



**UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERIA INDUSTRIAL**

**VALORIZACIÓN MARGINAL DEL PROGRAMA DE PASAJERO FRECUENTE EN
UNA LÍNEA AÉREA**

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL INDUSTRIAL

INFORME DE AVANCE FINAL

FELIPE ANTONIO ARES DARLIC

PROFESOR GUÍA:

José Mosquera

MIEMBROS DE LA COMISIÓN:

Ricardo San Martín

Marcelo Henríquez

SANTIAGO DE CHILE
DICIEMBRE 2008

RESUMEN EJECUTIVO

VALORIZACIÓN MARGINAL DEL PROGRAMA DE PASAJERO FRECUENTE EN UNA LÍNEA AÉREA

Debido a los recientes cambios experimentados en la industria aérea, la empresa LAN Airlines se ha visto en la necesidad de evaluar su programa de fidelización de clientes LANPASS. Específicamente desea obtener cuanto vale su programa de pasajero frecuente, de modo de incorporar estos resultados en la toma de decisión de tarifas y reservas llevadas a cabo por su departamento de *Revenue Management*.

El presente trabajo de título tuvo como objetivo cuantificar el valor agregado que sienten los clientes de las aerolíneas que operan en Chile sobre la existencia de algún programa de pasajero frecuente. Específicamente se estudió la valorización de los pasajeros hacia el programa de fidelización de LAN Airlines llamado LANPASS.

El estudio se centro en una muestra de 5 rutas, las cuales obedecen a los mercados específicos de LAN (Intercontinental, Larga distancia Sudamérica, Larga distancia EEUU, Mediana distancia y Corta distancia) y abarcan mas del 51% de los pasajeros movilizadas anualmente por la industria.

Utilizando la información sobre participaciones de mercado y tarifas medias mensuales en los años 2007 y 2008 sobre las aerolíneas que operan las rutas mencionadas anteriormente, se formuló un modelo predictivo de la probabilidad de algún pasajero cualquiera de escoger algunas de las líneas aéreas, para alguna ruta particular. Este modelo incorpora (además del precio y la participación de mercado) una variable extra que obedece a la valorización sentimental que sienten los clientes por cada línea aérea.

Una vez estimada esta valorización, utilizando modelos lineales, se logró determinar de forma convincente (con un ajuste del 99,3%) que dichas valorizaciones dependen de ciertas variables específicas de las aerolíneas en estudio, tales como la utilización de vuelos directos, las frecuencias de salidas, rutas en que operan y utilización de algún programa de pasajero frecuente. La incorporación de esta última variable permite conocer el valor agregado de dichos programas de fidelización en dólares.

Utilizando el modelo lineal de menor error de predicción, se llegó al resultado de que en el caso específico del programa de pasajero frecuente de LAN Airlines, el valor agregado de dicho programa es de hasta un 4,36% del valor del pasaje. Esto quiere decir por ejemplo, que de pagar 100 dólares por un vuelo a Buenos Aires, un pasajero estaría dispuesto a cancelar 95,64 dólares y quedar excluido de la acumulación de kilómetros.

Si bien los resultados obtenidos no están exentos de error, este estudio presenta una referencia del valor que entrega LANPASS a los resultados de la empresa en el corto plazo, quedando por investigar en estudios posteriores, los impactos a largo plazo que podría generar el programa de pasajero frecuente LANPASS.

ÍNDICE DE CONTENIDOS

1.- Introducción	5
2.- Antecedentes y definición del problema.....	5
2.1.- Sobre el programa de pasajero frecuente	5
2.2.- Razones del estudio	6
2.3.- Objetivos.....	7
2.3.1.- Objetivos principales	7
2.3.2.- Objetivos secundarios.....	7
2.4.- Alcances de la memoria	7
3.- Metodología y marco teórico	7
3.1.- Resolución del problema	7
3.2.- Marco teórico.....	8
3.2.1.- Sobre el cálculo de las valorizaciones.....	8
3.2.1.1.- Caso simple (dos aerolíneas):.....	8
3.2.1.2.- Caso General:	9
- Supuestos del modelo:	9
- Derivación de los S_{it} :.....	10
- Primera solución: aproximación lineal.....	11
- Obtención numérica de los valores α_1 y α_2	13
- Mejora de la solución lineal	14
- Obtención de la constante de escala M.....	17
- Comentarios sobre los supuestos.....	17
3.2.2.- Sobre el modelo a utilizar:.....	19
4.- Desarrollo.....	20
4.1.- Estudio y selección de las rutas	20
4.2.- Cálculo de las valorizaciones	21
4.3.- Selección de las variables explicativas.....	23
4.4.- Modelo de predicción	24
4.4.1.- Primer modelo: Ln(P-S)	24
4.4.2.- Segundo modelo: S/Pmedio.....	27
4.4.3.- Tercer modelo: S	28
4.4.4.- Elección del modelo a utilizar.....	30
3.4.5.- Resultados sobre el modelo elegido	31
5.- Conclusiones	34
6.- Bibliografía	36
7.- Material complementario	36
7.1.- Anexo A	36
7.2.- Anexo B.....	37
7.3.- Anexo C.....	41
7.4.- Anexo D.....	43

ÍNDICE DE ILUSTRACIONES Y CUADROS

Ilustración 1: Error absoluto medio vs. Iteración	16
Ilustración 2: Diferencias respecto a la constante M.....	17
Ilustración 3: Serie de valorizaciones para la ruta Santiago-Madrid.....	22
Ilustración 4: Serie general de valorizaciones	23
Ilustración 5: Errores del modelo Ln(P-S)	26
Ilustración 6: Frecuencia de errores modelo Ln(P-S).....	26
Ilustración 7: Comparación de S del modelo 1.....	27
Ilustración 8: Errores del modelo S/Pmed	28
Ilustración 9: Frecuencia de errores del modelo S/Pmed	28
Ilustración 10: Errores del modelo S	29
Ilustración 11: Frecuencia de errores del modelo S	30
Cuadro 1: Comparación de precios para distintas tarifas	18
Cuadro 2: Comparación porcentual de precios para distintas tarifas	18
Cuadro 3: Rutas seleccionadas para el estudio.....	21
Cuadro 4: Mercados objetivos y pasajeros transportados	21
Cuadro 5: Valorizaciones medias por ruta	21
Cuadro 6: Error medio por ruta	22
Cuadro 7: Variable Dependiente: Ln(P-S)	25
Cuadro 8: Variable Dependiente: S/Pmed.....	28
Cuadro 9: Variable Dependiente: S	29
Cuadro 10: Comparación entre los modelos.....	30
Cuadro 11: Varianzas-Covarianzas de las constantes del modelo logarítmico.....	32
Cuadro 12: Datos propios de LAN por rutas.....	33
Cuadro 13: Valorización debido al programa de pasajero frecuente.....	33
Cuadro 14: Resultados sobre los modelos rezagados.....	34

1.- Introducción

La industria aérea en Chile ha experimentado una serie de cambios en los últimos años. Al aumento considerable observado en el último año de pasajes vendidos se suma la entrada de nuevas firmas a la industria, aumento preocupante de los costos operacionales (variabilidad en el precio del petróleo), baja en los precios de los pasajes, etc. Es por esto que la empresa LAN Airlines se ha visto en la necesidad de definir de buena manera sus estrategias comerciales, de modo de afrontar de la mejor forma este mercado cambiante. Dentro de las estrategias comerciales que han identificado a la empresa esta el programa de pasajero frecuente LANPASS, el cual se ha visto en la necesidad de ser valorizado económicamente, de modo de estudiar su factibilidad y utilización a largo plazo.

El programa de pasajero frecuente de LAN Airlines consiste en el regalo de millas virtuales que potencialmente pueden ser cambiadas por pasajes o beneficios en los vuelos (como paso de clase económica a una superior). La cantidad de millas regaladas por vuelo dependen proporcionalmente del largo del trayecto volado, la clase en que se voló (turista, primera, etc.) y el estado del cliente (cliente Premium, Premium Silver, comodoro). Por otro lado, la cantidad necesaria de millas virtuales para ser intercambiadas por pasajes también dependerá del tamaño del trayecto o del precio del pasaje.

El problema fundamental de la aerolínea es como incorporar el valor de este programa con el departamento de *Revenue Management*, pues es este último el encargado de fijar tarifas y asientos, y para ellos, el valor de los pasajes regalados vía LANPASS carecen de valor en sus decisiones.

Es la intención de esta memoria entonces valorizar el programa de fidelización utilizado por la empresa, de modo de determinar cual es el valor agregado que este proporciona en los clientes de las aerolíneas en el momento de decidir su compra.

En este informe se presentarán los antecedentes al estudio, así también como el marco teórico en que se envuelve. En cuanto al trabajo de datos se presentarán claramente los métodos utilizados, resultados obtenidos, discusiones al respecto y conclusiones importantes hacia los objetivos planteados.

2.- Antecedentes y definición del problema

2.1.- Sobre el programa de pasajero frecuente

Un programa de pasajero frecuente es un sistema de fidelización de clientes empleado por muchas aerolíneas a nivel mundial. Generalmente estos programas se basan en la acumulación de kilómetros virtuales proporcionales a la cantidad de kilómetros viajados por los clientes, los cuales pueden ser canjeados por pasajes, beneficios o descuentos. El primer programa de pasajero frecuente creado fue Aadvantage de American Airlines en 1981, el cual fue imitado por las principales líneas aéreas hacia mitad de la década de los 80. Hoy en día, una gran cantidad de aerolíneas de calidad utilizan como método de fidelización dichos programas. En Chile, en el mercado de vuelos internacionales, un 65% de las aerolíneas que operan

rutas internacionales mantienen un programa de pasajero frecuente (Anexo A1).

LANPASS es el Programa de Pasajero Frecuente de LAN creado para premiar la preferencia y lealtad de sus pasajeros. Actualmente, cuenta con más de 2 millones de socios de Chile, Argentina, Perú, Ecuador y Estados Unidos, y son más de 110.000 los que vuelan gratis al año. LANPASS, que cuenta con más de 10 años de funcionamiento, consiste en el traspaso desde la aerolínea hacia sus clientes de kilómetros virtuales, los cuales pueden ser canjeados por pasajes o cambios de categoría dentro de la línea aérea, dentro de las aerolíneas del conglomerado ONE WORLD¹ o en aerolíneas asociadas a LANPASS².

Los socios del Programa LANPASS acumulan kilómetros cada vez que vuelan en LAN, en cualquiera de las líneas aéreas de oneworld o en las líneas aéreas asociadas a LANPASS. Además, existe la posibilidad de acumular kilómetros usando los servicios de las empresas asociadas al programa o por promociones puntuales que ofrece la aerolínea, tales como uso preferente de los servicios *online* de la empresa o promociones de temporada.

Cabe destacar además que de acuerdo a la cantidad de kilómetros viajados en un año, LANPASS cataloga sus clientes en tres categorías: Premium, Premium Silver y Comodoro, los cuales cuentan con distintos beneficios en vuelos, descuentos y acumulación de kilómetros.

2.2.- Razones del estudio

El programa de viajero frecuente LANPASS utilizado por LAN no es simple de analizar, debido a su tremenda envergadura y complejos nexos con los resultados de la empresa. Al ser LANPASS una pieza fundamental en la estrategia comercial de la empresa, LAN se ha visto con la necesidad de evaluar a grandes rasgos el valor agregado del programa, así como los distintos usos preferentes que se dan por parte de sus clientes, características de estos y principales sensibilidades en cuanto a su uso. Todo esto apunta a las posibles decisiones que se desean tomar en cuanto al futuro del programa y su correcto posicionamiento estratégico.

Este estudio le proporcionará una estimación del valor marginal que incorpora LANPASS, es decir, como la existencia de este programa de fidelización afecta el valor presente de cada cliente a través de la recompra de pasajes y preferencia de la aerolínea. Con esta valorización se facilitaría la venta de kilómetros LANPASS hacia el departamento de *Revenue Management*, puesto que ambos departamentos funcionan separadamente y hoy en día, este segundo no incorpora el verdadero valor de este programa en sus operaciones.

¹Oneworld es una alianza global de líneas aéreas que reúne a diez líneas aéreas del mundo, entre las cuales cuenta: American Airlines, British Airways, Cathay Pacific, Finnair, Iberia, Japan Airlines, LAN, Qantas, Maev y Royal Jordanian.

²Las líneas aéreas asociadas a LANPASS son: Alaska, Mexicana, Aeroméxico y TAM

2.3.- Objetivos

2.3.1.- Objetivos principales

El objetivo principal de la memoria es estimar cuantitativamente el valor agregado que genera el programa de pasajero frecuente en los clientes de la aerolínea.

2.3.2.- Objetivos secundarios

Valorizar monetariamente el valor del kilómetro regalado, o bien que porcentaje del pasaje comprado están pagando los clientes debido al programa de pasajero frecuente.

2.4.- Alcances de la memoria

El alcance de la memoria se acota a un orden de magnitud de la cuantificación del valor agregado que proporciona el programa en el valor presente de los clientes, por lo que contempla sólo el estudio primario de los clientes y sus sensibilidades a corto plazo y mediano plazo. El resultado entonces serviría para obtener un norte del valor real de los kilómetros regalados, de modo de incorporar esto a la toma de decisiones del *Revenue Management*.

En este estudio se cuantificará el valor agregado que generan los programas de pasajero frecuente en los clientes y no el valor total que dichos programas generan en los resultados de las aerolíneas.

3.- Metodología y marco teórico

3.1.- Resolución del problema

Debido a la calidad de la información manejada, y a la nula posibilidad de realizar encuestas para el estudio, se procederá a utilizar una metodología que incorpore la información agregada del mercado. Esta metodología contará de cuatro partes específicas.

En la primera fase se analizarán las distintas rutas en la que opera LAN, de modo de identificar una submuestra representativa de estas, las cuales den cuenta de la mayoría de las operaciones de LAN.

En una segunda fase se procederá a estudiar cual es el valor agregado que los clientes sienten por cada aerolínea para las rutas específicas elegidas en la primera etapa.

En la tercera fase, se estudiarán cuales son las variables que afectan dichas valorizaciones, estudiando trabajos anteriores y estimaciones de trabajadores de la empresa. Cabe destacar que dentro de estas variables tiene que incluirse "presencia de Programa de Pasajero Frecuente", pues será esta la que aportará información sobre los objetivos planteados.

En una última etapa se procederá a construir un modelo que vincule dichas variables con las valorizaciones estimadas anteriormente, de modo de cuantificar el cambio marginal que la variable “presencia de Programa de Pasajero Frecuente” pueda ocasionar.

3.2.- Marco teórico

3.2.1.- Sobre el cálculo de las valorizaciones.

