



Evaluación Formativa

Taller de métodos cuantitativos FMSA315

Primer Trimestre 2025

1 Instrucciones

- Cuenta con 180 minutos para realizar la actividad.
- Debe cargar un archivo PDF (incluir nombre y RUT en el documento), cualquier otro tipo de documento no será recibido y será calificado con nota mínima, sin apelación.
- Crear un breve informe con la solución de cada pregunta, comentando los principales resultados.
- No se asignarán puntos a respuestas sin justificación.
- Cualquier situación o problema en el transcurso de la evaluación debe ser informado de manera inmediata a su docente. Cualquier justificación una vez terminada la evaluación no será considerada.

2 Contextualización

El dataset **winequality-red.csv** contiene varias propiedades químicas del vino, como la acidez, el azúcar residual, los cloruros y los niveles de dióxido de azufre y calificaciones de calidad. A menudo se utiliza para predecir la calidad del vino en función de estas propiedades químicas.

Las columnas información relativa a este dataset, es la siguiente:

1. fixed acidity : concentración de todos los ácidos presentes en el vino, que no son volátiles (*gr/lt*)
2. volatile acidity : concentración de ácidos volátiles en *gr/lt*
3. citrid acid: concentración de ácido cítrico en *gr/lt*
4. residual sugar: concentración de azúcar residual en *gr/lt*
5. chlorides: Concentración de cloruros en *gr/lt*.
6. total sulfur dioxide: suma total de sulfuros en *gr/lt*
7. density: Densidad del vino, generalmente se mide en *gr/ml*
8. pH: Medida de acidez
9. sulphates: Compuestos añadidos para preservar el vino, y evitar la oxidación *gr/lt*
10. alcohol : Contenido de alcohol
11. quality: Valor numérico que representa la calidad del vino, determinada mediante pruebas de cata.

3 Pruebas de Hipótesis

Preguntas:

1. Se desea comprobar si el promedio de acidez fija (fixed acidity) es mayor en los vinos Cabernet Sauvignon que en los vinos Carmenere. Formule una prueba de hipótesis y estudie esta afirmación.
2. Compare el promedio de alcohol entre vinos de calidad superior (calidad ≥ 7) y vinos de calidad media (calidad 5 – 6). ¿Existe evidencia suficiente para afirmar que los vinos de calidad superior tienen mayor contenido de alcohol?
3. ¿La proporción de vinos Merlot con un pH mayor a 3.1 es significativamente distinta a 0.6? Comente que supuestos son necesarios para realizar esta prueba de hipótesis.

4 Regresión Lineal

El objetivo de esta parte de la evaluación es desarrollar un modelo de regresión lineal múltiple que permita predecir la calidad del vino blanco a partir de un conjunto de variables medibles en laboratorio. La calidad del vino se evalúa mediante una escala numérica, mientras que las propiedades físico-químicas se miden a partir de diferentes parámetros propios de la industria.

1. Elaborar un modelo de regresión lineal múltiple, en donde la variable regresada sea la calidad del vino y las variables regresoras sean la acidez volátil, la densidad, la concentración de azúcar, el contenido del alcohol del vino, el pH y la cantidad de sulfitos que son agregadas al vino para su preservación.
2. Determinar e interpretar el coeficiente de determinación R^2
3. **Análisis de los supuestos:**
 - Determinar si los residuos tienen varianza constante (test de homocedasticidad).
 - Determinar si los residuos siguen una distribución normal y complementar con un histograma de residuos.
 - Determinar si los residuos están correlacionados
4. Utilizando la ecuación del modelo, predecir la calidad del vino para las siguientes condiciones:
 - pH : 3.27, 4.0, 2.8
 - volatile acidity: 0.45, 0.37, 0.5
 - residual sugar: 0.91, 1.21, 2.01
 - density : 0.995, 0.999, 0.912
 - alcohol: 9.4, 11.01, 10.9
 - sulphates: 0.5, 0.54, 0.59

¿Qué se puede inferir con respecto a las estimaciones? ¿Son confiables?

Para todos sus cálculos considere $\alpha = 5\%$