Introdução à Probabilidade e Estatística

Universidade de Évora

Departamento de Matemática

Ano lectivo 2015/16

Ana Isabel Santos

Aulas 11 e 12

Testes de Hipóteses

Testes de Hipóteses Paramétricos

Teste de Hipóteses

Uma **hipótese estatística** é uma conjetura sobre uma determinada caraterística da população, X.

Um **teste de hipóteses** é um procedimento estatístico que permite validar ou não determinada afirmação acerca da população, com base na informação amostral, ou seja, se os dados sustentam determinada hipótese. Em particular, nos **testes de hipóteses paramétricos** a validação diz respeito apenas aos parâmetros da população.

Num teste há sempre duas hipóteses:

Hipótese nula - H_0 *vs* Hipótese alternativa - H_1

H_0 - estabelece que não há mudança. Enquanto que, H_1 - indica que houve mudança e especifica a natureza da mudança.

Tipos de testes de hipóteses

Quando se pretende testar hipóteses relativas a um dado parâmetro, θ, aplica-se um teste unilateral ou um teste bilateral.

Se $H_0 : \theta = \theta_0 \ \ vs \ \ H_1 : \theta \neq \theta_0$, o **teste é bilateral**;

Exemplo: $H_0 : \mu = 1 \ \ vs \ \ H_1 : \mu \neq 1$

Se $H_0 : \theta \leq \theta_0 \ \ vs \ \ H_1 : \theta > \theta_0$, o **teste é unilateral direito**;

Exemplo: $H_0 : \mu \leq 3,5 \ \ vs \ \ H_1 : \mu > 3,5$

Se $H_0 : \theta \geq \theta_0 \ \ vs \ \ H_1 : \theta < \theta_0$, o **teste é unilateral esquerdo**.

Exemplo: $H_0 : \mu \geq 10 \ \ vs \ \ H_1 : \mu < 10$

Logo, num **teste bilateral** a hipótese H_1 contempla possibilidades à esquerda **e** à direita de H_0, enquanto que num **teste unilateral** apenas contempla possibilidades à direita **ou** à esquerda de H_0.

Testes de hipóteses

A forma clássica de realização de um teste de hipóteses recorre a uma estatística $T = T(X_1, X_2, \ldots, X_n)$, designada por **estatística de teste**, para a qual conhecemos a distribuição de probabilidade. Trabalhamos então com o conjunto de todos os valores particulares $t_{obs} = T(x_1, x_2, \ldots, x_n)$.

Nestas circunstâncias, um teste de hipóteses estabelece uma regra que permite determinar quando devemos rejeitar, ou não, H_0. De facto, conhecendo a distribuição da estatística de teste, de acordo com um **nível de significância** fixo, α, definimos uma **região de rejeição** (**R.R.**), ou região crítica e uma **região de aceitação** (**R.A.**). A regra de decisão é:

1. Rejeitar H_0 se $t_{obs} \in R.R.$;

2. Não rejeitar H_0 se $t_{obs} \in R.A.$.

Testes de hipóteses - Metodologia

I - Procedimento com base na região de rejeição

1. Estabelecerem as hipóteses H_0 e H_1 ;

2. Escolher uma estatística de teste - T adequada, da qual se conhece a distribuição (admitindo que H_0 é verdadeira);

3. Fixar um nível de significância, α, que representa o risco máximo que se está disposto a correr de se rejeitar incorretamente H_0.

4. Determinar as regiões de aceitação e de rejeição.

5. Calcular t_{obs} que é o valor que T assume para todos os valores observados.

6. Tomar uma decisão.

7. Concluir.

Erros nos testes de hipóteses

A decisão de rejeitar a hipótese nula é uma decisão forte e construtiva, enquanto que a decisão de não a rejeitar é fraca e não permite avançar. Isto quer dizer que se deve colocar como hipótese H_1 aquela que se tem mais interesse em provar que é válida.

Face à decisão tomada relativamente à hipótese nula podemos estar a cometer um de dois tipos de erro correspondentes às decisões erradas:

	Realidade	
Decisão do teste	H_0 **verdadeira**	H_0 **falsa**
Não rejeitar H_0	decisão correta	decisão incorreta **(Erro de tipo II)**
Rejeitar H_0	decisão incorreta **(Erro de tipo I)**	decisão correta

Erros nos testes de hipóteses

Exemplo: H_0: A pessoa é inocente *vs* H_1: A pessoa é culpada

Erro de tipo I: A pessoa é condenada, mas está inocente.

Erro de tipo II: A pessoa é absolvida, mas é culpada.

Podemos calcular a probabilidade de cometer cada um destes erros:

1. A probabilidade de cometer um erro de tipo I

$$P\left(\text{Erro de tipo I}\right) = P\left(\text{rejeitar } \mathrm{H}_0 \mid \mathrm{H}_0 \text{ verdadeira}\right) < \alpha;$$

2. A probabilidade de cometer um erro de tipo II

$$\beta(\theta_1) = P\left(\text{Erro de tipo II}\right) = P\left(\text{não rejeitar } \mathrm{H}_0 \mid \mathrm{H}_0 \text{ falsa}\right)$$

$$= P\left(\text{não rejeitar } \mathrm{H}_0 \mid \mathrm{H}_1 \text{ verdadeira}\right) = P\left(\text{não rejeitar } \mathrm{H}_0 \mid \theta = \theta_1\right) = \beta;$$

O valor de β diminui à medida que o verdadeiro valor de $\theta = \theta_1$, se afasta de θ_0, visto que neste caso é menos provável que não se detete o verdadeiro valor.

