

## Ejercicios adicionales correspondientes a unidades 6 a 9

### Ejercicio 1

Se desea estimar el ingreso medio de los habitantes de una ciudad, para lo cual se tomó una muestra de 100 personas que arrojó una media de \$625 y un desvío estándar de \$576. Estime un intervalo de confianza adecuado con un nivel de confianza del 90% e interprete. ¿Qué supuesto debe realizar?

Rta:  $P(529,384 < \mu < 720,616) = 90\%$  Supuesto que X sigue una distribución normal con media 625 y desvío 576.

### Ejercicio 2

Indicar si las siguientes afirmaciones son verdaderas o falsas **justificando** su respuesta.

A) "Dados los siguientes intervalos de confianza estimados, en los cuales se utilizó una muestra de tamaño 20 en ambos casos:

$$(10,0 < \mu < 11,2) = 90\%$$

$$(9,9 < \mu < 11,3) = 95\%$$

Es posible concluir que, a igual tamaño de muestra y mayor nivel de confianza, se obtiene mayor precisión".

Rta: FALSO, justificar

B) "En un modelo de regresión lineal, el supuesto de homogeneidad de la varianza significa que la variabilidad de Y no es la misma con cada valor de X"

Rta: FALSO, justificar

C) "En el modelo de regresión lineal, el coeficiente de determinación viene dado por uno menos el cociente entre la suma de cuadrados del error y la suma de cuadrados total, brindando una medida de la proporción de la variación de Y que es explicada por el modelo"

Rta: VERDADERO, justificar

### Ejercicio 3

Se lleva a cabo un estudio para comparar el tiempo que tardan los operarios de la planta A y de la planta B en armar un producto determinado. Una muestra aleatoria de tiempos para 11 operarios de la planta A arrojó un desvío de 6,1 minutos y una muestra de 14 operarios de la planta B arrojó un desvío de 5,3 minutos. Probar con un nivel del 1% la hipótesis de que la varianza de los tiempos de los operarios de la planta A es mayor que la de los operarios de la planta B. ¿Qué supuesto debe realizar?

Rta: Especificar supuesto. No se rechaza  $H_0$ .

### Ejercicio 4

Un banco internacional realizó un estudio sobre el costo de vida, salarios, entre otras variables, en 73 ciudades del mundo. En la imagen se observan los resultados del análisis de regresión lineal realizado tomando como variable independiente al salario neto promedio por hora (en dólares) y como variable dependiente la cantidad de horas necesarias para comprar un iPhone (con un salario promedio).

```
=====
                        OLS Regression Results
=====
Dep. Variable:          iPhone 4S(hr)      R-squared:                0.652
Model:                  OLS                Adj. R-squared:         0.647
Method:                 Least Squares       F-statistic:             131.3
Date:                   Sun, 09 Jun 2024    Prob (F-statistic):       1.03e-17
Time:                   13:06:54            Log-Likelihood:          -391.88
No. Observations:       72                 AIC:                     787.8
Df Residuals:           70                 BIC:                     792.3
Df Model:                1
Covariance Type:        nonrobust
=====
                        coef    std err          t      P>|t|      [0.025    0.975]
-----
const          238.8875      12.607      18.949    0.000      213.744    264.031
x1             -2.5156       0.220     -11.457    0.000      -2.953    -2.078
=====
```

A) Especifique el modelo obtenido, definiendo cada componente e indicando que representa.

Rta:  $Y = 238,89 - 2,52 * X$ . Definir cada componente del modelo (Y, X y los betas). Indicar que significa cada coeficiente del modelo (interpretar en el contexto del problema).

B) ¿Son significativas la ordenada al origen y la pendiente de la recta? Plantee de forma completa los Test de Hipótesis y concluya.

Rta: la ordenada al origen y la pendiente son significativas.

C) ¿En cuánto por ciento la variación de Y es explicada por el modelo? Justifique.

Rta: la variación en las horas promedio para comprar un iPhone es explicada por el modelo de regresión lineal en un 65%.

D) Especifique cual es el intervalo de confianza al 95% para la pendiente de la recta de regresión.

Rta:  $P(-2,95 < \beta_1 < -2,08) = 95\%$

### Ejercicio 5

Un proceso de armado de autos requiere número diferente de horas para terminarlo. Por experiencias anteriores se sabe el número de horas empleadas sigue una distribución normal con media igual a 12 y varianza igual a  $4 \text{ hs}^2$ . Si se toman muestras aleatorias de 25 casos y se sabe que la media muestral está entre 11,2 y 12,8 ¿Cuál es el mínimo tamaño de muestra que debe tomarse si se quiere que la probabilidad de que la media  $\bar{X}$  esté entre esos mismos valores sea por lo menos del 80%?

