

Estatística e Delineamento Experimental 2024-25

4 Exercícios de Delineamento Experimental e de Análise de Variância de Efeitos Fixos

1. Pretende-se estudar o efeito de 4 concentrações de cálcio no solo (C1, C2, C3 e C4) no crescimento das raízes de estacas da variedade de oliveira Galega. Para tal dispõe-se de 20 vasos, cada um contendo uma planta. Assumindo condições externas homogêneas, o técnico responsável pelo estudo irá aplicar cada um dos 4 tratamentos aleatoriamente a 5 vasos.
 - (a) Indique as unidades experimentais do ensaio.
 - (b) Indique o número de repetições de cada tratamento.
 - (c) Indique o tipo de delineamento experimental do ensaio.
 - (d) Gere no R, com recurso ao *package agricolae*, uma casualização possível para este ensaio.
 - (e) Se o investigador concluir que será melhor fazer o ensaio em estufa e havendo uma variação de temperatura na bancada (menor temperatura perto da entrada e maior no seu interior), como poderá ele controlar essa fonte de variação? Indique um tipo de delineamento experimental que considere plausível e gere no R, com recurso ao *package agricolae*, uma possível aleatorização e apresente um esquema possível para essa experiência.
2. Foi efetuado um estudo para investigar os efeitos de quatro tipos de fungicidas (F1, F2, F3, F4) no rendimento da batata em determinadas parcelas de campo. Foi também incluído um tratamento de controlo sem fungicida (C) para permitir uma comparação de base. O campo designado para o ensaio, apresenta heterogeneidade, mas foi possível identificar seis diferentes parcelas de terreno homogêneas. Qual o delineamento experimental que deverá adotar? Gere no R, com recurso ao *package agricolae*, uma possível aleatorização e apresente um esquema possível para essa experiência.
3. Foi conduzida uma experiência para investigar os efeitos do tipo de solo (argila, G e areia, R) e da disponibilidade de água (baixa, B e alta, A) no crescimento do tremçoço. Foram, então, aplicados 4 quatro tratamentos, designados por GB, GA, RB, RA, representando diferentes combinações de tipo de solo (G e R) e de quantidade de água fornecida (B e A). A experiência foi conduzida em 16 vasos numa estufa, cada vaso com 2 plantas, aos quais foram aleatoriamente aplicados os tratamentos de modo a ter para cada tratamento o mesmo número de repetições.
 - (a) Indique as unidades experimentais do ensaio.
 - (b) Esta experiência tem pseudo-repetições?
 - (c) Qual o tipo de delineamento experimental desta experiência?
 - (d) Quantas repetições existem por tratamento?
 - (e) Gere no R, com recurso ao *package agricolae*, uma possível aleatorização e apresente um esquema para esta experiência.
 - (f) Após uma observação mais atenta das condições existentes na estufa, verificou-se que existe um gradiente de temperatura ao longo da largura da bancada e que existem diferentes níveis de luz ao longo do comprimento da bancada. As linhas e as colunas do conjunto de vasos deverão ser, por conseguinte, consideradas como fatores de controlo dessas fontes de variação.
 - i. Qual o tipo de delineamento experimental constante desta experiência?
 - ii. Gere no R, com recurso ao *package agricolae*, uma possível aleatorização e apresente um esquema para esta experiência.

