

BIOESTATÍSTICA

Parte 5 – Testes de Hipóteses

Aulas Teóricas de 05/05/2011 a 19/05/2011

5.1. Conceito de erro, estatística de teste, região de rejeição, nível de significância, valor de prova, potência do teste

[testes de hipóteses põem em confronto 2 hipóteses ou afirmações]

Objectivo de um teste de hipóteses:

determinar se uma hipótese ou conjectura que fazemos acerca de um parâmetro de uma população é plausível, i.e., tem razão de ser, com base na informação obtida a partir de uma amostra extraída dessa população.

Ex: admite-se que a quantidade de nicotina (em mg) existente nos cigarros de certa marca tem distribuição normal. O fabricante afirma que a quantidade média de nicotina, por cigarro, é de 13.5 mg. A DECO desconfia que a quantidade média de nicotina é superior a este valor.

Hipóteses a testar:

$$\mu = 13.5 \quad \text{versus} \quad \mu > 13.5 \\ \text{(contra)}$$

Para verificar se as suspeitas têm razão de ser, são recolhidos aleatoriamente vários cigarros dessa marca e regista-se a quantidade de nicotina em cada um. A decisão será tomada com base no valor de uma estatística que é função da média desses valores.

Hipóteses a testar:

$$H_0: \mu = 13.5 \quad \text{vs} \quad H_1: \mu > 13.5$$

Hipótese nula (H_0): corresponde frequentemente ao estado actual, ao que é tradicionalmente aceite. Reflete a situação em que não há mudança, sendo pois uma hipótese conservadora.

Hipótese alternativa (H_1): corresponde a uma situação em que existe uma alteração face ao que é habitual; exprime, por exemplo, aquilo que um investigador está a tentar estabelecer com um novo estudo sobre o assunto.

Notação
alternativa – H_a

[A hipótese nula tem de ter uma igualdade, ou seja pode ser estabelecida recorrendo a =, (hipótese simples) a \leq ou a \geq (hipóteses compostas). Na disciplina só vai ser abordado o caso em que a hipótese nula é simples ($\mu = k$)]

Com base na amostra, pretende-se tomar uma decisão:

- rejeitar H_0
- não rejeitar H_0 (i.e. manter H_0)

Rejeita-se H_0 , em favor da hipótese alternativa H_1 , apenas quando existe evidência significativa para tal, fornecida pelos dados.

[não rejeitar H_0 não implica provar que H_0 é verdadeira, porque continua-se a não conhecer o valor do parâmetro μ ; ausência de evidência de uma coisa não é evidência de outra]

Estatística de teste: é uma variável aleatória, função apenas da amostra, com base na qual será tomada a decisão de rejeitar ou não a hipótese nula. No caso de H_0 ser verdadeira, a sua distribuição é conhecida, pois não depende de parâmetros desconhecidos.

Ao tomar uma decisão, podemos cometer dois tipos de erros:

Erro de tipo I: rejeitar H_0 se H_0 é verdadeira Ou erro de 1ª espécie

Erro de tipo II: não rejeitar H_0 se H_1 é verdadeira

		Situação real	
		H_0 verdadeira	H_1 verdadeira
Decisão tomada	Não rejeitar H_0	Decisão correcta	Erro tipo II
	Rejeitar H_0	Erro tipo I	Decisão correcta

Nível de significância do teste:

$$\alpha = P(\text{rejeitar } H_0 \mid H_0 \text{ é verdadeira})$$

Os valores usuais para α são 0.01, 0.05 ou 0.1.

↓
É o valor mais típico dentro dos mais usuais

Fixado o nível de significância α e dada uma amostra proveniente da população, obtemos o valor observado da estatística de teste e comparamo-lo com o quantil adequado da distribuição, sob a validade de H_0 , dessa estatística.

Valor de Prova ou Valor-p

Ao realizar um teste de hipóteses, podemos não especificar à partida qual o nível de significância. Uma abordagem alternativa consiste em calcular a seguinte probabilidade, designada por valor-p:

Valor-p é a probabilidade de obter um valor tão ou mais extremo do que o valor observado da estatística de teste, supondo H_0 verdadeira.

Quanto menor for esta probabilidade, mais forte é a evidência contra a hipótese nula.

Valor-p é o menor valor do nível de significância que leva à rejeição da hipótese nula, para a amostra recolhida.

Na prática, se o valor-p for superior a 0.1 (ou até superior a 0.05), não se rejeita H_0 , considera-se que não há evidência contra a hipótese nula.

Se foi fixado o nível de significância do teste, podemos também optar por calcular o valor-p. Então, rejeita-se H_0 ao nível α se $\text{valor-p} \leq \alpha$.

[Exemplo: se o valor-p fosse 0,25, rejeitávamos H_0 para níveis de significância maiores que 0,25 (1%, 5%, 10%, etc)]

Testes de Hipóteses para o Valor médio (uma população)

É dada uma amostra de dimensão n proveniente de uma certa população.

1) População normal, com σ conhecido

Estatística de teste:

$$Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} \cap N(0,1) \quad \text{se } H_0 \text{ verdadeira}$$

\bar{X} - média amostral

n - dimensão da amostra

Fixado o nível de significância α ,

H_0	H_1	Rejeitar H_0 se
$\mu = \mu_0$	$\mu > \mu_0$	$z \geq z_{1-\alpha}$
$\mu = \mu_0$	$\mu < \mu_0$	$z \leq z_{\alpha}$
$\mu = \mu_0$	$\mu \neq \mu_0$	$ z \geq z_{1-\alpha/2}$

z representa o valor observado da estatística de teste Z

2) População normal, com σ desconhecido

Teste t

Estatística de teste:

$$T = \frac{\bar{X} - \mu_0}{S / \sqrt{n}} \cap t_{(n-1)} \quad \text{se } H_0 \text{ verdadeira}$$

