

Regresion Multivariante

Veremos un supuesto Basico de analisis de datos predictivo basado en Regresiones con multiples variables (lineal). Buscaremos que se cumplan las condiciones de los modelos de regresion que son: Normalidad, Linealidad y Homocedasticidad (Igualdad de las Varianzas) e Independencia de los Residuos.

Una empresa de Bingo desea conocer la ecuación que le pueda permitir predecir los ingresos en función de los gastos de publicidad en radio, publicidad en television y la inversión en publicidad en los periódicos . Se realiza un estudio en el que se reúnen los datos mensuales correspondientes a los últimos 20 meses . Estos datos se muestran en la siguiente tabla (todas las cantidades expresadas en miles de €):

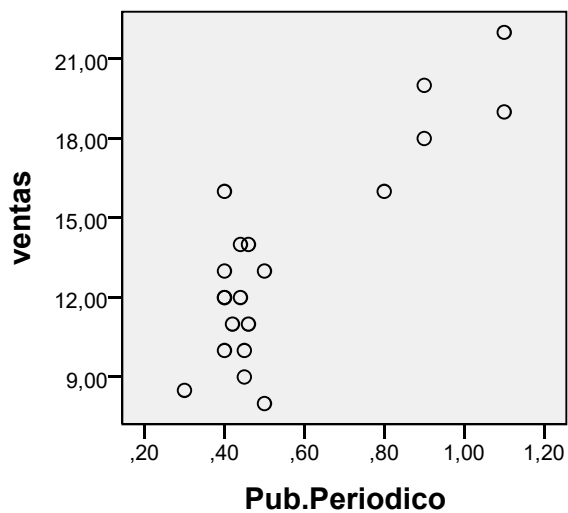
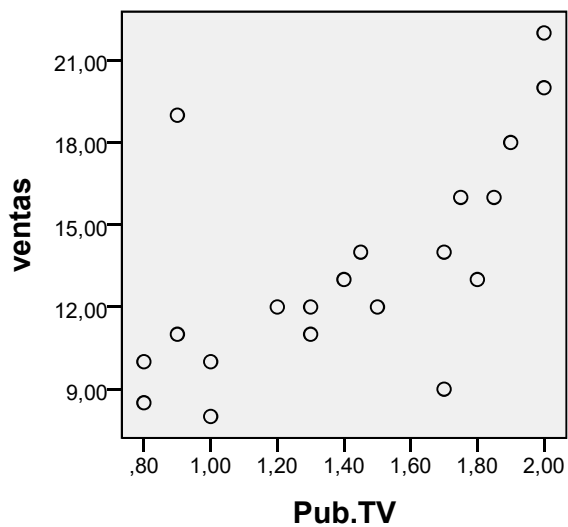
Ventas	Pub. en tv	Pub. en radio	Pub. en per.
10,0	1,30	56	,40
12,0	1,40	55	,40
11,0	1,50	60	,42
13,0	1,70	65	,50
12,0	1,75	69	,40
14,0	1,30	67	,44
16,0	1,45	68	,40
12,0	,90	67	,44
14,0	,80	97	,46
11,0	,90	66	,46
10,0	,80	65	,45
19,0	1,00	60	1,10
8,5	1,70	70	,30
8,0	1,80	110	,50
9,0	1,85	75	,45
13,0	1,90	80	,40
16,0	2,00	85	,80
18,0	2,00	90	,90
20,0	1,30	56	,90
22,0	1,40	55	1,10

