UNIVERSIDAD DE SEVILLA

DEPARTAMENTO DE ECONOMÍA APLICADA I

Especificaciones Condicionales para Vectores Bidimensionales, Discretos y Finitos.

Trabajo presentado por **Román A. Pérez Villalta** para la obtención del grado de doctor, dirigido por el doctor **D. José Antonio Camúñez Ruiz**

Sevilla, Octubre de 2015.

Agradecimientos.

Este trabajo no se habría podido realizar sin el apoyo de diversas personas que, aún sin saberlo, ayudaron. Como no podía ser de otra manera, cuentan con mi agradecimiento.

En primer lugar al director de este trabajo, doctor José Antonio Camúñez Ruiz, cuyos consejos y confianza han resultado cruciales.

A todos mis compañeros del departamento de Economía Aplicada I, que de una forma u otra me apoyaron durantes estos años que hemos compartido. En especial al Doctor José Luis Pérez Diez de los Rios por sus indicaciones finales, a Lola Gómez y Rosa María Paz, por su valiosa ayuda en la necesaria navegación por el proceloso e ignoto mar de la administración.

A mis padres y hermanos, por tanto sacrificio, tanta tristeza y tantas alegrías.

Y por último, y no por ello menos importante, a mi mujer, Magdala, y a mi hija, Magdala, por soportar, durante la realización de este trabajo, mis ausencias y, lo que a veces fue peor, mi presencia.

ÍNDICE.

Introducción.

Este trabajo está dedicado al estudio de la relación existente entre las distribuciones condicionadas y la distribución conjunta de un cierto vector aleatorio bidimensional. Más concretamente, nos planteamos si el conocimiento de las distribuciones condicionadas, determinan completamente la distribución conjunta de un vector aleatorio bidimensional. Nos restringiremos al caso en que el vector aleatorio es discreto y finito.

Los primeros estudios de este problema, para el caso general, pueden encontrarse en Besag (1972, 1974) y posteriormente, a mediados de la década de los ochenta, con el surgimiento de los métodos de Monte Carlo basados en Cadenas de Markov, especialmente el denominado Muestreo de Gibbs, donde la existencia de distribución conjunta cuando las condicionadas están dadas es una condición imprescindible para la convergencia del algoritmo, una magnífica introducción a este tema es Casella y George (1992).

Por otra parte, una extensa colección de trabajos como, Brucker (1979), Gouriroux y Monfort (1979), Abrahams y Thomas (1984), Arnold y Pourahmadi (1988), Arnold y Press (1989), Gupta y Vargas (1990), Arnold, Castillo y Sarabia (1992, 1999) tratan el problema en su forma general, lo que le confiere entidad propia. Pérez-Villalta (2000) trata el problema restringido a variables aleatorias bidimensionales discretas y finitas, proponiendo el método de extensión a rango unidad que formaliza y extienden Arnold, Castillo y Sarabia (2004) y Song, Li, Chen, Jiang y Kuo (2010). Además, proliferan trabajos dedicados al estudio del problema que nos ocupa para variables discretas, sin ánimo de ser exhaustivos podemos citar, Hürlimann (1991), Ip y Wang (2009), Tian et al. (2009), Wang y Kuo (2010), Kuo y Wang (2011), Wang (2012), Ghosh y Balakishan (2013) y Ghosh y Nadarajah (2014).

El trabajo está dividido en cinco capítulos y dos anexos. En el capítulo primero estudiamos el problema de la compatibilidad de distribuciones condicionadas, que para distribuciones discretas y finitas, caso que nos ocupa, se traducen en dos matrices no negativas verificando alguna condición adicional. Nos preguntamos entonces ¿bajo qué condiciones existe un vector aleatorio que las tenga como distribuciones condicionadas? Para ello se analizan diversos métodos con especial énfasis en el método de extensión a rango unidad del que, además se ofrecen algunas aplicaciones.

El capítulo dos estudia el problema de la unicidad de soluciones encontradas en el capítulo uno, caso de que existan. Se utilizan varios argumentos, convexidad, teoría de gráfos …

El capítulo tres está dedicado al estudio del caso de incompatibilidad. Esto es si el problema no tiene solución, no existe una distribución conjunta que tenga a las condicionadas dadas como sus propias distribuciones condicionadas. En este caso, podemos encontrar soluciones aproximadas que sean óptimas en algún sentido.

En el capítulo cuatro, se extiende el problema de compatibilidad estudiado a vectores aleatorios de rango infinito numerable.

En el capítulo cinco se propone una aplicación. Se trata de modelizar situaciones que reflejen la posibilidad de que el conocimiento probabilístico de dos agentes, sobre cierta situación competitiva, se reduzca a poder modelizar las reacciones de ambos ante la acción del otro.

En los dos anexos se recoge material que completa nuestro trabajo.

Capítulo 1

Compatibilidad.

Comenzamos desarrollando un tratamiento matricial de la probabilidad condicionada, que nos lleva a una generalización del Teorema de Bayes. Este tratamiento será de utilidad en este trabajo. Estudiamos, a lo largo del capítulo, el problema de la compatibilidad de distribuciones condicionadas, que para distribuciones discretas y finitas, caso que nos ocupa, se traducen en dos matrices no negativas verificando alguna condición adicional, nos preguntamos entonces ¿bajo qué condiciones existe un vector aleatorio que las tenga como distribuciones condicionadas? Tras estudiar los primeros resultados, estudiamos algunos métodos más generales, en concreto, la representación marginal uniforme, el uso de la fórmula de Besag y el método de extensión de rango unidad. Finalizamos estudiando algunas aplicaciones de este último método.

1.- Tratamiento matricial de probabilidades condicionadas.

Consideremos dos sistemas completos de sucesos $A_1, A_2, ..., A_i, ..., A_n$ con probabilidades no nulas y $B_1, B_2, ..., B_j, ..., B_m$ con probabilidades $P(B_1), P(B_2), ..., P(B_j), ..., P(B_m)$ no nulas.

Definimos los $P(A_1), P(A_2), \ldots, P(A_i), \ldots, P(A_n)$ vectores

$$\vec{\alpha} = \begin{bmatrix} P(A_1) \\ P(A_2) \\ \vdots \\ P(A_n) \end{bmatrix} \qquad \vec{\beta} = \begin{bmatrix} P(B_1) \\ P(B_2) \\ \vdots \\ P(B_m) \end{bmatrix}$$

Y las matrices

$$A = ((a_{ij})) \qquad a_{ij} = P(A_i \mid B_j), \quad i = 1,2,\ldots,n; \quad j = 1,2,\ldots,m$$

$$A = \begin{bmatrix} P(A_1 \mid B_1) & P(A_1 \mid B_2) & \ldots & P(A_1 \mid B_m) \\ P(A_2 \mid B_1) & P(A_2 \mid B_2) & \ldots & P(A_2 \mid B_m) \\ \vdots & \vdots & \ddots & \vdots \\ P(A_n \mid B_1) & P(A_n \mid B_2) & \ldots & P(A_n \mid B_m) \end{bmatrix}$$

$$B = ((b_{ij})) \qquad b_{ij} = P(B_j \mid A_i), \quad j = 1,2,\ldots,m; \quad i = 1,2,\ldots,n$$

$$B = \begin{bmatrix} P(B_1 \mid A_1) & P(B_2 \mid A_1) & \dots & P(B_m \mid A_1) \\ P(B_1 \mid A_2) & P(B_2 \mid A_2) & \dots & P(B_m \mid A_2) \\ \vdots & \vdots & \ddots & \vdots \\ P(B_1 \mid A_n) & P(B_2 \mid A_n) & \dots & P(B_m \mid A_n) \end{bmatrix}$$

Ambas matrices tienen dimensión $n \times m$.

Denominamos $\vec{1}_n$ al vector columna de dimensión n con todas sus componentes iguales a uno, y por $\vec{0}_n$ al vector columna de dimensión n con todas componentes iguales a cero.

Definición: Una matriz se dice que es estocástica por filas (columnas) si es no negativa y la suma de los elementos de cada fila (columna) vale uno.

Teorema 1.1.1. A es estocástica por columnas y B lo es por filas.

Observemos que $\vec{1}_n^t A = \vec{1}_m^t$ y $B\vec{1}_m = \vec{1}_n$.

Teorema 1.1.2 (Teorema de la probabilidad total, forma matricial).

$$A\vec{\beta} = \vec{\alpha} \qquad\qquad B^t \vec{\alpha} = \vec{\beta}$$

Demostración: nos basamos en el teorema de la probabilidad total en su forma habitual:

$$A\vec{\beta} = \begin{bmatrix} P(A_1 \mid B_1) & P(A_1 \mid B_2) & \dots & P(A_1 \mid B_m) \\ P(A_2 \mid B_1) & P(A_2 \mid B_2) & \dots & P(A_2 \mid B_m) \\ \vdots & \vdots & \ddots & \vdots \\ P(A_n \mid B_1) & P(A_n \mid B_2) & \dots & P(A_n \mid B_m) \end{bmatrix} \begin{bmatrix} P(B_1) \\ P(B_2) \\ \vdots \\ P(B_m) \end{bmatrix} =$$

$$= \begin{bmatrix} \sum_{j=1}^{m} P(A_1 \mid B_j) P(B_j) \\ \sum_{j=1}^{m} P(A_2 \mid B_j) P(B_j) \\ \vdots \\ \sum_{j=1}^{m} P(A_n \mid B_j) P(B_j) \end{bmatrix} = \begin{bmatrix} P(A_1) \\ P(A_2) \\ \vdots \\ P(A_n) \end{bmatrix} = \vec{\alpha}$$

Análogamente se demuestra $B^t \vec{\alpha} = \vec{\beta}$

□

Teorema 1.1.3. Las matrices AB^t y $B^t A$ tienen autovalor uno con $\vec{\alpha}$ y $\vec{\beta}$ autovectores asociados a ese valor y, respectivamente, a cada matriz.

Demostración: combinando las fórmulas de la proposición anterior se tiene:

$$A\vec{\beta} = \vec{\alpha} \rightarrow AB^t \vec{\alpha} = \vec{\alpha}$$

$$B^t \vec{\alpha} = \vec{\beta} \rightarrow B^t A \vec{\beta} = \vec{\beta}$$

□

Como consecuencia de lo anterior:

$$AB^t \vec{\alpha} = \vec{\alpha} \rightarrow (AB^t - I_n)\vec{\alpha} = \vec{0}_n \rightarrow \det(AB^t - I_n) = 0$$

$$B^t A \vec{\beta} = \vec{\beta} \rightarrow (B^t A - I_m)\vec{\beta} = \vec{0}_m \rightarrow \det(B^t A - I_m) = 0$$

Teorema 1.1.4.

$$\vec{\alpha}^t A\vec{\beta} = \|\vec{\alpha}\|_2^2 \qquad \vec{\beta}^t B^t\vec{\alpha} = \|\vec{\beta}\|_2^2 \qquad \vec{\alpha}^t B\vec{\beta} = \|\vec{\beta}\|_2^2$$

Demostración: Basta aplicar el teorema 1.1.2

$$\vec{\alpha}^t A\vec{\beta} = \vec{\alpha}^t\vec{\alpha} = \|\vec{\alpha}\|_2^2$$

$$\vec{\beta}^t B^t\vec{\alpha} = \vec{\beta}^t\vec{\beta} = \|\vec{\beta}\|_2^2$$

Por otra parte, la última igualdad se obtiene trasponiendo la segunda.

□

Sea M la matriz de las probabilidades de las intersecciones de los dos sistemas completos de sucesos:

$$M = \begin{bmatrix} P(A_1B_1) & P(A_1B_2) & \dots & P(A_1B_m) \\ P(A_2B_1) & P(A_2B_2) & \dots & P(A_2B_m) \\ \vdots & \vdots & \ddots & \vdots \\ P(A_nB_1) & P(A_nB_2) & \dots & P(A_nB_m) \end{bmatrix}$$

Teorema 1.1.5 (Teorema de la probabilidad total no condicional).

$$\vec{1}_n^t M = \vec{\beta}^t \qquad M^t\vec{1}_n = \vec{\beta} \qquad M\vec{1}_m = \vec{\alpha}$$

Demostración: Basta operar en el primer miembro de cada ecuación,

$$\vec{1}_n^t M = \begin{bmatrix} 1 & 1 & \dots & 1 \end{bmatrix} \begin{bmatrix} P(A_1B_1) & P(A_1B_2) & \dots & P(A_1B_m) \\ P(A_2B_1) & P(A_2B_2) & \dots & P(A_2B_m) \\ \vdots & \vdots & \ddots & \vdots \\ P(A_nB_1) & P(A_nB_2) & \dots & P(A_nB_m) \end{bmatrix} =$$

$$= \begin{bmatrix} \sum_{i=1}^{n} P(A_iB_1) & \sum_{i=1}^{n} P(A_iB_2) & \dots & \sum_{i=1}^{n} P(A_iB_m) \end{bmatrix} = \vec{\beta}^t$$

La segunda igualdad se obtiene por trasposición de la primera. La última se prueba como la primera.

□

Definamos ahora las matrices $D(\vec{\alpha})$ y $D(\vec{\beta})$ como las matrices diagonales con diagonal el vector que se indica.

$$D(\vec{\alpha}) = \begin{bmatrix} P(A_1) & 0 & \dots & 0 \\ 0 & P(A_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & P(A_n) \end{bmatrix} \quad D(\vec{\beta}) = \begin{bmatrix} P(B_1) & 0 & \dots & 0 \\ 0 & P(B_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & P(B_m) \end{bmatrix}$$

Teorema 1.1.6.

$$AD(\vec{\beta}) = M = D(\vec{\alpha})B$$

Demostración:

$$AD(\vec{\beta}) = A = \begin{bmatrix} P(A_1 | B_1) & P(A_1 | B_2) & \dots & P(A_1 | B_m) \\ P(A_2 | B_1) & P(A_2 | B_2) & \dots & P(A_2 | B_m) \\ \vdots & \vdots & \ddots & \vdots \\ P(A_n | B_1) & P(A_n | B_2) & \dots & P(A_n | B_m) \end{bmatrix} \begin{bmatrix} P(B_1) & 0 & \dots & 0 \\ 0 & P(B_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & P(B_m) \end{bmatrix} =$$

$$= \begin{bmatrix} P(A_1 \mid B_1)P(B_1) & P(A_1 \mid B_2)P(B_2) & \dots & P(A_1 \mid B_m)P(B_m) \\ P(A_2 \mid B_1)P(B_1) & P(A_2 \mid B_2)P(B_2) & \dots & P(A_2 \mid B_m)P(B_m) \\ \vdots & \vdots & \ddots & \vdots \\ P(A_n \mid B_1)P(B_1) & P(A_n \mid B_2)P(B_2) & \dots & P(A_n \mid B_m)P(B_m) \end{bmatrix} =$$

$$= \begin{bmatrix} P(A_1B_1) & P(A_1B_2) & \dots & P(A_1B_m) \\ P(A_2B_1) & P(A_2B_2) & \dots & P(A_2B_m) \\ \vdots & \vdots & \ddots & \vdots \\ P(A_nB_1) & P(A_nB_2) & \dots & P(A_nB_m) \end{bmatrix} = M$$

La demostración de la otra igualdad es análoga.

□

Teorema 1.1.7 (Teorema de Bayes, forma matricial).

$$A = D(\vec{\alpha})BD(\vec{\beta})^{-1} \qquad\qquad B = D(\vec{\alpha})^{-1}AD(\vec{\beta})$$

Demostración: Es consecuencia inmediata del teorema 1.1.6

Recuérdese que cualquier variable aleatoria discreta de rango finito X¸ con rango $R_X = \{x_1 x_2, \dots, x_n\}$ tiene asociado un sistema completo de sucesos que viene dado por:

$$A_k = [X = x_k], \qquad\qquad k = 1, 2, \dots, n$$

Este hecho será utilizado con asiduidad en lo que sigue.

2. Planteamiento del problema y notación.

En la sección anterior se considera el hecho de que dados dos sistemas completos de sucesos o dos variables aleatorias discretas finitas se pueden construir matrices A y B y vectores $\vec{\alpha}$ y $\vec{\beta}$ de probabilidades condicionadas y marginales respectivamente que verifican ciertas propiedades. En esta sección plantearemos el problema contrario, es decir, dadas un par de matrices de dimensiones y propiedades adecuadas ¿cuándo pueden ser consideradas como las matrices de probabilidades condicionadas de un vector aleatorio bidimensional discreto y finito? Concretemos.

Sean dos matrices $A=(a_{ij})$ y $B=(b_{ij})$ de orden $n\times m$ que verifican $a_{ij} \geq 0$, $b_{ij} \geq 0$ $\forall i = 1,\dots,n;\ \forall j = 1,\dots,m$ y además

$$\sum_{i=1}^{n} a_{ij} = 1, \quad \sum_{j=1}^{m} b_{ij} = 1.$$

En consecuencia, A y B^t son matrices estocásticas por columnas.

La cuestión a dilucidar es, bajo qué condiciones dos matrices A y B, de dimensiones $n\times m$, en las condiciones anteriores, pueden ser consideradas distribuciones condicionadas de cierto vector aleatorio (X,Y) de variables discretas con recorrido finito. Esto es si $x_1, x_2, \dots, x_n$ son los posibles valores de X e $y_1, y_2, \dots, y_m$ los de la variable Y

$$a_{ij} = P\left[X = x_i \,|Y = y_j \right] \qquad i = 1...n$$
$$b_{ij} = P\left[Y = y_j \,|X = x_i \right] \qquad j = 1...\,m$$

Equivalentemente, bajo qué condiciones existen dos vectores $\vec{\alpha}$ y $\vec{\beta}$, de dimensiones n y m respectivamente, positivos y de forma que $\alpha_i b_{ij} = \beta_j a_{ij} \quad i = 1,...,n;\ j = 1,...,m$ pues la versión normalizada (la suma de sus coordenadas es la unidad) de $\vec{\alpha}$ y $\vec{\beta}$ serían las distribuciones marginales correspondientes. La condición anterior indica que *A y* $\vec{\alpha}$ determinan la misma matriz de probabilidades conjuntas que *B* y $\vec{\beta}$. Nótese que podemos suponer $\alpha_i, \beta_j > 0 \qquad i = 1,...,n;\ j = 1,...,m$ ya que si algún α_i o algún β_j es nulo, el valor x_i o y_j tiene todas sus probabilidades asociadas nulas en *A* y *B* y, una vez identificado, puede ser eliminado del problema.

En forma matricial, la condición anterior viene dada por $D(\vec{\alpha})B = AD(\vec{\beta})$.

Esta primera cuestión recibe el nombre de problema de compatibilidad y, en este contexto, las matrices *A* y *B* se denominan candidatas a distribuciones condicionadas, además, si la respuesta al problema es afirmativa se dice que *A* y *B* son compatibles.

Ejemplo 1: Consideremos las matrices

$$A = \begin{bmatrix} \frac{1}{3} & \frac{1}{2} & \frac{3}{5} \\ \frac{2}{3} & \frac{1}{2} & \frac{2}{5} \end{bmatrix} \qquad B = \begin{bmatrix} \frac{1}{5} & \frac{1}{5} & \frac{3}{5} \\ \frac{2}{5} & \frac{1}{5} & \frac{2}{5} \end{bmatrix}$$

Que son compatibles, pues existen vectores, dados por:

$$\vec{\alpha} = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix} \qquad \vec{\beta} = \begin{bmatrix} \frac{3}{10} \\ \frac{2}{10} \\ \frac{5}{10} \end{bmatrix}$$

Que verifican la igualdad de las siguientes matrices:

$$AD(\vec{\beta}) = \begin{bmatrix} \frac{1}{3} & \frac{1}{2} & \frac{3}{5} \\ \frac{2}{3} & \frac{1}{2} & \frac{2}{5} \end{bmatrix} \begin{bmatrix} \frac{3}{10} & 0 & 0 \\ 0 & \frac{2}{10} & 0 \\ 0 & 0 & \frac{5}{10} \end{bmatrix} = \begin{bmatrix} \frac{1}{10} & \frac{1}{10} & \frac{3}{10} \\ \frac{2}{10} & \frac{1}{10} & \frac{2}{10} \end{bmatrix}$$

$$D(\vec{\alpha})B = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} \frac{1}{5} & \frac{1}{5} & \frac{3}{5} \\ \frac{2}{5} & \frac{1}{5} & \frac{2}{5} \end{bmatrix} = \begin{bmatrix} \frac{1}{10} & \frac{1}{10} & \frac{3}{10} \\ \frac{2}{10} & \frac{1}{10} & \frac{2}{10} \end{bmatrix}$$

Ejemplo 2: Consideremos las matrices

$$A = \begin{bmatrix} \frac{1}{3} & \frac{1}{2} & \frac{3}{5} \\ \frac{2}{3} & \frac{1}{2} & \frac{2}{5} \end{bmatrix} \qquad B = \begin{bmatrix} \frac{2}{5} & 0 & \frac{3}{5} \\ \frac{2}{5} & \frac{1}{5} & \frac{2}{5} \end{bmatrix}$$

Estas matrices son incompatibles pues para cualquier elección de los vectores $\vec{\alpha}$ y $\vec{\beta}$ sin componentes nulas, $\alpha_1 b_{12} = \alpha_1 0 = 0$ y $\beta_2 a_{12} \neq 0$.

Un segundo problema a tratar es, supuesto que existan, la unicidad de estas variables X e Y o vectores $\vec{\alpha}$ y $\vec{\beta}$. Veremos que los vectores marginales $\vec{\alpha}$ y $\vec{\beta}$, de existir, no tienen que ser únicos.

Ejemplo2: Consideremos las matrices

$$A = \begin{bmatrix} 0.25 & 0.50 & 0 & 0 \\ 0.75 & 0.50 & 0 & 0 \\ 0 & 0 & 0.50 & 0.50 \\ 0 & 0 & 0.50 & 0.50 \end{bmatrix} \qquad B = \begin{bmatrix} 0.50 & 0.50 & 0 & 0 \\ 0.75 & 0.25 & 0 & 0 \\ 0 & 0 & 0.50 & 0.50 \\ 0 & 0 & 0.50 & 0.50 \end{bmatrix}$$

Y la familia de vectores,

$$\vec{\alpha} = \begin{bmatrix} \frac{\lambda}{3\lambda+4} \\ \frac{2\lambda}{3\lambda+4} \\ \frac{2}{3\lambda+4} \\ \frac{2}{3\lambda+4} \end{bmatrix} \qquad \vec{\beta} = \begin{bmatrix} \frac{2\lambda}{3\lambda+4} \\ \frac{\lambda}{3\lambda+4} \\ \frac{2}{3\lambda+4} \\ \frac{2}{3\lambda+4} \end{bmatrix}$$

que se han seleccionado con suma unidad. Veamos que para cualquier $\lambda > 0$ hay compatibilidad.

$$AD(\vec{\beta}) = \begin{bmatrix} 0.25 & 0.50 & 0 & 0 \\ 0.75 & 0.50 & 0 & 0 \\ 0 & 0 & 0.50 & 0.50 \\ 0 & 0 & 0.50 & 0.50 \end{bmatrix} \begin{bmatrix} \frac{2\lambda}{3\lambda+4} & 0 & 0 & 0 \\ 0 & \frac{\lambda}{3\lambda+4} & 0 & 0 \\ 0 & 0 & \frac{2}{3\lambda+4} & 0 \\ 0 & 0 & 0 & \frac{2}{3\lambda+4} \end{bmatrix} =$$

$$= \begin{bmatrix} \frac{0.5\lambda}{3\lambda+4} & \frac{0.5\lambda}{3\lambda+4} & 0 & 0 \\ \frac{1.5\lambda}{3\lambda+4} & \frac{0.5\lambda}{3\lambda+4} & 0 & 0 \\ 0 & 0 & \frac{1}{3\lambda+4} & \frac{1}{3\lambda+4} \\ 0 & 0 & \frac{1}{3\lambda+4} & \frac{1}{3\lambda+4} \end{bmatrix}$$

$$D(\vec{\alpha})B = \begin{bmatrix} \frac{\lambda}{3\lambda+4} & 0 & 0 & 0 \\ 0 & \frac{2\lambda}{3\lambda+4} & 0 & 0 \\ 0 & 0 & \frac{2}{3\lambda+4} & 0 \\ 0 & 0 & 0 & \frac{2}{3\lambda+4} \end{bmatrix} \begin{bmatrix} 0.50 & 0.50 & 0 & 0 \\ 0.75 & 0.25 & 0 & 0 \\ 0 & 0 & 0.50 & 0.50 \\ 0 & 0 & 0.50 & 0.50 \end{bmatrix} =$$

$$= \begin{bmatrix} \frac{0.5\lambda}{3\lambda+4} & \frac{0.5\lambda}{3\lambda+4} & 0 & 0 \\ \frac{1.5\lambda}{3\lambda+4} & \frac{0.5\lambda}{3\lambda+4} & 0 & 0 \\ 0 & 0 & \frac{1}{3\lambda+4} & \frac{1}{3\lambda+4} \\ 0 & 0 & \frac{1}{3\lambda+4} & \frac{1}{3\lambda+4} \end{bmatrix}$$

Para las matrices *A* y *B* definimos:

$$N_A = \{(i,j):\ a_{ij} > 0\}\ , \quad N_B = \{(i,j):\ b_{ij} > 0\}$$

Cuando se verifique que $N_A = N_B$ los denotaremos, indistintamente, por *N*.

Así mismo, denotaremos por $N_{n\times m} = \{(i,j):\ i = 1,2,...,n; \quad j = 1,2,...,m\}$

3.- Compatibilidad. Primeros resultados.

Un primer resultado sobre compatibilidad puede encontrarse en cualquier referencia sobre este tema y refleja el hecho,

$$P(X = x, Y = y) = 0 \leftrightarrow P(X = x \mid Y = y) = 0$$

$$P(X = x, Y = y) = 0 \leftrightarrow P(Y = y \mid X = x) = 0$$

Teorema 1.3.1 (Condición Necesaria de Compatibilidad). Si las matrices *A* y *B*, en las condiciones de la sección anterior, son compatibles, entonces $N_A = N_B$

Es decir, si *A* y *B* son compatibles, sus ceros están en las mismas posiciones.

En vista de esta condición necesaria, salvo mención expresa de lo contrario asumiremos que $N_A = N_B = N$

Arnold y Press (1989) presentan el siguiente resultado:

Teorema 1.3.2 (Arnold y Press, 1989). Dos matrices *A* y *B* en las condiciones de la sección anterior son compatibles si y sólo si $N_A = N_B = N$ y además existen dos vectores, $\vec{u} = \left(u_1, u_2, ..., u_n\right)$ y $\vec{v} = \left(v_1, v_2, ..., v_m\right)$ tales que

$$c_{ij} = \frac{a_{ij}}{b_{ij}} = u_i v_j \;\; \forall \left(i, j\right) \in N$$

Nótese que si $\left(i, j\right) \in N_{n\times m} - N$ los cocientes $c_{ij} = \frac{a_{ij}}{b_{ij}}$ no están definidos pero pueden ser calculados bajo hipótesis de compatibilidad. Si estos valores se organizan en forma de tabla de doble entrada de forma que para $\left(i, j\right) \in N_{n\times m}$

$$c_{ij} = \begin{cases} \dfrac{a_{ij}}{b_{ij}} & \left(i, j\right) \in N \\ * & \left(i, j\right) \in N \end{cases}$$

Donde * es un valor desconocido a determinar.

Esta tabla se denomina, matriz incompleta de cocientes. Nótese que es una matriz, en el sentido habitual del término, si $N = N_{n\times m}$.

Por otra parte, el teorema 3.1 de Arnold y Press (1989) afirma:

Dos matrices *A* y *B* en las condiciones habituales son compatibles si y sólo si $N_A = N_B = N$ y además para cualquier (i_1, i_2, j_1, j_2) tales que $(i_1, j_1), (i_1, j_2), (i_2, j_1), (i_2, j_2) \in N$ se verifica que $c_{i_1 j_1} \times c_{i_2 j_2} = c_{i_1 j_2} \times c_{i_2 j_1}$.

Si bien la condición necesaria se verifica, pues:

si *A* y *B* son compatibles, existe un vector aleatorio (*X*,*Y*) de variables discretas con recorrido finito. $x_1, x_2, \ldots, x_n$ son los posibles valores de *X* e $y_1, y_2, \ldots, y_m$ los de *Y* tal que

$$a_{ij} = P\left[X = x_i \mid Y = y_j\right] \qquad i = 1 \ldots n$$
$$b_{ij} = P\left[Y = y_j \mid X = x_i\right] \qquad j = 1 \ldots m$$

Así, para (i_1, i_2, j_1, j_2) tales que $(i_1, j_1), (i_1, j_2), (i_2, j_1), (i_2, j_2) \in N$

$$c_{i_1 j_1} \times c_{i_2 j_2} = \frac{a_{i_1 j_1}}{b_{i_1 j_1}} \times \frac{a_{i_2 j_2}}{b_{i_2 j_2}} = \frac{P(X = x_{i_1} \mid Y = y_{j_1})}{P(Y = y_{j_1} \mid X = x_{i_1})} \times \frac{P(X = x_{i_2} \mid Y = y_{j_2})}{P(Y = y_{j_2} \mid X = x_{i_2})} = \frac{P(X = x_{i_1})}{P(Y = y_{j_1})} \times \frac{P(X = x_{i_2})}{P(Y = y_{j_2})}$$
$$c_{i_1 j_2} \times c_{i_2 j_1} = \frac{a_{i_1 j_2}}{b_{i_1 j_2}} \times \frac{a_{i_2 j_1}}{b_{i_2 j_1}} = \frac{P(X = x_{i_1} \mid Y = y_{j_2})}{P(Y = y_{j_2} \mid X = x_{i_1})} \times \frac{P(X = x_{i_2} \mid Y = y_{j_1})}{P(Y = y_{j_1} \mid X = x_{i_2})} = \frac{P(X = x_{i_1})}{P(Y = y_{j_2})} \times \frac{P(X = x_{i_2})}{P(Y = y_{j_1})}$$

No se verifica la condición suficiente, tal y cómo indican Song, Li, Chen, Jiang y Kuo (2010), que ofrecen el siguiente ejemplo

Ejemplo 1: (Song, Li, Chen, Jiang y Kuo, 2010)

Consideremos las matrices:

$$A = \begin{bmatrix} \frac{1}{2} & \frac{1}{3} & \frac{2}{3} & 0 \\ \frac{1}{2} & \frac{2}{3} & 0 & \frac{1}{3} \\ 0 & 0 & \frac{1}{3} & \frac{2}{3} \end{bmatrix} \qquad B = \begin{bmatrix} \frac{1}{2} & \frac{1}{6} & \frac{1}{3} & 0 \\ \frac{1}{2} & \frac{1}{3} & 0 & \frac{1}{6} \\ 0 & 0 & \frac{1}{6} & \frac{5}{6} \end{bmatrix}$$

En este caso sólo hay un (i_1, i_2, j_1, j_2) que verifique $(i_1, j_1), (i_1, j_2), (i_2, j_1), (i_2, j_2) \in N$ y resulta ser (1, 2, 1, 2): $(1,1), (1,2), (2,1), (2,2) \in N$ Para esos índices se verifica la condición $c_{11} \times c_{22} = c_{12} \times c_{21}$ Sin embargo son incompatibles.

Si fuesen compatibles existirían vectores $\vec{\alpha}$ y $\vec{\beta}$, no nulos y de dimensiones adecuadas, que verifican la siguiente ecuación matricial:

$$\begin{bmatrix}\alpha_1 & 0 & 0\\ 0 & \alpha_2 & 0\\ 0 & 0 & \alpha_3\end{bmatrix}\begin{bmatrix}\frac{1}{2} & \frac{1}{6} & \frac{1}{3} & 0\\ \frac{1}{2} & \frac{1}{3} & 0 & \frac{1}{6}\\ 0 & 0 & \frac{1}{6} & \frac{5}{6}\end{bmatrix}=\begin{bmatrix}\frac{1}{2} & \frac{1}{3} & \frac{2}{3} & 0\\ \frac{1}{2} & \frac{2}{3} & 0 & \frac{1}{3}\\ 0 & 0 & \frac{1}{3} & \frac{2}{3}\end{bmatrix}\begin{bmatrix}\beta_1 & 0 & 0 & 0\\ 0 & \beta_2 & 0 & 0\\ 0 & 0 & \beta_3 & 0\\ 0 & 0 & 0 & \beta_4\end{bmatrix}$$

$$\begin{bmatrix}\frac{\alpha_1}{2} & \frac{\alpha_1}{6} & \frac{\alpha_1}{3} & 0\\ \frac{\alpha_2}{2} & \frac{\alpha_2}{3} & 0 & \frac{\alpha_2}{6}\\ 0 & 0 & \frac{\alpha_3}{6} & \frac{5\alpha_3}{6}\end{bmatrix}=\begin{bmatrix}\frac{\beta_1}{2} & \frac{\beta_2}{3} & \frac{2\beta_3}{3} & 0\\ \frac{\beta_1}{2} & \frac{2\beta_2}{3} & 0 & \frac{\beta_4}{3}\\ 0 & 0 & \frac{\beta_3}{3} & \frac{2\beta_4}{3}\end{bmatrix}$$

De ahí, $\frac{5\alpha_3}{6}=\frac{2\beta_4}{3}\rightarrow\alpha_3=\frac{4\beta_4}{5}$ por la ecuación de la posición 3,4

Por otra parte:

$$\frac{\alpha_3}{\beta_4}\overset{3,3}{=}\frac{2\beta_3}{\beta_4}\overset{2,4}{=}\frac{2\beta_3}{\frac{\alpha_2}{2}}=\frac{4\beta_3}{\alpha_2}\overset{1,2;2,1}{=}\frac{4\beta_3}{\alpha_1}\overset{1,3}{=}\frac{4\beta_3}{2\beta_3}=2$$

Así, $\alpha_3=2\beta_4$ y $\alpha_3=\frac{4\beta_4}{5}$, por tanto $\alpha_3=\beta_4=0$ lo que es imposible pues, de ser así, la tercera fila de *A* sería nula (y la cuarta de *B*). Por tanto, *A* y *B* son incompatibles.

Por tanto, el resultado es la siguiente condición necesaria.

Teorema 1.3.3 (Arnold y Press , 1989). Dos matrices *A* y *B* en las condiciones habituales son compatibles entonces N_A = N_B = *N* y además para cualquier

(i_1, i_2, j_1, j_2) tales que $(i_1, j_1), (i_1, j_2), (i_2, j_1), (i_2, j_2) \in N$ se verifica que $c_{i_1 j_1} \times c_{i_2 j_2} = c_{i_1 j_2} \times c_{i_2 j_1}$.

Observemos que la condición es también suficiente si $N_{n\times m} = N$ pues, para este caso, sobre la condición $c_{i_1 j_1} \times c_{i_2 j_2} = c_{i_1 j_2} \times c_{i_2 j_1}$ sumando en i_2

$$\sum_{i_2=1}^{n} c_{i_1 j_1} \times c_{i_2 j_2} = \sum_{i_2=1}^{n} c_{i_1 j_2} \times c_{i_2 j_1} \rightarrow c_{i_1 j_1} \times c_{\bullet j_2} = c_{i_1 j_2} \times c_{\bullet j_1}$$

Y ahora, sumando en j_2

$$\sum_{j_2=1}^{m} c_{i_1 j_1} \times c_{\bullet j_2} = \sum_{j_2=1}^{m} c_{i_1 j_2} \times c_{\bullet j_1} \rightarrow c_{i_1 j_1} \times c_{\bullet\bullet} = c_{i_1 \bullet} \times c_{\bullet j_1}$$

De esta forma $c_{ij} \times c_{\bullet\bullet} = c_{i\bullet} \times c_{\bullet j} \quad \forall i, j \in N$ y por aplicación del teorema 1.3.1 se tiene la compatibilidad.