$t_{(n-1)}$ designa a distribuição t-Student com $n-1$ graus de liberdade

Fixado o nível de significância α ,

H_0	H_1	Rejeitar H_0 se
$\mu = \mu_0$	$\mu > \mu_0$	$t \geq t_{n-1, 1-\alpha}$
$\mu = \mu_0$	$\mu < \mu_0$	$t \leq t_{n-1, \alpha}$
$\mu = \mu_0$	$\mu \neq \mu_0$	$ t \geq t_{n-1, 1-\alpha/2}$

t representa o valor observado da estatística de teste T

Foram apresentados os resultados de uma investigação sobre o nível de hemoglobina (em g/dl) numa dada população. Com base nos dados, referentes a 20 indivíduos seleccionados aleatoriamente, obteve-se

$$\bar{x} = 15.33 \quad \text{e} \quad s^2 = 0.382$$

Teste a hipótese de que o nível médio de hemoglobina desta população é igual a 15, assumindo que o nível de hemoglobina segue uma distribuição normal. Considere $\alpha = 0.05$.

$$H_0: \mu = 15 \quad \text{vs} \quad H_1: \mu \neq 15$$

[Outro enunciado para a mesma situação: Verifique se o nível médio de hemoglobina é diferente de 15. Nunca se pode perguntar directamente se o nível médio de hemoglobina é 15.]

Sob a validade de H_0 , $T = \frac{\bar{X} - \mu_0}{S / \sqrt{n}} \cap t_{(n-1)}$

valor observado da estatística de teste :

$$t = \frac{15.33 - 15}{0.618 / \sqrt{20}} = 2.388$$

Rejeitar H_0 se $|t| \geq t_{19,0.975} \quad t_{19,0.975} = 2.093$

Logo, rejeitamos H_0 ao nível $\alpha=5\%$. Há evidência de que o nível médio de hemoglobina é diferente de 15.

3) População não normal, σ conhecido, grande amostra ($n \geq 30$)

Estatística de teste:

$$Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} \cap N(0,1) \quad \text{se } H_0 \text{ verdadeira}$$

4) População não normal, σ desconhecido, grande amostra ($n \geq 30$)

Estatística de teste:

$$Z = \frac{\bar{X} - \mu_0}{S / \sqrt{n}} \cap N(0,1) \quad \text{se } H_0 \text{ verdadeira}$$

Fixado o nível de significância α ,

H_0	H_1	Rejeitar H_0 se
$\mu = \mu_0$	$\mu > \mu_0$	$z \geq z_{1-\alpha}$
$\mu = \mu_0$	$\mu < \mu_0$	$z \leq z_{\alpha}$
$\mu = \mu_0$	$\mu \neq \mu_0$	$ z \geq z_{1-\alpha/2}$

Cálculo do valor-p

- População normal, com σ conhecido
- População não normal, grande amostra ($n \geq 30$), σ conhecido ou desconhecido

$$1) H_0: \mu = \mu_0 \quad \text{vs} \quad H_1: \mu > \mu_0$$

$$p = P(Z \geq z) = 1 - \Phi(z)$$

$$2) H_0: \mu = \mu_0 \text{ vs } H_1: \mu < \mu_0$$

$$p = P(Z \leq z) = \Phi(z)$$

$$3) H_0: \mu = \mu_0 \text{ vs } H_1: \mu \neq \mu_0$$

$$p = 2P(Z > |z|) = 2(1 - \Phi(|z|))$$

$Z \sim N(0,1)$ e z representa o valor observado da estatística de teste

Exemplo:

Observou-se uma amostra de 200 ampolas de determinado tipo e mediu-se o seu comprimento (em cm), tendo-se obtido

$$\bar{x} = 5.975 \text{ e } s = 0.36$$

Verifique se existe evidência para afirmar que o comprimento médio das ampolas é inferior a 6 cm.

$$H_0: \mu = 6 \text{ vs } H_1: \mu < 6$$

População não normal, σ desconhecido, grande amostra ($n=200$)

$$\text{Sob } H_0 \quad Z = \frac{\bar{X} - \mu_0}{S / \sqrt{n}} \overset{a}{\sim} N(0,1)$$

$$\text{valor observado da estatística de teste: } z = \frac{5.975 - 6}{0.36 / \sqrt{200}} = -0.98$$

$$\alpha = 5\% \quad \text{Rejeitar } H_0 \text{ se } z \leq z_{0.05} \quad z_{0.05} = -1.645$$

$$\text{valor -p} = P(Z \leq -0.98) = \Phi(-0.98) = 1 - 0.83646 = 0.16354$$

Não rejeitamos H_0 para $\alpha=5\%$. O valor-p é elevado, não há evidência para afirmar que o comprimento médio das ampolas é inferior a 6 cm.

Teste de Hipóteses para a variância (uma população)

É dada uma amostra de dimensão n proveniente de uma população normal, com μ desconhecido.

$$\text{Estatística de teste: } Q = \frac{(n-1)S^2}{\sigma_0^2} \sim \chi_{n-1}^2 \text{ se } H_0 \text{ verdadeira}$$

Fixado o nível de significância α ,

H_0	H_1	Rejeitar H_0 se
$\sigma^2 = \sigma_0^2$	$\sigma^2 > \sigma_0^2$	$q \geq \chi_{n-1, 1-\alpha}^2$
$\sigma^2 = \sigma_0^2$	$\sigma^2 < \sigma_0^2$	$q \leq \chi_{n-1, \alpha}^2$
$\sigma^2 = \sigma_0^2$	$\sigma^2 \neq \sigma_0^2$	$q \leq \chi_{n-1, \alpha/2}^2$ ou $q \geq \chi_{n-1, 1-\alpha/2}^2$

q representa o valor observado da estatística de teste Q

Inferência estatística sobre a diferença entre os valores médios de duas populações que há evidência

Os valores seguintes dizem respeito a duas formulações, A e B, de um certo fármaco e representam a concentração máxima (C_{\max}) do fármaco (em ng/ml) atingida na circulação sistémica a seguir à administração, para dois grupos de 24 indivíduos

Formulação A	
125	126
128	117
131	120
119	126
130	140
121	121
129	126
133	136
135	135
123	133
120	127
125	122

Formulação B	
130	149
141	132
133	142
129	136
128	151
145	141
136	130
150	138
127	130
132	129
138	122
148	140

Formulação A

média 127
desvio padrão 6.14

Formulação B

média 136.54
desvio padrão 8.09

Haverá evidência de que, em média, a concentração máxima do fármaco na formulação B é superior à concentração máxima na formulação A?

