

MODELO ESTADÍSTICO PARA OPTIMIZACIÓN DE ESPACIOS FÍSICOS



ADOLFO G. BOUILLON

— Profesor de Gestión de Operaciones
Área Académica de Administración
Facultad de Negocios, UPC

INTRODUCCIÓN¹

Los pronósticos juegan un papel importante en el proceso de planificación de toda organización, porque influyen en las decisiones actuales sobre operaciones a futuro. Por ejemplo, en la universidad, el pronóstico de los alumnos que se matricularían en las asignaturas permite determinar el número de secciones requeridas, así como el personal docente que impartiría clases en dichas secciones, es decir, permite optimizar el uso de capital físico como del capital humano.

En este sentido, cuando el pronóstico no es correcto se presentan dos situaciones: primero, el número de secciones estimadas sería superior al número real de secciones requeridas, debiendo cerrarse— por norma— aquellas secciones que tienen menos del mínimo establecido de alumnos, generando malestar en los docentes que ya no contarían con aquellas horas de dictado; segundo, el número de secciones podría ser menor al número real requerido, generando un déficit que deberá cubrirse con más profesores y más aulas; y en caso de no lograrlo, los alumnos terminarían perjudicándose al no poder matricularse en las asignaturas que ellos seleccionaron.

Por lo anteriormente expuesto, una adecuada estimación de los alumnos matriculados en los cursos tiene un impacto importante en la eficiencia de las operaciones de la universidad, así como, y sobre todo, en los indicadores de satisfacción de los estudiantes. En este contexto, el objetivo general de esta investigación es estimar y proyectar el número de alumnos matriculados y consecuentemente, el número de secciones de un curso seleccionado de una universidad privada local. Ello con el fin de tener mayor información para asignar óptimamente el capital físico (i.e. infraestructura: aulas y laboratorios) y humanos (i.e. número de docentes), y así resaltar la importancia del uso de modelos empíricos, para estimar, y finalmente proyectar, en la toma de decisiones empresariales.

Específicamente, en esta investigación, presentamos un modelo estadístico para estimar y proyectar la cantidad de alumnos matriculados para un curso de administración

de las operaciones, perteneciente a la Dirección de Administración de dicha universidad. Para cumplir con estos dos propósitos metodológicamente, adoptamos un modelo de regresión lineal en logaritmos (i.e., una estructura de doble logaritmo), estimada estadísticamente con la técnica de mínimos cuadrados ordinarios. Una vez estimado, el modelo permite predecir con la confianza esperada el número de alumnos que el curso en estudio debería tener a inicio de cada ciclo académico, desde el periodo 1996 hasta 2015. Para estimar el modelo, usamos Gauss y MS Excel.

Este trabajo se desarrolla en tres partes. En primer lugar, describimos brevemente el marco teórico del modelo de regresión lineal múltiple como principio organizacional. En segundo término, planteamos y desarrollamos el modelo, y finalmente, como tercer paso, usamos los resultados para inferir un pronóstico del número de alumnos que se matricularían en el curso antes referido, para los fines antes descritos.

EL FUNDAMENTO TEÓRICO

El objetivo de todo pronóstico es predecir el futuro afín. Para ello, tenemos que considerar aquellas variables relevantes en estimar su asociación estadística con la variable que queremos explicar para luego proyectarla, i.e., la variable de interés.

La regresión estadística tiene una importancia en todos los niveles de las organizaciones, tanto en lo estratégico como en lo táctico.

Entre los tipos de modelos cuantitativos, se tienen los modelos de series de tiempo y los modelos causales. Los primeros utilizan datos históricos de una variable de interés para pronosticar sus respectivos valores futuros. Los modelos causales, en cambio, suponen que la variable de interés se relaciona con uno o varios factores, los cuales permiten pronosticar el valor de la variable de interés. Tomando en cuenta que en este caso se identifican factores que pueden pronosticar el número de alumnos que se matricularían en el curso de Administración de las Operaciones y que sus relaciones con este número son lineales en logaritmos, es que se considera como mejor opción, el uso de un modelo de regresión de tipo causal.

1. Esta monografía fue publicada originalmente como documento de trabajo de la Working Paper Series de la Dirección de Administración y Finanzas, Facultad de Negocios, UPC. Este autor agradece al Dr. Andrés A. Escalante por sus valiosos comentarios a los borradores anteriores. Toda inquietud respecto a esta investigación, favor de remitirla a:

adolfo.bouillon@upc.edu.pe

Regresión Lineal

La regresión estadística trata del estudio de la relación de una variable (variable dependiente) respecto de una o más variables (variables explicativas) con el **objetivo de estimar o predecir el valor medio poblacional** de la primera en términos de los valores conocidos o fijos (en muestras repetidas) de las segundas. ²

Cuando se utilizan dos o más variables explicativas se tiene un modelo de regresión múltiple. Este puede ser lineal o no lineal; cuando el modelo es lineal se hace referencia a la linealidad que existe en los regresores o variables explicativas del modelo. Es decir, las variables independientes deben estar elevadas a la primera potencia cuando la regresión es lineal. Un modelo de regresión lineal es entonces una función estadística, donde se postula una relación lineal entre una variable dependiente y una o más variables independientes, estimada a partir de los datos históricos que se hayan recopilado en la muestra de información relevante.

Mínimos Cuadrados Ordinarios

El estimador que se utiliza en esta investigación para determinar los parámetros, mediante los que asociamos las variables independientes con la variable dependiente es el de mínimos cuadrados ordinarios (MCO). La teoría estadística, establece, en el caso que nos ocupa, que las variables independientes o regresoras, no siempre explican la totalidad del comportamiento de la variable dependiente, ya que existen fenómenos en la naturaleza que pueden no ser considerados o que se ignoran, dada la aleatoriedad presente en dichos fenómenos. Es así que suele existir un error estadístico, conocido como la diferencia entre el valor real de la variable dependiente y el valor estimado por el modelo. Lo que busca entonces el MCO, como estimador, **es minimizar el cuadrado de dicho error**, para lo que se resuelve un sistema de ecuaciones normales que resultan de la diferenciación parcial de la función de error cuadrático del modelo de regresión. Los valores estimados para los parámetros se utilizan a su vez dependiente, pudiendo así, tomar mejores decisiones.

