Teste de Hipóteses

VÍCTOR HUGO LACHOS DÁVILA

Exemplo 1. *Considere que uma industria compra de um certo fabricante, pinos cuja resistência média à ruptura é especificada em 60 kgf (valor nominal da especificação). Em um determinado dia, a indústria recebeu um grande lote de pinos e a equipe técnica da industria deseja verificar se o lote atende as especificações*.

H_0: O lote atende as especificações (Hipóteses nula)
H_1: O lote não atende as especificações (Hipóteses alternativa)

Seja a v.a *X :* resistência à ruptura

$X \sim N(\mu; 25)$

$H_0: \mu = 60$ (Hipóteses simples)
$H_1: \mu \neq 60$ (Hipóteses Composta bilateral)

Definição: Uma hipóteses estatística é uma afirmação ou conjetura sobre o parâmetro, ou parâmetros, da distribuição de probabilidades de uma característica, X, da população ou de uma v.a.

Definição: Um teste de uma hipóteses estatística é o procedimento ou regra de decisão que nos possibilita decidir por H_0 ou H_a, com base a informação contida na amostra.

Suponha que a equipe técnica da indústria tenha decidido retirar uma amostra aleatória de tamanho n=16, do lote recebido, medir a resistência de cada pino e calcular a resistência média $\bar{X}$ (estimador de μ)

$$\bar{X} \sim N\left(\mu, \frac{25}{16}\right)$$

Para quais valores de $\bar{x}$ a equipe técnica deve rejeitar Ho e portanto não aceitar o lote?

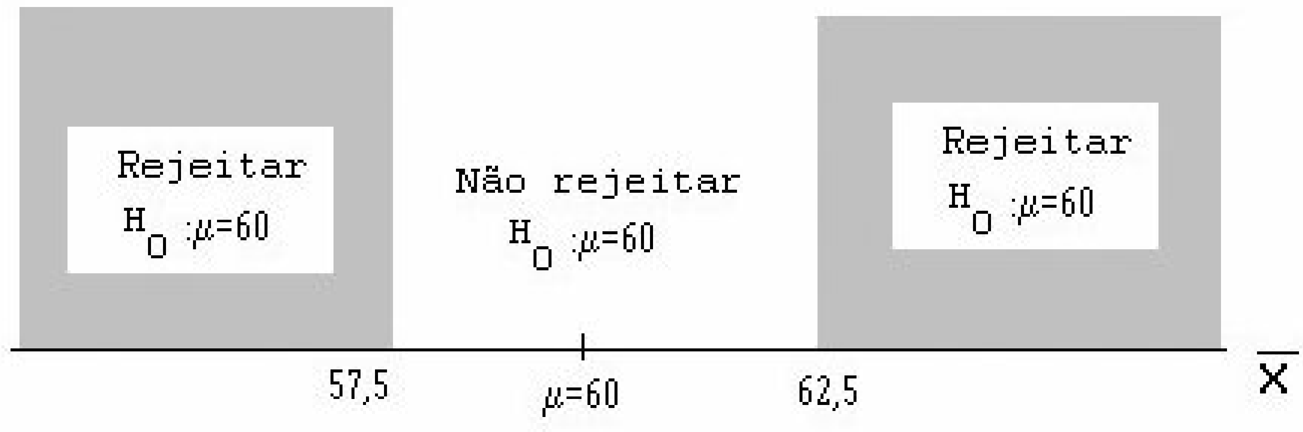
Definição: Região crítica (Rc) é o conjunto de valores assumidos pela variável aleatória ou estatística de teste para os quais a hipótese nula é rejeitada.

Se o lote está fora de especificação , isto é , H1:$\mu \neq 60$, espera-se que a média amostral seja inferior ou superior a 60 kgf

Suponha que equipe técnica tenha decidido adotar a seguinte regra:rejeitar Ho se $\bar{X}$ for maior que 62.5 kgf e ou menor que 57.5 kgf.

$R_c = \{\bar{X} > 62{,}5 \ ou \ \bar{X} < 57{,}5\}$ $\Rightarrow$Região de rejeição de Ho.

$\bar{R}_c = R_a = \{57{,}5 \leq \bar{X} \leq 62{,}4\}$ $\Rightarrow$Região de aceitação de Ho.



Procedimento (teste)

$$Se \;\; \bar{x} \in R_c \Rightarrow \text{Rejeita - se } \mathrm{H}_0$$
$$Se \;\; \bar{x} \notin R_c \Rightarrow Ace\text{ita - se } \mathrm{H}_0$$

Tipos de Erros

Erro tipo I: Rejeitar H_0 quando de fato H_0 é verdadeiro.

Erro tipo II: Não rejeitamos H_0 quando de fato H_0 é falsa.

Exemplo 2: Considere o exemplo 1.

H_0: Aceitar o lote

H_1: Não aceitar o lote

Erro tipo I: Não aceitar o lote sendo que ela está dentro das especificações.

Erro tipo II:Aceitar o lote sendo que ela está fora das especificações.

Decisão	Situação	
	Ho verdadeira	Ho falsa
Não rejeitar Ho	Decisção correta	Erro II
Rejeitar Ho	Erro I	Decisão correta

P(Erro tipo I)=α (nível de significância)

$$\alpha = P(\text{Rejeitar H}_0 \mid H_0 \text{ verdadeira})$$

$$P(Erro\ II) = \beta = P(\text{Não rejeitar H}_0 \mid H_0 \text{ falso}).$$

$$1-\beta = P(\text{Rejeitar} \mid H_0 \text{ é falso}).$$ $\rightarrow$Poder do teste

