

Clase: Regresión Múltiple

Orlando Joaqui Barandica
orlando.joaqui@javerianacali.edu.co

Pontificia Universidad Javeriana Cali

2023

Tabla de Contenido

1 Transformaciones

2 Modelos polinómicos

- Modelo cuadrático

El análisis de regresión describe la relación entre dos variables; un requisito es que esta relación sea lineal, y lo mismo aplica para el coeficiente de correlación, el cual mide la fuerza de una **relación lineal** entre dos variables.

El análisis de regresión describe la relación entre dos variables; un requisito es que esta relación sea lineal, y lo mismo aplica para el coeficiente de correlación, el cual mide la fuerza de una **relación lineal** entre dos variables.

¿Y que pasa si la relación no es lineal?

El análisis de regresión describe la relación entre dos variables; un requisito es que esta relación sea lineal, y lo mismo aplica para el coeficiente de correlación, el cual mide la fuerza de una **relación lineal** entre dos variables.

¿Y que pasa si la relación no es lineal?

La solución está en:

- Volver a hacer la escala de una o ambas variables para que la nueva relación sea lineal; por ejemplo, en vez de utilizar los valores actuales de la variable dependiente Y , se crearía otra variable dependiente calculando el $\text{Log}(Y)$, a este calculo se le llama transformación.
- Otras transformaciones conocidas son: Raíz cuadrada, calcular el recíproco, evaluar al cuadrado la variable.

Recuerde:

Dos variables pueden estar estrechamente relacionadas aunque su relación no sea lineal. Tenga cuidado cuando interprete el coef. de correlación o una ecuación de regresión. Estos estadísticos pueden indicar que no existe una relación lineal, pero quizá haya una relación de alguna otra forma no lineal.

Recuerde:

Dos variables pueden estar estrechamente relacionadas aunque su relación no sea lineal. Tenga cuidado cuando interprete el coef. de correlación o una ecuación de regresión. Estos estadísticos pueden indicar que no existe una relación lineal, pero quizá haya una relación de alguna otra forma no lineal.

Las transformaciones matemáticas también se aplican a casos donde el modelo de regresión lineal no cumple alguno de los supuestos, tal es el caso de la normalidad en los residuos, realizando la transformación apropiada se logra corregir esta situación. Incluso problemas de heterocedasticidad se pueden corregir.

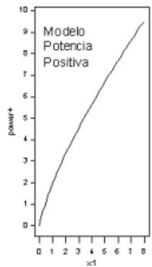
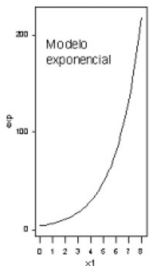
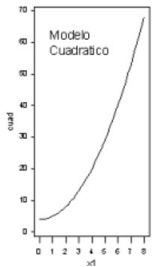
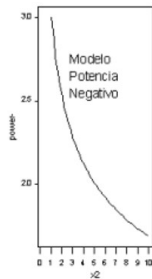
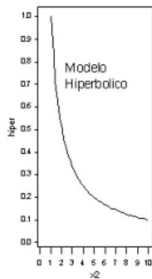
Procedimiento:

- Identificar la ecuación matemática apropiada (ecuación del modelo)
- Linealizar la ecuación matemática (modelo linealizado)
- Adecuar las variables al modelo linealizado
- Con las variables transformadas se debe construir un nuevo modelo de regresión lineal simple estimando sus coeficientes, se le realiza la prueba de linealidad y se calcula el coeficiente de determinación R^2

Transformaciones

Nombre del modelo	Ecuación del Modelo	Transformación	Modelo Linealizado
Exponencial	$Y = \alpha e^{\beta X}$	$Z = \log Y \quad X = X$	$Z = \log \alpha + \beta X$
Logaritmico	$Y = \alpha + \beta \log X$	$Y = Y \quad W = \log X$	$Y = \alpha + \beta W$
Doblemente Logaritmico o Potencia	$Y = \alpha X^{\beta}$	$Z = \log Y \quad W = \log X$	$Z = \log \alpha + \beta W$
Hiperbólico	$Y = \alpha + \beta/X$	$Y = Y \quad W = 1/X$	$Y = \alpha + \beta W$
Doblemente Inverso	$Y = 1/(\alpha + \beta X)$	$Z = 1/Y \quad X = X$	$Z = \alpha + \beta X$

Transformaciones



El director corporativo de marketing de GroceryLand Supermarkets desea estudiar el efecto del precio en las ventas semanales de las botellas de dos litros de su marca de refresco. Junto al director de ventas consideran razonable aumentar el precio de la botella de dos litros desde 0.5 hasta 2.00 dólares.

