

# APLICACIÓN DE LA CORRELACIÓN Y REGRESIÓN EN EL DESEMPEÑO INGENIERIL

**Ing. Ana Rosa Montané Cobas, Ing. Ruslán García Linares**

*Universidad de Matanzas – Sede “Camilo Cienfuegos”, Vía Blanca Km.3, Matanzas, Cuba.*

[ana.montane@umcc.cu](mailto:ana.montane@umcc.cu)

[ruslan.garcia@umcc.cu](mailto:ruslan.garcia@umcc.cu)

## **Resumen**

La ingeniería como ciencia se ha beneficiado con el desarrollo reciente de la Estadística. La correlación y regresión, como parte de la inferencia estadística, constituye una herramienta poderosa que ofrece al ingeniero la posibilidad de ajustar procesos, tomar decisiones y establecer políticas. Por esta causa resulta vital la perfección del proceso de enseñanza y aprendizaje siendo los profesores los principales protagonistas. La correcta aplicación de estos importantes métodos estadísticos mediante las nuevas tecnologías por parte de los ingenieros en la sociedad, no se puede medir como un hecho meramente investigativo, sino como un aporte esencialmente económico para desarrollar nuestro país.

*Palabras claves: Correlación; Regresión; Ingeniería.*

---

## Introducción

Vivimos en una época plagada de cambios donde ciencia, tecnología e innovación son términos difíciles de ignorar en todos los rincones del mundo. El desarrollo vertiginoso de las Tecnologías de la Información y las Comunicaciones (TIC) en la actualidad ha afectado todos los campos de nuestra sociedad, donde la Estadística no es una excepción. Según Helmholtz toda ciencia es medición, toda medición es Estadística. Esto ha facilitado a que esta ciencia en el siglo XXI muestre sus aplicaciones en diversos fenómenos ligada al desarrollo de software estadísticos más eficaces y eficientes.

Existen pocas áreas donde el impacto del desarrollo reciente de la estadística se haya hecho sentir más que en la ingeniería y en la administración industrial. En realidad sería muy difícil exagerar sus aportes a los problemas de producción, el uso eficiente de materiales y fuerza de trabajo, a la investigación básica y el desarrollo de nuevos productos. Como en las demás ciencias, la estadística ha venido a ser una herramienta vital para ingenieros. Les permite comprender fenómenos sujetos a variaciones y predecirlos o controlarlos eficazmente (Miller, et al., 1992).

Son disímiles los ejemplos donde el uso de métodos estadísticos en función del mejoramiento de la calidad ha hecho una revolución industrial y económica notable. Este puede ser el camino para mejorar los avances que ha hecho nuestro país en función de la economía. Ingenieros y científicos, armados no solo con sus conocimientos técnicos sino también con una buena base estadística, sería la clave para un avance vertiginoso en nuestra industria y en nuestra sociedad.

Por tal importancia en la Educación Cubana los contenidos relacionados con estadística no solo se quedan al nivel de la Educación Superior, en grados inferiores estos temas son tocados con menor intensidad pero con la intensidad que llegue a toda la sociedad.

El proceso de enseñanza y aprendizaje de la estadística no está exento de los mismos fenómenos que sufren todas las asignaturas de la disciplina de Matemáticas. Según (Barreto Villanueva, 2012) en su mayoría rechazada por los estudiantes pero reconocida y valorada por profesionales de diversas ramas. Dicho rechazo ligado en su mayoría por la falta de comprensión de la importancia de la estadística y sus aplicaciones. Donde esta parte es de total responsabilidad de los académicos que llevan tan importante tarea en sus hombros. Según (Johnson, 1996) “La estadística es el lenguaje universal de las ciencias”. Si los futuros ingenieros reciben la motivación correcta en cuanto a la gama de fenómenos que tienen solución con simple métodos estadísticos, pocos serían los problemas en la sociedad de manera global. En la Universidad de Matanzas se forman a la mayoría de los ingenieros que impulsan la economía de la provincia, dispersos en todos los sectores, desde una simple oficina de un registro civil hasta una industria tan grande como una termoeléctrica.

