Cross-Currents: An International Peer-Reviewed Journal on Humanities & Social Sciences
Abbreviated Key Title: Cross Current Int J Peer Reviewed J Human Soc Sci
ISSN: 2394-451X (Print) & Open Access
DOI: 10.36344/ccijhss.2020.v06i03.003

Volume-6 | Issue-3|Mar-2020 |

Case Report

Application of Linear Correlation to Evaluate a Process Production: Special Red Mole Study Case

Villarreal Ramírez Carmen Danae[1], Pérez Gutiérrez Néstor Marcel[1], Ortiz Zamora Concepción[1], José Celis Verónica Andrea[1], Verduzco Ramírez Patricia Ivonney[2] Villar Laguna Víctor[3]

[1]Departamento de Administración UniverMilenium, Campus Ixtapaluca Avenida Acozac 11, Santa Barbara, C. P. 56530 Ixtapaluca, Estado de México, México

[2]Universidad de la República Mexicana Av. Tláhuac 4761, El Vergel, 09880 Alcaldía de Iztapalapa, Ciudad de México, México

[3]ESIA Tecamachalco Instituto Politécnico Nacional Av. Fuente de Los Leones 28, Lomas de Tecamachalco, 53950 Naucalpan de Juárez, Estado de México, México

***Corresponding author:**
Villarreal Ramírez Carmen Danae

Abstract: The present investigation seeks to measure the degree of reliability of the production process of the special red mole, for which the following variables were measured: parts produced against detected defects, applied with a 95% reliability margin of reliability with a margin 0.05% error to achieve a more robust production system. And above all to achieve a more efficient process to avoid waste change as it is: raw material, time and money.

Keywords: linear correlation, time standardization & continuous improvement.

INTRODUCTION

Today organizations seek to transcend in a competitive market, wh ich globalization has been gradually modifying manufacturing procedures, to increase productivity, that is why SMEs must modify their processes to survive in this competitive market and that is why that the correlation between one variable and another must be known, to know the trend that sets the pace of production.

Linear regression models are widely used in engineering since they serve to analyze the behavior of the input (or regressor) and output (or response) variables by establishing predictions and estimates. A correlation coefficient measures the degree of relationship or association generally existing between two random variables (Restrepo & González, 2007).

Among The Advantages Of Using This Statistical Technique Are:

- It is used to forecast important events.
- Measure the reliability of a process.
- There is a relative comparison between a variable to another.

The Usefulness Of This Statistical Technique Is:

- Give a first contemplation to a set of data.
- When you have numerical data with many peculiarities.
- Perform econometric forecasts.
- Creating marketing responses.

While it is true it is used in time series, they try to predict the future based on the previous information. It is also used for causal correlation.

Linear regression is used for forecasts of both time series and causal relationships. when the dependent variable changes as a result of time. It is a time series analysis. If one variable changes due to changes in the other.

In the present investigation, the linear correlation was implemented to determine the degree of reliability in a production process focused on the food process in a SME company.

BODY OF MANUSCRIPT

Antecedentes

Linear correlation: as a linear correlation coefficient (from Pearson), it is a regression measure

Received: 23.02. 2020
Accepted: 25.02.2020
Published: 30.03.2020

Journal homepage:
http://crosscurrentpublisher.com/ccijhss/

Quick Response Code

that aims to quantify the degree of joint variation between two variables (Jaramillo, 2018).

Linear regression is a statistical tool widely used to analyze how a set of variables (independent or explanatory) influence another (dependent or explained), allowing the numerical estimation of the signs and magnitudes of the coefficients in a previously established linear relationship. One of the characteristics of the economic variables is the possible correlation between them due to the existence of common determinants, so the multicollinearity between explanatory variables in a multiple regression should be considered a usual situation (Salmerón & Rodríguez, 2017).

The mathematical linear regression tool was used to evaluate two case studies in forage corn; in the first one, the relation of the plate variables (of evaporation and applied water) to open sky and as a function of time was established; in the second, the production of dry matter (DM) was related to the foliar area index (IAF) in forage corn with and without plastic padding over time, in order to understand the phenomenon that occurs in both trials. The works were carried out during the period from 2006 to 2013, in the Lagunera Region of Coahuila and Durango, Mexico, where data groups of the variables evaporation (EV), applied irrigation sheet (LR), MS and IAF of the corn crop forage were used to propose mathematical functions of the linear regression type (Montemayor, Munguía, Segura, Yescas, Orozco & Woo, 2017).