Nível de significância

Na abordagem aos testes de hipóteses é dada maior importância ao facto de se evitar cometer um erro de tipo I, isto é, *rejeitar H_0 dado que é verdadeira*.

Portanto, fixa-se inicialmente o valor do **nível de significância** α, de forma a estabelecer a probabilidade máxima de estarmos a cometer um erro de tipo I.

Na prática, os níveis de significância mais utilizados são:

0.01, **0.05** e 0.1.

Testes de hipóteses

O valor-p (**p-value**) é o menor nível de significância a partir do qual se rejeita a hipótese nula, isto é, rejeitamos H_0 para $\alpha \geq$ valor-p.

$$\text{valor-p} = P\left(\text{rejeitar } H_0 \mid H_0 \text{ verdadeira}\right).$$

Seja $T = T(X_1, X_2, \ldots, X_n)$ a estatística de teste a utilizar para testar hipóteses sobre o parâmetro θ e t_{obs} o valor observado de T.

1. $H_0: \theta \leq \theta_0$ vs $H_1: \theta > \theta_0 \Rightarrow \text{valor-p} = P(T \geq t_{obs} | H_0)$

2. $H_0: \theta \geq \theta_0$ vs $H_1: \theta < \theta_0 \Rightarrow \text{valor-p} = P(T \leq t_{obs} | H_0)$

3. $H_0: \theta = \theta_0$ vs $H_1: \theta \neq \theta_0$

se a distribuição da estatística de teste é simétrica, então

$$\text{valor-p} = 2 \times P\left(T \geq |t_{obs}| \mid H_0\right);$$

se a distribuição da estatística de teste é assimétrica, então

$$\text{valor-p} = 2 \times min\left\{P\left(T \leq |t_{obs}| \mid H_0\right),\ P\left(T \geq |t_{obs}| \mid H_0\right)\right\}.$$

Testes de hipóteses - Metodologia

II - Procedimento com base no valor-p ou p-value

1. Estabelecerem as hipóteses H_0 e H_1;

2. Escolher a estatística de teste T adequada (admitindo que H_0 é verdadeira);

3. Fixar de um nível de significância, α.

4. Determinar o p-value do teste (dado pelo software estatístico SPSS).

5. Tomar uma decisão.

6. Concluir.

Potência do teste

A **função potência de teste**, $\pi(\theta_1)$, designa a probabilidade associada à decisão correta de *rejeição de H_0 quando esta hipótese é na realidade falsa*, ou seja,

$$\begin{aligned}\pi(\theta_1) &= P\left(\text{rejeitar } H_0 \mid H_0 \text{ é falsa}\right)\\ &= P\left(\text{rejeitar } H_0 \mid \theta = \theta_1\right)\\ &= 1 - P\left(\text{não rejeitar } H_0 \mid \theta = \theta_1\right)\\ &= 1 - \beta(\theta_1).\end{aligned}$$

Teste de hipóteses para a média

Os testes de hipóteses para a média são:

$H_0 : \mu \geq \mu_0 \;\; vs \, H_1 : \mu < \mu_0 \qquad H_0 : \mu = \mu_0 \; vs \, H_1 : \mu \neq \mu_0 \qquad H_0 : \mu \leq \mu_0 \; vs \, H_1 : \mu > \mu_0$

1. Se a distribuição da população é Normal e a variância σ² for conhecida, então a estatística de teste é

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

2. Se a distribuição da população não é Normal e a variância σ² for conhecida, mas a dimensão da amostra for grande ($n \geq 30$), então a estatística de teste é

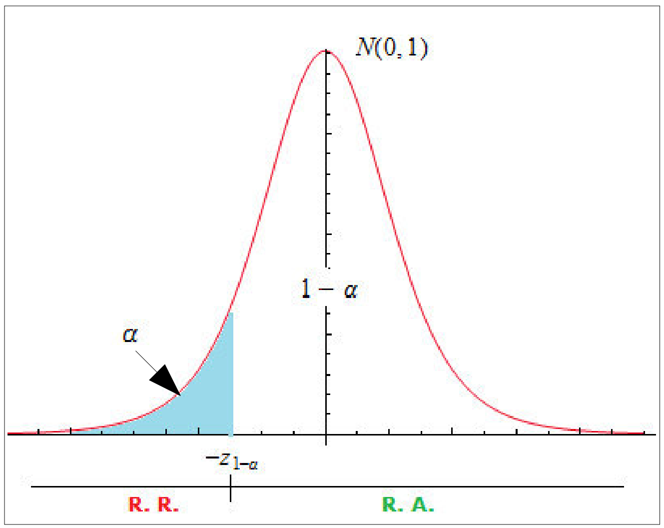
$$Z = \frac{\overline{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}} \overset{\circ}{\sim} N(0,1).$$