En esta memoria, se entenderá como valorización al valor agregado que sienten los clientes hacia las distintas aerolíneas. Esta valorización podrá ser tanto positiva como negativa, y será medida en unidades de dinero (dólares). A continuación se presenta la forma de cálculo de dichas valorizaciones. Primero para el caso simple de una ruta y dos aerolíneas y luego, el caso general.

3.2.1.1.- Caso simple (dos aerolíneas):

Supongamos el caso de dos aerolíneas (A y B) que operan en una ruta específica. Sea el precio de la aerolínea A y B, P_{At} y P_{Bt} en el mes t respectivamente. Sean además los valores agregados que sienten los clientes hacia dichas aerolíneas S_{At} y S_{Bt} (valorizaciones). El precio relativo percibido por algún cliente de dichas aerolíneas en el tiempo t es siguiente:

$$U_{At} = P_{At} - S_{At}$$

$$U_{Bt} = P_{Bt} - S_{Bt}$$

Un cliente racional cualquiera elegirá la aerolínea A si es que $U_{At} < U_{Bt}$. En caso contrario elegiría la aerolínea B.

Los precios son variables fijas, es decir son iguales y conocidos para dos clientes distintos en un mes t cualquiera. Debido a esto, las valorizaciones no pueden ser fijas e iguales para todos los clientes, puesto que cada cliente posee distintas sensibilidades a distintos atributos, además, si dichos valores fuesen fijos e iguales, todos los clientes elegirían la misma aerolínea, es decir, en cada periodo solo una aerolínea se llevaría todos los clientes.

Debido a esto y de modo de agregar la variabilidad en las decisiones se modelan las valorizaciones de la siguiente manera:

$$\begin{aligned} \tilde{S}_{At} &\rightarrow N(S_{At}, \sigma_{At}) \\ \tilde{S}_{Bt} &\rightarrow N(S_{Bt}, \sigma_{Bt}) \end{aligned} \quad \text{Variables aleatorias normales}$$

Conociendo las participaciones de mercado de dichas aerolíneas (X_{At} y X_{Bt}) para un mes t cualquiera, se puede decir que la probabilidad que un cliente haya elegido la aerolínea A en aquel mes t es X_{At} . Es decir:

$$P\left[P_{At} - \tilde{S}_{At} < P_{Bt} - \tilde{S}_{Bt}\right] = X_{At} \Leftrightarrow P\left[P_{At} - N(S_{At}, \sigma_{At}) < P_{Bt} - N(S_{Bt}, \sigma_{Bt})\right] = X_{At}$$

Con esto se obtiene una ecuación para las 4 variables desconocidas $S_{At}, \sigma_{At}, S_{Bt}, \sigma_{Bt}$. Estas variables son: media de la valorización para la aerolínea A, media de la valorización para la aerolínea B, desviación estándar de la valorización de la aerolínea A y desviación estándar de la valorización de la aerolínea B

Debido a la naturaleza del problema, uno de los valores agregados quedará libre, por lo que se asume que uno de ellos será cero. Esto es pues un cliente elige por la diferencia entre los precios relativos, y no por el valor absoluto de este.

Con esto se tiene una ecuación para tres variables ($S_{At}, \sigma_{At}, \sigma_{Bt}$). Al asumir que las desviaciones sean una fracción de los precios ($\sigma_{it} = P_{it}/M$) (M conocido³), se obtienen suficientes ecuaciones para la única variable desconocida (media de la valorización A), por lo que ya se puede cuantificar cual es el valor agregado que en general sienten los clientes para las dos aerolíneas.

3.2.1.2.- Caso General:

- Supuestos del modelo:

- 1) Cada cliente de las aerolíneas posee una valorización (S) hacia cada aerolínea en cada mes t.
- 2) Los clientes eligen aquella aerolínea que maximiza sus utilidades, es decir, aquella que en un mes t cualquiera posee el menor precio relativo ($P - S$). Esto quiere decir que los clientes poseen una función de utilidad lineal.
- 3) Las valorizaciones de los clientes para una aerolínea específica, en un mes t dado son variables aleatorias normales, independiente de las valorizaciones en meses anteriores y a las valorizaciones de las otras aerolíneas.
- 4) La desviación estándar de las valorizaciones en un mes t cualquiera es una fracción del precio de la aerolínea en aquel mes t: $\sigma_{it} = P_{it}/M$
- 5) Los precios a utilizar son fijos e iguales para alguna aerolínea en un mes t cualquiera, y todos los clientes eligen bajo estos mismos precios
- 6) El cambio que produce una alteración de los precios o de la valorización de una aerolínea, sobre las participaciones de mercado del resto de las aerolíneas que compiten directamente, es igual para todas.

³ La constante de escala M a utilizar queda definida mas adelante en la sección 2.2.3

- **Derivación de los S_{it} :**

En el caso de tener en competencia N aerolíneas para una ruta específica, en donde se tomaron los datos para T meses distintos, se define lo siguiente:

$$P = \begin{pmatrix} P_{11} & \dots & P_{1N} \\ \vdots & \ddots & \vdots \\ P_{T1} & \dots & P_{TN} \end{pmatrix} \rightarrow \text{Matriz que contiene los precios de las N aerolíneas durante los T meses de estudio. Las columnas contienen las series de precios para cada aerolínea específica.}$$

$$S = \begin{pmatrix} S_{11} & \dots & S_{1N} \\ \vdots & \ddots & \vdots \\ S_{T1} & \dots & S_{TN} \end{pmatrix} \rightarrow \text{Matriz que contiene las medias de las valorizaciones de las N aerolíneas durante los T meses de estudio.}$$

$$X = \begin{pmatrix} X_{11} & \dots & X_{1N} \\ \vdots & \ddots & \vdots \\ X_{T1} & \dots & X_{TN} \end{pmatrix} \rightarrow \text{Matriz que contiene las participaciones de mercado de las N aerolíneas durante los T meses de estudio.}$$

Además se denotarán las filas y columnas de alguna matriz A cualquiera como \bar{A}_t y \bar{A}_n respectivamente.

Cabe destacar que la matriz P y X es conocida, mientras que la matriz S es aquella que se desea calcular.

Volviendo a los supuestos del modelo, se tiene que para un mes t cualquiera, un cliente observará los precios de las N aerolíneas, de las cuales sentirá también N valorizaciones (una para cada aerolínea) independiente unas de otras. Bajo esto, aquel cliente elegirá la aerolínea que presente el menor precio observado (P-S) entre las N aerolíneas. Esto quiere decir que elegirá:

$$(1) \arg(\min_{i \in \{1..N\}} \{P_{it} - \tilde{S}_{it}\})$$

Al modelar las valorizaciones observadas como $\tilde{S}_{it} \rightarrow N(S_{it}, \frac{P_{it}}{M})$ (supuesto 2 y 3), es claro ver que la solución en (1) estará sujeta a una cierta ley de probabilidad, es decir, la elección de algún cliente cualquiera en el mes t de alguna de las aerolíneas es ahora también una variable aleatoria. Debido a que se conocen las participaciones de mercado de las aerolíneas en los meses de estudio, se puede saber cual es la proporción de clientes que eligieron dicha aerolínea, por lo que se puede crear la siguiente ecuación:

$$(2) P \left[\arg(\min_{i \in \{1..N\}} \{P_{it} - \tilde{S}_{it}\}) = j \right] = X_{tj} \quad \forall j \in \{1, \dots, N\}$$

Esto quiere decir que la probabilidad de algún cliente de elegir la aerolínea j en algún mes t dado (dentro de los meses de estudio) es igual a la participación de mercado de la aerolínea j en el mes t .

Debido a que las valorizaciones son modeladas como variables aleatorias normales, la ecuación en (2) puede escribirse de la siguiente forma:

$$(3) P \left[P_{ij} - N \left(S_{ij}, \frac{P_{ij}}{M} \right) < \underset{\substack{i \in \{1, 2, \dots, N\} \\ i \neq j}}{\text{Min}} \left\{ P_{ii} - N \left(S_{ii}, \frac{P_{ii}}{M} \right) \right\} \right] = X_{ij} \quad \forall j \in \{1, \dots, N\}$$

Como se puede ver, para un tiempo t , se pueden obtener N ecuaciones para las N valorizaciones medias⁴ incógnitas ($\bar{S}_{t\bullet}$). Debido a que la ecuación en (3) no es una ecuación de solución simple (ver ecuación (21)), no es posible llegar a alguna solución exacta del problema, por lo que se procede a encontrar alguna aproximación numérica de mínimo error.

- Primera solución: aproximación lineal

De modo de obtener una primera solución al problema dado en la ecuación (3), se procede a encontrar alguna aproximación lineal de la función (para alguna aerolínea j):

$$(4) f_{ij}(\bar{P}_{t\bullet}, \bar{S}_{t\bullet}) = P \left[P_{ij} - N \left(S_{ij}, \frac{P_{ij}}{M} \right) < \underset{\substack{i \in \{1, 2, \dots, N\} \\ i \neq j}}{\text{Min}} \left\{ P_{ii} - N \left(S_{ii}, \frac{P_{ii}}{M} \right) \right\} \right] \quad j \in \{1, \dots, N\}$$

Esta función es nada más que la probabilidad de elegir la aerolínea j en el mes t dado un vector de precios $\bar{P}_{t\bullet}$ (que contiene los precios de las aerolíneas que operan en la ruta estudiada en el mes t) y un vector de valorizaciones $\bar{S}_{t\bullet}$.

En general se desean conocer aproximaciones lineales de la función en (4) para todas las aerolíneas, es decir, para la función vectorial general:

$$(5) F_t(\bar{P}_{t\bullet}, \bar{S}_{t\bullet}) = \left[f_{t1}(\bar{P}_{t\bullet}, \bar{S}_{t\bullet}) \quad \dots \quad f_{tN}(\bar{P}_{t\bullet}, \bar{S}_{t\bullet}) \right]$$

La función vectorial en (5) es una función que calcula las N participaciones de mercado (N columnas) de las N aerolíneas que operan en la ruta de estudio en el mes t dado un vector de precios $\bar{P}_{t\bullet}$ y un vector de valorizaciones $\bar{S}_{t\bullet}$.

Tomando en cuenta que una aproximación lineal de alguna función $g(\vec{x})$ en torno a los valores $\{\vec{x}\}$ se puede representar como:

$$(6) g(\vec{x}) \approx g(\{\vec{x}\}) + \nabla g(\{\vec{x}\})(\vec{x} - \vec{x}) \quad g(\vec{x}) : R^n \rightarrow R$$

⁴ La valorización media es equivalente a la media de la valorización.

En el caso de la ecuación en (4), se generará una aproximación lineal en torno al precio medio de la ruta y a valorizaciones medias nulas. El precio medio de la ruta es el promedio de los precios de todas las aerolíneas que operan en dicha ruta a lo largo de todos los meses de estudio, es decir, es el promedio de todos los valores de la matriz P definida anteriormente. A este valor se le llamará \bar{P} . Se utiliza este valor \bar{P} para generar la aproximación lineal pues es el valor que está mas cerca al conjunto de precios, además es un valor general para todos los meses en estudio, por lo que la aproximación servirá entonces para todos los t posibles. Se definen entonces los vectores (en los cuales se generará la aproximación lineal) como:

$$\bar{\bar{P}} = [\bar{P} \quad \dots \quad \bar{P}] \quad \text{Vector de N términos}$$

$$\bar{0} = [0 \quad \dots \quad 0] \quad \text{Vector de N términos}$$

$$[\bar{\bar{P}} | \bar{0}] = [\bar{P}, \quad \dots, \quad \bar{P}, \quad 0, \quad \dots, \quad 0] \quad \text{Vector de 2N términos}$$

Con lo que queda:

$$(7) \quad f_{ij}(\bar{P}, \bar{S}) \approx f_{ij}(\bar{\bar{P}}, \bar{0}) + \nabla f_{ij}(\bar{\bar{P}}, \bar{0}) \left[[\bar{P} | \bar{S}] - [\bar{\bar{P}} | \bar{0}] \right]^T \quad \forall j \in \{1, \dots, N\}$$

$$(8) \quad \nabla f_{ij}(\bar{\bar{P}}, \bar{0}) = \left[\frac{\delta f_{ij}(\bar{\bar{P}}, \bar{0})}{\delta P_{i1}}, \quad \dots, \quad \frac{\delta f_{ij}(\bar{\bar{P}}, \bar{0})}{\delta P_{iN}}, \quad \frac{\delta f_{ij}(\bar{\bar{P}}, \bar{0})}{\delta S_{i1}}, \quad \dots, \quad \frac{\delta f_{ij}(\bar{\bar{P}}, \bar{0})}{\delta P_{iN}} \right]$$

Debido a que la función (4) se comporta de igual manera ante cambios marginales de precios y medias de valorizaciones en aerolíneas distintas a la j, se puede deducir lo siguiente:

$$(9) \quad \frac{\delta f_{ij}(\bar{\bar{P}}, \bar{0})}{\delta P_{ii}} = \frac{\delta f_{ij}(\bar{\bar{P}}, \bar{0})}{\delta P_{ik}} \quad \forall i, k \neq j$$

$$(10) \quad \frac{\delta f_{ij}(\bar{\bar{P}}, \bar{0})}{\delta S_{ii}} = \frac{\delta f_{ij}(\bar{\bar{P}}, \bar{0})}{\delta S_{ik}} \quad \forall i, k \neq j$$

Tomando (9) y (10), e incorporándolo a la ecuación en (7), se obtiene:

$$(11) \quad f_{ij}(\bar{P}, \bar{S}) \approx f_{ij}(\bar{\bar{P}}, \bar{0}) + \alpha_1 (\{\bar{P}\}_j - \bar{P}) + \alpha_2 (\{\bar{S}\}_j) + \alpha_3 \left(\sum_{i \neq j}^N \{\bar{P}\}_i - \bar{P} \right) + \alpha_4 \left(\sum_{i \neq j}^N \{\bar{S}\}_i \right)$$

Donde $\{\bar{A}\}_j$ es el elemento j del vector \bar{A} y:

$$(12) \quad \frac{\delta f_{ij}(\bar{\bar{P}}, \bar{0})}{\delta P_{ij}} = \alpha_1 \quad \frac{\delta f_{ij}(\bar{\bar{P}}, \bar{0})}{\delta S_{ij}} = \alpha_2 \quad \frac{\delta f_{ij}(\bar{\bar{P}}, \bar{0})}{\delta P_{i(i \neq j)}} = \alpha_3 \quad \frac{\delta f_{ij}(\bar{\bar{P}}, \bar{0})}{\delta S_{i(i \neq j)}} = \alpha_4$$

Incorporando en la ecuación el hecho de que un cambio marginal en el precio j provocaría un cambio marginal en la participación de j, la cual sería igual a la

suma de de los cambios provocados en las otras participaciones, se obtiene (lo mismo para las valorizaciones):

$$(13) \quad \frac{\delta f_{ij}(\bar{P}, \bar{0})}{\delta P_{ij}} = - \sum_{\substack{k=1 \\ k \neq j}}^N \frac{\delta f_{ik}(\bar{P}, \bar{0})}{\delta P_{ij}} \Leftrightarrow \alpha_3 = - \left(\frac{1}{N-1} \right) \alpha_1$$

$$(14) \quad \frac{\delta f_{ij}(\bar{S}, \bar{0})}{\delta S_{ij}} = - \sum_{\substack{k=1 \\ k \neq j}}^N \frac{\delta f_{ik}(\bar{S}, \bar{0})}{\delta S_{ij}} \Leftrightarrow \alpha_4 = - \left(\frac{1}{N-1} \right) \alpha_2$$

Por último, es claro ver que el valor de la función (4) en los valores en que se está generando la aproximación lineal es $(1/N)$, debido a que da cuenta de la participación de mercado de la aerolínea j en el caso en que todas las aerolíneas que operan presentaran el mismo precio, y todas presentaran una valorización cero, lo que llevaría (bajo los supuestos de este modelo) a que todas obtuvieran la misma participación de mercado. Es decir

$$(15) \quad f_{ij}(\bar{P}, \bar{0}) = 1/N$$

Debido a que la aproximación lineal se hizo en torno a un precio medio que rescata todos los meses en estudio, las constantes $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)$ son iguales para cada mes.

