

Ejemplos de Aplicación - Semana 2

Problema de costo mano de obra

En una gran carpintería se registraron los costos (hora-hombre) en función de la superficie de las piezas en m^2 .

Datos: *Ejemplo_Aplicacion_Semana_2.xlsx*

Cargamos las librerías

```
# Para importar datos de excel
library(readxl)
# Para graficos más profesionales
library(ggplot2)
```

Cargamos los datos del archivo

```
# Importamos los datos desde Excel
# Al archivo debemos agregarle la ruta completa
# Por ejemplo: datos <- read_excel("C:/micarpetalocal/Clase_1.xlsx")
datos <- read_excel("Ejemplo_Aplicacion_Semana_2.xlsx")

# Cambiamos el nombre a las columnas/variables para que sean más sencillos
colnames(datos) <- c("superficie", "costo")
```

a) Haga una evaluación visual de los datos. Genere un gráfico básico sin incluir el origen de coordenadas. ¿Le parece razonable el asumir que el intercepto del modelo es nulo? Realice un nuevo gráfico incluyendo el origen de coordenadas. ¿Le parece razonable el asumir que el intercepto del modelo es nulo?

Generamos el gráfico básico

```
# Gráfico de dispersion entre la variable explicativa y la de respuesta (mas profesional)
ggplot(datos, aes(x = superficie, y = costo)) +
  geom_point()
```

Generamos el gráfico con el origen de coordenadas

```
# Gráfico de dispersion entre la variable explicativa y la de respuesta (mas profesional)
ggplot(datos, aes(x = superficie, y = costo)) +
  geom_point() +
  # fijamos los límites inferiores de los ejes en 0.
  # NA = Not available. Dejamos que R decida la mejor opción
  xlim(0, NA)+
  ylim(0, NA)+
  # Establecemos tema bw (estilo de gráfico mas sobrio)
  theme_bw() +
  # Agregamos en gris ejes cartesianos
  geom_vline(xintercept = 0, color = "grey")+
  geom_hline(yintercept = 0, color = "grey")
```

b) Investigue la asociación lineal entre las variables, considerando que no hay costo fijo. Valide el modelo. Grafique el modelo junto con los datos obtenidos. ¿Utilizaría este modelo?

Generamos el modelo con intercepto nulo

```
# Definimos el modelo de regresión con intercepto nulo
# observe el "-1" en la fórmula. Este -1 le indica el R que el modelo debe estimarse sin i
modelo <- lm(formula = costo ~ superficie - 1, data = datos)
modelo
summary(modelo)
```

c) Proponga un nuevo modelo, considerando que si hay costo fijo. Valide el modelo. Grafique el modelo junto con los datos obtenidos. Grafique el modelo junto con los datos obtenidos. Comparado con el modelo anterior ¿cuál considera mejor?

Proponemos un nuevo modelo donde la ordenada al origen estimada por el modelo

```
modelo2 <- lm(formula = costo ~ superficie, data = datos)
modelo2
summary(modelo2)
```

d) A raíz de lo analizado, consultó con el gerente de costos y éste le informó que para los presupuestos se utiliza un costo estándar de 0.66 pesos/ m^2 . ¿Considera que el modelo es consistente con este costo estándar? Utilice un riesgo del 10% para responder.

f) Suponiendo que fuera válido el modelo lineal y que una pieza de 1,2 m^2 tuvo un costo de 3,4, ¿aceptaría Ud. que este valor es consistente este valor con su modelo? Asuma un riesgo del 5%. Realice el análisis con el modelo con intercepto nulo y con intercepto estimado. Compare y dicuta nuevmanete que modelo se comporta mejor.

```
# Armamos un nuevo data.frame con los valores que necesito predecir
# El costo lo dejamos NA
datos_nuevos =
  data.frame(superficie = 1.2,
             costo = NA)

predict(modelo2,
        newdata = datos_nuevos,
        interval = "predict",
        level = 0.95)
```

g) Asuma que le informa el gerente de costos que un costo fijo de 0,7 \$. Estime el modelo asumiendo que este costo es correcto, valide el modelo, grafique y seleccione cual de los tres modelos es mas adecuado.

```
# Restamos 0,70 a la variable Y para estimar el modelo
modelo_costo_fijo <- lm(formula = costo - 0.7 ~ superficie - 1, data = datos)
modelo_costo_fijo
summary(modelo_costo_fijo)
```

Graficamos modelo SIN costo fijo

```
ggplot(datos, aes(x = superficie, y = costo)) +  
  geom_point() +  
  geom_smooth(formula = y ~ x , method=lm , color="red", se=FALSE)+  
  geom_smooth(formula = y ~ x - 1, method=lm , color="blue", se=FALSE)+  
  geom_smooth(formula = y - 0.7 ~ x - 1, method=lm , color="green", se=FALSE)+  
  # fijamos los límites inferiores de los ejes en 0.  
  # NA = Not available. Dejamos que R decida la mejor opción  
  xlim(0, NA)+  
  ylim(0, NA)+  
  # Establecemos tema bw (estilo de gráfico mas sobrio)  
  theme_bw() +  
  # Agregamos en gris ejes cartesianos  
  geom_vline(xintercept = 0, color = "grey")+  
  geom_hline(yintercept = 0, color = "grey")
```