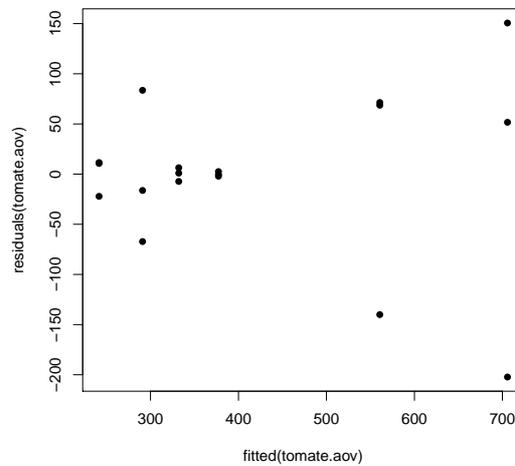
4. Um pomar tem 5 linhas seguidas, cada uma com 50 plantas de uma variedade de macieira (portanto, 5 variedades distintas em estudo). Pretende-se estudar o efeito da variedade na produção de fruto.
 - (a) Considera que tem condições que permitam responder ao objetivo do estudo?
 - (b) No seu entender como deverá delinear uma experiência que responda ao problema formulado?
5. Pretende-se instalar um ensaio de campo para estudar os efeitos competitivos de quatro espécies diferentes de infestantes (A, B, C, D) no milho sob dois níveis de stress hídrico (presença, I, ou ausência de irrigação, NI). Os tratamentos associados ao stress hídrico só podem ser aplicados a grandes áreas de terreno, enquanto que as diferentes espécies de infestantes podem ser aplicadas a pequenas parcelas. O terreno onde será instalado o ensaio apresenta 4 grandes manchas de solo associadas a diferentes níveis de fertilidade. Na sua opinião, qual deverá ser tipo de delineamento a adotar? Gere no R, com recurso ao *package agricolae*, uma possível aleatorização e apresente um esquema para esta experiência.
6. Pretende-se comparar o sucesso de enraizamento de estacas de 4 clones de eucalipto (E1, E2, E3, E4). De cada clone existem 80 estacas disponíveis para conduzir a experiência, sendo que o técnico A propõe colocar as 80 estacas de cada clone num único tabuleiro e o técnico B propõe fazer 8 tabuleiros por clone, cada um com 10 estacas.
 - (a) Qual dos técnicos propõe o procedimento correcto? Justifique.
 - (b) Assumindo que será seguido o processo correcto, os tabuleiros serão colocados num gabinete em ambiente controlado. Nesse gabinete existem diferentes prateleiras, pelo que se verifica que a disponibilidade de luz varia de acordo com a prateleira.
 - i. Esta experiência tem pseudo-repetições?
 - ii. Indique as unidades experimentais do ensaio.
 - iii. Quantas repetições tem a experiência?
 - iv. De acordo com a descrição das condições existentes para efetuar o estudo, qual o delineamento experimental que considera adequado de modo a comparar de forma não enviesada o efeito do clone no enraizamento? Gere no R, com recurso ao *package agricolae*, uma possível casualização e apresente um esquema para esta experiência.
7. Através de um estudo pretende-se comparar a acidez total do mosto de 3 variedades de videira (Touriga Nacional, V1; Touriga Fêmea, V2; Touriga Franca, V3) cultivadas numa quinta na região do Douro e dispostas no terreno segundo um delineamento experimental adequado a tal comparação. De cada variedade são levadas para o laboratório da adega 6 amostras de bagos (6 sacos com 100 bagos). Sabe-se que a precisão com que é feita a medição da acidez total é afetada pelo operador, no entanto, para maior rapidez na obtenção dos resultados, essas análises serão feitas por 6 técnicos diferentes (A, B, C, D, E, F). Qual o delineamento experimental que acha conveniente adotar no laboratório da adega? Gere no R, com recurso ao *package agricolae*, uma possível aleatorização e apresente um esquema para esta experiência.
8. Pretende-se comparar propriedades organolépticas de sumos de maçã de 8 marcas comerciais (S1, S2, S3, S4, S5, S6, S7, S8), organizando-se para tal uma sessão de avaliação dessas características com 10 provadores (P1, P2, P3, P4, P5, P6, P7, P8, P9, P10). Devido à subjetividade das respostas de cada provador, qual o delineamento experimental que considera adequado para garantir que os resultados da prova não sejam enviesados? Gere no R, com recurso ao *package agricolae*, uma possível aleatorização.
9. No melhoramento de variedades tradicionais de tomate, uma característica importante é a resistência da película. Esta característica foi avaliada em 6 variedades de tomate, num ensaio envolvendo 18 parcelas de tomateiros. Concretamente, cada variedade foi associada aleatoriamente a 3 parcelas de tomateiros. Cada observação individual é constituída pela resistência média dos frutos de cada

parcela (medida num texturómetro, em grama força, *gf*). Eis os valores obtidos em cada parcela (que se encontram no ficheiro disponível na página *web* da disciplina de nome `tomate.txt`), bem como as médias e variâncias para cada variedade e para a totalidade das observações:

Variedade	Observações			Média	Variância
18	632.04	629.30	420.59	560.6433	14 713.08
28	253.00	219.34	252.11	241.4833	367.9434
29	223.71	374.48	274.66	290.9500	5881.921
40C	503.51	757.44	856.39	705.7800	33 132.64
Ace	375.18	376.81	379.77	377.2533	5.414433
Roma	333.05	324.82	338.45	332.1067	47.11163

- A média global das observações é $\bar{y}_{..} = 418.0361$;
- a variância amostral da totalidade das observações é $s_y^2 = 34 517.82$.