Utilizando las ecuaciones (13), (14) y (15) se plantea el siguiente sistema lineal $(\tilde{F}_i(\bar{P}, \bar{S}))$ aproximación lineal de $F_i(\bar{P}, \bar{S})$ en (5)):

$$(16) \quad \tilde{F}_i(\bar{P}_i, \bar{S}_i) = \bar{X}_i$$

Al intentar solucionar el sistema (16), se obtienen infinitas soluciones, esto es debido a que una fila es linealmente dependiente del resto. Como se comentó anteriormente, esto se debe a que los clientes eligen debido a la diferencia entre sus distintas utilidades, y no por su valor absoluto. Para encontrar una solución del sistema se fija alguna valorización en cero. Al hacer esto, la solución del sistema queda de la siguiente forma:

$$(17) \quad S_{ii} = \frac{1}{N\alpha_2} [(N-1)(X_{ii} - X_{iN}) - N\alpha_1(P_{ii} - P_{iN})] \quad \forall i \in \{1, 2, \dots, N-1\}$$

$$S_{iN} = 0$$

Ya con esto, sólo falta conocer las variables (α_1, α_2) de modo de obtener alguna primera solución para el vector \bar{S}_i .

- **Obtención numérica de los valores α_1 y α_2**

Utilizando las ecuaciones en (12) es posible llegar a los valores requeridos. Para esto es necesario definir la función $f_{ji}(\bar{P}, \bar{S})$ para luego calcularles sus derivadas. Sea:

$$(18) \quad g(x, \mu_{ij}, \sigma_{ij}) = \frac{1}{\sqrt{2\pi\sigma_{ij}^2}} e^{-\frac{(x-\mu_{ij})^2}{2\sigma_{ij}^2}} \rightarrow \text{Función de densidad de una variable aleatoria normal de media y desviación estándar } (\mu_{ij}, \sigma_{ij}).$$

$$(19) \quad G(z, \mu_{jt}, \sigma_{jt}) = \int_{-\infty}^z \frac{1}{\sqrt{2\pi\sigma_{jt}^2}} e^{-\frac{(x-\mu_{jt})^2}{2\sigma_{jt}^2}} dx \rightarrow \text{Probabilidad de que una variable aleatoria normal de media } \mu_{jt} \text{ y desviación estándar } \sigma_{jt} \text{ sea menor a } z.$$

$$(20) \quad P(\underset{i \in \{1,2,3,\dots,N\}}{\text{Min}} \{N(\mu_{it}, \sigma_{it})\} < R) = 1 - \prod_{\substack{i=1 \\ i \neq j}}^N (1 - G(R, \mu_{it}, \sigma_{it})) = W_j(R, \{\mu_{it}\}_{\substack{i \in \{1,2,3,\dots,N\} \\ i \neq j}}, \{\sigma_{it}\}_{\substack{i \in \{1,2,3,\dots,N\} \\ i \neq j}})$$

La ecuación en (20) refiere a la probabilidad de que el mínimo de N-1 variables aleatorias normales independientes de medias $\{\mu_{it}\}_{\substack{i \in \{1..N\} \\ i \neq j}}$ y desviaciones estándares $\{\sigma_{it}\}_{\substack{i \in \{1..N\} \\ i \neq j}}$ sea menor a un valor R.

Uniendo (18), (19) y (20)

$$(21) \quad f_{ij}(\bar{P}, \bar{S}) = \int_{-\infty}^{\infty} \left(1 - W \left(z, \{\bar{P} - \bar{S}\}_{\substack{i \neq j \\ i \in \{1..N\}}}, \frac{\{\bar{P}_{i\bullet}\}_{\substack{i \neq j \\ i \in \{1..N\}}}}{M} \right) \right) g \left(z, \{\bar{P} - \bar{S}\}_j, \frac{\{\bar{P}_{i\bullet}\}_j}{M} \right) dz$$

La función (21) da cuenta de la probabilidad de un cliente cualquiera de elegir la aerolínea j (dados los vectores de precios y medias de valorizaciones en el mes t para todas las líneas aéreas). Esto significa que utilizando la ecuación (5), un vector de precios $\bar{P}_{i\bullet}$ y un vector de valorizaciones $\bar{S}_{i\bullet}$ se puede llegar a las participaciones de mercado que el modelo predeciría bajo los vectores mencionados anteriormente.

Finalmente utilizando el software matemático MAPLE, se pueden calcular las derivadas de f, en torno al precio promedio y valorizaciones cero discutidos anteriormente, obteniendo de esta manera los α_1 y α_2 explicados en la ecuación (12).

- Mejora de la solución lineal

Utilizando la solución descrita en (17) se obtiene un primer conjunto de solución para el vector $\bar{S}_{i\bullet}$ (el cual se denominará $\bar{S}_{i\bullet}^0$). Como estos pueden no ser la mejor solución para el problema (debido a que son calculados vía una aproximación lineal), se procede a mejorar dicha solución.

Para esto, se utilizó una búsqueda de mejores soluciones tipo punto fijo, partiendo de la solución $\bar{S}_{i\bullet}^0$. El método se inicia calculando el valor de

$F_t(\overline{P}_{t\bullet}, \overline{S}_{t\bullet}^0)$, utilizando la ecuación (21) y el programa MAPLE. Este vector (si es que la solución fuese perfecta) debería ser igual al vector de participaciones de mercado para el mes t ($\overline{X}_{t\bullet}$). Como claramente la aproximación lineal no será óptima, se procede a calcular el error asociado a esta primera solución (error de predicción del mes t en la iteración cero: EP_t^0).

$$(22) \quad EP_t^0 = \oplus \left| F_t(\overline{P}_{t\bullet}, \overline{S}_{t\bullet}^0) - \overline{X}_{t\bullet} \right| \quad \text{donde } \oplus |\vec{A}| \text{ refiere a la suma de los valores absolutos de los elementos del vector } \vec{A}$$

Este error consiste en la suma de los valores absolutos de las diferencias entre el vector $F_t(\overline{P}_{t\bullet}, \overline{S}_{t\bullet}^0)$ y $\overline{X}_{t\bullet}$, es decir, en cuanto se equivocó nuestra solución inicial en predecir las participaciones correspondientes. Una vez obtenido este error, se guarda para considerarlo mas adelante.

Luego de calcular las diferencias entre las participaciones reales y las modeladas para cada aerolínea, se cambia la primera solución de las valorizaciones $\overline{S}_{t\bullet}^0$ (para llegar a un segundo vector de soluciones $\overline{S}_{t\bullet}^1$) de la siguiente forma:

- Si la participación de mercado modelada para alguna aerolínea es menor que la real observada en el mes t , entonces la valorización media de aquella aerolínea se aumenta en un 10% (esto debido a que una menor participación a la real se debe a que su valorización media es menor a la verdadera).
- De la misma forma, si la participación modelada es mayor a la real, se disminuye la valorización media de aquella aerolínea en un 10%.

Lo anterior se puede expresar de la siguiente forma:

$$(23) \quad \left\{ \begin{array}{l} \left\{ \overline{S}_{t\bullet}^1 \right\}_i = (1+10\%) \left\{ \overline{S}_{t\bullet}^0 \right\}_i \Leftrightarrow \left\{ F_t(\overline{P}_{t\bullet}^0, \overline{S}_{t\bullet}^0) \right\}_i - \left\{ \overline{X}_{t\bullet} \right\}_i < 0 \\ \left\{ \overline{S}_{t\bullet}^1 \right\}_i = (1-10\%) \left\{ \overline{S}_{t\bullet}^0 \right\}_i \Leftrightarrow \left\{ F_t(\overline{P}_{t\bullet}^0, \overline{S}_{t\bullet}^0) \right\}_i - \left\{ \overline{X}_{t\bullet} \right\}_i > 0 \end{array} \right\}$$

Con esto, se obtiene un nuevo vector $\overline{S}_{t\bullet}^1$, el cual es distinto al anterior. Luego se prueba este nuevo vector solución en el programa MAPLE y se calcula el nuevo error de predicción EP_t^1 . Si este error es menor al calculado con la solución anterior, este nuevo vector $\overline{S}_{t\bullet}^1$ queda como la nueva solución, si no, se vuelve al primer vector solución. En el caso que la nueva solución sea peor a la anterior, se procede a efectuar de nuevo un cambio de solución, pero ahora en vez de cambiar las soluciones en un 10%, se cambian sólo en un 5% (la mitad que la vez anterior). En el caso de que el error sea menor al anterior, en la próxima iteración las soluciones se cambian de la misma forma que la vez anterior.

Este método se repite unas 10 veces (no mas debido al largo tiempo que toma el programa MAPLE en resolver lo pedido), llegando a una solución mejorada.

A continuación se presenta un grafo de cómo mejora la solución inicial (medida en como disminuye su error de predicción medio, es decir $\frac{EP_i}{N}$) en una muestra dada.

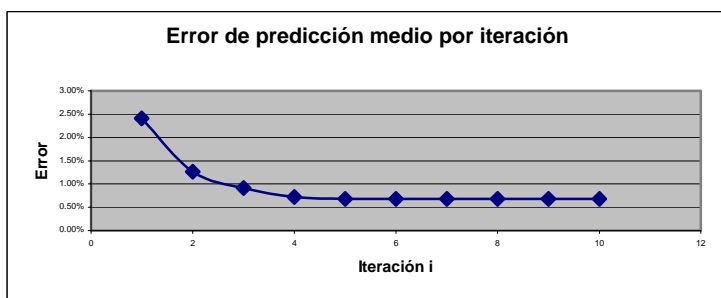


Ilustración 1: Error absoluto medio vs. Iteración

El error de predicción medio es la media de los valores absolutos de las diferencias entre las participaciones de mercado reales para un mes t , y las participaciones modeladas con las valorizaciones $\bar{S}_{i,t}$. Como se puede ver, se producen disminuciones del error considerables en las primeras iteraciones. Luego de la sexta iteración, en general, no se encontraron mejores soluciones. Esto es importante, pues con esto se puede inferir (por el método de iteración) que los valores de valorizaciones medias al que se llega no difieren en mas que un 0,3125%⁵ al valor exacto que se podría llegar (iteración 6)

En general se encontraron soluciones con errores medios de predicción (diferencia entre la participación de mercado real y la modelada) cercanos al 0.63%, lo que es mas que aceptable.

Una vez obtenido el vector solución final $\bar{S}_{i,t}$, se procede a “normalizar”, es decir, a cada elemento del vector se le resta el promedio de todos los elementos del mismo, de modo de no obtener siempre una aerolínea (en este caso la aerolínea N) con valorización cero.

Con esta metodología, se pueden calcular las valorizaciones medias de los clientes hacia las aerolíneas utilizando sólo la información agregada del precio por periodo de cada empresa aérea y su respectiva participación de mercado para cada ruta elegida.

⁵ Debido a que las valorizaciones se mueven en un $\frac{10\%}{2^{n-1}}$ en la iteración n . Si no se encuentran soluciones mas

allá de la iteración 6, la diferencia con la solución real es a lo mas $\frac{10\%}{2^5} = 0,3125\%$

- Obtención de la constante de escala M

Utilizando la metodología anterior, se procedió a calcular las distintas valorizaciones, por periodo y por aerolínea para un conjunto de datos. Al probar distintos valores del parámetro M (supuesto de que las desviaciones son $1/M$ del precio de cada aerolínea) se llegó a los siguientes resultados.

- Para valores del parámetro M entre 3 y 9, las diferencias entre las valorizaciones calculadas es menor al 3%.
- Para M=10 no es posible encontrar soluciones pues la desviación estándar es muy pequeña.
- Para M=2 (desviación estándar igual a la mitad del precio) los valores de las valorizaciones se hacen mayores, llegando hasta un 6% de diferencia con las calculadas para M=9.

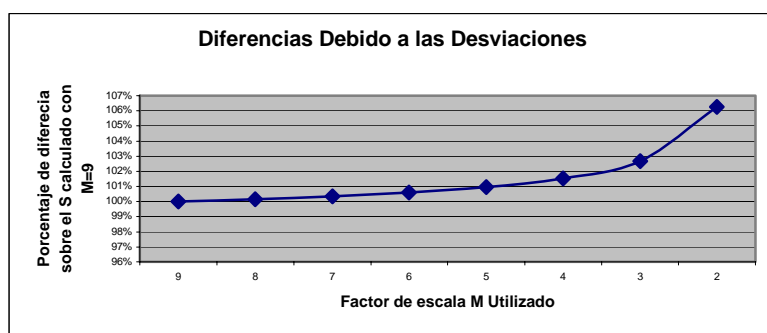


Ilustración 2: Diferencias respecto a la constante M

Analizando lo obtenido se puede observar que las diferencias entre las valorizaciones obtenidas para valores de M entre 3 y 9 son mínimas. Por otro lado, al utilizar un valor de M igual a dos, las diferencias son mayores (6%). Tomando en cuenta que las valorizaciones obtenidas son generalmente un 50% del precio de la aerolínea, al tomar M igual a 2 se estaría trabajando con desviaciones estándar iguales o mayores a los valores promedios. Esto generaría mucha incerteza en los resultados, y puede ser la razón por la que se notan tantas diferencias. Debido a esto, en adelante, se procederá a utilizar un valor de M igual a 6, de modo de trabajar con incertezas ni muy mayores ni menores a las valorizaciones obtenidas y por ser este un valor representativo del intervalo entre 3 y 9.