Rta: El mínimo tamaño de muestra para que la media muestral esté entre 11,2 y 12,8 es 11

### Ejercicio 6

Indicar si las siguientes afirmaciones son verdaderas o falsas justificando su respuesta:

A) "Si n (tamaño de la muestra) es lo suficientemente grande [ $n > 30$ ], entonces  $\bar{X}$  no tiene aproximadamente una distribución  $N(\mu; \frac{\sigma}{\sqrt{n}})$ "

Rta: FALSO, justificar

B) "En una estimación por intervalo de confianza, a medida que se aumenta el tamaño de la muestra, para un mismo nivel de confianza, se obtiene una estimación menos precisa".

Rta: FALSO, justificar

### Ejercicio 7

En un taller se cortan telas cuya longitud es una variable aleatoria con distribución normal. Se miden 25 cortes realizados al azar, obteniéndose  $\bar{X}=180 \text{ cm}$  y  $S=10 \text{ cm}$ .

a) Definir la variable sobre la cual se realizará la inferencia.

Rta: X: "longitud de las telas (en cm)"  $X \sim N(\mu; \sigma)$

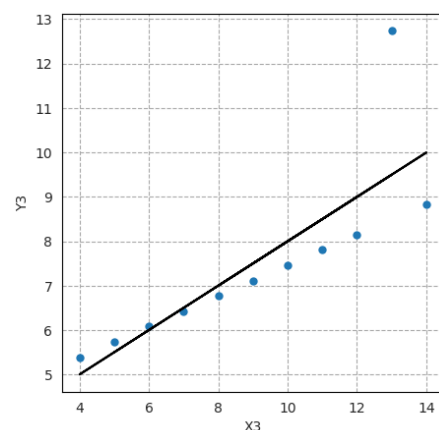
b) Hallar un intervalo de confianza del 0.95 para la varianza.

Rta: La varianza de la longitud de las telas se encontrará entre 60,97 y 193,53 con un 95% de confianza

### Ejercicio 8

Dadas dos variables Y3 y X3, se desea estudiar la relación lineal entre ellas. Para ello se estima un modelo de regresión lineal, obteniéndose los siguientes resultados:

OLS Regression Results						
=====						
Dep. Variable:	Y3	R-squared:	0.666			
Model:	OLS	Adj. R-squared:	0.629			
Method:	Least Squares	F-statistic:	17.97			
Date:	Sun, 09 Jun 2024	Prob (F-statistic):	0.00218			
Time:	13:23:19	Log-Likelihood:	-16.838			
No. Observations:	11	AIC:	37.68			
Df Residuals:	9	BIC:	38.47			
Df Model:	1					
Covariance Type:	nonrobust					
=====						
	coef	std err	t	P> t	[0.025	0.975]
-----						
const	3.0025	1.124	2.670	0.026	0.459	5.546
x1	0.4997	0.118	4.239	0.002	0.233	0.766
=====						



a) Especifique el modelo obtenido, indicando que significa cada coeficiente.

Rta:  $Y_3 = 3 + 0,5 * X_3$ . Indicar que significa cada coeficiente del modelo.

b) ¿Qué representa la pendiente de la recta de regresión obtenida?

$\beta_1=0,5$  hacer su interpretación

c) ¿Son significativas la ordenada al origen y la pendiente de la recta? Plantee de forma completa los Test de Hipótesis y concluya.

Rta: ambos son significativos

d) ¿En cuánto por ciento la variación de  $Y_3$  es explicada por el modelo? Justifique.

Rta: Dado que el  $R^2$  es igual a 0,67 entonces es posible establecer que el modelo explica el 67% de la proporción de la variación de  $Y_3$

## Ejercicio 9

Un banco quiere estimar el monto medio de los préstamos hipotecarios. Si los montos se distribuyen normalmente con desvío \$1200: ¿de qué tamaño deberá tomarse la muestra si se quiere con una confianza del 95% que la longitud del intervalo sea a lo sumo de \$3000?

Rta: El tamaño de muestra deberá ser de al menos 3 préstamos para estimar el monto medio de estos con un 95% de confianza.

## Ejercicio 10

Una fábrica de artículos de consumo masivo ha lanzado una campaña publicitaria sobre un producto, a efectos de incrementar las ventas. Se registraron las ventas durante los cuatro meses de la campaña y se obtuvo una media de 642 unidades (en miles) y un desvío de 85.25 unidades (en miles). Además, se dispone de registros históricos de los últimos doce meses previos a la campaña donde la media de ventas es 570 unidades (en miles) con un desvío de 45 unidades (en miles).

A) Defina las variables aleatorias. ¿Qué supuesto debe realizar? Realice el test que valide el mismo.

Rta: No rechaza  $H_0$ .

B) Para comprobar si la campaña ha sido efectiva (las ventas han aumentado) con un 95% de confianza se pide realizar el test de hipótesis adecuado.

Rta: Se rechaza  $H_0$ .

## Ejercicio 11

Se realizó un estudio sobre el gasto en bienes y servicios y el salario. En la imagen se observan los resultados del análisis de regresión lineal realizado tomando como variable **dependiente** al costo de bienes y servicios y como variable **independiente** al salario.