Tenemos por tanto,

Teorema 1.3.4 (Arnold y Press, 2002). Sean A y B dos matrices en las condiciones habituales con $N_{n\times m} = N$. A y B son compatibles si y sólo si para cualquier (i_1, i_2, j_1, j_2) tales que $(i_1, j_1), (i_1, j_2), (i_2, j_1), (i_2, j_2) \in N$ se verifica que $c_{i_1 j_1} \times c_{i_2 j_2} = c_{i_1 j_2} \times c_{i_2 j_1}$.

La clave de la validez de esta prueba en este caso, y su no validez en el caso en el que $N_{n\times m} \neq N$, estriba en que las sumas no dependen de los índices que se pretenden eliminar si los índices de sumación recorren todos los valores, cosa que no ocurre en el caso de $N_{n\times m} \neq N$.

En el caso que nos ocupa, variables finitas, es posible utilizar resultados de álgebra matricial para analizar la compatibilidad de las matrices *A* y *B*. En efecto, supongamos que *N*= $\{1,...,L\}\times\{1,...,M\}$ y denotemos por C la matriz de elementos $c_{i,j}=\dfrac{a_{i,j}}{b_{i,j}}$, Note que si $N_{n\times m}=N$ la matriz está completamente definida.

El uso de esta matriz nos permite obtener una condición necesaria de compatibilidad.

Teorema 1.3.5. Sean *A* y *B* en las condiciones habituales, con $N_{n\times m}=N$. Si *A* y B son compatibles, entonces $\sum_i c_{i\bullet}^{-1}=1$ donde $c_{i\bullet}=\sum_j c_{ij}$. O equivalentemente, $\sum_j c_{\bullet j}^{*-1}=1$ donde $c_{\bullet j}^{*}=\sum_i c_{ij}^{-1}$.

Demostración: Si las matrices son compatibles, existen dos vectores $\vec{\alpha}$ y $\vec{\beta}$, de dimensiones *n* y *m* respectivamente, no negativos, con suma de coordenadas unidad y de forma que $\alpha_i b_{ij}=\beta_j a_{ij}\quad i=1,...,n\quad ;j=1,...,m$, entonces $c_{ij}=\dfrac{\alpha_i}{\beta_j}\qquad i=1,...,n\quad ;j=1,...,m$ la comprobación de la fórmula es trivial.

□

Teorema 1.3.6 *(Pérez Villalta, 2000).* Sean *A* y *B* en las condiciones habituales, con $N_{n\times m}=N$. *A* y B son compatibles si y sólo si el rango de *C* es uno.

Demostración: Si rango de *C* es uno, existen dos matrices ortogonales *U y V* de órdenes *n* y *m* respectivamente, de forma que $C=U\Sigma V^t$ donde $\mathbf{\Sigma}$ es una matriz $n\times m$ cuyo elemento (1,1) es no nulo y el resto nulos. Obsérvese que esto no es más que la descomposición en valores singulares de la matriz *C* y el elemento no nulo de $\mathbf{\Sigma}$ el único valor singular de *C* (ver p.e. Schott, 1997). Dada la forma de la matriz $\mathbf{\Sigma}$ podemos escribir $C=\vec{u}\sigma\vec{v}^t$ donde $\vec{u}$ y $\vec{v}$ son las primeras columnas de *U y V* respectivamente. Así, $c_{i,j}=u_i\sigma v_j$, $\forall(i,j)\in N$ y por tanto, incluyendo el valor singular en uno de los vectores $\vec{u}$ y $\vec{v}$, se tiene la factorización $c_{i,j} = u_i.v_j$, $\forall (i,j) \in N$.

Por otra parte, puesto que $\vec{u}$ es autovector de CC^t , que es no negativa, y está asociado al autovalor $\sigma > 0$, puede ser elegido con coordenadas no negativas y puesto que $\vec{v} \propto C^t \vec{u}$ también es no negativo, las matrices A y B son compatibles y, en este caso, una adecuada elección de las distribuciones marginales es

$$P\left[X = x_i\right] = \alpha_i = \frac{u_i}{\sum_{i=1}^{L} u_i} \qquad i = 1,\ldots,n$$

$$P\left[Y = y_j\right] = \beta_j = \frac{v_j^{-1}}{\sum_{j=1}^{M} v_j^{-1}} \qquad j = 1\ldots,m$$

Recíprocamente, si las matrices A y B son compatibles, el rango de C es la unidad pues, si tomamos un menor de dimensión dos cualquiera,

$$\begin{vmatrix} c_{i,j} & c_{i,k} \\ c_{l,j} & c_{l,k} \end{vmatrix} = \begin{vmatrix} \dfrac{a_{i,j}}{b_{i,j}} & \dfrac{a_{i,k}}{b_{i,k}} \\ \dfrac{a_{l,j}}{b_{l,j}} & \dfrac{a_{l,k}}{b_{l,k}} \end{vmatrix} = \begin{vmatrix} u_i v_j & u_i v_k \\ u_l v_j & u_l v_k \end{vmatrix} = u_i v_j u_l v_k - u_i v_k u_l v_j = 0,$$

$$\forall i,l = 1,\ldots,n \text{ y } \forall j,k = 1,\ldots,m .$$

□

Observación: Es obvio que este resultado se puede demostrar fácilmente utilizando el teorema 1.3.2, sin embargo, el uso de la descomposición en valores singulares de la matriz de cocientes será de utilidad en capítulos posteriores.

Observemos que, en este caso, la condición de compatibilidad vía descomposición singular se simplifica notablemente pues, como observan Arnold y Press(1989) y Arnold, Castillo y Sarabia (1992), se verifica que:

$$c_{i,j}c_{\bullet\bullet} = c_{i\bullet}c_{\bullet j} \quad \forall (i,j) \in N \quad \text{donde}$$

$$c_{\bullet\bullet} = \sum_{i,j} c_{ij} \quad , \quad c_{\bullet j} = \sum_{i} c_{ij} \quad \text{y} \quad c_{i\bullet} = \sum_{j} c_{ij}$$

Y en consecuencia $u_i \propto c_{i\bullet}$ y $v_j \propto c_{\bullet j}$ $\forall (i,j) \in N$

Nótese que la condición $c_{ij}c_{\bullet} = c_{i\bullet}c_{\bullet j}$ $\forall (i,j) \in N$ es una condición necesaria y suficiente de compatibilidad en el caso $N_{n\times m} = N$.

Teorema 1.3.7 *(Arnold y Press, 1989).* Sean A y B dos matrices en las condiciones habituales con $N_{n\times m} = N$. A y B son compatibles si y sólo $c_{i,j}c_{\bullet\bullet} = c_{i\bullet}c_{\bullet j}$ $\forall (i,j) \in N$

En forma matricial, este resultado adopta la siguiente forma:

Teorema 1.3.8. Sean A y B en las condiciones habituales, con $N_{n\times m} = N$. A y B son compatibles si y sólo si $C\left(\vec{1}_n^t C \vec{1}_m\right) = C\vec{1}_m \vec{1}_n^t C = C 1_{m\times n} C$.

Demostración: Basta observar que

$$\vec{1}_n^t C = \begin{pmatrix} c_{\bullet 1} & c_{\bullet 2} & \ldots & c_{\bullet m} \end{pmatrix}$$

$$C\vec{1}_m = \begin{pmatrix} c_{1\bullet} \\ c_{2\bullet} \\ \vdots \\ c_{m\bullet} \end{pmatrix}$$

$$\vec{1}_n^t C \vec{1}_m = c_{\bullet\bullet}$$

Y transformar las condición $c_{i,j}c_{..} = c_{i.}c_{.j} \quad \forall (i,j) \in N$ en $C\left(\vec{1}_n^t C \vec{1}_m\right) = C\vec{1}_m \vec{1}_n^t C$.

□

Ejemplo 2. Consideremos las matrices:

$$A = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{4}{5} & \frac{1}{5} \end{bmatrix} \qquad B = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix}$$

La matriz de cocientes es $C = \begin{bmatrix} \frac{15}{10} & \frac{15}{20} \\ \frac{24}{20} & \frac{6}{10} \end{bmatrix} = \frac{3}{20}\begin{bmatrix} 10 & 5 \\ 8 & 4 \end{bmatrix}$, tiene rango unidad y por

tanto *A* y *B* son compatibles

Luego $v_1 \propto \frac{15}{10} + \frac{24}{20} = \frac{54}{20}$ y $v_2 \propto \frac{15}{20} + \frac{6}{10} = \frac{27}{20}$. Por otra parte, $u_1 \propto \frac{15}{10} + \frac{15}{20} = \frac{45}{20}$

y $u_2 \propto \frac{24}{20} + \frac{6}{10} = \frac{36}{20}$ podemos tomar por tanto

$(v_1, v_2) \propto (2,1)$ y $(u_1, u_2) \propto (5,4)$

El producto $(5,4)^t$ (2,1) debe ser una matriz que, elemento a elemento, es proporcional a *C*.

$$\begin{bmatrix} 4 \\ 5 \end{bmatrix} [2,1] = \begin{bmatrix} 10 & 5 \\ 8 & 4 \end{bmatrix}$$

Multiplicando esta matriz por $\frac{3}{20}$ se obtiene *C*.

Ejemplo 2: Las matrices:

$$A = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{3}{5} & \frac{2}{5} \end{bmatrix} \qquad B = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix}$$

son incompatibles pues la matriz de cocientes C tiene, claramente, rango dos:

$$C = \begin{bmatrix} \frac{15}{10} & \frac{15}{20} \\ \frac{18}{20} & \frac{12}{10} \end{bmatrix} = \frac{3}{20} \begin{bmatrix} 10 & 5 \\ 6 & 8 \end{bmatrix}$$

Definición. Dada una matriz *T* de dimensiones 2×2 con elementos positivos. Se denomina razón de productos cruzados, a $\frac{t_{11}t_{22}}{t_{12}t_{21}}$.

Teorema 9 (Arnold y Gokhale, 1994, 1998). Sean *A* y *B* en las condiciones habituales, con $N_{n\times m} = N$. *A* y *B* son compatibles si y sólo si todas las razones de productos cruzados son iguales en *A* y *B*.

Demostración: consideremos cualquier menor de orden dos en *A*:

$$\begin{bmatrix} a_{ij} & a_{ik} \\ a_{lj} & a_{lk} \end{bmatrix} \qquad \forall i,l = 1,\dots,n \text{ y } \forall j,k = 1,\dots,m$$

Su razón de producto cruzado es $\dfrac{a_{ij}a_{lk}}{a_{ik}a_{lj}}$. Para *B*

$$\begin{bmatrix} b_{ij} & b_{ik} \\ b_{lj} & b_{lk} \end{bmatrix} \qquad \forall i,l = 1,\dots,n \text{ y } \forall j,k = 1,\dots,m .$$

Su razón de producto cruzado es $\dfrac{b_{ij}b_{lk}}{b_{ik}b_{lj}}$.

Dividiendo ambas razones, obtenemos:

$$\frac{\dfrac{a_{ij}a_{lk}}{a_{ik}a_{lj}}}{\dfrac{b_{ij}b_{lk}}{b_{ik}b_{lj}}} = \frac{c_{ij}c_{lk}}{c_{ik}c_{lj}}$$

Suponiendo que $N_{n\times m} = N$ tenemos

A y B compatibles si y sólo si $rang\ C = 1 \leftrightarrow c_{ij}c_{lk} - c_{ik}c_{lj} = 0$

$\forall i, l = 1, \ldots, n$ y $\forall j, k = 1, \ldots, m \leftrightarrow \dfrac{c_{ij}c_{lk}}{c_{ik}c_{lj}} = 1\ \forall i, l = 1, \ldots, n$ y $\forall j, k = 1, \ldots, m \leftrightarrow$

$\leftrightarrow \dfrac{\dfrac{a_{ij}a_{lk}}{a_{ik}a_{lj}}}{\dfrac{b_{ij}b_{lk}}{b_{ik}b_{lj}}} = 1\ \forall i, l = 1, \ldots, n$ y $\forall j, k = 1, \ldots, m$ si y sólo si las razones de los productos cruzados son iguales en A y B.

□

Es importante señalar que la condición de positividad $N_{n\times m} = N$ es imprescindible, véase Arnold y Gokhale (1994).

El siguiente resultado relaja la condición de positividad de A y B,

Teorema 1.3.9 (Gupta y Vargas, 1990). Sean A y B dos matrices en las condiciones habituales $N_A = N_B = N$ y supongamos que $\exists\ (i_0, j_0) \in N$ tal que, para cada $j = 1, \ldots, M$ $(i_0, j) \in N$ y que para cada $i = 1, \ldots, L$ $(i, j_0) \in N$. Entonces, A y B son compatibles si y sólo si

$$b_{ij} \frac{a_{ij_0}}{b_{ij_0}} = \kappa\, a_{ij} \frac{b_{i_0 j}}{a_{i_0 j}} \qquad \forall (i, j) \in N$$

donde κ es una constante que depende únicamente de i_0 y de j_0 y es distinta para cada elección de estos índices.

Puede comprobarse que si A y B son compatibles la igualdad se tiene tomando

$$\kappa = \left[\frac{\sum_{j=1}^{M} c_{i_0 j}^{-1}}{\sum_{i=1}^{L} c_{ij_0}} \right]^{-1}$$

mientras que el recíproco se obtiene tomando la distribución conjunta

$$p_{ij} = \kappa' b_{ij} \frac{a_{ij_0}}{b_{ij_0}} \text{ para } (i,j) \in N$$

$$p_{ij} = 0 \qquad\qquad \text{para } (i,j) \notin N$$

siendo $$\kappa' = \left[\sum_{i=1}^{L} c_{ij_0} \right]^{-1}$$

Ejemplo 6. Consideramos las matrices

$$A = \begin{bmatrix} 0 & 0 & \frac{1}{3} & \frac{1}{2} \\ 0 & 0 & \frac{1}{3} & \frac{1}{4} \\ 1 & 1 & \frac{1}{3} & \frac{1}{4} \end{bmatrix} \quad \text{y} \qquad B = \begin{bmatrix} 0 & 0 & \frac{1}{3} & \frac{2}{3} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{2}{5} & \frac{2}{5} & \frac{1}{5} & \frac{1}{5} \end{bmatrix}$$

entonces, $N_A = N_B = N = \{(1,3),(1,4),(2,3),(2,4),(3,1),(3,2),(3,3),(3,4)\}$

La elección de i_0 y j_0 en este caso ofrece dos posibilidades. $(i_0, j_0) = (3,3)$ o bien $(i_0, j_0) = (3,4)$; tomemos la primera posibilidad

$$\kappa = \left[\frac{\sum_{j=1}^{4} c_{3\ j}^{-1}}{\sum_{i=1}^{3} c_{i3}} \right]$$

Comprobar las igualdades es trivial.

La distribución conjunta será:

$$p_{ij} = \kappa' b_{ij} \frac{a_{i3}}{b_{i3}} \text{ para } (i,j) \in N \text{ con } \kappa' = \left[\sum_{i=1}^{3} c_{i3} \right] = \frac{3}{10}$$

Obteniéndose, entonces,

$$(p_{ij}) = \frac{1}{10} \begin{bmatrix} 0 & 0 & 1 & 2 \\ 0 & 0 & 1 & 1 \\ 1 & 2 & 1 & 1 \end{bmatrix}$$

Ejemplo 7: Las siguientes matrices no pueden ser estudiadas utilizando el resultado anterior ya que en cada fila y columna hay, al menos un elemento nulo.

$$A = \begin{bmatrix} 0.5 & 0.125 & 0 & 0 \\ 0.5 & 0.875 & 0 & 0 \\ 0 & 0 & 0.5 & 0.125 \\ 0 & 0 & 0.5 & 0.125 \end{bmatrix} \qquad B = \begin{bmatrix} 0.5 & 0.5 & 0 & 0 \\ 0.125 & 0.875 & 0 & 0 \\ 0 & 0 & 0.5 & 0.5 \\ 0 & 0 & 0.125 & 0.875 \end{bmatrix}$$

4.- Representación marginal uniforme.

En el contexto del estudio de tablas de contingencia, Mosteller (1968) introduce la representación marginal uniforme (UMR), y la utiliza para estudiar los patrones de asociación entre las variables cualitativas que son difíciles de ver, posiblemente por acción de las marginales, en la tabla inicial.

La UMR se calcula sobre tablas de contingencia, con frecuencias absolutas o relativas y, obviamente, sobre matrices no negativas. Suponemos que en cada fila o columna hay, al menos, una entrada positiva. El procedimiento de cálculo para el caso de tablas de contingencia con frecuencias relativas o matrices con elementos en el intervalo unidad, es el siguiente:

Dada una tabla de contingencia o matriz no negativa, de dimensiones $n \times m$, y elementos en el intervalo unidad,

(i) Normalizar cada fila de forma que sumen $1/n$: si <u>*SF(i)*</u> representa la suma de todos los elementos de la fila i, dividir cada elemento de esa fila entre $nSF(i)$ (i=1,2,…,n).

(ii) Normalizar cada columna de forma que sumen $1/m$: si <u>*SC(j)*</u> representa la suma de todos los elementos de la columna j, dividir cada elemento de esa columna entre $mSC(j)$ (j=1,2,…,m).

(iii) Repetir los pasos anteriores hasta que el procedimiento converja (si es que lo hace).

No existen estudios sobre la convergencia del método.

Ejemplo 1:

Sea la matriz $D = \begin{bmatrix} 0.1 & 0.1 & 0.2 & 0.1 \\ 0.1 & 0.1 & 0.2 & 0.2 \end{bmatrix}$

La suma de las filas y columnas son $SF = \begin{bmatrix} 0.5, & 0.6 \end{bmatrix}$, $SC = \begin{bmatrix} 0.2, & 0.2, & 0.4, & 0.3 \end{bmatrix}$

Procedamos a aplicar el algoritmo para encontrar la representación marginal uniforme. En cada iteración aparece, como resultado una matriz, que corresponde a la normalización de filas y columnas, también aparece la suma de filas y columnas.

IT	Matriz.	Suma fila y columnas.
1	0.1363636 0.1363636 0.1363636 0.09375 0.1136364 0.1136364 0.1136364 0.15625	SF= 0.5028409 0.4971591 SC= 0.25 0.25 0.25 0.25
2	0.1356589 0.1356589 0.1356589 0.09308511 0.1143411 0.1143411 0.1143411 0.15691489	SF= 0.5000619 0.4999381 SC= 0.25 0.25 0.25 0.25
3	0.1356436 0.1356436 0.1356436 0.09307065 0.1143564 0.1143564 0.1143564 0.15692935	SF= 0.5000013 0.499998 SC= 0.25 0.25 0.25 0.25
4	0.1356432 0.1356432 0.1356432 0.09307034 0.1143568 0.1143568 0.1143568 0.15692966	SF= 0.5 0.5 SC= 0.25 0.25 0.25 0.25

Con el nivel de precisión utilizado. En definitiva, la representación de marginales uniformes es

$$D_{RMU} = \begin{bmatrix} 0.1356432 & 0.1356432 & 0.1356432 & 0.09307034 \\ 0.1143568 & 0.1143568 & 0.1143568 & 0.15692966 \end{bmatrix}$$

Una interpretación de la RMU de una matriz de probabilidad puede encontrarse en Arnold y Gokhale (1994). Se define la función de información de Kullback-Liebler como a la pseudo distancia entre dos distribuciones (bidimensionales) con matrices de probabilidad conjunta P y Q (de la misma dimensión) como:

$$I(P,Q) = \sum_{i,j} p_{ij} \ln\left(\frac{p_{ij}}{q_{ij}}\right)$$

La representación marginal uniforme de P, es la distribución con marginales uniformes que minimiza $I(P,Q)$, es decir

$$I(P,RMU) = \min\left\{I(P,Q) : Q \text{ distribución con marginales uniformes}\right\}$$

Además, Arnold y Gokhale (1994) relaciona la URM de las matrices candidatas a condicionadas con el problema de compatibilidad.

Teorema 1.4.1 (Arnold y Gokhale, 1994). Sean A y B en las condiciones habituales, con $N_{n\times m} = N$. A y B son compatibles si y sólo las URM de A y B coinciden.

Ejemplo 2.

Sean las siguientes matrices

$$A = \begin{bmatrix} \frac{1}{2} & \frac{3}{8} \\ \frac{1}{2} & \frac{5}{8} \end{bmatrix} \quad B = \begin{bmatrix} \frac{1}{4} & \frac{3}{4} \\ \frac{1}{6} & \frac{5}{6} \end{bmatrix}$$

Tras cuatro iteraciones para A y cinco para B ambas matrices presentan la representación de marginales uniformes (con el nivel de precisión utilizado):

$$A_{RMU} = \begin{bmatrix} 0.2817542 & 0.2182458 \\ 0.2182458 & 0.2817542 \end{bmatrix} = B_{RMU}$$

Por lo que las matrices son compatibles.

La hipótesis de positividad de A y B es indispensable. Arnold y Gokhale (2002) ofrecen ejemplos en los que $N_{n\times m} \neq N$ y:

- UMR no existe, pero son compatibles.
- UMR no existe, pero son incompatibles.
- UMR existe y son iguales, pero son incompatibles.

5.- Uso de la fórmula de Besag.

Besag (1972, 1974) presenta una fórmula que relaciona los valores de la función de densidad o cuantía en dos puntos del rango de un vector aleatorio, a través de densidades o cuantías condicionadas. Nosotros nos ocuparemos del caso bidimensional discreto, que es pertinente para este trabajo.

Supondremos que las probabilidades que intervienen son no nulas.

Sean $(x_1, y_1),(x_2, y_2) \in \mathrm{R}_{(X,Y)}$ distintos. Relacionaremos en primer lugar la función de cuantía en dos puntos en los que se ha cambiado una sola coordenada.

<u>*Desplazamiento horizontal*</u>*:* los puntos varían en la primera componente, es decir los puntos son de la forma $(x_1, y_1),(x_2, y_1) \in \mathrm{R}_{(X,Y)}$

Puesto que:

$$P(X = x_1, Y = y_1) = P(X = x_1 \mid Y = y_1)P(Y = y_1)$$
$$P(X = x_2, Y = y_1) = P(X = x_2 \mid Y = y_1)P(Y = y_1)$$

Se tiene que:

$$\frac{P(X = x_1, Y = y_1)}{P(X = x_2, Y = y_1)} = \frac{P(X = x_1 \mid Y = y_1)}{P(X = x_2 \mid Y = y_1)}$$

<u>*Desplazamiento vertical*</u>*:* los puntos varían en la segunda componente, es decir los puntos son de la forma $(x_2, y_1),(x_2, y_2) \in \mathrm{R}_{(X,Y)}$

Puesto que:

$$P(X = x_2, Y = y_1) = P(Y = y_1 \mid X = x_2)P(X = x_2)$$
$$P(X = x_2, Y = y_2) = P(Y = y_2 \mid X = x_2)P(X = x_2)$$

Se tiene que:

$$\frac{P(X = x_2, Y = y_1)}{P(X = x_2, Y = y_2)} = \frac{P(Y = y_1 \mid X = x_2)}{P(Y = y_2 \mid X = x_2)}$$

<u>*Desplazamiento diagonal*</u>*:* los puntos varían en ambas componentes, es decir los puntos son de la forma $(x_1, y_1),(x_2, y_2) \in \mathrm{R}_{(X,Y)}$

Descomponiendo en desplazamientos horizontales y verticales:

$$\frac{P(X=x_1,Y=y_1)}{P(X=x_2,Y=y_2)}=\frac{P(X=x_1,Y=y_1)}{P(X=x_2,Y=y_1)}\frac{P(X=x_2,Y=y_1)}{P(X=x_2,Y=y_2)}$$

Se obtienen:

$$\frac{P(X=x_1,Y=y_1)}{P(X=x_2,Y=y_2)}=\frac{P(X=x_1\,|\,Y=y_1)}{P(X=x_2\,|\,Y=y_1)}\frac{P(Y=y_1\,|\,X=x_2)}{P(Y=y_2\,|\,X=x_2)}$$

Que es la versión bidimensional de la fórmula que encontramos en Besag (1972 y 1974).

Utilizando esta misma técnica, se puede relacionar cualquier par de puntos mediante un número finito de desplazamientos horizontales y verticales, seleccionando así, un camino en el que se cumpla la condición de positividad de las probabilidades implicadas. Indiquemos, además, que las posiciones de los caminos no tienen por qué ser adyacentes, siendo la condición básica, que las posiciones correspondan a desplazamientos horizontales y/o verticales.

Ejemplo 1:

Consideremos las matrices de probabilidad condicionadas de las variables aleatorias X e Y que toman valores X=1, 2, 3 e Y=1,2,3,4.

$$A=\begin{bmatrix} 1 & 0 & 0.6 & 0 \\ 0 & 0.4 & 0.4 & 0.75 \\ 0 & 0.6 & 0 & 0.25 \end{bmatrix} \qquad B=\begin{bmatrix} 0.4 & 0 & 0.6 & 0 \\ 0 & 0.2 & 0.2 & 0.6 \\ 0 & 0.6 & 0 & 0.4 \end{bmatrix}$$

Podemos relacionar las probabilidades de los puntos (1,1) y (2,3), para ello debemos encontrar un camino que los una de forma que entre cada punto y el anterior sólo exista un desplazamiento vertical u horizontal, por ejemplo

$$(1,1)\rightarrow(1,3)\rightarrow(2,3)$$

Gráficamente, si representamos en una matriz las posiciones de cada elemento no nulo, el camino seguido entre las posiciones en negrilla es

$$\begin{bmatrix} \mathbf{1,1} \rightarrow & & 1,3 \downarrow & \\ & 2,2 & \mathbf{2,3} & 2,4 \\ & 3,2 & & 3,4 \end{bmatrix}$$

Así la relación de probabilidades es:

$$\frac{P(X=1,Y=1)}{P(X=2,Y=3)} = \frac{P(X=1,Y=1)}{P(X=1,Y=3)} \frac{P(X=1,Y=3)}{P(X=2,Y=3)} =$$
$$= \frac{P(Y=1|X=1)}{P(Y=3|X=1)} \frac{P(X=1|Y=3)}{P(X=2|Y=3)} = \frac{0.4}{0.6}\frac{0.6}{0.4} = 1$$

De la misma manera, se podría relacionar probabilidades utilizando caminos más largos, por ejemplo para relacionar (1,1) con (3,2) debemos usar el camino:

$$(1,1) \rightarrow (1,3) \rightarrow (2,3) \rightarrow (2,2) \rightarrow (3,2)$$

$$\frac{P(X=1,Y=1)}{P(X=3,Y=3)} = \frac{P(X=1,Y=1)}{P(X=1,Y=3)} \frac{P(X=1,Y=3)}{P(X=2,Y=3)} \frac{P(X=2,Y=3)}{P(X=2,Y=2)} \frac{P(X=2,Y=2)}{P(X=3,Y=2)} =$$
$$= \frac{P(Y=1|X=1)}{P(Y=3|X=1)} \frac{P(X=1|Y=3)}{P(X=2|Y=3)} \frac{P(Y=3|X=2)}{P(Y=2|X=2)} \frac{P(X=2|Y=2)}{P(X=3|Y=2)} = \frac{0.4}{0.6}\frac{0.6}{0.4}\frac{0.2}{0.2}\frac{0.4}{0.6} = \frac{2}{3}$$

Podemos utilizar caminos alternativos, por ejemplo:

$$(1,1) \rightarrow (1,3) \rightarrow (2,3) \rightarrow (2,4) \rightarrow (3,4) \rightarrow (3,2)$$

$$\frac{P(X=1,Y=1)}{P(X=3,Y=2)}=$$
$$=\frac{P(X=1,Y=1)}{P(X=1,Y=3)}\frac{P(X=1,Y=3)}{P(X=2,Y=3)}\frac{P(X=2,Y=3)}{P(X=2,Y=4)}\frac{P(X=2,Y=4)}{P(X=3,Y=4)}\frac{P(X=3,Y=4)}{P(X=3,Y=2)}=$$
$$=\frac{P(Y=1|X=1)}{P(Y=3|X=1)}\frac{P(X=1|Y=3)}{P(X=2|Y=3)}\frac{P(Y=3|X=2)}{P(Y=4|X=2)}\frac{P(X=2|Y=4)}{P(X=3|Y=4)}\frac{P(Y=4|X=3)}{P(Y=2|X=3)}=$$
$$=\frac{0.4}{0.6}\frac{0.6}{0.4}\frac{0.2}{0.6}\frac{0.75}{0.25}\frac{0.4}{0.6}=\frac{2}{3}$$

Como cabía esperar, este valor coincide con el obtenido con anterioridad.

Gráficamente los dos caminos utilizados han sido:

$$\begin{bmatrix} \mathbf{1,1}\rightarrow & & 1,3\downarrow & \\ & 2,2\downarrow & \leftarrow 2,3 & 2,4 \\ & \mathbf{3,2} & & 3,4 \end{bmatrix} \begin{bmatrix} \mathbf{1,1}\rightarrow & & 1,3\downarrow & \\ & 2,2 & 2,3\rightarrow & \downarrow 2,4 \\ & \mathbf{3,2} & & \leftarrow 3,4 \end{bmatrix}$$

Nótese que este método permite reconstruir la matriz de probabilidades conjuntas. Para ello, basta con relacionar todos los valores (x_i, y_j) que tengan probabilidades no nulas con uno de ellos, el valor de esta relación es independiente del camino utilizado para obtenerla. Además, las probabilidades condicionadas nulas implican la conjunta nula.

Ejemplo 1 (continuación): continuando con el ejemplo anterior, observemos que

$$P(X=1,Y=2)=P(X=1,Y=4)=P(X=2,Y=1)=P(X=3,Y=1)=P(X=3,Y=3)=0$$

Relacionaremos todas las probabilidades conjuntas no nulas con la del punto (1,1). Note que algunas ya han sido calculadas. Únicamente indicaremos el camino y el resultado.

$(1,1) \to (1,3)$ $\qquad P(X=1,Y=3)=\frac{3}{2}P(X=1,Y=1)$

$(1,1) \to (1,3) \to (2,3) \to (2,2)$ $\qquad P(X=2,Y=2)=P(X=1,Y=1)$

$(1,1) \to (1,3) \to (2,3)$ $\qquad P(X=2,Y=3)=P(X=1,Y=1)$

$(1,1) \to (1,3) \to (2,3) \to (2,4)$ $\qquad P(X=2,Y=4)=3P(X=1,Y=1)$

$(1,1) \to (1,3) \to (2,3) \to (2,4) \to (3,4) \to (3,2)$

$$P(X=3,Y=3)=\frac{3}{2}P(X=1,Y=1)$$

$(1,1) \to (1,3) \to (2,3) \to (2,4) \to (3,4)$ $\quad P(X=3,Y=4)=P(X=1,Y=1)$

Puesto que $\sum_{x,y} P(X=x,Y=y)=1$ se tiene $P(X=1,Y=1)=0.1$ y de ahí

$$M=\begin{bmatrix} 0.1 & 0 & 0.15 & 0 \\ 0 & 0.1 & 0.1 & 0.3 \\ 0 & 0.15 & 0 & 0.1 \end{bmatrix}$$

Estas ideas se utilizan para estudiar y resolver el problema de la compatibilidad de dos matrices en Wang y Kuo (2010) y Kuo y Wang (2011).

Dadas matrices $A=(a_{ij})$ y $B=(b_{ij})$ de orden $n\times m$ que verifican $a_{ij} \geq 0$, $b_{ij} \geq 0$ $\forall i=1,\dots,n;$ $\forall j=1,\dots,m,$ $N_A=N_B=N$ y en las condiciones habituales, observan que

(i) Si todo $(i,j)\in N$ se puede unir mediante, al menos, un camino con una posición fija, y los cocientes son iguales a través de cualquier camino que una la posición fija y cualquier otra, entonces las matrices son compatibles.

(ii) Si todo $(i,j)\in N$ se puede unir mediante un camino con una posición fija, y al menos dos cocientes son distintos, entonces las matrices son incompatibles.

A partir de ahí, el cálculo de la distribución conjunta siguen los mismos pasos que vimos en el ejemplo anterior.

Ejemplo 2:

Consideremos las siguientes matrices

$$A=\begin{bmatrix}1 & 0.4 & 0.5\\ 0 & 0.6 & 0.5\end{bmatrix} \qquad B=\begin{bmatrix}0.2 & 0.4 & 0.4\\ 0 & 0.6 & 0.4\end{bmatrix}$$

A partir de la posición (1,1) es posible construir caminos que la conecte con el resto de las posiciones de la matriz. Comprobaremos la compatibilidad de las matrices consideradas y calcularemos las distribuciones conjuntas y marginales de las variables aleatorias asociadas al problema. Supondremos que X toma los valores 1,2 e Y los valores 1,2,3.

(1,1)→(1,2) [único]

$$\frac{P(X=1,Y=1)}{P(X=1,Y=2)}=\frac{P(Y=1\,|\,X=1)}{P(Y=2\,|\,X=1)}=\frac{0.2}{0.4}=\frac{1}{2}$$

(1,1)→(1,2)→(1,3) [1 de 2]

$$\frac{P(X=1,Y=1)}{P(X=1,Y=3)}=\frac{0.2}{0.4}\frac{0.4}{0.4}=\frac{1}{2}$$

(1,1)→(1,2)→(2,2)→(2,3)→(1,3) [2 de 2]

$$\frac{P(X=1,Y=1)}{P(X=1,Y=3)}=\frac{0.2}{0.4}\frac{0.4}{0.6}\frac{0.6}{0.4}\frac{0.5}{0.5}=\frac{1}{2}$$

(1,1)→(1,2)→(2,2)[1 de 2]

$$\frac{P(X=1,Y=1)}{P(X=2,Y=2)}=\frac{0.2}{0.4}\frac{0.4}{0.6}=\frac{1}{3}$$

(1,1)→(1,2)→(1,3)→(2,3)→(2,2) [2 de 2]

$$\frac{P(X=1,Y=1)}{P(X=2,Y=2)}=\frac{0.2}{0.4}\frac{0.4}{0.4}\frac{0.5}{0.5}\frac{0.4}{0.6}=\frac{1}{3}$$

(1,1)→(1,2)→(1,3)→(2,3) [1 de 2]

$$\frac{P(X=1,Y=1)}{P(X=2,Y=3)}=\frac{0.2}{0.4}\frac{0.4}{0.4}\frac{0.5}{0.5}=\frac{1}{2}$$

(1,1)→(1,2)→(2,2)→(2,3)[2 de 2]

$$\frac{P(X=1,Y=1)}{P(X=2,Y=3)}=\frac{0.2}{0.4}\frac{0.4}{0.6}\frac{0.6}{0.4}=\frac{1}{2}$$

De lo anterior se deduce que las matrices son compatibles. Para el cálculo de las probabilidades conjuntas observemos que,

$P(X=2,Y=1)=0$, $P(X=1,Y=2)=2P(X=1,Y=1)$,

$P(X=1,Y=3)=2P(X=1,Y=1)$, $P(X=2,Y=2)=3P(X=1,Y=1)$,

$P(X=2,Y=3)=2P(X=1,Y=1)$

Puesto que $\sum_{i,j} P(X=i,Y=j)=1$ se tiene $\mathrm{P}(X=1,Y=1)=0.1$ y la matriz de probabilidades conjuntas es:

$$M=\begin{bmatrix}0.1 & 0.2 & 0.2\\ 0 & 0.3 & 0.2\end{bmatrix}$$

las marginales son, por tanto,

$$\vec{\alpha}=\begin{bmatrix}0.5\\ 0.5\end{bmatrix} \qquad \vec{\beta}=\begin{bmatrix}0.1\\ 0.5\\ 0.4\end{bmatrix}$$

Ejemplo 4: Consideremos las siguientes matrices

$$A=\begin{bmatrix}1 & 0.4 & 0.5\\ 0 & 0.6 & 0.5\end{bmatrix} \qquad B=\begin{bmatrix}0.2 & 0.7 & 0.1\\ 0 & 0.9 & 0.1\end{bmatrix}$$

Y calculemos el valor del cociente $\dfrac{P(X=1,Y=1)}{P(X=2,Y=2)}$ a través de los caminos:

(1,1)→(1,2)→(2,2)

$$\frac{P(X=1,Y=1)}{P(X=2,Y=2)}=\frac{0.2}{0.7}\frac{0.4}{0.6}=\frac{4}{21}$$

(1,1)→(1,2)→(1,3)→(2,3)→(2,2)

$$\frac{P(X=1,Y=1)}{P(X=2,Y=2)}=\frac{0.2}{0.7}\frac{0.7}{0.1}\frac{0.5}{0.5}\frac{0.1}{0.9}=\frac{2}{9}$$

Puesto que los valores obtenidos son distintos, las matrices son incompatibles.