- Indique qual o tipo de delineamento experimental em causa. Explícite o modelo correspondente e todas as hipóteses adicionais que sejam necessárias à consideração do problema em estudo.
- Construa a tabela-resumo da análise de variância correspondente a este caso,
 - utilizando apenas uma máquina de calcular e a informação disponível neste enunciado;
 - utilizando, no R, o comando `summary(aov(res.pel ~ variedade , data=tomate))`.
- Formalize e efectue o teste F adequado ao problema acima referido, com um nível de significância de 5%. Pode afirmar-se que a resistência média da película não é sempre igual, em todas as variedades?
- Qual o maior nível de significância α para o qual mudaria a sua resposta na alínea anterior? Como se designa esse valor?
- Utilize o comando `model.matrix` do R para inspeccionar a natureza da matriz do modelo, \mathbf{X} , neste contexto.
- Utilize o comando `fitted` do R para identificar os valores ajustados da variável resposta, nesta Análise de Variância.
- Calcule as estimativas dos parâmetros do modelo e interprete os valores obtidos. No final, utilize os comandos: `coef(tomate.aov)` e `tomate.lm<-lm(res.pel ~ variedade , data=tomate)`.
- Quais as variedades cuja resistência difere significativamente da resistência da variedade 40C? Justifique com um teste adequado, ao nível $\alpha = 0.05$. No final, execute o comando `library(agricolae)` do R e utilize os comandos: `tomate.aov<-aov(res.pel ~ variedade , data=tomate)` e `HSD.test(tomate.aov, "variedade", console=TRUE)`.
- O gráfico dos resíduos (usuais) das observações, contra os valores ajustados pelo modelo de análise de variância, é apresentado a seguir. Comente o gráfico e as suas possíveis implicações. Identifique a observação cujo resíduo é, em módulo, mais elevado.



10. Um estudo sobre três variedades de café, referenciadas por CA, CL e PR, efetuado de acordo com um delineamento totalmente casualizado, focou-se sobre os comprimentos dos estomas das respectivas folhas. De cada variedade foram seleccionadas 12 plantas, e para cada planta foi medido o comprimento médio dos estomas das suas folhas em condições ambientais controladas (variável Comprimento, em μm). São conhecidas apenas as médias e variâncias das 12 observações (plantas) de cada variedade:

	CA	CL	PR
Média	22.85833	19.49333	25.31583
Variância	13.69303	2.725424	9.388936

- Explicite pormenorizadamente o modelo ANOVA adequado ao estudo do problema.
 - Construa a tabela-resumo da ANOVA que indicou na alínea anterior.
 - Qual é a variância amostral dos comprimentos dos estomas na totalidade das 36 observações?
 - É possível afirmar que, na população, o comprimento médio dos estomas é igual nas três variedades, para um nível de significância $\alpha = 0.05$? Responda pormenorizadamente.
 - Compare todos os pares de médias através dum teste de Tukey ($\alpha = 0.05$). Comente.
11. Sabe-se que o dióxido de carbono tem um efeito crítico no crescimento de populações microbianas; pequenas quantidades de CO_2 podem estimular o crescimento de algumas espécies enquanto que, pelo contrário, grandes concentrações têm de forma geral uma acção inibitória. Este último efeito é usado comercialmente para preservar alimentos armazenados.

Realizou-se um estudo para investigar a acção de diferentes concentrações de CO_2 na taxa de crescimento de *Pseudomonas fragi*; os diferentes níveis (tratamentos) foram pré-fixados e a variável resposta medida foi a percentagem de variação na massa das culturas após uma hora de crescimento nas respectivas condições, originando os dados da seguinte tabela.

Concentração de CO_2				
0.0	.083	.29	.50	.86
62.6	50.9	45.5	29.5	24.9
59.6	44.3	41.1	22.8	17.2
64.5	47.5	29.8	19.2	7.8
59.3	49.5	38.3	20.6	10.5
58.6	48.5	40.2	29.2	17.8
64.6	50.4	38.5	24.1	22.1
50.9	35.2	30.2	22.6	22.6
56.2	49.9	27.0	32.7	16.8
52.3	42.6	40.0	24.4	15.9
62.8	41.6	33.9	29.6	8.8

Estes dados estão disponíveis na página *web* da disciplina no ficheiro de nome `C02.txt`, sendo as concentrações de CO_2 repetidas em duas colunas: numa sob a forma de factor e noutra sob a forma de variável numérica.

- (a) Pretende-se testar a hipótese nula $H_0 : \mu_1 = \mu_2 = \dots = \mu_5$, onde μ_i indica a taxa de crescimento esperada para a i -ésima concentração de CO_2 . É sugerida a utilização de uma Análise de Variância. Enuncie os pressupostos necessários para poder efectuar o teste referido.
 - (b) Haverá evidência suficiente para rejeitar H_0 com um nível de significância de $\alpha = .05$?
 - (c) Caso tenha respondido afirmativamente na alínea anterior, indique os pares de concentrações de CO_2 que têm efeitos diferentes sobre o crescimento do *Pseudomonas fragi*, recorrendo aos intervalos de confiança de Tukey.
 - (d) Estude a validade dos pressupostos do modelo ANOVA.
 - (e) Dada a natureza da variável preditora, também poderia ser considerada uma regressão linear das taxas de crescimento sobre as concentrações de dióxido de carbono, encaradas como uma variável numérica. Utilizando a coluna de CO_2 com as concentrações dadas como variáveis numéricas (isto é, a coluna `C02.numerico`), responda às seguintes questões.
 - i. Construa a nuvem de pontos da variação de massa sobre concentração de CO_2 .
 - ii. Ajuste a regressão linear simples referida, traçando a recta de regressão sobre a nuvem de pontos. Comente.
 - iii. Compare os resultados do teste F de ajustamento global obtidos, quer no modelo com o factor `C02`, quer no modelo com o preditor numérico `C02.numerico`.
12. Pretende-se avaliar se uma técnica de processamento de feijão verde altera a concentração de zinco neste alimento. Há especial interesse em avaliar a referida concentração em quatro fases do processamento: no início; imediatamente antes do chamado branqueamento; logo após o branqueamento; e no final do processamento.

