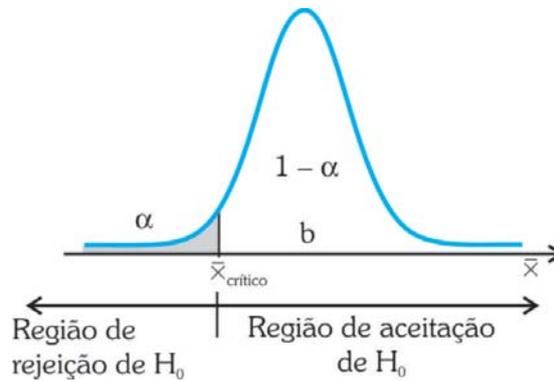


Figura 75 – **H<sub>1</sub>:  $\mu < b$** .

Fonte: elaborada pelo autor.

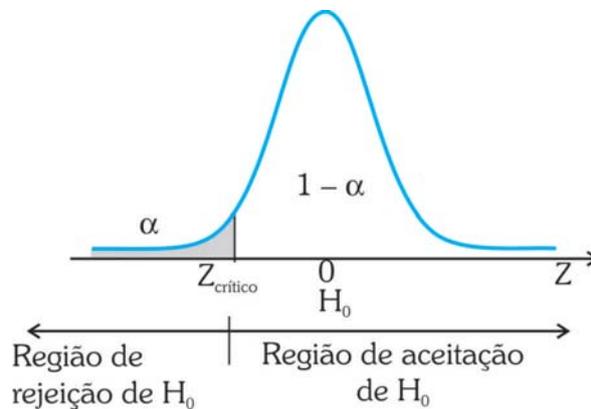


Figura 76 – **H<sub>1</sub>:  $\mu < 0$** .

Fonte: elaborada pelo autor.

Observe que há um valor crítico de  $\bar{x}$ : abaixo dele a Hipótese Nula será rejeitada, acima será aceita. A determinação do valor é feita com base no Nível de Significância, a área abaixo da curva normal até o valor crítico de  $\bar{x}$ . Geralmente obtém-se o valor crítico da variável de teste (**Z** neste caso, na segunda figura) através de uma tabela, que corresponde ao valor crítico de  $\bar{x}$ , faz-se a transformação de variáveis  $Z = \frac{(\bar{x} - \mu_0)}{\sigma}$  e obtém-se o valor crítico de  $\bar{x}$ .  $\mu_0$  é o valor

sob teste ( $\mathbf{b}$  no exemplo) e  $\sigma$  é um desvio padrão (cujo valor será explicitado posteriormente).

A decisão será tomada comparando valor da média amostral  $\bar{x}$  com o valor crítico desta mesma média: se for menor do que o valor crítico  $\bar{x}_{\text{crítico}}$ , ou seja, está na região de **Rejeição de  $H_0$**  então rejeita-se a Hipótese Nula. É muito comum também tomar a decisão comparando o valor da variável de teste ( $\mathbf{Z}$  neste caso), obtido com base nos dados da amostra, com o valor crítico  $Z_{\text{crítico}}$  desta mesma variável (obtido de uma tabela ou programa computacional): se for menor do que o valor crítico rejeita-se a Hipótese Nula. Observe que o valor do Nível de Significância  $\alpha$  é colocado na curva referente à Hipótese Nula, porque é esta que é aceita como válida até prova em contrário. Observe também que a faixa de valores da região de Rejeição pertence à curva da Hipótese Nula, assim o valor  $\alpha$  é a probabilidade de Rejeitar  $H_0$  sendo  $H_0$  verdadeira.

Neste ponto é importante ressaltar um ponto que costuma passar despercebido. Se a decisão for tomada com base na variável de teste ( $\mathbf{Z}$ , por exemplo) é interessante notar que, como o teste é Unilateral à esquerda o valor  $Z_{\text{crítico}}$  será NEGATIVO, uma vez que a região de Rejeição de  $H_0$  está à ESQUERDA de 0 (menor do que zero). No teste Unilateral à direita, que veremos a seguir, o valor de  $Z_{\text{crítico}}$  será positivo, pois a região de Rejeição de  $H_0$  estará à DIREITA de 0 (maior do que zero). Se por exemplo o Nível de Significância fosse de 5% (0,05) o valor de  $Z_{\text{crítico}}$  para o teste Unilateral à esquerda seria  $-1,645$ . Se houvesse interesse em obter o valor de  $\bar{x}_{\text{crítico}}$  correspondente bastaria usar a expressão  $Z = (\bar{x} - \mu_0)/\sigma$  substituindo  $\mathbf{Z}$  por  $-1,645$ .

Se a Hipótese Alternativa fosse  $H_1: \mu > \mathbf{b}$  ( $H_1: \mu > \mathbf{0}$ ), ou seja, se o teste fosse Unilateral à direita a Região de Rejeição de  $H_0$  seria (Figuras 77 e 78):

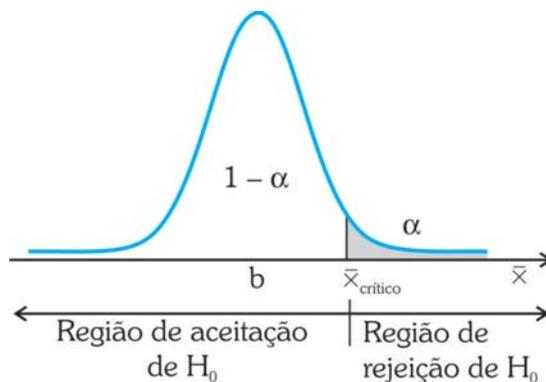


Figura 77 –  $H_1: \mu > \mathbf{b}$ .

Fonte: Elaborada pelo autor.

Probabilidade de tomar uma decisão errada fixada pelo pesquisador.

O sinal correto é importante para que o valor de coerente com a posição da região de Rejeição de  $H_0$ .

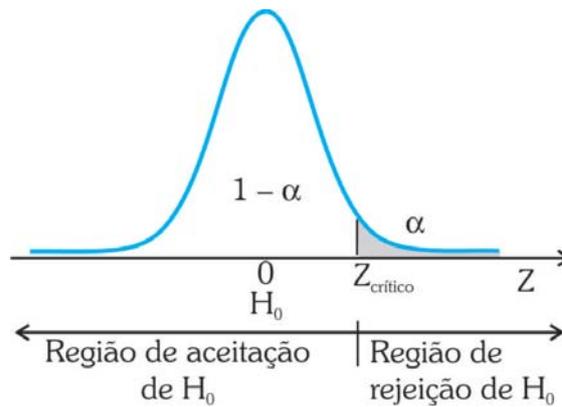


Figura 78 –  $H_1: \mu > 0$ .  
 Fonte: elaborada pelo autor.

Neste caso o valor crítico está à direita: se a média amostral  $\bar{x}$  ou a variável de teste  $Z$  tiverem valores superiores aos respectivos valores críticos, a Hipótese Nula será rejeitada, pois os valores “caíram” na região de Rejeição de  $H_0$ . Como foi notado anteriormente o valor de  $Z_{\text{crítico}}$  será positivo, pois é maior do que zero: usando o mesmo Nível de Significância de 5% o valor de  $Z_{\text{crítico}}$  seria 1,645, igual em módulo ao anterior, uma vez que a distribuição normal padrão é simétrica em relação à sua média que é igual a zero.