En relación a la condición "Si todo $(i,j)\in N$ se puede unir mediante un camino con una posición fija" señalemos que en ocasiones existe un elemento que se puede unir con el resto de elementos de *N* mediante un único camino, en ese caso las matrices son compatibles. Y por otra parte, en ocasiones tal elemento no existe. Veamos ejemplos de ambas situaciones.

Ejemplo 3:

$$A = \begin{bmatrix} 1 & 0.4 & 0 \\ 0 & 0.6 & 1 \end{bmatrix} \qquad B = \begin{bmatrix} 0.2 & 0.8 & 0 \\ 0 & 0.6 & 0.4 \end{bmatrix}$$

Note que la situación de los ceros, hace que para cada par de elementos de N exista un único camino que los une.

Ejemplo 4:

$$A = \begin{bmatrix} 0.5 & 0.125 & 0 & 0 \\ 0.5 & 0.875 & 0 & 0 \\ 0 & 0 & 0.5 & 0.125 \\ 0 & 0 & 0.5 & 0.125 \end{bmatrix} \qquad B = \begin{bmatrix} 0.5 & 0.5 & 0 & 0 \\ 0.125 & 0.875 & 0 & 0 \\ 0 & 0 & 0.5 & 0.5 \\ 0 & 0 & 0.125 & 0.875 \end{bmatrix}$$

En este caso no existe ninguna posición que pueda unirse con todas las demás. Así, por ejemplo, la posición (1, 1) no puede unirse con las posiciones del segundo bloque de la diagonal. Sin embargo son compatibles. Volveremos sobre esta cuestión en el próximo capítulo, obteniendo todas las soluciones.

Un algoritmo más pragmático que el que hemos discutido, puede encontrarse en Kuo y Wang (2011), al que denominan *onepath.* Básicamente consiste en utilizar un único camino para cada elemento de N y con posterioridad, comprobar las soluciones:

(i) Unir cada $(i, j) \in N$ mediante un camino con una posición fija, y calcular los cocientes de probabilidad.

(ii) Calcular la matriz de probabilidades conjuntas a partir de los cocientes obtenidos en (i).

(iii) Comprobar que las distribuciones condicionadas de (ii) coinciden con las iniciales.

Ejemplo 2 (continuación)

$$A = \begin{bmatrix} 1 & 0.4 & 0.5 \\ 0 & 0.6 & 0.5 \end{bmatrix} \qquad B = \begin{bmatrix} 0.2 & 0.4 & 0.4 \\ 0 & 0.6 & 0.4 \end{bmatrix}$$

A partir de la posición (1,1) es posible construir caminos que lo una con el resto de las posiciones de la matriz. Supondremos que *X* toma valores 1, 2 e *Y* 1, 2, 3.

(1,1)→(1,2)

$$\frac{P(X=1,Y=1)}{P(X=1,Y=2)} = \frac{P(Y=1 \mid X=1)}{P(Y=2 \mid X=1)} = \frac{0.2}{0.4} = \frac{1}{2}$$

(1,1)→(1,2)→(1,3)

$$\frac{P(X=1,Y=1)}{P(X=1,Y=3)} = \frac{0.2}{0.4}\frac{0.4}{0.4} = \frac{1}{2}$$

(1,1)→(1,2)→(2,2)

$$\frac{P(X=1,Y=1)}{P(X=2,Y=2)} = \frac{0.2}{0.4}\frac{0.4}{0.6} = \frac{1}{3}$$

(1,1)→(1,2)→(1,3)→(2,3)

$$\frac{P(X=1,Y=1)}{P(X=2,Y=3)} = \frac{0.2}{0.4}\frac{0.4}{0.4}\frac{0.5}{0.5} = \frac{1}{2}$$

Con los datos obtenidos se pueden calcular las probabilidades conjuntas, tal y como se hizo en el ejemplo 2, resultando

$$M = \begin{bmatrix} 0.1 & 0.2 & 0.2 \\ 0 & 0.3 & 0.2 \end{bmatrix}$$

El cálculo de las matrices de probabilidades condicionadas nos conduce a las matrices iniciales:

$$A = \begin{bmatrix} 1 & 0.4 & 0.5 \\ 0 & 0.6 & 0.5 \end{bmatrix} \qquad B = \begin{bmatrix} 0.2 & 0.4 & 0.4 \\ 0 & 0.6 & 0.4 \end{bmatrix}$$

6.- Extensión de rango unidad.

El objetivo de esta sección es mostrar la generalización del teorema 1.3.6 al caso en el que $N_A = N_B = N \neq N_{n\times m}$. Salvo mención expresa de lo contrario, supondremos, a lo largo de toda la sección que $N_A = N_B = N$ sin asumir la condición de positividad de las matrices *A* y *B*.

Observemos que si $N_A = N_B = N \neq N_{n\times m}$, la matriz de cocientes *C* no está completamente determinada por las matrices candidatas a distribuciones condicionadas *A* y *B*. Así si, por ejemplo, en

$$A = \begin{bmatrix} 0.25 & 0.25 & 0 \\ 0.5 & 0.25 & 0.5 \\ 0.25 & 0.5 & 0.5 \end{bmatrix} \qquad B = \begin{bmatrix} 0.5 & 0.5 & 0 \\ 0.5 & 0.25 & 0.25 \\ 0.25 & 0.5 & 0.25 \end{bmatrix}$$

el elemento c_{13} de C está indeterminado, al ser cociente de dos elementos nulos. Pérez-Villalta (2000) observa que, bajo la hipótesis de compatibilidad, ese elemento indeterminado de C está definido, pues es cociente de probabilidades marginales no nulas (pues, en otro caso el elemento se elimina), y puede ser determinado, bajo ciertas condiciones sobre C, a partir de A y B. En efecto, bajo hipótesis de compatibilidad, existen un vector aleatorio *(X, Y)* cuyas condicionadas son las matrices dadas y por tanto:

$$c_{ij} = \frac{a_{ij}}{b_{ij}} = \frac{P(X = x_i \mid Y = y_j)}{P(Y = y_j \mid X = x_i)} = \frac{P(X = x_i)}{P(Y = y_j)} \quad \forall (i, j) \in N$$

Y, también,

$$c_{ij} = \frac{P(X = x_i)}{P(Y = y_j)} \quad \forall (i, j) \in N_{n\times m}$$

Se trata, por tanto, de encontrar métodos para calcular C.

Teorema 1.6.1 (Pérez-Villalta, 2000). Sean A y B dos matrices en las condiciones habituales con $N_A = N_B = N$. Si A y B son compatibles entonces, la matriz C puede completarse de forma que tenga rango 1.

Demostración: Si las matrices dadas determinan la distribución conjunta de un vector (X,Y) discreto y finito, para cada $(i,j) \notin N$ nos basta tomar

$$c_{ij} = \frac{P\left[X = x_i\right]}{P\left[Y = y_j\right]} = \frac{\alpha_i}{\beta_j}$$

Obsérvese que estos c_{ij} están definidos pues P[Y=y_j] ≠ 0 pues, en otro caso, y_j se habría eliminado del problema.

□

Nótese que si C puede completarse de forma que tenga rango unidad podremos aplicarle el teorema 1.3.6 y tenemos el siguiente

Teorema 1.6.2 (Pérez-Villalta, 2000). En las condiciones anteriores, si C puede completarse de forma que tenga rango unidad entonces, A y B son compatibles.

Demostración: Basta aplicar la técnica del teorema 1.3.6 para obtener distribuciones marginales.

□

Ejemplo1: Considérese

$$A = 10^{-1}\begin{bmatrix} 1 & 2 & 1 & 1 \\ 0 & 4 & 2 & 2 \\ 5 & 0 & 2 & 4 \\ 4 & 4 & 5 & 3 \end{bmatrix} \qquad \text{y} \qquad B = 10^{-1}\begin{bmatrix} 2 & 1 & 6 & 1 \\ 0 & 4 & 2 & 4 \\ 4 & 0 & 3 & 3 \\ 1 & 2 & 3 & 4 \end{bmatrix}$$

Entonces, la matriz C viene dada (parcialmente) por

$$C = \begin{bmatrix} \frac{1}{2} & 2 & \frac{1}{6} & 1 \\ * & 1 & 1 & \frac{1}{2} \\ \frac{5}{4} & * & \frac{2}{3} & \frac{4}{3} \\ 4 & 2 & \frac{5}{3} & \frac{3}{4} \end{bmatrix}$$

A y *B* son incompatibles pues, por ejemplo

$$\begin{vmatrix} 2 & \frac{1}{6} \\ 1 & 1 \end{vmatrix} \neq 0$$

Ejemplo 2.

$$A = \begin{bmatrix} \frac{1}{2} & 0 & \frac{3}{8} & \frac{3}{4} \\ \frac{1}{4} & \frac{3}{4} & 0 & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{5}{8} & \frac{1}{4} \end{bmatrix} \qquad B = \begin{bmatrix} \frac{1}{4} & 0 & \frac{3}{8} & \frac{3}{8} \\ \frac{1}{4} & \frac{3}{4} & 0 & 0 \\ \frac{1}{8} & \frac{1}{8} & \frac{5}{8} & \frac{1}{8} \end{bmatrix}$$

Matriz de cocientes (incompleta) es:

$$C = \begin{bmatrix} 2 & * & 1 & 2 \\ 1 & 1 & * & * \\ 2 & 2 & 1 & 2 \end{bmatrix}$$

Calculemos una extensión positiva de rango uno de *C*, para ello calcularemos los elementos que faltan en *C* obligando que ciertos determinantes sean nulos, perfilando la posibilidad de que la matriz resultante sea uno.

$$\begin{vmatrix} 2 & a \\ 1 & 1 \end{vmatrix} = 0 \to a = 2 \qquad C_1 = \begin{bmatrix} 2 & 2 & 1 & 2 \\ 1 & 1 & * & * \\ 2 & 2 & 1 & 2 \end{bmatrix}$$

$$\begin{vmatrix} 2 & 1 \\ 1 & a \end{vmatrix} = 0 \to a = \frac{1}{2} \qquad C_2 = \begin{bmatrix} 2 & 2 & 1 & 2 \\ 1 & 1 & \frac{1}{2} & * \\ 2 & 2 & 1 & 2 \end{bmatrix}$$

$$\begin{vmatrix} 1 & 2 \\ \frac{1}{2} & a \end{vmatrix} = 0 \to a = 1 \qquad C_3 = \begin{bmatrix} 2 & 2 & 1 & 2 \\ 1 & 1 & \frac{1}{2} & 1 \\ 2 & 2 & 1 & 2 \end{bmatrix}$$

Note que es posible utilizar valores obtenidos en pasos anteriores. Una vez que se ha obtenido los valores faltantes, la matriz resultante debe ser de rango uno, si no lo es, las matrices *A* y *B* son incompatibles. En nuestro caso, puesto que $rang(C_3) = 1$se tiene $C_3 = \bar{C}$.

Cálculo de marginales: $(1 \quad 1 \quad 1)\begin{bmatrix} 2 & 2 & 1 & 2 \\ 1 & 1 & \frac{1}{2} & 1 \\ 2 & 2 & 1 & 2 \end{bmatrix} = (5 \quad 5 \quad \frac{5}{2} \quad 5)$

Invirtiendo elemento a elemento $\vec{\beta} = \begin{bmatrix} \frac{1}{5} \\ \frac{1}{5} \\ \frac{2}{5} \\ \frac{1}{5} \end{bmatrix}$. La otra marginal se puede obtener, por ejemplo, $A\vec{\beta} = \vec{\alpha}$ y, en este caso se obtiene $\vec{\alpha} = \begin{bmatrix} \frac{2}{5} \\ \frac{1}{5} \\ \frac{2}{5} \end{bmatrix}$.

Ejemplo 3. Consideramos las matrices:

$$A = \begin{bmatrix} 0 & \frac{2}{3} & 0 & \frac{1}{3} \\ \frac{1}{2} & 0 & \frac{1}{4} & 0 \\ 0 & \frac{1}{3} & 0 & \frac{2}{3} \\ \frac{1}{2} & 0 & \frac{3}{4} & 0 \end{bmatrix} \qquad B = \begin{bmatrix} 0 & \frac{4}{5} & 0 & \frac{1}{5} \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{4} & 0 & \frac{3}{4} & 0 \end{bmatrix}$$

La matriz (parcial) de cocientes es

$$C=\begin{bmatrix} * & \frac{5}{6} & * & \frac{5}{3} \\ 1 & * & \frac{1}{2} & * \\ * & \frac{2}{3} & * & \frac{4}{3} \\ 2 & * & 3 & * \end{bmatrix}$$

Si hacemos $c_{11}= \lambda \neq 0$ obtenemos

$$C=\begin{bmatrix} \lambda & \frac{5}{6} & \frac{\lambda}{2} & \frac{5}{3} \\ 1 & \frac{5}{6\lambda} & \frac{1}{2} & \frac{5}{3\lambda} \\ \frac{4}{5}\lambda & \frac{2}{3} & \frac{2}{5}\lambda & \frac{4}{3} \\ 2 & \frac{5}{3\lambda} & 1 & \frac{10}{3\lambda} \end{bmatrix}$$

sin más que imponer que tenga rango 1. Para cada valor de λ se obtienen distintas C y por tanto esta matriz no se puede completar de manera única de forma que tenga rango unidad. Además, para cada λ se tendrán distintos vectores $\vec{u}$ y $\vec{v}$ (donde $c_{ij}=u_i v_j \ \forall (i,j) \in N$). En este caso,

$$\vec{u} \propto \left(\frac{5}{2}+\frac{3}{2}\lambda, \frac{3}{2}+\frac{5}{2\lambda}, 2+\frac{6}{5}\lambda, 3+\frac{5}{\lambda} \right) \quad \vec{v} \propto \left(3+\frac{9}{5}\lambda, \frac{3}{2}+\frac{1}{\lambda}\frac{5}{2}, \frac{3}{2}+\frac{9}{10}\lambda, 3+\frac{5}{\lambda} \right)$$

El caso de no unicidad será estudiado en el próximo capítulo.

La idea de completar la matriz de cocientes C puede ser formalizada como sigue (Arnold, Castillo y Sarabia, 2004)

Definición: Una matriz $\bar{C}$ se dice que es una extensión positiva de la matriz C (incompleta) si todos los elementos de $\bar{C}$ son positivos y los elementos $\bar{C}$ coinciden con los de C sobre N, es decir, $\bar{c}_{ij} = c_{ij} \quad \forall (i, j) \in N$

De esta forma, los teoremas 1.6.1 y 1.6.2 adoptan la forma:

Teorema 1.6.3 (Pérez-Villalta, 2000; reformulado). Sean A y B dos matrices en las condiciones habituales con $N_A = N_B = N$. Entonces, A y B son compatibles sí y sólo si, la matriz C admite una extensión positiva de rango 1.

Nótese que si A y B son compatibles, una extensión positiva de rango unidad es $\bar{c}_{ij} = \dfrac{P(X = x_i)}{P(Y = y_j)} \quad \forall (i, j) \in N_{n\times m}$.

A partir del resultado anterior, Arnold, Castillo y Sarabia (2004), proporcionan los dos resultados siguientes, que son consecuencias inmediatas del anterior.

Teorema 1.6.4 (Arnold, Castillo y Sarabia, 2004). Sean A y B dos matrices en las condiciones habituales. A y B son compatibles sí y sólo si $N_A = N_B = N$ y existe una matriz positiva $\bar{C}$ tal que

$$\begin{vmatrix} \bar{c}_{ij} & \bar{c}_{im} \\ \bar{c}_{nj} & \bar{c}_{nm} \end{vmatrix} = 0 \;\; i = 1,2,...,n-1, \;\; j = 1,2,..., j-1 \;\; \text{y} \;\; c_{ij} = \frac{a_{ij}}{b_{ij}} \;\; \forall (i, j) \in N$$

Teorema 1.6.5 (Arnold, Castillo y Sarabia, 2004): Sean A y B dos matrices en las condiciones habituales tales que $N_A = N_B = N$ y existe una matriz positiva $\bar{C}$ tal que $c_{ij}=\dfrac{a_{ij}}{b_{ij}}\ \forall(i,j)\in N$ y $\bar{c}_{ij}=\dfrac{\bar{c}_{i\bullet}\bar{c}_{\bullet j}}{\bar{c}_{\bullet\bullet}}\ \forall(i,j)\in N_{n\times m}$ entonces A y B son compatibles.

Donde $\bar{c}_{i\bullet}=\sum_{j=1}^{m}\bar{c}_{ij},\ \bar{c}_{\bullet j}=\sum_{i=1}^{n}\bar{c}_{ij}\ y\ \bar{c}_{\bullet\bullet}=\sum_{(i,j)\in N_{n\times m}}\bar{c}_{ij}\ \ \forall(i,j)\in N_{n\times m}$

Si la extensión positiva $\bar{C}$ de C tiene rango unidad, sus filas y columnas son proporcionales y es posible considerar las matrices de las constantes de proporcionalidad por filas y por columnas

Definimos la matriz de constantes de proporcionalidad por filas F como la matriz $n\times n$ dada por $f_{il}=\dfrac{\bar{c}_{lj}}{\bar{c}_{ij}}$. Nótese que está bien definida si las matrices A y B son compatibles

$$f_{il}=\frac{\bar{c}_{lj}}{\bar{c}_{ij}}=\frac{P(X=x_l)\Big/P(Y=y_j)}{P(X=x_i)\Big/P(Y=y_j)}=\frac{P(X=x_l)}{P(X=x_i)},\quad i,l=1,2,\dots,n$$

La matriz es, por tanto

$$F=\begin{bmatrix} \frac{P(X=x_1)}{P(X=x_1)} & \frac{P(X=x_2)}{P(X=x_1)} & \cdots & \frac{P(X=x_n)}{P(X=x_1)} \\ \frac{P(X=x_1)}{P(X=x_2)} & \frac{P(X=x_2)}{P(X=x_2)} & \cdots & \frac{P(X=x_n)}{P(X=x_2)} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{P(X=x_1)}{P(X=x_n)} & \frac{P(X=x_2)}{P(X=x_n)} & \cdots & \frac{P(X=x_n)}{P(X=x_n)} \end{bmatrix}$$

El siguiente resultado resume las propiedades de F y, por tanto, las de las constantes de proporcionalidad por filas:

Teorema 1.6.7. Sean A y B dos matrices en las condiciones habituales y compatibles (por lo tanto existe $\bar{C}$) y sea F la matriz de constantes de proporcionalidad por filas, entonces:

a.- F está bien definida.

b.- $f_{ii}=1,\ f_{il}=f_{li}^{-1}\ \ i,l=1,2,...,n$

c.- $rang(F)=1$

d.- $F\vec{1}_n=\begin{bmatrix} \frac{1}{P(X=x_1)} \\ \frac{1}{P(X=x_2)} \\ \vdots \\ \frac{1}{P(X=x_n)} \end{bmatrix}$

e.- $$F = \begin{bmatrix} \dfrac{1}{P(X = x_1)} \\ \dfrac{1}{P(X = x_2)} \\ \vdots \\ \dfrac{1}{P(X = x_n)} \end{bmatrix} \begin{bmatrix} P(X = x_1) & P(X = x_2) & \ldots & P(X = x_n) \end{bmatrix}$$

(descomposición en valores singulares).

La demostración es evidente.

Ejemplo 2 (continuación)

$$A = \begin{bmatrix} \frac{1}{2} & 0 & \frac{3}{8} & \frac{3}{4} \\ \frac{1}{4} & \frac{3}{4} & 0 & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{5}{8} & \frac{1}{4} \end{bmatrix} \qquad B = \begin{bmatrix} \frac{1}{4} & 0 & \frac{3}{4} & \frac{3}{8} \\ \frac{1}{4} & \frac{3}{4} & 0 & 0 \\ \frac{1}{8} & \frac{1}{8} & \frac{5}{8} & \frac{1}{8} \end{bmatrix}$$

Matriz de cocientes (incompleta) es:

$$C = \begin{bmatrix} 2 & * & 1 & 2 \\ 1 & 1 & * & * \\ 2 & 2 & 1 & 2 \end{bmatrix}$$

Una extensión positiva de rango uno de C,

$$\bar{C} = \begin{bmatrix} 2 & 2 & 1 & 2 \\ 1 & 1 & \frac{1}{2} & 1 \\ 2 & 2 & 1 & 2 \end{bmatrix}$$

La matriz de constantes de proporcionalidad por filas bien dada por:

$f_{il} = \dfrac{\bar{c}_{lj}}{\bar{c}_{ij}}$, por ejemplo, $f_{12} = \dfrac{\bar{c}_{2j}}{\bar{c}_{1j}} = \dfrac{1}{2}$ por tanto $F = \begin{bmatrix} 1 & \frac{1}{2} & 1 \\ 2 & 1 & 2 \\ 1 & \frac{1}{2} & 1 \end{bmatrix}$

Podemos utilizar el apartado del teorema anterior para calcular una de las marginales.

$$F\vec{1}_3 = \begin{bmatrix} 1 & \frac{1}{2} & 1 \\ 2 & 1 & 2 \\ 1 & \frac{1}{2} & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{5}{2} \\ 5 \\ \frac{5}{2} \end{bmatrix}$$

Invirtiendo elemento a elemento $\vec{\alpha} = \begin{bmatrix} \frac{2}{5} \\ \frac{1}{5} \\ \frac{2}{5} \end{bmatrix}$ la otra marginal puede ser calculada

mediante $B^t\vec{\alpha} = \vec{\beta}$.

De la misma forma, definimos la matriz de constantes de proporcionalidad por filas G como la matriz $m\times m$ dada por $g_{jk} = \frac{\bar{c}_{ik}}{\bar{c}_{ij}}$. Note que está bien definida si las matrices A y B son compatibles

$$g_{jk} = \frac{\bar{c}_{ik}}{\bar{c}_{ij}} = \frac{P(Y=y_i)\big/P(Y=y_k)}{P(Y=y_i)\big/P(Y=y_j)} = \frac{P(Y=y_j)}{P(Y=y_k)},\ \ j,k = 1,2,\ldots,m$$

La matriz es, por tanto

$$G = \begin{bmatrix} \frac{P(Y=y_1)}{P(Y=y_1)} & \frac{P(Y=y_1)}{P(Y=y_2)} & \cdots & \frac{P(Y=y_1)}{P(Y=y_m)} \\ \frac{P(Y=y_2)}{P(Y=y_1)} & \frac{P(Y=y_2)}{P(Y=y_2)} & \cdots & \frac{P(Y=y_2)}{P(Y=y_2)} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{P(Y=y_m)}{P(Y=y_1)} & \frac{P(Y=y_m)}{P(Y=y_2)} & \cdots & \frac{P(Y=y_m)}{P(Y=y_m)} \end{bmatrix}$$

El siguiente resultado resume las propiedades de G y, por tanto, las de las constantes de proporcionalidad por columnas:

Teorema 1.6.8. Sean A y B dos matrices en las condiciones habituales y compatibles (por lo tanto existe $\bar{C}$) y sea G la matriz de constantes de proporcionalidad por columnas, entonces:

a.- G está bien definida.

b.- $g_{jj}=1,\ g_{kj}=g_{jk}^{-1}\quad j,k=1,2,...,m$

c.- $rang(G)=1$

d.- $.\vec{1}_m^t G=\begin{bmatrix}\dfrac{1}{P(Y=y_1)} & \dfrac{1}{P(Y=y_2)} & \cdots & \dfrac{1}{P(Y=y_m)}\end{bmatrix}.$

e.- $F=\begin{bmatrix}P(Y=y_1)\\ P(Y=y_2)\\ \vdots\\ P(Y=y_m)\end{bmatrix}\begin{bmatrix}\dfrac{1}{P(Y=y_1)} & \dfrac{1}{P(Y=y_2)} & \cdots & \dfrac{1}{P(Y=y_m)}\end{bmatrix}$

(descomposición en valores singulares).

La demostración es evidente.

Ejemplo 2 (continuación)

$$A=\begin{bmatrix}\frac{1}{2} & 0 & \frac{3}{8} & \frac{3}{4}\\ \frac{1}{4} & \frac{3}{4} & 0 & 0\\ \frac{1}{4} & \frac{1}{4} & \frac{5}{8} & \frac{1}{4}\end{bmatrix}\qquad B=\begin{bmatrix}\frac{1}{4} & 0 & \frac{3}{8} & \frac{3}{8}\\ \frac{1}{4} & \frac{3}{4} & 0 & 0\\ \frac{1}{8} & \frac{1}{8} & \frac{5}{8} & \frac{1}{8}\end{bmatrix}$$

Matriz de cocientes (incompleta) es:

$$C=\begin{bmatrix}2 & * & 1 & 2\\ 1 & 1 & * & *\\ 2 & 2 & 1 & 2\end{bmatrix}$$

Una extensión positiva de rango uno de *C*,

$$\bar{C} = \begin{bmatrix} 2 & 2 & 1 & 2 \\ 1 & 1 & \frac{1}{2} & 1 \\ 2 & 2 & 1 & 2 \end{bmatrix}$$

La matriz *G* de constantes de proporcionalidad por columnas puede ser calculada mediante $g_{jk} = \frac{\bar{c}_{ik}}{\bar{c}_{ij}}$, por ejemplo, $g_{12} = \frac{\bar{c}_{i2}}{\bar{c}_{i1}} = \frac{2}{2} = 1$, por tanto $G = \begin{bmatrix} 1 & 1 & \frac{1}{2} & 1 \\ 1 & 1 & \frac{1}{2} & 1 \\ 2 & 2 & 1 & 2 \\ 1 & 1 & \frac{1}{2} & 1 \end{bmatrix}$

Y utilizando el apartado d del teorema anterior, podemos calcular la marginal

$$\vec{1}_4^t G = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & \frac{1}{2} & 1 \\ 1 & 1 & \frac{1}{2} & 1 \\ 2 & 2 & 1 & 2 \\ 1 & 1 & \frac{1}{2} & 1 \end{bmatrix} = \begin{bmatrix} 5 & 5 & \frac{5}{2} & 5 \end{bmatrix}$$

Los recíprocos de esos elementos forman el vector la marginal $\vec{\beta} = \begin{bmatrix} \frac{1}{5} \\ \frac{1}{5} \\ \frac{2}{5} \\ \frac{1}{5} \end{bmatrix}$.

Los siguientes resultados (Song, Li, Chen, Jiang y Kuo, 2010), serán de gran importancia en el próximo capítulo, representan un avance en el método que tratamos.

Definición. Con la notación habitual, una matriz de cocientes *C* se dice que es reducible si después de intercambiar algunas filas y algunas columnas se puede escribir como la siguiente matriz por bloques:

$$\begin{bmatrix} T_1 & * \\ * & T_2 \end{bmatrix}$$

Donde las entradas fuera de la diagonal por bloques son todos elementos desconocidos (*) y las entradas de los bloques diagonales no son todas desconocidas. Una matriz de cocientes es irreducible si no es reducible.

Lema 1.6.1 (Song, Li, Chen, Jiang y Kuo, 2010). Para cualquier matriz de cocientes *C*, por intercambio de filas y/o columnas se puede reescribir como una matriz irreducible con $M \geq 1$ bloques en la diagonal irreducibles y bloque fuera de la diagonal formados por elementos desconocidos.

$$T(C) = \begin{bmatrix} T_1 & * & \cdots & * \\ * & T_2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \cdots & T_M \end{bmatrix}$$

Notación: $T(C) = Diag(T_1, T_2, ..., T_M)$, $M \geq 1$ a esta descomposición la llamaremos descomposición en bloques irreducibles (DBI)

Esta descomposición no es única

Ejemplo 4:

Consideremos las matrices

$$A = \begin{bmatrix} 0.6 & 0 & 0 & 0.8 & 0 & 0 & 0 \\ 0 & 0 & 0.1 & 0 & 0 & 0.9 & 0 \\ 0 & 1 & 0 & 0 & 0.3 & 0 & 0.4 \\ 0 & 0 & 0 & 0 & 0.7 & 0 & 0.6 \\ 0 & 0 & 0.9 & 0 & 0 & 0.1 & 0 \\ 0.4 & 0 & 0 & 0.2 & 0 & 0 & 0 \end{bmatrix}$$

$$B = \begin{bmatrix} 0.6 & 0 & 0 & 0.4 & 0 & 0 & 0 \\ 0 & 0 & 0.1 & 0 & 0 & 0.9 & 0 \\ 0 & 0.5 & 0 & 0 & 0.3 & 0 & 0.2 \\ 0 & 0 & 0 & 0 & 0.7 & 0 & 0.3 \\ 0 & 0 & 0.9 & 0 & 0 & 0.1 & 0 \\ 0.8 & 0 & 0 & 0.2 & 0 & 0 & 0 \end{bmatrix}$$

La matriz de cocientes es

$$C = \begin{bmatrix} 1 & * & * & 2 & * & * & * \\ * & * & 1 & * & * & 1 & * \\ * & 2 & * & * & 1 & * & 2 \\ * & * & * & * & 1 & * & 2 \\ * & * & 1 & * & * & 1 & * \\ 0.5 & * & * & 1 & * & * & * \end{bmatrix}$$

Intercambiando las filas 2 y 6

$$C_1 = \begin{bmatrix} 1 & * & * & 2 & * & * & * \\ 0.5 & * & * & 1 & * & * & * \\ * & 2 & * & * & 1 & * & 2 \\ * & * & * & * & 1 & * & 2 \\ * & * & 1 & * & * & 1 & * \\ * & * & 1 & * & * & 1 & * \end{bmatrix}$$

Intercambiamos las columnas 2 y 4

$$C_2 = \begin{bmatrix} 1 & 2 & * & * & * & * & * \\ 0.5 & 1 & * & * & * & * & * \\ * & * & * & 2 & 1 & * & 2 \\ * & * & * & * & 1 & * & 2 \\ * & * & 1 & * & * & 1 & * \\ * & * & 1 & * & * & 1 & * \end{bmatrix}$$

Y finalmente, intercambiamos las columnas 3 y 7,

$$C_3 = \begin{bmatrix} 1 & 2 & * & * & * & * & * \\ 0.5 & 1 & * & * & * & * & * \\ * & * & 2 & 2 & 1 & * & * \\ * & * & 2 & * & 1 & * & * \\ * & * & * & * & * & 1 & 1 \\ * & * & * & * & * & 1 & 1 \end{bmatrix}$$

En vista de la forma de C_3 concluimos que $C_3 = T(C)$ siendo los bloques diagonales

$$T_1 = \begin{bmatrix} 1 & 2 \\ 0.5 & 1 \end{bmatrix} \; T_2 = \begin{bmatrix} 2 & 2 & 1 \\ 2 & * & 1 \end{bmatrix} \; T_3 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

Por otra parte es claro que se puede hacer otras transformaciones para obtener otra matriz que siendo DBI de C no es la que hemos obtenido, por ejemplo se puede obtener

$$\begin{bmatrix} 1 & 2 & * & * & * & * & * \\ 0.5 & 1 & * & * & * & * & * \\ * & * & 1 & 1 & * & * & * \\ * & * & 1 & 1 & * & * & * \\ * & * & * & * & 2 & 2 & 1 \\ * & * & * & * & 2 & * & 1 \end{bmatrix} \quad \text{o bien} \quad \begin{bmatrix} 0.5 & 1 & * & * & * & * & * \\ 1 & 2 & * & * & * & * & * \\ * & * & 1 & 1 & * & * & * \\ * & * & 1 & 1 & * & * & * \\ * & * & * & * & 2 & 1 & 2 \\ * & * & * & * & 2 & 1 & * \end{bmatrix}$$

Llegados a este punto, debemos recordar (ver por ejemplo Schott, 1997) que las transformaciones de filas y columnas se pueden almacenar en matrices de dimensiones adecuadas y tendríamos

$$T(C) = U_f C U_c$$

Donde U_f y U_c almacenan las transformaciones filas y columnas que se han realizado en *C*. Admitiendo las siguientes reglas para las operaciones con el símbolo $*$

$$0 + * = * + 0 = *,\ 0 \times * = * \times 0 = 0,\ 1 \times * = * \times 1 = *$$

Ejemplo 4 (continuación)

En nuestro caso, se ha realizado una transformación por filas, intercambio de la fila dos por la seis. Realizando esta misma transformación en filas y columnas a la matriz identidad del mismo número de filas que *C*, se obtiene U_f

$$U_f = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Hemos realizado dos transformaciones por columnas, intercambios de la dos y la cuatro e intercambio de la tres y la siete. Realizando esta misma transformación en las columnas de la matriz identidad del mismo número de columnas que *C*, se obtiene U_c

$$
U_c = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}
$$

Obtenemos *T(C)* por multiplicación de estas matrices:

$$
\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}
\begin{bmatrix} 1 & * & * & 2 & * & * & * \\ * & * & 1 & * & * & 1 & * \\ * & 2 & * & * & 1 & * & 2 \\ * & * & * & * & 1 & * & 2 \\ * & * & 1 & * & * & 1 & * \\ 0.5 & * & * & 1 & * & * & * \end{bmatrix}
\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} =
$$

$$
= \begin{bmatrix} 1 & 2 & * & * & * & * & * \\ 0.5 & 1 & * & * & * & * & * \\ * & * & 2 & 2 & 1 & * & * \\ * & * & 2 & * & 1 & * & * \\ * & * & * & * & * & 1 & 1 \\ * & * & * & * & * & 1 & 1 \end{bmatrix}
$$

Lema 1.6.2 **(**Song, Li, Chen, Jiang y Kuo, 2010). Una matriz de cocientes C es de rango uno si y sólo si en cualquier DBI $T(C) = Diag(T_1, T_2, ..., T_M)$, $M \geq 1$ de ella, cada T_k $k = 1, 2, ..., M$ es de rango uno.

