

**DANIEL HERNANDEZ TAPIA**

**MODELO ECONOMETRICO PARA EL ESTUDIO DE DRIVERS QUE IMPACTEN EN LA SATISFACCION DE CLIENTES EN UN PROCESO DE RECLAMOS**

# **SOLUCION PROBLEMA 1**

## Introducción

El análisis de regresión es una técnica estadística para investigar y modelar la relación entre variables. Hay un sinnúmero de aplicaciones de la regresión, y las hay en cualquier campo (ingeniería, ciencias físicas y químicas, entre otras).

Una de las aplicaciones que vamos a realizar es el estudio de una base de datos que contiene variables que miden la satisfacción de los clientes a lo largo de los procesos de reclamos.

En este trabajo vamos a construir el modelo, ver la significancia del modelo tanto global como individualmente, realizar artificios para que cumplan con los supuestos del modelo si es que no lo cumple y finalmente validarlo. Además, realizaremos el análisis de residuos para detectar puntos atípicos.

## Objetivo

Mediante una serie de criterios, averiguar qué variables influyen significativamente en cuanto a la satisfacción del cliente y ver qué medidas tomar para que cumpla con los supuestos de normalidad, homocedasticidad y multicolinealidad.

## Planteamiento del problema

Se necesita estudiar todo el proceso de reclamo de los clientes para mejorar en los servicios que se brinda.

Se necesita saber cuál de estas variables son las que influyen en la satisfacción del cliente.

IR01= "AMABILIDAD AL MOMENTO DE INGRESAR RECLAMOS"

IR02= "INTERES DEL FUNCIONARIO AL MOMENTO DE INGRESAR EL RECLAMO"

IR03= "RAPIDEZ DEL FUNCIONARIO AL MOMENTO DE INGRESAR EL RECLAMO

IR04= "FACILIDAD Y SIMPLEZA DEL FUNCIONARIO AL MOMENTO DE INGRESAR EL RECLAMO

IR05="NIVEL DE CONOCIMEINTOS DEL FUNCIONARIO AL MOMENTO DE INGRESAR EL RELCAMO"

IR06="CANTIDAD DE INFORMACION BRINDADA AL MOMENTO DE INGRESAR EL RECLAMO"

IR07= "TIEMPO DE ESPERA PARA SER ATENDIDO AL MOMENTO DE INGRESAR EL RECLAMO"

IR08= "TIEMPO DE ATENCION AL MOMENTO DE INGRESAR EL RECLAMO"

RR01= "CANTIDAD DE INFORMACION Y SUSTENTO EN LA RESPUESTA DEL RECLAMO

RR02= "RAPIDEZ EN LA RESPUESTA DEL RECLAMO"

RR03= "AMABILIDAD MOSTRADA EN LA RESPUESTA DEL RECLAMO"

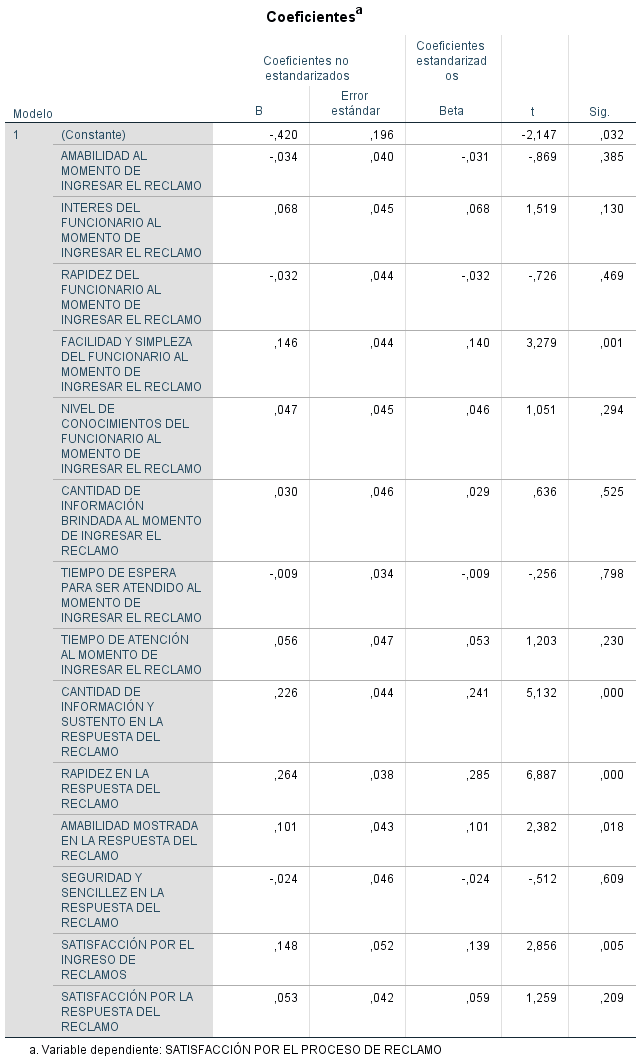
RR04= "SEGURIDAD Y SENCILLEZ EN LA RESPUESTA DEL RECLAMO"

