

UNIDADE 4

1) Uma variável tem média 200 e desvio padrão 12 na população, com uma distribuição bastante assimétrica.

a) Imagine que serão retiradas amostras aleatórias de 2 elementos desta população.

a.1 – Encontre a média das médias amostrais. (R. 200)

a.2 – Encontre o desvio padrão das médias amostrais (R. 8,485)

a.3 – A distribuição das médias amostrais será aproximadamente normal? JUSTIFIQUE.

b) Imagine que serão retiradas amostras aleatórias de 225 elementos desta população.

b.1 – Encontre a média das médias amostrais. (R. 200)

b.2 – Encontre o desvio padrão das médias amostrais (R. 0,8)

b.3 – A distribuição das médias amostrais será aproximadamente normal? JUSTIFIQUE.

a) *Serão retiradas amostras de 2 elementos. Pode ser considerado um tamanho pequeno de amostra para uma variável quantitativa.*

a.1 – *A teoria nos mostra (Unidade 4, página 109) que o valor esperado das médias amostrais (média das médias amostrais) será igual à média populacional, independente do tamanho da amostra. Então, como a média populacional vale 200, a média das médias amostrais será igual a 200 também: $E(\bar{x}) = \mu = 200$*

a.2 – *A teoria também nos mostra (Unidade 4, página 109) que o desvio padrão das médias amostrais será igual ao desvio padrão populacional dividido pela raiz quadrada do tamanho da amostra. Então: $s_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{12}{\sqrt{2}} = 8,485$*

a.3 – *A distribuição das médias amostrais para amostras de 2 elementos não deverá ser aproximadamente normal, porque como a distribuição da variável na população é bastante assimétrica e as amostras tem apenas 2 elementos, número provavelmente insuficiente para que o teorema central do limite seja atendido.*

b) *Serão retiradas amostras de 225 elementos. Pode ser considerado um tamanho grande de amostra para uma variável quantitativa.*

b.1 – *A teoria nos mostra (Unidade 4, página 109) que o valor esperado das médias amostrais (média das médias amostrais) será igual à média populacional, independente do tamanho da amostra. Então, como a média populacional vale 200, a média das médias amostrais será igual a 200 também: $E(\bar{x}) = \mu = 200$*

b.2 – *A teoria também nos mostra (Unidade 4, página 109) que o desvio padrão das médias amostrais será igual ao desvio padrão populacional dividido pela raiz quadrada do tamanho da amostra. Então: $s_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{12}{\sqrt{225}} = 0,8$*

b.3 – *A distribuição das médias amostrais para amostras de 200 elementos deverá ser aproximadamente normal, porque como a distribuição da variável na população é bastante assimétrica e as amostras tem 225 elementos, número provavelmente suficiente para que o teorema central do limite seja atendido.*

2) O censo indicou que 60% dos homens de um município são casados. Se fossem retiradas amostras aleatórias de 200 elementos da população de homens.

a) Qual é a média da proporção amostral de casados? (R. 0,60).

b) Qual é o desvio padrão da proporção amostral de casados? (R. 0,0346).

c) A distribuição das proporções amostrais será aproximadamente normal? JUSTIFIQUE.

d) Supondo que a distribuição das proporções amostrais possa ser considerada normal, qual é a probabilidade de uma proporção de uma das amostras aleatórias diferir em mais de 5% (para mais ou para menos) da proporção populacional? (R. aproximadamente 0,1484)

a) *A teoria nos mostra (Unidade 4, página 111) que o valor esperado das proporções amostrais (média das proporções amostrais) de um dos valores de uma variável qualitativa será igual a*

proporção populacional do mesmo valor na população. Então se forem retiradas amostras aleatórias de 200 elementos, e registradas as proporções amostrais de casados:

$$E(p) = \pi = 0,60 \text{ (60\%)}$$

b) A teoria nos mostra (Unidade 4, página 111) que a variância das proporções amostrais será igual à proporção populacional do valor da variável qualitativa, vezes o seu complementar dividido pelo tamanho da amostra. Como o desvio padrão é a raiz quadrada da variância:

$$s_p = \sqrt{\frac{\pi \times (1 - \pi)}{n}} = \sqrt{\frac{0,6 \times (1 - 0,6)}{200}} = 0,0346$$

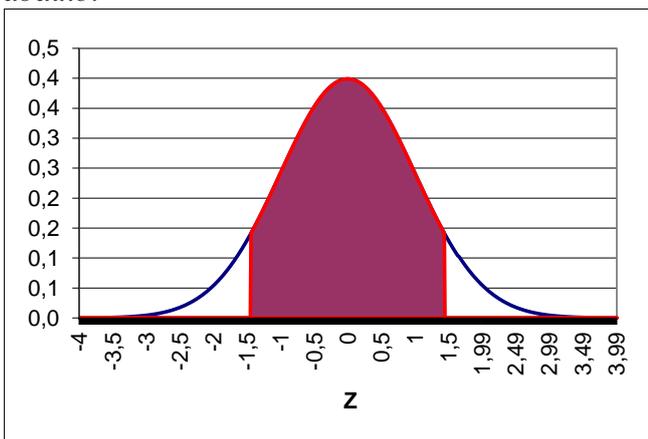
c) A teoria nos mostra que o modelo binomial pode ser aproximado pelo normal (ver Unidade 2, Unidade 9, página 111). No modelo binomial usa-se contagem de ocorrências de um valor de uma variável qualitativa, mas se usarmos proporção de ocorrências do mesmo valor também será possível aproximar pela normal. Na Unidade 2 mostrou-se que a distribuição de X (número de ocorrências de um valor) poderá ser aproximada pela normal se $n \times \pi \geq 5$ e $n \times (1 - \pi) \geq 5$, então: $n \times \pi = 200 \times 0,6 = 120 \geq 5$ e $n \times (1 - \pi) = 200 \times (1 - 0,6) = 80 \geq 5$. Como ambos são maiores do que 5 é possível aproximar pela normal a distribuição de X , e a distribuição de p (X/n) também pode ser aproximada por uma normal.

d) Sabemos que a distribuição amostral da proporção é normal, e conhecemos seu desvio padrão (0,0346, calculado na letra b). Desejamos calcular a probabilidade de que a proporção amostral difira da proporção populacional em mais do que 5% (0,05), para mais ou para menos. Para calcular probabilidades em uma distribuição normal (ver Unidade 2) geralmente transformamos o valor de interesse em um valor da variável Z , da distribuição normal padrão (com média zero e desvio padrão igual a um), e obtemos a probabilidade através da tabela. Precisamos calcular Z :

$$Z = \frac{\text{"valor"} - \text{média}}{\text{desviopadrão}}$$

Conhecemos o desvio padrão, é o desvio padrão da proporção amostral (0,0346). A diferença "valor" – média deve ser igual a 0,05 (5%). Sendo assim: $Z = \frac{0,05}{0,0346} = 1,45$

Queremos encontrar a seguinte probabilidade: $P(-1,45 > Z > 1,45)$, a probabilidade de que Z seja menor do que -1,45 e maior do que 1,45 (correspondente a 0,05 de diferença). Veja a figura abaixo:



Repare que estamos interessados nos valores das caudas. Devido à simetria da distribuição normal padrão em relação a sua média zero $P(Z > 1,45) = P(Z < -1,45)$. Procurando na tabela da distribuição normal padrão:

$P(Z > 1,45) = 0,0735 = P(Z < -1,45)$. Então, a probabilidade procurada $P(-1,45 > Z > 1,45)$ será igual a $2 \times 0,0735 = 0,147$.

UNIDADE 5

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:

- Faça uma estimação por intervalo para o tempo médio de utilização para toda a população de clientes da agência lotérica, utilizando um nível de confiança de 95%. R. 190,35 a 199,65 segundos
- Se a legislação estabelecesse que em média o tempo seja de 180 segundos para atendimento, a lotérica está atendendo ao padrão? JUSTIFIQUE.
- A amostra utilizada seria suficiente para uma precisão de 1 minuto? R. Sim, $n = 1$.

a) A variável sob análise (tempo de atendimento) é **QUANTITATIVA**. Portanto serão feitas inferências sobre a **MÉDIA**.

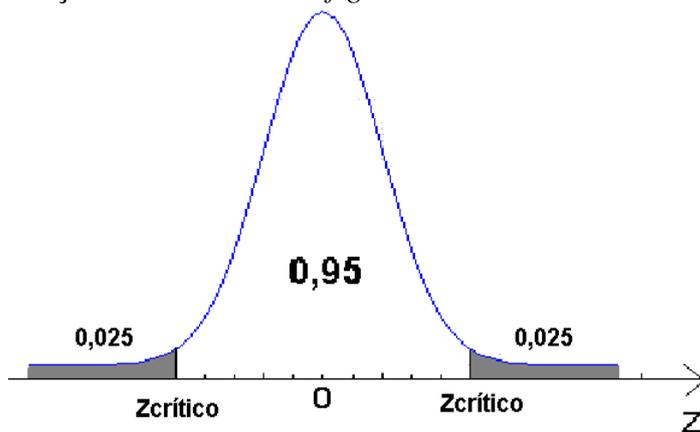
O parâmetro de interesse é a média populacional μ do tempo de atendimento.

Adotou-se um nível de confiança de 95%, então $1 - \alpha = 0,95$ $\alpha = 0,05$ $\alpha/2 = 0,025$. 3)

Estatísticas disponíveis são: **média amostral** = 195 segundos $s = 15$ segundos $n = 40$

Definição da 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 40 elementos (portanto mais de 30) a variável de teste será **Z** da distribuição normal.

Encontrar o valor de **Z_{crítico}** : como o Intervalo de Confiança para a média é bilateral, teremos uma situação semelhante à da figura abaixo:



Para encontrar o valor crítico devemos procurar na tabela da distribuição normal padrão pela probabilidade 0,025 e 0,975 ($0,95 + 0,025$) O valor da probabilidade pode ser visto na figura ao lado: os valores críticos serão **Z_{0,025}** e **Z_{0,975}** os quais serão iguais em módulo.

$P(Z > Z_{\text{crítico}}) = 0,025$. Então **Z_{crítico}** será igual a 1,96 (em módulo).

Ou, usar a função do Excel/Calc: **INV.NORMP.N(0,975)**, usa-se 0,975 porque é a probabilidade acumulada até o **Z_{crítico}** superior ($0,025 + 0,95$).

Passa-se agora a determinação dos limites do intervalo, através da expressão abaixo (cujo resultado será somado e subtraído da média amostral) para determinar os limites do intervalo:

$$e_0 = \frac{Z_{\text{crítico}} \times s}{\sqrt{n}} = \frac{1,96 \times 15}{\sqrt{40}} = 4,65 \text{ segundos}$$

$$L_I = \bar{x} - e_0 = 195 - 4,65 = 190,35 \text{ segundos}$$

$$L_S = \bar{x} + e_0 = 195 + 4,65 = 199,65 \text{ segundos}$$

Então o intervalo de 95% de confiança para a média populacional do tempo de atendimento é [190,35;199,65] segundos. Interpretação: há 95% de probabilidade de que a média populacional do tempo de atendimento esteja entre 190,35 e 199,65 segundos.

b) **NÃO**, não está atendendo porque há 95% de probabilidade da média populacional do tempo de atendimento estar entre 190,35 e 199,65 segundos, portanto mais de 10 segundo acima do valor limite de 180 segundos.

c) Como a variância populacional é **DESCONHECIDA**, e o tamanho da amostra é maior do que 30 elementos, pode ser usada a variável de teste **Z** da distribuição normal padrão. Assim será

empregada a seguinte expressão para calcular o tamanho mínimo de amostra para a estimação por intervalo da média populacional.

$$n = \left(\frac{Z_{\text{crítico}} \times s}{e_0} \right)^2$$

O nível de significância é o mesmo do item a. Sendo assim, o valor crítico continuará sendo o mesmo: $Z_{\text{crítico}} = 1,96$. O desvio padrão amostral vale 15 segundos, e o valor de e_0 , a precisão, foi fixado em 1 minuto, ou seja, 60 segundos. Basta então substituir os valores na expressão:

$$n = \left(\frac{Z_{\text{crítico}} \times s}{e_0} \right)^2 = \left(\frac{1,96 \times 15}{60} \right)^2 = 0,24 \cong 1 \text{ elementos}$$

Observe que o tamanho mínimo de amostra necessário para atender a 95% de confiança e precisão de 60 segundos deveria ser de 1 elemento. Como a amostra coletada possui 40 elementos ela é plenamente SUFICIENTE para a significância e precisão exigidas.

2) O tempo de montagem de determinados conectores utiliza um processo já há algum tempo, que dura em média 3,5 segundos. Está sendo analisada a possibilidade de troca deste processo para um outro que se afirma possuir um tempo de montagem menor. Para esta análise foram observados os tempos de montagem de conectores por um operário padrão utilizando o novo processo e foram anotados os seguintes valores (em segundos): 2,5 2,5 2,6 3,0 3,2 3,5 3,7 3,7 2,1 2,4 2,7 2,8 3,1 3,1 3,6 3,6 2,5 2,9 2,8 3,8

Considerando a situação exposta acima e utilizando um nível de confiança de 95% :

a) Estime o tempo médio de montagem dos conectores utilizando o novo processo. R. 2,767 a 3,243 s.

b) Considerando que o tempo médio aceitável seja de 3 minutos, o novo processo atende ao padrão? JUSTIFIQUE.

c) Calcule o tamanho mínimo da amostra que seria necessária para estimar a média com 95% de confiança e precisão de 0,5 segundos. R. $n = 5$

A variável sob análise (tempo de montagem) é QUANTITATIVA. Portanto serão feitas inferências sobre a MÉDIA.

a) Calculando a média e o desvio padrão amostrais obtêm-se: $\bar{x} = 3,005 \text{ min}$ e $s = 0,5083 \text{ min}$

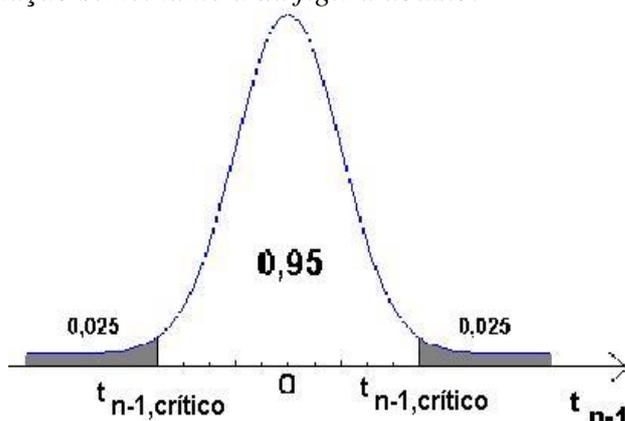
O parâmetro de interesse é a média populacional μ do tempo de montagem do novo processo.

Adotou-se um nível de confiança de 95%, então $1 - \alpha = 0,95$ $\alpha = 0,05$ $\alpha/2 = 0,025$.

Estatísticas: **média amostral** = 3,005 segundos **s** = 0,5083 segundos **n** = 20

Definição da variável de teste: uma vez que a variância populacional da variável é DESCONHECIDA (o valor fornecido é o desvio padrão AMOSTRAL), e a amostra retirada apresenta 20 elementos (portanto menos de 30) a distribuição amostral da média será **t** de Student, e a variável de teste será **t_{n-1}** .

Encontrar o valor de **$t_{n-1, \text{crítico}}$** : como o Intervalo de Confiança para a média é bilateral, teremos uma situação semelhante à da figura abaixo:



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 $20 - 1 = 19$ graus de liberdade. O valor da probabilidade pode ser visto na figura ao lado: $P(t > t_{n-1, \text{crítico}}) = 0,025$ e $P(t > t_{n-1, \text{crítico}}) = 0,975$ (os valores são iguais em módulo).

