

# ELEMENTOS DE ESTATÍSTICA MULTIVARIADA |

**21163**

**Atividades Formativas (Exercícios e Resoluções)**

***Catarina S. Nunes***  
**Secção de Matemática, Departamento de Ciências e  
Tecnologia  
Universidade Aberta**



Março de 2020



## ATIVIDADE FORMATIVA 1

### Temas:

- Populações Multivariadas
- Variáveis Aleatórias Multidimensionais

### Objetivos

- Conceito de Estatística Multivariada
- Tipos de dados
- Medidas descritivas de populações multivariadas e para amostras multivariadas
- Conceito de variável aleatória
- Função de distribuição e Funções de distribuições marginais
- Independência de variáveis aleatórias e Combinações lineares de variáveis aleatórias
- Propriedades das matrizes de variância/covariância

### Descrição da Atividade (Exercícios a Resolver)

1. Um empresário proprietário de quatro empresas de diferentes ramos, obteve os seguintes resultados líquidos (em milhões de euros) ao longo de seis anos consecutivos.

Empresa	Anos					
	1996	1997	1998	1999	2000	2001
E1	0,80	0,75	0,90	1,20	0,95	0,82
E2	0,90	0,80	0,85	1,05	0,70	0,90
E3	0,70	0,79	1,00	0,90	0,85	0,83
E4	0,75	0,80	0,87	1,10	0,92	0,92

Com base nos valores indicados do quadro, determine nas quatro empresas durante os seis anos:

- 1.1 As médias dos resultados líquidos anuais.
- 1.2 As variâncias e covariâncias dos resultados líquidos anuais.
- 1.3 As correlações entre os resultados líquidos.

2. Para a admissão de novos alunos numa escola de natação são registadas as seguintes características:

$X_1$  : aptidão do novo praticante ( 1 = sabe nadar ; 2 = não sabe nadar )  
 $X_2$  : sexo ( 1 = masculino ; 2 = feminino )

A função de probabilidade conjunta da variável bidimensional  $(X_1, X_2)$  é a seguinte:

$f(X_1, X_2)$	$X_1 = 1$	$X_1 = 2$
$X_2 = 1$	4/15	5/15
$X_2 = 2$	4/15	2/15

- 2.1 Obtenha as funções de probabilidade marginal de  $X_1$  e  $X_2$  .  
 2.2 Calcule a probabilidade de um novo aluno ser do sexo feminino.  
 2.3 Calcule a probabilidade de um novo praticante não saber nadar, sabendo-se que é homem.  
 2.4 Qual a probabilidade de uma futura aluna já ter experiência em natação?
3. Seja  $X_1$  uma variável aleatória com média  $\mu_1 = 5$  e variância  $\sigma_{11} = 10$ , e  $X_2$  uma segunda variável aleatória com média  $\mu_2 = 6$ , variância  $\sigma_{22} = 6$  e  $\sigma_{12} = \sigma_{21} = 8$ . Determine:

- 3.1  $E(2X_1)$   
 3.2  $V(3X_1)$   
 3.3  $Cov(2X_1, 3X_2)$   
 3.4  $E(5X_1 + 2X_2)$   
 3.5  $V(X_1 + 3X_2)$

4. Três variáveis aleatórias,  $X$ ,  $Y$  e  $Z$ , têm função de densidade de probabilidade conjunta:

$$f(x, y, z) = \begin{cases} kxyz^2 & \text{se } 0 < x, y < 1; 0 < z < 3 \\ 0 & \text{se } \dots \end{cases}$$

- 4.1 Mostre que a constante  $k$  é 4/9.  
 4.2 Mostre que as variáveis aleatórias  $X$ ,  $Y$  e  $Z$  são independentes.  
 4.3 Mostre que a função de densidade marginal conjunta de  $X$  e  $Z$  é dada por:

$$f(x, z) = \begin{cases} 2xz^2/9 & \text{se } 0 < x < 1, 0 < z < 3 \\ 0 & \text{se } \dots \end{cases}$$

- 4.4 Mostre que  $E(X) = 2/3$ .  
 4.5 Determina e função distribuição de  $X$  dado  $Y = 1/2$  e  $Z = 1$ .  
 4.6 Mostre que a covariância de  $X$  e  $Z$  é zero.



UAb

## ATIVIDADE FORMATIVA 2

### Temas:

- Distribuição Normal Multivariada
- Distribuições Amostrais
- Testes de Hipóteses Multivariados

### Objetivos

- Conceito e estrutura de distribuição normal multivariada
- Representação da distribuição normal multivariada
- Combinações lineares de variáveis aleatórias normais multivariadas
- Introdução a distribuições amostrais
- Estimadores de máxima verosimilhança para a distribuição normal multivariada
- Distribuição da média amostral
- Distribuições amostrais relacionadas com as matrizes de variância/covariância (distribuição de Wishart e distribuição de  $T^2$  de Hotelling)
- Método da união intersecção (testes para o vetor se médias com a matriz de variância/covariância conhecida ou desconhecida; teste para a matriz de variância/covariância)
- Método do quociente de verosimilhança (testes para o vetor se médias com a matriz de variância/covariância conhecida ou desconhecida; teste para a matriz de variância/covariância)
- Comparação entre dois vetores de médias (teste multivariado como generalização do teste univariado; teste para igualdade de dois vetores de médias quando a matriz de variância/covariância é desconhecida)

### Descrição da Atividade (Exercícios a Resolver)

1. Considere o vector aleatório  $(X_1, X_2)$  que segue uma distribuição normal bivariada  $N_2(\mu, \Sigma)$  em que  $\mu = (\mu_1, \mu_2) = (0, 2.5)$ ,  $\sigma_1 = \sigma_2 = 1$  e em que a matriz de correlações  $R$  é dada por:

$$R = \begin{bmatrix} 1 & 0.65 \\ 0.65 & 1 \end{bmatrix}$$

- 1.1 Calcule o valor da covariância entre  $X_1$  e  $X_2$ . Comente ainda o valor da intensidade da correlação entre as duas variáveis. Justifique o seu comentário.
- 1.2 Considere uma terceira variável aleatória  $X_3$ , dada por  $X_3 = \frac{X_2-1}{2}$ , com distribuição normal com valor médio  $\mu_3$  e desvio padrão  $\sigma_3$ . Utilizando as propriedades do valor esperado e da covariância da combinação linear de variáveis aleatórias, determine o vector das médias e a correlação do par aleatório  $(X_1, X_3)$ .
- 1.3 Escreva a expressão da função de densidade de probabilidade conjunta do par aleatório dado inicialmente, isto é,  $f(x_1, x_2)$ .
2. Seja  $X \sim N_2(\mu, \Sigma)$  com  $\Sigma$  desconhecida. A partir de uma amostra de 20 observações obtiveram-se as seguintes estimativas:

$$\bar{X} = \begin{bmatrix} 52.35 \\ 55.76 \end{bmatrix}, S = \begin{bmatrix} 12 & 0.5 \\ 0.5 & 18 \end{bmatrix}, S^{-1} = \begin{bmatrix} 0.0834 & -0.0023 \\ -0.0023 & 0.0556 \end{bmatrix}$$

- 2.1 Use o método do quociente de verosimilhança para testar a hipótese  $H_0 : \mu' = [51.5 \ 54.6]$  considerando  $\alpha=5\%$ . Indique todos os pressupostos de aplicabilidade. Justifique todos os cálculos e apresente as suas conclusões.
- 2.2 Teste, ao nível  $\alpha=5\%$ , a hipótese nula

$$H_0 : \Sigma = \begin{bmatrix} 10 & 0 \\ 0 & 16 \end{bmatrix}$$

indique todos os pressupostos que achar necessário.