[O que é equivalente a perguntar se a média da concentração máxima das duas formulações é suficientemente diferente para se considerar que há evidência que o valor médio da concentração do fármaco B é realmente superior à do fármaco A]

Geralmente, ao fazer um estudo comparativo, utiliza-se o termo estatístico **tratamento** para designar as características que estão a ser comparadas.

O planeamento da experiência pode, geralmente, produzir:

- amostras independentes
- amostra emparelhada (pares de observações)

Amostras independentes:

Os indivíduos ou unidades experimentais que vão ser expostos a cada tratamento são escolhidos aleatoriamente e distribuídos por dois grupos, atribuindo-se a um dos grupos o tratamento 1 e ao outro o tratamento 2. Os resultados obtidos em cada grupo de tratamento constituem uma amostra de uma certa população, pelo que falamos em comparação de duas populações.

Frequentemente, tem-se um grupo experimental que recebe o tratamento sob investigação e um grupo de controlo, que recebe um placebo ou o tratamento tradicional.

Amostra emparelhada:

Os indivíduos ou unidades experimentais são escolhidos aos pares, de modo que os membros de cada par sejam semelhantes (por ex: quanto ao sexo, idade, história clínica); cada elemento do par é então exposto a um dos tratamentos.

A situação ideal em experimentação biomédica seria dispor de gémeos idênticos.

O emparelhamento mais simples consiste em usar cada unidade experimental como o seu próprio controlo, i.e., proceder-se a duas medições, uma sob a situação experimental e a outra sob a situação de controlo.

[Os pares de gémeos assegurariam que a diferença entre os resultados se devem única e exclusivamente à variável estudada e não a outras variabilidades individuais]

Amostras Independentes:

Consideremos duas amostras aleatórias independentes

- uma amostra de dimensão n_1 proveniente da população X_1 , de valor médio μ_1 e desvio padrão σ_1
- uma amostra de dimensão n_2 proveniente da população X_2 , de valor médio μ_2 e desvio padrão σ_2

Os dois “tratamentos” têm o mesmo efeito médio? Pretendemos então fazer inferência sobre $\mu_1 - \mu_2$

Um estimador pontual de $\mu_1 - \mu_2$ é $\bar{X}_1 - \bar{X}_2$

$$E(\bar{X}_1 - \bar{X}_2) = \mu_1 - \mu_2$$

$$\text{var}(\bar{X}_1 - \bar{X}_2) = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} \quad \text{visto que } \bar{X}_1 \text{ e } \bar{X}_2 \text{ são v.a.'s independentes}$$

[Note-se que como de costume $H_0: \mu_1 = \mu_2$, o que é equivalente a $H_0: \mu_1 - \mu_2 = 0$, o que é o objectivo do teste]

1. Populações normais e variâncias conhecidas

Como a combinação linear de v.a.'s independentes com distribuição normal ainda tem distribuição normal, vem que

$$\frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \cap N(0,1)$$

Estatística de teste: $Z = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \cap N(0,1)$ se H_0 verdadeira

H_0	H_1	Ao nível α , rejeitar H_0 se
$\mu_1 - \mu_2 = 0$	$\mu_1 - \mu_2 > 0$	$z \geq z_{1-\alpha}$
$\mu_1 - \mu_2 = 0$	$\mu_1 - \mu_2 < 0$	$z \leq z_{\alpha}$
$\mu_1 - \mu_2 = 0$	$\mu_1 - \mu_2 \neq 0$	$ z \geq z_{1-\alpha/2}$

z representa o valor observado da estatística de teste Z

Note-se que, habitualmente, as hipóteses nula e alternativa costumam ser escritas na seguinte forma:

H_0	H_1
$\mu_1 = \mu_2$	$\mu_1 > \mu_2$
$\mu_1 = \mu_2$	$\mu_1 < \mu_2$
$\mu_1 = \mu_2$	$\mu_1 \neq \mu_2$

[Muito Importante: na hipótese alternativa não trocar a ordem de μ_1 com μ_2 , têm de estar pela mesma ordem de H_0]

Intervalo de 100(1-α)% de confiança para $\mu_1 - \mu_2$

$$\left(\bar{x}_1 - \bar{x}_2 - z_{1-\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}, \bar{x}_1 - \bar{x}_2 + z_{1-\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \right)$$

[Se o intervalo de confiança contém o zero é possível que $\mu_1 = \mu_2$ (contudo não temos garantias que assim o seja). Mas se o intervalo não conter o valor zero, com 100%(1-α) de confiança pode-se concluir que $\mu_1 \neq \mu_2$]

[Variâncias Desconhecidas]

Para fazer testes de hipóteses para a diferença de valores médios entre duas populações com variâncias populacionais desconhecidas é necessário saber se estas variâncias são iguais ou diferentes entre si. Este dado pode ser conhecido à partida ou não. Neste último caso é necessário realizar um teste de hipóteses para a variância, em que:

$$H_0: \sigma_1^2 = \sigma_2^2 \text{ vs } H_1: \sigma_1^2 \neq \sigma_2^2$$

(a desenvolver mais à frente neste capítulo)]

2. Populações normais e variâncias desconhecidas e iguais

teste t para amostras independentes

Como $\sigma_1^2 = \sigma_2^2 = \sigma^2$ é desconhecida, é necessário estimar esta variância comum. O estimador utilizado é a média ponderada das variâncias amostrais (*pooled variance*):

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

[A estimação de S_p^2 é uma espécie de média ponderada, pois dá maior peso à variância da amostra com maior dimensão]

Então

$$\frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \cap t_{(n_1+n_2-2)}$$

Quando as variâncias populacionais são iguais ($\sigma_1^2 = \sigma_2^2$) diz-se que as populações são **homocedásticas** ou que existe **homocedasticidade**.