Teste de hipóteses para a média com σ² conhecida

Teste unilateral esquerdo

$$H_0 : \mu \geq \mu_0 \;\; vs \; H_1 : \mu < \mu_0$$

Regiões críticas:



$$\text{R.A.:}\left]-z_{1-\alpha};+\infty\right[$$

$$\text{R.R.:}\left]-\infty;-z_{1-\alpha}\right]$$

Regra de decisão:

Rejeitar H_0 se $z_{obs} \leq -z_{1-\alpha}$

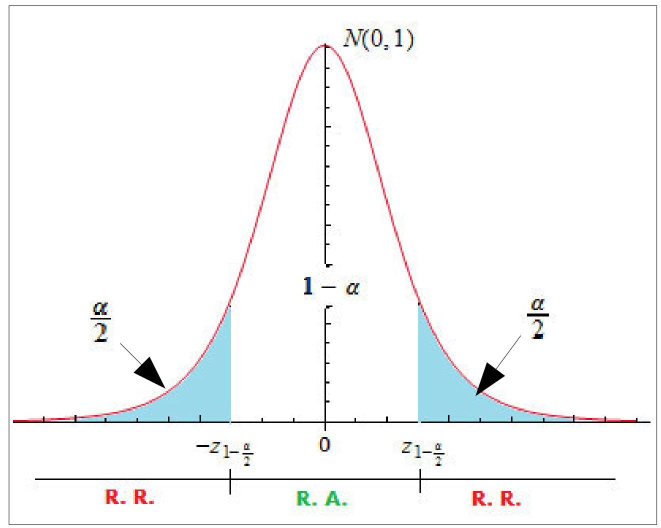
Cálculo do p-value

$$\text{p-value} = P(Z \leq z_{obs}) = \Phi(z_{obs})$$

Teste bilateral

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

Regiões críticas:



$$\text{R.A.:}\left]-z_{1-\frac{\alpha}{2}};z_{1-\frac{\alpha}{2}}\right[$$

$$\text{R.R.:}\left]-\infty;-z_{1-\frac{\alpha}{2}}\right] \cup \left[z_{1-\frac{\alpha}{2}};+\infty\right[$$

Regra de decisão:

Rejeitar H_0 se $|z_{obs}| \geq z_{1-\frac{\alpha}{2}}$

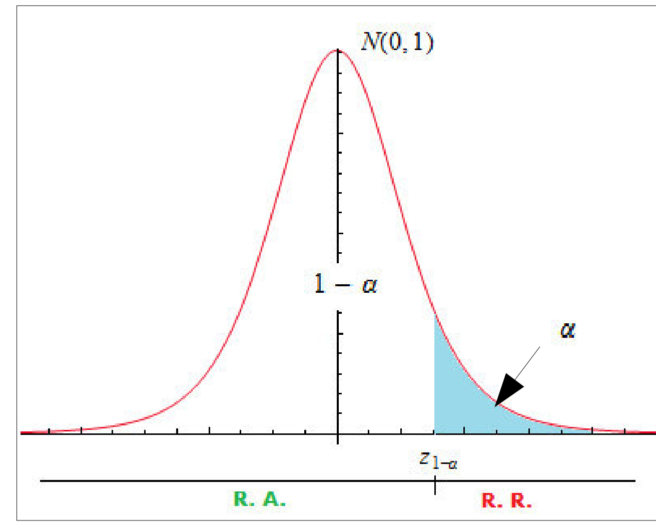
Cálculo do p-value

$$\text{p-value} = 2 \times P(Z \geq |z_{obs}|)$$
$$= 2[1 - \Phi(|z_{obs}|)]$$

Teste unilateral direito

$$H_0 : \mu \leq \mu_0 \;\; vs \; H_1 : \mu > \mu_0$$

Regiões críticas:



$$\text{R.A.:}\left]-\infty;z_{1-\alpha}\right[$$

$$\text{R.R.:}\left[z_{1-\alpha};+\infty\right[$$

Regra de decisão:

Rejeitar H_0 se $z_{obs} \geq z_{1-\alpha}$

Cálculo do p-value

$$\text{p-value} = P(Z \geq z_{obs})$$
$$= 1 - \Phi(z_{obs})$$

Teste de hipóteses para a média

3. Se a distribuição da população é Normal e a variância σ^2 for desconhecida, então a estatística de teste é

$$T = \frac{\overline{X} - \mu_0}{\frac{S}{\sqrt{n}}} \sim t_{n-1}.$$

4. Se a distribuição da população não é Normal e a variância σ^2 for desconhecida, mas a dimensão da amostra for grande ($n \geq 30$), então a estatística de teste é

