

FUNCIÓN DE DEMANDA PARA GALLETAS ARTESANALES EN EL CONJUNTO RESIDENCIAL PORTAL DEL BOSQUE, FLORIDABLANCA

Hugo Nicolás Castro Suárez *

Revisado por: Henry Sebastián Rangel Quiñonez †

1 Introducción

Este documento fue presentado como trabajo final de clase de la materia de Econometría I de la Facultad de Economía de la Universidad Santo Tomás. El objetivo general es estimar la función de demanda de galletas artesanales¹ a partir de herramientas econométricas de regresión múltiple, considerando cuatro variables explicativas (precio del bien, género del consumidor, ingreso semanal promedio de los consumidores y número de habitantes del hogar) y como endógena la cantidad de galletas consumidas, tomando como base investigaciones ya realizadas por Caraballo (2003), Cajamarca Pucha and Criollo Lima (2012), Silva (2001) y Marcillo and López (2012), quienes analizaron y estimaron la oferta y demanda de diferentes productos alimenticios.

2 Metodología

2.1 Recolección de datos.

Los datos que se utilizarán para estimar la función de demanda serán de construcción propia con información suministrada a partir de una breve encuesta que se realizará aproximadamente a 82 personas (muestra representativa) del conjunto residencial Portal del Bosque en Floridablanca, de forma aleatoria.

Tamaño de muestra para proporción:

$$n = \frac{N \cdot Z_{\alpha}^2 \cdot p(1-p)}{e^2(N-1) + Z_{\alpha}^2 \cdot p(1-p)} \quad (1)$$

$$n = \frac{(512)(1,926^2)(0,25)}{(0,10^2)(8512-1) + (1,96^2)(0,25)} = 81,00369 \cong 82 \quad (2)$$

La población total del conjunto residencial es cercana a 512 individuos, divididos en 8 torres de 4 pisos y cada piso con 4 apartamentos. El promedio de personas por familia es de 4 habitantes por apartamento, de acuerdo con la administración del conjunto. Se aproxima hacia arriba porque los individuos no se pueden fraccionar. Las preguntas que tendrá la encuesta serán las siguientes (Realizada en un formato web)(el consumidor probará una muestra estándar de galleta artesanal):

1. ¿Cuál es su nombre?
2. ¿Cuál su género?
3. ¿Cuántas personas viven con usted en su casa?
4. ¿Cuántas galletas consumiría semanalmente si el precio fuera de 500 a 5000 pesos (Se muestra una matriz de relación cantidad precio que va en las filas desde 500 pesos a 5000 pesos y en las columnas desde 0 a 20 galletas)?
5. ¿Cuál es su ingreso semanal promedio?

*Estudiante Facultad de Economía. Universidad Santo Tomás, Bucaramanga

†Docente, Facultad de Economía. Universidad Santo Tomás, Bucaramanga

¹Las galletas artesanales son idea emprendedora, cuya receta es propia del autor de este documento.

2.2 Estimación del modelo.

Según medidas básicas expresadas en Gujarati and Porter (2010) se decidieron tomar como variables a tener en cuenta para la estimación del modelo:

- Género: Categorizado en 0 o 1, donde 0 hacía referencia al género masculino y 1 al género femenino.
- Precio: Se estableció un rango de precios para los que los consumidores encuestados establecían la cantidad por consumir.
- Número de habitantes del hogar: Se refiere a la cantidad de personas que habitan con el entrevistado en el mismo apartamento.
- Ingreso semanal promedio: Se tomó en cuenta el ingreso semanal promedio, es decir, cuánto reciben las personas encuestadas (en pesos colombianos) semanalmente, resultado de sus actividades productivas o como subvención por parte de sus familiares (algunos de los encuestados vivían solos y eran estudiantes).

3 Resultados

3.1 Modelo final.

Se procedió a aplicar raíz cuadrada al modelo con mayor sesgo estandarizado (todas las estimaciones se relizaron en el programa R).

Tabla 1: *Resumen Estadístico para Cantidad*

Resumen Estadístico	Precio	SQRT (Cantid) (Y)	Sexo	SQRT (Ingreso)	N. de Personas (X)
Recuento	850	850	850	850	850
Promedio	2750,0	1,11184	0,576471	407,42	3,90588
Desviación Estándar	1436,99	1,07122	0,494409	172,378	1,19504
Coficiente de Variación	52,254%	96,3463%	85,7648%	42,3097%	30,5959%
Mínimo	500,0	0	0	173,205	1,0
Máximo	5000,0	4,47214	1,0	774,597	7,0
Rango	4500,0	4,47214	1,0	601,392	6,0
Sesgo Estandarizado	0	7,02856	-3,69059	5,49948	1,67191
Curtosis Estandarizada	-7,28653	-2,11431	-11,357	-6,08301	-2,43815

Fuente: Elaboración propia.

