

Modelos basados en grafos: una aplicación al estudio del gasto de cruceristas en Uruguay

BRIDA, JUAN GABRIEL

Grupo de Investigación en Dinámica Económica, Facultad de Ciencias Económicas y de Administración

Universidad de la República (Uruguay)

Correo electrónico: gbrida@ccee.edu.uy

SANTIÑAQUE, FLORENCIA

Grupo de Investigación en Dinámica Económica, Facultad de Ciencias Económicas y de Administración

Universidad de la República (Uruguay)

Correo electrónico: fsantinaque@iesta.edu.uy

LANZILOTTA, BIBIANA

Grupo de Investigación en Dinámica Económica, Facultad de Ciencias Económicas y de Administración

Universidad de la República (Uruguay)

Correo electrónico: bibiana@iecon.ccee.edu.uy

RESUMEN

Este trabajo introduce el uso de grafos en el análisis de los determinantes del gasto de los turistas y lo aplica al estudio del comportamiento de los pasajeros de cruceros que desembarcan en Uruguay en los puertos de Montevideo y Punta del Este. Este enfoque ofrece un abordaje alternativo a los usualmente empleados para analizar los determinantes del gasto turístico al introducir una representación gráfica que permite sintetizar y visualizar las relaciones entre el conjunto de variables que caracterizan a los turistas o grupos de turistas y los determinantes de su nivel de gasto. En particular, la metodología permite representar las relaciones de dependencia entre variables (entre nodos adyacentes) e independencias condicionadas (entre nodos no adyacentes). Para su aplicación, se consideran datos individuales de las encuestas realizadas a cruceristas correspondientes a la temporada de cruceros comprendida entre noviembre de 2014 y abril de 2015, ambos incluidos, en Uruguay. Se estudia la existencia de dependencias de las variables relativas a registrar el gasto de los cruceristas con otras variables (sociodemográficas, de contexto, etc.). El estudio muestra que las variables que mejor explican el comportamiento de los visitantes son las vinculadas al puerto de desembarco y al gasto. Asimismo, los resultados muestran que las variables socioeconómicas no están vinculadas al gasto en forma directa.

Palabras claves: modelos de grafos; gasto de los turistas; cruceros; Uruguay.

Clasificación JEL: C19; C21; D12; L83.

MSC2010: 03H10; 62P20.

Artículo recibido el 31 de octubre de 2016 y aceptado el 7 de agosto de 2017.

Graph-Based Models: An Application to the Study of Cruise Passengers' Expenditure in Uruguay

ABSTRACT

This paper introduces the use of graphs in the analysis of the determinants of tourists' expenditure, providing an alternative approach to the methods usually employed. Graph-based analysis is applied to study the behavior of cruise passengers arriving at the ports of Montevideo and Punta del Este in Uruguay. The graphical representation allows to synthesize and visualize the relationships between the set of variables that characterize tourists or tourist groups and the determinants of their level of spending. The graph-based methodology is particularly suitable to represent the dependency relationships among variables (between adjacent nodes) and conditional independence (between non-adjacent nodes). For the empirical application, individual data from the surveys of cruise passengers corresponding to the cruise season between November 2014 and April 2015 in Uruguay are considered. The existence of dependencies between spending variables and other variables that may be of interest (sociodemographic, contextual, etc.) is studied. The exercise shows that the variables that better explain visitors' behavior are linked to the port of disembarkation and the expenditure. The results also show that socioeconomic variables are not directly linked to spending.

Keywords: graph-based models; tourists' expenditure; cruise; Uruguay.

JEL classification: C19; C21; D12; L83.

MSC2010: 03H10; 62P20.



1. Introducción

El impacto económico de los flujos turísticos en la economía es a menudo significativo y actúa como motor de crecimiento económico. Con el fin de mejorar los efectos e impactos económicos de la visita de turistas a un destino, se necesitan datos y herramientas apropiadas para analizar los mercados y así dirigir la oferta del sector privado y planificar las acciones de los formuladores de políticas. En consecuencia, el tener buenos modelos analíticos para entender el comportamiento de los turistas es crucial para entender los determinantes del gasto para el consumidor-turista. En este contexto, la estadística y la econometría ofrecen herramientas valiosas para el estudio de las relaciones entre el gasto turístico y las características de los consumidores.

El análisis y la interpretación de elementos tales como la elasticidad de la demanda o el gasto medio de los turistas asociado con características específicas, pueden proporcionar información valiosa para el desarrollo de la economía del turismo. Sin embargo, los trabajos académicos que analizan los determinantes del gasto turístico a nivel individual (microeconómicos) han sido relativamente menos que los que estudian la demanda agregada (macroeconómicos), considerando tanto el número de contribuciones y la heterogeneidad de los métodos estadísticos y modelos econométricos que se aplican (Brida & Scuderi, 2013; Wang & Davidson, 2010). Usualmente estos trabajos han utilizado regresiones a los efectos de cuantificar la significatividad e impacto de cada regresor en la variable de respuesta. Este trabajo contribuye a la rama de la literatura acerca de los determinantes del gasto turístico, introduciendo una manera diferente y un enfoque complementario a las regresiones clásicas, mediante la aplicación de teoría de grafos, para estudiar el gasto de los pasajeros de cruceros que desembarcan en Uruguay en los puertos de Montevideo y Punta del Este.

La literatura en economía del turismo ha puesto creciente interés en el análisis del turismo de cruceros. Ello responde en buena medida a la acelerada expansión de este tipo de turismo a nivel mundial en los últimos 15 años, crecimiento que ha conducido a que se constituyera en uno de los más dinámicos dentro del sector turístico. Para 2015, de acuerdo a la información de la *Florida-Caribbean Cruise Association* (FCCA, 2016), el número de cruceristas habría alcanzado los 23 millones. Uruguay no ha sido ajeno a esta dinámica. Aun cuando el turismo de cruceros representa una proporción reducida del turismo receptivo total –12% en 2014, según Ministerio de Turismo (MINTUR, 2015)–, la cantidad de personas arribadas ha aumentado significativamente en la última década.¹ Estas han pasado de algo más

¹ En rigor, dicho aumento no ha sido sostenido. En las últimas dos temporadas de cruceros se ha notado un descenso en la cantidad de buques arribados al puerto de Punta del Este, lo que ha ocasionado una cierta retracción en el gasto y los turistas desde el máximo experimentado en la temporada 2012-2013.

de 56 mil visitantes (en la temporada 2004-2005) a más de 400 mil en las temporadas 2011-2012 y 2012-2013. Y si bien en las últimas temporadas los visitantes se redujeron, continuaron superando holgadamente los 300 mil.