Estatística de teste:
$$T = \frac{\bar{X}_1 - \bar{X}_2}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \cap t_{(n_1+n_2-2)} \quad \text{se } H_0 \text{ verdadeira}$$

H_0	H_1	Ao nível α , rejeitar H_0 se
$\mu_1 = \mu_2$	$\mu_1 > \mu_2$	$t \geq t_{1-\alpha}$
$\mu_1 = \mu_2$	$\mu_1 < \mu_2$	$t \leq t_{\alpha}$
$\mu_1 = \mu_2$	$\mu_1 \neq \mu_2$	$ t \geq t_{1-\alpha/2}$

T é uma v.a. com distribuição t-Student com (n_1+n_2-2) graus de liberdade
 t_{α} designa o quantil de probabilidade α dessa distribuição
 t representa o valor observado da estatística de teste

Intervalo de 100(1- α)% de confiança para $\mu_1 - \mu_2$

$$\left(\bar{x}_1 - \bar{x}_2 - t_{1-\alpha/2} S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}, \bar{x}_1 - \bar{x}_2 + t_{1-\alpha/2} S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \right)$$

3. Populações normais e variâncias desconhecidas e diferentes

Para investigar a igualdade das variâncias de uma forma rigorosa deve-se usar o teste F (apenas no caso de populações normais) para testar

$$H_0 : \sigma_1^2 = \sigma_2^2 \quad \text{vs} \quad H_1 : \sigma_1^2 \neq \sigma_2^2$$

Caso seja rejeitada a hipótese nula, no caso de populações normais e pequenas amostras podemos usar a aproximação de Welch-Satterthwaite para testar $H_0 : \mu_1 = \mu_2$ contra a hipótese alternativa conveniente.

[Ou seja não foi dado o teste de hipóteses para a diferença de valores médios para duas populações normais, com variâncias desconhecidas e diferentes]

4. Populações não normais e amostras de grande dimensão ($n_1 \geq 30$, $n_2 \geq 30$)

(i) Se as variâncias forem conhecidas

$$\text{Estatística de teste: } \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \overset{a}{\cap} N(0,1) \quad \text{se } H_0 \text{ verdadeira}$$

A inferência estatística sobre $\mu_1 - \mu_2$ é idêntica à descrita para populações normais.

(ii) Se as variâncias forem desconhecidas (iguais ou diferentes)

Intervalo de $100(1-\alpha)\%$ de confiança para $\mu_1 - \mu_2$

$$\left(\bar{x}_1 - \bar{x}_2 - z_{1-\alpha/2} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}, \bar{x}_1 - \bar{x}_2 + z_{1-\alpha/2} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \right)$$

$$\text{Estatística de teste: } Z = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \overset{a}{\cap} N(0,1) \quad \text{se } H_0 \text{ verdadeira}$$

H_0	H_1	Ao nível α , rejeitar H_0 se
$\mu_1 = \mu_2$	$\mu_1 > \mu_2$	$z \geq z_{1-\alpha}$
$\mu_1 = \mu_2$	$\mu_1 < \mu_2$	$z \leq z_\alpha$
$\mu_1 = \mu_2$	$\mu_1 \neq \mu_2$	$ z \geq z_{1-\alpha/2}$

z representa o valor observado da estatística de teste Z

Amostras Emparelhadas:

Ao comparar dois tratamentos, seria desejável que os indivíduos sujeitos a esses tratamentos fossem tão semelhantes quanto possível, de modo que as diferenças entre os resultados dos dois grupos pudessem ser atribuídas apenas às diferenças entre os tratamentos.

Por outro lado, a imposição de que todos os indivíduos sejam semelhantes é demasiado restritiva e pode impedir a recolha de amostras de dimensão suficiente para determinadas análises estatísticas.

- A. Uma opção consiste em escolher os indivíduos por pares, de modo que dentro de cada par tenham características semelhantes e sejam diferentes, para pares distintos. Depois de seleccionados os pares, aplica-se a cada elemento do par o tratamento 1 ou 2 de forma aleatória.
- B. Outra opção é que o mesmo indivíduo seja sujeito aos dois “tratamentos” e forneça portanto um par de valores correspondentes aos resultados dos tratamentos 1 e 2.

Exemplos:

A. Foi realizado um estudo destinado a comparar a percentagem de retenção de partículas na traqueia e nos brônquios de fumadores e não fumadores. Para tal foi medida a percentagem de retenção em 6 pares de gémeos idênticos, em que um é fumador e o outro não fumador:

Fumador:	60.6	12.0	56.0	75.2	12.5	29.7
Não fumador:	47.5	13.3	33.0	55.2	21.9	27.9

Os dados sustentam a hipótese de que a percentagem de retenção é maior nos fumadores?

B. Sete indivíduos saudáveis ofereceram-se para realizar uma experiência envolvendo uma nova droga. Os sete indivíduos submeteram-se, antes e depois de ingerir a tal droga, a um exame em que foi medido (em centésimos de segundo) o tempo de resposta a um sinal sonoro. Os resultados desse exame encontram-se na tabela seguinte:

Tempo de resposta depois:	17	27	39	27	30	21	36
Tempo de resposta antes:	19	22	34	21	27	24	29

Verifique se existe evidência para afirmar que a nova droga provoca o aumento do tempo de resposta a estímulos auditivos.