Teorema 1.6.9 **(**Song, Li, Chen, Jiang y Kuo, 2010). Sea *C* la matriz de cocientes de *A* y *B* y $T(C) = Diag(T_1, T_2, ..., T_M)$, $M \geq 1$ una DBI de ella entonces *A* y *B* son compatibles si y sólo si $N_A = N_B = N$ y cada T_k $k = 1, 2, ..., M$ es de rango uno mejor tiene una extensión positiva de rango uno.

Ejemplo 4 (continuación)*:*

Consideremos, de nuevo

$$T(C) = \begin{bmatrix} 1 & 2 & * & * & * & * & * \\ 0.5 & 1 & * & * & * & * & * \\ * & * & 2 & 2 & 1 & * & * \\ * & * & 2 & * & 1 & * & * \\ * & * & * & * & * & 1 & 1 \\ * & * & * & * & * & 1 & 1 \end{bmatrix}$$

cuyos bloques diagonales (incompletos) son

$$T_1 = \begin{bmatrix} 1 & 2 \\ 0.5 & 1 \end{bmatrix} \; T_2 = \begin{bmatrix} 2 & 2 & 1 \\ 2 & * & 1 \end{bmatrix} \; T_3 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

Notemos que los bloque diagonales de *T(C)* pueden tener elementos desconocidos, de hecho los tienen, y debemos calcular una extensión positiva de rango uno. En este caso, puesto que T_3 debe tener rango dos, las filas deben ser proporcionales y, de esta forma, el elemento desconocido vale 2.

$$\bar{T}_1 = \begin{bmatrix} 1 & 2 \\ 0.5 & 1 \end{bmatrix} \; \bar{T}_2 = \begin{bmatrix} 2 & 2 & 1 \\ 2 & 2 & 1 \end{bmatrix} \; \bar{T}_3 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

Puesto que todos los bloques diagonales tienen rango unidad, concluimos que *A* y *B* son compatibles.

7.- Aplicaciones del método de extensión a rango uno.

La potencia del método de extensión positiva de *C* con rango uno puede verse en las siguientes aplicaciones, no sólo a nivel práctico, también en términos teóricos.

i.- Especificación incompleta.

Arnold, Castillo y Sarabia (2004) estudia el caso en el que las matrices *A* y *B* no están completamente especificadas, tengan o no ceros, en este último caso deben ser conocidos. La idea básica del tratamiento de esta situación consiste en construir la matriz de cocientes *C* y deducir el valor de los parámetros bajo la condición $rang(C)=1$.

Ejemplo 1:

Consideremos las matrices dadas parcialmente por

$$A=\begin{bmatrix} \frac{1}{4} & \frac{2}{3} & a_{13} \\ a_{21} & \frac{1}{3} & a_{23} \end{bmatrix} \qquad B=\begin{bmatrix} b_{11} & b_{12} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{6} & b_{23} \end{bmatrix}$$

El problema consiste en encontrar esos valores de forma que las matrices sean compatibles.

Observemos, en primer lugar, que A es estocástica por columnas, por lo que $a_{21} = \frac{3}{4}$ y al ser B estocástica por filas $b_{23} = \frac{1}{3}$. Además $a_{13} + a_{23} = 1$ y $b_{11} + b_{12} + \frac{1}{4} = 1$

Construimos la matriz C de cocientes:

$$C = \begin{bmatrix} \frac{1}{4b_{11}} & \frac{2}{3b_{12}} & 4a_{13} \\ \frac{3}{2} & 2 & 3a_{23} \end{bmatrix}$$

De aquí, puesto que $rang(C) = 1$ tenemos:

$$\begin{vmatrix} \frac{1}{4b_{11}} & \frac{2}{3b_{12}} \\ \frac{3}{2} & 2 \end{vmatrix} = 0 \qquad \begin{vmatrix} \frac{1}{4b_{11}} & 4a_{13} \\ \frac{3}{2} & 3a_{23} \end{vmatrix} = 0$$

Que junto a $a_{13} + a_{23} = 1$ y $b_{11} + b_{12} + \frac{1}{4} = 1$ forman un sistema de ecuaciones cuya solución es $a_{13} = \frac{1}{6}$, $a_{23} = \frac{2}{3}$ y $b_{11} = \frac{1}{4}$, $b_{12} = \frac{1}{2}$ lo que nos permite obtener A y B.

Un caso particular de este problema no requiere la partición del método de extensión a rango unidad, se trata del caso en el que una de las matrices es completamente conocida y la otro lo es parcialmente. Aquí es interesante la reducción del número de incógnitas al pasar por la distribución conjunta del vector (supongamos A completamente conocida) mediante $a_{ij}\beta_j$ a partir de ahí $b_{ij} = \frac{a_{ij}\beta_j}{\sum_{j=1}^{m} a_{ij}\beta_j}$. Veamos un ejemplo.

Ejemplo 2:

Supongamos que la matriz A esta completamente especificada y viene dada por

$$A = \begin{bmatrix} \frac{1}{4} & \frac{2}{3} & \frac{1}{3} \\ \frac{3}{4} & \frac{1}{3} & \frac{2}{3} \end{bmatrix}$$

Mientras que de la matriz B sólo conocemos $b_{11} = \frac{1}{4}$ y $b_{22} = \frac{1}{6}$

La matriz M de la distribución conjunta viene dada por

$$M = \begin{bmatrix} \frac{1}{4}\beta_1 & \frac{2}{3}\beta_2 & \frac{1}{3}\beta_3 \\ \frac{3}{4}\beta_1 & \frac{1}{3}\beta_2 & \frac{2}{3}\beta_3 \end{bmatrix} \qquad \beta_1 + \beta_2 + \beta_3 = 1$$

Y la matriz B es

$$B = \begin{bmatrix} \frac{\frac{1}{4}\beta_1}{\frac{1}{4}\beta_1 + \frac{2}{3}\beta_2 + \frac{1}{3}\beta_3} & \frac{\frac{2}{3}\beta_2}{\frac{1}{4}\beta_1 + \frac{2}{3}\beta_2 + \frac{1}{3}\beta_3} & \frac{\frac{1}{3}\beta_3}{\frac{1}{4}\beta_1 + \frac{2}{3}\beta_2 + \frac{1}{3}\beta_3} \\ \frac{\frac{3}{4}\beta_1}{\frac{3}{4}\beta_1 + \frac{1}{3}\beta_2 + \frac{2}{3}\beta_3} & \frac{\frac{1}{3}\beta_2}{\frac{3}{4}\beta_1 + \frac{1}{3}\beta_2 + \frac{2}{3}\beta_3} & \frac{\frac{2}{3}\beta_3}{\frac{3}{4}\beta_1 + \frac{1}{3}\beta_2 + \frac{2}{3}\beta_3} \end{bmatrix}$$

Ahora bien, nuestro conocimiento de B, nos lleva a

$$\frac{\frac{1}{4}\beta_1}{\frac{1}{4}\beta_1 + \frac{2}{3}\beta_2 + \frac{1}{3}\beta_3} = \frac{1}{4}$$
$$\frac{\frac{1}{3}\beta_2}{\frac{3}{4}\beta_1 + \frac{1}{3}\beta_2 + \frac{2}{3}\beta_3} = \frac{1}{6}$$

Que junto a $\beta_1 + \beta_2 + \beta_3 = 1$ forman un sistema de ecuaciones lineales que tiene solución única y es $\beta_1 = \frac{4}{10}$, $\beta_2 = \beta_3 = \frac{3}{10}$. De esta forma

$$M = \begin{bmatrix} \frac{1}{10} & \frac{2}{10} & \frac{1}{10} \\ \frac{3}{10} & \frac{1}{10} & \frac{2}{10} \end{bmatrix}$$

Y la matriz B viene dada por

$$B = \begin{bmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{6} & \frac{1}{3} \end{bmatrix}$$

Los siguientes resultados muestran la potencia del método rango unidad en la demostración de resultados teóricos. Concretamente estudiaremos condiciones sobre las matrices A y B para que las marginales tengan alguna propiedad probabilística de interés.

ii.- Distribuciones condicionalmente especificadas con marginales uniformes.

Teorema 1.7.1 (marginales uniformes). Sean A y B dos matrices $n\times m$ en las condiciones habituales. Entonces A y B son compatibles y admiten marginales uniformes si y sólo si $nA=mB$.

Demostración:

Si A y B son compatibles y admiten marginales uniformes, entonces admite una extensión de rango unidad que verifica,

$$\overline{c}_{ij} = \frac{a_{ij}}{b_{ij}} = \frac{P(X = x_i \mid Y = y_j)}{P(Y = y_j \mid X = x_i)} = \frac{P(X = x_i)}{P(Y = y_j)} = \frac{1/n}{1/m} = \frac{m}{n} \quad \forall (i,j) \in N_{n\times m}$$

Por lo tanto $na_{ij} = mb_{ij} \quad \forall (i,j) \in N_{n\times m} \rightarrow nA = mB$

Recíprocamente, supongamos *nA=mB*

$$nA = mB \rightarrow na_{ij} = mb_{ij} \quad \forall (i,j) \in N \rightarrow c_{ij} = \frac{m}{n} \quad \forall (i,j) \in N$$

Y por tanto *C* admite una extensión de rango unidad definida por:

$$\overline{c}_{ij} = \frac{m}{n} \quad \forall (i,j) \in N_{n\times m}$$

De esta forma, las matrices *A* y *B* son compatibles y se tiene,

$$\vec{1}_n^t \overline{C} = \vec{1}_n^t \frac{m}{n} 1_{n\times m} = \frac{m}{n} \vec{1}_n^t 1_{n\times m} = \frac{m}{n} n \vec{1}_m^t$$

Y por tanto $\vec{\beta} = \begin{pmatrix} \frac{1}{m} \\ \frac{1}{m} \\ \vdots \\ \frac{1}{m} \end{pmatrix}$ y de la misma manera $\vec{\alpha} = \begin{pmatrix} \frac{1}{n} \\ \frac{1}{n} \\ \vdots \\ \frac{1}{n} \end{pmatrix}$.

□

Corolario 1.7.1. Sean *A* y *B* dos matrices *n*×*n* en las condiciones habituales. Entonces *A* y *B* son compatibles y admiten marginales uniformes si y sólo si *A*=*B*.

iii.- Distribuciones condicionalmente especificadas con marginales idénticamente distribuidas.

Los siguientes resultados exploran el caso en el que las marginales están idénticamente distribuidas, por tanto supondremos que *n*=*m* y que, los rango de las variables aleatorias son, sin pérdida de generalidad, $R_X = R_Y = \{1,2,...,n\}$.

Teorema 1.7.2 (marginales idénticamente distribuidas). Sean A y B dos matrices $n \times n$ en las condiciones habituales. Supongamos además, que las matrices son compatibles y admiten marginales idénticamente distribuidas, entonces

(i) $diag(A) = diag(B)$

(ii) $\frac{a_{ij}}{b_{ij}} = \frac{b_{ji}}{a_{ji}} \ \forall (i,j),(j,i) \in N$

Demostración:

Supongamos que las matrices admiten marginales idénticamente distribuidas, sean estas X e Y.

Probaremos que $diag(A) = diag(B)$. Para cada i=1,2,..,n

$$a_{ii} = P(X = i \mid Y = i) = \frac{P(X = i, Y = i)}{P(Y = i)} = \frac{P(X = i, Y = i)}{P(X = i)} = P(Y = i \mid X = i) = b_{ii}$$

Probaremos que $\frac{a_{ij}}{b_{ij}} = \frac{b_{ji}}{a_{ji}} \ \forall (i,j),(j,i) \in N$

$$\frac{a_{ij}}{b_{ij}} = \frac{P(X = i \mid Y = j)}{P(Y = j \mid X = i)} = \frac{P(X = i)}{P(Y = j)} = \frac{P(Y = i)}{P(X = j)} = \frac{P(X = i \mid X = j)}{P(X = j \mid Y = i)} = \frac{b_{ji}}{a_{ji}} \ \forall (i,j),(j,i) \in N$$

□

Teorema 1.7.3 (marginales idénticamente distribuidas). Sean A y B dos matrices $n \times n$ en las condiciones habituales, tales que ningún elemento de la diagonal es nulo y compatibles. Si $diag(A) = diag(B)$ entonces las marginales están idénticamente distribuidas.

Demostración:

Si $diag(A)=diag(B)$ y los elementos son no nulos entonces $c_{ii}=1\ \forall i=1,2,...,n$ y por tanto
$$c_{ii}=\frac{P(X=i)}{P(Y=i)}=\frac{P(Y=i)}{P(X=i)}=1\ \forall i=1,2,...,n\rightarrow P(X=i)=P(Y=i)\ \ \forall i=1,2,...,n$$

Y las marginales están idénticamente distribuidas.

□

Corolario 1.7.2. Sean *A* y *B* dos matrices $n\times n$ en las condiciones habituales, tales que admiten una extensión positiva de rango unidad de su matriz de cocientes con elementos de la diagonal iguales a uno entonces son compatibles y las marginales están idénticamente distribuidas

La condición sobre la diagonal en el teorema anterior es esencial

Ejemplo:

Sean las matrices

$$A=\begin{bmatrix}\frac{1}{4} & \frac{1}{2} & \frac{1}{2}\\ \frac{1}{2} & 0 & \frac{1}{2}\\ \frac{1}{4} & \frac{1}{2} & 0\end{bmatrix}\qquad B=\begin{bmatrix}\frac{1}{4} & \frac{1}{4} & \frac{1}{2}\\ \frac{1}{2} & 0 & \frac{1}{2}\\ \frac{1}{2} & \frac{1}{2} & 0\end{bmatrix}$$

Son compatibles, tienen diagonales iguales, sin embargo no están idénticamente distribuidas pues su matriz de cocientes es

$$C=\begin{bmatrix}1 & 2 & 1\\ 1 & * & 1\\ \frac{1}{2} & 1 & *\end{bmatrix}$$

Y sólo admite la extensión positiva de rango unidad, que tiene elementos distintos de unos en la diagonal

$$\bar{C} = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 2 & 1 \\ \frac{1}{2} & 1 & \frac{1}{2} \end{bmatrix}$$

Las marginales son:

$$\vec{\alpha} = \begin{bmatrix} \frac{2}{5} \\ \frac{2}{5} \\ \frac{1}{5} \end{bmatrix} \qquad \vec{\beta} = \begin{bmatrix} \frac{2}{5} \\ \frac{1}{5} \\ \frac{2}{5} \end{bmatrix}$$

Denotemos por $N_e = \{(i,j) \in N_{n\times n} : i \neq j\}$ y $N_d = \{(i,j) \in N_{n\times n} : i = j\}$ esto es, los elementos extra diagonales o diagonales de $N_{n\times n}$.

Teorema 1.7.4 (marginales idénticamente distribuidas). Sean *A* y *B* dos matrices $n\times n$ en las condiciones habituales, tales que $N_e \subseteq N$ y compatibles. Si $\frac{a_{ij}}{b_{ij}} = \frac{b_{ji}}{a_{ji}}$ $\forall (i,j),(j,i) \in N$ entonces las marginales están idénticamente distribuidas.

Demostración:

Por ser compatibles, la matriz de cocientes admite una extensión positiva de rango unidad. Probaremos que esa extensión tiene unos en la diagonal.

$$\begin{vmatrix} \bar{c}_{11} & \bar{c}_{12} \\ \bar{c}_{21} & \bar{c}_{22} \end{vmatrix} = 0 \to \bar{c}_{11}\bar{c}_{22} - \bar{c}_{21}\bar{c}_{12} = 0 \to \bar{c}_{11}\bar{c}_{22} - 1 = 0 \to \bar{c}_{11}\bar{c}_{22} = 1$$

De la misma forma $\bar{c}_{22}\bar{c}_{33} = 1, \bar{c}_{33}\bar{c}_{44} = 1, \bar{c}_{(n-1)(n-1)}\bar{c}_{nn} = 1$ y considerando ahora el menor $\begin{vmatrix} \bar{c}_{11} & \bar{c}_{13} \\ \bar{c}_{31} & \bar{c}_{33} \end{vmatrix} = 0 \to \bar{c}_{11}\bar{c}_{33} - \bar{c}_{31}\bar{c}_{13} = 0 \to \bar{c}_{11}\bar{c}_{33} - 1 = 0 \to \bar{c}_{11}\bar{c}_{33} = 1$ es fácil ver que $\bar{c}_{11} = \bar{c}_{22} = \ldots = \bar{c}_{nn} = 1$, por tanto, aplicando el corolario anterior se tiene el resultado.

□

Corolario 1.7.3. Sean A y B dos matrices $n \times n$ en las condiciones habituales, supongamos además, que $N = N_{n \times n}$ y son compatibles. Son equivalentes:

(i) $diag(A) = diag(B)$

(ii) $\frac{a_{ij}}{b_{ij}} = \frac{b_{ji}}{a_{ji}} \ \forall (i,j),(j,i) \in N$

(iii) Las marginales están idénticamente distribuidas.

Demostración:

(i)→(iii) basta aplicar el teorema 3

(iii)→(ii) basta aplicar el teorema 2

(ii)→(i) está probado en la demostración del teorema 4.

□

iv.- Distribuciones condicionalmente especificadas con marginales independientes.

Recordemos que en este trabajo suponemos que en cada fila y columna de A y B hay al menos un elemento no nulo, esto implica que, de ser compatibles, las distribuciones marginales no producen probabilidades nulas.

Teorema 1.7.5 (condición necesaria de independencia). Sean A y B dos matrices $n \times m$ en las condiciones habituales. Supongamos además, que las matrices son compatibles. Si A y B admiten marginales independientes entonces ..

Demostración:

Denotamos por X e Y las marginales independientes. Supongamos que $N \neq N_{n \times m}$ entonces existe

$$a_{ij}=0 \rightarrow P(X=x_i \mid Y=y_j)=0 \rightarrow P(X=x_i Y=y_j)=0 \rightarrow P(X=x_i)P(Y=y_j)=0$$

Y en consecuencia $P(X=x_i)=0$ o $P(Y=y_j)=0$ lo que es absurdo.

□

En consecuencia, sólo nos plantearemos la independencia de las marginales si $N=N_{n\times m}$

Teorema 1.7.6 (marginales independientes). Sean *A* y *B* dos matrices $n\times m$ en las condiciones habituales con $N=N_{n\times m}$. Estas matrices son compatibles y admiten marginales independientes, si y sólo si las columnas *A* son iguales entre si y las filas de *B* son iguales entre sí.

Demostración:

Supongamos que las matrices son compatibles y admiten marginales independientes, entonces

$$\begin{aligned} a_{ij} &= P(X=x_i \mid Y=y_j)=P(X=x_i) \quad \forall (i,j)\in N_{n\times m} \\ b_{ij} &= P(Y=y_j \mid X=x_i)=P(Y=y_j) \quad \forall (i,j)\in N_{n\times m} \end{aligned}$$

Y las matrices son

$$A=\begin{bmatrix} P(X=x_1) & P(X=x_1) & \dots & P(X=x_1) \\ P(X=x_2) & P(X=x_2) & \dots & P(X=x_2) \\ \vdots & \vdots & \ddots & \vdots \\ P(X=x_n) & P(X=x_n) & \dots & P(X=x_n) \end{bmatrix}$$

$$B=\begin{bmatrix} P(Y=y_1) & P(Y=y_2) & \dots & P(Y=y_m) \\ P(Y=y_1) & P(Y=y_2) & \dots & P(Y=y_m) \\ \vdots & \vdots & \ddots & \vdots \\ P(Y=y_1) & P(Y=y_2) & \dots & P(Y=y_m) \end{bmatrix}$$

Recíprocamente, si las columnas A son iguales entre si y las columnas de B son iguales entre sí, estas matrices adoptan la forma

$$A=\begin{bmatrix} a_1 & a_1 & \dots & a_1 \\ a_2 & a_2 & \dots & a_2 \\ \vdots & \vdots & \ddots & \vdots \\ a_n & a_n & \dots & a_n \end{bmatrix} \qquad B=\begin{bmatrix} b_1 & b_2 & \dots & b_m \\ b_1 & b_2 & \dots & b_m \\ \vdots & \vdots & \ddots & \vdots \\ b_1 & b_2 & \dots & b_m \end{bmatrix}$$

Y la matriz C es $c_{ij}=\dfrac{a_{ij}}{b_{ij}}=\dfrac{a_i}{b_j}\ \ \forall (i,j)\in N$. Veamos su rango,

Sean $i_1,i_2=1,2,...,n$ y $j_1,j_2=1,2,...,m$ entonces el menor de la matriz C vale

$$\begin{vmatrix} c_{i_1 j_1} & c_{i_1 j_2} \\ c_{i_2 j_1} & c_{i_2 j_2} \end{vmatrix}=\begin{vmatrix} \dfrac{a_{i_1}}{b_{j_1}} & \dfrac{a_{i_1}}{b_{j_2}} \\ \dfrac{a_{i_2}}{b_{j_1}} & \dfrac{a_{i_2}}{b_{j_2}} \end{vmatrix}=0$$

Y por tanto, las matrices son compatibles y existe *(X,Y)* que verifica,

$$\begin{aligned} a_{ij} &= P(X=x_i \mid Y=y_j)=a_i=P(X=x_1 \mid Y=y_j)= \\ &= P(X=x_2 \mid Y=y_j)=...=P(X=x_m \mid Y=y_j) \quad \forall j=1,2,...,m \end{aligned}$$

Luego X e Y son independientes.

□

v.- Distribuciones condicionalmente especificadas con marginales independientes e idénticamente distribuidas.

Combinemos las condiciones obtenidas más arriba.

Condiciones derivadas de la idéntica distribución:

$n=m$

$diag(A)=diag(B)$

$$\frac{a_{ij}}{b_{ij}} = \frac{b_{ji}}{a_{ji}} \;\; \forall (i,j),(j,i) \in N$$

Condiciones derivadas de la independencia:

$$N = N_{n\times n}$$

Las columnas *A* son iguales entre si y las columnas de *B* son iguales entre sí, es decir:

$$A = \begin{bmatrix} a_1 & a_1 & \dots & a_1 \\ a_2 & a_2 & \dots & a_2 \\ \vdots & \vdots & \ddots & \vdots \\ a_n & a_n & \dots & a_n \end{bmatrix} \qquad B = \begin{bmatrix} b_1 & b_2 & \dots & b_m \\ b_1 & b_2 & \dots & b_m \\ \vdots & \vdots & \ddots & \vdots \\ b_1 & b_2 & \dots & b_m \end{bmatrix}$$

Con $a_i, b_i > 0,\ i = 1,2,...,n$

Partiendo de las condiciones de independencia, $\frac{a_{ij}}{b_{ij}} = \frac{b_{ji}}{a_{ji}} \;\; \forall (i,j),(j,i) \in N$ equivale a $\frac{a_i}{b_j} = \frac{b_i}{a_j} \;\; \forall (i,j) \in N_{n\times n} \leftrightarrow a_i a_j = b_i b_j \;\; \forall (i,j),(j,i) \in N_{n\times n}$ y la condición $diag(A) = diag(B)$ equivale a $a_i = b_i \;\; i = 1,2,...,n$, por tanto la condición anterior es superflua.

En definitiva la estructura de las matrices *A* y *B* que verifican todas las condiciones es

$$A = \begin{bmatrix} a_1 & a_1 & \dots & a_1 \\ a_2 & a_2 & \dots & a_2 \\ \vdots & \vdots & \ddots & \vdots \\ a_n & a_n & \dots & a_n \end{bmatrix} \qquad B = \begin{bmatrix} a_1 & a_2 & \dots & a_m \\ a_1 & a_2 & \dots & a_m \\ \vdots & \vdots & \ddots & \vdots \\ a_1 & a_2 & \dots & a_m \end{bmatrix}$$

Con $a_i > 0,\ i = 1,2,...,n$. La matriz *A* tiene todas sus columnas iguales y $B = A^t$.

En definitiva, por aplicación del teorema 6, el corolario 3 y las observaciones anteriores, se tiene:

Teorema 1.7.7 (Marginales independientes e idénticamente distribuidas)

Sean A y B dos matrices $n\times n$ en las condiciones habituales con $N = N_{n\times n}$. Estas matrices son compatibles y admiten marginales independientes e idénticamente distribuidas, si y sólo si las columnas A son iguales entre si y $B = A^t$.

vi.- Especificación de modelos: condicionadas binomiales.

Existe una extensa literatura sobre especificación de modelos mediante condicionadas, en especial modelos continuos. La idea es fijar las condicionadas, en alguna familia de modelos (suficientemente amplia) y estudiar las condiciones para que sean compatibles, a título de ejemplo, estudiamos el caso en el que ambas condicionadas son binomiales, para ello utilizaremos el método de extensión a rango unidad.

Sean A y B dos matrices en las condiciones habituales y supongamos que corresponden a distribuciones condicionadas binomiales, esto es, para la matriz A sus elementos son

$a_{xy} = \binom{n}{x} p(y)^x (1-p(y))^{n-x}$, $y = 0,1,...,n;\quad x = 0,1,...,n$ es decir si las matrices fuesen compatibles, entonces $X \mid Y = y \sim B(n, p(y))\quad y = 0,1,...,n$. Análogamente, para B, $b_{xy} = \binom{n}{y} q(x)^y (1-q(x))^{n-y}$ $x = 0,1,...,n;\quad y = 0,1,...,n$.

Si las matrices fuesen compatibles, $Y \mid X = x \sim B(n, q(x))$ $x = 0,1,...,n$.

Queremos encontrar condiciones sobre *p(y)* y *q(x)* que hagan que las matrices sean compatibles. Utilizaremos el método de extensión a rango unidad.

En este caso, puesto que las matrices *A* y *B* no tienen elementos nulos, basta encontrar condiciones sobre *p(y)* y *q(x)* que hagan que el rango de la matriz de cocientes sea uno.

$$c_{xy} = \frac{a_{xy}}{b_{xy}} = \frac{\binom{n}{x} p(y)^x (1-p(y))^{n-x}}{\binom{n}{y} q(x)^y (1-q(x))^{n-y}} = \frac{\binom{n}{x} p(y)^x \,\overline{p}(y)^{n-x}}{\binom{n}{y} q(x)^y \,\overline{q}(x)^{n-y}}$$

Donde $\overline{p} = 1 - p$ y $\overline{q} = 1 - q$.

Para que la condición sobre el rango se verifique, todos los menores de orden dos deben ser nulos.

$$\begin{vmatrix} c_{x_1 y_1} & c_{x_1 y_2} \\ c_{x_2 y_1} & c_{x_2 y_2} \end{vmatrix} =$$

$$= \frac{\binom{n}{x_1}\binom{n}{x_2}}{\binom{n}{y_1}\binom{n}{y_2}} \left[\frac{p(y_1)^{x_1}\,\overline{p}(y_1)^{n-x_1}\, p(y_2)^{x_2}\,\overline{p}(y_2)^{n-x_2}}{q(x_1)^{y_1}\,\overline{q}(x_1)^{n-y_1}\, q(x_2)^{y_2}\,\overline{q}(x_2)^{n-y_2}} - \frac{p(y_2)^{x_1}\,\overline{p}(y_2)^{n-x_1}\, p(y_1)^{x_2}\,\overline{p}(y_1)^{n-x_2}}{q(x_1)^{y_2}\,\overline{q}(x_1)^{n-y_2}\, q(x_2)^{y_1}\,\overline{q}(x_2)^{n-y_1}} \right] = 0$$

Lo que ocurre si,

$$\frac{p(y_1)^{x_1}\,\overline{p}(y_1)^{-x_1}\, p(y_2)^{x_2}\,\overline{p}(y_2)^{-x_2}}{q(x_1)^{y_1}\,\overline{q}(x_1)^{-y_1}\, q(x_2)^{y_2}\,\overline{q}(x_2)^{-y_2}} = \frac{p(y_2)^{x_1}\,\overline{p}(y_2)^{-x_1}\, p(y_1)^{x_2}\,\overline{p}(y_1)^{-x_2}}{q(x_1)^{y_2}\,\overline{q}(x_1)^{-y_2}\, q(x_2)^{y_1}\,\overline{q}(x_2)^{-y_1}}$$

Que operando se obtiene que el rango de la matriz de cocientes C tiene rango unidad si y sólo si

$$\left[\frac{p(y_1)}{\bar{p}(y_1)}\frac{\bar{p}(y_2)}{p(y_2)}\right]^{x_1-x_2}=\left[\frac{q(x_1)}{\bar{q}(x_1)}\frac{\bar{q}(x_2)}{q(x_2)}\right]^{y_1-y_2} \text{ para todo } x_1,x_2,y_1,y_2=0,1,...,n$$

Que para el caso $n=1$ (Bernoulli) la condición se convierte en $\frac{a_1a_2}{\bar{a}_1\bar{a}_2}=\frac{b_1b_2}{\bar{b}_1\bar{b}_2}$ donde $0<a_1,a_2,b_1,b_2<1$ y se ha utilizado la parametrización de Hürlimann (1993),

$$A=\begin{bmatrix} a_1 & 1-a_2 \\ 1-a_1 & a_2 \end{bmatrix} \qquad B=\begin{bmatrix} b_1 & 1-b_1 \\ 1-b_2 & b_2 \end{bmatrix}$$

Véase Arnold, Castillo y Sarabia (1999) y Hürlimann (1993) para más información sobre condicionadas binomiales, Bernoulli y dicotómicas.

Capítulo 2

Unicidad.

Dadas dos matrices $A=(a_{ij})$ y $B=(b_{ij})$ de orden $n\times m$ que verifican:

$$a_{ij}\geq 0,\ b_{ij}\geq 0 \quad \forall\, i=1,\dots,n;\ \forall\, j=1,\dots,m \text{ y además}$$

$$\sum_{i=1}^{n} a_{ij}=1, \quad \sum_{j=1}^{m} b_{ij}=1.$$

Sabemos que pueden ser compatibles o no. Además, se han visto ejemplos en los que son compatibles y tiene una única solución y en los que la solución no es única. Recordar los ejemplos de la sección dos del capítulo anterior, entre otros. Aquí entendemos por solución al problema, bien la distribución de probabilidad conjunta o la pareja de marginales compatibles con las matrices condicionadas.

En el capítulo anterior proporcionamos métodos para dilucidar cuándo un par de matrices, en las condiciones anteriores, tiene solución. En este capítulo daremos métodos que nos permitan distinguir cuándo A y B proporcionan una o varias soluciones al problema de compatibilidad.

En la primera sección estudiamos la estructura topólogica de las soluciones, en las siguientes, estudiaremos diversos métodos para discutir la unicidad de la solución, en particula r, volveremos sobre la fórmula de Besag y la extensión de rango uno. También obtendremos algunos resultados basados en teorías de grafos.

1.- Estructura de las soluciones.

El principal resultado de esta sección comprueba el carácter convexo del conjunto de soluciones.

Definición 1: Sean A y B dos matrices en las condiciones anteriores. Llamaremos conjunto de soluciones del problema de compatibilidad entre A y B a

$$S(A,B)=\left\{\begin{pmatrix}\vec{\alpha}\\ \vec{\beta}\end{pmatrix}\in\mathbb{R}^{n+m}:\vec{\alpha}\text{ y }\vec{\beta}\text{ son marginales compatibles con }A\text{ y }B\right\}$$

O equivalentemente,

$$S(A,B)=\left\{\begin{pmatrix}\vec{\alpha}\\ \vec{\beta}\end{pmatrix}\in\mathbb{R}^{n+m}:D(\vec{\alpha})B=AD(\vec{\beta})\text{ y }\sum_{i=1}^{n}\alpha_i=1,\ \sum_{j=1}^{m}\beta_j=1,\ \alpha_i,\beta_j>0,\ \forall i,j\right\}$$

Si no hay lugar a confusión, denotaremos $S(A,B)=S$

Es claro que, A y B son compatibles si y sólo si $S(A,B)\neq\varnothing$

Teorema 1.2.1. Dadas dos matrices A y B en las condiciones habituales, $S(A,B)$ es convexo.

Demostración:

Si las matrices son incompatibles $S(A,B)=\varnothing$ y por lo tanto convexo.

Si las matrices producen una única solución entonces $S(A,B)=\left\{\begin{pmatrix}\vec{\alpha}_0\\ \vec{\beta}_0\end{pmatrix}\right\}$ es unitario y por lo tanto convexo.

Supongamos que el cardinal de *S(A,B)*, $card\left(S(A,B)\right)\geq 2$ y sean $\begin{pmatrix}\vec{\alpha}_1\\ \vec{\beta}_1\end{pmatrix},\begin{pmatrix}\vec{\alpha}_2\\ \vec{\beta}_2\end{pmatrix}\in S(A,B)$ distintos y $\lambda\in[0,1]$ notemos, en primer lugar, que $\lambda\begin{pmatrix}\vec{\alpha}_1\\ \vec{\beta}_1\end{pmatrix}+(1-\lambda)\begin{pmatrix}\vec{\alpha}_2\\ \vec{\beta}_2\end{pmatrix}$ tiene todas sus coordenadas positivas. Por otra parte, se tiene:

$$\begin{pmatrix}\vec{\alpha}_1\\ \vec{\beta}_1\end{pmatrix},\begin{pmatrix}\vec{\alpha}_2\\ \vec{\beta}_2\end{pmatrix}\in S(A,B)\leftrightarrow\begin{cases}D(\vec{\alpha}_1)B=AD(\vec{\beta}_1)\leftrightarrow\lambda D(\vec{\alpha}_1)B=\lambda AD(\vec{\beta}_1)\\ D(\vec{\alpha}_2)B=AD(\vec{\beta}_2)\leftrightarrow(1-\lambda)D(\vec{\alpha}_2)B=(1-\lambda)AD(\vec{\beta}_2)\end{cases}\leftrightarrow$$

$$\leftrightarrow\begin{cases}D(\lambda\vec{\alpha}_1)B=AD(\lambda\vec{\beta}_1)\\ D((1-\lambda)\vec{\alpha}_2)B=AD((1-\lambda)\vec{\beta}_2)\end{cases}\rightarrow$$

$$\rightarrow D(\lambda\vec{\alpha}_1)B+D((1-\lambda)\vec{\alpha}_2)B=AD(\lambda\vec{\beta}_1)+AD((1-\lambda)\vec{\beta}_2)\leftrightarrow$$

$$\leftrightarrow\left(D(\lambda\vec{\alpha}_1)+D((1-\lambda)\vec{\alpha}_2)\right)B=A\left(D(\lambda\vec{\beta}_1)+D((1-\lambda)\vec{\beta}_2)\right)\leftrightarrow$$

$$\leftrightarrow D\left(\lambda\vec{\alpha}_1+(1-\lambda)\vec{\alpha}_2\right)B=AD\left(\lambda\vec{\beta}_1+(1-\lambda)\vec{\beta}_2\right)$$

Por tanto $\begin{pmatrix}\lambda\vec{\alpha}_1+(1-\lambda)\vec{\alpha}_2\\ \lambda\vec{\beta}_1+(1-\lambda)\vec{\beta}_2\end{pmatrix}\in S(A,B)$ y evidentemente cada componente de ese vector es una distribución de probabilidad. En consecuencia $S(A,B)$ es convexo.

□

Corolario 1.2.1. Un problema de compatibilidad puede no tener solución, tener exactamente una o tener infinitas.

3.- Primeros resultados.