3. Foram analisados 10 equipamentos de uma fábrica ao longo de 5 anos, tendo-se obtido os seguintes resultados:

$$\bar{X} = \begin{bmatrix} 52.1 \\ 12.3 \\ 20.2 \end{bmatrix}, S = \begin{bmatrix} 5.5 & & \\ 10.1 & 50.2 & \\ -1.8 & -1.2 & 20.1 \end{bmatrix}$$

Seja  $X_1$  - custo do equipamento (u.m.);  $X_2$  - custos de reparação (u.m.);  $X_3$  - lucros obtidos (u.m.).

Teste a hipótese para  $\alpha = 5\%$ :

$$H_0 : \mu = \begin{bmatrix} 50 \\ 15 \\ 20 \end{bmatrix} \quad \text{versus} \quad H_1 : \mu \neq \begin{bmatrix} 50 \\ 15 \\ 20 \end{bmatrix}$$

4. Considere  $X \sim N_2(\mu, \Sigma)$  com  $\Sigma$  desconhecida. A partir de uma amostra de 75 observações obtiveram-se as seguintes estimativas:

$$\bar{X} = \begin{bmatrix} 52.37 \\ 55.76 \end{bmatrix}, S = \begin{bmatrix} 23.44 & 15.62 \\ 15.62 & 27.39 \end{bmatrix}$$

Teste a hipótese, que  $H_0 : \mu' = [52, 52]$ , considerando  $\alpha = 5\%$ .

5. Teste a seguinte hipótese sobre a matriz de variância/covariância de uma população com distribuição normal bivariada:

$$H_0 : \Sigma = \begin{bmatrix} 20 & 0 \\ 0 & 20 \end{bmatrix}$$

para  $\alpha = 1\%$ , sabendo que  $n = 75$  e

$$S = \begin{bmatrix} 22.56 & 15.86 \\ 15.86 & 17.43 \end{bmatrix}$$

6. Durante três anos consecutivos registaram-se resultados relativos aos lucros obtidos por 36 clubes de futebol. Com base nesses resultados obtiveram-se as seguintes estimativas:

$X_1$  resultados do ano 1 (u.m.);  $X_2$  resultados do ano 2 (u.m.);  $X_3$  resultados do ano 3 (u.m.)

$$\bar{X} = \begin{bmatrix} 0.2 \\ -0.5 \\ -0.9 \end{bmatrix} \quad S = \begin{bmatrix} 1.2 & & \\ -0.9 & 2 & \\ 1.1 & -0.8 & 5.1 \end{bmatrix}$$

6.1 Teste a hipótese de que os resultados nos últimos três anos foram nulos, ou seja

$H_0 : \mu = 0$ , com  $\alpha = 1\%$ .

6.2 Obtenha a condição que define a região de confiança a 99% para  $\mu$ .





UAb

### ATIVIDADE FORMATIVA 3

#### Tema:

- Análise de Variância Multivariada (MANOVA)

#### Objetivos

- Teste de igualdade de k vetores de médias (Oneway MANOVA)
- Teste para a igualdade de matrizes de variância/covariância
- Verificação de pressupostos de normalidade
- Transformações para a normalidade

#### Descrição da Atividade (Exercícios a Resolver)

1. Considere 2 quintas de produção de carne bovina. Em cada quinta são recolhidos o peso (em Kg) e a altura (em cm) de 25 animais. Os dados amostrais recolhidos foram os seguintes:

	Quinta 1		Quinta 2	
	Peso	Altura	Peso	Altura
Média	137.8	98.5	112.3	111.6
Desvio padrão	23.7	12.4	22.8	11.9
Correlação	0.43		0.49	

- 1.1 Formule e teste, a 5% de significância, a hipótese de as matrizes de covariância dos 2 grupos serem iguais.
  - 1.2 Formule e teste, a 5% de significância, a hipótese de as 2 quintas criarem animais similares.
2. Após análise dos dados recolhidos de 10 veículos duma frota de uma empresa de transportes relativos a 4 anos de actividade, obtiveram-se os seguintes resultados:

$$\bar{X} = \begin{bmatrix} 850 \\ 160 \\ 110 \end{bmatrix} \quad S = \begin{bmatrix} 15.0 & & \\ 10.1 & 10.2 & \\ -1.5 & -1.8 & 6.1 \end{bmatrix}$$

Sendo  $X_1$  Kilómetros percorridos ( $\times 10^3$ );  $X_2$  custos de manutenção (u.m.);  $X_3$  lucros obtidos (u.m.). Teste a hipótese para  $\alpha = 5\%$ :

$$H_0 : \mu = \begin{bmatrix} 851 \\ 159 \\ 108 \end{bmatrix} \quad \text{versus} \quad H_1 : \mu \neq \begin{bmatrix} 851 \\ 159 \\ 108 \end{bmatrix}$$

3. Os valores que se apresentam referem-se às classificações de (0 a 20 valores) obtidas por alunos em dois exames de diferentes matérias. A amostra aleatória refere-se a 11 alunos sendo 4 do sexo masculino e 7 do sexo feminino. Teste a hipótese:

$$H_0 : \mu = \begin{bmatrix} \mu_{11} \\ \mu_{21} \end{bmatrix} = \begin{bmatrix} \mu_{12} \\ \mu_{22} \end{bmatrix}$$

considerando um nível de significância  $\alpha = 5\%$ .

Alunos		Alunas	
Exame 1	Exame 2	Exame 1	Exame 2
9	11	6	14
13	14	12	13
8	9	8	12
11	13	10	8
$\bar{x}_{11}=10.25$	$\bar{x}_{21}=11.75$	13	9
		11	10
		10	15
		$\bar{x}_{12}=10$	$\bar{x}_{22}=12$

4. Apresente de forma sucinta duas razões que justifiquem a preferência por um teste multivariado em vez de testes univariados.
5. Obtiveram-se duas amostras relativas a condutores de um certo modelo de veículos ligeiros: uma amostra respeitante a condutores não profissionais ( $n_1=14$ ) e outra a condutores de Táxi profissionalizados ( $n_2=10$ ). Para essas amostras mediram-se as seguintes características:

$X_1$ : custos com a manutenção do veículo (em milhares de euros)

$X_2$ : número de quilómetros percorridos ( $\times 10^3$ )

para as quais se obtiveram os seguintes valores:

Condutores não profissionais:

Custos	0,75	0,56	0,87	0,95	1,32	0,32	0,45	0,47	0,55	0,87	0,98	1,05	1,03	1,50
Quilómetros	65	56	78	67	109	57	52	72	50	60	88	102	92	110

Condutores de Táxi profissionalizados:

Custos	0,85	1,56	1,65	1,32	0,91	2,02	1,78	1,65	2,98	1,98
Quilómetros	85	156	178	167	102	304	204	156	450	234

Proceda aos testes das seguintes hipóteses:

5.1 As duas amostras são provenientes de populações com idêntica dispersão:  $H_0 : \Sigma_1 = \Sigma_2$ .

5.2 Os condutores não profissionais e os condutores de Táxi têm médias idênticas para as características medidas.

6. Considerem-se os seguintes dados respeitantes a duas variáveis ( $X_1$  e  $X_2$ ) e três grupos ( $n_1=4$ ,  $n_2=3$  e  $n_3=5$ ) recolhidos através de um processo aleatório:<sup>1</sup>