A amostra obtida é constituída por n pares de observações

$$(x_1, y_1), \dots, (x_n, y_n)$$

mas em vez de realizar a análise estatística fazendo uso dessas observações individuais, considera-se a amostra das diferenças entre cada par de valores

$$(d_1, \dots, d_n) \text{ em que } d_i = x_i - y_i$$

Para a amostra das diferenças (d_1, \dots, d_n) defina-se

$$\bar{d} = \frac{1}{n} \sum_{i=1}^n d_i \quad s_d^2 = \frac{1}{n-1} \sum_{i=1}^n (d_i - \bar{d})^2$$

Representemos por X_1 e X_2 as respostas ou resultados dos “tratamentos” 1 e 2. Seja $D = X_1 - X_2$.

A inferência estatística sobre a diferença $\mu_1 - \mu_2$ entre os valores médios dos resultados devidos aos dois tratamentos coincide com a inferência sobre $\mu = \mu_1 - \mu_2$, seguindo-se os processos descritos anteriormente para fazer inferência sobre o valor médio de uma população.

1. Populações normais

Supondo que as populações são normais, tem-se que $D = X_1 - X_2$ é uma população normal de valor médio $\mu = \mu_1 - \mu_2$ e desvio padrão σ_D desconhecido.

Note-se que, para que os métodos de inferência seguintes sejam válidos, basta que as diferenças tenham uma distribuição normal. Nas aplicações práticas, ocorre frequentemente que a amostra das diferenças pode ser considerada como proveniente de uma população normal.

Intervalo de 100(1- α)% de confiança para $\mu = \mu_1 - \mu_2$

$$\left(\bar{d} - t_{n-1, 1-\alpha/2} \frac{s_d}{\sqrt{n}}, \bar{d} + t_{n-1, 1-\alpha/2} \frac{s_d}{\sqrt{n}} \right)$$

Teste t emparelhado (t-paired test)

Estatística de teste:

$$T = \frac{\bar{D}}{S_D / \sqrt{n}} \cap t_{(n-1)} \quad \text{se } H_0 \text{ verdadeira}$$

H_0	H_1	Ao nível α , rejeitar H_0 se
$\mu_1 = \mu_2$	$\mu_1 > \mu_2$	$t \geq t_{n-1, 1-\alpha}$
$\mu_1 = \mu_2$	$\mu_1 < \mu_2$	$t \leq t_{n-1, \alpha}$
$\mu_1 = \mu_2$	$\mu_1 \neq \mu_2$	$ t \geq t_{n-1, 1-\alpha/2}$

T é uma v.a. com distribuição t-Student com (n-1) graus de liberdade
 t representa o valor observado da estatística de teste

2. Populações não normais, amostra de grande dimensão (n≥30)

Intervalo de 100(1-α)% de confiança para $\mu = \mu_1 - \mu_2$

$$\left(\bar{d} - z_{1-\alpha/2} \frac{s_d}{\sqrt{n}}, \bar{d} + z_{1-\alpha/2} \frac{s_d}{\sqrt{n}} \right)$$

Testes de hipóteses sobre $\mu = \mu_1 - \mu_2$

Estatística de teste: $Z = \frac{\bar{D}}{S_D / \sqrt{n}} \overset{a}{\sim} N(0,1)$ se H_0 verdadeira

H_0	H_1	Ao nível α , rejeitar H_0 se
$\mu_1 = \mu_2$	$\mu_1 > \mu_2$	$z \geq z_{1-\alpha}$
$\mu_1 = \mu_2$	$\mu_1 < \mu_2$	$z \leq z_{\alpha}$
$\mu_1 = \mu_2$	$\mu_1 \neq \mu_2$	$ z \geq z_{1-\alpha/2}$

Testes de Hipóteses para a igualdade de variâncias

Em [2] populações normais

Consideremos duas amostras aleatórias independentes

- uma amostra de dimensão n_1 proveniente da população X_1 , normal de valor médio μ_1 e desvio padrão σ_1
- uma amostra de dimensão n_2 proveniente da população X_2 , normal de valor médio μ_2 e desvio padrão σ_2

em que μ_1, μ_2, σ_1 e σ_2 são parâmetros desconhecidos

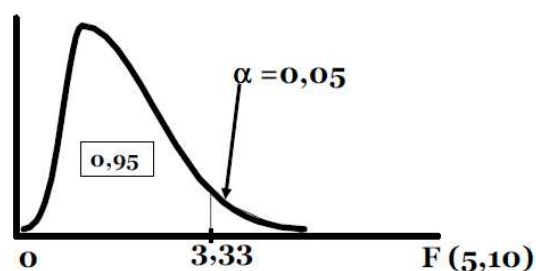
$$F = \frac{S_1^2 / \sigma_1^2}{S_2^2 / \sigma_2^2}$$

tem uma distribuição F com (n_1-1) graus de liberdade no numerador e (n_2-1) graus de liberdade no denominador e que se representa por $F(n_1-1; n_2-1)$

[É um método muito pouco robusto: só funciona para populações Normais. No caso de outros tipos de populações recorre-se a outro tipo de testes]

[Distribuição F – Snedecor:

- Semelhante à Distribuição do qui-Quadrado:
 - Assimétrica;
 - Só toma valores positivos;
- Possui dois parâmetros:
 - Graus de liberdade do numerador;
 - Graus de liberdade do denominador.]



Obs: $F_{\alpha}(m;n) = \frac{1}{F_{1-\alpha}(n;m)}$ onde F_{α} designa o quantil de probabilidade α

Exemplo: $F_{0,05}(5,12) = \frac{1}{F_{0,05}(12,5)}$

Supondo verdadeira a hipótese nula $H_0: \sigma_1^2 = \sigma_2^2$

a estatística de teste $F = \frac{S_1^2}{S_2^2}$ tem distribuição $F(n_1-1; n_2-1)$

[note-se mesmo que a hipótese nula se verifique ($\sigma_1 = \sigma_2$), $F \neq 1$, porque sendo as amostras diferentes $s_1^2 \neq s_2^2$]