La industria de cruceros posee efectos de distinto signo sobre los países y economías de destino (Brida *et al.*, 2010). Los beneficios se vinculan principalmente al gasto que realizan los pasajeros y tripulantes en el destino de desembarco, así como a la demanda de servicios de navegación y portuarios, todo lo cual dinamiza el empleo local, incentiva inversiones asociadas a los servicios portuarios y turísticos y genera aumento de recaudación. Como sector transable e internacionalmente competitivo, potencialmente puede generar externalidades positivas. Entre otras que mencionan Dwyer & Forsyth, (1998), se encuentran mejoras en los estándares ambientales, seguridad, amenidades turísticas, transporte, así como la posibilidad de mejoras en términos de intercambio, aprovechamiento de economías de escala, etc.

El puerto de desembarco incide en el impacto económico que el crucero tenga sobre la región. En un puerto de escala (aquel donde los pasajeros desembarcan temporalmente), el gasto de los turistas se circunscribe a transporte, *tours*, compras, alimentación y otros gastos. En el caso de los puertos de embarque el impacto del crucerismo puede ser mayor ya que, por ejemplo, incluye gastos de alojamiento lo cual es, en general, uno de los epígrafes de gasto turístico más importante (Brida *et al.*, 2014a; 2014b). En Uruguay, los puertos que reciben cruceros son el puerto de Montevideo y el de Punta del Este. El puerto de Montevideo es principalmente de escala², mientras que el de Punta del Este es únicamente puerto de escala de cruceros. En su gran mayoría, los visitantes cruceristas son de nacionalidad brasileña, siguiéndoles los argentinos en importancia. También tienen una participación significativa aquellos de nacionalidad europea y norteamericana. El gasto de los cruceristas que descienden del buque se dirige principalmente a compras y servicios gastronómicos –44% y 33%, según el Anuario Estadístico de Turismo 2015, (MINTUR, 2016) –, no incluyendo en ningún caso el epígrafe “alojamiento”, dada la condición de puertos de escala tanto para Montevideo como para Punta del Este.

La dinámica del gasto turístico general y de cruceros en particular, así como de sus determinantes, ha sido analizada desde diversos puntos de vista y enfoques estadísticos y econométricos. Desde un punto de vista macroeconómico, se ha analizado principalmente asociándolo al impacto que esta actividad tiene sobre el país de destino (Dwyer & Forsyth, 1998; Seidl *et al.*, 2006; Brida & Zapata-Aguirre, 2010). Desde un punto de vista microeconómico, se ha indagado en los determinantes del gasto turístico y específicamente de

² Desde 2011, el Puerto de Montevideo es considerado puerto de embarque solo para el grupo de cruceros Costa. (MINTUR, 2015).

cruceros desde esquemas empíricos y econométricos diversos (Henthorne, 2000; Morrison *et al.*, 1996; Seidl *et al.*, 2007; Brida *et al.*, 2010; Brida *et al.*, 2012a, 2012b; Brida *et al.*, 2014a, 2014b; Brida & Scuderi, 2013, entre otros). Específicamente para Uruguay, Risso (2012) y Brida *et al.* (2010) analizaron el gasto de los cruceristas desde un punto de vista microeconómico mediante un modelo Heckit. Risso (2012) se centró en el estudio de las temporadas 2008 a 2010, aplicando un modelo de selección de Heckman (véase Cameron & Trivedi, 2005) con el objetivo de determinar las variables que explican los diferentes tipos de gastos e identificar los factores que inciden favorablemente (o desfavorablemente) en el gasto. En particular, este trabajo encuentra que el turista argentino reduce la probabilidad del gasto por persona mientras que el brasileño la aumenta. La probabilidad de gasto es alta (más de 80%) y el desagrado por los precios reduce esta probabilidad. Brida *et al.* (2010) aplicaron los modelos *tobit* y *probit* para estimar el efecto de distintas variables en el monto y la probabilidad de gasto en varias categorías (o rubros). De hecho, mostraron que el tamaño del grupo con el que viajan los cruceristas, así como la movilidad que estos tienen dentro del país, son variables importantes para explicar su patrón de gasto. Se desprende de la revisión de la literatura que el estudio sobre las diferentes metodologías utilizadas en la medición del gasto de cruceristas es relevante y de reciente desarrollo y que cada vez gana más interés en el plano académico, pero que aún el número de casos de estudio y de metodologías utilizadas es poco numeroso.

El presente trabajo se basa en la metodología propuesta en Abbruzzo *et al.* (2014) para el análisis del gasto de los turistas internacionales en Uruguay y tiene como objetivo estudiar el comportamiento de consumo de los cruceristas a partir de un enfoque metodológico alternativo al de sus antecedentes locales. Es de señalar que, de la revisión bibliográfica realizada, se desprende que el trabajo de Abbruzzo es el primero donde aparece un enfoque metodológico que está basado en la teoría de grafos. Mediante esta metodología se identifican y analizan las variables que inciden en el gasto de los cruceristas, visualizando las mismas desde una perspectiva de red (Edwards, 2000; Jackson, 2008; Lauritzen, 1996; Whittaker, 1990). En el presente estudio se amplía y detalla la metodología utilizada y, para su aplicación, se consideran datos individuales de la encuesta realizada a cruceristas por parte del Ministerio de Turismo correspondiente a la temporada de cruceros que se extiende entre noviembre 2014 y abril 2015. (MINTUR, 2015).

En las secciones segunda y tercera, se explica con mayor profundidad esta metodología y se describen los datos respectivamente. A continuación, se presenta una cuarta sección con los resultados de la investigación y, por último, se comentan las principales conclusiones.

2. Estrategia empírica: metodología

La metodología empírica empleada en este trabajo se basa en la teoría de grafos. Los modelos basados en grafos permiten identificar relaciones entre el conjunto de variables que caracterizan a los turistas o grupos de turistas y los determinantes de su nivel de gasto. Cada una de las variables aleatorias involucradas en el modelo se representa por un nodo en el grafo G ; mientras que cada arista representa la interacción entre las variables del modelo que están asociadas a sus nodos. En este sentido, es posible obtener una simplificación gráfica tanto de las relaciones de dependencia entre variables, así como también de las independencias condicionadas entre nodos que tiene algún tipo de conexión no directa. Es así que los modelos basados en grafos tienen la ventaja de sintetizar y detectar visualmente las relaciones que se producen dentro de grandes conjuntos de variables aleatorias, a través de una salida fácil de interpretar, dada por la representación gráfica del grafo. (Abbruzzo *et al.*, 2014).