Sean dos matrices $A= (a_{ij})$ y $B = (b_{ij})$ de orden $n\times m$ que verifican:

$$a_{ij} \geq 0,\ b_{ij} \geq 0 \quad \forall i= 1,\dots,n;\ \forall j = 1,\dots,m \text{ y además}$$

$$\sum_{i=1}^{n} a_{ij} = 1, \quad \sum_{j=1}^{m} b_{ij} = 1.$$

En esta sección, salvo mención expresa de lo contrario, supondremos que son compatibles lo que implica, en particular que sus elementos nulos están situados en las mismas posiciones, $N_A= N_B =N$.

Un primer resultado de unicidad, por otra parte trivial, es el siguiente:

Teorema 2.3.1: Sean A y B dos matrices en las condiciones anteriores. Supongamos además que A y B son compatibles. Si $N = N_{n\times m}$ entonces, la solución es única.

Su contrarrecíproco puede ser de utilidad.

Corolario 2.3.1: Sean A y B dos matrices en las condiciones anteriores. Supongamos además que A y B son compatibles. Si la solución no es única, entonces $N \neq N_{n\times m}$.

En otras palabras, para que la solución no sea única, debe ocurrir $N \neq N_{n\times m}$.

La condición $N = N_{n\times m}$ puede relajarse, por ejemplo, tal y como indica el siguiente resultado.

Teorema 2.3.2 (Gupta y Varga 1990)*:* Si A y B tienen al menos una fila y una columna de elementos no nulos, la distribución conjunta que determinan A y B es única.

Ejemplo 1: Las siguientes matrices fueron estudiadas en el capítulo anterior y se obtuvo que son compatibles.

$$A = \begin{bmatrix} 0 & 0 & \frac{1}{3} & \frac{1}{2} \\ 0 & 0 & \frac{1}{3} & \frac{1}{4} \\ 1 & 1 & \frac{1}{3} & \frac{1}{4} \end{bmatrix} \quad \text{y} \qquad B = \begin{bmatrix} 0 & 0 & \frac{1}{3} & \frac{2}{3} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{2}{5} & \frac{2}{5} & \frac{1}{5} & \frac{1}{5} \end{bmatrix}$$

Al tener una fila y una columna formada con elementos no nulos, podemos concluir que la distribución conjunta que definen (o pares de marginales) es única.

4.- Uso de la fórmula de Besag.

Es posible utilizar el método basado en la fórmula de Besag para estudiar la unicidad de la distribución conjunta (o de la marginales), observemos que

Teorema 2.4.3 (Wang y Kuo, 2010). Sean A y B dos matrices en las condiciones anteriores. Supongamos además que A y B son compatibles. Si existe una posición que puede ser unida mediante, al menos, un camino de recorrido horizontal y/o vertical con el resto de posiciones que tienen probabilidades no nulas, la distribución conjunta es única.

Como ejemplo puede verse cualquiera de los estudiados por este método en el capítulo anterior.

Ejemplo 3:

Consideremos las matrices

$$A=\begin{bmatrix}0.5 & 0.5 & 0\\ 0.5 & 0.5 & 0\\ 0 & 0 & 1\end{bmatrix} \qquad B=\begin{bmatrix}0.25 & 0.75 & 0\\ 0.25 & 0.75 & 0\\ 0 & 0 & 1\end{bmatrix}$$

Estas matrices son compatibles pues proceden de la distribución conjunta dada por la matriz,

$$M=\begin{bmatrix}0.1 & 0.3 & 0\\ 0.1 & 0.3 & 0\\ 0 & 0 & 0.2\end{bmatrix}$$

Observe que no es aplicable el método basado en la fórmula de Besag, pues el elemento 3,3 de la matriz no se puede conectar con ninguno mediante caminos de recorrido horizontal y/o vertical. Observemos sin embargo que es posible modificar el método de forma que se puedan obtener la(s) distribución(es) conjunta(s).

Supondremos, sin pérdida de generalidad que los rangos de X e Y son 1,2,3 Comencemos desde la posición 1,1 que trataremos de unir con el resto de posiciones de probabilidades no nulas.

1,1→1,2 $\qquad \dfrac{P(X=1,Y=2)}{P(X=1,Y=1)}=\dfrac{P(Y=2\mid X=1)}{P(Y=1\mid X=1)}=\dfrac{0.75}{0.25}=3$

1,1→2,1 $\qquad \dfrac{P(X=2,Y=1)}{P(X=1,Y=1)}=\dfrac{P(X=2\mid Y=1)}{P(X=1\mid Y=1)}=\dfrac{0.5}{0.5}=1$

1,1→2,2 $\qquad \dfrac{P(X=2,Y=2)}{P(X=1,Y=1)}=\dfrac{P(X=2\mid Y=1)}{P(X=1\mid Y=1)}\dfrac{P(Y=2\mid X=2)}{P(Y=1\mid X=2)}=\dfrac{0.5}{0.5}\dfrac{0.75}{0.25}=3$

No hay un camino de recorrido vertical/horizontal que una la posición 1,1 con la 3,3 a través de posiciones con probabilidad no nula, sin embargo el cociente

$\dfrac{P(X=3,Y=3)}{P(X=1,Y=1)}$ existe pues tanto numerador como denominador son no nulos, denotémoslo $\lambda > 0$. Así,

$1,1 \to 3,3 \qquad \dfrac{P(X=3,Y=3)}{P(X=1,Y=1)} = \lambda$

Si $p = P(X=1,Y=1)$, tenemos

$$P(X=1,Y=2)=3p \qquad\qquad P(X=2,Y=1)=p$$

$$P(X=2,Y=2)=3p \qquad\qquad P(X=3,Y=3)=\lambda p$$

Sumando e igualando a uno obtenemos que $p = \dfrac{1}{8+\lambda}$

Podemos obtener la matriz de las distribuciones conjuntas compatibles con *A* y *B*:

$$M_\lambda = \begin{bmatrix} \frac{1}{8+\lambda} & \frac{3}{8+\lambda} & 0 \\ \frac{1}{8+\lambda} & \frac{3}{8+\lambda} & 0 \\ 0 & 0 & \frac{\lambda}{8+\lambda} \end{bmatrix}$$

Note que para $\lambda = 2$ se obtiene la distribución *M*.

Ejemplo 4:

Sean las matrices

$$A = \begin{bmatrix} \frac{1}{3} & 0 & \frac{1}{3} & 0 & \frac{1}{3} \\ 0 & 1 & 0 & 1 & 0 \\ \frac{2}{3} & 0 & \frac{2}{3} & 0 & \frac{2}{3} \end{bmatrix} \qquad B = \begin{bmatrix} \frac{1}{4} & 0 & \frac{1}{4} & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ \frac{1}{4} & 0 & \frac{1}{4} & 0 & \frac{1}{2} \end{bmatrix}$$

Partiendo de la posición 1,1 es fácil, como en el ejemplo precedente, obtener las siguientes relaciones:

$$\frac{P(X=1,Y=3)}{P(X=1,Y=1)}=1 \qquad \frac{P(X=1,Y=4)}{P(X=1,Y=1)}=2 \qquad \frac{P(X=3,Y=1)}{P(X=1,Y=1)}=2$$

$$\frac{P(X=3,Y=3)}{P(X=1,Y=1)}=2 \qquad \frac{P(X=3,Y=4)}{P(X=1,Y=1)}=4$$

No es posible unir la posición 1,1 con la 2,2 por un camino en las condiciones requeridas, pero la razón $\frac{P(X=2,Y=2)}{P(X=1,Y=1)}$ existe, aunque es desconocida, pongamos $\frac{P(X=2,Y=2)}{P(X=1,Y=1)}=\lambda>0$, en este caso diremos que la posición 2,2 se usa como posición puente. A partir de ahí se puede conectar 1,1 con 2,4, aunque el camino no tenga las condiciones requeridas, pero la posición 2,2 hace, digamos de puente,

$$\frac{P(X=2,Y=4)}{P(X=1,Y=1)}=\frac{P(X=2,Y=2)}{P(X=1,Y=1)}\frac{P(X=2,Y=4)}{P(X=2,Y=2)}=\lambda\frac{\frac{1}{2}}{\frac{1}{2}}=\lambda$$

Haciendo $p=P(X=1,Y=1)$, tenemos

$$P(X=1,Y=3)=p \qquad P(X=1,Y=4)=2p$$

$$P(X=3,Y=1)=2p$$

$$P(X=3,Y=3)=2p \qquad P(X=3,Y=4)=4p$$

$$P(X=2,Y=2)=\lambda p$$

$$P(X=2,Y=4)=\lambda p$$

Sumando e igualando a uno obtenemos que $p=\frac{1}{12+2\lambda}$

Y la matriz de las distribuciones conjuntas es

$$M_\lambda = \begin{bmatrix} \frac{1}{12+2\lambda} & 0 & \frac{1}{12+2\lambda} & 0 & \frac{2}{12+2\lambda} \\ 0 & \frac{\lambda}{12+2\lambda} & 0 & \frac{\lambda}{12+2\lambda} & 0 \\ \frac{2}{12+2\lambda} & 0 & \frac{2}{12+2\lambda} & 0 & \frac{4}{12+2\lambda} \end{bmatrix}$$

Es fácil ver que para cada λ, las distribuciones condicionadas coinciden con *A* y *B*, por tanto *A* y *B* son compatibles y las distribuciones que determinan no son únicas.

Ejemplo 5:

$$A = \begin{bmatrix} \frac{1}{3} & \frac{1}{2} & 0 & 0 & 0 & 0 \\ \frac{2}{3} & \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{3} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{2}{3} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \qquad B = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 \\ \frac{2}{3} & \frac{1}{3} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{2}{3} & \frac{1}{3} & 0 & 0 \\ 0 & 0 & \frac{4}{5} & \frac{1}{5} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

Comencemos por 1,1 y, como en los ejemplos anteriores.

$$\frac{P(X=1,Y=2)}{P(X=1,Y=1)} = 1 \qquad \frac{P(X=2,Y=1)}{P(X=1,Y=1)} = 2 \qquad \frac{P(X=2,Y=2)}{P(X=1,Y=1)} = 1$$

Puesto que no podemos unir 1,1 con 3,3 pongamos $\frac{P(X=3,Y=3)}{P(X=1,Y=1)} = \lambda > 0$ (3,3 posición puente) y a partir de ahí,

$$\frac{P(X=3,Y=4)}{P(X=1,Y=1)} = \frac{P(X=3,Y=3)}{P(X=1,Y=1)} \frac{P(X=3,Y=4)}{P(X=3,Y=3)} = \frac{\lambda}{2}$$

De la misma forma

$$\frac{P(X=4,Y=3)}{P(X=1,Y=1)}=2\lambda \qquad \frac{P(X=4,Y=4)}{P(X=1,Y=1)}=\frac{\lambda}{2}$$

Para acceder al último bloque de la matriz, pongamos $\dfrac{P(X=5,Y=5)}{P(X=1,Y=1)}=\mu>0$

Y, a partir de él, se obtiene que $\dfrac{P(X=5,Y=6)}{P(X=1,Y=1)}=\mu$ (5,6 nueva posición puente)

y como en otros ejemplos, sumando e igualando a uno, se tiene

$P(X=1,Y=1)=\dfrac{1}{5+4\lambda+2\mu}$ y de ahí:

$$M_{\lambda,\mu}=\begin{bmatrix} \frac{1}{5+4\lambda+2\mu} & \frac{1}{5+4\lambda+2\mu} & 0 & 0 & 0 & 0 \\ \frac{2}{5+4\lambda+2\mu} & \frac{1}{5+4\lambda+2\mu} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{\lambda}{5+4\lambda+2\mu} & \frac{\frac{1}{2}\lambda}{5+4\lambda+2\mu} & 0 & 0 \\ 0 & 0 & \frac{\lambda}{5+4\lambda+2\mu} & \frac{\frac{1}{2}\lambda}{5+4\lambda+2\mu} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{\mu}{5+4\lambda+2\mu} & \frac{\mu}{5+4\lambda+2\mu} \end{bmatrix}$$

Es fácil ver que para cada λ, las distribuciones condicionadas coinciden con *A* y *B*, por tanto *A* y *B* son compatibles y las distribuciones que determinan no son únicas. Observe que, en esta ocasión, la familia de soluciones es biparamétrica.

De la sección 5 del capítulo anterior, así como de lo anterior se deduce.

Teorema 2.4.4. Dadas dos matrices *A* y *B* en las condiciones habituales.

i.- Todas las posiciones de *N*, pueden unirse mediante, al menos, un camino de recorrido horizontal y/o vertical o mediante el uso de un número finito de posiciones puente.

Supongamos además que las matrices *A* y *B* son compatibles.

ii.- Si existe una posición que puede ser unida mediante, al menos, un camino de recorrido horizontal y/o vertical con el resto de posiciones que tienen probabilidades no nulas, la distribución conjunta es única.

iii.- Si no existe una posición que puede ser unida mediante, al menos, un camino de recorrido horizontal y/o vertical con el resto de posiciones que tienen probabilidades no nulas, entonces es necesario el uso de posiciones puente y la distribución conjunta no es única.

5.- Uso de Teoría de Grafos.

Una lectura detallada de la sección 5 del capítulo uno y de la sección anterior, sugiere una conexión entre el problema unicidad con la Teoría de Grafos, más concretamente, definiremos un grafo bipartito asociado al problema y conectaremos las propiedades de ambos.

Recordemos que un grafo es un par *(V, E),* donde *V* es un conjunto finito a cuyos elementos llamamos vértices y *E*, es un conjunto de pares no ordenados con elementos en *V*, estos pares se denominan aristas y exigiremos que sus vértices sean distintos. Un grafo se dice que es bipartito si *V* se puede descomponer como $V = R_X \cup R_Y,\ R_X \cap R_Y = \varnothing$ de forma que no existe ninguna arista que une elementos de R_X entre sí, ni ninguna arista que une elementos de R_Y entre sí. Un camino entre dos vértices v, w es una sucesión de aristas y vértices de la forma: $v = v_{i_1}, \{v_{i_1}, v_{i_2}\}, v_{i_2}, \{v_{i_2}, v_{i_3}\}, v_{i_3}, \ldots, v_{i_{r-1}}, \{v_{i_{r-1}}, v_{i_r}\}, v_{i_r} = w$ diremos que el camino une a *v* y *w*. Lo denotaremos $\overline{vw}$ y diremos que *v* y *w* están conectados. Un grafo se dice que es conexo si para cualquier par de vértices existe un camino que los une. Es claro que si existe un vértice v^* tal que, para cualquier $w \in V$ existe un camino que los une, el grafo es conexo y recíprocamente.

Dadas dos matrices A y B en las condiciones habituales, definimos su grafo asociado como el grafo bipartito con $R_X = \{x_1, x_2, ..., x_n\}$, $R_Y = \{y_1.y_2, ..., y_m\}$ todos esos valores distintos y el conjunto de aristas $E = \{\{x_i, y_j\} : (i, j) \in N\}$ es decir $\{x_i, y_j\}$ es una arista si y sólo si el elemento (*i*,*j*) es no nulo en A y en B .

Lema 2.5.1. *A* y *B* en las condiciones habituales. Dos posiciones con probabilidades no nulas *i*,*j* y *l*,*k* pueden ser unidas mediante, al menos, un camino de recorrido horizontal y/o vertical si y sólo si existe un camino en el grafo asociado que une x_i con y_k .

Demostración:

$i, j \to i, k \to l, k$ es un camino en la matriz si y sólo si $y_j \{x_i, y_j\} x_i \{x_i, y_k\} y_k \{x_l, y_k\} x_l$ es un camino en el grafo bipartito.

Análogamente, $i, j \to l, j \to l, k$ es un camino en la matriz si y sólo si $x_i \{x_i, y_j\} y_j \{x_l, y_j\} x_l \{x_l, y_k\} y_k$ es un camino en el grafo bipartito.

Para caminos de mayor longitud, basta repetir lo anterior un número finito de veces.

En todo caso tenemos que *i*,*j* y *l*,*k* pueden ser unidos como posiciones de la matriz si y sólo si x_i e y_k están unidos como vértices del grafo.

□

En vista del lema y de los teoremas 3 y 4 de la sección anterior,

Teorema 2.5.1. Dadas dos matrices *A* y *B* en las condiciones habituales y compatibles. La solución es única si y sólo si el grafo asociado es conexo.

Demostración:

Si la solución es única, cada posición de elementos de probabilidad no nula puede unirse mediante caminos adecuados con una posición fija, entonces, en el grafo bipartito puede unirse mediante un camino cualquier y_k con un x_i fijo. Por tanto es conexo.

Si el grafo asociado es conexo, se puede unir y_k con x_i. Sean *i,j* y *l,k* dos posiciones de las matrices con probabilidades no nulas, en virtud del lema 1 existe un camino, en la matriz, que los une, tomando uno de ellos, por ejemplo, *i,j* fijo y haciendo que *l,k* recorran *N*, tendremos una posición que puede ser unida mediante, al menos, un camino de recorrido horizontal y/o vertical con el resto de posiciones que tienen probabilidades no nulas, por tanto la solución es única.

□

Se llama matriz de adyacencia de un grafo a la matriz $s\times s$:

$$K = (k_{ij}) \quad k_{ij} = \begin{cases} 1 & \text{si } \overline{v_i v_j} \in E \\ 0 & \text{si } \overline{v_i v_j} \notin E \end{cases}$$

donde *s* es el cardinal de *V*.

Cuando el grafo es bipartito, que supondremos asociado a un problema de compatibilidad, la matriz de adyacencia puede describirse por bloques, y es de la forma:

$$K = \begin{bmatrix} 0_{n\times n} & H_{n\times m} \\ H^t_{m\times n} & 0_{m\times m} \end{bmatrix}$$

Donde la matriz *H* está definida por:

$$h_{ij} = \begin{cases} 1 & \text{si } \overline{x_i y_j} \in E \\ 0 & \text{si } \overline{x_i y_j} \notin E \end{cases} \qquad i = 1,2,\ldots,n; \;\; j = 1,2,\ldots,m$$

Es conocido que si un grafo con *s* vértices tiene matriz de adyacencia *K*, entonces es conexo si y sólo si $\sum_{r=1}^{s-1} K^r$ es una matriz positiva (ver por ejemplo Noble y Daniels, 1989). En el caso que nos ocupa basta estudiar el comportamiento del bloque 1,2 de la matriz *K*. Por inducción puede probarse sin dificultad que, para *r>0* vale

$$K_{12}^{2r} = 0_{n\times m} \qquad\qquad K_{12}^{2r+1} = \left(HH^t\right)^r H$$

y por tanto debemos comprobar la positividad de

$$\sum_{r=0}^{ent(\frac{m+n}{2})} K_{12}^{2r+1} = \sum_{r=0}^{ent(\frac{m+n}{2})} (HH^t)^r H$$

Donde *ent(.)* es la función parte entera.

Ejemplo 1:

$$A = \begin{bmatrix} \frac{1}{3} & \frac{3}{4} & 0 & 0 \\ \frac{2}{3} & \frac{1}{4} & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} \qquad\qquad B = \begin{bmatrix} \frac{1}{4} & \frac{3}{4} & 0 & 0 \\ \frac{2}{3} & \frac{1}{3} & 0 & 0 \\ 0 & 0 & \frac{2}{3} & \frac{1}{3} \end{bmatrix}$$

Puede demostrarse, usando métodos del capítulo anterior, que son compatibles. Su matriz de adyacencia viene dada por:

$$H = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} \qquad\qquad K = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Calculamos $(HH^t)^r H$ para $r = 0,1,2,3$

r	$(HH^t)^r H$
0	*1 1 0 0* *1 1 0 0* *0 0 1 1*
1	*4 4 0 0* *4 4 0 0* *0 0 2 2*
2	*16 16 0 0* *16 16 0 0* *0 0 4 4*
3	*64 64 0 0* *64 64 0 0* *0 0 8 8*
Suma:	*85 85 0 0* *85 85 0 0* *0 0 15 15*

Como puede observarse, la matriz suma no tiene todas sus entradas positivas, por lo que el grafo asociado al problema no es conexo y existen infinitas distribuciones conjuntas compatibles con las matrices *A* y *B*.

Ejemplo 2:

Consideremos las matrices:

$$A=\begin{bmatrix}1 & 0 & 0.25\\ 0 & 1 & 0.75\end{bmatrix} \qquad B=\begin{bmatrix}0.5 & 0 & 0.5\\ 0 & 0.625 & 0.375\end{bmatrix}$$

Que puede comprobarse que son compatibles.

Sus matrices de adyacencia son

$$H=\begin{bmatrix}1 & 0 & 1\\ 0 & 1 & 1\end{bmatrix} \qquad K=\begin{bmatrix}0 & 0 & 1 & 0 & 1\\ 0 & 0 & 0 & 1 & 1\\ 1 & 0 & 0 & 0 & 0\\ 0 & 1 & 0 & 0 & 0\\ 1 & 1 & 0 & 0 & 0\end{bmatrix}$$

Calculamos $(HH^t)^r H$ para $r = 0,1,2$

r	$(HH^t)^r H$
0	*1 0 1* *0 1 1*
1	*2 1 3* *1 2 3*
2	*5 4 9* *4 5 9*
Suma:	*8 5 13* *5 8 13*

Como puede observarse, la matriz suma tiene todas sus entradas positivas, por lo que el grafo asociado al problema es conexo y existe una única distribución conjunta compatible con las matrices *A* y *B*. Nótese que, para nuestros propósitos, nos podríamos haber detenido en el paso *r=1* pues la matriz resultante ya tiene todos sus elementos positivos.

6.- Método de extensión a rango unidad.

Otro enfoque de este problema es el siguiente. Supóngase que *C* puede ser completada de manera única y de forma que tenga rango unidad, entonces aplicando la descomposición singular obtenemos

$$C = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} \begin{pmatrix} v_1 & v_2 & \dots & v_m \end{pmatrix}$$

donde el valor singular lo hemos incluido en uno de los vectores $\vec{u}$ o $\vec{v}$. Puesto que el vector $\vec{u}$ es autovector asociado al único autovalor no nulo de $C^t C$, y puesto que este autovector es simple, $\vec{u}$ es único salvo proporcionalidad (recordar que se toma no negativo) pero, la marginal que corresponde a $\vec{u}$ está obviamente normalizada y por consiguiente es única, de ahí que la distribución conjunta también. En definitiva se ha establecido el siguiente

Teorema 2.6.1 (Pérez-Villalta, 2000). Sean *A* y *B* dos matrices compatibles. Si la matriz de cocientes puede ser completada de manera única de forma que tenga rango unidad, entonces, la distribución que determinan *A* y *B* es única.

En el lenguaje de Arnold, Castillo y Sarabia (2004) este resultado queda:

Teorema 2.6.1 (Pérez-Villalta, 2000; reformulado). Sean *A* y *B* dos matrices compatibles. Si existe una única extensión positiva de rango uno de la matriz de cocientes entonces, la distribución que determinan *A* y *B* es única.

Ejemplo 1:

El siguiente ejemplo se utilizó en el capítulo anterior (ejemplo 2 de la sección 3)

.

$$A = \begin{bmatrix} \frac{1}{2} & 0 & \frac{3}{8} & \frac{3}{4} \\ \frac{1}{4} & \frac{3}{4} & 0 & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{5}{8} & \frac{1}{4} \end{bmatrix} \qquad B = \begin{bmatrix} \frac{1}{4} & 0 & \frac{3}{8} & \frac{3}{8} \\ \frac{1}{4} & \frac{3}{4} & 0 & 0 \\ \frac{1}{8} & \frac{1}{8} & \frac{5}{8} & \frac{1}{8} \end{bmatrix}$$

Matriz de cocientes (incompleta) es:

$$C = \begin{bmatrix} 2 & * & 1 & 2 \\ 1 & 1 & * & * \\ 2 & 2 & 1 & 2 \end{bmatrix}$$

La extensión positiva de rango uno de *C*,

$$\bar{C} = \begin{bmatrix} 2 & 2 & 1 & 2 \\ 1 & 1 & \frac{1}{2} & 1 \\ 2 & 2 & 1 & 2 \end{bmatrix}$$

y vimos que es compatible. Ahora podemos afirmar, además, que la distribución conjunta es única.

Como ya vimos en la sección precedente no siempre es posible completar *C* de manera única con la condición $rang(C) = 1$. Así pues, llegados a este punto, parece evidente la necesidad de caracterizar dichas matrices. Para ello, y con el fin de aligerar la notación, supondremos que mediante cambios de filas y columnas se han llevado el mayor número posible de elementos desconocidos de C a una submatriz de ella situada en la esquina superior izquierda. Obsérvese que esto no conlleva ninguna pérdida de generalidad supone únicamente una reordenación de los valores del vector aleatorio.

Supondremos además, que todos aquellos valores desconocidos de C que se puedan calcular unívocamente vía la condición $rang(C)=1$, han sido calculados. Esto implica que esos elementos han sido determinados mediante menores de dimensión dos con un único elemento desconocido**.** Estos elementos serán denominados *determinables o calculables.* En definitiva, la forma de la matriz C es

$$C=\begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}$$

Donde, C_{11} es una matriz de orden $r \times s$, con todos sus elementos indeterminados. Además de los elementos de C_{11} **,** es posible que existan otros elementos indeterminados en C_{12} **,**C_{21} y C_{22}. En estas condiciones se tiene

Teorema 2.6.2 (Pérez-Villalta, 2000). Si C_{22} y C_{11} tiene todos sus elementos indeterminados y C_{12} y C_{21} tienen todos sus elementos determinados, entonces, la distribución que determinan A y B no es única.

Demostración. Utilizando la descomposición singular para la matriz C_{12} ($rang(C_{12})$=1) se puede determinar un par de vectores $\begin{pmatrix} u_1 & u_2 & \dots & u_r \end{pmatrix}^t$ y $\begin{pmatrix} v_1 & v_2 & \dots & v_m \end{pmatrix}^t$ de forma que

$$C_{12}=\begin{pmatrix}u_1\\u_2\\\vdots\\u_n\end{pmatrix}\begin{pmatrix}v_1 & v_2 & \dots & v_m\end{pmatrix}$$

pero también, si $\lambda > 0$:

$$C_{12}=\begin{pmatrix}\lambda u_1\\\lambda u_2\\\vdots\\\lambda u_r\end{pmatrix}\left(\frac{1}{\lambda}v_{s+1},\frac{1}{\lambda}v_{s+2},\dots\frac{1}{\lambda}v_M\right)$$

Análogamente utilizando C_{21}

$$C_{21}=\begin{pmatrix}u_{r+1}\\u_{r+2}\\\vdots\\u_n\end{pmatrix}(v_1,v_2,\dots,v_s)$$

Pero también si $\mu > 0$

$$C_{21}=\begin{pmatrix}\mu u_{r+1}\\\mu u_{r+2}\\\vdots\\\mu u_M\end{pmatrix}\left(\frac{1}{\mu}v_1,\frac{1}{\mu}v_2,\dots,\frac{1}{\mu}v_s\right)$$

Ahora los vectores

$$\vec{u} = \left(\lambda u_1, \lambda u_2, \ldots \lambda u_r, \mu u_{r+1}, \mu u_{r+2}, \ldots, \mu u_n\right)^t$$

$$\vec{v} = \left(\frac{1}{\mu} v_1, \frac{1}{\mu} v_2, \ldots \frac{1}{\mu} v_s, \frac{1}{\lambda} v_{s+1}, \frac{1}{\lambda} v_{s+2}, \ldots, \frac{1}{\lambda} v_m\right)^t$$

son proporcionales a las distribuciones marginales y si $\lambda \neq \mu$ es evidente que no son únicos aún normalizándolos. En consecuencia la distribución conjunta no es única.

□

Teorema 2.6.3 (Pérez-Villalta, 2000)**.** Si la distribución conjunta que determinan A y B no es única, la matriz de cocientes C puede escribirse, permutando filas y columnas, como

$$C = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}$$

con C_{11} y C_{22} matrices de elementos indeterminados.

Demostración. Supongamos que por permutación de filas y columnas la matriz C se ha escrito en la forma

$$C = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}$$

con C_{11} de elementos desconocidos. Esta forma de escribir C no es única, nosotros seleccionaremos la siguiente: C_{11} tiene el mayor número posible de columnas de entre todas las particiones de C del tipo anterior y, para ese número de columnas, el mayor número de filas posibles. Supongamos que las dimensiones resultantes son $r \times s$. Probaremos que entonces, C_{22} está formada por elementos desconocidos. En efecto, supongamos, sin pérdida de generalidad, que el elemento $c_{r+1,s+1}$ es conocido (o ha sido calculado en algún paso anterior) entonces, algún elemento de su columna c_{ij} $i = 1,\ldots,r$ $j = s + 1,\ldots,n$ es conocido pues, si no, las dimensiones de C_{11} no serían máximas. Supongamos, sin pérdida de generalidad, que es $c_{r,s+1}$. Análogamente trabajando con la fila $r+1$ tendremos que podemos suponer que $c_{r+1,s}$ es conocido. Así pues son conocidos $c_{r+1,s}$, $c_{r,s+1}$, $c_{r+1,s+1}$, pero entonces como el determinante

$$\begin{vmatrix} c_{r,s} & c_{r,s+1} \\ c_{r+1,s} & c_{r+1,s+1} \end{vmatrix} = 0$$

tenemos que $c_{r,s}$ es conocido, lo que es absurdo.

Obsérvese que C_{11} existe realmente (i.e. $r > 0$ y $s > 0$) pues la distribución conjunta no es única y en consecuencia existe, al menos, un elemento no determinable ya que de no ser así, todos los elementos de C se pueden determinar y en consecuencia, aplicando resultados anteriores la distribución conjunta es única.

□

Veamos algunos ejemplos:

Ejemplo 8: Retomando el ejemplo 3 de la sección 3 en el capítulo anterior, donde ningún elemento desconocido es determinable, y permutando filas y columnas

$$C = \begin{bmatrix} * & \frac{5}{6} & * & \frac{5}{3} \\ 1 & * & \frac{1}{2} & * \\ * & \frac{2}{3} & * & \frac{4}{3} \\ 2 & * & 3 & * \end{bmatrix} \to \begin{bmatrix} * & * & \frac{5}{6} & \frac{5}{3} \\ 1 & \frac{1}{2} & * & * \\ * & * & \frac{2}{3} & \frac{4}{3} \\ 2 & 3 & * & * \end{bmatrix} \to \begin{bmatrix} * & * & \frac{5}{6} & \frac{5}{3} \\ * & * & \frac{2}{3} & \frac{4}{3} \\ 1 & \frac{1}{2} & * & * \\ 2 & 3 & * & * \end{bmatrix}$$

con lo que, aplicando resultados anteriores, se concluye que, aún siendo compatibles, la distribución conjunta no es única.

El siguiente ejemplo muestra la utilización de la descomposición singular en submatrices para obtener las distribuciones marginales.

Ejemplo 9. Las siguientes matrices pueden considerarse como distribuciones condicionadas de un vector aleatorio finito:

$$A = \begin{bmatrix} 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$B = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ \frac{1}{2} & 1 & \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 \end{bmatrix}$$

La matriz de cocientes viene dada por:

$$C=\begin{bmatrix} * & * & * & * & \frac{1}{2} & \frac{1}{2} & * \\ \frac{2}{3} & \frac{1}{3} & \frac{2}{3} & * & * & * & * \\ * & * & * & \frac{1}{2} & * & * & \frac{1}{2} \\ 1 & * & 1 & * & * & * & * \end{bmatrix}$$

Intercambiando la 2ª y 3ª fila y las columnas 5ª y 7ª obtenemos:

$$C\rightarrow\begin{bmatrix} * & * & * & * & * & \frac{1}{2} & \frac{1}{2} \\ * & * & * & \frac{1}{2} & \frac{1}{2} & * & * \\ \frac{2}{3} & \frac{1}{3} & \frac{2}{3} & * & * & * & * \\ 1 & * & 1 & * & * & * & * \end{bmatrix}$$

El elemento $c_{4,2}$ es determinable mediante la condición $rang(C)=1$, obteniéndose que $c_{42}=\frac{1}{2}$. Los elementos conocidos forman menores de orden dos nulo y, en consecuencia, las matrices *A* y *B* se pueden considerar distribuciones condicionadas si *C* se puede completar de forma que tenga rango unidad. La matriz $C_{1,2}$ reproduce la estructura por cajas que estudiamos en el teorema 3 y además *C* también la reproduce.

Completaremos en primer lugar las submatrices de la submatriz $C_{1,2}$:

La submatriz que ocupa el lugar 1,2 en $C_{1,2}$ es $(\frac{1}{2},\frac{1}{2})$. Aplicando la descomposición en valores singulares (como siempre incluimos el valor singular en los vectores) obtenemos:

$$\begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}^t = \vec{u}'\vec{v}'^t$$

donde podemos tomar $\vec{u}' = \lambda,\ \vec{v}' = \begin{bmatrix} \frac{1}{2\lambda} \\ \frac{1}{2\lambda} \end{bmatrix}$ para cualquier número real λ > 0.

Observamos que $\vec{u}'$ es la componente u_1 del vector $\vec{u}$ que resuelve el problema en la matriz C y $\vec{v}'$ contiene los componentes v_6, v_7 del vector $\vec{v}$ que resuelve el problema en la matriz C, así se tiene:

$$u_1 = \lambda \qquad v_6 = \frac{1}{2\lambda} \qquad v_7 = \frac{1}{2\lambda}$$

Trabajando ahora con la submatriz que ocupa el lugar 2,1 en $C_{1,2}$ obtenemos,

$$\begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}^t = \mu \begin{bmatrix} \frac{1}{2\mu} \\ \frac{1}{2\mu} \end{bmatrix}^t$$

y como antes se tiene

$$u_2 = \mu \qquad v_4 = \frac{1}{2\mu} \qquad v_5 = \frac{1}{2\mu}$$

con estos vectores es posible reconstruir $C_{1,2}$.

$$C_{1,2} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} (v_4 v_5 v_6 v_7) = \begin{bmatrix} \lambda \\ \mu \end{bmatrix} \left(\frac{1}{2\mu}, \frac{1}{2\mu}, \frac{1}{2\lambda}, \frac{1}{2\lambda} \right) = \begin{bmatrix} \frac{\lambda}{2\mu} & \frac{\lambda}{2\mu} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{\mu}{2\lambda} & \frac{\mu}{2\lambda} \end{bmatrix}$$

En consecuencia, la matriz C adopta la forma

$$C=\begin{bmatrix} * & * & * & \frac{\lambda}{2\mu} & \frac{\lambda}{2\mu} & \frac{1}{2} & \frac{1}{2} \\ * & * & * & \frac{1}{2} & \frac{1}{2} & \frac{\mu}{2\lambda} & \frac{\mu}{2\lambda} \\ \frac{2}{3} & \frac{1}{3} & \frac{2}{3} & * & * & * & * \\ 1 & \frac{1}{2} & 1 & * & * & * & * \end{bmatrix}$$

Para completar $\vec{u}$ y $\vec{v}$ necesitamos u_3, u_4, v_1, v_2, y v_3, los obtendremos de la descomposición en valores singulares de la matriz $C_{2,1}$.

$$\begin{bmatrix} \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ 1 & \frac{1}{2} & 1 \end{bmatrix}=\begin{bmatrix} \frac{1}{3} \\ \frac{1}{2} \end{bmatrix}(2,1,2)=\begin{bmatrix} \frac{1}{3}\omega \\ \frac{1}{2}\omega \end{bmatrix}\left(\frac{2}{\omega},\frac{1}{\omega},\frac{2}{\omega}\right)$$

en consecuencia, $u_3=\frac{1}{3}\omega,\ u_4=\frac{1}{2}\omega,\ v_1=\frac{2}{\omega},\ v_2=\frac{1}{\omega}$, y $v_3=\frac{2}{\omega}$

$$\vec{u}^t=\left(\lambda,\mu,\frac{1}{3}\omega,\frac{1}{2}\omega\right),$$

$$\vec{v}^t=\left(\frac{2}{\omega},\frac{1}{\omega},\frac{2}{\omega},\frac{1}{2\mu},\frac{1}{2\mu},\frac{1}{2\lambda},\frac{1}{2\lambda}\right)$$

Así se obtiene:

$$C=\begin{bmatrix}\lambda\\ \mu\\ \frac{1}{3}\omega\\ \frac{1}{2}\omega\end{bmatrix}\left(\frac{2}{\omega},\frac{1}{\omega},\frac{2}{\omega},\frac{1}{2\mu},\frac{1}{2\mu},\frac{2}{2\lambda},\frac{2}{2\lambda}\right)=\begin{bmatrix}\frac{2\lambda}{\omega} & \frac{\lambda}{\omega} & \frac{2\lambda}{\omega} & \frac{\lambda}{2\mu} & \frac{\lambda}{2\mu} & \frac{1}{2} & \frac{1}{2}\\ \frac{2\mu}{\omega} & \frac{\mu}{\omega} & \frac{2\mu}{\omega} & \frac{1}{2} & \frac{1}{2} & \frac{\mu}{2\lambda} & \frac{\mu}{2\lambda}\\ \frac{2}{3} & \frac{1}{3} & \frac{2}{3} & \frac{\omega}{6\mu} & \frac{\omega}{6\mu} & \frac{\omega}{6\lambda} & \frac{\omega}{6\lambda}\\ 1 & \frac{1}{2} & 1 & \frac{\omega}{4\mu} & \frac{\omega}{4\mu} & \frac{\omega}{4\lambda} & \frac{\omega}{4\lambda}\end{bmatrix}$$

Para obtener distribuciones marginales, basta con normalizar los vectores $\vec{u}$ y $\vec{v}$, sin embargo, aún en ese caso, dando valores positivos a λ, μ *y* ω obtenemos distintas distribuciones compatibles, con lo que se tiene la no unicidad.

