**Introducción a la econometría**

1. En un modelo de regresión lineal múltiple para estimar el precio de un departamento en la CDMX se obtiene la siguiente tabla de resultados sobre las variables independientes del modelo.

|  |  |  |  |
| --- | --- | --- | --- |
| **Variable** | **Coeficiente** | **t-score** | **p-value** |
| Superficie habitable (m2) | 70 | 3.2 | 0.001 |
| Número de habitaciones | 50 | 2.5 | 0.02 |
| Proximidad a escuelas (km) | 4 | 0.6 | 0.55 |
| Antigüedad (años) | -13 | -1.8 | 0.07 |

¿Cuál de las variables anteriores es la menos significativa en el modelo?

* 1. Superficie habitable
  2. Número de habitaciones
  3. Proximidad a escuelas
  4. Antigüedad

1. En un modelo de regresión lineal múltiple se aplica el método de regularización Ridge, con el fin de resolver el problema de:
   1. Sobreajuste (overfitting)
   2. Subajuste (underfitting)
   3. Variables con distribuciones muy sesgadas
   4. Ninguna de las anteriores
2. Al analizar una serie de tiempo, se suele separar en diferentes componentes, uno de ellos la tendencia. ¿Cómo podría definirse la tendencia?
   1. Un patrón repetitivo a corto plazo en la serie de tiempo
   2. Un patrón repetitivo a largo plazo en la serie de tiempo
   3. La dirección general (creciente o decreciente) a largo plazo de los datos
   4. La parte que no se puede modelar de los datos (el ruido)
3. Un modelo ARIMA es una de las herramientas más comunes para el análisis de series de tiempo, caracterizado por tres parámetros: p, d y q. Para un caso de ventas mensuales de videojuegos, se obtuvo un modelo ARIMA (p = 3, d = 2, q = 2). ¿Qué implica el valor p = 3 en dicho modelo?
   1. Se utilizarán 3 términos de media móvil para el cálculo actual
   2. Se utilizarán 3 periodos previos para el cálculo de autorregresión
   3. Se realizarán 3 diferenciaciones para hacer la serie estacionaria
   4. La serie de tiempo tiene una estacionalidad (seasonality) de 3 meses
4. ¿Cuál de las siguientes expresiones presenta correctamente todos los componentes de un modelo SARIMA?
   1. SARIMA (p, d, q)
   2. SARIMA (P, D, Q)
   3. SARIMA (p, d, q)
   4. SARIMA (p, d, q) (P, D, Q) (s)

**Diagnóstico para líneas de acción**

1. Para un estudio socioeconómico en México, se analiza la variable de “Ingresos totales del hogar”. Al realizar su histograma se observa que dicha variable tiene un sesgo a la derecha (asimetría positiva). ¿Qué medida representaría mejor los ingresos totales de la mayoría de los hogares?
   1. Media aritmética
   2. Mediana
   3. Desviación estándar
   4. Varianza
2. Para verificar el buen funcionamiento de una máquina, se quiere verificar si la media de edulcolorante en una muestra de 100 refrescos es significativamente diferente a la media de edulcolorante que se ha reportado el último año en los refrescos de esa máquina (40,000 refrescos). ¿Cuál sería la prueba más apropiada para verificar esto?
   1. Prueba de hipótesis de 1 muestra
   2. Prueba de hipótesis de 2 muestras
   3. Análisis de correlación de Pearson
   4. Prueba de Ji Cuadrada
3. Se ha recopilado información sobre el color de las etiquetas de refrescos de cinco sabores distintos: limón, toronja, uva, naranja y fresa. Queremos determinar si hay alguna asociación significativa entre el color la etiqueta y el sabor del refresco. ¿Cuál sería la prueba estadística más apropiada para verificar esto?
   1. Prueba de hipótesis de 1 muestra
   2. Prueba de hipótesis de 2 muestras
   3. Análisis de correlación de Pearson
   4. Prueba de Ji Cuadrada
4. Se aplica un análisis de regresión lineal simple, para predecir las ventas de un automóvil utilizando el gasto en publicidad como predictor. En dicho modelo se obtiene un error estándar muy cercano a cero y un R2 muy cercano a 1. ¿Qué indica esto sobre el modelo?
   1. El modelo no se ajusta a los datos
   2. El modelo se ajusta bien a los datos
   3. El gasto en publicidad no se relaciona con la venta de automóviles
   4. Ninguna de las anteriores
5. Para predecir una variable Y, se utiliza un modelo de regresión lineal simple. Se obtuvo un p-value de 0.002 para la variable independiente X. ¿Cuál de las siguientes afirmaciones es correcta?
   1. La variable X tiene una distribución gaussiana (normal)
   2. El modelo no se ajusta bien a los datos
   3. No existe una relación significativa entre la variable X y Y
   4. Existe una relación significativa entre la variable X y Y