Generalmente, para estudiar asociaciones o interacciones entre dos o más variables categóricas representadas mediante tablas de contingencia, suelen aplicarse modelos log-lineales. Los modelos log-lineales se usan para analizar la relación entre dos o más variables categóricas, sin realizar distinción de variables independientes y dependientes (todas las variables se consideran variables respuesta). La estrategia básica en el modelado consiste en ajustar un modelo a las frecuencias observadas en la tabla de contingencia. Los modelos permiten representar las probabilidades conjuntas mediante frecuencias esperadas.

Sean tres variables aleatorias categóricas (X,Y,Z) , cuyo número de dimensiones posibles de respuesta son $\{i = 1, \dots, I\}$, $\{j = 1, \dots, J\}$ y $\{k = 1, \dots, K\}$ respectivamente y cuyas frecuencias observadas se pueden representar en una tabla de contingencia tridimensionales de orden $I \times J \times K$, en la que se estudian n individuos. Entonces el modelo saturado generalizado (que correspondería al ajuste perfecto) puede representarse como:

$$\log(\mu_{ijk}) = \theta + \theta_i^X + \theta_j^Y + \theta_k^Z + \theta_{ij}^{XY} + \theta_{ik}^{XZ} + \theta_{jk}^{YZ} + \theta_{ijk}^{XYZ}$$

siendo μ_{ijk} la frecuencia esperada de la casilla ijk , que se modela como la suma: del logaritmo del número de unidades experimentales n (representado por el parámetro θ); los siguientes tres términos representan el logaritmo de las probabilidades marginales de cada variable, respectivamente; los siguientes tres términos representan el logaritmo de las probabilidades conjuntas bivariadas (representa el nivel de interacción entre variables dos a dos); y el último término representa el logaritmo de la probabilidad conjunta de las tres variables en cuestión. Adicionalmente, se realiza una prueba de hipótesis estadística para testear si cada término de interacción entre variables es significativamente no negativo. Se testea si hay evidencia estadística que indique que dicho término es no nulo; en ese caso, conlleva a la aparición de una arista en el grafo para el par de nodos asociados a las variables en cuestión.

En este modelo todas las variables son condicionalmente dependientes.

Este tipo de modelos admite una representación mediante grafos. Dicho de otra forma, un grafo representa gráficamente la interacción entre un conjunto de variables, mostrando como ciertos modelos probabilísticos pueden factorizarse en distribuciones probabilísticas marginales más simples. Los modelos probabilísticos que admiten una factorización completa se denominan modelos descomponibles. Un modelo de probabilidad será descomponible cuando pueda expresarse como el producto de varios factores (modelos probabilísticos condicionados sobre la distribución conjunta). En estos casos, la estimación de parámetros se simplifica significativamente.

En el grafo G que modela las variables categóricas y la relación entre éstas, el término de interacción $\theta_{ij}^{XY} \neq 0$ representaría la existencia de una arista entre el par de nodos asociados a las variables X e Y (y además, se dirá que dichos nodos son adyacentes). Por tanto, la presencia de conexiones entre pares de nodos en el grafo indica una dependencia entre las variables asociadas a dichos nodos, condicionada al resto. En cambio, un par de nodos no adyacentes en el grafo reflejan una relación de independencia condicional de las dos variables analizadas al resto de las variables. Esta propiedad se conoce como propiedad de Markov para pares de variables. Además, condicionado solo a los nodos adyacentes, cualquier variable asociada a dichos nodos es independiente de todas las restantes; propiedad que se conoce como la propiedad local de Markov (Hojsgaard *et al.* 2012, Cap. 1). Si un subconjunto de variables cualesquiera se ve separada por una tercera variable, entonces se dice que presentan una relación de independencia condicional sobre esta tercera variable. En resumen, los modelos basados en grafos resultan útiles a la hora de encontrar tanto dependencias condicionales entre variables, reflejadas por la presencia de nodos adyacentes, como independencias condicionales entre las mismas, representado por nodos no adyacentes.

Una de las principales dificultades frente al estudio de tablas de contingencia múltiples es encontrar el mejor modelo que ajuste a los datos. En este sentido, los grafos proporcionan una ventaja por su representación e interpretación gráfica.

Los tipos de grafos que representan las interacciones entre variables mencionadas anteriormente se denominan grafos de “independencia condicional”. En particular, se estudiará la formación de bosques y árboles. Los árboles son grafos no dirigidos, conexos y acíclicos³. La unión disjunta de árboles se denomina bosque. En particular, la ausencia de dirección en las aristas indica que no se podrá establecer relaciones de causalidad entre variables y solo se podrá

³ Es decir, siempre existe un camino (sucesión de aristas) que permite alcanzar un nodo desde cualquier otro y no es posible hallar tal sucesión de aristas de modo que el nodo de comienzo y de fin sea el mismo.

visualizar si dos variables poseen algún tipo de relación de dependencia condicional al resto de las variables del grafo.

En general, si $P_{X_1, X_2, \dots, X_n}(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n)$ es la distribución conjunta de n variables aleatorias discretas $\{X_i\}_{i=1}^n$ tomando respectivos valores $\{x_i\}_{i=1}^n$, entonces encontrar el árbol óptimo de independencia condicional consiste en aplicar un algoritmo de búsqueda “voraz” entre los n^{n-2} árboles que existen en un grafo con n nodos etiquetados (fórmula de Cayley). Para obtener de forma eficiente dicho árbol óptimo, Chow & Liu (1968) describieron un algoritmo de búsqueda de un árbol de expansión mínima basándose esencialmente en el algoritmo de Kruskal (1956). Dicho algoritmo consiste en un método iterativo por el que se va agregando una arista y un nodo adicional del grafo al árbol que se ha obtenido en la iteración previa conectando dicho nodo al árbol ya existente, optimizando algún tipo de ponderador (esencialmente el peso acumulado al sumar los pesos de todas las aristas en el árbol resultante) y asegurando que la inclusión de esa arista no supone la aparición de ciclos en el árbol resultante. Edwards *et al.*, (2010) describieron un algoritmo para la obtención de un árbol de expansión mínima a partir de considerar el “criterio de información bayesiano” (BIC) como los pesos/ponderadores de cada arista del grafo. Bajo dicho criterio, el algoritmo devuelve un bosque de expansión mínima (tal y como ocurre al considerar el algoritmo generalizado de Kruskal suponiendo pesos positivos). El ponderador BIC puede escribirse como sigue:

$$BIC(E) = -2.Ln(\hat{L}) + k.Ln(n)$$

siendo E la arista a seleccionar, \hat{L} el máximo de la función de verosimilitud del modelo agregando la arista correspondiente, k la cantidad de parámetros libres a estimar y n representando el tamaño de muestra. Dicho ponderador, por un lado, busca maximizar la función verosimilitud del modelo; pero, por otro lado, penaliza la inclusión de un mayor número de parámetros a estimar. Foygel & Drton (2010) demostraron que minimizar el BIC equivale a encontrar un árbol de expansión mínima.