Se a Hipótese Alternativa fosse  $H_1: \mu \neq b$  ( $H_1: \mu \neq 0$ ), ou seja, o teste fosse Unilateral à direita a Região de Rejeição de  $H_0$  seria (Figuras 79 e 80):

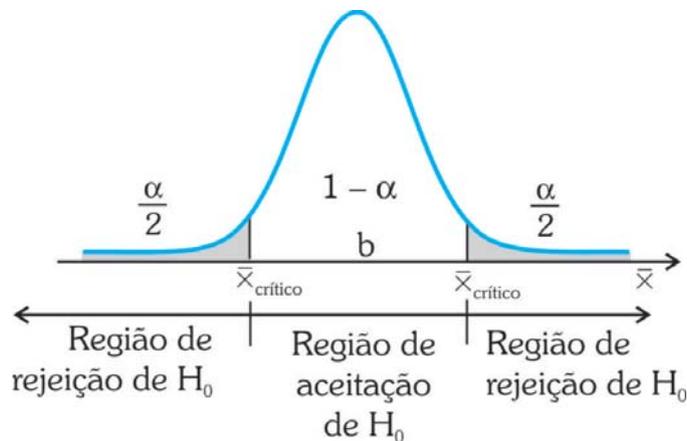


Figura 79 –  $H_1: \mu \neq b$ .  
 Fonte: elaborada pelo autor.

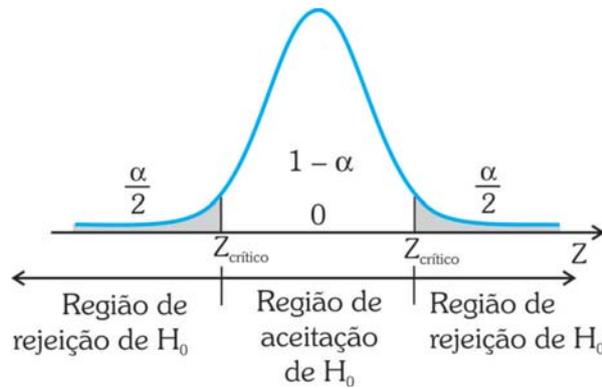


Figura 80 –  $H_1: \mu \neq 0$ .

Fonte: elaborada pelo autor.

Neste caso a região de Rejeição se divide em duas iguais (probabilidades iguais à metade do Nível de Significância  $\alpha$ ), semelhante ao que acontece na Estimação por Intervalo. Existirão dois valores críticos, um abaixo do valor de teste e outro acima: se a média amostral  $\bar{x}$  ou a variável de teste  $Z$  tiverem valores acima do valor crítico “superior” ou abaixo do valor crítico “inferior” a Hipótese Nula será rejeitada, pois os valores “caíram” em uma das duas regiões de Rejeição. Se for usada a variável de teste  $Z$  os valores críticos serão iguais em módulo, pois estão à mesma distância do valor sob teste (zero).

Recordando as três situações que foram abordadas anteriormente, seria interessante definir completamente as Hipóteses Estatísticas. Nos dois primeiros casos optou-se por um Teste Unilateral e no terceiro por um Teste Bilateral.

- a) Um novo protocolo de atendimento foi implementado no Banco RMG, visando reduzir o tempo que as pessoas passam na fila do caixa. O protocolo será considerado satisfatório, se a média do tempo de fila for menor do que 30 minutos. Um teste **Unilateral** seria o adequado. Mas Unilateral à Esquerda ou à Direita? Como está grifado na frase anterior haverá problema se a média do tempo fosse menor do que 30, resultando:

Teste **unilateral à esquerda**:

$H_0: \mu = 30$  onde  $\mu_0 = 30$  (valor de teste); e

$H_1: \mu < 30$  Teste Unilateral à Esquerda.

- b) Cerca de 2.000 formulários de pedidos de compra estão sendo analisados. Os clientes podem ficar insatisfeitos se houver erros nos formulários. Nesse caso, admite-se que a proporção máxima de formulários com erros seja de 5%. Ou seja, um valor maior do que 5% causaria problemas. Um teste **Unilateral** seria o adequado. Nesse caso, um teste de proporção, o problema será um valor maior do que 5%, resultando:

**Teste unilateral à direita:**

$H_0: \pi = 5\%$ , onde  $\pi_0 = 5\%$  (valor de teste); e

$H_1: \pi > 5\%$ .

- c) Uma peça automotiva precisa ter 100 mm de diâmetro, exatamente. Nesse caso, a dimensão não pode ser maior ou menor do que 100 mm (em outras palavras não pode ser diferente de 100 mm), pois isso indicará que a peça não está de acordo com as especificações. Um teste **Bilateral** seria o adequado, resultando:

**Teste Bilateral:**

$H_0: \mu = 100$  mm, onde  $\mu_0 = 100$  mm (valor de teste); e

$H_1: \mu \neq 100$  mm.

Para a definição correta das hipóteses é imprescindível a correta identificação do valor de teste, pois se trata de um dos aspectos mais importantes do teste: o resultado da amostra será comparado ao valor de teste.

Lembrando novamente que a tomada de decisão depende da correta determinação da região de Rejeição (e, por conseguinte, de Aceitação) da Hipótese Nula, e isso, por sua vez, depende diretamente da formulação das Hipóteses Estatísticas.

## Testes de Hipóteses sobre a Média de uma Variável em uma População

Nesse caso, há interesse em testar a hipótese de que o parâmetro média populacional ( $\mu$ ) de uma certa variável quantitativa seja maior, menor ou diferente de um certo valor. Para a realização desse teste é necessário que uma das duas condições seja satisfeita:

- sabe-se, ou é razoável supor, que a variável de interesse segue um modelo normal na população: isso significa que a distribuição amostral da média também será normal, permitindo realizar a inferência estatística paramétrica;
- a distribuição da variável na população é desconhecida, mas a amostra retirada desta população é considerada “suficientemente grande” o que, de acordo com o Teorema Central do Limite, permite concluir que a distribuição amostral da média é normal; e
- supõe-se também que a amostra é representativa da população e foi retirada de forma aleatória.

Tal como na Estimação de Parâmetros por Intervalo existirão diferenças nos testes, dependendo do conhecimento ou não da variância populacional da variável.

a) Se a variância populacional ( $\sigma^2$ ) da variável (cuja média populacional se quer testar) for conhecida.

Nesse caso, a variância amostral da média poderá ser calculada através da expressão:

$$V(\bar{x}) = \frac{\sigma^2}{n}, \text{ e, por conseguinte, o “desvio padrão” será } \frac{\sigma}{\sqrt{n}}.$$

A variável de teste será a variável Z da distribuição normal padrão, lembrando que:

$$Z = \frac{\text{valor} - \text{“média”}}{\text{“desvio-padrão”}}$$

A “**média**” será o valor de teste (suposto verdadeiro até prova em contrário), denotado por  $\mu_0$ . O **valor** (obtido pela amostra) será a média amostral (que é o melhor estimador da média populacional)

Há muita controvérsia a respeito do que seria uma amostra “suficientemente grande”, mas geralmente uma amostra com pelo menos 30 elementos costuma ser considerada grande o bastante para que a distribuição amostral da média possa ser aproximada por uma normal.

O desvio padrão é a raiz quadrada positiva da variância.

Neste caso  $Z_{\text{crítico}}$  será negativo, já que a região de Rejeição de  $H_0$  está à esquerda de zero.

