III- Ensaios paramétricos

O objectivo dos ensaios de hipóteses é possibilitar validar ou não determinadas afirmações que se fazem sobre determinadas características (parâmetros) de uma ou mais populações. O contributo assenta no auxílio à tomada de decisão entre opções (hipóteses) alternativas.

3.1- Hipóteses e erros

Considere-se o julgamento de uma pessoa acusada de ter cometido um crime. O julgamento consiste em apreciar os factos apresentados pela acusação e pela defesa e decidir em função deles e da lei. Duas hipóteses se podem colocar:

H_o: o réu é inocente

H_a: o réu é culpado

Se as provas apresentadas forem incompatíveis com a manutenção de H_o, a decisão é considerar o réu como culpado, e, rejeitar H_o; se as provas forem compatíveis com H_o, a decisão é considerá-lo inocente, i. e., não se rejeita H_o.
Contudo, associado a estas duas decisões existe o risco de se estar a tomar uma decisão errada.

- Se o réu é na verdade inocente e for dado como culpado.
- Se o réu é na verdade culpado e for dado como inocente.

Ocorrem dois tipos de erros:

Erro tipo I- ocorre quando se rejeita indevidamente uma hipótese verdadeira. Ou seja, rejeita-se H_o quando H_o é verdadeira.

Erro tipo II- ocorre quando não se rejeita H_o sendo H_o falsa.

Decisão	**Estados da natureza**	
	H_0 **é verdadeira** (o réu é inocente)	H_0 **é falsa** (o réu é culpado)
Rejeitar H_0	Decisão incorrecta **Erro tipo I-** α	Decisão correcta
Não Rejeitar H_0	Decisão correcta	Decisão incorrecta **Erro tipo II-** β

No domínio estatístico, a decisão é tomada em face da evidência dos dados recolhidos em amostras aleatórias. Assim, em qualquer processo de decisão deve-se ter presente estes tipos de erros que se podem cometer para que possam ser **minimizados.**

3.2 – Como se deve proceder a um ensaio de hipóteses

A realização de um "bom" ensaio de hipóteses parte de uma correcta formulação de hipóteses a qual se obtém pelo estudo do problema proposto. A recolha de dados é também determinante na realização de um "bom" ensaio de hipóteses. Os erros de amostragem impedem que a amostra represente convenientemente a população em causa.

Realização de um ensaio de hipóteses

1° passo: Formulação das hipóteses

As hipóteses podem ser **simples,** quando o parâmetro assume um único valor, ou **compostas,** quando o parâmetro pode assumir vários valores. Por exemplo, "a média de produtos empacotados por minuto pela máquina A é de 60", i. e.:

$$\underbrace{\begin{array}{l} H_0 : \mu = 60 \\ H_a : \mu \neq 60 \end{array}}_{\text{Teste bilateral}} \quad \text{ou} \quad \underbrace{\begin{array}{l} H_0 : \mu = 60 \\ H_a : \mu < 60 \end{array}}_{\text{Teste unilateral à esquerda}_*} \quad \text{ou} \quad \underbrace{\begin{array}{l} H_0 : \mu = 60 \\ H_a : \mu > 60 \end{array}}_{\text{Teste unilateral à direita}}$$

Consoante a hipótese que se dispõe relativamente à hipótese alternativa, o teste pode ser bilateral ou unilateral (à direita ou à esquerda).

De qualquer forma H_o é uma hipótese simples, enquanto H_a é uma hipótese composta.

