Soluciones a los ejercicios propuestos del Tema 6

6.1. Es un diseño de un factor (la estirpe). La tabla ANOVA es:

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrados medios	F
Factor (entre tratamientos)	489740.16	3	163246.72	$f = 12.7281076$
Error (dentro de los tratamientos)	153908.24	12	12825.68667	
Total	643648.4	15		

usando que tenemos $a = 4$ tratamientos, con $n_i = n = 4$ observaciones en cada tratamiento (en total, $N = 16$), y que $\sum_{i=1}^a (\bar{y}_i)^2 = 34504533.19$, $\sum_{i=1}^a \sum_{j=1}^n y_{ij}^2 = 138172041$ y $N(\bar{y}_..)^2 = 137528392.6$. Entonces, como $f \cong 12.72811 > F_{0.95}^{3,12} = 3.49$, podemos decir que el factor “estirpe” influye en la cantidad del compuesto químico.

Vamos a usar el método de separación de medias de Tukey. Para ello calculamos $TSD = Q_{0.95}^{4,12} \sqrt{\frac{12825.68667}{4}} = 237.75254$, ya que $Q_{0.95}^{4,12} = 4.1987$. Aplicando el método de Tukey nos da:

raza 4 (A), raza 3 (B), raza 1 (B), raza 2 (B),

luego las diferencias son debidas a la diferencia entre la raza 4 por un lado (es la que da la respuesta más baja) y las otras tres razas por otro.

6.2. Es un diseño de un factor (el tipo de anestesia, con $a = 3$ niveles). Se toman $n_i = n = 5$ observaciones con cada nivel, luego el número total de observaciones es $N = 15$. La tabla ANOVA es:

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrados medios	F
Factor (entre tratamientos)	543.6	2	271.8	$f = 16.08284$
Error (dentro de los tratamientos)	202.8	12	16.9	
Total	746.4	14		

con $\sum_{i=1}^a (\bar{y}_i)^2 = 680.04$, $\sum_{i=1}^a \sum_{j=1}^n y_{ij}^2 = 3603$ y $N(\bar{y}_{..})^2 = 2856.6$. Como $f \cong 16.08284 > F_{0.99}^{2,12} = 6.93$, vemos que sí que se puede decir que se observan diferencias significativas en la concentración de epinefrina según el tipo de anestesia utilizada. Calculamos ahora la diferencia significativa de Tukey: $TSD = Q_{0.99}^{3,12} \sqrt{\frac{16.9}{5}} = 9.276774$, ya que $Q_{0.99}^{3,12} = 5.0459$. Aplicando el método de Tukey de separación de medias tenemos que:

anestesia 3 (A), anestesia 1 (A), anestesia 2 (B),

luego las diferencias son debidas a la diferencia entre la anestesia 2 por un lado (es la que da la respuesta más alta) y las otras dos anestесias por otro.

6.3. Es un diseño de un factor (el químico, con $a = 4$ niveles). Se toman $n_i = n = 3$ observaciones con cada nivel, luego el número total de observaciones es $N = 12$. La tabla ANOVA es:

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrados medios	F
Factor (entre tratamientos)	1.04459	3	$0.34819\widehat{6}$	$f = 3.24579$
Error (dentro de los tratamientos)	0.85821	8	0.10727625	
Total	1.9028	11		

con $\sum_{i=1}^a (\bar{y}_i)^2 = 28661.70953$, $\sum_{i=1}^a \sum_{j=1}^n y_{ij}^2 = 85985.9868$ y $N(\bar{y}_.)^2 = 85984.084$.

Como $f \cong 3.24579 < F_{0.95}^{3,8} = 4.07$. No podemos decir que haya diferencias, así que no es adecuado utilizar el método de Tukey de seraración de medias.

