

# ANÁLISIS ESTADÍSTICO

MA1031

Examen Argumentativo

Febrero-Junio 2023

**Instrucciones:** Lea cuidadosamente el problema antes de empezar a resolverlo. Cada problema debe tener un **procedimiento** adecuado y la respuesta correcta para ser considerado correcto. No se aceptan dos resultados diferentes para un problema ya que si escribe dos respuestas será anulado el reactivo. Marque cada resultado en un **recuadro**. Todos los resultados deben estar **simplificados**. Al terminar subir sus procedimientos a ambos apartados. (El examen está sobre 100 **puntos**.)

“Apegándome al Código de Ética de los Estudiantes del Tecnológico de Monterrey, me comprometo a que mi actuación en este examen esté regida por la honestidad académica”.

Nombre: gabriel Eduardo Melendez Zavala Matrícula: A01638293

- I. (40 puntos). El Camry de Toyota es uno de los automóviles más vendidos en Estados Unidos. El precio de un Camry usado depende de varios factores, cómo ser modelo, las millas recorridas y las condiciones generales. Con el fin de investigar la relación entre las millas recorridas “x” (Miles) y el precio de venta “y” (Price, miles de dólares) de un Camry modelo 2015, se presentan los datos de las millas recorridas y el precio de operaciones de venta. (datos tomados de PriceHub 2020).

Miles (1000s)	Price (1000s de dólares)
22	16.2
29	16.0
36	13.8
47	11.5
63	12.5
77	12.9
73	11.2
87	13.0
92	11.8
101	10.8
110	8.3
28	12.5
59	11.1
68	15.0
68	12.2
91	13.0
42	15.6
65	12.7
110	8.3

En un problema de millas recorridas, queremos, a partir de mediciones hechas en muestras standard, estudiar la relación entre las mediciones y las ventas. Esta relación permitirá en el futuro, medir una muestra desconocida y conocer aproximadamente su verdadero valor. Cada dato, gráfica, o resultado obtenido, debe ser interpretado con base en las variables y contexto del problema planteado.

Para resolver este caso, se le ha pedido realizar un análisis y debe responder a las siguientes peticiones, argumentando cada una de tus respuestas usando lo conveniente y un máximo de 100 palabras:

1. Qué modelo matemático consideras que se ajusta mejor a la muestra de datos.

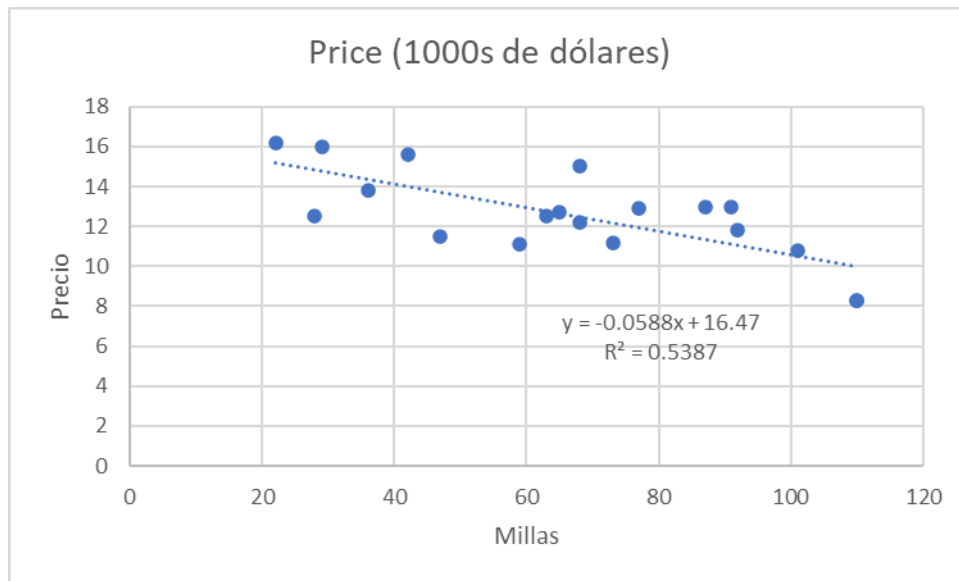
**Para generar el modelo matemático que mejor se ajusta a la muestra tenemos que conocer la gráfica de dispersión de la muestra y visualmente observar que tendencias tiene. En este caso como no es una serie de tiempo regresión lineal puede ser optimo.**

2. Generando un modelo lineal ¿Qué me indica la pendiente estimada en este contexto?

**La pendiente es:  $y = -0.0588x + 16.47$**

**Que estima el precio basado en el numero de millas del carro**

**Está pendiente indica que por cada milla que el carro contenga su evaluación disminuye 0.0588 miles de dólares.**



3. Pretendes comprar un Camry 2015 usado que tiene 60000 millas recorridas. El precio que predice, ¿lo ofrecerías al vendedor?