El director corporativo de marketing de GroceryLand Supermarkets desea estudiar el efecto del precio en las ventas semanales de las botellas de dos litros de su marca de refresco. Junto al director de ventas consideran razonable aumentar el precio de la botella de dos litros desde 0.5 hasta 2.00 dólares.

El director selecciona una muestra aleatoria de 20 tiendas y asigna al azar un precio de venta entre 0.50 dólares y 2.00 dólares en cada una. Al final de la semana los resultados de las ventas de botellas son reportados.

¿Existe alguna relación entre el precio de venta y las ventas semanales?

Tabla: Datos de ventas y precio de GroceryLand

Tienda	Precio (US)	Ventas	Tienda	Precio (US)	Ventas
17	0,5	181	30	0,76	91
121	1,35	33	127	1,79	13
227	0,79	91	266	1,57	22
135	1,71	13	117	1,27	34
6	1,38	34	132	0,96	74
282	1,22	47	120	0,52	164
172	1,03	73	272	0,64	129
196	1,84	11	120	1,05	55
143	1,73	15	194	0,72	107
66	1,62	20	105	0,75	119

En un modelo de regresión lineal simple se supone apropiada una relación considerando solo a X en su expresión básica, pero existen relaciones de orden superior tales como X^2 que pueden mejorar el ajuste de los datos. Un modelo expresado en términos de X^p es denominado modelo de polinómico de orden p .

$$Y = \beta_0 + \beta_1 X + \beta_2 X^2 + \beta_3 X^3 + \dots + \beta_p X^p + \varepsilon \quad (1)$$

En un modelo de regresión lineal simple se supone apropiada una relación considerando solo a X en su expresión básica, pero existen relaciones de orden superior tales como X^2 que pueden mejorar el ajuste de los datos. Un modelo expresado en términos de X^p es denominado modelo de polinómico de orden p .

$$Y = \beta_0 + \beta_1 X + \beta_2 X^2 + \beta_3 X^3 + \dots + \beta_p X^p + \varepsilon \quad (1)$$

En la práctica modelos de orden superior a tres ($p > 3$) son muy raros por lo que solo se consideraran en este documento los modelos cuadráticos ($p = 2$) y cúbicos ($p = 3$)

El modelo cuadrático o parabólico

$$Y = \beta_0 + \beta_1 X + \beta_2 X^2 + \varepsilon \quad (2)$$

El modelo cuadrático o parabólico

$$Y = \beta_0 + \beta_1 X + \beta_2 X^2 + \varepsilon \quad (2)$$

El análisis para una situación dada es igual al que se hace para la regresión lineal simple:

- Gráfica de dispersión
- Significancia general del modelo (tabla anova)
- Prueba de la significancia del termino cuadrático
- Interpretación de cada uno de los coeficientes del modelo.

Cuando se introducen en un modelo de regresión términos como X , X^2 , etc. Se corre el riesgo de introducir problemas de multicolinealidad, problema que se analizará en el capítulo de regresión lineal múltiple.

Cuando se introducen en un modelo de regresión términos como X , X^2 , etc. Se corre el riesgo de introducir problemas de multicolinealidad, problema que se analizará en el capítulo de regresión lineal múltiple.

Para estimar los coeficientes del modelo cuadrático, se recurre al método de los mínimos cuadrados a través del programa SPSS mediante la opción *Estimación Curvilínea*.

Modelo cuadrático

El departamento de mercadeo de una gran cadena de supermercados quiere estudiar el efecto de los Precios sobre las Ventas de paquetes de máquinas de afeitar desechables. Se formaron paquetes de tres cantidades diferentes con precios de \$7,900, \$9,900 y \$11,900. A continuación se muestran los resultados:

Precios $\times 100\$$	No. Paquetes Vendidos
79	142
79	151
79	163
79	168
79	176
99	91
99	100
99	107
99	115
99	126
119	77
119	86
119	95
119	100
119	106

Suponga que hay una relación cuadrática entre precio y ventas.

- Elabore un diagrama de dispersión para precio y ventas.
- Estime la ecuación de regresión cuadrática.
- Elabore un pronóstico de las ventas semanales para un precio de \$ 7,900.
- Efectúe un análisis residual de los resultados y determine la idoneidad del modelo.
- Con un nivel de significancia de 0.05. ¿Existe una relación cuadrática significativa entre ventas y precio?
- Calcule e interprete el R^2 .