

Departamento de Estadística y Matemáticas
Facultad de Ciencias Económicas
Estadística II
Parcial IV

Nombre: _____ Cédula: _____

1. **(5 puntos)** El acceso al agua potable es esencial para la salud, un derecho humano básico y un componente de una política eficaz de protección de la salud. Esto es importante como cuestión de salud y desarrollo a nivel nacional, regional y local. En algunas regiones, se ha demostrado que las inversiones en abastecimiento de agua y saneamiento pueden generar un beneficio económico neto, ya que las reducciones de los efectos adversos para la salud y los costos de atención médica superan los costos de llevar a cabo las intervenciones.

Por tal motivo se ha decidido recolectar la información de una base de datos que contiene algunas métricas de calidad del agua para 3276 cuerpos de agua diferentes.

El diccionario de variables se presenta a continuación

- **ph:** El pH es un parámetro importante para evaluar el equilibrio ácido-base del agua. También es el indicador de la condición ácida o alcalina del estado del agua. La OMS ha recomendado un límite máximo permitido de pH de 6,5 a 8,5.
- **dureza:** La dureza es causada principalmente por las sales de calcio y magnesio. Estas sales se disuelven en los depósitos geológicos a través de los cuales viaja el agua. El tiempo que el agua está en contacto con el material que produce dureza ayuda a determinar cuánta dureza hay en el agua cruda. La dureza se definió originalmente como la capacidad del agua para precipitar el jabón provocada por el calcio y el magnesio.
- **solidos:** El agua tiene la capacidad de disolver una amplia gama de minerales o sales inorgánicos y algunos orgánicos, como potasio, calcio, sodio, bicarbonatos, cloruros, magnesio, sulfatos, etc. Estos minerales produjeron un sabor no deseado y un color diluido en la apariencia del agua. Este es el parámetro importante para el uso del agua. El agua con un valor alto de solidos indica que el agua está altamente mineralizada. El límite deseable para los Solidos del agua es de 500mg/l y el límite máximo es de 1000mg/l , el cual se prescribe para beber.
- **cloraminas:** El cloro y la cloramina son los principales desinfectantes utilizados en los sistemas públicos de agua. Las cloraminas se forman con mayor frecuencia cuando se agrega amoníaco al cloro para tratar el agua potable. Los niveles de cloro de hasta 4 miligramos por litro (mg/L o 4 partes por millón (ppm)) se consideran seguros en el agua potable.
- **sulfato:** Los sulfatos son sustancias naturales que se encuentran en minerales, suelos y rocas. Están presentes en el aire ambiente, el agua subterránea, las plantas y los alimentos. El principal uso comercial del sulfato es en la industria química. La concentración de sulfato en el agua de mar es de aproximadamente 2700 miligramos por litro (mg/L). Varía de 3 a 30mg/L en la mayoría de los suministros de agua dulce, aunque en algunas ubicaciones geográficas se encuentran concentraciones mucho más altas (1000mg/L).
- **conductividad:** El agua pura no es un buen conductor de la corriente eléctrica, sino un buen aislante. El aumento de la concentración de iones mejora la conductividad eléctrica del agua. Generalmente, la cantidad de sólidos disueltos en el agua determina la conductividad eléctrica. La conductividad eléctrica (CE) en realidad mide el proceso iónico de una solución que le permite transmitir corriente. Según los estándares de la OMS, el valor de CE no debe exceder los $400\mu\text{S/cm}$.
- **carbono_organico:** El carbono orgánico total (COT) en las fuentes de agua proviene de materia orgánica natural (NOM) en descomposición, así como de fuentes sintéticas.

COT es una medida de la cantidad total de carbono en compuestos orgánicos en agua pura. Según la EPA de EE. UU., $< 2mg/L$ como *COT* en agua tratada/potable y $< 4mg/L$ en agua de origen que se utiliza para tratamiento.

- **trihalometanos:** Los *THM* son sustancias químicas que se pueden encontrar en el agua tratada con cloro. La concentración de THM en el agua potable varía según el nivel de materia orgánica en el agua, la cantidad de cloro necesaria para tratar el agua y la temperatura del agua que se está tratando. Se consideran seguros niveles de THM de hasta $80ppm$ en el agua potable.
- **turbidez:** La turbiedad del agua depende de la cantidad de materia sólida presente en estado suspendido. Es una medida de las propiedades emisoras de luz del agua y la prueba se utiliza para indicar la calidad de la descarga de desechos con respecto a la materia coloidal. El valor medio de turbidez obtenido para Wondo Genet Campus ($0,98NTU$) es inferior al valor recomendado por la OMS de $5,00NTU$.
- **potabilidad:** Indica el porcentaje de potabilidad del agua, es decir, la probabilidad de que sea segura para el consumo humano donde entre más cercana de 1 habrá más probabilidad de que sea Potable y entre más cercana de 0 tendrá menos probabilidad de ser Potable.

- a) **(0.5 puntos)** Si tuviera que plantear una relación lineal entre el porcentaje de potabilidad del agua y otra variable de la forma

$$Y = \beta_0 + \beta_1 X + \varepsilon$$

cómo plantearía dicha relación desde sus racionalidad y conocimientos en Economía?, es decir, cuál variable escogería usted como variable explicativa X , y por qué?

- b) **(0.5 puntos)** Basado en el planteamiento que realizó en el inciso anterior, justifique cómo esperaría usted que fuese el signo de los parámetros β_0 y β_1 del modelo, desde un punto de vista económico.
- c) **(0.5 puntos)** Realice el cálculo de los estimadores para los parámetros β_0 y β_1 , e interprete éstos en el contexto de los datos. Los resultados obtenidos fueron consistentes con lo que esperaba en el inciso anterior?
- d) **(0.5 puntos)** Pruebe la significancia estadística de los parámetros β_0 y β_1 , empleando para ello un nivel de significancia del 95 %, e interprete en el contexto de los datos.
- e) **(0.5 puntos)** Construya intervalos de confianza para los parámetros β_0 y β_1 , empleando para ello un nivel de confianza del 95 %, e interprete en el contexto de los datos.
- f) **(1 puntos)** Verifique si se cumplen o no los supuestos del modelo de regresión lineal planteado, e interprete los resultados.
- g) **(0.5 puntos)** Pruebe la significancia estadística de la regresión lineal planteada, empleando para ello un nivel de significancia del 5 %, e interprete en el contexto de los datos.
- h) **(0.5 puntos)** Realice el cálculo del coeficiente de determinación $R^2 = \frac{SCR}{SCT}$ asociado a la regresión lineal planteada, e interprete el resultado obtenido. Dicho resultado es consistente con lo que se concluyó en el inciso anterior?
- i) **(0.5 puntos)** Seleccione tres valores para x_0 entre los posibles que considera que puede tomar la variable que escogió como X , y con éste, construya un intervalo de predicción del 95 % para la variable que escogió como Y e interprete en el contexto de los datos.