Lo anterior puede escribirse matemáticamente así:

Donde Y_i : Valor verdadero del i ésimo valor de Y ; \hat{Y}_i : Valor estimado del i ésimo valor de Y ; \hat{u}_i Error estadístico

$$Y_i = \hat{Y}_i + \hat{u}_i$$

de la i ésima observación.

Al generalizar el modelo de regresión a k variables, con la variable dependiente Y , n datos, y $k-1$ variables

explicativas X_2, X_3, \dots, X_k , tenemos la siguiente función de regresión muestral:

$$Y_i = \hat{\beta}_1 + \hat{\beta}_2 X_{2i} + \hat{\beta}_3 X_{3i} + \dots + \hat{\beta}_k X_{ki} + \hat{u}_i \quad i = 1, 2, 3, \dots, n$$

Que se traduce en el siguiente conjunto de n ecuaciones simultáneas:

$$Y_1 = \hat{\beta}_1 + \hat{\beta}_2 X_{21} + \hat{\beta}_3 X_{31} + \dots + \hat{\beta}_k X_{k1} + \hat{u}_1$$

$$Y_2 = \hat{\beta}_1 + \hat{\beta}_2 X_{22} + \hat{\beta}_3 X_{32} + \dots + \hat{\beta}_k X_{k2} + \hat{u}_2$$

.....

$$Y_n = \hat{\beta}_1 + \hat{\beta}_2 X_{2n} + \hat{\beta}_3 X_{3n} + \dots + \hat{\beta}_k X_{kn} + \hat{u}_n$$

Este conjunto de ecuaciones puede reescribirse matricialmente, para fines de estimación y gestión analítica en general, de la siguiente forma:

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_{21} & X_{31} & \dots & X_{k1} \\ 1 & X_{22} & X_{32} & \dots & X_{k2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{2n} & X_{3n} & \dots & X_{kn} \end{bmatrix} \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \vdots \\ \hat{\beta}_k \end{bmatrix} + \begin{bmatrix} \hat{u}_1 \\ \hat{u}_2 \\ \vdots \\ \hat{u}_n \end{bmatrix}$$

Es decir,

$$\mathbf{Y} = \mathbf{X}\hat{\boldsymbol{\beta}} + \mathbf{u} \quad (1)$$

Donde \mathbf{Y} = vector columna $n \times 1$ de observaciones sobre la variable Y ; \mathbf{X} = matriz de $n \times k$, con n observaciones sobre las $k-1$ variables explicativas X_2, X_3, \dots, X_k , y la primera columna de unos representa al intercepto ($\hat{\beta}_1$); $\hat{\boldsymbol{\beta}}$ = vector columna de $k \times 1$ de los parámetros desconocidos $\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_k$; \mathbf{u} = vector columna $n \times 1$ de n errores u_i .

El estimador MCO tiene como objetivo minimizar la suma de los errores cuadráticos, el que matricialmente es:

$$\hat{\mathbf{u}}' \hat{\mathbf{u}} = \begin{bmatrix} \hat{u}_1 & \hat{u}_2 & \dots & \hat{u}_n \end{bmatrix} \begin{bmatrix} \hat{u}_1 \\ \hat{u}_2 \\ \vdots \\ \hat{u}_n \end{bmatrix} = \hat{u}_1^2 + \hat{u}_2^2 + \dots + \hat{u}_n^2 = \sum \hat{u}_i^2$$

de la **ecuación (1)** tenemos $\mathbf{u} = \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}$

2. Gujarati, D. N., & Porter, D. C. (2010). Econometría. México, DF: McGraw-Hill/Interamericana Editores, S.A de C.V, página 15.

por consiguiente, $\hat{u}'\hat{u} = (y - X\hat{\beta})(y - X\hat{\beta})$, es decir,

$$\hat{u}'\hat{u} = \hat{y}'\hat{y} - 2\hat{\beta}'X'y + \hat{\beta}'X'X\hat{\beta}$$

Diferenciando $\hat{u}'\hat{u}$ con respecto a $\hat{\beta}$,

$$\frac{\partial(\hat{u}'\hat{u})}{\partial\hat{\beta}} = -2X'y + 2X'X\hat{\beta}$$
, e igualando a cero, obtenemos

$$(X'X)\hat{\beta} = X'y$$
, de donde obtenemos:

$$\hat{\beta} = (X'X)^{-1}X'y \quad (2)$$

Siempre que exista la matriz inversa.³

La **ecuación (2)** representará el mejor estimador linealmente insesgado si se cumplen los siguientes supuestos:⁴

1. El modelo de regresión es lineal en los parámetros, aunque puede no ser lineal en las variables.
2. Los valores de X son fijos en muestreo repetido o son independientes del término de error.
3. El valor medio de la perturbación u_i es igual a cero.
4. Homoscedasticidad o varianza constante de u_i : La varianza del término de error es la misma independientemente del valor de X.
5. No hay autocorrelación entre las perturbaciones u_i .
6. El número de observaciones n debe ser mayor al número de parámetros a estimar en el modelo.
7. Naturaleza de las variables X: No todos los valores X en una muestra deben ser iguales, la varianza de estos valores deben ser mayores que cero, y además no deben tener valores atípicos o muy diferentes del resto de valores en la muestra.
8. No debe haber colinealidad exacta entre las variables X
9. No hay sesgo de especificación: Es decir, el modelo está correctamente especificado.

