Ayudantía 11 MAT033

Eduardo Rubio Marín

Correo: eduardo.rubiom@sansano.usm.cl

Julio, 2021

Pregunta 1

A partir de datos del simce se quiere estudiar si existe una relación entre el desempeño y el cambió de escuelas la muestre es de estudiantes de octavo año básico, se define el grupo 1 de estudiantes que asistieron a dos o más escuelas y el grupo 2 de estudiantes que permanecieron en la misma escuela.

Grupo 1	$n_1 = 15$	$\mu_1 = 85$	$S_1^2 = 30$
Grupo 2	$n_2 = 22$	$\mu_2 = 87$	$S_2^2 = 25$

a. Verifique si hay relación entre las varianzas de ambos grupos mediante Un test que involucre Intervalos de confianza de 95% de esta.
b. Concluya mediante un test adecuado si es correcto asumir igualdad entre las medias de ambos grupos.

Solución: Veamos que nos interesa ver si el desempeño promedio entre estudiantes difiere o no, por lo que planetamos las hipotesis.

Solución: Veamos que nos interesa ver si el desempeño promedio entre estudiantes difiere o no, por lo que planetamos las hipotesis.

$$\begin{array}{c} H_0 : \mu_1 = \mu_2 \\ \text{VS} \\ H_1 : \mu_1 \neq \mu_2 \end{array} \quad \Leftrightarrow \quad \begin{array}{c} H_0 : \mu_1 - \mu_2 = 0 \\ \text{VS} \\ H_1 : \mu_1 - \mu_2 \neq 0 \end{array}$$

Solución: Veamos que nos interesa ver si el desempeño promedio entre estudiantes difiere o no, por lo que planetamos las hipotesis.

$$\begin{array}{c} H_0 : \mu_1 = \mu_2 \\ \text{VS} \\ H_1 : \mu_1 \neq \mu_2 \end{array} \quad \Leftrightarrow \quad \begin{array}{c} H_0 : \mu_1 - \mu_2 = 0 \\ \text{VS} \\ H_1 : \mu_1 - \mu_2 \neq 0 \end{array}$$

Solución: Veamos que nos interesa ver si el desempeño promedio entre estudiantes difiere o no, por lo que planetamos las hipotesis.

$$\begin{array}{ccc} H_0 : \mu_1 = \mu_2 & & H_0 : \mu_1 - \mu_2 = 0 \\ \text{VS} & \Leftrightarrow & \text{VS} \\ H_1 : \mu_1 \neq \mu_2 & & H_1 : \mu_1 - \mu_2 \neq 0 \end{array}$$

Tenemos la significancia $\alpha = 0,05$.

Solución: Veamos que nos interesa ver si el desempeño promedio entre estudiantes difiere o no, por lo que planetamos las hipotesis.

$$\begin{array}{c} H_0 : \mu_1 = \mu_2 \\ \text{VS} \\ H_1 : \mu_1 \neq \mu_2 \end{array} \Leftrightarrow \begin{array}{c} H_0 : \mu_1 - \mu_2 = 0 \\ \text{VS} \\ H_1 : \mu_1 - \mu_2 \neq 0 \end{array}$$

Tenemos la significancia $\alpha = 0,05$. supondremos que las poblaciones distribuyen normalmente, donde la varianza es desconocida y conocemos las varianzas muéstrales, diremos que las muestras son independientes,

Solución: Veamos que nos interesa ver si el desempeño promedio entre estudiantes difiere o no, por lo que planetamos las hipotesis.

$$\begin{matrix} H_0 : \mu_1 = \mu_2 \\ \text{VS} \\ H_1 : \mu_1 \neq \mu_2 \end{matrix} \quad \Leftrightarrow \quad \begin{matrix} H_0 : \mu_1 - \mu_2 = 0 \\ \text{VS} \\ H_1 : \mu_1 - \mu_2 \neq 0 \end{matrix}$$

Tenemos la significancia $\alpha = 0,05$. supondremos que las poblaciones distribuyen normalmente, donde la varianza es desconocida y conocemos las varianzas muéstrales, diremos que las muestras son independientes, para estudiar si σ^2 es igual en ambas poblaciones haremos un test de hipótesis sobre la división de varianzas.