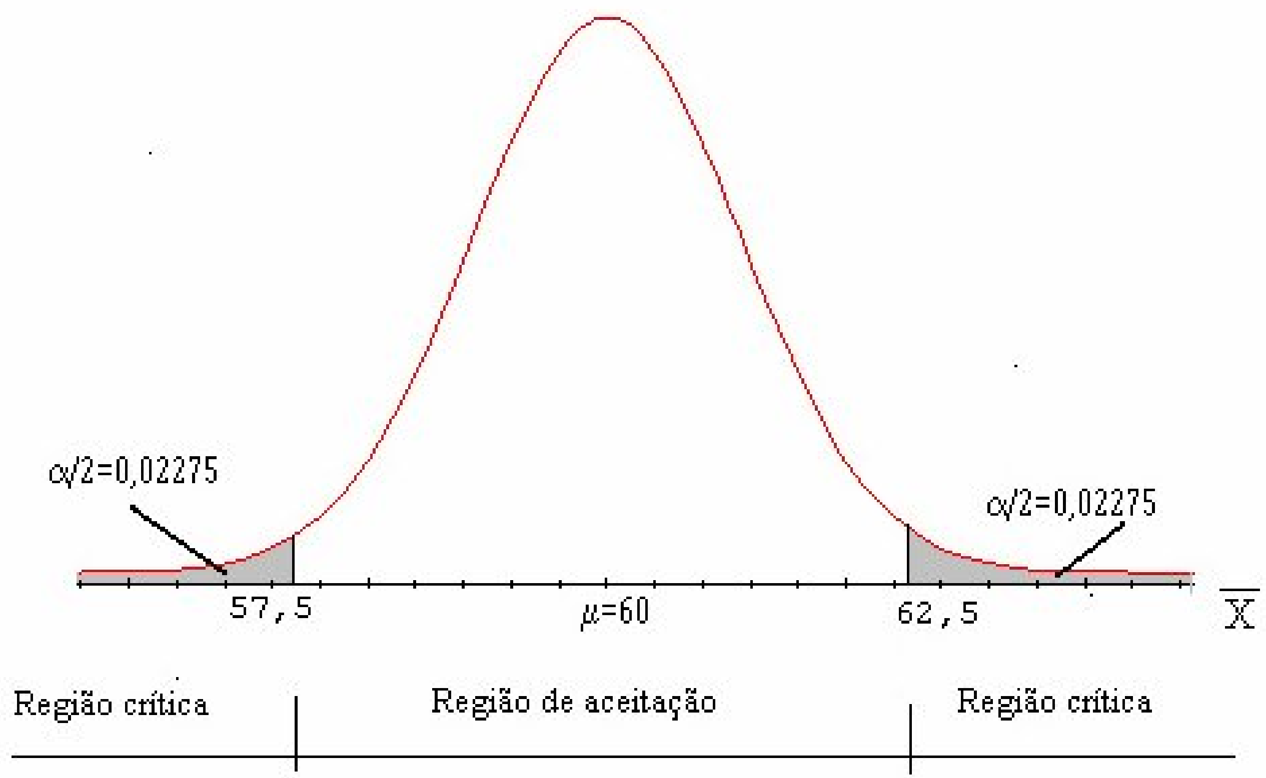
Exemplo 3: Considerando as hipóteses do exemplo 1: H0: μ = 60 contra H1: $\mu \neq$ 60.

$$\alpha = P\left[\bar{X} > 62{,}5 \ ou\ \bar{X} < 57{,}5 \mid H_0 : \mu = 60\right] \quad \Rightarrow Sob\ H_0,\ \bar{X} \sim N(60{,}25/16).$$

$$\alpha = P\left[\bar{X} > 62{,}5 \mid H_0 : \mu = 60\right] + P\left[\bar{X} < 57{,}5 \mid H_0 : \mu = 60\right] =$$

$$P\left[\frac{\bar{X}-60}{\sqrt{25/16}} > \frac{62{,}5-60}{\sqrt{25/16}}\right] + P\left[\frac{\bar{X}-60}{\sqrt{25/16}} < \frac{57{,}5-60}{\sqrt{25/16}}\right] =$$

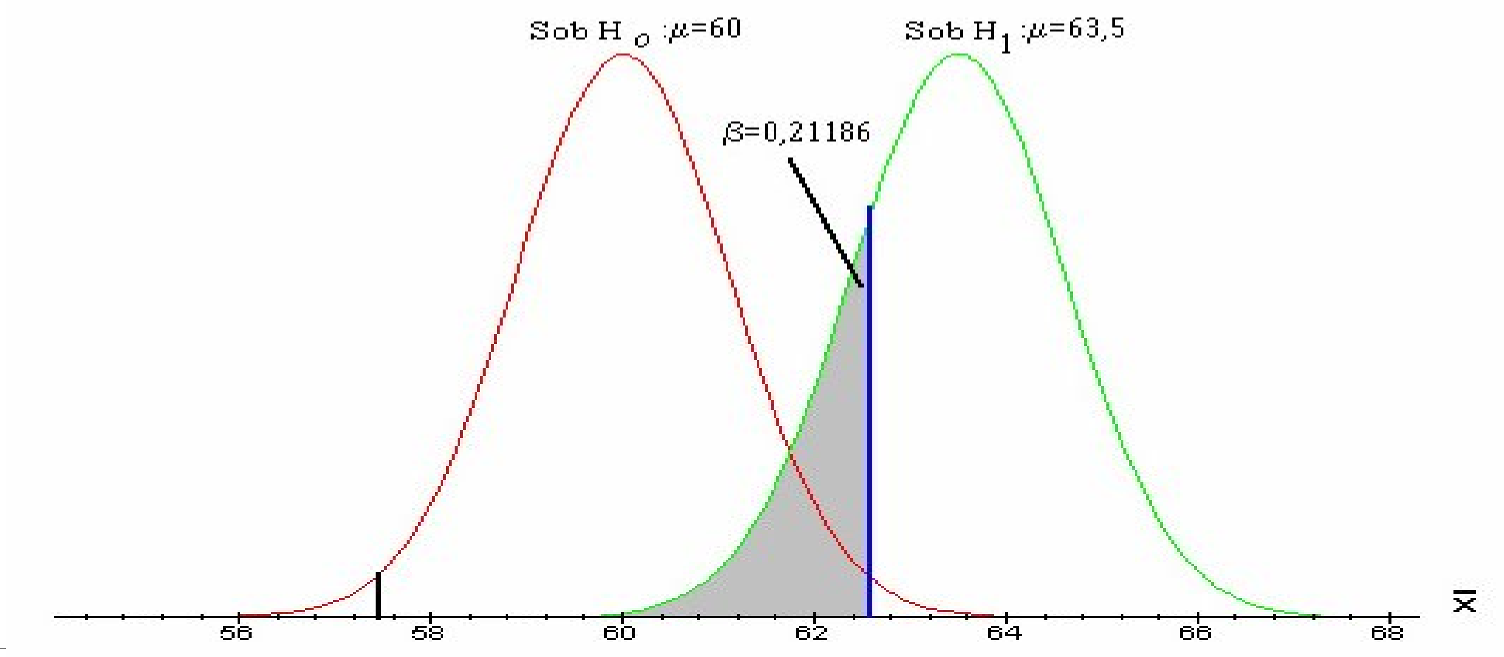
$$P[Z > 2] + P[Z < -2] = 0{,}02275 + 0{,}02275 = 0{,}0445$$


