CAPITULO 3.

3. MODELO DE REGRESIÓN MÚLTIPLE PARA LA DETERMINACIÓN DE LA UTILIDAD.

3.1 Introducción

La regresión lineal y la correlación son las dos herramientas estadísticas más poderosas y versátiles que se puede utilizar para solucionar problemas, por este motivo la utilizaremos en esta tesis.

Para plantear el modelo de regresión lineal múltiple para proyectar la utilidad de la empresa distribuidora de celulares y accesorios, es necesario definir la variable dependiente y las variables independientes.

Debido a que la utilidad depende de los ingresos, gastos y costos de venta; se ha definido a la *Utilidad* como la *Variable Dependiente* y a los *Ingresos, Gastos y Costos de Venta* como las *Variables Independientes,* quedando el modelo planteado de la siguiente manera:

$$\boxed{\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 + \hat{\beta}_3 X_3}$$

Donde :

$\hat{\beta}_0$, $\hat{\beta}_1$, $\hat{\beta}_3$ = Constantes.

X_1 = Variable Ingresos

X_2 = Variable Gastos

X_3 = Variable Costos de Ventas

$\hat{Y}$ = Variable Utilidad Estimada.

3.2 Modelo de Regresión Lineal Múltiple para la determinación de la Utilidad Diaria

Debido a que tanto los costos de venta como los gastos originados diariamente son valores de $0 como se indicó en el análisis de los datos en el Capitulo II, no se pudo realizar el modelo de regresión lineal múltiple para la determinación de la utilidad diaria, debido a que el programa estadístico utilizado SPSS considera estos datos como constantes impidiendo implementar el modelo, debido a que ciertos factores necesarios para validarlo no aparecen en las tablas que son utilizadas para aceptar o rechazar el modelo. Por este motivo se ha considerado implementar un modelo de regresión lineal múltiple para la determinación de la utilidad pero de manera mensual.

También podemos corroborar los resultados obtenidos en SPSS para el desarrollo del modelo de regresión lineal múltiple de la utilidad diaria mediante el gráfico 3.1 de dispersión entre las variables de estudios, llegando a la conclusión de que no existe una relación entre las variables. A continuación se muestra el gráfico de dispersión:

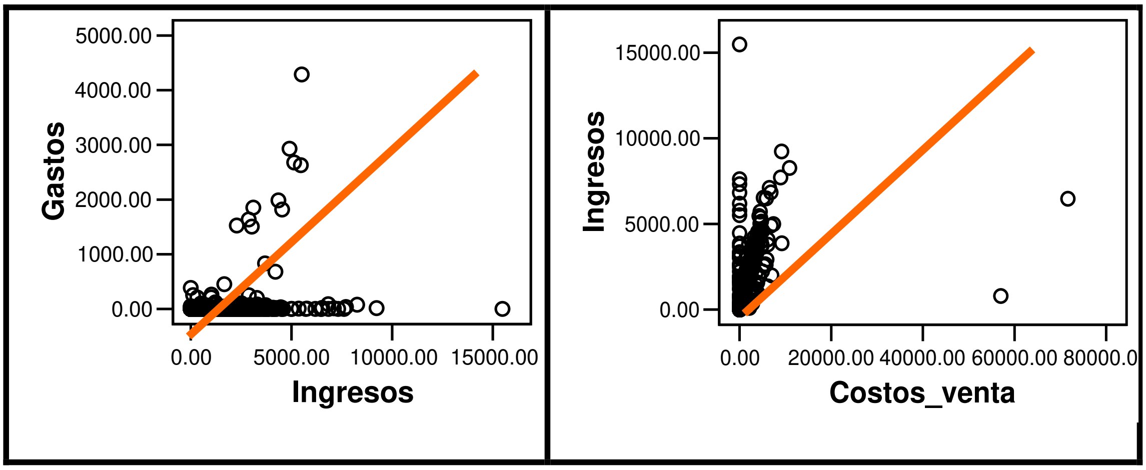


Gráfico 3.1 Gráfico de Dispersión entre las variables de estudio (Diaria)

Luego de ver que no existe una relación entre las variables de forma diaria, a continuación se propone un modelo de regresión de la utilidad de manera mensual.

