Miscelánea Matemática **59** (2014) 91-102 SMM

Modelos gráficos en datos discretos

Ricardo Ramírez Aldana
Instituto de Geografía,
Universidad Nacional Autónoma de México,
ricardoramirezaldana@ciencias.unam.mx

Ruth Fuentes García
Facultad de Ciencias, Universidad Nacional Autónoma de México
Circuito exterior s/n, Ciudad Universitaria, México DF, 04510, México.
rfuentes@ciencias.unam.mx

Resumen

En este trabajo se introducen modelos estadísticos que involucran datos discretos y que se relacionan con distintos aspectos de la teoría de gráficas.

1. Introducción

Los datos discretos son aquellos en los que se tienen mediciones que toman un número finito de valores. Por ejemplo, si se tiene una encuesta y se pregunta el género del encuestado, entonces esta es una variable discreta con dos posibles valores. A los valores que toma una variable discreta se les conoce como niveles o categorías de la variable.

Buscaremos describir los modelos gráficos, que se pueden ver como modelos lineales, por ello hablaremos en primer lugar de estos últimos. En Estadística con frecuencia son de interés los modelos lineales [1], que son aquellos en los cuales una variable, llamada dependiente, es una función lineal de otra u otras llamadas explicativas. En particular, cuando la variable dependiente involucra un conteo y las independientes corresponden a datos discretos se tienen los llamados modelos loglineales. Los modelos gráficos son un subconjunto de estos últimos, cuya ventaja es que pueden representarse mediante gráficas, tal como son definidas en matemáticas discretas, permitiendo el desarrollo de propiedades estadísticas de interés de manera eficiente a través de resultados y definiciones dadas en el contexto de teoría de gráficas. La

importancia de la representación gráfica de los modelos loglineales, tanto para abordar aspectos de estimación como para su interpretación, se ha discutido ampliamente en la literatura [3]. Más aún, en los últimos diez años se han presentado resultados relevantes en el área de la estadística algebraica, misma que utilizando herramientas de la geometría algebraica permite abordar la estimación en los casos en los que algunas categorías tienen conteos iguales a cero [1]. Además, interesan con frecuencia propiedades estadísticas como la asociación entre variables y la dependencia o independencia que existe entre ellas y la representación gráfica permite una mejor visualización de estas asociaciones.

2. Contexto general

En esta sección presentamos un conjunto de ejemplos con el fin de dar una motivación a los llamados modelos loglineales, para después abordar con detalle el caso particular de los llamados modelos gráficos loglineales.

Ejemplo 1

Supóngase que tenemos el número de personas que se transladan de una localidad i a una j en una región o país. A esta variable la llamaremos variable dependiente o variable respuesta. Podemos tratar de explicar la estructura de estos conteos a través de una serie de variables que llamamos variables independientes o explicativas. Por ejemplo, la cantidad de población existente en cada localidad, $P(i)$ y $P(j)$ para la localidad i y j, respectivamente, y la distancia euclideana que hay entre las poblaciones i y j, variable d_{ij}. No es descabellado pensar que la atracción esperada entre localidades [2] $T(i,j)$ sigue una ley similar a la de gravitación universal, para la cual la atracción entre ellas es directamente proporcional a su masa dada en términos de la cantidad de población y a una función de decaemiento de la distancia, $F(d_{ij})$, la cual indica que a mayor distancia la atracción disminuye. De tal forma que se tiene:

$$T(i,j) \propto P(i)P(j)F(d_{ij}).$$

[1] Se refiere a que existan combinaciones de valores de algunas variables que ningún individuo en la muestra tome, e.g. al analizar sexo y estado civil que no haya hombres viudos en una muestra.

[2] La atracción es una forma de medir la proximidad o cercanía entre lugares. Por ejemplo, si se quisiera medir la relación entre dos localidades en México, una de ellas una ciudad grande y otra a poca distancia, entonces el número de viajes o gente que se translada diariamente entre ambos lugares podría usarse para medir esa atracción.