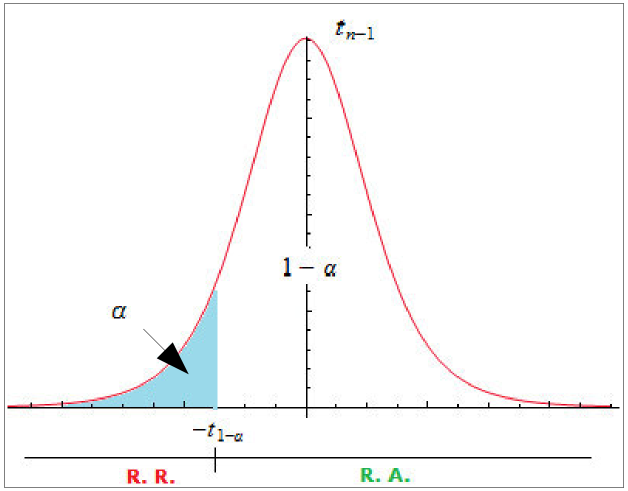
$$Z = \frac{\overline{X} - \mu_0}{\frac{S}{\sqrt{n}}} \overset{\circ}{\sim} N(0,1).$$

Teste de hipóteses para a média com σ² desconhecida

Teste unilateral esquerdo

$H_0 : \mu \geq \mu_0 \ \ vs \ H_1 : \mu < \mu_0$

Regiões críticas:



R.A.:$\left]-t_{n-1;1-\alpha};+\infty\right[$

R.R.:$\left]-\infty;-t_{n-1;1-\alpha}\right]$

Regra de decisão:

Rejeitar H_0 se $t_{obs} \leq -t_{n-1;1-\alpha}$

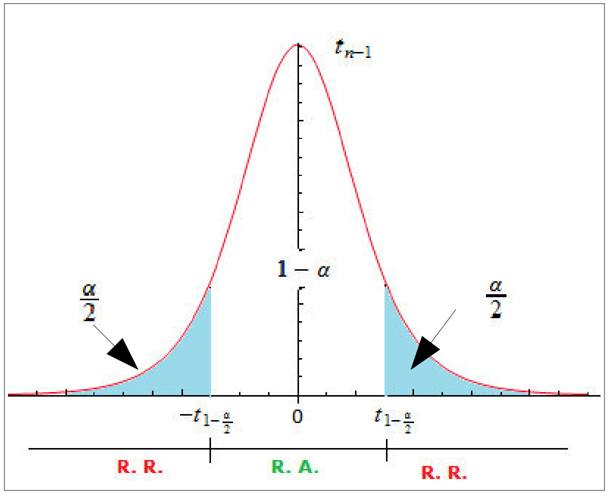
Cálculo do p-value

p-value $= P(T \leq t_{obs})$

Teste bilateral

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

Regiões críticas:



R.A.:$\left]-t_{n-1;1-\frac{\alpha}{2}};t_{n-1;1-\frac{\alpha}{2}}\right[$

R.R.:$\left]-\infty;-t_{n-1;1-\frac{\alpha}{2}}\right] \cup \left[t_{n-1;1-\frac{\alpha}{2}};+\infty\right[$

Regra de decisão:

Rejeitar H_0 se $|t_{obs}| \geq t_{n-1;1-\frac{\alpha}{2}}$

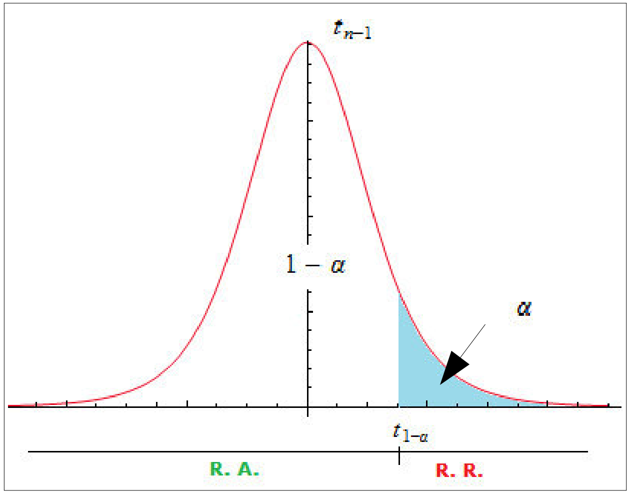
Cálculo do p-value

p-value $= 2 \times P(T \geq |t_{obs}|)$

Teste unilateral direito

$H_0 : \mu \leq \mu_0 \ vs \ H_1 : \mu > \mu_0$

Regiões críticas:



R.A.:$\left]-\infty;t_{n-1;1-\alpha}\right[$

R.R.:$\left[t_{n-1;1-\alpha};+\infty\right[$

Regra de decisão:

Rejeitar H_0 se $t_{obs} \geq t_{n-1;1-\alpha}$

Cálculo do p-value

p-value $= P(T \geq t_{obs})$

Teste de hipóteses para a diferença de médias

Os testes de hipóteses para a diferença de médias são:

$H_0 : \mu_X - \mu_Y = \mu_0 \;vs\; H_1 : \mu_X - \mu_Y \neq \mu_0$ **Teste bilateral**

$H_0 : \mu_X - \mu_Y \leq \mu_0 \;vs\; H_1 : \mu_X - \mu_Y > \mu_0$ **Teste unilateral direito**

$H_0 : \mu_X - \mu_Y \geq \mu_0 \;vs\; H_1 : \mu_X - \mu_Y < \mu_0$ **Teste unilateral esquerdo**

