

Inferência para a Distribuição Normal Multivariada: parte 2

Prof. Caio Azevedo

Suposições e dados

- Novamente, seja $\mathbf{X} = (X_1, \dots, X_p)' \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, em que

$$\boldsymbol{\mu} = (\mu_1, \dots, \mu_p)' \text{ e } \boldsymbol{\Sigma} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \dots & \sigma_{1p} \\ \sigma_{12} & \sigma_2^2 & \dots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{1p} & \sigma_{2p} & \dots & \sigma_p^2 \end{bmatrix}.$$

- Considere uma amostra aleatória (a.a.) de tamanho n do vetor \mathbf{X} , ou seja, $\mathbf{X}_1, \dots, \mathbf{X}_n \stackrel{i.i.d.}{\sim} N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$

- Matriz de dados: $\mathbf{X}_{(n \times p)} = \begin{bmatrix} X_{11} & X_{12} & \dots & X_{1p} \\ X_{21} & X_{22} & \dots & X_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ X_{n1} & X_{n2} & \dots & X_{np} \end{bmatrix}$

Distribuição t de Student não-central ([link](#))

- Defina $T = \frac{Z+\mu}{\sqrt{V/\nu}}$, em que $Z \perp V$, $Z \sim N(0, 1)$ e $V \sim \chi_{(\nu)}^2$. Dizemos que T tem distribuição t de student não central, com ν graus de liberdade e parâmetro de não centralidade μ .
- Uma forma de se apresentar a densidade é:

$$f_T(t) = \frac{1}{2^{\frac{\nu-1}{2}} \sqrt{\pi\nu} \Gamma(\nu/2)} \exp\left\{\frac{-\nu\mu^2}{\nu+t^2}\right\} \left(\frac{\nu}{\nu+t^2}\right)^{(\nu-1)/2} \\ \times \int_0^\infty y^\nu \exp\left\{-\frac{1}{2}\left(y - \frac{\mu t}{\sqrt{t^2+\nu}}\right)^2\right\} dy$$

- Se $\mu = 0$, temos a distribuição t de Student central, ou seja $T \sim t_{(\nu)}$.

Distribuição qui-quadrado não central ([link](#))

- Sejam $X_1, X_2, \dots, X_n \stackrel{ind}{\sim} N(\mu_i, \sigma_i^2)$.
- Defina $Y = \sum_{i=1}^n \left(\frac{X_i}{\sigma_i}\right)^2$. Dizemos então que Y tem distribuição qui-quadrado não central com n graus de liberdade e parâmetro de não-centralidade $\delta = \sum_{i=1}^n \left(\frac{\mu_i}{\sigma_i}\right)^2$
- Notação $Y \sim \chi_{(n, \delta)}^2$, cuja fdp é dada por

$$f_Y(y) = \sum_{i=0}^{\infty} \frac{e^{-\delta/2} (\delta/2)^i}{i!} f_{W_{n+2i}}(y) \mathbb{1}_{(0, \infty)}(y),$$

em que $W_{n+2i} \sim \chi_{(n+2i)}^2$

- Se $\delta = 0$, então $Y \sim \chi_{(n)}^2$.

Distribuição F não central ([link](#))

- Seja V uma outra v.a., independente de Y , $V \sim \chi^2_{(m)}$.
- Defina $F = \frac{Y/n}{V/m}$. Então, F tem distribuição F não central com graus de liberdade, n e m e parâmetro de não centralidade δ .
- Notação $F \sim \chi^2_{(n,m,\delta)}$, cuja fdp é dada por

$$f_F(f) = \sum_{i=0}^{\infty} \frac{e^{-\delta/2} (\delta/2)^i}{\beta(m/2, n/2 + i) i!} \frac{f^{n/2-1+i}}{(1+f)^{i+(n+m)/2}} \times \mathbb{1}_{(0,\infty)}(f)$$

em que $\beta(a, b) = \int_0^1 x^{a-1} (1-x)^{b-1} dx$

No programa R

- As distribuições t de Student, qui-quadrado e F não centrais estão disponíveis no R, através das funções (conhecidas):
 - t de Student: `*t`.
 - Qui-quadrado: `*chisq`.
 - F: `*f`.

em que `*` =d,p,q,r.