H_0	H_1	Ao nível α , rejeitar H_0 se
$\sigma_1^2 = \sigma_2^2$	$\sigma_1^2 > \sigma_2^2$	$f \geq F_{1-\alpha}(n_1-1; n_2-1)$
$\sigma_1^2 = \sigma_2^2$	$\sigma_1^2 < \sigma_2^2$	$f \leq F_{\alpha}(n_1-1; n_2-1)$
$\sigma_1^2 = \sigma_2^2$	$\sigma_1^2 \neq \sigma_2^2$	$f \geq F_{1-\alpha/2}(n_1-1; n_2-1)$ ou $f \leq F_{\alpha/2}(n_1-1; n_2-1)$

F_{α} designa o quantil de probabilidade α da distribuição $F(n_1-1; n_2-1)$
 f representa o valor observado da estatística de teste F

[A última hipótese alternativa é usada com pré-teste antes de realizar um teste de hipóteses para o valor médio, para determinar se as variâncias desconhecidas são iguais ou não]

Resumo dos Intervalos de Confiança

Uma população:

<u>Intervalos a 100(1-α)% de confiança para o valor médio, μ</u>	
Populações Normais:	Amostra de dimensão "grande":
<ul style="list-style-type: none"> σ conhecido $\left(\bar{x} \pm z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}} \right)$	<ul style="list-style-type: none"> σ conhecido $\left(\bar{x} \pm z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}} \right)$
<ul style="list-style-type: none"> σ desconhecido $\left(\bar{x} \pm t_{n-1; 1-\alpha/2} \frac{s}{\sqrt{n}} \right)$	<ul style="list-style-type: none"> σ desconhecido $\left(\bar{x} \pm z_{1-\alpha/2} \frac{s}{\sqrt{n}} \right)$
<u>Intervalos a 100(1-α)% de confiança para a variância populacional, σ², em populações Normais</u>	
<ul style="list-style-type: none"> μ conhecido $\left(\frac{\sum_{i=1}^n (x_i - \mu)^2}{\chi_{n-1; 1-\alpha/2}^2}, \frac{\sum_{i=1}^n (x_i - \mu)^2}{\chi_{n-1; \alpha/2}^2} \right)$	<ul style="list-style-type: none"> μ desconhecido $\left(\frac{(n-1)s^2}{\chi_{n-1; 1-\alpha/2}^2}, \frac{(n-1)s^2}{\chi_{n-1; \alpha/2}^2} \right)$

Duas Populações:

<u>Intervalos a 100(1-α)% de confiança para a diferença de valores médios, μ₁-μ₂</u>	
Populações Normais:	Amostras de dimensão "grande":
<ul style="list-style-type: none"> σ₁ e σ₂ conhecidos $\left(\bar{x}_1 - \bar{x}_2 \pm z_{1-\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \right)$	<ul style="list-style-type: none"> σ₁ e σ₂ conhecidos $\left(\bar{x}_1 - \bar{x}_2 \pm z_{1-\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \right)$
<ul style="list-style-type: none"> σ₁ = σ₂ desconhecidos $\left(\bar{x}_1 - \bar{x}_2 \pm t_{n_1+n_2-2; 1-\alpha/2} s_{12} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \right)$ <p>com</p> $s_{12} = \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1 + n_2 - 2}}$	<ul style="list-style-type: none"> σ₁ e σ₂ desconhecidos $\left(\bar{x}_1 - \bar{x}_2 \pm z_{1-\alpha/2} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \right)$

Resumo dos testes de Hipóteses:

Testes de Hipóteses			
Teste	Estatística de Teste	Hipóteses	Região de rejeição
Sobre o valor médio μ	Populações Normais ou Amostra de dimensão "grande": σ conhecido $Z = \frac{\bar{X} - \mu_0}{\sigma} \sqrt{n} \underset{n_0}{\sim} N(0,1)$	1. $H_0: \mu = \mu_0$ vs. $H_1: \mu \neq \mu_0$ 2. $H_0: \mu = \mu_0$ vs. $H_1: \mu > \mu_0$ 3. $H_0: \mu = \mu_0$ vs. $H_1: \mu < \mu_0$	1. $z_0 \leq -z_{1-\alpha/2}$ ou $z_0 \geq z_{1-\alpha/2}$ 2. $z_0 \geq z_{1-\alpha}$ 3. $z_0 \leq -z_{1-\alpha}$
	Populações Normais: σ desconhecido $T = \frac{\bar{X} - \mu_0}{S} \sqrt{n} \underset{n_0}{\sim} t_{(n-1)}$	1. $H_0: \mu = \mu_0$ vs. $H_1: \mu \neq \mu_0$ 2. $H_0: \mu = \mu_0$ vs. $H_1: \mu > \mu_0$ 3. $H_0: \mu = \mu_0$ vs. $H_1: \mu < \mu_0$	1. $t_0 \leq -t_{\alpha/2, n-1}$ ou $t_0 \geq t_{\alpha/2, n-1}$ 2. $t_0 \geq t_{1-\alpha, n-1}$ 3. $t_0 \leq -t_{1-\alpha, n-1}$
	Amostra de dimensão "grande": σ desconhecido $Z = \frac{\bar{X} - \mu_0}{S} \sqrt{n} \underset{n_0}{\sim} N(0,1)$	1. $H_0: \mu = \mu_0$ vs. $H_1: \mu \neq \mu_0$ 2. $H_0: \mu = \mu_0$ vs. $H_1: \mu > \mu_0$ 3. $H_0: \mu = \mu_0$ vs. $H_1: \mu < \mu_0$	1. $z_0 \leq -z_{1-\alpha/2}$ ou $z_0 \geq z_{1-\alpha/2}$ 2. $z_0 \geq z_{1-\alpha}$ 3. $z_0 \leq -z_{1-\alpha}$
Sobre a variância populacional, σ^2	Populações Normais μ conhecido $\chi^2 = \frac{(X_i - \mu)^2}{\sigma_0^2} \underset{n_0}{\sim} \chi_{(n)}^2$	1. $H_0: \sigma^2 = \sigma_0^2$ vs. $H_1: \sigma^2 \neq \sigma_0^2$ 2. $H_0: \sigma^2 = \sigma_0^2$ vs. $H_1: \sigma^2 > \sigma_0^2$ 3. $H_0: \sigma^2 = \sigma_0^2$ vs. $H_1: \sigma^2 < \sigma_0^2$	1. $\chi_0^2 \leq \chi_{\alpha/2, n}^2$ ou $\chi_0^2 \geq \chi_{1-\alpha/2, n}^2$ 2. $\chi_0^2 \geq \chi_{1-\alpha, n}^2$ 3. $\chi_0^2 \leq \chi_{\alpha, n}^2$
	Populações Normais μ desconhecido $Q = \frac{(n-1)S^2}{\sigma_0^2} \underset{n_0}{\sim} \chi_{(n-1)}^2$	1. $H_0: \sigma^2 = \sigma_0^2$ vs. $H_1: \sigma^2 \neq \sigma_0^2$ 2. $H_0: \sigma^2 = \sigma_0^2$ vs. $H_1: \sigma^2 > \sigma_0^2$ 3. $H_0: \sigma^2 = \sigma_0^2$ vs. $H_1: \sigma^2 < \sigma_0^2$	1. $q_0 \leq \chi_{\alpha/2, n-1}^2$ ou $q_0 \geq \chi_{1-\alpha/2, n-1}^2$ 2. $q_0 \geq \chi_{1-\alpha, n-1}^2$ 3. $q_0 \leq \chi_{\alpha, n-1}^2$
Sobre a diferença de valores médios, $\mu_1 - \mu_2$	Populações Normais ou Amostras de dimensão "grande": σ_1 e σ_2 conhecidos $Z = \frac{\bar{X}_1 - \bar{X}_2 - D_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \underset{n_0}{\sim} N(0,1)$	1. $H_0: \mu_1 - \mu_2 = D_0$ vs. $H_1: \mu_1 - \mu_2 \neq D_0$ 2. $H_0: \mu_1 - \mu_2 = D_0$ vs. $H_1: \mu_1 - \mu_2 > D_0$ 3. $H_0: \mu_1 - \mu_2 = D_0$ vs. $H_1: \mu_1 - \mu_2 < D_0$	1. $z_0 \leq -z_{1-\alpha/2}$ ou $z_0 \geq z_{1-\alpha/2}$ 2. $z_0 \geq z_{1-\alpha}$ 3. $z_0 \leq -z_{1-\alpha}$
	Populações Normais $\sigma_1 = \sigma_2$ desconhecidos $T = \frac{\bar{X}_1 - \bar{X}_2 - D_0}{S_D \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \underset{n_0}{\sim} t_{(n_1+n_2-2)}$ com: $S_D = \sqrt{\frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1 + n_2 - 2}}$	1. $H_0: \mu_1 - \mu_2 = D_0$ vs. $H_1: \mu_1 - \mu_2 \neq D_0$ 2. $H_0: \mu_1 - \mu_2 = D_0$ vs. $H_1: \mu_1 - \mu_2 > D_0$ 3. $H_0: \mu_1 - \mu_2 = D_0$ vs. $H_1: \mu_1 - \mu_2 < D_0$	1. $t_0 \leq -t_{\alpha/2, n_1+n_2-2}$ ou $t_0 \geq t_{\alpha/2, n_1+n_2-2}$ 2. $t_0 \geq t_{1-\alpha, n_1+n_2-2}$ 3. $t_0 \leq -t_{1-\alpha, n_1+n_2-2}$
	Amostras de dimensão "grande": σ_1 e σ_2 desconhecidos $Z = \frac{\bar{X}_1 - \bar{X}_2 - D_0}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \underset{n_0}{\sim} N(0,1)$	1. $H_0: \mu_1 - \mu_2 = D_0$ vs. $H_1: \mu_1 - \mu_2 \neq D_0$ 2. $H_0: \mu_1 - \mu_2 = D_0$ vs. $H_1: \mu_1 - \mu_2 > D_0$ 3. $H_0: \mu_1 - \mu_2 = D_0$ vs. $H_1: \mu_1 - \mu_2 < D_0$	1. $z_0 \leq -z_{1-\alpha/2}$ ou $z_0 \geq z_{1-\alpha/2}$ 2. $z_0 \geq z_{1-\alpha}$ 3. $z_0 \leq -z_{1-\alpha}$

Análise da Variância (ANOVA)

Objectivo: comparar os valores médios de k populações, sob o pressuposto de que são populações normais e homocedásticas (variâncias iguais)

As k populações resultam da classificação dos indivíduos de uma população global, dando origem a k grupos (ou estratos).

- Usa-se quando existem 3 ou mais grupos;
 - A comparação destes grupos dois a dois, através da realização de vários testes t , aumenta imenso (na ordem dos 90%) a probabilidade de erro tipo 1, pelo que se recorre a um método específico - ANOVA;
- É válido para populações normais ou (aproximadamente) para grandes amostras;
- É um método *robusto* (também funciona para populações razoavelmente simétricas)
- Com amostras pequenas de populações não normais recorrem-se a métodos não paramétricos;
- Inclui vários métodos, dos quais o mais simples é a Análise de Variância Simples.

Quando existe apenas um critério (designado por **factor**) que serve de base à definição das populações



Análise da variância simples (one-way ANOVA)

[factor ou critério caracteriza os indivíduos da amostra]

Exemplo: Ensaio clínico de 3 fármacos (A, B e C) antipiréticos para testar a sua eficácia numa dada população de indivíduos com febre.

População global: indivíduos com febre;

Variável tratamento: fármaco A, B e C;

- 3 estratos (subpopulações) consoante o fármaco administrado;
- 1 único factor: tipo de fármaco

Variável resposta: números de horas necessárias para os indivíduos atingirem a temperatura normal.]

Cada uma das populações corresponde a um nível do factor, que é uma variável qualitativa. É a variável “tratamento”.
Qual o efeito do “tratamento” na variável resposta?