Grupo	1	Grupo	2	Grupo	3
	$X_1$		$X_2$		$X_1$
	2		4		7
	3		5		8
3	5		6		10
	2		7		9
	3		8		7
	4		9		8
	5		10		5
	6		7		6
	7		8		7
	8		9		8
	9		10		9
	10		11		10
	11		12		11
	12		13		12
	13		14		13
	14		15		14
	15		16		15
	16		17		16
	17		18		17
	18		19		18
	19		20		19
	20		21		20
	21		22		21
	22		23		22
	23		24		23
	24		25		24
	25		26		25
	26		27		26
	27		28		27
	28		29		28
	29		30		29
	30		31		30
	31		32		31
	32		33		32
	33		34		33
	34		35		34
	35		36		35
	36		37		36
	37		38		37
	38		39		38
	39		40		39
	40		41		40
	41		42		41
	42		43		42
	43		44		43
	44		45		44
	45		46		45
	46		47		46
	47		48		47
	48		49		48
	49		50		49
	50		51		50
	51		52		51
	52		53		52
	53		54		53
	54		55		54
	55		56		55
	56		57		56
	57		58		57
	58		59		58
	59		60		59
	60		61		60
	61		62		61
	62		63		62
	63		64		63
	64		65		64
	65		66		65
	66		67		66
	67		68		67
	68		69		68
	69		70		69
	70		71		70
	71		72		71
	72		73		72
	73		74		73
	74		75		74
	75		76		75
	76		77		76
	77		78		77
	78		79		78
	79		80		79
	80		81		80
	81		82		81
	82		83		82
	83		84		83
	84		85		84
	85		86		85
	86		87		86
	87		88		87
	88		89		88
	89		90		89
	90		91		90
	91		92		91
	92		93		92
	93		94		93
	94		95		94
	95		96		95
	96		97		96
	97		98		97
	98		99		98
	99		100		99
	100		101		100
	101		102		101
	102		103		102
	103		104		103
	104		105		104
	105		106		105
	106		107		106
	107		108		107
	108		109		108
	109		110		109
	110		111		110
	111		112		111
	112		113		112
	113		114		113
	114		115		114
	115		116		115
	116		117		116
	117		118		117
	118		119		118
	119		120		119
	120		121		120
	121		122		121
	122		123		122
	123		124		123
	124		125		124
	125		126		125
	126		127		126
	127		128		127
	128		129		128
	129		130		129
	130		131		130
	131		132		131
	132		133		132
	133		134		133
	134		135		134
	135		136		135
	136		137		136
	137		138		137
	138		139		138
	139		140		139
	140		141		140
	141		142		141
	142		143		142
	143		144		143
	144		145		144
	145		146		145
	146		147		146
	147		148		147
	148		149		148
	149		150		149
	150		151		150
	151		152		151
	152		153		152
	153		154		153
	154		155		154
	155				



## ATIVIDADE FORMATIVA 1 - Proposta de Resolução

### 1.

#### 1.1

$$\mu = \frac{1}{N} \sum_{i=1}^N X_i$$

$$\begin{aligned} \mu &= \frac{1}{6} \sum_{i=1}^6 X_i = \frac{1}{6} \left[ \begin{pmatrix} 0.8 \\ 0.9 \\ 0.7 \\ 0.75 \end{pmatrix} + \begin{pmatrix} 0.75 \\ 0.8 \\ 0.79 \\ 0.8 \end{pmatrix} + \begin{pmatrix} 0.9 \\ 0.85 \\ 1 \\ 0.87 \end{pmatrix} + \begin{pmatrix} 1.2 \\ 1.05 \\ 0.9 \\ 1.1 \end{pmatrix} + \begin{pmatrix} 0.95 \\ 0.7 \\ 0.85 \\ 0.92 \end{pmatrix} + \right. \\ &\left. + \begin{pmatrix} 0.82 \\ 0.9 \\ 0.83 \\ 0.92 \end{pmatrix} \right] = \frac{1}{6} \begin{bmatrix} 5.42 \\ 5.2 \\ 5.07 \\ 5.36 \end{bmatrix} = \begin{bmatrix} 0.903 \\ 0.867 \\ 0.845 \\ 0.893 \end{bmatrix} \end{aligned}$$

$\mu_1$ - média dos resultados líquidos anuais em milhares de euros para  $E1 = 0.903$ ;

$\mu_2$ - média dos resultados líquidos anuais em milhares de euros para  $E2 = 0.867$ ;

$\mu_3$ - média dos resultados líquidos anuais em milhares de euros para  $E3 = 0.845$ ;

$\mu_4$ - média dos resultados líquidos anuais em milhares de euros para  $E4 = 0.893$ ;

#### 1.2

$$\Sigma = \frac{1}{N} \sum_{i=1}^N (X_i - \mu)(X_i - \mu)'$$

$$\Sigma = \frac{1}{6} \sum_{i=1}^6 (X_i - \mu)(X_i - \mu)' =$$

$$= \frac{1}{6} \left[ \begin{pmatrix} 0.8 - 0.903 \\ 0.9 - 0.867 \\ 0.7 - 0.845 \\ 0.75 - 0.893 \end{pmatrix} (0.8 - 0.903 \quad 0.9 - 0.867 \quad 0.7 - 0.845 \quad 0.75 - 0.893) + \right.$$

$$+ \begin{pmatrix} -0.153 \\ -0.067 \\ -0.055 \\ -0.093 \end{pmatrix} (-0.153 \quad -0.067 \quad -0.055 \quad -0.093) +$$

$$+ \begin{pmatrix} -0.0033 \\ -0.0167 \\ 0.1550 \\ -0.0233 \end{pmatrix} (-0.0033 \quad -0.0167 \quad 0.1550 \quad -0.0233) +$$

$$\begin{aligned}
& + \begin{pmatrix} 0.2967 \\ 0.1833 \\ 0.0550 \\ 0.2067 \end{pmatrix} ( 0.2967 \ 0.1833 \ 0.0550 \ 0.2067 ) + \\
& + \begin{pmatrix} 0.0467 \\ -0.1667 \\ 0.0050 \\ 0.0267 \end{pmatrix} ( 0.0467 \ -0.1667 \ 0.0050 \ 0.0267 ) + \\
& + \begin{pmatrix} -0.0833 \\ 0.0333 \\ -0.0150 \\ 0.0267 \end{pmatrix} ( -0.0833 \ 0.0333 \ -0.0150 \ 0.0267 ) = \\
& = \begin{bmatrix} 0.0219 & 0.0084 & 0.0068 & 0.0149 \\ 0.0084 & 0.0114 & 0.0008 & 0.0060 \\ 0.0068 & 0.0008 & 0.0086 & 0.0056 \\ 0.0149 & 0.006 & 0.0056 & 0.0123 \end{bmatrix}
\end{aligned}$$

1.3

$$\begin{aligned}
& P = \Delta^{-1/2} \Sigma \Delta^{-1/2} \\
& P = \begin{bmatrix} \frac{1}{\sqrt{0.0219}} & 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{0.0114}} & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{0.0086}} & 0 \\ 0 & 0 & 0 & \frac{1}{\sqrt{0.0123}} \end{bmatrix} \Sigma \begin{bmatrix} \frac{1}{\sqrt{0.0219}} & 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{0.0114}} & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{0.0086}} & 0 \\ 0 & 0 & 0 & \frac{1}{\sqrt{0.0123}} \end{bmatrix} = \\
& = \begin{bmatrix} 1 & 0.5348 & 0.4956 & 0.9086 \\ 0.5348 & 1 & 0.0844 & 0.5088 \\ 0.4956 & 0.0844 & 1 & 0.5421 \\ 0.9086 & 0.5088 & 0.5421 & 1 \end{bmatrix}
\end{aligned}$$

2.