6.4. Se trata de un diseño factorial de dos factores, $A =$ temperatura y $B =$ presión, con $a = 3$ niveles para A , $b = 3$ niveles para B , y $n = 2$ réplicas para cada condición experimental. El número total de observaciones es $N = 3 \times 3 \times 2 = 18$. La tabla ANOVA es:

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrados medios	F
A	0.08998	2	0.04499	$f_A = 0.88024$
B	0.42334	2	0.21167	$f_B = 4.14137$
AB	0.24668	4	0.06167	$f_{AB} = 1.20659$
Error	0.46	9	$0.05\widehat{1}$	
Total	1.22	17		

$$\text{con } \sum_{i=1}^a (\bar{y}_{i..})^2 = 24552.66833, \sum_{j=1}^b (\bar{y}_{.j})^2 = 24552.72389, N(\bar{y}_{...})^2 = 147315.92,$$

$$\sum_{i=1}^a \sum_{j=1}^b (\bar{y}_{ij.})^2 = 73658.34, \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n y_{ijk}^2 = 147317.14.$$

Como $f_A \cong 0.88204 < F_{0.95}^{2,9} = 4.26$, $f_B \cong 4.14137 < F_{0.95}^{2,9} = 4.26$, $f_{AB} \cong 1.20659 < F_{0.95}^{4,9} = 3.63$, vemos que no hay interacción significativa ni efecto de ninguno de los dos factores en la lectura de la variable asociada al proceso químico.

6.5. Se trata de un diseño factorial de dos factores, $A =$ región y $B =$ variedad del fruto. Se tiene $a = 3$ niveles para A , $b = 4$ niveles para B , y $n = 2$ réplicas para cada condición experimental. El número total de observaciones es $N = 3 \times 4 \times 2 = 24$. La tabla ANOVA es:

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrados medios	F
A	160.33334	2	80.16667	$f_A = 21.14286$
B	12.4583	3	4.15276	$f_B = 1.09524$
AB	44.6	6	7.4	$f_{AB} = 1.96337$
Error	45.5	12	3.7916	
Total	262.9583	23		

$$\text{con } \sum_{i=1}^a (\bar{y}_{i..})^2 = 37848.29688, \sum_{j=1}^b (\bar{y}_{.j})^2 = 50439.75, N(\bar{y}_{...})^2 = 302626.0417,$$

$\sum_{i=1}^a \sum_{j=1}^b (\bar{y}_{ij.})^2 = 151421.75$, $\sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n y_{ijk}^2 = 302889$. Como $f_A \cong 21.14286 > F_{0.99}^{2,12} = 6.93$, $f_B \cong 1.09524 < F_{0.99}^{3,12} = 5.95$, $f_{AB} \cong 1.96337 < F_{0.99}^{6,12} = 4.82$, podemos decir que no hay interacción significativa ni efecto del factor B , pero

sí hay efecto significativo del factor A , que es la región de procedencia del fruto, en su valor nutritivo (la región C da frutos con un valor nutritivo mayor, independientemente de la variedad).

6.6. Tenemos un diseño factorial de dos factores, $A =$ período (con $a = 2$ niveles) y $B =$ medio (también con $b = 2$ niveles), con $n = 6$ réplicas por condición experimental. En total, tenemos $N = 2 \times 2 \times 6 = 24$ observaciones.

Los cálculos necesarios para construir la tabla ANOVA son:

$$\sum_{i=1}^a (\bar{y}_{i..})^2 = 1804.45139, \quad \sum_{j=1}^b (\bar{y}_{.j.})^2 = 1756.0625, \quad N(\bar{y}_{...})^2 = 21063.375,$$

$$\sum_{i=1}^a \sum_{j=1}^b (\bar{y}_{ij.})^2 = 3625.80555, \quad \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n y_{ijk}^2 = 21857.$$

La tabla es:

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrados medios	F
A	590.04168	1	590.04168	$f_A = 115.50571$
B	9.375	1	9.375	$f_B = 1.83524$
AB	92.041688	1	92.04168	$f_{AB} = 18.01795$
Error	$102.1\widehat{6}$	20	$5.108\widehat{3}$	
Total	793.625	23		

$f_A \cong 115.50571 > F_{0.95}^{1,20} = 4.35$, $f_B \cong 1.83524 < F_{0.95}^{1,20} = 4.35$, $f_{AB} \cong 18.01795 > F_{0.95}^{1,20} = 4.35$, luego resultan significativos tanto la interacción con el efecto del factor A .

La combinación “medio de cultivo-período de tiempo” que da mejor resultado es: Medio 1 y 18 horas de tiempo (haciendo gráfica).

6.7 Consideraremos un diseño factorial de dos factores, $A =$ alimentación y $B =$ raza, con $a = 2$, $b = 3$ y $n = 3$ réplicas por condición experimental, con $N = 18$ observaciones en total.

