

Testes de Hipóteses para uma Única Amostra

Roteiro

1. Introdução
2. Testes para a Média
 - i. População normal com variância conhecida
 - ii. População normal com variância desconhecida
3. Testes para a Variância e o Desvio-padrão
4. Testes para a Proporção
5. Roteiro para Procedimentos de Inferência
6. Testes de Adequação de Ajuste
7. Testes para Tabela de Contingência
8. Referências

Testes de Hipóteses

Teste de Hipóteses

- Procedimento de tomada de decisão sobre hipóteses envolvendo a população
- Aspecto bastante útil da Inferência Estatística:
 - √ Problemas de tomada de decisão, teste ou experimentos podem ser formulados em termos de teste hipóteses
- Há uma conexão entre testes de hipóteses e intervalos de confiança
 - √ Estimação de parâmetros, com teste estatístico de hipóteses e com intervalo de confiança são fundamentais na análise de dados

Hipótese Estatística

- Afirmação sobre os parâmetros de uma ou mais populações
 - ✓ A hipótese geralmente envolve um ou mais parâmetros da distribuição de probabilidades que modela a população

Exemplo

- Taxa de queima de propelente sólido, usado para fornecer energia aos sistemas de escapamento de aeronaves
 - ✓ Taxa de queima é uma variável aleatória descrita por distribuição de probabilidades
 - ✓ Interesse:
 - Taxa média de queima (parâmetro da distribuição)
 - Taxa média de queima é ou não 50 cm/s?

✓ Hipóteses:

- $H_0: \mu = 50 \text{ cm/s}$
- $H_1: \mu \neq 50 \text{ cm/s}$

✓ Denominação:

- H_0 : hipótese nula
- H_1 : hipótese alternativa

✓ No exemplo a hipótese alternativa é bilateral

✓ Pode-se formular hipótese alternativa unilateral

- $H_0: \mu = 50 \text{ cm/s}$
- $H_1: \mu > 50 \text{ cm/s}$

ou

- $H_0: \mu = 50 \text{ cm/s}$
- $H_1: \mu < 50 \text{ cm/s}$

• Importante:

- ✓ Hipóteses são sempre afirmações sobre a população ou distribuição sobre estudo
 - Não são afirmações sobre a amostra!

Teste de Hipóteses

- Procedimento que leva a uma decisão acerca de uma hipótese particular
 - ✓ Apoiam-se no uso de informações de amostras aleatórias da população de interesse
 - ✓ Se a informação amostral for consistente com a hipótese, essa não será rejeitada
 - ✓ Se a informação amostral for inconsistente com a hipótese, conclui-se que a hipótese é falsa

- Não sabemos com certeza sobre a verdade ou falsidade de uma hipótese
 - ✓ Só se pudermos examinar toda a população!
- O procedimento de teste de hipóteses deve ser desenvolvido considerando a probabilidade de alcançar uma conclusão errada!

Teste de Hipóteses

- Estruturação:
 - ✓ Estabelecer H_0 de modo que ela especifique um valor exato do parâmetro
 - $\mu = 50 \text{ cm/s}$
 - ✓ Selecionar uma amostra aleatória
 - ✓ Calcular uma estatística de teste a partir dos dados amostrais
 - ✓ Tomar uma decisão a respeito de H_0 .

Exemplo

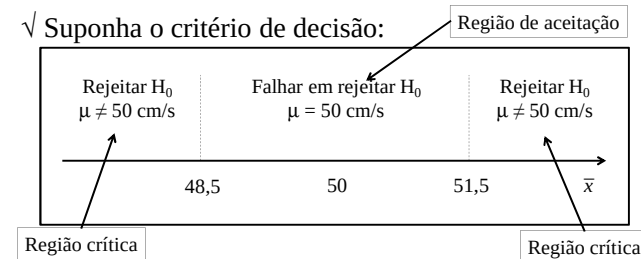
- Taxa de queima de propelente
- Hipóteses:
 - ✓ $H_0: \mu = 50 \text{ cm/s}$ vs. $H_1: \mu \neq 50 \text{ cm/s}$
- Amostra de tamanho $n = 10$
- Estatística de teste: $\bar{x}$
 - ✓ Estimativa da média verdadeira

- Critério para decisão:

- √ Se a estimativa cair próxima ao valor da hipótese ($\mu=50$) é evidência que a média verdadeira é realmente 50 cm/s.
- √ Estimativa consideravelmente diferente de 50 cm/s é uma evidência que H_1 é válida.

- A média amostral pode assumir muitos valores diferentes

√ Suponha o critério de decisão:



- Não rejeitaremos H_0 se:
 $48,5 \leq \bar{x} \leq 51,5$
- Rejeitaremos H_0 em favor de da hipótese alternativa H_1 se:
 $\bar{x} < 48,5$ ou $\bar{x} \geq 51,5$

- Valores críticos:

- √ Limites entre as regiões críticas e de aceitação
- √ No exemplo:
 - 48,5 e 51,5

- Conclusões referem-se à H_0 :

- √ Rejeição de H_0 em favor de H_1 .
 - Estatística de teste cai na região crítica
- √ Falha em rejeitar H_0 :
 - Estatística de teste cai na região de aceitação

- Conclusões erradas:

- √ Taxa média de queima do propelente é 50 cm/s e média amostral cai na região crítica
 - A decisão de rejeitar H_0 está errada
- √ Taxa média de queima do propelente não é 50 cm/s e média amostral cai na região de aceitação
 - A decisão de não rejeitar H_0 está errada

Decisões no Teste de Hipóteses

- Erro tipo I:
√ Rejeitar H_0 quando ela for realmente verdadeira
- Erro tipo II:
√ Falhar em rejeitar H_0 quando ela for realmente falsa

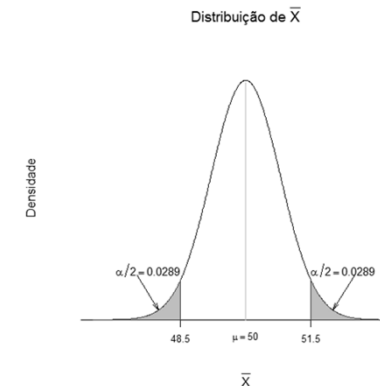
Decisão	Situação Real	
	H_0 é Verdadeira	H_0 é Falsa
Falhar em rejeitar H_0	Nenhum erro	Erro tipo II
Rejeitar H_0	Erro tipo I	Nenhum erro

- Decisão está baseada em variáveis aleatórias
√ Podem-se associar probabilidade com os erros
- Probabilidade de erro tipo I:
√ $\alpha = P\{\text{erro tipo I}\} = P\{\text{rejeitar } H_0 \mid H_0 \text{ é verdadeira}\}$
√ Nível de significância do teste ou erro α ou tamanho do teste

- No exemplo:
√ Taxa de queima com média verdadeira $\mu = 50$ cm/s e desvio-padrão $\sigma = 2,5$ cm/s
√ Taxa de queima tem distribuição para a qual se aplicam as condições do TCL

$$\begin{aligned}
 \bar{X} &\sim \text{Normal}\left(\mu = 50, \sigma = \frac{2,5}{\sqrt{10}}\right) \\
 \alpha &= P\{\bar{X} < 48,5 \mid \mu = 50\} + P\{\bar{X} > 51,5 \mid \mu = 50\} \\
 &= P\left\{Z < \frac{48,5 - 50}{0,79}\right\} + P\left\{Z > \frac{51,5 - 50}{0,79}\right\} \\
 &= P\{Z < -1,90\} + P\{Z > 1,90\} \\
 &= 0,0287 + 0,0287 = 0,0574
 \end{aligned}$$

- √ 5,74% de todas as amostras aleatórias conduziram à rejeição de H_0 . (se μ verdadeiro for 5 cm/s)



- É possível reduzir α .
√ Como?

- Alargamento da região de aceitação

✓ Valores críticos: 48 e 52

$$\begin{aligned}\alpha &= P\{\bar{X} < 48 | \mu = 50\} + P\{\bar{X} > 52 | \mu = 50\} \\ &= P\left\{Z < \frac{48 - 50}{0,79}\right\} + P\left\{Z > \frac{52 - 50}{0,79}\right\} \\ &= P\{Z < -2,53\} + P\{Z > 2,53\} \\ &= 0,0057 + 0,0057 = 0,0114\end{aligned}$$

- Aumento no tamanho amostral

✓ $n = 16$

$$\begin{aligned}\alpha &= P\{\bar{X} < 48,5 | \mu = 50\} + P\{\bar{X} > 51,5 | \mu = 50\} \\ &= P\left\{Z < \frac{48,5 - 50}{0,625}\right\} + P\left\{Z > \frac{51,5 - 50}{0,625}\right\} \\ &= P\{Z < -2,40\} + P\{Z > 2,40\} \\ &= 0,0082 + 0,0082 = 0,0164\end{aligned}$$

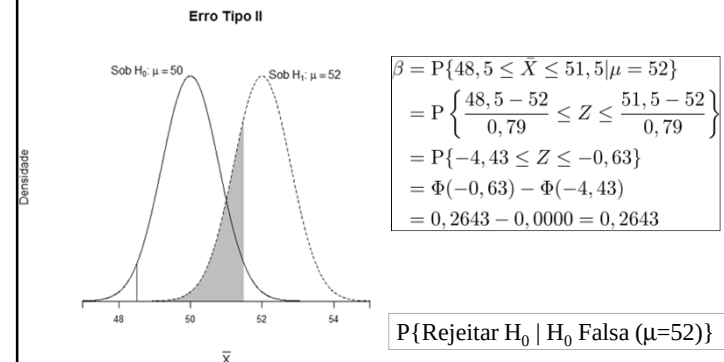
- Probabilidade de erro tipo II:

✓ $\beta = P\{\text{erro tipo II}\} = P\{\text{falhar em rejeitar } H_0 \mid H_0 \text{ é falsa}\}$

✓ Erro β .

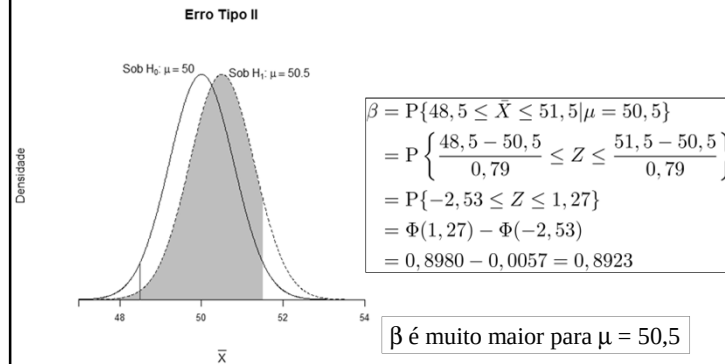
✓ Para calcular β temos de ter uma hipótese alternativa específica

- Erro tipo II quando $\mu = 52$ e $n = 10$



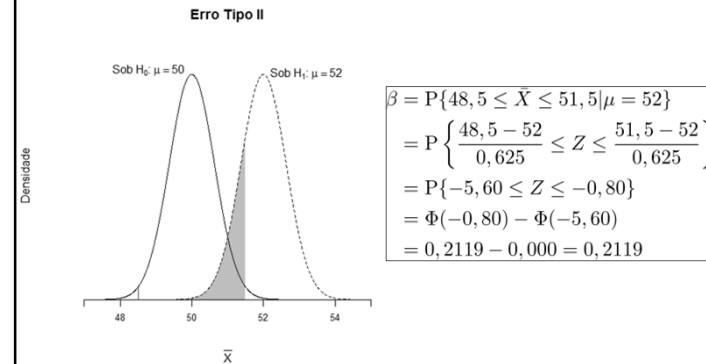
✓ β será também 0,264 se a média verdadeira for $\mu=48$

- Erro tipo II quando $\mu = 50,5$ e $n = 10$



✓ β aumenta rapidamente à medida que a média verdadeira μ se aproxima do valor de H_0 .

- Erro tipo II quando $\mu = 52$ e $n = 16$



✓ β depende também do tamanho da amostra
– Aumentar a amostra, implica diminuição de β .

- Tabela comparativa

Região de Aceitação	Tamanho amostral	α	β em $\mu=52$	β em $\mu=50,5$
$48,5 < \bar{x} < 51,5$	10	0,0576	0,2643	0,8923
$48 < \bar{x} < 52$	10	0,0114	0,5000	0,9705
$48,81 < \bar{x} < 51,19$	16	0,0576	0,0966	0,8606
$48,42 < \bar{x} < 51,58$	16	0,0114	0,2515	0,9578

- ✓ Escolha limites região crítica controla α .
- ✓ Erros tipo I e tipo II estão relacionados
 - Diminuir um deles implicar aumentar o outro
- ✓ Aumento amostral reduz β (c/ α constante)
- ✓ β aumenta quando μ se aproxima de H_0 .
(se H_0 for falsa)

Comentários

- α , em geral, é controlada com escolha de pontos críticos
- β depende do verdadeiro valor do parâmetro
 - ✓ Não é constante
 - ✓ Depende do tamanho da amostra
 - ✓ Depende da extensão com que H_0 é falsa

- Rejeitar H_0 é uma conclusão forte
 - ✓ O erro de rejeitar H_0 equivocadamente é controlado.
- Aceitar H_0 é uma conclusão fraca
 - ✓ Pode indicar a necessidade de mais dados para se obter uma conclusão forte
 - ✓ Deve-se dizer “Falha em rejeitar H_0 ”.

Poder

- É a probabilidade de rejeitar H_0 quando H_0 é falsa.

$$P_d = 1 - \beta$$

- ✓ É a probabilidade de rejeitar corretamente H_0 .
- ✓ Em geral, testes estatísticos são comparados através de seu poder
- ✓ Se o poder for julgado muito baixo, o analista poderá tanto α como o tamanho da amostra

- Poder é uma medida muito descritiva e concisa da **sensibilidade** de um teste estatístico
 - ✓ Sensibilidade é a habilidade em perceber diferenças
 - ✓ Exemplo:
Se a média verdadeira for realmente 52 cm/s, o teste rejeitará H_0 corretamente (perceberá a diferença) em 73,57% das vezes.

Hipóteses Unilaterais e Bilaterais

- H_0 é estabelecida sempre como uma igualdade
 - ✓ Deseja-se controlar α em um valor especificado
- H_1 pode ser bilateral ou unilateral

Exemplo 9-1

- Taxa de queima de propelente
 - ✓ Deseja-se que a conclusão forte seja quando a taxa de queima for menor do que 50 cm/s
 - ✓ Hipóteses:
 - $H_0: \mu = 50$
 - $H_1: \mu < 50$
 - ✓ Falhar em rejeitar H_0 não significa que $\mu = 50$
 - Significa que não temos forte evidência para suportar H_0 .

Exemplo – Situação difícil de ser formulada

- ✓ Engarrafador compra garrafas mas quer estar certo de que elas atendem especificação

– Resistência mínima é 200 psi

- ✓ Formulações possíveis:

- $H_0: \mu = 200$ psi Rejeitar $H_0 \rightarrow$ garrafas são satisfatórias
- $H_1: \mu > 200$ psi Fabricante deve demonstrar conformidade

– Há probabilidade de não rejeitar H_0 mesmo que $\mu > 200$

- $H_0: \mu = 200$ psi Rejeitar $H_0 \rightarrow$ garrafas não conformes
- $H_1: \mu < 200$ psi Pequenos desvios são aceitáveis

– Há probabilidade de aceitar H_0 mesmo que $\mu < 200$

- Rejeitar H_0 é sempre uma conclusão forte
 - ✓ O que é importante para fazer uma conclusão forte em H_1 ?
 - Depende do ponto de vista e da experiência na situação

Resultado do Teste

- Conclusão sobre resultado do teste
 - ✓ H_0 foi ou não rejeitado a um nível de significância α .
 - ✓ Não dá ideia a respeito da posição da estatística de teste na região de rejeição

Probabilidade de Significância (Valor-p)

- Expressa a conclusão final de um teste
- É a probabilidade de se obter resultados mais extremos do observado, sob H_0 .
 - ✓ $p = P\{\text{estatísticas iguais ou superiores} \mid H_0\}$
 - ✓ Qual a probabilidade de outra amostra apresentar resultados mais favoráveis (estatísticas de teste maiores) para rejeitar H_0 ?

- Se o p-valor é pequeno há duas conjecturas:
 - ✓ H_0 é verdadeira e ocorreu um evento extremamente raro
 - ✓ H_0 não deve ser verdadeira, pois a conjectura inicial não parece ser plausível
- Quanto menor o p-valor, maior a evidência para se rejeitar H_0
 - ✓ De maneira geral $p \leq 0,05$ indica que há diferenças significativas entre grupos

- p-valor carrega muita informação sobre o peso da evidência contra H_0 .
- p-valor é o menor nível de significância que conduz à rejeição de H_0 , com os dados amostrais
- Estatística de teste signifiante:
 - ✓ H_0 é rejeitado
 - ✓ p-valor é o menor nível α em que os dados são significantes
(ponto crítico é a estatística de teste)

Exemplo – Lançamento Moeda

- Você lança uma moeda 20 vezes e obtém 16 caras e 4 coroas.
 - ✓ H_0 : a moeda é honesta ($p_c = p_k = 0,50$)
 - ✓ Qual é a chance de obter 16 ou mais caras (ou 4 ou menos caras) lançando uma moeda honesta?
 - ✓ Resp.: $2 \times 0,59\% = 1,18\%$
 - É o mesmo que obter 4 coroas ou menos

- 200 alunos lançarão uma moeda 20 vezes e anotarão a quantidade de caras
 - √ Espera-se obter pelo menos 16 caras (ou coroas) em 1,18% dos experimentos.
 - √ Não seria surpresa encontrar um mais experimento com pelo menos 16 caras (ou coroas)
 - √ Esse é um caso de comparação múltipla!