It is important to mention that One of the important fields of application of linear regression and correlation is that of teaching in middle school. This paper aims to make a didactic analysis following the approach proposed by Fernández, Monroy & Rodríguez (1998) according to his book, where they conduct a study from three perspectives: content, teaching and learning. Depending on the analysis, various teaching situations that can be useful to carry out in the classrooms, together with the students Lavalle, Micheli, & Rubio (2006), are shown. The results show linear equations of first and second order of predictive type with R2 greater than 0.8, which allow us to understand and interpret the phenomena of water consumption and dry matter accumulation of forage corn. Montemayor, Munguía, Segura, Yescas, Orozco & Woo (2017).

A mathematical model is the simplification of a real system, generated to understand, explain, estimate or predict reality. Thornley (1976) mentions two types of models to explain the physiological processes of plants: the "mechanistic" type and the "empirical" type; for the first one, a thorough knowledge of the various processes involved (subsystems) is required to reach a final objective, in which each process is expressed as an equation, and each of these equations is concentrated in a system of equations that are they simplify to obtain a final equation that is subsequently verified with real data.

Linear regression models are widely used in engineering since they serve to analyze the behavior of the input (or regressor) and output (or response) variables by establishing predictions and estimates (Montgomery, 2006). In this work the return variable corresponds to the individual harmonic distortion of current and the response variable corresponds to the individual harmonic distortion of voltage. The equation shows the representation of a simple linear regression model, where Y is the answer, X is the regression variable, 0 and 1 are the parameters of the model or regression coefficients and it is the model error (Astorga, 2014).

It is important to mention that multiple linear regression has two basic hydrological applications: (1) expand short registers based on nearby long series, and (2) deduce empirical equations that allow estimating, in sites of interest without gaps, increasing design (QTr). As both applications are carried out in a regional context, multicollinearity is always present in the first case, and the lack of homocedasticity in the second. To correct the non-uniformity of the variances of the dependent variable (Yi), a weighted function (wi) is used in the adjustment of least squares, which leads to the technique of weighted least squares (MCP) (Campos, 2016).

One of the main assumptions of the linear regression analysis is the existence of a causality relationship between the analyzed variables, without the regression analysis allowing it to be demonstrated. This research demonstrates the causality between the variables analyzed through the construction and feedback analysis between the variables under study, embodied in a causal diagram and validated through dynamic simulation. One of the main contributions of this research is the proposal to use a systems dynamics approach, to develop a transition method from a predictive multiple linear regression model to an explanatory simple nonlinear regression model, which increases the level of model prediction (Baeza & Vázquez, 2014).

PURPOSE OF THE RESEARCH

Apply linear correlation to evaluate a production process: case study special mole red.

Specific objective

- Know the process of the special red mole and the topic of linear correlation.
- Understand the issue of linear correlation and the composition of the activities of the special red mole.
- Analyze the times of each activity of the special red mole process.

- Synthesize the information to be able to develop a study in the special red mole.
- Evaluate the process of the special red mole by applying linear correlation.

METHODOLOGY

Multiple regression analysis is a statistical method used in many areas of knowledge (Ramos, Bautista & Valdez, 2010).

That is why a sample of the following product was taken: Special red mole.

Image-1 Special red mole

For the study, the following data were taken:

Where:

X: Parts Produced
Y: Defects detected

The linear correlation was calculated in the complement of data analysis, to know if there is a relationship between temperature and mole consistency. For this, one of the Excel® (Data Analysis) add-ons was used, which favored the calculation of the linear correlation and the analysis of variance.

Linear regression analysis is an extremely important tool in the world of Finance, because it allows projections and forecasts of a dependent variable explained by one or more independent variables (Brenes, 2017). The objective of this work was to determine an equation that allowed estimating the average monthly price of the shares of the company Facebook, Inc., through a simple linear regression model. The estimated beta coefficients for the model were significant for both the constant and the slope, measured through the statistical test t. Likewise, the global test of significance of the beta coefficients was performed, determined through the statistical test F, this proved to be extremely significant, which allowed the validation of the model (Brenes, 2017).