A continuación, se desarrolla el algoritmo de búsqueda de un bosque óptimo basado en el algoritmo de Chow-Liu condicional (Kang *et al.*, 1997): se ordena de forma decreciente las aristas según los ponderadores mencionados anteriormente y se comienza a construir el árbol graficando los dos nodos adyacentes con el valor óptimo del ponderador seleccionado (i.e. mayor dependencia condicional). En cada paso se agregará aquella arista que produzca un mayor decremento de la función BIC y que no forme ciclo con las aristas previamente seleccionadas por el algoritmo.

- 1) Calcular el BIC para todas las aristas posibles (i.e. para las $n(n-1)/2$ aristas).

- 2) Seleccionar arista que optimice el valor BIC del modelo, suprimiéndola de la lista de posibles aristas a seleccionar en posteriores etapas.
- 3) Comprobar si la arista seleccionada en el Paso 2 no forma ciclo con las aristas incluidas en el bosque construido hasta esta etapa; en tal caso, añadir la nueva arista (y nodo) al bosque.
- 4) Si no existen más aristas para seleccionar, parar; en otro caso volver al Paso 2.

En una segunda etapa, se procede a implementar un método de selección del modelo hacia adelante (*forward selection*); partiendo del árbol de expansión mínima, en cada etapa se adicionará la arista que provoque la mayor reducción del BIC (si hubiera). En este caso, se está siendo menos restrictivo a la hora de agregar aristas, expandiendo la búsqueda al espacio del modelo descomponible. Estos modelos son los que se pueden descomponer sucesivamente en una secuencia de modelos más pequeños, donde cada uno de estos corresponde a un modelo marginal sin restricciones (véase Frydenberg & Steffen, 1989; Lauritzen, 1992). La mayor utilidad de este paso consiste en verificar si existen interacciones entre variables que resulten significativas y que dada la restricción respecto de la no formación de ciclos en el grafo del algoritmo anterior, no han sido consideradas en el bosque de expansión mínima. Si bien se consigue un grafo menos disperso, podría obtenerse interacciones entre variables que resulten de interés e incluso mejorar la posible interpretación del grafo generado.

Otra ventaja que aportan los modelos de grafos es la siguiente: si un bosque tiene n vértices y k aristas, entonces el bosque consiste en la unión disjunta de $n - k$ árboles (Chow & Liu, 1968). Por tanto, un grafo que cuente con estas características puede dar la pauta de la existencia de grupos o clústeres. Dicha información puede ser utilizada como insumo para determinar la cantidad de grupos *a priori* que requieren determinadas metodologías multivariadas (como pueden ser análisis de clústeres k -medias o PAM, entre otros).

Para la implementación de las funciones mencionadas anteriormente, se utilizaron la función *minForest* y *stepw* del paquete *gRapHD* del programa *R*.

3. Datos

Los datos sobre los cuales se aplica el método anteriormente descrito corresponden a la encuesta de cruceros correspondiente a la temporada 2014-2015 (MINTUR, 2015). El diseño de la muestra utilizada en el estudio fue realizado en base a dos muestras independientes, una para Montevideo y otra para Punta del Este (MINTUR, 2015), y comprende un total de 3275 encuestas. En Montevideo se realizaron 1927 entrevistas en 51 buques (de los 108 que arribaron en total), mientras que en Punta del Este se realizaron 1348 entrevistas en 35 buques (de un total de 64 que arribaron).

Tabla 1. Variables empleadas para la construcción de los grafos

Denominación	Descripción	Tipo	Valores
<i>Variables de gasto</i>			
TourEx	Gasto en <i>tours</i>	Cualitativa	Cero gasto, Bajo gasto, Medio gasto, Alto gasto
AlimEx	Gasto en alimentación	Cualitativa	Cero gasto, Bajo gasto, Medio gasto, Alto gasto
ShoppEx	Gasto en compras	Cualitativa	Cero gasto, Bajo gasto, Medio gasto, Alto gasto
TranEx	Gasto en transporte	Cualitativa	Cero gasto, Bajo gasto, Medio gasto, Alto gasto
RestoEx	Otros gastos	Cualitativa	Cero gasto, Bajo gasto, Medio gasto, Alto gasto
TotEx	Gasto total	Cualitativa	Cero gasto, Bajo gasto, Medio gasto, Alto gasto
<i>Variables socio-económicas</i>			
Res	Residencia	Cualitativa	Argentina, Brasil, USA y Canadá, Europa y otros, Resto América Latina
Occ	Ocupación	Cualitativa	Crucelistas, Tripulantes
<i>Variables de viaje</i>			
Nvis	Asiduidad del viaje	Cualitativa	"Primera vez" ; "Poco Leal"; "Leal"; "Muy Leal"
Puerto	Puerto	Cualitativa	Montevideo; Punta del Este
VisitMtdo	Visitó Montevideo	Dicotómica	Sí/No
Visitpta	Visitó Punta del Este	Dicotómica	Sí/No
VisitCol	Visitó Colonia	Dicotómica	Sí/No
VisitOtros	Visitó Otros	Dicotómica	Sí/No
<i>Tipología de grupo</i>			
Totpersonas	Total personas grupo	Cualitativa	Solo, En pareja, Más de 2 personas en el grupo
TipoGrupo ^a	Tipo de grupo	Cualitativa	Mixto, solo Mujeres, solo Hombres
<i>Variables de gusto y desagrado</i>			
GustaTodo /GustaCueros/GustaEdif/ GustaPaisaje/GustaHigiene/GustaPrecios GustaMtdo/GustaMercPuerto/GustaVinos GustaTiendas/GustaGente/GustaCasapueblo GustaCiudadVieja/GustaOrganizacion/GustaDedos GustaTranq/GustaPlaya/GustaComida GustaCasino/GustaPunta/GustaNada/GustaOtros DesagradSD/DesagradPrecios/DesagradNoTarjeta DesagradInseg/DesagradTaxis/DesagradTransito DesagradAtencion/DesagradPobreza DesagradNoTiendas/DesagradDeterioro DesagradNoBanios/DesagradGrafitti DesagradNoTaxis/DesagradPlaySuc/DesagradHigiene DesagradfaltaServPlaya/DesagradSenias DesagradTipoCambio/DesagradNada DesagradInfraest /DesagradFaltaEspect	Dicotómicas	Sí/No	

a) Composición por sexo del grupo.