E talvez este seja o motivo de se considerar mais de 30 elementos como sendo uma amostra suficientemente grande.

denotada por  $\bar{x}$ , e o “desvio padrão” será o valor obtido anteriormente. Sendo assim, a expressão da variável de teste  $Z$ :

$$Z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$$

Compara-se o valor da variável de teste com o valor crítico ( $Z_{\text{crítico}}$  que depende do Nível de Significância adotado) de acordo com o tipo de teste (as expressões abaixo também estão no apêndice):

Se  $H_1 \mu: > \mu_0$  Rejeitar  $H_0$  se  $Z > Z_{\text{crítico}}$  ( $\bar{x} > \bar{x}_{\text{crítico}}$ );

Se  $H_1 \mu: < \mu_0$  Rejeitar  $H_0$  se  $Z < Z_{\text{crítico}}$  ( $\bar{x} < \bar{x}_{\text{crítico}}$ ); e

Se  $H_1 \mu: \neq \mu_0$  Rejeitar  $H_0$  se  $|Z| \neq |Z_{\text{crítico}}|$ .

b) Se a variância populacional  $\sigma^2$  da variável for desconhecida.

Naturalmente esse é o caso mais encontrado na prática. Como se deve proceder? Dependerá do tamanho da amostra.

b.1 – Grandes amostras (mais de 30 elementos):

Nestes casos procede-se como no item anterior, apenas fazendo com que  $\sigma = s$ , ou seja, considerando que o desvio padrão da variável na população é igual ao desvio padrão da variável na amostra (suposição razoável para grandes amostras);

b.2 – Pequenas amostras (até 30 elementos):

Nestes casos a aproximação do item b.1 não será viável. Terá que ser feita uma correção na distribuição normal padrão ( $Z$ ) através da distribuição **t de Student**. Essa distribuição já é conhecida (ver Unidades 6 e 9). Trata-se de uma distribuição de probabilidades que possui média zero (como a distribuição normal padrão, variável  $Z$ ), mas sua variância é igual a  $n/(n-2)$ , ou seja, a variância depende do tamanho da amostra. Quanto maior for o tamanho da amostra mais o quociente acima se aproxima de 1 (a variância da distribuição normal padrão), e mais a distribuição t de Student aproxima-se da distribuição normal padrão. A partir de  $n = 30$ , já é possível considerar a variância da distribuição t de Student aproximadamente igual a 1.

A variável de teste será então  $t_{n-1}$  ( $t$  com  $n - 1$  graus de liberdade):

$$t_{n-1} = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

onde  $s$  é o desvio padrão amostral e os outros valores têm o mesmo significado da expressão anterior.

Compara-se o valor da variável de teste com o valor crítico ( $t_{n-1,crítico}$  que depende do Nível de Significância adotado) de acordo com o tipo de teste (as expressões abaixo também estão no Apêndice):

Se  $H_1: \mu > \mu_0$  Rejeitar  $H_0$  se  $t_{n-1} > t_{n-1,crítico}$  ( $\bar{x} > \bar{x}_{crítico}$ );

Se  $H_1: \mu < \mu_0$  Rejeitar  $H_0$  se  $t_{n-1} < t_{n-1,crítico}$  ( $\bar{x} < \bar{x}_{crítico}$ ); e

Se  $H_1: \mu \neq \mu_0$  Rejeitar  $H_0$  se  $|t_{n-1}| \neq |t_{n-1,crítico}|$ .

A correta identificação dos valores críticos, decorrente da correta identificação da região de rejeição de  $H_0$ , por sua vez decorrente da adequada formulação das hipóteses estatísticas, é indispensável para que o resultado obtido seja coerente.

Uma peça automotiva precisa ter 100 mm de diâmetro, exatamente. Foram medidas 15 peças, aleatoriamente escolhidas. Obteve-se média de 100,7 mm e variância de 0,01 mm<sup>2</sup>. Supõe-se que a dimensão segue distribuição normal na população. A peça está dentro das especificações? Usar 1% de significância? Vejamos neste primeiro exemplo:

Enunciar as hipóteses.

Conforme visto anteriormente o teste mais adequado para este caso é um Teste Bilateral:

$H_0: \mu = 100$  mm, onde  $\mu_0 = 100$  mm (valor de teste); e

$H_1: \mu \neq 100$  mm.

Nível de significância.

O problema declara que é necessário usar 1% de significância (se não fosse especificado, outro valor poderia ser usado).

Variável de teste.

Uma vez que a variância populacional da variável é DESCONHECIDA (o valor fornecido é a variância amostral), e a amostra retirada apresenta apenas 15 elementos (portanto menos de 30) a variável de teste será  $t_{n-1}$  da distribuição  $t$  de Student.

neste caso  $t_{n-1,crítico}$  será negativo, já que a região de Rejeição de  $H_0$  está à esquerda de zero.

Definir a região de aceitação de  $H_0$  (Figura 81).

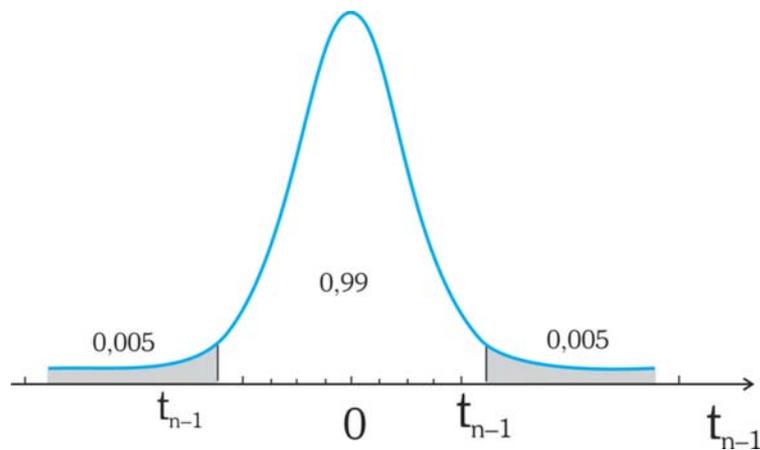


Figura 81: Regiões de rejeição e aceitação da hipótese nula – Teste bilateral de média.

Fonte: elaborada pelo autor.

Observe que por ser um teste Bilateral o Nível de Significância  $\alpha$  foi dividido em dois, metade para cada região de rejeição de  $H_0$ . Para encontrar o valor crítico devemos procurar na tabela da distribuição de *Student*, na linha correspondente a  $n-1$  graus de liberdade, ou seja, em  $15 - 1 = 14$  graus de liberdade. O valor da probabilidade pode ser visto na figura ao lado: os valores críticos serão  $t_{14;0,005}$  e  $t_{14;0,995}$  os quais serão iguais em módulo. E o valor de  $t_{n-1,critico}$  será igual a 2,977 (em módulo).

Através dos valores da amostra, avaliar o valor da variável.

Neste ponto é preciso encontrar o valor da variável de teste:

$$t_{n-1} = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

O valor de teste  $\mu_0$  é igual a 100 mm, a média amostral  $\bar{x}$  vale 100,7 mm, o tamanho de amostra  $n$  é igual a 15 e o desvio padrão amostral  $s$  é a raiz quadrada de 0,01 mm<sup>2</sup>. Substituindo na equação acima:

$$t_{n-1} = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = t_{15-1} = t_{14} = \frac{100,7 - 100}{\sqrt{0,01}/\sqrt{15}} = 27,11 \text{ então } |t_{14}| = 27,11.$$

Decidir pela aceitação ou rejeição de  $H_0$ .

