Pontificia Universidad Católica de Chile
Facultad de Matemáticas
Departamento de Estadística

Segundo Semestre 2022

Curso : Probabilidad y Estadística
Sigla : EYP1113
Profesores : Ricardo Aravena C., Ricardo Olea O. y Felipe Ossa M.

PAUTA INTERROGACIÓN 4

Pregunta 1

Durante el último tiempo se han observado fuertes bajas en los valores de los departamentos a la venta. Un especialista afirma que estás rebajas, por m^2 construido, son superiores a 10 UF. Usted, con el fin de comprobar la afirmación, selecciona 28 departamentos a la venta en el sector oriente de la capital, y registra las rebajas ofrecidas (en UF).

Un breve resumen se presenta a continuación:

```
summary(rebaja)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  3.368   8.749  11.264  10.945  14.198  16.631

sd(rebaja)
3.643366
```

(a) **[4.0 Ptos.]** Asumiendo que las rebajas siguen una distribución Normal, lleve a cabo el respectivo test de hipótesis, determine el valor-p e indique su decisión para un nivel de significancia del 10 %.

(b) **[2.0 Ptos.]** Dado el resultado del análisis previo, un especialista indica que el tamaño de muestra es insuficiente y le indica que el tamaño debe ser tal que el error de estimación no sea superior a 1UF con un 95 % de confianza. Utilizando los resultados previos, ¿cuál es el tamaño de muestra requerido?

Solución

(a) **Logro 1: Hipótesis**

Test de hipótesis con respecto a μ.

$$\text{H}_0 : \mu = \mu_0 \quad \text{vs} \quad \text{H}_a : \mu > \mu_0, \quad \textbf{[0.8 Ptos.]}$$

con $\mu_0 = 10$. **[0.2 Ptos.]**

Logro 2: Estadístico de prueba y distribución de probabilidad

Bajo H_0 se tiene que

$$T_0 = \frac{\overline{X}_n - \mu_0}{S/\sqrt{n}} \sim \text{t-Student}(n-1) \quad \textbf{[0.5 Ptos.]}$$

Reemplazando

$$T_0 = \frac{10.945 - 10}{3.643366/\sqrt{28}} = 1.372486. \quad \textbf{[0.5 Ptos.]}$$

Logro 3: Valor-p

valor-p $= P(T > 1.372486) = 1 - P(T \leq 1.372486) \rightarrow 5\,\% <$ valor-p $< 10\,\%$. **[1.0 Ptos.]**,

con $T \sim$ t-Student$(28 - 1)$.

Logro 4: Conclusión

Por lo tanto, existe suficiente evidencia para rechazar H_0 y apoyar la hipótesis que las rebajas superan las 10 UF por m^2. **[1.0 Ptos.]**

(b) **Logro 5: Usar resultado previos**

Usando los resultados previos se tiene que

$$n = \left(\frac{S \cdot k_{0.975}}{1}\right)^2 \quad \textbf{[1.0 Ptos.]}$$

Logro 6: Resultado

$$n = \left(\frac{3.643366 \cdot 1.96}{1}\right)^2 = 50.99384 \rightarrow n = 51 \quad \textbf{[1.0 Ptos.]}$$

+ 1 Punto Base

Pregunta 2

Esta semana los futuros estudiantes universitarios rindieron la PAES y un especialista afirma que las mujeres sienten una mayor dificultad que los hombres en la prueba de Matemáticas, para ello selecciona al azar una muestra de hombres y una muestra de mujeres, consultando si cree que la prueba de Matemáticas estaba más difícil de lo esperado. Los resultados son:

	Hombres	Mujeres
Encuestados	72	48
¿Más difícil de lo esperado? SI	18	20

Lleve a cabo la respectiva prueba de hipótesis que le permita confirmar o refutar la afirmación del especialista. Comente considerando un nivel de significancia del 5 %.

Solución

Definamos como

$$X\text{: Hombres},\quad Y\text{: Mujeres}$$

Logro 1: Hipótesis

Se pide

$$\text{H}_0 : p_X = p_Y \quad \text{vs} \quad \text{H}_a : p_X < p_Y \quad \textbf{[1.0 Ptos.]}$$

Logro 2: Estadístico de prueba y su distribución

Bajo H_0 se tiene que

$$Z_0 = \frac{\overline{X}_n - \overline{Y}_m}{\sqrt{\hat{p}\,(1-\hat{p})}\sqrt{\frac{1}{n}+\frac{1}{m}}} \sim \text{Normal}(0,1) \quad \textbf{[0.5 Ptos.]}$$

con $\hat{p} = \dfrac{n \cdot \overline{X}_n + m \cdot \overline{Y}_m}{n+m}$, estimador máximo verosímil bajo H_0. **[0.5 Ptos.]**

Logro 3: Valor estadístico de prueba

Del enunciado se tiene que

$$n = 72, \quad m = 48, \quad \overline{X}_n = \frac{18}{72}, \quad \overline{Y}_m = \frac{20}{48}.$$

Reemplazando

$$\textbf{[0.5 Ptos.]} \quad \hat{p} = 0.3166667 \to Z_0 = -1.922771 \quad \textbf{[0.5 Ptos.]}$$

Logro 4: Dirección valor-p o valor critico.

Alternativa 1: valor-p $= \Phi(Z_0)$. **[1.0 Ptos.]**

Alternativa 2: Valor critico k_α. **[1.0 Ptos.]**

Logro 5: Valor-p o valor critico.