E o valor de **$t_{n-1, \text{crítico}}$** será igual a **2,093** (em módulo)

Ou, usar a função do Excel/Calc: $INV.T.BC(0,05;19)$. Coloca-se como argumento a significância (0,05) e o número de graus de liberdade (19). Esta função só pode ser usada para obter valor crítico em casos bilaterais (intervalo de confiança ou testes de hipóteses bilaterais)

Passa-se agora a determinação dos limites do intervalo, através da expressão abaixo (cujo resultado será somado e subtraído da média amostral) para determinar os limites do intervalo:

$$e_0 = \frac{t_{n-1,crítico} \times s}{\sqrt{n}} = \frac{2,093 \times 0,5083}{\sqrt{20}} = 0,238 \text{ segundos}$$

$$L_I = \bar{x} - e_0 = 3,005 - 0,238 = 2,767 \text{ segundos}$$

$$L_S = \bar{x} + e_0 = 3,005 + 0,238 = 3,243 \text{ segundos}$$

Então o intervalo de 95% de confiança para a média populacional do tempo de montagem pelo novo processo é [2,767;3,243] segundos. Interpretação: há 95% de probabilidade de que a verdadeira média populacional do tempo de montagem pelo novo processo esteja entre 2,767 e 3,243 segundos.

b) Sim, o tempo médio de montagem é aceitável, porque há 95% de probabilidade de que a média populacional esteja entre 2,767 e 3,243 segundos, o que inclui o valor especificado de 3 segundos praticamente no centro do intervalo.

c) Como a variância populacional é DESCONHECIDA, e o tamanho da amostra é menor do que 30 elementos a distribuição amostral da média será **t** de Student, e a variável de teste será **t_{n-1}**. Assim será usada a seguinte expressão para calcular o tamanho mínimo de amostra para a estimação por intervalo da média populacional.

$$n = \left(\frac{t_{n-1,crítico} \times s}{e_0} \right)^2$$

O nível de significância é o mesmo do item a. Sendo assim, o valor crítico continuará sendo o mesmo: **t_{n-1,crítico}** = 2,093. O desvio padrão amostral vale 0,5083 segundos, e o valor de **e₀**, a precisão, foi fixado em 0,5 minutos. Basta então substituir os valores na expressão:

$$n = \left(\frac{t_{n-1,crítico} \times s}{e_0} \right)^2 = \left(\frac{2,093 \times 0,5083}{0,5} \right)^2 = 4,53 \cong 5 \text{ elementos}$$

Observe que o tamanho mínimo de amostra necessário para atender a 95% de confiança e precisão de 0,5 segundos deveria ser de 5 elementos. Como a amostra coletada possui 20 elementos ela é plenamente **SUFICIENTE** para a significância e precisão exigidas.

3) 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

a) Estime o atraso médio na entrega dos pedidos com confiança de 90%.

b) Um dos clientes da empresa propôs romper o contrato, pois reclama que os atrasos são muito grandes, ele aceitaria em média 2 dias. Com base nos resultados do item a, a empresa deve se preocupar com a possibilidade de rompimento do contrato? JUSTIFIQUE.

c) Para a situação do item a o tamanho da amostra é suficiente, se é necessária uma precisão de 0,5 dias, para o mesmo nível de confiança? R. Não, $n = 109$

A variável sob análise (tempo de atraso) é **QUANTITATIVA**. Portanto serão feitas inferências sobre a **MÉDIA**.

a)

O parâmetro de interesse é a média populacional **μ** do tempo de atraso nas entregas

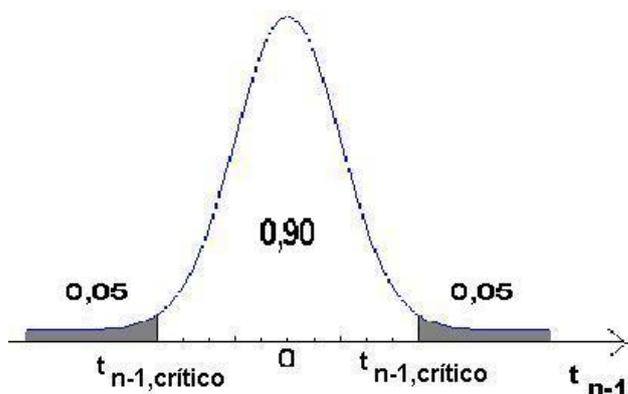
Adotou-se um nível de confiança de 90%, então $1 - \alpha = 0,90$ $\alpha = 0,10$ $\alpha/2 = 0,05$.

Estatísticas disponíveis são: **média amostral** = 3,3 dias **s** = 3,0105 dias **n** = 20

Definição da variável de teste: uma vez que a variância populacional da variável é **DESCONHECIDA** (o valor fornecido é o desvio padrão **AMOSTRAL**), e a amostra retirada

apresenta 20 elementos (portanto menos de 30) a distribuição amostral da média será t de Student, e a variável de teste será t_{n-1} .

Encontrar o valor de $t_{n-1,crítico}$: como o Intervalo de Confiança para a média é bilateral, teremos uma situação semelhante à da figura abaixo:



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 $20 - 1 = 19$ graus de liberdade. O valor da probabilidade pode ser visto na figura ao lado: $P(t > t_{n-1,crítico}) = 0,05$ e

$P(t > t_{n-1,crítico}) = 0,95$ (os valores são iguais em módulo).

E o valor de $t_{n-1,crítico}$ será igual a **1,729** (em módulo)

Ou, usar a função do Excel/Calc: $INV.T.BC(0,10;19)$. Coloca-se como argumento a significância (0,10) e o número de graus de liberdade (19). Esta função só pode ser usada para obter valor crítico em casos bilaterais (intervalo de confiança ou testes de hipóteses bilaterais)

Passa-se agora a determinação dos limites do intervalo, através da expressão abaixo (cujo resultado será somado e subtraído da média amostral) para determinar os limites do intervalo:

$$e_0 = \frac{t_{n-1,crítico} \times S}{\sqrt{n}} = \frac{1,729 \times 3,0105}{\sqrt{20}} = 1,164 \text{ dias}$$

$$L_I = \bar{x} - e_0 = 3,3 - 1,164 = 2,136 \text{ dias}$$

$$L_S = \bar{x} + e_0 = 3,3 + 1,164 = 4,464 \text{ dias}$$

Então o intervalo de 90% de confiança para a média populacional do tempo de atraso é [2,136;4,464] dias.

Interpretação: há 90% de probabilidade de que a verdadeira média populacional do tempo de atraso na entrega dos pedidos esteja entre 2,136 e 4,464 dias.

b) Sim, a empresa deve ficar preocupada com a quebra de contrato: há 90% de probabilidade de que a média do tempo de atraso esteja entre 2,136 e 4,464 dias, portanto acima do limite máximo aceito pelo cliente que vale 2 dias.

c) Para a mesma situação do item a.

Como a variância populacional é DESCONHECIDA, e o tamanho da amostra é menor do que 30 elementos a distribuição amostral da média será t de Student, e a variável de teste será t_{n-1} . Assim será usada a seguinte expressão para calcular o tamanho mínimo de amostra para a estimação por

intervalo da média populacional. $n = \left(\frac{t_{n-1,crítico} \times S}{e_0} \right)^2$

O nível de significância é o mesmo do item a. Sendo assim, o valor crítico continuará sendo o mesmo: $t_{n-1,crítico} = 1,729$. O desvio padrão amostral vale 3,0105 dias, e o valor de e_0 , a precisão, foi fixado em 0,5 dias. Basta então substituir os valores na expressão:

$$n = \left(\frac{t_{n-1,crítico} \times S}{e_0} \right)^2 = \left(\frac{1,729 \times 3,0105}{0,5} \right)^2 = 108,37 \cong 109 \text{ elementos}$$

Observe que o tamanho mínimo de amostra necessário para atender a 90% de confiança e precisão de 0,5 dias deveria ser de 109 elementos. Como a amostra coletada possui 20 elementos ela é INSUFICIENTE para a significância e precisão exigidas, seria necessário obter mais 89 medidas.

4) A satisfação da população em relação a determinado governo foi pesquisada através de uma amostra com a opinião de 1000 habitantes do estado. Destes, 585 se declararam insatisfeitas com a administração estadual. Admitindo-se um nível de significância de 5%, solucione os itens abaixo.

a) Estime o percentual da população que está insatisfeita com a administração estadual. R. 55,45% a 61,55%

b) Com base no resultado do item a você considera a população do estado satisfeita com o governo? JUSTIFIQUE.

c) Qual o tamanho da amostra necessária para a estimação se a empresa responsável pela pesquisa estipulou uma folga máxima de 2,5% ? R. $n = 1493$

A variável sob análise (opinião sobre a administração estadual) é **QUALITATIVA**, e somente admite dois resultados: *satisfeita ou insatisfeita*. Portanto serão feitas inferências sobre a proporção de pessoas insatisfeitas ou satisfeitas.

a)

O parâmetro de interesse é a proporção populacional π de pessoas insatisfeitas com a administração estadual.

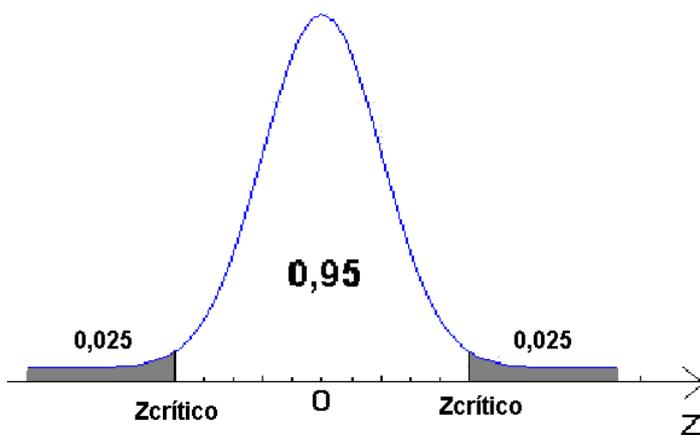
Adotou-se um nível de significância de 5%, então $\alpha = 0,05$ $\alpha/2 = 0,025$ $1 - \alpha = 0,95$

As estatísticas são: **proporção amostral de pessoas insatisfeitas** $p = 585/1000 = 0,585$, o seu complementar $1 - p = 0,415$ e $n = 1000$ elementos.

Definição da variável de teste: precisamos verificar se é possível fazer a aproximação pela normal, então, $n \times p = 1000 \times 0,585 = 585 > 5$ e $n \times (1 - p) = 1000 \times 0,415 = 415 > 5$.

Como ambos os produtos satisfazem as condições para a aproximação podemos usar a variável **Z** da distribuição normal padrão.

Encontrar o valor de **Z_{crítico}** : como o Intervalo de Confiança para a média é bilateral, teremos uma situação semelhante à da figura abaixo:



Para encontrar o valor crítico devemos procurar na tabela da distribuição normal padrão pela probabilidade 0,025 e 0,975 (0,95 + 0,025) O valor da probabilidade pode ser visto na figura ao lado: os valores críticos serão **Z_{0,025}** e **Z_{0,975}** os quais serão iguais em módulo. $P(Z > Z_{\text{crítico}}) = 0,025$. Então **Z_{crítico}** será igual a 1,96 (em módulo).

Ou, usar a função do Excel/Calc: $\text{INV.NORMP.N}(0,975)$, usa-se 0,975 porque é a probabilidade acumulada até o **Z_{crítico}** superior (0,025 + 0,95).

Passa-se agora a determinação dos limites do intervalo, através da expressão abaixo (cujo resultado será somado e subtraído da proporção amostral de pessoas insatisfeitas) para determinar os limites do intervalo:

$$e_0 = Z_{\text{crítico}} \times \sqrt{\frac{p \times (1 - p)}{n}} = 1,96 \times \sqrt{\frac{0,585 \times 0,415}{1000}} = 0,0305$$

$$L_i = p - e_0 = 0,585 - 0,0305 = 0,5545 \quad L_s = p + e_0 = 0,585 + 0,0305 = 0,6155$$

Então, o intervalo de 95% de confiança para a proporção populacional de pessoas insatisfeitas com a administração estadual é [55,45%;61,55%].

Interpretação: há 95% de probabilidade de que a verdadeira proporção populacional de pessoas insatisfeitas esteja entre 55,45% e 61,55%.

b) Não, a população está insatisfeita com o governo, há 95% de probabilidade de que a proporção populacional de Insatisfeitos esteja entre 55,45% e 61,55%, portanto acima da maioria absoluta

(51%), e a consequência é a que há 95% de probabilidade de que a proporção de satisfeitos esteja entre 38,45% e 44,55%, portanto abaixo da maioria absoluta.

c) De acordo com o item a é possível utilizar a aproximação pela distribuição normal. Assim, a expressão para o cálculo do tamanho mínimo de amostra para a proporção populacional será:

$$n = \left(\frac{Z_{\text{crítico}}}{e_0} \right)^2 \times p \times (1 - p)$$

Os valores de p e $1 - p$ já são conhecidos: $p = 0,585$ $1 - p = 0,415$

O nível de confiança exigido é de 95%: para encontrar o valor crítico devemos procurar na tabela da distribuição normal padrão pela probabilidade 0,025 e 0,975 (0,95+0,025); os valores críticos serão $Z_{0,025}$ e $Z_{0,975}$ os quais serão iguais em módulo. $P(Z > Z_{\text{crítico}}) = 0,025$. Então $Z_{\text{crítico}}$ será igual a 1,96 (em módulo). A precisão foi fixada em 2,5% (0,025). Substituindo os valores na expressão acima:

$$n = \left(\frac{Z_{\text{crítico}}}{e_0} \right)^2 \times p \times (1 - p) = \left(\frac{1,96}{0,025} \right)^2 \times 0,585 \times 0,415 = 1492,23 \cong 1493$$

Observe que o tamanho mínimo de amostra necessário para atender a 95% de confiança e precisão de 2,5% deveria ser de 1493 elementos. Como a amostra coletada possui apenas 1000 elementos ela é INSUFICIENTE para a confiança e precisão exigidas. Recomenda-se o retorno à população para a retirada aleatória de mais 493 pessoas.

5) Uma fábrica está convertendo as máquinas que aluga para uma versão mais moderna. Até agora foram convertidas 40 máquinas. O tempo médio de conversão foi de 24 horas, com desvio padrão de 3 horas.

a) Determine um intervalo de 98% de confiança para o tempo médio de conversão. R. 22,895 h a 25,105 h

b) A direção da fábrica esperava uma média de 20 h para a conversão. A equipe responsável atingiu este padrão? JUSTIFIQUE.

Adaptado de STEVENSON, W.J. Estatística Aplicada à Administração, São Paulo: Harper do Brasil, 2001.

A variável sob análise (tempo de conversão) é QUANTITATIVA, portanto serão feitas inferências sobre a média.

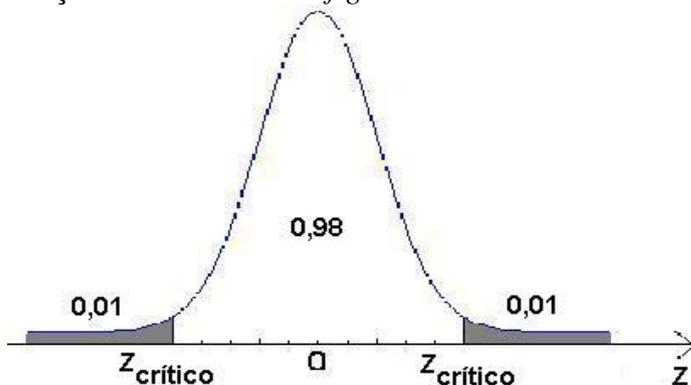
O parâmetro de interesse é a média populacional μ do tempo de conversão.