De hecho, también podríamos usar funciones de la cantidad de población. Un modelo adecuado sería:

$$T(i,j) = k\frac{P(i)^{\mu}P(j)^{\alpha}}{\exp(\delta d_{ij})}, \tag{1}$$

en donde k es una constante, μ, α y β son parámetros, con $\delta > 0$. Tomando logaritmos en ambos lados de la ecuación (1),

$$\ln(T(i,j)) = u + \mu\ln(P(i)) + \alpha\ln(P(j)) + \beta d_{ij}, \tag{2}$$

con $u = \ln(k)$ y $\beta = -\delta$. Sería de interés calcular el valor de los parámetros para así poder establecer una ley de atracción que rija los translados entre localidades. Para ello necesitaríamos los conteos del número de translados entre localidades.

El modelo (2) puede verse como un caso particular de una llamada regresión loglineal. Esta trata de explicar un fenómeno medido a través de conteos por medio de una serie de variables. Desde el punto de vista estadístico se les conoce como regresiones loglineales debido a la presencia del logaritmo aplicado a la variable respuesta [3]. Este tipo de modelos son un caso particular de los llamados modelos lineales generalizados, en los cuales el valor esperado de una variable continua o discreta es explicado a través de una función lineal de variables explicativas.

Sea $m(i,j)$ el número esperado de observaciones cuando una variable discreta S toma el valor i y otra variable discreta G toma el valor j, entonces un modelo loglineal puede expresarse como:

$$\ln(m(i,j)) = u + S(i) + G(j). \tag{3}$$

El modelo anterior solo contempla un efecto constante u y los efectos que cada variable tiene sobre los conteos, $S(i)$ y $G(j)$, a estos se les llama efectos principales. Puede existir un efecto conjunto entre ambas variables a lo que se le conoce como interacción. Las interacciones son una forma de ver la asociación que existe entre variables discretas.

Ejemplo 2

Considere que $m(i,j)$ es el número esperado de muertes por cáncer en pacientes y denótese como S al sexo del paciente y como G al grupo de edad al que pertenece. El modelo de la ecuación (3) indica que el número esperado de defunciones depende de los efectos de sexo y grupo de edad. Pudiera considerarse incluso un efecto conjunto $SG(ij)$ o interacción entre las combinaciones de sexo y edad que ayude a determinar el número esperado de defunciones.

[3] El logaritmo o en general a la función aplicada a la respuesta en un modelo lineal generalizado se le conoce como función liga.

En los modelos loglineales se hacen suposiciones sobre la distribución probabilística de los conteos. Esto es, existen leyes probabilísticas dadas a través de funciones conocidas (funciones de densidad o de distribución) que modelan el comportamiento de valores aleatorios y en este caso se utilizan algunas en particular. Usualmente, se supone que los conteos siguen una distribución Poisson con parámetro $m(i, j)$, cuando es así, al modelo se le conoce también como regresión Poisson y puede incluir variables explicativas tanto discretas como continuas. Sin embargo; para los conteos en un modelo loglineal pueden suponerse esta u otras distribuciones, como la multinomial, y además solo se permiten variables explicativas discretas. Afortunadamente, sin importar cual distribución sea usada los resultados que se obtienen son generalmente análogos.

Ejemplo 2 (continuación)

En el Ejemplo 2, la actividad laboral L también podría ayudar a explicar las defunciones esperadas y si además se considera la interacción entre el sexo y esta variable $SL(ik)$, con i y k valores específicos de las variables, se tendría el modelo

$$\ln(m(i, j, k)) = u + S(i) + G(j) + L(k) + SL(ik), \tag{4}$$

donde las defunciones esperadas para un género i, un grupo de edad j y una actividad laboral k específicos $m(i, j, k)$ dependen de tres variables a través de sus efectos principales y de una interacción. Como puede verse las interacciones y efectos que pueden entrar en un modelo loglineal aumentan en función del número de variables. Además puede haber interacciones de distinto orden. Las interacciones vistas son de orden uno porque incluyen dos variables, pero pueden ser de orden mayor, por ejemplo una interacción de orden dos sería $SLG(ijk)$, la interacción entre el sexo, actividad laboral y grupo de edad, y así sucesivamente.