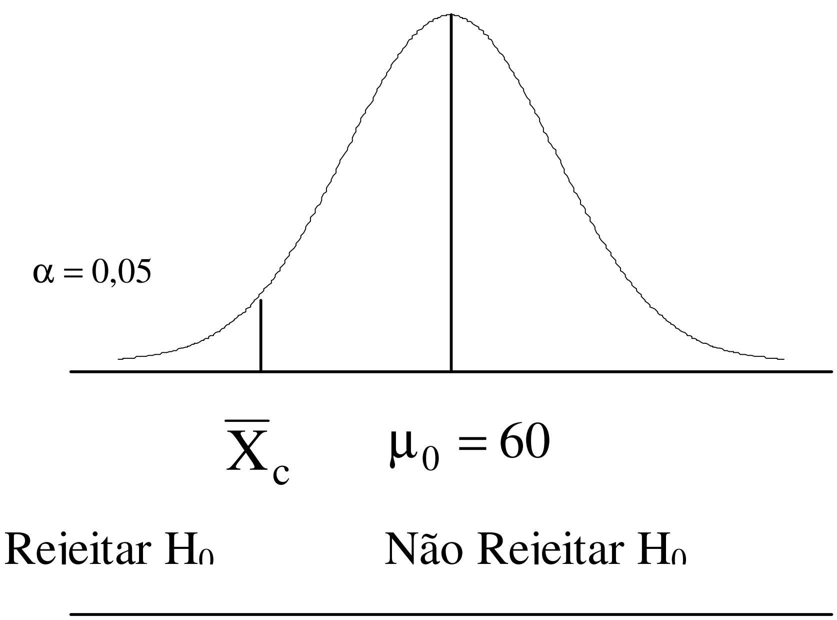
- Um procedimento usual é colocar em H_o a situação de igualdade já que essa igualdade permite identificar sem ambiguidade quais os valores da estatística que levam a rejeitar ou não a hipótese nula.

- Se H_o não for rejeitada, isso não quer dizer que H_o seja verdadeira. O que se pode dizer é que ela é provavelmente verdadeira, embora também se possa ter cometido um erro tipo II. Por isso, não se deve dizer que se "aceita" H_o.

2º passo: Fixação do nível de significância

O nível de significância é complementar do nível de confiança e deve ser o mais baixo possível já que é um erro e é através dele que se estabelece a região crítica, i. e., aquela que permite a rejeição de H_o.

Para o exemplo do caso anterior, suponha-se que a população é normal e, portanto, a distribuição da média amostral é também normal e o seu valor esperado é igual ao da população. Então, se o nível de significância, α, for de 0,05 e para um **teste unilateral à esquerda**, o problema pode ser representado como



Se a média amostral for inferior a $\overline{X}_c$, a decisão tende a rejeitar H_o. Quanto mais significativa for a diferença entre μ e $\overline{X}$, tanto mais se acredita que a população que gerou aquela amostra não é a que figura em H_o e decide-se rejeitar H_o. O ponto de separação entre uma diferença significativa de uma não significativa é designado por ponto crítico e significa o risco que o decisor está disposto a correr de rejeitar H_o quando H_o é verdadeiro. Logo,

$$\alpha = \text{erro tipo I} = \text{nível de significância}$$

$\therefore$ É desejável que o nível de significância seja tão baixo quanto possível, já que é um tipo de erro, o erro tipo I.

3° passo: <u>Escolha do teste de ensaio a usar e estabelecimento da regra de decisão</u>

O teste de ensaio a utilizar depende da estatística adequada ao problema. A regra de decisão consiste em definir a acção a tomar face ao resultado da amostra. É no espaço amostral que se vai trabalhar, definido-se duas regiões:

- a região crítica
- a região de aceitação

Se a estatística a usar é $\overline{X}$, se a variância populacional for conhecida e igual a 3, e, a dimensão da amostra for 25, o teste de ensaio a utilizar é

$$\frac{\overline{X}-\mu_o}{\sigma/\sqrt{n}} \cap N(0;1) \quad \text{e} \quad \overline{X}_c = \mu_o - 1{,}645\frac{\sigma}{\sqrt{n}} \approx 59{,}01$$

As regiões crítica e de aceitação serão:

$$RC = \left]-\infty; \overline{X}_c\right] = \left]-\infty; 59{,}01\right]$$

$$RA = \left]\overline{X}_c; +\infty\right[= \left]59{,}01; +\infty\right[$$

ou, em termos da distribuição normal padrão:

$$RC = \left]-\infty; -1{,}645\right] \text{ e } RA = \left]-1{,}645; +\infty\right[.$$

Em resumo, a regra de decisão, em termos do valor da média amostral, será:

1- Se $\overline{X} \leq 59{,}01$ produtos empacotados, **rejeita-se** H_o (a máquina não cumpre a especificação).
2- Se $\overline{X} > 59{,}01$ produtos empacotados, **não se rejeita** H_o (a máquina cumpre com a especificação).