Solución: Veamos que nos interesa ver si el desempeño promedio entre estudiantes difiere o no, por lo que planetamos las hipotesis.

$$\begin{array}{c} H_0 : \mu_1 = \mu_2 \\ \text{VS} \\ H_1 : \mu_1 \neq \mu_2 \end{array} \Leftrightarrow \begin{array}{c} H_0 : \mu_1 - \mu_2 = 0 \\ \text{VS} \\ H_1 : \mu_1 - \mu_2 \neq 0 \end{array}$$

Tenemos la significancia $\alpha = 0,05$. supondremos que las poblaciones distribuyen normalmente, donde la varianza es desconocida y conocemos las varianzas muéstrales, diremos que las muestras son independientes, para estudiar si σ^2 es igual en ambas poblaciones haremos un test de hipótesis sobre la división de varianzas.

Al desconocer los valores de las varianzas poblacionales respectivas, tendremos que utilizar un intervalo de confianza sobre el cual hacer inferencia.

Al desconocer los valores de las varianzas poblacionales respectivas, tendremos que utilizar un intervalo de confianza sobre el cual hacer inferencia.

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

$$\text{vs} \quad \Leftrightarrow \text{vs}$$

$$H_1 : \frac{\sigma_1^2}{\sigma_2^2} \neq 1 \qquad H_1 : \sigma_1^2 \neq \sigma_2^2$$

$$IC\left[\frac{S_1^2}{S_2^2} \cdot F_{(\frac{\alpha}{2}, n_1-1, n_2-1)} \leq \frac{\sigma_1^2}{\sigma_2^2} \leq \frac{S_1^2}{S_2^2} \cdot F_{(1-\frac{\alpha}{2}, n_1-1, n_2-1)}\right]$$

$$IC\left[0,468 \leq \frac{\sigma_1^2}{\sigma_2^2} \leq 3,4124\right]$$

Al desconocer los valores de las varianzas poblacionales respectivas, tendremos que utilizar un intervalo de confianza sobre el cual hacer inferencia.

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

$$\text{vs} \qquad \Leftrightarrow \text{vs}$$

$$H_1 : \frac{\sigma_1^2}{\sigma_2^2} \neq 1 \qquad H_1 : \sigma_1^2 \neq \sigma_2^2$$

$$IC\left[\frac{S_1^2}{S_2^2} \cdot F_{\left(\frac{\alpha}{2}, n_1-1, n_2-1\right)} \leq \frac{\sigma_1^2}{\sigma_2^2} \leq \frac{S_1^2}{S_2^2} \cdot F_{\left(1-\frac{\alpha}{2}, n_1-1, n_2-1\right)}\right]$$

$$IC\left[0,468 \leq \frac{\sigma_1^2}{\sigma_2^2} \leq 3,4124\right]$$

Como podemos observar que el 1 pertenece al intervalo, entonces no rechazamos H_0

Ahora procedemos a calcular la diferencia de medias muéstrales a través del estadístico t de student con $n_1 + n_2 - 2$ grados de libertad. Obteniendo que.

Ahora procedemos a calcular la diferencia de medias muéstrales a través del estadístico t de student con $n_1 + n_2 - 2$ grados de libertad. Obteniendo que.

$$t_{(0,05)}(35) = 2,0301$$

$$S_p^2 = \frac{14 \cdot 30 + 21 \cdot 25}{35} = 27$$

$$T = \frac{85 - 87}{1,76} = \frac{-2}{1,76} = -1,14$$

Ahora procedemos a calcular la diferencia de medias muéstrales a través del estadístico t de student con $n_1 + n_2 - 2$ grados de libertad. Obteniendo que.

$$t_{(0,05)}(35) = 2,0301$$

$$S_p^2 = \frac{14 \cdot 30 + 21 \cdot 25}{35} = 27$$

$$T = \frac{85 - 87}{1,76} = \frac{-2}{1,76} = -1,14$$

Como tenemos que $|-1,14| < 2,03$ no rechazamos H_0. Con lo que no podemos concluir que haya una diferencia entre las dos poblaciones.