Un modelo loglineal es jerárquico si al incluir una interacción entonces también se incluyen todas las interacciones de orden inferior, así como los efectos principales correspondientes y un término constante[4]. De lo anterior puede inferirse que hay un conjunto de interacciones que determinan al modelo porque a partir de ellas se obtienen todos los parámetros. A este conjunto A se le conoce como conjunto generador. En la ecuación (4) el conjunto generador sería A= $\{\{S, L\}, \{G\}\}$.

[4]Por ejemplo, el modelo dado en (4) contiene a la interacción $SL(ik)$, así que para ser jerárquico debe contener los efectos principales $S(i)$ y $L(k)$, si alguno de ellos no estuviera, entonces el modelo es aún loglineal pero no jerárquico.

3. Modelos gráficos

Dado un modelo jerárquico se le puede asociar una gráfica, llamada gráfica de interacción. En ésta a cada variable se le asocia un punto y dos puntos se unen entre sí mediante una línea si existe una interacción de primer orden que contenga a esas dos variables. La figura 1 corresponde al gráfico de interacción asociado al modelo dado en el ejemplo 2 (continuación).

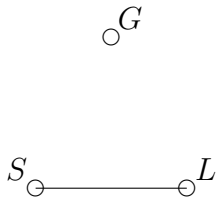

Figura 1. Gráfica de interacción para el modelo del ejemplo 2 (continuación).

Un modelo gráfico loglineal es un caso particular de modelo jerárquico donde pueden determinarse cuáles son las independencias marginales y condicionales que existen entre las variables. En el ejemplo 2 (continuación) podría ocurrir que el sexo y la actividad laboral estén asociados solo en función del grupo de edad, esto quiere decir que S es condicionalmente independiente de L dado G, o bien $S \perp L|G$. A través del modelo gráfico loglineal pueden verse estas independencias condicionales entre dos variables dadas otras e incluso independencias condicionales más generales o independencias marginales. Además hay una relación uno a uno entre ellas y el modelo gráfico. Para definir estos modelos se tienen algunos requisitos adicionales en un modelo loglineal jerárquico y para ello nos auxiliaremos de algunos resultados de teoría de gráficas.

Una gráfica G es una dupla que depende de un conjunto de vértices V y de un conjunto de aristas E, *i.e.* $G = (V, E)$. Los vértices son puntos $V = \{v_1, \ldots, v_k\}$, $k \in \mathbb{N}$ y las aristas son un subconjunto de parejas no ordenadas de $V \times V$ de la forma $\{v_i, v_j\}$ para alguna i, j=$1, 2, \ldots, k$. Estas se representan como líneas que unen los vértices correspondientes. Una gráfica es completa si todos los distintos pares de vértices que la conforman están unidos por una arista. La gráfica completa trivial consiste de un vértice aislado [5]. Una subgráfica G'= (V', E') de G= (V, E) es una gráfica en la cual $V' \subseteq V$ y $E' \subseteq E$. Una subgráfica inducida por un conjunto de vértices $S \subseteq V$, denotada como $G[S]$, es una subgráfica con vértices S y cuyo conjunto de aristas está formado por aquellas aristas cuyos extremos se encuentran en

[5]Si se tienen dos vértices la gráfica completa es una arista, si son tres es un triángulo y así sucesivamente.

elementos en S [6]. Un *clique* para G es una subgráfica inducida por un conjunto de vértices S, $G[S]$, que es completa y maximal respecto a su contención, *i.e.* no solo es una subgráfica que es completa sino que no debe estar contenida en otra gráfica inducida de G que sea completa. Por lo anterior, se pueden obtener todos los *cliques* de una gráfica. Por ejemplo, los *cliques* de la gráfica G dada en la figura 2 son $\{1,2,3\}, \{2,4\}$ y $\{4,5,6\}$ ya que cada una es una subgráfica de G completa no contenida en otra subgráfica de G completa.