En diversas actividades de la vida humana percibir, interpretar y predecir los sucesos que se presentan depende de habilidades y destrezas para detectar relaciones entre los acontecimientos. Sin embargo, los estudios relacionados con el razonamiento covaracional sugieren pobres capacidades en los adultos para estimar la correlación, o realizar predicciones de una variable en función de otra, en ausencia de enseñanza específica sobre el tema. Asimismo se han descrito numerosas dificultades y errores conceptuales que continúan después de la enseñanza. Una explicación es que el razonamiento humano en situaciones de incertidumbre está regido por valores y creencias del propio individuo. Debido a ello, el procesamiento de la información difiere de un proceso algorítmico, que produce una solución única para cualquier problema, dentro de una clase dada (Gea, et al., 2014).

Esta situación plantea a los profesores el reto de mejorar la enseñanza de la correlación y regresión, que según (Gea, et al., 2018) matemáticamente recoge el estudio de la dependencia funcional a situaciones aleatorias.

## DESARROLLO

El estudio de la estadística es realmente necesario debido a que hoy en día esta rama de la matemática se ha convertido en un método efectivo para describir con gran margen de fiabilidad las tendencias y valores de datos económicos, políticos, sociales, psicológicos, biológicos y físicos, y sirve como herramienta para relacionar y analizar dichos datos. Es necesario destacar que junto a esta se encuentra la parte inferencial que nos permite describir los aspectos probabilísticos de la vida real, esto es, aquellos fenómenos no determinísticos en los que está presente de una u otra manera la utilización de valores al azar (Belfiori, 2014) (Nuñez, 2009).

En la mayoría de las investigaciones –sin importar el campo del conocimiento en las que se desarrollen– en las cuales se realicen mediciones, observaciones o experimentos de donde se obtengan datos de diferentes variables; es fundamental determinar algún tipo de relación de dependencia entre las variables con el fin de hacer predicciones o pronósticos de eventos futuros de acuerdo con el comportamiento de ellas. Algunos profesionales, sin importar su especialidad, confían en su intuición para juzgar como se relacionan dos variables. Debido a ello, hacen pronósticos a tientas e incluso temerarios; sin embargo, si dichos profesionales tienen la posibilidad de tomar datos y utilizar un procedimiento estadístico de análisis para determinar cómo lo conocido se relaciona con el evento futuro, podrían ayudar considerablemente en el mejoramiento de los procesos que administran o en la solución eficaz de los problemas que se les presentan. El procedimiento estadístico que se utiliza para este fin se conoce como análisis de regresión que permite establecer la relación funcional o ecuación matemática que relaciona las variables, así como la fuerza de esa relación (Cardona, et al., 2013).

O sea, en este sentido (Vargas) expresa que cuando se realizan estudios que involucran varias variables evaluadas simultáneamente para cumplir un objetivo específico, se puede

analizar la relación inherente a ellas. Definir relaciones que posibiliten predecir una o más variables en términos de otras es uno de los objetivos fundamentales de muchas investigaciones, lo cual se puede abordar a través de las técnicas estadísticas de modelos de regresión. Los modelos de regresión se usan para estimar "la mejor" relación funcional entre una variable dependiente y una o varias variables independientes, mientras que los métodos de correlación se utilizan para medir el grado de asociación de las distintas variables.

A pesar de lo importante que resulta ser para cualquier profesional el conocimiento y uso del análisis de regresión, (Cardona, et al., 2013) expresan que es una herramienta muy poco aprovechada como lo demuestran un gran número de trabajos de grado, a nivel de posgrado y trabajos de investigación en los cuales el desarrollo estadístico solo se limita a la parte descriptiva y no a la inferencial. Esto hace necesario que se requiera profundizar en métodos correctos de enseñanza y aprendizaje del tema; promoviendo la motivación hacia la aplicación de estas técnicas en respuesta a muchos fenómenos de la vida práctica.

