

Ejemplo de análisis de la Varianza (ANOVA).

temp	TVBN
1	18,3
1	15,92
1	18,71
1	17,92
1	15,66
1	17,14
1	15,21
1	19,92
1	17,61
1	13,43
2	11,7
2	12,87
2	11,77
2	12,23
2	13,62
2	13,24
2	14,02
2	13,66
2	12,27
2	12,45
3	16,64
3	17,83
3	19,01
3	17,33
3	17,06
3	18,04
3	17,51
3	19,11
3	17,75
3	19,36

Las condiciones de conservación del pescado se evalúan a través de la concentración de TVBN (Total Volatile Base Nitrogen). A mayor concentración de este elemento, peor es el estado de conservación de la pieza. Con objeto de determinar la temperatura que produce la menor concentración de TVBN, se eligen al azar 30 atunes recién pescados, todos de idéntico peso y características generales. Se separan en tres grupos de 10 piezas cada uno; el primer grupo se congela a -4°C , el segundo a -20°C y el tercero a -40°C . La tabla de la derecha muestra la concentración de TVBN en cada pieza después de 2 semanas de congelación. La variable *temp* corresponde a los tres valores de temperatura señalados, codificados, respectivamente como 1, 2 y 3.

A partir de estos datos:

1. ¿Existen diferencias significativas en las concentraciones medias de TVBN a las tres temperaturas? Responder a esta pregunta utilizando el método del análisis de la varianza.
2. Estima la concentración media de TVBN a cada temperatura.
3. Estima la diferencia entre la concentración media de TVBN a cada temperatura y la concentración media global.
4. ¿Se cumplen las condiciones para la aplicación del análisis de la varianza?

5. Utiliza los tests de Scheffé para decidir, en caso de que se hayan detectado diferencias significativas, a qué temperatura se produce la menor concentración de TVBN.

Nota: Los datos pueden descargarse directamente con R desde la web mediante:

```
> atunes = read.table(file = "http://dl.dropbox.com/u/7610774/Datos/atunes.csv",
  header = T, sep = ";", dec = ",")
> attach(atunes)
```

1. Para determinar si existen diferencias significativas en las concentraciones de TVBN hemos de resolver el contraste de hipótesis:

$$\begin{cases} H_0 : \mu_1 = \mu_2 = \mu_3 \\ H_1 : \exists i, j : \mu_i \neq \mu_j \end{cases}$$

siendo μ_i la concentración media de TVBN cuando el pescado se conserva a la temperatura i .

Para ello utilizaremos la función `aov()` (acrónimo de *analysis of variance*). Pero primero hemos de comprobar la estructura del conjunto de datos:

```
> str(atunes)

'data.frame':      30 obs. of  2 variables:
 $ temp: int  1 1 1 1 1 1 1 1 1 1 ...
 $ tvbn: num  18.3 15.9 18.7 17.9 15.7 ...
```

Importante: si la variable que clasifica los grupos a comparar no es un factor, el resultado que muestra R al aplicar `aov()` es incorrecto.

En este caso, como la variable `temp` no es un factor, creamos una variable factor con los mismos valores:

```
> ftemp = factor(temp, levels = c(1, 2, 3), labels = c("-4",
  "-20", "-40"))
```

Podemos efectuar ahora el análisis de la varianza:

```
> adeva = aov(tvbn ~ ftemp)
> summary(adeva)
```

```
          Df Sum Sq Mean Sq F value    Pr(>F)
ftemp      2 151.462   75.731  43.518 3.575e-09 ***
Residuals 27  46.986    1.740
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Si elegimos como nivel de significación el habitual $\alpha = 0,05$, como vemos, el p-valor asociado al contraste ($3.575e-09$) es mucho menor que α , por lo que se rechaza la hipótesis nula H_0 anterior. Consecuentemente, existen diferencias significativas entre las medias, esto es, al menos dos de las temperaturas dan lugar a valores medios distintos de TVBN.

2. Para estimar la concentración media μ_i para cada temperatura ejecutamos la función:

```
> model.tables(adeva, "means")
```

```
Tables of means
Grand mean
```

```
15.90967
```

```
  ftemp
ftemp
  -4    -20    -40
16.982 12.783 17.964
```

Grand mean representa la media global de todos los datos (15.91). El resto de valores nos da la concentración media de TVBN para cada temperatura.

3. Asimismo, para estimar los efectos $\hat{\beta}_i = \hat{\mu}_i - \hat{\mu}$, empleamos:

```
> model.tables(adeva, "effects")
```

```
Tables of effects
```

```
  ftemp
ftemp
   -4   -20   -40
1.0723 -3.1267 2.0543
```

Así, a la temperatura 1 se produce, por término medio, una concentración de 1.0723 unidades de TVBN superior a la media global; a la temperatura 2 se produce una concentración -3.1267 unidades inferior a la media, y a la temperatura 3 una concentración 2.0543 unidades mayor que la media. Por tanto, aparentemente la mayor reducción de TVBN se consigue a la temperatura 2.

4. Para comprobar la validez de la aplicación del ANOVA hemos de verificar las condiciones de homoscedasticidad (igualdad de varianzas en todos los grupos) y normalidad de los residuos.

a) **Homoscedasticidad:** hemos de resolver el contraste:

$$\begin{cases} H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_p^2 \\ H_1 : \exists i, j : \sigma_i^2 \neq \sigma_j^2 \end{cases}$$

Para ello aplicamos el test de Levene:

```
> require(car)
> levene.test(tvbn ~ ftemp)

Levene's Test for Homogeneity of Variance
  Df F value Pr(>F)
group 2  3.5977 0.04120 *
  27
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

(**nota:** en versiones antiguas del paquete `car` el test de Levene se ejecuta mediante una sintaxis ligeramente distinta, `leveneTest()`).

El p-valor inferior a $\alpha = 0,05$ nos lleva a rechazar H_0 . Por tanto los datos no son homoscedásticos y el análisis de la varianza que se ha realizado en el apartado 1 no es válido.

b) **Normalidad:** para contrastar la normalidad de los residuos utilizamos el test de Shapiro-Wilk:

```
> shapiro.test(residuals(adeva))
      Shapiro-Wilk normality test

data:  residuals(adeva)
W = 0.9741, p-value = 0.6569
```

Los residuos son normales, pero como ha fallado la homoscedasticidad convertimos los datos a logaritmo y comprobamos si los datos transformados son homoscedásticos:

```
> tvbn.transf = log(tvbn)
> levene.test(tvbn.transf ~ ftemp)

Levene's Test for Homogeneity of Variance
      Df F value Pr(>F)
group  2  2.8155 0.07752 .
      27
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

En este caso el p-valor es mayor que $\alpha=0.05$, por lo que podemos aceptar la igualdad de las varianzas (homoscedasticidad) para los datos transformados a escala logarítmica. Procedemos entonces a repetir el análisis de la varianza, ahora sobre estos datos transformados:

```
> adeva2 = aov(tvbn.transf ~ ftemp)
> summary(adeva2)

      Df Sum Sq Mean Sq F value    Pr(>F)
ftemp    2  0.66106  0.33053  48.815 1.078e-09 ***
Residuals 27  0.18282  0.00677
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Vemos, pues, que el análisis de la varianza muestra también diferencias significativas entre las concentraciones de TVBN a las distintas temperaturas ($p\text{-valor} < \alpha$) en esta nueva escala. Ahora el análisis puede ser válido, ya que los datos son homoscedásticos. Faltaría comprobar la normalidad de los residuos en este nuevo modelo:

```
> shapiro.test(residuals(adeva2))
```

```
Shapiro-Wilk normality test
```

```
data: residuals(adeva2)
```

```
W = 0.9668, p-value = 0.4548
```

Como el p-valor es mayor que $\alpha = 0,05$ puede aceptarse la normalidad de los residuos. Consecuentemente la aplicación del ANOVA es correcta para los datos de TVBN transformados a escala logarítmica. Como el ANOVA nos ha conducido a decidir que existen diferencias significativas en la concentración de TVBN con los datos transformados, concluimos que tales diferencias deben existir también en la escala original.