A regra de decisão, em função da variável Z, será:

1- Se $z \leq -1{,}645$, **rejeita-se** H_o.
2- Se $z > -1{,}645$, **não se rejeita** H_o.

4º passo: Tomada de decisão

- Se a média amostral for $\overline{X} = 58 < 59{,}01$, $\overline{X} \in RC$. Quer isto dizer que existe evidência estatística que permite rejeitar H_o. A diferença entre o valor encontrado para a média amostral e o valor proposto para a média da população é significativamente grande. Logo, para este nível de significância e com base nos dados amostrais, concluímos que se rejeita H_o.

  A mesma decisão será tomada se utilizarmos a variável normal padrão.

  Dado que $z = \dfrac{58-60}{0{,}6} = -3,(3) < -1{,}645$, deve-se rejeitar H_o.

- Se $\overline{X} = 59{,}5 > 59{,}01$ ou se $Z = \dfrac{59{,}5-60}{0{,}6} = -0{,}8(3) > -1{,}645$, não se deve rejeitar H_o.

3.3- Análise dos erros

No exemplo que se está a seguir, se $\overline{X} = 58$, o erro que se pode estar a incorrer é, no máximo, o erro tipo I. Se $\overline{X} = 59{,}5$, pode-se estar a cometer um erro tipo II, β, já que se pode estar a não rejeitar H_o e esta ser falsa.

- Probabilidade do erro tipo I:

$$\alpha = P[\text{erro tipo I}] = P[\text{rejeitar } H_o | H_o \text{ verdadeira}]$$
$$= P[\overline{X}_c \le 59{,}01 | \mu_o = 60] = P\left[\frac{\overline{X}_c - 60}{0{,}6} \le \frac{59{,}01 - 60}{0{,}6}\right] =$$
$$= P[Z \le -1{,}645] = 0{,}05$$

O erro tipo I é função do valor de μ_o, i. e.,

$$\alpha(\mu_o) = P[\text{erro tipo I}]$$

E esta probabilidade é no máximo igual ao valor de α já que se está perante um teste unilateral[1], ou seja

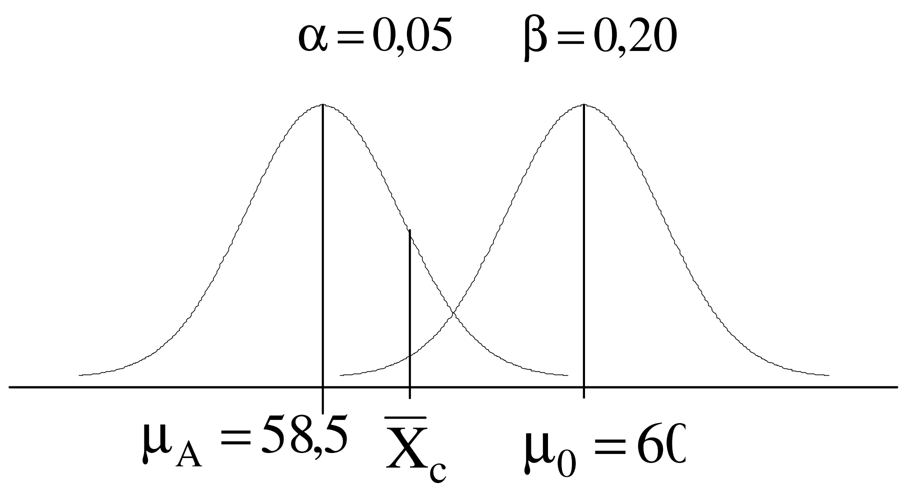
[1] Se estivéssemos perante um teste bilateral, este erro seria exactamente igual a α.