### **Ideas fundamentales en la enseñanza y aprendizaje de la correlación y regresión.**

Según (Burrill y Biehler, 2011) para un adecuado sentido estadístico se requiere del conocimiento de las ideas estadísticas fundamentales, que para el caso de correlación y regresión se analiza y desglosa algunas de ellas, según experiencias profesionales en la universidad cubana:

**Tablas y Gráfica:** Por su papel esencial en la organización, descripción y análisis de datos, las tablas y gráficos son un instrumento esencial en el análisis estadístico, y su conocimiento es parte de la cultura estadística así lo expresan (Arteaga, et al., 2011). Por lo tanto es necesario que el estudiante grafique ya que brinda una buena interpretación de la relación de las variables y es un método de comprobación para los resultados de los coeficientes y parámetros necesarios. La gráfica, que no siempre tiene que ser un diagrama de dispersión, permite visualizar si existe correlación, el tipo de dicha correlación (directa o inversa) y hasta cierta medida el grado de la misma según la mayor o menor dispersión de la nube de puntos. Se debe de reconocer además la organización de los datos a partir de tablas incluso para el cálculo más eficiente de los términos de las fórmulas.

**Supuestos:** Es fundamental además la demostración de los supuestos con su vinculación a temas anteriores y que a su vez hacen de un problema de correlación y regresión un caso de estudio.

**Fórmulas y Notación:** El desglose de las fórmulas debe partir de su concepto teórico relacionando cada uno de los términos del tema. Por ejemplo el cálculo del coeficiente de correlación a partir de la covarianza y el cálculo de la pendiente a partir del coeficiente de correlación. Una de las dificultades en la enseñanza y aprendizaje del tema de correlación y regresión como de otros temas de estadística es el poco uso del libro básico de la asignatura que por parte de los mismos profesores le intuimos a los estudiantes. En clases se debe de

ganar mejor acercamiento a la notación que usa el libro, para que de esa forma el estudiante tenga más independencia en su autoestudio.

**Coefficiente de Determinación:** No debe de faltar junto al coeficiente de correlación lineal la interpretación y cálculo del coeficiente de determinación; no solo dentro del tema de correlación sino también con el tema de regresión con las sumas de cuadrados. (Gea, et al., 2014) plantean que el coeficiente de determinación, además de medir la bondad del ajuste de los datos al modelo, también puede interpretarse como una medida de variabilidad: La proporción de la varianza de la variable dependiente Y explicada por el modelo de regresión.

**Estimación por intervalos:** La correcta interpretación del intervalo de predicción de la variable dependiente y el intervalo de confianza de dicha variable es fundamental y se le brinda poca importancia donde es fundamental que el estudiante entienda que en estadística a cada valor de la variable independiente le corresponde una distribución de valores de la variable dependiente.

**Ecuación de Regresión Estimada:** No se puede perder el sentido de la regresión donde consiste en tratar de encontrar una función matemática que relacione las variables X e Y de una distribución bidimensional, de forma que, si se conoce el valor de una variable, se puede calcular el correspondiente de la otra, con mayor o menor aproximación, así lo afirman (Monteagudo y Paz, 2008).

**Clasificación de las variables:** (Gea, et al., 2014) muestran que la diferencia entre variable dependiente e independiente ocupa un lugar central en el análisis de la regresión, ya que, una vez aceptada la dependencia entre las variables del estudio, y con objeto de expresar en forma de ecuación o modelo una variable en función de otra, se necesita seleccionar qué variable servirá como dependiente. La definición explícita de variable dependiente e independiente se encuentra en pocos textos. Según (Vizmanos, et al., 2008) la variable dependiente es aquella que se quiere estimar, y la variable que se utiliza para ello se denomina variable independiente. No obstante, es necesario que el estudiante conozca que en caso de querer cambiar la predicción e invertir la clasificación que realizó inicialmente de las variables, se hace explícita la existencia de dos rectas de regresión diferentes:

Si X se considera la variable independiente e Y la variable dependiente, la ecuación de la recta de regresión es:

$$y - \bar{y} = \frac{S_{xy}}{S_x^2} (x - \bar{x})$$

Esta recta se denomina recta de regresión de Y sobre X. A partir de ella, conocidos los valores de X, y sustituyéndolos en la ecuación, se pueden

calcular con una cierta aproximación los valores de Y.

Si se considera Y como variable independiente y X como variable dependiente, se obtiene la recta de regresión de X sobre Y, cuya ecuación es:

$$x - \bar{x} = \frac{S_{xy}}{S_y^2} (y - \bar{y})$$

Igual que en el caso anterior, conocidos los valores de Y, y sustituyéndolos en la ecuación de la recta, se obtienen con cierta aproximación los valores de X (Monteagudo y Paz, 2008).