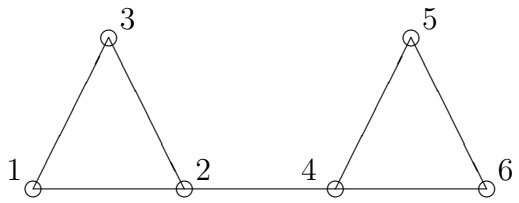

Figura 2. Gráfica con la que se ilustra el concepto de *clique*.

Un modelo loglineal es gráfico si su conjunto generador coincide con el conjunto de cliques de su gráfica asociada, ver e.g. [4]. En el modelo dado en la ecuación (4) del Ejemplo 2 (continuación) se tiene una gráfica asociada formada por un vértice aislado y una arista (figura 1), esto implica que el conjunto de cliques es $\{\{S,L\}, \{G\}\}$. Este conjunto, según lo visto, coincide con el conjunto generador del modelo. Por lo anterior, el modelo asociado a esta ecuación es un modelo gráfico loglineal.

Una propiedad importante en un modelo gráfico loglineal es que pueden derivarse las independencias condicionales y marginales que existen entre las variables que lo conforman. La más simple es aquella que indica que si en la gráfica de interacción asociada no hay una arista entre dos variables, entonces estas dos variables son independientes condicionalmente dadas el resto. Una propiedad más general indica que si se tiene un subconjunto de vértices A y otro subconjunto de vértices B, ambos unidos a través de un conjunto de vértices S, *i.e.* la única forma de acceder de A a B (por medio de aristas) es a través de los vértices en S, entonces se tendría independencia condicional entre las variables contenidas en A y las variables en B dadas aquellas en S, *i.e.* $A \perp B|S$. Desde el punto de vista de teoría de gráficas al conjunto S se le denomina conjunto de vértices separador de la gráfica G y separa al conjunto de vértices en A de los de B[7]. En particular si no hay aristas

[6]Por ejemplo, la subgráfica inducida por los vértices S= {1, 2, 3}, $G[S]$, en la figura 2 corresponde al triángulo a la izquierda de la gráfica.

[7]Por ejemplo, en la figura 2, el conjunto de vértices S= {2, 4} separa a los vértices en A= {1, 3} de los de B= {5, 6}. La definción formal de conjunto separador se basa en trayectorias, si todas

entre dos conjuntos (en teoría de gráficas estos son los llamados componentes conexos), entonces estos dos conjuntos son independientes entre sí, esta sería una independencia marginal.

De acuerdo a la distribución de los conteos se puede obtener la función de verosimilitud [8] asociada al modelo, la cual en escencia es la misma (el *kernel* es el mismo) sin importar si la distribución asociada es Poisson o multinomial [9]. De acuerdo al principio de máxima verosimilitud, pueden estimarse los parámetros asociados y obtener los conteos esperados bajo el modelo al maximizar esa función. En general no existen fórmulas cerradas para resolver las ecuaciones de verosimilitud, así que se usan métodos iterativos [10].