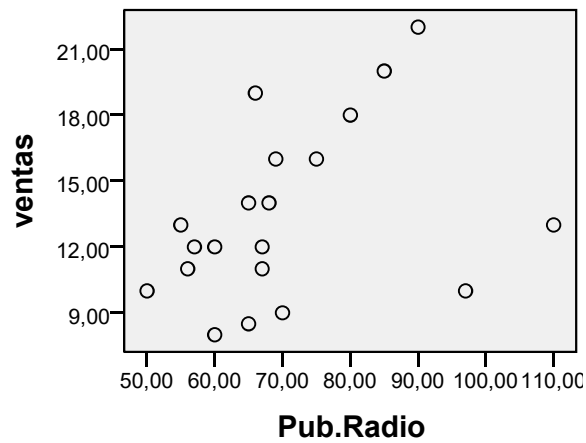
1) La variable dependiente o resultado (Y) son los ingresos/ventas y las variables independientes/explicativas del modelo son las diferentes inversiones en publicidad correspondiendo con X_1 (publi en tv), x_2 (Publi radio) y x_3 (publi periodico). Lo primero que realizaremos sera un diagrama de dispersion para observar a groso modo si hay algun tipo de correlacion entre la var

iable dependiente y cada una de las variables independientes La expresion del modelo multivariable propuesto es del tipo $Y = B_0 + B_1x_1 + B_2x_2 + B_3x_3 + E(\text{error})$

A simple vista podemos observar una mayor correlacion Cuando aumentan las ventas y cuando aumenta la inversion en publi Tv y Publi Radio y menos correlación entre aumento de ventas y aumento de inversion en Publi Periodico.

Gráfico





2) Ejecutaremos el modelo multivariable. Si nos fijamos en la Tabla ANOVA, el p-valor asociado al contraste es menor que 0.001, (Sig = 0.000), por lo que rechazamos la hipótesis nula de no linealidad entre las variables. Esto implica que al menos una de las variables independientes contribuye de forma significativa a la explicación de la variable respuesta. Si nos fijamos

Regresión

Estadísticos descriptivos

	Media	Desviación estándar	N
ventas	13,4250	3,95792	20
Pub.TV	1,4125	,41323	20
Pub.Radio	70,6000	15,13240	20
Pub.Periodico	,5610	,24787	20

Correlaciones

		ventas	Pub.TV	Pub.Radio	Pub.Periodico
Correlación de Pearson	ventas	1,000	,632	,371	,849
	Pub.TV	,632	1,000	,402	,382
	Pub.Radio	,371	,402	1,000	,312
	Pub.Periodico	,849	,382	,312	1,000
Sig. (unilateral)	ventas	.	,001	,054	,000
	Pub.TV	,001	.	,039	,048
	Pub.Radio	,054	,039	.	,090
	Pub.Periodico	,000	,048	,090	.
N	ventas	20	20	20	20
	Pub.TV	20	20	20	20
	Pub.Radio	20	20	20	20
	Pub.Periodico	20	20	20	20

Variables entradas/eliminadas^a

Modelo	Variables entradas	Variables eliminadas	Método
1	Pub. Periodico, Pub.Radio, Pub.TV ^b	.	Entrar

a. Variable dependiente: ventas

b. Todas las variables solicitadas introducidas.

Resumen del modelo^b

Modelo	R	R cuadrado	R cuadrado ajustado	Error estándar de la estimación	Durbin-Watson
1	,912 ^a	,832	,801	1,76707	1,225

a. Predictores: (Constante), Pub.Periodico, Pub.Radio, Pub.TV

b. Variable dependiente: ventas

ANOVA^a

Modelo		Suma de cuadrados	gl	Media cuadrática	F	Sig.
1	Regresión	247,677	3	82,559	26,440	,000 ^b
	Residuo	49,960	16	3,123		
	Total	297,638	19			

a. Variable dependiente: ventas

b. Predictores: (Constante), Pub.Periodico, Pub.Radio, Pub.TV

Coeficientes^a

Modelo	Coeficientes no estandarizados		Coeficientes estandarizados	t	Sig.
	B	Error estándar	Beta		
1 (Constante)	2,108	2,055		1,026	,320
Pub.TV	3,432	1,121	,358	3,060	,007
Pub.Radio	,001	,030	,006	,050	,961
Pub.Periodico	11,347	1,802	,711	6,298	,000

a. Variable dependiente: ventas

4) Nos fijamos en la tabla Coeficientes: La recta de regresión sería la siguiente: $2,108 = 3,432 x_1 + 0.001x_2 + 11,347x_3$ Las ventas estimadas son iguales a 2108 euros si no se produce inversión en publicidad (ni en televisión, ni en radio ni en periódicos).