Como se trata de um teste bilateral:

Rejeitar  $H_0$  se  $|t_{n-1}| \neq |t_{n-1,critico}|$ ; e

Como  $|t_{14}| = 27,11 > |t_{n-1,critico}| = |t_{14,0,995}| = 2,977$ .

Rejeitar  $H_0$  a 1% de Significância (há 1% de chance de erro).

Interpretar a decisão no contexto do problema. Há provas estatísticas suficientes de que a dimensão da peça não está dentro das especificações.

Um novo protocolo de atendimento foi implementado no Banco RMG, visando reduzir o tempo que as pessoas passam na fila do caixa. O protocolo será considerado satisfatório se a média do tempo de fila for **menor** do que 30 minutos. Suponha que o tempo que 35 clientes (selecionados aleatoriamente) passaram na fila foi monitorado, resultando em uma média de 29 minutos e desvio padrão de 5 minutos. O protocolo pode ser considerado satisfatório a 5% de significância? Vejamos neste segundo exemplo:

Esse problema já foi estudado anteriormente.

Enunciar as hipóteses. Conforme visto anteriormente o teste mais adequado para este caso é um Teste Unilateral à Esquerda:

$H_0: \mu = 30$ , onde  $\mu_0 = 30$  (valor de teste); e

$H_1: \mu < 30$ .

Nível de significância. O problema declara que é necessário usar 5%.

Variável de teste. Uma vez que a variância populacional da variável é DESCONHECIDA (o valor fornecido é o desvio padrão AMOSTRAL), mas a amostra retirada apresenta 35 elementos (portanto mais de 30) a variável de teste será **Z** da distribuição normal.

Definir a região de aceitação de  $H_0$  (Figura 82).

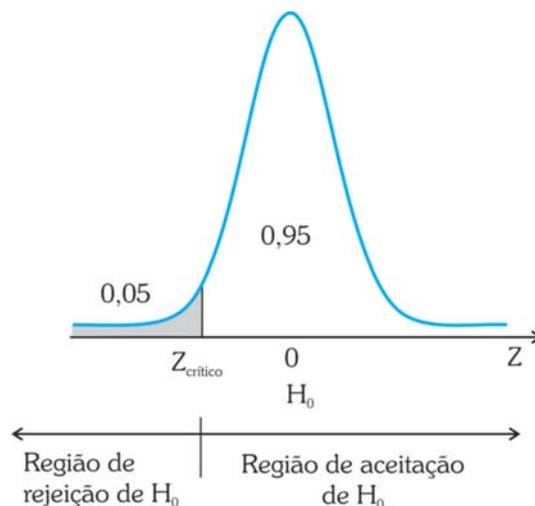


Figura 82: Regiões de aceitação e de rejeição – Teste unilateral à esquerda.

Fonte: elaborada pelo autor.

Cuidado com os casos de FRONTEIRA, em que o valor da variável de teste é muito próximo do valor crítico. Nesses casos a rejeição ou aceitação de  $H_0$  pode ocorrer por acaso. Sempre que apresentar o resultado recomende que uma nova amostra seja retirada para avaliar novamente o problema. Mas neste caso rejeita-se  $H_0$  com folga.

Observe que por ser um teste Unilateral à Esquerda o Nível de Significância  $\alpha$  está todo concentrado em um dos lados da distribuição, definindo a região de rejeição de  $H_0$ . Para encontrar o valor crítico devemos procurar na tabela da distribuição normal, pela probabilidade acumulada 0,95. Ou procurar a probabilidade complementar 0,05 e mudar o sinal do valor encontrado, pois o  $Z_{\text{crítico}}$  aqui é menor do que zero. O valor crítico será igual a  $-1,645$ .

Através dos valores da amostra avaliar o valor da variável.

Neste ponto é preciso encontrar o valor da variável de teste:

$$Z = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

O valor de teste  $\mu_0$  é igual a 30, a média amostral  $\bar{x}$  vale 29, o tamanho de amostra  $n$  é igual a 35 e o desvio padrão amostral  $s$  é 5. Substituindo na equação acima:

$$Z = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{29 - 30}{5/\sqrt{35}} = -1,183.$$

Decidir pela aceitação ou rejeição de  $H_0$ . Como se trata de um teste Unilateral à Direita:

Rejeitar  $H_0$  se  $Z < Z_{\text{crítico}}$ . Como  $Z = -1,185 > Z_{\text{crítico}} = -1,645$ .

Aceitar  $H_0$  a 5% de Significância (há 5% de chance de erro).

Interpretar a decisão no contexto do problema. Não há provas estatísticas suficientes para concluir que o protocolo tem um desempenho satisfatório.

## Testes de Hipóteses sobre a Proporção de uma Variável em uma População

Nesse caso há interesse em testar a hipótese de que o parâmetro proporção populacional ( $\pi$ ) de um dos valores de uma certa variável seja maior, menor ou diferente de um certo valor. Para a realização desse teste, tal como será descrito é necessário que duas condições sejam satisfeitas:

- que o produto  $n \times \pi_0$  seja maior ou igual a 5; e

- que o produto  $n \times (1 - \pi_0)$  seja maior ou igual a 5.

Onde  $n$  é o tamanho da amostra e  $\pi_0$  é a proporção sob teste (de um dos valores da variável). Se ambas as condições forem satisfeitas, a distribuição amostral da proporção que é binomial (uma Bernoulli repetida  $n$  vezes) pode ser aproximada por uma normal. Obviamente supõe-se que a amostra é representativa da população e foi retirada de forma aleatória, e que a variável pode assumir apenas dois valores, aquele no qual há interesse e o seu complementar.

Se as condições acima forem satisfeitas, a distribuição amostral da proporção poderá ser aproximada por uma normal com:

$$\text{Média} = \pi_0 \quad \text{Desvio Padrão} = \sqrt{\frac{\pi_0 \times (1 - \pi_0)}{n}}$$

Lembrando da expressão da variável  $Z$ :

$$Z = \frac{\text{valor} - \text{“média”}}{\text{“desvio-padrão”}}$$

O **valor** será a proporção amostral (que é o melhor estimador da proporção populacional) do valor da variável denotada por  $p$ . A “média” e o “desvio padrão” são os valores definidos acima, então a expressão de  $Z$  será:

$$Z = \frac{p - \pi_0}{\sqrt{\frac{\pi_0 \times (1 - \pi_0)}{n}}}$$

Compara-se o valor da variável de teste com o valor crítico ( $Z_{\text{crítico}}$  que depende do Nível de Significância adotado) de acordo com o tipo de teste:

- Se  $H_1 : \mu > \mu_0$  Rejeitar  $H_0$  se  $Z > Z_{\text{crítico}}$  ( $p > p_{\text{crítico}}$ );
- Se  $H_1 : \mu < \mu_0$  Rejeitar  $H_0$  se  $Z < Z_{\text{crítico}}$  ( $p < p_{\text{crítico}}$ ); e
- Se  $H_1 : \mu \neq \mu_0$  Rejeitar  $H_0$  se  $|Z| \neq |Z_{\text{crítico}}|$ .

Cerca de 2.000 formulários de pedidos de compra estão sendo analisados. Os clientes podem ficar insatisfeitos se houver erros nos formulários. Neste caso admite-se que a proporção máxima de formulários com erros seja de 5%. Suponha que dentre os 2.000 formulários 7% apresentavam erros. A proporção máxima foi ultrapassada a 1% de significância? Vejamos neste terceiro exemplo:

Neste caso  $Z_{\text{crítico}}$  será negativo, já que a região de Rejeição de  $H_0$  está à esquerda de zero.