- Comentarios sobre los supuestos

Los supuestos 1 y 2 del modelo explicado anteriormente dan cuenta de la racionalidad de cómo los clientes toman sus decisiones. Estos suponen que cada uno de los pasajeros que vuelan en las aerolíneas estudiadas toma sus decisiones bajo los mismos términos y conocimientos que el resto, y elige aquella que mejor responda a sus utilidades personales. Tal vez en la realidad estos supuestos no se cumplan cabalmente, pero debido al nivel de precios de los pasajes, y a la poca competencia en las rutas, estos supuestos no difieren mucho de la realidad.

El supuesto 3, que da cuenta de la aleatoriedad en las valorizaciones de los clientes hacia las aerolíneas es un supuesto fuerte pues considera que la distribución de probabilidades de dichas variables es normal e independiente. Este supuesto, al igual que el supuesto 4, son supuestos considerados de modo de simplificar el modelo, y poder llegar a alguna solución. Se utilizó una distribución normal, pues es una distribución ampliamente conocida, continua y simétrica.

Por último, el supuesto 5 esta basado en la calidad de la información con que se trabaja. En el caso de este estudio se pudo trabajar solamente con los precios spot de las aerolíneas, es decir, una especie de precio base con los que trabajan al mes para las distintas tarifas utilizadas, y si bien no son los precios exactos que finalmente se cancelaron, si reflejan una magnitud de orden de cuanto cobra cada aerolínea y las diferencias entre ellas.

Se recopiló información de precios sobre 4 distintas tarifas con las que operan las aerolíneas. Estas tarifas se distribuyen generalmente en económica flexible, económica, Business flexible, Business y primera. Dentro de estas tarifas, la más general (aquella que mejor refleja el nivel de precios de la aerolínea) es la tarifa económica. Debido a esto, en este estudio se trabajará con la tarifa económica ida que ofrecen las aerolíneas (YOW). Si bien trabajar sólo con esta tarifa es una generalización, lo importante en este estudio es llegar a alguna valorización sobre lo pagado como tarifa, es decir, mas allá de saber cuanto vale en dinero el hecho de incorporarse o no a algún programa de pasajero frecuente, es saber cual es el porcentaje de cambio sobre el precio pagado.

A modo de ejemplo, se muestra a continuación para la ruta Santiago-Madrid el precio medio para 16 meses de estudio para tres tipos de tarifas distintas: Económica, Business flexible y Business en US\$:

Santiago-Madrid	LAN	A. Argentinas	AIR COMET	IBERIA
YOW	2560,11	1482,33	1066,67	2524,83
DOW	4028,94	2140,67	1822,22	4000,65
COW	4231,00	2399,00	1933,33	4166,83

Cuadro 1: Comparación de precios para distintas tarifas

Se puede ver de mejor manera la diferencia de precios con respecto a LAN en el siguiente cuadro:

Santiago-Madrid	LAN	A. Argentinas	AIR COMET	IBERIA
YOW	100,00%	57,90%	41,66%	98,62%
DOW	100,00%	53,13%	45,23%	99,30%
COW	100,00%	56,70%	45,69%	98,48%

Cuadro 2: Comparación porcentual de precios para distintas tarifas

Aquí se puede apreciar claramente que la diferencia porcentual con respecto a la tarifa que presenta LAN es prácticamente la misma. Comportamientos similares se pueden apreciar en otras rutas (Anexo A2).

3.2.2.- Sobre el modelo a utilizar:

Una vez calculada las medias de las valorizaciones de las aerolíneas por rutas y mes, se generará un modelo lineal con variables a definir, entre las cuales contará la pertenencia o no a algún programa de pasajero frecuente. Es decir, se desea generar un modelo de la forma:

$$(24) S_{ij} = B_0 + B_1 Y_{ij1} + B_2 Y_{ij2} + \dots + B_k Y_{ijk}$$

En donde las variables Y_{ijk} son variables observables de cada aerolínea i , en cada mes t y para cada ruta j , tales como Frecuencia de salidas, Pertenencia a Programa de pasajero frecuente, Tipo de ruta, Pertenencia a otros programas, Tiempo de vuelo promedio (escalas), etc. Los B_i son constantes propias del modelo, las cuales deben ser calculadas.

Con esto, si es que la regresión es significativa, se podrá saber como afecta el hecho de poseer o no un programa de pasajero frecuente en las apreciaciones de los clientes.

Un punto importante para tomar en cuenta sobre los modelos de regresiones lineales es que las constantes B_i no pueden ser calculadas exactamente, puesto que generalmente los modelos no son perfectos y hay un cierto error vinculado. Lo más cercano para calcular estos B_i es estimarlos vía el método de mínimos cuadrados, el cual propone lo siguiente:

- Debido a que el modelo presenta un cierto error, la regresión en (24) se puede representar de la siguiente forma:

$$(25) S_{ij} = \beta_0 + \beta_1 Y_{ij1} + \beta_2 Y_{ij2} + \dots + \beta_k Y_{ijk} + \mu_{ij}$$

En donde los β_i son los estimadores de B_i , y μ_{ij} es el error asociado a que dichas constantes no siempre explican exactamente la variable dependiente (valorizaciones medias).

El procedimiento de los mínimos cuadrados minimiza el cuadrado del error, calculando las constantes β_i . En esta memoria, estos cálculos se harán a través del software estadístico SPSS.

Un valor importante que será usado para estimar que tan bien se ajusta el modelo a la realidad es el coeficiente de determinación R^2 . Este coeficiente determina la bondad de ajuste de la regresión, o que tan bien se ajusta la recta de regresión muestral a los datos. El R^2 queda determinado por la cantidad de varianza que el modelo de regresión lineal representa en la variable dependiente, el cual se calcula de la siguiente manera:

$$(26) R^2 = 1 - \frac{\sum_{t,i,j} (\mu_{ij})^2}{\sum_{t,i,j} (S_{t,i,j} - \bar{S}_{t,i,j})^2}$$

Donde \overline{S}_{ij} es el promedio de las valorizaciones medias calculadas.

Debido a que los valores β_i a calcular son estimadores de las constantes B_i , estos incorporan una cierta aleatoriedad, es decir, tienen una media y desviación estándar. Esto es importante, puesto que nos dice que tan significativos son cada uno de estos β_i . Mediante un simple test de hipótesis, se puede saber si estas variables son significativas o no, o sea, aportan información relevante al modelo. Estos tests de hipótesis serán una importante herramienta en el estudio, pues nos dirán que variables incorporar o no al modelo principal.

La razón principal de realizar estos modelos es saber cuanto aporta a las valorizaciones medias el β_i asociado a los programas de pasajeros frecuentes.

4.- Desarrollo

4.1.- Estudio y selección de las rutas

Debido a la presencia dominante de LAN en las rutas domésticas, la poca competencia percibida y la falta de información, el estudio se focalizará en las rutas internacionales.

Utilizando los datos obtenidos en la JAC (Junta Aeronáutica Civil), se estudiaron las participaciones de las distintas aerolíneas en las diferentes rutas. Se escogieron aquellas rutas internacionales en donde se presentaran un gran número de salidas, competencia importante con otras aerolíneas y variaciones de participación no tan grandes. Además, se escogieron rutas características, es decir, que representen un mercado específico. Luego de estudiar los datos y recibiendo recomendaciones de la gente de LAN⁶, se escogieron las siguientes rutas:

Ruta	Kilómetros	Aerolíneas Principales	Participación
Santiago - Madrid	10.710	LAN	38,55%
		AEROLINEAS ARGENTINAS (AR)	4,25%
		AIR COMET (AC)	14,92%
		IBERIA (IB)	43,93%
Santiago - Miami	6.653	LAN	54,31%
		AEROLINEAS ARGENTINAS (AR)	1,08%
		AMERICAN AIRLINES (AA)	43,93%
		TACA (TAC)	0,68%
Santiago - México	6.589	LAN	75,73%
		AEROMEXICO (AM)	24,27%
Santiago - B. Aires	1.139	LAN	72,20%
		AEROLINEAS ARGENTINAS (AR)	13,34%
		GOL	10,81%
		AIR CANADA (CC)	3,65%
Santiago - Sao Paulo	2.612	LAN	47,41%
		TAM	25,20%

⁶ Gabriela Godoy: Jefe de Proyectos Comerciales y Control de Gestión de RM

	GOL	16,69%
	SWISS (SW)	10,70%

Cuadro 3: Rutas seleccionadas para el estudio

Ruta	Mercado	Pasajeros por año 2007
Santiago - Madrid	Intercontinental	355.881
Santiago - Miami	Larga distancia EEUU	262.483
Santiago - México	Larga distancia Sudamérica	169.299
Santiago - B. Aires	Corta distancia	1.172.093
Santiago - Sao Paulo	Media distancia	552.925

Cuadro 4: Mercados objetivos y pasajeros transportados

De las 56 rutas internacionales con salidas desde Santiago en que opera LAN, estas cinco rutas escogidas corresponden al 51.13% de los pasajeros transportados en el año 2007 (total de pasajeros 4.914.490). De las rutas rezagadas, ninguna aporta más del 5% de los pasajeros transportados, a excepción de la ruta Santiago-Lima, la cual no se utilizó pues correspondía al mismo mercado que la ruta Santiago-Buenos Aires (Resumen de rutas internacionales año 2007 en Anexo B1).

4.2.- Cálculo de las valorizaciones

Una vez elegidas las rutas en donde se realizará el estudio se procedió a calcular las valorizaciones de los clientes hacia las aerolíneas por rutas, mes (año 2007-2008) y aerolínea. Utilizando la información sobre precios y participaciones y utilizando la metodología explicada en 2.2.1 se obtuvo lo siguiente:

Ruta	Precio Medio (US \$)	Meses	Valorización Media (US \$)			
Santiago - Madrid	1.951	16	LAN	AR	AC	IB
			745	-683	-791	728
Santiago - Miami	2.020	18	LAN	AR	AA	TAC
			396	-575	266	-87
Santiago - México	2.248	18	LAN	AM		
			216	-216		
Santiago - B. Aires	243	18	LAN	AR	GOL	CC
			117	24	-45	-96
Santiago - Sao Paulo	594	12	LAN	TAM	GOL	SW
			54	153	-103	-103

Cuadro 5: Valorizaciones medias por ruta

En el cuadro 2 se muestra el promedio de las valorizaciones medias a lo largo de sus meses de estudio. Esto quiere decir, que en todas las rutas se calculó una valorización media para cada mes dentro los meses de estudio y luego a estos valores se les sacó un promedio. A continuación, y a modo de ejemplo, se muestra la serie de valorizaciones medias para la ruta Santiago-Madrid, para las cuatro aerolíneas principales.

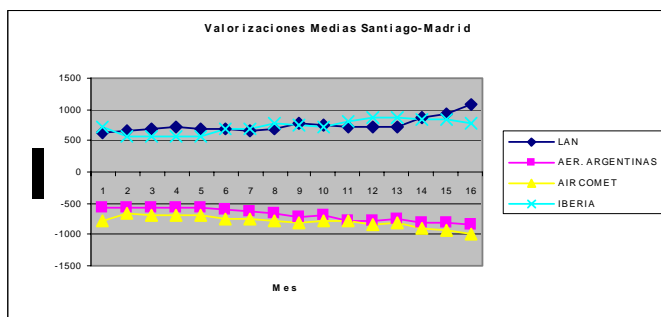


Ilustración 3: Serie de valorizaciones para la ruta Santiago-Madrid

Como se puede ver en el gráfico 3, existen dos aerolíneas con valorizaciones mayores, y dos con valorizaciones menores. En el caso de esta ruta en particular se puede ver que las valorizaciones estimadas de las aerolíneas mas grandes se alejan en el tiempo, esto es por que en este caso en particular se presento una alza de tarifas en las líneas aéreas mas grandes, implicando un aumento en las valorizaciones de los clientes al permanecer las participaciones de mercado en los mismos niveles.

Los errores de predicción comentados en 2.2.1 que se generaron en el cálculo de las valorizaciones anteriores son los siguientes:

Ruta	Error Medio de predicción
Santiago - Madrid	0,68%
Santiago - Miami	0,70%
Santiago - México	0,61%
Santiago - B. Aires	0,54%
Santiago - Sao Paulo	0,61%

Cuadro 6: Error medio por ruta

El error medio de predicción da cuenta de la diferencia promedio entre la participación de mercado real de las aerolíneas con la participación modelada bajo los supuestos del modelo y las valorizaciones calculadas. Esto quiere decir por ejemplo que si la aerolínea LAN obtuvo una participación de mercado real de un 38% el mes de abril del 2008, el modelo predijo que LAN obtuvo una participación de un 38% más menos 0,68%. Como se puede ver, el error de pronóstico es mínimo.

La serie general con las valorizaciones medias calculadas se puede ver a continuación:

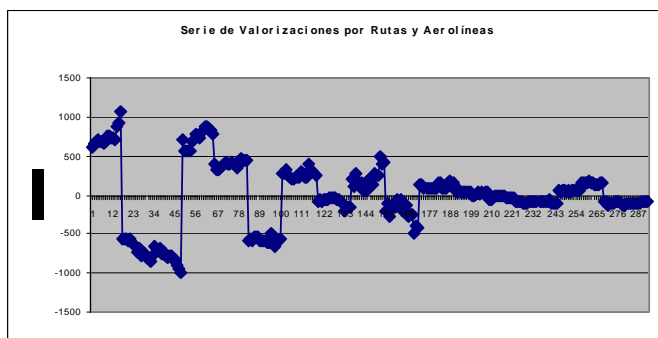


Ilustración 4: Serie general de valorizaciones

El gráfico da cuenta de todas las valorizaciones medias por ruta y aerolínea, esto quiere decir que los puntos mostrados se ordenan de la siguiente forma: Primero las valorizaciones de la ruta Santiago-Madrid, partiendo por las valorizaciones de LAN, luego AER. ARGENTINAS, luego AIR COMET y finalmente IBERIA. Después la serie sigue con la ruta Santiago-Madrid, partiendo por LAN, AER. ARGENTINAS, etc. Cabe destacar que esta no es una serie de tiempo, pues si bien las valorizaciones tienen una característica temporal, esta se toma solamente durante los meses de estudio de aquella ruta específica. La forma de la serie se da principalmente debido a que se presentan las rutas de mayor valor (mayor precio) primero, por lo que sus valorizaciones son de mayor valor absoluto, por esto la forma decreciente (en valor absoluto) que se puede apreciar.

Esta serie es de importancia pues será utilizada para comparar los distintos modelos a utilizar mas adelante, de modo de poder elegir cual es el de mayor exactitud.