**El precio predicho de un Camry 2015 con 60000 millas recorridas es:**

$$y = -0.0588(60) + 16.47 = 12.94$$

**\$12,940 es el monto que nuestro análisis de regresión predice.**

**Al compararlo con otras ventas similares que ya se han hecho, por ejemplo, de 59000 millas su precio es de \$11,100. Nuestra predicción es un poco alta para ofrecerlo dada las previas ventas, sin embargo, no estamos tomando en cuenta otros factores como las condiciones del carro que afectan la evaluación de un carro. Teniendo estas en cuenta, podemos tener una mejor idea si ofrecer o no el carro al vendedor.**

4. ¿Que puedes sugerir a la empresa? ¿Qué puedes sugerir a los clientes?

**Lo que le puedo sugerir a la empresa es tener una mejor recolección de base de datos para sus vehículos vendidos ya que las millas del carro no son el único factor que afecta el precio de venta de un carro. Pueden agregar condición general del carro.**

**Para los clientes, tener una idea de su presupuesto al comprar un Toyota Camry 2015 y que sean conscientes que por lo menos el precio y las millas tienen una relación.**

5. Calcula e interpreta la proporción de la variación y la proporción con respecto al modelo estimado.

**La proporción de la variación o coeficiente de determinación con respecto al modelo es de,**

$$R^2 = 0.5389.$$

**Lo que significa que podemos explicar un 53.89% de la variación del modelo ajustado a los datos reales. En otras palabras, esto es una medida de precisión del modelo. Cuanto más el valor se acerque a 1, significa que los valores están siendo predichos con mayor certeza.**

**Adicionalmente, la proporción o la correlación del modelo estimado, R, nos indica la relación entre dos variables continuas. Si la correlación es menor a cero, significa que es negativa, es decir, que las variables se relacionan inversamente. Si la correlación es positiva significa que las variables se relacionan directamente.**

$$R = -0.734$$

**En este caso las variables se relacionan inversamente.**

II. (40 puntos). La demanda diaria de un cierto periódico durante los últimos 10 días ha sido: 12, 13, 16, 15, 12, 18, 14, 12, 11, 14 (desde los menos recientes hasta los más recientes). Calcule el pronóstico de ventas para el próximo día usando:

1. Un promedio móvil simple con 3 periodos.

**Excel**

2. Un promedio móvil ponderado con 3 periodos con pesos 3, 1, y 1 (el valor más alto es para el periodo más reciente).

**Excel**

3. Una Proyección de tendencia.

**Excel**

4. Calcula los errores de cada uno y argumenta cuál es el mejor método según estos resultados.

**Una vez calculado y pronosticado la demanda diaria de un cierto periódico, el mejor método con el menor error porcentual absoluto medio (MAPE) es la proyección de tendencia con un MAPE de 12.65%**

III. (10 puntos). Use el método de suavización exponencial con  $\alpha = 0.5$  y  $\alpha = 0.7$  para pronosticar las ventas de baterías de una compañía para los últimos tres meses. Obtenga los errores cada valor de  $\alpha$  e indique cuál es el valor de  $\alpha$  que produce el mejor pronóstico y **por qué**. Asuma que el pronóstico para enero fue de 20.

Mes	Ventas
Enero	22
Febrero	21
Marzo	15
Abril	19

**El valor de  $\alpha$  que produce el mejor pronóstico es el de 0.5 con un error de 15.09%**

IV. (10 puntos) Dado que la  $DMA = \left( \frac{\sum |y - \hat{y}|}{n} \right)$  para el siguiente pronóstico es 4.0, ¿cuál fue el pronóstico en el periodo 4?

Periodo	Valor real	Pronóstico
1	11	15
2	13	20
3	21	25
4	23	

x	y	Pronóstico	Error de Pronóstico	Valor absoluto del Error de Pronóstico	Error cuadrado de Pronóstico	Error porcentual %	Valor absoluto del Error porcentual
---	---	------------	------------------------	---	---------------------------------------	--------------------------	--

## CRITERIOS DE EVALUACIÓN

Los criterios de evaluación del examen se presentan en la siguiente tabla.

SINGO203B Desarrollo de escenarios. Interpreta interacciones entre variables relevantes en un problema, como base para la construcción de modelos multivariados basados en datos de un fenómeno investigado que le permita reproducir la respuesta del mismo. Es capaz de construir modelos multivariados que expliquen el comportamiento estudiado.

SEGO501A Pensamiento sistémico. Identifica variables inter y transdisciplina que participan en una problemática de su profesión. Identifica la interconexión entre las variables que participan en una problemática en un enfoque inter y transdisciplina. Asocia las condiciones que intervienen en una problemática, manifestando comprensión de las mismas.