- Não é sempre fácil calcular o valor exato de p para um teste
 - √ Pacotes estatísticos reportam o p -valor

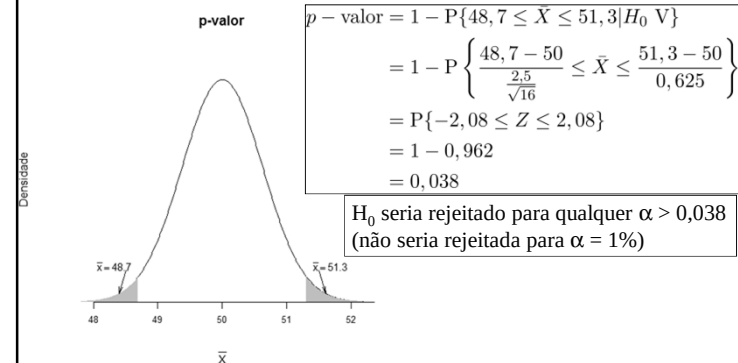
Exemplo

- $H_0: \mu = 50$ vs. $H_1: \mu \neq 50$

√ População com $\sigma = 2,5$ cm/s

$$\bar{x} = 51,3 \text{ cm/s}$$

$$n = 16$$



Testes de Hipóteses e Intervalos de Confiança

- Há relação entre Teste de Hipóteses sobre um parâmetro θ e o Intervalo de Confiança para θ
 - √ Intervalo de $(1 - \alpha) \times 100\%$ de confiança para θ :
 - √ $[l, u]$
 - √ Teste de Hipóteses com nível de significância α para θ
 - $H_0: \theta = \theta_0$.
 - $H_1: \theta \neq \theta_0$.
 - √ O Teste de Hipóteses conduzirá à rejeição de H_0 se, e somente se, θ_0 não pertencer ao IC $[l, u]$.

Exemplo

- Taxa de queima de propelente ($\sigma = 2,5$ cm/s)

✓ $H_0: \mu = 50$ vs. $H_1: \mu \neq 50$

✓ Dados amostrais: $\bar{x} = 51,3$ cm/s e $n = 16$

✓ p-valor: 0,038

✓ H_0 é rejeitada ao nível $\alpha = 5\%$

✓ Intervalo bilateral com 95% de confiança para μ .

$$51,3 - 1,96 \frac{2,5}{\sqrt{16}} \leq \mu \leq 51,3 + 1,96 \frac{2,5}{\sqrt{16}}$$

$$50,075 \leq \mu \leq 52,525$$

$$50 \notin [50,075; 52,525]$$

H_0 é rejeitada

- Embora equivalentes, o Teste de Hipóteses e o Intervalo de Confiança fornecem conhecimento diferentes.

✓ Intervalo de confiança:

- Faixa de valores prováveis para μ em um nível estabelecido de confiança

✓ Teste de Hipóteses:

- Estrutura para dispor os níveis de risco associados com decisão específica

Procedimento Geral para Teste de Hipóteses

1. A partir do contexto do problema, identifique o parâmetro de interesse
2. Estabeleça a hipótese nula H_0 .
3. Estabeleça uma hipótese alternativa adequada, H_1 .
4. Escolha um nível de significância, α .
5. Determine uma estatística de teste apropriada
6. Estabeleça a região de rejeição para a estatística
7. Calcule a estatística de teste observada
8. Decida se H_0 deve ou não ser rejeitada e conclua no contexto do problema

- Uma vez que o experimento tenha sido planejado (passos de 1 a 4), são requeridas as seguintes etapas:

1. Especificar a estatística de teste a ser usada
2. Especificar a localização da região crítica (bilateral, unilateral inferior, unilateral superior)
3. Especificar os critérios de rejeição (o valor de a ou o p-valor no qual a região deveria ocorrer)

Significância Estatística vs. Significância Prática

- Valor-p pequeno indica estatisticamente significativo
 - ✓ H_0 deve ser rejeitado em favor de H_1 .
 - ✓ O desvio real de H_0 detectado pode ter pouca significância prática
(isto é particularmente verdade quando n é grande)

- Para n grandes, o valor observado 50,5 sugere que $H_0: \mu = 50$ seja rejeitada
 - ✓ Mesmo que do ponto de vista prático, μ não difira muito do valor usado na hipótese ($\mu_0 = 50$)
- Amostra grande quase sempre conduzirá à rejeição de H_0 (a um nível fixo de significância)
 - ✓ Mesmo se houver pouca diferença prática entre a média verdadeira e o valor usado na hipótese

Exemplo

- Valor-p para testar $H_0: \mu = 50$
 - ✓ Para valor observado: $\bar{x} = 50,5 \text{ cm/s}$

Tamanho amostral (n)	p -valor	Poder ⁽¹⁾
10	0,527	0,097
25	0,317	0,170
50	0,157	0,293
100	0,046	0,516
400	$6,3 \times 10^{-5}$	0,979
1000	$2,5 \times 10^{-10}$	1,000

- ⁽¹⁾Quando μ verdadeiro = 50,5 ($\alpha = 5\%$)

- “Seja cuidadoso ao interpretar os resultados do teste de hipóteses quando a amostra tiver tamanho grande, visto que qualquer pequeno desvio do valor usado em H_0 (μ_0) será provavelmente detectado, mesmo quando a diferença for de pouca ou nenhuma significância prática”

Testes para a Média – População Normal com Variância Conhecida

Teste de Hipóteses para a Média

- Seja uma amostra aleatória $X_1, X_2, \dots, X_n$ proveniente de uma única população **normal** com variância σ^2 conhecida.
- Hipóteses:
 $H_0: \mu = \mu_0$
 $H_1: \mu \neq \mu_0$
- Nível de significância do teste: α .
- Estatística de teste: $Z_0 = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}} \sim N(0, 1)$

Média Amostral

- Considere uma amostra aleatória $X_1, X_2, \dots, X_n$ proveniente de uma população qualquer, com média μ e variância σ^2 .

✓ Média amostral é estimador não-viciado de μ .

✓ Distribuição amostral da média amostral:

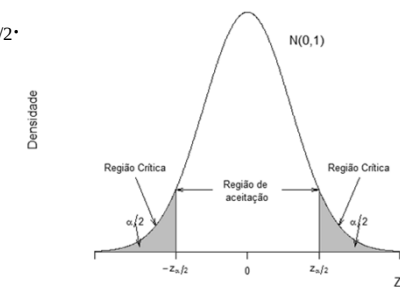
- Para amostras pequenas dependerá da distribuição da população, suas esperança e variância serão:

$$E[\bar{X}] = \mu$$
$$\text{Var}[\bar{X}] = \frac{\sigma^2}{n}$$

- Região de rejeição:

✓ $H_1: \mu \neq \mu_0$. (alternativa bilateral)

✓ $z_0 > z_{\alpha/2}$ ou $z_0 < -z_{\alpha/2}$.



✓ Pontos críticos: $-z_{\alpha/2}$ e $z_{\alpha/2}$.

✓ Regra de decisão:

- Rejeitar H_0 se $z_0 > z_{\alpha/2}$ ou $z_0 < -z_{\alpha/2}$.

- Região crítica expressa em termos do valor observado da média amostral:

✓ Rejeite $H_0: \mu = \mu_0$ se: $\bar{x} > a$ ou $\bar{x} < b$
em que:

$$a = \mu_0 + z_{\alpha/2} \frac{\sigma}{\sqrt{n}}, \text{ e}$$
$$b = \mu_0 - z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

Exemplo 9-2

- Taxa de queima de propelente é característica importante do produto
 - ✓ Especificações exigem que taxa de queima seja 50 cm/s
 - ✓ Sabe-se que desvio-padrão taxa de queima é $\sigma = 2$ cm/s
 - ✓ Analista especifica nível de significância $\alpha = 5\%$
 - ✓ Seleciona-se amostra aleatória de $n = 25$
 - ✓ Média amostral observada: $\bar{x} = 51,3$
 - ✓ Que conclusão o analista toma?

- Aplicação de Procedimento:

1. Parâmetro de interesse:
 - ✓ μ : taxa média de queima
2. $H_0: \mu = 50$ cm/s
3. $H_1: \mu \neq 50$ cm/s
4. $\alpha = 0,05$
5. Estatística de teste: $z_0 = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$
6. Rejeite H_0 se $z_0 > 1,96$ ou $z_0 < -1,96$
 - ✓ Valores críticos para $\alpha = 5\%$: $z_{\alpha/2} = 1,96$ e $-z_{\alpha/2} = -1,96$
7. Cálculo estatística de teste observada

$$z_0 = \frac{51,3 - 50}{\frac{2}{\sqrt{25}}} = 3,25$$

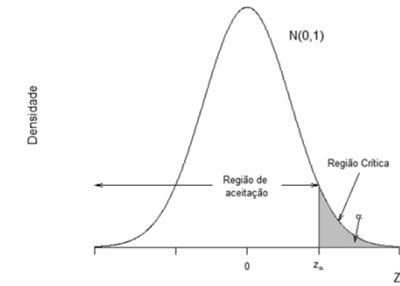
8. Conclusão:

- ✓ $z_0 = 3,25 > 1,96$, logo rejeitamos $H_0: \mu = 50$ cm/s, com nível de significância de 5%
- ✓ Concluimos que a taxa média de queima difere de 50 cm/s, baseados em uma amostra de 25 medidas
- ✓ Há forte evidência de que a taxa média de queima exceda 50 cm/s.

Hipótese Alternativa Unilateral Superior

- Hipóteses:
 - ✓ $H_0: \mu = \mu_0$.
 - ✓ $H_1: \mu > \mu_0$.
- Comentários:
 - ✓ Valor negativo de Z_0 não conduz à rejeição de H_0 .
 - ✓ Região crítica na extremidade superior
 - ✓ Rejeição de H_0 :
 - $z_0 > z_{\alpha}$.

- Região de rejeição:
 - ✓ $H_1: \mu > \mu_0$. (alternativa unilateral superior)
 - ✓ $z_0 > z_{\alpha}$.

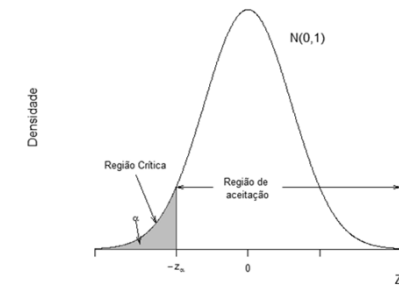


- ✓ Pontos crítico: z_{α} .
- ✓ Regra de decisão:
 - Rejeitar H_0 se $z_0 > z_{\alpha}$.

Hipótese Alternativa Unilateral Inferior

- Hipóteses:
 - ✓ $H_0: \mu = \mu_0$.
 - ✓ $H_1: \mu < \mu_0$.
- Comentários:
 - ✓ Valores pequenos de Z_0 conduzem à rejeição de H_0 .
 - ✓ Região crítica na extremidade inferior
 - ✓ Rejeição de H_0 :
 - $z_0 < -z_{\alpha}$.

- Região de rejeição:
 - ✓ $H_1: \mu < \mu_0$. (alternativa unilateral inferior)
 - ✓ $z_0 < -z_{\alpha}$.



- ✓ Pontos crítico: z_{α} .
- ✓ Regra de decisão:
 - Rejeitar H_0 se $z_0 < -z_{\alpha}$.

Testes para Média (Variância Conhecida)

- Hipótese nula:

✓ $H_0: \mu = \mu_0$.

- Estatística de teste:

$$Z_0 = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

Hipótese alternativa	Critério de rejeição
$H_1: \mu \neq \mu_0$	$z_0 > z_{\alpha/2}$ ou $z_0 < -z_{\alpha/2}$.
$H_1: \mu > \mu_0$	$z_0 > z_{\alpha}$.
$H_1: \mu < \mu_0$	$z_0 < -z_{\alpha}$.

Valor-p

- Para testes com a estatística Z_0 :

✓ Cálculo do valor-p:

p	Teste	Hipóteses	
$2[1 - \Phi(z_0)]$	Bilateral	$H_1: \mu = \mu_0$	$H_1: \mu \neq \mu_0$
$1 - \Phi(z_0)$	Bilateral superior	$H_1: \mu = \mu_0$	$H_1: \mu > \mu_0$
$\Phi(z_0)$	Bilateral inferior	$H_1: \mu = \mu_0$	$H_1: \mu < \mu_0$

✓ $\Phi(z)$: função de distribuição acumulada da $N(0, 1)$

$$\Phi(z) = \Pr(Z \leq z)$$

- No exemplo:

✓ Hipótese alternativa bilateral

✓ Estatística de teste observada:

- $z_0 = 3,25$

✓ Valor-p:

$$p = 2[1 - \Phi(z_0)] = 2[1 - \Phi(3,25)] = 0,0012$$

✓ $H_0: \mu = 50$ seria rejeitada com qualquer nível de significância $\alpha \geq 0,0012$

– H_0 seria rejeitada se $\alpha = 0,01$

– H_0 não seria rejeitada se $\alpha = 0,001$

Erro Tipo II

- Em um teste de hipóteses:

✓ Analista seleciona diretamente α .

✓ Probabilidade de erro tipo II (β) depende da escolha do tamanho da amostra

- Considere a hipótese bilateral:

$$\checkmark H_0: \mu = \mu_0$$

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

- Suponha que o verdadeira valor da média seja

$$\checkmark \mu = \mu_0 + \delta, \delta > 0 \text{ (} H_0 \text{ é falsa!)}$$

- Estatística de teste Z_0 é:

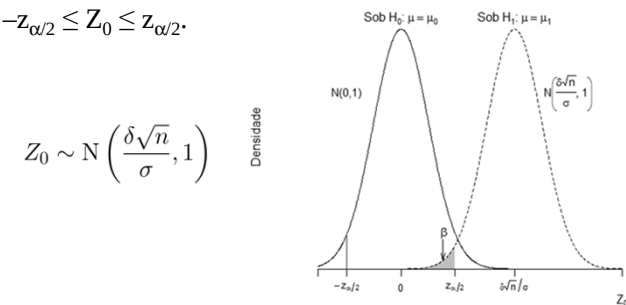
$$Z_0 = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{\bar{X} - (\mu_0 + \delta)}{\frac{\sigma}{\sqrt{n}}} + \frac{\delta\sqrt{n}}{\sigma}$$

- Quando H_1 for verdadeira ($\delta > 0$):

$$Z_0 \sim N\left(\frac{\delta\sqrt{n}}{\sigma}, 1\right)$$

- Comete-se erro tipo II quando H_1 é verdadeira e:

$$\checkmark -z_{\alpha/2} \leq Z_0 \leq z_{\alpha/2}.$$



- Probabilidade de erro tipo II para teste bilateral para a média (variância conhecida)

$$\begin{aligned} \beta &= \Pr \left[-z_{\alpha/2} \leq Z_0 \leq z_{\alpha/2} | Z_0 \sim N\left(\frac{\delta\sqrt{n}}{\sigma}, 1\right) \right] \\ &= \Pr \left(-z_{\alpha/2} - \frac{\delta\sqrt{n}}{\sigma} \leq Z_0 - \frac{\delta\sqrt{n}}{\sigma} \leq z_{\alpha/2} - \frac{\delta\sqrt{n}}{\sigma} \right) \end{aligned}$$

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

- A equação se mantém se $\delta < 0$

- Se $\delta = 0$, tem-se a probabilidade de não rejeitar H_0 quando ela for verdadeira ($1 - \alpha$)

- Considere a hipótese unilateral superior:

$$\checkmark H_0: \mu = \mu_0$$

$$\checkmark H_1: \mu > \mu_0$$

- Suponha que o verdadeira valor da média seja

$$\checkmark \mu = \mu_0 + \delta, \delta > 0 \text{ (} H_0 \text{ é falsa!)}$$

- Estatística de teste Z_0 é:

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

- Quando H_1 for verdadeira ($\delta > 0$):

$$Z_0 \sim N\left(\frac{\delta\sqrt{n}}{\sigma}, 1\right)$$

- ✓ Probabilidade de erro tipo II para teste unilateral superior para a média (variância conhecida)

$$\begin{aligned}\beta &= \Pr \left[Z_0 \leq z_\alpha | Z_0 \sim N \left(\frac{\delta\sqrt{n}}{\sigma}, 1 \right) \right] \\ &= \Pr \left(Z_0 - \frac{\delta\sqrt{n}}{\sigma} \leq z_\alpha - \frac{\delta\sqrt{n}}{\sigma} \right) \\ \beta &= \Phi \left(z_\alpha - \frac{\delta\sqrt{n}}{\sigma} \right)\end{aligned}$$

- Considere a hipótese unilateral inferior:

$$\checkmark H_0: \mu = \mu_0$$

$$\checkmark H_1: \mu < \mu_0$$

- Suponha que o verdadeira valor da média seja

$$\checkmark \mu = \mu_0 + \delta, \delta < 0 \text{ (} H_0 \text{ é falsa!)}$$

- Estatística de teste Z_0 é:

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

- ✓ Quando H_1 for verdadeira ($\delta < 0$):

$$Z_0 \sim N \left(\frac{\delta\sqrt{n}}{\sigma}, 1 \right)$$

- ✓ Probabilidade de erro tipo II para teste unilateral inferior para a média (variância conhecida)

$$\begin{aligned}\beta &= \Pr \left[Z_0 > -z_\alpha | Z_0 \sim N \left(\frac{\delta\sqrt{n}}{\sigma}, 1 \right) \right] \\ &= \Pr \left(Z_0 - \frac{\delta\sqrt{n}}{\sigma} > -z_\alpha - \frac{\delta\sqrt{n}}{\sigma} \right) \\ \beta &= 1 - \Phi \left(-z_\alpha - \frac{\delta\sqrt{n}}{\sigma} \right)\end{aligned}$$

Escolha do Tamanho da Amostra

- Tamanho da amostra para se obter um valor particular de β , para um dado δ e α .