Principais teoremas de formas quadráticas normais

- Seja $\mathbf{Y} \sim N_p(\mathbf{0}, \mathbf{I})$ e \mathbf{A} uma matriz simétrica. Então

$$\mathbf{Y}'\mathbf{A}\mathbf{Y} \sim \chi^2_{[r(\mathbf{A})]},$$

se, e somente se \mathbf{A} for idempotente.

- Seja $\mathbf{Y} \sim N_p(\mathbf{0}, \Sigma)$ e \mathbf{A} uma matriz simétrica. Então

$$\mathbf{Y}'\mathbf{A}\mathbf{Y} \sim \chi^2_{[r(\mathbf{A})]},$$

se, e somente se $\mathbf{A}\Sigma$ for idempotente.

Cont.

- Seja $\mathbf{Y} \sim N_p(\boldsymbol{\mu}, \mathbf{I})$ e \mathbf{A} uma matriz simétrica. Então

$$\mathbf{Y}'\mathbf{A}\mathbf{Y} \sim \chi^2_{[r(\mathbf{A}), \delta=\boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}]},$$

se, e somente se \mathbf{A} for idempotente.

- Seja $\mathbf{Y} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ e \mathbf{A} uma matriz simétrica. Então

$$\mathbf{Y}'\mathbf{A}\mathbf{Y} \sim \chi^2_{[r(\mathbf{A}), \delta=\boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}]},$$

se, e somente se $\mathbf{A}\boldsymbol{\Sigma}$ for idempotente.

- Naturalmente, se $\delta = \boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu} = 0$, as distribuições passam a ser qui-quadrados centrais.

Referências sobre distribuições de forma quadráticas

- Notas de aula.
- Rencher and Schaalje (2008), *Linear models in Statistics*, second edition.
- Searle and Gruber (2016), *Linear Models*, second edition.

Testes de Hipótese para o vetor de médias (μ)

- Desejamos testar as hipóteses $H_0 : \mu = \mu_0$ vs $H_1 : \mu \neq \mu_0$, com Σ conhecida e desconhecida, em que $\mu_0 = (\mu_{01}, \dots, \mu_{0p})' \in \mathbb{R}^p$.
- Como visto anteriormente, sob H_0 ,
$$Q = n(\bar{\mathbf{X}} - \mu_0)' \Sigma^{-1} (\bar{\mathbf{X}} - \mu_0) \sim \chi^2_{(p)}$$
 e $F = \frac{n-p}{(n-1)p} T \sim F_{(p, n-p)}$,
em que $T = n(\bar{\mathbf{X}} - \mu_0)' (\mathbf{S}^2)^{-1} (\bar{\mathbf{X}} - \mu_0)$.
- Sob H_1 , entretanto, usando resultados apresentados anteriormente, temos que $Q \sim \chi^2_{(p, \delta)}$, e $F \sim F(n, n-p, \delta)$ em que $\delta = n(\mu - \mu_0)' \Sigma^{-1} (\mu - \mu_0)$.
- Em suma: Se Σ for conhecida, usa-se a estatística Q , caso contrário, usa-se a Estatística F .

Testes de Hipótese para o vetor de médias (μ) (Cont.)

■ Resumo sobre a estatística Q :

- Nível descritivo: $p = P(Q > q_{calc} | \mu = \mu_0)$, sob H_0 , $Q \sim \chi^2_{(p)}$, em que $q_{calc} = n(\bar{x} - \mu_0)' \Sigma^{-1} (\bar{x} - \mu_0)$.
- Função poder: $\phi = 1 - \beta = P(Q > q_c | \mu \neq \mu_0, \alpha)$, sob H_1 , $Q \sim \chi^2_{(p, \delta)}$, em que q_c é o valor crítico para um dado α (nível de significância).
- Poder do teste estimado: $\tilde{\phi} = \widetilde{1 - \beta} = P(\tilde{Q} > q_c | \mu \neq \mu_0, \alpha)$, em que $\tilde{Q} \sim \chi^2_{(p, \tilde{\delta})}$, $\tilde{\delta} = n(\bar{x} - \mu_0)' \Sigma^{-1} (\bar{x} - \mu_0)$.