α/2=0,02275
α/2=0,02275
57,5
μ=60
62,5
$\overline{X}$
Região crítica
Região de aceitação
Região crítica

$$\beta = P(\text{Aceitar } \mathrm{H}_0 \mid H_1 \text{ verdadeiro}) = P\left[57{,}5 \le \bar{X} \le 62{,}5 \mid H_1 : \mu \ne 60\right]$$

Para o cálculo de β considerar H_1:μ=63,5. Sob H_1, $\bar{X} \sim N\left(63{,}5; \frac{25}{16}\right)$.

$$\beta = P\left[57{,}5 \le \bar{X} \le 62{,}5 \mid H_1 : \mu = 63{,}5\right] = P\left[\bar{X} \le 62{,}5\right] - P\left[\bar{X} \le 57{,}5\right]$$
$$= P\left[Z \le -0{,}8\right] - P\left[Z \le -4{,}8\right] = 0{,}21186 + 0{,}00 = 0{,}21186.$$



Testes bilaterais e unilaterais

Se a hipótese nula e alternativa de um teste de hipóteses são:

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

onde μ_0 é uma constante conhecida, o teste é chamada de *teste bilateral.*

Em muitos problemas tem-se interesse em testar hipótese do tipo:

$$H_0 : \mu = \mu_0$$
$$H_1 : \mu < \mu_0$$

o teste é chamado de *teste unilateral esquerdo. E quando*

$$H_0 : \mu = \mu_0$$
$$H_1 : \mu > \mu_0$$

o teste é chamada de *teste unilateral direito.*

Exemplo 4: Uma região do país é conhecida por ter uma população obesa. A distribuição de probabilidade do peso dos homens dessa região entre 20 e 30 anos é normal com média de 90 kg e desvio padrão de 10 kg. Um endocrinologista propõe um tratamento para combater a obesidade que consiste de exercícios físicos, dietas e ingestão de um medicamento. Ele afirma que com seu tratamento o peso médio da população da faixa em estudo diminuirá num período de três meses.

Neste caso as hipóteses que deverão ser testados são:

$$H_0 : \mu = 90$$

$$H_1 : \mu < 90$$

onde μ é a média dos pesos do homens em estudo após o tratamento.

Exemplo 5: Um fabricante de uma certa peça afirma que o tempo médio de vida das peças produzidas é de 1000 horas. Suponha que os engenheiros de produção têm interesse em verificar se a modificação do processo de fabricação aumenta a duração das peças

$$H_0 : \mu = 1000$$
$$H_1 : \mu > 1000$$

sendo μ o tempo médio das peças produzidas pelo novo processo.

Procedimento básico de teste de hipóteses

O procedimento básico de teste de hipóteses relativo ao parâmetro θ de uma população, será decomposto em 4 passos:

(i) Definição as hipóteses:

$$H_0 : \theta = \theta_0$$
$$H_1 : \theta < \theta_0 \ ou \ \theta > \theta_0 \ ou \ \theta \neq \theta_0$$

(ii) Identificação da estatística do teste e caracterização da sua distribuição.

(iii) Definição da regra de decisão, com a especificação do nível de significância do teste.

(iv) Cálculo da estatística de teste e tomada de decisão.

Teste de hipóteses para uma média populacional

Considere uma amostra aleatória de tamanho n de uma população normal com média μ (desconhecida) e variância σ^2(conhecida) Inicialmente, considera-se o caso do teste unilateral esquerdo. Suponha que tem-se interesse em verificar as seguintes hipóteses:

(i)

$$H_0 : \mu = \mu_0$$

$$H_1 : \mu < \mu_0$$

(ii) A estatística do teste é a média amostral $\overline{X}$. Se população é normal (ou se amostra é grande n $\geq$ 30, mesmo que a população não é normal) a distribuição de $\overline{X}$ é $N\left(\mu, \sigma^2 / n\right)$ e a variável aleatória sob H_0

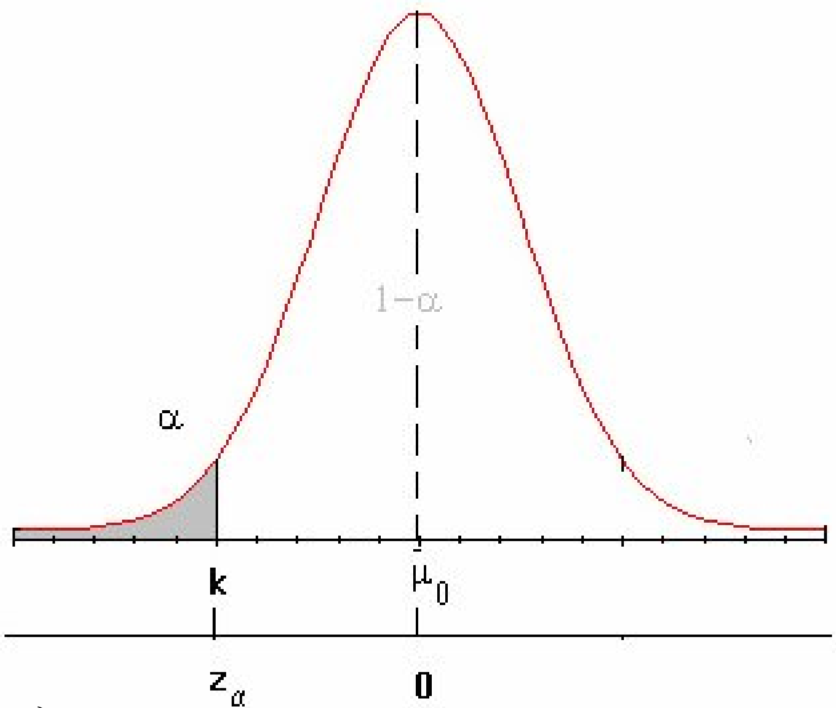
$$Z = \frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}} \sim N(0,1)$$

(iii) É razoável, rejeitar H_0 em favor de H_1, se a média amostral $\bar{X}$ é demasiado pequena em relação μ_0. A região crítica, então poderia ser obtido, selecionando um k da média amostral, de maneira que Rc={ $\bar{X} \leq$ k } onde k é tal que $P(\bar{X} \leq k \mid H_0 : \mu = \mu_0) = \alpha$. Ou seja sob H_0

$$P\left(\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} \leq \frac{k-\mu_0}{\sigma/\sqrt{n}}\right) = P\left(z \leq \frac{k-\mu_0}{\sigma/\sqrt{n}}\right) = \alpha$$

$$\Rightarrow \frac{k-\mu_0}{\sigma/\sqrt{n}} = z_\alpha \Rightarrow k = \mu_0 + z_\alpha \times \frac{\sigma}{\sqrt{n}}$$

$$\Rightarrow Rc = \left\{\bar{X} \leq \mu_0 + z_\alpha \times \frac{\sigma}{\sqrt{n}}\right\}$$



(iv) Conclusão: se $\bar{x} \in Rc = \left\{\bar{X} \leq \mu_0 + z_\alpha \times \frac{\sigma}{\sqrt{n}}\right\}$, rejeita-se H_0 em caso contrário não se rejeita H_0.

