

**Sección Teórica:** Responda de forma clara y concisa las preguntas que se le presentan a continuación.

- a. ¿Por qué es necesario hacer prueba de hipótesis a los coeficientes estimados de un modelo de regresión?

Para determinar si las relaciones que reflejan los coeficientes estimados utilizando datos muestrales, pueden generalizarse a nivel de la población. Generalmente en las pruebas de hipótesis se evalúan si los coeficientes estimados son estadísticamente diferentes de cero.

- b. ¿Qué es un intervalo de confianza y cómo se relaciona con la prueba de hipótesis?

Son estimaciones probabilísticas que permiten determinar el rango de valores que podría tomar un parámetro de interés. En el caso de un modelo de regresión, el IC proporciona una estimación (dado cierto grado de confiabilidad) alrededor de la cual se encuentra el parámetro de interés  $\beta$ . La lectura del IC permite llegar a la misma conclusión de una prueba de hipótesis. Si el IC contiene el cero eso significa que no se rechazó la  $H_0$  y por lo tanto el coeficiente no es significativo.

- c. Describa el procedimiento que se debe seguir para determinar la significancia estadística de un coeficiente.

Primero debe plantearse la hipótesis, luego debe calcularse el error estándar y determinarse el estadístico  $t$  o su respectivo valor  $p$ . Luego considerando un nivel de significancia, debe calcularse el valor crítico  $t$  considerando el número de observaciones del modelo y el número de parámetros estimados para determinar los grados de libertad. Si el valor del estadístico  $t$  es mayor al valor crítico (o si el valor  $p$  es inferior al nivel de significancia, teniendo presente si es una prueba unilateral o bilateral), se rechaza  $H_0$ , concluyendo que el coeficiente es estadísticamente diferente de cero.

- d. “Un modelo econométrico en donde sus coeficientes son altamente significativos (valor  $p$  cercano a cero), puede llevarnos a establecer causalidad entre la variable dependiente y las independientes” (verdadero o falso, justifique claramente su respuesta)

Falso. La relación de causalidad debe evaluarse a través del marco teórico bajo el cual se construye el modelo de regresión. En caso contrario únicamente podemos estar evaluando una correlación a través de la relación econométrica.

- e. ¿Qué información nos proporciona el rechazar  $H_0$  al realizar una prueba de hipótesis sobre un coeficiente estimado?

Depende de la forma en que esté planteada  $H_0$ . Generalmente, al plantear  $H_0: \beta_i = 0$ , nos indica que el coeficiente estimado es estadísticamente diferente de cero. Por lo tanto, podemos generalizar la relación encontrada en la muestra a nivel de la población.

- f. ¿Qué es más importante al evaluar un modelo de regresión: un  $R^2$  cercano a 1 o que sus coeficientes sean altamente significativos (valor p cercano a cero)?

Ambos, ya que generalmente tiende a coincidir una regresión con un  $R^2$  elevado (mayor a 0.5) con coeficientes altamente significativos (valor p cercano a cero). No obstante, el investigador tiene relaciones particulares que desea probar, desea establecer si ciertas variables tienen un impacto particular. Por ello, termina teniendo más importancia revisar la significancia estadística de los coeficientes. El  $R^2$  debe examinarse con detenimiento, existen casos particulares en donde se obtiene un  $R^2$  cercano a 1 pero que ningún coeficiente es significativo, este es un caso de multicolinealidad.

**Sección Práctica:** Resuelva el siguiente ejercicio, deje constancia de todo su procedimiento.

1. Se quiere encontrar si existe una relación lineal entre el costo total de producción (medido en Quetzales) y el número de unidades producidas por distintos vendedores de películas pirata.

**Tabla 1: Costos y Unidades producidas**

Obs.	Costo	Q
1	800	250
2	750	300
3	1,500	500
4	950	400
5	666	199
6	2,000	425
7	750	300
8	700	450
9	742	154
10	1,000	450

- a. A partir de los datos de la tabla 1 estime  $B_0$  y  $B_1$  e interprete los resultados a la luz del contexto del problema.

$B_0 = 273.52$ . Es el costo fijo. Es decir, es costo total cuando el número de unidades producidas es igual a cero.

$B_1 = 2.08$ . Es el costo variable. Es decir, cuando número de unidades producidas aumenta en una unidad, el costo total se incrementa en 2.08 quetzales.

- b. Estime el  $R^2$  e interprete el resultado obtenido.