La encuesta de cruceristas revela una serie de características para cada uno de los turistas que desciende del crucero, así como para su grupo de pertenencia. También revela un conjunto de variables vinculadas a las actividades que los cruceristas desarrollan en su desembarco, a variables de gusto asociados a la experiencia turística, gastronómica, de compras y naturalmente acerca del gasto efectuado. En la Tabla 1 se presentan las variables empleadas en

la aplicación empírica, su denominación, descripción y valores. Nótese que las variables de gusto y desagrado son dicotómicas y solo admiten como respuesta “Sí” o “No”. A los encuestados les consultan sobre gustos y disgustos de las actividades realizadas o sitios visitados.

4. Resultados empíricos

Como marco introductorio e interpretativo de los resultados, previamente se presenta una breve descriptiva de las principales características socio-demográficas y de comportamiento de los cruceristas arribados durante la temporada 2014-2015.

Como se indicó anteriormente, más de 330 mil personas desembarcaron de los 172 buques que arribaron a Uruguay en la temporada que se estudia. De ellas, más de 200 mil lo hicieron en el puerto de Montevideo. En ambos puertos, aproximadamente el 70% de las personas embarcadas descendieron, de las cuales el 35% aproximadamente eran brasileñas. Las nacionalidades argentinas, europeas y norteamericanas le siguieron en importancia. La Tabla 2 resume algunas de estas estadísticas.

Tabla 2. Indicadores de gasto, socio-demográficos y de gusto de cruceristas

Temporada de cruceros 2014-2015				
	Buques	Personas desemb.	Gasto (USD¹)	Gasto per cápita (USD¹)
Montevideo	108	201085	7222690	35,9
Punta del Este	64	131033	3720780	28,4
Total	172	332118	10943470	33,0
Puerto de desembarco				
Nacionalidades	Montevideo		Punta del Este	
	Personas	Gasto medio (USD¹)	Personas	Gasto medio (USD¹)
Norteamericanos	17638	37,1	12729	45,3
Otros América	2901	49,7	2740	34,8
Chilenos	1849	201,3	1921	27,6
Europeos	20683	34	6397	37,2
Otros	3188	28,8	1813	42,5
Brasileños	117610	39,3	70014	28,2
Argentinos	37216	18,3	35419	18,8
Agrado/desagrado	Montevideo		Punta del Este	
	Agrado	Desagrado	Agrado	Desagrado
Todo	23,0%		25,0%	
Nada		63,4%		75,1%
Tranquilidad, segur.	9,4%	1,5%	7,0%	
Higiene		1,7%	10,0%	
Gente, cord., atención	25,1%		13,0%	1,0%
Playa, rambla, costa	14,4%		19,0%	
Otros	28,4%		26,0%	6,3%
Precios/Tipo de Cambio		12,6%		16,7%
Señalización tránsito				0,9%

¹ USD refiere a dólares corrientes.

Fuente: Anuario Turístico (MINTUR, 2015)

Un elemento que caracterizó a esta temporada es que el gasto descendió significativamente respecto de las temporadas previas (medido en dólares corrientes, cayó aproximadamente el 40% y, en moneda nacional a precios constantes, el 39%). Probablemente el deterioro de la situación macroeconómica regional y, en particular, de Brasil, haya sido determinante en la explicación de esa caída. En ese contexto, en términos de gasto per cápita, las nacionalidades del resto de América del Sur mostraron un nivel de gasto muy por encima del promedio, destacándose especialmente, la de los chilenos. Como sucedió en otras temporadas, el principal rubro de gasto fue Compras y en segundo lugar Alimentación.

Las variables que recogen la percepción de los visitantes, las variables de gusto y disgusto, presentan algunas tendencias comunes y otras disímiles en función del lugar de desembarco. Si bien hay pequeñas diferencias en los porcentajes, es común el agrado total o el nulo desagrado, así como el agrado por playas, rambla o costa y el desagrado por precios o tipo de cambio. En las variables de gusto por la atención, la gente o cordialidad, las diferencias a favor de Montevideo son notorias. Lo contrario sucede con higiene, en donde el agrado es nulo en Montevideo y positivo en Punta del Este.

La elaboración de los grafos utilizados en este estudio se efectuó, en una primera instancia, sin restricción respecto de una distancia entre nodos respecto de las variables de interés⁴. Como resultado se obtuvo el bosque que se presenta en la Figura 1. Como puede apreciarse, este bosque posee tantos nodos como variables se están analizando y no todas se conectan directamente entre sí. Es el caso de las variables Gusta Precios, Visitó Colonia u Otros, Gusta Vino y Gusta Otros. De acuerdo con estos resultados, las variables centrales (i.e. esos nodos que concentran mayor cantidad de nodos adyacentes a su alrededor) son Gasto total (TotEx), Puerto, la variable de agrado total (Gusta Todo) y la de desagrado nulo (DesagradNada), siendo esta última la que mayor concentración presenta en relación a las otras dos variables y observando que no hay relación directa entre dichas tres variables centrales. Esta centralidad muestra dos puntos importantes: por un lado, la importancia de estas variables por su conexión con las demás y, por otro lado, indican las relaciones más fuertes de dependencias condicionadas con otras variables (nodos adyacentes).

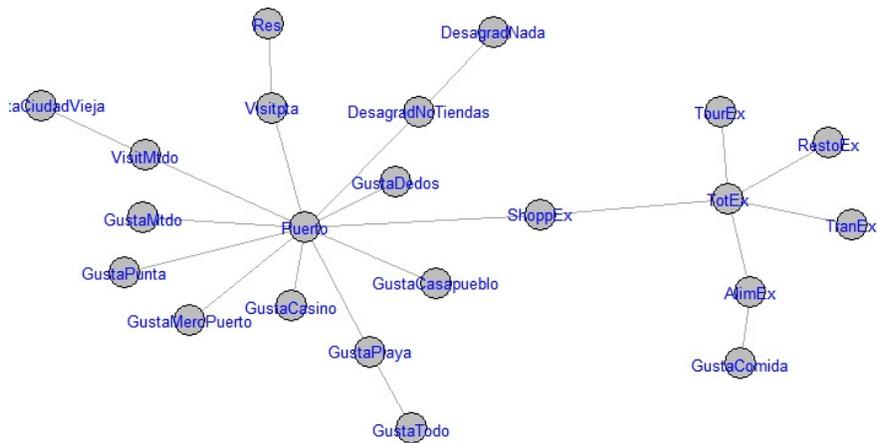
En la Tabla 3, se presentan las primeras cinco aristas que se graficaron según el algoritmo de búsqueda de bosque de expansión mínima detallado anteriormente.

⁴ Se define la distancia entre dos nodos como el número de aristas existentes entre dichos dos nodos (téngase en cuenta que solo existe una forma dentro de un árbol de alcanzar un nodo desde otro). La bola de centro el nodo v y de radio $r > 0$ es el conjunto de nodos que están a distancia menor o igual que r .