2.1

$$f_1(x_1) = \begin{cases} 8/15 & x_1 = 1 \\ 7/15 & x_1 = 2 \\ 0 & \text{outros valores} \end{cases}$$

$$f_2(x_2) = \begin{cases} 9/15 & x_2 = 1 \\ 6/15 & x_2 = 2 \\ 0 & \text{outros valores} \end{cases}$$

2.2

$$P[X_2 = 2] = f_2(2) = 6/15$$

2.3

$$P[X_1 = 2|X_2 = 1] = \frac{f(2,1)}{f_2(1)} = \frac{5/15}{9/15} = 5/9$$

2.4

$$P[X_1 = 1|X_2 = 2] = \frac{f(1,2)}{f_2(2)} = \frac{4/15}{6/15} = 4/6$$

3.

$$X_1: \mu_1 = 5 \quad \sigma_{11} = 10$$

$$X_2: \mu_2 = 6 \quad \sigma_{22} = 6$$

$$\sigma_{12} = \sigma_{21} = 8$$

3.1

$$E[2X_1] = 2E[X_1] = 2 \times 5 = 10$$

3.2

$$V(3X_1) = 3^2V(X_1) = 9 \times 10 = 90$$

3.3

$$\text{Cov}(2X_1, 3X_2) = 2 \times 3\text{Cov}(X_1, X_2) = 6\sigma_{12} = 6 \times 8 = 48$$

3.4

$$E[5X_1 + 2X_2] = 5E[X_1] + 2E[X_2] = 5\mu_1 + 2\mu_2 = 5 \times 5 + 2 \times 6 = 37$$

3.5

$$V(X_1 + 3X_2) = V(X_1) + 3^2V(X_2) + 2 \times 3\text{Cov}(X_1, X_2) = \sigma_{11} + 9\sigma_{22} + 6\sigma_{12} = 10 + 54 + 48 = 112$$

4.

$$f(x, y, z) = \begin{cases} kxyz^2 & \text{se } 0 < x, y < 1, 0 < z < 3 \\ 0 & \text{se } \dots \end{cases}$$

4.1

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y, z) dx dy dz = 1$$

$$\int_0^1 \int_0^1 \int_0^3 kxyz^2 dx dy dz = 1$$

$$\int_0^1 \int_0^1 kxy \left[ \frac{z^3}{3} \right]_0^3 dx dy = 1 \Leftrightarrow \int_0^1 \int_0^1 kxy 9 dx dy = 1$$

$$9 \int_0^1 kx \left[ \frac{y^2}{2} \right]_0^1 dx = 1 \Leftrightarrow 9 \int_0^1 \frac{1}{2} kx dx = 1$$

$$\frac{9}{2} k \left[ \frac{x^2}{2} \right]_0^1 = 1 \Leftrightarrow \frac{9}{4} k = 1 \Leftrightarrow k = \frac{4}{9}$$

4.2 Independência: f.d.p. conjunta = produto das f.d.p. marginais

$$f(x, y, z) = f_1(x)f_2(y)f_3(z)$$

$$\begin{aligned} f_1(x) &= \int_0^1 \int_0^3 f(x, y, z) dy dz = \int_0^1 \int_0^3 \frac{4}{9} xyz^2 dy dz = \int_0^1 \frac{4}{9} xy \left[ \frac{z^3}{3} \right]_0^3 dy = \\ &= \int_0^1 \frac{4}{9} xy 9 dy = 4x \left[ \frac{y^2}{2} \right]_0^1 = 4x \frac{1}{2} = 2x \end{aligned}$$

$$f_1(x) = \begin{cases} 2x & 0 < x < 1 \\ 0 & \dots \end{cases}$$

$$f_2(y) = \int_0^1 \int_0^3 f(x, y, z) dx dz = \int_0^1 \int_0^3 \frac{4}{9}xyz^2 dx dz = \int_0^1 4xy dx = 2y$$

$$f_2(y) = \begin{cases} 2y & 0 < y < 1 \\ 0 & \dots \end{cases}$$

$$f_3(z) = \int_0^1 \int_0^1 f(x, y, z) dx dy = \int_0^1 \int_0^1 \frac{4}{9}xyz^2 dx dy = \int_0^1 \frac{4}{9}xz^2 \left[ \frac{y^2}{2} \right]_0^1 dx = \\ = \int_0^1 \frac{2}{9}xz^2 dx = \frac{2}{9}z^2 \left[ \frac{x^2}{2} \right]_0^1 = \frac{z^2}{9}$$

$$f_3(z) = \begin{cases} \frac{z^2}{9} & 0 < z < 3 \\ 0 & \dots \end{cases}$$

$$f_1(x)f_2(y)f_3(z) = \begin{cases} 2x2y\frac{z^2}{9} & 0 < x, y < 1; 0 < z < 3 \\ 0 & \dots \end{cases} = \\ \begin{cases} \frac{4}{9}xyz^2 & 0 < x, y < 1; 0 < z < 3 \\ 0 & \dots \end{cases} = f(x, y, z)$$

$\Rightarrow X, Y$  e  $Z$  são independentes.

**4.3** Se  $X$  e  $Z$  são independentes então:

$$f(x, z) = f_1(x)f_3(z) = \begin{cases} \frac{2}{9}xz^2 & 0 < x < 1; 0 < z < 3 \\ 0 & \dots \end{cases}$$

**4.4**

$$E[X] = \mu_x = \int_0^1 xf_1(x) dx = \int_0^1 x2x dx = \int_0^1 2x^2 dx = \frac{2}{3} x^3 \Big|_0^1 = \frac{2}{3}$$

**4.5** A função distribuição condicional de  $X$  dado  $Y = 1/2$  e  $Z = 1$  é igual à função de distribuição marginal de  $X$ ,  $F_1(x)$ . Dado que as variáveis são independentes.

$$F_1(x) = P(X \leq x) = \int_0^x f_1(x) dx = \int_0^x 2x dx = x^2 \Big|_0^x = x^2$$

$$F_1(x) = \begin{cases} 0 & x \leq 0 \\ x^2 & 0 < x < 1 \\ 1 & x \geq 1 \end{cases}$$

**4.6** Se as variáveis aleatórias  $X$  e  $Z$  são independentes então  $Cov(X, Z) = 0$ . Como verificar:

$$E[Z] = \mu_z = \int_0^3 zf_3(z) dz = \int_0^3 z\frac{z^2}{9} dz = \frac{1}{9} \int_0^3 z^3 dz = \frac{1}{9} \left[ \frac{z^4}{4} \right]_0^3 = \frac{9}{4}$$

$$Cov(X, Z) = \sigma_{xz} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - \mu_x)(z - \mu_z)f(x, z) dx dz = \\ = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \left(x - \frac{2}{3}\right) \left(z - \frac{9}{4}\right) f_1(x)f_3(z) dx dz = \\ = \int_0^1 \int_0^3 \left(x - \frac{2}{3}\right) \left(z - \frac{9}{4}\right) 2x\frac{z^2}{9} dx dz =$$

$$\begin{aligned}
&= \int_0^3 \left( z - \frac{9}{4} \right) \frac{z^2}{9} \int_0^1 \left( x - \frac{2}{3} \right) 2x \, dx \, dz = \\
&= \int_0^3 \left( \frac{z^3}{4} - \frac{z^2}{4} \right) \int_0^1 \left( 2x^2 - \frac{4}{6}x \right) \, dx \, dz = \\
&= \int_0^3 \left( \frac{z^3}{4} - \frac{z^2}{4} \right) \left[ \frac{2x^3}{3} - \frac{2x^2}{6} \right]_0^1 \, dz = \\
&= \int_0^3 \left( \frac{z^3}{4} - \frac{z^2}{4} \right) \left( \frac{2}{3} - \frac{2}{6} \right) \, dz = \\
&= \frac{2}{6} \int_0^3 \left( \frac{z^3}{4} - \frac{z^2}{4} \right) \, dz = \frac{2}{6} \left[ \frac{z^4}{4 \times 4} - \frac{z^3}{4 \times 3} \right]_0^3 = \frac{2}{6} \left[ \frac{9}{4} - \frac{9}{4} \right] = 0
\end{aligned}$$



**ATIVIDADE FORMATIVA 2 - Proposta de Resolução**

1.

