

Análisis de Regresión y Correlación Lineal

Regression and Linear Correlation Analysis

Fabiola Leal-Cornejo^a, Robert E. López-García^a, Mónica G. Martínez-Montiel^a, Delia I. Tapia-Castillo^b, Irma I. de León-Vázquez^c

Abstract:

Regression and correlation is used in various disciplines such as sociology, biomedicine, engineering, economics, among others. They are closely related, understanding a form of estimation, they are usually used to solve a large number of problems, while the regression is responsible for the relationship between variables, the correlation measures the degree of linear relationship between two or more variables, resulting in the strength and meaning of the relationship.

In other sources consulted, it tells us that the correlation analysis results in a number that summarizes the degree of the correlation between two variables, and the regression analysis gives a mathematical equation that describes that relationship, the data that are necessary come from observations of related variables.

To make a correlation, it is necessary to know what type the variables are, be they quantitative or qualitative, since different methods were used, the advantages of the correlation have their own interpreted language.

In the case of regression, for the result to be correct it is necessary to make an adequate selection of the variables, because if we take variables that do not have any relation to the practice, it will result in illogical, that is, it will not make sense, this is a useful tool for planning.

It is also called regression analysis to any statistical method that establishes an equation allowing estimating the unknown value of a variable from the known value of one or more variables, this term was used for the first time in 1877 by the English statistician Francis Galton, based on a study he did, showing that the height of the tall children tended to recede, or return. It is always developed in an estimation equation, relating the unknown variables with the known variables.

Almost always in the correlation analysis with the regression analysis are used together to measure the effectiveness that the regression line explains the variation of the dependent variable, Y

Keywords:

Mathematics, regression, correlation, linear

Resumen:

La regresión y la correlación se emplea en diversas disciplinas como la sociología, la biomedicina, las ingenierías, la economía, entre otras., están estrechamente relacionadas comprendiendo una forma de estimación, se suelen utilizar para resolver una gran cantidad de problemas, mientras que la regresión se encarga de la relación que hay entre variables, la correlación mide el grado de relación lineal entre dos o más variables, dando como resultado la fuerza y sentido de la relación.

En otras fuentes consultadas, nos dice que el análisis de correlación da como resultado un número que resume el grado de la correlación entre dos variables, y el análisis de regresión da una ecuación matemática que describe dicha relación, los datos que son necesarios provienen de observaciones de variables relacionadas.

Para hacer una correlación se requiere saber de qué tipo son las variables, sean cuantitativas o cualitativas, pues se utilizaron distintos métodos, las ventajas de la correlación tienen su propio lenguaje interpretado.

^a Alumno de la Maestría en Educación, Universidad Interamericana para el Desarrollo, Sede Tula. Email: Autor de correspondencia: Roberto López-García roberthlogar@outlook.com; Mónica Martínez moni_03@outlook.com; Fabiola Leal fabileal@outlook.com

^b Profesor de Tiempo Completo, Universidad Autónoma del Estado de Hidalgo, Escuela Superior de Tlahuelilpan. Iliana Tapia Castillo itapia@uaeh.edu.mx

^c Autor de Correspondencia: Profesor por asignatura, Universidad Autónoma del Estado de Hidalgo, Escuela Superior de Tlahuelilpan. Email: Irma Isabel de León Vázquez irmalv@uaeh.edu.mx

En el caso de la regresión, para que el resultado sea correcto es necesario tomar hacer una selección adecuada de las variables, pues si tomamos variables que no tengan ninguna relación con la práctica nos dará como resultado ilógico, es decir no tendrá sentido, esta es una herramienta útil para la planeación.

También se le denomina análisis de regresión a cualquier método estadístico que establezca una ecuación permitiendo estimar el valor desconocido de una variable a partir del valor conocido de una o más variables, este término fue utilizado por primera vez en 1877 por el estadístico inglés Francis Galton, en base a un estudio que realizó, demostrando que la altura de los hijos altos, tendía a retroceder, o regresar. Esta siempre se desarrolla en una ecuación de estimación, relacionando las variables desconocidas con las variables conocidas.

Casi siempre en el análisis de correlación con el análisis de regresión se utilizan en conjunto para medición de la eficacia que la línea de regresión explica la variación de la variable dependiente, Y.

Palabras Clave:

Matemáticas, regresión, correlación, lineal

Introducción

La presente investigación se refiere a elementos necesarios de ser conocidos para cualquier tipo de estudio donde se tienen datos. El análisis de regresión analiza la relación de variables, ya sea una variable dependiente y una o varias variables independientes.

El análisis de correlación consiste en estudiar el grado de asociación de las variables ya sea de una o varias variables.

En este tema se describirá y medirá el tipo de relaciones, donde aparecen en una gran cantidad de problemas.

Marco Referencial

En el presente trabajo de investigación se quiere dar a conocer dos técnicas usadas en la rama de la estadística que son regresión y correlación, esperando generar un amplio conocimiento en cuanto al tema y sus características generales.

Metodología

En el presente trabajo de investigación es conveniente utilizar el programa de office Excel y el programa minitab para la elaboración de tablas de datos, graficas de dispersión.

Para la realizació de la tabla de datos proseguimos a utilizar minitab.