- ✓ Hipótese alternativa bilateral ($H_1: \mu \neq \mu_0$)

- Probabilidade de erro tipo II:

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

$$\checkmark \text{ Se } \delta > 0 \quad \Phi \left(-z_{\alpha/2} - \frac{\delta\sqrt{n}}{\sigma} \right) \approx 0, \text{ então: } \beta \approx \Phi \left(z_{\alpha/2} - \frac{\delta\sqrt{n}}{\sigma} \right)$$

- ✓ Seja z_β o percentil 100β superior da normal padrão

$$\beta = \Phi(-z_\beta)$$

$$-z_\beta \approx z_{\alpha/2} - \frac{\delta\sqrt{n}}{\sigma}$$

$$-z_\beta \approx z_{\alpha/2} - \frac{\delta\sqrt{n}}{\sigma}$$

- Tamanho da amostra para um teste bilateral para a média (variância conhecida)

$$n \approx \frac{(z_\beta + z_{\alpha/2})^2}{d^2}, \text{ em que } d = \frac{\mu - \mu_0}{\sigma} = \frac{|\delta|}{\sigma}$$

- ✓ Convencionou-se arredondar o tamanho da amostra para o maior inteiro mais próximo
- ✓ Essa aproximação é boa quando $\Phi(-z_{\alpha/2} - \delta\sqrt{n}/\sigma)$ é pequena comparada a β .

- Tamanho da amostra para se obter um valor particular de β , para um dado δ e α .

✓ Teste unilateral superior ($H_1: \mu > \mu_0, \delta > 0$)

- Probabilidade de erro tipo II:

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

$$\beta = \Phi(-z_\beta) \text{ e } -z_\beta = z_\alpha - \frac{\delta\sqrt{n}}{\sigma}$$

- ✓ Tamanho da amostra para um teste unilateral superior para a média (variância conhecida)

$$n = \frac{(z_\beta + z_\alpha)^2}{d^2}$$

- Tamanho da amostra para se obter um valor particular de β , para um dado δ e α .

✓ Teste unilateral inferior ($H_1: \mu < \mu_0, \delta > 0$)

- Probabilidade de erro tipo II:

$$\beta = 1 - \Phi\left(-z_\alpha - \frac{\delta\sqrt{n}}{\sigma}\right)$$

$$\beta = \Phi(z_\beta) \text{ e } z_\beta = -z_\alpha - \frac{\delta\sqrt{n}}{\sigma}$$

- ✓ Tamanho da amostra para um teste unilateral inferior para a média (variância conhecida)

$$n = \frac{(z_\beta + z_\alpha)^2}{d^2}$$

- Para **qualquer** hipótese unilateral:

- ✓ Tamanho da amostra para produzir um erro especificado do tipo II, com probabilidade β , dados δ e α :

$$n = \frac{(z_\beta + z_\alpha)^2}{d^2}, \text{ em que } \delta = \mu - \mu_0$$

Exemplo 9-3

- Taxa de queima de propelente

✓ Taxa verdadeira de queima é 49 cm/s ($\delta = 1$)

✓ $n = 25$

✓ $\sigma = 2$ cm/s

✓ Valor de β para teste bilateral

$$\begin{aligned}\beta &= \Phi\left(z_{\alpha/2} - \frac{\delta\sqrt{n}}{\sigma}\right) - \Phi\left(-z_{\alpha/2} - \frac{\delta\sqrt{n}}{\sigma}\right) \\ &= \Phi\left(1,96 - \frac{\sqrt{25}}{2}\right) - \Phi\left(-1,96 - \frac{\sqrt{25}}{2}\right) \\ &= \Phi(-0,54) - \Phi(-4,46) = 0,295\end{aligned}$$

- Probabilidade de cerca de 30% de que essa diferença não seja percebida

✓ Planejamento do teste de maneira que teste detectará, com alta probabilidade (90%), diferenças entre a média verdadeira e $\mu_0 = 50$ de não mais que $\delta = 1$ cm/s

$$\begin{aligned}n &\approx \frac{(z_\beta + z_{\alpha/2})^2 \sigma^2}{\delta^2} \\ &\approx \frac{(1,28 + 1,96)^2 2^2}{(1)^2} \approx 42\end{aligned}$$

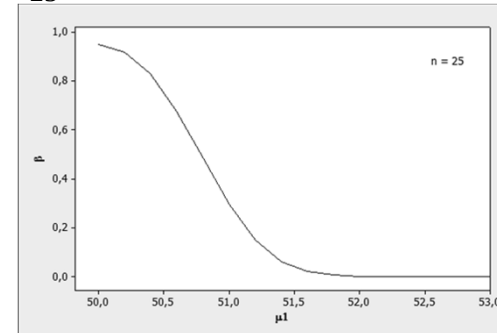
– A aproximação é boa.

$$\Phi\left(-z_{\alpha/2} - \frac{\delta\sqrt{n}}{\sigma}\right) = \Phi\left(-1,96 - \frac{(1)\sqrt{42}}{2}\right) = \Phi(-5,20) \approx 0$$

é pequena em relação a β .

- Probabilidade não rejeitar H_0 vs. média verdadeira

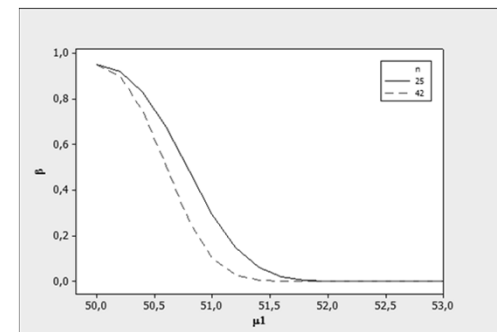
✓ $n = 25$



✓ A curva é simétrica com relação às médias verdadeiras menores que 50

- Probabilidade não rejeitar H_0 vs. média verdadeira

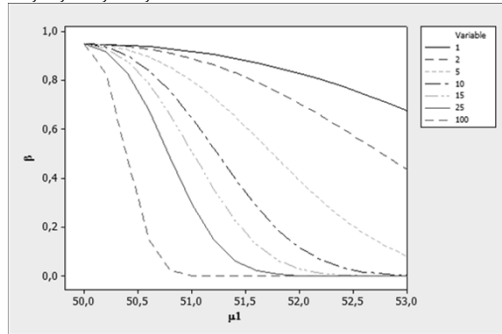
✓ $n = 42$



✓ Aumento do tamanho amostral melhora poder do teste

- Probabilidade não rejeitar H_0 vs. média verdadeira

✓ $n = 1, 2, 10, 15, 25$ e 100



✓ β é calculado em uma escala de σ !

- Probabilidade de erro tipo II:

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

✓ Considerando uma escala σ :

$$d = \frac{\mu_1 - \mu_0}{\sigma} = \frac{\delta}{\sigma}$$

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

✓ Teremos β calculado para qualquer μ_1 .

Curvas Características de Operação

- Curvas de β vs. um parâmetro d , para vários tamanhos amostrais

✓ Gráficos VIIa e VIIb – Apêndice

✓ Curvas para $\alpha = 0,05$ e $\alpha = 0,01$

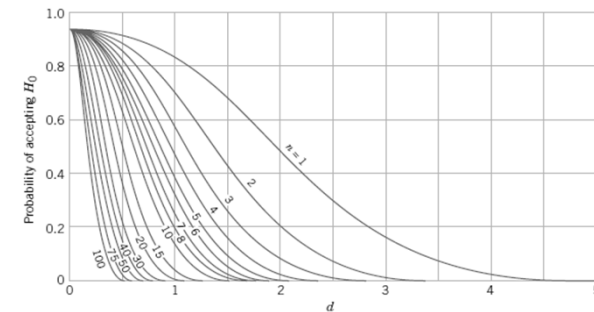
- Parâmetro d : $d = \frac{|\mu_1 - \mu_0|}{\sigma} = \frac{|\delta|}{\sigma}$

✓ d : Fator de não centralidade

✓ Conjunto de curvas características pode ser usado para todos os problemas, independente dos valores de μ_0 e σ .

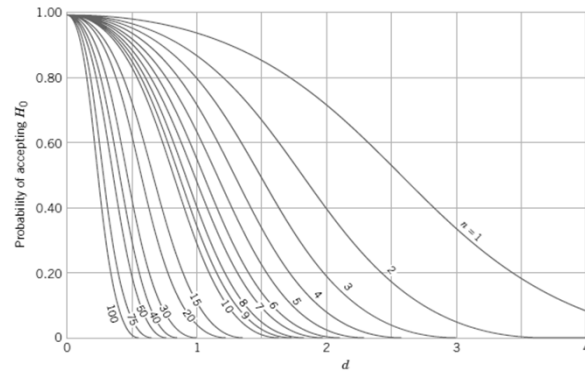
- Teste bilateral e nível de significância $\alpha = 0,05$

Chart VI Operating Characteristic Curves



(a) O.C. curves for different values of n for the two-sided normal test for a level of significance $\alpha = 0.05$.

- Teste bilateral e nível de significância $\alpha = 0,01$



(b) O.C. curves for different values of n for the two-sided normal test for a level of significance $\alpha = 0.01$.

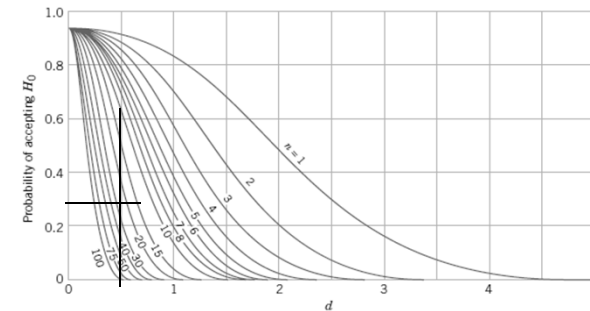
- Comentários:

- ✓ Quanto mais distante a média verdadeira estiver de μ_0 , menor será o valor de β , para um dado n e α .
- ✓ Para n e α especificados, são mais fáceis de detectar grandes diferenças na média
- ✓ Para um dado δ e α , a probabilidade β diminui à medida que n aumenta
 - Podemos tornar um teste mais potente aumentando n .

Exemplo 9-4

- Taxa de queima de propelente
 - ✓ $n = 25$
 - ✓ $\sigma = 2$ cm/s
 - ✓ $\alpha = 0,05$
 - ✓ Analista quer perceber se a taxa média verdadeira de queima for $\mu = 51$ cm/s
 - ✓ Erro tipo II a partir da Curva CO
 - $\delta = 51 - 50 = 1$
 - $d: d = \frac{|\mu_1 - \mu_0|}{\sigma} = \frac{|51 - 50|}{2} = 0,5$

Chart VI Operating Characteristic Curves



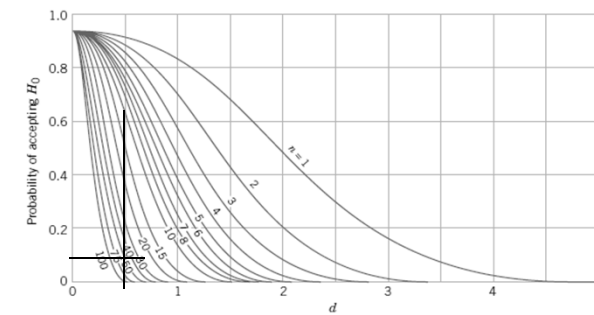
(a) O.C. curves for different values of n for the two-sided normal test for a level of significance $\alpha = 0.05$.

- ✓ Para $n = 25$, $\beta \approx 0,30$.

Exemplo 9-5

- Taxa de queima de propelente
 - √ Analista deseja que teste detecte diferença de não mais que 1 cm/s, com probabilidade alta (0,90).
 - $d = 0,5$
 - $\alpha = 0,05$
 - $\beta = 0,10$

Chart VI Operating Characteristic Curves

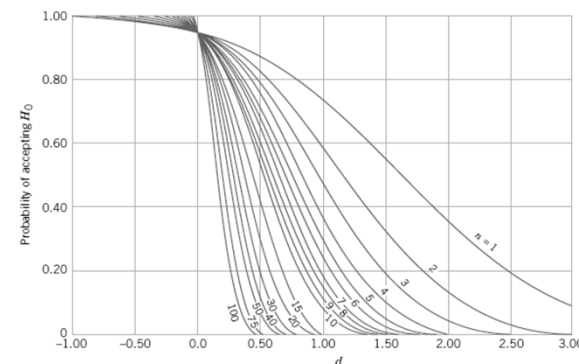


(a) O.C. curves for different values of n for the two-sided normal test for a level of significance $\alpha = 0.05$.

√ Tamanho amostra: $n \approx 40$

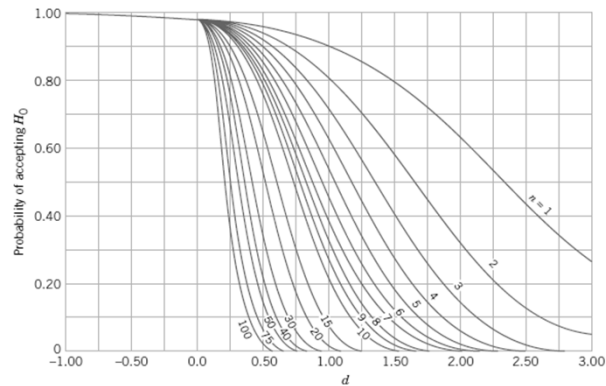
- Em geral, as curvas CO envolvem três parâmetros β , d e n .
- Aplicações típicas das curvas:
 - √ Para um dado n e d , encontre β .
 - Analista preocupado sobre sensibilidade de experimento já realizado ou com tamanho amostral restrito
 - √ Para um dado d e β , encontre n
 - Analista pode selecionar tamanho amostral no início do experimento

- Teste unilateral e nível de significância $\alpha = 0,05$



(c) O.C. curves for different values of n for the one-sided normal test for a level of significance $\alpha = 0.05$.

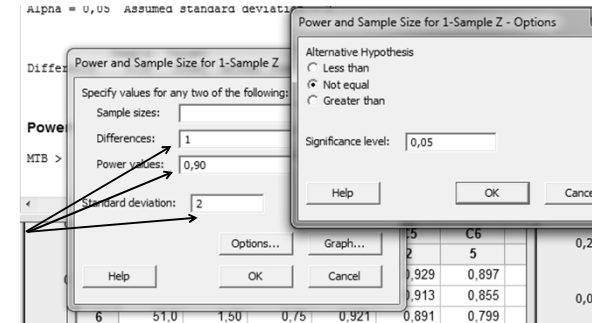
- Teste unilateral e nível de significância $\alpha = 0,01$



(d) O.C. curves for different values of n for the one-sided normal test for a level of significance $\alpha = 0.01$.

- Cálculo tamanho da amostra - Minitab

Stat > Power and Sample Size > 1-Sample Z

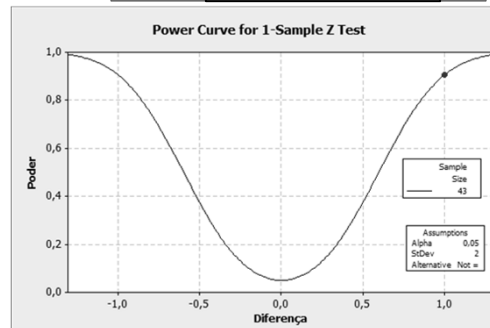


- Saída – Minitab

√ Dados:

- Diferença: 1
- Poder: 0,90

Power and Sample Size				
1-Sample Z Test				
Testing mean = null (versus not = null)				
Calculating power for mean = null + difference				
Alpha = 0,05 Assumed standard deviation = 2				
Difference	Sample Size	Target Power	Actual Power	
1	43	0,9	0,906375	



- Saída – Minitab

√ Dados:

- Diferença: 1
- Poder: 0,75

Power and Sample Size				
1-Sample Z Test				
Testing mean = null (versus not = null)				
Calculating power for mean = null + difference				
Alpha = 0,05 Assumed standard deviation = 2				
Difference	Sample Size	Target Power	Actual Power	
1	28	0,75	0,753578	

√ Tamanho amostral diminuiu para $n = 28$

- Saída – Minitab

√ Dados:

- Diferença: 1
- Tamanho amostra: 25

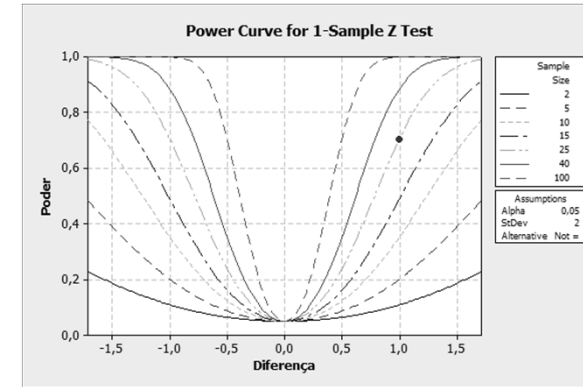
Power and Sample Size		
1-Sample Z Test		
Testing mean = null (versus not = null)		
Calculating power for mean = null + difference		
Alpha = 0,05 Assumed standard deviation = 2		
Difference	Sample Size	Power
1	25	0,705418

√ Poder = 0,7054 → $\beta = 0,2946$

- Os cálculos do computador são mais acurados que os valores da leitura visualizada a partir da curva CO.