$R^2 = 32.40\%$ . Significa que la variación del número de unidades producidas logran explicar en un 32.4% la variación de los costos.

- c. Si el error estándar de  $B_0$  es 91.43 y el de  $B_1$  es 1.05, ¿Son significativos los coeficientes con un  $\alpha=5\%$ ?, evidencie claramente sus hipótesis de trabajo y los cálculos que realiza para responder a la pregunta.

Con  $GL=8$ , un Alfa de 5% y  $T_{Crítico}=2.30$  (se usó alfa medios)

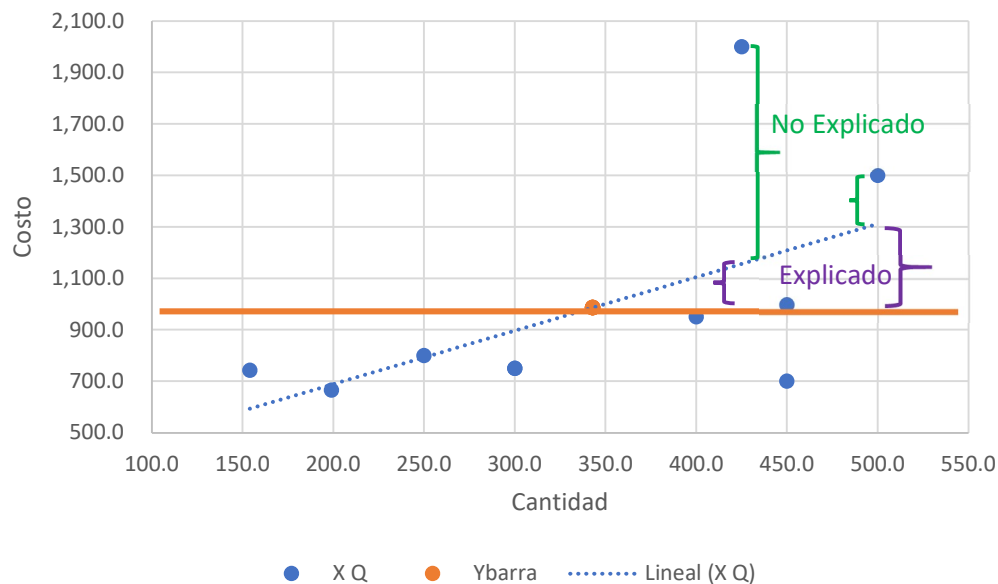
Estadístico  $t_{B1} = 1.98$ . No se rechaza  $H_0$ .

Estadístico  $t_{B0} = 2.99$ . Rechaza  $H_0$ .

- d. Construya un intervalo de confianza para  $B_1$  con un 95% de certeza y determine la relación que existe entre sus resultados y la conclusión del inciso anterior.

		Límite Superior	Límite Inferior
<b>b1</b>	2.08	4.5251	-0.3694
<b>b0</b>	273.52	484.36	62.68

- e. Para dos de los diez datos de la regresión, en el siguiente diagrama de dispersión identifique la variación total, explicada y no explicada.



2. En Guatemala se realizó un estudio econométrico para determinar el consumo mensual de las personas (en quetzales) en función del ingreso mensual (en quetzales). En la tabla 2 se muestran los resultados del modelo econométrico. Sin embargo, éstos se presentan con información incompleta.

**Tabla 2: Resultados econométricos de la función del consumo**

Variable Dependiente: Consumo			
Variable	Coefficiente	Error Estándar	Estadístico "t"
Constante	2.13	.94654	2.2
Ingreso	0.80	.40601	1.98
R <sup>2</sup>			0.8620
Sumatoria de cuadrados de la variación total			995.7
Sumatoria de cuadrados de los errores			137.41111
Tamaño de la muestra			15

- a) Complete la tabla 2 e interprete el significado de cada uno de los elementos que aparecen en la tabla y determine si los coeficientes son estadísticamente significativos.

Constante: Si el ingreso es igual a cero, el consumo mensual será 2.13 quetzales.

¿Consumo permanente?

Coefficiente: Si el ingreso aumenta en una unidad, el consumo lo hará en .80 centavos.

Con GL= 13, un Alfa de 5% y TCrítico= 2.16

Estadístico t B<sub>0</sub> = 2.2 Se rechaza Ho.

Estadístico t B<sub>1</sub> = 1.98 No se rechaza Ho.