Uma vez que se sabe que diferentes lotes de feijão diferem nas concentrações iniciais de zinco, decidiu-se utilizar nove diferentes lotes, medindo as respectivas concentrações nas quatro fases do processamento acima indicadas. As medições obtidas estão indicadas na Tabela (e no ficheiro disponível na página *web* da disciplina de nome `zinco.txt`).

Lote	Fases do processamento			
	1	2	3	4
1	2.23	3.71	2.53	5.46
2	2.20	4.67	2.87	5.19
3	2.44	3.45	2.83	5.51
4	2.11	2.73	2.33	4.82
5	2.30	2.58	2.19	6.63
6	1.72	1.85	1.80	2.39
7	1.78	1.81	1.75	2.09
8	2.36	2.32	1.83	2.27
9	2.91	2.50	1.97	2.39

- (a) Identifique o delineamento experimental utilizado e descreva o respectivo modelo ANOVA (de efeitos fixos).
- (b) Construa a tabela de Análise de Variância para o caso em estudo.
- (c) Efectue o teste F para saber se diferentes fases do processamento afectam as concentrações médias de zinco ($\alpha = 0.05$). Comente.
- (d) Calcule o valor ajustado para a concentração de zinco no início do processamento, para o primeiro lote. Compare com o valor observado e comente.
- (e) Por engano, um analista considerou que existia apenas um factor no estudo, correspondendo às diferentes fases do processamento. Compare as tabelas-resumo da ANOVA obtidas em cada caso e comente.
13. Em 2015 foi efectuado um ensaio em Vila Real envolvendo 6 clones da casta Aragonez, também conhecida no Douro por Tinta Roriz, num trabalho do ISA e da PORVID (Associação Portuguesa para a Diversidade da Videira). Pretendeu-se medir o chamado Índice de Fertilidade Potencial (variável *if*, definida como o número de cachos por número de olhos abrolhados) e saber se este difere consoante os génotipos estudados. Uma vez que se sabe que a fertilidade está também dependente do tipo de solo, foram escolhidos para o ensaio seis terrenos com solos diferentes. Cada terreno foi dividido em seis parcelas, sendo cada uma dessas parcelas associada, aleatoriamente, a um diferente génotipo. O ficheiro com os dados encontra-se disponível na página *web* da disciplina (*rzif.txt*). Foi ajustado um modelo ANOVA com os seguintes resultados:

	Df	Sum Sq	Mean Sq	F value
clone	??	??	??	4.204
terreno	??	??	??	2.691
Residuals	??	??	0.3660	

- (a) Indique o tipo de delineamento experimental utilizado, descrevendo em pormenor o modelo ANOVA correspondente à tabela-resumo acima indicada.
- (b) Complete a tabela-resumo, indicando como obtém os valores em falta.
- (c) Seguidamente, ajuste o modelo de análise de variância correspondente no R.
- (d) Diga, justificando formalmente, que tipo de efeitos são significativos, ao nível $\alpha = 0.01$.
- (e) Um analista que apenas conhecia as ANOVAs a um factor, decidiu estudar os efeitos de génotipo, com os mesmos dados. Construa a tabela-resumo resultante e indique a que conclusão chegaria (ao mesmo nível $\alpha = 0.01$). Comente.
- (f) Outro analista afirma que há génotipos que se dão particularmente bem em determinados terrenos estudados, e sugere que essa possibilidade deve ser tida em conta pelo modelo. Com base na experiência realizada é possível prever essa eventualidade?

14. Pretende-se comparar o rendimento obtido com quatro variedades de trigo. Identificaram-se 13 terrenos com características de solos diferentes, que correspondem aos tipos de terrenos nos quais se pretende fazer as culturas. Os 13 terrenos são então divididos em quatro parcelas de igual dimensão. Em cada terreno associa-se, de forma aleatória, uma parcela a cada uma das quatro variedades. Após a colheita registam-se os rendimentos obtidos (em t/ha) na tabela (e disponíveis na página *web* da disciplina no ficheiro `terrenos.txt`).