Adotou-se um nível de confiança de 98%, então $1 - \alpha = 0,98$ $\alpha = 0,02$ $\alpha/2 = 0,01$.

Estatísticas disponíveis são: **média amostral** = 24 horas $s = 3$ horas $n = 40$

Definição da 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 40 elementos (portanto mais de 30) a variável de teste será Z da distribuição normal.

Encontrar o valor de $Z_{\text{crítico}}$: como o Intervalo de Confiança para a média é bilateral, teremos uma situação semelhante à da figura abaixo:



Para encontrar o valor crítico devemos procurar na tabela da distribuição normal padrão pela probabilidade 0,01 e 0,99 (0,98+0,01) O valor da probabilidade pode ser visto na figura ao lado: os valores críticos serão $Z_{0,01}$ e $Z_{0,99}$ os quais serão iguais em módulo.

$P(Z > Z_{\text{crítico}}) = 0,01$. Então $Z_{\text{crítico}}$ será igual a 2,33 (em módulo).

Ou, usar a função do Excel/Calc: $INV.NORMP.N(0,99)$, usa-se 0,99 porque é a probabilidade acumulada até o $Z_{\text{crítico}}$ superior (0,01 + 0,98).

Passa-se agora a determinação dos limites do intervalo, através da expressão abaixo (cujo resultado será somado e subtraído da média amostral) para determinar os limites do intervalo:

$$e_0 = \frac{Z_{\text{crítico}} \times s}{\sqrt{n}} = \frac{2,33 \times 3}{\sqrt{40}} = 1,105 \text{ horas}$$

$$L_I = \bar{x} - e_0 = 24 - 1,105 = 22,895 \text{ horas}$$

$$L_S = \bar{x} + e_0 = 24 + 1,105 = 25,105 \text{ horas}$$

Então o intervalo de 98% de confiança para a média populacional do tempo de conversão das máquinas é [22,895;25,105] horas. Interpretação: há 98% de probabilidade de que a verdadeira média populacional do tempo de conversão esteja entre 22,895 e 25,105 horas.

b) Não, a equipe não atingiu o padrão, há 98% de probabilidade de que a média do tempo de conversão esteja entre 22,895 e 25,105 horas, portanto acima do limite estabelecido de 20 horas.

6) Um banco possui 800 terminais de auto-atendimento instalados no estado de SC. Avaliando 48 terminais, 6 apresentaram defeitos.

a) Estime a proporção de terminais com defeitos. R. 3,144% a 21,86%

b) Você considera o intervalo de confiança obtido na letra a preciso para estimar a proporção de terminais com defeitos? JUSTIFIQUE.

Adaptado de STEVENSON, W.J. Estatística Aplicada à Administração, São Paulo: Harper do Brasil, 2001.

A variável sob análise (estado do terminal) é **QUALITATIVA**, e somente admite dois resultados: sem defeito ou com defeito. Portanto serão feitas inferências sobre a proporção de terminais com defeito ou sem defeito.

O parâmetro de interesse é a proporção populacional π de terminais com defeitos.

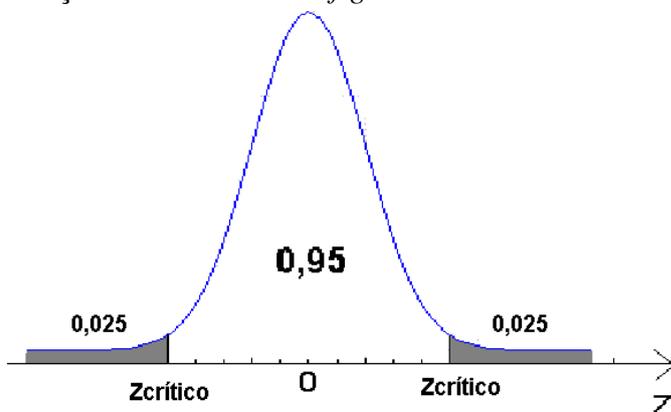
Adotou-se um nível de significância de 5%, então $\alpha = 0,05$ $\alpha/2 = 0,025$ $1 - \alpha = 0,95$

As estatísticas são: **proporção amostral de terminais com defeitos** $p = 6/48 = 0,125$, o seu complementar $1 - p = 0,875$ e $n = 48$ elementos.

Definição da variável de teste: precisamos verificar se é possível fazer a aproximação pela normal, então $n \times p = 48 \times 0,125 = 6 > 5$ e $n \times (1 - p) = 48 \times 0,875 = 42 > 5$.

Como ambos os produtos satisfazem as condições para a aproximação podemos usar a variável **Z** da distribuição normal padrão

Encontrar o valor de $Z_{\text{crítico}}$: como o Intervalo de Confiança para a média é bilateral, teremos uma situação semelhante à da figura abaixo:



Para encontrar o valor crítico devemos procurar na tabela da distribuição normal padrão pela probabilidade 0,025 e 0,975 (0,95 + 0,025) O valor da probabilidade pode ser visto na figura ao lado: os valores críticos serão $Z_{0,025}$ e $Z_{0,975}$ os quais serão iguais em módulo. $P(Z > Z_{\text{crítico}}) = 0,025$. Então $Z_{\text{crítico}}$ será igual a 1,96 (em módulo).

Ou, usar a função do Excel/Calc: $INV.NORMP.N(0,975)$, usa-se 0,975 porque é a probabilidade acumulada até o $Z_{\text{crítico}}$ superior (0,025 + 0,95).

Passa-se agora a determinação dos limites do intervalo, através da expressão abaixo (cujo resultado será somado e subtraído da proporção amostral de terminais com defeito) para determinar os limites do intervalo:

$$e_0 = Z_{\text{crítico}} \times \sqrt{\frac{p \times (1-p)}{n}} = 1,96 \times \sqrt{\frac{0,125 \times 0,875}{48}} = 0,09356$$

$$L_i = p - e_0 = 0,125 - 0,09356 = 0,03144 \quad L_s = p + e_0 = 0,125 + 0,09356 = 0,2186$$

Então, o intervalo de 95% de confiança para a proporção populacional de terminais com defeitos é [3,144%;21,86%]. Interpretação: há 95% de probabilidade de que a verdadeira proporção populacional de terminais com defeito esteja entre 3,144% e 21,86%.

b) Não, ele não é preciso porque a diferença entre os limites é muito grande: o limite superior é quase sete vezes maior do que o inferior.

7) Em uma pesquisa de mercado, acerca da preferência pelo produto X, 300 consumidores foram entrevistados, sendo que 100 declararam consumir o produto.

a) O fabricante quer que você determine um intervalo de 95% para a proporção populacional de pessoas que consomem o produto. R. 28% a 38,67%

b) Um dos diretores do fabricante exige que o intervalo de confiança para a proporção populacional tenha 99% de confiança, com um erro máximo de 2,5%. Qual seria o tamanho de amostra necessário para atingir tais critérios? R. 2358

Adaptado de BUSSAB, W.O., MORETTIN, P. A. Estatística Básica, 8a ed. São Paulo: Saraiva, 2013.

A variável sob análise (consumo do produto) é QUALITATIVA, e só admite dois resultados: consome o produto ou não consome o produto. Então serão feitas inferências sobre a proporção populacional de pessoas que consomem ou não consomem o produto.

a) O parâmetro de interesse é a proporção populacional π de pessoas que consomem o produto.

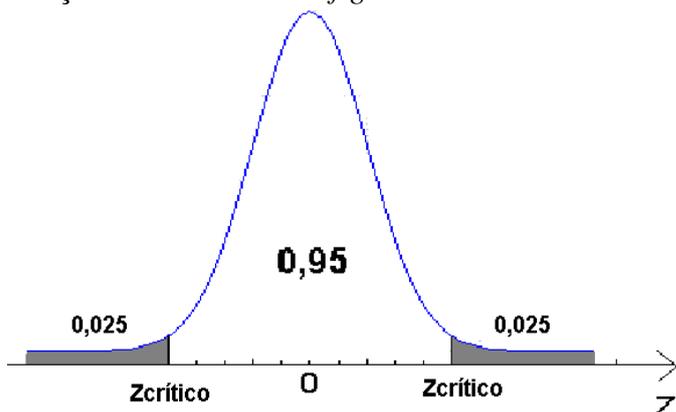
O problema exige uma confiança de 95%, então $1 - \alpha = 0,95$ $\alpha = 0,05$ $\alpha/2 = 0,025$

As estatísticas são: **proporção amostral de pessoas que consomem o produto** $p = 100/300$, o seu complementar $1 - p = 200/300$ e $n = 300$ elementos.

Definição da variável de teste: precisamos verificar se é possível fazer a aproximação pela normal, então $n \times p = 300 \times (100/300) = 100 > 5$ e $n \times (1 - p) = 300 \times (200/300) = 200 > 5$.

Como ambos os produtos satisfazem as condições para a aproximação podemos usar a variável Z da distribuição normal padrão

Encontrar o valor de $Z_{\text{crítico}}$: como o Intervalo de Confiança para a média é bilateral, teremos uma situação semelhante à da figura abaixo:



Para encontrar o valor crítico devemos procurar na tabela da distribuição normal padrão pela probabilidade 0,025 e 0,975 (0,95 + 0,025) O valor da probabilidade pode ser visto na figura ao lado: os valores críticos serão $Z_{0,025}$ e $Z_{0,975}$ os quais serão iguais em módulo. $P(Z > Z_{\text{crítico}}) = 0,025$. Então $Z_{\text{crítico}}$ será igual a 1,96 (em módulo).

Ou, usar a função do Excel/Calc: INV.NORMP.N(0,975), usa-se 0,975 porque é a probabilidade acumulada até o $Z_{\text{crítico}}$ superior (0,025 + 0,95).

Passa-se agora a determinação dos limites do intervalo, através da expressão abaixo (cujo resultado será somado e subtraído da proporção amostral de pessoas que consomem o produto) para determinar os limites do intervalo:

$$e_0 = Z_{\text{crítico}} \times \sqrt{\frac{p \times (1-p)}{n}} = 1,96 \times \sqrt{\frac{(100/300) \times (200/300)}{300}} = 0,0533$$

$$L_i = p - e_0 = (100/300) - 0,0533 = 0,2800 \quad L_s = p + e_0 = (100/300) + 0,0533 = 0,3867$$

Então, o intervalo de 95% de confiança para a proporção populacional de pessoas que consomem o produto é [28%;38,67%]. Interpretação: há 95% de probabilidade de que a verdadeira proporção populacional de pessoas que consomem o produto esteja entre 28% e 38,67%.

b) De acordo com o item anterior é possível utilizar a aproximação pela distribuição normal. Assim, a expressão para o cálculo do tamanho mínimo de amostra para a proporção populacional será:

$$n = \left(\frac{Z_{\text{crítico}}}{e_0} \right)^2 \times p \times (1 - p)$$

Os valores de p e $1 - p$ já são conhecidos: $p = 100/300$ $1 - p = 200/300$

O nível de confiança exigido é de 99%: para encontrar o valor crítico devemos procurar na tabela da distribuição normal padrão pela probabilidade 0,005 e 0,995 (0,99+0,005); os valores críticos serão $Z_{0,005}$ e $Z_{0,995}$ os quais serão iguais em módulo. $P(Z > Z_{\text{crítico}}) = 0,005$. E o valor de $Z_{\text{crítico}}$ será igual a 2,575 (em módulo). A precisão foi fixada em 2,5% (0,025). Substituindo os valores na expressão acima:

$$n = \left(\frac{Z_{\text{crítico}}}{e_0} \right)^2 \times p \times (1 - p) = \left(\frac{2,575}{0,025} \right)^2 \times (100 / 300) \times (200 / 300) = 2357,55 \cong 2358$$

Observe que o tamanho mínimo de amostra necessário para atender a 99% de confiança e precisão de 2,5% deveria ser de 2358 elementos. Como a amostra coletada possui apenas 300 elementos ela é INSUFICIENTE para a confiança e precisão exigidas. Recomenda-se o retorno à população para a retirada aleatória de mais 2058 pessoas.

8) A Polícia Rodoviária Estadual fez recentemente uma pesquisa secreta sobre as velocidades desenvolvidas na SC 401 das 23h às 2h. No período de observação, 100 carros passaram por um aparelho de radar a uma velocidade média de 112 km/h, com desvio padrão de 22 km/h.

a) Construa um intervalo de 95% de confiança para a média da população. R. 107,69 km/h a 116,31 km/h

b) O comando da Polícia divulgaria os resultados do item a apenas se a margem de erro fosse inferior a 10 km/h. Na sua opinião os resultados podem ser divulgados? JUSTIFIQUE.

Adaptado de STEVENSON, W.J. Estatística Aplicada à Administração, São Paulo: Harper do Brasil, 2001.

A variável sob análise (velocidade dos automóveis em km/h) é QUANTITATIVA, então será feita uma inferência sobre a média.

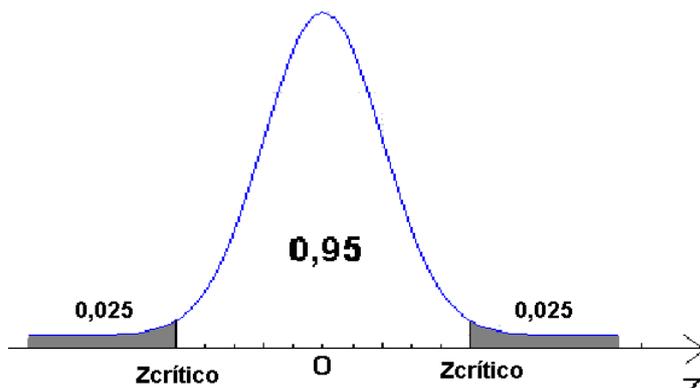
O parâmetro de interesse é a média populacional μ da velocidade dos carros.

O problema exigiu confiança de 95%, então $1 - \alpha = 0,95$ $\alpha = 0,05$ $\alpha/2 = 0,025$.

Estatísticas disponíveis são: **média amostral** = 112 km/h $s = 22$ km/h $n = 100$

Definição da 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 100 elementos (portanto mais de 30) a variável de teste será Z da distribuição normal.

Encontrar o valor de $Z_{\text{crítico}}$: como o Intervalo de Confiança para a média é bilateral, teremos uma situação semelhante à da figura abaixo:



Para encontrar o valor crítico devemos procurar na tabela da distribuição normal padrão pela probabilidade 0,025 e 0,975 (0,95+0,025) O valor da probabilidade pode ser visto na figura ao lado: os valores críticos serão $Z_{0,025}$ e $Z_{0,975}$ os quais serão iguais em módulo.

$P(Z > Z_{\text{crítico}}) = 0,025$. Então $Z_{\text{crítico}}$ Z será igual a 1,96 (em módulo).

Ou, usar a função do Excel/Calc: $\text{INV.NORMP.N}(0,975)$, usa-se 0,975 porque é a probabilidade acumulada até o $Z_{\text{crítico}}$ superior (0,025 + 0,95).