$$\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 2.5 \end{bmatrix}$$

$$R = \begin{bmatrix} 1 & 0.65 \\ 0.65 & 1 \end{bmatrix}$$

$\rho = 0.65$  : coeficiente de correlação entre  $X_1$  e  $X_2$

1.1

$$Cov(X_1, X_2) = \sigma_{12} = \rho\sigma_1\sigma_2 = 0.65$$

$\rho = 0.65 > 0$  (relação positiva entre  $X_1$  e  $X_2$ ) e  $\rho^2 = 0.45$

1.2

$$X_3 = \frac{X_2 - 1}{2} \quad X_3 \sim N(\mu_3, \sigma_3)$$

$$\mu_3 = E\left[\frac{X_2 - 1}{2}\right] = \frac{E[X_2]}{2} - \frac{1}{2} = \frac{2.5}{2} - 0.5 = 0.75$$

$$\sigma_3 = Var[X_3] = \frac{1}{4}Var[X_2] = \frac{1}{4}$$

$$(X_1, X_3) \rightarrow \mu = \begin{bmatrix} \mu_1 \\ \mu_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0.75 \end{bmatrix}$$

$$\sigma_{13} = Cov(X_1, X_3) = Cov\left(X_1, \frac{X_2 - 1}{2}\right) = Cov(X_1, X_2) \frac{1}{4} = \frac{1}{4} \times 0.65 = 0.1625$$

$$\rho_{13} = \frac{\sigma_{13}}{\sigma_1\sigma_3} = \frac{0.1625}{1 \times \frac{1}{4}} = 0.65$$

1.3

$$\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 2.5 \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{bmatrix} = \begin{bmatrix} 1 & 0.65 \\ 0.65 & 1 \end{bmatrix}$$

$$f(x_1, x_2) = \frac{1}{2\pi|\Sigma|^{1/2}} \exp\left[-\frac{1}{2}(X - \mu)' \Sigma^{-1}(X - \mu)\right] =$$

$$= \frac{1}{2\pi\sqrt{\sigma_1^2\sigma_2^2 - \sigma_{12}\sigma_{21}}} \exp\left[-\frac{1}{2} \begin{bmatrix} X_1 - \mu_1 & X_2 - \mu_2 \end{bmatrix} \Sigma^{-1} \begin{bmatrix} X_1 - \mu_1 \\ X_2 - \mu_2 \end{bmatrix}\right] =$$

$$= \frac{1}{2\pi \cdot 0.7599} \exp\left[-\frac{1}{2} \begin{bmatrix} X_1 & X_2 - 2.5 \end{bmatrix} \Sigma^{-1} \begin{bmatrix} X_1 \\ X_2 - 2.5 \end{bmatrix}\right]$$

$$\Sigma^{-1} = \begin{bmatrix} 1.7316 & -1.1255 \\ -1.1255 & 1.7316 \end{bmatrix}$$

2.

2.1  $H_0 : \mu = \mu_0 = \begin{bmatrix} 51.5 \\ 54.6 \end{bmatrix}$  com  $\Sigma$  desconhecida.  $X \sim N_2(\mu, \Sigma)$   
 $H_1 : \mu \neq \mu_0$  e  $\alpha = 5\%$

Estatística de teste resultante do quociente de verosimilhança:

$$-2 \ln \Lambda = n \ln \left[ 1 + (\bar{X} - \mu_0)' S^{-1} (\bar{X} - \mu_0) \right]$$

$$(\bar{X} - \mu_0)' S^{-1} (\bar{X} - \mu_0) = \begin{bmatrix} 0.85 & 1.16 \end{bmatrix} \begin{bmatrix} 0.0834 & -0.0023 \\ -0.0023 & 0.0556 \end{bmatrix} \begin{bmatrix} 0.85 \\ 1.16 \end{bmatrix} =$$

$$= \begin{bmatrix} 0.0682 & 0.0625 \end{bmatrix} \begin{bmatrix} 0.85 \\ 1.16 \end{bmatrix} = 0.1305$$

$$-2 \ln \Lambda = 20 \ln [1 + 0.1305] = 2.4532$$

$$-2 \ln \Lambda \cap T_{(2,19)}^2$$

$$\approx \frac{(n-p)}{p(n-1)} T_{(2,19)}^2 \cap F_{(2,18)}$$

$$\frac{(n-p)}{p(n-1)} (-2 \ln \Lambda) \cap F_{(2,18)}$$

$$\frac{(n-p)}{p(n-1)} (-2 \ln \Lambda) = \frac{20-2}{2(20-1)} 2.4532 = 1.162$$

$F_{(2,18)}$  para  $\alpha = 0.05$  é  $\approx 3.5 > 1.162 \Rightarrow$  Aceitar  $H_0$

2.2 Ao nível  $\alpha=5\%$  testar  $H_0 : \Sigma = \Sigma_0 = \begin{bmatrix} 10 & 0 \\ 0 & 16 \end{bmatrix}$  versus  $H_1 : \Sigma \neq \Sigma_0$ .

Estatística de teste resultante do quociente de verosimilhança:

$$-2 \ln \Lambda = n \operatorname{Tr} (\Sigma_0^{-1} S) - n \ln |\Sigma_0^{-1} S| - np$$

$$\Sigma_0 = \begin{bmatrix} 10 & 0 \\ 0 & 16 \end{bmatrix} \quad \Sigma_0^{-1} = \begin{bmatrix} 0.1 & 0 \\ 0 & 0.0625 \end{bmatrix}$$

$$\Sigma_0^{-1} S = \begin{bmatrix} 0.1 & 0 \\ 0 & 0.0625 \end{bmatrix} \begin{bmatrix} 12 & 0.5 \\ 0.5 & 18 \end{bmatrix} = \begin{bmatrix} 1.2 & 0.05 \\ 0.0313 & 1.125 \end{bmatrix}$$

$$-2 \ln \Lambda = 20 [12 + 1.125] - 20 \ln(1.3484) - 20 \times 2 = 20 \times 2.3250 - 20 \times 0.2989 - 40 = 0.5220$$

Para  $\alpha = 0.05$

$$\chi_{\frac{1}{2}p(p+1)}^2 = \chi_{\frac{1}{2} \times 2 \times 3}^2 = \chi_3^2 = 7.81 > 0.5220$$

logo, não devemos rejeitar  $H_0$ .

3.

$$\bar{X} = \begin{bmatrix} 52.1 \\ 12.3 \\ 20.2 \end{bmatrix} \quad S = \begin{bmatrix} 5.5 & & \\ 10.1 & 50.2 & \\ -1.8 & -1.2 & 20.1 \end{bmatrix}$$

$$H_0 : \mu = \mu_0 = \begin{bmatrix} 50 \\ 15 \\ 20 \end{bmatrix} \quad \text{versus} \quad H_1 : \mu \neq \mu_0$$

Tendo  $n = 10$  e  $p = 3$ .

$$S^{-1} = \begin{bmatrix} 0.2988 & -0.0596 & 0.0232 \\ -0.0596 & 0.0318 & -0.0034 \\ 0.0232 & -0.0034 & 0.0516 \end{bmatrix}$$

$$T^2 = n(\bar{X} - \mu_0)' S^{-1} (\bar{X} - \mu_0) = 10 \begin{bmatrix} 2.1 & -2.7 & 0.2 \end{bmatrix} S^{-1} \begin{bmatrix} 2.1 \\ -2.7 \\ 0.2 \end{bmatrix} = 22.5028$$

$$\frac{(n-p)}{p(n-1)} T^2 = \frac{10-3}{3(10-1)} 22.5028 = 5.8341$$

Para  $\alpha = 0.05$ ,  $F_{(p,n-p)} = F_{(3,7)} = 4.35 < 5.8341$

A decisão será rejeitar a hipótese nula.