Respuestas: Junto con la interpretación estadística de cada contenido es de suma importancia las respuestas prácticas, dichas respuestas siempre en función de la situación problémica que se está analizando. Esto pedagógicamente da una visión de la aplicación del tema en la vida práctica; no solo son procedimientos que responden a fórmulas y pasos sino a dar respuestas objetivas y claras sobre el fenómeno en cuestión.

Por otro lado es vital en este tema la posibilidad de trabajar con proyectos, en que el estudiante parte de un problema de investigación y completa todos los pasos de una investigación (Pratt, et al., 2011). Este tipo de trabajo fomenta los componentes del razonamiento estadístico: reconocimiento de la necesidad de datos, percepción de la variación, transnumeración, uso de modelos estadísticos e integración de la estadística y el contexto (Wild y Pfannkuch, 1999). De esta forma, si se trabaja con datos reales se potencia la interdisciplinariedad en las clases de estadística, permitiendo aprender contenidos que no se adquieren habitualmente con problemas tomados de los libros de texto (Hall, 2011).

Además, la introducción de resultados investigativos obtenidos de la aplicación correcta de los modelos estadísticos ayuda considerablemente al desarrollo del proceso docente, tanto en pregrado como en postgrado, siendo además una herramienta de motivación que muestra al estudiante como proceder de forma correcta en las investigaciones relacionadas con su especialidad que requieran del uso de técnicas estadísticas (Fraga y Brito, 2006).

(Gea, et al., 2018) afirman que un recurso para ayudar a los estudiantes a superar muchas de estas dificultades es la tecnología. Hoy en día, existe una gran variedad de recursos tecnológicos como la calculadora, hoja de cálculo, applets y programas de ordenadores y teléfonos específicos, que pueden facilitar la realización de cálculos y gráficos. Además, proporciona al estudiante recursos de visualización y exploración de las ideas estadísticas.

### **Las TIC en la enseñanza y aprendizaje de la estadística.**

Según (Mason, 1998), no se inventan nuevas metodologías, sino que la utilización de las Tecnologías de la Información y las Comunicaciones (TIC) en la educación abre nuevas perspectivas respecto a una enseñanza mejor.

Los rápidos cambios, el aumento de los conocimientos y las demandas de una educación de alto nivel constantemente actualizada se convierten en una exigencia permanente (Uriza, 2015). Algunas universidades valoran que esta situación ha sido la oportunidad perfecta para situar el uso de las TIC como un elemento diferencial de valor añadido, factor clave para la innovación y la mejora de la calidad educativa en sus instituciones (Sangrá, 2002).

La educación superior no se queda atrás en el proceso de integración de las TIC y en pos de garantizar la calidad de la docencia requiere la transformación de las estructuras institucionales hacia modelos de la sociedad del conocimiento, hacer uso eficiente y apropiado de las tecnologías facilitando la docencia y la investigación y promover la alfabetización digital para el desarrollo de la disciplina y la profesión (Durall, 2012).

A su vez, (Cabrera, 2008) plantea que la integración de las TIC es un proceso contextualizado, sistémico, continuo y reflexivo, orientado a la transformación de la práctica pedagógica tomando en cuenta las posibilidades de las mismas con la finalidad de incorporarlas armónicamente al proceso de enseñanza y aprendizaje, para satisfacer los objetivos educativos.

La reducción del tiempo de cálculo y la ampliación del tipo de gráficos que el estudiante puede realizar interactivamente es según (Pratt, et al., 2011) una de las aplicaciones de la tecnología en la enseñanza de la estadística. Igualmente, al facilitar el cálculo y la representación gráfica, la tecnología disminuye el problema tradicional en la enseñanza de la estadística en cuanto al desfase entre la comprensión de los conceptos y los medios técnicos de cálculo para poder aplicarlos (Batanero y Borovcnik, 2016).