Los cálculos necesarios para construir la tabla ANOVA son:

$$\sum_{i=1}^a (\bar{y}_{i..})^2 = 1008837.79, \quad \sum_{j=1}^b (\bar{y}_{.j.})^2 = 1669620.361, \quad N(\bar{y}_{...})^2 = 9072380.056,$$

$$\sum_{i=1}^a \sum_{j=1}^b (\bar{y}_{ij.})^2 = 3341900.111, \quad \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n y_{ijk}^2 = 10031071.$$

A partir de estos cálculos tenemos la tabla:

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrados medios	F
A	7160.054	1	7160.054	$f_A = 15.99813$
B	945342.11	2	472671.055	$f_B = 1056.11699$
AB	818.113	2	409.0565	$f_{AB} = 0.91398$
Error	5370.667	12	447.5556	
Total	958690.944	17		

Como $f_A \cong 15.99813 > F_{0.95}^{1,12} = 4.75$, $f_B \cong 1056.11699 > F_{0.95}^{2,12} = 3.89$, $f_{AB} \cong 0.91398 < F_{0.95}^{2,12} = 3.89$, podemos decir que no hay interacción significativa entre los factores, pero que sí hay efecto de ambos en el peso.

El efecto es en el sentido de que los de la raza B tienen mayor peso, y también los de madres a las que se ha suministrado la alimentación de tipo I.

6.8. Se trata de un diseño de un factor (el factor “agente químico”, con $a = 4$ niveles), con un bloque (el “rollo”, con $b = 3$ niveles).

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrados medios	F
Factor (A)	9	3	3	$f_A = 1.5$
Bloque (B)	104	2	52	$f_B = 26$
Error	12	6	2	
Total	125	11		

donde tenemos una réplica por condición experimental ($n = 1$) y un total de $N = 12$ observaciones. Los cálculos necesarios para hacer la tabla ANOVA han sido:

$$\sum_{i=1}^4 (\bar{y}_{i.})^2 = 21028, \quad \sum_{j=1}^3 (\bar{y}_{.j})^2 = 15794.75, \quad 12 (\bar{y}_{..})^2 = 63075,$$

$$\sum_{i=1}^4 \sum_{j=1}^3 y_{ij}^2 = 63200.$$

Como $f_A = 1.5 < F_{0.95}^{3,6} = 4.76$ y $f_B = 26 > F_{0.95}^{2,6} = 5.14$, concluimos que el diseño en bloque ha sido eficiente y que no hay “efecto agente químico”.

6.9. En primer lugar hacemos un diseño de un factor (“catalizador”) con $a = 4$ niveles y un bloque (“lote de materia prima”) con $b = 3$ niveles. Tenemos $n = 2$ réplicas y, en total, $N = 24$ observaciones. La tabla ANOVA es:

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrados medios	F
Factor (A)	4.8333	3	1.6111	$f_A = 0.23835$
Bloque (B)	1.3333	2	0.66665	$f_B = 0.09863$
Error	121.6667	18	6.75926	
Total	127.8333	23		

con $\sum_{i=1}^4 (\bar{y}_{i..})^2 = 21560.8\widehat{3}$, $\sum_{j=1}^3 (\bar{y}_{.j})^2 = 16170.1875$, $24 (\bar{y}_{...})^2 = 129360.1667$,
 $\sum_{i=1}^4 \sum_{j=1}^3 \sum_{k=1}^2 y_{ijk}^2 = 129488$. Tomamos $\alpha = 0.05$ como nivel de significación.

Entonces, como $f_A = 0.23835 < F_{0.95}^{3,18} = 3.16$ y $f_B = 0.09863 < F_{0.95}^{2,18} = 3.55$, resulta que concluimos que el diseño en bloque NO ha sido eficiente. Por ello, vamos a aprovechar los cálculos ya realizados para llevar a cabo el ANOVA de un factor (sin bloque), incorporando la suma de cuadrados y los grados de libertad del bloque al error, de la siguiente manera:

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrados medios	F
Factor (A)	4.8333	3	1.6111	$f_A(nueva) = 0.26197$
Error	123	20	6.15	
Total	127.8333	23		

Como $f_A(nueva) = 0.26197 < F_{0.95}^{3,20} = 3.10$, ahora concluimos que NO hay efecto del factor “catalizador” sobre la variable respuesta.