4.  $n = 75$ ,  $p = 2$

$$\bar{X} = \begin{bmatrix} 52.37 \\ 55.76 \end{bmatrix} \quad S = \begin{bmatrix} 23.44 & 15.62 \\ 15.62 & 27.39 \end{bmatrix}$$

Testar  $H_0 : \mu = \mu_0 = \begin{bmatrix} 52 \\ 52 \end{bmatrix}$  versus  $H_1 : \mu \neq \mu_0$  com  $\alpha = 0.05$

$$S^{-1} = \begin{bmatrix} 0.0688 & -0.0392 \\ -0.0392 & 0.0589 \end{bmatrix}$$

$$T^2 = n(\bar{X} - \mu_0)' S^{-1} (\bar{X} - \mu_0) = 75 \begin{bmatrix} 0.37 & 3.76 \end{bmatrix} \begin{bmatrix} 0.0688 & -0.0392 \\ -0.0392 & 0.0589 \end{bmatrix} \begin{bmatrix} 0.37 \\ 3.76 \end{bmatrix} = 7.3278$$

$$\frac{(n-p)}{p(n-1)} T^2 = \frac{75-2}{2(75-1)} 7.3278 = 3.6144$$

Para  $\alpha = 0.05$ ,  $F_{(p,n-p)} = F_{(2,73)} \approx 3.135 < 3.6144$

A decisão será rejeitar a hipótese nula.

5.  $X \sim N_2(\mu, \Sigma)$ ,  $n = 75$ ,  $p = 2$ ,  $S = \begin{bmatrix} 22.56 & 15.86 \\ 15.86 & 17.43 \end{bmatrix}$

Com  $\alpha = 1\%$  testar:  $H_0 : \Sigma = \Sigma_0 = \begin{bmatrix} 20 & 0 \\ 0 & 20 \end{bmatrix}$  versus  $H_1 : \Sigma \neq \Sigma_0$ .

$$\Sigma_0^{-1} = \begin{bmatrix} 0.05 & 0 \\ 0 & 0.05 \end{bmatrix} \quad \Sigma_0^{-1} S = \begin{bmatrix} 1.1280 & 0.7930 \\ 0.7930 & 0.8715 \end{bmatrix}$$

Estatística de teste resultante do quociente de verossimilhança:

$$\begin{aligned} -2 \ln \Lambda &= n \operatorname{Tr}(\Sigma_0^{-1} S) - n \ln |\Sigma_0^{-1} S| - np = \\ &= 75(1.1280 + 0.8715) - 75 \ln(0.3542) - 75 \times 2 = 77.8039 \end{aligned}$$

Para  $\alpha = 0.01$

$$\chi_{\frac{1}{2}p(p+1)}^2 = \chi_{\frac{1}{2} \times 2 \times 3}^2 = \chi_3^2 = 11.34 < 77.8039$$

logo, devemos rejeitar  $H_0$ .

6.  $n = 36$ ,  $p = 3$ ,  $\bar{X} = \begin{bmatrix} 0.2 \\ -0.5 \\ -0.9 \end{bmatrix}$  e  $S = \begin{bmatrix} 1.2 & -0.9 & 1.1 \\ -0.9 & 2 & -0.8 \\ 1.1 & -0.8 & 5.1 \end{bmatrix}$

**6.1** Com  $\alpha = 0.01$  testar  $H_0 : \mu = \mu_0 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$  versus  $H_1 : \mu \neq \mu_0$

$$T^2 = n (\bar{X} - \mu_0)' S^{-1} (\bar{X} - \mu_0)$$

$$S^{-1} = \begin{bmatrix} 1.4696 & 0.5703 & -0.2275 \\ 0.5703 & 0.7548 & -0.0046 \\ -0.2275 & -0.0046 & 0.2444 \end{bmatrix}$$

$$T^2 = 36 \begin{bmatrix} 0.2 & -0.5 & -0.9 \end{bmatrix} S^{-1} \begin{bmatrix} 0.2 \\ -0.5 \\ -0.9 \end{bmatrix} = 14.7298$$

$$\frac{(n-p)}{p(n-1)} T^2 = \frac{36-3}{3(36-1)} 14.7298 = 4.6294$$

Para  $\alpha = 0.01$ ,  $F_{(p,n-p)} = F_{(3,33)} \approx 4.50 < 4.6294$

A decisão será rejeitar a hipótese nula de  $\mu = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

**6.2** Região de confiança a 99% para  $\mu$

$$n (\bar{X} - \mu_0)' S^{-1} (\bar{X} - \mu_0) \cap T_{(p,n-1)}^2$$

Para um nível de significância  $\alpha$

$$P \left[ n (\bar{X} - \mu_0)' S^{-1} (\bar{X} - \mu_0) < T_{(p,n-1)}^2 \right] = 1 - \alpha$$

ou

$$P \left[ n (\bar{X} - \mu_0)' S^{-1} (\bar{X} - \mu_0) < \frac{(n-1)p}{(n-p)} F_{(p,n-p)} \right] = 1 - \alpha$$

A região de confiança para  $\mu$  será dada por todos os valores de  $\left[ n (\bar{X} - \mu_0)' S^{-1} (\bar{X} - \mu_0) \right]^2$  menores que  $c = \frac{(n-1)p}{(n-p)} F_{(p,n-p)}$ .

O elipsoide define a região de confiança para  $\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix}$ :

$$36 \begin{bmatrix} 0.2 - \mu_1 & -0.5 - \mu_2 & -0.9 - \mu_3 \end{bmatrix} S^{-1} \begin{bmatrix} 0.2 - \mu_1 \\ -0.5 - \mu_2 \\ -0.9 - \mu_3 \end{bmatrix} \leq \frac{35 \times 3}{33} \times 4.5$$

<sup>2</sup>Forma quadrática que define um elipsoide com centro em  $\bar{X}$ .

### ATIVIDADE FORMATIVA 3 - Proposta de Resolução

1.  $n_1 = 25, n_2 = 25$  e  $p = 2$

$$\bar{X}_1 = \begin{bmatrix} 137.8 \\ 98.5 \end{bmatrix} \quad \bar{X}_2 = \begin{bmatrix} 112.3 \\ 111.6 \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$$

$$S_1 = \begin{bmatrix} 23.7^2 & 0.43 \times 23.7 \times 12.4 \\ 0.43 \times 23.7 \times 12.4 & 12.4^2 \end{bmatrix} = \begin{bmatrix} 561.7 & 126.4 \\ 126.4 & 153.7 \end{bmatrix}$$

$$S_2 = \begin{bmatrix} 519.8 & 132.9 \\ 132.9 & 141.6 \end{bmatrix}$$

Admitindo que as duas amostras foram retiradas de populações com as seguintes características  $N_2(\mu_j, \Sigma_j)$ .

1.1  $\alpha = 0.05$

$H_0 : \Sigma_1 = \Sigma_2$  versus  $H_1 : \Sigma_1 \neq \Sigma_2$

Matriz de covariância total:

$$\hat{\Sigma} = \frac{\sum_{i=1}^k (n_i - 1) S_i}{n - k} = \frac{W}{n - k} = S$$

com  $n = 50$  e  $k = 2$ .

$$S = \frac{24}{48} \begin{bmatrix} 516.7 & 126.4 \\ 126.4 & 153.7 \end{bmatrix} + \frac{24}{48} \begin{bmatrix} 519.8 & 132.9 \\ 132.9 & 141.6 \end{bmatrix} = \begin{bmatrix} 540.75 & 129.65 \\ 129.65 & 147.65 \end{bmatrix}$$

Teste M de Box

$$M = (n - k) \ln |S| - \sum_{i=1}^k (n_i - 1) \ln |S_i|$$

$$|S| = 63933 \text{ e } \ln |S| = 11.0514$$

$$|S_1| = 70356 \text{ e } \ln |S_1| = 11.1613$$

$$|S_2| = 55941 \text{ e } \ln |S_2| = 10.9321$$

$$C = 1 - \frac{2p^2 + 3p - 1}{6(p+1)(k-1)} \left( \sum_{i=1}^k \frac{1}{n_i - 1} - \frac{1}{n - k} \right) =$$

$$= 1 - \frac{(2 \times 4) + 3 \times 2 - 1}{6 \times 3 \times 1} \left( \frac{1}{24} + \frac{1}{24} - \frac{1}{48} \right) = 0.9549$$

$$M.C = [(48 \times 11.0514) - (24 \times 11.1613) - (24 \times 10.9321)] \times 0.9549 = 0.2154$$

$$M.C \sim \chi_{\frac{1}{2}p(p+1)(k-1)}^2$$

Para  $\alpha = 0.05$ :  $\chi_3^2 = 7.81 > 0.2154 \Rightarrow$  Não rejeitar  $H_0$ : igualdade de matrizes de covariância.