Passa-se agora a determinação dos limites do intervalo, através da expressão abaixo (cujo resultado será somado e subtraído da média amostral) para determinar os limites do intervalo:

$$e_0 = \frac{Z_{\text{crítico}} \times s}{\sqrt{n}} = \frac{1,96 \times 22}{\sqrt{100}} = 4,31 \text{ km/h}$$

$$L_1 = \bar{x} - e_0 = 112 - 4,31 = 107,69 \text{ km/h} \quad L_5 = \bar{x} + e_0 = 112 + 4,31 = 116,31 \text{ km/h}$$

Então o intervalo de 95% de confiança para a média populacional velocidade dos carros é [107,69;116,31] km/h. Interpretação: há 95% de probabilidade de que a verdadeira média populacional da velocidade dos carros esteja entre 107,69 e 116,31 km/h.

b) A Polícia Rodoviária Federal deve divulgar os resultados, porque a margem de erro do intervalo de confiança foi igual a 4,31 km/h, portanto abaixo do limite estabelecido de 10 km/h

9) 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 para a linha de produção para rearranjo dos equipamentos envolvidos quando o percentual de defeituosos é superior a 3%.

a) Determinar um intervalo de 95% de confiança para a proporção de peças defeituosas. R. 2,361% a 4,639%

b) Com base nos resultados do item a o controle de qualidade deve parar a produção? JUSTIFIQUE.

C) Se há interesse em obter um intervalo de 95% de confiança, com precisão de 1,5%, para a proporção de peças defeituosas, qual deve ser o tamanho mínimo de amostra? R. 577

A variável sob análise (classificação das peças) é QUALITATIVA, e só pode assumir dois valores: boa ou defeituosa. Portanto, serão feitas inferências sobre a proporção (percentual) de peças defeituosas ou boas.

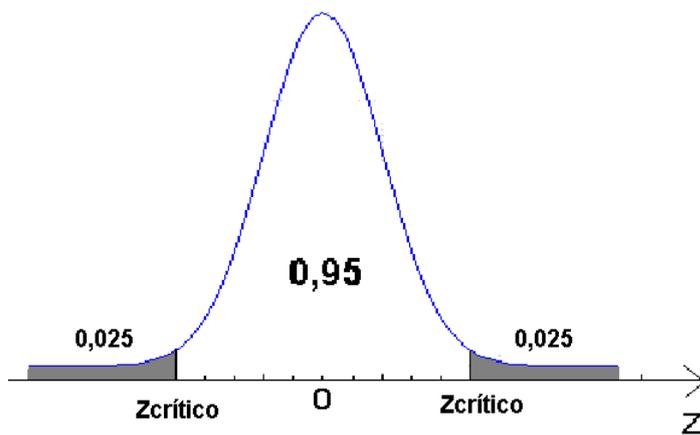
a) O parâmetro de interesse é a proporção populacional π de peças defeituosas

O problema exige uma confiança de 95%, então $1 - \alpha = 0,95$ $\alpha = 0,05$ $\alpha/2 = 0,025$

As estatísticas são: **proporção amostral de peças defeituosas** $p = 35/1000$, o seu complementar $1-p = 965/1000$ e $n = 1000$ elementos.

Definição da variável de teste: precisamos verificar se é possível fazer a aproximação pela normal, então $n \times p = 1000 \times (35/1000) = 35 > 5$ e $n \times (1-p) = 1000 \times (965/1000) = 965 > 5$. Como ambos os produtos satisfazem as condições para a aproximação podemos usar a variável Z da distribuição normal padrão.

Encontrar o valor de $Z_{\text{crítico}}$: como o Intervalo de Confiança para a média é bilateral, teremos uma situação semelhante à da figura abaixo:



Para encontrar o valor crítico devemos procurar na tabela da distribuição normal padrão pela probabilidade 0,025 e 0,975 (0,95+0,025) O valor da probabilidade pode ser visto na figura ao lado: os valores críticos serão $Z_{0,025}$ e $Z_{0,975}$ os quais serão iguais em módulo. $P(Z > Z_{\text{crítico}}) = 0,025$. Então $Z_{\text{crítico}}$ será igual a 1,96 (em módulo).

Ou, usar a função do Excel/Calc: $\text{INV.NORMP.N}(0,975)$, usa-se 0,975 porque é a probabilidade acumulada até o $Z_{\text{crítico}}$ superior (0,025 + 0,95).

Passa-se agora a determinação dos limites do intervalo, através da expressão abaixo (cujo resultado será somado e subtraído da proporção amostral de peças defeituosas) para determinar os limites do intervalo:

$$e_0 = Z_{\text{crítico}} \times \sqrt{\frac{p \times (1-p)}{n}} = 1,96 \times \sqrt{\frac{(35/1000) \times (965/1000)}{1000}} = 0,01139$$

$$L_i = p - e_0 = (35/1000) - 0,01139 = 0,02361 \quad L_s = p + e_0 = (35/1000) + 0,01139 = 0,04639$$

Então, o intervalo de 95% de confiança para a proporção populacional de peças defeituosas é [2,361%;4,639%]. Interpretação: há 95% de probabilidade de que a verdadeira proporção populacional de peças defeituosas esteja entre 2,361% e 4,639%.

b) O controle de qualidade deve pensar em parar a produção, porque há 95% de probabilidade da proporção de defeituosas estar entre 2,361% e 4,639%, e o limite inferior está bem próximo do valor máximo aceitável (3%), indicando que há alta probabilidade de encontrar valores acima de 3% (na realidade valerá $1 - 0,1949 = 0,8051 \Rightarrow 80,51\%$).

c) De acordo com o item a é possível utilizar a aproximação pela distribuição normal. Assim, a expressão para o cálculo do tamanho mínimo de amostra para a proporção populacional será:

$$n = \left(\frac{Z_{\text{crítico}}}{e_0} \right)^2 \times p \times (1-p)$$

Os valores de p e $1-p$ já são conhecidos: $p = 35/1000$ $1-p = 965/1000$

O nível de confiança exigido é de 95%: para encontrar o valor crítico devemos procurar na tabela da distribuição normal padrão pela probabilidade 0,025 e 0,975 (0,95+0,025); os valores críticos serão $Z_{0,025}$ e $Z_{0,975}$ os quais serão iguais em módulo. $P(Z > Z_{\text{crítico}}) = 0,025$. Então $Z_{\text{crítico}}$ será igual a 1,96 (em módulo). A precisão foi fixada em 1,5% (0,015). Substituindo os valores na expressão acima:

$$n = \left(\frac{Z_{\text{crítico}}}{e_0} \right)^2 \times p \times (1-p) = \left(\frac{1,96}{0,015} \right)^2 \times \frac{35}{1000} \times \frac{965}{1000} = 576,67 \cong 577$$

Observe que o tamanho mínimo de amostra necessário para atender a 95% de confiança e precisão de 1,5% deveria ser de 578 elementos. Como a amostra coletada possui 1000 elementos ela é **SUFICIENTE** para a confiança e precisão exigidas.

10) Os índices apresentados pelos alunos do curso de Economia e de Administração estão sendo questionados pelos alunos, no sentido de definirem se há diferença entre os cursos. Para tanto foram analisados os índices de 10 alunos de cada curso, escolhidos aleatoriamente dentre os regularmente matriculados e anotados seus valores, onde se obteve:

Economia	- média 7,3	desvio padrão 2,6
Administração	- média 7,1	desvio padrão 3,1

a) Estime os valores médios dos índices de cada curso com 95% de confiança. R. Economia: 5,44 a

9,16; Administração: 4,88 a 9,32.

b) Com base nos resultados do item a, há diferenças significativas entre as médias dos índices dos dois cursos? JUSTIFIQUE.

c) Para o mesmo nível de confiança de a. Será que 10 alunos é uma amostra suficiente, em ambos os cursos, para estimar seus índices médios, com uma precisão igual a 1? Quais seriam os tamanhos de amostra necessários? R. Economia: insuficiente, $n = 35$. Administração: insuficiente, $n = 50$.

A variável de interesse, nota dos alunos, é QUANTITATIVA. Então serão feitas inferências sobre as MÉDIAS nos cursos de economia e administração.

a) Os parâmetros de interesse são as médias populacionais μ das notas dos cursos de economia e administração.

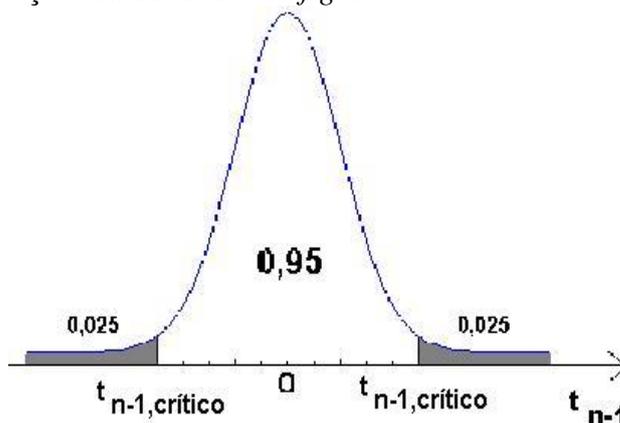
Adotou-se um nível de confiança de 95%, então $1 - \alpha = 0,95$ $\alpha = 0,05$ $\alpha/2 = 0,025$.

Estatísticas: **média amostral de economia** = 7,3 **s** = 2,6 **n** = 10

média amostral de administração = 7,1 **s** = 3,1 **n** = 10

Definição da variável de teste: uma vez que a variância populacional da variável é DESCONHECIDA (o valor fornecido é o desvio padrão AMOSTRAL), e as amostras retiradas apresentam 10 elementos em cada grupo (portanto menos de 30) a distribuição amostral da média será **t** de Student, e a variável de teste será t_{n-1} (tanto para o curso de economia quanto para o de administração).

Encontrar o valor de $t_{n-1,crítico}$: como o Intervalo de Confiança para a média é bilateral, teremos uma situação semelhante à da figura abaixo:



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 $10 - 1 = 9$ graus de liberdade. O valor da probabilidade pode ser visto na figura ao lado: $P(t > t_{n-1,crítico}) = 0,025$ e $P(t > t_{n-1,crítico}) = 0,975$ (os valores são iguais em módulo).

E o valor de $t_{n-1,crítico}$ será igual a **2,262** (*em módulo*)

Ou, usar a função do Excel/Calc: $INV.T.BC(0,05;19)$. Coloca-se como argumento a significância (0,05) e o número de graus de liberdade (9). Esta função só pode ser usada para obter valor crítico em casos bilaterais (intervalo de confiança ou testes de hipóteses bilaterais)

Podemos usar o valor obtido para os intervalos de confiança das médias dos dois cursos.

Passa-se agora a determinação dos limites dos intervalos, através da expressão abaixo (cujo resultado será somado e subtraído das médias amostrais) para determinar os limites dos intervalos:

$$\text{Economia: } e_0 = \frac{t_{n-1,crítico} \times s}{\sqrt{n}} = \frac{2,262 \times 2,6}{\sqrt{10}} = 1,86$$

$$L_1 = \bar{x} - e_0 = 7,3 - 1,86 = 5,44$$

$$L_S = \bar{x} + e_0 = 7,3 + 1,86 = 9,16$$

$$\text{Administração: } e_0 = \frac{t_{n-1,crítico} \times s}{\sqrt{n}} = \frac{2,262 \times 3,1}{\sqrt{10}} = 2,22$$

$$L_1 = \bar{x} - e_0 = 7,1 - 2,22 = 4,88$$

$$L_S = \bar{x} + e_0 = 7,1 + 2,22 = 9,32$$

Economia: o intervalo de 95% de confiança para a média populacional das notas em economia é [5,44;9,16]. Interpretação: há 95% de probabilidade de que a verdadeira média populacional das notas em economia esteja entre 5,44 e 9,16.

Administração: o intervalo de 95% de confiança para a média populacional das notas em administração é [4,88;9,32]. Interpretação: há 95% de probabilidade de que a verdadeira média populacional das notas em economia esteja entre 4,88 e 9,32. Observe que os intervalos são relativamente sobrepostos, o que caracterizaria um “empate técnico” entre as notas dos dois cursos.

b) Não há evidência de diferença entre as médias dos índices dos dois cursos porque há grande sobreposição entre os intervalos de confiança para as médias: o intervalo de economia está contido dentro do de administração, o limite inferior de economia (5,44) é maior do que o de administração (4,88) e o superior de economia é menor do que o de administração (9,32).

c) Como a variância populacional é DESCONHECIDA, e o tamanho da amostra é menor do que 30 elementos a distribuição amostral da média será **t** de Student, e a variável de teste será **t_{n-1}**. Assim será usada a seguinte expressão para calcular o tamanho mínimo de amostra para a estimação por intervalo da média populacional.

$$n = \left(\frac{t_{n-1, \text{crítico}} \times S}{e_0} \right)^2$$

O nível de significância é o mesmo do item a. Sendo assim, o valor crítico continuará sendo o mesmo: **t_{n-1,crítico} = 2,262**. Vamos calcular o tamanho mínimo de amostra para os dois cursos. O desvio padrão amostral das notas em economia vale 2,6, e em administração vale 3,1, e o valor de **e₀**, a precisão, foi fixado em 1 para ambos os cursos. Basta então substituir os valores nas expressões.

Economia:
$$n = \left(\frac{t_{n-1, \text{crítico}} \times S}{e_0} \right)^2 = \left(\frac{2,262 \times 2,6}{1} \right)^2 = 34,59 \cong 35 \text{ elementos}$$

Observe que o tamanho mínimo de amostra necessário para atender a 95% de confiança e precisão de 1 deveria ser de 35 elementos. Como a amostra coletada possui 10 elementos ela é **INSUFICIENTE** para a significância e precisão exigidas, para o curso de Economia.

Administração:
$$n = \left(\frac{t_{n-1, \text{crítico}} \times S}{e_0} \right)^2 = \left(\frac{2,262 \times 3,1}{1} \right)^2 = 49,17 \cong 50 \text{ elementos}$$

Observe que o tamanho mínimo de amostra necessário para atender a 95% de confiança e precisão de 1 deveria ser de 50 elementos. Como a amostra coletada possui 10 elementos ela é **INSUFICIENTE** para a significância e precisão exigidas, para o curso de Administração.

11) O CRA de SC está conduzindo uma pesquisa sobre a opinião dos acadêmicos de administração sobre seus respectivos cursos. Suspeita-se que haja diferença entre as proporções de satisfeitos de instituições públicas e privadas: os acadêmicos das públicas seriam mais satisfeitos. Para avaliar esta suposição foi conduzida uma pesquisa por amostragem, entrevistando alunos de duas instituições públicas, SHUFSC e GASE, e de três privadas, PATÁPIO de SÁ, UNIMALI e UNILUS. Os resultados estão na tabela a seguir:

Medidas	UNIVERSIDADES				
	SHUFSC	GASE	PATÁPIO	UNIMALI	UNILUS
n	120	165	185	194	189
p	0,55	0,48	0,32	0,49	0,25
N (população)	890	900	1500	1200	1800

Usando 1% de significância responda os itens a seguir:

a) Estime a proporção populacional de satisfeitos com o seu curso, em cada universidade.

R. SHUFSC 44,11% a 65,89%, GASE 38,94% a 57,06%, PATÁPIO 23,73% a 40,27%, UNIMALI 40,53% a 57,47%, UNILUS 17,32% a 32,68%.

b) De acordo com os resultados do item a, a suposição do CRA é confirmada? JUSTIFIQUE.

c) Para uma margem de erro de 2% qual deveria ser o tamanho de amostra para estimar a proporção de satisfeitos em cada universidade? R. SHUFSC 732, GASE 740, PATÁPIO 1060, UNIMALI 931, UNILUS 1141.

A variável sob análise (satisfação) é **QUALITATIVA**, e pode assumir apenas dois valores: *satisfeito ou insatisfeito*. Portanto, serão feitas inferências sobre a proporção de acadêmicos de administração satisfeitos ou insatisfeitos.

a) O parâmetro de interesse é a proporção populacional π de acadêmicos satisfeitos nos cursos de administração da SHUFSC, GASE, PATÁPIO de SÁ, UNIMALI e UNILUS (calcularemos então 5 intervalos de 99% de confiança).