$\alpha(\mu) \leq \alpha, \quad \forall \mu \in H_o$. Por isso se diz que α é o *supremo* dos α's.

- Probabilidade do erro tipo II:

Sabe-se que:

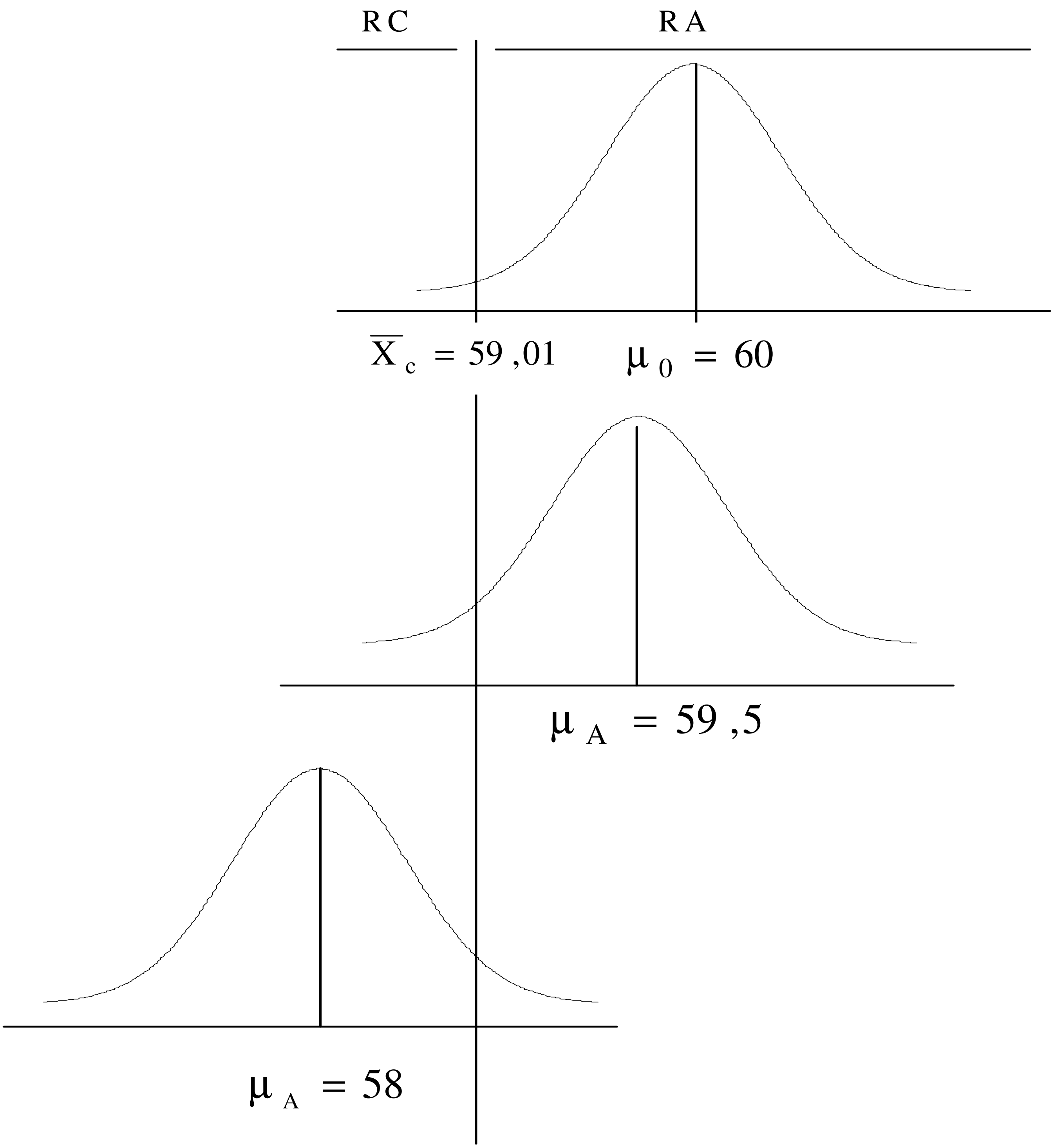
$$\beta = P[\text{erro tipo II}] = P[\text{não rejeitar } H_o | H_o \text{ falsa}]$$
$$= P[\overline{X}_c > 59{,}01 | \mu_a]$$



