INSTITUTO SUPERIOR DE AGRONOMIA

ESTATÍSTICA E ANÁLISE DE DADOS EM ZOOTECNIA – 2025-26

Exercício de Análise de Variância

1. Um estudo efectuado pela Secção de Produção Animal do ISA visou estudar se existem efeitos na digestibilidade em leitões, associados à natureza da fibra (factor *Fibra*, com 2 níveis) e à adição, ou não de enzimas digestivas (factor *Enzima*, tendo sido atribuídos os códigos “1” à ausência de enzima e “2” à presença de enzima) no alimento dos leitões. Foi considerada como variável resposta o Coeficiente de Utilização Digestiva para a celulose (variável *CEL*). A experiência utilizou doze repetições por célula. Os dados obtidos são indicados na tabela.

Fibra	Enzima	
	1 (não)	2 (sim)
1	0.44 0.43 0.44 0.53 0.36 0.29	0.54 0.49 0.53 0.45 0.47 0.37
	0.12 0.52 0.39 0.53 0.40 0.41	0.38 0.43 0.46 0.48 0.51 0.48
2	0.50 0.35 0.25 0.26 0.13 0.08	0.64 0.49 0.38 0.35 0.36 0.32
	-0.05 0.47 0.38 0.47 0.63 0.30	0.42 0.52 0.65 0.54 0.43 0.51

Em todos os testes de hipóteses que efectuar, utilize um nível de significância $\alpha = 0.05$. Os resultados do ajustamento do modelo de análise de variância adequado a este estudo estão no slide 349 das aulas teóricas.

(a) Identifique o delineamento experimental utilizado e descreva o modelo ANOVA correspondente.

(b) É admissível considerar que os valores médios da variável *CEL* são afectados por interacção entre os factores *Fibra* e *Enzima*?

(c) Existem efeitos principais do factor *Fibra*? E do factor *Enzima*?

(d) Independentemente da sua resposta nas alíneas anteriores, utilize a distribuição de Tukey para dizer se existe alguma combinação de *Fibra/Enzima* que esteja associada a um Coeficiente de Utilização Digestiva da celulosa médio significativamente mais elevado, a um nível de significância $\alpha = 0.05$. A sua resposta é coerente com os testes F correspondentes?

RESOLUÇÃO

a) Trata-se dum delineamento factorial a dois factores: Fibra (Factor A, com $a = 2$ níveis) e Enzima (Factor B, com $b = 2$ níveis). Em cada uma destas $ab = 4$ células há $n_c = 12$ repetições, pelo que se trata dum delineamento equilibrado. A variável resposta é CEL, o Coeficiente de Utilização Digestiva (CUD) da celulose. Representando por Y_{ijk} a k-ésima observação desta variável resposta CEL, correspondente ao nível i de Fibra e j de Enzima, tem-se o seguinte modelo ANOVA a dois factores, com interacção:

$Y_{ijk} = \mu_{11} + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk}, \ \forall\, i = 1,2\ ,\ j = 1,2\ ,\ k = 1,2,..,12,$
com $\alpha_1 = 0$, $\beta_1 = 0$ e $(\alpha\beta)_{ij} = 0$ se i ou j tomarem o valor 1. Neste caso concreto, e tendo em conta que cada factor tem apenas dois níveis, só existe um efeito de cada tipo: α_2, β_2 e $(\alpha\beta)_{22}$. Na equação,

- μ_{11} indica o CUD médio (populacional) para a celulose, na célula $(1,1)$;
- α_i indica o efeito principal do nível i do Factor A (*Fibra*);
- β_j indica o efeito principal do nível j do Factor B (*Enzima*);
- $(\alpha\beta)_{ij}$ indica o efeito de interacção na célula (i,j); e
- ϵ_{ijk} indica o erro aleatório associado à observação Y_{ijk}.

- $\epsilon_{ijk} \frown \mathcal{N}(0,\sigma^2)$, $\forall\, i,j,k$.
- $\{\epsilon_{ijk}\}_{i,j,k}$ constituem um conjunto de variáveis aleatórias independentes.

b) Pede-se a realização dum teste F à existência dos efeitos de interacção previstos no modelo.Eis o teste pedido (escrevendo as hipóteses da forma especial que resulta de terem-se apenas dois níveis em cada factor):

Hipóteses: $H_0 : (\alpha\beta)_{22} = 0$ vs. $H_1 : (\alpha\beta)_{22} \neq 0$.
Estatística do teste: $F = \frac{QMAB}{QMRE} \frown F_{[(a-1)(b-1)\,,\,n-ab]}$, sob H_0.
Nível de significância: $\alpha = 0.05$.
Região Crítica (Unilateral Direita): Rejeitar H_0 se $F_{calc} > f_{0.05(1,44)} \approx 4.06$.
Conclusões: O valor da estatística do teste foi já calculado: $F_{calc} = 1.560 < 4.06$, pelo que não se rejeita H_0, não havendo motivo para admitir a existência de efeitos de interacção.