El estimador MCO permite estimar los parámetros del modelo de regresión lineal. Sin embargo, el objetivo es utilizar la función de regresión muestral para realizar inferencias sobre la función de regresión poblacional. Los parámetros son una función lineal del error estadístico, que también es una variable aleatoria; de esta manera, la distribución de probabilidad de β_1 dependerá del supuesto que se haga de la distribución de probabilidad del error estadístico u_i .

En el contexto de la regresión, por lo general se establece que el error estadístico tiene una distribución normal. Al agregar esta suposición, tenemos el modelo clásico de regresión lineal normal (MCRLN), donde se establece que $u_i \sim N(0, \sigma^2)$, con media 0 y varianza σ^2 .

3. Ver Greene (1998), para mayor detalle.

4. Ver Gujarati & Porter para mayor detalle.

5. Por ciclo regular se entiende los semestres 1 y 2, excluyéndose de la muestra estadística, por tanto, los ciclos de verano.

EL MODELO: ESPECIFICACIÓN, ESTIMACIÓN Y ANÁLISIS DE RESULTADOS

Análisis de los Datos Históricos

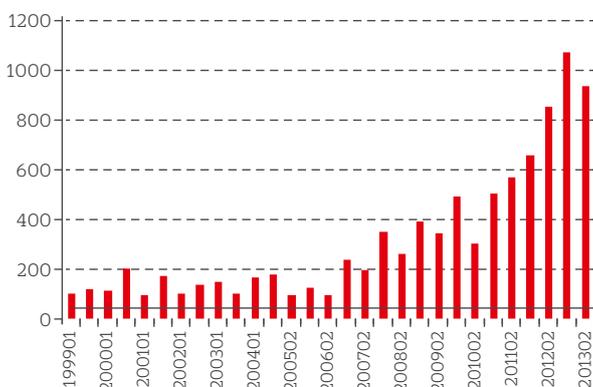
La tabla 1 muestra los datos históricos de la cantidad de alumnos matriculados en el curso de Administración de las Operaciones durante los ciclos regulares académicos desde el periodo 1999-1 hasta el periodo 2013-2.⁵ La figura 1, presenta en términos gráficos esta serie de tiempo para una mayor ilustración.

Tabla 1. Datos históricos de la demanda de alumnos para el curso Administración de las Operaciones

Periodo	Alumnos matriculados	Periodo	Alumnos matriculados
201302	382	200601	50
201301	440	200502	37
201202	349	200501	71
201201	268	201401	67
201102	233	200302	40
201101	206	200301	59
201002	124	200202	56
201001	201	200201	41
200902	141	200102	69
200901	159	200101	37
200802	105	200002	82
200801	142	200001	46
200702	79	199902	48
200701	96	199901	41
200602	39		

Fuente: elaboración propia.

Figura 1. Datos históricos normalizados de la demanda de alumnos, curso Administración de las Operaciones (Base 19901 = 100)



Fuente: elaboración propia.

De la **Figura 1** se puede ver que el número de alumnos se mantuvo relativamente estable entre 1999-1 y 2010-1. Este aumento de forma gradual hasta llegar a su valor más alto en el ciclo 2013-1. Este aumento implica una mayor cantidad de secciones, volviéndose crítico, por ejemplo, el proceso de planificación del equipo de profesores para la debida atención de los estudiantes.

Inicialmente, propusimos un modelo con diferentes variables independientes para explicar la matrícula de alumnos en el curso en estudio (i.e. la variable de interés). Entre ellos, por ejemplo, consideramos la de los alumnos hábiles para matricularse, ya que el universo al cual pertenece el número de alumnos inscritos corresponde a todos los alumnos que están hábiles para matricularse en el curso analizado, es decir, aquellos alumnos que han aprobado los pre-requisitos del curso bajo estudio. Asimismo, se consideraron las matrículas de cursos que tienen el mismo pre-requisito del curso analizado, para tomar en cuenta un efecto de sustitución entre un curso y otro. También se recopiló información de los cursos que están situados espacialmente cerca al curso analizado en la malla curricular, para analizar un posible efecto de complementariedad.

Otra de las variables explicativas fue considerar la del número de alumnos que aprobaron el curso en el ciclo anterior, ya que estos alumnos dejan de pertenecer al conjunto de alumnos hábiles para llevar el curso en el siguiente ciclo.

Para estos efectos, la tabla 2 contiene el número de variables contempladas inicialmente:

Tabla 2. Variables iniciales consideradas en el modelo de pronóstico

X_1	Alumnos hábiles para matricularse
X_2	Alumnos aprobados en el curso el ciclo anterior
X_3	Alumnos inscritos en el curso de Comportamiento Organizacional
X_4	Alumnos inscritos en el curso de Sistemas de Información
X_5	Alumnos inscritos en el curso de Investigación de Mercados
X_6	Número de carreras en la Facultad de Negocios
d_1	Variable dicótoma respecto al ciclo regular, 1= ciclo regular, 0 = en otro caso
d_2	Variable dicótoma respecto a la estructura del curso, 1= Estructura actual, 0 = en otro caso

Fuente: elaboración propia.

Al realizarse un análisis de significancia, a través del estadístico t , eliminamos 6 de las 8 variables que contempla la **Tabla 2**, quedando el modelo propuesto con el número de alumnos hábiles al inicio del ciclo académico, y los alumnos que aprobaron la asignatura en el ciclo académico regular anterior, como variables explicativas del grueso de la variación de los alumnos matriculados al inicio de ciclo, la variable dependiente o explicada.