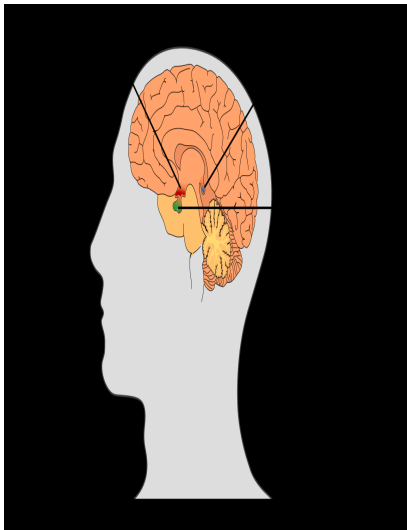
Testes de Hipótese para o vetor de médias (μ) (Cont.)

■ Resumo sobre a estatística F :

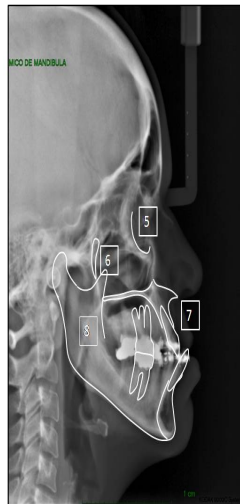
- Nível descritivo: $p = P(F > f_{calc} | \mu = \mu_0)$, sob H_0 , $F \sim F_{(p, n-p)}$, em que $f_{calc} = \frac{n-p}{(n-1)p} n(\bar{x} - \mu_0)' (s^2)^{-1} (\bar{x} - \mu_0)$.
- Função poder: $\phi = 1 - \beta = P(F > f_c | \mu \neq \mu_0, \alpha)$, sob H_1 , $F \sim F_{(p, n-p, \delta)}$, em que f_c é o valor crítico para um dado α (nível de significância).
- Poder do teste estimado: $\tilde{\phi} = \widetilde{1 - \beta} = P(\tilde{F} > f_c | \mu \neq \mu_0, \alpha)$, em que $\tilde{F} \sim F_{(p, n-p, \tilde{\delta})}$, $\tilde{\delta} = n(\bar{x} - \mu_0)' (s^2)^{-1} (\bar{x} - \mu_0)$.

Exemplo 2: Distância do centro da glândula pituitária para a fissura pterigomaxilar (Potthoff and Roy (1964))

- Este conjunto de dados corresponde aos famosos dados de **Potthoff-Roy**, usado para demonstrar a utilização da MANOVA em dados de medidas repetidas (comparação entre grupos, embora comparação entre variáveis seja possível).
- O estudo considerou 16 meninos e 11 meninas, nos quais, nas idades 8, 10, 12 e 14 anos tiveram a distância (mm) do centro da glândula pituitária para a fissura pterigomaxilar medidas.



- 5 – Borda inferior da órbita
- 6 – Fissura pterigomaxilar
- 7 – Maxila
- 8 – Mandíbula
- 9 – Dentes (incisivos e 1º. molares)



Exemplo 2: cont.

- Mudanças nas distâncias pituitária-pterigomaxilar durante o crescimento são importantes na terapia ortodôntica.
- Os objetivos do estudo foram descrever a distância em função da idade e comparar esse desenvolvimento (temporal) entre sexos.
- Por enquanto, em termos de testes de hipótese, vamos desconsiderar os grupos.
- Então temos que:

$\mathbf{X}_i = (X_{i1}, X_{i2}, X_{i3}, X_{i4}) \stackrel{i.i.d}{\sim} N_4(\boldsymbol{\mu}, \boldsymbol{\Sigma}), i = 1, 2, \dots, 27$, ou seja, X_{ij} é a distância no tempo j ($j=1,2,3,4$) do indivíduo i .