Método alternativo

Um método alternativo prático é trabalhar diretamente na escala Z

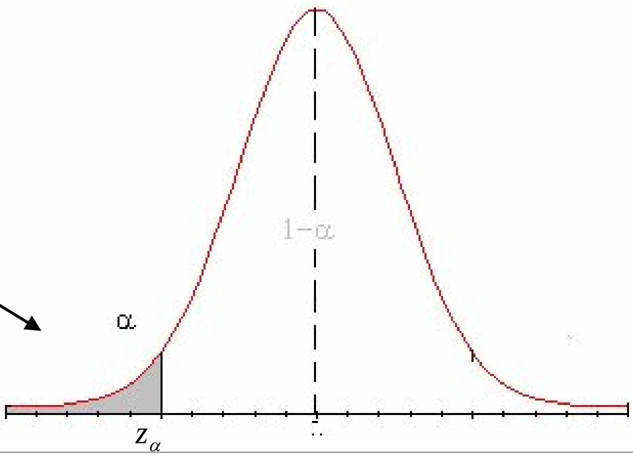
$$(i)\ H_0 : \mu = \mu_0 \ \text{ contra } \ H_1 : \mu < \mu_0$$

(ii) A estatística de teste

$$Z = \frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}} \underset{sob\ H_0}{\sim} N(0,1)$$

(iii) A região crítica para um nível de significância α fixado

$$Rc = \{z \in R; Z \leq z_\alpha\}$$



iv) se $z_{obs} \in Rc = \{Z \leq z_\alpha\}$, rejeita-se H_0 em caso contrário não se rejeita H_0.

Exemplo

Um comprador de tijolos acha que a qualidade dos tijolos está diminuindo. De experiências anteriores, considera-se a resistência média ao desmoronamento de tais tijolos é igual a 200 kg, com um desvio padrão de 10 kg. Uma amostra de 100 tijolos, escolhidos ao acaso, forneceu uma média de 195 kg. Ao nível de significância de 5%, pode-se afirmar que a resistência média ao desmoronamento diminuiu?

(i) As hipóteses de interesse são :

$$H_0 : \mu = 200\,Kg$$

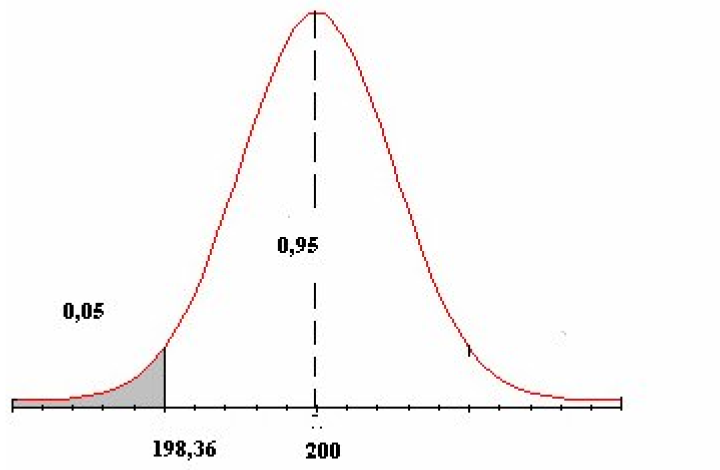
$$H_1 : \mu < 200\,Kg$$

(ii) A estatística do teste é a média amostral $\overline{X}$. Já que n=100 $\geq$ 30, tem-se que sob H_0 $\overline{X} \sim N\left(200, \frac{100}{100}\right)$.

(iii) A região crítica, então poderia ser obtido, selecionando um k da média amostral, de maneira que Rc={ $\overline{X} \leq$ k } onde k é tal que $P(\overline{X} \leq k \mid H_0 : \mu = \mu_0) = \alpha = 0{,}05$. Ou seja sob H_0

$$P\left(\frac{\overline{X}-200}{10/\sqrt{100}} \leq \frac{k-200}{10/\sqrt{100}}\right) = P\left(z \leq \frac{k-200}{1}\right) = \alpha = 0{,}05 \Rightarrow k-200 = -1{,}64 \Rightarrow k = 198{,}36$$

$$\Rightarrow Rc = \left\{\overline{X} \leq 198{,}36\right\}$$



(iv) Do enunciado tem-se $\overline{x} = 195 \in Rc = \left\{\overline{X} \leq 198{,}36\right\}$, $\Rightarrow$ rejeita-se H_0 ao nível de 5% de significância.

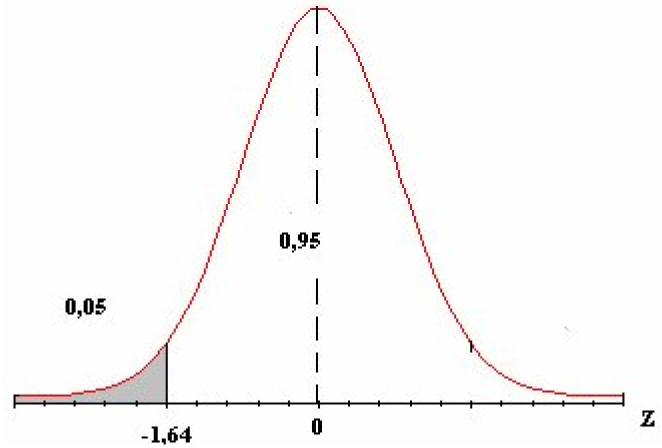
Método alternativo

$$(i)\ H_0 : \mu = 200 \ \text{ contra } \ H_1 : \mu < 200$$

(ii) A estatística de teste $$Z = \frac{\overline{X} - 200}{\sigma / \sqrt{n}} \underset{sob\ H_0}{\sim} N(0,1)$$

(iii) A região crítica para um nível de significância α=0,05 fixado

$$Rc = \{z \in R; R \leq -1{,}64\}$$



iv) Do enunciado temos: $z_{obs} = \frac{195 - 200}{10/\sqrt{100}} = -5 \in R_c$ $\Rightarrow$ rejeita-se H_0. ao nível de 5% de significância.