SAT\_1= "SATISFACCION POR EL INGRESO DE RECLAMOS"

SAT\_2= "SATISFACCION POR LA RESPUESTA DEL RECLAMO"

Y= SAT= "SATISFACCION POR EL PROCESO DE RECLAMO"

En primer lugar, se estiman los parámetros por el método de MCO para todas las variables independientes.



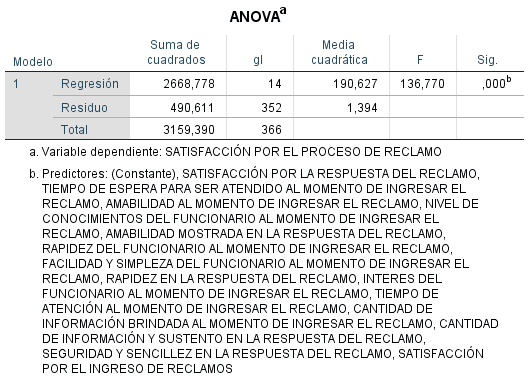
En el cuadro se aprecia que las variables que son significativas son IR04, RR01, RR02, RR03 Y SAT\_1

Una vez estimados los parámetros del modelo, nos preguntamos cual es la adecuación general del modelo y cuáles son los regresoras específicos que parecieran importantes, para ello hacemos la prueba de significancia de la regresión.

Significancia de la regresión:

Ho:

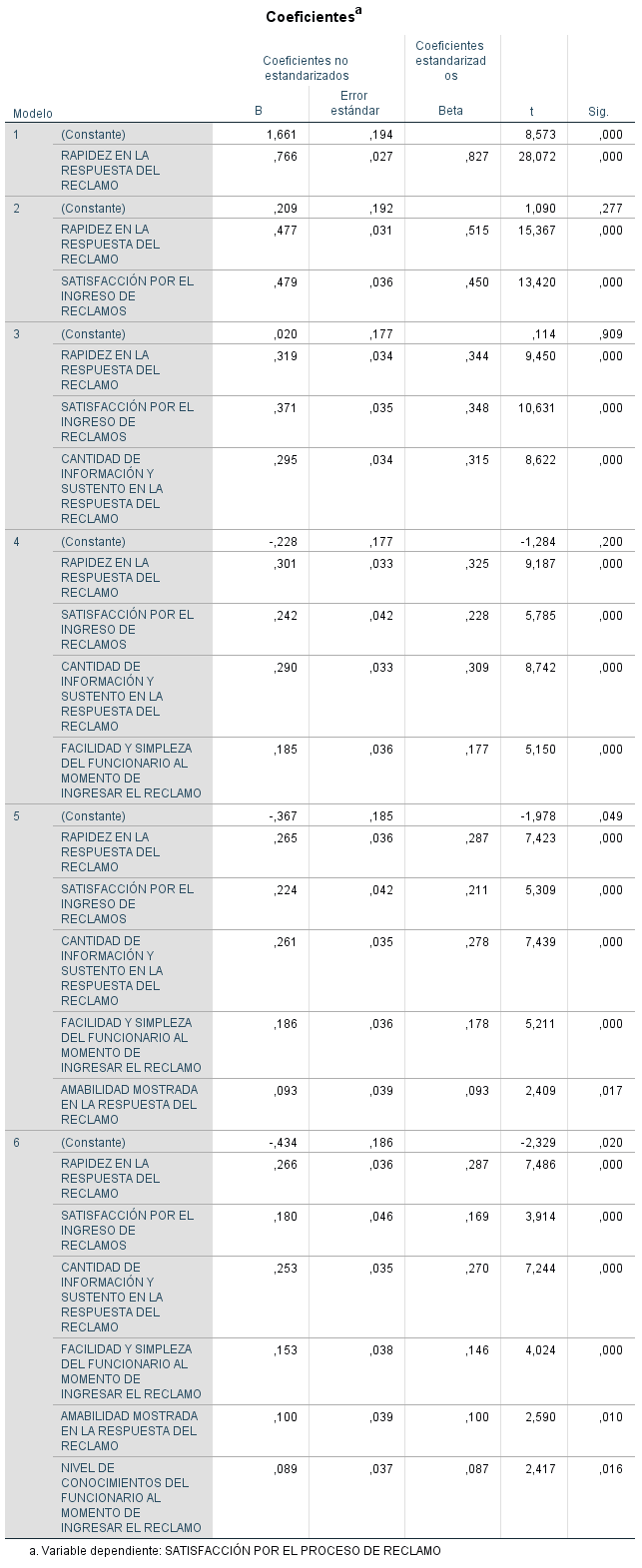
Ha: al menos un j



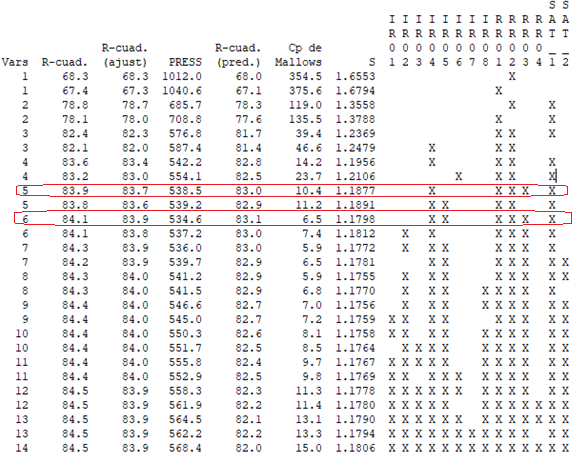
Una vez que determinamos que al menos uno de los regresoras es importante (REGRESION SIGNIFICATIVA), la pregunta lógica es ¿Cuál de todas ellas es?

1. Ingresaremos las variables con el método de pasos sucesivos

Con el SPSS, aplicamos una regresión con selección de variables mediante pasos sucesivos



Al comparar con el modelo anterior, en este método que utiliza el método de stepwise, solo escoge 6 variables de las 14 candidatas lo que nos indica que este puede ser el modelo con el que trabajaremos, sin embargo, validaremos con el método de todas las regresiones posibles:



Se puede apreciar que entre los subconjuntos candidatos se encuentra nuestro modelo, pero tenemos que escoger aquel que tenga el menor S y menor Cp Mallows en este caso nos quedamos con el método de pasos sucesivos que hicimos en el paso anterior, por tener entre los menores valores de S y Cp, lo cual valida nuestro procedimiento y tenemos suficiente evidencia estadística para decir que el mejor modelo escogido es el siguiente:

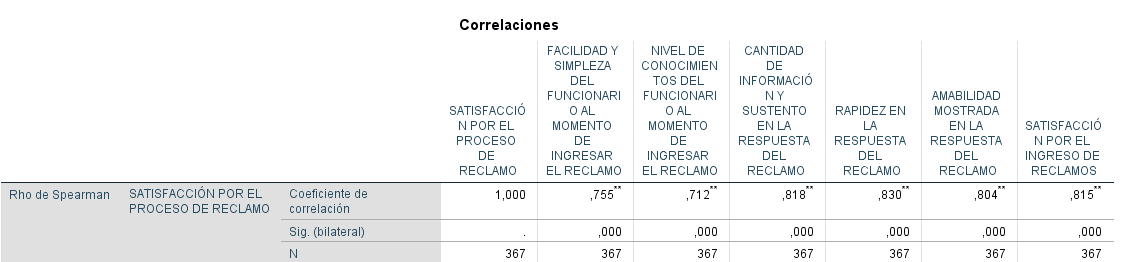
**SAT = -0.434 + 0.1525 IR04 + 0.0894 IR05 + 0.2531 RR01 + 0.2660 RR02 + 0.0998 RR03 + 0.1797 SAT\_1**

## Linealidad

Para la demostración de la linealidad tendremos que comprobar si cada una de las variables con la variable dependiente tiene comportamiento lineal, con el método gráfico.