La tabla 1 muestra los estadísticos de resumen para todas las variables en el modelo. Incluye medidas de tendencia central, medidas de variabilidad y medidas de forma. De particular interés aquí son el sesgo estandarizado y la curtosis estandarizada, las cuales pueden utilizarse para determinar si la muestra proviene de una distribución normal. Valores de estos estadísticos fuera del rango de -2 a +2 indican desviaciones significativas de la normalidad, lo que tendería a invalidar cualquier prueba estadística con referencia a la desviación estándar. En este caso, el valor de sesgo estandarizado en cantidad, sexo e ingreso no se encuentra dentro del rango esperado para datos provenientes de una distribución normal, pero a comparación del anterior modelo, los valores aquí presentados son más bajos y tienen mejor distribución.

Estimadores:

$$q = \beta_0 + \beta_1 \text{precio} + \beta_2 \text{sexo} + \beta_3 \sqrt{\text{ingreso}} + \beta_4 \text{personas} \quad (3)$$

Tabla 2: *Resumen Estadístico para Cantidad*

Parámetro	Estimación	Valor -P
Constante	2,36799	0
Precio	-0,00053	0
Sexo	-0,353877	0
SQRT(Ingreso)	-0,000494	0,0007
X	0,155573	0

Fuente: Elaboración propia.

La salida muestra los resultados de la transformación del modelo anterior. La ecuación del modelo mejorado es:

$$SQRT(Cantidad) = 2,36799 - 0,000530442 \cdot Precio - 0,353877 \quad (4)$$

$$\cdot Sexo - 0,000493544 \cdot SQRT(Ingreso) + 0,155573 \cdot \#Personas \quad (5)$$

Puesto que el valor-P en la tabla ANOVA es menor que 0,05, existe una relación estadísticamente significativa entre las variables con un nivel de confianza del 95,0%. Los Betas para cada una de las variables quedan de la siguiente forma:

Tabla 3: *ANOVA*

Parámetro	Estimación	Estándar	T	Valor-P
Constante	2,36799	0,115439	20,5129	0,0000
Precio	-0,000530442	0,0000170283	-31,1505	0,0000
Sexo	-0,353877	0,0507969	-6,96651	0,0000
SQRT(Ingreso)	-0,000493544	0,000145807	-3,38491	0,0007
X	0,155573	0,0205671	7,56417	0,0000

Fuente: Elaboración propia.

La constante es igual a B_0 , es decir, el intercepto de la función con el eje Y, en este caso 2,36799. El estadístico R-Cuadrada indica que el modelo así ajustado explica 55,909% de la variabilidad en Cantidad. El estadístico R-Cuadrada ajustada, que es más apropiada para comparar modelos con diferente número de variables independientes, es 55,7003%. El R de correlación indica si la variable dependiente está relacionada positiva o negativamente con las variables explicativas. Para este caso el R de correlación indica que las variables están relacionadas positivamente en un 75%.

Puesto que el valor-P es menor que 0,05, hay indicación de una posible correlación serial con un nivel de confianza del 95,0%. Para determinar si el modelo puede simplificarse, hay que notar que el valor-P más alto de las variables independientes es 0,0067, que corresponde a Ingreso. Puesto que el valor-P del F estadístico (Significancia global) es menor que 0,05 ($2,2e-16$), ese término es estadísticamente significativo con un nivel de confianza del 95,0%. Consecuentemente, probablemente no se quisiera eliminar ninguna variable del modelo.

3.2 Correlación Serial entre variables explicativas

La tabla 4 muestra las correlaciones estimadas entre los coeficientes en el modelo ajustado. Estas correlaciones pueden usarse para detectar la presencia de multicolinealidad severa, es decir, correlación entre las variables predictoras. En este caso, no hay correlaciones con valores absolutos mayores muy grandes, todos son cercanos a 0,5 (sin incluir el término constante).

Tabla 4: Matriz de correlación para las estimaciones de los coeficientes

	Cantidad (constante)	Precio	Sexo	SQRT (ingreso)	X
Cantidad (constante)	1,0000	-0,4057	-0,3193	-0,5157	-0,6383
Precio	-0,4057	1,0000	0,0000	0,0000	0,0000
Sexo	-0,3193	0,0000	1,0000	0,2194	-0,0678
SQRT (ingreso)	-0,5157	0,0000	0,2194	1,0000	-0,0784
X	-0,6383	0,0000	-0,0678	-0,0784	1,0000

Fuente: Elaboración propia.