Procedimento Geral

A seguir é apresentado o procedimento geral de teste de hipóteses para uma média populacional considerando o procedimento alternativo descrito acima.

(i)

$$\underbrace{\begin{array}{l} H_0 : \mu = \mu_0 \,(ou \ \geq \mu_0) \\ H_1 : \mu < \mu_0 \end{array}}_{U.Esquerdo} \qquad \underbrace{\begin{array}{l} H_0 : \mu = \mu_0 \ (ou \ \leq \mu_0) \\ H_1 : \mu > \mu_0 \end{array}}_{U.Direito} \qquad \underbrace{\begin{array}{l} H_0 : \mu = \mu_0 \\ H_1 : \mu \neq \mu_0 \end{array}}_{Bilateral}$$

(ii) A estatística de teste

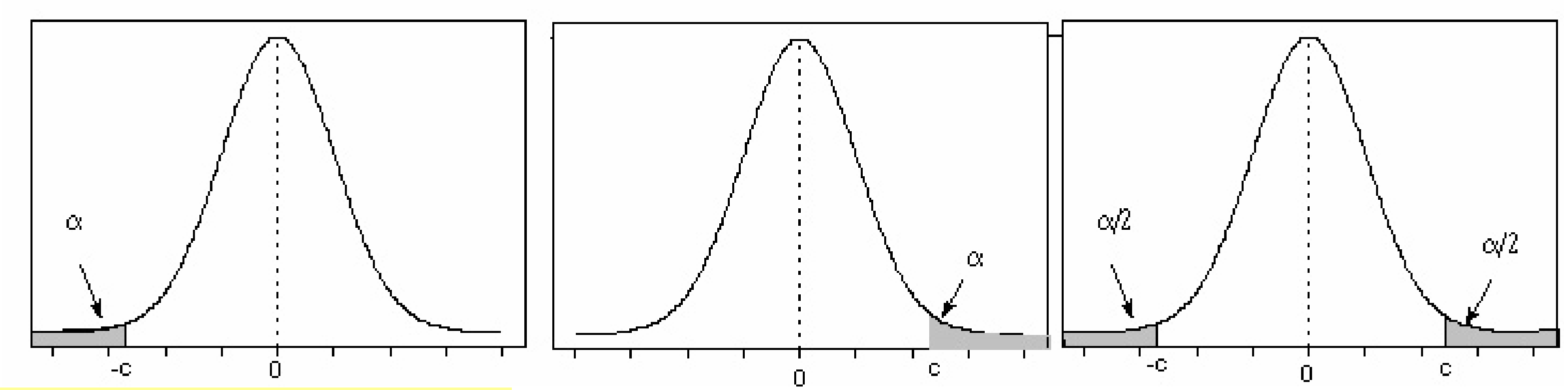
(a) Quando a variância e conhecida

$$Z = \frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}} \underset{sob\ H_0}{\sim} N(0,1)$$

(b) Quando a variância é desconhecida e amostra pequenas

$$T = \frac{\overline{X} - \mu_0}{S / \sqrt{n}} \underset{sob\ H_0}{\sim} t(n-1)$$

(iii) A região crítica para um nível de significância α fixado



$$R_c^{(Z)} = \{z \in R; Z \leq -c\} \quad R_c^{(Z)} = \{z \in R; Z \geq c\} \quad R_c^{(Z)} = \{z \in R; |Z| \geq c\}$$

$$R_c^{(T)} = \{z \in T; T \leq -c\} \quad R_c^{(T)} = \{z \in T; T \geq c\} \quad R_c^{(T)} = \{z \in T; |T| \geq c\}$$

(iv) Se a $ET_{obs} \in R_C$., rejeita-se H_o em caso contrário não se rejeita H_0.

Exemplo

Os registros dos últimos anos de um colégio atestam para calouros admitidos uma nota média 115 (teste vocacional). Para testar a hipóteses de que a média de uma nova turma é a mesma das turmas anteriores, retirou-se, ao acaso, uma amostra de 20 notas, obtendo-se média 118 desvio padrão 20. Use α=0,05

Supondo que as notas dos novos calouros tem distribuição normal com média μ e desvio padrão σ

(i) As hipóteses de interesse são :

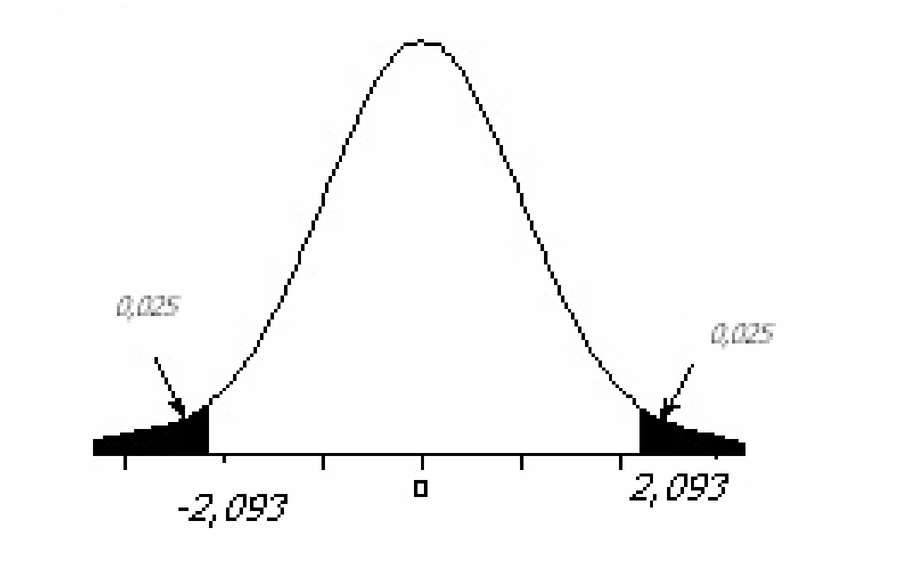
$$H_0 : \mu = 115$$

$$H_1 : \mu \neq 115$$

(ii) A estatística de teste

$$T = \frac{\bar{X} - 115}{S / \sqrt{n}} \underset{sob\ H_0}{\sim} t(n-1)$$

(iii) A região crítica para um nível de significância α=0,05 fixado



$$Rc = \{z \in T; |T| \geq 2,093\}$$

iv) Do enunciado temos: $T_{obs} = \dfrac{118-115}{20/\sqrt{20}} = 0,67 \notin R_c \Rightarrow$ não rejeita-se H_0.

ao nível de 5% de significância.