Pregunta 2

Una empresa de neumáticos afirma que un nuevo juego de neumáticos en promedio dura al menos 28.000 km. Las pruebas con 64 neumáticos dan como resultado una duración media de 27.800 km, con una desviación estándar de 1.000 km. Bajo la normalidad de los neumáticos, se pide:

a. Comprobar si hay evidencia suficiente para rechazar la afirmación de la empresa, con un nivel de signifancia del 5%
b. ¿Cuál es el p-valor?

Solución:

a. Debemos definir la variable aleatoria sobre la población de neumáticos.

Solución:

a. Debemos definir la variable aleatoria sobre la población de neumáticos.
X="Duración en kilometros del juego".

Solución:

a. Debemos definir la variable aleatoria sobre la población de neumáticos.
X="Duración en kilometros del juego".
Donde sabemos que $X \sim N(28.000, \sigma)$.

Solución:

a. Debemos definir la variable aleatoria sobre la población de neumáticos.
X="Duración en kilometros del juego".
Donde sabemos que $X \sim N(28.000, \sigma)$.
Ahora las hipótesis sobre la media poblacional con μ y σ^2 desconocida,

Solución:

a. Debemos definir la variable aleatoria sobre la población de neumáticos.
X="Duración en kilometros del juego".
Donde sabemos que $X \sim N(28.000, \sigma)$.
Ahora las hipótesis sobre la media poblacional con μ y σ^2 desconocida,

$$H_0 : \mu \geq 28.000$$
$$vs$$
$$H_1 : \mu < 28.000$$

Solución:

a. Debemos definir la variable aleatoria sobre la población de neumáticos.
X="Duración en kilometros del juego".
Donde sabemos que $X \sim N(28.000, \sigma)$.
Ahora las hipótesis sobre la media poblacional con μ y σ^2 desconocida,

$$H_0 : \mu \geq 28.000$$
$$vs$$
$$H_1 : \mu < 28.000$$

La regla de decisión para un k sería:

Solución:

a. Debemos definir la variable aleatoria sobre la población de neumáticos.
X="Duración en kilometros del juego".
Donde sabemos que $X \sim N(28.000, \sigma)$.
Ahora las hipótesis sobre la media poblacional con μ y σ^2 desconocida,

$$H_0 : \mu \geq 28.000$$
$$vs$$
$$H_1 : \mu < 28.000$$

La regla de decisión para un k sería:
Sí $\overline{x} \geq k$ NO rechazamos H_0, esto se conoce como región de aceptación.

Solución:

a. Debemos definir la variable aleatoria sobre la población de neumáticos.
X="Duración en kilometros del juego".
Donde sabemos que $X \sim N(28.000, \sigma)$.
Ahora las hipótesis sobre la media poblacional con μ y σ^2 desconocida,

$$H_0 : \mu \geq 28.000$$
$$vs$$
$$H_1 : \mu < 28.000$$

La regla de decisión para un k sería:
Sí $\overline{x} \geq k$ NO rechazamos H_0, esto se conoce como región de aceptación.
Sí $\overline{x} < k$ rechazamos H_0, esto se conoce como región de rechazo.

Como estamos en el caso de una muestra de una población norma con varianza desconocida, tendremos que la media muestral $\overline{x} \sim N\left(\mu, \frac{S_x}{\sqrt{n}}\right)$.

Como estamos en el caso de una muestra de una población norma con varianza desconocida, tendremos que la media muestral $\overline{x} \sim N\left(\mu, \frac{S_x}{\sqrt{n}}\right)$.

Que bajo la hipotesis nula sería.