A partir del modelo gráfico ajustado y su conjunto generador podemos obtener la gráfica correspondiente y determinar las independencias condicionales y marginales que existen en los datos. Además, a través de la gráfica, podemos identificar las asociaciones que existen entre las variables. Para poder hacer inferencia y así determinar si las independencias obtenidas son genuinas, es necesario determinar que tan bien se ajusta el modelo a los datos. Para ello se utiliza un estadístico, es decir una función de los datos y de los valores estimados, conocido como devianza residual. Con esto se prueba la hipótesis nula de que el modelo se ajusta bien a los datos contra la alternativa de que no. La devianza residual corresponde a una función del cociente de las funciones de verosimilitudes del modelo que se quiere ajustar y un modelo con ajuste perfecto, o sea un modelo bajo el cual los valores estimados corresponden con los valores reales, llamado modelo saturado. Cuando el valor obtenido es grande quiere decir que el modelo saturado y el de interés difieren mucho entre sí y conforme disminuye implica que no es así, por lo anterior cuando el valor es grande se rechaza la hipótesis nula. En realidad todo esto aplica a modelos loglineales en general y se adapta a los modelos gráficos loglineales, obteniendo expresiones particulares, por ejemplo se puede obtener la expresión de la devianza resultante al quitar una arista (devianza que resulta de comparar el modelo con y sin la arista) y así probar si conviene o no quitarla del modelo.

Existe software específico que permite hacer el ajuste de los modelos gráficos loglineales, incluyendo métodos de selección de modelos, *i.e.* métodos que a partir de una base de datos proporcionan el mejor modelo gráfico que se ajusta a los datos de acuerdo a procesos iterativos,

---

las trayectorias entre los vértices de A y los de B pasan por los elementos en S, entonces este es un conjunto separador.

[8] La verosimilitud es la función de distribución asociada al modelo evaluada en valores específicos, en este caso aquella asociada a los conteos.

[9] Esto significa que el máximo sobre los parámetros de interés obtenido es el mismo, ya que la verosimilitud y su *kernel* son proporcionales.

[10] Métodos numéricos que bajo un nivel de precisión aceptable tratan de resolver de forma aproximada un conjunto de ecuaciones, e.g. *iterative proportional fitting*.

algunos de ellos basados en la devianza. Un ejemplo de este tipo de software es *MIM*, el cual además es de uso libre.

A continuación se proporciona un ejemplo de un ajuste de un modelo gráfico loglineal en datos reales.

Ejemplo 3

Los datos utilizados en este ejemplo corresponden a factores de riesgo para desarrollar enfermedades coronarias en trabajadores de fábricas de coches en Checoslovaquia, ver [2] y [5]. Es de interés determinar la asociación existente entre: a) la presencia de un historial de enfermedades coronarias en la familia del trabajador (Hc), b) si la presencia de una razón entre las beta y alfa lipoproteínas es inferior a tres (rba), c) si la presión sistólica del trabajador es inferior a 140 mm (Ps), d) si el trabajo desarrollado es extenuante físicamente (Tf), e) si este es extenunate mentalmente (Tm) y e) si la persona fuma (F). Todas las variables tienen dos posibles valores, Sí y No.

Basándose en métodos de selección disponibles en el software *MIM* se pudo determinar un modelo gráfico loglineal asociado que se ajusta de manera satisfactoria a los datos. Este modelo ajusta bien a los datos porque de acuerdo al estadístico correspondiente a la devianza residual no se rechaza la hipótesis nula correspondiente (el valor del estadístico fue de 51.36 con 46 grados de libertad y el nivel más bajo en que se rechaza la hipótesis nula H_0 o p-valor es de 0.27, por lo cual a un nivel de confianza de 0.05 no se rechaza H_0). El modelo seleccionado tiene el conjunto generador A= $\{\{rba, Ps, F\}, \{rba, Tf, F\}, \{Tm, Hc\}, \{Tf, Tm, F\}\}$ cuya gráfica asociada está dada en la figura 3.