O problema exige uma confiança de 99%, então $1 - \alpha = 0,99$ $\alpha = 0,01$ $\alpha/2 = 0,005$

As estatísticas são: **proporção amostral de acadêmicos satisfeitos** p , o seu complementar $1 - p$, o tamanho de amostra n , e o tamanho da população N , para cada universidade:

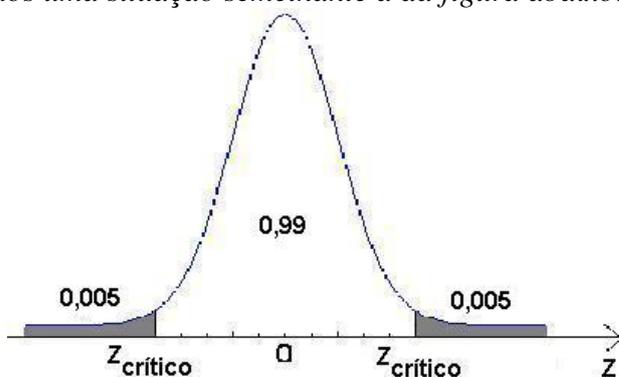
Medidas	UNIVERSIDADES				
	SHUFSC	GASE	PATÁPIO	UNIMALI	UNILUS
n	120	165	185	194	189
p	0,55	0,48	0,32	0,49	0,25
1-p	0,45	0,52	0,68	0,51	0,75
N (população)	890	900	1500	1200	1800

Definição da variável de teste: precisamos verificar se é possível fazer a aproximação pela normal, então multiplicamos $n \times p$ e $n \times (1-p)$ em cada universidade:

Medidas	UNIVERSIDADES				
	SHUFSC	GASE	PATÁPIO	UNIMALI	UNILUS
$n \times p$	66	79,2	59,2	95,06	47,25
$n \times (1 - p)$	54	85,8	125,8	98,94	141,75

Como todos os produtos satisfazem as condições para a aproximação podemos usar a variável Z da distribuição normal padrão em todos os intervalos de confiança.

Encontrar o valor de $Z_{\text{crítico}}$: como o Intervalo de Confiança para a proporção é bilateral, teremos uma situação semelhante à da figura abaixo:



Para encontrar o valor crítico devemos procurar na tabela da distribuição normal padrão pela probabilidade 0,005 e 0,995 (0,99+ 0,005) O valor da probabilidade pode ser visto na figura ao lado: os valores críticos serão $Z_{0,005}$ e $Z_{0,905}$ os quais serão iguais em módulo. $P(Z > Z_{\text{crítico}}) = 0,005$. Então $Z_{\text{crítico}}$ será igual a 2,575 (em módulo).

Ou, usar a função do Excel/Calc: $\text{INV.NORMP.N}(0,995)$, usa-se 0,995 porque é a probabilidade acumulada até o $Z_{\text{crítico}}$ superior (0,005 + 0,99).

Passa-se agora a determinação dos limites dos intervalos, através das expressões abaixo (cujos resultados serão somados e subtraídos das proporções amostrais de acadêmicos satisfeitos, em cada curso) para determinar os limites do intervalo:

SHUFSC	$e_0 = Z_{\text{crítico}} \times \sqrt{\frac{p \times (1 - p)}{n}} = 2,575 \times \sqrt{\frac{0,55 \times 0,45}{120}} = 0,116981$
GASE	$e_0 = Z_{\text{crítico}} \times \sqrt{\frac{p \times (1 - p)}{n}} = 2,575 \times \sqrt{\frac{0,48 \times 0,52}{165}} = 0,100184$
PATÁPIO	$e_0 = Z_{\text{crítico}} \times \sqrt{\frac{p \times (1 - p)}{n}} = 2,575 \times \sqrt{\frac{0,32 \times 0,68}{185}} = 0,088341$

UNIMALI	$e_0 = Z_{\text{critico}} \times \sqrt{\frac{p \times (1-p)}{n}} = 2,575 \times \sqrt{\frac{0,49 \times 0,51}{194}} = 0,092448$
UNILUS	$e_0 = Z_{\text{critico}} \times \sqrt{\frac{p \times (1-p)}{n}} = 2,575 \times \sqrt{\frac{0,25 \times 0,75}{189}} = 0,081131$

Como o tamanho da população N é conhecido é preciso **corrigir** o valor do e_0 , para calcular os limites dos intervalos de confiança:

SHUFSC	$e_{0c} = e_0 \times \sqrt{\frac{N-n}{N-1}} = 0,116981 \times \sqrt{\frac{890-120}{890-1}} = 0,108870$
GASE	$e_{0c} = e_0 \times \sqrt{\frac{N-n}{N-1}} = 0,100184 \times \sqrt{\frac{900-165}{900-1}} = 0,090586$
PATÁPIO	$e_{0c} = e_0 \times \sqrt{\frac{N-n}{N-1}} = 0,088341 \times \sqrt{\frac{1500-185}{1500-1}} = 0,082741$
UNIMALI	$e_{0c} = e_0 \times \sqrt{\frac{N-n}{N-1}} = 0,092448 \times \sqrt{\frac{1200-194}{1200-1}} = 0,084682$
UNILUS	$e_{0c} = e_0 \times \sqrt{\frac{N-n}{N-1}} = 0,081131 \times \sqrt{\frac{1800-189}{1800-1}} = 0,076775$

Os limites então serão:

SHUFSC	$L_I = p - e_0 = 0,45 - 0,108870 = 0,441130$	$L_S = p + e_0 = 0,45 + 0,108870 = 0,658870$
GASE	$L_I = p - e_0 = 0,48 - 0,090586 = 0,389414$	$L_S = p + e_0 = 0,48 + 0,090586 = 0,570586$
PATÁPIO	$L_I = p - e_0 = 0,32 - 0,082741 = 0,237259$	$L_S = p + e_0 = 0,32 + 0,082741 = 0,402741$
UNIMALI	$L_I = p - e_0 = 0,49 - 0,084682 = 0,405318$	$L_S = p + e_0 = 0,49 + 0,084682 = 0,574682$
UNILUS	$L_I = p - e_0 = 0,25 - 0,076775 = 0,173225$	$L_S = p + e_0 = 0,25 + 0,076775 = 0,326775$

Então, os intervalos de 99% de confiança para a proporção populacional de satisfeitos com os cursos são:

SHUFSC	[44,11%; 65,89%]
GASE	[38,94%; 57,06%]
PATÁPIO	[23,73%; 40,27%]
UNIMALI	[40,53%; 57,47%]
UNILUS	[17,32%; 32,68%]

Interpretação: há 99% de probabilidade de que as verdadeiras proporções populacionais de satisfeitos com os cursos de administração estejam de: 44,11% a 65,89% na SHUFSC, 38,94% a 57,06% na GASE, 23,73% a 40,27% na PATÁPIO, 40,53% a 57,47% na UNIMALI e 17,32% a 32,68% na UNILUS.

b) A suposição do CRA é confirmada em parte: realmente os alunos da SHUFSC apresentam maior satisfação: os limites inferior e superior do intervalo de confiança da proporção de satisfeitos na SHUFSC (44,11% e 65,89%) são os maiores entre todas as universidades. Mas os alunos da UNIMALI (privada) tem proporção de satisfeitos muito semelhante aos da GASE: limites inferiores de 40,53% versus 38,94% e limites superiores de 57,47% versus 57,06%, demonstrando grande sobreposição. Os alunos de PATÁPIO e UNILUS têm menor satisfação, abaixo dos limites inferiores de SHUFSC e UNIMALI.

c) De acordo com o item a é possível utilizar a aproximação pela distribuição normal. Assim, a expressão para o cálculo do tamanho mínimo de amostra para a proporção populacional será:

$$n = \left(\frac{Z_{\text{critico}}}{e_0} \right)^2 \times p \times (1-p)$$

Precisamos realizar o procedimento para cada uma das proporções. Além disso, como o tamanho da população é conhecido o tamanho de amostra calculado pela equação acima deve ser corrigido.

O nível de confiança exigido é de 99%: para encontrar o valor crítico devemos procurar na tabela da distribuição normal padrão pela probabilidade 0,005 e 0,995 (0,99+0,005); os valores críticos serão $Z_{0,005}$ e $Z_{0,995}$ os quais serão iguais em módulo. $P(Z > Z_{\text{crítico}}) = 0,005$. Então $Z_{\text{crítico}}$ será igual a 2,575 (em módulo). A precisão foi fixada em 2% (0,02).

Os valores de p e $1 - p$ já são conhecidos para cada universidade. Substituindo todos os valores na expressão do tamanho de amostra vamos obter os resultados da tabela abaixo:

SHUFSC	$n_0 = \left(\frac{Z_{\text{crítico}}}{e_0} \right)^2 \times p \times (1-p) = \left(\frac{2,575}{0,02} \right)^2 \times 0,55 \times 0,45 = 4105,342$
GASE	$n_0 = \left(\frac{Z_{\text{crítico}}}{e_0} \right)^2 \times p \times (1-p) = \left(\frac{2,575}{0,02} \right)^2 \times 0,48 \times 0,52 = 4140,175$
PATÁPIO	$n_0 = \left(\frac{Z_{\text{crítico}}}{e_0} \right)^2 \times p \times (1-p) = \left(\frac{2,575}{0,02} \right)^2 \times 0,32 \times 0,68 = 3609,384$
UNIMALI	$n_0 = \left(\frac{Z_{\text{crítico}}}{e_0} \right)^2 \times p \times (1-p) = \left(\frac{2,575}{0,02} \right)^2 \times 0,49 \times 0,51 = 4145,152$
UNILUS	$n_0 = \left(\frac{Z_{\text{crítico}}}{e_0} \right)^2 \times p \times (1-p) = \left(\frac{2,575}{0,02} \right)^2 \times 0,25 \times 0,75 = 3110,108$

É óbvio que os tamanhos de amostra precisam ser corrigidos, uma vez que todos são maiores que a própria população.

SHUFSC	$n = \frac{N \times n_0}{N + n_0} = \frac{890 \times 4105,342}{890 + 4105,342} \cong 732$
GASE	$n = \frac{N \times n_0}{N + n_0} = \frac{900 \times 4140,175}{900 + 4140,175} \cong 740$
PATÁPIO	$n = \frac{N \times n_0}{N + n_0} = \frac{1500 \times 3609,384}{1500 + 3609,384} \cong 1060$
UNIMALI	$n = \frac{N \times n_0}{N + n_0} = \frac{1200 \times 4145,152}{1200 + 4145,152} \cong 931$
UNILUS	$n = \frac{N \times n_0}{N + n_0} = \frac{1800 \times 3110,108}{1800 + 3110,108} \cong 1141$

Observe que os tamanhos mínimos de amostra necessários para atender a 99% de confiança e precisão de 2% são todos superiores aos das amostras coletadas, indicando que elas são INSUFICIENTES para os fins propostos.

UNIDADE 6

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%?

Observe que é preciso tomar uma decisão: com base nos dados da amostra a afirmação do dono da agência é verdadeira ou ele deve contratar mais um atendente? Trata-se então de um **teste de hipóteses**. A amostra foi coletada para avaliar se o tempo médio de atendimento de 3 minutos (180 segundos) ainda é válido: não haverá problema algum se o tempo for igual ou menor do que 180

segundos, mas se for **maior**, o dono da agência precisaria contratar um novo atendente. Então faremos um teste **unilateral à direita**.

Enunciar as hipóteses.

Conforme visto acima o teste mais adequado para este caso é um Teste Unilateral à Direita:

$$H_0 : \mu = 180 \text{ onde } \mu_0 = 180 \text{ segundos (valor de teste)}$$

$$H_1 : \mu > 180$$

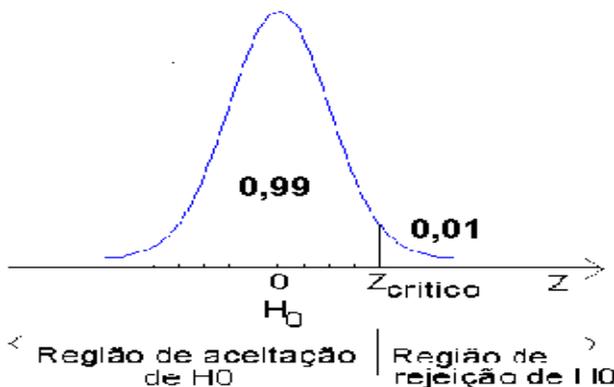
Nível de significância.

O problema declara que é necessário usar 1%. Então $\alpha = 0,01$ e $1 - \alpha = 0,99$

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 40 elementos (portanto mais de 30) a variável de teste será **Z** da distribuição normal.

Definir a região de aceitação de H_0 .



Por ser um teste Unilateral à Direita o Nível de Significância α 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. Repare que o **Z crítico** aqui é maior do que zero:

$$P(Z > \mathbf{Z}_{\text{crítico}}) = 0,01.$$

$$\text{Então } \mathbf{Z}_{\text{crítico}} \cong 2,33$$

Ou, usar a função do Excel/Calc: $\text{INV.NORMP.N}(0,99)$.

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 μ_0 é igual a 180, a média amostral \bar{X} vale 195, o tamanho de amostra n é igual a 40 e o desvio padrão amostral s é 15. Substituindo na equação acima:

$$Z = \frac{\bar{x} - \mu_0}{s / \sqrt{n}} = \frac{195 - 180}{15 / \sqrt{40}} = 6,32$$

Decidir pela aceitação ou rejeição de H_0 .

Como se trata de um teste Unilateral à Direita:

Abordagem clássica

Rejeitar H_0 se $Z > Z_{\text{crítico}}$ Como $Z = 6,32 > Z_{\text{crítico}} = 2,33$

REJEITAR H_0 a 1% de Significância (há 1% de chance de erro)

Abordagem do valor-p

Rejeitar H_0 se $P(Z > 6,32) < \alpha$

Procurando na tabela da distribuição normal padrão observa-se que o maior valor de Z é 5, e que $P(Z > 5) = 0,000000287$, que é praticamente igual a zero. Como $6,32 > 5$, $P(Z > 6,32)$ deve ser ainda menor.

Através da função do Excel/Calc: $1 - \text{DIST.NORMP.N}(6,32; \text{VERDADEIRO})$, usa-se $1 - \text{DIST.NORMP.N}$ por que as planilhas eletrônicas calculam os valores acumulados até um valor, e neste caso queremos a probabilidade de a normal padrão assumir valores ACIMA de 6,32 (porque é teste unilateral à direita). Para o presente caso $1 - \text{DIST.NORMP.N}(6,32; \text{VERDADEIRO}) = 0,00000000013 \cong 0$

Como $P(Z > 6,32) \cong 0 < 0,01$

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 o tempo médio de atendimento é maior do que 180 segundos. A afirmação do dono da agência não é verdadeira, um novo atendente deveria ser contratado.