Song et al. (2010) aplican la descomposición en bloques irreducibles al estudio de la unicidad del problema de unicidad,

Teorema 2.6.4 (Song et al, 2010). Sea *C* la matriz de cocientes de dos matrices $n\times m$ *A* y *B* compatibles. Son equivalentes,

i.- *M*=1 para cualquier descomposición por bloque irreducibles (DBI) $T(C)=Diag(T_1,T_2,\ldots,T_M)$

ii.- *C* es irreducible.

iii.- *C* tiene una única extensión positiva.

iv.- Existe una única distribución de conjunta con matriz de cocientes *C* y condicionadas *A* y *B.*

Además, asocia biunívocamente a cada extensión positiva de rango unidad, una distribución conjunta que tiene a *A* y *B* como condicionadas.

Teorema 2.6.5 (Song et al, 2010). Sea *C* la matriz de cocientes de dos matrices $n\times m$, *A* y *B* en las condiciones habituales y compatibles. Denotamos **C** al conjunto de todas las extensiones positivas de rango uno de C y **F** el conjunto de todas las posibles distribuciones de probabilidad conjuntas que tienen a A y B por condicionadas. Entonces, la aplicación $\Phi:\mathbf{C}\rightarrow\mathbf{F}$ definida por $\Phi(\bar{C})=f$ con $\bar{C}=\vec{u}\vec{v}^t$ y $f(x_i,y_j)=\dfrac{b_{ij}u_i}{\sum_{i=1}^{n}u_i}$ para todo $(i,j)\in N$ está bien definida y es biyectiva.

En virtud del resultado anterior, el cardinal de las extensiones positivas de rango unidad de una matriz de cocientes C es el mismo del de las distribuciones que tienen a A y B como condicionadas tenemos, por tanto, el siguiente,

Corolario 2.6.1. Sea C la matriz de cocientes de dos matrices $n\times m$, A y B en las condiciones habituales. El cardinal del conjunto de extensiones positivas de rango uno es, cero, uno o infinito (no numerable).

En el mismo trabajo, proporcionan un método para calcular todas las distribuciones conjuntas que tienen a A y B como condicionadas

Teorema 2.6.6 (Song et al, 2010). Sea C la matriz de cocientes de dos matrices $n\times m$ A y B compatibles, y $T(C)=Diag(T_1,T_2,...,T_M)$ una descomposición por bloque irreducibles (DBI) de C para alguna matrices de permutación E y F. Para cada $s=1,2,\ldots,M$, consideremos $\bar{T}_s=\vec{u}_s\vec{v}_s^t$ una extensión positiva de rango unidad de T_s. Para cada $\vec{k}=(k_1,k_2,...,k_M)^t\in\mathbb{R}^M$ con $k_s>0$ y sea $\vec{u}_{\vec{k}}^t=(k_1\vec{u}_1^t,k_2\vec{u}_2^t,...,k_M\vec{u}_M^t)$

Definamos ahora $\vec{p}_{\vec{k}}=E^t\vec{u}_{\vec{k}}$ y $f_{\vec{k}}(x_i,\mathrm{y}_j)=\dfrac{b_{ij}p_{\vec{k}i}}{\sum_{i=1}^{n}p_{\vec{k}i}}$ donde $\vec{p}_{\vec{k}}^t=(p_{\vec{k}1},p_{\vec{k}2},...,p_{\vec{k}n})$.

Entonces, la familia $\left\{f_{\vec{k}}:\vec{k}^t=(k_1,k_2,...,k_M),\ k_s>0\ 1\le s\le M\right\}$ es el conjunto de todas las posibles distribuciones de probabilidad conjunta de (X,Y) para una C dada.

Nótese que este método requiere *M* parámetros, sin embargo, a nuestro parecer, únicamente se requieren *M*-1, pues basta observar la evolución de la extensión positiva cuando se van introduciendo parámetros en *T(C)*:

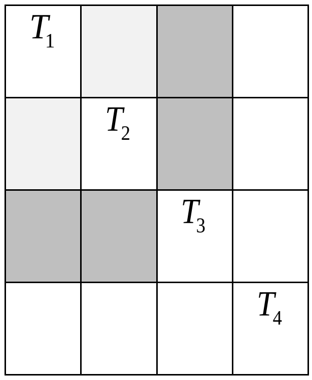
Consideremos la siguiente DBI de *C*, para *M* =4

T_1			
	T_2		
		T_3	
			T_4

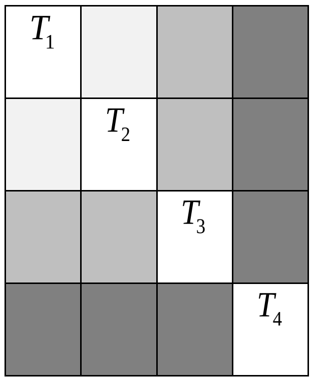
Supongamos que introducimos un parámetro en la esquina inferior izquierda del bloque 1,2; entonces, podemos completar los bloques 1,2 y 2,1

T_1			
	T_2		
		T_3	
			T_4

Si introducimos un nuevo parámetro en la esquina inferior izquierda del bloque 2,3; entonces, podemos completar los bloques que marcamos,

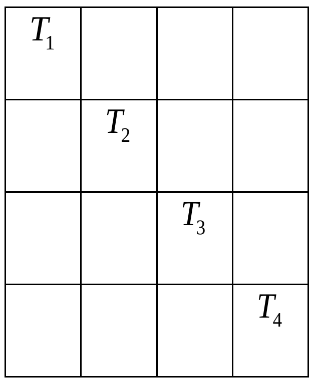


Haciendo lo mismo con el bloque 3,4 completamos la extensión positiva

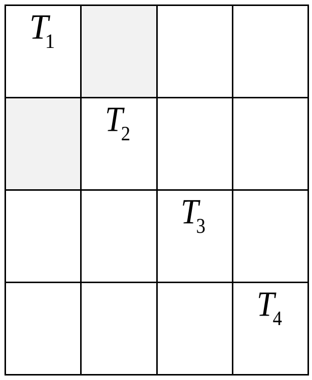


Lo que supone tres parámetros y constituyen todas las extensiones posibles.

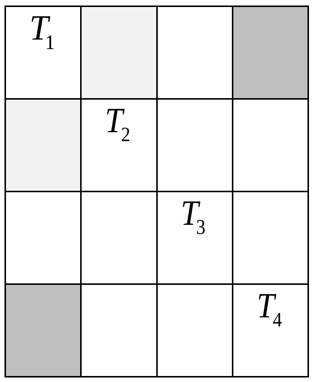
A título de ejemplo mostraremos otra estrategia de extensión de *T(C)*.



Comenzando, como anteriormente introduciendo un parámetro en la esquina inferior izquierda del bloque 1,2



Introduciendo un nuevo parámetro en la esquina inferior izquierda del bloque 1,4



Aún se necesita otro parámetro para completar la extensión y basta con uno, cualquier situación nos lleva a completar la extensión. Es fácil, aunque tedioso, probar que ocurre lo mismo para cualquier *M*.

7.- Más sobre grafos.

Un grafo (V',E') se dice que es un subgrafo de (V,E) si y sólo si $V'\subseteq V$ y $E'\subseteq E$

Una componente conexa de un vértice *v*, es el subgrafo que se obtiene tomando todos los vértices conectados con *v* y todas las aristas que conectan con dichos vértices. Una componente conexa de un grafo (V,E) es la componente conexa de un vértice. Una componente conexa de un grafo (V,E) es, en sí misma conexa. Un grafo es conexo si sólo tiene una componente conexa.

Consideremos una DBI de la matriz *C* en un problema de compatibilidad y sean *H* y *K* las matrices de adyacencia del grafo bipartito asociado, supongamos además, que están en la forma dictada por *T(C)* y que se ha extendido cada bloque diagonal, es decir a partir de

$$T(C)=\begin{bmatrix} T_1 & * & \cdots & * \\ * & T_2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \cdots & T_M \end{bmatrix}$$

Construimos

$$\bar{T}(C)=\begin{bmatrix} \bar{T}_1 & * & \cdots & * \\ * & \bar{T}_2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \cdots & \bar{T}_M \end{bmatrix}$$

Y a partir de ella la matriz de adyacencia $(n\times m)$

$$H=\begin{bmatrix} H_1 & * & \cdots & * \\ * & H_2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \cdots & H_M \end{bmatrix}$$

Observemos que al haberse construido *H* a partir de la extensión de cada bloque de *T(C),* sus bloques diagonales están formado por unos, por tanto, todos los vértices que están involucrados en un bloque están conectados entre sí (subgrafo completo) y por tanto el subgrafo que definen es conexo. Esto se puede repetir con todos los bloque diagonales de *H*. Y, en vista de la forma de *H*, no existe ningún camino entre pares de vértices que conciernen a cada bloque, en consecuencia cada bloque define una componente conexa del grafo asociado al problema de compatibilidad.

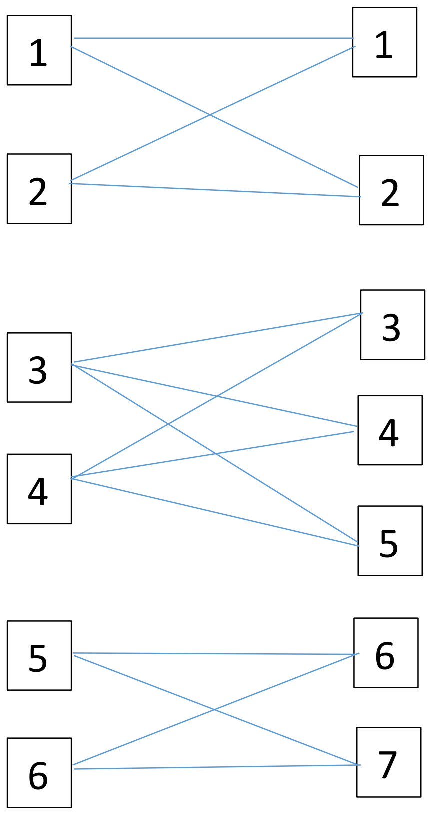
Para conectar dos componentes conexas, basta añadir una arista (arista puente), así para conectar el grafo asociado a un problema de compatibilidad con *M* componentes conexas (*M* bloques en *T(C))* debemos añadir *M-1* aristas. Cada arista añadida se corresponde con un parámetro en la extensión de *T(C)* por tanto, las soluciones de un problema de compatibilidad con $T(C) = Diag(T_1, T_2, ..., T_M)$ depende de *M-1* parámetros.

Ejemplo 1:

Consideremos la matriz *T(C)* del ejemplo 4, sección 3 en el capítulo anterior.

$$T(C) = \begin{bmatrix} \lceil 1 & 2 & * & * & * & * & * \\ 0.5 & 1\rfloor & * & * & * & * & * \\ * & * & \lceil 2 & 2 & 1 & * & * \\ * & * & 2 & * & 1\rfloor & * & * \\ * & * & * & * & * & \lceil 1 & 1 \\ * & * & * & * & * & 1 & 1\rfloor \end{bmatrix}$$

Completando los bloques,

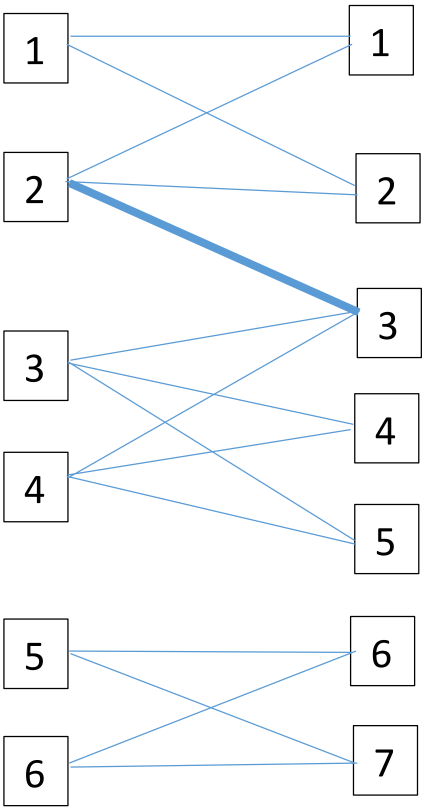


$$\bar{T}(C)=\begin{bmatrix} 1 & 2 & * & * & * & * & * \\ 0.5 & 1 & * & * & * & * & * \\ * & * & 2 & 2 & 1 & * & * \\ * & * & 2 & 2 & 1 & * & * \\ * & * & * & * & * & 1 & 1 \\ * & * & * & * & * & 1 & 1 \end{bmatrix}$$

La matriz de adyacencia del problema es

$$H=\begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

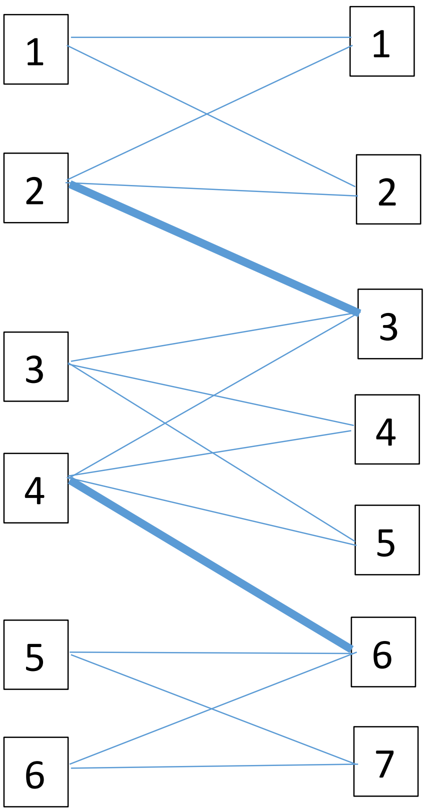
En el grafo asociado hay tres componentes conexas que corresponden a $\{x_1, x_2, y_1, y_2\}$ que son los vértices del primer bloque, las aristas son todas las posibles, el segundo bloque tiene vértices $\{x_3, x_4, y_3, y_4, y_5\}$, y en el tercer bloque $\{x_5, x_6, y_6, y_7\}$, con el mismo comentario sobre las aristas. Para conectar los dos primeros bloque podemos usar la arista $\{x_2, y_3\}$ que corresponde a la posición 2,3 en $\bar{T}(C)$ y se traduce en



$$\bar{T}(C)=\begin{bmatrix} 1 & 2 & * & * & * & * & * \\ 0.5 & 1 & \lambda & * & * & * & * \\ * & * & 2 & 2 & 1 & * & * \\ * & * & 2 & 2 & 1 & * & * \\ * & * & * & * & * & 1 & 1 \\ * & * & * & * & * & 1 & 1 \end{bmatrix}$$

$$\bar{T}(C)=\begin{bmatrix} 1 & 2 & 2\lambda & 2\lambda & \lambda & * & * \\ 0.5 & 1 & \lambda & \lambda & \frac{\lambda}{2} & * & * \\ \frac{1}{\lambda} & \frac{2}{\lambda} & 2 & 2 & 1 & * & * \\ \frac{1}{\lambda} & \frac{2}{\lambda} & 2 & 2 & 1 & * & * \\ * & * & * & * & * & 1 & 1 \\ * & * & * & * & * & 1 & 1 \end{bmatrix}$$

Y para conectar los bloques dos y tres usamos la arista $\{x_4, y_6\}$ que corresponde a posición 4,6 de $\bar{T}(C)$ y se traduce en



$$\bar{T}(C)=\begin{bmatrix} 1 & 2 & * & * & * & * & * \\ 0.5 & 1 & \lambda & * & * & * & * \\ * & * & 2 & 2 & 1 & * & * \\ * & * & 2 & 2 & 1 & \mu & * \\ * & * & * & * & * & 1 & 1 \\ * & * & * & * & * & 1 & 1 \end{bmatrix}$$

$$\bar{T}(C)=\begin{bmatrix} 1 & 2 & 2\lambda & 2\lambda & \lambda & * & * \\ 0.5 & 1 & \lambda & \lambda & \frac{\lambda}{2} & * & * \\ \frac{1}{\lambda} & \frac{2}{\lambda} & 2 & 2 & 1 & * & * \\ \frac{1}{\lambda} & \frac{2}{\lambda} & 2 & 2 & 1 & \mu & * \\ * & * & * & * & * & 1 & 1 \\ * & * & * & * & * & 1 & 1 \end{bmatrix}$$

$$\bar{T}(C)=\begin{bmatrix} 1 & 2 & 2\lambda & 2\lambda & \lambda & \lambda\mu & \lambda\mu \\ 0.5 & 1 & \lambda & \lambda & \frac{\lambda}{2} & \frac{\lambda\mu}{2} & \frac{\lambda\mu}{2} \\ \frac{1}{\lambda} & \frac{2}{\lambda} & 2 & 2 & 1 & \mu & \mu \\ \frac{1}{\lambda} & \frac{2}{\lambda} & 2 & 2 & 1 & \mu & \mu \\ \frac{1}{\lambda\mu} & \frac{2}{\lambda\mu} & \frac{2}{\mu} & \frac{2}{\mu} & \frac{1}{\mu} & 1 & 1 \\ \frac{1}{\lambda\mu} & \frac{2}{\lambda\mu} & \frac{2}{\mu} & \frac{2}{\mu} & \frac{1}{\mu} & 1 & 1 \end{bmatrix}$$

Con esas aristas puente es suficiente para conectar el grafo, y con esos dos parámetros es suficiente para encontrar una extensión positiva de rango uno para $\bar{T}(C)$. Note que el grafo bipartito obtenido es conexo.

Note la relación entre el número de bloque y el número de parámetros.

Nº de parámetros = Nº de Bloques -1

Y puesto que el número de bloque es el número de componentes conexas,

Nº de parámetros = Nº de comp. conexas -1

8.- Uso de Cadenas de Markov.

Una vía alternativa en el estudio de la unicidad es el uso de las cadenas de Markov tal y como proponen Arnold y Press (1989). Este enfoque surge de la siguiente observación:

$$\vec{\alpha} = A\vec{\beta} \qquad \vec{\beta} = B^t\vec{\alpha}$$

Combinando lo anterior

$$\vec{\alpha} = A\vec{\beta} = AB^t\vec{\alpha}$$

puesto que AB^t es una matriz estocástica por columnas, al serlo A y B^t, AB^t puede ser considerada como la matriz de transición de una cadena de Markov de n estados y, consecuentemente, $\vec{\alpha}$ se interpreta como una distribución invariante o estacionaria de la cadena.

Arnold, y Press (1989) observan que si A y por tanto B tienen sus elementos no nulos se tiene la unicidad, resultado que obtuvimos anteriormente. Por su parte, Nervole y Press (1986) prueban que la distribución estacionaria es única si y solo si existe $k>0$ tal que $\sum_{i=1}^{n}\left(BA^t\right)^{k+i}$ tiene todos sus elementos positivos.

Capítulo 3

Incompatibilidad

En este capítulo se estudia el caso de incompatibilidad, concretamente se dan métodos para obtener soluciones aproximadas, en algún sentido que se precisará, al problema que nos ocupa. Se estudia la ε-compatibilidad y la compatibilidad aproximada o cuasi compatibilidad. Para finalizar, utilizaremos la descomposición en valores singulares para encontrar soluciones aproximadas al problema, esta forma de proceder se basa en el método, ya estudiado, de extensión a rango unidad.

1.- Uso de sistemas lineales con restricciones en compatibilidad.

Arnold, Castillo y Sarabia (2002) utilizan sistemas lineales para calcular las probabilidades conjuntas o las marginales asociadas a las matrices *A* y *B*.

Método A. Buscando las probabilidades conjuntas, $p_{ij} \geq 0 \ \forall (i, j) \in N$

$$p_{ij} - a_{ij}\sum_{k=1}^{n} p_{kj} = 0, \quad p_{ij} - b_{ij}\sum_{k=1}^{m} p_{ik} = 0, \quad \forall (i,j) \in N \qquad \sum_{(i,j)\in N} p_{ij} = 1$$

Número de ecuaciones: $2card(N)+1$, número de incognitas $card(N)$

Método B. Buscando los dos vectores de probabilidades marginales, $\vec{\alpha}$ y $\vec{\beta}$ no negativos.

$$\beta_j a_{ij} - \alpha_i \beta_j = 0, \quad \forall (i,j) \in N \qquad \sum_{i=1}^{n} \alpha_i = 1 \qquad \sum_{j=1}^{m} \beta_j = 1$$

Número de ecuaciones: $card(N)+2$, número de incognitas $n+m$

Método C. Buscando un único vector de probabilidad marginal, $\vec{\alpha}$ no negativo.

$$a_{ij}\sum_{k=1}^{n} \alpha_k b_{kj} - \alpha_i b_{ij} = 0, \quad \forall (i,j) \in N \qquad \sum_{i=1}^{n} \alpha_i = 1$$

Número de ecuaciones: $card(N)+1$, número de incognitas n

Nótese que todos los métodos requieren la identificación de soluciones no negativas.

Teorema 3.1.1 (Arnold, Castillo y Sarabia, 2002). Las probabilidades conjuntas y marginales obtenidas por los métodos A, B y C coinciden.

Seguimos a Castillo et al. (2002)

Definición. Sea *A* una matriz real de dimensiones $n\times m$. El cono poliédrico convexo engendrado por *A*, que denotamos por $\pi(A)$, es el conjunto de todos los vectores de $\mathbb{R}^n$ que pueden expresarse como combinación lineal, con coeficientes no negativos, de las columnas de *A*.

Definición. Para cada cono poliédrico convexo $\pi(A)$, definimos su dual o cono polar $\Omega(\pi(A))$ como el conjunto

$$\Omega(\pi(A)) = \left\{ \vec{u} \in \mathbb{R}^n : \vec{v}^t \vec{u} \leq 0, \ \forall v \in \pi(A) \right\} = \left\{ \vec{u} \in \mathbb{R}^n : \mathrm{A}^t \vec{u} \leq 0 \right\}$$

Se puede comprobar que $\Omega(\pi(A))$ es, en sí mismo, un cono poliédrico convexo. Así se puede ver como generado por algún conjunto finito de vectores de $\mathbb{R}^n$. En cualquier cono π pueden existir algunos vectores tales que $\vec{u} \in \pi$ y $-\vec{u} \notin \pi$ y algunos vectores tales que $\vec{u}, -\vec{u} \in \pi$.

Un cono polar admite la representación $\Omega(\pi(A)) = \rho(V) + \pi(W)$ donde $\rho(V)$ denota el espacio lineal generado por las columnas de una matriz de dimensiones $n \times k_1$, V y $\pi(W)$ es el cono generado por las columnas de una matriz de dimensiones $n \times k_2$, *W*. Las matrices *V* y *W* se denominan generadores del cono $\Omega(\pi(A))$. Juegan un papel crucial en la solución de la clase de soluciones del sistema lineal de ecuaciones e inecuaciones. El cálculo de estas matrices no es fácil, ver Castillo et al (2002) para detalles y algoritmos.

Teorema 3.1.2. Sea *C* una matriz real de dimensiones $m \times n$, $\vec{a} \in \mathbb{R}^m$ y $\vec{x}$ un vector columna de *n* incógnitas. El sistema de ecuaciones lineales sujeto a $\vec{x} \geq \vec{0}$ tiene alguna solución si y sólo si $V^t\vec{a} = \vec{0}$ y $W^t\vec{a} \leq \vec{0}$ donde *V* y *W* son los generadores de $\Omega(\pi(C))$.

Los sistemas que hemos considerado métodos A, B y C tienen la particularidad de que el vector de términos independientes es de la forma $\vec{a}^t = (0\ 0 \ldots 0\ 1)$ para A y C $\vec{a}^t = (0\ 0 \ldots 1\ 1)$ para B. Se tiene entonces,

El método A, tiene solución si la última fila de *V* es nula y la última fila de *W* es no positiva.

El método B, tiene solución si la suma de las dos últimas *V* es nula y la misma suma en las filas de *W* es no positiva.

El método C, tiene solución si la última fila de *V* es nula y la última fila de *W* es no positiva.

Ejemplo (Rodrigues Correia, 2007):

Consideremos las matrices,

$$A=\begin{bmatrix} 0 & 0 & \frac{1}{3} & \frac{1}{2} \\ 0 & 0 & \frac{1}{3} & \frac{1}{4} \\ 1 & 1 & \frac{1}{3} & \frac{1}{4} \end{bmatrix} \quad \text{y} \quad B=\begin{bmatrix} 0 & 0 & \frac{1}{3} & \frac{2}{3} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{2}{5} & \frac{2}{5} & \frac{1}{5} & \frac{1}{5} \end{bmatrix}$$

Utilizando el método C, para $\vec{\alpha} \geq \vec{0}$,

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ -\frac{2}{9} & \frac{1}{6} & \frac{1}{15} \\ -\frac{1}{3} & \frac{1}{4} & \frac{1}{10} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ \frac{1}{9} & -\frac{1}{3} & \frac{1}{15} \\ \frac{1}{6} & -\frac{3}{8} & \frac{1}{20} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ \frac{1}{9} & \frac{1}{6} & -\frac{2}{15} \\ \frac{1}{6} & \frac{1}{8} & -\frac{3}{20} \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

Para determinar los generadores V y W del cono polar, utiliza el algoritmo gamma (Castillo et al, 1999, Castillo et al, 2002) y obtiene,

$$V=\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -\frac{3}{2} & 0 & 0 & 0 & \frac{1}{4} & 0 & 0 & 1 & \frac{5}{4} & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad W=\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 12 & 2 & -8 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 3 & 13 & -7 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix}$$

La última fila de V es nula, y la última fila de W es negativa, por tanto, las matrices son compatibles.

Resolviendo el sistema se obtiene que $\vec{\alpha}=\begin{bmatrix} \frac{3}{10} \\ \frac{1}{5} \\ \frac{1}{2} \end{bmatrix}$

y premultiplicando por B,

$$M=\begin{bmatrix} 0 & 0 & \frac{1}{10} & \frac{1}{5} \\ 0 & 0 & \frac{1}{10} & \frac{1}{10} \\ \frac{1}{10} & \frac{1}{5} & \frac{1}{10} & \frac{1}{10} \end{bmatrix}$$

Nótese la dificultad del método frente a otros estudiados.

Pueden verse más ejemplos en Arnold y Gokhale (1998) y Arnold, Castillo y Sarabia (1999 y 2002).

2.- ε-compatibilidad.

En esta sección se estudia un concepto alternativo al de la compatibilidad, siguiendo a Arnold, Castillo y Sarabia (1999, 2002). La idea que subyace es la posibilidad de aceptar una suerte de compatibilidad aproximada, en el sentido que se precisará. Supongamos que se tiene una matriz $n \times m$, W que indica la importancia de la acuracidad relativa del valor de cada elemento p_{ij} de M en la determinación de p_{ij}, de forma que un w_{ij} pequeño indica que se desea una alta acuracidad en el cálculo de p_{ij}, y un w_{ij} grande indica que es suficiente una baja acuracidad en dicho cálculo. El método consiste en utilizar los sistemas de ecuaciones de la sección anterior y resolverlos de forma aproximada. Se tienen entonces, las tres opciones siguientes.

Opción 1. Usando el método A, encontrar la matriz de probabilidad conjunta $M = (p_{ij})$ que verifica,

$$p_{ij} \geq 0 \;\; \forall (i,j) \in N$$

$$\left| p_{ij} - a_{ij} \sum_{k=1}^{n} p_{kj} \right| \leq \varepsilon w_{ij}, \quad \forall (i,j) \in N$$

$$\left| p_{ij} - b_{ij} \sum_{k=1}^{m} p_{ik} \right| \leq \varepsilon w_{ij}, \quad \forall (i,j) \in N$$

$$\sum_{(i,j) \in N} p_{ij} = 1$$

$$p_{ij} \geq 0 \;\; \forall (i,j) \in N$$

Opción 2. Usando el método B, encontrar los vectores de probabilidad marginal $\vec{\alpha}$ y $\vec{\beta}$ que verifican,

$$\left|\beta_j a_{ij} - \alpha_i \beta_j\right| \leq \varepsilon w_{ij}, \quad \forall (i,j) \in N$$

$$\sum_{i=1}^{n} \alpha_i = 1$$

$$\sum_{j=1}^{m} \beta_j = 1$$

$$\alpha_i, \beta_j \geq 0, \; i = 1,...,n; j = 1,...,m$$

Opción 3. Usando el método C, encontrar un único vector de probabilidad marginal, $\vec{\alpha}$ no negativo.

$$\left|a_{ij} \sum_{k=1}^{n} \alpha_k b_{kj} - \alpha_i b_{ij}\right| \leq \varepsilon w_{ij}, \quad \forall (i,j) \in N$$

$$\sum_{i=1}^{n} \alpha_i = 1$$

$$\alpha_i \geq 0, \; i = 1,...,n$$

Suponga conocida la matriz *W* de elementos w_{ij} .

Definición. Sean *A* y *B* dos matrices en las condiciones habituales. Se dice que *A* y *B* son ε_1-compatibles, si y sólo si los sistemas definidos en las opciones 1, 2 o 3 tienen solución para todo $\varepsilon \geq \varepsilon_1$ y no la tiene para $\varepsilon < \varepsilon_1$. Es decir ε_1 es el menor ε para el que los sistemas de inecuaciones anteriores tienen solución.

Note que *A* y *B* son compatibles si y solo si son 0-compatibles.

Esto no agota las posibilidades de opciones, Arnold, Castillo y Sarabia (2002) proporcionan otras posibilidades.

Para analizar la ε-compatibilidad caben diversas metodologías, algunas involucran conos polares lo que dificulta enormemente la práctica. Aquí utilizaremos otra, basada en programación lineal. Consideremos ε como una variable y decidamos la opción a utilizar. Resolver

$$\min \varepsilon \text{ Sujeto a la opción elegida.}$$

Así, si seleccionamos la opción tres, y $W = 1_{n\times m}$ el programa es:

$$\min \varepsilon$$

S.A.

$$a_{ij}\sum_{k=1}^{n}\alpha_k b_{kj} - \alpha_i b_{ij} \leq \varepsilon,$$

$$\alpha_i b_{ij} - a_{ij}\sum_{k=1}^{n}\alpha_k b_{kj} \leq \varepsilon, \ \forall (i,j) \in N$$

$$\sum_{i=1}^{n}\alpha_i = 1$$

$$\alpha_i \geq 0, \ i = 1,...,n$$

Ejemplo:

Consideremos las matrices:

$$A = \begin{bmatrix} \frac{1}{2} & 0 & \frac{3}{5} \\ \frac{1}{2} & \frac{1}{5} & \frac{2}{5} \end{bmatrix} \qquad B = \begin{bmatrix} \frac{3}{5} & 0 & \frac{2}{5} \\ \frac{1}{5} & \frac{2}{5} & \frac{2}{5} \end{bmatrix}$$

Estas matrices son incompatibles, pues la matriz, incompleta, de cocientes,

$$C=\begin{bmatrix}\frac{5}{6} & * & \frac{3}{2}\\ \frac{5}{2} & \frac{5}{2} & 1\end{bmatrix}$$

no puede completarse de forma que tenga rango unidad pues

$$\begin{vmatrix}\frac{5}{6} & \frac{3}{2}\\ \frac{5}{2} & 1\end{vmatrix}\neq 0$$

Si denotamos $\vec{\alpha}=\begin{bmatrix}x_1\\ x_2\end{bmatrix}$, el conjunto de restricciones viene dado por:

$$x_1+x_2=1,\qquad x_1\geq 0,\qquad x_2\geq 0$$

$$\left|\frac{1}{2}\left(x_1\frac{3}{5}+x_2\frac{1}{5}\right)-x_1\frac{3}{5}\right|\leq\varepsilon$$

$$\left|\frac{3}{5}\left(x_1\frac{2}{5}+x_2\frac{2}{5}\right)-x_1\frac{2}{5}\right|\leq\varepsilon$$

$$\left|\frac{1}{2}\left(x_1\frac{3}{5}+x_2\frac{1}{5}\right)-x_2\frac{1}{5}\right|\leq\varepsilon$$

$$\left|\left(x_1 0 + x_2 \frac{2}{5}\right) - x_2 \frac{2}{5}\right| \le \varepsilon$$

$$\left|\frac{2}{5}\left(x_1 \frac{2}{5} + x_2 \frac{2}{5}\right) - x_2 \frac{2}{5}\right| \le \varepsilon$$

Operando, se reduce a

$$x_1 + x_2 = 1, \qquad x_1 \ge 0, \qquad x_2 \ge 0$$

$$-3x_1 + x_2 \le 10\varepsilon \qquad 3x_1 - x_2 \le 10\varepsilon$$

$$-4x_1 + 6x_2 \le 25\varepsilon \qquad 4x_1 - 6x_2 \le 25\varepsilon$$

$$3x_1 - x_2 \le 10\varepsilon \qquad -3x_1 + x_2 \le 10\varepsilon$$

$$\varepsilon \ge 0$$

$$4x_1 - 6x_2 \le 25\varepsilon \qquad -4x_1 + 6x_2 \le 25\varepsilon$$

En definitiva, se tiene el programa lineal:

$$\min \ \varepsilon$$

S.A: $x_1 + x_2 = 1, \qquad x_1 \ge 0, \qquad x_2 \ge 0 \qquad \varepsilon \ge 0$

$$-3x_1 + x_2 \le 10\varepsilon \qquad 3x_1 - x_2 \le 10\varepsilon$$

$$-4x_1 + 6x_2 \le 25\varepsilon \qquad 4x_1 - 6x_2 \le 25\varepsilon$$

$$3x_1 - x_2 \le 10\varepsilon \qquad -3x_1 + x_2 \le 10\varepsilon$$

Resolviendo (se ha utilizado el programa WINQSB 2.0):

$$\vec{\alpha}_{\min} = \begin{bmatrix} 0.425 \\ 0.575 \end{bmatrix} \qquad \varepsilon = 0.07$$

3.- Mínima incompatibilidad.