Enunciar as hipóteses. Conforme visto anteriormente, o teste mais adequado para esse caso é um Teste Unilateral à Direita:

$H_0: \mu = 5\%$ , onde  $\mu_0 = 5\%$  (valor de teste); e

$H_1: \mu > 5\%$ .

Nível de significância. O problema declara que é necessário usar 1% de significância (se não fosse especificado, outro valor poderia ser usado).

Variável de teste. Como se trata de um teste de proporção é necessário verificar o valor dos produtos:

$n \times \pi_0 = 2000 \times 0,05 = 100$  e  $n \times (1 - \pi_0) = 2000 \times 0,95 = 1900$ . Como ambos são maiores do que 5 é possível fazer uma aproximação pela normal, e a variável de teste será  $Z$ .

Definir a região de aceitação de  $H_0$  (Figura 83).

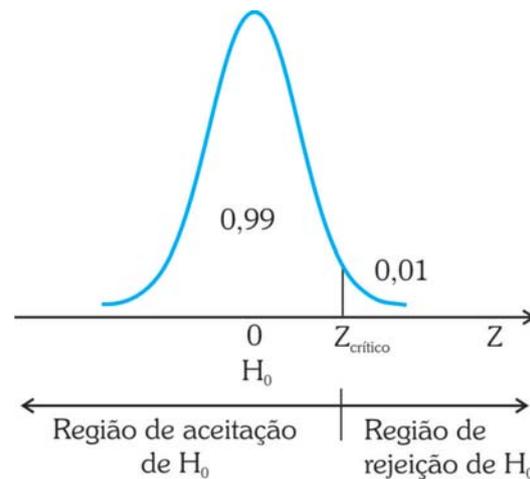


Figura 83: Regiões de aceitação e de rejeição – Teste unilateral à direita.

Fonte: elaborada pelo autor.

Observe que por ser um teste Unilateral à Direita, o Nível de Significância  $\alpha$  está todo concentrado em um dos lados da distribuição, definindo a região de rejeição de  $H_0$ . Para encontrar o valor crítico devemos procurar na tabela da distribuição normal, pela probabilidade acumulada 0,01 (o  $Z_{\text{crítico}}$  aqui é maior do que zero). O valor crítico será igual a 2,326.

Através dos valores da amostra avaliar o valor da variável. Nesse ponto é preciso encontrar o valor da variável de teste:

$$Z = \frac{p - \pi_0}{\sqrt{\frac{\pi_0 \times (1 - \pi_0)}{n}}}$$

O valor de teste  $\pi_0$  é igual a 0,05 (5%), a proporção amostral  $p$  vale 0,07 (7%), e o tamanho de amostra  $n$  é igual a 2.000. Substituindo na equação acima:

$$Z = \frac{p - \pi_0}{\sqrt{\frac{\pi_0 \times (1 - \pi_0)}{n}}} = \frac{0,07 - 0,05}{\sqrt{\frac{0,05 \times (0,95)}{2000}}} = 4,104.$$

Decidir pela aceitação ou rejeição de  $H_0$ . Como se trata de um teste Unilateral à Direita:

**Rejeitar  $H_0$  se  $Z > Z_{\text{crítico}}$**  Como  $Z = 4,104 > Z_{\text{crítico}} = 2,326$ .

Rejeitar  $H_0$  a 1% de Significância (há 1% de chance de erro).

Interpretar a decisão no contexto do problema. Há provas estatísticas suficientes de que a proporção está acima do máximo admitido. Provavelmente os vendedores/compradores precisarão passar por novo treinamento.

Este não é um caso de fronteira

Agora vamos ver um tipo de teste estatístico muito utilizado pelos administradores, para avaliar o relacionamento entre duas variáveis qualitativas: o teste de associação (independência de quiquadrado).

## Teste de associação de quiquadrado

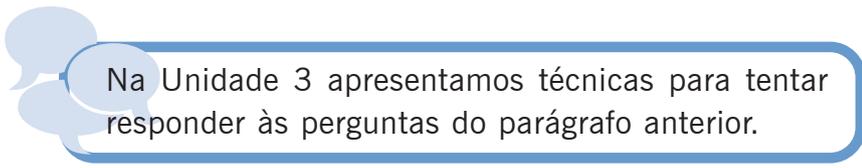
O teste do quiquadrado, também chamado de teste de independência de quiquadrado, está vinculado à análise de duas variáveis qualitativas. Vamos ver alguns conceitos antes de apresentar o teste de associação de quiquadrado.

**Teste de associação (independência) de quiquadrado** – teste que permite avaliar se duas variáveis qualitativas, cujas frequências estão dispostas em uma tabela de contingências, apresentam associação significativa ou não. Fonte: Barbetta, Reis e Borna (2008).

## Variáveis qualitativas e tabelas de contingência

É comum haver interesse em saber se duas variáveis quaisquer estão relacionadas, e o quanto estão relacionadas, seja na vida prática, seja em trabalhos de pesquisa, por exemplo:

- se a satisfação com um produto está relacionada à faixa etária do consumidor; e
- se a função exercida por uma pessoa em uma organização está associada a seu gênero.



Na Unidade 3 apresentamos técnicas para tentar responder às perguntas do parágrafo anterior.

Variáveis Qualitativas são as variáveis cujas realizações são atributos, categorias (Unidades 1 e 3). Como exemplo de variáveis qualitativas temos: sexo de uma pessoa (duas categorias, masculino e feminino), grau de instrução (analfabeto, primeiro grau incompleto, etc.), opinião sobre um assunto (favorável, desfavorável, indiferente).

Em estudos sobre variáveis qualitativas é extremamente comum registrar as frequências de ocorrência de cada valor que as variáveis podem assumir, e quando há duas variáveis envolvidas é comum registrar-se a frequência de ocorrência dos cruzamentos entre valores: por exemplo, quantas pessoas do sexo masculino são favoráveis a uma certa proposta de lei, quantas são desfavoráveis, quantas pessoas do sexo feminino são favoráveis. E, para facilitar a análise dos resultados, eles costumam ser dispostos em uma Tabela de Contingências. A Tabela de Contingências relaciona os possíveis valores de uma variável qualitativa com os possíveis valores da outra, registrando quantas ocorrências foram verificadas de cada cruzamento.

Exemplo 4 – O Quadro 23 mostra uma tabela de contingências relacionando as funções exercidas e o sexo de 474 funcionários de uma organização.

SEXO	FUNÇÃO			
	ESCRITÓRIO	SERVIÇOS GERAIS	GERÊNCIA	TOTAL
Masculino	157	27	74	258
Feminino	206	0	10	216
Total	363	27	84	474

Quadro 23: Tabela de contingências de Função por Sexo.  
 Fonte: elaborado pelo autor

Podemos apresentar os percentuais calculados em relação aos totais das colunas no Quadro 24:

SEXO	FUNÇÃO			
	ESCRITÓRIO	SERVIÇOS GERAIS	GERÊNCIA	TOTAL
Masculino	43,25%	100%	88,10%	54%
Feminino	56,75%	0%	11,90%	46%
Total	100%	100%	100%	100%

Quadro 24: Tabela de contingências de Função por Sexo:  
 percentuais por colunas.  
 Fonte: elaborado pelo autor.

