

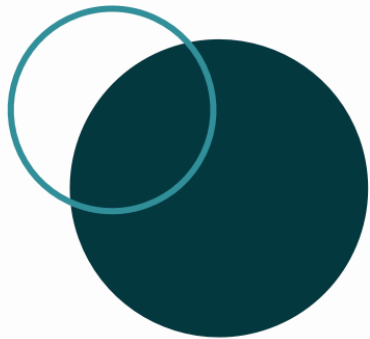


Sesión 2

Regresión Lineal Múltiple - Supuestos



AGENDA DEL DÍA DE HOY



Repaso del Modelo de Regresión Lineal
Supuestos

ECUACIÓN

$$Y = B0 + B1 * X1 + e$$

Y = Variable dependiente

X = Variable independiente

B0 = Intercepto

B1 = Pendiente de la variable

e = error residual

PASOS PARA EL ANÁLISIS



¿El modelo es
válido?

¿Las VI aportan al modelo?

Valor explicativo

Ecuación de la recta

Validación de supuestos

Supuestos	¿Qué se espera?	¿Qué ocurre si no se cumple?
Linealidad	Que la relación entre la variable D y las variables I sea lineal. No encontrar un patrón en los errores.	La interpretación del modelo sería limitada, se podría optar por modificarlo.
Normalidad de residuos	Que los residuos del modelo sigan una distribución normal	Influye en los estimadores de mínimos cuadrados
Homocedasticidad	Que la varianza de los residuos sea constante y homogénea	Método de mínimos cuadrados no es eficiente
Ausencia de multicolinealidad	Que las variables independientes no estén relacionadas entre sí	El modelo será inestable. Los coeficientes oscilarán ante cambios mínimos.
Independencia de residuos	Que los residuos no estén correlacionados	Puede influenciar en el cálculo de los estimadores de R2

Supuestos	¿Cómo se verifica?	Comando	¿Qué revisamos en el gráfico?	Hipótesis
Linealidad	Gráficos Prueba de correlación	Op1: library(ggfortify) plot(modelo,1) Op2: cor.test(d\$v1, d\$v2)	La línea roja debe de estar lo más cerca posible a la punteada.	H0: No hay correlación/relación lineal
Normalidad de residuos	Gráficos Pruebas de normalidad (shapiro, kolmogorov, lillie)	Op1: plot(modelo,2) Op2: shapiro.test(modelo\$resid)	Todos los puntos deben de estar sobre la diagonal	<u>H0: La distribución es normal</u>
Homocedasticidad	Gráficos Test Breusch Pagan	Op1: plot(modelo,3) Op2: library(lmtest) bptest(modelo1)	La linea roja debe seguir una tendencia horizontal.	H0: Hay homocedasticidad/modelo es homocedástico
Ausencia de multicolinealidad	Factor de Inflación de Varianza	library(car) vif(modelo)	-	Si el VIF es menor a 5 no hay multicolinealidad
Independencia de residuos	Prueba de Durbin Watson	durbinWatsonTest(modelo)	-	H0: Los residuos son independientes