A partir de la gráfica asociada al modelo pueden obtenerse independencias condicionales de acuerdo a conjuntos de vértices separadores. Por ejemplo, se puede determinar que la variable que identifica presión sistólica superior a 140mm y aquella para la razón entre las beta y alfa lipoproteínas son independientes condicionalmente al historial coronario dadas las variables correspondientes a trabajo físico y mental extenuantes y la variable que identifica si el trabajador fuma, $Ps, rba \perp Hc | Tf, Tm, F$. Esto es porque $\{Tf, Tm, F\}$ es un conjunto de vértices que separa a los conjuntos de vértices $\{Ps, rba\}$ de $\{Hc\}$. Esto significa que el historial coronario se asocia con las variables correspondientes a mediciones médicas tomadas a los trabajadores Ps y rba a través de las variables que tienen que ver con el tipo de labor y hábitos (si el sujeto fuma). Las asociaciones en la gráfica son lógicas, por un lado se tienen asociaciones entre fumar y variables relacionadas con medidas que determinan posibles problemas cardiovasculares y por otro lado se tienen asociaciones entre

fumar y el tipo de trabajo que tiene que ver con aspectos sociales. Se tiene que la presión sistólica, la variable relacionada con fumar, la razón entre las lipoproteínas y el trabajo físico son condicionalmente independientes del historial coronario dada la variable asociada a trabajo mental $Ps, F, rba, Tf \perp Hc|Tm$. Se pueden obtener otras independencias condicionales, como por ejemplo $Ps, rba, F \perp Hc|Tm, Tf$; $Ps, rba \perp Hc|Tm, Tf, F$ y $Ps, rba \perp Hc, Tm|Tf, F$.

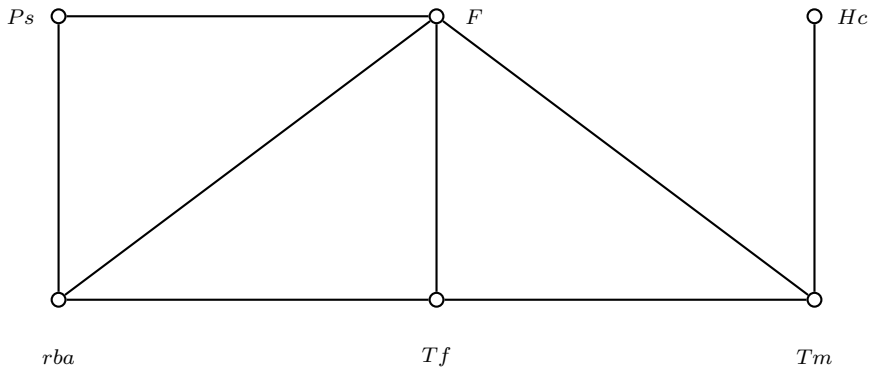

Figura 3. Modelo gráfico loglineal asociado a los datos correspondientes a factores de riesgo para desarrollar enfermedad coronaria introducidos en el ejemplo 3.

Cuando se tiene un conjunto de variables discretas, estas pueden representarse mediante tablas similares a las de frecuencias, en las cuales se representan de forma anidada todas las posibles combinaciones de categorías o niveles que pueden tomar las variables discretas y en cada combinación se incluyen el número de casos que toman esa combinación de valores. A estas tablas se les conoce como tablas de contingencia.

Utilizando los datos presentados por [1, p. 423] consideramos una variable discreta correspondiente la región de residencia de ciudadanos estadounidenses en 1980 con cuatro posibles valores: Noreste, Medio Oeste, Sur y Oeste y una segunda variable correspondiente al lugar de residencia en 1985 para las mismas regiones. Podemos obtener una tabla de contingencia correspondiente al cambio de residencia entre 1980 y 1985 como se muestra en el cuadro 1. En cada entrada se tiene el número de observaciones para cada posible combinación de regiones en ambos años.