Otra aplicación, característica del uso de las TIC, es la simulación y visualización. Las representaciones icónicas de conceptos (por ejemplo, la desviación de los puntos de un diagrama de dispersión a la recta de regresión) pueden ayudar a los estudiantes a pensar a través de modelos concretos, cuando todavía no son capaces de generalizar sus ideas matemáticas. Por otro lado, permiten explorar objetos abstractos, creando micromundos virtuales donde los estudiantes pueden experimentar con las diferentes variables que intervienen. Una simulación es ya un modelo matemático (al haber simplificado la realidad), pero es menos abstracto que el modelo matemático y constituye un puente entre éste y la realidad (Gea, et al., 2018).

La correlación y la regresión no se queda atrás en el uso de las TIC y como en la estadística en general el propósito fundamental debe ser encaminado a la utilización de software estadísticos cada vez más eficientes. El estudiante fuera de conocer todos los conceptos y terminologías referentes a estas técnicas debe de ser de vital importancia el fomento de su habilidad en el uso de software y en la interpretación de todas sus salidas. En la práctica el

cálculo que se realiza en clases sería totalmente engorroso e innecesario, dado que para lograr mejor predicción del modelo de regresión es preciso que el tamaño de la muestra sea significativamente mayor que el de los pequeños ejercicios que se resuelven en clases.

Se debe de cambiar el esquema de prioridad de las habilidades donde la memorización de las grandes fórmulas estadísticas no debe de ser tan importante, sino el manejo de software para la solución y la interpretación de los resultados. Es importante además, potenciar el uso de las herramientas informáticas que más se utilizan en la práctica por profesionales de la rama a la que se le imparte el tema. Se ha ganado mucho en el enfoque de los ejercicios en función de la especialidad pero hay que continuar trabajando en las motivaciones a partir de las reales aplicaciones de las técnicas en la vida profesional.

### **La correlación y regresión en la ingeniería.**

Dado el desarrollo acelerado de la Ciencia y la Técnica en estos tiempos, cada día crecen de forma continua las investigaciones en la rama de las ingenierías y con ello los problemas en donde es necesaria la utilización de técnicas estadísticas para el análisis y obtención de resultados en forma racional. Por consiguiente, los futuros ingenieros y científicos deben tener una amplia y adecuada preparación en esta ciencia, que comprende además de los conocimientos teóricos que abarca esta rama de la matemática, una correcta interpretación de los resultados de la investigación y la utilización de las herramientas de cómputo tan eficaces en la actualidad (Fraga y Brito, 2006).

La introducción de los laboratorios con el uso de programas estadísticos como Statgraphics y la posibilidad de interacción constante entre los propios estudiantes, ha propiciado un reforzamiento en el trabajo en equipo, lo cual es muy importante en las diferentes disciplinas. Además, los estudiantes emprenden una búsqueda y replanteamiento continuo de contenidos y procedimientos, aumentando la implicación en sus tareas y el desarrollo de iniciativas.

El correcto enfoque de la ciencia, la tecnología, y la innovación como papel protagonista en el crecimiento económico, en la productividad, en la competitividad, en el desarrollo sostenible y en el mejoramiento de la calidad de vida de una sociedad, es fundamental para la orientación de las investigaciones, de acuerdo con las necesidades y capacidades propias de cada región, de tal manera que se desarrollen, fortalezcan y mantengan las capacidades científicas adquiridas (Alcázar y Lozano, 2009).

Desde esa perspectiva es de vital importancia fomentar la correcta enseñanza y aprendizaje de la estadística en todas las ingenierías, enfocándonos principalmente en la correlación y regresión por su notable aplicación en la práctica.

### **Impacto de la enseñanza y aprendizaje de la correlación y regresión en la economía y la sociedad.**

Sin una correcta medición, no se puede planificar, prever e innovar. Por ello, el siguiente estudio está encaminado a demostrar la importancia de la enseñanza y aprendizaje de la correlación y regresión para una sociedad en constantes cambios y avances en su economía y sociedad. Asimismo, se enfatiza sobre la importancia y necesidad de medir el impacto económico y social generado por la instrucción de la estadística en la formación de los ingenieros de nuestro país.