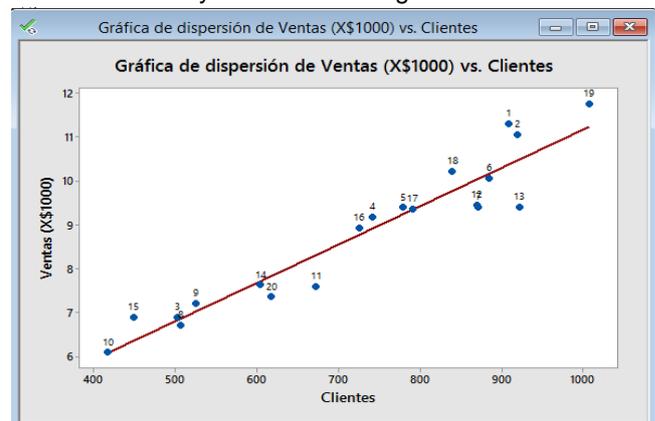
Descripción del primer paso, agregando los datos tomados de un ejemplo:

Tienda	Clientes	Ventas (X\$1000)
1	909	11.3
2	920	11.06
3	503	6.89
4	742	9.18

5	779	9.39
6	885	10.05
7	872	9.4
8	507	6.71
9	526	7.21
10	418	6.09
11	673	7.6
12	870	9.44
13	922	9.41
14	605	7.63
15	449	6.9
16	726	8.93
17	791	9.35
18	840	10.22
19	1007	11.76
20	618	7.36

Resultados

Como resultados se presenta la gráfica de dispersión entre las ventas y el análisis de regresión correlacional.



Gráfica de dispersión de Ventas (X\$1000) vs. Clientes
Análisis de regresión: Ventas (X\$1000) vs. Clientes

Fuente	GL	SC Ajust	MC Ajust	Valor F	Valor P
Regresión	1	47.183	47.1833	175.16	0.000
Clientes	1	47.183	47.1833	175.16	0.000
Error	18	4.849	0.2694		
Total	19	52.032			

Resumen de modelo

S	R.cuad	R-cuad. (ajustado)	R-cuad (pred)
0.519005	90.68%	90.16%	88.39%

Coefficientes

Término	Coef	EE del Coef	Valor T	Valor P	VIP
Constante	2.414	0.496	4.87	0.000	
Clientes	0.008763	0.000662	13.23	0.000	1.00

Ecuación de regresión:

$$\text{Ventas (X\$1000)} = 2.414 + 0.008763 \text{ Clientes}$$

Ajustes y diagnósticos para observaciones poco comunes:

Obs	Ventas (X\$1000)	Ajuste	Resid	Resid est
13	9.410	10.493	-1.083	-2.21 R

Residuo grande R

Correlación: Clientes, Ventas (X\$1000)

Correlación de Pearson de Clientes y Ventas (X\$1000)=0.952

Valor p=0.000

Rho de Spearman: Clientes y Ventas (X\$1000)

Rho de Spearman para Clientes y Ventas (X\$1000)=0.950

Valor p=0.000

Discusión

Se puede observar en los resultados obtenidos en la gráfica de dispersión de ventas que la tienda con mayor número de ventas es la tienda número 19 así mismo la tienda con menor número de clientes que genera escasas de ventas que es la tienda número 10.

Tomando estos datos se puede hacer un análisis para identificar el problema, y plantear una posible solución.

Se calculó el coeficiente de correlación de Pearson y el coeficiente de correlación de Spearman arrojando datos específicos que nos ayudan a la correcta interpretación de los datos generados al inicio del problema.

Conclusiones

Es importante saber sobre estos dos temas, pues son de suma importancia para el estudio de las variables, pues dan a conocer el efecto, relación que hay entre una o más variables, también se puede predecir en casos futuros la correlación de estas, a partir de otra.

Además, ayudan considerablemente la toma de decisiones, si los responsables determinan como lo conocido se relaciona con un evento futuro, por lo tanto el análisis de regresión lineal, es el proceso general de predecir una variable (Y) a partir de otra (X), su relación puede ser directa o inversa, directa cuando la pendiente crece en conjunto con la variable X e inversa porque a media que aumenta la variable Y, va disminuyendo la variable X.

Referencias

- [1] Anónimo. (s.f.). *FAO*. Recuperado el 08 de Agosto de 2018, de <http://www.fao.org/docrep/003/x6845s/x6845s02.htm>
- [2] Mendiburu, F. d. (s.f.). *Análisis de regresión y correlación*. Recuperado el 08 de Agosto de 2018, de <https://tarwi.lamolina.edu.pe/~fmendiburu/index-filer/academic/metodos1/Regresion.pdf>
- [3] Mora, F. A. (18 de Enero de 2017). *Youtube*. Recuperado el 08 de Agosto de 2018, de https://www.youtube.com/watch?v=RI_8gMLnEus
- [4] UNAM. (08 de 2018). Obtenido de http://asesorias.cuautitlan2.unam.mx/Laboratoriovirtualdeestadistica/CARPETA%203%20INFERENCIA_ESTADISTICA/DOC_%20INFERENCIA/TEMA%204/09%20REGRESION%20Y%20CORRELACION%20LINEAL%20SIMPLE.pdf