En una tabla de contingencia de dos variables, cada una con el mismo número de categorías o niveles, podría ocurrir que los conteos arriba de la diagonal son iguales a los de abajo de la diagonal. A esto se le conoce como simetría y puede verse como un caso particular de modelo

Residencia en 1980	**Residencia en 1985**			
	Noreste	Medio oeste	Sur	Oeste
Noreste	11607	100	366	124
Medio oeste	87	13677	515	302
Sur	172	225	17819	270
Oeste	63	176	286	10192

Cuadro 1. Región de residencia, 1980 y 1985, en una muestra de 55,981 residentes estadounidenses.

loglineal. En el cuadro 1, la simetría indicaría que el número de individuos que emigran de una región i a una región j es el mismo número de los que emigran de la región j a la región i, así que sería como una población cerrada y sin diferencias de migración [11]. El caso de simetría con dos variables es el más simple, cuando se tienen más variables se han desarrollado distintas definiciones que generalizan el concepto de simetría, ya que no existe una única forma de definirla. Algunas por ejemplo, se refieren a un tipo de simetría en la cual la distribución es la misma a pesar de permutar los niveles de las variables de todas las formas posibles.

En el trabajo desarrollado por [6] se han establecido nuevos modelos que proporcionan una liga entre los modelos gráficos loglineales y que combinan distintos conceptos de simetría cuando se tienen dos o más variables. Abarcan la simetría correspondiente a valores en una tabla de contingencia que son iguales en distintas celdas de la tabla, similar al caso bidimensional. También se considera la simetría en términos de una distribución que se preserva al intercambiar [12] vértices en la gráfica asociada al modelo. Así mismo, se agrega el concepto de simetría en términos de la gráfica a través del concepto de automorfismo, esto es, transformaciones que preservan la gráfica al permutar ciertos vértices [13].

Con los modelos que incluyen estos conceptos pueden modelarse por ejemplo datos asociados a gemelos. Supóngase que se tienen los datos correspondientes a presencia o no de alcoholismo y depresión en parejas de gemelos, a través de estos modelos pudo determinarse que el alcoholismo y la depresión están relacionados en cada elemento de la pareja y entre las parejas. Al ajustar un modelo pudo inferirse que la información que aporta uno de los gemelos es suficiente para entender

[11] Para estos datos en realidad el modelo que ajusta bien es el de cuasisimetría. Este corresponde a un modelo que indica que no hay simetría en su totalidad porque marginalmente no se tiene el mismo número de observaciones entre regiones en los años estudiados.

[12] El intercambio está dado a través de permutaciones.

[13] Por ejemplo, si en la figura 1 se intercambian S y L, la gráfica es la misma ya que los vértices que eran vecinos siguen siéndolo después del intercambio. Sin embargo; si se intercambian G y L la gráfica no es la misma puesto que ahora G estaría unido a S.

la distribución asociada ya que el número esperado de observaciones es el mismo aunque se intercambie la información de los gemelos. Estos modelos incorporan aún más conceptos algebraicos y de teoría de gráficas.

4. Discusión

Hemos introducido a los modelos gráficos para datos discretos, enfatizando que su definición y propiedades ligan conceptos probabilísticos y estadísticos con conceptos de teoría de gráficas. La representación gráfica de estos modelos puede ayudar a visualizar datos discretos e identificar propiedades, sobre todo en lo que se refiere a independencias, asociaciones e incluso, bajo ciertas modificaciones en los modelos, algunas relaciones de simetría. Sin embargo, a medida que el número de variables incrementa, la representación gráfica puede volverse más difícil de visualizar. La ventaja de los modelos permanece porque los conceptos gráficos que los determinan no necesariamente requieren de la gráfica para poder aplicarlos, conservando su significado. El trabajo a desarrollar en modelos gráficos discretos es muy amplio, sobre todo porque la dificultad que representa tomar en cuenta los niveles de las variables trae como consecuencia que se dé preferencia a los modelos gráficos continuos. En estos últimos los conceptos pueden definirse de forma más directa. Así mismo, el estudio del concepto de simetría, tanto en el área de datos discretos como continuos es también un tema de interés que sigue siendo estudiado.