Desagrada Nada, que, como ya se mencionó, son algunos de los nodos que toman importancia dado su conexión con los demás.

Asimismo, se generó un subgrafo a partir del primero, extrayendo los nodos cuya distancia respecto de Gasto Total (TotEx) no superase las 4 aristas (bola de centro Gasto total y radio 4). Dicha bola se delimita según aquellos nodos que se encuentran condicionalmente asociados en no más de 4 aristas, lográndose una representación más parsimoniosa respecto de las variables de interés (véase Figura 2).

Figura 2. Árbol de expansión mínima según bola de centro Gasto total y radio 4 aristas



Fuente: elaboración propia

En la Tabla 4 se presentan cinco medidas alternativas de centralidad o cercanía para caracterizar el grafo de la Figura 2. Estas medidas, que surgen de la vertiente teórica de redes sociales y económicas, permiten determinar la ubicación o posición de los nodos en las redes (indicador de grado de centralidad) y la cercanía considerada a través de 4 indicadores (Jackson, 2008). El grado de centralidad indica la cantidad de conexiones de cada uno de los nodos. En la Tabla 4 se presentan los nodos que presentan mayor número de conexiones, es decir: Gasto total, Puerto y Gasto en compras. Por su parte, el grado de centralidad es un indicador de cuán fácilmente un nodo puede alcanzar a otro. Los 4 indicadores empleados para representar el grado de cercanía son: Cercanía ($C(g)$) y Cercanía con factor descendente ($C_i^d(g)$), con tres variantes de este factor. El primero se define como:

$$C(g) = \frac{(n - 1)}{\sum_{j \neq i} l(i, j)}$$

donde n representa el número de nodos y $l(i, j)$ representa el número de aristas correspondiente al camino más corto entre el nodo i y el nodo j .

La cercanía con factor descendente es una forma más rica de medir la cercanía ya que dicho factor opera como ponderador (inverso) de la distancia entre los nodos. Es decir, los nodos más distantes ponderan menos que los más cercanos. Este indicador se define como:

$$C_i^d(g) = \frac{\sum_{j \neq i} \delta^{l(i,j)}}{[(n-1)\delta]}$$

donde n representa el número de nodos, $l(i,j)$ representa el número de aristas correspondiente al camino más corto entre el nodo i y el nodo j y δ representa el factor descendente o declinante y varía entre $1 > \delta > 0$. Es posible notar que cuando este factor se acerca a 0, esta medida de centralidad otorga infinitamente mayor peso a los nodos más cercanos.

Tabla 4. Medidas de centralidad y cercanía (bola de centro Gasto total)

Variables centrales	Grado de centralidad	Cercanía	Cercanía con factor descendente		
			$\delta=0,5$	$\delta=0,75$	$\delta=0,25$
Gasto total	5	0,382	0,488	0,751	0,326
Puerto	11	0,568	0,771	0,996	0,623
Gasto en compras	2	0,467	0,482	0,776	0,257

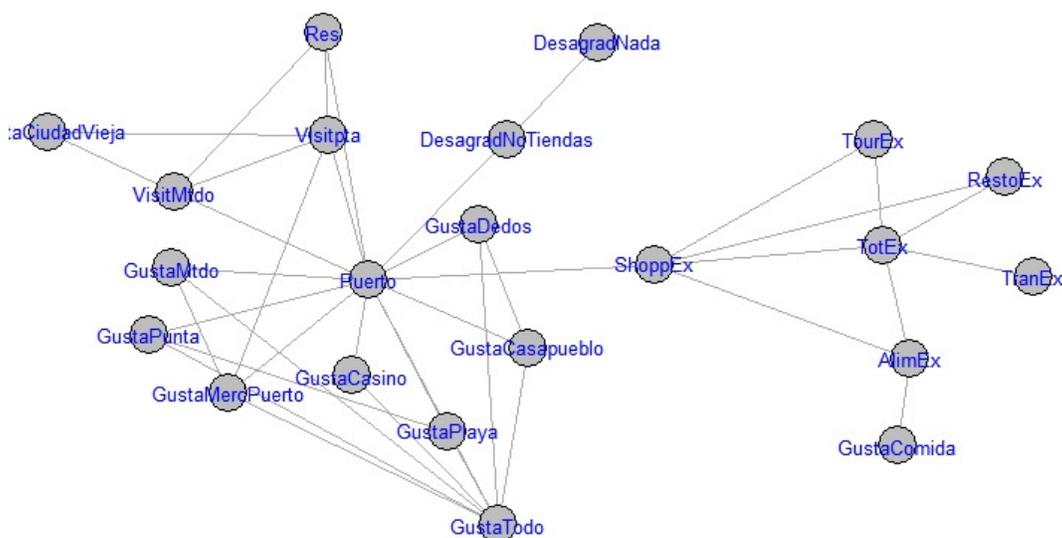
Fuente: elaboración propia

Nota: Calculado sobre la base de árbol de expansión mínima según bola de centro Gasto total (Figura 2); δ : factor descendente según lejanía ($l(i,j)$) de los nodos

Acorde a este análisis, la variable Puerto de desembarco es, bajo todos los indicadores, la de mayor centralidad y cercanía, por lo que su importancia sería significativa en el análisis del gasto. De hecho, como se visualiza en las Figuras 1 y 2, las variables de gasto se vinculan a las variables de percepción, comportamiento y características del grupo a través de la variable Puerto, a la que se unen mediante el gasto en compras.

A partir de árbol de expansión mínima generado anteriormente, se realizó *forward selection*. En la Figura 3, se destaca que la mayor cantidad de conexiones que resultaron significativas (lograron una disminución del BIC total del árbol) se encuentran en los vecindarios de los nodos asociados a las variable de Puerto y Gasto total. Específicamente se puede observar que la incorporación de conexiones entre el nodo asociado a la variable Gasto en compras y los nodos asociadas a las variables Gasto en *tour*, Gasto en alimentación y Otros gastos (*RestoEx*) provocó una disminución del BIC de dicho árbol. Sin embargo, desde el nodo asociado a la variable Puerto, solo se agregó una arista que lo conecta con el nodo asociado a Residencia. Demás interacciones que resultaron significativas se han dado entre las variables de visitas y gustos entre las que se destacan: Visita Montevideo y visita Punta del Este, Gusta Montevideo y Gusta Mercado del Puerto, Gusta Punta del Este y Gusta Playas, Gusta Ciudad Vieja y Visita Punta del Este.