Dadas dos matrices en las condiciones habituales, no compatibles. Trataremos de encontrar una matriz $M = (p_{ij})$ de dimensión $n \times m$, no negativa, cuyos elementos suman uno y de forma que estén próxima, en algún sentido a las matrices *A* y *B*, esto es

$$\sum_{(i,j)\in N} p_{ij} = 1 \qquad p_{ij} \ge 0 \;\; \forall (i,j) \in N$$

$$\frac{p_{ij}}{\sum_{i=1}^{n} p_{ij}} \simeq a_{ij} \qquad \frac{p_{ij}}{\sum_{j=1}^{m} p_{ij}} \simeq b_{ij} \qquad \forall (i,j) \in N$$

Para precisar esto, necesitaremos una medida de la discrepancia (vía distancia o pseudo distancia) entre las matrices *M, A* y *B*, $\delta(M; A, B)$ que se minimizará entre las matrices *M* no negativas cuyos elementos sumen uno.

$$\min \delta(M;A,B)$$

$$\text{S.A.} \sum_{(i,j)\in N} p_{ij} = 1$$

$$p_{ij} \geq 0 \ \ \forall (i,j) \in N$$

Denotaremos M^* a ese mínimo, y $\delta(M^*;A,B)$ puede ser utilizado como medida de incompatibilidad, de forma que si $\delta(M^*;A,B)$ es una medida del grado de proximidad entre M^* y el par (A,B). Por otra parte, si (A,B) y (A',B') son pares de matrices de la misma dimensión y $\delta(M_1^*;A,B) < \delta(M_2^*;A',B')$ diremos que (A,B) son menos incompatibles que (A',B').

La primera medida de discrepancia (Arnold y Gokhale, 1998) utilizada fue la divergencia de Kullback-Leiber aprovechando que es una pseudo distancia entre distribuciones .

$$D(M;A,B) = \sum_{i=1}^{n}\sum_{j=1}^{m} b_{ij} \log\left(\frac{b_{ij}\, p_{i\bullet}}{p_{ij}}\right) + \sum_{i=1}^{n}\sum_{j=1}^{m} a_{ij} \log\left(\frac{a_{ij}\, p_{\bullet j}}{p_{ij}}\right)$$

Convenimos que *0log0=0*

Así el problema quedaría

$$\min D(M;A,B)$$

$$\text{S.A.} \sum_{(i,j)\in N} p_{ij} = 1$$

$$p_{ij} \geq 0 \ \ \forall (i,j) \in N$$

Teorema 3.3.1 (Arnold y Gokhale, 1998).

Con la notación anterior, la matriz $M^* = (p_{ij}^*)$ que minimiza $D(M;A,B)$ verifica:

$$\frac{p_{ij}^*}{p_{i\bullet}^*} + \frac{p_{ij}^*}{p_{\bullet j}^*} = a_{ij} + b_{ij}, \quad i = 1,2,\ldots,n; \ \ j = 1,2,\ldots,m$$

Un algoritmo iterado para resolver el sistema viene dado por:

$$p_{ij}^{(n+1)} = \frac{(a_{ij}+b_{ij})\left[\dfrac{1}{p_{i\bullet}^{(n)}}+\dfrac{1}{p_{\bullet j}^{(n)}}\right]^{-1}}{\sum_{i,j}(a_{ij}+b_{ij})\left[\dfrac{1}{p_{i\bullet}^{(n)}}+\dfrac{1}{p_{\bullet j}^{(n)}}\right]^{-1}}$$

Tomando como valores iniciales, $p_{ij}^{(0)} = \dfrac{1}{nm}, \quad i = 1,2,...,n;\ j = 1,2,...,m$

Este método iterado parece converger rápidamente, aunque no se dispone de prueba de convergencia. Puede verse una solución alternativa en Arnold y Gokhale (1998) y Arnold, Castillo y Sarabia (1999). En esas referencias pueden verse ejemplos.

Existen otras medidas de discrepancia que pueden ser utilizada en el mismo sentido que se ha utilizado *D*, por ejemplo,

$$Q(M;A,B) = \sum_{i=1}^{n}\sum_{j=1}^{m}\left[(p_{ij}-b_{ij}p_{i\bullet})^2+(p_{ij}-a_{ij}p_{\bullet j})^2\right]$$

Los valores p_{ij} que minimizan *Q*. puede obtenerse resolviendo el sistema:

$$p_{i\bullet} = 2\sum_{j=1}^{m} p_{\bullet j}a_{ij}, \quad i = 1,2,...,n$$

$$p_{\bullet j} = 2\sum_{j=1}^{m} p_{i\bullet}b_{ij}, \quad j = 1,2,...,m$$

Una vez calculados $p_{i\bullet}$ y $p_{\bullet j}$ basta aplicar:

$$p_{ij} = \frac{1}{2}(p_{\bullet j}a_{ij} + p_{i\bullet}b_{ij}) \qquad \forall i, j$$

Pueden verse más medidas de discrepancia y su aplicación a este problema en Arnold y Gokhale (1998), Arnold, Castillo y Sarabia (1999) y Ghosh y Balakrishnan (2013).

4.- Fórmula de Besag e incompatibilidad.

En el capítulo 1, estudiamos un método basado en la fórmula de Besag para comprobar la compatibilidad o incompatibilidad de dos matrices *A* y *B*, posibles distribuciones condicionadas de un vector aleatorio bidimensional. Este método también puede utilizarse (Wang y Kuo, 2010) para encontrar una distribución que podamos considerar próxima a la compatibilidad. La idea es asignar a cada cociente de probabilidad la media geométrica de su valor a través de todos los caminos posibles, entre la posición indicada por el denominador, y la posición indicada por el denominador.

Es claro, que si las matrices son compatibles, existe un vector aleatorio cuyas distribuciones condicionadas coinciden con *A* y *B*, y este método obtendría la distribución conjunta de dicho vector.

Ejemplo 1:

Consideremos las siguientes matrices,

$$A = \begin{bmatrix} \frac{1}{4} & \frac{1}{3} \\ \frac{3}{4} & \frac{2}{3} \end{bmatrix} \qquad B = \begin{bmatrix} \frac{1}{5} & \frac{4}{5} \\ \frac{3}{5} & \frac{2}{5} \end{bmatrix}$$

Todos los caminos que utilizaremos comenzarán en la posición (1,1) y utilizaremos la notación habitual, es decir la notación que se utilizaría si fuesen compatibles.

Unimos (1,1) con (1,2). Existen dos caminos, cuyos cocientes de probabilidad valen:

(1,1)→(1,2)

$$\frac{P(X=1,Y=2)}{P(X=1,Y=1)}=\frac{P(Y=2\,|\,X=1)}{P(Y=1\,|\,X=1)}=\frac{\frac{4}{5}}{\frac{1}{5}}=4$$

(1,1)→(2,1)→(2,2)→(1,2)

$$\frac{P(X=1,Y=2)}{P(X=1,Y=1)}=\frac{P(X=2,Y=1)}{P(X=1,Y=1)}\frac{P(X=2,Y=2)}{P(X=2,Y=1)}\frac{P(X=1,Y=2)}{P(X=2,Y=2)}=1$$

Asignamos el valor

$$\frac{P(X=1,Y=2)}{P(X=1,Y=1)}=\sqrt{4\cdot 1}=2$$

Unimos (1,1) con (2,1). Existen dos caminos, cuyos cocientes de probabilidad valen:

(1,1)→(2,1) $$\frac{P(X=2,Y=1)}{P(X=1,Y=1)}=3$$

(1,1)→(1,2)→(2,2)→(2,1) $$\frac{P(X=2,Y=1)}{P(X=1,Y=1)}=\frac{16}{3}$$

Asignamos el valor:

$$\frac{P(X=2,Y=1)}{P(X=1,Y=1)}=\sqrt{3\cdot\frac{16}{3}}=4$$

Unimos (1,1) con (2,2). Existen dos caminos, cuyos cocientes de probabilidad valen:

(1,1)→(1,2)→(2,2) $\dfrac{P(X=2,Y=2)}{P(X=1,Y=1)}=2$

(1,1)→(2,1)→(2,2) $\dfrac{P(X=2,Y=2)}{P(X=1,Y=1)}=8$

Asignamos el valor:

$$\frac{P(X=2,Y=2)}{P(X=1,Y=1)}=\sqrt{2\cdot 8}=4$$

Haciendo $P(X=1,Y=1)=p$ y obligando a que la suma de todas las probabilidades sea la unidad, obtenemos $p=\dfrac{1}{11}$ lo que nos lleva a la matriz de probabilidad conjunta,

$$M_{aprox}=\frac{1}{11}\begin{bmatrix}1 & 2\\ 4 & 4\end{bmatrix}$$

Y las matrices de probabilidad condicionada,

$$A_{aprox}=\begin{bmatrix}\frac{1}{5} & \frac{1}{3}\\ \frac{4}{5} & \frac{2}{3}\end{bmatrix} \qquad B_{aprox}=\begin{bmatrix}\frac{1}{3} & \frac{2}{3}\\ \frac{1}{2} & \frac{1}{2}\end{bmatrix}$$

Uno de las dificultades del método, es que no se tiene una medida adecuada de la bondad de la distribución conjunta obtenida. Una posibilidad, que puede ser aplicada para la comparación de la bondad de la aproximación de otros pares de matrices incompatibles, es la medida

$$\varphi=\frac{1}{2}\left(\left\|A-A_{aprox}\right\|+\left\|B-B_{aprox}\right\|\right)$$

Donde, podemos utilizar cualquier norma de matriz.

Ejemplo1 (continuación)

Utilizaremos el ejemplo anterior y la norma de Frobenius.

$$\varphi = \frac{1}{2}\left(\left\|A - A_{aprox}\right\|_F + \left\|B - B_{aprox}\right\|_F\right) = \frac{1}{2}\left(\left\|\begin{bmatrix} \frac{1}{4} & \frac{1}{3} \\ \frac{3}{4} & \frac{2}{3} \end{bmatrix} - \begin{bmatrix} \frac{1}{5} & \frac{1}{3} \\ \frac{4}{5} & \frac{2}{3} \end{bmatrix}\right\|_F + \left\|\begin{bmatrix} \frac{1}{5} & \frac{4}{5} \\ \frac{3}{5} & \frac{2}{5} \end{bmatrix} - \begin{bmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}\right\|_F\right) =$$

$$= \frac{1}{2}\left(\left\|\begin{bmatrix} \frac{1}{20} & 0 \\ \frac{-1}{20} & 0 \end{bmatrix}\right\|_F + \left\|\begin{bmatrix} \frac{-2}{15} & \frac{2}{15} \\ \frac{1}{10} & \frac{-1}{10} \end{bmatrix}\right\|_F\right) = \frac{13\sqrt{2}}{120} \simeq 0.1532...$$

5.- Uso de la descomposición en valores singulares.

En esta sección aplicaremos los teoremas de aproximación de rango bajo a la matriz de cocientes *C* para encontrar distribuciones marginales que, combinadas con las matrices condicionadas nos permitan obtener una distribución conjunta aproximada, en algún sentido que precisaremos. En la primera parte de esta sección seguimos a Schot (1997).

En virtud de la descomposición en valores singulares de una matriz, toda matriz T de dimensiones $n\times m$ y rango r, puede escribirse como

$$T = UDV^t = \sum_{i=1}^{r} \sigma_i \vec{u}_i \vec{v}_i$$

Donde $\vec{u}_i$ y $\vec{v}_i$ son los vectores columnas de las matrices U y V respectivamente y los valores singulares de T están ordenados $\sigma_1 \geq \sigma_2 \geq ... \geq \sigma_r$. Es decir, toda matriz puede escribirse como suma de matrices de rango uno.

Se llama aproximación de rango $k<r$ de la matriz T a la matriz

$$T_k = \sum_{i=1}^{k} \sigma_i \vec{u}_i \vec{v}_i = T \sum_{i=1}^{k} \vec{v}_i \vec{v}_i^t$$

Se tiene

(a) El rango de T_k es k.

(b) T_k es la matriz de rango *k* más próxima en el sentido de la norma matricial euclidea (norma matricial inducida por la norma vectorial euclidea), es decir

$$\|T-T_k\|_2 = \min\left\{\|T-S\|_2 : S = S_{n\times m},\ rang(S) = k\right\} = \sigma_{k+1}$$

(c) T_k es la matriz de rango *k* más próxima en el sentido de la norma matricial de Frobenius , $\|T\|_F = \sqrt{\sum_{i,j} t_{ij}^2}$,es decir

$$\|T-T_k\|_F = \min\left\{\|T-S\|_F : S = S_{n\times m},\ rang(S) = k\right\} = \sqrt{\sigma_{k+1}^2 + \sigma_{k+2}^2 + \ldots + \sigma_r^2}$$

Estos resultados pueden ser aplicados a la matriz de cocientes *C* para obtener distribuciones próximas en el caso de incompatibilidad. Evidentemente, las aproximaciones que usaremos serán de rango unidad. Nótese que es necesario que $N = N_{n\times m}$.

Ejemplo 1.

Sean las matrices

$$A = \begin{bmatrix} 0.2 & 0.6 \\ 0.8 & 0.4 \end{bmatrix} \qquad B = \begin{bmatrix} 0.25 & 0.75 \\ 0.25 & 0.75 \end{bmatrix}$$

Se tienen los siguientes resultados obtenidos con MATLAB,

$$C = \begin{bmatrix} 0.8000 & 0.8000 \\ 3.2000 & 0.5333 \end{bmatrix} \qquad \vec{u} = \begin{bmatrix} 0.2829 \\ 0.9591 \end{bmatrix} \qquad \vec{v} = \begin{bmatrix} 0.9758 \\ 0.2185 \end{bmatrix}$$

Los valores singulares son 3.3772 y 0.6317

De ahí, se obtiene la matriz aproximación de rango unidad $C_1 = \begin{bmatrix} 0.9324 & 0.2088 \\ 3.1610 & 0.7077 \end{bmatrix}$

que tiene rango unidad y está a una distancia euclidea (en este caso a la misma distancia de Frobenius) $\|C-C_1\|_2 = \|C-C_1\|_F = 0.6317$ nos permitirá obtener las marginales próximas, utilizando $\vec{u}$ y $\vec{v}$ obteniendo:

$$\vec{\alpha} = \begin{bmatrix} 0.2278 \\ 0.7722 \end{bmatrix} \qquad \vec{\beta} = \begin{bmatrix} 0.1829 \\ 0.8171 \end{bmatrix}$$

Es claro que $\|C-C_1\|_2 = \|C-C_1\|_F = 0.6317$ tiene una difícil interpretación en término de las distribuciones que hemos denominado próximas, sin embargo, puede servir para comparar el grado de incompatibilidad de pares de matrices incompatibles *A* y *B*.

Puesto que las matrices *A* y *B* son incompatibles la condición $N_A = N_B$ no tiene por qué verificarse. En este caso, podemos extender el método al caso en que *A* tenga elementos nulos y *B* no los tenga, si fuese al revés basta cambiar el papel de las variables para poder aplicarlos.

Capítulo 4

Rango infinito numerable.

En este capítulo discutimos el caso en el que al menos una de las variables aleatorias discretas implicadas, tiene rango infinito numerable. Este es el caso menos estudiado en la literatura. Apenas se cuentan con fuentes sobre el tema. Aquí daremos una breve introducción al tema, modificando el método de extensión positiva de rango unidad.

Estudiamos, separadamente dos casos:

Caso finita vs infinita, en el que una de las variables tiene rango finito y la otra rango infinito. Sin pérdida de generalidad, supondremos que es X la que tiene rango finito.

Caso infinita vs infinita, en el que ambas variables tiene rango infinito.

1.- Caso finita vs infinita.

Dadas las colecciones de números

$$A=\left\{a_{ij}:i=1,\dots,n;\quad j=1,2,\dots\right\} \quad \text{y} \quad B=\left\{b_{ij}:i=1\dots.n;\quad j=1,2,\dots\right\}$$

Todos ellos no negativos y $\sum_{i=1}^{n} a_{ij} = 1$ y $\sum_{j=1}^{+\infty} b_{ij} = 1$.

Supondremos que estos conjuntos son compatibles, en el sentido de que existen dos variables aleatorias discretas, *X* e *Y*, cuyos valores posibles son $R_X = \{x_1, x_2, ..., x_n\}$ $R_Y = \{y_1, y_2, ...\}$ y tales que:

$$a_{ij} = P\left[X = x_i \mid Y = y_j\right] \qquad i = 1, 2, ..., n; \ j = 1, 2, ...$$

$$b_{ij} = P\left[Y = y_j \mid X = x_i\right] \qquad i = 1, 2, ..., n; \ j = 1, 2, ...$$

supondremos además que ninguna distribución condicionada es idénticamente nula pues, en otro caso, el valor de *X* (o de *Y*) correspondiente tiene probabilidad nula en cualquier distribución y, consecuentemente, puede ser eliminado del problema. Y además, se sigue verificando la condición necesaria de compatibilidad, esto es, si los conjuntos numéricos son compatibles,

$$A = \left\{a_{ij} : i = 1, ..., n; \quad j = 1, 2, ...\right\} \quad \text{y} \quad B = \left\{b_{ij} : i = 1....n; \quad j = 1, 2, ...\right\}$$

Entonces, los conjuntos, $N_A = \left\{(i, j) : a_{ij} > 0 \quad i = 1, ..., n; \quad j = 1, 2, ...\right\}$ y $N_B = \left\{(i, j) : b_{ij} > 0 \quad i = 1....n; \quad j = 1, 2, ...\right\}$ verifican $N_A = N_B$ en cuyo caso lo denominamos *N*.

La condición de unicidad de Gupta y Vargas (1990) queda, en este contexto, como sigue:

Existe $(i_o, j_0) \in N$ tal que:

$$\forall j \in N_B, \qquad (i_0, j) \in N \quad \text{esto es } a_{i_0 j} > 0 \quad \text{y} \quad b_{i_0 j} > 0$$

$$\forall i \in N_A, \qquad (i, j_0) \in N \quad \text{esto es } a_{ij_0} > 0 \quad \text{y} \quad b_{ij_0} > 0$$

Esto quiere decir que, al menos, una distribución condicionada de cada tipo tiene todas sus probabilidades no nulas.

En particular si $a_{ij} > 0 \;\; \forall i = 1,2,...,n \;\; \forall j = 1,2...$ se tiene la unicidad.

Por otra parte, el uso de la matriz de cocientes y la metodología utilizada en la sección 5 del primer capítulo, no es directamente aplicable aquí, sin embargo, es posible su generalización en **el caso en que el número de ceros en las distribuciones condicionadas es finito**. En efecto, supongamos que $a_{ij} > 0 \;\; \forall j \geq r$ y consideramos los cocientes $c_{ij} = \dfrac{a_{ij}}{b_{ij}}$ para aquellos $(i, j) \in N$ construyamos ahora la siguiente tabla:

$$\begin{array}{lllll}
c_{11} & c_{12} \ldots\ldots & c_{1r} & c_{1r+1} \ldots\ldots & c_{1j} \ldots\ldots \\
c_{21} & c_{22} \ldots\ldots & c_{2r} & c_{2r+1} \ldots\ldots & c_{2j} \ldots\ldots \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
c_{i1} & c_{i2} & c_{ir} & c_{ir+1} \ldots\ldots & c_{ij} \ldots\ldots \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
c_{n,1} & c_{n,2} & c_{nr} & c_{nr} \ldots\ldots & c_{nj} \ldots\ldots
\end{array}$$

Las primeras r columnas, forman una matriz C, de orden $n \times r$, con algunos de sus elementos desconocidos, mientras que las restantes columnas (que son infinitas) contiene todos sus elementos conocidos.

Obsérvese que la columna r de C es de elementos conocidos. Con esta notación es evidente que A y B determinan una única distribución conjunta si y sólo si C se puede completar de manera única de forma que su rango sea la unidad.

Para verificar esta condición se pueden usar los resultados expuestos en el primer capítulo.

Respecto a la determinación efectiva de la distribución conjunta es recomendable determinar la sucesión $v_1, v_2, \ldots$ mediante la conocida condición

$$v_j = \sum_{i=1}^{L} c_{ij}$$

de ahí la distribución marginal de Y, β_j, pasa por la normalización de la sucesión v_j.

$$\beta_j = \frac{v_j^{-1}}{\sum_{j=1}^{\infty} v_j^{-1}} \qquad j = 1, 2, \ldots$$

y la distribución conjunta viene dada por

$$p_{ij} = a_{ij}\beta_j \qquad i = 1, \ldots, n; \quad j = 1, 2, \ldots$$

La razón por la que se elige esta forma de trabajar es que la serie .$\sum_{j=1}^{\infty} c_{ij}$ .puede ser divergente, aunque también puede serlo $\sum_{j=1}^{\infty} v_j^{-1}$, en cuyo caso no existe la marginal, y no serían compatibles.

Ejemplo:

Consideremos los siguientes conjuntos que, puede probarse sin dificultad que son compatibles.

$$A=\left\{ a_{ij} : \quad i=1,2,3; \qquad j=1,2,\ldots \right\}$$

$$a_{11} = 1/2 \quad , \quad a_{21} = 0 \quad , \quad a_{3,1} = 1/2$$

$$a_{12} = 1/3 \quad , \quad a_{22} = 2/3 \quad , \quad a_{32} = 0$$

$$a_{13} = 0 \quad , \quad a_{23} = 1/2 \quad , \quad a_{33} = 1/2$$

$$a_{1j} = 1/3 \quad , \quad a_{2j} = 1/3 \quad , \quad a_{3j} = 1/3 \quad \text{para } j \geq 4$$

$$B=\left\{ b_{ij}: \quad i=1,2,3; \qquad j=1,2,\ldots \right\}$$

$$b_{11}=1/3 \quad , \quad b_{12}=1/3 \quad , \quad b_{13}=0 \qquad b_{ij}=\frac{1}{3\times 2^{j-3}} \quad \text{para} \quad j\geq 4$$

$$b_{21}=0, \qquad b_{22}=1/2, \qquad b_{23}=1/4, \qquad b_{2j}=\frac{1}{4}\times\frac{1}{2^{j-3}} \quad \text{para} \quad j\geq 4$$

$$b_{31}=1/3, \qquad b_{32}=0, \qquad b_{33}=1/3, \qquad b_{3j}=\frac{1}{3}\times\frac{1}{2^{j-3}} \quad \text{para} \quad j\geq 4$$

Como puede observarse el número de probabilidades nulas es finito e igual a tres.

La tabla (incompleta) de los cocientes $c_{ij}=\dfrac{a_{ij}}{b_{ij}}$ queda:

$$\begin{array}{cccccc}
\frac{3}{2} & 1 & * & 2 & 4\cdots\cdots & 2^{j-3}\cdots\cdots \\
* & \frac{4}{3} & 2 & \frac{8}{3} & \frac{16}{3}\cdots\cdots & \frac{4}{3}2^{j-3}\cdots\cdots \\
\frac{3}{2} & * & \frac{3}{2} & 2 & 4\cdots\cdots & 2^{j-3}\cdots\cdots
\end{array}$$

Aquí r = 4 y la matriz C incompleta es

$$C=\begin{bmatrix} \frac{2}{3} & 1 & * & 2 \\ * & \frac{4}{3} & 2 & \frac{8}{3} \\ \frac{3}{2} & * & \frac{3}{2} & 2 \end{bmatrix}$$

completando vía *rang* C = 1

$$C = \begin{bmatrix} \frac{2}{3} & 1 & \frac{3}{2} & 2 \\ 2 & \frac{4}{3} & 2 & \frac{8}{3} \\ \frac{3}{2} & 1 & \frac{3}{2} & 2 \end{bmatrix}$$

El cálculo de los v_j es ahora inmediato

$$v_1 = 5, \quad v_2 = \frac{10}{3}, \quad v_3 = 5, \quad v_j = \frac{10}{3} 2^{j-3} \qquad j \geq 4$$

de ahí es inmediata la obtención de la marginal

$$\beta_1 = \frac{1}{5} \qquad \beta_2 = \frac{3}{10} \qquad \beta_3 = \frac{1}{5} \qquad \beta_j = \frac{3}{10} \frac{1}{2^{j-3}} \quad j \geq 4$$

y utilizando el conjunto a_{ij} se obtiene la siguiente distribución conjunta que es única

	1	2	3	4	5…..	J
1	$\frac{1}{10}$	$\frac{1}{10}$	0	$\frac{1}{10}\frac{1}{2}$	$\frac{1}{10}\frac{1}{2^2}$ ….	$\frac{1}{10}\frac{1}{2^{j-3}}$ ….
2	0	$\frac{1}{5}$	$\frac{1}{10}$	$\frac{1}{10}\frac{1}{2}$	$\frac{1}{10}\frac{1}{2^2}$ ….	$\frac{1}{10}\frac{1}{2^{j-3}}$ ….
3	$\frac{1}{10}$	0	$\frac{1}{10}$	$\frac{1}{10}\frac{1}{2}$	$\frac{1}{10}\frac{1}{2^2}$ ….	$\frac{1}{10}\frac{1}{2^{j-3}}$ ….

Es importante señalar que merced a la estructura de la matriz *C*, cuya última columna es de elementos no nulos, se tiene **que si los conjuntos de *A* y *B* son compatibles y el número de probabilidades condicionadas nulas es finito, entonces, la distribución conjunta que determinan es única.** El recíproco es falso, como tendremos ocasión de comprobar.

La metodología de los cocientes no es, en general, viable para el caso en que el número de probabilidades condicionadas nulas sea infinito.

Si el número de probabilidades condicionadas nulas es infinito, es posible utilizar la teoría de Cadenas de Markov para resolver el problema. En efecto, sea $\alpha_i,\ i = 1, 2, \dots, n$ la distribución marginal de la variable aleatoria *X*, entonces

$$\alpha_i = \sum_{j=1}^{\infty} a_{ij}\beta_j = \sum_{j=1}^{\infty} a_{ij} \sum_{k=1}^{L} b_{kj}\alpha_k = \sum_{j=1}^{\infty}\sum_{k=1}^{L} a_{ij} b_{kj}\alpha_k$$

Intercambiando el orden de sumación (ver, por ejemplo, Isaacson y Madsen, 1976).

$$\alpha_i = \sum_{k=1}^{L}\left[\sum_{j=1}^{\infty} a_{ij} b_{kj}\right]\alpha_k$$

en consecuencia los números $d_{ik} = \sum_{j=1}^{\infty} a_{ij} b_{kj}$ pueden ser considerados probabilidades de transición de una cadena de Markov con *n* estados.

$$\sum_{i=1}^{L} d_{ik} = \sum_{i=1}^{L}\sum_{j=1}^{\infty} a_{ij} b_{kj} = \sum_{j=1}^{\infty}\sum_{i=1}^{L} a_{ij} b_{kj} = \sum_{j=1}^{\infty} b_{kj} = 1$$

y $\vec{\alpha}$ es una distribución estacionaria de dicha cadena; al ser una cadena de Markov finita, la cuestión de la unicidad se reduce a un problema de cálculo de autovalores.

Veamos algunos ejemplos:

Ejemplo 1 (Ng, 1997)*:*

Este ejemplo muestra una tabla con infinitos ceros.

Supongamos que se tienen repeticiones independientes de un experimento de Bernoulli de parámetro *p*. Es sabido que esto se puede visualizar como una sucesión de unos (éxito) y ceros (fracaso). Consideremos las sucesiones finitas que terminan en uno. Sean X el número de unos en la sucesión y N la longitud de la misma. Es claro que, excluyendo el último uno, podemos modelizar la distribución condicionada $X-1\,|\,(N-1=n-1) \sim B(n-1,p)$. Por otra parte, es natural modelizar el número de elementos de la sucesión como $N\,|\,(X=x) \sim B^{-}(x,p), \quad n=x,x+1,\ldots$

$$c_{xn} = \frac{P(X=x\,|\,N=n)}{P(N=n\,|\,X=x)} = \frac{P(X-1=x-1\,|\,N=n)}{P(N=n\,|\,X=x)} = \frac{\binom{n-1}{x-1}p^{x-1}(1-p)^{n-x}}{\binom{n-1}{x-1}p^{x}(1-p)^{n-x}} = \frac{1}{p},\ n \geq x$$

Y, por tanto, se tiene que

$$\begin{vmatrix} c_{x_1 n_1} & c_{x_1 n_2} \\ c_{x_2 n_1} & c_{x_2 n_2} \end{vmatrix} = \begin{vmatrix} \frac{1}{p} & \frac{1}{p} \\ \frac{1}{p} & \frac{1}{p} \end{vmatrix} = 0$$

Con una adecuada elección de los valores x_1, x_2, n_1, n_2.

Obsérvese que, $\sum_{n=x}^{+\infty} \frac{1}{p}$ diverge y por lo tanto, no existen marginales compatibles con las condicionadas anteriores.

Ejemplo 2:

Consideremos los siguientes conjuntos compatibles:

$$.\ a_{1,j} = \begin{cases} 0 & \text{si } j \text{ es par} \\ \dfrac{r}{1+r} & \text{si } j \text{ es impar} \end{cases}.$$

$$a_{2,j} = \begin{cases} \dfrac{r}{1+r} & \text{si } j \text{ es par} \\ \dfrac{1}{1+r} & \text{si } j \text{ es impar} \end{cases}$$

$$a_{3,j} = \begin{cases} \dfrac{1}{1+r} & \text{si } j \text{ es par} \\ 0 & \text{si } j \text{ es impar} \end{cases}$$

con $j = 1,2,\ldots$ y r un número real, $0 < r < 1$.

Así obtenemos el conjunto A= $\{a_{ij}$: $i = 1,2,3$; $j = 1,2,\ldots\}$

Por su parte el conjunto B= $\{b_{ij}$: $i = 1,2,3$; $j = 1,2,\ldots\}$ está definido por:

$$b_{1j} = \begin{cases} 0 & \text{si } j \text{ es par} \\ r^{j-1}\left(1-r^2\right) & \text{si } j \text{ es impar} \end{cases}$$

$$b_{2j} = r^{j-1}\left(1-r\right) \qquad \text{para todo } j$$

$$b_{3j} = \begin{cases} r^{j-2}\left(1-r^2\right) & \text{si } j \text{ es par} \\ 0 & \text{si } j \text{ es impar} \end{cases}$$

Calculemos las probabilidades d_{lk} l,k = 1,2,3.

$$d_{1,1} = \sum_{j=1}^{\infty} a_{1,j} b_{1,j} = \sum_{h=0}^{\infty} r^{2h}\left(1-r^2\right)\frac{r}{1+r} = \frac{r}{1+r}$$

$$d_{1,2} = \sum_{j=1}^{\infty} a_{1,j} b_{2,j} = \sum_{h=0}^{\infty} \frac{r}{1+r} r^{2h}\left(1-r\right) = \frac{r}{(1+r)^2}$$

$$d_{1,3} = \sum_{j=1}^{\infty} a_{1,j} b_{3,j} = \sum_{h=0}^{\infty} \frac{r}{1+r} 0 = 0$$

Análogamente se obtiene

$$d_{21} = \frac{1}{1+r} \quad , \qquad d_{22} = \frac{r^2+1}{(1+r)^2} \quad , \qquad d_{23} = \frac{r}{1+r}$$

$$d_{31} = 0 \quad , \qquad d_{32} = \frac{r}{(1+r)^2} \quad , \qquad d_{33} = \frac{1}{1+r}$$

Obteniéndose así la matriz:

$$D = \begin{bmatrix} \frac{r}{1+r} & \frac{r}{(1+r)^2} & 0 \\ \frac{1}{1+r} & \frac{r^2+1}{(1+r)^2} & \frac{r}{1+r} \\ 0 & \frac{r}{(1+r)^2} & \frac{1}{1+r} \end{bmatrix}$$

Cuya ecuación característica es

$$-\frac{r^2+1}{(r+1)^2}\lambda + \left[\frac{r^2+1}{(r+1)^2}+1\right]\lambda^2 - \lambda^3 = 0$$

y los autovalores son:

$$\lambda_1 = 1 \qquad \lambda_2 = \frac{1+r^2}{(1+r)^2} \qquad \lambda_3 = 0$$

puesto que $\lambda_2 < 1$ para $r > 0$ se tiene que la cadena de Markov con matriz de transición D tiene una única distribución estacionaria, esta es proporcional al autovector asociado al uno, éste es

$$\vec{\alpha} = \left[\frac{r}{2(1+r)}, \frac{1}{2}, \frac{1}{2(1+r)}\right]^t$$

y así, la distribución conjunta viene dada por la siguiente tabla:

	1	*2*	*3*	*4 ...*	*2h*	*2h+1*	*2h+2 ...*
1	αr^2	0	αr^4	$0 \ldots$	0	αr^{2h+2}	$0 \ldots$
2	αr	αr^2	αr^3	$\alpha r^4 \ldots$	αr^{2h}	αr^{2h+1}	$\alpha r^{2h+2} \ldots$
3	0	αr	0	$\alpha r^3 \ldots$	αr^{2h-1}	0	$\alpha r^{2h+1} \ldots$

donde la constante α toma el valor $\alpha = \dfrac{1-r}{2r}$.

El siguiente ejemplo, es de no unicidad.

Ejemplo 3:

Los conjuntos de números *A* y *B* vienen dados por

$$a_{1j} = \begin{cases} 0 & \text{si } j \text{ es par} \\ 1 & \text{si } j \text{ es impar} \end{cases}$$

$$a_{2j} = \begin{cases} 1 & \text{si } j \text{ es par} \\ 0 & \text{si } j \text{ es impar} \end{cases}$$

$$b_{1j} = \begin{cases} 0 & \text{si } j \text{ es par} \\ (1-r^2)\, r^{j-1} & \text{si } j \text{ es impar} \end{cases}$$

$$b_{2j} = \begin{cases} (1-r^2)\, r^{j-2} & \text{si } j \text{ es par} \\ 0 & \text{si } j \text{ es impar} \end{cases}$$

donde $j = 1,2,\ldots$ y r es un número real $0 < r < 1$.

Calculamos la matriz D

$$d_{11} = \sum_{j=1}^{\infty} a_{1j} b_{1j} = \sum_{h=0}^{\infty} \frac{1-r^2}{r^2} r^{2h+2} = 1$$

$$d_{12} = \sum_{j=1}^{\infty} a_{1j} b_{2j} = 0$$

$$d_{21} = \sum_{j=1}^{\infty} a_{2j} b_{1j} = 0$$

$$d_{22} = \sum_{j=1}^{\infty} a_{2j} b_{2j} = \sum_{h=1}^{\infty} \frac{1-r^2}{r^3} r^{2h+1} = 1$$

Así

$$D=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$$

y cualquier distribución es estacionaria, es decir

$$\vec{\alpha}=(\beta,1-\beta)^t \qquad \text{con } 0<\beta<1$$

es una posible distribución marginal, y la distribución conjunta que se obtiene es

$$p_{1j}=\begin{cases}0 & \text{si } j \text{ es par}\\ \beta r^{j-1}\left(1-r^2\right) & \text{si } j \text{ es impar}\end{cases}$$

$$p_{2j}=\begin{cases}(1-\beta)\left(1-r^2\right)r^{j-2} & \text{si } j \text{ es par}\\ 0 & \text{si } j \text{ es impar}\end{cases}$$

2.- Caso infinita vs infinita.