5. Bibliografía básica

1. A. Agresti, *Categorical data analysis*, 2a.ª ed., John Wiley and Sons, Nueva York, 2002.
2. D. Edwards, *Introduction to graphical modelling*, 2a.ª ed., Springer-Verlag, Nueva York, 2000.
3. S. E. Fienberg y A. Rinaldo, «Three centuries of categorical data analysis: log-linear models and maximum likelihood estimation», *Journal of Statistical Planning and Inference*, vol. 137, 2007, 3430–3445.
4. S. L. Lauritzen, *Graphical models*, Clarendon Press, Oxford, Inglaterra, 1996.
5. R. Ramírez-Aldana, «Restricted or coloured graphical log-linear models», tesis de doctorado, Tesis Doctoral, Posgrado en Ciencias Matemáticas, UNAM, 2010.
6. R. Ramírez-Aldana y G. Eslava-Gómez, «Restricted or coloured graphical log-linear models.», , 2015, , Por aparecer *Austrian Journal of Statistics*.

6. Bibliografía complementaria

7. E. Andersen, *The statistical analysis of categorical data*, 2a.ª ed., Springer-Verlag, Heidelberg, 1991.
8. Y. M. M. Bishop, S. E. Fienberg y P. W. Holland, *Discrete multivariate analysis*, MIT Press, Cambridge, Massachusetts, 1975, Reimpresión Springer, Nueva York, 2007.
9. J. A. Bondy y U. S. R. Murty, *Graph theory with applications*, MacMillan Press, Londres, 1976.
10. R. Christensen, *Log-linear models and logistic regression*, 2a.ª ed., Springer-Verlag, Nueva York, 1997.
11. J. N. Darroch, S. L. Lauritzen y T. P. Speed, «Markov fields and log-linear interaction models for contingency tables», *The Annals of Statistics*, vol. 8 (3), 1980, 522–539.
12. R. Diestel, *Graph theory*, 3a.ª ed., Springer-Verlag, Nueva York, 2005.
13. D. Edwards y T. Havránek, «A fast model selection procedure for large families of models», *Journal of the American Statistical Association*, vol. 83 (397), 1987, 205–213.
14. S. E. Fienberg, *The analysis of cross-classified categorical data*, 2a.ª ed., MIT Press, Cambridge, Massachusetts, 1980.
15. F. A. Graybill, *An introduction to statistical linear models vol. 1*, Mc Graw Hill Book Co., Nueva York, 1961.
16. S. J. Haberman, «Algorithm AS 51: log-linear fit for contingency tables», *Applied Statistics*, vol. 21 (2), 1972, 218–225.
17. S. L. Lauritzen, *Lectures on contingency tables*, 3a.ª ed., Department of Mathematics, Aalborg, Dinamarca, 1989, Versión electrónica 2002 disponible en `http://www.stats.ox.ac.uk/~steffen/papers/cont.pdf`.
18. S. L. Lauritzen y N. Wermuth, «Graphical models for associations between variables, some of which are qualitative and some quantitative», *The Annals of Statistics*, vol. 17 (1), 1989, 31–57.
19. P. McCullagh y J. A. Nelder, *Generalized linear models*, 2a.ª ed., Chapmann and Hall, Londres, 1989.
20. R. L. Plackett, *The analysis of categorical data*, 2a.ª ed., Griffin, Londres, 1981.
21. R. Ramírez-Aldana, «Una aplicación de modelos gráficos probabilísticos en investigación médica», tesis de maestría, Tesis de Maestría, Posgrado en Ciencias Matemáticas, UNAM, 2005.
22. R. Ramírez-Aldana y G. Eslava-Gómez, «*Restricted or coloured graphical log-linear models: Definition and applications*», Reporte de investigación 5-10, Departmento de Matemáticas, Facultad de Ciencias, UNAM, 2010.
23. J. Whittaker, *Graphical models in applied multivariate statistics*, John Wiley and Sons, Chichester, Inglaterra, 1990, Edición en pasta blanda 2009 por Wiley.