3.3 Análisis Residual

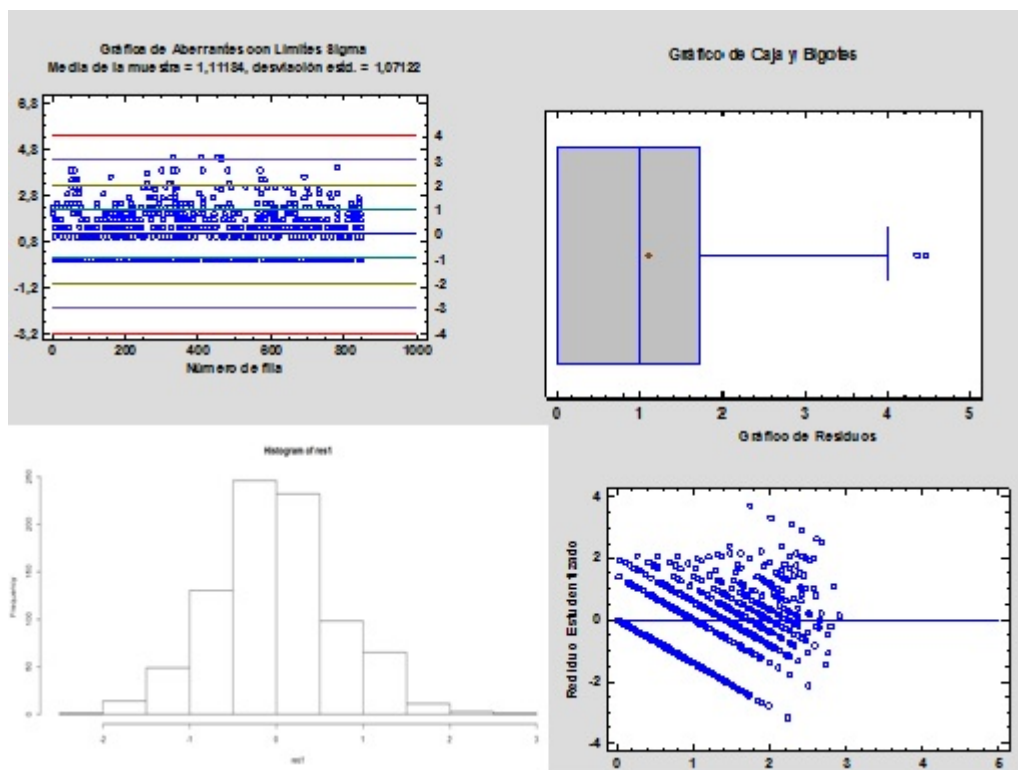


Figura 1. Análisis Residual. Fuente: Elaboración propia.

Tabla 5: Studentized Breusch-Pagan test

Data: reg2

BP=36.481 df=4 p-value=20303e-07

Fuente: Elaboración propia.

Tabla 6: Shapiro-Wilk normality test

Data: resl

W=0.99409 p-value=0.002034

Fuente: Elaboración propia.

Para el análisis residual de este modelo transformado también se realizaron dos pruebas, la primera, la Shapiro-Wilk, para determinar la normalidad de los errores y, la segunda, la prueba Breusch-Pagan, para determinar la heterocedasticidad y si los errores tienen varianza constante. Los resultados fueron los siguientes: Para la prueba de Shapiro-Wilk el valor-P más bajo de las pruebas realizadas es igual a 0,002034; debido a que el valor-P para esta es menor que 0,05, se puede rechazar la idea de que Cantidad proviene de una distribución normal con 95% de confianza.

Para la de Breusch-Pagan el valor-P más bajo de las pruebas realizadas es igual a 2,303e-7, debido a que el valor-P para esta es menor que 0,05, se puede rechazar la idea de que Cantidad proviene de una distribución homocedástica.

Las pruebas son concluyentes, pero es válido aclarar, que entre los dos modelos analizados, la transformación presenta tanto menor heterocedasticidad como una distribución más normal. En la figura 1 se muestra una distribución normal de los residuales; en la caja y bigotes existen datos atípicos, pero no tanto como en el anterior modelo; en la gráfica de aberrantes los datos ahora están dentro del límite establecido y los residuos estudentizados muestran valores que se ubican mejor y describen mejor lo que es nube de puntos, concordando con el análisis de los test para el análisis residual antes descritos.