Especificación del Modelo

El modelo de pronóstico propuesto queda expresado de la siguiente manera:

$$Y = \alpha X_1^{\beta_1} X_2^{\beta_2} \quad (3)$$

La transformación monótona nos permite reescribir la **ecuación (3)** en logaritmos:

$$\ln Y = \alpha + \beta_1 \ln X_1 + \beta_2 \ln X_2$$

donde Y : Alumnos matriculados al inicio del ciclo; X_1 : Alumnos hábiles para matricularse; X_2 : Alumnos aprobados en el curso el ciclo anterior; α : Coeficiente de intercepción del modelo y β_i es la elasticidad de la variable Y respecto a X_i .⁶

El modelo, trabajado con los datos históricos⁷, está especificado en logaritmos. Esta especificación puntualmente permite estimar el modelo de manera lineal en logaritmos y facilita el análisis de los parámetros, relacionando estos a las elasticidades de Y respecto a cada regresor X_i .⁸ Así, β_i representa la elasticidad de los alumnos matriculados respecto a los alumnos hábiles, es decir, el cambio porcentual en la cantidad de alumnos matriculados cuando los alumnos hábiles varían en 1%. De la misma forma, β_2 representa el cambio porcentual en los alumnos matriculados cuando el número de alumnos que aprobaron el ciclo anterior cambia en 1%. Asimismo, una transformación monótona de corte logarítmico tiende a reducir el problema potencial de heteroscedasticidad, debido a que comprime las escalas en las cuales se miden las variables, reduciendo una relación entre dos variables de diez veces a una diferencia de 2 veces.⁹

6. El coeficiente de elasticidad en notación de cálculo es

$$\epsilon = \frac{dy}{dx} \cdot \frac{x}{y}, \text{ en el modelo propuesto } \frac{d \ln y}{d \ln x_i} = \beta_i,$$

considerando que $\frac{d \ln y}{d \ln x_i} = \frac{1}{y}$; y $\frac{d \ln x_i}{d x_i} = \frac{1}{x_i}$; se tiene que

$$\frac{d \ln y}{d \ln x_i} = \epsilon$$

7. La data histórica está disponible de requerirse.

8. Ver Chiang (1984).

9. Ver Gujarati y Porter.

La muestra contiene 29 observaciones. Por tanto, en términos matriciales, nuestro modelo (**ver ecuación 1**) se dimensiona de la siguiente forma:

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_{29} \end{bmatrix} = \begin{bmatrix} 1 & X_{11} & X_{21} \\ 1 & X_{12} & X_{22} \\ \vdots & \vdots & \vdots \\ 1 & X_{129} & X_{229} \end{bmatrix} \begin{bmatrix} \hat{\alpha} \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_{29} \end{bmatrix}$$

Donde Y es un vector de 29 x 1, X es una matriz de 29 x 3, que representa a las variables explicativas, $\hat{\beta}$ es un vector de 3 x 1, que representa a los parámetros del modelo, y \hat{u} es vector de 29 x 1, que representa al error estadístico.

Estimación del Modelo

Los valores estimados de los parámetros se exhiben en la **tabla 3**:

Tabla 3. Parámetros del modelo

Coefficientes	Valor	Estadístico t	Probabilidad	Error
$\hat{\alpha}$	0.1952	0.7587	0.4548	0.2573
$\hat{\alpha}_1$	0.6968	9.5084	0.0000	0.0733
$\hat{\alpha}_2$	0.2151	2.5601	0.0166	0.0840
Coefficiente de Determinación:	0.9228			
Test Anderson-Darling:	0.2308			
p-value (AD):	0.8048			

Fuente: Estadísticas de GAUSS.

Elaboración: Propia. Modelo estadístico programado y estimado en GAUSS

Al realizar el análisis de significancia con el estadístico t se determinó que ambos parámetros ($\hat{\beta}_1, \hat{\beta}_2$) son significativos al 95% de confianza (**ver Tabla 2**), validando el modelo propuesto.

Respecto a la bondad de ajuste el coeficiente de determinación r^2 indica que el modelo explica las variaciones de la variable dependiente en un 92.28%, es decir, que el 92.28% de la variación de Y se explica por los cambios de todas las variables independientes, dejando un 7.72% de la variación de aquella a variables no contempladas explícitamente, pero que se capturan en el error estadístico. Al aplicar la prueba de normalidad **Anderson – Darling** a los residuos, se obtuvo un p -value de 0.8048, es decir, que asumiendo un 5% de riesgo no se rechaza la hipótesis nula de que los errores tienen una distribución normal. Esto es importante porque el hecho de que el error estadístico tenga distribución normal

permite realizar inferencias sobre la distribución de probabilidad de los parámetros del modelo, debido a que éstos son una función lineal del error estadístico.

La validez de estos resultados se basan parcialmente en que el modelo no presente multicolinealidad, autocorrelación y heteroscedasticidad. Por multicolinealidad entendemos una relación lineal exacta entre las variables explicativas X_i . Autocorrelación supone que no hay correlación entre las perturbaciones u_i . La heteroscedasticidad se presenta cuando la varianza condicional de u_i para los diferentes valores de las variables X_i no es la misma. Ver supuestos de la página 10.

Este punto es importante porque de mostrar multicolinealidad, autocorrelación y heteroscedasticidad, el uso de los MCO como estimador no garantizaría obtener el mejor estimador linealmente insesgado.¹⁰

Respecto a la multicolinealidad, es importante ver el grado de correlación que existe entre las variables explicativas X_i . Para tal efecto, la **tabla 4** presenta la matriz de correlación del modelo:

Tabla 4. Matriz de Correlación

	LnX_1	LnX_2
LnX_1	1	0.7550
LnX_2	0.7550	1

Fuente: elaboración propia.

El coeficiente de correlación entre los regresores es 0.75, que es menor a 0.80; asimismo el estadístico r^2 es 0.57, mucho menor al que el modelo registra, de 0.9229, con lo que se concluye, según la regla práctica de Klein¹¹, que el grado de multicolinealidad no representa un problema potencial en esta oportunidad.

