

## Práctica Nº8

R: practica\_stepwise.R

**1)** Un productor está interesado en invertir más de su capital en la producción de soja. Sin embargo, le gustaría tener un modelo que le permita conocer el rendimiento de este cultivo para lo cual recurre a un Meteorólogo. Este profesional le solicita la mayor cantidad de datos posibles que disponga para poder estudiar el rendimiento de este cultivo. El productor le ofrece datos de Rendimiento de la soja(qq/ha) en los meses de enero, febrero y marzo, y datos sobre la cantidad de fertilizante de fósforo (kg) y de azufre (kg) que utiliza en su campo. Los registros de estas variables comprenden los meses de enero, febrero y marzo del período 1955-2018.

El meteorólogo considerando que el rendimiento no depende solo del fertilizante empleado, recurre al SMN para solicitar datos temperatura (°C) y precipitación acumulada mensual (mm) de estos meses en la región del sur de la provincia de Córdoba, donde se ubica el campo en cuestión.

**a)** Estudie el comportamiento de cada una de las variables predictoras con respecto a la variable a predecir para la estación de verano. ¿Qué análisis simple puede realizar? Si tuviese que proponer un modelo tridimensional, ¿qué variables consideraría? Puede basarse en aproximaciones físicas como también estadísticas.

**b)** Grafique el rendimiento observado en conjunto con el estimado. Analice sus resultados.

**2)** De acuerdo con el problema anterior, las aproximaciones previas que usted consideró no convencieron en forma completa al productor y le pidió nuevamente que realice un análisis más exhaustivo. Teniendo en cuenta los siguientes puntos que a continuación se enumeran.

**a)** ¿Qué variables cree usted que representan el mejor modelo posible para evaluar el rendimiento de la soja en el sur de Córdoba en los meses de verano? Justifique.

**b)** Determine el modelo seleccionado. Justifique.

**c)** ¿Si el meteorólogo solo hubiese considerado los datos del productor, hubiese obtenido un buen modelo? Justifique.

**d)** Compare el presente análisis con el del punto 1).

**3)** Se desea realizar un estudio agro-climático cuya finalidad es estimar la mejor relación de la variable rendimiento (sin tendencia) de soja a partir de variables meteorológicas, en la localidad Marcos Juárez. Para ello, se utiliza el método Stepwise y se consideran las siguientes variables meteorológicas: precipitación mensual y temperaturas máximas y mínimas mensuales dentro del ciclo promedio del cultivo (Noviembre a Mayo del año siguiente).

Luego de un estudio inicial de selección de las variables se dispone de la información mencionada en los siguientes meses en el período (1974-2000):

ARCHIVO: M JUÁREZ SIN T SOJA\_CLIMA.STA

- Totales mensuales de precipitación para los meses de diciembre, enero, febrero y abril.
- Temperatura máxima de los meses de diciembre, enero y abril.
- Temperatura mínima de los meses diciembre, enero y abril.

Para llevar a cabo el estudio, determinar:

- ¿Qué variables están correlacionadas (nivel de significancia = 5%)?
- A partir de las 10 variables de entrada calcular la mejor regresión (rendimiento sin tendencia = función (variables meteorológicas)). ¿Qué variables participan? Justificar su participación. ¿Qué porcentaje de la varianza del rendimiento es explicada por esta regresión?
- Graficar rendimiento real y estimado en función del tiempo. Eliminar del estudio la temperatura máxima de todos los meses y repetir los puntos b) y c). A partir de los resultados obtenidos en los 2 ajustes, ¿cuál resultó ser el mejor ajuste?
- Realizar una regresión lineal múltiple estándar entrando las 10 variables predictoras de entrada. Analizar los resultados y compararlos con los obtenidos con el método forward stepwise.