Seria interessante saber se as duas variáveis são estatisticamente dependentes, e o quão forte é essa associação. Repare que os percentuais de homens e mulheres em cada função são diferentes dos percentuais marginais (de homens e mulheres no total de funcionários), sendo que em duas funções as diferenças são bem grandes.

O teste de associação de quiquadrado é uma das ferramentas estatísticas mais utilizadas quando se deseja estudar o relacionamento entre duas variáveis qualitativas. Permite verificar se duas variáveis qualitativas são independentes, se as proporções de ocorrência dos valores das variáveis observadas estão de acordo com o que era esperado, etc. Neste texto haverá interesse em usar o teste para avaliar se duas variáveis qualitativas são independentes.

Como todo teste de hipóteses o teste de associação de quiquadrado consiste em comparar os valores observados em uma amostra com os valores de uma referência (referência esta que supõe que a hipótese nula seja válida).

As frequências observadas da variável são representadas em uma tabela de contingências, e a Hipótese Nula ( $H_0$ ) do teste será a

de que as duas variáveis não se diferem em relação às frequências com que ocorre uma característica particular, ou seja, as variáveis são independentes, que será testada contra a Hipótese Alternativa ( $H_1$ ) de que as variáveis não são independentes.

O teste pode ser realizado porque o grau de dependência pode ser quantificado descritivamente através de uma estatística, que se chama justamente quiquadrado ( $\chi^2$ ) na população, mas na amostra é chamada de  $q^2$ , cuja expressão é:

$$q^2 = \sum_{i=1}^L \sum_{j=1}^C \left[ \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \right]$$

Sendo  $E_{ij} = \frac{\text{total da linha } i \times \text{total da coluna } j}{\text{total geral}}$ .

Onde:

- $E_{ij}$  é a frequência esperada, sob a condição de independência entre as variáveis, em uma célula qualquer da tabela de contingências. Todas as frequências esperadas precisam ser maiores ou iguais a 5 para que o resultado do teste seja válido.
- $O_{ij}$  é a frequência observada em uma célula qualquer da tabela de contingências.
- $L$  é o número total de linhas da tabela de contingências (número de valores que uma das variáveis pode assumir).
- $C$  é o número total de colunas da tabela (número de valores que a outra variável pode assumir).

Então, para cada célula da tabela de contingências calcula-se a diferença entre a frequência observada e a esperada. Para evitar que as diferenças positivas anulem as negativas, as diferenças são elevadas ao quadrado. E para evitar que uma diferença grande em termos absolutos, mas pequena em termos relativos, “inflacione” a estatística, ou que uma diferença pequena em termos absolutos, mas grande em termos relativos, tenha sua influência reduzida, divide-se o quadrado da diferença pela frequência esperada. Somam-se os valores de todas as células e obtêm-se o valor da estatística: quanto maior  $q^2$ , mais o Observado se afasta do Esperado, portanto maior a dependência.

Se isso não ocorrer recomenda-se agrupar as categorias (de uma ou outra variável, ou de ambas) até obter todas as frequências pelo menos iguais a 5.

Sob a hipótese de independência a estatística  $q^2$  seguirá o modelo quiquadrado, que estudamos na Unidade 7, prometendo usá-la aqui na Unidade 10.

O Teste do quiquadrado para avaliar se duas variáveis são independentes será unilateral: ou seja a Hipótese Nula será rejeitada se  $q^2 > q^2_{\text{crítico}}$ , para um certo número de graus de liberdade. Por exemplo, para o caso em que há três graus de liberdade, e o Nível de Significância fosse de 5% (a região de Rejeição de  $H_0$  ficará À DIREITA), o valor crítico seria (lembre-se da Unidade 7) (Figura 84):

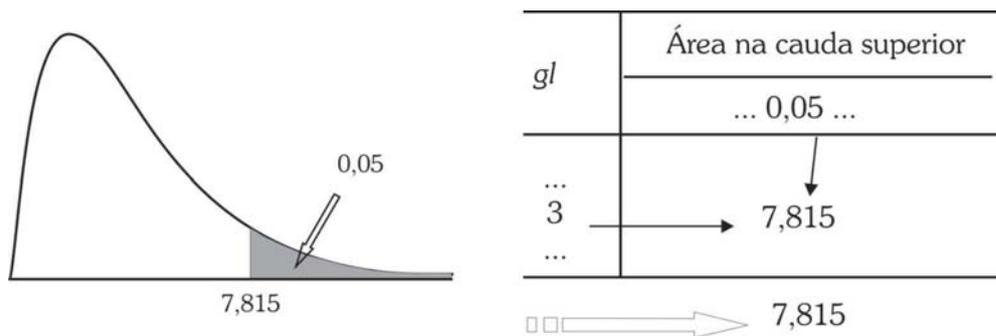


Figura 84: Uso da tabela da distribuição quiquadrado. Ilustração com  $gl = 3$  e área na cauda superior de 5%.

Fonte: adaptada de Barbeta, Reis e Bornia (2008).

O número de graus de liberdade da estatística é calculado da seguinte forma:

$$\text{graus de liberdade} = (L - 1) \times (C - 1).$$

Sendo o número de linhas e o número de colunas referentes à tabela de contingências (são os números de valores que cada variável pode assumir).

O número de graus de liberdade assume esse valor porque para calcular as frequências esperadas não é necessário calcular os valores de todas as células, as últimas podem ser calculadas por diferença já que os totais são fixos. Por exemplo, para duas variáveis que somente podem assumir dois valores cada, o número de graus de liberdade seria igual a 1  $[(2-1) \times (2-1)]$ : bastaria calcular a frequência esperada de uma célula e obter as outras por diferença em relação ao total.

Para o conjunto do Exemplo 4, supondo que os resultados são provenientes de uma amostra aleatória, vamos verificar neste quinto exemplo se as variáveis são independentes a 1% de significância.

Enunciar as Hipóteses:

$H_0$ : as variáveis sexo e função são independentes; e

$H_1$ : as variáveis sexo e função não são independentes.

Nível de significância: determinado pelo problema,  $\alpha = 0,01$ ;  $\beta = 0,99$ .

Retirar as amostras aleatórias e montar a tabela de contingências (isso já foi feito) (Quadro 25):

SEXO	FUNÇÃO			TOTAL
	ESCRITÓRIO	SERVIÇOS GERAIS	GERÊNCIA	
Masculino	157	27	74	258
Feminino	206	0	10	216
Total	363	27	84	474

Quadro 25: Tabela de contingências de Função por Sexo.

Fonte: elaborado pelo autor.

No Quadro 25 encontramos os totais marginais e o total geral:

$$L_1 = \text{total Masculino} = 258 \quad L_2 = \text{total Feminino} = 216$$

$$C_1 = \text{total Escritório} = 157.$$

$$C_2 = \text{total S.Gerais} = 27 \quad C_3 = \text{total gerência} = 84$$

$$N = \text{total geral} = 474.$$

Repare que somando os totais das linhas o resultado é o total geral, e que somando os totais das colunas o resultado é o total geral também.

Calcular as frequências esperadas:

Calculando as frequências esperadas de acordo com a fórmula vista anteriormente:

$$\text{Masculino - Escritório} \quad E = (258 \times 363) / 474 = 197,58.$$

$$\text{Masculino - Serviços Gerais} \quad E = (258 \times 27) / 474 = 14,70.$$

$$\text{Masculino - Gerência} \quad E = (258 \times 84) / 474 = 45,72.$$

$$\text{Feminino - Escritório} \quad E = (216 \times 363) / 474 = 165,42.$$

$$\text{Feminino - Serviços Gerais} \quad E = (216 \times 27) / 474 = 12,30.$$

$$\text{Feminino - Gerência} \quad E = (216 \times 84) / 474 = 38,28.$$

Observe que os resultados são os mesmos obtidos no Exemplo 3.2.