2) O tempo de montagem de determinados conectores utiliza um processo já há algum tempo, que dura em média 3,5 segundos. Está sendo analisada a possibilidade de troca deste processo para um outro que se afirma possuir um tempo de montagem menor. Para esta análise foram observados os tempos de montagem de conectores por um operário padrão utilizando o novo processo e foram anotados os seguintes valores (em segundos):

2,5 2,5 2,6 3,0 3,2 3,5 3,7 3,7 2,1 2,4 2,7 2,8 3,1 3,1 3,6 3,6 2,5 2,9 2,8 3,8

Aplicando o teste estatístico apropriado, considerando a situação acima, com um nível de confiança de 95%, a empresa deve mudar para o novo processo ou manter o atual? R. Deve mudar. $t = -4.36$

Observe que é preciso tomar uma decisão: com base nos dados da amostra deve-se mudar para o novo processo? Trata-se então de um **teste de hipóteses**. A amostra foi coletada para avaliar se o tempo médio de montagem do novo processo é de 3,5 segundos: se o tempo for igual ou maior não há razão para mudar, mas se for **menor**, a mudança será interessante pois haverá um ganho de produtividade. Então faremos um teste **unilateral à esquerda**.

Enunciar as hipóteses.

Conforme visto acima o teste mais adequado para este caso é um Teste Unilateral à esquerda:

$H_0 : \mu = 3,5$ onde $\mu_0 = 3,5$ segundos (valor de teste)

$H_1 : \mu < 3,5$

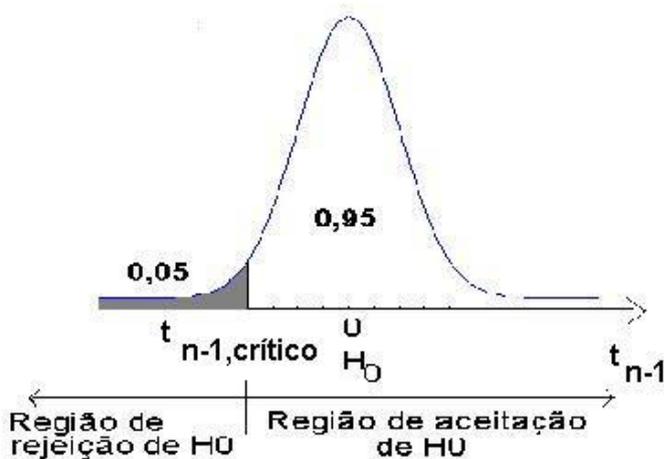
Nível de significância.

O problema pede para usar 5%, então $\alpha = 0,05$ e $1 - \alpha = 0,95$

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 20 elementos (portanto menos de 30) a variável de teste será t_{n-1} da distribuição **t** de Student.

Definir a região de aceitação de H_0 .



Observe que por ser um teste Unilateral à Esquerda o Nível de Significância α 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 de Student, na linha correspondente a **n-1** graus de liberdade, ou seja em $20 - 1 = 19$ graus de liberdade. O valor da probabilidade pode ser visto na figura ao lado: $P(t > t_{n-1, crítico}) = 0,95$. Deve-se procurar a probabilidade complementar 0,05 e mudar o sinal do valor encontrado, pois o $t_{n-1, crítico}$ aqui é menor do que zero.

O valor crítico será igual a $-1,729$.

Ou, usar a função do Excel/Calc: $INV.T(0,05;19)$.

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 μ_0 é igual a 3,5, a média amostral \bar{X} vale 3,005, o tamanho de amostra n é igual a 20 e o desvio padrão amostral s é 0,5083. Substituindo na equação acima:

$$t_{n-1} = \frac{\bar{x} - \mu_0}{s / \sqrt{n}} = \frac{3,005 - 3,5}{0,5083 / \sqrt{20}} = -4,36$$

Decidir pela aceitação ou rejeição de H_0 .

Como se trata de um teste Unilateral à Esquerda:

Abordagem clássica

Rejeitar H_0 se $t_{n-1} < t_{n-1crítico}$

Como $t_{n-1} = -4,36 < t_{n-1crítico} = -1,729$

REJEITAR H_0 a 5% de Significância (há 5% de chance de erro)

Abordagem do valor-p

Rejeitar H_0 se $P(t_{19} < -4,36) < \alpha$

Procurando na tabela da distribuição t de Student para 19 graus de liberdade observa-se que o maior valor de t é 3,883, e que $P(t_{19} > 3,883) = 0,0005$.

Devido à simetria da t de Student em relação à média:

$P(t_{19} > 3,883) = P(t_{19} < -3,883) = 0,0005$.

Como $-4,36 < -3,883$, $P(t_{19} < -4,36)$ deve ser ainda menor.

Através da função do Excel/Calc: $DIST.T(-4,36;19;VERDADEIRO)$.

Para o presente caso $DIST.T(-4,36;19;VERDADEIRO) = 0,000168518 \cong 0$

Como $P(t_{19} < -4,36) \cong 0 < 0,05$

REJEITAR H_0 a 5% de Significância (há 5% de chance de erro)

Interpretar a decisão no contexto do problema.

Há provas estatísticas suficientes de que o tempo médio de montagem dos conectores pelo novo processo é menor do que o atual processo. A empresa deve mudar para o novo processo, pois, terá ganho de produtividade.

3) 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?

Observe que é preciso tomar uma decisão: com base nos dados da amostra deve-se confiar na declaração da empresa sobre o tempo de atraso? Trata-se então de um **teste de hipóteses**. A amostra foi coletada para avaliar se o tempo médio de atraso na entrega dos pedidos é maior do que 1 dia: se o tempo for igual ou menor não há razão para o cliente reclamar, mas se for **maior**, a reclamação tem fundamento. Então faremos um teste **unilateral à direita**.

Enunciar as hipóteses.

Conforme visto acima o teste mais adequado para este caso é um Teste Unilateral à direita:

H_0 : $\mu = 1$ onde $\mu_0 = 1$ dia (valor de teste)

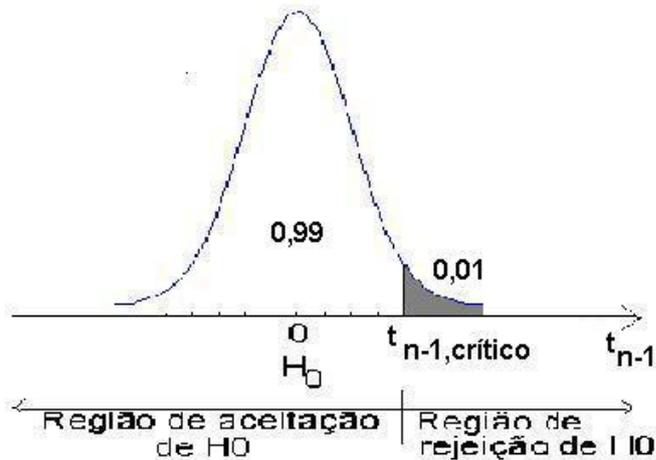
H_1 : $\mu > 1$

Nível de significância.

O problema declara que é necessário usar uma confiança de 99%, então $1 - \alpha = 0,99$ e $\alpha = 0,01$
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 20 elementos (portanto menos de 30) a variável de teste será t_{n-1} da distribuição **t** de Student.

Definir a região de aceitação de H_0 .



Observe que por ser um teste Unilateral à Esquerda o Nível de Significância α 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 de Student, na linha correspondente a $n-1$ graus de liberdade, ou seja em $20 - 1 = 19$ graus de liberdade. O valor da probabilidade pode ser visto na figura ao lado: $P(t > t_{n-1,critico}) = 0,01$. Então $t_{n-1,critico}$ será igual a 2,539.

Ou, usar a função do Excel/Calc: $INV.T(0,99;19)$.

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 μ_0 é igual a 1, a média amostral \bar{X} vale 3,3, o tamanho de amostra n é igual a 20 e o desvio padrão amostral s é 3,0105. Substituindo na equação acima:

$$t_{n-1} = \frac{\bar{x} - \mu_0}{s / \sqrt{n}} = \frac{3,3 - 1}{3,0105 / \sqrt{20}} = 3,42$$

Decidir pela aceitação ou rejeição de H_0 .

Como se trata de um teste Unilateral à Direita:

Abordagem clássica

Rejeitar H_0 se $t_{n-1} > t_{n-1,critico}$

Como $t_{n-1} = 3,42 > t_{n-1,critico} = 2,539$

Abordagem do valor-p

Rejeitar H_0 se $P(t_{19} > 3,42) < \alpha$

Procurando na tabela da distribuição t de Student para 19 graus de liberdade observa-se que 3,42 está entre 3,174 e 3,579, portanto,
 $0,001 < P(t_{19} > 3,42) < 0,0025$

Através da função do Excel/Calc: $1-DIST.T(3,42;19;VERDADEIRO)$ pode-se obter o valor-p exato, usa-se $1-DIST.T$ por que as planilhas eletrônicas calculam os valores acumulados até um valor, e neste caso queremos a probabilidade de a t de Student com 19 graus de liberdade assumir valores ACIMA de 3,42 (porque é teste unilateral à direita).

Para o presente caso $1-DIST.T(3,42;19;VERDADEIRO) = 0,001435768$

Como $P(t_{19} > 3,42) \cong 0,001436 < 0,01$

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 o tempo médio de atraso na entrega dos pedidos é maior do que 1 dia. O cliente tem razão na sua reclamação.

4) A satisfação da população em relação a determinado governo foi pesquisada através de uma amostra com a opinião de 1000 habitantes do estado. Destes, 585 se declararam insatisfeitas com a administração estadual. Admitindo-se um nível de significância de 5%, solucione os itens abaixo.

A atual administração decidiu que se o percentual de descontentamento fosse superior a 50% deveria ser redirecionado o plano governamental. Aplicando o teste estatístico apropriado, utilizando a informação amostral, o que você conclui? R. Redirecionar o plano. $Z = 5,375$.

Observe que é preciso tomar uma decisão: com base nos dados da amostra deve ser redirecionado o plano governamental? Trata-se então de um **teste de hipóteses**. A amostra foi coletada para avaliar se a proporção de insatisfeitos com a administração estadual é igual a 50% (0,5): se a proporção for igual ou menor não haveria razão para redirecionar o plano, mas se for **maior** significa que a maioria da população está insatisfeita, e algo precisa ser feito. Então faremos um teste **unilateral à direita**.

Enunciar as hipóteses.

Conforme visto acima o teste mais adequado para este caso é um Teste Unilateral à Direita:

$$H_0 : \pi = 0,5 \text{ (50\%)} \quad \text{onde } \pi_0 = 0,5 \text{ (valor de teste)}$$

$$H_1 : \pi > 0,5 \text{ (50\%)}$$

Nível de significância.

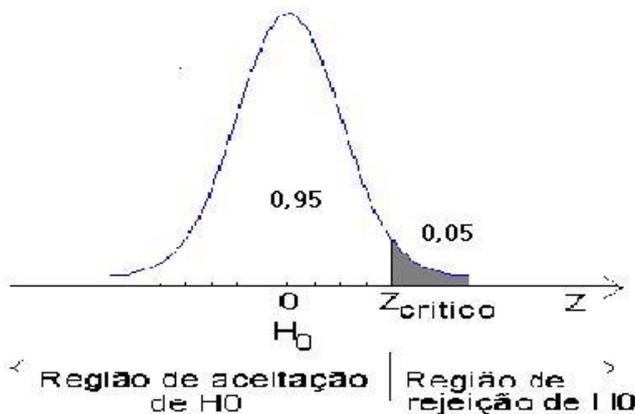
O problema pede para usar uma significância de 5%, então $\alpha = 0,05$ e $1 - \alpha = 0,95$.

Variável de teste.

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

$n \times \pi_0 = 1000 \times 0,5 = 500$ e $n \times (1 - \pi_0) = 1000 \times 0,5 = 500$. 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 **H₀**.



Observe que por ser um teste Unilateral à Direita o Nível de Significância α está todo concentrado em um dos lados da distribuição, definindo a região de rejeição de **H₀**. Para encontrar o valor crítico devemos procurar na tabela da distribuição normal, pela probabilidade acumulada 0,05 (o **Z_{crítico}** aqui é maior do que zero). $P(Z > Z_{crítico}) = 0,05$. Então **Z_{crítico}** será igual a 1,645.

Ou, usar a função do Excel/Calc: $INV.NORMP.N(0,95)$.

Através dos valores da amostra avaliar o valor da variável.

Neste 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 π_0 é igual a 0,5 (50%), a proporção amostral **p** vale 0,585, e o tamanho de amostra **n** é igual a 1000. Substituindo na equação acima:

$$Z = \frac{p - \pi_0}{\sqrt{\frac{\pi_0 \times (1 - \pi_0)}{n}}} = \frac{0,585 - 0,5}{\sqrt{\frac{0,5 \times 0,5}{1000}}} = 5,375$$

Decidir pela aceitação ou rejeição de **H₀**. Como se trata de um teste Unilateral à direita:

Como se trata de um teste unilateral à direita

Abordagem clássica

Rejeitar H₀ se $Z > Z_{crítico}$

Como $Z = 5,375 > Z_{crítico} = 1,645$

Abordagem do valor-p

Rejeitar H₀ se $P(Z > 5,375) < \alpha$

Procurando na tabela da distribuição normal padrão observa-se que o maior valor de Z é 5, e que $P(Z > 5) = 0,000000287$, que é praticamente igual a zero. Como $5,375 > 5$, $P(Z > 6,32)$ deve ser ainda menor.

Através da função do Excel/Calc:

1-DIST.NORMP.N(5,375;VERDADEIRO), usa-se 1-DIST.NORMP.N por que as planilhas eletrônicas calculam os valores acumulados até um valor, e neste caso queremos a probabilidade de a normal padrão assumir valores ACIMA de 5,375 (porque é teste unilateral à direita). Para o presente caso $1-DIST.NORMP.N(5,375;VERDADEIRO) = 0,00000003829 \cong 0$

Como $P(Z > 5,375) \cong 0 < 0,05$

REJEITAR H_0 a 5% de Significância (há 5% de chance de erro)

Interpretar a decisão no contexto do problema.

Há provas estatísticas suficientes para recomendar o redirecionamento do plano governamental, mais de 50% das pessoas estão insatisfeitas com a administração estadual.

5) Uma firma está convertendo as máquinas que aluga para uma versão mais moderna. Até agora foram convertidas 40 máquinas. O tempo médio de conversão foi de 24 horas, com desvio padrão de 3 horas. O fabricante das novas máquinas afirma que a conversão em média dura no máximo 25 horas. Aplicando o teste estatístico apropriado, com base nas conversões feitas até o momento, a 1% de significância, a afirmação do fabricante é verdadeira? R. Sim. $Z = -2,1082$

Adaptado de STEVENSON, W.J. Estatística Aplicada à Administração, São Paulo: Harper do Brasil, 2001.

Observe que é preciso tomar uma decisão: com base nos dados da amostra a afirmação do fabricante das máquinas é verdadeira? Trata-se então de um teste de hipóteses. A amostra foi coletada para avaliar se o tempo médio de conversão de 25 horas é válido: não haverá problema algum se o tempo for igual ou menor do que 25 horas, mas se for maior, a afirmação do fabricante não é correta. Então faremos um teste unilateral à direita.

Enunciar as hipóteses.

Conforme visto o teste mais adequado para este caso é um Teste Unilateral à Direita:

$H_0 : \mu = 25$ onde $\mu_0 = 25$ horas (valor de teste)

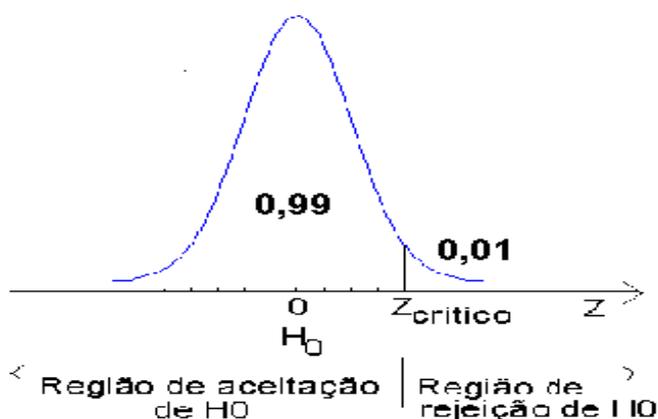
$H_1 : \mu > 25$

Nível de significância. É necessário usar 1%. Então $\alpha = 0,01$ e $1 - \alpha = 0,99$

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 40 elementos (portanto mais de 30) a variável de teste será Z da distribuição normal.