1. Sejam $X_1, X_2, \ldots, X_n$ e $Y_1, Y_2, \ldots, Y_m$ duas a. a. independentes, de dimensão n e m, respetivamente, retiradas de duas populações Normais $X \sim N(\mu_X; \sigma_X)$ e $Y \sim N(\mu_Y; \sigma_Y)$ com variâncias σ_X^2 e σ_Y^2 conhecidas. Então, a estatística de teste é:

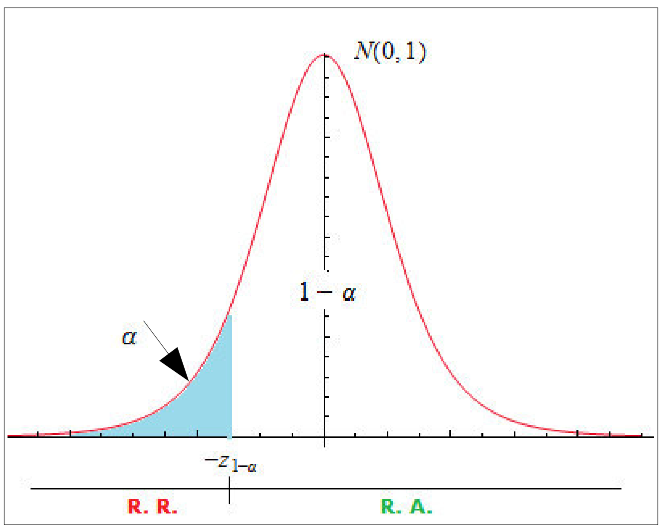
$$Z = \frac{\left(\overline{X} - \overline{Y}\right) - \mu_0}{\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}} \sim N(0;1).$$

Teste de hipóteses para a diferença de médias

Teste unilateral esquerdo

$$H_0 \ : \mu_X - \mu_Y \geq \mu_0$$
$$vs$$
$$H_1 \ : \mu_X - \mu_Y < \mu_0$$

Regiões críticas:



$$\text{R.A.:}\left]-z_{1-\alpha};+\infty\right[$$

$$\text{R.R.:}\left]-\infty;-z_{1-\alpha}\right]$$

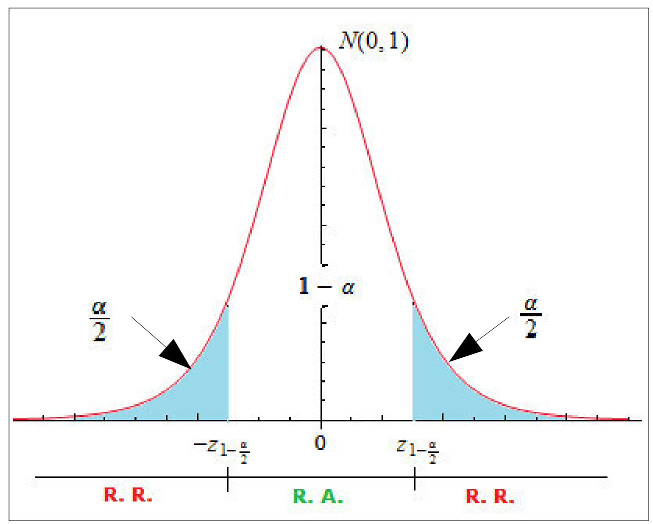
Regra de decisão:

Rejeitar H_0 se $z_{obs} \leq -z_{1-\alpha}$

Teste bilateral

$$H_0 \ : \mu_X - \mu_Y = \mu_0$$
$$vs$$
$$H_1 \ : \mu_X - \mu_Y \neq \mu_0$$

Regiões críticas:



$$\text{R.A.:}\left]-z_{1-\frac{\alpha}{2}};z_{1-\frac{\alpha}{2}}\right[$$

$$\text{R.R.:}\left]-\infty;-z_{1-\frac{\alpha}{2}}\right] \cup \left[z_{1-\frac{\alpha}{2}};+\infty\right[$$

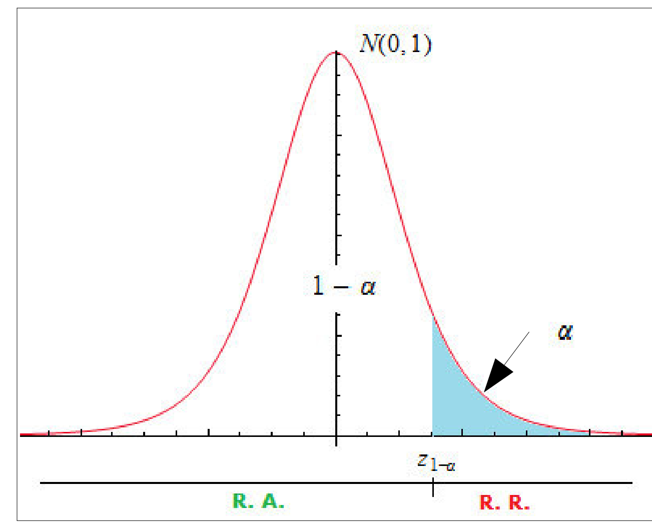
Regra de decisão:

Rejeitar H_0 se $|z_{obs}| \geq z_{1-\frac{\alpha}{2}}$

Teste unilateral direito

$$H_0 \ : \mu_X - \mu_Y \leq \mu_0$$
$$vs$$
$$H_1 \ : \mu_X - \mu_Y > \mu_0$$

Regiões críticas:



$$\text{R.A.:}\left]-\infty;z_{1-\alpha}\right[$$

$$\text{R.R.:}\left[z_{1-\alpha};+\infty\right[$$

Regra de decisão:

Rejeitar H_0 se $z_{obs} \geq z_{1-\alpha}$

Teste de hipóteses para a diferença de médias

2. Se as populações não forem Normais e as variâncias σ_X^2 e σ_Y^2 forem conhecidas, mas as amostras forem de grande dimensão ($n \geq 30$ e $m \geq 30$), então a estatística de teste é:

$$Z \overset{\circ}{\sim} N(0,1).$$

3. Sejam $X_1, X_2, \ldots, X_n$ e $Y_1, Y_2, \ldots, Y_m$ duas a. a. independentes, de dimensão n e m, respetivamente, retiradas de duas populações Normais $X \sim N(\mu_X; \sigma_X)$ e $Y \sim N(\mu_Y; \sigma_Y)$ com variâncias σ_X^2 e σ_Y^2 desconhecidas mas iguais. Então, a estatística de teste é

$$T = \frac{\left(\overline{X} - \overline{Y}\right) - \mu_0}{\sqrt{\frac{(n-1)S_X^2 + (m-1)S_Y^2}{n+m-2}}\sqrt{\frac{1}{n} + \frac{1}{m}}} \sim t_{n+m-2}$$

Teste de hipóteses para a diferença de médias

Teste bilateral

$$H_0 : \mu_X - \mu_Y = \mu_0$$
$$vs$$
$$H_1 : \mu_X - \mu_Y \neq \mu_0$$

Regiões críticas:

$$\text{R.A.:}\left]-t_{n+m-2;1-\frac{\alpha}{2}} ; t_{n+m-2;1-\frac{\alpha}{2}}\right[$$

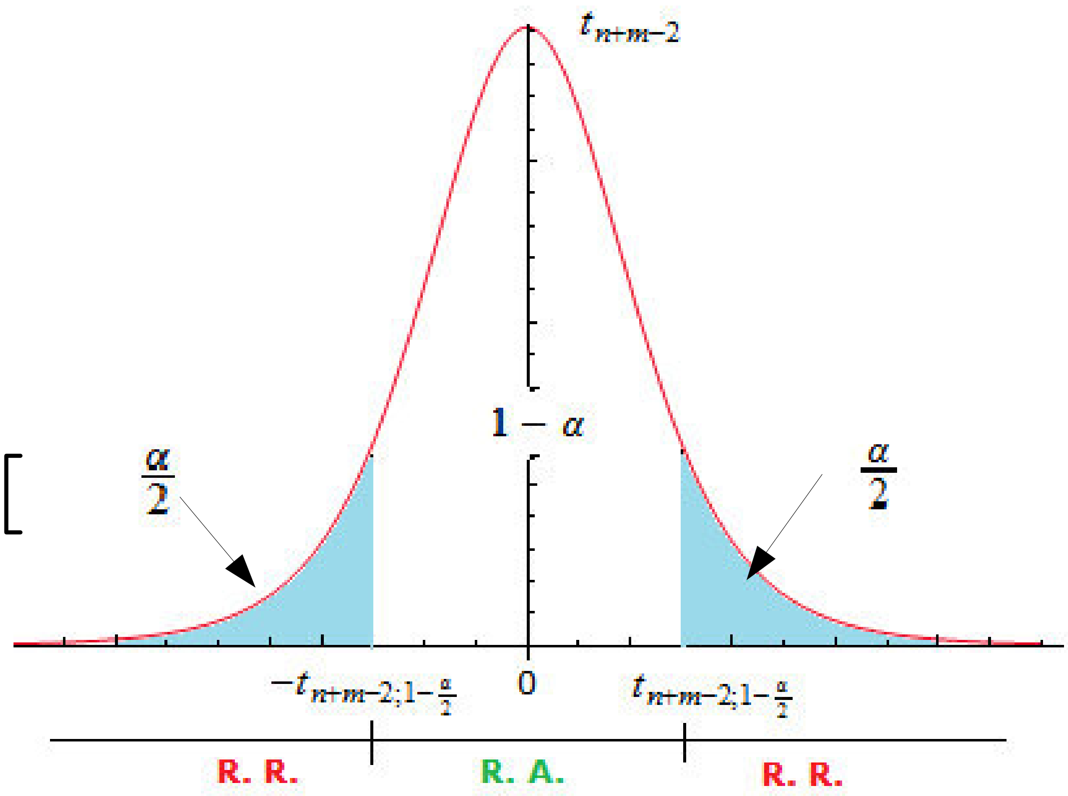
$$\text{R.R.:}\left]-\infty ; -t_{n+m-2;1-\frac{\alpha}{2}}\right] \cup \left[t_{n+m-2;1-\frac{\alpha}{2}} ; +\infty\right[$$

Regra de decisão:

Rejeitar H_0 se $|t_{obs}| \geq t_{n+m-2;1-\frac{\alpha}{2}}$

Cálculo do p-value

$$\text{p-value} = 2 \times P(T \geq |t_{obs}|)$$



Teste de hipóteses para a diferença de médias

Teste unilateral direito

$$H_0 : \mu_X - \mu_Y \leq \mu_0$$
$$vs$$
$$H_1 : \mu_X - \mu_Y > \mu_0$$

Regiões críticas:

$$\text{R.A.:} \left] -\infty ; t_{n+m-2;1-\alpha} \right[$$

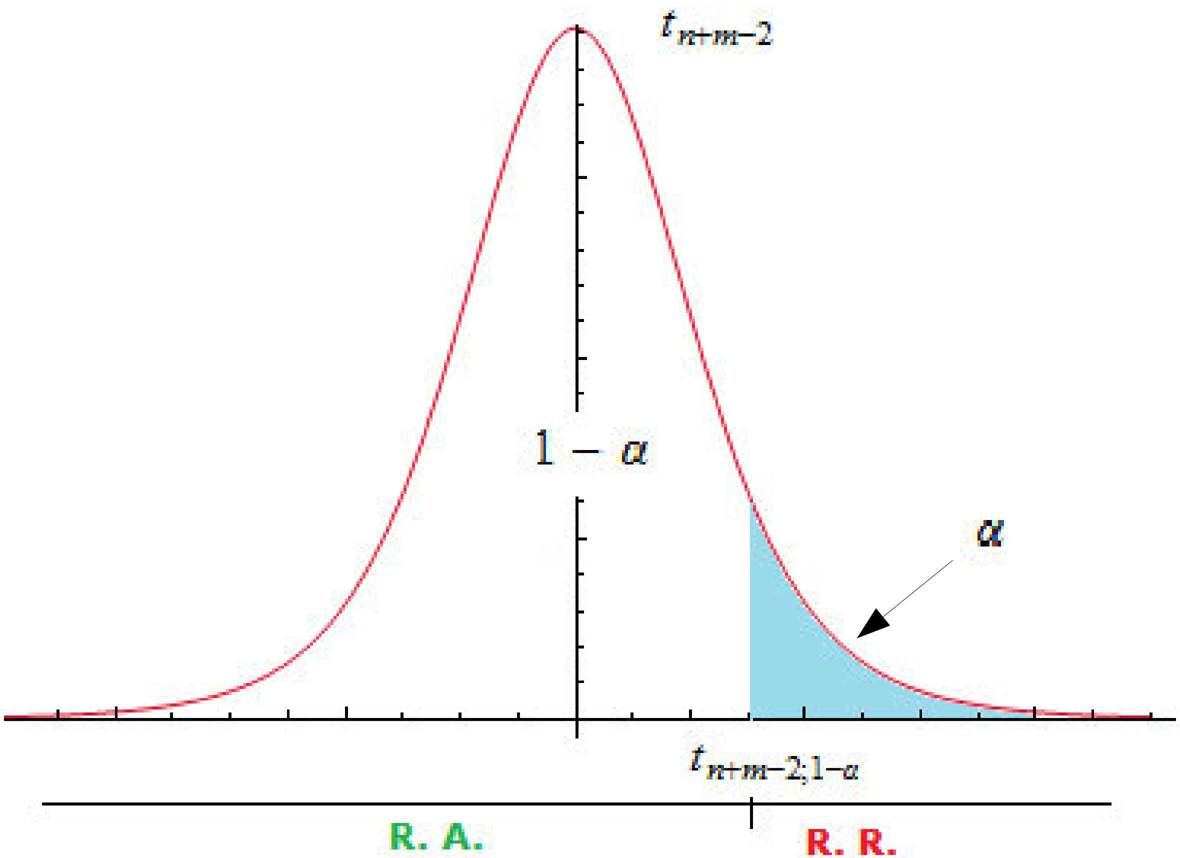
$$\text{R.R.:} [t_{n+m-2;1-\alpha} ; +\infty[$$

Regra de decisão:

Rejeitar H_0 se $t_{obs} \geq t_{n+m-2;1-\alpha}$

Cálculo do p-value

$$\text{p-value} = P(T \geq t_{obs})$$



Teste de hipóteses para a diferença de médias

Teste unilateral esquerdo

$$H_0 : \mu_X - \mu_Y \geq \mu_0$$
$$vs$$
$$H_1 : \mu_X - \mu_Y < \mu_0$$

Regiões críticas:

$$\text{R.A.:} \left] -t_{n+m-2;1-\alpha}; +\infty \right[$$

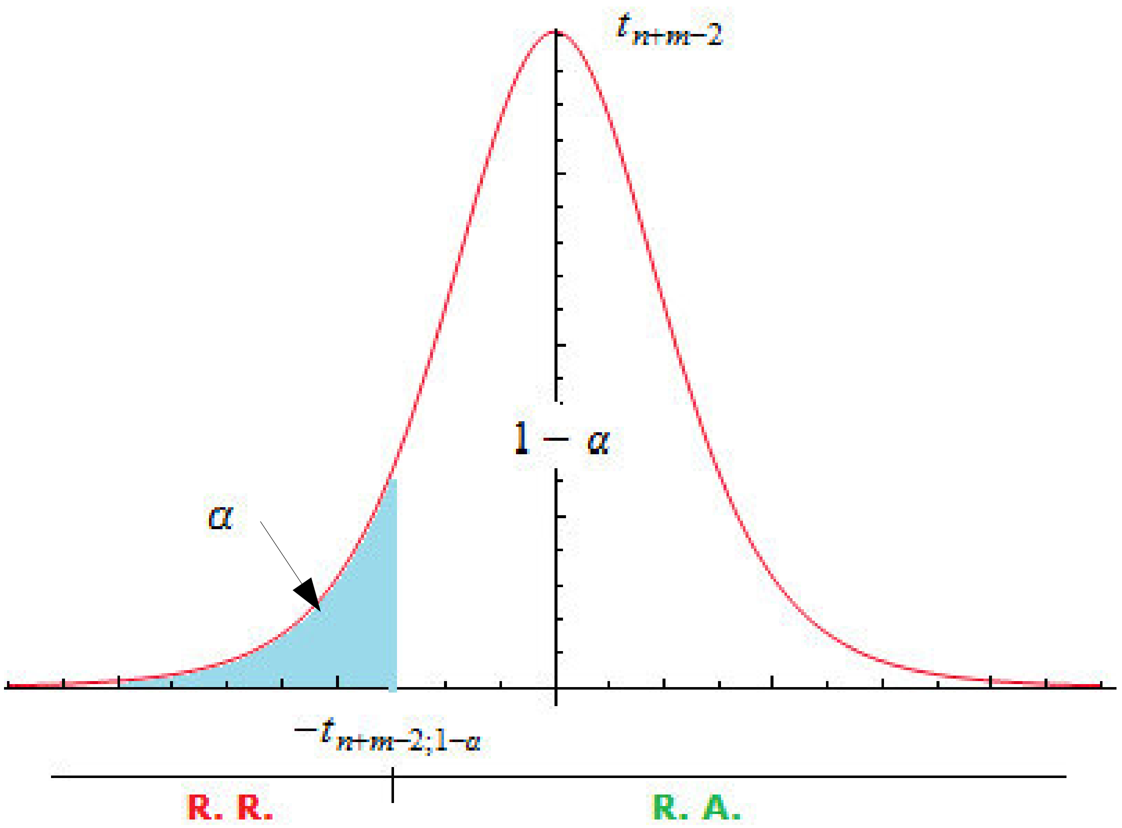
$$\text{R.R.:} \left] -\infty; -t_{n+m-2;1-\alpha} \right]$$

Regra de decisão:

Rejeitar H_0 se $t_{obs} \leq -t_{n+m-2;1-\alpha}$

Cálculo do p-value

$$\text{p-value} = P(T \leq t_{obs})$$



Teste de hipóteses para a diferença de médias

4. Se as populações não forem Normais e as variâncias σ_X^2 e σ_Y^2 forem desconhecidas mas iguais $(\sigma_X^2 = \sigma_Y^2)$ e as amostras forem de grande dimensão ($n \geq 30$ e $m \geq 30$), então a estatística de teste é

$$Z = \frac{\left(\overline{X} - \overline{Y}\right) - \mu_0}{\sqrt{\dfrac{(n-1)S_X^2 + (m-1)S_Y^2}{n+m-2}}\sqrt{\dfrac{1}{n} + \dfrac{1}{m}}} \overset{\circ}{\sim} N(0,1).$$

Teste de hipóteses para amostras emparelhadas

Consideremos, agora, o caso em que temos duas amostras aleatórias de igual dimensão $(X_{11}, X_{12}, \ldots, X_{1n})$ e $(X_{21}, X_{22}, \ldots, X_{2n})$ que formam uma amostra emparelhada constituída por n pares, (X_{1i}, X_{2i}), $i = 1, 2, \ldots, n$. Cada par é reduzido a um único valor, construindo-se uma nova variável

$$D_i = X_{1i} - X_{2i}, \quad i = 1, 2, \ldots, n,$$

cuja média e desvio-padrão amostrais são, respetivamente,

$$\overline{D} = \overline{X}_1 - \overline{X}_2 = \sum_{i=1}^{n} \frac{D_i}{n}, \qquad S_D^2 = \sum_{i=1}^{n} \frac{\left(D_i - \overline{D}\right)^2}{n-1}.$$

Teste de hipóteses para amostras emparelhadas

Assim, as hipóteses a testar podem ser escritas da seguinte forma:

$H_0 : \mu_D = \mu_0$ vs $H_1 : \mu_D \neq \mu_0$ **Teste bilateral**

$H_0 : \mu_D \leq \mu_0$ vs $H_1 : \mu_D > \mu_0$ **Teste unilateral direito**

$H_0 : \mu_D \geq \mu_0$ vs $H_1 : \mu_D < \mu_0$ **Teste unilateral esquerdo**

O que significa que estamos perante um teste de hipóteses para a média no caso em que a população segue uma distribuição Normal da qual se desconhece a variância, pelo que a estatística de teste será

$$T = \frac{\overline{D} - \mu_0}{\frac{s_D}{\sqrt{n}}} \sim t_{n-1}.$$