- (a) Identifique o delineamento experimental adotado e ajuste no R o correspondente modelo de análise de variância.
- (b) As médias amostrais de cada variedade sugerem que há variedades com desempenho superior. Mas serão essas diferenças significativas? A fim de responder, efectue uma Análise da Variância adequada, construindo a tabela-resumo correspondente. Comente as suas conclusões.
- (c) Teste se, entre terrenos, existem diferenças significativas, como seria de supôr. Comente.
- (d) Calcule o rendimento ajustado para a variedade A no terreno I. Compare com o rendimento observado e comente.

Terreno	Variedade			
	A	B	C	D
I	1.800	2.457	0.722	0.789
II	1.709	1.839	1.546	1.304
III	1.277	1.293	1.515	1.273
IV	1.675	1.745	0.800	0.846
V	1.814	1.833	1.678	1.732
VI	1.896	1.203	1.192	1.580
VII	1.078	1.689	1.583	1.168
VIII	1.740	1.518	1.050	1.305
IX	1.200	1.133	0.778	1.033
X	1.500	0.722	0.636	0.925
XI	1.932	1.700	1.203	0.850
XII	1.169	1.209	1.112	0.986
XIII	1.438	1.577	1.355	1.525
Médias	1.556	1.532	1.167	1.178
Variâncias	0.0879	0.1855	0.1266	0.0934

15. Num estudo sobre características de crescimento de pinheiro manso, conduzido em Sines e em Tavira pelo Instituto Nacional de Investigação Agrária e Veterinária (INIAV), avaliou-se a altura média de pinheiros de cinco diferentes proveniências (Marrocos, Grécia, Portugal e duas proveniências de Itália), dois anos após a plantação. Quer em Sines, quer em Tavira, foram plantados seis talhões com árvores de cada proveniência, gerando assim $n=60$ valores de alturas (variável `alt2`, em cm), cuja variância amostral é $s^2 = 34.49584$ (dados disponíveis na página *web* da disciplina no ficheiro `pmanso.txt`). Eis algumas médias resultantes.

prov					local	
Grecia	Italia-1	Italia-2	Marrocos	Portugal	Sines	Tavira
28.81	32.75	30.23	35.13	31.90	28.14	35.38
prov:local						
	local		Grand mean			
prov	Sines	Tavira	31.76298			
Grecia	22.52	35.10				
Italia-1	31.03	34.46				
Italia-2	26.91	33.56				

Marrocos	31.16	39.09
Portugal	29.09	34.70

- (a) Identifique o delineamento experimental utilizado e o modelo ANOVA adequado. Descreva pormenorizadamente o modelo.
- (b) Sabendo que o Quadrado Médio Residual é 16.59 e que a Soma de Quadrados associada às cinco diferentes proveniências é 280.61, construa a tabela-resumo do modelo ANOVA adequado.
- (c) Ajuste o modelo de análise de variância descrito na alínea a) no R.
- (d) Use um teste F para avaliar a existência de efeitos de interação entre proveniência dos pinheiros e local do ensaio. Comente as suas conclusões. Indique brevemente que outros tipos de efeitos devem ser considerados significativos. Considere $\alpha = 0.05$.
- (e) Na amostra, a maior altura média em Sines é inferior à menor altura média em Tavira. Independentemente das suas respostas nas alíneas anteriores, use o teste de Tukey para indicar se igual afirmação se pode estender à população. Comente. No final execute os comandos do R necessários para o obter os resultados do teste de Tukey.
16. Uma experiência visa estudar o rendimento duma variedade de trigo em função de diferentes formas de aplicar dois adubos, um com fósforo (um adubo fosfatado), e outro com potássio. Consideram-se três dosagens de aplicação do adubo fosfatado, designadas por Baixa, Média e Elevada. Igualmente, consideram-se três dosagens de aplicação do fertilizante com potássio, igualmente designadas por Baixa, Média e Elevada. A experiência realiza-se num terreno com 27 parcelas de igual dimensão. Repartem-se, de forma totalmente casualizada, três parcelas por cada combinação de dosagem de um e outro fertilizante. Os resultados obtidos (em t/ha) foram os seguintes:

		Potássio (K)									Média	Variância
		Baixa			Média			Elevada				
Fósforo (P)	Baixa	4.6	4.9	4.3	6.3	6.1	6.4	6.6	6.7	6.9	5.8667	0.9775
	Média	5.4	5.6	5.2	6.8	5.7	6.7	7.5	8.0	7.3	6.4667	1.045
	Elevada	5.3	5.7	5.1	7.5	7.0	7.2	7.1	7.4	6.1	6.4889	0.88861
Média		5.1222			6.6333			7.0667				
Variância		0.20944			0.3200			0.3175				

As médias observadas para cada combinação de dosagens de cada tipo de fertilizante foram as seguintes:

		Potássio		
		Baixa	Média	Elevada
Fósforo	Baixa	4.600	6.267	6.733
	Média	5.400	6.400	7.600
	Elevada	5.367	7.233	6.867