Definir a região de aceitação de H_0 .



*Observe que por ser um teste Unilateral à Direita o Nível de Significância α 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. Repare que o **Zcrítico** aqui é maior do que zero:*

$P(Z > Z_{\text{crítico}}) = 0,01.$

Então $Z_{\text{crítico}} \cong 2,33$

Ou, usar a função do Excel/Calc: INV.NORMP.N(0,99).

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 μ_0 é igual a 25, a média amostral \bar{X} vale 24, o tamanho de amostra n é igual a 40 e o desvio padrão amostral s é 3. Substituindo na equação acima:

$$Z = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{24 - 25}{3/\sqrt{40}} = -2,1082 \cong -2,11$$

Decidir pela aceitação ou rejeição de H_0 .

Como se trata de um teste unilateral à direita

Abordagem clássica

Rejeitar H_0 se $Z > Z_{\text{crítico}}$ Como $Z = -2,11 < Z_{\text{crítico}} = 2,33$

Abordagem do valor-p

Rejeitar H_0 se $P(Z > -2,11) < \alpha$

Procurando na tabela da distribuição normal padrão observa-se que $P(Z > 2,11) = 0,0174$, mas procura-se $P(Z > -2,11)$, que será igual a $1 - P(Z > 2,11)$ devido à simetria da normal padrão, então $P(Z > -2,11) = 1 - 0,0174 = 0,9826$.

Através da função do Excel/Calc:

1-DIST.NORMP.N(-2,11;VERDADEIRO), usa-se 1-DIST.NORMP.N porque as planilhas eletrônicas calculam os valores acumulados até um valor, e neste caso queremos a probabilidade de a normal padrão assumir valores ACIMA de -2,11 (porque é teste unilateral à direita). Para o presente caso

1-DIST.NORMP.N(-2,11;VERDADEIRO) = 0,9826

Como $P(Z > -2,11) \cong 0,9826 > 0,01$

ACEITAR H_0 a 1% de Significância (há 1% de chance de erro)

Interpretar a decisão no contexto do problema.

NÃO há provas estatísticas suficientes de que o tempo médio de conversão das máquinas é maior do que 25 horas. A afirmação do fabricante é verdadeira.

6) Em uma pesquisa de mercado, acerca da preferência pelo produto X, 300 consumidores foram entrevistados, sendo que 100 declararam consumir o produto.

O fabricante do produto X afirma que é a marca líder no mercado, que mais de 40% dos consumidores a preferem. Aplique o teste estatístico apropriado e com base nos dados verifique se a afirmação é válida. Usar 1% de significância. R. Não. $Z = -2,35$

Adaptado de BUSSAB, W.O., MORETTIN, P. A. Estatística Básica, 8a ed. São Paulo: Saraiva, 2013.

Observe que é preciso tomar uma decisão: com base nos dados da amostra marca ainda é líder de mercado? Trata-se então de um teste de hipóteses. A amostra foi coletada para avaliar se a proporção de pessoas que consomem o produto é igual a 40% (0,4): se a proporção for igual ou maior não haverá problemas, mas se for menor a marca não tem mais a liderança do mercado e algo precisa ser feito. Então faremos um teste unilateral à esquerda.

Enunciar as hipóteses.

Conforme visto o teste mais adequado para este caso é um Teste Unilateral à Esquerda:

H_0 : $\pi = 0,4$ (40%) onde $\pi_0 = 0,4$ (valor de teste)

H_1 : $\pi < 0,4$ (40%)

Nível de significância.

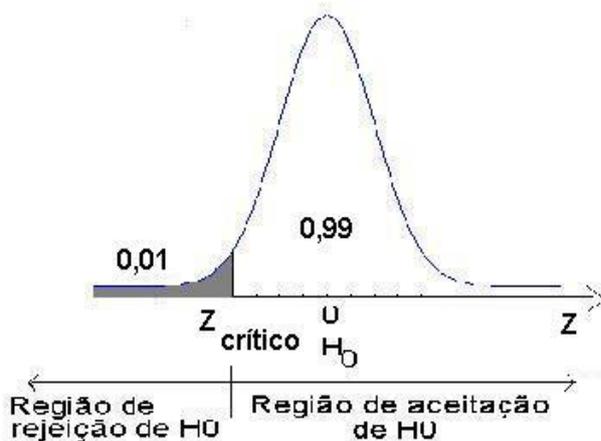
O problema exige uma significância de 1%, então $\alpha = 0,01$ e $1 - \alpha = 0,99$.

Variável de teste.

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

$n \times \pi_0 = 300 \times 0,4 = 120$ e $n \times (1 - \pi_0) = 300 \times 0,6 = 180$. Como ambos os produtos são maiores do que 5, as condições para a aproximação são satisfeitas e podemos usar a variável Z da distribuição normal padrão.

Definir a região de aceitação de H_0 .



Observe que por ser um teste Unilateral à Esquerda o Nível de Significância α 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 0,99. $P(Z > Z_{\text{crítico}}) = 0,99$. Deve-se procurar a probabilidade complementar 0,01 e mudar o sinal do valor encontrado, pois o $Z_{\text{crítico}}$ aqui é menor do que zero. Então $Z_{\text{crítico}}$ será igual a $-2,33$.

Ou, usar a função do Excel/Calc: $INV.NORMP.N(0,01)$.
Através dos valores da amostra avaliar o valor da variável.
Neste 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 π_0 é igual a 0,4 (40%), a proporção amostral p vale 100/300, e o tamanho de amostra n é igual a 300. Substituindo na equação acima:

$$Z = \frac{p - \pi_0}{\sqrt{\frac{\pi_0 \times (1 - \pi_0)}{n}}} = \frac{(100 / 300) - 0,4}{\sqrt{\frac{0,4 \times 0,6}{300}}} = -2,35$$

Decidir pela aceitação ou rejeição de H_0 .
Como se trata de um teste Unilateral à esquerda:
Abordagem clássica

Rejeitar H_0 se $Z < Z_{\text{crítico}}$ Como $Z = -2,35 < Z_{\text{crítico}} = -2,33$

Abordagem do valor-p

Rejeitar H_0 se $P(Z < -2,35) < \alpha$

Procurando na tabela da distribuição normal padrão observa-se que $P(Z > 2,35) = 0,0094$, mas procura-se $P(Z < -2,35)$, mas devido à simetria da normal padrão, então $P(Z > 2,35) = P(Z < -2,35) = 0,0094$.

Através da função do Excel/Calc:

$DIST.NORMP.N(-2,35;VERDADEIRO)$, Para o presente caso

$DIST.NORMP.N(-2,35;VERDADEIRO) = 0,0094$

Como $P(Z < -2,35) \cong 0,0094 < 0,01$

REJEITAR H_0 a 1% de Significância (há 1% de chance de erro)

Interpretar a decisão no contexto do problema. Há provas estatísticas suficientes para considerar que a marca não detém mais a liderança no mercado, que a proporção de pessoas que consomem o produto é menor do que 40%. Contudo, é um caso de fronteira!

7) 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?

Observe que é preciso tomar uma decisão: com base nos dados da amostra é preciso parar a linha de produção? Trata-se então de um teste de hipóteses. A amostra foi coletada para avaliar se a proporção de peças defeituosas é igual a 3% (0,03): se a proporção for igual ou menor não haverá

problemas, mas se for **maior** há problemas de qualidade, e a linha de produção precisa ser parada. Então faremos um teste **unilateral à direita**.

Enunciar as hipóteses.

Conforme visto acima o teste mais adequado para este caso é um Teste Unilateral à Direita: $H_0 : \pi = 0,03$ (3%) onde $\pi_0 = 0,03$ (valor de teste)

$H_1 : \pi > 0,03$ (3%)

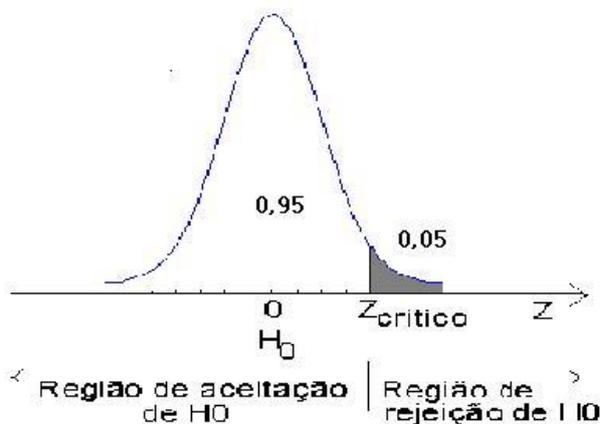
Nível de significância. Não informado pelo problema. Vamos usar uma significância de 5%, então $\alpha = 0,05$ e $1 - \alpha = 0,95$.

Variável de teste.

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

$n \times \pi_0 = 1000 \times 0,03 = 30$ e $n \times (1 - \pi_0) = 1000 \times 0,97 = 970$. Como ambos os produtos são maiores do que 5, as condições para a aproximação são satisfeitas e podemos usar a variável **Z** da distribuição normal padrão.

Definir a região de aceitação de H_0 .



Observe que por ser um teste Unilateral à Esquerda o Nível de Significância α 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 0,05 (o **Z crítico** aqui é maior do que zero). $P(Z > Z_{\text{crítico}}) = 0,05$. Então **Z crítico** será igual a 1,645.

Ou, usar a função do Excel/Calc: $INV.NORMP.N(0,95)$.

Através dos valores da amostra avaliar o valor da variável.

Neste 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 π_0 é igual a 0,03 (3%), a proporção amostral **p** vale 35/1000, e o tamanho de amostra **n** é igual a 1000. Substituindo na equação acima:

$$Z = \frac{p - \pi_0}{\sqrt{\frac{\pi_0 \times (1 - \pi_0)}{n}}} = \frac{(35/1000) - 0,03}{\sqrt{\frac{0,03 \times 0,97}{1000}}} = 0,9268 \cong 0,93$$

Decidir pela aceitação ou rejeição de H_0 .

Como se trata de um teste Unilateral à direita:

Abordagem clássica

Rejeitar H_0 se $Z > Z_{\text{crítico}}$

Como $Z = 0,93 < Z_{\text{crítico}} = 1,645$

Abordagem do valor-p

Rejeitar H_0 se $P(Z > 0,93) < \alpha$

Procurando na tabela da distribuição normal padrão observa-se que $P(Z > 0,93) = 0,1762$.

Através da função do Excel/Calc:

$1-DIST.NORMP.N(0,93;VERDADEIRO)$, usa-se $1-DIST.NORMP.N$ porque as planilhas eletrônicas calculam os valores acumulados até um valor, e neste caso queremos a probabilidade de a normal padrão assumir valores ACIMA de 0,93 (porque é teste unilateral à direita). Para o presente caso

1-DIST.NORMP.N(0,93;VERDADEIRO) = 0,1762

Como $P(Z > 0,9268) \cong 0,1762 > 0,05$

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 recomendar a parada da linha de produção, apenas 3% das peças são defeituosas.

8) Em 600 lançamentos de um dado obteve-se a face 6 em 123 lançamentos.

a) Aplique o teste estatístico apropriado e verifique se a 5% de significância há razão para desconfiar que o dado é viciado quanto a face 6 ? R. Sim. $Z = 2,519$

b) E a 1% de significância? R. Não.

a) *Observe que é preciso tomar uma decisão: com base nos dados da amostra deve-se desconfiar que o dado está viciado? Trata-se então de um teste de hipóteses. A amostra foi coletada para avaliar se a proporção de faces 6 é 1/6: se a proporção for menor ou maior haverá problemas, pois o dado poderá estar viciado. Então faremos um teste bilateral.*

Seguindo o roteiro da apostila:

1) Enunciar as hipóteses.

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

$H_0 : \pi = 1/6$ onde $\pi_0 = 1/6$ (valor de teste)

$H_1 : \pi \neq 1/6$

2) Nível de significância.

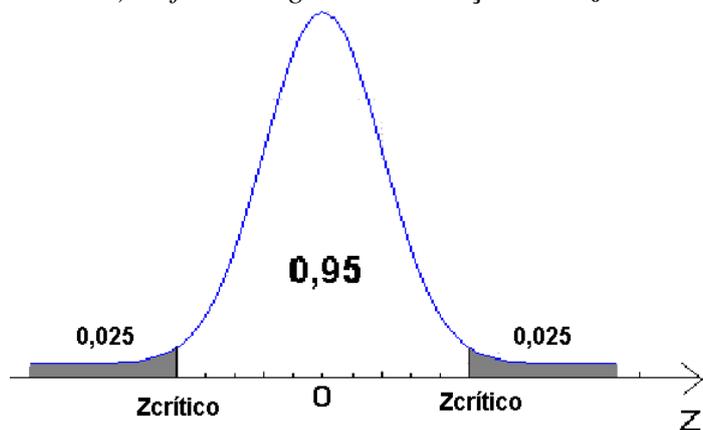
O problema exige uma significância de 5%. Então $\alpha = 0,05$ $\alpha/2 = 0,025$ $1 - \alpha = 0,95$.

3) Variável de teste.

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

$n \times \pi_0 = 600 \times (1/6) = 100$ e $n \times (1 - \pi_0) = 600 \times (5/6) = 500$. Como ambos os produtos são maiores do que 5, portanto satisfazem as condições para a aproximação pela normal, podemos usar a variável **Z** da distribuição normal padrão.

4) Definir a região de aceitação de H_0 .



Observe que por ser um teste Bilateral o Nível de Significância α foi dividido em dois, metade para cada região de rejeição de H_0 . Para encontrar o valor crítico devemos procurar na tabela da distribuição normal padrão pela probabilidade 0,025 e 0,975 (0,95+0,025) O valor da probabilidade pode ser visto na figura ao lado: os valores críticos serão $Z_{0,025}$ e $Z_{0,975}$ os quais serão iguais em módulo. $P(Z > Z_{crítico}) = 0,025$. Então $Z_{crítico}$ será igual a 1,96 (em módulo).

Ou, usar a função do Excel/Calc: $INV.NORMP.N(0,975)$, usa-se 0,975 porque é a probabilidade acumulada até o $Z_{crítico}$ superior (0,025 + 0,95).

Através dos valores da amostra avaliar o valor da variável.

Neste 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 π_0 é igual a 1/6, a proporção amostral **p** vale 123/600, e o tamanho de amostra **n** é igual a 600. Substituindo na equação acima:

$$Z = \frac{p - \pi_0}{\sqrt{\frac{\pi_0 \times (1 - \pi_0)}{n}}} = \frac{(123 / 600) - (1 / 6)}{\sqrt{\frac{(1 / 6) \times (5 / 6)}{600}}} = 2,519 \cong 2,52$$

Decidir pela aceitação ou rejeição de H_0 .

Como se trata de um teste Bilateral:

Abordagem clássica:

Rejeitar H_0 se $|Z| > |Z_{\text{crítico}}|$

Como $|Z| = 2,52 > |Z_{\text{crítico}}| = 1,96$

Abordagem do valor-p

Rejeitar H_0 se $2 \times P(Z > 2,52) < \alpha$

Procurando na tabela da distribuição normal padrão observa-se que $P(Z > 2,52) = 0,0059$.