5. Para ver entre qué medias están las diferencias significativas utilizamos el test de Scheffe (ojo, sobre los datos transformados, en los que hemos validado el ANOVA; si el ANOVA hubiese sido válido con los datos originales, habríamos utilizado aquellos.):

```
> require(agricolae)
```

```
> scheffe.test(adeva2, "ftemp")
```

```
Study:
```

```
Scheffe Test for tvbn.transf
```

```
Mean Square Error : 0.006771033
```

```
ftemp, means
```

	tvbn.transf	std.err	replication
-20	2.546255	0.02032692	10
-4	2.826161	0.03686694	10
-40	2.887200	0.01609208	10

```
alpha: 0.05 ; Df Error: 27
```

```
Critical Value of F: 3.354131
```

```
Minimum Significant Difference: 0.09531197
```

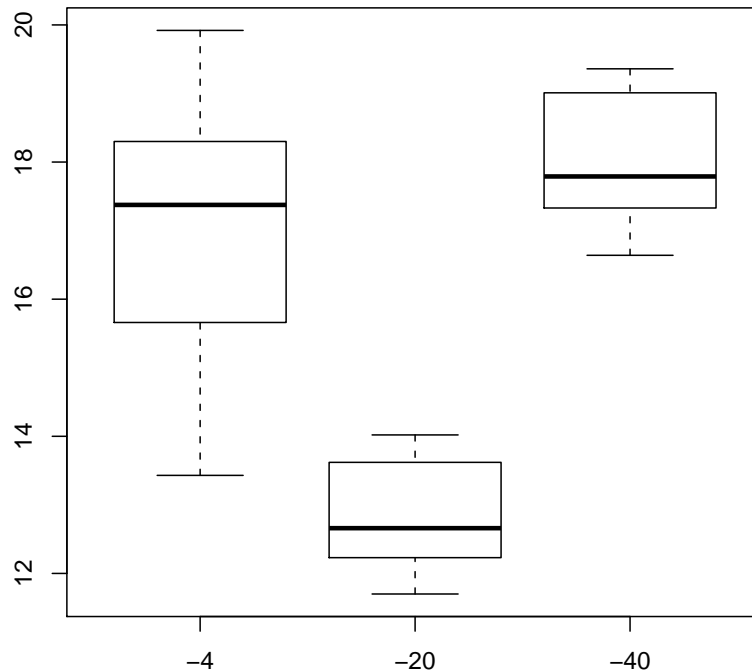
Means with the same letter are not significantly different.

Groups, Treatments and means

a	-40	2.8872
a	-4	2.826161
b	-20	2.546255

Vease que en el resultado se nos informa que "las medias con la misma letra no son significativamente diferentes". En este caso, los *tratamientos* (temperaturas) 3 y 1 van acompañados de la misma letra (a), y el tratamiento (temperatura) 2 lleva la letra (b). Eso significa que entre los valores medios de TVBN que se producen a las temperaturas 1 y 3 no se observan diferencias significativas, que sí se observan con la temperatura 2 que muestra una media de TVBN inferior (las medias están en escala logarítmica, si bien sus valores en la escala original se han mostrado en el apartado 2). El resultado queda claramente ilustrado con el gráfico siguiente (datos en escala original):

```
> boxplot(tvbn ~ ftemp)
```



En el gráfico se aprecia que la media de TVBN a la temperatura 2 es inferior al resto (y todo el análisis que hemos realizado nos prueba que es *significativamente inferior al resto*, esto es, que no es inferior por azar). Asimismo se percibe que los grupos no son homoscedásticos y que el primero es más variable que el resto.

En este caso, con la transformación logarítmica hemos conseguido la homoscedasticidad y por tanto hemos podido utilizar el ANOVA. Si ninguna transformación hubiese producido datos homoscedásticos, podríamos haber utilizado en cualquier caso el test de Kruskal-Wallis:

```
> kruskal.test(tvbn ~ ftemp)

Kruskal-Wallis rank sum test

data:  tvbn by ftemp
Kruskal-Wallis chi-squared = 18.831, df = 2, p-value =
8.145e-05
```

Como vemos, también este test nos confirma la existencia de diferencias significativas entre los valores centrales de concentración de TVBN a distintas temperaturas. El problema en este caso es que, de no cumplirse las condiciones del ANOVA, los contrastes a posteriori, para decidir entre qué grupos están las diferencias no podrían realizarse mediante el método de Scheffé (habría que utilizar varios test de Wilcoxon corrigiendo el nivel de significación de los contrastes, lo que queda pendiente de explicar).