Calculando a estatística  $q^2$  para cada célula:

Agora podemos calcular as diferenças entre as frequências e as demais operações, que serão mostradas nos Quadros 26 , 27 e 28).

O – E	FUNÇÃO		
SEXO	ESCRITÓRIO	SERVIÇOS GERAIS	GERÊNCIA
Masculino	157 – 197,58	27 – 14,70	74 – 45,72
Feminino	206 – 165,42	0 – 12,30	10 – 38,28

Quadro 26: Diferença entre frequências observadas e esperadas de Função por Sexo.

Fonte: elaborado pelo autor.

(O – E) <sup>2</sup>	FUNÇÃO		
SEXO	ESCRITÓRIO	SERVIÇOS GERAIS	GERÊNCIA
Masculino	1646,921	151,383	799,672
Feminino	1646,921	151,383	799,672

Quadro 27: Diferença entre frequências observadas e esperadas de Função por Sexo elevadas ao quadrado.

Fonte: elaborado pelo autor.

Finalmente:

(O – E) <sup>2</sup> /E	FUNÇÃO		
SEXO	ESCRITÓRIO	SERVIÇOS GERAIS	GERÊNCIA
Masculino	8,336	10,301	17,490
Feminino	9,956	12,304	20,891

Quadro 28: Estatísticas  $q^2$  de Função por Sexo.

Fonte: elaborado pelo autor.

Agora podemos somar os valores:

$$q^2 = 8,336 + 10,301 + 17,490 + 9,956 + 12,304 + 20,891 = 79,227.$$

Os graus de liberdade:  $(\text{número de linhas} - 1) \times (\text{número de colunas} - 1) = (2 - 1) \times (3 - 1) = 2$ .

Então  $q^2_2 = 79,227$ .

O  $q^2$  crítico será: procurando na Tabela 3 do ambiente, ou em um programa, para 2 graus de liberdade e 99% de confiança (1% de significância):  $q^2_{2,\text{crítico}} = 6,63$ .

Ver Figura 85:

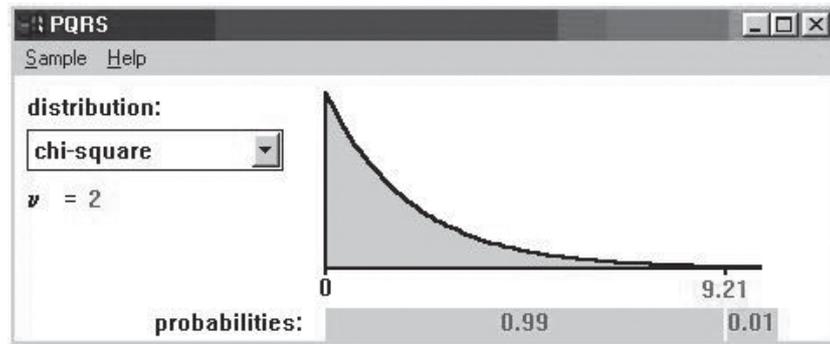


Figura 85: Valor crítico de  $q^2$  para 2 graus de liberdade e 1% de significância.  
Fonte: adaptada pelo autor de PQRS®.

Como  $q^2_2$  é maior do que  $q^2_{2,\text{crítico}}$  rejeitamos  $H_0$  a 1% de significância. Há evidência estatística suficiente que indica que as variáveis: função e sexo não são independentes. Isso confirma nossas suspeitas iniciais, devido às grandes diferenças nas frequências da tabela.

No tópico Tô a fim de saber... você terá indicações de vários outros tipos de hipóteses que não foram mencionados nesta Unidade. As referências lá citadas serão extremamente valiosas se você tiver que:

- aplicar testes para avaliar se há diferenças entre médias de duas ou mais populações;
- aplicar testes para avaliar se há diferenças entre proporções de duas populações; e
- aplicar testes não paramétricos, por exemplo, testes de aderência dos dados a um determinado modelo probabilístico.

Com este tópico terminamos nossa jornada... Agora é com vocês. Boa sorte!

## Saiba Mais

Sobre tipos de erro, poder, em testes de hipóteses: BARBETTA, Pedro A.; REIS, Marcelo M.; BORNIA, Antonio C. *Estatística para Cursos de Engenharia e Informática*. 2. ed. São Paulo: Atlas, 2008, Capítulo 8; STEVENSON, Willian J. *Estatística Aplicada à Administração*. São Paulo: Harbra, 2001, Capítulo 10.

Sobre testes de uma variância: BARBETTA, Pedro A.; REIS, Marcelo M.; BORNIA, Antonio C. *Estatística para Cursos de Engenharia e Informática*. 2. ed. São Paulo: Atlas, 2008, Capítulo 8; TRIOLA, Mário. *Introdução à Estatística*. Rio de Janeiro: LTC, 1999, Capítulo 7.

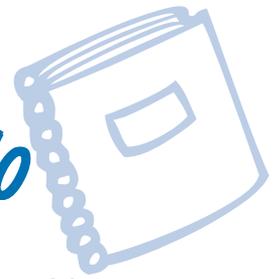
Sobre testes de comparação de duas médias: BARBETTA, Pedro A.; REIS, Marcelo M.; BORNIA, Antonio C. *Estatística para Cursos de Engenharia e Informática*. 2. ed. São Paulo: Atlas, 2008, Capítulo 9.

Sobre testes de comparação de duas proporções: MOORE, David S.; McCABE, George P.; DUCKWORTH, William M.; SCLOVE, Stanley L. *A prática da estatística empresarial: como usar dados para tomar decisões*. Rio de Janeiro: LTC, 2006, Capítulo 8.

Sobre Análise de Variância, comparação de várias médias: BARBETTA, Pedro A.; REIS, Marcelo M.; BORNIA, Antonio C. *Estatística para Cursos de Engenharia e Informática*. 2. ed. São Paulo: Atlas, 2008, Capítulo 9; STEVENSON, Willian J. *Estatística Aplicada à Administração*. São Paulo: Harbra, 2001, Capítulo 11; MOORE, David S.; McCABE, George P.; DUCKWORTH, William M.; SCLOVE, Stanley L. *A prática da estatística empresarial: como usar dados para tomar decisões*. Rio de Janeiro: LTC, 2006, Capítulos 14 e 15; TRIOLA, Mário. *Introdução à Estatística*. Rio de Janeiro: LTC, 1999, Capítulo 11.

Sobre testes não paramétricos: BARBETTA, Pedro A.; REIS, Marcelo M.; BORNIA, Antonio C. *Estatística para Cursos de Engenharia e Informática*. 2. ed. São Paulo: Atlas, 2008, Capítulo 10; STEVENSON, Willian J. *Estatística Aplicada à Administração*. São Paulo: Harbra, 2001, Capítulo 13; SIEGEL, Stanley. *Estatística Não Paramétrica (para as Ciências do Comportamento)*. São Paulo: McGraw-Hill, 1975.