Respecto a la autocorrelación, se presentan los resultados de la prueba d de Durbin – Watson:

Tabla 5. Resultados de la prueba Durbin – Watson

Elemento	Valor
Estadístico d	2.3395
N	29.0000
Grados de libertad	2.0000
dl	1.2700
du	1.5600

Fuente: elaboración propia.

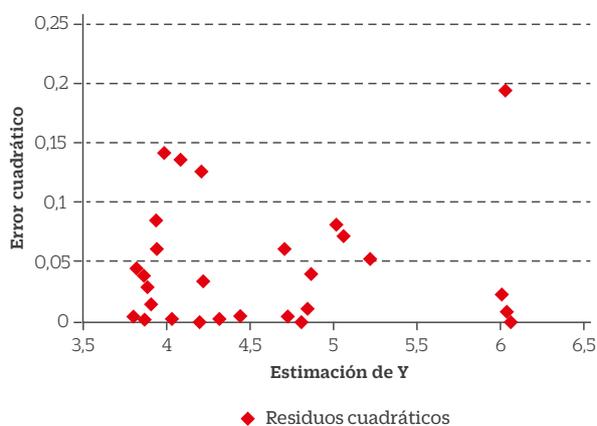
10. Ver Greene.

11. Ver Gujarati y Porter.

El estadístico d es de 2.33, al considerar un nivel de confianza de 95% y 2 grados de libertad; y dado que el estadístico d se encuentra entre du y $(4-du)$, no se rechaza la hipótesis nula de que no hay ni autocorrelación positiva ni autocorrelación negativa en el modelo.

Para el análisis de la heteroscedasticidad, se preparó el análisis de los residuos cuadráticos versus el valor estimado de Y (Ver figura 2).

Figura 2. Diagrama de dispersión, residuos cuadráticos - estimación de Y



Fuente: Modelo estadístico programado en GAUSS.
Elaboración: propia.

Esta figura muestra que los errores cuadráticos están dispersos, con lo que se atenúa el problema de heteroscedasticidad.

Sin embargo, para corroborar lo anterior, se realizó la prueba general de White, obteniéndose los siguientes resultados (ver Tabla 6).

Tabla 6. Prueba general de heteroscedasticidad de

Estadístico Observado	Grados de Libertad	P - value (Chi Cuadrada)
11.0744	5	0.0499

Fuente: Estadísticas entregada por el programa GAUSS.
Elaboración: propia.

El valor de probabilidad de esta prueba es igual a 0.0499, lo cual indica que con un 96% de confianza, no se rechaza la hipótesis nula de que no hay heteroscedasticidad en los datos; de esta manera, considerando aceptable el nivel de confianza de 96%, se concluye que no hay heteroscedasticidad en el error.

Análisis de Resultados del Modelo

Pudiendo inferir que el error estadístico satisface sus supuestos, que podemos hacer de MCO el mejor estimador lineal insesgado, tiene sentido ahora interpretar los valores de los parámetros.

De la tabla 3, $\beta_1 = 0.6968$ y $\beta_2 = 0.2151$. Como se dijo anteriormente, estos valores representan las elasticidades de la variable dependiente respecto a las dos variables independientes propuestas. De esta manera, al interpretar el coeficiente β_1 , se tiene que un aumento del 1% en la cantidad de alumnos hábiles al inicio del ciclo actual t respecto del ciclo anterior $t-1$, en igualdad de condiciones incrementaría en 0.7% la cantidad de alumnos inscritos en el curso en el ciclo actual t . Asimismo, β_2 significa que un aumento del 1% en la cantidad de alumnos aprobados en el ciclo anterior $t-1$ respecto al ciclo $t-2$, incrementará en 0.2% la cantidad de alumnos inscritos para el ciclo regular actual t . (Los alumnos que aprueban el curso ya no forman parte del conjunto de alumnos hábiles para matricularse en dicho curso).

Es así que, al tener la información de inicio de ciclo y del ciclo anterior, se puede determinar cuántos alumnos estarán dispuestos a matricularse en el curso de Administración de las Operaciones; y por lo tanto, cuántas secciones se requerirán para el ciclo académico de interés, cuya planificación se desea realizar adecuadamente.

PROYECCIÓN DE LA MATRÍCULA

Estimación del Número de Alumnos Matriculados en Periodos de Tiempo Pasados y Análisis de los Errores del Modelo

Con el modelo propuesto, se calculó el número de alumnos que se deberían haber matriculado y se comparó con el número real de alumnos matriculados a lo largo de los 29 periodos de estudio. La tabla N° 7 muestra los resultados obtenidos en la estimación de periodos pasados:

Tabla 7. Estimación de periodos históricos, periodos 1999-1 a 2013-2

Periodo	Alumnos estimados	Alumnos reales	Error	Error absoluto	Secciones absoluto	Error en secciones enteras
201302	420.07	382	38.07	38.07	1.27	1.0
201301	430.22	440	-9.78	9.78	0.33	0.0
201202	407.19	349	58.19	58.19	1.94	2.0
201201	416.45	268	148.45	148.45	4.95	5.0
201102	185.06	233	-47.94	47.94	1.20	-1.0
201101	157.45	206	-48.55	48.55	1.21	-1.0
201002	122.48	124	-1.52	1.52	0.04	0.0
201001	151.02	201	-49.98	49.98	1.25	-1.0
200902	127.06	141	-13.94	13.94	0.35	-1.0
200901	129.98	159	-29.02	29.02	0.73	-1.0
200802	112.81	105	7.81	7.81	0.20	0.0
200801	110.79	142	-31.21	31.21	0.78	-1.0
200702	84.88	79	5.88	5.88	0.15	0.0
200701	67.27	96	-28.73	28.73	0.72	-1.0
200602	47.54	39	8.54	8.54	0.21	0.0
200601	47.65	50	-2.35	2.35	0.06	0.0
200502	53.93	37	16.93	16.93	0.42	1.0
200501	74.85	71	3.85	3.85	0.10	0.0
200401	66.51	67	-0.49	0.49	0.01	0.0
200302	51.25	40	11.25	11.25	0.28	0.0
200301	56.32	59	-2.68	2.68	0.07	0.0
200202	49.53	56	-6.47	6.47	0.16	0.0
200201	48.68	41	7.68	7.68	0.19	0.0
200102	51.52	69	-17.48	17.48	0.44	-1.0
200101	45.79	37	8.79	8.79	0.22	0.0
200002	68.07	82	-13.93	13.93	0.35	-1.0
200001	47.89	46	1.89	1.89	0.05	0.0
199902	44.71	48	-3.29	3.29	0.08	0.0
199901	59.31	41	18.31	18.31	0.46	1.0

Fuente: elaboración propia.