Por cada mil euros invertidos en publicidad en televisión las ventas esperadas aumentan en **3432 euros**, supuesto que permanecen constantes las otras variables.

Por cada mil euros invertidos en publicidad en radio, las ventas estimadas aumentan únicamente en **1 euro**, suponiendo que se mantienen constantes las otras variables independientes.

Por cada mil euros invertidos en publicidad en periódicos se produce un incremento en las ventas esperadas de **11347 euros**, supuestas constantes las restantes variables predictivas. Para la Variable Publi_Radio su nivel de significacion es 0.961 > 0.05 del nivel de significacion por tanto se acepta la hipotesis nula de no linealidad entre pub.Radio y las ventas, por tanto esta variable explicativa se puede excluir de lmodelo propuesto.

A la vista de estos resultados parece recomendable la inversión en publicidad en periódicos frente a la publicidad en televisión o en radio.

Estadísticas de residuos^a

	Mínimo	Máximo	Media	Desviación estándar	N
Valor pronosticado	8,3532	21,5859	13,4250	3,61049	20
Residuo	-4,15122	3,24603	,00000	1,62157	20
Valor pronosticado estándar	-1,405	2,260	,000	1,000	20
Residuo estándar	-2,349	1,837	,000	,918	20

a. Variable dependiente: ventas

Gráficos

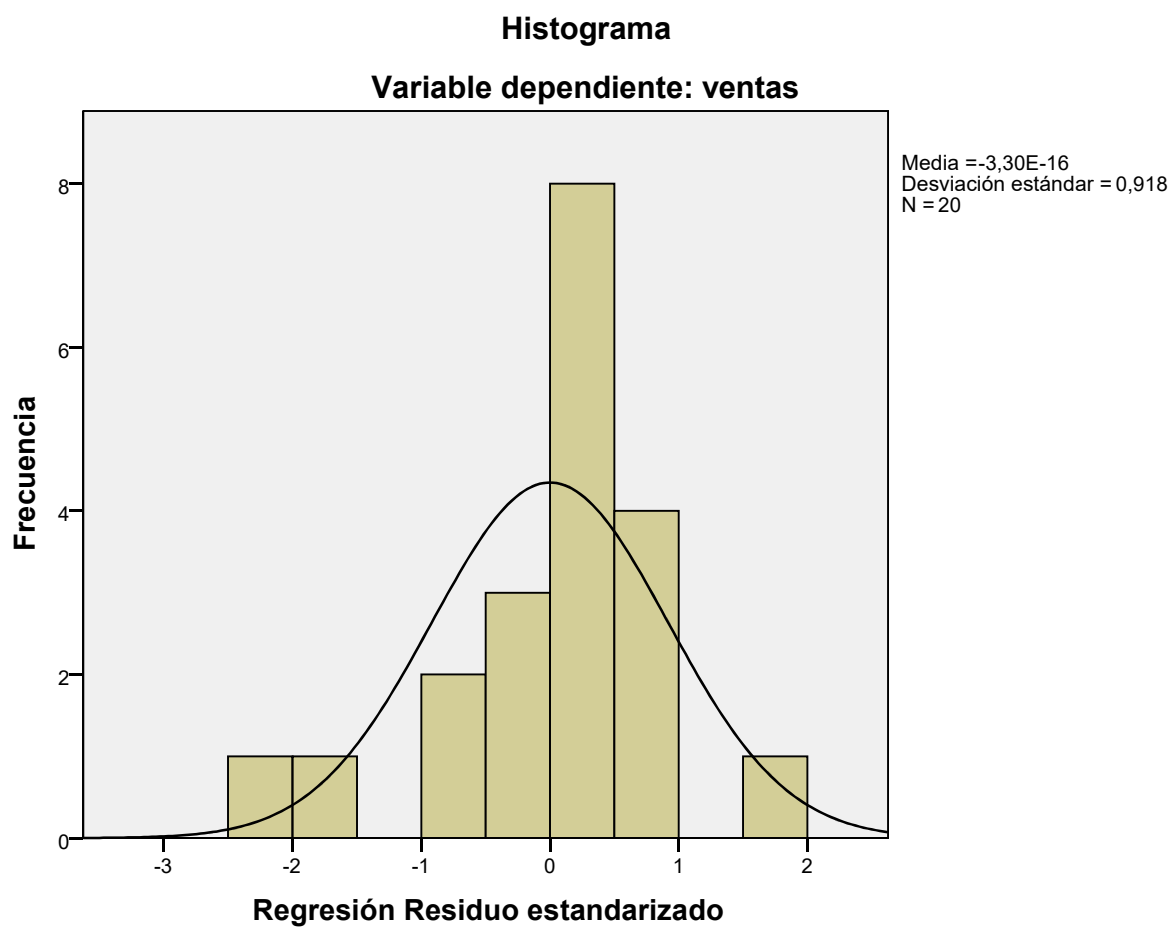
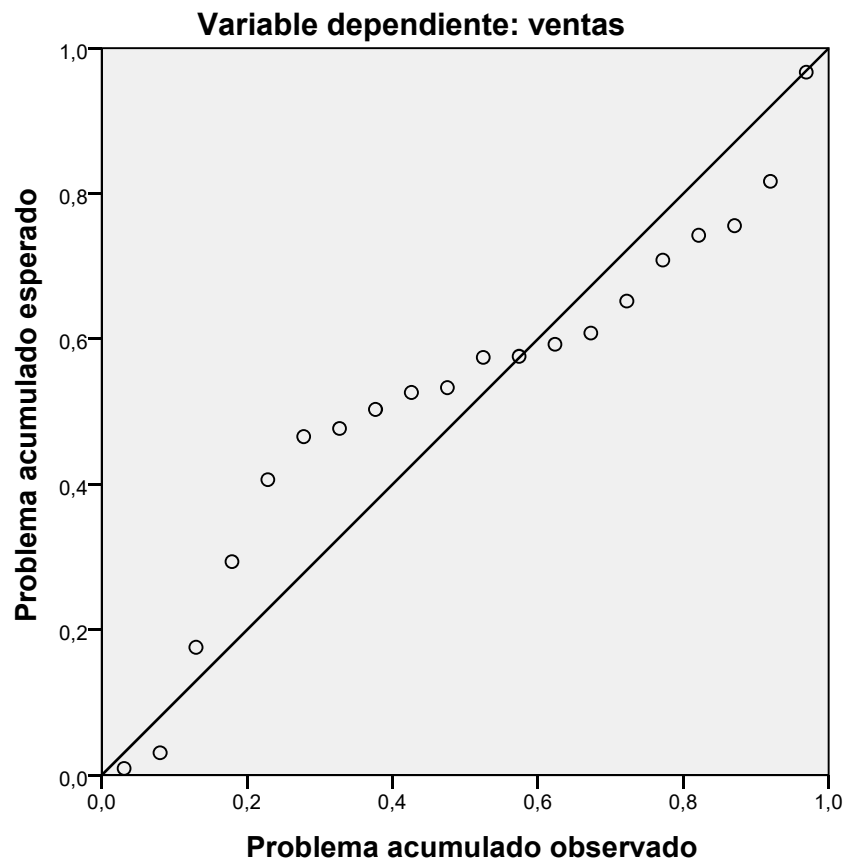


Gráfico P-P normal de regresión Residuo estandarizado



Regresión

Diagrama de dispersión
Variable dependiente: ventas