**1.2** Testar  $H_0 : \mu_1 = \mu_2$  (ou  $H_0 : \mu_1 - \mu_2 = 0$ ) versus  $H_1 : \mu_1 \neq \mu_2$ .  
Pela alínea anterior temos:

$$S = \begin{bmatrix} 540.75 & 129.65 \\ 129.65 & 147.65 \end{bmatrix}$$

e

$$(\bar{X}_1 - \bar{X}_2) = \begin{bmatrix} 137.8 \\ 98.5 \end{bmatrix} - \begin{bmatrix} 112.3 \\ 111.6 \end{bmatrix} = \begin{bmatrix} 25.5 \\ -13.1 \end{bmatrix}$$

$$\bar{X}_1 - \bar{X}_2 \cap N_3 \left( \mu_1 - \mu_2, \left( \frac{1}{25} + \frac{1}{25} \right) \Sigma \right)$$

$$(n_1 + n_2 - 2)S \cap W_3(48, \Sigma)$$

$$\frac{n_1 n_2}{n_1 + n_2} (\bar{X}_1 - \bar{X}_2)' S^{-1} (\bar{X}_1 - \bar{X}_2) \cap T_{(p, n_1 + n_2 - 2)}^2$$

$$T_{(p, n_1 + n_2 - 2)}^2 \cong \frac{(n_1 + n_2 - 2)p}{n_1 + n_2 - p - 1} F_{(p, n_1 + n_2 - p - 1)}$$

$$T^2 = \frac{25 \times 25}{50} \begin{bmatrix} 25.5 & -13.1 \end{bmatrix} \begin{bmatrix} 0.0023 & -0.0021 \\ -0.0021 & 0.0086 \end{bmatrix} \begin{bmatrix} 25.5 \\ -13.1 \end{bmatrix} = 54.6199$$

Com  $\alpha = 0.05$  temos:

$$\frac{(n_1 + n_2 - 2)p}{n_1 + n_2 - p - 1} F_{(p, n_1 + n_2 - p - 1)} = \frac{48 \times 2}{47} F_{(2, 47)} = 2.0426 \times 3.18 = 6.4953$$

Decisão: Como  $6.4953 < 54.6199 \Rightarrow$  Rejeitar  $H_0 : \mu_1 = \mu_2$ .

**2.**  $n = 10$  e  $p = 3$

$$\bar{X} = \begin{bmatrix} 850 \\ 160 \\ 110 \end{bmatrix} \quad S = \begin{bmatrix} 15 & 10.1 & -1.5 \\ 10.1 & 10.2 & -1.8 \\ -1.5 & -1.8 & 6.1 \end{bmatrix}$$

Considerando  $\alpha = 0.05$ , pretendemos testar:

$$H_0 : \mu = \mu_0 = \begin{bmatrix} 851 \\ 159 \\ 108 \end{bmatrix} \quad \text{versus} \quad H_1 : \mu \neq \mu_0$$

Estatística de teste:

$$T^2 = n (\bar{X} - \mu_0)' S^{-1} (\bar{X} - \mu_0) =$$

$$= 10 \begin{bmatrix} -1 & 1 & 2 \end{bmatrix} \begin{bmatrix} 0.2006 & -0.2004 & -0.0098 \\ -0.2004 & 0.3035 & 0.0403 \\ -0.0098 & 0.0403 & 0.1734 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix} = 17.9890$$

$$\frac{n-p}{p(n-1)} T^2 = \frac{7}{27} \times 17.9890 = 4.6638$$

Temos

$$\frac{n-p}{p(n-1)} T^2 \sim F_{(p, n-p)}$$

para  $\alpha = 0.05$ ,  $F_{(3,7)} = 4.35 < 4.6638 \Rightarrow$  Rejeitar  $H_0$ .

3.  $n_1 = 4$ ,  $n_2 = 7$  e  $p = 2$ . Com  $\alpha = 0.05$ , testar:

$$H_0 : \mu_1 = \mu_2 \text{ (ou } H_0 : \mu_1 - \mu_2 = 0) \text{ versus } H_1 : \mu_1 \neq \mu_2$$

com

$$\mu_1 = \begin{bmatrix} \mu_{11} \\ \mu_{21} \end{bmatrix} \text{ e } \mu_2 = \begin{bmatrix} \mu_{12} \\ \mu_{22} \end{bmatrix}$$

$$\bar{X}_1 = \begin{bmatrix} 10.25 \\ 11.75 \end{bmatrix} \quad \bar{X}_2 = \begin{bmatrix} 10 \\ 12 \end{bmatrix}$$

Estatística de teste:

$$T^2 = \frac{n_1 n_2}{n_1 + n_2} (\bar{X}_1 - \bar{X}_2)' S^{-1} (\bar{X}_1 - \bar{X}_2) \cap T^2_{(p, n_1 + n_2 - 2)}$$

$$S_1 = \frac{1}{n_1 - 1} \sum_{i=1}^{n_1} (\bar{X}_1 - \bar{X}_2) (\bar{X}_1 - \bar{X}_2)'$$

Idêntico para  $S_2$ . Supondo que  $\Sigma_1 = \Sigma_2 = \Sigma$  desconhecido, temos

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

$$S_1 = \begin{bmatrix} 4.9167 & 4.7500 \\ 4.7500 & 4.9167 \end{bmatrix} \quad S_2 = \begin{bmatrix} 5.6667 & -2.8333 \\ -2.8333 & 6.9529 \end{bmatrix} \quad S = \begin{bmatrix} 5.4167 & -0.3056 \\ -0.3056 & 6.2738 \end{bmatrix}$$

$$(\bar{X}_1 - \bar{X}_2) = \begin{bmatrix} 0.25 \\ -0.25 \end{bmatrix} \quad S^{-1} = \begin{bmatrix} 0.1851 & 0.0090 \\ 0.0090 & 0.1598 \end{bmatrix} \quad T^2 = 0.0520$$

$$\frac{n_1 + n_2 - p - 1}{(n_1 + n_2 - 2)p} T^2 \cap F_{(p, n_1 + n_2 - p - 1)}$$

$$\frac{4 + 7 - 2 - 1}{(4 + 7 - 2) \times 2} \times 0.0520 = 0.0231$$

Para  $\alpha = 0.05$ , temos  $F_{(2,8)} = 4.46 > 0.0231 \Rightarrow$  Aceitar  $H_0$ .

4. Podemos destacar os seguintes pontos:

- O elevado erro de Tipo I com a utilização de vários testes separados, ou seja, a probabilidade de rejeitar a hipótese nula sendo esta verdadeira é mais elevada, ultrapassando o de  $\alpha$  global estabelecido (ou necessário);
- Os testes multivariados incorporam informação entre as variáveis (e.g. correlações), considerando a matriz de variância / covariância na análise e decisão;
- Quando as variáveis são analisadas em conjunto poderão resultar em diferenças significativas e tal não acontecer quando são testadas separadamente. Neste caso os testes multivariados são mais potentes, diminuindo o erro de Tipo II;
- As decisões com testes multivariados podem ser diferentes do que se forem utilizados testes univariados (e mesmo contraditórias). O melhor processo pode ser realizar primeiro testes multivariados e em caso de se observar diferenças significativas, fazer de seguida testes univariados para ter uma melhor percepção das variáveis que contribuíram para a decisão.