Table-1. Project data

x	y
2	1
3	2.5
2.5	1
3.5	4
4	4
4	5
6	5
6.5	5
7	4
7	6
8	6
7	6.5

The Excel® Software Data Analysis tool was used. Using the Data Analysis option in the Excel Tools menu, a Linear Regression Analysis is made that allows obtaining the slope and the intercept of the adjusted line that relates the system variables, as well as their respective typical errors, and the correlation coefficient linear between these variables (Cuenca, Tamayo & Tamayo, 2010).

The results of the data analysis tool are shown in table 2.

Table-2 Results of the project

Estadísticas de la regresión	
Coeficiente de correlación múltiple	0.85276092
Coeficiente de determinación R^2	0.72720119
R^2 ajustado	0.69992131
Error típico	1.00692712
Observaciones	12

It is observed that the adjusted R ^ 2 coefficient has a value of 0.69, which according to the literature, Martínez, Tuya, Martínez, Pérez & Cánovas (2009) indicate in their article that a result of 0.69 is equivalent to a regular scale.

Table 3 represents the results of the analysis of variance.

Table-3 Results of the project

	Grados de liberta	na de cuadra	dio de los cua	F	alor crítico de F
Regresión	1	27.0276444	27.0276444	26.657052	0.00042335
Residuos	10	10.1390223	1.01390223		
Total	11	37.1666667			

e

The result of the ANOVA table in the value of "F" or the variation of the regression is 27.02. In order to know if the result of the project is significant, the value of "F" is then compared, which is 27.02 and is much greater than the critical value for F (0.0004). This concludes that the result is significant. In other words, there is a relationship between the variables.

Table-4 Results of the coefficients

	Coeficientes
Intercepción	0.37276299
Variable X 1	0.75250982

Finally, the results of the intercept coefficients and Variable X1 are reviewed. These results will serve to form a forecast model.

Likewise, this model was tested in Minitab® in its student version 18, in which a statistical model of multiple linear regression was proposed for decision-making as shown below (Alcántar, Treviño & Martínez, 2015).

In real life, there are statistical variables that cannot be accurately observed or that have intrinsically interval character. This is the case, for example, of the measures provided by certain instruments that have some margin of error or tolerance in their measurements or the temperature range of a city over several days.

Regarding the nature of the intervals, some authors consider that the uncertainty arises from observing a real random variable with some imprecision. Thus, the true value of the variable is contained in the interval. Other authors work with intervals obtained through data censorship. So-called symbolic data also includes intervals, which are used to summarize large data sets. Finally, there are intervals, which arise naturally when trying to represent a magnitude. For example, fluctuations or ranges. These are the type of intervals that are considered during this thesis. Within the different studies that exist to work with intervals in Statistics, the objective is focused on regression (García, 2016).

Linear regression models are widely used in engineering since they serve to analyze the behavior of the input (or regressor) and output (or response) variables by establishing predictions and estimates In this work the regression variable corresponds to the individual harmonic distortion of current and the response variable corresponds to the individual harmonic voltage distortion. The equation shows the representation of a simple linear regression model, where Y is the answer, X is the regression variable, 0 and 1 are the parameters of the model or regression coefficients and it is the model error (Pizarro, 2003).

One of the main assumptions of the linear regression analysis is the existence of a causality relationship between the analyzed variables, without the regression analysis allowing it to be demonstrated. This research demonstrates the causality between the variables analyzed through the construction and feedback analysis between the variables under study, embodied in a causal diagram and validated through dynamic simulation. One of the main contributions of this research is the proposal to use a systems dynamics approach, to develop a transition method from a predictive multiple linear regression model to an explanatory simple nonlinear regression model, which increases the level of model prediction (Baeza & Vázquez, 2014).

One of the most characteristic signs of today's society is undoubtedly the explicit concern for quality, productivity and costs in all areas of the international economy. However, this concern has acquired different nuances depending on the sector of the economy in question and the national context in question (Rincon, 2001).