Como estamos en el caso de una muestra de una población norma con varianza desconocida, tendremos que la media muestral $\overline{x} \sim N\left(\mu, \frac{S_x}{\sqrt{n}}\right)$.
Que bajo la hipotesis nula sería.

$$\overline{X} \sim N\left(28.000, \frac{1.000}{\sqrt{64}}\right) \equiv N(28.000, 125)$$

Como estamos en el caso de una muestra de una población norma con varianza desconocida, tendremos que la media muestral $\overline{x} \sim N\left(\mu, \frac{S_x}{\sqrt{n}}\right)$.
Que bajo la hipotesis nula sería.

$$\overline{X} \sim N\left(28.000, \frac{1.000}{\sqrt{64}}\right) \equiv N(28.000, 125)$$

El valor que nos interesa de k, bajo la hipótesis nula, se determina con el nivel de signifación $\alpha = 0.05$

Como estamos en el caso de una muestra de una población norma con varianza desconocida, tendremos que la media muestral $\overline{x} \sim N\left(\mu, \frac{S_x}{\sqrt{n}}\right)$.
Que bajo la hipotesis nula sería.

$$\overline{X} \sim N\left(28.000, \frac{1.000}{\sqrt{64}}\right) \equiv N(28.000, 125)$$

El valor que nos interesa de k, bajo la hipótesis nula, se determina con el nivel de signifación $\alpha = 0.05$

$$\begin{aligned}\alpha &= P[\text{Rechazar } H_0/H_0 \text{ cierta}] = P[\overline{x} < k/H_0 \text{ cierta}]\\ &= P[\overline{x} < k/\mu = 28.000] = P[\frac{\overline{x} - 28.000}{125} < \frac{k - 28.000}{125}]\\ &= P[Z < \frac{k - 28.000}{125}] = 0,05 \Leftrightarrow P[Z > -\frac{k - 28.000}{125}] = 0,05\end{aligned}$$

Utilizando el resultado anterior, por tabla obtenemos que

$$\frac{k - 28.000}{125} = 1,645 \Rightarrow k = 28.000 - 125 * 1.645 = 27794,375$$

Como $\overline{x} = 27.800 > 27794,375$ no se rechaza H_0 con una confianza del 95%.

- b El p-valor (α_p) es el menor nivel de significancia para el cual se rechaza H_0.

- (b) El p-valor (α_p) es el menor nivel de significancia para el cual se rechaza H_0.

$$\alpha_p = \text{p-valor} = P[\text{Rechazar el estadístico muestral}/H_0 \text{ es cierta}]$$

- b) El p-valor (α_p) es el menor nivel de significancia para el cual se rechaza H_0.

$$\alpha_p = \text{p-valor} = P[\text{Rechazar el estadístico muestral}/H_0 \text{ es cierta}]$$

Si $\alpha_p > \alpha$ no se rechaza la hipótesis nula H_0.

- b) El p-valor (α_p) es el menor nivel de significancia para el cual se rechaza H_0.

$\alpha_p =$ p-valor $= P[$Rechazar el estadístico muestral$/H_0$ es cierta$]$

Si $\alpha_p > \alpha$ no se rechaza la hipótesis nula H_0.

$$\alpha_p = P[\overline{x} < 27.800/H_0 cierta] = P[\overline{x} < 27.800/N(28.000, 125)] =$$
$$= P\left[\frac{\overline{x} - 28.000}{125} < \frac{27.800 - 28.000}{125}\right] = P[Z < -1,6]$$
$$= P[Z > 1,6] = 0,0548$$

- b. El p-valor (α_p) es el menor nivel de significancia para el cual se rechaza H_0.

$\alpha_p =$ p-valor $= P[$Rechazar el estadístico muestral$/H_0$ es cierta$]$

Si $\alpha_p > \alpha$ no se rechaza la hipótesis nula H_0.

$$\begin{aligned}\alpha_p &= P[\overline{x} < 27.800/H_0 cierta] = P[\overline{x} < 27.800/N(28.000, 125)] = \\ &= P\left[\frac{\overline{x} - 28.000}{125} < \frac{27.800 - 28.000}{125}\right] = P[Z < -1,6] \\ &= P[Z > 1,6] = 0,0548\end{aligned}$$

Con lo anterior cualquier $\alpha < \alpha_p$ no se rechaza H_0

Pregunta 3

Se ha realizado un estudio sobre la resistencia de dos materiales el primero X es un polimero sintetico de goma y el segundo Y un compuesto de carbono. La taza de exito de una muestra es la siguiente.

$$
\begin{aligned}
n_x &= 250 \\
n_y &= 250 \\
p_x &= \frac{101}{250} = 0,404 \\
p_y &= \frac{75}{250} = 0,300
\end{aligned}
$$

Se sabe que el material X cuesta un poco más, por lo que es de esperar que tenga un desempeño superior al material Y, corroboré esto a travez de un test de hipotesis adecuado.