- Taxa de queima de propelente

√ Curvas de poder vs. distância entre média verdadeira e H_0 .



Teste para Amostras Grandes

- Na maioria das situações práticas:
 - √ σ^2 é desconhecida.
 - √ Não podemos estar certo que a população seja bem modelada por uma distribuição normal
- Se a amostra for grande ($n > 40$):
 - √ Pode-se substituir σ pelo desvio-padrão amostral
 - √ Usa-se a estatística Z_0 para teste de hipótese sobre média da população
 - Baseado no TCL

Testes para a Média – População Normal com Variância Desconhecida

Teste de Hipóteses para a Média

- Seja uma amostra aleatória $X_1, X_2, \dots, X_n$ proveniente de uma única população **normal** com variância σ^2 **desconhecida**.

- Hipóteses:

$$H_0: \mu = \mu_0$$

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

- Nível de significância do teste: α .

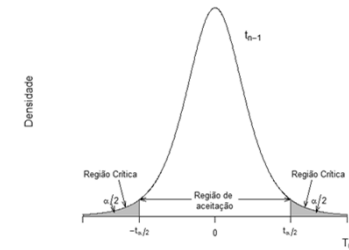
- Estatística de teste:
$$T_0 = \frac{\bar{X} - \mu_0}{\frac{S}{\sqrt{n}}} \stackrel{H_0}{\sim} t_{n-1}$$

✓ S: desvio-padrão amostral

- Região de rejeição:

✓ $H_1: \mu \neq \mu_0$. (alternativa bilateral)

✓ $t_0 > t_{\alpha/2, n-1}$ ou $t_0 < -t_{\alpha/2, n-1}$.



✓ Pontos críticos: $-t_{\alpha/2, n-1}$ e $t_{\alpha/2, n-1}$.

✓ Regra de decisão:

- Rejeitar H_0 se $t_0 > t_{\alpha/2, n-1}$ ou $t_0 < -t_{\alpha/2, n-1}$.

Hipótese Alternativa Unilateral Superior

- Hipóteses:

✓ $H_0: \mu = \mu_0$.

✓ $H_1: \mu > \mu_0$.

- Comentários:

✓ Valor negativo de T_0 não conduz à rejeição de H_0 .

✓ Região crítica na extremidade superior

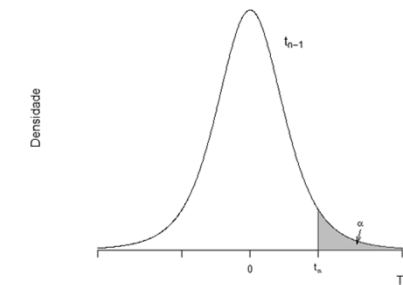
✓ Rejeição de H_0 :

- $t_0 > t_{\alpha, n-1}$.

- Região de rejeição:

✓ $H_1: \mu > \mu_0$. (alternativa unilateral superior)

✓ $t_0 > t_{\alpha, n-1}$.



✓ Pontos crítico: t_{α} .

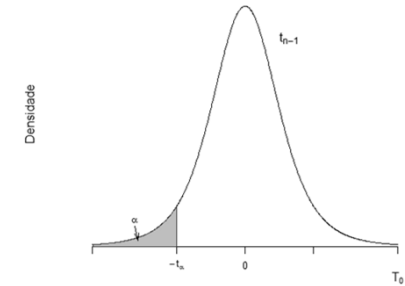
✓ Regra de decisão:

- Rejeitar H_0 se $t_0 > t_{\alpha, n-1}$.

Hipótese Alternativa Unilateral Inferior

- Hipóteses:
 - ✓ $H_0: \mu = \mu_0$.
 - ✓ $H_1: \mu < \mu_0$.
- Comentários:
 - ✓ Valores pequenos de T_0 conduzem à rejeição de H_0 .
 - ✓ Região crítica na extremidade inferior
 - ✓ Rejeição de H_0 :
 - $t_0 < -t_{\alpha, n-1}$.

- Região de rejeição:
 - ✓ $H_1: \mu < \mu_0$. (alternativa unilateral inferior)
 - ✓ $t_0 < -t_{\alpha, n-1}$.



- ✓ Pontos crítico: $t_{\alpha, n-1}$.
- ✓ Regra de decisão:
 - Rejeitar H_0 se $t_0 < -t_{\alpha, n-1}$.

Testes para Média (Variância Desconhecida)

- Hipótese nula:
 - ✓ $H_0: \mu = \mu_0$.
- Estatística de teste: $T_0 = \frac{\bar{X} - \mu_0}{\frac{S}{\sqrt{n}}}$

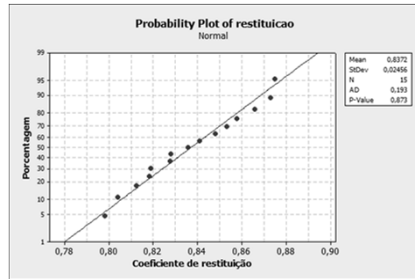
Hipótese alternativa	Critério de rejeição
$H_1: \mu \neq \mu_0$	$t_0 > t_{\alpha/2, n-1}$ ou $t_0 < -t_{\alpha/2, n-1}$.
$H_1: \mu > \mu_0$	$t_0 > t_{\alpha, n-1}$.
$H_1: \mu < \mu_0$	$t_0 < -t_{\alpha, n-1}$.

Exemplo 9-6

- Projeto do Taco de Golfe
 - ✓ Quantificação da razão entre velocidade de saída e velocidade de chegada da bola (coeficiente de restituição)
 - ✓ Experimento para determinar se há evidência (com $\alpha=5\%$) de suporte a afirmação de que o coeficiente médio de restituição exceda 0,82
 - ✓ Tamanho amostral: 15 tacos

- Valores observados:

Coef. Restituição		
n	$\bar{x}$	s
15	0,83725	0,02456



✓ Gráfico de probabilidade suporta a suposição de que o coeficiente de restituição é normalmente distribuído

- Procedimento de teste:

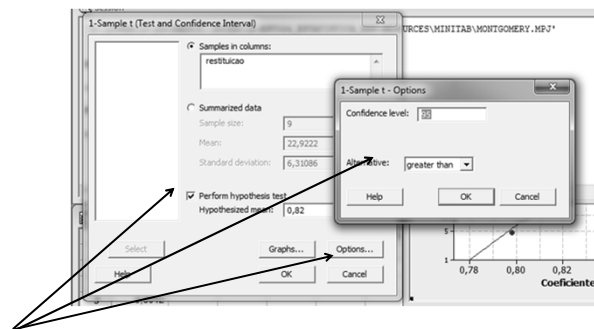
- Parâmetro de interesse: coeficiente de restituição (μ)
- $H_0: \mu = 0,82$
- $H_1: \mu > 0,82$
- $\alpha = 0,05$
- Estatística de teste: $T_0 = \frac{\bar{X} - \mu_0}{\frac{s}{\sqrt{n}}} \underset{H_0}{\sim} t_{n-1}$
- Rejeite H_0 se $t_0 > t_{0,05; 14} = 1,761$
- Estatística de teste observada: $t_0 = \frac{0,83725 - 0,82}{\frac{0,02456}{\sqrt{15}}} = 2,72$
- Conclusão:

- $t_0 = 2,72 > 1,761 \rightarrow$ Rejeitamos H_0
- Afirmamos, em um nível de 5% de significância, que o coeficiente médio de restituição excede 0,82.

- Minitab:

✓ Dados: BD_producao.xls/ guia: restituicao

Stat > Basic Statistics > 1-Sample t



- Saída:

One-Sample T: restituicao					
Test of $\mu = 0,82$ vs $\mu > 0,82$					
Variable	N	Mean	StDev	SE Mean	95% Lower Bound
restituicao	15	0,83724	0,02456	0,00634	0,82607
					T
					2,72
					P
					0,008

✓ Limite inferior com 95% de confiança excede 0,82

- Rejeita-se $H_0: \mu = 0,82$, concluindo-se que $H_1: \mu > 0,82$ é verdadeira
- Valor-p confirma conclusão

Valor-p

- Menor nível de significância no qual H_0 seria rejeitada
- Para testes com a estatística T_0 :
 √ Cálculo do valor-p:

p	Teste	Hipóteses	
$2P\{T_{n-1} > t_0 \}$	Bilateral	$H_1: \mu = \mu_0$	$H_1: \mu \neq \mu_0$
$P\{T_{n-1} > t_0\}$	Bilateral superior	$H_1: \mu = \mu_0$	$H_1: \mu > \mu_0$
$P\{T_{n-1} < t_0\}$	Bilateral inferior	$H_1: \mu = \mu_0$	$H_1: \mu < \mu_0$

√ T_{n-1} : distribuição t com n-1 graus de liberdade

- Cálculo do p-valor – Minitab:

√ No exemplo: $H_1: \mu > 0,82$

√ Estatística de teste observada:

• $t_0 = 2,72$

```
MTB > CDF 2,72;
SUBC> t 14.

Cumulative Distribution Function

Student's t distribution with 14 DF

x      P( X <= x )
2,72    0,991703
```

√ $p = P\{T_{14} > 2,72\} = 1 - 0,9917 = 0,0083$

- $H_0: \mu = 0,82$ seria rejeitada com qualquer nível de significância $\alpha \geq 0,0083$

- Cálculo do p-valor – Tabela:

t distribution critical values										
df	Upper-tail probability p									
	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025
1	1.000	1.376	1.963	3.078	6.314	12.71	15.89	31.82	63.66	127.3
2	0.816	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925	14.09
3	0.765	0.978	1.250	1.638	2.353	3.182	3.482	4.541	5.841	7.453
4	0.741	0.941	1.190	1.533	2.132	2.776	2.999	3.747	4.604	5.598
5	0.727	0.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032	4.773
6	0.718	0.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707	4.317
7	0.711	0.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499	4.029
8	0.706	0.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833
9	0.703	0.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250	3.690
10	0.700	0.879	1.093	1.372	1.812	2.228	2.359	2.764	3.169	3.581
11	0.697	0.876	1.088	1.363	1.796	2.201	2.328	2.718	3.106	3.497
12	0.695	0.873	1.083	1.356	1.782	2.179	2.303	2.681	3.055	3.428
13	0.694	0.870	1.079	1.350	1.771	2.160	2.282	2.650	3.012	3.372
14	0.692	0.868	1.076	1.345	1.761	2.145	2.264	2.624	2.977	3.326
15	0.691	0.866	1.074	1.341	1.753	2.131	2.249	2.602	2.947	3.286
16	0.690	0.865	1.071	1.337	1.746	2.120	2.235	2.583	2.921	3.252

√ $t_0 = 2,72$ está entre 2,624 e 2,977

– Limites superior e inferior para o p-valor: 0,005 e 0,01

- Se $H_1: \mu \neq 0,82$

√ Limite inferior para p: $2(0,005) = 0,01$

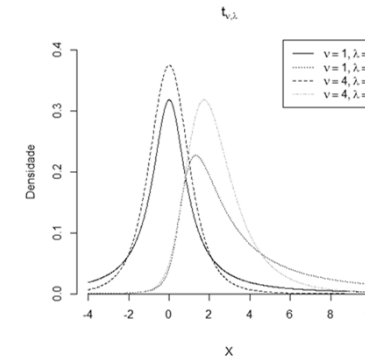
√ Limite superior para p: $2(0,01) = 0,02$

√ $0,01 < p < 0,02$

Erro Tipo II

- Quando média verdadeira $\mu = \mu_0 + \delta$, $\delta > 0$.
 - ✓ $H_0: \mu = \mu_0$ é falsa
 - ✓ $T_0 \sim$ distribuição t não-central, com $n - 1$ graus de liberdade e parâmetro de não-centralidade $\frac{\delta\sqrt{n}}{S} = d\sqrt{n}$
 - ✓ Se $\delta = 0$, a distribuição t não central é a distribuição t usual.

- Distribuição t não central:



- ✓ Dois parâmetros:

- v : graus de liberdade ($v > 0$)
- λ : não centralidade

Cálculo da Probabilidade de Erro Tipo II

- Supondo que a média verdadeira seja μ_1 .
 - ✓ $\mu_1 = \mu_0 + \delta$.
- Hipótese nula:
 - ✓ $H_0: \mu = \mu_0$.
- Hipóteses alternativas:
 - $H_1: \mu \neq \mu_0$.
 - $H_1: \mu > \mu_0$.
 - $H_1: \mu < \mu_0$.
- ✓ Parâmetro de não centralidade λ :

$$\lambda = \frac{(\mu_1 - \mu_0)\sqrt{n}}{S} = \frac{\delta\sqrt{n}}{S}$$

- Considere a hipótese bilateral:

- ✓ $H_0: \mu = \mu_0$
- ✓ $H_1: \mu \neq \mu_0$

- Suponha que o verdadeira valor da média seja

- ✓ $\mu = \mu_0 + \delta$, $\delta > 0$ (H_0 é falsa!)

- Estatística de teste: $T_0 = \frac{\bar{X} - \mu_0}{\frac{S}{\sqrt{n}}}$

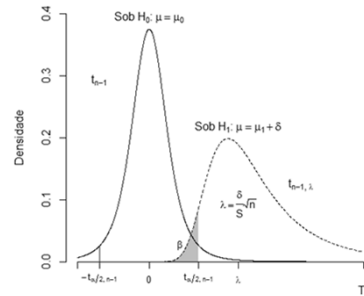
- ✓ Quando H_1 for verdadeira ($|\delta| > 0$): $T_0 \sim t'_{n-1,\lambda}$

- Distribuição t não central, com $n - 1$ graus de liberdade e parâmetro de não centralidade λ .

$$\lambda = \frac{\delta\sqrt{n}}{S} = d\sqrt{n}$$

- Comete-se erro tipo II quando H_1 é verdadeira e:

$$\sqrt{-t_{\alpha/2, n-1} \leq T_0 \leq t_{\alpha/2, n-1}}$$



- Probabilidade de erro tipo II para teste bilateral:

$$\begin{aligned}\beta &= \Pr[-t_{\alpha/2, n-1} \leq T_0 \leq t_{\alpha/2, n-1} | \delta \neq 0] \\ &= \Pr[-t'_{\alpha/2, n-1} \leq T'_0 \leq t'_{\alpha/2, n-1}]\end{aligned}$$

- Supondo que a média verdadeira seja:

$$\sqrt{\mu = \mu_0 + \delta, \delta > 0 \text{ (} H_0 \text{ é falsa!)}$$

- Para a hipótese unilateral superior:

$$\sqrt{H_0: \mu = \mu_0 \text{ vs. } H_1: \mu > \mu_0}$$

- Probabilidade de erro tipo II (β):

$$\begin{aligned}\beta &= \Pr[T_0 \leq t_{\alpha, n-1} | \delta \neq 0] \\ &= \Pr[T'_0 \leq t'_{\alpha, n-1}]\end{aligned}$$

- Distribuição t não central, com $n - 1$ graus de liberdade e parâmetro de não centralidade λ .

$$\lambda = \frac{\delta \sqrt{n}}{S} = d \sqrt{n}$$

- Supondo que a média verdadeira seja:

$$\sqrt{\mu = \mu_0 + \delta, \delta < 0 \text{ (} H_0 \text{ é falsa!)}$$

- Para a hipótese unilateral inferior:

$$\sqrt{H_0: \mu = \mu_0 \text{ vs. } H_1: \mu < \mu_0}$$

- Probabilidade de erro tipo II (β):

$$\begin{aligned}\beta &= \Pr[T_0 \geq -t_{\alpha, n-1} | \delta \neq 0] \\ &= \Pr[T'_0 \geq -t'_{\alpha, n-1}]\end{aligned}$$

- Distribuição t não central, com $n - 1$ graus de liberdade e parâmetro de não centralidade λ .

$$\lambda = \frac{\delta \sqrt{n}}{S} = d \sqrt{n}$$

Cálculo Aproximado de β – Bilateral

(Sem usar a t não central)

- Supondo que a média verdadeira seja:

$$\sqrt{\mu = \mu_0 + \delta, \delta > 0 \text{ (} H_0 \text{ é falsa!)}$$

- Hipótese bilateral:

$$\sqrt{H_0: \mu = \mu_0}$$

$$\sqrt{H_1: \mu \neq \mu_0}$$

- Nível de significância do teste: α .

• Procedimento:

1. Determinação dos valores críticos:

- Ponto crítico à esquerda: $C_E = \mu_0 - t_{\alpha/2, n-1} \frac{s}{\sqrt{n}}$
- Ponto crítico à direita: $C_D = \mu_0 + t_{\alpha/2, n-1} \frac{s}{\sqrt{n}}$

2. Cálculo aproximado de β :

$$\beta = \Pr [C_E \leq \bar{X} \leq C_D | \mu = \mu_0 + \delta = \mu_1]$$

$$= \Pr \left[\frac{C_E - \mu_1}{s/\sqrt{n}} \leq T_{n-1} \leq \frac{C_D - \mu_1}{s/\sqrt{n}} \right]$$

$$\beta = F_{t_{n-1}} \left(\frac{C_D - \mu_1}{s/\sqrt{n}} \right) - F_{t_{n-1}} \left(\frac{C_E - \mu_1}{s/\sqrt{n}} \right)$$

- Função de distribuição acumulada de uma distribuição t com $n - 1$ graus de liberdade

Cálculo Aproximado de β – Unilateral Superior

(Sem usar a t não central)

- Supondo que a média verdadeira seja:
 $\sqrt{\mu = \mu_0 + \delta, \delta > 0}$ (H_0 é falsa!)
- Hipótese unilateral superior:
 $\sqrt{H_0: \mu = \mu_0}$
 $\sqrt{H_1: \mu > \mu_0}$
- Nível de significância do teste: α .