5.  $n_1 = 14, n_2 = 10, p = 2, n = 24$  e  $k = 2$

$$\bar{X}_1 = \begin{bmatrix} 0.8336 \\ 75.5714 \end{bmatrix} \quad \bar{X}_2 = \begin{bmatrix} 1.67 \\ 203.60 \end{bmatrix}$$

$$S_1 = \frac{1}{n_1 - 1} \sum_{i=1}^{n_1} (X_{1i} - \bar{X}_1) (X_{1i} - \bar{X}_1)' = \begin{bmatrix} 0.1161 & 6.2970 \\ 6.2970 & 446.8791 \end{bmatrix}$$

$$S_2 = \begin{bmatrix} 0.3702 & 62.3744 \\ 62.3744 & 11404 \end{bmatrix}$$

$$S = \frac{(n_1 - 1)S_1 + (n_2 - 1)S_2}{n_1 + n_2 - 2} = \frac{13}{22}S_1 + \frac{9}{22}S_2 = \begin{bmatrix} 0.2201 & 29.2378 \\ 29.2378 & 4929.2 \end{bmatrix}$$

$$S^{-1} = \begin{bmatrix} 21.4437 & -0.1272 \\ -0.1272 & 0.0010 \end{bmatrix} \quad (\bar{X}_1 - \bar{X}_2) = \begin{bmatrix} -0.8364 \\ -128.0286 \end{bmatrix}$$

5.1 Com  $\alpha = 0.05$  testar  $H_0 : \Sigma_1 = \Sigma_2$  versus  $H_1 : \Sigma_1 \neq \Sigma_2$ .

Teste M de Box:

$$M = (n - k) \ln |S| - \sum_{i=1}^k (n_i - 1) \ln |S_i| =$$

$$= 22 \times 5.4375 - 13 \times 2.5034 - 9 \times 5.8038 = 34.8468$$

$$C = 1 - \frac{2p^2 + 3p - 1}{6(p+1)(k-1)} \left( \sum_{i=1}^k \frac{1}{n_i - 1} - \frac{1}{n - k} \right) =$$

$$= 1 - \frac{(2 \times 4) + 3 \times 2 - 1}{6 \times 3 \times 1} \left( \frac{1}{13} + \frac{1}{9} - \frac{1}{22} \right) = 0.8891$$

$$M.C = 34.8468 \times 0.8891 = 30.9825$$

$$M.C \cap \chi^2_{\left(\frac{1}{2}p(p+1)(k-1)\right)}$$

Para  $\alpha = 0.05, \chi^2_{(3)} = 7.81 < 30.9825 \Rightarrow$  Rejeitar  $H_0$ .

5.2 Testar  $H_0 : \mu_1 = \mu_2$  (ou  $\mu_1 - \mu_2 = 0$ ) versus  $H_1 : \mu_1 \neq \mu_2$ .

Estatística de teste:

$$T^2 = \frac{n_1 n_2}{n_1 + n_2} (\bar{X}_1 - \bar{X}_2)' S^{-1} (\bar{X}_1 - \bar{X}_2) \cap T^2_{(p, n_1 + n_2 - 2)}$$

$$T^2 = \frac{14 \times 10}{24} \begin{bmatrix} -0.8364 & -128.0286 \end{bmatrix} S^{-1} \begin{bmatrix} -0.8364 \\ -128.0286 \end{bmatrix} = 20.1399$$

Para  $\alpha = 0.05$

$$\frac{(n_1 + n_2 - 2)p}{n_1 + n_2 - p - 1} F_{(p, n_1 + n_2 - p - 1)} = \frac{22 \times 2}{21} F_{(2, 21)} = 2.0952 \times 3.49 = 7.3122$$

Decisão:  $7.3122 < 20.1399 \Rightarrow$  Rejeitar  $H_0$

6. <sup>3</sup>  $p = 2, k = 3, e n = n_1 + n_2 + n_3 = 12$ . Com  $\alpha = 0.05$  testar

$$H_0 : \begin{bmatrix} \mu_{11} \\ \mu_{21} \end{bmatrix} = \begin{bmatrix} \mu_{12} \\ \mu_{22} \end{bmatrix} = \begin{bmatrix} \mu_{13} \\ \mu_{23} \end{bmatrix} \text{ versus } H_1 : \mu_i \neq \mu_j (i \neq j)$$

Estatística de teste:

$$\Lambda = \frac{|W|}{|T|}$$

Onde  $W$  - matriz agregada da variabilidade dentro dos grupos

$$W = W_1 + W_2 + W_3$$

$$\begin{aligned} W_1 &= \sum_{i=1}^4 (X_{1i} - \bar{X}_1) (X_{1i} - \bar{X}_1)' = \\ &= \begin{bmatrix} -1 & -1 \end{bmatrix} \begin{bmatrix} -1 & -1 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \end{bmatrix} + \begin{bmatrix} 2 \\ 0 \end{bmatrix} \begin{bmatrix} 2 & 0 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} \begin{bmatrix} -1 & 1 \end{bmatrix} = \\ &= \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix} + \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 6 & 0 \\ 0 & 2 \end{bmatrix} \end{aligned}$$

De forma idêntica:

$$W_2 = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \quad W_3 = \begin{bmatrix} 6.8 & 2.6 \\ 2.6 & 5.2 \end{bmatrix}$$

$$W = W_1 + W_2 + W_3 = \begin{bmatrix} 14.8 & 1.6 \\ 1.6 & 9.2 \end{bmatrix}$$

Matriz de variabilidade total =  $T = W + B$ .

$$\bar{X}_1 = 68/12 = 5.67 \text{ e } \bar{X}_2 = 69/12 = 5.75 \text{ e } \bar{X} = \begin{bmatrix} 5.67 \\ 5.75 \end{bmatrix}.$$

$$\begin{aligned} B &= \sum_{j=1}^3 n_j (X^-_j - \bar{X}) (X^-_j - \bar{X})' = \\ &= 4 \begin{bmatrix} 3 - 5.67 \\ 4 - 5.75 \end{bmatrix} \begin{bmatrix} -2.67 & -1.75 \end{bmatrix} + 3 \begin{bmatrix} -0.67 \\ 1.25 \end{bmatrix} \begin{bmatrix} -0.67 & 1.25 \end{bmatrix} + 5 \begin{bmatrix} 2.53 \\ 0.65 \end{bmatrix} \begin{bmatrix} 2.53 & 0.65 \end{bmatrix} = \\ &= \begin{bmatrix} 61.87 & 24.40 \\ 24.40 & 19.05 \end{bmatrix} \end{aligned}$$

$$T = W + B = \begin{bmatrix} 76.67 & 26.00 \\ 26.00 & 28.25 \end{bmatrix}$$

$$\Lambda = \frac{|W|}{|T|} = \frac{14.8 \times 9.2 - (1.6)^2}{76.67 \times 28.25 - (26)^2} = 0.0897$$

Aproximação à distribuição  $\chi^2$  com  $p(k-1)$  graus de liberdade

$$-[(n-1) - 0.5(p+k)] \ln \Lambda = -[(12-1) - 0.5(2+3)] \ln(0.0897) = 20.4987$$

Para  $\alpha = 0.05, \chi_{p(k-1)}^2 = \chi_4^2 = 9.49 < 20.4987 \Rightarrow$  Rejeitar  $H_0$

FIM

<sup>3</sup> Adaptado de Reis, E. (2001) "Estatística Multivariada Aplicada", 2ª Edição, Edições Sílabo, Lisboa, Portugal