La aterosclerosis coronaria (AC) es una de las principales causas de mortalidad en los países occidentales. Se cree que la oxidación del colesterol de baja densidad (ver, por ejemplo, Steinberg et al. (1989), *New Engl. J. Med.*, 320, 915-924) es un importante mecanismo en el desarrollo de la AC. Hay una interesante polémica científica sobre los supuestos efectos antioxidantes (y, por tanto, beneficiosos contra la AC) de las bebidas alcohólicas consumidas en cantidades moderadas. Algunos autores han puesto en duda estos efectos cardioprotectores, otros los han atribuido al alcohol en sí mismo y, por último, otros (ver, por ejemplo, Gorinstein et al., 2002, *Nutrition Research* 22, 659-666, para una visión más amplia de este tema) atribuyen la mayor parte de los efectos beneficiosos a las sustancias fenólicas que están contenidas en el vino tinto en mucha mayor medida que en otras bebidas alcohólicas. Según los defensores de esta última teoría, el vino tinto (consumido siempre en cantidades muy moderadas) sería mucho más cardiosaludable que las demás bebidas alcohólicas.
Se han determinado los valores de epicatequina (una sustancia fenólica) en 10 muestras de vino tinto, encontrando que la media muestral era 195.1 mg/l y el error típico 10.051. Los correspondientes valores para 10 muestras de cerveza fueron 65.5 mg/L y 3.4184. En vista de estos resultados ¿puede aceptarse que las varianzas de la epicatequina en el vino tinto y en la cerveza son iguales?, ¿puede decirse que son significativamente diferentes los contenidos medios de esta sustancia en el vino tinto y en la cerveza?
Responde a la misma pregunta para el caso del ácido ferúlico en el que los resultados obtenidos fueron, para el vino tinto, 7.2 mg/l (media) y 0.4541 mg/l (error típico) (tamaño muestral = 10) y para la cerveza 6.8 mg/l (media) y 0.3571 mg/l (error típico) (tamaño muestral = 10).

Solución:

Epicatequina**:** Dado que los tamaños muestrales para el vino tinto y la cerveza son pequeños ($n_1 = n_2 = 10$), para resolver el problema suponemos que el valor de epicatequina en el vino tinto (resp. cerveza) es una variable aleatoria X (resp. Y) con distribución $N(\mu_1, \sigma_1)$ (resp $N(\mu_2, \sigma_2)$) y que las dos v.a. son independientes entre sí. Queremos contrastar

$$\begin{array}{ll} H_0: & \sigma_1 = \sigma_2 \\ H_1: & \sigma_1 \neq \sigma_2 \end{array}$$

al nivel de significación $\alpha = 0.1$ y

$$\begin{array}{ll} H_0: & \mu_1 = \mu_2 \\ H_1: & \mu_1 \neq \mu_2 \end{array}$$

al nivel de significación $\alpha = 0.05$. La información muestral proporcionada en el enunciado es

$$\begin{array}{ll} n_1 = 10 & n_2 = 10 \\ \bar{x} = 195.1 & \bar{y} = 65.5 \\ \dfrac{s_1}{\sqrt{n_1}} = 10.051 & \dfrac{s_2}{\sqrt{n_2}} = 3.4184. \end{array}$$

De aquí deducimos que $s_1 = 10.051\sqrt{10} = 31.784$ y $s_2 = 3.4184\sqrt{10} = 10.810$. Por tanto, $s_1^2/s_2^2 = 8.645$. Como $F_{9;9;0.05} = 3.18$ y $F_{9;9;0.95} = 1/F_{9;9;0.05} = 0.31$, hay suficiente evidencia muestral para rechazar la homocedasticidad de X e Y.

Para el contraste de igualdad de medias, si suponemos que $\sigma_1 = \sigma_2$, calculamos la varianza combinada

$$s_p^2 = \frac{9 \cdot 31.784^2 + 9 \cdot 10.810^2}{18} = 563.539.$$

La región de rechazo de H_0 en este contraste es $R = \left\{|t| > t_{n_1+n_2-2,\alpha/2}\right\}$, donde

$$t = \frac{\bar{x} - \bar{y}}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = 12.208$$

es el estadístico del contraste. Como $t_{18;0.025} = 2.101$, rechazo H_0 a nivel 0.05. A este nivel de significación hay evidencia de que los niveles medios de epicatequina en el vino tinto y la cerveza son diferentes.

Si no suponemos igualdad de varianzas (es decir, $\sigma_1 \neq \sigma_2$) una opción es usar la región de rechazo $R = \{|t| \geq t_{f;\alpha/2}\}$, donde f es el entero más próximo a

$$\frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{(s_1^2/n_1)^2}{n_1-1} + \frac{(s_2^2/n_2)^2}{n_2-1}} = 11.1 \to t_{f;\alpha/2} = t_{11;0.025} = 2.201$$

y el estadístico del contraste es

$$t = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = 12.208.$$

Por tanto, también rechazamos $H_0 : \mu_1 = \mu_2$ a nivel 0.05, bajo heterocedasticidad.

Ácido ferúlico**:** En este caso, las variables X e Y son respectivamente los niveles de ácido ferúlico en vino y cerveza respectivamente. También suponemos normalidad y homocedasticidad de las dos variables. Ahora μ_1 y μ_2 son los niveles medios de ácido ferúlico en el vino tinto y la cerveza respectivamente. Queremos hacer el contraste

$$\begin{array}{ll} H_0: & \mu_1 = \mu_2 \\ H_1: & \mu_1 \neq \mu_2 \end{array}$$

Se puede comprobar que el estadístico del contraste ahora toma el valor $t = 0.692$, que es menor que $t_{18;0.025} = 2.101$. Por tanto, a nivel $\alpha = 0.05$ no hay evidencia estadística suficiente para rechazar H_0, es decir, el contenido medio de ácido ferúlico en vino tinto y cerveza es el mismo.