$$\beta(\mu_a = 59{,}5) = P\left[\overline{X}_c > 59{,}01 \middle| \mu_a = 59{,}5\right] =$$
$$= P\left[\frac{\overline{X}_c - 59{,}5}{0{,}6} > \frac{59{,}01 - 59{,}5}{0{,}6}\right] =$$
$$= P\left[Z > -0{,}81(6)\right] = P\left[Z < 0{,}81(6)\right] \approx 0{,}7939$$

$$\beta(\mu_a = 59) = P\left[\overline{X}_c > 59{,}01 \middle| \mu_a = 59\right] =$$
$$= P\left[\frac{\overline{X}_c - 59}{0{,}6} > \frac{59{,}01 - 59}{0{,}6}\right] =$$
$$= P\left[Z > 0{,}01(6)\right] = 1 - \phi(0{,}01(6) \approx 0{,}492$$

$$(\mu_a = 58) = P\left[\overline{X}_c > 59{,}01 \middle| \mu_a = 58\right] = P\left[\frac{\overline{X}_c - 58}{0{,}6} > \frac{59{,}01 - 58}{0{,}6}\right] =$$
$$= P\left[Z > 1{,}68(3)\right] = 1 - \phi(1{,}68(3)) \approx 0{,}0465$$


RC
RA
$\overline{X}_c = 59{,}01$
$\mu_0 = 60$
$\mu_A = 59{,}5$
$\mu_A = 58$

Conclusão: a probabilidade de se incorrer neste tipo de erro decresce à medida que os valores de μ_a se afastam de μ_o já que é mais fácil de se cometer um erro quando se tem de decidir entre duas hipóteses muito parecidas.

Qual o novo α a partir do qual a decisão a tomar se altera?

Como já se viu, se $\overline{X} = 59{,}5$ a decisão a tomar é não rejeitar H_o para um $\alpha = 0{,}05$. Neste caso, pode-se determinar qual o novo α a partir do qual a decisão a tomar se altera.

$$P[Z \leq -0{,}8(3)] = ?$$
$$P[Z \leq -0{,}8(3)] = 1 - \phi(0{,}8(3)) \approx 1 - 0{,}7967 \approx 0{,}2$$

Logo, para $\alpha \geq 0{,}2$, a decisão alterar-se-ia.

Minimização dos erros

Como já devem ter verificado, os erros tipo I e II não são complementares já que o erro tipo I é condicionado à hipótese de que H_o é verdadeira, enquanto que o erro tipo II é condicionado à hipótese de que H_a é verdadeira. Contudo, como fazem parte do mesmo problema de decisão, estes erros variam inversamente, i. e., quando um aumenta o outro diminui.

A única maneira de se minimizar estes dois tipos de erros simultaneamente é aumentar-se a dimensão da amostra.

3.4 - Função potência de ensaio

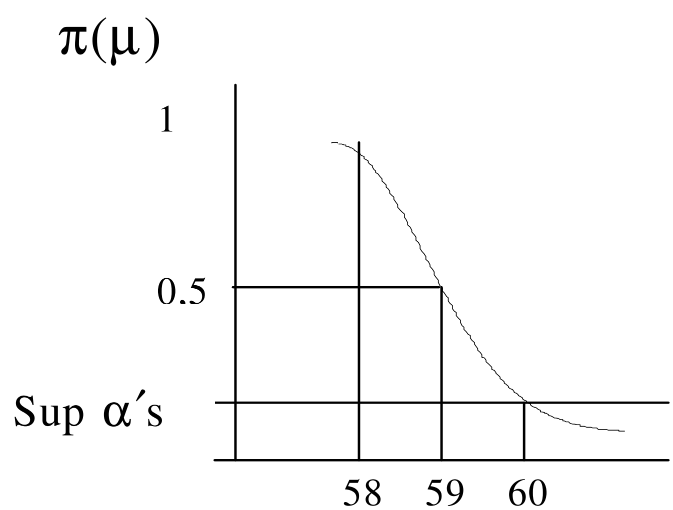
A função potência do ensaio permite calcular-se a probabilidade de se rejeitar H_o, i. e.,

$$\pi = \begin{cases} \pi(\mu_o) = 1 - \beta(\mu_o) = \alpha \\ \pi(\mu_a) = 1 - \beta(\mu_a) = P\left[\text{rejeitar } H_o \middle| \mu_a\right] \end{cases}$$