Sobre a utilização do Microsoft Excel para realizar testes de hipóteses: LEVINE, David M.; STEPHAN, David; KREHBIEL, Timothy C.; BERENSON, Mark L. *Estatística: Teoria e Aplicações – Usando Microsoft Excel em Português*. 5. ed. Rio de Janeiro: LTC, 2000, Capítulo 6.



O resumo desta Unidade está demonstrado na Figura 86:

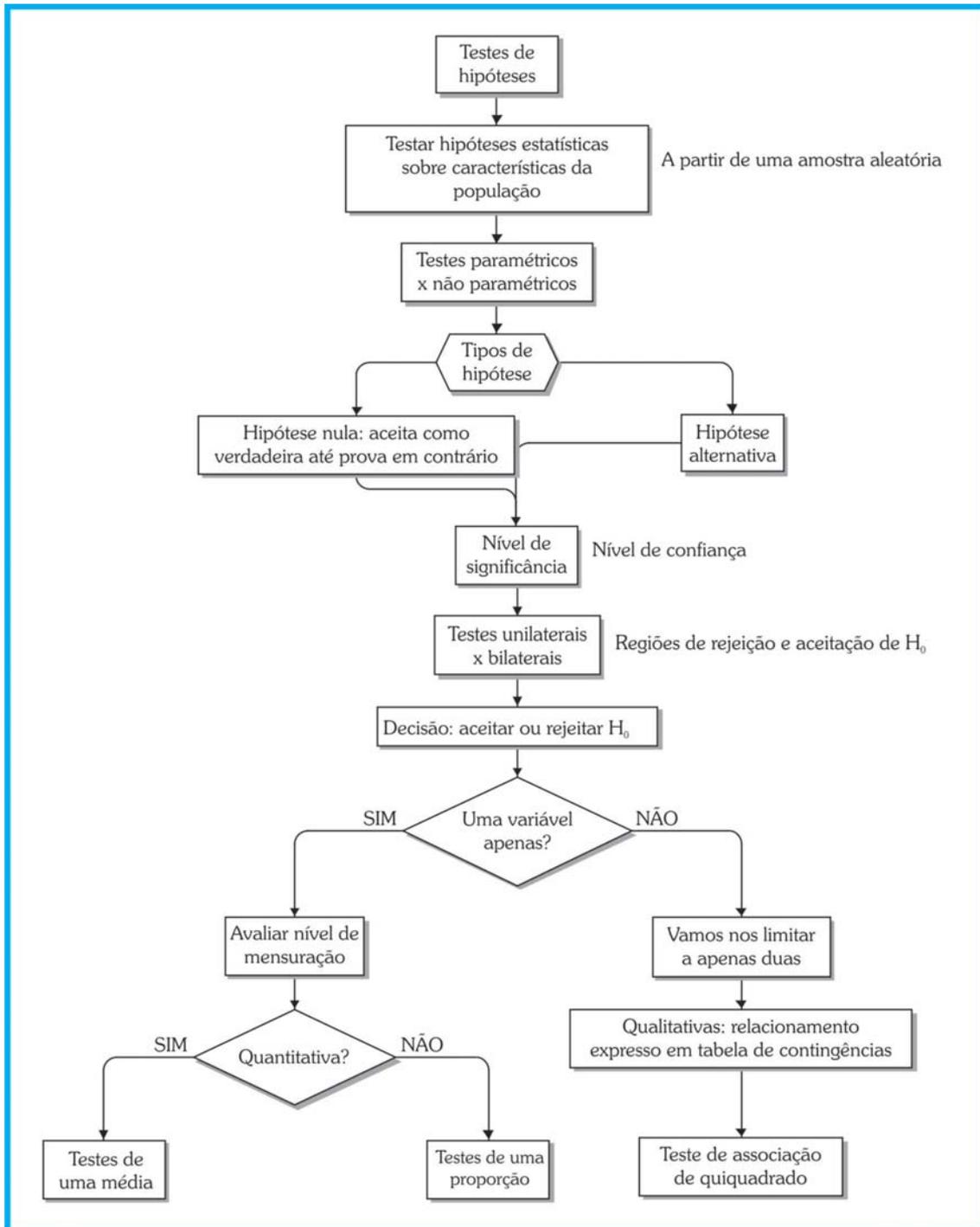


Figura 86: Resumo da Unidade 10.

Fonte: elaborado pelo autor.

Chegamos ao final da disciplina de Estatística Aplicada à Administração. Estudamos nesta última Unidade os testes de hipóteses, tipos de hipóteses e suas variáveis. A Unidade foi explorada com Figuras, exemplos e Quadros para melhor representar o conteúdo oferecido. Além do material produzido pelo professor, você tem em mãos uma riquíssima fonte de referências para saber mais sobre o assunto. Explore os conhecimentos propostos. Não tenha esta Unidade como fim, mas sim como o começo de uma nova trajetória em sua vida acadêmica. Bons estudos e boa sorte!



## Atividades de aprendizagem

- 1) O tempo médio de atendimento em uma agência lotérica está sendo analisado por técnicos. Uma amostra de 40 clientes foi sistematicamente monitorada em relação ao tempo que levavam para serem atendidos, obtendo-se as seguintes estatísticas: tempo médio de atendimento de 195 segundos e desvio padrão de 15 segundos.

Considerando que o tempo de utilização segue uma distribuição normal: O dono da agência garante que o tempo médio de atendimento é de 3 minutos (se for maior ele se compromete a contratar mais um atendente). Com base nos dados da amostra a afirmação do dono é verdadeira, ou ele deve contratar um novo atendente? Use um nível de significância de 1%?

- 2) Buscando melhorar a qualidade do serviço, uma empresa estuda o tempo de atraso na entrega dos pedidos recebidos. Supondo que o tempo de atraso se encontra normalmente distribuído, e conhecendo o tempo de atraso dos últimos 20 pedidos, descritos abaixo (em dias), determine:

5 1 0 3 6 10 2 3 4 1 5 3 1 6 6 9 0 0 1 0

Um cliente enfurecido quer testar estatisticamente a hipótese (declarada pela empresa) de que o atraso médio será de no máximo 1 dia. Ele argumenta que deve ser maior, e quer uma confiança de 99%. O cliente tem razão na sua reclamação?

- 3) Uma máquina produz peças classificadas como boas ou defeituosas. Retirou-se uma amostra de 1000 peças da produção, verificando-se que 35 eram defeituosas. O controle de qualidade pára a linha de produção para rearranjo dos equipamentos envolvidos quando o percentual de defeituosos é superior a 3%.

Baseado nos dados amostrais, a linha de produção deve ser parada?

- 4) Queremos saber se há diferenças significativas entre três meios de comunicação, em termos de lembrança do consumidor da propaganda veiculada. O resultado de um estudo sobre propaganda mostrou:

CAPACIDADE DE LEMBRANÇA	MEIO DE COMUNICAÇÃO			
	REVISTA	TV	RÁDIO	TOTAL
Lembram da propaganda	25	93	7	125
Não lembram da propaganda	73	10	108	191
Total	98	103	115	316

- a) Usando 1% de significância é possível concluir que há associação entre a capacidade de lembrança e o meio de comunicação usado?
- b) Observando os resultados acima, qual meio de comunicação você recomendaria para veicular uma propaganda, para maximizar a capacidade de lembrança do público alvo? Por quê?

Adaptado de LEVINE, David M.; STEPHAN, David; KREHBIEL, Timothy C.; BERENSON, Mark L. *Estatística: Teoria e Aplicações usando Microsoft Excel® em Português*. Rio de Janeiro: LTC, 2000.