La columna de Error muestra la diferencia entre los alumnos estimados por el modelo y los alumnos que efectivamente se inscribieron en el curso en los periodos analizados. Si estos números son negativos, el modelo estaría subestimando la matrícula real, lo que a su vez, generaría una oferta insatisfecha para los alumnos que desean matricularse en el curso. Si los números son positivos, tendríamos el caso opuesto.

Para obtener el error expresado en número de secciones se tomó el error en términos absolutos y se dividió entre el número de alumnos por sección. Desde el periodo 1999-1 hasta el periodo 2011-2 el número de alumnos por sección fue de 40. En los periodos que van desde el 2012-1 hasta el 2013-2 el número de alumnos por sección fue de 30.

Inicialmente el error en número de secciones se obtuvo en números decimales. Con el objetivo de obtener el error en número entero de secciones se consideró la política de la universidad respecto al mínimo requerido de alumnos para aperturar una sección, que es de doce alumnos. De esta manera, en aquellos casos donde la fracción no atendida de alumnos no superó dicho número se redondeó el número al entero inmediato inferior; asimismo, cuando la fracción supera dicho número la estimación se redondea al número entero inmediato superior. Estos resultados se muestran en la columna siete de la **tabla N° 9**. Los números negativos representan una oferta insuficiente en términos de secciones; los números positivos representan una oferta superior al número de secciones requeridas. Cuando el número es cero significa un ajuste perfecto entre oferta y secciones requeridas.

Desde el punto de vista de la planificación, estos resultados generan tres situaciones: el escenario ideal es cuando el ajuste es perfecto, es decir, el número del error en secciones es de cero; cuando la oferta es superior, aquellas secciones que no llegan al mínimo establecido de doce alumnos deben cerrarse, perdiendo los profesores de dichas secciones las horas asignadas previamente; y cuando la oferta es inferior, se deben buscar profesores, ya sean actuales o nuevos, para cubrir el déficit establecido de horas.

Al analizar los resultados de la **tabla N° 9**, se puede determinar que la desviación absoluta media del pronóstico fue de 0.63 secciones, es decir, menos de una sección en promedio. Asimismo, de la muestra observada de 29 periodos, hubo ajuste perfecto en 15 periodos, que representan el 51.7% del total. El número de periodos donde la oferta fue superior a las secciones requeridas fue de cinco, representando el 17.3% del total. En el caso de una oferta insuficiente, el número de periodos fue de nueve, que representan el 31% del total.

Cuando la oferta de secciones fue mayor a lo requerido, el promedio de secciones a cerrar fue de dos. Cuando hubo oferta insuficiente, el número promedio de secciones faltantes fue de uno.

De esta manera, se puede observar que el error es pequeño, sobre todo en los periodos donde hubo una oferta menor a la demanda real de secciones, ya que el promedio de secciones faltantes fue de uno. Esta situación es importante porque la subestimación de secciones, *vis-à-vis* la demanda efectiva implica un número disponible, tanto de profesores, como de salones, inadecuado; y mientras el número de secciones faltantes sea más pequeño, será más fácil conseguir los profesores necesarios para otorgar una oferta adecuada al número de secciones requeridas por los estudiantes.

Proyección de los alumnos matriculados

Al término de esta investigación, se tuvo la oportunidad de contrastar la estimación del modelo con los datos reales para los periodos 201401 y 201402. Los resultados se presentan en la siguiente tabla:

Tabla 8. Pronóstico de los periodos 201401 y 201402

Periodo	Variables		Logaritmos		Pronóstico	Alumnos pronosticados	Alumnos realmente matriculados
	X_1	X_2	X_1	X_2			
201401	812	346	6.70	5.85	6.12	455.21	425
201402	878	299	6.78	5.70	6.14	465.82	454

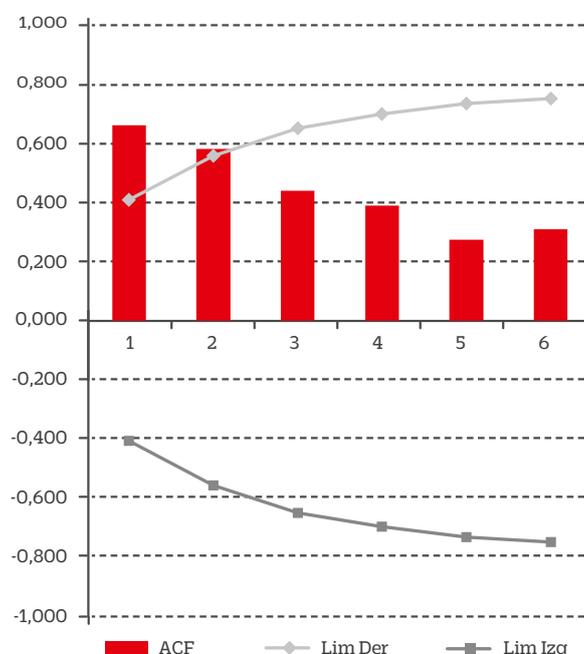
Fuente: elaboración propia.