A Tabela-Resumo associada a esta experiência é a seguinte:

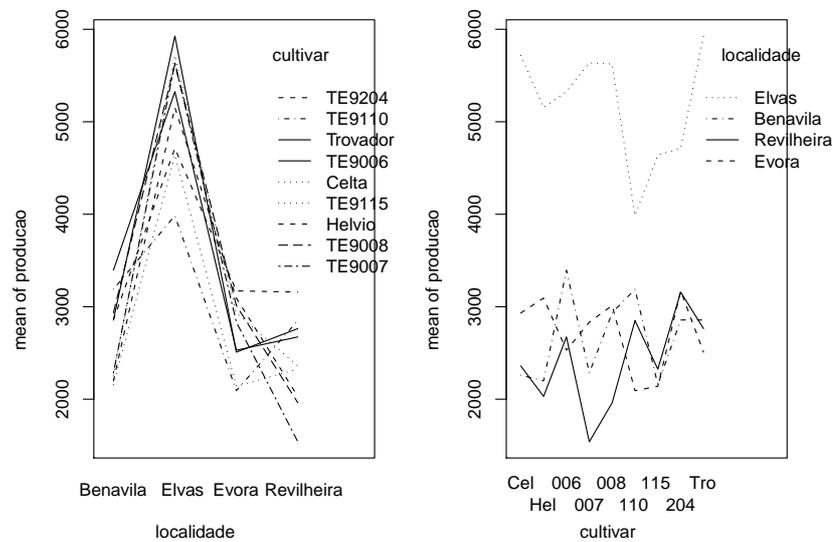
	Df	Sum of Sq	Mean Sq	F Value	Pr(F)
fosforo	?	2.24296	1.121481	?	0.00366530
potassio	2	18.75630	?	?	0.00000001
fosforo:potassio	?	1.93926	0.484815	3.36504	0.03187154
Residuals	18	2.59333	?		

Responda às seguintes questões, utilizando a informação disponível no enunciado.

- (a) Complete a Tabela-Resumo, indicando como obtém cada um dos valores omissos.
 - (b) Que *tipo* de efeitos do modelo associado a este caso devem ser considerados significativos? Justifique, explicitando as hipóteses dos testes que efectuou, as estatísticas dos testes e os níveis de significância utilizados, bem como a natureza das regiões críticas, os valores obtidos e as conclusões.
 - (c) Diga se considera que os rendimentos justificam uma opção entre as duas hipóteses seguintes: (i) Elevada dosagem de potássio com Baixa dosagem de fósforo; e (ii) Elevada dosagem de potássio com Média dosagem de fósforo? Justifique.
 - (d) Ajuste agora um modelo a dois factores, mas que não preveja os efeitos de interacção. Construa a tabela-resumo correspondente. Identifique as diferenças entre esta tabela e a que se indicou acima (associada ao modelo que prevê efeitos de interacção). Comente as diferenças e identifique as consequências de não prever a existência de efeitos de interacção quando na realidade esses efeitos parecem existir.
17. Uma engenheira agrónoma pretende seleccionar cultivares de trigo para as quatro explorações agrícolas pelas quais é responsável, que se localizam em Elvas, Évora, Benavila e Revilheira. Em cada uma destas explorações, definem-se 36 parcelas de terra, associando aleatoriamente quatro parcelas a cada uma de nove cultivares: Celta, Helvio, TE9006, TE9007, TE9008, TE9110, TE9115, TE9204 e Trovador. Em cada parcela foi medido o rendimento, em kg/ha. A variância da totalidade dos rendimentos observados é $s^2 = 1\,714\,242$.
- (a) Especifique o delineamento experimental utilizado, e descreva em pormenor o modelo ANOVA adequado a esta experiência.
 - (b) Foi ajustado um modelo ANOVA, com o programa R. Resultados parciais desse ajustamento são dados de seguida.

	Df	Sum Sq	Mean Sq	F value
localidade	???	183759916	???	234.9531
cultivar	???	???	964060	???
localidade:cultivar	???	???	???	4.0768
Residuals	???	28156076	260704	

- i. Complete a tabela, indicando como obtém cada um dos valores omissos.
- ii. Qual o valor estimado da variância dos erros aleatórios do modelo, e quais as suas unidades de medida?
- iii. Teste formalmente (a um nível de significância $\alpha = 0.01$) quais os tipos de efeitos do modelo que devem ser considerados significativos. Descreva um teste em pormenor e discuta os restantes de forma sintética.
- iv. Os gráficos de interacção associados a esta experiência são os seguintes. Comente-os, relacionando as suas conclusões das alíneas anteriores com os gráficos.