To achieve sustainability, the company needs to be productive and face potential losses in a creative and innovative way and this can be done from knowledge management. Organizations that can deploy their capabilities and resources with speed and flexibility and that also manage intelligence, creativity and intuition effectively and productively develop the ability to innovate quickly, achieving growth and development in a changing environment with dynamic and permanent evolution.

Productivity is a strategic objective of companies, because without it the products or services do not reach the necessary levels of competitiveness in the globalized world (Medina, 2010).

The results in the Minitab® software were the following:

Figure 1 represents the summary of the model, where the adjusted R ^ 2 coefficient stands out has a value of 0.69. Which is the same result as the Excel® software data analysis tool (figure 1).

Resumen del modelo

S	R-cuad.	R-cuad. (ajustado)	R-cuad. (pred)
1.00693	72.72%	69.99%	60.66%

Figure-1 Results of the project

Then the Minitab® software yielded the results of the variance analysis. You can see in figure two, the section of Value F, where it gives a value of 26.66, same value as the result of table 3 that the data analysis tool of Excel® software (figure 2).

Análisis de regresión: y vs. x

Análisis de Varianza

Fuente	GL	SC Ajust.	MC Ajust.	Valor F	Valor p
Regresión	1	27.028	27.0276	26.66	0.000
x	1	27.028	27.0276	26.66	0.000
Error	10	10.139	1.0139		
Falta de ajuste	7	6.139	0.8770	0.66	0.709
Error puro	3	4.000	1.3333		
Total	11	37.167			

Following the calculation of the problem, the Minitab® software calculated the coefficients of the regression model, giving the same results as Table 4 as the Excel® software data analysis tool (figure 3).

Coeficientes

Término	Coef	EE del coef.	Valor T	Valor p	FIV
Constante	0.373	0.790	0.47	0.647	
x	0.753	0.146	5.16	0.000	1.00

Figure-3 Results of the project

An advantage of the Minitab® software is that it offers the regression equation equation. Figure 4 represents the equation for the project (figure 4).

Ecuación de regresión

$$y = 0.373 + 0.753\ x$$

Figure-4 Regression equation

CONCLUSIONS

Regression and correlation techniques help quantify the statistical association between two or more variables. The analysis of this work constitutes an approach to knowledge about the teaching and learning of regression and correlation. This study helped to know the relationship of the variables of the red special mole process.

The research achieved the purpose, which was: apply linear correlation to evaluate a production process: case study red mole special. It was also possible to meet each of the specific objectives. Technology tools could be used for analysis and interpretation.

It is recommended to train employees in statistics and senior managers in the use of specialized statistics software.

Acknowledgment

- Arturo González Torres who advised and worked together with the project team.
- Vianey Ríos Romero who advised and worked together with the project team.