Figura 3. Grafo generado mediante *forward selection*



Fuente: elaboración propia

En torno a la variable Puerto se sitúan variables de comportamiento del crucerista y del grupo, como, por ejemplo, Visita a Montevideo o Visita a Punta del Este. También se hallan conectados a este nodo las variables de percepción: Gusto (Gusta Playa, Gusta Montevideo, Gusta Casa Pueblo, Gusta Punta del Este, Gusta Casino, etc.) y de Disgusto por insuficiente oferta de tiendas. Conectado con el nodo asociado a la variable Visita a Punta del Este (véase Figura 1), a través del nodo asociado a la residencia, se halla la mayor cantidad de nodos asociados a variables que caracterizan al grupo, como ser indicador del número de visitas, ocupación, tamaño, y tipo, asociadas entre sí en ese orden. No obstante, todas ellas se encuentran relativamente alejadas del nodo central al que estarían asociadas (Puerto) y de hecho desaparecen al calcularse el árbol con criterios más restrictivos en cuanto a distancia respecto del nodo asociado a Gasto total.

En resumen, se puede decir que las variables asociadas a la variable de interés, Gasto total, son otras variables de gasto (en compras, *tours*, alimentación, otros y transporte), lo cual resulta intuitivo, y la variable de gusto por comida. Este vecindario se conecta con las restantes variables (de comportamiento, características del grupo y percepción) a través de la variable Puerto (mediante su vínculo con gasto en compras).

A la hora de aplicar un modelo de regresión que explique los componentes del gasto, se deberían tomar como variables explicativas aquellas que se encuentran adyacentes al nodo que representa la variable de interés. Por ejemplo, si se quisiera explicar el comportamiento del Gasto Total, según la Figura 3, las variables potencialmente significativas para el modelo son las que representan a los demás componentes del gasto (Gasto en compras, *tours*, transporte y alimentación), así como los términos de interacciones entre Gasto en compras: Gasto en *tours*,

serían Gasto en alimentación, *tours* y otros gastos; y un segundo grupo relacionado a la visita por Montevideo que se relaciona con variables de gusto, por ejemplo Gusta Ciudad Vieja, Gusta Cueros, Gusta Gente pero también incluye variables de desagrado de playas.

5. Conclusiones

En este trabajo se introdujo el uso de modelos basados en grafos para el estudio de los determinantes del gasto turístico. Como aplicación empírica de esta metodología, se consideran datos individuales de las encuestas realizadas a cruceristas correspondiente a la temporada de cruceros entre noviembre 2014 y abril 2015 en Uruguay. Como muestra la revisión bibliográfica, los trabajos estadísticos y econométricos anteriores al presente se acercaron a este problema mediante la aplicación de modelos econométricos clásicos. El enfoque de los modelos basados en grafos comprende un amplio conjunto de metodologías que ofrecen varias ventajas en términos de su uso y de la interpretación de los resultados. Son adecuados para resumir las interrelaciones dentro de grandes conjuntos de variables aleatorias, como es el caso de la encuesta para los cruceristas en Uruguay. Proporcionan resultados fácilmente interpretables, incluso en presencia de relaciones no lineales entre las variables. La metodología es particularmente útil para la construcción de mapas conceptuales basadas en datos y con el fin de visualizar los complejos vínculos que ocurren dentro de un conjunto de variables.

En comparación con los modelos econométricos clásicos, que pueden ser muy útiles para otros objetivos, los modelos basados en grafos ofrecen información útil y complementaria, que puede ser utilizada para la selección del modelo econométrico. Los grafos pueden ser utilizados para analizar la existencia de relaciones entre las variables consideradas. Por otra parte, a partir del grafo de independencia condicional, uno puede darse cuenta si dos o más variables están vinculadas directamente, o más bien si su conexión depende de alguna manera de otros elementos. Esto puede ser de ayuda tanto para los investigadores y académicos como para los operadores del mercado turístico. El comprobar que una relación entre dos variables se produce no de manera directa (i. e. los nodos asociados a las variables son adyacentes) sino que tiene lugar por mediación de otras variables explicativas (i. e. existe una sucesión de aristas que permiten conectar los dos nodos asociados, pero es necesario recorrer nodos adicionales y, por tanto, considerar las variables asociadas a dichos nodos), puede tener profundas implicaciones sobre las acciones a tomar hacia los conductores de una variable de respuesta. La estructura de enlaces hace que sea posible identificar aquellos elementos que son cruciales en el gobierno de las interrelaciones dentro de un conjunto de nodos. Esto se verifica por el hecho de que en algunas redes los nodos asumen una posición central con respecto a los demás. Esto ocurre para las variables Gasto total, Puerto, las variables de agrado total y la de desagrado nulo.

Otra ventaja que aportan los modelos basados en grafos consiste en obtener una visualización gráfica de un gran conjunto de variables que permita observar posibles agrupaciones o patrones de comportamiento similares entre variables; es decir, podría ofrecer una metodología complementaria al análisis de clústeres así como a otras metodologías multivariadas como pudiera ser el análisis de componentes principales (ACP).

Los resultados empíricos vinculados a las interacciones entre las variables de gasto y las de comportamiento pueden brindar interesantes sugerencias para los decisores en políticas de turismo del país y destinos. Tómese en cuenta que las variables asociadas a la variable de interés (Gasto total) son algunas de las variables de gastos específicos (compras, *trips*, alimentación, por ejemplo), así como a la variable de gusto por Comida. A su vez, recuérdese que este vecindario se conecta con las restantes variables (de comportamiento, características del grupo y percepción) a través de la variable Puerto, mediante su vínculo con Gasto en compras.

Ello permite hipotetizar que, dado que los cruceristas tienen casi todo su consumo incluido en el barco, los restaurantes y bares locales podrían obtener mejores beneficios si tratan de diferenciar su oferta del consumo en alimentos y bebidas que los cruceristas ya tienen incluido en su paquete del crucero; por ejemplo, proponiendo productos locales. Asimismo, es posible pensar que los cruceristas que deciden ir de compras en el destino, debido al “efecto riqueza”, puedan dar valor a participar en actividades culturales; dado que la cultura y las compras podrían estar compitiendo como actividades a ser realizadas por los turistas, sobre todo en este caso de los cruceristas donde el tiempo es una restricción a la cantidad de actividades que pueden hacerse. Por lo tanto, la promoción de actividades culturales debe estructurarse con el fin de captar el interés de este tipo de personas que tienen poco tiempo para pasar en el destino.