4.3.- Selección de las variables explicativas

Luego de leer una serie de papers sobre investigaciones similares⁷ se llegó a la determinación de utilizar las siguientes variables como base del modelo:

- 1) Kilómetros Vuelo: Distancia de la ruta en kilómetros.
- 2) Competencia: Número de aerolíneas en la ruta específica.
- 3) Directo: Si la aerolínea opera con vuelos directos en la ruta.
- 4) Principal (FLAG): Variable que toma en cuenta si la aerolínea es la aerolínea principal (de mayor tamaño) en su país de origen, y opere en una ruta que contemple su país.
- 5) FFP: Variable que describe si la aerolínea posee un programa de pasajero frecuente (frequent flyer program).

⁷ “Airline choice, switching costs and frequent flyer program”, “Predicting airline choices a decision support”, “Modeling the repatronage behavior of business airline travelers”, “Impact of frequent flyer programs”. Detalles bibliográficos en la sección de bibliografía

- 6) Asociados: Si la aerolínea pertenece a algún plan conjunto con LAN (One World o líneas asociadas. Detalle en página 6).
- 7) LAN: Variable que describe si la aerolínea es LAN o no.
- 8) FREC%: porcentaje de la frecuencia de vuelos al mes que tiene una cierta aerolínea en comparación con la frecuencia total de vuelos en una ruta en particular.
- 9) FREC a: Frecuencia de salida al mes de la aerolínea en una ruta particular.
- 10) FREC c: Frecuencia de salida de la competencia al mes en una ruta particular.
- 11) Opera: Esta variable en particular describe si la aerolínea en cuestión opera la ruta específica como parte de su negocio principal o no. Esta variable describe aerolíneas como SWISS o AIR CANADA, las cuales son grandes aerolíneas, pero el fuerte de su negocio no está en las rutas chilenas.
- 12) Ruta: Por último se incorporan variables dummies por rutas.

Con estas variables ya definidas se procede a realizar el modelo de predicción. Cabe señalar que las variables más importantes en este estudio son aquellas que describen la presencia o no de algún programa de pasajero frecuente. En este caso tanto las variables FFP como la variable asociados describen la utilización o no de algún programa de pasajero frecuente, debido a que las asociaciones de LAN vienen acompañadas siempre de la cooperación con otros programas de pasajero frecuente.

En el Anexo B2 se pueden ver estadísticas descriptivas de las variables mencionadas anteriormente.

4.4.- Modelo de predicción

Utilizando las variables descritas en 3.3 se generaron distintos modelos que pudiesen incorporar la variabilidad de las valorizaciones medias descritas anteriormente. De todos los modelos posibles, se rescataron 3, los cuales eran aquellos que presentaban una mejor predicción de los datos utilizados y una mayor explicación de la varianza de estos mismos.

A continuación se presentarán dichos modelos, sus errores de predicción, comparación entre ellos y conclusión sobre cual modelo elegir.

4.4.1.- Primer modelo: Ln(P-S)

Este primer modelo toma como variable dependiente el logaritmo natural del precio de cierta aerolínea en una ruta y mes específica y le resta su valorización media calculada. La ventaja de esto es que modela el precio “relativo” percibido por los clientes de las aerolíneas, además de normalizar

los valores, pues al sumarle el precio, se notan mucho menos las diferencias de valor absoluto entre las valorizaciones.

El modelo descrito anteriormente es el siguiente:

$$(27) \quad \ln(P_{ij} - S_{ij}) = \beta_1 + \beta_2 V_{ij}^1 + \dots + \beta_m V_{ij}^m + \mu$$

En donde (P_{ij}, S_{ij}) se refieren al precio y la valorización media en el mes t, en la aerolínea i, en la ruta j respectivamente. Por otro lado la variable V_{ij}^k se refiere al valor que toma la variable k-esima (dentro de las variables comentadas en 3.3) en el mes t, en la aerolínea i, en la ruta j. Los β_i son constantes del modelo a predecir, mientras que la variable μ es una variable de error que refleja la imperfección del modelo.

Utilizando el programa SPSS se realiza una regresión lineal para el modelo en (27), llegando a los siguientes resultados:

Variable	Coefficiente	Error Estándar	t	Sigma
Constante	652,31%	2,04%	320,19	0,00%
Opera	3,53%	1,78%	1,98	4,78%
FFP	2,46%	1,41%	1,74	8,21%
Asoc	-7,12%	1,97%	-3,61	0,04%
Madrid	122,33%	1,71%	71,61	0,00%
Miami	124,42%	1,74%	71,38	0,00%
México	154,79%	1,99%	77,84	0,00%
Buenos Aires	-90,70%	1,50%	-60,39	0,00%
Frec%	-62,73%	6,45%	-9,72	0,00%
N	292			
R^2	0,993			

Cuadro 7: Variable Dependiente: Ln(P-S)

Como se puede ver en el cuadro 7, no se muestran todas las variables descritas en 3.3, esto debido a que no todas las variables dieron estadísticamente significativas. Esto quiere decir que muchas de las variables utilizadas en el modelo tenían una alta probabilidad de ser nulas, por lo que su presencia en el modelo final no era necesaria. La bondad de ajuste R^2 del modelo, que infiere la cantidad de varianza que es descrita por el modelo es bastante alto (0,993), lo que habla muy bien del modelo logarítmico, pero no es razón necesaria para tomar una decisión sobre que modelo utilizar.

También se puede observar que las variables FFP y Opera poseen un sigma no menor (4,78% y 8,21%), pero debido a que la exclusión de estas variables disminuía la bondad de ajuste, además de la importancia que es mantener la variable FFP dentro del modelo se procedió a dejar dichas variables dentro del estudio. De todas formas dichas variables son significativas a un nivel del 10%.

A continuación se muestra la serie de la variable μ (error), así también como su distribución a lo largo de la serie.

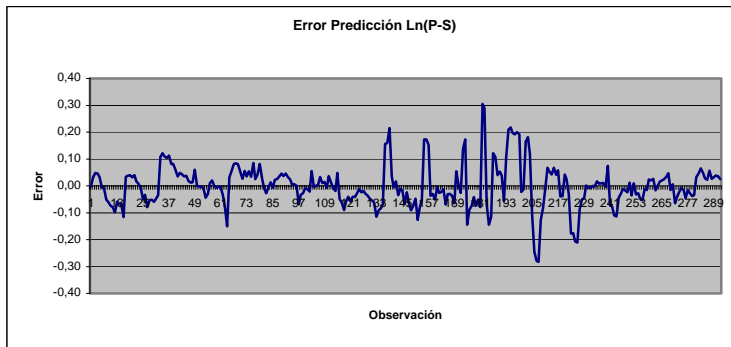


Ilustración 5: Errores del modelo Ln(P-S)

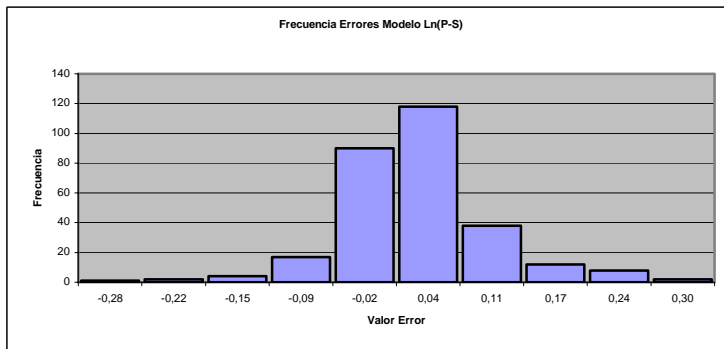


Ilustración 6: Frecuencia de errores modelo Ln(P-S)

La desviación estándar de la variable error μ es de 0,08, y analizando la ilustración 6 se puede inferir que posee una distribución parecida a la Normal. De hecho, al realizar un test de normalidad de Kolmogorov-Smirnov, este arroja una buena probabilidad de que las diferencias entre esta distribución y una normal sean debido a alteraciones aleatorias ($\sigma=0,191$ ver anexo C1).

Si bien este es un modelo de bajo error, la variable que se desea predecir en este estudio es la valorización media, y no su logaritmo arreglado. Es por esto que hay que analizar también cual es el error involucrado al intentar predecir la valorización. Para rescatar el valor de la valorización media hay que transformar el modelo en (25) de la siguiente forma (aplicando exponencial a ambos lados):

$$(28) \quad S_{ij} = P_{ij} - e^{\beta_1 + \beta_2 V_{ij}^1 + \dots + \beta_m V_{ij}^m + \mu}$$

Al comparar esta nueva formula con la serie original en las valorizaciones medias presentada en 3.2 (Ilustración 4) se observa lo siguiente:

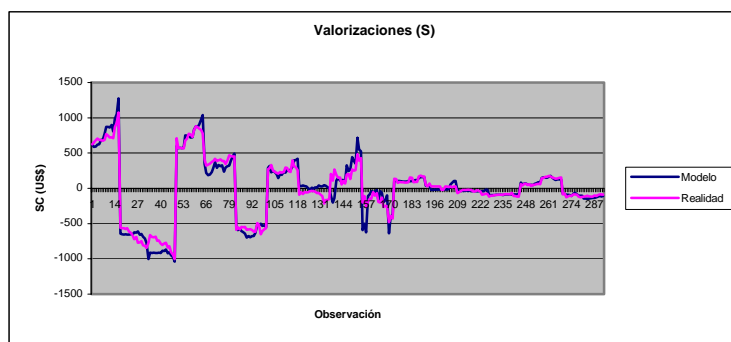


Ilustración 7: Comparación de S del modelo 1

Como se puede ver hay un buen ajuste a la curva original. Si se calcula la serie de diferencias entra el modelo y la serie original se puede llegar a un error de predicción de las valorizaciones. Este error presenta ahora una desviación estándar de 95,36 US\$. Ahora si se tiene una mejor noción del error que conlleva este modelo, pues la desviación está en dólares.

4.4.2.- Segundo modelo: S/Pmedio

Este segundo modelo caracteriza las valorizaciones medias calculadas divididas por el precio medio entre las aerolíneas para una ruta y mes específicos. Lo bueno de este modelo es que arroja como resultado final el porcentaje de las valorizaciones con respecto a los precios medios, resultado que se puede utilizar directamente para sacar conclusiones sobre la valorización del programa de pasajero frecuente (el primer modelo debía ser transformado de modo de poder sacar alguna conclusión interesante). Otra ventaja de utilizar esta fórmula es que normaliza de buena manera lo estudiado, ya que como se había descrito anteriormente, las valorizaciones de ciertas rutas eran mucho mayores en valor absoluto, lo que cambia al utilizar porcentajes sobre el precio de la ruta. Este modelo se define de la siguiente forma.

$$(29) \frac{S_{ij}}{P_{ij}^{med}} = \beta_1 + \beta_2 V_{ij}^1 + \dots + \beta_m V_{ij}^m + \mu$$

En este nuevo modelo, las variables, constantes y valorizaciones se definen de la misma forma que en el modelo (27), a excepción de la variable P_{ij}^{med} que refiere al precio medio para el mes t en la ruta j.

Utilizando SPSS para realizar una regresión lineal al modelo (29) se obtienen los siguientes parámetros:

Variable	Coficiente	Error Estándar	t	Sigma
Constante	-62,37%	5,25%	-11,88	0,00%
Opera	-29,37%	2,88%	-10,20	0,00%
FFP	19,61%	1,61%	12,14	0,00%
Asoc	-11,41%	3,69%	-3,09	0,22%
Madrid	17,19%	4,01%	4,28	0,00%
Miami	10,79%	3,95%	2,73	0,66%
México	-17,80%	4,82%	-3,69	0,03%

frec%	39,92%	10,28%	3,88	0,01%
frec a	0,16%	0,03%	5,10	0,00%
frec c	-0,05%	0,02%	-2,59	0,99%
Directo	43,41%	3,95%	10,97	0,00%
N	292			
R^2	0,851			

Cuadro 8: Variable Dependiente: S/Pmed

Como se puede observar, hay mas variables que dieron estadísticamente significativas, y dentro de las que quedaron son todas significativas al 1%. El ajuste de este modelo es bueno ($R^2=0,851$) pues explica alrededor de un 85% de la varianza original.

El error μ que presenta esta regresión se caracteriza a continuación.

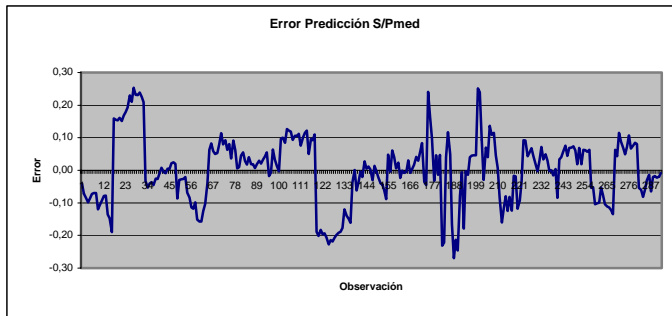


Ilustración 8: Errores del modelo S/Pmed

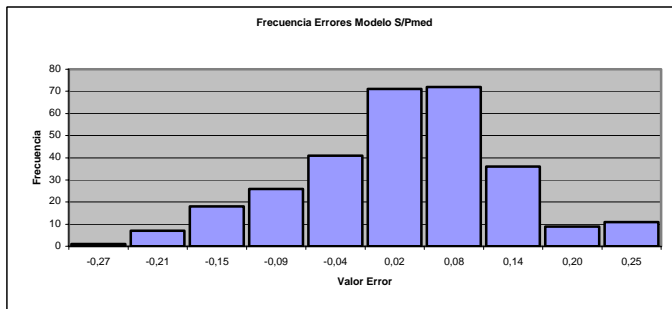


Ilustración 9: Frecuencia de errores del modelo S/Pmed

La ilustración 8 muestra la serie del error del modelo, mientras que la ilustración 9 caracteriza la distribución de este. Esta distribución no es normal bajo el test de Kolmogorov-Smirnov (ver anexo C2). La desviación estándar de este error es de un 10,24%. Si se toma en cuenta que el precio medio de toda la serie (todos los meses, todas las rutas, todas las aerolíneas) es cercano a 1360 US\$, entonces se puede inferir que esta desviación estándar de 10,24% podría afectar la valorización en 139,26 US\$.