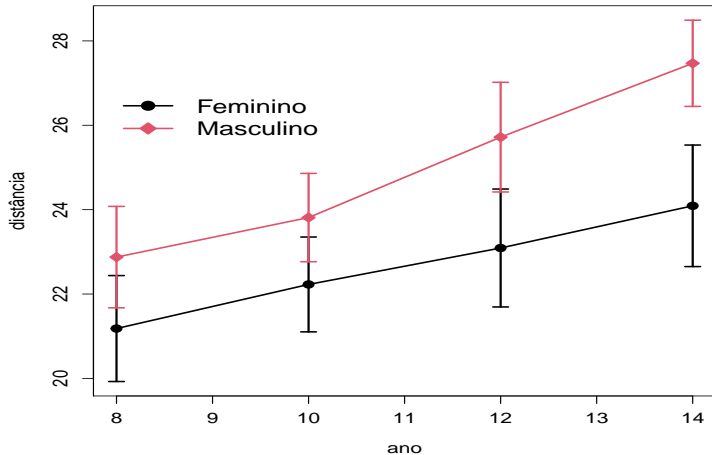
Banco de dados

Indivíduo	Sexo	idade			
		8	10	12	14
1	Feminino	21,0	20,0	21,5	23,0
2	Feminino	21,0	21,5	24,0	25,5
⋮	⋮	⋮	⋮	⋮	⋮
11	Feminino	24,5	25,0	28,0	28,0
1	Masculino	26,0	25,0	29,0	31,0
2	Masculino	21,5	22,5	23,0	26,5
⋮	⋮	⋮	⋮	⋮	⋮
16	Masculino	22,0	21,5	23,5	25,0

Medidas resumo por sexo (S) e ano (A)

S	A	Média	DP	Var.	CV(%)	Min.	Med.	Max.	Curt.	CA	n
F	8	21,18	2,12	4,51	10,03	16,50	21,00	24,50	-0,12	-0,49	11
	10	22,23	1,90	3,62	8,56	19,00	22,50	25,00	-1,37	-0,12	11
	12	23,09	2,36	5,59	10,24	19,00	23,00	28,00	-0,37	0,31	11
	14	24,09	2,44	5,94	10,12	19,50	24,00	28,00	-0,98	-0,22	11
M	8	22,88	2,45	6,02	10,72	17,00	23,00	27,50	0,24	-0,34	16
	10	23,81	2,14	4,56	8,97	20,50	23,50	28,00	-0,88	0,41	16
	12	25,72	2,65	7,03	10,31	22,50	25,00	31,00	-0,59	0,82	16
	14	27,47	2,09	4,35	7,59	25,00	26,75	31,50	-1,07	0,59	16

Perfis médios por sexo



Medidas resumo (desconsiderando sexo)

Ano	Media	DP	Var.	CV(%)	Min.	Med.	Max.	Curt.	CA	n
8	22,19	2,43	5,93	10,97	16,50	22,00	27,50	0,22	-0,22	27
10	23,17	2,16	4,65	9,31	19,00	23,00	28,00	-0,41	0,30	27
12	24,65	2,82	7,94	11,43	19,00	24,00	31,00	0,05	0,57	27
14	26,09	2,77	7,65	10,60	19,50	26,00	31,50	-0,14	-0,15	27

Perfil médio (desconsiderando sexo)

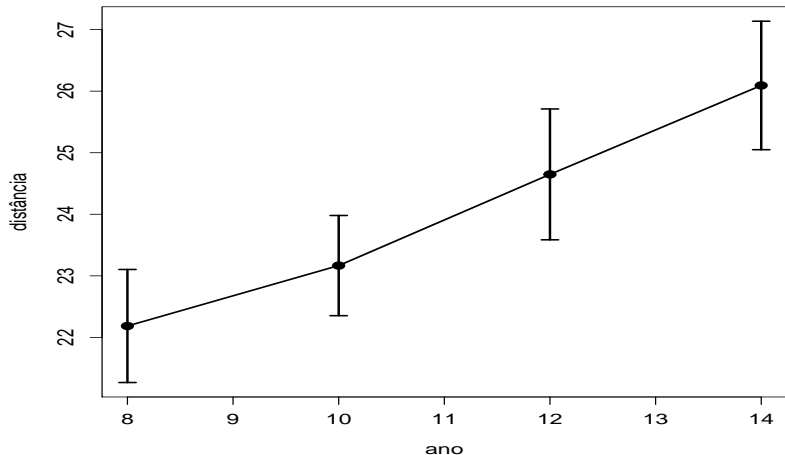
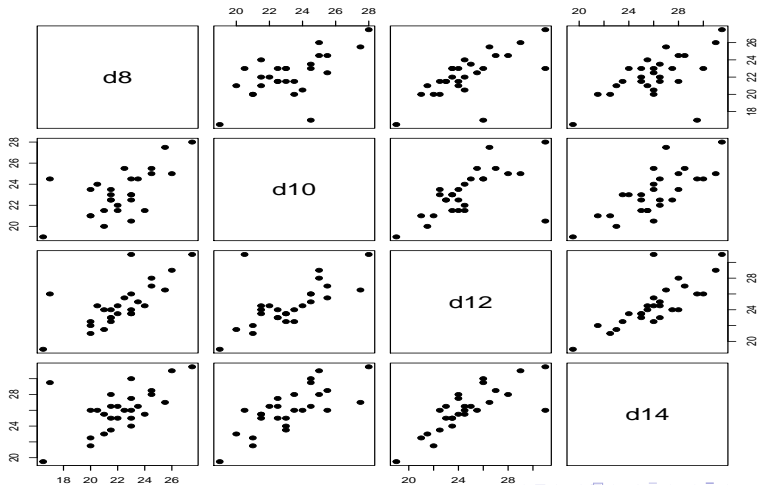
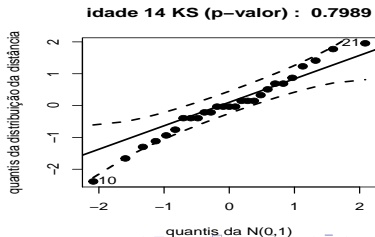
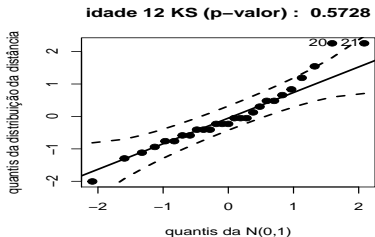
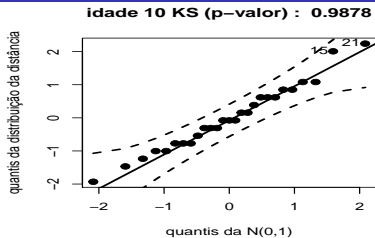
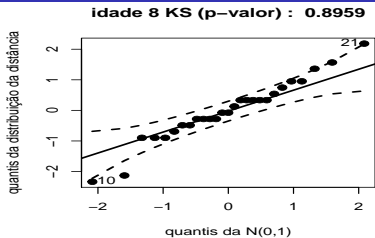


