

## Tarea 1

1. Los datos `ChemReact.csv` contiene información de un experimento químico. Se tiene como objetivo determinar las condiciones de temperatura ( $x_1$ ) y tiempo ( $x_2$ ) que proporcionan un alto rendimiento ( $y$ ). Para ello, se propone un modelo cuadrático completo:

$$y_i = \beta_0 + x_{1i}\beta_1 + x_{2i}\beta_2 + x_{1i}^2\beta_3 + x_{2i}^2\beta_4 + x_{1i}x_{2i}\beta_5 + \varepsilon_i,$$

donde  $\varepsilon_i \sim N(0, \sigma^2)$  y  $cov(\varepsilon_i, \varepsilon_k) = 0$ .

- a) Ajuste el modelo y verifique el cumplimiento de los supuestos del modelo. En caso que no se cumplan los supuestos, realice transformaciones sobre la variable respuesta para corregirlo.
  - b) ¿Cómo afectan las covariables a la variable respuesta? (esto lo puede hacer gráficamente y/o interpretando los coeficientes).
  - c) Evalúe si el efecto interacción es significativo. ¿Qué implicaciones tiene esto en la relación de las covariables con la variable respuesta?
  - d) Encuentre la combinación de temperatura y tiempo que maximiza el rendimiento. Calcule e interprete un intervalo de predicción del rendimiento en este punto.
2. Considere los datos `oldfaith` de la librería `alr4` sobre las erupciones del geyser Old Faithful durante octubre de 1980. Las variables observada son: la duración en segundos de la erupción actual ( $x$ ), y el intervalo de tiempo en minutos hasta la próxima erupción ( $y$ ).

La relación entre las variables es aproximadamente lineal, sin embargo, se cree que hay un cambio de pendiente. Por esta razón se propone el siguiente modelo:

$$y_i = \beta_0 + x_i\beta_1 + (x_i - 180)_+\beta_2 + \varepsilon_i,$$

donde:

$$(x_i - 180)_+ = \begin{cases} 0 & \text{si } x_i \leq t, \\ (x_i - 180) & \text{si } x_i > t. \end{cases}$$

- a) Ajuste el modelo e interprete los coeficientes.
- b) ¿el cambio de pendiente es significativo?
- c) compare el modelo anterior con un modelo lineal simple. ¿cual proporciona mejor ajuste?

Para la entrega, tenga en cuenta lo siguiente:

- El reporte no debe exceder 6 páginas (texto en una sola columna). No incluir códigos o salidas de R (cree sus propias tablas).
- En el reporte, todas las tablas y figuras que incluyan deben estar enumeradas y referenciadas en texto (por ejemplo, “como se observa en la Figura 1, ...” o “la Tabla 2 presenta ...”). Solo incluya tablas y figura que sean relevantes (es decir, todas las tablas y figuras debe tener una referencia en el texto).
- Fecha de entrega: XX de marzo de 2023 al empezar la clase (en físico y a través del campus virtual).