Profesor: Vicente Acuña
Auxiliar: Gonzalo Maturana

Auxiliar 10 MA34B-03

Problema 1
Un profesor desea estimar cuanta es la variabilidad del tiempo en que sus alumnos se demoran en responder un problema "tipo control" de estadística. Toma una muestra de 20 alumnos y obtiene que el tiempo promedio de respuesta es de 25 min. Además observa que $S_n^2 = 6.2$. Suponiendo que el tiempo de respuesta se distribuye como una $N(\mu, \sigma^2)$. (Hace unos años, el profesor había determinado que para ese problema $\sigma^2 = 5$). Determine si el aparente aumento de la dispersión es significativo o no.

Se propone:

$H_0 : \sigma^2 = 5$
$H_1 : \sigma^2 > 5$

$\alpha = 0.05$
Sol:

La región de rechazo será de la forma
$W = \{(x_1,...,x_n) tq : S_n^2 > C\}$

P(rechazar H_0 / H_0 cierto) = 0.05

$\Rightarrow P(S_n^2 > C / \sigma^2 = 5) = 0.05$

$$P(\frac{nS_n^2}{\sigma^2} > \frac{nC}{\sigma^2} / \sigma^2 = 5) = 0.05$$

$P(\chi^2_{20-1} > 4C) = 0.05$

De la tabla obtenemos entonces que $4C = 30.1 \Rightarrow C = 7.525$

$\Rightarrow W = \{S_n^2 > 7.525\}$ y $S_n^2 = 6.2 \notin W$

$\Rightarrow$ No se rechaza H_0.

Problema 2

Se tiene X→ $N(\mu_1,\sigma_1^2)$, con n_1= 622 ; Y→ $N(\mu_2,\sigma_2^2)$, con n_2=694 , indep.

En las muestras, las medias de cada grupo son $\bar{x} = 34.643$, $\bar{y} = 35.062$ y las varianzas $S_1^2 = 2.18$, $S_2^2 = 2.16$.

¿Podemos admitir que $\sigma_1^2 = \sigma_2^2 = \sigma^2$?

Sol.:

En este caso, las hipótesis son.

$$H_0 : \frac{\sigma_1^2}{\sigma_2^2} = 1$$

$$H_1 : \frac{\sigma_1^2}{\sigma^2} > 1$$

, pues conocemos una estimación de las varianzas $\frac{S_1^2}{S_2^2} = \frac{2.18}{2.16} > 1$

- Alternativa 1) calculando el pvalor

$$P(rechazarHo/Hocierto) = P\left(\frac{S_1^2}{S_2^2} > 1 / \frac{\sigma_1^2}{\sigma_2^2} = 1\right) = P\left(\frac{n_1 S_1^2}{\sigma_1^2(n_1-1)} * \frac{\sigma_2^2(n_2-1)}{n_2 S_2^2} > \frac{n_1(n_2-1)\sigma_2^2}{n_2(n_1-1)\sigma_1^2} / \frac{\sigma_1^2}{\sigma_2^2} = 1\right)$$

$$= P\left(\frac{n_1 S_1^2}{(n_1-1)} * \frac{(n_2-1)}{n_2 S_2^2} > \frac{n_1(n_2-1)}{n_2(n_1-1)}\right) = P\left(F_{n_1-1,n_2-1}, > F\right)$$

Una aproximación del estadístico F_{n_1-1,n_2-1} es $= \left(\frac{n_1 S_1^2}{(n_1-1)} * \frac{(n_2-1)}{n_2 S_2^2}\right) = \frac{622 * 693 * 2.18}{621 * 694 * 2.16} = 1.0094$

Entonces, tenemos que calcular $P(F_{621,693} > 1.009)$

De la tabla de Fischer, vemos que $P(F_{621,693} > 1.08) = 0.05$ y que $P(F_{621,693} > 1.12) = 0.01$

Por lo tanto $P(F_{621,693} > 1.009) > 0.05 \Rightarrow$ aceptamos Ho.

- Alternativa 2) Con la región de rechazo, es decir ,encontrar C tal que

$$= P\left(\frac{S_1^2}{S_2^2} > C / \frac{\sigma_1^2}{\sigma_2^2} = 1\right) = \alpha,$$

tomando $\alpha = 0.05$ (peor de los casos)

$$\Rightarrow P\left(\frac{n_1 S_1^2}{\sigma_1^2(n_1-1)} * \frac{\sigma_2^2(n_2-1)}{n_2 S_2^2} > \frac{C * n_1(n_2-1)\sigma_2^2}{n_2(n_1-1)\sigma_1^2} / \frac{\sigma_1^2}{\sigma_2^2} = 1\right) = P\left(\frac{n_1 S_1^2}{(n_1-1)} * \frac{(n_2-1)}{n_2 S_2^2} > \frac{C * n_1(n_2-1)}{n_2(n_1-1)}\right)$$

$$\Rightarrow P\left(F_{621,693} > F\right) = 0.05$$

De la tabla F= 1.08 $= \left(\frac{C * n_1(n_2-1)}{n_2(n_1-1)}\right) = \frac{C * 622 * 693}{694 * 621} = 1.00017 * C \qquad \Rightarrow C = 1.079$