Através da função do Excel/Calc:

1-DIST.NORMP.N(2,52;VERDADEIRO), usa-se 1-DIST.NORMP.N porque as planilhas eletrônicas calculam os valores acumulados até um valor, e neste caso queremos a probabilidade de a normal padrão assumir valores ACIMA de 2,52, e tal probabilidade será multiplicada por 2 (porque é teste bilateral). Para o presente caso $1-DIST.NORMP.N(2,52;VERDADEIRO) = 0,0059$

Como $2 \times P(Z > 2,519) \cong 0,0118 < 0,05$

REJEITAR H_0 a 5% de Significância (há 5% de chance de erro)

Interpretar a decisão no contexto do problema.

Há provas estatísticas suficientes de que a proporção populacional de faces 6 é diferente de 1/6, o que é razão para desconfiar que o dado é viciado.

b) A única diferença do item anterior é o nível de significância, que passou a ser de 1%.

Observe que é preciso tomar uma decisão: com base nos dados da amostra deve-se desconfiar que o dado está viciado? Trata-se então de um teste de hipóteses. A amostra foi coletada para avaliar se a proporção de faces 6 é 1/6: se a proporção for **menor** ou **maior** haverá problemas, pois o dado poderá estar viciado. Então faremos um teste **bilateral**.

Enunciar as hipóteses.

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

$H_0 : \pi = 1/6$ onde $\pi_0 = 1/6$ (valor de teste)

$H_1 : \pi \neq 1/6$

Nível de significância.

O problema exige uma significância de 1%. Então $\alpha = 0,01$ $\alpha/2 = 0,005$ $1 - \alpha = 0,99$.

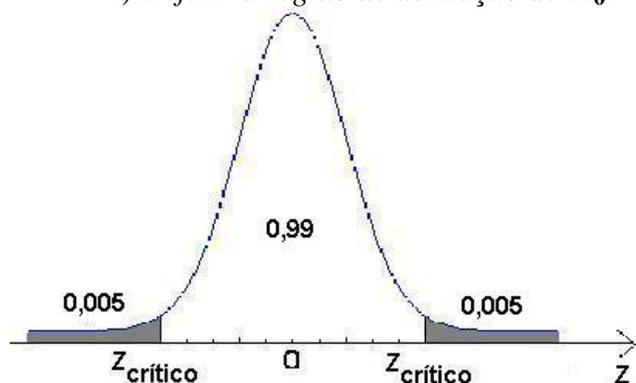
Variável de teste.

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

$n \times \pi_0 = 600 \times (1/6) = 100$ e $n \times (1 - \pi_0) = 600 \times (5/6) = 500$. Como ambos os produtos são maiores do que 5, portanto satisfazem as condições para a aproximação pela normal, podemos usar a variável Z da distribuição normal padrão.

4) Definir a região de aceitação de H_0 .

Observe que por ser um teste Bilateral o Nível de Significância α foi dividido em dois, metade para cada região de rejeição de H_0 . Para encontrar o valor crítico devemos procurar na tabela da distribuição normal padrão pela probabilidade 0,005 e 0,995 (0,99 + 0,005) O valor da probabilidade pode ser visto na figura ao lado: os valores críticos serão $Z_{0,005}$ e $Z_{0,995}$ os quais serão iguais em módulo. $P(Z > Z_{\text{crítico}}) = 0,005$. Então $Z_{\text{crítico}}$ será igual a 2,575 (em módulo).



Ou, usar a função do Excel/Calc: $INV.NORMP.N(0,995)$, usa-se 0,995 porque é a probabilidade acumulada até o Zcrítico superior (0,005 + 0,99).

Através dos valores da amostra avaliar o valor da variável.

Neste 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 π_0 é igual a 1/6, a proporção amostral p vale 123/600, e o tamanho de amostra n é igual a 600. Substituindo na equação acima:

$$Z = \frac{p - \pi_0}{\sqrt{\frac{\pi_0 \times (1 - \pi_0)}{n}}} = \frac{(123 / 600) - (1 / 6)}{\sqrt{\frac{(1 / 6) \times (5 / 6)}{600}}} = 2,519 \cong 2,52$$

Decidir pela aceitação ou rejeição de H_0 .

Como se trata de um teste Bilateral:

Abordagem clássica:

Rejeitar H_0 se $|Z| > |Z_{\text{crítico}}|$ Como $|Z| = 2,519 < |Z_{\text{crítico}}| = 2,575$

Abordagem do valor-p

Rejeitar H_0 se $2 \times P(Z > 2,519) < \alpha$ Procurando na tabela da distribuição normal padrão observa-se que $P(Z > 2,52) = 0,0059$.

Através da função do Excel/Calc:

1-DIST.NORMP.N(2,52;VERDADEIRO), usa-se 1-DIST.NORMP.N porque as planilhas eletrônicas calculam os valores acumulados até um valor, e neste caso queremos a probabilidade de a normal padrão assumir valores ACIMA de 2,52, e tal probabilidade será multiplicada por 2 (porque é teste bilateral). Para o presente caso 1-DIST.NORMP.N(2,52;VERDADEIRO) = 0,0059

Como $2 \times P(Z > 2,519) \cong 0,0118 > 0,01$

ACEITAR H_0 a 1% de Significância (há 1% de chance de erro)

Interpretar a decisão no contexto do problema.

NÃO há provas estatísticas suficientes de que a proporção populacional faces 6 é diferente de 1/6, então ainda não há razão para desconfiar que o dado é viciado.

Observe que a conclusão foi oposta a do item a, devido ao menor nível de significância adotado, que exige evidências estatísticas mais fortes para rejeitar a hipótese nula.

9) Uma amostra aleatória entre homens e mulheres foi analisada com o objetivo de pesquisar-se o comportamento de “fumar cigarros”. Verificou-se que de 27 homens, 15 eram fumantes, e que de 33 mulheres, 12 tinham o hábito de fumar.

Teste a hipótese de que o sexo influencia o comportamento de fumar, a um nível de 5% de significância. R. Não associada. $q^2 = 2,210$

Enunciar as Hipóteses:

H_0 : as variáveis são independentes

H_1 : as variáveis não são independentes

Nível de significância: determinado pelo problema, $\alpha = 0,05$; $1 - \alpha = 0,95$

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

Sexo	Hábitos		Total
	Fumante	Não fumante	
Masculino	15	12	27
Feminino	12	21	33
Total	27	33	60

Fonte: hipotética

Calcular as frequências esperadas: devemos calculá-las para todas as células da tabela (4 no presente problema). Os resultados estão na tabela abaixo:

Eij	Hábitos	
Sexo	Fumante	Não fumante
Masculino	12,15	14,85
Feminino	14,85	18,15

Calculando a estatística q^2 para cada célula:

Agora podemos calcular as diferenças entre as frequências e as demais operações. Os valores finais estão na tabela abaixo:

(O-E) ² /E	Hábitos	
Sexo	Fumante	Não fumante
Masculino	0,6685185	0,54697
Feminino	0,5469697	0,447521

Agora podemos somar os valores:

$$q^2 = 2,209979$$

Os graus de liberdade: (número de linhas - 1) x (número de colunas - 1) = (2 - 1) x (2 - 1) = 1

Então $q^2_1 = 2,209979$

O q^2 crítico será: procurando na tabela da distribuição Chi-Quadrado, ou em um programa, para 1 grau de liberdade e 95% de confiança (5% de significância):

$$q^2_{1,\text{crítico}} = 3,84$$

Ou usando a função do Excel/Calc $INV.QUIQUA(0,95;1)$.

Abordagem clássica

Rejeitar H_0 se $q^2_1 > q^2_{1,\text{crítico}}$

Como $q^2_1 = 2,210 < q^2_{1,\text{crítico}} = 3,84$

Abordagem do valor-p

Rejeitar H_0 se $P(q^2_1 > 2,210) < \alpha$

Procurando na tabela da distribuição qui-quadrado para 1 grau de liberdade observa-se que 2,21 está entre 1,32 e 2,71, portanto, $0,1 < P(q^2_1 > 2,21) < 0,25$

Através da função do Excel/Calc: $1-DIST.QUIQUA(2,21;1;VERDADEIRO)$ pode-se obter o valor-p exato, usa-se $1-DIST.QUIQUA$ por que as planilhas eletrônicas calculam os valores acumulados até um valor, e neste caso queremos a probabilidade de a qui-quadrado com 1 grau de liberdade assumir valores ACIMA de 2,21 (porque é assim no teste de associação de qui-quadrado).

Para o presente caso $1-DIST.QUIQUA(2,21;1;VERDADEIRO) = 0,1371$

Como $P(q^2_1 > 2,210) \cong 0,1371 > 0,05$

ACEITAR H_0 a 1% de Significância (há 1% de chance de erro)

NÃO HÁ evidência estatística suficiente que indica que as variáveis sexo e comportamento de fumar são dependentes.

10) Dentre os alunos de uma sala alguns não frequentavam as aulas, apenas comparecendo às provas. Na tabela abaixo estão apresentados seus resultados:

	Aprovados	Reprovados	Total
“frequentadores”	22	8	30
“ausentes”	10	18	28
Total	32	26	58

Utilizando $1 - \alpha = 99\%$. Aplique o teste apropriado para verificar se a presença nas aulas está associada aos resultados finais dos alunos?

R. Associada. $q^2 = 8,287$

Enunciar as Hipóteses:

H_0 : as variáveis são independentes

H_1 : as variáveis *não* são independentes

Nível de significância: determinado pelo problema, $\alpha = 0,01$; $1 - \alpha = 0,99$

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

	Aprovação		
Frequência	Aprovados	Reprovados	Total
“frequentadores”	22	8	30
“ausentes”	10	18	28
Total	32	26	58

Fonte: hipotética

Calcular as frequências esperadas: devemos calculá-las para todas as células da tabela (4 no presente problema). Os resultados estão na tabela abaixo:

	Aprovação		
Frequência	Aprovados	Reprovados	
“frequentadores”	16,55172414	13,44827586	
“ausentes”	15,44827586	12,55172414	

Calculando a estatística q^2 para cada célula:

Agora podemos calcular as diferenças entre as frequências e as demais operações. Os valores finais estão na tabela abaixo:

(O-E) ² /E	Aprovação		
Frequência	Aprovados	Reprovados	
“frequentadores”	1,793390805	2,207250221	
“ausentes”	1,921490148	2,364910951	

Agora podemos somar os valores:

$$q^2 = 8,2870421$$

Os graus de liberdade: (número de linhas -1)x(número de colunas - 1) = (2 -1)x(2-1)= 1

Então $q^2_1 = 8,2870421$

O q^2 crítico será: procurando na tabela da distribuição Chi-Quadrado, ou em um programa, para 1 grau de liberdade e 99% de confiança (1% de significância): $q^2_{1,\text{crítico}} = 6,63$

Ou usando a função do Excel/Calc INV.QUIQUA(0,99;1).

Abordagem clássica

Rejeitar H_0 se $q^2_1 > q^2_{1,\text{crítico}}$

Como $q^2_1 = 8,29 > q^2_{1,\text{crítico}} = 6,63$

Abordagem do valor-p

Rejeitar H_0 se $P(q^2_1 > 8,29) < \alpha$

Procurando na tabela da distribuição qui-quadrado para 1 grau de liberdade observa-se que 8,29 está entre 7,88 e 9,14, portanto, $0,0025 < P(q^2_1 > 8,29) < 0,005$

Através da função do Excel/Calc: 1-DIST.QUIQUA(8,29;1;VERDADEIRO) pode-se obter o valor-p exato, usa-se 1-DIST.QUIQUA por que as planilhas eletrônicas calculam os valores acumulados até um valor, e neste caso queremos a probabilidade de a qui-quadrado com 1 grau de liberdade assumir valores ACIMA de 8,29 (porque é assim no teste de associação de qui-quadrado).

Para o presente caso 1-DIST.QUIQUA(8,29;1;VERDADEIRO) = 0,004

Como $P(q^2_1 > 8,29) \cong 0,004 < 0,01$

REJEITAR H_0 a 1% de Significância (há 1% de chance de erro)

HÁ evidência estatística suficiente que indica que as variáveis frequência às aulas e aprovação são dependentes.

11) 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, D.M., BERENSON, M.L., STEPHAN, D., Estatística: Teoria e Aplicações usando Microsoft® Excel em Português. Rio de Janeiro: LTC, 2000.

a) Trata-se de um exercício de teste do qui-Quadrado de independência: queremos saber se a variável lembrança do consumidor está associada à variável meio de comunicação (e vice-versa).

Enunciar as Hipóteses:

H_0 : as variáveis são independentes

H_1 : as variáveis *não* são independentes

Nível de significância: determinado pelo problema, $\alpha = 0,01$; $1 - \alpha = 0,99$

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

Lembrança	Meio de comunicação			
	Revista	TV	Rádio	Total
Lembram	25	93	7	125
Não lembram	73	10	108	191
Total	98	103	115	316

Fonte: hipotética

Calcular as frequências esperadas: devemos calculá-las para todas as células da tabela (6 no presente problema). Os resultados estão na tabela abaixo:

Lembrança	Meio de comunicação			
	Revista	TV	Rádio	Total
Lembram	38,76582	40,74367	45,490506	
Não lembram	59,23418	62,25633	69,509494	

Calculando a estatística q^2 para cada célula:

Agora podemos calcular as diferenças entre as frequências e as demais operações. Os valores finais estão na tabela abaixo:

(O-E) ² /E	Meio de comunicação			
	Revista	TV	Rádio	Total
Lembram	4,888272	67,02204	32,56765	
Não lembram	3,199131	43,86259	21,31391	

Agora podemos somar os valores:

$$q^2 = 172,8536$$

Os graus de liberdade: (número de linhas - 1) x (número de colunas - 1) = (2 - 1) x (3 - 1) = 2

Então $q^2_2 = 172,8536$

O q^2 crítico será: procurando na tabela da distribuição Chi-Quadrado, ou em um programa, para 2 graus de liberdade e 99% de confiança (1% de significância): $q^2_{2,\text{crítico}} = 9,21$

Ou usando a função do Excel/Calc INV.QUIQUA(0,99;2).

Abordagem clássica

Rejeitar H_0 se $q^2_2 > q^2_{2,\text{crítico}}$

Como $q^2_2 = 172,8536 > q^2_{2,\text{crítico}} = 9,21$

Abordagem do valor-p

Rejeitar H_0 se $P(q^2_2 > 172,8536) < \alpha$

Procurando na tabela da distribuição qui-quadrado para 2 graus de liberdade observa-se que o maior valor é 13,82, e que $P(q^2_2 > 13,82) = 0,001$, portanto $P(q^2_2 > 172,8536)$ deve ser ainda menor.

Através da função do Excel/Calc: `1-DIST.QUIQUA(172,8536;2;VERDADEIRO)` pode-se obter o valor-p exato, usa-se `1-DIST.QUIQUA` por que as planilhas eletrônicas calculam os valores acumulados até um valor, e neste caso queremos a probabilidade de a qui-quadrado com 2 graus de liberdade assumir valores ACIMA de 172,8536 (porque é assim no teste de associação de qui-quadrado).

Para o presente caso `1-DIST.QUIQUA(172,8536;1;VERDADEIRO) \cong 0`

Como $P(q^2_2 > 172,8536) \cong 0 < 0,01$

REJEITAR H_0 a 1% de Significância (há 1% de chance de erro)

HÁ evidência estatística suficiente que indica que as variáveis meio de comunicação da propaganda e lembrança do consumidor são dependentes.

b) Há associação entre meio de comunicação e lembrança da peça publicitária. Como apresentou as maiores contribuições ao valor do q^2_2 a melhor escolha deve ser a TV.