La estimación realizada para el periodo 201401 fue de 455.21 alumnos. Al revisar los datos reales, se tuvo en dicho ciclo un total de 426 alumnos realmente matriculados; de esta manera el pronóstico sobreestimó la demanda real en 30 alumnos, o el equivalente de una sección.

Para el periodo 201402 la estimación fue de 466 alumnos, mientras que los alumnos realmente matriculados fueron 454; lo que indica que el pronóstico sobreestimó la demanda real en 12 alumnos, lo que representaría una sección.

Lógicamente, no se cuenta con información respecto a los periodos 201501 en adelante, razón por la cual se realizó un análisis de series de tiempo para las variables X_1 y X_2 . Para ello, determinamos si existen patrones de estacionalidad o tendencia en las variables. Para tal efecto, se presentan los autocorrelogramas y sus análisis correspondientes:

Figura 3. Autocorrelograma de la variable X_1 (alumnos hábiles)



Fuente: elaboración propia.

Al observar la **figura 3**, se puede ver que los datos presentan tendencia, mas no estacionalidad, ya que los coeficientes se van haciendo pequeños y no significativos al estar dentro del intervalo de confianza al 95%. De acuerdo a esta información, se eligieron dos métodos¹², la regresión lineal simple y el modelo de ajuste exponencial de Holt, debido a que ambos métodos toman en cuenta la tendencia para realizar el pronóstico con series de tiempo; es así que se obtienen los siguientes resultados:

Tabla 9. Comparación de métodos de pronóstico, variable X1 (alumnos hábiles al inicio de ciclo)

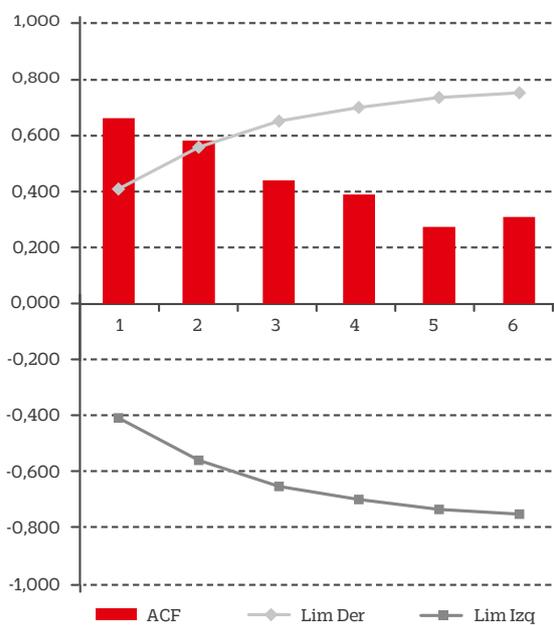
Modelo	Parámetros	Desviación Absoluta Media (MAD)
Regresión Lineal Simple	a = -106.13 b = 20.99	134.64
Modelo de Holt	Alfa = 0.87 Beta = 0.01	60.44

Fuente: elaboración propia.

Debido a que el modelo de Holt presenta una menor desviación absoluta media (MAD), es decir, una menor diferencia entre los datos reales y estimados, es que se elige este método para realizar la proyección de la variable X_1 .

Respecto a la variable X_2 , se tiene la siguiente información:

Figura 4. Autocorrelograma de la variable X2 (alumnos aprobados en el ciclo anterior)



Fuente: elaboración propia.

Observamos lo mismo que vimos en cuanto a la variable anterior: los datos presentan tendencia, mas no estacionalidad, ya que los coeficientes se van haciendo pequeños y no significativos al estar dentro del intervalo de confianza al 95%. Como en el primer caso, elegimos los mismos métodos para estimar esta segunda variable: la regresión lineal simple y el modelo de ajuste exponencial de Holt, obteniéndose los siguientes resultados:

Tabla 10. Comparación de métodos de pronóstico de la variable X2 (alumnos aprobados en el ciclo anterior)

Modelo	Parámetros	Desviación Absoluta Media (MAD)
Regresión Lineal Simple	a = -17.17 b = 8.05	33.68
Modelo de Holt	Alfa = 0.55 Beta = 0.01	28.58

Fuente: elaboración propia.

Nuevamente, debido a que el modelo de Holt presenta la menor desviación absoluta media, se eligió este modelo para realizar la proyección de la variable X_2 . Se presenta entonces la información proyectada con el método de ajuste exponencial de Holt:

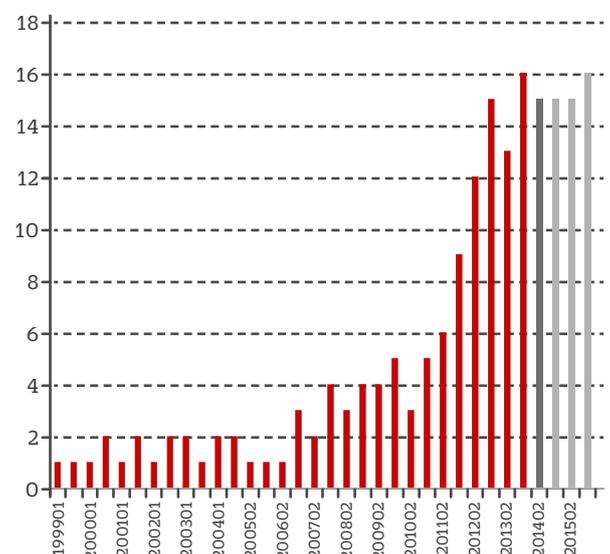
Tabla 11. Pronóstico de los periodos 201501 a 201502

Periodo	Variables		Logaritmos		Pronóstico	Alumnos pronosticados	Secciones enteras
	X_1	X_2	X_1	X_2			
201501	793.93	345.48	6.68	5.84	6.10	447.98	15.0
201502	817.17	354.56	6.71	5.87	6.13	459.63	15.0
201601	840.40	363.65	6.73	5.90	6.16	471.26	16.0

Fuente: elaboración propia.