• Procedimento:

1. Determinação dos valor crítico:

- Ponto crítico à direita: $C_D = \mu_0 + t_{\alpha, n-1} \frac{s}{\sqrt{n}}$

2. Cálculo aproximado de β :

$$\beta = \Pr [\bar{X} \leq C_D | \mu = \mu_0 + \delta = \mu_1]$$

$$= \Pr \left[T_{n-1} \leq \frac{C_D - \mu_1}{s/\sqrt{n}} \right]$$

$$\beta = F_{t_{n-1}} \left(\frac{C_D - \mu_1}{s/\sqrt{n}} \right)$$

- $F_{t_{n-1}}$: Função de distribuição acumulada de uma distribuição t com $n - 1$ graus de liberdade

Cálculo Aproximado de β – Unilateral Inferior

(Sem usar a t não central)

- Supondo que a média verdadeira seja:
 $\sqrt{\mu = \mu_0 + \delta, \delta > 0}$ (H_0 é falsa!)
- Hipótese unilateral inferior:
 $\sqrt{H_0: \mu = \mu_0}$
 $\sqrt{H_1: \mu < \mu_0}$
- Nível de significância do teste: α .

• Procedimento:

1. Determinação dos valor crítico:

- Ponto crítico à esquerda: $C_E = \mu_0 - t_{\alpha, n-1} \frac{s}{\sqrt{n}}$

2. Cálculo aproximado de β :

$$\beta = \Pr [\bar{X} \geq C_E | \mu = \mu_0 + \delta = \mu_1]$$

$$= \Pr \left[T_{n-1} \geq \frac{C_E - \mu_1}{s/\sqrt{n}} \right]$$

$$\beta = 1 - F_{t_{n-1}} \left(\frac{C_E - \mu_1}{s/\sqrt{n}} \right)$$

- $F_{t_{n-1}}$: Função de distribuição acumulada de uma distribuição t com n – 1 graus de liberdade

Curvas Características de Operação

- Curvas de β vs. parâmetro d, para vários tamanhos amostrais

✓ Gráficos VIIe e VIIh – Apêndice

✓ Curvas para $\alpha = 0,05$ e $\alpha = 0,01$

- Parâmetro d: $d = \frac{|\mu_1 - \mu_0|}{\sigma} = \frac{|\delta|}{\sigma}$

✓ d: Fator de não centralidade que depende do parâmetro desconhecido σ^2 .

✓ Conjunto de curvas características pode ser usado para todos os problemas, independente dos valores de μ_0 e σ .

• Determinação do fator de não centralidade d:

✓ Utilização de estimativa prévia de σ^2 através de experimento piloto

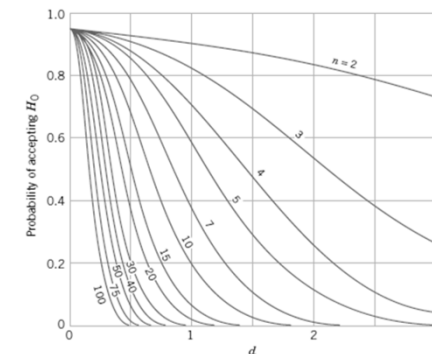
✓ Uso da variância amostral s^2 para estimar σ^2 .

- Para avaliar desempenho do teste após coleta dos dados

✓ Definição da diferença relativa que se quer detectar

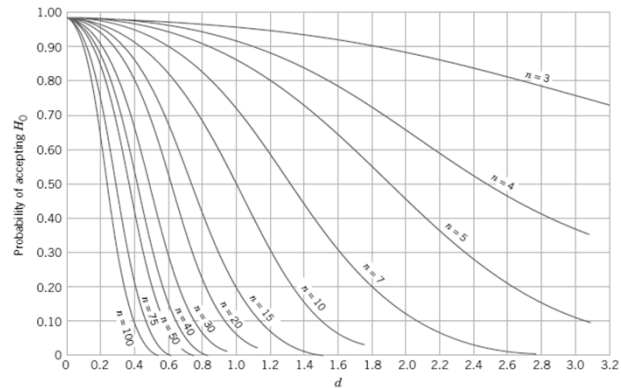
- Caso não haja experiência prévia que possa ser usada

• Teste bilateral e nível de significância $\alpha = 0,05$



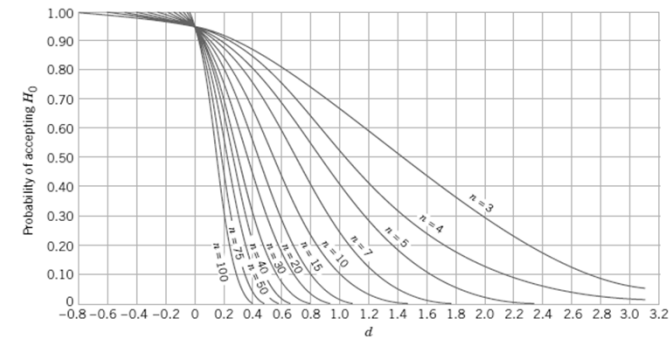
(e) O.C. curves for different values of n for the two-sided t-test for a level of significance $\alpha = 0.05$.

- Teste bilateral e nível de significância $\alpha = 0,01$



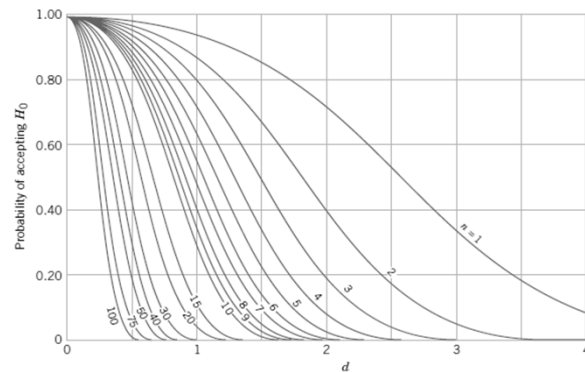
(f) O.C. curves for different values of n for the two-sided t -test for a level of significance $\alpha = 0.01$.

- Teste unilateral e nível de significância $\alpha = 0,05$



(g) O.C. curves for different values of n for the one-sided t -test for a level of significance $\alpha = 0.05$.

- Teste unilateral e nível de significância $\alpha = 0,01$



(b) O.C. curves for different values of n for the two-sided normal test for a level of significance $\alpha = 0.01$.

Exemplo 9-7

- Projeto do Taco de Golfe
 - ✓ Quantificação da razão entre velocidade de saída e velocidade de chegada da bola (coeficiente de restituição)
 - ✓ Experimento para determinar se há evidência (com $\alpha=5\%$) de suporte a afirmação de que o coeficiente médio de restituição exceda 0,82
 - ✓ Tamanho amostral: 15 tacos é adequado para assegurar que H_0 será rejeitada para uma diferença de 0,02, com probabilidade no mínimo 0,90?

- $H_0: \mu = 0,82$ vs. $H_1: \mu > 0,82$
- Desvio-padrão amostral; $s = 0,02456$
- Distância padronizada:

$$d = \frac{\delta}{S} = \frac{0,02}{0,02456} = 0,81$$

- Poder do teste para $\delta = 0,02$ (cálculo exato):

$$\begin{aligned} 1 - \beta &= 1 - \Pr [T_0 \leq t_{\alpha, n-1} | \delta \neq 0] \\ &= \Pr [T'_0 \leq 1,76131 | n - 1 = 14 \text{ e } \lambda = 0,81\sqrt{15}] \\ &= 1 - 0,0884 = 0,9116 \end{aligned}$$

- Erro tipo II – Cálculo aproximado:

1. Determinação do valor crítico à direita:

$$C_D = 0,82 + 1,761 \frac{0,02456}{\sqrt{15}} = 0,831$$

2. Cálculo aproximado de β :

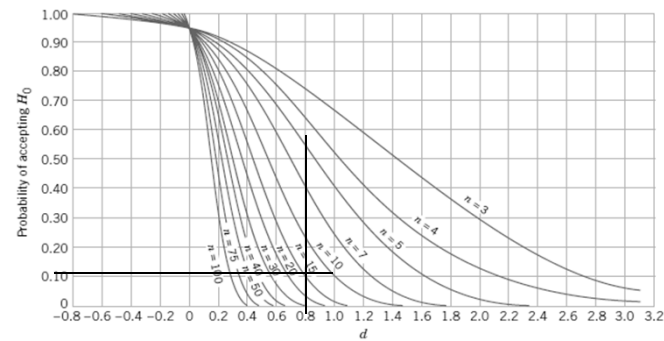
$$\beta = \Pr [\bar{X} \leq 0,831 | \mu = 0,82 + 0,02]$$

$$= \Pr \left[T_{n-1} \leq \frac{0,831 - 0,841}{(0,02456)/\sqrt{15}} \right]$$

$$\beta = F_{t_{n-1}}(-1,393) = 0,0927$$

$$Pd = 1 - \beta = 1 - 0,0927 = 0,9073$$

✓ Valor exato (t não central): $\beta=0,0884$ (Pd = 0,9116)

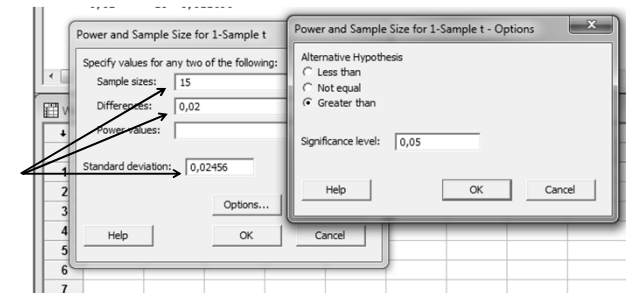


(g) O.C. curves for different values of n for the one-sided t -test for a level of significance $\alpha = 0.05$.

✓ Para $n = 15$, $\beta \approx 0,10$.

- Cálculo Poder - Minitab

Stat > Power and Sample Size > 1-Sample t



• Saída – Minitab

✓ Dados:

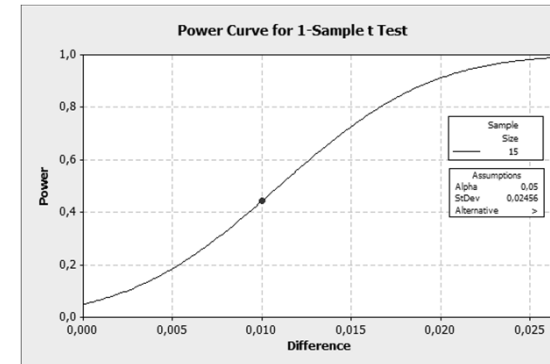
- Diferença: 0,02
- Tamanho amostra: 15

Power and Sample Size		
1-Sample t Test		
Testing mean = null (versus > null)		
Calculating power for mean = null + difference		
Alpha = 0,05 Assumed standard deviation = 0,02456		
Difference	Sample Size	Power
0,02	15	0,911696

✓ Poder = 0,9117 → $\beta = 0,0883$

- Os cálculos do computador são mais precisos que os valores da leitura visualizada a partir da curva CO.

- O tamanho da amostra $n = 15$ é adequado para fornecer a sensibilidade desejada



• Saída – Minitab

✓ Dados:

- Diferença: 0,01
- Tamanho amostra: 15

Power and Sample Size		
1-Sample t Test		
Testing mean = null (versus > null)		
Calculating power for mean = null + difference		
Alpha = 0,05 Assumed standard deviation = 0,02456		
Difference	Sample Size	Power
0,01	15	0,442346

✓ Poder = 0,4425 → $\beta = 0,5575$

- Tamanho amostral inadequado para fornecer a sensibilidade desejada.

Escolha Tamanho Amostral – Variância Desconhecida

- Deve-se determinar (ou conhecer):
 - ✓ Estrutura do experimento
 - ✓ Método de análise
 - ✓ Nível de significância (α) escolhido
 - ✓ Poder desejado
 - ✓ Variabilidade das medidas:
 - Realizar um estudo piloto, ou
 - Usar dados de experimentos anteriores ou publicados
 - ✓ Efeito significativo mínimo

- O cálculo de n (e de β) dependem do parâmetro desconhecido σ^2 .
- Maneiras para determinar o tamanho amostral no caso em que σ^2 é desconhecido:
 - ✓ Aproximação pela normal
 - ✓ Aproximação pela t não central
 - ✓ Curvas CO
 - ✓ Procedimentos computacionais
- Notação:
 - ✓ s_0 : estimativa prévia do desvio-padrão da população

Aproximação pela Normal

- ✓ Usada com amostras grandes
- ✓ z_η o percentil 100η superior da normal padrão
- ✓ Fator de não centralidade $d = \frac{\mu - \mu_0}{s_0} = \frac{|\delta|}{\sigma}$
- Teste bilateral
 - ✓ $H_0: \mu = \mu_0$ vs. $H_1: \mu \neq \mu_0$ $n \approx \left(\frac{z_\beta + z_{\alpha/2}}{d} \right)^2$
- Teste unilateral
 - ✓ $H_0: \mu = \mu_0$
 - ✓ $H_1: \mu > \mu_0$ ou $H_1: \mu < \mu_0$ $n \approx \left(\frac{z_\beta + z_\alpha}{d} \right)^2$

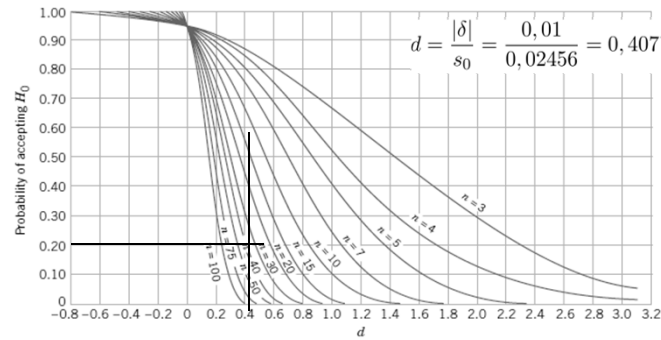
Aproximação pela t Não Central

- ✓ $t_{\eta, n_0 - 1}$ o percentil 100η superior da t com $n_0 - 1$ graus de liberdade
- ✓ n_0 : tamanho de amostra piloto
- Teste bilateral
 - ✓ $H_0: \mu = \mu_0$ vs. $H_1: \mu \neq \mu_0$ $n \approx \left(\frac{t_{\beta, n_0 - 1} + t_{\alpha/2, n_0 - 1}}{d} \right)^2$
- Teste unilateral
 - ✓ $H_0: \mu = \mu_0$
 - ✓ $H_1: \mu > \mu_0$ ou $H_1: \mu < \mu_0$ $n \approx \left(\frac{t_{\beta, n_0 - 1} + t_{\alpha, n_0 - 1}}{d} \right)^2$
 - ✓ n não pode ser “isolado” na expressão

Exemplo 9-7 (continuação)

- Projeto do Taco de Golfe
 - ✓ Quantificação da razão entre velocidade de saída e velocidade de chegada da bola (coeficiente de restituição)
 - ✓ Nível de significância (α): 5%
 - ✓ Hipóteses:
 - $H_0: \mu = 0,82$ vs. $H_1: \mu > 0,82$.
 - ✓ Tamanho amostral (n): 15 tacos
 - ✓ Qual tamanho amostral para detectar uma diferença de 0,01, com probabilidade no mínimo 0,80?

• Curvas CO:



(g) O.C. curves for different values of n for the one-sided t -test for a level of significance $\alpha = 0.05$.

✓ Tamanho amostra: $n \approx 40$

• Aproximação pela normal:

✓ $z_{0,05}$: 1,645

✓ $z_{0,20}$: 0,842

✓ Fator de não centralidade: $d = 0,407$

$$n \approx \left(\frac{z_\beta + z_\alpha}{d} \right)^2$$

$$\approx \left(\frac{0,842 + 1,645}{0,407} \right)^2 = 37,3$$

✓ Tamanho amostra: $n \approx 38$

• Aproximação pela t não-central:

✓ Tamanho de amostra piloto: $n_0 = 15$

✓ $t_{0,05,15-1}$: 1,761

✓ $t_{0,20,15-1}$: 0,868

✓ Fator de não centralidade: $d = 0,407$

$$n \approx \left(\frac{t_{\beta,n_0-1} + t_{\alpha,n_0-1}}{d} \right)^2$$

$$\approx \left(\frac{0,868 + 1,761}{0,407} \right)^2 = 41,72$$

✓ Tamanho amostra: $n \approx 42$

• Minitab:

Power and Sample Size			
1-Sample t Test			
Testing mean = null (versus > null)			
Calculating power for mean = null + difference			
Alpha = 0,05 Assumed standard deviation = 0,02456			
Difference	Sample Size	Target Power	Actual Power
0,01	39	0,8	0,802949

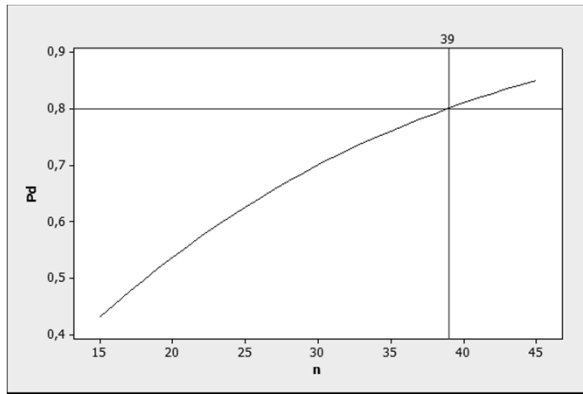
✓ Tamanho amostra: $n = 39$

• Se a aproximação pela t for usada com os percentis de uma t com 38 graus de liberdade

$$n \approx \left(\frac{t_{0,20,39-1} + t_{0,05,39-1}}{d} \right)^2$$

$$\approx \left(\frac{0,851 + 1,685}{0,407} \right)^2 = 38,82$$

✓ O Minitab tem procedimento iterativo para cálculo



✓ $n=39$ é o menor tamanho amostral tal que $\beta \geq 0,80$

Testes para Variância e Desvio-padrão – População Normal

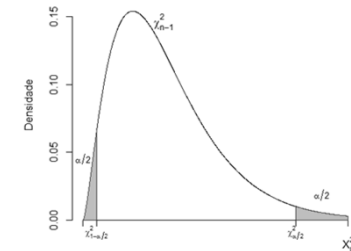
Teste de Hipóteses para a Variância

- Seja uma amostra aleatória $X_1, X_2, \dots, X_n$ proveniente de uma única população **normal**.
 - Hipóteses:
 - $H_0: \sigma^2 = \sigma_0^2$
 - $H_1: \sigma^2 \neq \sigma_0^2$
 - Nível de significância do teste: α .
 - Estatística de teste: $X_0^2 = \frac{(n-1)S^2}{\sigma_0^2} \stackrel{H_0}{\sim} \chi_{n-1}^2$
- ✓ S : desvio-padrão amostral

- Região de rejeição:

✓ $H_1: \sigma^2 \neq \sigma_0^2$. (alternativa bilateral)

✓ $X_0^2 > \chi_{\alpha/2, n-1}^2$ ou $X_0^2 < \chi_{1-\alpha/2, n-1}^2$.