Solución:

Realicemos una prueba de hiptesis unilateral para poder determinar un mejor desempeño de X sobre Y.

Solución:
Realicemos una prueba de hiptesis unilateral para poder determinar un mejor desempeño de X sobre Y.

$$H_0 : p_x \leq p_y$$
$$vs$$
$$H_1 : p_x > p_y$$

Solución:
Realicemos una prueba de hiptesis unilateral para poder determinar un mejor desempeño de X sobre Y.

$$H_0 : p_x \leq p_y$$
$$vs$$
$$H_1 : p_x > p_y$$

Tenemos que bajo H_0.

Solución:
Realicemos una prueba de hiptesis unilateral para poder determinar un mejor desempeño de X sobre Y.

$$H_0 : p_x \leq p_y$$
$$vs$$
$$H_1 : p_x > p_y$$

Tenemos que bajo H_0.

$$Z = \frac{\hat{p}_x - \hat{p}_y}{\sqrt{\tilde{p}(1-\tilde{p})(\frac{1}{n_X} + \frac{1}{n_Y})}} \sim N(0,1)$$

Solución:
Realicemos una prueba de hiptesis unilateral para poder determinar un mejor desempeño de X sobre Y.

$$H_0 : p_x \leq p_y$$
$$vs$$
$$H_1 : p_x > p_y$$

Tenemos que bajo H_0.

$$Z = \frac{\hat{p}_x - \hat{p}_y}{\sqrt{\tilde{p}(1-\tilde{p})(\frac{1}{n_X} + \frac{1}{n_Y})}} \sim N(0,1)$$

Donde $\tilde{p} = 0,352$ y $Z = 2,43$,

Solución:
Realicemos una prueba de hiptesis unilateral para poder determinar un mejor desempeño de X sobre Y.

$$H_0 : p_x \leq p_y$$
$$vs$$
$$H_1 : p_x > p_y$$

Tenemos que bajo H_0.

$$Z = \frac{\hat{p}_x - \hat{p}_y}{\sqrt{\tilde{p}(1-\tilde{p})(\frac{1}{n_X} + \frac{1}{n_Y})}} \sim N(0,1)$$

Donde $\tilde{p} = 0,352$ y $Z = 2,43$, con lo que tenemos que el p-valor=$P[Z \geq z] = P[Z \geq 2,43] = 0,0075$.

Solución:
Realicemos una prueba de hiptesis unilateral para poder determinar un mejor desempeño de X sobre Y.

$$H_0 : p_x \leq p_y$$
$$vs$$
$$H_1 : p_x > p_y$$

Tenemos que bajo H_0.

$$Z = \frac{\hat{p}_x - \hat{p}_y}{\sqrt{\tilde{p}(1-\tilde{p})(\frac{1}{n_X} + \frac{1}{n_Y})}} \sim N(0,1)$$

Donde $\tilde{p} = 0,352$ y $Z = 2,43$, con lo que tenemos que el p-valor=$P[Z \geq z] = P[Z \geq 2,43] = 0,0075$. Con lo que rechazamos la hipotesis nula en cualquiera de los niveles de significancia tipicos,

Solución:
Realicemos una prueba de hiptesis unilateral para poder determinar un mejor desempeño de X sobre Y.

$$H_0 : p_x \leq p_y$$
$$vs$$
$$H_1 : p_x > p_y$$

Tenemos que bajo H_0.

$$Z = \frac{\hat{p}_x - \hat{p}_y}{\sqrt{\tilde{p}(1-\tilde{p})(\frac{1}{n_X} + \frac{1}{n_Y})}} \sim N(0,1)$$

Donde $\tilde{p} = 0,352$ y $Z = 2,43$, con lo que tenemos que el p-valor=$P[Z \geq z] = P[Z \geq 2,43] = 0,0075$. Con lo que rechazamos la hipotesis nula en cualquiera de los niveles de significancia tipicos, con lo que es factible considerar que el material X tiene un desempeño superior.