Contrariamente ao que foi concluído para β, quanto mais próximo do valor de μ_o estiver o valor de μ_a, menos potente é o teste. Isto é, menos capacidade tem para distinguir os verdadeiros valores do parâmetro dos falsos. Quanto mais afastados estiverem os valores de μ_a relativamente a μ_o, mais capaz é o teste de tomar decisões correctas.

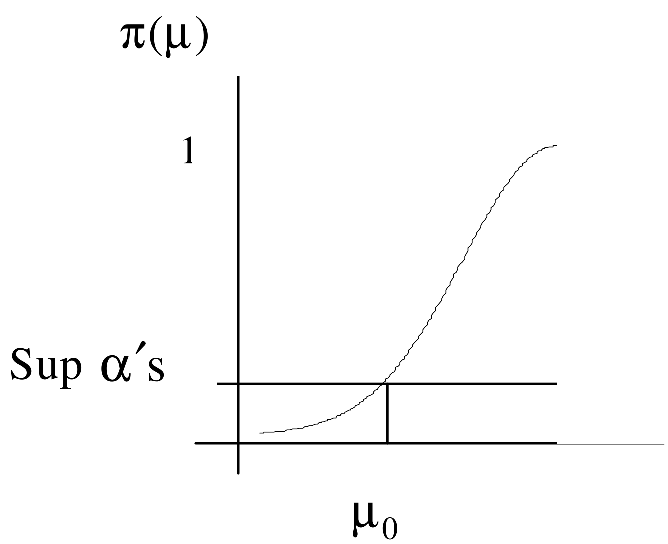
Para o exercício tipo, o gráfico correspondente à função potência é o seguinte:

RCUE

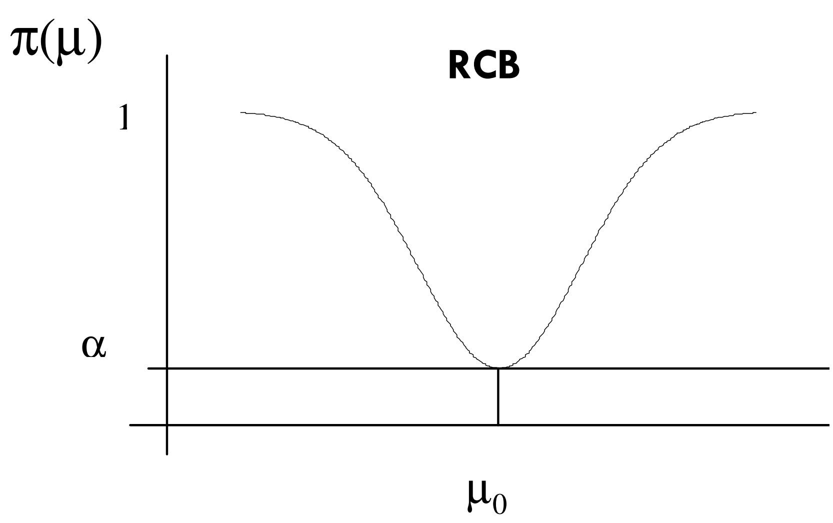


Nota: Só em ensaios de hipóteses bilaterais é que o α é único.