3.3 Modelo de Regresión Múltiple para la determinación de la Utilidad Mensual.

3.3.1 Análisis de Diagramas de dispersión de las variables de estudio

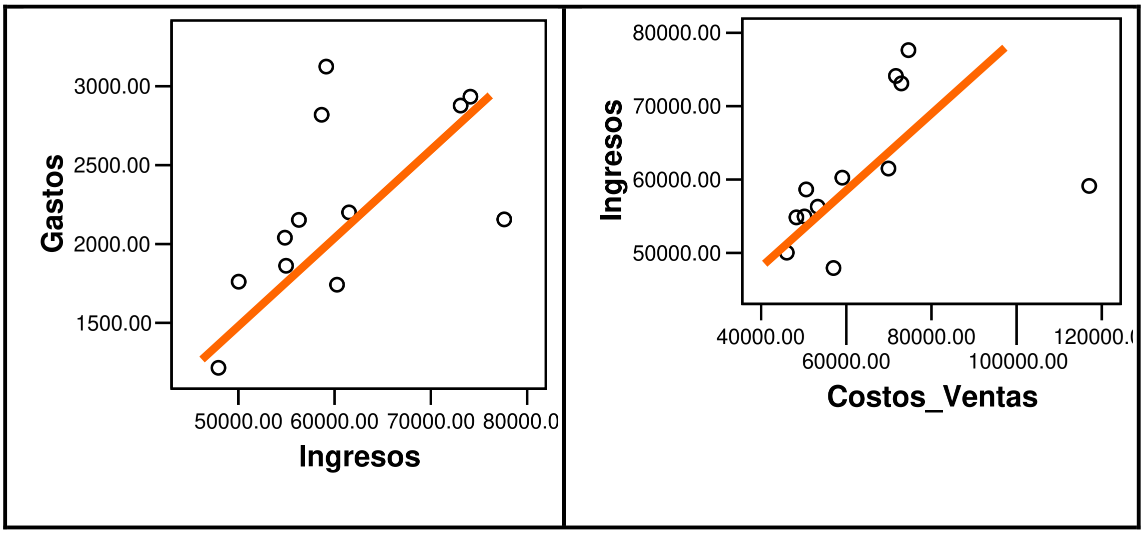


Gráfico 3.2. Diagrama de dispersión de las variables de estudio de (Mensual)

Como se puede apreciar en la gráfico 3.2. las variables de estudio llegan a tener una relación lineal. Este caso no se lo pudo apreciar en el gráfico 3.1 donde su diagrama de dispersión nos mostraba que no existía ninguna relación entre las variables.

Por tal motivo se ha llegado a proponer un modelo de regresión lineal múltiple de manera mensual, que a continuación se detallará.

3.3.2 Resumen del Modelo

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación	Durbin-Watson
1	1,00	,999	,999	575,26892	2,277

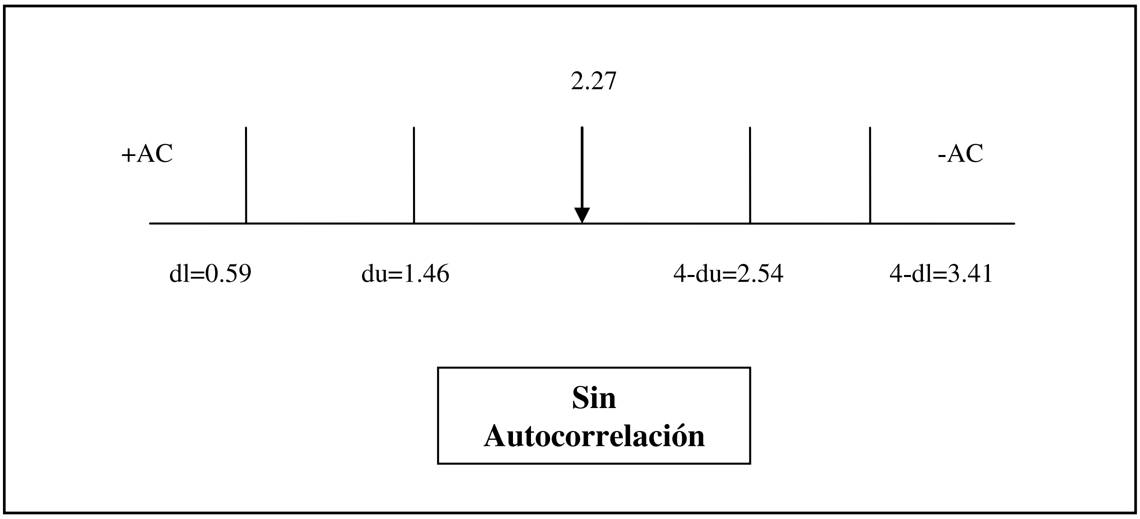
Tabla XIV. Resumen del Modelo planteado.

3.3.2.1 Autocorrelación

Los diagramas residuales nunca son tan obvios o tan fáciles de leer, por lo que se ha calculado mediante la Prueba de Durbin-Watson un valor para detectar una autocorrelación. Para esto he supuesto dos hipótesis para probar si existe o no correlación entre términos de errores sucesivos:

$$H_o = \rho_{e_i, e_{i-1}} = 0 \quad \text{(No existe autocorrelación)}$$

$$H_A = \rho_{e_i, e_{i-1}} \neq 0 \quad \text{(Existe autocorrelación)}$$



Dibujo 3. Prueba de Durbin-Watson

En este caso como Durbin-Watson está entre du y 4-du, no se rechaza la hipótesis nula, por lo tanto concluyo que la correlación no existe. Este tema de autocorrelación lo podemos clarificar cuando se analice la matriz de covarianza y varianza.