**Caso de Estudio:** Se lleva a cabo un estudio en 10 empresas de la provincia de Matanzas con el propósito de entrevistar a los ingenieros que laboran en ellas; aplicando la siguiente encuesta de forma anónima:

**¿Qué conocimientos tienes de Correlación y Regresión?**

\_\_\_\_\_ **Mal**                      \_\_\_\_\_ **Regular**                      \_\_\_\_\_ **Bien**

**Especialidad:** \_\_\_\_\_

Las empresas que se tomaron al azar fueron las siguientes: Labiofam Matanzas, Termoeléctrica “Antonio Guiteras Holmes”, Empresa Eléctrica Provincial de Matanzas, Empresa de Conformación de Cubos “Noel Fernández”, Empresa Ferroviaria “José Valdés Reyes”, EPEP-C, Hotel Melia Varadero, ECOA-47, Instituto provincial de suelos y Empresa Recuperadora de Equipos Automotrices “Camilo Cienfuegos”.

En la siguiente tabla se muestra la recogida de datos de las encuestas ordenadas por especialidad para un total de 109 encuestas válidas:

**Especialidades**

<b>Industrial</b>	B	B	R	B	B	M	B	B	R	M	B	R	B	B	B	B
<b>Informática</b>	R	M	B	M	M	R	B	M	M	M	B					
<b>Mecánica</b>	B	M	B	M	B	R	R	M	M	M	R	B	R	B		
<b>Agronomía</b>	B	R	M	M	M	M	R	M	B							
<b>Civil</b>	B	B	M	B	M	B	R	R	M	M	R					
<b>Automática</b>	B	M	R	M	B	R	M	M	M	M						
<b>Química</b>	B	M	M	M	B	R	B	R	R	R	M	B				
<b>Eléctrica</b>	B	M	M	M	R	R	M	M	M	B	M					
<b>Ciencias Informáticas</b>	M	B	R	M	R	B	M	R	M	M	R	M	R	M	B	

*Primer Diseño Experimental:* Se realiza un diseño completamente al azar en función de demostrar que los conocimientos del tema en cuestión dependen del programa de estudio de cada carrera de ingeniería.

- VAO.  $Y_{ij}$  ~ índice de conocimientos de los ingenieros del tema de correlación y regresión.
- Factor Docimado: Especialidad.
- Niveles: Industrial, Informática, Mecánica, Agronomía, Civil, Automática, Química, Eléctrica, Ciencias Informática.
- Método de solución: Análisis de Varianza. Método Kruskal-Wallis

Kruskal-Wallis Test for Conocimiento by Especialidad

Especialidad	Sample Size	Average Rank
Agronomía	9	46.7222
Automática	10	44.4
C_Informática	15	49.2667
Civil	11	58.1818
Eléctrica	11	42.5
Industrial	16	77.0938
Informática	11	48.6818
Mecánica	14	58.2143
Química	12	58.3333

Test statistic = 14.423    P-Value = 0.0713841

The StatAdvisor

The Kruskal-Wallis test tests the null hypothesis that the medians of Conocimiento within each of the 9 levels of Especialidad are the same. The data from all the levels is first combined and ranked from smallest to largest. The average rank is then computed for the data at each level. Since the P-value is greater than or equal to 0.05, there is not a statistically significant difference amongst the medians at the 95.0% confidence level.

Fig. 1. Salida del software. Kruskal-Wallis.

- Respuesta: Según el análisis del experimento la salida del software Statgraphics muestra que no hay diferencias significativas ( $P\text{-Value} > 0.05$ ) entre las medias de las especialidades con un nivel de confianza del 95%, o sea, el tipo de especialidad con su respectivo plan de estudio no influye significativamente en el conocimiento que tienen los ingenieros sobre el tema en cuestión.

*Segundo Diseño Experimental:* Este diseño está encaminado a medir de alguna forma el impacto económico y social que tiene la correcta enseñanza y aprendizaje del tema de correlación y regresión para los ingenieros. Para ello valoramos la posible influencia de las buenas prácticas de correlación y regresión en la solución al banco de problemas de las empresas.

- Variable Dependiente ( $Y_i$ ) ~ porcentaje de solución al banco de problemas de la empresa  $i$  a partir de las técnicas de correlación y regresión (%).

- Variable Independiente ( $X_i$ ) ~ cantidad de ingenieros con nivel avanzado de la empresa  $i$  en el tema de correlación y regresión.
- Método de Solución: Se aplicaron las técnicas de correlación y regresión lineal.