$\Rightarrow W = \left(\frac{S_1^2}{S_2^2} > 1.079\right)$, como $\frac{S_1^2}{S_2^2} = \frac{2.18}{2.16} = 1.0092 \Rightarrow \notin W \Rightarrow aceptamosHo$

Problema 3

Se desea comprobar la eficiencia de un nuevo tratamiento médico, mediante el análisis de cierto examen. Se espera que si el tratamiento es eficaz los exámenes arrojen valores mayores. Se tienen los resultados del examen antes (X) y después (Y) del tratamiento en n=50 personas. La media del examen antes del tratamiento es $\overline{X} = 20$ y después $\overline{Y} = 22.5$. Las varianzas $S_x^2 = \frac{1}{n}\sum(x_i - \overline{x})^2 = 75$ y después $S_y^2 = \frac{1}{n}\sum(y_i - \overline{y})^2 = 125$. La covarianza entre el resultado antes y después del tratamiento es $S_{xy}^2 = \frac{1}{n}\sum(x_i - \overline{x})(y_i - \overline{y}) = 20$.

a) Se define la diferencia del resultado antes y después del tratamiento por $d = x - y$. Suponiendo la normalidad para $X \to N(\mu_1, \sigma_1^2)$ y para $Y \to N(\mu_2, \sigma_2^2)$. Determine la distribución de d y estime los parámetros $\delta = E(d)$ y $\tau^2 = Var(d)$.

b) Realice el test de hipótesis $\begin{matrix} H_0 : \delta = 0 \\ H_1 : \delta < 0 \end{matrix}$, con un error tipo I igual a $\alpha = 0.05$

Solución:

a) Si $X \to N(\mu_1, \sigma_1^2)$ e $Y \to N(\mu_2, \sigma_2^2)$

$\Rightarrow d = X - Y \to N(\delta, \tau^2)$

Por otro lado, $\delta = E(X - Y) = E(X) - E(Y) = \mu_1 - \mu_2$

Entonces la estimación de δ es

$$\hat{\delta} = \overline{X} - \overline{Y} = -2.5$$

Por otro lado, $\tau^2 = Var(X) + Var(Y) - 2Cov(X, Y) = \sigma_1^2 + \sigma_2^2 - 2Cov(X, Y)$

Entonces la estimación de τ^2 es

$$\hat{\tau}^2 = S_x^2 + S_y^2 - 2 * S_{xy} = 75 + 125 - 40 = 160$$

Observación: No conozco ni δ ni τ^2, sólo tengo estimaciones de ellos.

b) Tenemos el test $\begin{matrix} H_0 : \delta = 0 \\ H_1 : \delta < 0 \end{matrix}$, con $\alpha = 0.05$

Para este caso, la región crítica o re rechazo es de la forma $W = \{\hat{\delta} < C\}$

Y tenemos que el error tipo I es

$$P(\hat{\delta} < C / \delta = 0) = 0.05$$

Además sabemos por normalidad que $\frac{(\hat{\delta}-\delta)\sqrt{n}}{\tau} \to N(0,1)$ y $\frac{n\hat{\tau}^2_{n-1}}{\tau^2} \to x^2_{n-1}$

$$\Rightarrow \frac{\frac{(\hat{\delta}-\delta)\sqrt{n}}{\tau}}{\sqrt{\frac{n\hat{\tau}^2_{n-1}}{(n-1)\tau^2}}} \to t_{n-1} \qquad \Rightarrow \frac{(\hat{\delta}-\delta)\sqrt{n-1}}{\hat{\tau}} \to t_{n-1}$$

Como n=50 $\Rightarrow \frac{(\hat{\delta}-\delta)\sqrt{49}}{\hat{\tau}} \to t_{49}$

$$\Rightarrow P(\hat{\delta} < C / \delta = 0) = P\left(\frac{(\hat{\delta}-\delta)\sqrt{n-1}}{\hat{\tau}} < \frac{(C-\delta)\sqrt{n-1}}{\hat{\tau}} / \delta = 0\right) =$$

$$= P\left(\frac{\hat{\delta}\sqrt{n-1}}{\hat{\tau}} < \frac{C\sqrt{n-1}}{\hat{\tau}}\right) = 0.05$$

$$\Rightarrow P(t_{49} < \frac{C\sqrt{n-1}}{\hat{\tau}}) = 0.05$$

$$\Rightarrow P(t_{49} < t) = 0.05$$

Pero en la tabla no está la estimación para ese valor de n, entonces hay que estimarlo.

Tenemos que $P(t_{40} < t_1) = 0.05 \Rightarrow t_1 = -1.684$

$P(t_{60} < t_2) = 0.05 \Rightarrow t_2 = -1.671$

Tomando estos dos puntos calculamos un valor intermedio que nos da $t = -1.678$

$$\Rightarrow C = \frac{-1.678 * \hat{\tau}}{\sqrt{n-1}} = -3.03$$

De la parte a) $\hat{\tau}^2 = 160 \Rightarrow \hat{\tau} = 12.6$

En nuestro caso $\hat{\delta} = -2.5$. Por lo tanto $\hat{\delta} > C$ y no está en la región de rechazo W.

$\Rightarrow$ no se rechaza H_0.