Gráfico de dispersão múltipla

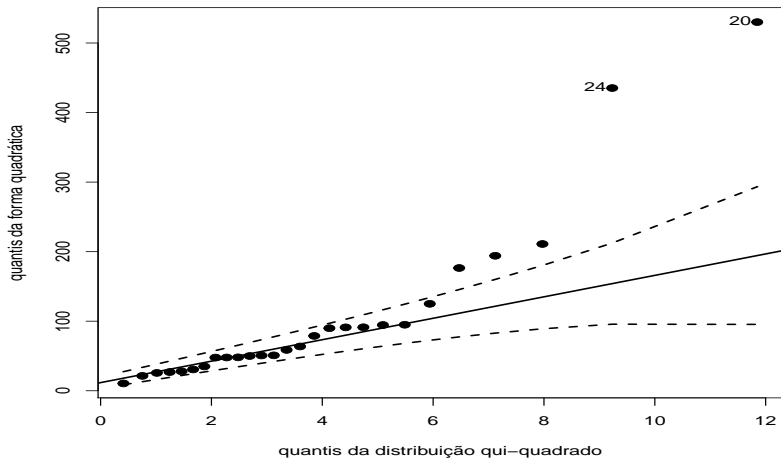


Gráficos de quantis-quantis com envelope (por variável)



Gráficos QQ com envelope (forma quadrática)

KS (p-valor) : < 0,0001



Continuação da análise

- Aplicando os dois testes nos dados (slides de 10 a 12) do Exemplo 2, considerando $\mu_0 = (22, 23, 25, 26)'$, temos os seguintes resultados:
 - Σ conhecido: $q_{calc} = 3,01$ ($p = 0,5558$), $\tilde{\phi} = \widetilde{1 - \beta} = 0,2451$ (para $\alpha = 0,05$).
 - Σ desconhecido: $f_{calc} = 0,67$ ($p = 0,6220$), $\tilde{\phi} = \widetilde{1 - \beta} = 0,2053$ (para $\alpha = 0,05$).
 - Dessa forma, em ambas as situações, não se rejeita H_0 , ou seja, o vetor de médias populacional não é diferente de μ_0 .