RCUD



RCB



NOTA: Escolha da estatística adequada ao ensaio

A escolha adequada desta estatística, bem como a correspondente distribuição amostral, a situações mais habituais é feita tendo em conta o tipo de população, o conhecimento que se tem da respectiva variância e da dimensão da amostra.
Primeiro, começa-se com ensaios relativos a uma amostra; depois, tratar-se-ão dos ensaios relativos a duas amostras independentes e não independentes (amostras emparelhadas, por exemplo); e, por fim tratar-se-ão dos ensaios relativos a mais que duas amostras (análise da variância).

A tabela usada para a construção de intervalos de confiança será a mesma que deve ser utilizada para os ensaios de hipóteses, com excepção para a diferença de proporções. Assim, num ensaio de hipóteses, a estatística adequada será

$$T = \frac{(\overline{X}_1 - \overline{X}_2) - (p_1 - p_2)_o}{\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}} = \frac{(\overline{X}_1 - \overline{X}_2) - (p_1 - p_2)_o}{\sqrt{\overline{X}_p (1 - \overline{X}_p)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \overset{\circ}{\cap} N(0;1)$$

em que $\overline{X}_p = \frac{n_1 \overline{X}_1 + n_2 \overline{X}_2}{n_1 + n_2}$ representa a média ponderada das proporções no total das duas amostras.

Amostras emparelhadas

Se estivermos perante duas amostras (não independentes) que são formadas por pares de observações feitas sobre os mesmos elementos e retiradas de populações normais, diz-se que se está perante amostras emparelhadas.

Se a hipótese a testar for a igualdade de médias entre os pares de observações, i. e.,

$$H_o : \mu_D = 0$$
$$H_a : \mu_D \neq 0$$

então a estatística adequada a este teste será:

$$\overline{D} = \frac{\sum D_i}{n} \quad \text{e} \quad s_D'^2 = \frac{\sum (D_i - \overline{D})^2}{n-1}$$

$$T = \frac{\overline{D} - \mu_D}{s_D' / \sqrt{n}} \cap t_{(n-1)}$$

Ensaio para a diferença de k médias ou análise da variância simples (ANOVA)

Considerem-se k amostras independentes de populações normais com variâncias desconhecidas embora iguais. As hipóteses a testar são

$$H_o : \mu_1 = \cdots = \mu_k$$

$$H_a : \mu_r \neq \mu_j \quad \text{para algum par } (r, j) \text{ com } r \neq j$$

É um ensaio conhecido por "análise da variância simples" já que a estatística deste teste assente na decomposição da variância total na soma de duas parcelas, a variação explicada pelo factor independente (por exemplo, o local de exposição de um produto) e a variação devida a erro (ou não explicada por esse factor).

- Variância total ou soma total de quadrados (SST)

  A soma total dos quadrados dos desvios dos valores observados em torno da média global é igual a

$$SST = \sum_{j=1}^{k} \sum_{i=1}^{n_j} (X_{ij} - \overline{X})^2$$

- Soma de quadrados dentro dos grupos (SSW) funciona como um indicador da variabilidade ou heterogeneidade existente dentro dos grupos

$$SSW = \sum_{j=1}^{k} \sum_{i=1}^{n_j} (X_{ij} - \overline{X}_j)^2$$

- Soma de quadrados entre os grupos (SSB) funciona como um indicador das diferenças entre os grupos.

$$SSB = \sum_{j=1}^{k} n_j (\overline{X}_j - \overline{X})^2$$

em que: -k é o número de grupos

-n_j é a dimensão da amostra j

- X_{ij} é a observação para o indivíduo i do grupo j
- $\overline{X}_j$ é a média amostral do grupo i
- $\overline{X}$ é a média global e todas as observações

Assim,

$$SST = SSW + SSB$$

Os respectivos graus de liberdade são

$$(n-1) = (n-k) + (k-1)$$

A estatística do teste e a respectiva distribuição são

$$T = \frac{SSB/(k-1)}{SSW/(n-k)} = \frac{MSSB}{MSSW} \cap F_{(k-1;n-k)}$$

Só faz sentido rejeitar H_o para valores elevados deste teste, valores esses que ocorrem quando a variação entre os grupos for relativamente elevada quando comparada com a variação verificada dentro dos grupos $\Leftrightarrow$ RCUD.