Ahora lo analizaremos de manera analítica.



Podemos ver que la variable dependiente tiene relación lineal con cada variable independiente, esto se ha confirmado por la correlación de Spearman.

## Normalidad

Luego, haremos la prueba de Anderson Darling para demostrar la normalidad en el modelo.

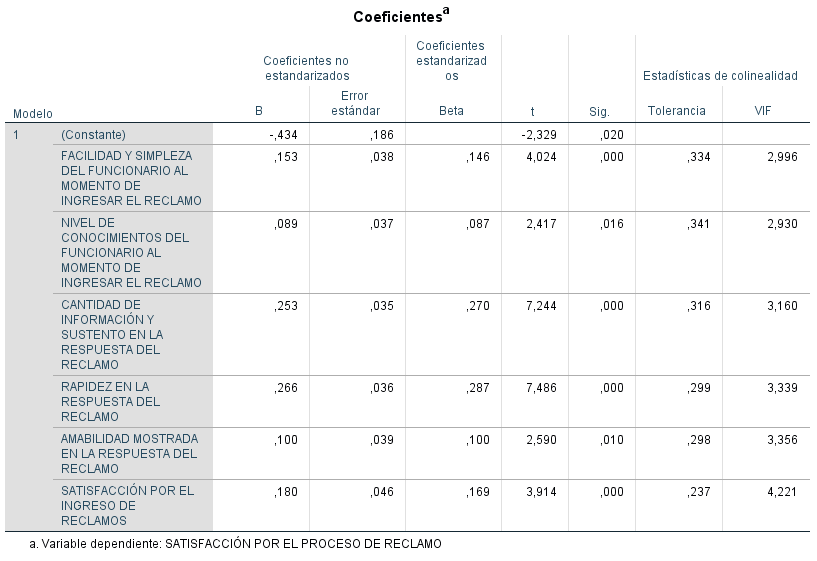
Ho: Los e son normales

Ha: Los e no son normales



Concluimos que existe suficiente evidencia estadística para rechazar la Hipótesis nula, es decir los residuos no presentan normalidad ya que el valor de p es <0,005.

## Multicolinealidad



Se puede apreciar que en todas las variables regresoras el VIF es menor que 5 lo que indica que no hay problema de multicolinealidad.

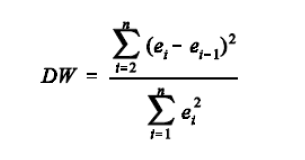
## Independencia

La independencia, está relacionado con la aleatoriedad de los errores

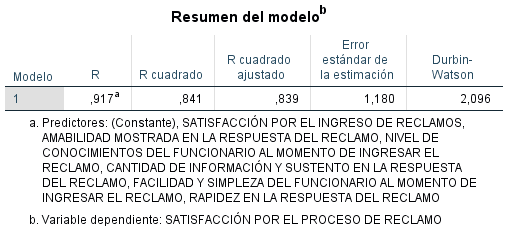
Este supuesto se usa sobre todo cuando se tiene series temporales, un estadístico para este fin es Durbin Watson que proporciona información sobre el grado de independencia existente entre los errores, sin embargo, en este caso lo utilizaremos para nuestro ejemplo de motores:

Ho: p = 0

Ha: p

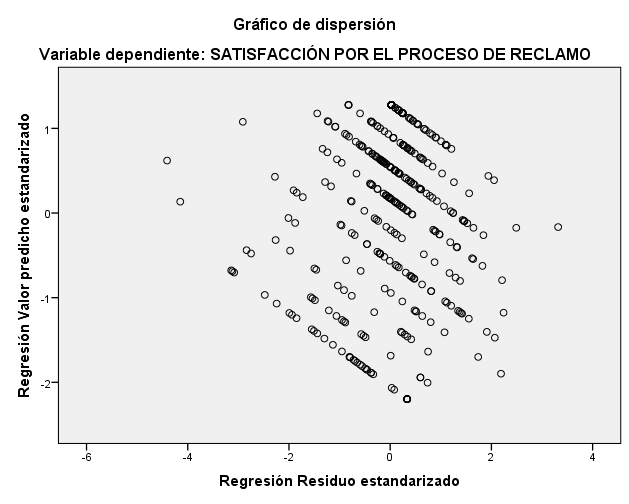


Su valor va de cero a cuatro, se asume independencia si el estadístico está entre 1.5 y 2.5.

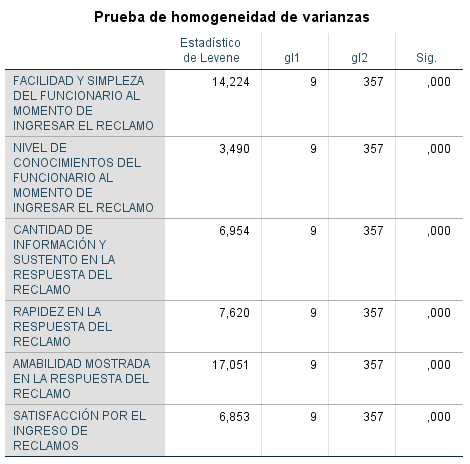


En este caso como el estadístico de Durbin Watson es 2.096 se puede decir que hay independencia.

## Homocedasticidad



Gráficamente se observa que no existe homocedasticidad debido a que los residuales al parecer siguen un patrón gráfico de comportamiento. Por tanto, analizaremos mediante la prueba de Levene para tener más certeza.



La significancia en todos los casos es menor a 0.05, por tanto, se rechaza la hipótesis nula de homocedasticidad.

## ANALISIS RESIDUAL

Por SPSS se ha obtenido los indicadores de DCOOK, LEVERAGE, DDFITS y COVRATIO con el fin de analizar la influencia de los residuos. Luego del análisis, obtenemos los siguientes gráficos con los límites de cada indicador:

Notamos en las diferentes graficas los valores que escapan de los límites, por tanto, podemos ver qué puntos son. Ver el archivo Excel (ANALISIS RESIDUAL) adjunto.

## Conclusiones

1.- Construimos un modelo con Stepwise y lo comparamos con todos los posibles modelos guiándonos con el Cp de Mallows. Además, vimos otros indicadores como el R^2, la S y la significancia de las variables y hemos resultado con un modelo de 6 variables. Ver indicadores del modelo más arriba.

2.- Analizando la linealidad de las variables exógenas, todas ellas presentan fuerte correlación positiva con la variable de estudio.

3.- Analizando el supuesto de normalidad con el estadístico Anderson-Darling podemos concluir que los residuos no presentan normalidad.

4.- Analizando la multicolinealidad, se logra ver en forma analítica que dicho problema no existe en este modelo.

5.- Analizando la independencia de las variables, con el estadístico Durbin Watson podemos concluir que, efectivamente, presentan independencia.

6.- Analizando el supuesto de homocedasticidad podemos concluir que, en el modelo no existe homocedasticidad.

## RECOMENDACIONES

1. La entidad financiera debe prestar más atención en las variables de FACILIDAD Y SIMPLEZA DEL FUNCIONARIO AL MOMENTO DE INGRESAR EL RECLAMO, NIVEL DE CONOCIMIENTOS DEL FUNCIONARIO AL MOMENTO DE INGRESAR EL RELCAMO, CANTIDAD DE INFORMACION Y SUSTENTO EN LA RESPUESTA DEL RECLAMO, RAPIDEZ EN LA RESPUESTA DEL RECLAMO, AMABILIDAD MOSTRADA EN LA RESPUESTA DEL RECLAMO y SATISFACCION POR EL INGRESO DE RECLAMOS.
2. Podemos ver que las variables están relacionadas con en primera instancia con el funcionario, por tanto, la entidad debe capacitar a esta persona en su forma de atención y el grado de conocimientos que este debe tener para guiar al cliente de una forma adecuada a lo largo del proceso de reclamo.