Teste de hipóteses para uma proporção populacional

O procedimento para os testes de hipóteses para proporção populacional é basicamente igual ao procedimento para o teste para uma média populacional. Considere o problema de testar a hipótese que a proporção de sucessos de um ensaio de Bernoulli é igual a valor especifico, p_0. Isto é, testar as seguintes hipóteses:

(i)

$$\underbrace{\begin{array}{l} H_0 : p = p_0 \,(ou \ \geq p_0) \\ H_1 : p < p_0 \end{array}}_{U.Esquerdo} \quad \underbrace{\begin{array}{l} H_0 : p = p_0 \ (ou \ \leq \mu_0) \\ H_1 : p > p_0 \end{array}}_{U.Direito} \quad \underbrace{\begin{array}{l} H_0 : p = p_0 \\ H_1 : p \neq p_0 \end{array}}_{Bilateral}$$

(ii) A estatística de teste

$$Z = \frac{\hat{p} - p_o}{\sqrt{\dfrac{p_0(1 - p_o)}{n}}} \underset{sob\ H_0}{\sim} N(0,1)$$

Exemplo

Um estudo é realizado para determinar a relação entre uma certa droga e certa anomalia em embriões de frango. Injetou-se 50 ovos fertilizados com a droga no quarto dia de incubação. No vigésimo dia de incubação, os embriões foram examinados e 7 apresentaram a anomalia. Suponha que deseja-se averiguar se a proporção verdadeira é inferior a 25% com um nível de significância de 0,05.

(i) As hipóteses de interesse são:

$H_0 : p = 0{,}25$

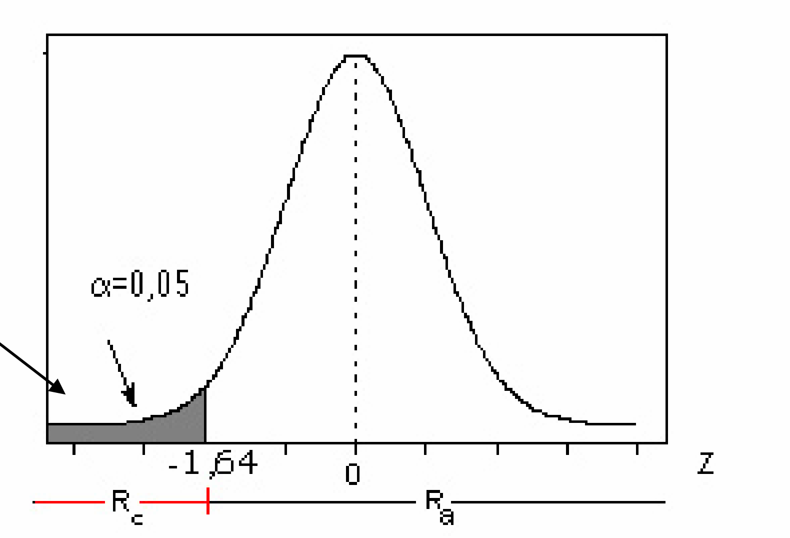
$H_1 : p < 0{,}25$

(ii) A estatística de teste

$$Z = \frac{\hat{p} - 0{,}25}{\sqrt{\dfrac{0{,}25(1-0{,}25)}{50}}} \underset{sob\ H_0}{\sim} N(0{,}1)$$

(iii) A região crítica para um nível de significância α=0,05 fixado

$$Rc = \{z \in R; R \leq -1,64\}$$



iv) Do enunciado temos n=50, $\hat{p} = \frac{7}{50} = 0,14$; $z_{obs} = \frac{0,14 - 0,25}{\sqrt{\frac{025 \times 0,75}{50}}} = -1,7963 \in R_c \Rightarrow$

rejeita-se $H_{0.}$ ao nível de 5% de significância.

Inferência Para Duas Amostras

$X \sim N(\mu_1, \sigma_1^2)$

População 1

$X_1, \cdots, X_n$

$$\bar{X} \sim N\left(\mu_1, \frac{\sigma_1^2}{n}\right)$$

$Y \sim N(\mu_2, \sigma_2^2)$

População 2

$Y_1, \cdots, Y_m$

$$\bar{Y} \sim N\left(\mu_2, \frac{\sigma_m^2}{m}\right)$$

$$\bar{X} - \bar{Y} \sim N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}\right)$$

Teste de hipóteses e intervalo de confiança para $\mu_1 - \mu_2$

Suponha que $X_1,\ldots,X_n$ é uma amostral aleatória de tamanho n de uma população com característica X, que tem distribuição normal com média μ_1 e variância σ_1^2. Considere que $Y_1,\ldots,Y_m$ é uma amostra aleatória de tamanho m, de uma população com característica Y que tem distribuição normal com média μ_2 e variância σ_2^2, alem disso, X e Y são independentes. Suponha que tem-se interesse em verificar se existe ou não uma diferença significativa entre as médias populacionais μ_1 e μ_2. O procedimento básico de teste, neste caso é a seguinte:

(i)

$$\underbrace{\begin{array}{l} H_0 : \mu_1 - \mu_2 = \Delta (ou\ \geq \Delta) \\ H_1 : \mu_1 - \mu_2 < \Delta \end{array}}_{U.Esquerdo} \quad \underbrace{\begin{array}{l} H_0 : \mu = \mu_0\ (ou\ \leq \Delta) \\ H_1 : \mu_1 - \mu_2 > \Delta \end{array}}_{U.Direito} \quad \underbrace{\begin{array}{l} H_0 : \mu_1 - \mu_2 = \Delta \\ H_1 : \mu_1 - \mu_2 \neq \Delta \end{array}}_{Bilateral}$$

onde Δ é constante conhecida no caso Δ=0, temos teste de hipóteses para a igualdade de 2 médias populacionais

(ii) A estatística de teste

(a) Quando $\sigma_1^2, e\ \sigma_2^2$ são conhecidos

$$Z = \frac{\overline{X} - \overline{Y} - \Delta}{\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}} \underset{sob\ H_0}{\sim} N(0,1)$$

(b) Quando $\sigma_1^2 = \sigma_2^2 = \sigma^2$ desconhecidos

$$T = \frac{\overline{X} - \overline{Y} - \Delta}{\sqrt{S_p^2\left(\frac{1}{n} + \frac{1}{m}\right)}} \underset{sob\ H_0}{\sim} t(n+m-2)$$

onde $S_p^2 = \dfrac{(n-1)S_1^2 + (m-1)S_2^2}{n+m-2}$

Exemplo 1: Estuda-se o conteúdo de nicotina de duas marcas de cigarros (A e B), obtendo-se os seguintes resultados.