Datos:

$Y_i$	44	49.5	9.1	0	3.2	12	46.8	10.3	40	4.2
$X_i$	6	5	2	3	1	2	7	1	5	3

Salidas del Software (Solución):

```

Regression Analysis - Linear model: Y = a + b*X
-----
Dependent variable: porcentaje
Independent variable: ingenieros
-----
Parameter      Estimate      Standard      T
                Error              Statistic      P-Value
-----
Intercept      -7.52025     6.62803      -1.13461     0.2894
Slope          8.40864     1.64169       5.12195     0.0009
-----

                        Analysis of Variance
-----
Source          Sum of Squares  Df  Mean Square  F-Ratio  P-Value
-----
Model           2863.56         1   2863.56     26.23    0.0009
Residual       873.226         8   109.153
-----
Total (Corr.)  3736.79         9
-----
Correlation Coefficient = 0.875395
R-squared = 76.6316 percent
R-squared (adjusted for d.f.) = 73.7106 percent
Standard Error of Est. = 10.4476
Mean absolute error = 7.18973
Durbin-Watson statistic = 2.00869 (P=0.4698)
Lag 1 residual autocorrelation = -0.109442
The StatAdvisor
-----
The output shows the results of fitting a linear model to describe
the relationship between porcentaje and ingenieros. The equation of
the fitted model is

porcentaje = -7.52025 + 8.40864*ingenieros

Since the P-value in the ANOVA table is less than 0.01, there is a
statistically significant relationship between porcentaje and
ingenieros at the 99% confidence level.

The R-Squared statistic indicates that the model as fitted explains
76.6316% of the variability in porcentaje. The correlation
coefficient equals 0.875395, indicating a moderately strong
relationship between the variables. The standard error of the
estimate shows the standard deviation of the residuals to be 10.4476.
This value can be used to construct prediction limits for new
observations by selecting the Forecasts option from the text menu.

The mean absolute error (MAE) of 7.18973 is the average value of
the residuals. The Durbin-Watson (DW) statistic tests the residuals
to determine if there is any significant correlation based on the
order in which they occur in your data file. Since the P-value is
greater than 0.05, there is no indication of serial autocorrelation in
the residuals.

```

Fig. 2. Salida del software. Correlación y regresión lineal simple.

- Respuestas: Después de aplicadas las técnicas podemos concluir que ambas variables tienen una relación directa significativa donde el 76.6316% de la variabilidad total de la variable dependiente se explica por la influencia de los valores de la variable independiente. Además, se puede predecir el porcentaje de solución del banco de problemas de las empresas a partir del conocimiento que tienen los ingenieros contratados del tema de correlación y regresión ( $P\_Value < 0.01$ ) utilizando la ecuación de regresión estimada:

$$\hat{y} = -7.52025 + 8.40864x.$$

Este estudio resulta significativo para la toma de decisiones en las empresas en cuanto a la contratación de su personal, siendo el conocimiento de la correlación y regresión un parámetro notable. Al influir estos conocimientos en la solución de muchos fenómenos del banco de problemas de las empresas impacta de forma satisfactoria tanto en la economía como en la sociedad.

## CONCLUSIONES

El trabajo con proyectos, la introducción de resultados investigativos obtenidos de la aplicación correcta de los modelos estadísticos y el uso de las tecnologías son elementos claves para motivar a los estudiantes en el campo de la Estadística, dentro de esta la correlación y regresión. Donde este tema por su impacto directo en la economía y la sociedad, como se demostró en los casos de estudio, es necesario que sea recibido por los ingenieros de manera amplia y con obligatoria vinculación con las demás disciplinas del plan de estudio que reciben los mismos.

## Bibliografía

ALCÁZAR, E. AND LOZANO, A. *Desarrollo histórico de los indicadores de Ciencia y Tecnología, avances en América Latina y México*. s.l.: Revista Española de Documentación Científica, 2009. pp. 119-126. Vol. 32 (3).

ARTEAGA, P., ET AL. *Las tablas y gráficos estadísticos como objetos culturales*. s.l.: NÚMEROS: Revista de Didáctica de las Matemáticas, 2011. pp. 55-67. Vol. 76.