- v. Discuta o efeito de mudar as unidades de medida da variável resposta de kg/ha para toneladas por hectare. Quais os valores da tabela que se alteram, e quais os que ficam iguais? Quais os efeitos da mudança de unidades nas conclusões dos testes F ?
- (c) Os rendimentos médios das quatro parcelas correspondentes a cada combinação de localidade e cultivar foram os seguintes.

localidade	cultivar								
	Celta	Helvio	TE9006	TE9007	TE9008	TE9110	TE9115	TE9204	Trovador
Benavila	2260	2198	3395	2283	2935	3188	2157	2857	2857
Elvas	5719	5154	5326	5636	5624	3984	4640	4713	5927
Evora	2932	3094	2530	2833	3013	2092	2140	3171	2509
Revilheira	2360	2030	2673	1540	1962	2850	2325	3158	2763

Utilizando o teste de Tukey, indique quais as cultivares que, em Elvas, não diferem significativamente do melhor rendimento observado nessa localidade ($\alpha = 0.10$). Comente.

18. Com o objectivo de analisar as alterações no conteúdo em taninos da polpa de sapotis (frutos do sapotizeiro, *Manilkara achras*) provocadas pela temperatura de conservação (alta/baixa) e pelo tempo de armazenamento (0, 3, 6 ou 9 dias) foi efectuado um estudo que forneceu os seguintes dados:

Temperatura	Tempo							
	0 dias		3 dias		6 dias		9 dias	
alta	20.8	19.7	26.5	27.5	26.5	26.4	26.5	26.9
	18.0	19.5	27.0	26.4	27.0	24.0	25.9	26.3
baixa	32.3	34.1	20.8	20.5	16.4	15.7	10.3	9.7
	30.7	31.8	21.0	20.9	15.9	16.0	7.8	9.8

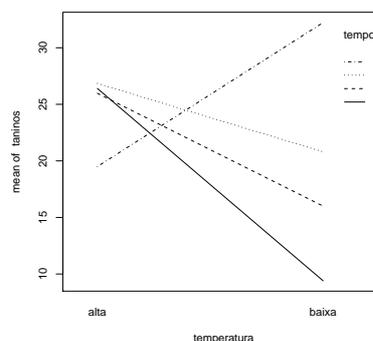
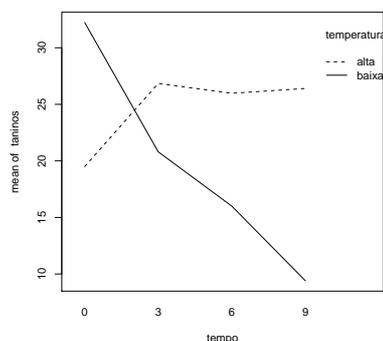
A média e a variância do conjunto das 32 observações são 22.14375 e 47.83222, respectivamente. As médias associadas a cada tempo de armazenamento, cada temperatura e cada combinação de tempo e temperatura, são:

Tables of means

tempo				tempo:temperatura	
0	3	6	9	temperatura	
25.862	23.825	20.987	17.900	tempo	alta baixa
				0	19.50 32.23
				3	26.85 20.80
				6	25.97 16.00
				9	26.40 9.40



- Identifique o delineamento experimental utilizado no estudo e descreva de forma pormenorizada o melhor modelo ANOVA que lhe está associado.
- Sabendo que a Soma dos Quadrados dos Resíduos é 20.72 e que o Quadrado Médio associado aos diferentes tempos de armazenamento é 96.01, construa o Quadro-Resumo da Análise de Variância associado a esta experiência.
- Pode considerar-se que os diferentes tempos de armazenamento influenciam o teor de taninos na polpa destes frutos? Responda a esta questão utilizando testes de hipóteses.
- Interprete os seguintes gráficos de interacção e compare as conclusões obtidas com os resultados do Quadro-Resumo desta ANOVA.



- Um investigador afirma que se a temperatura de conservação é alta, o tempo de armazenamento não afecta o teor médio de taninos na polpa dos sapos. Realize uma análise estatística que lhe permita afirmar se concorda, ou não, com esta afirmação.

19. Num estudo sobre um genótipo de uma variedade de videira (casta Tinta Francisca), estudou-se o rendimento (em kg/planta) em dois diferentes locais (Tabuaço e Régua), e em diferentes anos. Os dados disponíveis para o Tabuaço referem-se aos anos 1999 e 2003, e para a Régua aos anos 1999, 2000 e 2002. Em cada local e ano foram observadas 8 unidades experimentais (parcelas), tendo sido obtidos os resultados indicados na tabela em baixo, que estão presentes no ficheiro `TabRegua.txt`.

Tabuaço		Régua		
1999	2003	1999	2000	2002
1.292	1.188	0.181	0.992	1.594
1.025	0.458	0.917	0.881	2.315
0.958	0.258	0.435	0.835	2.125
0.625	0.892	0.195	0.758	0.835
0.625	0.225	0.142	0.488	1.675
0.125	0.244	0.235	0.925	0.995
0.425	0.444	0.029	0.488	0.679
0.092	0.092	0.019	0.125	0.594

Pretende-se saber se existem efeitos de local e/ou efeitos de ano no rendimento do referido genótipo.