Esta investigación puede proseguir mediante la comparación de esta metodología con otras que se utilizan a la hora de agrupar variables como son el análisis de componentes principales o el análisis de clústeres de variables. Asimismo, la metodología deja el terreno preparado para poder hacer regresiones a los efectos de cuantificar la significatividad e impacto de las variables identificadas por el modelo en el gasto del turista. Además, se podría utilizar como insumo a la hora de modelar cierta variable de interés mediante un modelo de regresión, ya que el grafo proporciona información útil a la hora de seleccionar el mejor modelo de ajuste. Finalmente, y en el caso específico del estudio del gasto de los cruceristas en Uruguay, visto que se cuenta con datos uniformes para diversas temporadas, se podría analizar cada temporada con la metodología y estudiar la evolución de los determinantes del gasto.

Bibliografía

- Abbruzzo, A.; Brida, J.G. & Scuderi, R. (2014): "Determinants of individual tourist expenditure as a network: Empirical findings from Uruguay". *Tourism Management*, 43, 36–45.
- Brida, J.G.; Bukstein, D.; Garrido, N.; Tealde, E. & Zapata-Aguirre, S. (2010): "Cruise passengers expenditure in the Caribbean port of call Cartagena de Indias: A cross-section data analysis". *Estudios y Perspectivas de Turismo*, 19, 607–634.
- Brida, J.G.; Bukstein, D. & Tealde, E. (2013): "Exploring cruise ship passenger spending patterns in two Uruguayan ports of call". *Current Issues in Tourism*, 18(7), 684–700.
- Brida, J.G.; Fasone, V.; Scuderi, R. & Zapata-Aguirre, S. (2014a): "ClustOfVar and the segmentation of cruise passengers from mixed data: Some managerial implications". *Knowledge-Based Systems* 70: 128–136.
- Brida, J.G.; Garrido, N. & Such Devesa, M.J. (2012a): "Cruise passengers satisfaction: Cartagena de Indias". *Benchmarking: An International Journal*, 19(1), 55–69.
- Brida, J.G.; Pulina, M.; Riaño, E. & Zapata-Aguirre, S. (2012b): "Cruise passenger's experience embarking in a Caribbean Home Port." *Ocean and Coastal Management*, 55, 135–145.
- Brida, J.G. & Scuderi, R. (2013): "Determinants of tourist expenditure: A review of microeconomic models". *Tourism Management Perspectives*, 6, 28–40.
- Brida, J.G.; Scuderi, R. & Seijas, M.N. (2014b): "Segmenting cruise passengers visiting Uruguay: A factor-cluster analysis". *International Journal of Tourism Research*, 16, 209–222.
- Brida, J.G. & Zapata-Aguirre, S. (2010): "Cruise tourism: economic, socio-cultural and environmental impacts". *International Journal Leisure and Tourism Marketing*, 1 (3), 205–226.
- Cameron, A.C. & Trivedi, P.K. (2005): *Microeconometrics: Methods and applications*. Cambridge: Cambridge University Press.
- Chow, C. & Liu, C. (1968): "Approximating discrete probability distributions with dependence trees". *IEEE Transactions on Information Theory*, 14(3), 462–467.
- Dwyer, L. & Forsyth, P. (1998): "Economic significance of cruise tourism". *Annals of Tourism Research*, 25 (2), 393–415.
- Edwards, D. (2000): *Introduction to graphical modeling*. New York: Springer.

- Edwards, D.; de Abreu, G.C.G. & Labouriau, R. (2010): “Selecting high-dimensional mixed graphical models using minimal AIC or BIC forests”. *BMC Bioinformatics*, 11(18), 13 pp.
- FCCA (2016): *Cruise Industry Overview 2016*. Pembroke Pines: Florida-Caribbean Cruise Association. Disponible en <http://www.f-cca.com/downloads/2016-FCCA-Cruise-Industry-Overview-Cruises-Statistics.pdf> <acceso 31/07/2017>.
- Foygel, R. & Drton, M. (2010): “Extended Bayesian Information Criteria for Gaussian Graphical Models”. *Advances in Neural Information Processing Systems*, 23, 2020–2028.
- Frydenberg, M. & Steffen, L.L. (1989): “Decomposition of maximum likelihood in mixed graphical interaction models”. *Biometrika*, 76(3), 539–555.
- Henthorne, T.L. (2000): “An Analysis of Expenditures by Cruise Ship Passengers in Jamaica”. *Journal of Travel Research*, 38(3), 246–250.
- Hojsgaard, S.; Edwards, D. & Lauritzen, S. (2012): *Graphical models with R*. New York: Springer Verlag.
- Jackson, M. (2008): *Social and Economic Networks*. New Jersey: Princeton University Press.
- Kang, H.J.; Kim, K. & Kim, J.H. (1997): “Optimal approximation of discrete probability distribution with k th-order dependency and its application to combining multiple classifiers”. *Pattern Recognition Letters*, 18(6), 515–523.
- Kruskal, J. (1956): “On the shortest spanning subtree of a graph and the traveling salesman problem”. *Proceedings of the American Mathematical Society*, 7(1), 48–50.
- Lauritzen, S.L. (1992): “Propagation of probabilities, means, and variances in mixed graphical association models”. *Journal of the American Statistical Association*, 87(420), 1098–1108.
- Lauritzen, S.L. (1996): *Graphical models*. Oxford Statistical Science Series Vol. 17. New York: Oxford University Press.
- MINTUR (2015): *Temporada de Cruceros 2104-2015*. Montevideo: Ministerio de Turismo de Uruguay. Disponible en: <http://mintur.gub.uy/index.php/component/jdownloads/summary/11-turismo-de-cruceros/37-2014-2015> <acceso 1/12/2017>.
- MINTUR (2016): *Anuario de Turismo 2015*. Montevideo: Área de Investigación y Estadísticas. Ministerio de Turismo de Uruguay.

- Morrison, A.M.; Chunghui, C.; O’Leary, J.T. & Nadkarni, N. (1996): “Comparative profiles of travelers on cruises and land-based resort vacations”. *Journal of Tourism Studies*, 7(2), 15–27.
- Risso, W.A. (2012): “El gasto de los cruceristas en Uruguay 2008–2010”. *Revista de Turismo y Patrimonio*, 10 (3), 393–406.
- Seidl, A.; Guiliano, F. & Pratt, L. (2006): “Cruise tourism and community economic development in Central America and the Caribbean: The case of Costa Rica”. *Revista de Turismo y Patrimonio Cultural*, 4 (2), 213–224.
- Seidl, A.; Guiliano, F. & Pratt, L. (2007): “Cruising for colones: cruise tourism economics in Costa Rica”. *Tourism Economics*, 13(1), 67–85.
- Wang, Y. & Davidson, M. (2010): “A review of micro-analyses of tourist expenditure”. *Current Issues in Tourism*, 13(6), 507–524.
- Whittaker, J. (1990): *Graphical models in applied multivariate statistics*. New York: Wiley.