A: 17; 20; 23; 20

B: 18; 20; 21; 22; 24

Admitindo que o conteúdo de nicotinas das duas marcas tem distribuição normal e que as variâncias populacionais são iguais, com α=0,05, pode-se afirmar que existe alguma diferença significativa no conteúdo médio de nicotina nas duas marcas?

Sejam X: O conteúdo de nicotina da marca A $X \sim N(\mu_1, \sigma_1^2)$

Y: : O conteúdo de nicotina da marca B $Y \sim N(\mu_2, \sigma_2^2)$

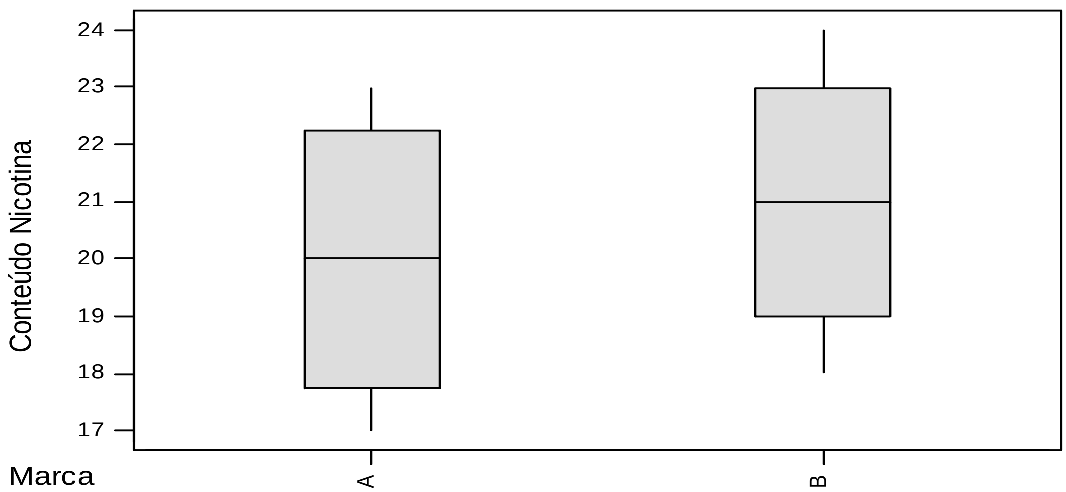
Nosso interesse é testar as seguintes hipóteses:

(i)

$$H_0 : \mu_1 = \mu_2 \qquad H_0 : \mu_1 - \mu_2 = 0$$

$$H_1 : \mu_1 \neq \mu_2 \qquad H_1 : \mu_1 - \mu_2 \neq 0$$

Boxplots do Conteúdo de Nicotina por Marca



$$n = 4,\ \overline{X} = 20\ S_1^2 = 6$$

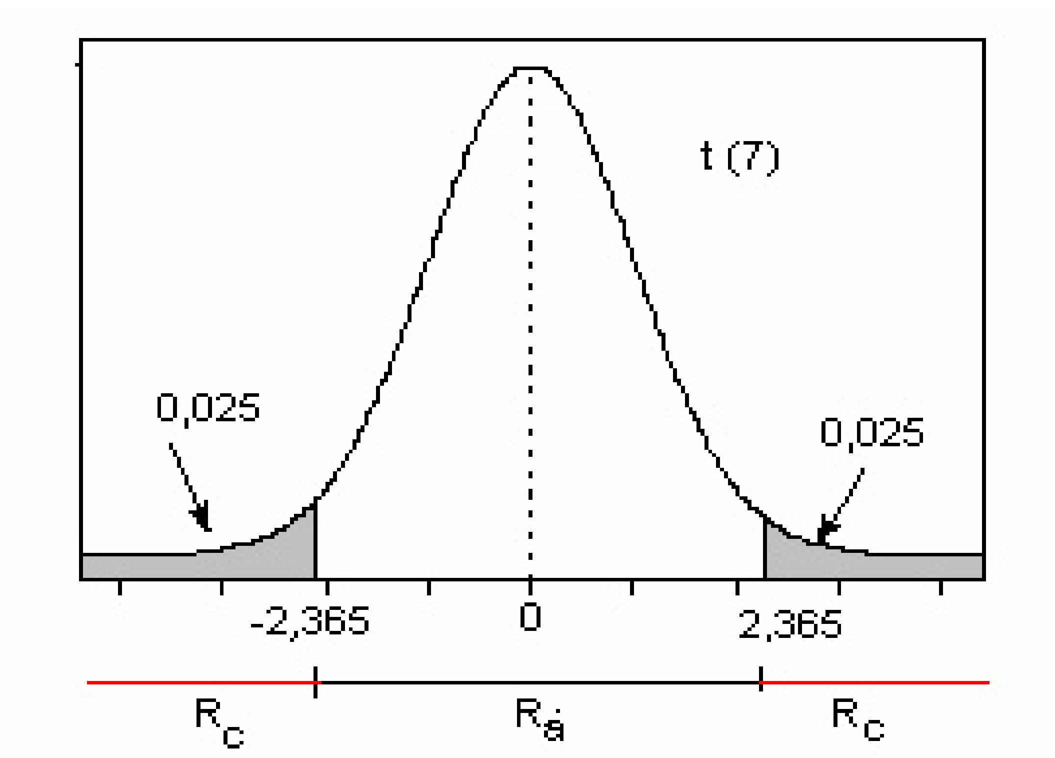
$$m = 5,\ \overline{Y} = 21\ S_2^2 = 5$$

A estatística de teste é dada por:

(ii)

$$T = \frac{\overline{X} - \overline{Y}}{\sqrt{S_p^2\left(\frac{1}{n} + \frac{1}{m}\right)}} \underset{sob\ H_0}{\sim} t(n+m-2)$$