- (a) Indique o tipo de delineamento experimental utilizado, e um modelo ANOVA adequado para esta experiência.
- (b) Efectue os testes necessários para indicar se considera que existem efeitos de local e/ou efeitos de ano.
- (c) Utilize os testes de Tukey para estudar pares de combinações de local e ano que tenham rendimentos médios significativamente diferentes.
20. O interesse em introduzir em Palmela castas exteriores à região, conduziu à organização (numa colaboração ISA/PORVID) dum ensaio com duas castas: Malvasia Fina (MF) e Antão Vaz (AN). Decidiu-se trabalhar com quatro genótipos de cada casta (designados MF201, MF1035, MF1420 e MF1426 no caso da Malvasia Fina e AN105, AN142, AN145 e AN510 para a Antão Vaz), sendo o objectivo escolher as combinações de casta e genótipo associadas aos maiores rendimentos na região. Foi utilizado um delineamento equilibrado com oito repetições em cada situação experimental. O rendimento médio da totalidade das observações foi $4.467625 \text{ kg/planta}$, e a respectiva variância amostral $5.389415 \text{ (kg/planta)}^2$.
- (a) Descreva em pormenor o modelo ANOVA mais adequado à experiência agora descrita.
- (b) Construa a tabela-resumo correspondente a este modelo, sabendo que o Quadrado Médio Residual é 2.873782 e que a Soma de Quadrados associada às castas é 79.73597.
- (c) Foi importante prever a possibilidade de os rendimentos serem diferentes por efeito dos genótipos? Justifique a resposta através dum teste formal.
- (d) Para além de eventuais efeitos de genótipos, pode-se falar em diferenças entre as duas castas? Justifique formalmente a sua resposta.
- (e) Sabendo que os rendimentos médios nas oito situações experimentais são os abaixo indicados, teste se é possível concluir que o rendimento médio do genótipo MF201 é diferente de todos os outros.

AN105	AN142	AN145	AN510	MF1035	MF1420	MF1426	MF201
2.925	2.208	3.593	4.680	3.451	5.367	5.839	7.678

21. Após um ensaio para selecção de variedades de alho (numa colaboração ISA/INIAV), um agricultor de Braga foi aconselhado a utilizar um genótipo comercializado com a designação V7088. Sabendo que o estudo que levava ao aconselhamento fora feito em Elvas nos anos de 2000 e 2004, o referido agricultor decidiu repetir o ensaio nos seus terrenos, em condições semelhantes às do ensaio inicial. Durante três anos (de 2007 a 2009) seleccionou aleatoriamente quatro parcelas, tendo utilizado o referido genótipo. Tal como no estudo inicial, o agricultor optou por medir a variável, comercialmente importante, do peso dos bolbos de alho. Os pesos médios (em g) obtidos nas 4 parcelas de cada situação referida foram os seguintes.

1	Elvas	Ano2000	13.9200
2	Elvas	Ano2004	29.8200
3	Braga	Ano2007	15.4500
4	Braga	Ano2008	19.9325
5	Braga	Ano2009	12.9425

Tendo pedido apoio estatístico, o agricultor foi confrontado com o seguinte modelo ANOVA ajustado:

```
> summary(aov(peso.bolbo ~ Local/Ano, data=exame300113))
              Df Sum Sq Mean Sq F value    Pr(>F)
Local          1  159.34  159.345   13.072 0.002543
Local:Ano      3   605.94  201.980   16.570 5.005e-05
Residuals     15   182.84   12.189
```

- (a) Indique que tipo de delineamento experimental descreve a experiência realizada, e descreva em pormenor o modelo ANOVA que foi ajustado, explicando o significado de todas as parcelas da equação do modelo, bem como os pressupostos adicionais admitidos.
- (b) Efectue os testes F que considerar apropriados. Comente os resultados e, com base nesses testes, diga se é possível afirmar que os pesos médios dos bolbos do referido genótipo são diferentes em Braga e Elvas.
- (c) Há diferenças significativas nos pesos médios anuais em Braga? Justifique mediante um teste de Tukey ($\alpha = 0.05$).
- (d) Confuso com os resultados, o agricultor, que havia frequentado um curso de introdução à Estatística onde apenas fora leccionada a Análise de Variância a um factor, decidiu ajustar um modelo desse tipo, utilizando todos os dados disponíveis, mas ignorando os anos a que os dados se referem.
 - i. Construa, justificando, a tabela resumo da ANOVA que o agricultor obteve.
 - ii. Descreva sinteticamente os resultados obtidos através da ANOVA do agricultor ($\alpha = 0.05$). Compare esses resultados com os obtidos através do estudo do modelo anterior e comente.