3.4 Determinación de mejor modelo

Para determinar qué modelo se ajusta mejor se utilizó el Criterio de información de Akaike, en el que suponemos que se presenta menor estimación para el mejor modelo.

$$AIC(reg2) : 1844.073 \tag{6}$$

$$AIC(reg) : 3981.162 \tag{7}$$

Es evidente entonces, que la mejora logró el objetivo de ser un mejor modelo.

3.5 Comparación de normalidades

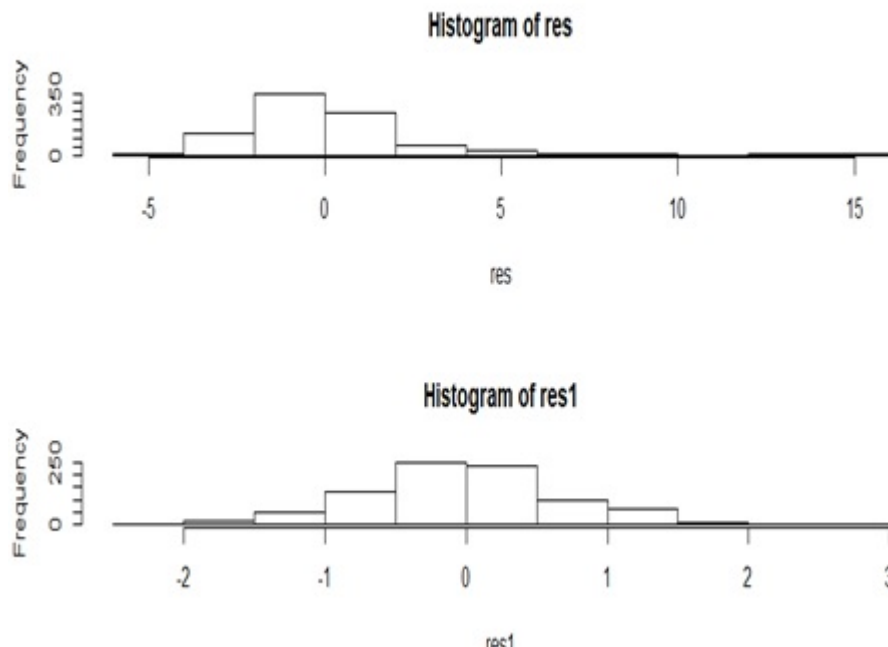


Figura 1. Histograma. Fuente: Elaboración propia.

Res= Residuales Regresión 1.

Res1=Residuales Regresión 2 (Modelo con transformación). La normalidad de los modelos varió bastante para establecer una mejoría en el modelo final.

4 Conclusiones

Las conclusiones se realizarán para el modelo transformado, ya que explica mejor el modelo y está más normalizado.

- El modelo logra explicar la variable dependiente en un casi 56% y es significativo a un nivel del 5%.
- Las variables afectan negativamente al precio, excepto el número de personas en el hogar, esto quiere decir que entre más personas en el hogar, mayor consumo, sin embargo, y como contrapartida, un incremento en las otras variables disminuyen el consumo.
- Las variables significativas que afectan al modelo de mayor influencia a menor influencia son Precio y esto responde a la razón de un aumento en 500 pesos del producto, disminuirá su consumo en 0,000530442 unidades de galletas, Si aumenta la Raíz(Ingreso) disminuirá el consumo de galletas en -0,000493544, Sexo disminuye en -0,353877 si el consumidor es mujer; para el número de personas, que es el único que afecta positivamente, la variación por número aumentó de unidad de persona que habite la casa será de 0,155573 unidades de galleta.

References

- Cajamarca Pucha, A. L. and Criollo Lima, A. M. (2012). Proyecto de microempresa, análisis y factibilidad de producir y distribuir galletas artesanales a tres colegios de cuenca. B.S. thesis.
- Caraballo, L. J. (2003). ¿ cómo estimar una funcion de demanda? caso: Demanda de carne de res en colombia. *Geoenseñanza*, 8(2).
- Gujarati, D. and Porter, D. (2010). *Econometría* (quinta edición).
- Marcillo, G. and López, L. E. (2012). Estimación lineal de las funciones de oferta y demanda de la papa en la ciudad de pasto, 1999-2010. *Tendencias*, 13(2):84–102.
- Silva, A. (2001). Estimación de la función de demanda de pollo a nivel minorista para la república argentina. *Expo LEAA. Licenciatura en Economía y Administración Agrarias. Facultad de Agronomía. Universidad de Buenos Aires. Págs*, 11.