4.4.3.- Tercer modelo: S

En este tercer modelo se pretende caracterizar simplemente a la variable valorización media. Lo bueno de este modelo particular es que su resultado es

el más directo, ya que se caracteriza en dinero. Una desventaja para esta caracterización es que no se puede disminuir la diferencia en valores absolutos entre las distintas rutas para sus valorizaciones. El modelo a construir es el siguiente:

$$(30) S_{ij} = \beta_1 + \beta_2 V_{ij}^1 + \dots + \beta_m V_{ij}^m + \mu$$

Utilizando el programa SPSS se llegó a los siguientes resultados:

Variable	Coefficiente	Error Estándar	t	Sigma
Constante	-1272,377789	86,96	-14,63	0,00%
Opera	-184,8051155	47,69	-3,87	0,01%
FFP	255,2653002	34,97	7,30	0,00%
Asoc	-125,4549201	61,17	-2,05	4,12%
Madrid	416,3961039	66,45	6,27	0,00%
Miami	352,5660681	65,40	5,39	0,00%
México	-287,0821296	79,87	-3,59	0,04%
frec%	1184,23499	170,40	6,95	0,00%
frec a	-1,531520939	0,52	-2,92	0,38%
Directo	862,1270635	65,54	13,15	0,00%
N		292		
R^2		0,829		

Cuadro 9: Variable Dependiente: S

En este modelo también se excluyeron ciertas variables que no eran estadísticamente significativas, y por lo tanto, no aportaban a la explicación de la varianza de los datos originales. La bondad de ajuste R^2 en esta instancia es de un 0,829, la cual es buena. La serie de errores y su distribución se muestran a continuación:

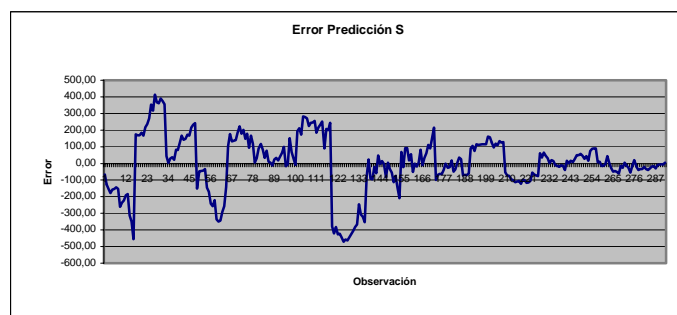


Ilustración 10: Errores del modelo S

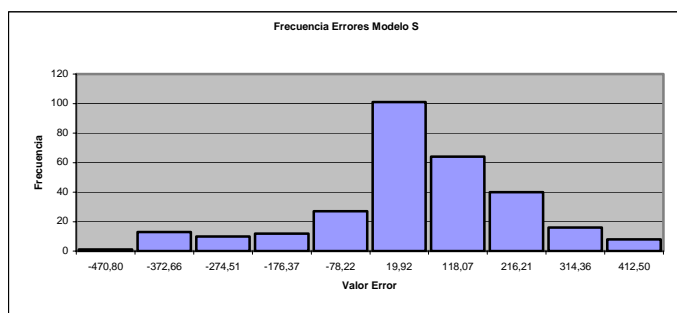


Ilustración 11: Frecuencia de errores del modelo S

La desviación estándar del error μ en este modelo es de 169,7 US\$. Esta distribución no es normal bajo el test de Kolmogorov-Smirnov (ver anexo C3).

4.4.4.- Elección del modelo a utilizar

Analizando los tres modelos descritos anteriormente se puede construir el siguiente cuadro:

Modelo	Ln(P-S)	S/Pmed	S
R^2	0,993	0,861	0,829
Error medio	0	0	0
Desviación estándar del error	0,08	10,24%	169,70
Normalidad de residuos	Si	No	No
Desviación del error en US\$	65,36	139,27	169,70
Directo	No	Si	Si
Lineal	No	Si	Si

Cuadro 10: Comparación entre los modelos

Si bien se puede ver claramente que el modelo logarítmico posee el mayor ajuste R^2 , este no puede ser un parámetro de comparación, pues esta valor sólo es comparable cuando la variable dependiente es la misma en todos los modelos. Como en este caso la variable dependiente es distinta, se deben mirar otras propiedades de los modelos para lograr llegar a alguna conclusión.

El error medio en todos los casos es cero, lo que es esperado pues generalmente las medias de los errores en regresiones lineales son nulas. Por otro lado analizar las desviaciones estándar de los errores no llevaría a mucho, pues no están medidas todas en el mismo valor. Un valor que si puede ser comparado es la desviación estándar del error llevado a dólares. En el caso del modelo S esto es simple, pues los valores predichos ya vienen evaluados en dinero. Para los otros dos modelos el traspaso no es tan complicado (El cual se explica en 3.4.1 y 3.4.2). Analizando estos valores se puede observar claramente que el modelo logarítmico es el de menor error, de hecho, su desviación estandar es practicamente la mitad que los otros modelos (en el anexo C4 es posible ver como estos tres modelos predicen las valorizaciones).

Otro punto importante que rescatar es que los resultados para los modelos S/Pmed y S son directos para la conclusión del estudio, pues no deben ser transformados para saber cuales son las repercusiones en dinero en caso de

tener o no programa de pasajero frecuente. Por otro lado el modelo logarítmico no es directo, y su transformación genera incremento en los errores.

Por último, un punto muy importante a considerar es que los modelos S y S/Pmed son lineales con respecto a sus variables explicativas, esto quiere decir que los cambios marginales que suceden al poseer o no programa de pasajero frecuente son iguales, independiente del valor que tomen las otras variables explicativas. Por otro lado, en el modelo logarítmico, los cambios marginales en la variable de programa de pasajero frecuente depende del valor que tomen el resto de sus variables, esto es por ejemplo, que el cambio que podría provocar la variable del programa es diferente en el caso de ser una aerolínea con rutas directas o indirectas. Esto último permite al modelo logarítmico estimar las distintas elasticidades que producen las distintas variables en juego. Este tipo de resultados genera conclusiones más reales, pues nunca las variables predicativas están demasiado desligadas.

Considerando esto último, además del menor error en cuanto a dinero, se considerará como mas apto el modelo logarítmico.

3.4.5.- Resultados sobre el modelo elegido

Utilizando el modelo elegido, se desea saber cual es la diferencia en las valorizaciones medias de los clientes con respecto a alguna aerolínea que contenga o no programa de pasajero frecuente, es decir, cuanto de la valorización de un cliente es debido a la operación de un programa de pasajero frecuente. Para esto, se trabajará con la fórmula transformada del modelo 1, pues esta da cuenta del valor de valorización media.

$$(31) S_{ij} = P_{ij} - e^{C + \beta_1 opera_{ij} + \beta_2 FFP_{ij} + \beta_3 asoc_{ij} + \beta_4 madrid_{ij} + \beta_5 miami_{ij} + \beta_6 mexico_{ij} + \beta_7 b.aires_{ij} + \beta_8 frec\%_{ij} + \mu}$$

En este caso específico, se desea saber como cambia la variable S cuando las variables FFP (programa de pasajero frecuente) y Asociados (variable que da cuenta si la aerolínea esta asociado a otro programa de pasajero frecuente) son uno (presencia) o cero (ausencia de la variable).

En el caso que haya presencia de estas dos variables, la valorización S se modela de la siguiente forma (dejando como variables los demás inputs del modelo):

$$(32) S = P - e^{C + \beta_1 opera + \beta_2 \times 1 + \beta_3 \times 1 + \beta_4 madrid + \beta_5 miami + \beta_6 mexico + \beta_7 b.aires + \beta_8 frec\% + \mu_1}$$

Como se puede ver las variables FFP y Asociados pasaron a valer 1. En el caso de no presentar estas variables, estas valdrían cero, y la valorización quedaría representada de la siguiente forma:

$$(33) S = P - e^{C + \beta_1 opera + \beta_4 madrid + \beta_5 miami + \beta_6 mexico + \beta_7 b.aires + \beta_8 frec\% + \mu_2}$$

Restando (32) con (33), para saber cual es la diferencia entre poseer o no estas variables se llega a lo siguiente:

$$(34) \Delta S = e^{C+\beta_1 opera+\beta_4 madrid+\beta_5 miami+\beta_6 mexico+\beta_7 b.aires+\beta_8 frec\%} (e^{\mu_2} - e^{\beta_2+\beta_3+\mu_1})$$

Donde μ_1 y μ_2 son errores de igual comportamiento, equivalentes al error del modelo logarítmico explicado en 3.4.1 (ilustración 5 y 6) y de comportamiento Normal.

Asumiendo el comportamiento normal de las constantes β , y utilizando el cuadro de varianzas-covarianzas de dichas constantes para el modelo logarítmico:

	Opera	frec%	FFP	Asoc	Madrid	Miami	México	BAires
Opera	0,04%	0,05%	-0,01%	0,00%	0,01%	0,01%	0,00%	0,00%
frec%	0,05%	0,42%	-0,01%	-0,10%	0,04%	0,04%	-0,01%	0,00%
FFP	-0,01%	-0,01%	0,03%	-0,15%	0,00%	-0,01%	0,00%	0,00%
Asoc	0,00%	-0,10%	-0,15%	0,04%	-0,01%	-0,01%	0,00%	0,00%
Madrid	0,01%	0,04%	0,00%	-0,01%	0,03%	0,02%	0,02%	0,01%
Miami	0,01%	0,04%	-0,01%	-0,01%	0,02%	0,03%	0,02%	0,01%
México	0,00%	-0,01%	0,00%	0,00%	0,02%	0,02%	0,04%	0,01%
BAires	0,00%	0,00%	0,00%	0,00%	0,01%	0,01%	0,01%	0,02%

Cuadro 11: Varianzas-Covarianzas de las constantes del modelo logarítmico

Tomando en cuenta además que lo que se quiere medir es como afectan estas variables en la aerolínea LAN, entonces las variables del modelo deben ser aquellas que sean propias de esta aerolínea, es decir, la variable Opera es siempre cero para LAN puesto que todas las rutas están en zonas donde opera como primer negocio esta línea aérea. Las variables binarias por rutas no pueden ser elegidas a priori puesto que estas tomarán valor de acuerdo a la ruta que se desee estudiar (lo mismo con la variable frecuencia)

Además, utilizando la propiedad de que si X es una variable aleatoria de distribución normal $X \rightarrow N(\mu_X, \sigma_X^2)$, entonces e^X es una variable aleatoria de media $e^{\mu_X + \frac{\sigma_X^2}{2}}$ y de desviación estándar $e^{2\mu_X + \sigma_X^2} (e^{\sigma_X^2} - 1)$ (ver anexo C5) se obtienen los siguientes resultados:

$$(35) E(\Delta S) = e^{A_1 + \frac{A_2}{2}} - e^{B_1 + \frac{B_2}{2}}$$

Con:

$$(36) A_1 = C + \beta_4 madrid + \beta_5 miami + \beta_6 mexico + \beta_7 b.aires + \beta_8 frec\%$$

$$(37) A_2 = VAR(C) + VAR(\beta_4)madrid^2 + VAR(\beta_5)miami^2 + VAR(\beta_6)mexico^2 + VAR(\beta_7)b.aires^2 + VAR(\beta_8)frec\%^2 + VAR(\mu_2) + 2COVAR(\beta_8, \beta_4)frec\% \square madrid + 2COVAR(\beta_8, \beta_5)frec\% \square miami + 2COVAR(\beta_8, \beta_6)frec\% \square mexico + 2COVAR(\beta_8, \beta_7)frec\% \square b.aires$$

$$(38) B_1 = C + \beta_2 + \beta_3 + \beta_4 \text{madrid} + \beta_5 \text{miami} + \beta_6 \text{mexico} + \beta_7 \text{b.aires} + \beta_8 \text{frec\%}$$

$$(39) B_2 = \text{VAR}(C) + \text{VAR}(\beta_2) + \text{VAR}(\beta_3) + \text{VAR}(\beta_4) \text{madrid}^2 + \text{VAR}(\beta_5) \text{miami}^2 + \text{VAR}(\beta_6) \text{mexico}^2 + \text{VAR}(\beta_7) \text{b.aires}^2 + \text{VAR}(\mu_2) + 2\text{COVAR}(\beta_2, \beta_3) + 2\text{COVAR}(\beta_2, \beta_4) \text{madrid} + 2\text{COVAR}(\beta_2, \beta_5) \text{miami} + 2\text{COVAR}(\beta_2, \beta_6) \text{mexico} + 2\text{COVAR}(\beta_2, \beta_7) \text{b.aires} + 2\text{COVAR}(\beta_2, \beta_8) \text{frec\%} + 2\text{COVAR}(\beta_3, \beta_4) \text{madrid} + 2\text{COVAR}(\beta_3, \beta_5) \text{miami} + 2\text{COVAR}(\beta_3, \beta_6) \text{mexico} + 2\text{COVAR}(\beta_3, \beta_7) \text{b.aires} + 2\text{COVAR}(\beta_3, \beta_8) \text{frec\%} + 2\text{COVAR}(\beta_4, \beta_5) \text{frec\%} \square \text{madrid} + 2\text{COVAR}(\beta_4, \beta_6) \text{frec\%} \square \text{miami} + 2\text{COVAR}(\beta_4, \beta_7) \text{frec\%} \square \text{b.aires}$$

Tomando en cuenta la fórmula en (35), y utilizándola para los siguientes escenarios (cinco rutas distintas):

Ruta	Opera	frec% media	Madrid	Miami	México	BAires	Precio Medio
Madrid	0	35,33%	1	0	0	0	1951
Miami	0	41,32%	0	1	0	0	2020
México	0	64,89%	0	0	1	0	2248
Baires	0	52,67%	0	0	0	1	242
Spaulo	0	36,61%	0	0	0	0	589

Cuadro 12: Datos propios de LAN por rutas

Se obtienen las siguientes diferencias en valorizaciones, debido a la presencia de algún programa de pasajero frecuente:

Ruta	E(ΔS) US\$	E(ΔS)/Precio medio	Desviación ΔS US\$
Madrid	85,17	4,36%	104,55
Miami	83,97	4,16%	143,70
México	98,46	4,38%	168,06
Baires	9,09	3,76%	15,57
Spaulo	24,80	4,21%	42,61

Cuadro 13: Valorización debido al programa de pasajero frecuente

Como se puede ver en el Cuadro 13, el efecto que conlleva poseer algún programa de pasajero frecuente bordea el 4% del precio medio del ticket en una cierta ruta. Un punto importante que no se puede dejar de lado es el hecho que la desviación de esta diferencia es mucho mayor que los valores de esta. Esto se debe principalmente a que en el momento de restar dos valores de este modelo, se incorpora el doble del error dentro de la fórmula, llevando a obtener grandes desviaciones con respecto a la media. La desviación media obtenida en el cuadro 13 es de aproximadamente 85 US\$

Los resultados obtenidos al utilizar los modelos rezagados (27) y (28) son más fáciles de calcular, debido a la linealidad comentada anteriormente. Debido a que para calcular los incrementos marginales de poseer o no algún programa de pasajero frecuente es necesario restar dos escenarios posibles,

los errores arrastrados en estos modelos serán cercanos a la raíz de dos veces la desviación estándar de los residuos de los modelos, por lo tanto, se obtienen los siguientes resultados:

Modelo	$E(\Delta S)$ US\$	$E(\Delta S)/\text{Precio medio}$	Desviación ΔS US\$
S/ Pmed	111,52	8,20%	204,54
S	129,81	9,54%	249,40

Cuadro 14: Resultados sobre los modelos rezagados

Como se puede ver las valorizaciones correspondientes a los programas de pasajeros frecuentes no varían dependiendo la ruta. Los resultados de estos modelos dieron diferente al modelo escogido. Una razón de esto podría ser el gran error de predicción que poseen estos modelos (mas de 2,5 veces la desviación del modelo logarítmico). Por otro lado se puede observar que estos modelos no se diferencian mucho uno de otro en cuanto a resultados, lo que probablemente se debe a que ambos recogen las mismas variables de predicción, a diferencia del modelo logarítmico.