BARRETO VILLANUEVA, A. *El progreso de la Estadística y su utilidad en la evaluación del desarrollo*. s.l.: Papeles de Población, 2012. pp. 1-31. Vol. 18 (73).

BATANERO, C. AND BOROVCNIK, M. *Statistics and probability in high school*. s.l.: Rotterdam: Sense Publishers, 2016.

BELFIORI, L. V. *Enseñanza de estadística con recursos TIC*. Buenos Aires, Argentina : Congreso Iberoamericano de Ciencia, Tecnología, Innovación y Educación, 2014.

BURRILL, G. AND BIEHLER, R. *Fundamental statistical ideas in the school curriculum and in training teachers*. 2011. pp. 57-69.

CABRERA, J. *Modelo de Centro Virtual de Recursos para contribuir a la integración de las TIC en el proceso de enseñanza-aprendizaje*. La Habana, Cuba : Instituto Superior Politécnico José Antonio Echeverría, 2008. Doctorado.

CARDONA, D. F., ET AL. *Aplicación de la regresión lineal en un problema de pobreza*. [ed.] Universidad Libre. Facultad de Ciencias de la Educación. Colombia : Interacción, 2013. pp. 73-84. Vol. 12.

DURALL, E. G. *Perspectivas tecnológicas: educación superior en Iberoamérica*. Austin, Texas : The New Media Consortium, 2012.

FRAGA, E. AND BRITO, M. L. *Papel de la estadística en la formación del ingeniero mecánico*. Ciudad de la Habana, Cuba : Instituto Superior Politécnico “José Antonio Echeverría”, Facultad de Ingeniería Mecánica, Departamento de Matemática, 2006.

GEA, M. M., BATANERO, C. AND ROA, R. *El sentido de la correlación y regresión*. s.l.: NÚMEROS: Revista de Didáctica de las Matemáticas, 2014. pp. 25-35. Vol. 87.

GEA, M. M., ET AL. *Conocimiento Tecnológico sobre la Correlación y Regresión: un estudio exploratorio con Futuros Profesores*. [ed.] Universidad de Granada. s.l.: Bolema, 2018. pp. 134 - 155. Vol. 32 (60).

HALL, J. *Engaging teachers and students with real data: Benefits and challenges*. New York : Springer, 2011.

- JOHNSON, R. *Elementary Statistics*. 7th. USA : s.n., 1996.
- MASON, R. *Models of online courses*. s.l. : ALN Magazine, 1998. Vol. 2.
- MILLER, I. R., FREUND, J. E. AND JOHNSON, R. *Probabilidad y Estadísticas para Ingenieros*. México : s.n., 1992. Vol. I.
- MONTEAGUDO, M. F. AND PAZ, J. *1º Bachillerato. Matemáticas. Ciencias y tecnología*. Zaragoza : Edelvives Editorial Luis Vives, 2008. Vol. Parte 2.
- . *1º Bachillerato. Matemáticas. Ciencias y tecnología*. Zaragoza : Edelvives Editorial Luis Vives, 2008. Vol. Parte1.
- NUÑEZ, A. *Recursos TIC para el bloque de estadística y probabilidad en la ESO y Bachillerato*. Granada, España : s.n., 2009.
- PRATT, D., DAVIES, N. AND CONNOR, D. *The role of technology in teaching and learning statistics*. New York, USA : Springer, 2011. pp. 97-107.
- SANGRÁ, A. *Educación a distancia, educación presencial y usos de la tecnología: una tríada para el progreso educativo*. España : Universitat Oberta de Catalunya (UOC), 2002.
- URIZA, T. *Las Tics. Diapositivas de Investigación*. 2015. Disponible en: <http://www.slideshare.net/theresithaza/diapositivas-de-las-tics-teresa-uriza-riobueno>.
- VARGAS, V. Capítulo 5. Modelos de Regresión. *Estadística Descriptiva para Ingeniería Ambiental con SPSS*. pp. 123-156.
- VIZMANOS, J. R., ET AL. *Matemáticas aplicadas a las Ciencias Sociales*. Madrid, España : Madrid: SM, 2008. Vol. I.
- WILD, C. J. AND PFANNKUCH, M. *Statistical thinking in empirical enquiry*. s.l. : Voorburg, 1999. pp. 223-265. Vol. 67(3).