3. Si el salario anual de un ejecutivo se estima por el siguiente modelo:

$$(\text{Salario})_i = \alpha + \beta_1 (\text{Promedio})_i + \beta_2 (\text{Experiencia})_i + \beta_3 (\text{Edad})_i + \varepsilon_i$$

Donde el salario está medido en dólares, el promedio es el promedio universitario con que se graduó, en valores de 1 a 100, la experiencia y la edad se miden en años. La estimación del modelo utilizando OLS da el siguiente resultado:

$$(\text{Salario})_i = 11,583 + 51.9(\text{Promedio})_i + 70.64 (\text{Experiencia})_i + 6.1(\text{Edad})_i + \varepsilon_i$$

(2,000)                      (30.5)                      (50.13)                      (4.8)

En paréntesis se muestran errores estándar.

- a. A partir del contexto del problema, interprete el significado de cada coeficiente, incluyendo la constante.

Constante: Cuando todas las variables independientes son cero, el salario medio es igual a 11,583 dólares.

Promedio: Cuando el promedio unitario aumenta en una unidad el salario medio lo hace en 51.9 dólares.

Experiencia: Cuando la experiencia aumenta en un año el salario lo hace en 70.64 dólares.

Edad: Cuando la edad aumenta en un año el salario lo hace en 6.1 dólares.

- b. ¿Son significativos los coeficientes de forma individual?, utilice un alfa del 10% (n=20)

	Coeficiente	ee	T Estadístico	Significativo?
Constante	11,583.0	2,000.0	5.79	TRUE
Promedio	51.9	30.5	1.70	FALSE
Experiencia	70.6	50.1	1.41	FALSE
Edad	6.1	4.8	1.27	FALSE
GL	16.0			
Parametros Est.	3			
Alfa	10.00%			
T Critico	1.75			

4. Utilice los datos del archivo **HPRICE1.RAW** (Base de datos de Wooldridge) para estimar el modelo:

$$price = \beta_0 + \beta_1 sqrft + \beta_2 bdrms + u,$$

donde *price* es el precio de casas dado en miles de dólares.

- a. Estime el modelo de regresión e interprete el significado de los coeficientes dentro del contexto del problema.

Model 1: OLS, using observations 1-88  
Dependent variable: price

	<i>Coefficient</i>	<i>Std. Error</i>	<i>t-ratio</i>	<i>p-value</i>	
const	-19.3150	31.0466	-0.6221	0.5355	
sqrft	0.128436	0.0138245	9.291	<0.0001	***
bdrms	15.1982	9.48352	1.603	0.1127	
Mean dependent var	293.5460	S.D. dependent var		102.7134	
Sum squared resid	337845.4	S.E. of regression		63.04484	
R-squared	0.631918	Adjusted R-squared		0.623258	
F(2, 85)	72.96353	P-value(F)		3.57e-19	
Log-likelihood	-487.9989	Akaike criterion		981.9978	
Schwarz criterion	989.4298	Hannan-Quinn		984.9919	

Const: Cuando el tamaño de la casa es cero y el número de habitación también, el valor esperado del precio de la casa es -19.32 mil US\$.

Sqrft: Cuando el tamaño de la casa aumenta un pie cuadrado, el valor esperado del precio de la casa aumente en 0.13 mil US\$.

Bdrms: Cuando el aumenta en una unidad, el valor esperado del precio de la casa aumente en 15.2 mil US\$.

- b. Evalúe el R<sup>2</sup> y significancia de los coeficientes presentando una breve reflexión de lo que encontró.

La variación del tamaño de la casa y el número de habitación explica el 63.19% de la variación del precio de las casas. La variable Sqrft tiene un alto nivel de significancia, y es la única variable significativa dentro de esta estimación.

- c. La primera casa en la muestra tiene  $sqrft=2,438$  y  $bdrms=4$ . Determine el precio de venta estimado para esta casa con la línea de regresión de estimada a través de MCO.

	<i>Coefficient</i>
const	(19.32)
sqrft	0.13
bdrms	15.20
sqrft=	2,437.00
bdrms=	4.00
Price (2437,4)=	354.48

- d. El precio de venta de la primera casa en la muestra fue \$300,000 (así que  $price = 300$ ). Determine el residual para esta casa. ¿Sugiere esto que el comprador pagó de más o de menos por la casa?

Valor Esperado	354.48
Precio de Venta	300.00
<b>Diferencia</b>	<b>54.48</b>

Si el comprador 300 mil US\$ por la casa, dado que su valor esperado según modelo es 354 mil US\$, se concluye que el comprador se ahorró 54 mil US\$ en la compra de la casa.