Testes de hipótese para combinações lineares de μ

- Já vimos como testar $\mu = \mu_0$ vs $\mu \neq \mu_0$, com Σ conhecida e desconhecida, em que $\mu_0 = (\mu_{01}, \dots, \mu_{0p})' \in \mathbb{R}$.
- Suponha que queremos testar as hipóteses $H_0 : \mathbf{R}\mu = \mathbf{b}$ vs $H_1 : \mathbf{R}\mu \neq \mathbf{b}$, com Σ conhecida e desconhecida, em que \mathbf{b} é um vetor conhecido ($c \times 1$) e \mathbf{R} é uma matriz conhecida de dimensão $c \times p$, $c \leq p$ (esta restrição ajuda a garantir que as linhas de \mathbf{R} não correspondam a hipóteses redundantes), de posto linha completo.
- Primeiramente, consideraremos Σ conhecida.

Testes de hipótese para $R\mu$

- Sabemos que $\bar{\mathbf{X}} \sim N_p(\mu, \frac{1}{n}\Sigma)$, logo, **por resultados anteriores**, temos que $\mathbf{Y} = R\bar{\mathbf{X}} \sim N_c(R\mu, \frac{1}{n}(R\Sigma R'))$, ou seja ,
 $\mathbf{Y} \sim N_c(\mu^*, \Sigma^*)$.

- Considere, agora:

$$\begin{aligned} Q^* &= n(\mathbf{Y} - \mathbf{b})'(\Sigma^*)^{-1}(\mathbf{Y} - \mathbf{b}) \\ &= n(R\bar{\mathbf{X}} - \mathbf{b})'(R\Sigma R')^{-1}(R\bar{\mathbf{X}} - \mathbf{b}) \end{aligned}$$

- Sob H_0 , ou seja, se $R\mu = \mathbf{b}$, então $Q^* \sim \chi_{(c)}^2$ (resultados anteriores, destes slides).

Testes de hipótese para $R\boldsymbol{\mu}$ (cont.)

- Considerando $\boldsymbol{\Sigma}$ desconhecida.
- Além de termos que $\mathbf{Y} = \mathbf{R}\bar{\mathbf{X}} \sim N_c(\mathbf{R}\boldsymbol{\mu}, \frac{1}{n}(\mathbf{R}\boldsymbol{\Sigma}\mathbf{R}'))$, pode-se provar que $\mathbf{R}\mathbf{S}^2\mathbf{R}' \sim W_c(n-1, \mathbf{R}\boldsymbol{\Sigma}\mathbf{R}')$ e que $\mathbf{Y} \perp \mathbf{R}\mathbf{S}^2\mathbf{R}'$.
- Considere, agora:

$$\begin{aligned} F^* &= \frac{(n-c)}{(n-1)c} n(\mathbf{Y} - \mathbf{b})' (\mathbf{R}\mathbf{S}^2\mathbf{R}')^{-1} (\mathbf{Y} - \mathbf{b}) \\ &= \frac{(n-c)}{(n-1)c} n(\mathbf{R}\bar{\mathbf{X}} - \mathbf{b})' (\mathbf{R}\mathbf{S}^2\mathbf{R}')^{-1} (\mathbf{R}\bar{\mathbf{X}} - \mathbf{b}) \end{aligned}$$

- Sob H_0 , ou seja, se $\mathbf{R}\boldsymbol{\mu} = \mathbf{b}$, então $F^* \sim F_{(c, n-c)}$ (resultados anteriores, destes slides).

Testes de hipótese para $R\mu$ (cont.)

- Resumo sobre a estatística Q^* :
 - Nível descritivo: $p = P(Q^* > q_{calc}^* | R\mu = b_0)$, sob H_0 , $Q^* \sim \chi_{(c)}^2$, em que $q_{calc}^* = n(R\bar{x} - b_0)' (R\Sigma R')^{-1} (R\bar{x} - b_0)$.
 - Função poder: $\phi = 1 - \beta = P(Q^* > q_c | R\mu \neq b_0, \alpha)$, sob H_1 , $Q^* \sim \chi_{(c,\delta)}^2$, $\delta = n(R\mu - b_0)' (R\Sigma R')^{-1} (R\mu - b_0)$, em que q_c é o valor crítico para um dado α (nível de significância).
 - Poder do teste estimado: $\tilde{\phi} = \widetilde{1 - \beta} = P(\tilde{Q}^* > q_c | R\mu \neq b_0, \alpha)$, em que $\tilde{Q}^* \sim \chi_{(c,\tilde{\delta})}^2$, $\tilde{\delta} = n(R\bar{x} - b_0)' \Sigma^{-1} (R\bar{x} - b_0)$.