c) Pedem-se agora os testes aos efeitos principais de cada factor. Eis o teste ao efeito do Factor A que, havendo apenas dois níveis no factor, é um teste a que α_2 seja nulo:

Hipóteses: $H_0 : \alpha_2 = 0$ vs. $H_1 : \alpha_2 \neq 0$.
Estatística do teste: $F = \frac{QMA}{QMRE} \frown F_{[a-1\,,\,n-ab]}$, sob H_0.
Nível de significância: $\alpha = 0.05$.
Região Crítica (Unilateral Direita): Rejeitar H_0 se $F_{calc} > f_{0.05(1,44)} \approx 4.06$.
Conclusões: O valor da estatística do teste é dado na tabela-resumo: $F_{calc} = 1.450 < 4.06$, pelo que não se rejeita H_0, não havendo motivo para admitir que a natureza da fibra afecte a digestibilidade.

Seguidamente, o teste ao efeito da presença de enzimas nas dietas:

Hipóteses: $H_0 : \beta_2 = 0$ vs. $H_1 : \beta_2 \neq 0$.
Estatística do teste: $F = \frac{QMB}{QMRE} \frown F_{[b-1\,,\,n-ab]}$, sob H_0.
Nível de significância: $\alpha = 0.05$.
Região Crítica (Unilateral Direita): Rejeitar H_0 se $F_{calc} > f_{0.05(1,44)} \approx 4.06$.
Conclusões: O valor da estatística do teste é calculado: $F_{calc} = 8.364 > 4.06$, pelo que se rejeita H_0, concluindo-se pela existência de efeitos principais associados à presença de enzimas no alimento.

Assim, conclui-se (ao nível $\alpha = 0.05$) que a adição de enzimas introduz alterações na digestibilidade média dos alimentos, não havendo no entanto efeitos significativos associados ao factor Fibra, nem de interacção.

d) Repare-se que as conclusões da alínea anterior permitem responder à pergunta através duma via alternativa à utilização de testes de Tukey. Uma vez que apenas se concluiu pela existência de efeitos principais do factor B, e este só tem dois níveis, conclui-se que as médias de célula apenas diferem entre si caso pertençam a diferentes níveis do factor Enzima. De facto, recorde-se que $\mu_{21} = \mu_{11} + \alpha_2$, pelo que ao se admitir que $\alpha_2 = 0$, está-se a admitir que $\mu_{21} = \mu_{11}$. De igual modo, $\mu_{12} = \mu_{11} + \beta_2$, pelo que ao rejeitar-se a hipótese $\beta_2 = 0$, se está a concluir que $\mu_{12} \neq \mu_{11}$. Finalmente, $\mu_{22} = \mu_{11} + \alpha_2 + \beta_2 + (\alpha\beta)_{22}$. Uma vez que se admite $\alpha_2 = 0$ e $(\alpha\beta)_{22} = 0$, admite-se $\mu_{22} = \mu_{11} + \beta_2 = \mu_{12}$.

No entanto, efectuaremos os teste de Tukey, como pedido no enunciado. O facto de a teoria subjacente a testes de Tukey e testes F da ANOVA não ser idêntica pode fazer surgir alguma discrepância nas respectivas conclusões. O termo de comparação do teste de Tukey, utilizando um nível de significância global $\alpha = 0.05$, é dado por

$$q_{\alpha(ab,n-ab)} \sqrt{\frac{QMRE}{n_c}} \;=\; q_{0.05\,(4,44)} \sqrt{\frac{0.01645}{12}} \;\approx\; 3.78 \times 0.03702477 \;=\; 0.1399536 \;.$$

Ora, as quatro médias amostrais de célula são:

```
          Fibra:Enzima
            Enzima
Fibra      1      2
     1 0.4050 0.4658
     2 0.3142 0.4675
```

As médias de célula são indicadas na tabela final. Dos seis possíveis pares de médias de células, apenas em dois casos as médias de célula diferem por mais do que o termo de comparação: $|\overline{Y}_{21.} - \overline{Y}_{12.}| = 0.1516 > 0.1400$ e $|\overline{Y}_{21.} - \overline{Y}_{22.}| = 0.1533 > 0.1400$. Logo, e ordenando as quatro médias de célula por ordem crescente, tem-se:

$$
\begin{array}{cccc}
\overline{y}_{21.} & \overline{y}_{11.} & \overline{y}_{12.} & \overline{y}_{22.} \\
0.3142 & 0.4050 & 0.4658 & 0.4675 \\
\cline{1-2}
\end{array}
$$

$$
\begin{array}{cccc}
 & 0.4050 & 0.4658 & 0.4675 \\
\cline{2-4}
\end{array}
$$

As conclusões não são inteiramente coerentes com as conclusões obtidas através dos testes F, uma vez que não se conclui que μ_{11} seja diferente das duas médias de célula associadas ao nível 2 do factor *Enzima*.