(iii) A região crítica, para α=0,05, (parte achurada) representa os valores correspondente da distribuição t-Student com n+m-2=4+5-2=7 graus de liberdade com mostra a figura



$$Rc = \{t \in t(7); |T| \geq 2{,}365\}$$

(iv) Dos dados do exemplo temos:

$$S_p^2 = \frac{(n-1)S_1^2 + (m-1)S_2^2}{n+m-2} = \frac{(4-1)(6)+(5-1)5}{4+5-2} = \frac{38}{7}$$

Daí temos, que estatística observada ou calculada é:

$$T_{obs} = \frac{\bar{X} - \bar{Y}}{\sqrt{S_p^2\left(\frac{1}{n} + \frac{1}{m}\right)}} = \frac{20-21}{\sqrt{\frac{38}{7}\left(\frac{1}{4} + \frac{1}{5}\right)}} = -0,641$$

Como $T_{obs} \notin Rc \Rightarrow$ Não se rejeita H_0

Bernoulli(p_1)

população 1

$X_1, \cdots, X_n$

$$\hat{p}_1 \sim N\left(p_1, \frac{p_1(1-p_2)}{n} \right)$$

Bernoulli(p_2)

população 2

$Y_1, \cdots, Y_m$

$$\hat{p}_2 \sim N\left(p_2, \frac{p_2(1-p_2)}{2} \right)$$

$$\hat{p}_1 - \hat{p}_2 \sim N\left(p_1 - p_2, \frac{p_1(1-p_1)}{n} + \frac{p_2(1-p_2)}{m} \right)$$

Teste de hipóteses para $p_1 - p_2$

Suponha que tem-se duas amostras independentes de tamanhos n e m suficientemente grandes (n>30 e m>30), de duas populações Bernoulli, com probabilidades de sucessos p_1 e p_2 respectivamente. E sejam X: o número de sucessos na amostra de tamanho n e Y: o número de sucessos na amostra de tamanho m. Portanto, X~B(n,p_1 e Y~ B(m,p_2). Há interesse em verificar as seguintes hipóteses estatística:

(i)

$$H_0: p_1 = p_2\,(ou\ \geq p_2) \quad H_0: p_1 = p_2\ (ou\ \leq p_2) \quad H_0: p_1 = p_2$$

$$\underbrace{H_1: p_1 < p_2}_{U.Esquerdo} \qquad \underbrace{H_1: p_1 > p_2}_{U.Direito} \qquad \underbrace{H_1: p_1 \neq p_2}_{Bilateral}$$

(ii) A estatística de teste

$$Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\overline{p}(1-\overline{p})\left(\frac{1}{n}+\frac{1}{m}\right)}} \underset{Sob\ H_0}{\sim} N(0,1)$$

onde $\hat{p}_1 = \dfrac{x}{n}, \; \hat{p}_2 = \dfrac{y}{m}; \;\; \overline{p} = \dfrac{x+y}{n+m} = \dfrac{n\hat{p}_1 + m\,\hat{p}_2}{n+m}$

Os passos (iii) e (iv) são equivalentes ao procedimento de teste para uma média populacional.

Exemplo 3: Dois tipos de solução de polimento estão sendo avaliados para possível uso em uma operação de polimento na fabricação de lentes intra-oculares usadas no olho humano depois de uma operação de catarata. Trezentas lentes foram polidas usando a primeira solução de polimento e, desse número 253 não tiveram defeitos induzidos pelo polimento. Outras 300 lentes foram polidas, usando a segunda solução de polimento sendo 196 lentes consideradas satisfatórios. Há qualquer razão para acreditar que as duas soluções diferem? Use α=0,01.

X: o número de lentes sem defeito das 300 polidas com a 1ª solução, $\Rightarrow X \sim B(300, p_1)$

Y: o número de lentes sem defeito das 300 polidas com a 2ª solução $\Rightarrow Y \sim B(300, p_2)$.

Nosso interesse é testar as seguintes hipóteses:

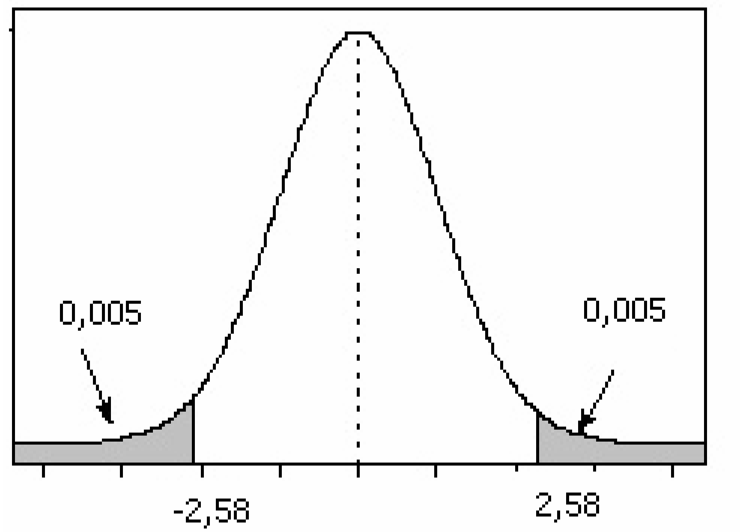
$$H_0 : p_1 = p_2$$
$$H_1 : p_1 \neq p_2$$

(ii) A estatística de teste

$$Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\bar{p}(1-\bar{p})\left(\frac{1}{n} + \frac{1}{m}\right)}} \underset{Sob\ H_0}{\sim} N(0,1)$$

(iii) A região crítica, para α=0,01, (parte achurada) representa os valores correspondente da distribuição norma padrão com mostra a figura

$$Rc = \{t \in Z; |Z| \geq 2{,}58\}$$



(iv) Dos dados do exemplo temos:

$$\hat{p}_1 = \frac{253}{300} = 0{,}8433;\ \hat{p}_2 = \frac{196}{300};\ n = m = 300;\ \overline{p} = \frac{253+196}{300} = 0{,}7483$$

$$Z_{obs} = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\overline{p}(1-\overline{p})\left(\frac{1}{n}+\frac{1}{m}\right)}} = \frac{0{,}8433 - 0{,}6533}{\sqrt{0{,}7483(0{,}2517)\left(\frac{1}{300}+\frac{1}{300}\right)}} = 5{,}36$$

Como $Z_{obs} \in Rc \Rightarrow$ rejeita-se H_0