✓ Pontos críticos: $\chi_{\alpha/2, n-1}^2$ e $\chi_{1-\alpha/2, n-1}^2$.

✓ Regra de decisão:

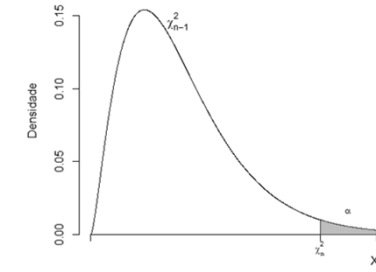
- Rejeitar H_0 se $X_0^2 > \chi_{\alpha/2, n-1}^2$ ou $X_0^2 < \chi_{1-\alpha/2, n-1}^2$.

Hipótese Alternativa Unilateral Superior

- Hipóteses:
 - ✓ $H_0: \sigma^2 = \sigma_0^2$.
 - ✓ $H_1: \sigma^2 > \sigma_0^2$.
- Comentários:
 - ✓ Valores grandes de X_0^2 conduzem à rejeição de H_0 .
 - ✓ Região crítica na extremidade superior
 - ✓ Rejeição de H_0 :
 - $X_0^2 > \chi_{\alpha, n-1}^2$.

- Região de rejeição:

- ✓ $H_1: \sigma^2 > \sigma_0^2$. (alternativa unilateral superior)
- ✓ $X_0^2 > \chi_{\alpha, n-1}^2$.



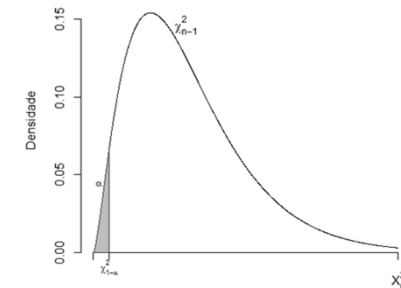
- ✓ Pontos crítico: $\chi_{\alpha, n-1}^2$.
- ✓ Regra de decisão:
 - Rejeitar H_0 se $X_0^2 > \chi_{\alpha, n-1}^2$.

Hipótese Alternativa Unilateral Inferior

- Hipóteses:
 - ✓ $H_0: \sigma^2 = \sigma_0^2$.
 - ✓ $H_1: \sigma^2 < \sigma_0^2$.
- Comentários:
 - ✓ Valores pequenos de X_0^2 conduzem à rejeição de H_0 .
 - ✓ Região crítica na extremidade inferior
 - ✓ Rejeição de H_0 :
 - $X_0^2 < \chi_{1-\alpha, n-1}^2$.

- Região de rejeição:

- ✓ $H_1: \sigma^2 < \sigma_0^2$. (alternativa unilateral inferior)
- ✓ $X_0^2 < \chi_{1-\alpha, n-1}^2$.



- ✓ Pontos crítico: $\chi_{1-\alpha, n-1}^2$.
- ✓ Regra de decisão:
 - Rejeitar H_0 se $X_0^2 < \chi_{1-\alpha, n-1}^2$.

Testes para Média (Variância Desconhecida)

- Hipótese nula:
 $\sqrt{H_0: \sigma^2 = \sigma_0^2}$
- Estatística de teste: $X_0^2 = \frac{(n-1)S^2}{\sigma_0^2}$

Hipótese alternativa	Critério de rejeição
$H_1: \sigma^2 \neq \sigma_0^2$	$X_0^2 > \chi_{\alpha/2, n-1}^2$ ou $X_0^2 < \chi_{1-\alpha/2, n-1}^2$
$H_1: \sigma^2 > \sigma_0^2$	$X_0^2 > \chi_{\alpha, n-1}^2$
$H_1: \sigma^2 < \sigma_0^2$	$X_0^2 < \chi_{1-\alpha, n-1}^2$

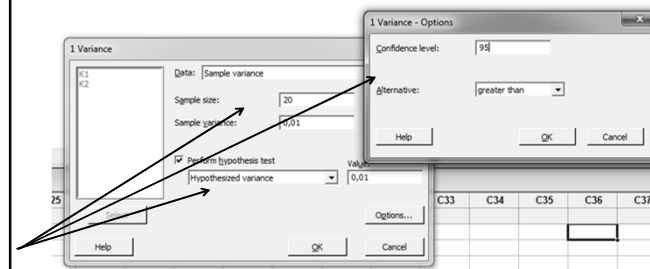
- Procedimento de teste:
 - Parâmetro de interesse: coeficiente de restituição (μ)
 - $H_0: \sigma^2 = 0,01$
 - $H_1: \sigma^2 > 0,01$
 - $\alpha = 0,05$
 - Estatística de teste: $X_0^2 = \frac{(n-1)s^2}{\sigma_0^2} \sim \chi_{n-1}^2$
 - Rejeite H_0 se $X_0^2 > \chi_{0,05; 19}^2 = 30,144$
 - Estatística de teste observada: $X_0^2 = \frac{19(0,0153)}{0,01} = 29,07$
 - Conclusão:
 - $X_0^2 = 29,07 < 30,144 \rightarrow$ Falhamos em rejeitar H_0
 - Concluimos, em um nível de 5% de significância, que não há evidência forte de que a variância no volume de enchimento excede 0,01 (onça fluida)².

Exemplo 9-8

- Enchimento automático
 - ✓ Volume de garrafas de detergente
 - ✓ Amostra: 20 garrafas
 - ✓ Variância amostral: $s^2 = 0,0153$ (onça fluida)².
 - ✓ Proporção inaceitável de garrafas não conformes
 - $\sigma^2 > 0,01$ (onça fluida)².
 - ✓ Há evidências amostrais que sugira problemas na produção?
 - Nível de significância do teste(α): 5%
 - Suponha que o volume de enchimento populacional tenha distribuição normal

- Minitab:

Stat > Basic Statistics > 1-Variance



- Saída:

Test and CI for One Variance

Method

Null hypothesis Sigma-squared = 0,01
 Alternative hypothesis Sigma-squared > 0,01

The chi-square method is only for the normal distribution.
 The Bonett method cannot be calculated with summarized data.

Statistics

N	StDev	Variance
20	0,124	0,0153

95% One-Sided Confidence Intervals

Method	Lower Bound for StDev	Lower Bound for Variance
Chi-Square	0,098	0,0096

Tests

Method	Statistic	DF	P-Value
Chi-Square	29,07	19	0,065

Limite inferior com
95% de confiança

Estadística
 X_0^2

✓ Limite inferior não excede 0,01 (95% de confiança)

- Não há evidência amostral para rejeitar $H_0: \sigma^2 = 0,01$
- Valor-p confirma conclusão

- Cálculo do p-valor – Minitab:

✓ No exemplo: $H_1: \sigma^2 > 0,01$

✓ Estatística de teste observada:

- $X_0^2 = 29,07$

MTB > CDF 29,07;	
SUBC> Chisq 19.	
Cumulative Distribution Function	
Chi-Square with 19 DF	
x	P (X <= x)
29,07	0,935108

✓ $p = P\{\chi^2_{19} > 29,07\} = 1 - 0,9351 = 0,0649$

- $H_0: \sigma^2 = 0,01$ seria rejeitada com qualquer nível de significância $\alpha \geq 0,0649$

Valor-p

- Menor nível de significância no qual H_0 seria rejeitada
- Para testes com a estatística X_0^2 :
✓ Cálculo do valor-p:

p	Teste	Hipóteses	
$2P\{\chi^2_{n-1} > (\text{ou } <) X_0^2\}^{(1)}$	Bilateral	$H_1: \sigma^2 = \sigma_0^2$	$H_1: \sigma^2 \neq \sigma_0^2$
$P\{\chi^2_{n-1} > X_0^2\}$	Bilateral superior	$H_1: \sigma^2 = \sigma_0^2$	$H_1: \sigma^2 > \sigma_0^2$
$P\{\chi^2_{n-1} < X_0^2\}$	Bilateral inferior	$H_1: \sigma^2 = \sigma_0^2$	$H_1: \sigma^2 < \sigma_0^2$

✓ χ^2_{n-1} : distribuição χ^2 com n-1 graus de liberdade

✓ $^{(1)}\text{Mínimo}(P\{\chi^2_{n-1} > X_0^2\}; P\{\chi^2_{n-1} < X_0^2\})$

- Cálculo do p-valor – Tabela:

df	$\chi^2_{.995}$	$\chi^2_{.990}$	$\chi^2_{.975}$	$\chi^2_{.950}$	$\chi^2_{.900}$	$\chi^2_{.800}$	$\chi^2_{.700}$	$\chi^2_{.600}$	$\chi^2_{.500}$	$\chi^2_{.400}$	$\chi^2_{.300}$	$\chi^2_{.200}$	$\chi^2_{.100}$	$\chi^2_{.050}$	$\chi^2_{.025}$	$\chi^2_{.010}$	$\chi^2_{.005}$
1	0.000	0.000	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879							
2	0.010	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.597							
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345	12.838							
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860							
5	0.412	0.554	0.831	1.145	1.610	9.236	11.070	12.833	15.086	16.750							
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812	18.548							
7	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475	20.278							
8	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090	21.955							
9	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666	23.589							
10	2.156	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188							
11	2.603	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.725	26.757							
12	3.074	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217	28.300							
13	3.565	4.107	5.009	5.892	7.042	19.812	22.362	24.736	27.688	29.819							
14	4.075	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141	31.319							
15	4.601	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.578	32.801							
16	5.142	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000	34.267							
17	5.697	6.408	7.564	8.672	10.085	24.769	27.587	30.191	33.409	35.718							
18	6.265	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805	37.156							
19	6.844	7.633	8.907	10.117	11.651	27.204	30.144	32.852	36.191	38.582							
20	7.434	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566	39.997							

✓ $X_0^2 = 29,07$ está entre 27,20 e 30,14

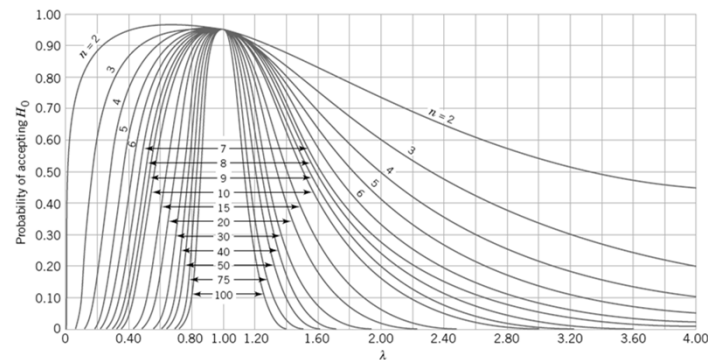
– Limites superior e inferior para o p-valor: 0,05 e 0,10

- Se $H_1: \sigma^2 \neq 0,01$
 - ✓ Limite inferior para p: $2(0,05) = 0,10$
 - ✓ Limite superior para p: $2(0,10) = 0,20$
 - ✓ $0,10 < p < 0,20$

Curvas Características de Operação

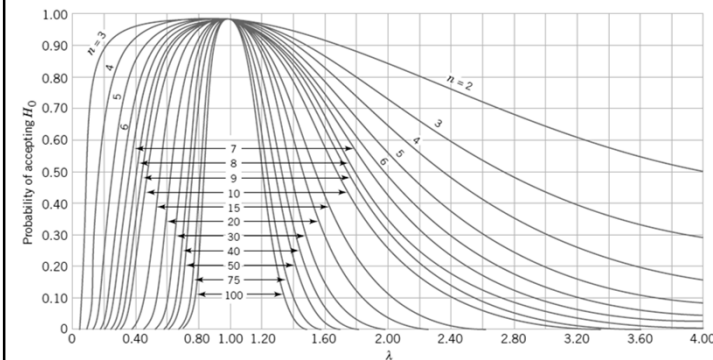
- Curvas de β vs. parâmetro λ , para vários tamanhos amostrais
 - ✓ Gráficos VIIi e VIIn – Apêndice
 - ✓ Curvas para $\alpha = 0,05$ e $\alpha = 0,01$
- Parâmetro λ : $\lambda = \frac{\sigma}{\sigma_0}$
 - ✓ λ : Fator de não centralidade.
 - ✓ Conjunto de curvas características pode ser usado para todos os problemas, independente dos valores de σ .

- Teste bilateral e nível de significância $\alpha = 0,05$



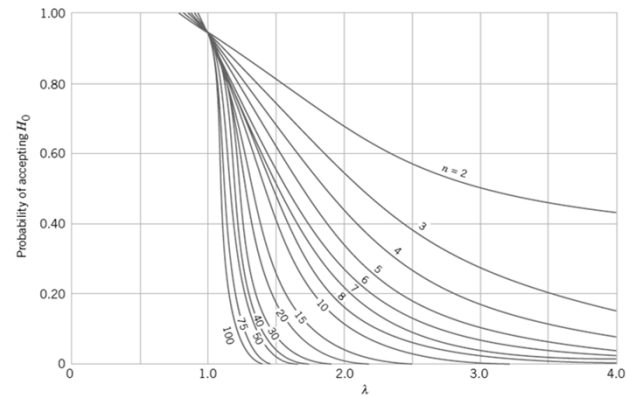
(i) O.C. curves for different values of n for the two-sided chi-square test for a level of significance $\alpha = 0.05$.

- Teste bilateral e nível de significância $\alpha = 0,01$



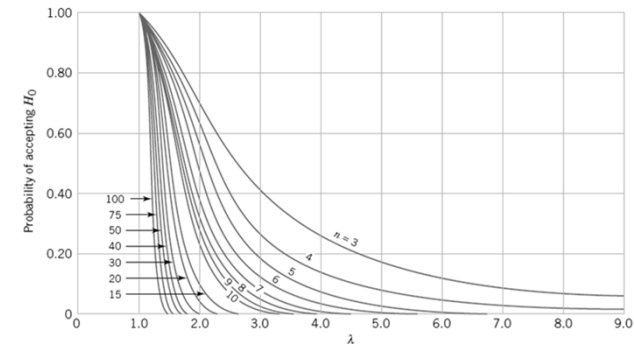
(j) O.C. curves for different values of n for the two-sided chi-square test for a level of significance $\alpha = 0.01$.

• Teste unilateral superior ($\alpha = 0,05$)



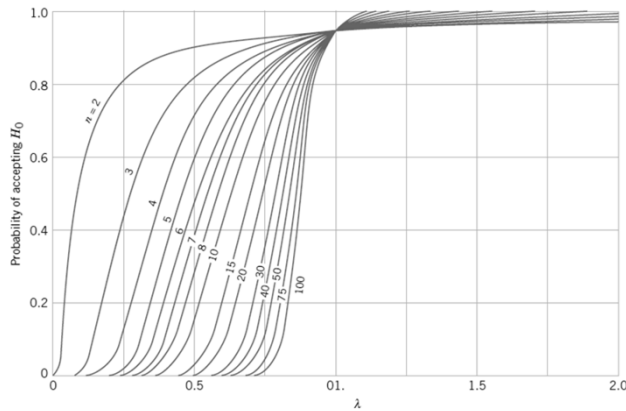
(k) O.C. curves for different values of n for the one-sided (upper tail) chi-square test for a level of significance $\alpha = 0.05$.

• Teste unilateral superior ($\alpha = 0,01$)



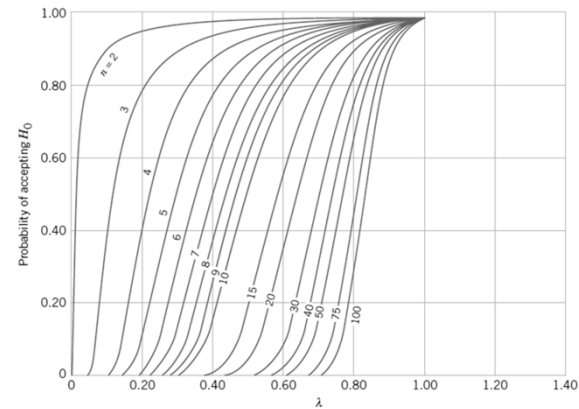
(l) O.C. curves for different values of n for the one-sided (upper tail) chi-square test for a level of significance $\alpha = 0.01$.