OLS Regression Results						
=====						
Dep. Variable:	Bienes y servicios(\$)		R-squared:	0.807		
Model:	OLS		Adj. R-squared:	0.804		
Method:	Least Squares		F-statistic:	291.8		
Date:	Sun, 09 Jun 2024		Prob (F-statistic):	1.14e-26		
Time:	13:30:24		Log-likelihood:	-516.87		
No. Observations:	72		AIC:	1038.		
Df Residuals:	70		BIC:	1042.		
Df Model:	1					
Covariance Type:	nonrobust					
=====						
	coef	std err	t	P> t	[0.025	0.975]
-----						
const	1683.2904	71.547	23.527	0.000	1540.595	1825.986
x1	21.2852	1.246	17.082	0.000	18.800	23.770
-----						

A) Especifique el modelo obtenido, definiendo conceptualmente cada componente en términos del problema.

Rta:  $Y = 1683,29 + 21,29 * X$ . Definir cada componente del modelo. Indicar que significa cada coeficiente del modelo (interpretar en el contexto del problema).

B) ¿Son significativas la **ordenada al origen y la pendiente** de la recta? Plantee de forma completa los Test de Hipótesis y concluya. Rta: la ordenada al origen y la pendiente son significativas.

C) ¿En cuánto por ciento la variación de Y es explicada por el modelo? Justifique. Rta: Dado que  $R^2$  es 0,81, implica que la variación en las horas promedio para solventar el gasto en bienes y servicios es explicada por el modelo de regresión lineal en un 81%.

D) Especifique cual es el intervalo de confianza al 95% para la pendiente de la recta de regresión.

Rta:  $P(18,8 < \beta_1 < 23,77) = 95\%$

### Ejercicio 12

El encargado del departamento de producción de una fábrica recibe un lote de 2000 piezas necesarias para el montaje de un artículo. El fabricante de las piezas asegura que en este lote no hay más de 100 piezas defectuosas.

A) ¿Cuántas piezas hay que examinar para que, con un nivel de confianza del 95%, el error que se cometa en la estimación de la proporción de piezas defectuosas sea 0.05?

Rta: La cantidad de piezas que se deben examinar es 73.

B) Si se toma una muestra de 100 piezas elegidas al azar y se encuentran 4 defectuosas, determinar un intervalo de confianza para la proporción de defectuosas al nivel del 95%.

Rta:  $P(0,0016 < \mu < 0,0784) = 95\%$

### Ejercicio 13

En una empresa se toma el monto promedio de consumos realizados con tarjeta de crédito por el cliente A que resulta ser \$1490 con una dispersión de \$150 sobre una muestra de 13 transacciones. Para el cliente B, el monto promedio de sus consumos con tarjeta de crédito resulto ser \$1600 con una dispersión de \$94,23 sobre 15 transacciones.

A) ¿Existe suficiente evidencia para suponer que las varianzas son iguales? Utilice un  $\alpha=0,05$

Rta: No se rechaza  $H_0$ .

B) ¿Existe suficiente evidencia para asegurar que la media del cliente A difiere significativamente en 110 respecto de la media del cliente B? Utilice un  $\alpha=0,05$

Rta: Se rechaza  $H_0$ .

### Ejercicio 14

Dado los siguientes índices de cotización de dólar blue y oficial:

fecha	I_blue	I_Doficial
3/1/2022	100	
18/7/2022	141,26	
19/7/2022	146,12	
20/7/2022	153,88	
21/7/2022	163,59	100
22/7/2022	164,08	100,12
25/7/2022	156,31	100,63
26/7/2022	156,80	100,85
27/7/2022	158,25	101,03
28/7/2022	152,43	101,19
29/7/2022	143,69	101,30

Comparar la variación de cotización del dólar blue versus el oficial en el día 29/7/2022 respecto de 21/7/2022. ¿Cuál cotización tuvo mayor incremento en el período? Rta: el índice del dólar oficial sufrió mayor aumento.

¿Cuál cotización tuvo mayor incremento en 29/7/2022 respecto de 28/7/2022? Rta: el índice del dólar oficial sufrió mayor aumento.

### Ejercicio 15

Se quiere estudiar la vida útil de unas nuevas pilas que se van a lanzar al mercado. Para ello se examina la duración de 40 de ellas, resultando una media de 63 horas. Suponiendo que el tiempo de vida de las pilas sigue una distribución normal, y que la varianza se puede tomar la misma que las fabricadas anteriormente que era 38,44, se pide:

a) Intervalo de confianza del 95% de la duración media de las nuevas pilas. Redondear a 4 decimales el resultado final. Rta:  $P(61,0786 \leq \mu \leq 64,9214) = 95\%$

b) Intervalo de confianza del 99% de la duración media de las nuevas pilas. Redondear a 4 decimales el resultado final. ¿Qué conclusión puede extraer del resultado obtenido en A en comparación con el obtenido en B?

Rta:  $P(60,4747 \leq \mu \leq 65,5253) = 99\%$ . Realizar la conclusión.