REFERENCES

1. Alcántar Ruiz, R. A., Treviño Treviño, F. E., & Martínez Flores, J. L. (2015). Modelo estadístico que permite observar el impacto de los factores que inciden en el rendimiento de combustible. Nova Scientia [Internet].7 (14), 236-253.
2. Astorga Gómez, J. M. (2014). Aplicación de modelos de regresión lineal para determinar las armónicas de tensión y corriente. *Ingeniería Energética, XXXV* (3), septiembre-diciembre, pp. 234-241
3. Baeza Serrato, R., & Vázquez López, J. A. (2014). Transición de un modelo de regresión lineal múltiple predictivo, a un modelo de regresión no lineal simple explicativo con mejor nivel de predicción: Un enfoque de dinámica de sistemas. *Revista Facultad de Ingeniería Universidad de Antioquia*, (71).
4. Brenes González, H. A. (2017). Aplicación del análisis de regresión lineal simple para la estimación de los precios de las acciones de Facebook, Inc. REICE: Revista Electrónica de Investigación en Ciencias Económicas.
5. Campos Aranda, D. F. (2016). Una aplicación hidrológica de la regresión lineal múltiple ponderada. *Tecnología y ciencias del agua, 7*(4), 161-173.
6. Cardozo, O. D. (2011). Integración de modelos para la estimación de la demanda en el transporte público. Aplicación de sistemas de información geográfica, análisis de regresión múltiple y funciones distance-decay al metro de *Madrid (Universidad de Alcalá) España*.
7. Cuenca Rodríguez, D., Tamayo Cuenca, R. & Tamayo Pupo, J. (2010). Aplicación del Programa Microsoft Excel para resolver problemas experimentales de Física Ciencias Holguín, *XVI (3), julio-septiembre, pp. 1-10. Centro de Información y Gestión Tecnológica de Santiago de Cuba. Holguín, Cuba*
8. Fernández, F., Monroy, O. L., & Rodríguez, L. (1998). Diseño, desarrollo y evaluación de situaciones problemáticas de estadística. *Una empresa docente®. Universidad de los Andes.*
9. García Bárzana, M. (2016) Estimación de modelos de regresión lineales múltiples con restricciones aleatorias lineale (*Universidad de Oviedo). España.*
10. Instituto Superior Politécnico José Antonio Echeverría. La Habana, Cuba.
11. Jaramillo Sánchez, D. (2018). Correlación Lineal. Actividad. Materia. Probabilidad y Estadística. Semestre: agosto – diciembre 2018. Facultad de Ingeniería Mecánica y Eléctrica. *Universidad Autonoma de Nuevo Leon. San Nicolás de Los Garza, Nuevo León.*
12. Lavalle, A. L., Micheli, E. B., & Rubio, N. (2006). Análisis didáctico de regresión y correlación para la enseñanza media. Revista Latinoamericana de Investigación en Matemática Educativa, *RELIME*, 9(3).
13. Martínez Ortega, R. M., Tuya Pendás, L. C. Martínez Ortega, M., Pérez Abreu, A., & Cánovas, A. M. (2009). El coeficiente de correlación de los rangos de spearman caracterizacion. Revista Habanera de Ciencias Médicas, 8, (2), abril-junio. *Universidad de Ciencias Médicas de La Habana. Ciudad de La Habana, Cuba.*
14. Medina Fernández de Soto, J. E (2010). Modelo integral de productividad, aspectos importantes para su implementación. Revista Escuela de Administración de Negocios, (69).
15. Montemayor Trejo, J. A., Munguía López, J., Segura Castruita, M. Á., Yescas Coronado, P., Orozco Vidal, J. A., & Woo Reza, J. L. (2017). La regresión lineal en la evaluación de variables de ingeniería de riego agrícola y del cultivo de maíz forrajero. *Acta Universitaria [Internet]*. *27*(1), 40-44.
16. Montgomery, D. (2006). Introducción al análisis de regresión lineal. *México: Ed. Limusa Wiley*, 588 p.
17. Nagles García, N. (2006) Productividad: una propuesta desde la gestión del conocimiento. Universidad EAN Bogóta, Colombia.
18. Pizarro Junquera, J. (2003). Selección de modelos en regresión mediante la respuesta al ruido En la *Universidad de Cádiz. España*.
19. Ramos Herrera, S., Bautista Margulis, R., & Valdez Manzanilla, A. (2010). Estudio estadístico de la correlación entre contaminantes atmosféricos y variables meteorológicas en la zona norte de Chiapas, México. *Universidad y ciencia, 26*(1), 65-80.
20. Restrepo B., L. F., & González, L. J. (2007). De Pearson a Spearman. Revista Colombiana de Ciencias Pecuarias, 20 (2), abril-junio, pp. 183-192. Universidad de Antioquia. Medellín, Colombia.
21. Rincón de Parra, H. (2001). Calidad, Productividad y Costos: Análisis de Relaciones entre estos Tres Conceptos. Actualidad Contable Faces, 4(4).
22. Salmerón Gómez, R., & Rodríguez Martínez, E. (2017). Métodos cuantitativos para un modelo de regresión lineal con multicolinealidad. Aplicación a rendimientos de letras del tesoro. Revista de Métodos Cuantitativos para la Economía y la Empresa [Internet]. 24. Pp.169-189.
23. Thornley, J. H. M. (1976). Mathematical Models in Plant Physiology. *Ed. Academic Press. England*. 318.