• Teste unilateral inferior ($\alpha = 0,05$)



(m) O.C. curves for different values of n for the one-sided (lower tail) chi-square test for a level of significance $\alpha = 0.05$.

• Teste unilateral inferior ($\alpha = 0,01$)

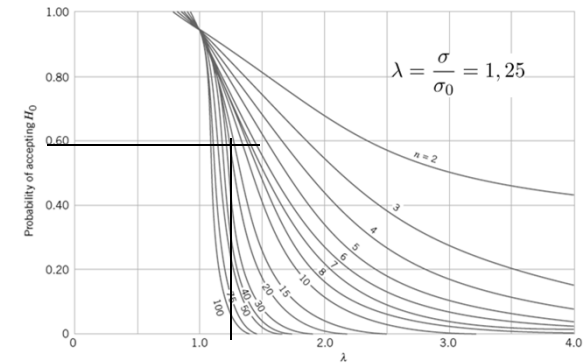


(n) O.C. curves for different values of n for the one-sided (lower tail) chi-square test for a level of significance $\alpha = 0.01$.

Exemplo 9-9

- Enchimento automático – Tamanho de amostra
 - ✓ Volume de garrafas de detergente
 - ✓ Amostra: 20 garrafas
 - ✓ Variância amostral: $s^2 = 0,0153$ (onça fluida)².
 - ✓ Proporção inaceitável de garrafas não conformes
 - $\sigma^2 > 0,01$ (onça fluida)².
 - ✓ Deseja-se detectar excesso de 25% do desvio-padrão verdadeiro, com probabilidade de no mínimo 80%?
 - O tamanho amostral de 20 garrafas é adequado?
 - Qual o menor tamanho amostral adequado?

• Curvas CO:



(k) O.C. curves for different values of n for the one-sided (upper tail) chi-square test for a level of significance $\alpha = 0.05$.

✓ Probabilidade de erro tipo II: $\beta \approx 0,60$

• Minitab:

```
Cumulative Distribution Function
Chi-Square with 19 DF
    x  P( X <= x )
29,07  0,935108

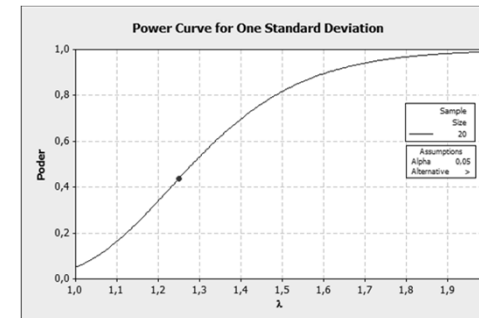
Power and Sample Size

Test for One Standard Deviation
Testing StDev = null (versus > null)
Calculating power for (StDev / null) = ratio
Alpha = 0,05
```

Ratio	Sample Size	Power
1,25	20	0,438262

✓ Poder = 0,4383 → $\beta = 0,5617$

- Há somente cerca de 40% de chance de rejeitar H_0 se o desvio-padrão verdadeiro for tão alto quanto $\sigma = 0,125$ onça fluida.



✓ Amostra de tamanho amostral detecta excesso de aproximadamente 50% com probabilidade de 0,80.

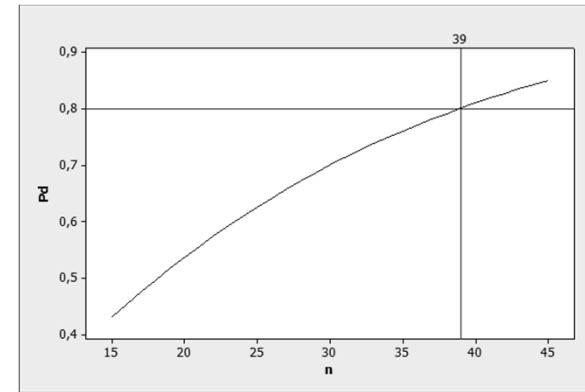
- Minitab:

Power and Sample Size				
Test for One Standard Deviation				
Testing StDev = null (versus > null)				
Calculating power for (StDev / null) = ratio				
Alpha = 0,05				
Ratio	Sample Size	Target Power	Actual Power	
1,25	61	0,8	0,800805	

✓ $n = 61$ é o menor tamanho amostral tal que $\beta \geq 0,80$ para detectar um excesso de pelo menos 25% no desvio-padrão

✓ As curvas CO indicam tamanho amostral $n=75$

- (são menos precisas)



✓ $n = 39$ é o menor tamanho amostral tal que $\beta \geq 0,80$

Testes para a Proporção Populacional

Teste de Hipóteses para a Proporção de População

- Seja uma amostra aleatória de tamanho n , retirada de uma grande população (possivelmente infinita).
 - ✓ p : proporção da população que pertence a categoria de interesse
 - ✓ X : número de observações amostrais que pertencem a essa categoria de interesse.
 - ✓ $X \sim \text{binomial}(n, p)$

- Hipóteses:

$$H_0: p = p_0$$

$$H_1: p \neq p_0$$

- Nível de significância do teste: α .

- Se H_0 for verdadeira:

$$\sqrt{X \sim \text{binomial}(n, p_0)}$$

$\sqrt{X}$: quantidade de itens da amostra que pertencem a categoria de interesse.

- Estatística para testar H_0 :

$$Z_0 = \frac{X - np_0}{\sqrt{np_0(1 - p_0)}} \stackrel{as.}{\sim} N(0, 1)$$

$\sqrt{\text{Aproximação será boa para amostras grandes e se } p \text{ não estiver muito próximo de } 0 \text{ ou } 1.}$

Testes Aproximados para Proporção Binomial

- Hipótese nula:

$$\sqrt{H_0: p = p_0.}$$

- Estatística de teste: $Z_0 = \frac{X - np_0}{\sqrt{np_0(1 - p_0)}}$

Hipótese alternativa	Critério de rejeição
$H_1: p \neq p_0$	$z_0 > z_{\alpha/2}$ ou $z_0 < -z_{\alpha/2}$.
$H_1: p > p_0$	$z_0 > z_{\alpha}$.
$H_1: p < p_0$	$z_0 < -z_{\alpha}$.

Exemplo 9-10

- Controlador de Motor de Automóveis

$\sqrt{\text{Consumidor exige que fração defeituosos em etapa crítica de produção não exceda } 0,05}$

$\sqrt{\text{Fabricante deve demonstrar capacidade de processo com } \alpha = 5\%.}$

$\sqrt{\text{Amostra com } 200 \text{ aparelhos}}$

– 4 são defeituosos.

$\sqrt{\text{Fabricante pode demonstrar que processo é capaz?}}$

- Procedimento de teste:

1. Parâmetro de interesse: fração de defeituosos (p)

2. $H_0: \mu = 0,05$

3. $H_1: \mu < 0,05$

4. $\alpha = 0,05$

5. Estatística de teste: $Z_0 = \frac{X - np_0}{\sqrt{np_0(1 - p_0)}} \stackrel{H_0}{\sim} N(0, 1)$

6. Rejeite H_0 se $z_0 < -z_{0,05} = -1,645$

7. Estatística observada: $z_0 = \frac{4 - 200(0,05)}{\sqrt{200(0,05)(1 - 0,05)}} = -1,95$

8. Conclusão:

• $z_0 = -1,95 > -1,645 \rightarrow \text{Rejeitamos } H_0$

• Concluímos, em um nível de 5% de significância, que a fração de defeituosos do processo é menor do que 0,05.

Forma Alternativa da Estatística de Teste

✓ Estimador pontual de p: $\hat{p} = \frac{X}{n}$, com:

$$\mu_{\hat{p}} = E[\hat{p}] = p \quad \text{e} \quad \sigma_{\hat{p}}^2 = \text{Var}[\hat{p}] = \frac{p(1-p)}{n}$$

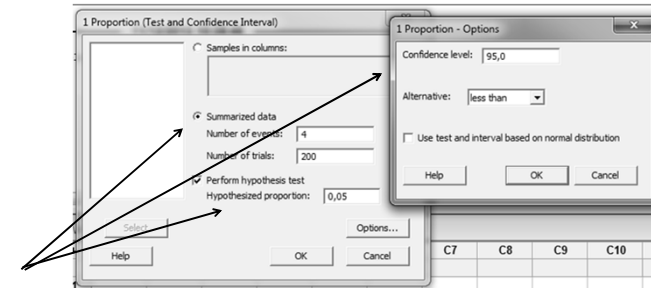
✓ Distribuição amostral aproximada da proporção amostral

$$\hat{p} \stackrel{as.}{\sim} N[\mu_{\hat{p}}, \sigma_{\hat{p}}^2]$$

- Se p não estiver muito próximo de 0 ou 1 e se n for relativamente grande
- Em geral, necessita-se que np e $np(1-p) \geq 5$
- Pacotes estatísticos trabalham em geral com proporção amostral

- Minitab:

Stat > Basic Statistics > 1-Proportion



- Saída – Distribuição exata:

Limite superior com 95% de confiança

Test of p = 0,05 vs p < 0,05				
Sample	X	N	Sample p	95% Upper Bound
1	4	200	0,020000	0,045180

Exact P-Value: 0,026

- ✓ Minitab utilizou a distribuição exata
- ✓ Limite superior com 95% de confiança inferior a 0,05
 - Rejeita-se $H_0: p = 0,05$, concluindo-se que $H_1: p < 0,05$ é verdadeira
 - Valor-p confirma conclusão

- Saída – Distribuição aproximada

Limite superior com 95% de confiança

Test of p = 0,05 vs p < 0,05				
Sample	X	N	Sample p	95% Upper Bound
1	4	200	0,020000	0,036283

Z-Value: -1,95
P-Value: 0,026

Using the normal approximation.
The normal approximation may be inaccurate for small samples.

- ✓ Minitab utilizou a distribuição exata na construção IC
- ✓ Limite superior com 95% de confiança é menor que aquele obtido pela distribuição exata
 - Rejeita-se $H_0: p = 0,05$, concluindo-se que $H_1: p < 0,05$ é verdadeira

Erro Tipo II

- Equações para aproximação da probabilidade de erro tipo II (β):

✓ Probabilidade de erro tipo II (β) depende da escolha do tamanho da amostra

✓ Probabilidade de erro tipo II para teste bilateral para a proporção binomial:

- $H_1: p \neq p_0$.

$$\begin{aligned}\beta &= \Pr(-z_{\alpha/2} \leq Z_0 \leq z_{\alpha/2} | p \neq p_0) \\ &= \Pr\left(p_0 - z_{\alpha/2} \sqrt{\frac{p_0(1-p_0)}{n}} \leq \hat{p} \leq p_0 + z_{\alpha/2} \sqrt{\frac{p_0(1-p_0)}{n}} \mid p \neq p_0\right) \\ &= \Pr\left(\frac{p_0 - p - z_{\alpha/2} \sqrt{\frac{p_0(1-p_0)}{n}}}{\sqrt{\frac{p(1-p)}{n}}} \leq \frac{\hat{p} - p_0}{\sqrt{\frac{p(1-p)}{n}}} \leq \frac{p_0 - p + z_{\alpha/2} \sqrt{\frac{p_0(1-p_0)}{n}}}{\sqrt{\frac{p(1-p)}{n}}}\right) \\ \beta &= \Phi\left(\frac{p_0 - p + z_{\alpha/2} \sqrt{\frac{p_0(1-p_0)}{n}}}{\sqrt{\frac{p(1-p)}{n}}}\right) - \Phi\left(\frac{p_0 - p - z_{\alpha/2} \sqrt{\frac{p_0(1-p_0)}{n}}}{\sqrt{\frac{p(1-p)}{n}}}\right)\end{aligned}$$

✓ Probabilidade de erro tipo II para teste unilateral superior para a proporção binomial:

- $H_1: p > p_0$.

$$\begin{aligned}\beta &= \Pr(Z_0 \leq z_{\alpha} | p \neq p_0) \\ &= \Pr\left(\hat{p} \leq p_0 + z_{\alpha} \sqrt{\frac{p_0(1-p_0)}{n}} \mid p \neq p_0\right) \\ \beta &= \Phi\left(\frac{p_0 - p + z_{\alpha} \sqrt{\frac{p_0(1-p_0)}{n}}}{\sqrt{\frac{p(1-p)}{n}}}\right)\end{aligned}$$

✓ Probabilidade de erro tipo II para teste unilateral inferior para a proporção binomial:

- $H_1: p > p_0$.

$$\begin{aligned}\beta &= \Pr(Z_0 > -z_{\alpha} | p \neq p_0) \\ &= \Pr\left(\hat{p} > p_0 - z_{\alpha} \sqrt{\frac{p_0(1-p_0)}{n}} \mid p \neq p_0\right) \\ \beta &= 1 - \Phi\left(\frac{p_0 - p - z_{\alpha} \sqrt{\frac{p_0(1-p_0)}{n}}}{\sqrt{\frac{p(1-p)}{n}}}\right)\end{aligned}$$

Escolha do Tamanho da Amostra

- Tamanho aproximado da amostra:

✓ Hipótese alternativa bilateral ($H_1: p \neq p_0$)

- Probabilidade de erro tipo II:

$$\beta = \Phi\left(\frac{p_0 - p + z_{\alpha/2}\sqrt{\frac{p_0(1-p_0)}{n}}}{\sqrt{\frac{p(1-p)}{n}}}\right) - \Phi\left(\frac{p_0 - p - z_{\alpha/2}\sqrt{\frac{p_0(1-p_0)}{n}}}{\sqrt{\frac{p(1-p)}{n}}}\right)$$

✓ Se $p > p_0$, então:

$$\beta \approx \Phi\left(\frac{p_0 - p + z_{\alpha/2}\sqrt{\frac{p_0(1-p_0)}{n}}}{\sqrt{\frac{p(1-p)}{n}}}\right)$$

✓ Seja z_β o percentil 100β superior da normal padrão

$$\beta = \Phi(-z_\beta)$$

$$-z_\beta \approx \frac{p_0 - p + z_{\alpha/2}\sqrt{\frac{p_0(1-p_0)}{n}}}{\sqrt{\frac{p(1-p)}{n}}}$$

- Tamanho aproximado da amostra para um teste bilateral para a proporção binomial

$$n \approx \left[\frac{z_{\alpha/2}\sqrt{p_0(1-p_0)} + z_\beta\sqrt{p(1-p)}}{p - p_0} \right]^2$$

- Tamanho aproximado da amostra para um teste unilateral para a proporção binomial

$$n \approx \left[\frac{z_\alpha\sqrt{p_0(1-p_0)} + z_\beta\sqrt{p(1-p)}}{p - p_0} \right]^2$$

Exemplo 9-11

(Continuação Ex. 9-10)

- Controlador de Motor de Automóveis

✓ Consumidor exige que fração defeituosos em etapa crítica de produção não exceda 0,05

✓ Nível de significância do teste (α): 5%

✓ Amostra com 200 aparelhos

✓ Fração defeituosa do processo é realmente $p = 0,03$.

✓ Qual a probabilidade de erro tipo II (β)?

- Probabilidade de erro tipo II (β):

$$\begin{aligned} \beta &= 1 - \Phi\left(\frac{p_0 - p - z_{\alpha}\sqrt{\frac{p_0(1-p_0)}{n}}}{\sqrt{\frac{p(1-p)}{n}}}\right) \\ &= 1 - \Phi\left(\frac{0,05 - 0,03 - (1,645)\sqrt{\frac{0,05(1-0,05)}{200}}}{\sqrt{\frac{0,03(1-0,03)}{200}}}\right) \\ &= 1 - \Phi(-0,44) = 0,67 \end{aligned}$$

✓ Poder do teste: $Pd = 1 - \beta = 0,33$

✓ A diferença entre p_0 e p é relativamente pequena

✓ A amostra não é particularmente grande

- Fabricante aceita a aceitar erro b tão grande quanto 0,10, se o valor verdadeiro for $p = 0,03$
- Qual o tamanho da amostra?

$$n \approx \left[\frac{z_{\alpha} \sqrt{p_0(1-p_0)} + z_{\beta} \sqrt{p(1-p)}}{p - p_0} \right]^2$$

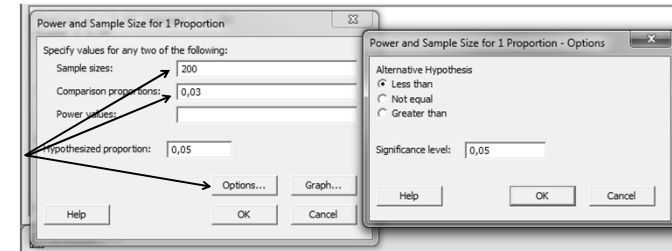
$$\approx \left[\frac{(1,645) \sqrt{0,05(0,95)} + 1,28 \sqrt{0,03(0,97)}}{0,03 - 0,05} \right]^2$$

$$\approx 832$$

✓ Tenta-se detectar desvio razoavelmente pequeno → a amostra tem tamanho grande

- Cálculo Poder - Minitab

Stat > Power and Sample Size > 1 Proportion



- Saída – Minitab

✓ Dados:

- Proporção verdadeira: 0,03
- Tamanho amostra: 200

✓ Poder = 0,3287 → $\beta = 0,6713$

- Minitab realiza cálculo exato de β .

Power and Sample Size			
Test for One Proportion			
Testing p = 0,05 (versus < 0,05)			
Alpha = 0,05			
Comparison p	Sample Size	Power	
0,03	200	0,328725	

- Saída – Minitab

✓ Dados:

- Proporção verdadeira: 0,03
- Poder: 0,90

✓ Tamanho amostra n = 833

✓ Dados:

- Proporção verdadeira: 0,03
- Poder: 0,75

✓ Tamanho amostra n = 561

- Tamanho amostrais grandes pois diferença entre proporções é pequena.