5. - Conclusiones

Suponiendo que los pasajeros de las distintas aerolíneas que operan en Chile eligen la línea aérea en que viajaran de modo racional, es decir, utilizando aquella que mejor los beneficie, se logró establecer un modelo que pudiere predecir la probabilidad de algún cliente de escoger alguna de las aerolíneas de interés. Este modelo se basa exclusivamente en los precios observados por los clientes y el valor sentimental que estos asignan a las distintas líneas aéreas (medidos en dolares). Este valor que no se refleja en los precios, da cuenta de las diferenciaciones percibidas por los clientes hacia sus aerolíneas, y mide de buena manera como la forma de operar de estas influye en las apreciaciones de los pasajeros.

Este valor asignado por los clientes, denominado valorización en esta memoria, varía de acuerdo a una serie de variables propias de las aerolíneas y rutas en que operan. Dentro de estas variables se encuentra aquella que da cuenta si es que la línea aérea opera algún programa de pasajero frecuente, la cual afecta positivamente la valorización percibida por los pasajeros hacia las aerolíneas que contengan dichos programas.

Estableciendo una serie de modelos que predigan de mejor manera la variable dependiente (valorización) de acuerdo a las variables independientes mencionadas anteriormente se llegó a una serie de conclusiones importantes para este estudio.

- Las valorizaciones de los clientes hacia las aerolíneas respectivas dependen de variables tales como: frecuencia de salida, tipo de ruta, programas de pasajero frecuente, asociaciones entre líneas aéreas, vuelos directos y centro de operación de manera significativa. Esto quiere decir que dichas variables explican la varianza de la variable dependiente en un nivel mayor al 99% (para el mejor modelo de predicción).
- Las variables que se relacionan con la utilización de algún programa de pasajero frecuente (tales como utilización de dicho programa o asociaciones con otros programas de líneas aéreas externas) predicen

que estos generan un incremento positivo en las valorizaciones de los pasajeros, generando un aumento en las participaciones de mercado.

- Utilizando como modelo de predicción aquel que de menor error en las estimaciones es posible concluir que dichos incrementos aludidos a la operación de programas de pasajero frecuente bordea el 4,17% del valor del pasaje medio pagado para el caso específico de LAN Airlines, es decir, en dicho porcentaje los pasajeros valorizan el hecho de que LAN posea el programa de pasajero frecuente LANPASS.
- El valor agregado que genera el programa de pasajero frecuente LANPASS depende de la ruta en que opere y el valor del pasaje, siendo menor para el caso de rutas mas cortas con tarifas menores.

Es importante destacar que estos resultados incorporan un cierto margen de error no menor al intentar predecir cuantitativamente el valor agregado que genera el programa de pasajero frecuente. Estos errores difieren de acuerdo al modelo utilizado, al igual que sus resultados. El modelo que mejor explica las valorizaciones de los clientes hacia las aerolíneas obtiene un error (desviación estandar) cercano a los 85 US\$, muy por debajo de los 204 US\$ y 249 US\$ de los otros modelos. Si bien los dos modelos rezagados obtienen resultados similares entre ellos y diferentes al modelo elegido (estiman en 8,2% y 9,5% de la tarifa media la valorización marginal de los programas de pasajero frecuente), su error estandar es tan grande que dichos resultados pueden deberse mas a la variabilidad de los modelos y al hecho que ambos incorporan las mismas variables de predicción. Bajo esta misma línea, el modelo logarítmico (aquel de menor error) incorpora en sus resultados a las demas variables explicativas, lo que no sucede con los otros dos modelos que predicen que la valorización marginal de algún pasajero hacia una aerolínea debido a su programa de pasajero frecuente no depende de la ruta en que opera dicha línea aérea, si posee vuelos directo, de sus frecuencias de salida, etc.

El error estándar mencionado anteriormente podría disminuirse utilizando en las regresiones mayores y mejores resultados, esto es, incorporando más rutas al estudio y obteniendo una serie de precios más cercana a la realidad. La primera solución es más factible, pero significaría un gran uso de tiempo de cálculo pues la forma de calcular las valorizaciones no es inmediata. La segunda solución requiere de información no pública de todas las aerolíneas en estudio, lo cual es difícil de obtener al menos que los resultados del estudio beneficiaran a todas las líneas aéreas

Se concluye en este estudio entonces que el programa de pasajero frecuente LANPASS afecta la decisión de compra de los clientes de dicha aerolínea, generando un valor agregado de hasta 4,38% del valor del pasaje. Con esto se puede obtener entonces un norte del valor total que LANPASS genera en los resultados totales de la empresa, evaluando de aquella manera el valor del kilómetro regalado. No se calcula en esta memoria este valor, pues LANPASS es más que la valorización marginal de los clientes, pues posee también una serie de vínculos con empresas de otros rubros, las cuales generan distintos tipos de valor. Además, este estudio calcula el valor agregado hoy, es decir, lo que los clientes sienten a corto plazo. La exclusión de LANPASS generaría también consecuencias a largo plazo, las cuales podrían llegar a ser bastante significativas.

6.- Bibliografía

CARLSSON, F. y LOFGREN, A. 2004. Airline choice, switching costs and frequent flyer programs, Working Papers in Economics no. 123. Suecia, Gothenburg University, Department of Economics.

CHIN, A. 2002. Impact of frequent flyer programs on the demand for air travel. Singapore, National University of Singapore, Department of Economics. Journal of Air Transportation Vol. 7. No 2.

GARCÍA, J. y REINARES, P. 2007. The importance of rewards in the management of multisponsor loyalty programmes. Madrid, Universidad Rey Juan Carlos, Area de Marketing. Journal of Database Marketing & Customer Management 15, 37-48.

GREENE, W. Análisis Econométrico, Tercera edición.

GUJARATI, D. Econometría, Cuarta edición.

HARRIS, J. y UNCLES, M. 2007. Modeling the repatronage behavior of business airline travelers. Journal of Service Research 9; 297.

HUANG, Z., SUZUKI, Y., ZHAO, H. y ZHU, D. 2006. Predicting airline choices: a decision support. Perspective and alternative approaches. EEUU, Iowa State University and Pennsylvania State University.

7.- Material complementario

7.1.- Anexo A

7.1.1.- Anexo A1: Aerolíneas que operan rutas internacionales en Chile

Aerolínea	Programa de pasajero frecuente
Aerolíneas Argentinas	SI
American Airlines	SI
Aerolínea Principal	NO
Aeromexico Aerovías de México	SI
Air Canada	SI
Air Comet	NO
Air France - KLM	SI
Avianca Aerovías del continente americano	SI
Copa Compañía panameña de aviación	SI
Cubana de Aviación	NO
Delta Airlines	SI
Gol	NO
Lan Airlines	SI
Lineas aéreas de España Iberia	SI
Pluna Líneas aéreas uruguayas	NO
SKY Airlines	NO
Swiss International Airlines	SI
Transportes aéreos del continente americano	NO
Transportes aéreos del MERCOSUR	SI
VARIG	SI

7.1.2.- Anexo A2: Diferencias entre tarifas para las rutas de interes. Tarifas Económica Flexible, Económica y Business Premium.

Tarifas en US \$ Santiago-Miami				
Tarifa	LAN	Aerolíneas Argentinas	American Airlines	TACA
BOW	1384.06	1177.89	1508.77	786.45
YOW	2058.47	1709.00	2059.27	1548.10
JOW	3126.78	2775.67	3116.33	2142.07

Diferencias con respecto a la tarifa LAN Santiago-Miami				
Tarifas	LAN	Aerolíneas Argentinas	American Airlines	TACA
BOW	100.00%	85.10%	109.01%	56.79%
YOW	100.00%	83.02%	100.04%	75.20%
JOW	100.00%	88.77%	99.67%	68.51%

Tarifas en US\$ Santiago-C. De Mexico		
Tarifas	LAN	Aeromexico
BOW	2063.67	2006.56
YOW	2269.06	2226.61
JOW	3512.50	3268.17

Diferencias con la tarifa LAN Santiago-C. De Mexico		
Tarifas	LAN	Aeromexico
BOW	100.00%	97.23%
YOW	100.00%	98.13%
JOW	100.00%	93.04%

Tarifas en US \$ Santiago-B. Aires				
Tarifas	LAN	Aerolíneas Argentinas	Gol	Air Canada
BOW	188.50	140.42	159.22	139.5
YOW	303.89	258.67	215.22	237.5
JOW	687.50	612.56	SIN TARIFA	486

Diferencias con respecto a la tarifa LAN Santiago-B. Aires				
Tarifas	LAN	Aerolíneas Argentinas	Gol	Air Canada
BOW	100.00%	74.49%	84.47%	74.01%
YOW	100.00%	85.12%	70.82%	78.15%
JOW	100.00%	89.10%		70.69%

7.2.- Anexo B

7.2.1.- Anexo B1: Resumen de rutas internacionales año 2007. Pasajeros movilizados.

Origen	Destino	Total Movilizado	Participación
SANTIAGO	BARILOCHE	16.087	0,33%
SANTIAGO	BUENOS AIRES	1.172.097	23,85%
SANTIAGO	CORDOBA	102.167	2,08%
SANTIAGO	MENDOZA	135.895	2,77%
SANTIAGO	PUERTO MADRYN	135	0,00%
SANTIAGO	RIO GALLEGOS ARG	6	0,00%

SANTIAGO	ROSARIO, ARG.	20.316	0,41%
SANTIAGO	SALTA, ARG.	69	0,00%
SANTIAGO	TUCUMAN	124	0,00%
SANTIAGO	USHUAIA, ARG.	10.487	0,21%
SANTIAGO	AREQUIPA	195	0,00%
SANTIAGO	LIMA	373.459	7,60%
SANTIAGO	COCHABAMBA	43	0,00%
SANTIAGO	LA PAZ	28.458	0,58%
SANTIAGO	SANTA CRUZ, BOL.	14.090	0,29%
SANTIAGO	BELO HORIZONTE	226	0,00%
SANTIAGO	BRASILIA	112	0,00%
SANTIAGO	FLORIANAPOLIS	27.417	0,56%
SANTIAGO	IGUASSU, BRASIL	785	0,02%
SANTIAGO	NAVEGANTES BRA.	3	0,00%
SANTIAGO	PORTO ALEGRE	8.304	0,17%
SANTIAGO	PORTO SEGURO	642	0,01%
SANTIAGO	RIO DE JANEIRO	204.698	4,17%
SANTIAGO	SALVADOR BAHIA	6.276	0,13%
SANTIAGO	SAO PAULO	552.925	11,25%
SANTIAGO	ASUNCION	37.483	0,76%
SANTIAGO	MONTEVIDEO	111.453	2,27%
SANTIAGO	PTA. DEL ESTE	4.261	0,09%
SANTIAGO	GUAYAQUIL	30.251	0,62%
SANTIAGO	QUITO	51.283	1,04%
SANTIAGO	BOGOTA	110.735	2,25%
SANTIAGO	CARACAS	50.098	1,02%
SANTIAGO	PTO.STANLEY	5.399	0,11%
SANTIAGO	SAN JOSE, C.R.	25.890	0,53%
SANTIAGO	LA HAVANA	18.327	0,37%
SANTIAGO	C.DE PANAMA	128.272	2,61%
SANTIAGO	PUNTA CANA, R.D.	18.427	0,37%
SANTIAGO	SANTO DOMINGO	5.260	0,11%
SANTIAGO	C. DE MEXICO	169.299	3,44%
SANTIAGO	CANCUN, MEX.	12.148	0,25%
SANTIAGO	TORONTO	47.254	0,96%
SANTIAGO	ATLANTA	113.361	2,31%
SANTIAGO	DALLAS, US.	109.398	2,23%
SANTIAGO	LOS ANGELES	93.488	1,90%
SANTIAGO	MIAMI	262.483	5,34%
SANTIAGO	NUEVA YORK	66.654	1,36%
SANTIAGO	FRANKFURT	33.097	0,67%
SANTIAGO	BARCELONA ESP.	18.164	0,37%
SANTIAGO	LAS PALMAS, ESP.	345	0,01%
SANTIAGO	MADRID	355.881	7,24%
SANTIAGO	PARIS	113.153	2,30%
SANTIAGO	ROMA	2.172	0,04%
SANTIAGO	ZURICH	36.320	0,74%
SANTIAGO	SYDNEY	75.414	1,53%
SANTIAGO	AUCKLAND N.Z.	70.141	1,43%
SANTIAGO	PAPEETE	16.801	0,34%
TOTAL	GENERAL	4.914.490	100,00%

7.2.2.- Anexo B2: Estadísticas descriptivas de las variables a utilizar por aerolíneas

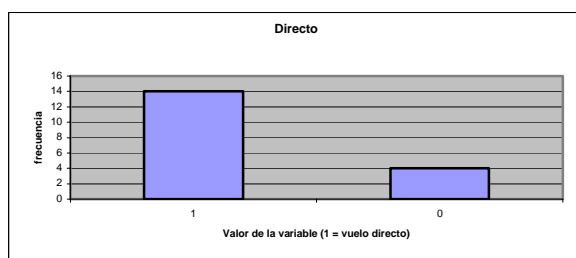
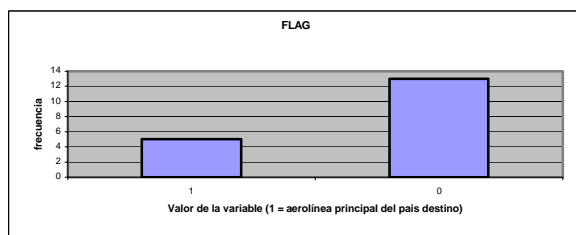
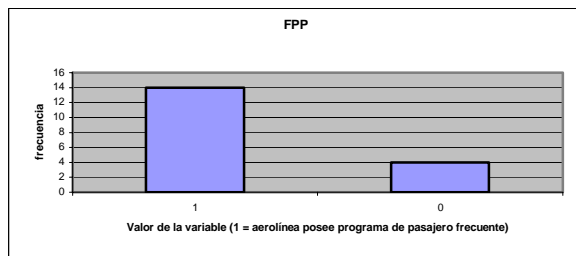
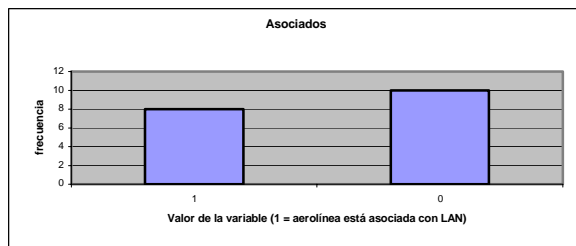
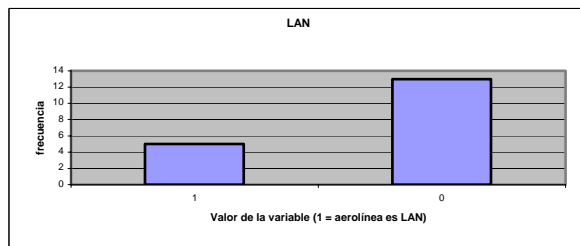
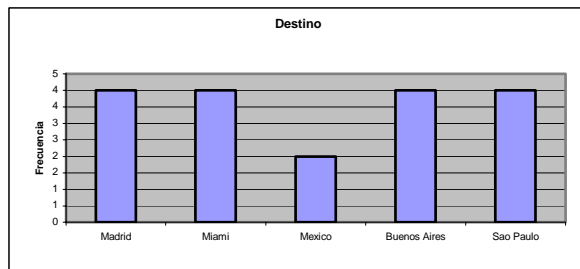
- Resumen de las variables continuas por aerolíneas y rutas:

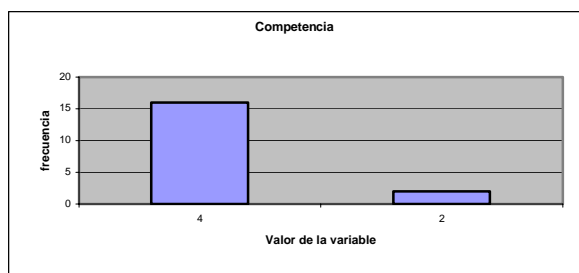
Aerolínea - Ruta	Kilómetros Vuelo	Comp	FREC% media	FREC a media	FREC c media
LAN Madrid	10710	4	35,13%	30,50	56,31
AR Madrid	10710	4	15,69%	13,63	73,19
AC Madrid	10710	4	14,11%	12,25	74,56
IB Madrid	10710	4	35,06%	30,44	56,38
LAN Miami	6653	4	41,04%	39,17	56,28
AR Miami	6653	4	8,32%	7,94	87,50
AA Miami	6653	4	37,08%	35,39	60,06
TAC Miami	6653	4	13,56%	12,94	82,50
LAN Mexico	6589	2	64,33%	31,06	17,22
AM Mexico	6589	2	35,67%	17,22	31,06
LAN B. Aires	1139	4	52,73%	244,39	219,06
AR B. Aires	1139	4	23,64%	109,56	353,89
GOL B. Aires	1139	4	15,75%	73,00	390,44
CC B. Aires	1139	4	7,88%	36,50	426,94
LAN S. Paulo	2612	4	36,40%	70,58	123,33
TAM S. Paulo	2612	4	24,97%	48,42	145,50
GOL S. Paulo	2612	4	24,15%	46,83	147,08
SW S. Paulo	2612	4	14,48%	28,08	165,83

- Resumen de las variables binarias por aerolíneas y rutas:

Aerolínea	Directo	FLAG	FFP	Asociados	LAN	Madrid	Miami	Mexico	B. Aires	S. Paulo
LAN	1	0	1	1	1	1	0	0	0	0
AR	0	0	1	0	0	1	0	0	0	0
AC	0	0	0	0	0	1	0	0	0	0
IB	1	1	1	1	0	1	0	0	0	0
LAN	1	0	1	1	1	0	1	0	0	0
AR	0	0	1	0	0	0	1	0	0	0
AA	1	1	1	1	0	0	1	0	0	0
TAC	0	0	1	0	0	0	1	0	0	0
LAN	1	0	1	1	1	0	0	1	0	0
AM	1	1	1	1	0	0	0	1	0	0
LAN	1	0	1	1	1	0	0	0	1	0
AR	1	1	1	0	0	0	0	0	1	0
GOL	1	0	0	0	0	0	0	0	1	0
CC	1	0	0	0	0	0	0	0	1	0
LAN	1	0	1	1	1	0	0	0	0	1
TAM	1	1	1	0	0	0	0	0	0	1
GOL	1	0	0	0	0	0	0	0	0	1
SW	1	0	1	0	0	0	0	0	0	1

- Gráficos de frecuencias





7.3.- Anexo C

7.3.1.- Anexo C1: Test de Kolmogorov – Smirnov para los residuos del modelo 1 realizado en SPSS

One-Sample Kolmogorov-Smirnov Test		
		error
N		292
Normal Parameters(a,b)	Mean	-0,000171233
	Std. Deviation	0,079558305
Most Extreme Differences	Absolute	0,110034202
	Positive	0,110034202
	Negative	-0,077197155
Kolmogorov-Smirnov Z		1,083
Asymp. Sig. (2-tailed)		0,191
a	Test distribution is Normal.	
b	Calculated from data.	

7.3.2.- Anexo C2: Test de Kolmogorov – Smirnov para los residuos del modelo 2 realizado en SPSS

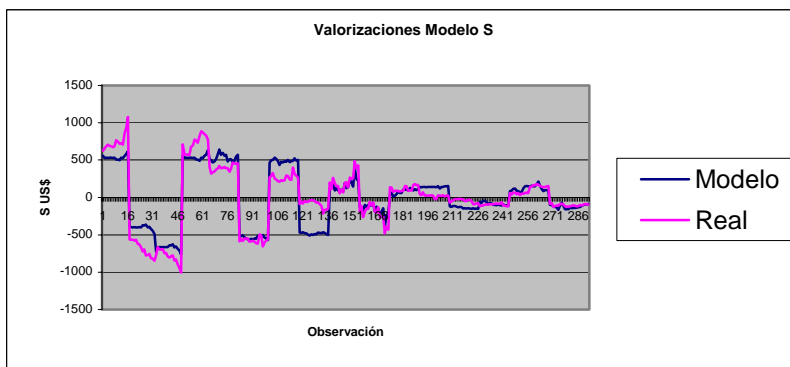
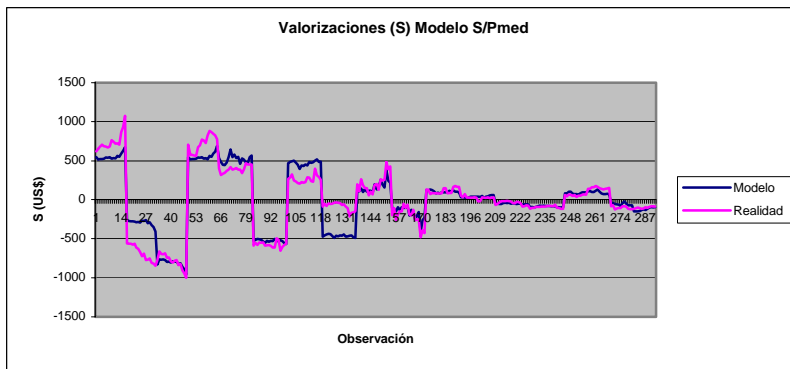
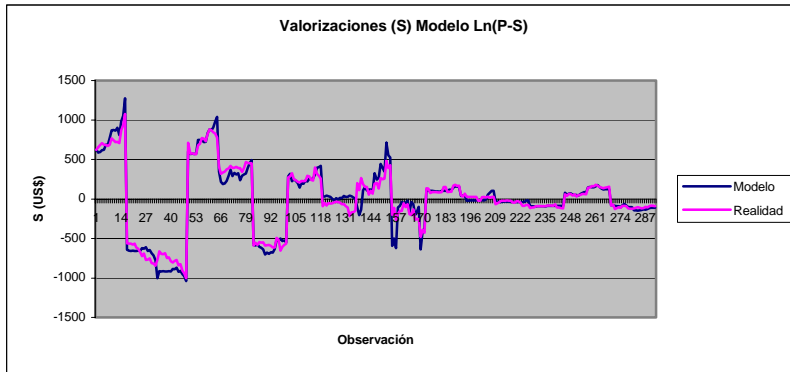
One-Sample Kolmogorov-Smirnov Test		
		error
N		292
Normal Parameters(a,b)	Mean	1E-04
	Std. Deviation	0,102
Most Extreme Differences	Absolute	0,063
	Positive	0,05
	Negative	-0,063
Kolmogorov-Smirnov Z		1,754
Asymp. Sig. (2-tailed)		0,041
a	Test distribution is Normal.	
b	Calculated from data.	

7.3.3.- Anexo C3: Test de Kolmogorov – Smirnov para los residuos del modelo 3 realizado en SPSS

One-Sample Kolmogorov-Smirnov Test		
		error
N		292
Normal Parameters(a,b)	Mean	8,502
	Std. Deviation	159,2
Most Extreme Differences	Absolute	0,129
	Positive	0,067
	Negative	-0,13

Kolmogorov-Smirnov Z	2,212
Asymp. Sig. (2-tailed)	1E-04
a	Test distribution is Normal.
b	Calculated from data.

7.3.6.- Anexo C4: Predicciones de las valorizaciones por los tres modelos:



7.3.5.- Anexo C5: Demostración de la media y varianza de una variable aleatoria exponencial sobre una normal

Sea X una variable aleatoria Normal de media μ_x y desviación estándar σ_x , entonces:

$$(C1) \quad E(e^X) = \int_{-\infty}^{\infty} e^x \cdot \frac{1}{\sqrt{2\pi\sigma_x^2}} e^{-\frac{(x-\mu_x)^2}{2\sigma_x^2}} dx$$

Como:

$$(C2) \quad e^x \cdot e^{-\frac{(x-\mu_x)^2}{2\sigma_x^2}} = e^{x-\frac{(x-\mu_x)^2}{2\sigma_x^2}} = e^{\frac{2\sigma_x^2 x - (x-\mu_x)^2}{2\sigma_x^2}} = e^{\frac{-(x-(\sigma_x^2 + \mu_x))^2 + 2\sigma_x^2 \mu_x + \sigma_x^4}{2\sigma_x^2}} = e^{\frac{(x-(\sigma_x^2 + \mu_x))^2}{2\sigma_x^2}} \cdot e^{\mu_x + \frac{\sigma_x^2}{2}}$$

Entonces:

$$(C3) \quad E(e^X) = \int_{-\infty}^{\infty} e^x \cdot \frac{1}{\sqrt{2\pi\sigma_x^2}} e^{-\frac{(x-\mu_x)^2}{2\sigma_x^2}} dx = e^{\mu_x + \frac{\sigma_x^2}{2}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma_x^2}} e^{-\frac{(x-(\sigma_x^2 + \mu_x))^2}{2\sigma_x^2}} dx = e^{\mu_x + \frac{\sigma_x^2}{2}}$$

Por otro lado, la varianza se puede expresar como:

$$(C4) \quad Var(e^X) = E[(e^X)^2] - E[e^X]^2 = E[e^{2X}] - (e^{\mu_x + \frac{\sigma_x^2}{2}})^2$$

Incorporando en la ecuación el hecho que $2X$ es una variable aleatoria normal de media $2\mu_x$ y desviación estándar $2\sigma_x$, entonces:

$$(C5) \quad E[e^{2X}] = e^{2\mu_x + \frac{4\sigma_x^2}{2}} = e^{2\mu_x + 2\sigma_x^2}$$

Por lo que:

$$(C6) \quad Var(e^X) = e^{2\mu_x + 2\sigma_x^2} - (e^{\mu_x + \frac{\sigma_x^2}{2}})^2 = e^{2\mu_x + \sigma_x^2} (e^{\sigma_x^2} - 1)$$

7.4.- Anexo D

7.4.1 Anexo D1: Resultado de la estimación de regresión completa para el modelo Ln(P-S):

Variables	Coficiente	Error Estándar	t	Sigma
(Constant)	643,58%	2,01%	320,19	0,00%
Asoc	-7,12%	1,97%	-3,62	0,04%
Buenos Aires	-106,29%	1,76%	-60,39	0,00%
Directo	3,90%	7,64%	0,51	61,08%
FFP	3,03%	1,59%	1,91	5,62%
FLAG	1,31%	2,05%	0,64	52,22%
frec%	-62,27%	6,40%	-9,73	0,00%
frecuencia	1,97%	2,17%	0,91	36,28%
frecuencia competencia	1,57%	1,60%	0,98	32,70%
LAN	1,60%	1,98%	0,81	41,81%
Madrid	123,19%	1,72%	71,62	0,00%
Mexico	161,13%	2,07%	77,84	0,00%
Miami	129,21%	1,81%	71,39	0,00%
Opera	3,84%	1,67%	2,30	4,24%
N	292			
R ²	0,994			

7.4.2 Anexo D2: Resultado de la estimación de regresión completa para el modelo P-S:

Variable	Coficiente	Error Estándar	t	Sigma
Constante	-66,93%	5,38%	-12,44	0,00%
Asoc	-10,29%	3,42%	-3,81	0,02%
Buenos Aires	-2,64%	8,53%	-0,31	75,66%
Directo	43,49%	3,88%	11,21	0,00%
FFP	16,75%	1,27%	13,19	0,00%
FLAG	2,97%	4,87%	0,61	54,18%
frec%	55,99%	9,98%	5,61	0,00%
frecuencia	0,17%	0,03%	5,7	0,00%
frecuencia competencia	-0,07%	0,02%	-3,31	0,01%
LAN	4,66%	6,21%	0,75	45,32%
Madrid	17,37%	4,03%	4,31	0,00%
México	-18,11%	4,55%	-3,98	0,00%
Miami	11,39%	3,90%	2,92	0,36%
Opera	-29,68%	2,87%	-10,34	0,00%
N	292			
R^2	0,857			

7.4.3 Anexo D3: Resultado de la estimación de regresión completa para el modelo S:

Variable	Coficiente	Error Estándar	t	Sigma
Constante	-1324,96	86,43	-15,33	0,00%
Asoc	-126,96	59,89	-2,12	3,40%
Buenos Aires	-18,96	70,22	-0,27	78,72%
Directo	844,13	64,44	13,1	0,00%
FFP	249,73	34,64	7,21	0,00%
FLAG	17,54	40,79	0,43	66,72%
frec%	1179,84	169,27	6,97	0,00%
frecuencia	-1,14	0,37	-3,11	0,19%
frecuencia competencia	2,34	-1,80	-1,3	19,36%
LAN	38,32	58,96	0,65	51,57%
Madrid	420,54	66,44	6,33	0,00%
México	-316,68	78,97	-4,01	0,01%
Miami	361,04	63,79	5,66	0,00%
Opera	-172,34	45,84	-3,76	0,02%
N	292			
R^2	0,833			