

**UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS**

**FACULTAD DE CIENCIAS MATEMÁTICAS**

**ESCUELA PROFESIONAL DE ESTADÍSTICA**

**Comparación de dos métodos de regresión al modelar  
los gastos mensuales brutos, en presencia de valores  
atípicos, de los hogares de la sierra central del Perú –  
2013**

**TESINA**

Para optar el Título Profesional de Licenciada en Estadística

Modalidad Examen de Suficiencia Profesional

**AUTOR**

**Sandy Giulianna CARBAJAL ORIHUELA**

Lima - Perú

2016

## RESUMEN

### “COMPARACIÓN DE DOS MÉTODOS DE REGRESION AL MODELAR LOS GASTOS MENSUALES BRUTOS, EN PRESENCIA DE VALORES ATÍPICOS, DE LOS HOGARES DE LA SIERRA CENTRAL DEL PERÚ - 2013”

Sandy Carbajal Orihuela  
Setiembre 2016

---

En la presente investigación se comparan los resultados de la bondad de ajuste al aplicar dos métodos de la Regresión Múltiple, el método de Máxima verosimilitud y el método robusto, cuyo objetivo es el de conocer cuál de los dos muestra mejores resultados al aplicarlos al modelo de regresión que explica los gastos mensuales de los peruanos, en presencia de observaciones atípicas e incumpliendo el supuesto de la normalidad de los residuos. Se realiza un análisis exploratorio previo de los datos para las diferentes variables y se analizarán usando los paquetes estadísticos SPSS y MATLAB; los datos usados en este trabajo corresponden a la Sumaria del Dominio Sierra Centro de la Encuesta Nacional de Hogares (ENAHO) del año 2013.

**PALABRAS CLAVE:** Regresión Lineal Múltiple, Regresión Robusta, redes Neuronales.

## ABSTRAC

### “COMPARACIÓN DE DOS MÉTODOS DE REGRESION AL MODELAR LOS GASTOS MENSUALES BRUTOS, EN PRESENCIA DE VALORES ATÍPICOS, DE LOS HOGARES DE LA SIERRA CENTRAL DEL PERÚ - 2013”

Sandy Carbajal Orihuela  
Setiembre 2016

---

In the present research the results of goodness of fit are compared by applying two methods of Multiple Regression, the Maximum likelihood method and the robust method, whose objective is to know which of the two shows better results when applied to the model Of regression that explains the monthly expenses of the Peruvians, in the presence of atypical observations and failing the assumption of the normality of the residues. A previous exploratory analysis of the data for the different variables is performed and analyzed using the statistical packages SPSS and MATLAB; The data used in this work correspond to the Sumaria of the Sierra Centro Domain of the National Survey of Households (ENAHO) of the year 2013.

**KEYWORDS:** Multiple Linear Regression, Robust Regression, Neural networks.