3.3.3 Análisis de Correlación

3.3.3.1 Coeficiente de Correlación

Para el caso de estudio el coeficiente de correlación obtenido en la Tabla 3.1, salió +1, lo que nos indica una relación positiva perfecta entre la variable dependiente (Utilidad) y las variables independientes (Ventas, Gastos y Costo de Venta).

3.3.3.2 Coeficiente de Determinación

Para el caso de estudio, se obtuvo un coeficiente de determinación R^2 de 0.99, lo que nos indica que el modelo se encuentra bien planteado, ya que se considera que un R^2 superior al 75% es aceptable.

3.3.4 Anova

	Suma de cuadrados	gl	Media cuadrática	F	F_α
Regresión	3.594.253.606,29	3	1.198.084.535,43	3.620,31	4,07
Residual	2.647.474,61	8	330.934,33		
Total	3.596.901.080,90	11			

Tabla XV. ANOVA

Mediante el ANOVA vamos a probar las siguientes hipótesis:

$$H_0 : \beta_o = \beta_1. = \beta_3 = \beta_4 = 0$$
$$H_1 : \exists i, \beta_i \neq 0$$

En este caso el estadístico de prueba F_{obs}=3620,309 y $F_\alpha(3{,}8)$ donde α =0,05 es igual a 4,07.Para rechazar la hipótesis nula, el estadístico de prueba debe ser mayor a F_α. En este caso de estudio se rechaza la hipótesis nula, es decir que al menos una de las variables X_i contribuye significativamente al modelo.

Al nivel del 5% parece existir una relación entre ventas, gastos, costo de ventas con las utilidades de la empresa, por lo que se puede decir que el modelo planteado cumpliría las futuras expectativas de la empresa para el año 2007.

3.3.5 Coeficientes.

	Coeficientes no estandarizados		**Coeficientes estandarizados**	**t**	**Sig.**
	B	**Error típ.**	**Beta**		
Cte.	1.298,16	1.128,81		1,15	0,28
Ventas	0,976	0,023	0,512	43,31	0,0
Gastos	-1,166	0,422	-0,038	-2,76	0,03
Costos_ Ventas	-0,995	0,011	-1,079	-88,79	0,0

Tabla XVI. Coeficientes

En la Tabla XVI se puede observar las estimaciones de los $\beta_0, \beta_1, \beta_2, \beta_3$ con su respectivos errores estándares.

El modelo de regresión lineal múltiple, está dado por la siguiente ecuación:

$$Y = 1298{,}156 + 0{,}976x_1 - 1{,}166x_2 - 0{,}995x_3$$

Los signos negativos para las variables: gastos y costos de ventas, son correctos, ya que como sabemos, estas dos variables disminuyen los ingresos, y consecuentemente se obtendrá la utilidad esperada.

De acuerdo a la tabla XVI nos podemos dar cuenta que β_0 no es significativa para el modelo, ya que con 8 grado de libertad y con un alfa del 5%, no se rechaza la hipótesis nula que es que β_0 = 0, debido a que $t_{8,0.05}$ = 2,306 y este valor es mayor a 1. Cabe recalcar que podemos confirmar lo expuesto en el párrafo anterior, por medio del valor p, ya que como se muestra en la tabla 3.2.5 el valor p de β_0 es 0,28 y este valor es mayor que el valor alfa 0,05, por lo tanto se concluye también que no se rechaza la hipótesis de que β_0 no es significativo para el modelo, es decir que es igual a cero. Pero para nuestro vamos a considerar el valor de β_0 en nuestra ecuación.

3.3.6 Correlaciones de los Coeficientes

	Variables	Costos_Ventas	Ventas	Gastos
Correlaciones	Costos_Ventas	1,000	-,090	-,499
	Ventas	-,090	1,000	-,456
	Gastos	-,499	-,456	1,000
Covarianzas	Costos_Ventas	,000	,000	-,002
	Ventas	,000	,001	-,004
	Gastos	-,002	-,004	,178

Tabla XVII. Correlación de Coeficientes.

Mediante la Tabla XVII, podemos decir que la variable Costo de Venta y la variable Ventas poseen una relación inversa, con un coeficiente de correlación de -0,090. Lo mismo podemos concluir de la variable Costos de Ventas con la variable Gastos, con un coeficiente de correlación de -0,499.

Ahora, entre las variables ventas y gastos en relación con los costos de ventas, la variable Gastos tiene una relación lineal más fuerte con los costos de ventas ya que el valor de -0,499 es más cercano a -1.

La variable Ventas posee una relación inversa también con variable Gastos, con un coeficiente de correlación de -0,456, además entre estas 2 variables existe un relación lineal ya que el valor de correlación está cerca de -0,456.

Claro está que, las variables que poseen una relación lineal fuerte son los costos de ventas con los gastos.