Alternativa 1: valor-p $= \Phi(-1.922771) \approx 1 - \Phi(1.92) = 1 - 0.9726 = 0.0274$. **[1.0 Ptos.]**

Alternativa 2: $k_{0.05} = -1.654$. **[1.0 Ptos.]**

Logro 6: Conclusión

Alternativa 1: Dado que el valor-p $< \alpha = 5\,\%$, entonces existe suficiente evidencia para rechazar H_0 y apoyar la hipótesis del especialista. **[1.0 Ptos.]**

Alternativa 2: Dado que el $Z_0 < k_{0.05}$, entonces existe suficiente evidencia para rechazar H_0 y apoyar la hipótesis del especialista. **[1.0 Ptos.]**

+ 1 Punto Base

Pregunta 3

Considere una muestra aleatoria Exponencial(ν) de tamaño n. Determine la distribución asintótica del estimador máximo verosímil de la mediana.

Solución

Logro 1: EMV de ν

El estimador máximo verosimil de ν está dado por $\dfrac{1}{\overline{X}_n}$. **[1.0 Ptos.]**

Logro 2: Mediana teórica

Si $X \sim$ Exponencial(ν), entonces la mediana es $\dfrac{\ln(2)}{\nu} = g(\nu)$. **[1.0 Ptos.]**

Logro 3: EMV mediana

El estimador máximo verosimil de la mediana, por invarianza, está dado por

$$\widehat{\text{mediana}} = g(\hat{\nu}) = \ln(2) \cdot \overline{X}_n \quad \textbf{[1.0 Ptos.]}$$

Logro 4: Aproximación de 1er orden valor esperado

$$\text{E}(\widehat{\text{mediana}}) = \frac{\ln(2)}{\nu} \quad \textbf{[1.0 Ptos.]}$$

Logro 5: Aproximación de 1er orden varianza

Usando CCR

$$\text{Var}(\widehat{\text{mediana}}) = [g(\nu)']^2 \cdot \text{CCR} = \left[-\frac{\ln(2)}{\nu^2}\right]^2 \cdot \frac{\nu^2}{n} = \frac{[\ln(2)]^2}{n\,\nu^2} \quad \textbf{[1.0 Ptos.]}$$

Usando directamente Var$(\overline{X}_n)$

$$\text{Var}(\widehat{\text{mediana}}) = [\ln(2)]^2 \cdot \text{Var}(\overline{X}_n) = \frac{[\ln(2)]^2}{n\,\nu^2} \quad \textbf{[1.0 Ptos.]}$$

Logro 6: Distribución asintótica EMV

$$\widehat{\text{mediana}} \overset{\text{aprox}}{\sim} \text{Normal}\left(\frac{\ln(2)}{\nu}, \frac{\ln(2)}{\sqrt{n}\,\nu}\right) \quad \textbf{[1.0 Ptos.]}$$

+ 1 Punto Base

Pregunta 4

En general, los rendimientos de los alumnos en sucesivas evaluaciones se encuentran correlacionadas. Más aún, se piensa que el rendimiento de la 2da interrogación puede ser explicada en parte por el rendimiento de la 1era interrogación.

(a) **[3.0 Ptos.]** Para evaluar la hipótesis del enunciado, se ajusta un modelo de regresión a una muestra aleatoria de 25 alumnos cuya salida parcial es la siguiente:

```
lm(formula = I2 ~ I1)

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   1.6112     1.1810   1.364   0.1857
I1            0.4962     0.2554   XXXXX   0.0644

Residual standard error: 1.219 on XXXXX degrees of freedom
Multiple R-squared:  XXXXX, Adjusted R-squared:  0.1036
F-statistic: XXXXX on 1 and XXXXX DF,  p-value: 0.0644
```

Complete los campos faltantes y comente utilizando un nivel de significancia del 5 %.

(b) **[3.0 Ptos.]** Uno de los supuestos del modelo de regresión es la normalidad de lo datos. El estadístico chi-cuadrado para las notas I2, sobre una muestra aleatoria de 200 alumnos, es igual a 1.536167 para los siguientes intervalos:

```
[1,3] (3,4] (4,5] (5,7]
   14    49    70    67
```

Por otra parte, el valor-p del estadístico KS aplicado a estos datos para el supuesto antes mencionado es 0.4700453. Comente.

Solución

(a) **Logro 1: t value y degrees of freedom**

$$\textbf{[0.5 Ptos.]}\quad \texttt{t value} = \frac{\texttt{Estimate}}{\texttt{Std. Error}} = 1.9428$$
$$\texttt{degrees of freedom} = 25 - 2 = 23 \quad \textbf{[0.5 Ptos.]}$$

Logro 2: F y Multiple R-squared

$$\texttt{F-statistic} = \texttt{t value}^2 = 3.7746 \quad \textbf{[0.5 Ptos.]}$$
$$\texttt{Multiple R-squared} = \frac{\texttt{F-statistic}}{25 - 2 - \texttt{F-statistic}} = 0.141 \quad \textbf{[0.5 Ptos.]}$$

Logro 3: Comentario

- El modelo explica un 14.1 % de la variabilidad presente en los datos. **[0.3 Ptos.]**
- Una décima en la I1, implica aproximadamente 0.5 puntos en la I2. **[0.3 Ptos.]**
- Pendiente no resulta significativo al 5 %. **[0.4 Ptos.]**

(b) **Logro 4: Distribución estadístico de prueba χ^2**

Dado que se deben estimar los parámetros para construir el estadístico:

$$X^2 = 1.536167 \sim \chi^2(4-1-2) \quad \textbf{[1.0 Ptos.]}$$

Logro 5: Valor p estadístico χ^2

$$20\,\% < \texttt{p value} < 30\,\% \quad \textbf{[1.0 Ptos.]}$$

Logro 6: Comentario

- Estadístico de prueba χ^2 no rechaza la normalidad. **[0.5 Ptos.]**
- Valor p estadístico KS no rechaza la normalidad. **[0.5 Ptos.]**

+ 1 Punto Base