Con esta información se tiene que la proyección es de 15 secciones para los periodos 201501 y 201502 y 16 secciones para el periodo 201601. No es recomendable proyectar más periodos ya que el uso del método de ajuste exponencial de Holt está pensado en el mediano plazo, es decir, no más de 18 meses. El número de secciones históricas y estimadas se muestra gráficamente en la **figura 5**:

Figura 5. Número de secciones en periodos históricos (199901-201402) y proyección para los periodos 201501-201601



Fuente: elaboración propia.

Uso de las elasticidades para realizar la proyección de periodos futuros

Se utilizaron las elasticidades las variables entregadas por el modelo para calcular la proyección futura y su comparación con los resultados obtenidos en la **tabla 11**. Los datos trabajados se presentan en las **tablas 12 y 13**:

Tabla 12. Pronóstico de los periodos 201501 a 201502 con elasticidades, variable X1

Periodo	Variación porcentual X_1	Variación porcentual según modelo X_1	Alumnos pronosticados	Diferencia en alumnos	Diferencia en secciones
201501	3.01%	2.10%	442.89	5.09	0.17
201502	2.93%	2.04%	451.92	7.71	0.26
201601	2.84%	1.98%	460.87	10.38	0.35

Fuente: elaboración propia.

Tabla 13. Pronóstico de los periodos 201501 a 201502 con elasticidades, variable X2

Periodo	Variación porcentual X_2	Variación porcentual según modelo X_2	Alumnos pronosticados	Diferencia en alumnos	Diferencia en secciones
201501	2.70%	0.58%	429.59	18.38	0.61
201502	2.63%	0.57%	432.02	27.61	0.92
201601	2.56%	0.55%	434.40	36.85	1.23

Fuente: elaboración propia.

Como se puede apreciar en ambas tablas, si bien existen una diferencias en la estimación del número de alumnos, aquéllas representan en casi todos los casos una cantidad menor a 1 sección; de esta manera no representa una diferencia significativa respecto a la estimación del número de alumnos utilizando la ecuación completa del modelo propuesto.

CONCLUSIONES

El objetivo de este estudio ha sido estimar y proyectar el número de alumnos y secciones de un curso de administración de las operaciones de una universidad privada local.

Para este fin se obtuvo una muestra de los alumnos matriculados en dicho curso desde el ciclo 1999-1 hasta el ciclo 2013-2. Se utilizó un modelo de pronóstico asociativo, a través de la regresión lineal múltiple; y el método para estimar los parámetros fue el de MCO. El modelo constó de dos variables: número de alumnos hábiles al inicio de ciclo, y el número de alumnos que aprobaron el curso en el ciclo académico anterior. En la especificación del modelo se utilizaron logaritmos, ya que estos facilitan el uso de un modelo lineal en logaritmos y por ende el empleo de MCO como estimador. Además, un modelo lineal en logaritmos permite calcular la elasticidad de la variable dependiente, respecto de las variables independientes; como también reducir el problema potencial de heteroscedasticidad.

El modelo satisfizo todos los supuestos del modelo clásico de regresión lineal normal. Es decir, con la debida confianza, podemos inferir que no hay heteroscedasticidad, autocorrelación ni multicolinealidad.

Al utilizar el modelo para estimar los periodos históricos, se obtuvo un ajuste perfecto entre el número de secciones reales y el número estimado de secciones en un 51.7% del total de la muestra; en el caso de una estimación inferior al número real de secciones ésta se dio en el 31% de los periodos; y en el caso de una estimación superior se obtuvo en 17.3% del total de periodos.

Una estimación inferior al número real de secciones en alrededor de 30% de los periodos podría ser preocupante, en la medida en que ello se traduzca en un gran número de secciones faltantes. Afortunadamente, esta subestimación se traduce en un promedio de secciones subestimadas de 1. Cuando la estimación fue superior al número real de secciones, el promedio fue de dos secciones sobreestimadas.

Una de las limitaciones del presente trabajo ha sido la cantidad de datos para la muestra de estudio, que fue de 29 observaciones, debido a que el curso como tal se creó en el periodo 1999-1. Otra limitación afín ha sido que la estructura y el pre-requisito del curso cambiaron en el ciclo 2012-1, pudiendo haber otros factores que hayan influido en la cantidad de alumnos inscritos y que en este modelo están representados por el error estadístico.

Los resultados obtenidos validan el uso del modelo propuesto como un paso en la dirección correcta para estimar el número de alumnos inscritos en los periodos futuros de un curso de administración de las operaciones. Generalizando, estos resultados validan el uso de este tipo de modelos, relativamente sencillos computacionalmente, aunque intelectualmente exigentes, para tomar mejores decisiones empresariales, en cuanto a organización y uso óptimo de recursos tangibles, como intangibles.

12. Ver Gujarati & Porter, y Hanke (2010), para mayor detalle.

Como agenda de trabajo a futuro, se podría emplear este modelo en el resto de cursos de una facultad y de la universidad, para que, una vez validados, sirvan con su aplicación en la mejora del proceso de planificación de inicio de ciclo, que es el objetivo ulterior de esta investigación.

REFERENCIAS BIBLIOGRÁFICAS

Chiang, A. C. (1984). *Fundamental Methods of Mathematical Economics*. Nueva York: McGraw-Hill.

Greene, W. H. (1998). *Análisis Econométrico*. Madrid: Prentice Hall.

Gujarati, D. N., & Porter, D. C. (2010). *Econometría*. México, DF: McGraw-Hill/Interamericana Editores, S.A de C.V.

Hanke, J. E. (2010). *Pronósticos en los negocios*. México, DF: Pearson Educación.

Heizer, J., & Render, B. (2009). *Administración de Operaciones*. México: Pearson Educación.