Testes de hipótese para $R\mu$ (cont.)

■ Resumo sobre a estatística F^* :

- Nível descritivo: $p = P(F^* > f_{calc}^* | R\mu = \mathbf{b}_0)$, sob H_0 , $F^* \sim F_{(c, n-c)}$, em que $f_{calc}^* = \frac{n-c}{(n-1)c} n(\mathbf{R}\bar{\mathbf{x}} - \mathbf{b}_0)' (\mathbf{R}\mathbf{s}^2 \mathbf{R}')^{-1} (\mathbf{R}\bar{\mathbf{x}} - \mathbf{b}_0)$.
- Função poder: $\phi = 1 - \beta = P(F^* > f_c | R\mu \neq \mathbf{b}_0, \alpha)$, sob H_1 , $F^* \sim F_{(c, n-c, \delta)}$, $\delta = n(\mathbf{R}\mu - \mathbf{b}_0)' (\mathbf{R}\Sigma \mathbf{R}')^{-1} (\mathbf{R}\mu - \mathbf{b}_0)$, em que f_c é o valor crítico para um dado α (nível de significância).
- Poder do teste estimado: $\tilde{\phi} = 1 - \tilde{\beta} = P(\tilde{F}^* > f_c | R\mu \neq \mathbf{b}_0, \alpha)$, em que $\tilde{F}^* \sim F_{(c, n-c, \tilde{\delta})}$, $\tilde{\delta} = n(\mathbf{R}\bar{\mathbf{x}} - \mathbf{b}_0)' (\mathbf{R}\mathbf{s}^2 \mathbf{R}')^{-1} (\mathbf{R}\bar{\mathbf{x}} - \mathbf{b}_0)$.

Teste a respeito do efeito do tempo, em anos “consecutivos”

- Vamos testar se há diferenças entre as médias de tempos consecutivos, ou seja:
 - (1) $H_0 : \mu_1 = \mu_2$ vs $H_1 : \mu_1 \neq \mu_2$; (2) $H_0 : \mu_2 = \mu_3$ vs $H_1 : \mu_2 \neq \mu_3$;
(3) $H_0 : \mu_3 = \mu_4$ vs $H_1 : \mu_3 \neq \mu_4$.
 - Em cada caso (1): $\mathbf{R} = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \end{bmatrix}$, $\mathbf{b} = \mathbf{0}$; $\mathbf{R} = \begin{bmatrix} 0 & 0 & 1 & -1 \end{bmatrix}$, $\mathbf{b} = \mathbf{0}$.

Resultados numéricos

■ (1) (tempos 1 e 2):

- Σ conhecido: $q_{calc}^* = 6,49$ ($p = 0,0109$), $\tilde{\phi} = \widetilde{1 - \beta} = 0,7214$ (para $\alpha = 0,05$).
- Σ desconhecido: $f_{calc}^* = 6,48$ ($p = 0,0171$), $\tilde{\phi} = \widetilde{1 - \beta} = 0,6887$ (para $\alpha = 0,05$).

■ (2) (tempos 2 e 3):

- Σ conhecido: $q_{calc}^* = 12,16$ ($p = 0,0005$), $\tilde{\phi} = \widetilde{1 - \beta} = 0,9365$ (para $\alpha = 0,05$).
- Σ desconhecido: $f_{calc}^* = 12,16$ ($p = 0,0018$), $\tilde{\phi} = \widetilde{1 - \beta} = 0,9185$ (para $\alpha = 0,05$).

Resultados

- (3) (tempos 3 e 4):
 - Σ conhecido: $q_{calc}^* = 17,61$ ($p < 0,0001$), $\tilde{\phi} = \widetilde{1 - \beta} = 0,9873$ (para $\alpha = 0,05$).
 - Σ desconhecido: $f_{calc}^* = 17,61$ ($p = 0,0003$), $\tilde{\phi} = \widetilde{1 - \beta} = 0,9811$ (para $\alpha = 0,05$).
- Exercício: Escrever (e testar) a hipótese estatística correspondente as seguintes hipóteses relativas ao problema: H_0 : há efeito do tempo vs H_1 : não há efeito de tempo.
- Exercício: Colocar dentro da estrutura vista ($R\mu = b_0$) o teste para dados pareados (visto nos cursos de inferência).
- Exercício: Repetir os testes feitos, para cada um dos sexos em separado.