A análise da variância simples é uma generalização do teste t para amostras independentes.

μ_i é o valor médio da população i

$$X_{ij} = \mu_i + \varepsilon_{ij} \quad i = 1, \dots, k \quad j = 1, \dots, n_i$$

são i.i.d com distribuição $N(0, \sigma)$

valor da variável resposta para o indivíduo j do grupo i

$$X_{ij} = \mu + \alpha_i + \varepsilon_{ij}$$

valor médio global
erro
efeito do nível i do factor

As hipóteses a testar são:

$$H_0 : \mu_1 = \mu_2 = \dots = \mu_k \quad \text{vs} \quad H_1 : \text{nem todos os } \mu_j \text{ são iguais}$$

Uma vez que admitimos igualdade das variâncias, então, se a hipótese nula for verdadeira, a população global tem distribuição normal com os mesmos parâmetros que cada um dos k grupos.

	Factor				
	Nível 1	Nível 2	...	Nível k	
	X_{11}	X_{21}		X_{k1}	
	X_{12}	X_{22}		X_{k2}	
	
	X_{1,n_1}	X_{2,n_2}		X_{k,n_k}	
Médias em cada grupo	\bar{X}_1	\bar{X}_2	...	\bar{X}_k	$\bar{\bar{X}}$ média global

$$N = \sum_{i=1}^k n_i \quad \text{dimensão da amostra conjunta}$$

[No exemplo anterior:

Nível 1 – fármaco A

Nível 2 – fármaco B

Nível 3 – fármaco B]

Soma de quadrados dentro dos grupos e entre grupos

$$WSS = \sum_{j=1}^{n_1} (X_{1j} - \bar{X}_1)^2 + \dots + \sum_{j=1}^{n_k} (X_{kj} - \bar{X}_k)^2 \quad \text{Within groups Sum of Squares}$$

$$BSS = \sum_{i=1}^k n_i (\bar{X}_i - \bar{\bar{X}})^2 \quad \text{Between groups Sum of Squares}$$

$$TSS = \sum_{j=1}^{n_1} (X_{1j} - \bar{\bar{X}})^2 + \dots + \sum_{j=1}^{n_k} (X_{kj} - \bar{\bar{X}})^2 \quad \text{Total Sum of Squares}$$

[soma de quadrados \Rightarrow medida de variabilidade entre amostras e dentro das amostras]

$$TSS = BSS + WSS$$

Soma de Quadrados Total =
= Soma de Quadrados entre grupos + Soma de Quadrados dentro dos grupos

Para facilitar os cálculos, podem ser usadas as seguintes expressões:

$$\begin{aligned} I &= \sum \sum x_{ij}^2 \\ II &= \frac{(\sum \sum x_{ij})^2}{N} \\ III &= \frac{(\sum x_{i1})^2}{n_1} + \dots + \frac{(\sum x_{ik})^2}{n_k} \end{aligned}$$

Então tem-se que

$$\begin{aligned} BSS &= III - II \\ WSS &= I - III \\ TSS &= I - II \end{aligned}$$

Tabela ANOVA

Graus de liberdade

Fonte de variação	Soma de quadrados	g.l.	Soma de quadrados média	Estatística F
Entre grupos	BSS	k-1	$MBS = \frac{BSS}{k-1}$	$F = \frac{MBS}{MWS}$
Dentro dos grupos	WSS	N-k	$MWS = \frac{WSS}{N-k}$	
Total	TSS	N-1		

K – número de grupos

N – amostra global

A estatística de teste é

$$F = \frac{MBS}{MWS} = \frac{BSS/(k-1)}{WSS/(N-k)} \sim F(k-1, N-k) \quad \text{se } H_0 \text{ verdadeira}$$

Ao nível de significância α , rejeitar H_0 se $f \geq F_{1-\alpha}(k-1, N-k)$

onde $F_{1-\alpha}(k-1, N-k)$ é o quantil de probabilidade $1-\alpha$ da distribuição $F(k-1, N-k)$ e f representa o valor observado da estatística F

[Quanto maior F , maior a evidência que algum dos valores médios é diferente, ou seja, maior a evidência contra H_0]

Exemplo:

Durante a produção de um determinado comprimido revestido, são recolhidos periodicamente comprimidos provenientes de linhas de produção localizadas em 3 fábricas diferentes, que usam o mesmo processo.

São registados os pesos (em mg) de 15 comprimidos de cada fábrica. Haverá diferença significativa entre os pesos médios dos comprimidos produzidos em cada fábrica?

Fábrica A		Fábrica B		Fábrica C	
277,3	278,4	271,6	275,5	275,5	272,3
280,3	272,9	274,8	274,0	274,2	273,4
279,1	274,7	271,2	274,9	267,5	275,1
275,2	276,8	277,6	269,2	274,2	273,7
273,6	269,1	274,5	283,2	270,5	268,7
276,7	276,3	275,7	280,6	284,4	275,0
281,7	273,1	276,1	274,6	275,6	268,3
278,7		275,9		277,1	

Admita que os dados provêm de populações normais e homocedásticas (variâncias iguais).

$$H_0: \mu_1 = \mu_2 = \mu_3 \quad \text{vs} \quad H_1: \text{nem todos os } \mu_j \text{ são iguais}$$

Desvio
Padrão

				95% Confidence Interval for Mean	
	N	Mean	Std. Deviation	Lower Bound	Upper Bound
A	15	276,260	3,2651	274,452	278,068
B	15	275,293	3,4634	273,375	277,211
C	15	273,700	4,1571	271,398	276,002
Total	45	275,084	3,7211	273,966	276,202

ANOVA

Peso_comprimidos

	Sum of Squares	df	Mean Square	F	P-value
Between Groups	50,134	2	25,067	1,883	,165
Within Groups	559,125	42	13,313		
Total	609,259	44			

Ao nível $\alpha=5\%$, não rejeitamos H_0 . Não há evidência de que haja diferenças entre os pesos médios dos comprimidos produzidos em cada fábrica.