Power and Sample Size			
Test for One Proportion			
Testing p = 0,05 (versus < 0,05)			
Alpha = 0,05			
Comparison p	Sample Size	Target Power	Actual Power
0,03	833	0,9	0,900135

Power and Sample Size			
Test for One Proportion			
Testing p = 0,05 (versus < 0,05)			
Alpha = 0,05			
Comparison p	Sample Size	Target Power	Actual Power
0,03	561	0,75	0,750302

Testes de Adequação de um Ajuste

Teste χ^2 de Adequação de Ajuste

- Teste de Adequação de Ajuste baseado na distribuição χ^2 :
 - ✓ Amostra aleatória de tamanho n oriunda de população com distribuição desconhecida
 - ✓ Construção de histograma de frequências das n observações, com k classes

- Notação:
 - ✓ O_i : frequência observada na i -ésima classe
 - ✓ E_i : frequência esperada na i -ésima classe (sob H_0)
- Estatística de Teste para Adequação de Ajuste

$$X_0^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i}$$

- Distribuição amostral:
$$X_0^2 \stackrel{H_0}{\sim} \chi_{k-p-1}^2$$
 - ✓ k : quantidade de classes
 - ✓ p : número de parâmetros desconhecidos da distribuição (sob H_0)

- Regra de decisão:
 - ✓ Rejeitar H_0 se $X_0^2 > \chi_{\alpha, k-p-1}^2$.
- Comentários:
 - ✓ Se as frequências esperadas forem muito pequenas, X_0^2 não refletirá o desvio entre o observado e o esperado
 - ✓ Os valores de 3, 4 e 5 são usados como mínimo

Exemplo 9-12

- Defeitos em placa de circuito impresso
 - ✓ Supõe-se que os defeitos seguem uma distribuição de Poisson
 - Parâmetro desconhecido tem de ser estimado a partir dos dados amostrais
 - ✓ Amostra aleatória com 60 placas impressas

✓ Dados amostrais e cálculo das frequências esperadas

Número de Defeitos	Frequência Observada	Probabilidade	Frequência Esperada
0	32	0,472	28,32
1	15	0,354	21,24
2	9	0,133	7,98
3 (ou mais)	4	0,041	2,46

$$E_i = np_i = 60 p_i$$

– Quantidade média de defeitos por placa:

$$\hat{\lambda} = \frac{(32)(0) + (15)(1) + (9)(2) + (4)(3)}{60} = 0,75$$

– Cálculo probabilidades (sob H_0):

$$p_1 = P\{X = 0\} = \frac{e^{-0,75}(0,75)^0}{0!} = 0,472 \quad p_3 = P\{X = 2\} = \frac{e^{-0,75}(0,75)^2}{2!} = 0,133$$

$$p_2 = P\{X = 1\} = \frac{e^{-0,75}(0,75)^1}{1!} = 0,354 \quad p_4 = P\{X \geq 3\} = 1 - (p_1 + p_2 + p_3) = 0,041$$

✓ Tabela com combinação das duas última células:

Número de Defeitos	Frequência Observada	Probabilidade	Frequência Esperada
0	32	0,472	28,32
1	15	0,354	21,24
2 (ou mais)	13	0,174	10,44

• Procedimento de teste:

- Variável de interesse: número de defeitos em placas
- H_0 : forma da distribuição de defeitos é Poisson
- H_1 : forma da distribuição de defeitos não é Poisson
- $\alpha = 0,05$

$$5. \text{ Estatística de teste: } X_0^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i}$$

$$6. \text{ Rejeite } H_0 \text{ se } X_0^2 > \chi_{0,05;1}^2 = 3,84$$

7. Estatística observada:

$$X_0^2 = \frac{(32 - 28,32)^2}{28,32} + \frac{(15 - 21,24)^2}{21,24} + \frac{(13 - 10,44)^2}{10,44} = 2,94$$

8. Conclusão:

- $X_0^2 = 2,94 < 3,84 \rightarrow$ Falhamos em rejeitar H_0 de que a distribuição de defeitos nas placas de circuito impresso é Poisson.

• Cálculo do p-valor – Minitab:

✓ Estatística de teste observada:

• $X_0^2 = 2,94$

```
MTB > CDF 2,94;
SUBC> Chisq 1.

Cumulative Distribution Function

Chi-Square with 1 DF

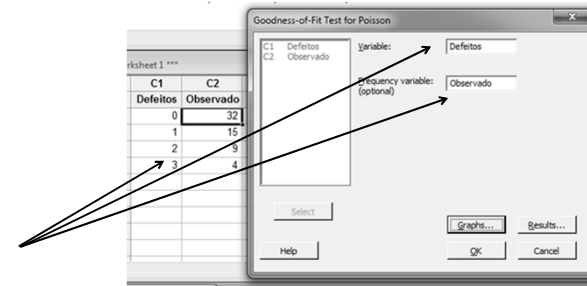
x    P( X <= x )
2,94  0,913589
```

✓ $= P\{\chi^2_1 > 2,94\} = 1 - 0,913589 = 0,0864$

• Não há evidência amostral para se rejeitar H_0 .

• Teste de Adequação de Ajuste - Poisson

Stat > Basic Statistics > Goodness-of-Fit Test for Poisson



• Saída – Minitab:

```
Goodness-of-Fit Test for Poisson Distribution

Data column: Defeitos
Frequency column: Observado

Poisson mean for Defeitos = 0,75

Defeitos  Observed  Poisson  Expected  Contribution
0          32      0,472367  28,3420   0,47213
1          15      0,354275  21,2565   1,84149
2           9      0,132853   7,9712   0,13279
>=3         4      0,040505   2,4303   1,01380

N  N*  DF  Chi-Sq  P-Value
60  0   2   3,46021  0,177

1 cell(s) (25,00%) with expected value(s) less than 5.
```

✓ Procedimento prejudicado

– Célula com frequência esperada menor que 5

• Saída – Minitab:

✓ Combinação duas últimas células

– Minitab executa teste apenas por meio de comando manual no *prompt*.

```
MTB > PGGoodness 'Defeitos';
SUBC> Frequencies 'Observado';
SUBC> mean 0,75;
SUBC> GBar;
SUBC> GChiSq;
SUBC> Pareto;
SUBC> RTable.

Goodness-of-Fit Test for Poisson Distribution

Data column: Defeitos
Frequency column: Observado

Poisson mean for Defeitos = 0,75

Defeitos  Observed  Poisson  Expected  Contribution
0          32      0,472367  28,3420   0,47213
1          15      0,354275  21,2565   1,84149
>=2         13      0,173359  10,4015   0,64915

N  N*  DF  Chi-Sq  P-Value
60  0   2   2,96277  0,227
```

✓ Importante:

– O p-valor deve ser recalculado pois dessa maneira a rotina considera o parâmetro como conhecido!

- Correção dos graus de liberdade da distribuição:

✓ Estatística de teste observada:

- $X_0^2 = 2,96277$

```
MTB > CDF 2,96277;
SUBC> ChiSq 1.

Cumulative Distribution Function

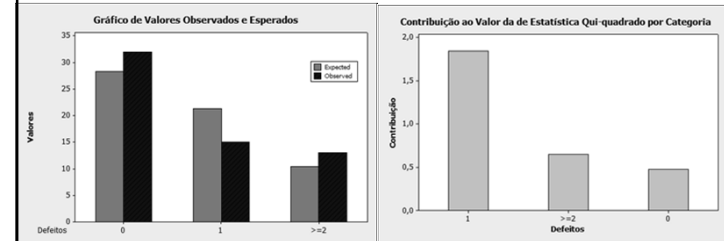
Chi-Square with 1 DF

      x      P( X <= x )
2,96277      0,914798
```

✓ $= P\{\chi^2_1 > 2,94\} = 1 - 0,914798 = 0,0852$

- Não há evidência amostral para se rejeitar H_0 .

✓ Gráficos do procedimento de teste



Exemplo 9-13

- Teste em suprimento de energia em notebook

✓ Suposição: tensão de saída tem distribuição normal

✓ Amostra com $n = 100$ unidades

– Média amostral: 5,04 V

– Desvio-padrão amostral: 0,08 V

- Prática comum:

✓ Escolher limites das células de modo que as frequências esperadas sejam iguais

$$(p_1 = p_2 = \dots = p_k)$$

$$p_i = P\{a_{i-1} \leq X \leq a_i\} = \int_{a_{i-1}}^{a_i} f(x) dx$$

✓ Dados amostrais e cálculo das frequências esperadas

Intervalo de classe	Frequência Observada	Probabilidade	Frequência Esperada
$x < 4,948$	12	$1/8 = 0,125$	12,5
$4,948 \leq x < 4,986$	14	0,125	12,5
$4,986 \leq x < 5,014$	12	0,125	12,5
$5,014 \leq x < 5,040$	13	0,125	12,5
$5,040 \leq x < 5,066$	12	0,125	12,5
$5,066 \leq x < 5,094$	11	0,125	12,5
$5,094 \leq x < 5,132$	12	0,125	12,5
$x \geq 5,132$	14	0,125	12,5
Total	100	0,125	100

– Intervalos igualmente prováveis da normal padrão:

▪ $[0; 0,32)$, $[0,32; 0,675)$, $[0,675; 1,15)$ e $[1,15; \infty)$ e seus ‘espelhos’

– Limite do primeiro intervalo: $\bar{x} - 1,15s = 4,948$

– Limites do segundo intervalo: $[\bar{x} - 1,15s, \bar{x} - 0,675s] = [4,948; 4,986]$

• Procedimento de teste:

1. Variável de interesse: voltagem do suprimento
2. H_0 : forma da distribuição de defeitos é normal
3. H_1 : forma da distribuição de defeitos não é normal
4. $\alpha = 0,05$

5. Estatística de teste: $X_0^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i}$

6. Rejeite H_0 se $X_0^2 > \chi^2_{0,05;5} = 11,07$

7. Estatística observada:

$$X_0^2 = \frac{(12 - 12,5)^2}{12,5} + \frac{(14 - 12,5)^2}{12,5} + \dots + \frac{(14 - 12,5)^2}{12,5} = 0,64$$

8. Conclusão:

- $X_0^2 = 0,64 < 11,07 \rightarrow$ Falhamos em rejeitar H_0 e não evidência forte para indicar que a voltagem de saída não seja normalmente distribuída.

• Cálculo do p-valor – Minitab:

✓ Estatística de teste observada:

• $X_0^2 = 0,64$

```
MTB > CDF 0,64;
SUBC> ChiSq 5.

Cumulative Distribution Function

Chi-Square with 5 DF

    x    P( X <= x )
0,64    0,0139013
```

✓ $= P\{\chi^2_5 > 0,64\} = 1 - 0,01391 = 0,9861$

- Não há evidência amostral para se rejeitar H_0 .

Testes para Tabela de Contingência

Tabelas de Contingência

- n elementos de amostra provenientes de uma população classificados de acordo com dois critérios diferentes

✓ Interesse: saber se os dois métodos de classificação são estatisticamente independentes

✓ Exemplo:

- Indivíduos de amostra classificados por sexo (M ou F) e por habilidade manual (destro ou canhoto)
- Há associação entre as duas categorias?

- Suposições:
 - ✓ 1º método de classificação: r níveis
 - ✓ 2º método de classificação: c níveis
 - ✓ O_{ij} : frequência observada para o nível i do 1º método de classificação e nível j do 2º método
- Dados gerais – Tabela de contingência $r \times c$:

		Colunas			
		1	2	...	c
Linhas	1	O_{11}	O_{12}	...	O_{1c}
	2	O_{21}	O_{22}	...	O_{2c}
	...	...	...	...	...
	r	O_{r1}	O_{r2}	...	O_{rc}

- Teste de Hipóteses:
 - ✓ H_0 : Os métodos de classificação são independentes (não há associação entre as variáveis categóricas)
 - ✓ H_1 : Os métodos de classificação não são independentes (há associação entre as variáveis categóricas)
- Para um elemento selecionado ao acaso (sob H_0):

$$p_{ij} = u_i v_j$$

- ✓ p_{ij} : probabilidade de cair na ij -ésima célula
- ✓ u_i : probabilidade de cair na linha i
- ✓ v_j : probabilidade de cair na coluna j

- Procedimento de teste para amostra grande:

- ✓ Estatística de teste:

$$X_0^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

- ✓ Distribuição amostral:

$$X_0^2 \stackrel{H_0}{\sim} \chi_{(r-1) \times (c-1)}^2$$

- ✓ Regra de decisão:

- Rejeitar H_0 se $X_0^2 > \chi_{\alpha, (r-1) \times (c-1)}^2$.

- Cálculo da frequência esperada (E_{ij}): (supondo independência entre as variáveis)

- ✓ Estimadores de u_i e v_j :

$$\hat{u}_i = \frac{1}{n} \sum_{j=1}^c O_{ij} \quad \hat{v}_j = \frac{1}{n} \sum_{i=1}^r O_{ij}$$

- ✓ Frequência esperada de cada célula

$$E_{ij} = n \hat{u}_i \hat{v}_j = \frac{1}{n} \sum_{j=1}^c O_{ij} \sum_{i=1}^r O_{ij}$$

Exemplo 9-14

- Preferência para Plano de Pensão:

- ✓ Interesse:

- Verificar se a escolha entre 3 planos de pensão é independente da classificação do trabalho

- ✓ Amostra aleatória de 500 empregados

- ✓ Nível de significância do teste (α): 5%

- Opiniões dos 500 empregados da amostra:

Classificação do Trabalho	Plano de Pensão			Total
	1	2	3	
Assalariado	160	140	40	340
Diarista	40	60	60	160
Total	200	200	100	500

- Estimativas das probabilidades marginais:

$$\hat{u}_1 = \frac{340}{500} = 0,68$$

$$\hat{v}_1 = \frac{200}{500} = 0,40$$

$$\hat{u}_2 = \frac{160}{500} = 0,32$$

$$\hat{v}_2 = \frac{200}{500} = 0,40$$

$$\hat{v}_3 = \frac{100}{500} = 0,20$$

- Estimativa das probabilidades conjuntas (p_{ij}):

Classificação do Trabalho	Plano de Pensão			Total
	1	2	3	
Assalariado	0,272	0,272	0,136	0,68
Diarista	0,128	0,128	0,064	0,32
Total	0,40	0,40	0,20	1,00

(0,20)(0,68)

- Frequências esperadas:

Classificação do Trabalho	Plano de Pensão			Total
	1	2	3	
Assalariado	136	136	68	340
Diarista	64	64	32	160
Total	200	200	100	500

(500)(0,136)

- Procedimento de teste:

1. Variável de interesse: preferência entre planos

2. H_0 : a preferência é independente do trabalho

3. H_1 : a preferência não é independente do trabalho

4. $\alpha = 0,05$

5. Estatística de teste: $X_0^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$

6. Rejeite H_0 se $X_0^2 > \chi^2_{0,05; 2} = 5,99$

7. Estatística observada:

$$X_0^2 = \frac{(160 - 136)^2}{136} + \frac{(140 - 136)^2}{136} + \dots + \frac{(60 - 32)^2}{32} = 49,63$$

8. Conclusão:

- $X_0^2 = 49,63 > 5,99 \rightarrow$ Rejeitamos H_0 e concluímos que a preferência para planos de pensão não é independente da classificação de trabalho.

- Cálculo do p-valor – Minitab:

✓ Estatística de teste observada:

- $X_0^2 = 49,63$

```
MTB > CDF 49,63;
SUBC> Chisq 2.

Cumulative Distribution Function

Chi-Square with 2 DF

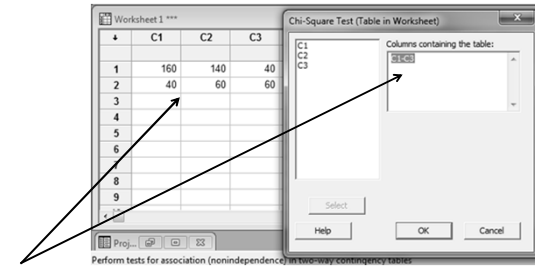
      x      P( X <= x )
49,63      1,00000
```

✓ $= P\{\chi^2_2 > 49,63\} = 1 - 1,00000 \approx 0$

- Confirma a rejeição de H_0 .

- Tabela de Contingência – Teste de Independência:

Stat > Tables > Chi-Square Test



- Saída – Minitab:

```
Chi-Square Test: C1; C2; C3
Expected counts are printed below observed counts
Chi-Square contributions are printed below expected counts

      C1      C2      C3      Total
1      160      140      40      340
  136,00  136,00  68,00
   4,235   0,118  11,529
2       40       60       60      160
   64,00   64,00  32,00
   9,000   0,250  24,500
Total     200     200     100     500

Chi-Sq = 49,632; DF = 2; P-Value = 0,000
```

✓ Rejeita-se H_0 .

- Há associação entre preferência para planos de pensão e classificação de trabalho

Teste de Homogeneidade

- Situação:
 - ✓ Há r populações de interesse e cada população é dividida em c categorias
 - ✓ Toma-se uma amostra da i -ésima população e as contagens são colocadas nas colunas da i -ésima linha
- Objetivo:
 - ✓ Investigar se as proporções das c categorias são as mesma para todas as populações

- Hipóteses:
 - ✓ H_0 : populações são homogêneas com relação às categorias
 - ✓ H_1 : populações não são homogêneas em relação às categorias
- Procedimento de teste:
 - ✓ Idêntico ao teste de independência

Referências

Bibliografia Recomendada

- Montgomery, D. C. (LTC)
Estatística Aplicada e Probabilidade para Engenheiros
- Pinheiro, J. I. D et al. (Campus)
Probabilidade e Estatística: Quantificando a Incerteza