En el caso en que tanto *X* como *Y* son discretas e infinitas, algunos resultados particulares pueden deducirse del trabajo de Gupta y Vargas (1990). Así mismo, es posible el uso de cadenas de Markov.

Con la notación e hipótesis habituales, recordemos la condición necesaria, los conjuntos,

$$N_A = \left\{(i,j) : a_{ij} > 0 \quad i = 1,2...; \quad j = 1,2,...\right\}$$

$$N_B = \left\{(i,j) : b_{ij} > 0 \quad i = 1,2,...; \quad j = 1,2,...\right\}$$

verifican $N_A = N_B$ en cuyo caso lo denominamos *N*.

El resultado de Gupta y Vargas (1990) puede escribirse de la siguiente forma:

Sean *A* y *B* dos conjuntos compatibles, si existe un $(i_0, j_0) \in N$ tal que:

$$\forall j \in N_B, \qquad (i_0, j) \in N \qquad \text{esto es } a_{i_0,j} > 0 \text{ y } b_{i_0 j} > 0$$

$$\forall i \in N_A, \qquad (i, j_0) \in N \qquad \text{esto es } a_{ij_0} > 0 \text{ y } b_{ij_0} > 0$$

$$\forall j \in N_B, \qquad (i_0, j) \in N \qquad \text{esto es } a_{i_0 j} > 0 \text{ y } b_{i_0 j} > 0$$

$$\forall i \in N_A, \qquad (i, j_0) \in N \qquad \text{esto es } a_{ij_0} > 0 \text{ y } b_{ij_0} > 0$$

Entonces, la distribución conjunta que definen *A* y *B* es única.

Esto quiere decir que existen, al menos, dos distribuciones condicionadas, una de cada tipo, con todas sus probabilidades no nulas, en particular si $a_{ij} b_{ij} \neq 0$ para cualesquiera enteros positivos *i* y *j*, se tiene la unicidad. Más aún, si el número de probabilidades condicionadas nulas es finito, entonces, los conjuntos *A* y *B* definen una única distribución de probabilidad conjunta. El recíproco no es cierto.

Ejemplo 1:

Sea p_{ij} una sucesión de números reales verificando

(i) $\quad \sum_{i,j} p_{ij} = 1 \qquad 0 \le p_{ij} \le 1 \qquad i, j = 1, 2, \ldots$

(ii) $\quad$ Para i, j pares $\quad p_{ij} = 0$ y para i o j impar $p_{ij} \neq 0$ con $i, j > 1$

(iii) $\quad p_{i1} p_{1j} \neq 0 \qquad \forall i = 2, 3, \ldots \quad \forall j = 2, 3, \ldots$

$$p_{i\bullet} = \sum_{j=1}^{\infty} p_{ij} \qquad\qquad p_{\bullet j} = \sum_{i=1}^{\infty} p_{ij}$$

Definimos

$$a_{ij} = \frac{p_{ij}}{p_{\bullet j}} \qquad\qquad i, j = 1, 2, \ldots$$

$$b_{ij} = \frac{p_{ij}}{p_{i\bullet}} \qquad\qquad i, j = 1, 2, \ldots$$

Los conjuntos A y B están formados por los elementos a_{ij} y b_{ij} respectivamente y son compatibles por construcción. Estos conjuntos, definen una única distribución conjunta, por aplicación de Gupta y Vargas (1990), y hay una cantidad infinita de probabilidades nulas.

El siguiente ejemplo muestra el caso de no unicidad.

Ejemplo 2: (Arnold, Castillo y Sarabia 1992).

Considérese los conjuntos A y B dados por

$$a_{ij} = b_{ij} = 1 \qquad \text{si} \qquad i = j$$

$$a_{ij} = b_{ij} = 0 \qquad \text{si} \qquad i \neq j$$

Son compatibles y no definen una distribución conjunta única pues si consideramos la distribución bidimensional

$$p_{ij} = p_i \qquad \text{si} \qquad i = j = 0,1,2,...$$

$$p_{ij} = 0 \qquad \text{si} \qquad i \neq j$$

siendo p_i una sucesión real que verifica

$$0 \leq p_i \leq 1 \qquad \sum_{i=0}^{\infty} p_i = 1$$

Tiene por distribuciones condicionadas a los conjuntos *A* y *B*.

El estudio de la unicidad, vía cadenas de Markov, tiene un interés teórico innegable, aunque su operatividad es sumamente restringida. La justificación del uso de estos conceptos es como sigue:

Para cada $i = 1,2,\ldots$

$$\alpha_i = \sum_{j=1}^{\infty} a_{ij}\beta_j = \sum_{j=1}^{\infty} a_{ij} \sum_{k=1}^{L} b_{kj}\alpha_k = \sum_{j=1}^{\infty}\sum_{k=1}^{\infty} a_{ij}b_{kj}\alpha_k$$

Puesto que la serie doble es absolutamente convergente podemos cambiar, por aplicación del teorema de Fubini, el orden de sumación (ver Isaacsen y Madsen, 1976).

Así obtenemos:

$$\alpha_i = \sum_{k=1}^{\infty} \left[\sum_{j=1}^{\infty} a_{i,j} b_{k,j} \right] \alpha_k$$

y los números

$$d_{ik} = \sum_{j=1}^{\infty} a_{ij} b_{kj}$$

pueden considerarse las probabilidades de transición de una cadena de Markov con espacio de estados $\{1,2,\ldots\}$ puesto que

$$\sum_{i=1}^{\infty} d_{ik} = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} a_{ij} b_{kj} = \sum_{j=1}^{\infty} \sum_{i=1}^{\infty} a_{ij} b_{kj} = \sum_{j=1}^{\infty} b_{kj} \sum_{i=1}^{\infty} a_{ij} = \sum_{j=1}^{\infty} b_{kj} = 1$$

De esta forma, la distribución marginal η es una distribución estacionaria de la cadena de Markov. Sin embargo, las condiciones bajo las cuáles una cadena de Markov posee una única distribución estacionaria es difícil de comprobar en la práctica.

En el caso en que la tabla no tenga ceros, es posible utilizar la matriz de cocientes. Los siguientes ejemplos muestran cómo podemos deducir si dos condicionadas, procederían de un vector aleatorio. Los siguientes son especificaciones condicionales de modelos, que ya comentamos anteriormente.

Ejemplo 3 (Condicionadas de Poisson)*:*

Sean X e Y dos variables aleatorias, que verifican,

$X \mid Y = y \sim P(\lambda(y)),\ y = 0,1,....$ y $Y \mid X = x \sim P(\mu(x)),\ x = 0,1,....$

$$c_{xy} = \frac{P(X = x \,|\, Y = y)}{P(Y = y \,|\, X = x)} = \frac{e^{-\lambda(y)} \lambda(y)^x \, y!}{e^{-\mu(x)} \mu(x)^y \, x!}$$

$$\begin{vmatrix} c_{x_1 y_1} & c_{x_1 y_2} \\ c_{x_2 y_1} & c_{x_2 y_2} \end{vmatrix} = \frac{e^{-\lambda(y_1)} e^{-\lambda(y_2)}}{e^{-\mu(x_1)} e^{-\mu(x_2)}} \frac{y_1! y_2!}{x_1! x_2!} \left[\frac{\lambda(y_1)^{x_1} \lambda(y_2)^{x_2}}{\mu(x_1)^{y_1} \mu(x_2)^{y_2}} - \frac{\lambda(y_1)^{x_2} \lambda(y_2)^{x_1}}{\mu(x_1)^{y_2} \mu(x_2)^{y_1}} \right]$$

Que es nulo si y sólo si

$$\frac{\lambda(y_1)^{x_1} \lambda(y_2)^{x_2}}{\mu(x_1)^{y_1} \mu(x_2)^{y_2}} - \frac{\lambda(y_1)^{x_2} \lambda(y_2)^{x_1}}{\mu(x_1)^{y_2} \mu(x_2)^{y_1}} = 0$$

Operando, es nulo si y sólo si

$$\left[\frac{\lambda(y_1)}{\lambda(y_2)} \right]^{x_1 - x_2} = \left[\frac{\mu(x_1)}{\mu(x_2)} \right]^{y_1 - y_2}$$

Por tanto las condicionadas son compatibles sí y sólo si

$$\left[\frac{\lambda(y_1)}{\lambda(y_2)} \right]^{x_1 - x_2} = \left[\frac{\mu(x_1)}{\mu(x_2)} \right]^{y_1 - y_2} \quad x_1, x_2, y_1, y_2 = 0,1,2,.... \quad x_1 \neq x_2 \;\; y_1 \neq y_2$$

Ejemplo 4 (Condicionadas geométricas)*:*

Sean X e Y dos variables aleatorias, que verifican,

$X \,|\, Y = y \sim G(p(y)),\; y = 1,2,....$ y $Y \,|\, X = x \sim G(q(x)),\; x = 1,2,....$ donde p y q son funciones que toman valores en (0,1).

$$c_{xy}=\frac{P(X=x\,|\,Y=y)}{P(Y=y\,|\,X=x)}=\frac{(1-p(y))^{x-1}\,p(y)}{(1-q(x))^{y-1}q(x)}$$

Denominamos $\overline{p}=1-p,\quad \overline{q}=1-q$ y tenemos,

$$\begin{vmatrix} c_{x_1y_1} & c_{x_1y_2} \\ c_{x_2y_1} & c_{x_2y_2} \end{vmatrix}=\frac{p(y_1)p(y_2)}{q(x_1)q(x_2)}\left[\frac{\overline{p}(y_1)^{x_1-1}\,\overline{p}(y_2)^{x_2-1}}{\overline{q}(x_1)^{y_1-1}\overline{q}(x_2)^{y_2-1}}-\frac{\overline{p}(y_1)^{x_2-1}\,\overline{p}(y_2)^{x_1-1}}{\overline{q}(x_1)^{y_2-1}\overline{q}(x_2)^{y_1-1}}\right]$$

Es nulo si y sólo si $\dfrac{\overline{p}(y_1)^{x_1-1}\,\overline{p}(y_2)^{x_2-1}}{\overline{q}(x_1)^{y_1-1}\overline{q}(x_2)^{y_2-1}}-\dfrac{\overline{p}(y_1)^{x_2-1}\,\overline{p}(y_2)^{x_1-1}}{\overline{q}(x_1)^{y_2-1}\overline{q}(x_2)^{y_1-1}}=0$ si y sólo si

$$\left[\frac{\overline{p}(y_1)}{\overline{p}(y_2)}\right]^{x_1-x_2}=\left[\frac{\overline{q}(x_1)}{\overline{q}(x_2)}\right]^{y_1-y_2}$$

Por tanto las variables condicionadas son compatibles si y sólo si

$$\left[\frac{\overline{p}(y_1)}{\overline{p}(y_2)}\right]^{x_1-x_2}=\left[\frac{\overline{q}(x_1)}{\overline{q}(x_2)}\right]^{y_1-y_2}\quad x_1,x_2,y_1,y_2=1,2,\ldots.\qquad x_1\neq x_2\quad y_1\neq y_2$$

Capítulo 5

Acción y Reacción Probabilística.

En este capítulo, se ponen las bases para una aplicación del problema estudiado. Concretamente se estudia el problema de la compatibilidad y unicidad para el caso en que dos agentes (personas, empresas,…) enfrentados en algún sentido, pueden tener cierto conocimiento de las probabilidades de las distintas opciones de que dispongan, cuando suponen que ya ha actuado la otra, las preguntas que se plantean son por ejemplo ¿existe coherencia en los agentes? (compatibilidad) ¿existe una única distribución de probabilidad que ha generado las probabilidades condicionadas? (unicidad)

1.- Modelo de acción y reacción.

Consideremos dos agentes (empresas, organizaciones, personas…) α y β que en cierta situación competitiva pueden realizar diferentes *acciones.* Sean $A_1, A_2, ..., A_i, ..., A_n$ con probabilidades $P(A_1), P(A_2), ..., P(A_i), ..., P(A_n)$ las acciones que puede realizar α y $B_1, B_2, ..., B_j, ..., B_m$ con probabilidades $P(B_1), P(B_2), ..., P(B_j), ..., P(B_m)$ las acciones para β.

Asumimos que, para cada agente, el conjunto de acciones forman un sistema completo de sucesos con probabilidades no nulas, es decir, $A_1, A_2, ..., A_i, ..., A_n$ cubren todas las posibilidades para α y no se pueden realizar dos acciones simultáneamente, además las probabilidades de realizar cualquier acción es positiva. Análogamente para las acciones $B_1, B_2, ..., B_j, ..., B_m$ del agente β.

Las probabilidades definidas en los párrafos anteriores dependen, entre otras, del agente y de las condiciones externas.

Supongamos que la situación estudiada admite respuesta por parte de cada agente ante una acción del otro, a estas respuestas se les denominará *reacciones* y estarán dentro del conjunto de acciones. Así, si α realiza A_i, el agente β reaccionará efectuando una de las acciones $B_1, B_2, ..., B_j, ..., B_m$. De esta forma se pueden considerar, las probabilidades condicionadas $P(B_j \mid A_i),\ j = 1, 2, ..., m; i = 1, 2, ..., n$ que denominaremos *probabilidades de reacción.* De la misma forma, para una acción de β las probabilidades de las reacciones de α serán $P(A_i \mid B_j), \quad i = 1, 2, ..., n; \quad j = 1, 2, ..., m$.

Como hemos visto con anterioridad, estas probabilidades pueden ordenarse en forma matricial:

$$A = ((a_{ij})) \qquad a_{ij} = P(A_i \mid B_j), \quad i = 1, 2, ..., n; \quad j = 1, 2, ..., m$$

$$B = ((b_{ij})) \qquad b_{ij} = P(B_j \mid A_i), \quad j = 1,2,...,m; \quad i = 1,2,...,n$$

Ambas matrices tienen dimensión $n \times m$ y A es estocástica por columnas y B lo es por filas.

Ejemplo 1: Supongamos que los agentes α y β son dos empresas competidoras que producen el mismo producto. En un determinado momento, las acciones para α son

Subir el precio de su producto (A_1)

Bajar el precio de su producto (A_2)

No efectuar modificación en el precio (A_3).

Las acciones para la empresa β son las mismas aplicada a su producto.

Subir el precio de su producto (B_1)

Bajar el precio de su producto (B_2)

No efectuar modificación en el precio (B_3).

Así las acciones y sus probabilidades vienen dadas en la siguiente tabla.

Empresa α			**Empresa** β	
Acción	Probabilidad de la acción.		Acción	Probabilidad de la acción.
A_1	$P(A_1)$		B_1	$P(B_1)$
A_2	$P(A_2)$		B_2	$P(B_2)$
A_3	$P(A_3)$		B_3	$P(B_3)$

Por otra parte, si α decide subir el precio de su producto (A_1), la empresa β tiene las siguientes opciones sobre el precio de su producto: subir el precio (B_1), bajar el precio (B_2), o no efectuar modificación en el precio (B_3). Estas son las posibles reacciones de β a A_1 de α.

Las probabilidades de las posibles reacciones vienen dadas por la siguiente tabla.

Reacciones de α	Acciones de β		
	B_1	B_2	B_3
A_1	$P(A_1 \mid B_1)$	$P(A_1 \mid B_2)$	$P(A_1 \mid B_3)$
A_2	$P(A_2 \mid B_1)$	$P(A_2 \mid B_2)$	$P(A_2 \mid B_3)$
A_3	$P(A_3 \mid B_1)$	$P(A_3 \mid B_2)$	$P(A_3 \mid B_3)$

Acciones de α	Reacciones de β		
	B_1	B_2	B_3
A_1	$P(B_1 \mid A_1)$	$P(B_1 \mid A_2)$	$P(B_1 \mid A_3)$
A_2	$P(B_2 \mid A_1)$	$P(B_2 \mid A_2)$	$P(B_2 \mid A_3)$
A_3	$P(B_3 \mid A_1)$	$P(B_3 \mid A_2)$	$P(B_3 \mid A_3)$

Nótese que los cuerpos de las tablas son las matrices A y B .

Es sabido que conocidas las probabilidades de las acciones (probabilidades marginales) $P(A_i), \quad i = 1, 2, ..., n$ y $P(B_j), \quad j = 1, 2, ..., m$ siempre podemos determinar probabilidades de las intersecciones $P(A_i B_j), \quad i = 1, 2, ..., n; \quad j = 1, 2, ..., m$ (probabilidades conjuntas) de forma que, a su vez, reproduzcan las marginales, es decir

$$\sum_{i=1}^{n} P(A_i B_j) = P(B_j) \quad j = 1, 2, ..., m$$

$$\sum_{j=1}^{m} P(A_i B_j) = P(A_i) \quad i = 1, 2, ..., n$$

si bien esta determinación de probabilidades conjunta no es única. El estudio de este problema entra de lleno en la teoría de cópulas, tópico que no abordamos aquí.

A partir de las probabilidades conjuntas podemos determinar probabilidades de reacción.

Supongamos ahora que conocemos las probabilidades de reacción, $P(A_i \mid B_j),\quad i=1,2,\ldots,n;\quad j=1,2,\ldots,m$ y $P(B_j \mid A_i),\quad j=1,2,\ldots,m;\quad i=1,2,\ldots,n$ como sabemos, es posible determinar, bajo ciertas condiciones, las probabilidades de las acciones bajo ciertas circunstancias, más aún, las probabilidades de las acciones no tienen que ser únicas.

La compatibilidad entre las probabilidades de las reacciones revelaría que, entre los agentes, existe cierta coherencia respecto a las condiciones externas, digamos que revelaría cierta empatía entre ello, en cuanto a su percepción de la situación estudiada. Coherencia que no existe cuando las probabilidades son incompatibles.

La no unicidad de las acciones, revela el hecho de que distintas acciones pueden provocar las mismas reacciones.

Ejemplo 2:

Supongamos que en el ejemplo 1, se tienen las siguientes matrices de probabilidades de reacción.

$$A = \begin{bmatrix} 0.25 & 0.25 & 0.25 \\ 0.25 & 0.25 & 0.25 \\ 0.50 & 0.50 & 0.50 \end{bmatrix} \qquad B = \begin{bmatrix} 0.20 & 0.40 & 0.40 \\ 0.20 & 0.40 & 0.40 \\ 0.20 & 0.40 & 0.40 \end{bmatrix}$$

A partir de ahí, la matriz de cocientes viene dada por,

$$C = \begin{bmatrix} 1.250 & 0.625 & 0.625 \\ 1.250 & 0.625 & 0.625 \\ 2.500 & 1.250 & 1.250 \end{bmatrix}$$

Es fácil ver que $rang(C) = 1$y por lo tanto las reacciones son compatibles.

Siguiendo cualquiera de los métodos expuestos en los capítulos anteriores, se puede obtener que, las probabilidades de las acciones para α y β son:

$$\vec{\alpha} = \begin{bmatrix} 0.25 \\ 0.25 \\ 0.50 \end{bmatrix} \qquad \vec{\beta} = \begin{bmatrix} 0.20 \\ 0.20 \\ 0.40 \end{bmatrix}$$

Ejemplo 3:

Suponga que las probabilidades de reacción han sido,

$$A = \begin{bmatrix} 0.30 & 0.30 & 0.20 \\ 0.20 & 0.20 & 0.20 \\ 0.50 & 0.50 & 0.60 \end{bmatrix} \qquad B = \begin{bmatrix} 0.20 & 0.40 & 0.40 \\ 0.20 & 0.40 & 0.40 \\ 0.20 & 0.40 & 0.40 \end{bmatrix}$$

La matriz *C* es ahora,

$$C = \begin{bmatrix} 1.50 & 0.75 & 0.50 \\ 1.00 & 0.50 & 0.50 \\ 2.50 & 1.25 & 1.50 \end{bmatrix}$$

Cuyo rango es dos, por lo que no son compatibles. No existen probabilidades de las acciones compatibles con las probabilidades de reacción propuestas.

2.- Algunos casos particulares.

i.- Modelos 2×2

Pese a su sencillez, de todos es conocido, la enorme utilidad de los modelos definidos recurriendo a tablas dos por dos. En este contexto tendríamos, dos sucesos *S* y *T* y sus negaciones. Y las matrices *A* y *B* son

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \qquad B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$$

Caso 1.- Supongamos que todas las entradas de estas matrices son positivas, entonces la condición de compatibilidad es $\det(C) = 0$

$$\det(C) = \begin{vmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{vmatrix} = \begin{vmatrix} \dfrac{a_{11}}{b_{11}} & \dfrac{a_{12}}{b_{12}} \\ \dfrac{a_{21}}{b_{21}} & \dfrac{a_{22}}{b_{22}} \end{vmatrix} = 0 \leftrightarrow \frac{a_{11}}{b_{11}}\frac{a_{22}}{b_{22}} - \frac{a_{12}}{b_{12}}\frac{a_{21}}{b_{21}} = 0$$

o equivalentemente,

$$\frac{a_{11}a_{22}}{a_{12}a_{21}} = \frac{b_{11}b_{22}}{b_{12}b_{21}}$$

que es la condición de productos cruzados, estudiada en el primer capítulo.

Bajo el supuesto de compatibilidad, se tiene también la unicidad pues estamos suponiendo que las matrices A y B son estrictamente positivas, entonces se puede obtener fácilmente y por métodos elementales las probabilidades de las acciones:

$$P(S|T)P(T) = P(ST) = P(T|S)P(S) \rightarrow \frac{P(S|T)}{P(T|S)} = \frac{P(S)}{P(T)}$$

Análogamente,

$$P(\bar{S}|T)P(T) = P(\bar{S}T) = P(T|\bar{S})P(\bar{S}) \rightarrow \frac{P(\bar{S}|T)}{P(T|\bar{S})} = \frac{P(\bar{S})}{P(T)}$$

Sumando ambas igualdades:

$$\frac{P(\bar{S}\,|\,T)}{P(T\,|\,\bar{S})}+\frac{P(S\,|\,T)}{P(T\,|\,S)}=\frac{P(\bar{S})}{P(T)}+\frac{P(S)}{P(T)}=\frac{1}{P(T)}$$

Y por tanto,

$$P(T)=\left[\frac{P(\bar{S}\,|\,T)}{P(T\,|\,\bar{S})}+\frac{P(S\,|\,T)}{P(T\,|\,S)}\right]^{-1}$$

Una vez obtenido $P(T)$, la probabilidad $P(S)$ puede obtenerse de la siguiente manera

$$P(ST)=P(S\,|\,T)P(T)$$

de ahí

$$P(ST)=P(T\,|\,S)P(S)\rightarrow P(S)=\frac{P(S\,|\,T)}{P(T\,|\,S)}P(T)$$

$$P(S)=\frac{P(S\,|\,T)}{P(T\,|\,S)}\left[\frac{P(\bar{S}\,|\,T)}{P(T\,|\,\bar{S})}+\frac{P(S\,|\,T)}{P(T\,|\,S)}\right]^{-1}$$

Ejemplo 4:

Suponga que se tienen las matrices *A* y *B* correspondiente a un modelo 2×2

$$A = \begin{bmatrix} 0.20 & 0.20 \\ 0.80 & 0.80 \end{bmatrix} \qquad B = \begin{bmatrix} 0.50 & 0.50 \\ 0.50 & 0.50 \end{bmatrix}$$

La matriz *C* de cocientes puede comprobarse que tiene rango uno.

El cálculo de las probabilidades de las acciones se pueden calcular, utilizando métodos discutidos en capítulos anteriores, o bien por métodos elementales, utilizando las fórmulas que acabamos de deducir, para ello debemos observar:

$$A = \begin{bmatrix} P(S \mid T) & P(S \mid \bar{T}) \\ P(\bar{S} \mid T) & P(\bar{S} \mid \bar{T}) \end{bmatrix} = \begin{bmatrix} 0.20 & 0.20 \\ 0.80 & 0.80 \end{bmatrix}$$

$$B = \begin{bmatrix} P(T \mid S) & P(\bar{T} \mid S) \\ P(T \mid \bar{S}) & P(\bar{T} \mid \bar{S}) \end{bmatrix} = \begin{bmatrix} 0.50 & 0.50 \\ 0.50 & 0.50 \end{bmatrix}$$

Sustituyendo en las fórmulas anteriores, obtenemos $P(S) = 0.2,\, P(T) = 0.5$.

Caso 2.- Una única entrada de *A* es nula. Sin pérdida de generalidad supondremos que $a_{12} = 0$ y el resto de elementos de *A* son no nulos.

La condición necesaria de compatibilidad obliga a que $b_{12} = 0$ y el resto de elementos de *B* sean no nulos. Así obtenemos:

$$A = \begin{bmatrix} a_{11} & 0 \\ a_{21} & 1 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 0 \\ b_{21} & b_{22} \end{bmatrix}$$

Por su parte, la matriz *C* tiene un elemento indeterminado

$$C = \begin{bmatrix} c_{11} & * \\ c_{21} & c_{22} \end{bmatrix}$$

Completando de forma que tenga rango uno,

$$\det \begin{bmatrix} c_{11} & \lambda \\ c_{21} & c_{22} \end{bmatrix} = 0 \leftrightarrow \lambda = \frac{c_{11}c_{22}}{c_{21}}$$

Para ese valor, las probabilidades de las reacciones son compatibles y las probabilidades de las acciones únicas.

Bajo las hipótesis anteriores, $P(S \mid \bar{T}) = P(\bar{T} \mid S) = 0$ y el cálculo de las probabilidades de las acciones, siguiendo métodos elementales, verificarán:

$$P(S) = P(S \mid T)P(T)$$

$$P(T) = P(\bar{T} \mid S)P(S) + P(T \mid \bar{S})$$

De donde:

$$P(T) = \left(1 - P(\bar{T} \mid \bar{S})P(T \mid S)\right)^{-1} P(T \mid \bar{S})$$

$$P(S) = \left(1 - P(\bar{T} \mid \bar{S})P(T \mid S)\right)^{-1} P(T \mid \bar{S})P(S \mid T)$$

Ejemplo 5:

Suponga que se tienen las matrices A y B correspondiente a un modelo 2×2

$$A = \begin{bmatrix} 0.40 & 0 \\ 0.60 & 1.00 \end{bmatrix} \qquad B = \begin{bmatrix} 1.000 & 0 \\ 0.375 & 0.625 \end{bmatrix}$$

La matriz C de cocientes

$$C = \begin{bmatrix} 0.40 & * \\ 1.60 & 1.60 \end{bmatrix}$$

puede ampliarse de manera única de forma que tenga rango unidad:

$$C = \begin{bmatrix} 0.40 & 0.40 \\ 1.60 & 1.60 \end{bmatrix}$$

Por tanto A y B son compatibles.

El cálculo de las probabilidades de las acciones se pueden efectuarse, utilizando métodos discutidos en capítulos anteriores, o bien por métodos elementales, utilizando las fórmulas que acabamos de deducir, para ello debemos observar:

$$A = \begin{bmatrix} P(S \mid T) & P(S \mid \bar{T}) \\ P(\bar{S} \mid T) & P(\bar{S} \mid \bar{T}) \end{bmatrix} = \begin{bmatrix} 0.40 & 0 \\ 0.60 & 1.00 \end{bmatrix}$$

$$B = \begin{bmatrix} P(T \mid S) & P(\bar{T} \mid S) \\ P(T \mid \bar{S}) & P(\bar{T} \mid \bar{S}) \end{bmatrix} = \begin{bmatrix} 1.000 & 0 \\ 0.375 & 0.625 \end{bmatrix}$$

Sustituyendo en las fórmulas anteriores, obtenemos $P(S) = 0.2,\ P(T) = 0.5$.

Caso 3.- Dos elementos nulos. Los únicos casos posibles tienen los elementos nulos en distinta línea pues, en otro caso, las matrices *A* y *B* no serían estocásticas. Así, los únicos casos son:

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = B$$

O bien,

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = B$$

Sin embargo, cada caso se obtiene del otro por intercambio de filas y columnas, por lo que, desde un punto de vista teórico, son iguales. Analizamos el primero de ellos.

La matriz C está incompleta

$$C = \begin{bmatrix} 1 & * \\ * & 1 \end{bmatrix}$$

Puede ser completada con elementos positivos de forma que tenga rango unidad.

$$C = \begin{bmatrix} 1 & \lambda \\ \frac{1}{\lambda} & 1 \end{bmatrix}$$

Por lo que se tiene compatibilidad, aunque no se tiene unicidad. Es interesante hacer notar que, necesariamente se tiene: $P(S) = P(T)$ pues, bajo compatibilidad, $P(S) = P(S \mid T)P(T) + P(S \mid \bar{T})P(\bar{T}) = P(T)$. Cualquier distribución de probabilidad de las acciones es compatible con las probabilidades de la reacción propuestas. Obteniéndose la matriz M de las intersecciones:

$$M = \begin{bmatrix} P(ST) & P(S\bar{T}) \\ P(\bar{S}T) & P(\bar{S}\bar{T}) \end{bmatrix} = \begin{bmatrix} P(T) & 0 \\ 0 & P(\bar{T}) \end{bmatrix}$$

_ii_.- Modelo de comportamiento independiente

La siguiente situación corresponde, como es obvio, al caso en el que las acciones de los agentes son independientes.

$$P(A_i \mid B_j) = a_i, \quad i = 1,2,\dots,n; \quad j = 1,2,\dots,m$$

$$P(B_j \mid A_i) = b_j, \quad j = 1,2,\dots,m; \quad i = 1,2,\dots,n$$

$$A = \begin{bmatrix} a_1 & a_1 & \cdots & a_1 \\ a_2 & a_2 & \cdots & a_2 \\ \vdots & \vdots & \ddots & \vdots \\ a_n & a_n & \cdots & a_n \end{bmatrix} \qquad B = \begin{bmatrix} b_1 & b_2 & \cdots & b_m \\ b_1 & b_2 & \cdots & b_m \\ \vdots & \vdots & \ddots & \vdots \\ b_1 & b_2 & \cdots & b_m \end{bmatrix}$$

Esas probabilidades serían no nulas, pues en otro caso, las matrices tendrían una columna y una fila nulas (condición necesaria de compatibilidad) y la marginal, de existir, tendría un elemento con probabilidad nula y hubiese sido eliminado del problema**.**

Por otra parte, la compatibilidad está garantizada porque, la matriz *C* tiene rango uno, en efecto:

$$C = \begin{bmatrix} \frac{a_1}{b_1} & \frac{a_1}{b_2} & \cdots & \frac{a_1}{b_m} \\ \frac{a_2}{b_1} & \frac{a_2}{b_2} & \cdots & \frac{a_2}{b_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{a_n}{b_1} & \frac{a_n}{b_2} & \cdots & \frac{a_n}{b_m} \end{bmatrix}$$

Cualquier menor de orden dos tiene determinante nulo:

$$\begin{vmatrix} c_{ij} & c_{ik} \\ c_{lj} & c_{lk} \end{vmatrix} = \begin{vmatrix} \frac{a_i}{b_j} & \frac{a_i}{b_k} \\ \frac{a_l}{b_j} & \frac{a_l}{b_k} \end{vmatrix} = \frac{a_i}{b_j}\frac{a_l}{b_k} - \frac{a_l}{b_j}\frac{a_i}{b_k} = 0 \text{ para } i,l = 1,2,\ldots,n; \quad j,k = 1,2,\ldots,m$$

La unicidad es consecuencia de que los elementos de *A* y *B* no tienen elementos nulos.

Las probabilidades de las acciones son, obviamente:

$$P(A_i) = a_i, \quad i = 1,2,...,n;$$

$$P(B_j) = b_j, \quad j = 1,2,...,m$$

***iii*.- Modelo espejo.**

Con la notación anterior, supongamos que *n*=*m* y además que los A_i y B_j representan las mismas opciones para los dos agentes. Diremos que dos agentes siguen un modelo espejo si las reacciones de cada agente son coincidentes con las acciones del otro, esto es:

$$P(A_i \mid B_j) = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \quad i = 1,2,...,n$$

Si exigimos compatibilidad debe ser

$$P(B_i \mid A_j) = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \quad i = 1,2,...,n$$

Note que, en este caso, $A = B = I_n$

Las hipótesis anteriores implican que las distribuciones condicionadas son compatibles. En efecto, la matriz de cocientes en este caso tiene unos en la diagonal principal, mientras que el resto de elementos son desconocidos, pues los elementos correspondientes en las matrices *A* y *B* son nulos, para completar *C,* basta sustituir los elementos de la línea por encima de la diagonal principal por elementos no nulos desconocidos $x_i \in \mathbb{R}^+ \quad i = 1, 2, ..., n-1$

$$C = \begin{bmatrix} 1 & x_1 & * & ... & * & * \\ * & 1 & x_2 & ... & * & * \\ * & * & 1 & ... & * & * \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ * & * & * & ... & 1 & x_{n-1} \\ * & * & * & ... & * & 1 \end{bmatrix}$$

a partir de aquí completar la matriz bajo la condición de que tenga rango unidad. Se obtiene que los elementos de *C* verifican

$$c_{ij} = \begin{cases} 1 & si \quad 1 \le i = j \le n \\ \prod_{h=i}^{j-1} x_h & si \quad 1 \le i < j \le n \\ \dfrac{1}{c_{ji}} & si \quad 1 \le j < i \le n \end{cases}$$

Por tanto las condicionadas son compatibles. Por otra parte, puesto que son compatibles,

$$P(A_i) = \sum_{j=1}^{n} P(A_i \mid B_j) P(B_j) = P(A_i \mid B_i) P(B_i) = P(B_i)$$

Es decir, las probabilidades de acción son iguales para ambos agentes.

Las probabilidades conjuntas quedan:

$$P(A_i B_j) = P(A_i \mid B_j) P(B_j) = \begin{cases} P(B_j) & si \quad i = j \\ 0 & si \quad i \neq j \end{cases} \quad \forall i, j = 1,2,\dots,n$$

$$P(A_i B_j) = P(B_j \mid A_i) P(A_i) = \begin{cases} P(A_i) & si \quad i = j \\ 0 & si \quad i \neq j \end{cases} \quad \forall i, j = 1,2,\dots,n$$

Que obviamente coinciden.

Así, en el modelo espejo, las probabilidades de las reacciones son compatibles, aunque no determinan las probabilidades de acción de manera única, cualquier distribución de probabilidad sobre el conjunto de acciones es válida.

Ejemplo 6:

Consideremos el caso $n=4$, entonces las matrices A y B son la identidad 4×4. La matriz de cocientes inicial es

$$C = \begin{bmatrix} 1 & * & * & * \\ * & 1 & * & * \\ * & * & 1 & * \\ * & * & * & 1 \end{bmatrix}$$

Completando a rango 1, para x, y, z positivos:

$$C = \begin{bmatrix} 1 & x & xy & xyz \\ \frac{1}{x} & 1 & y & yz \\ \frac{1}{xy} & \frac{1}{y} & 1 & z \\ \frac{1}{xyz} & \frac{1}{yz} & \frac{1}{z} & 1 \end{bmatrix}$$

Cualquier distribución de probabilidad definida sobre las acciones puede ser considerada marginal.

Este modelo representa un comportamiento determinista de los agentes, siendo este, idéntico.

Observe que el caso en el que uno de los agentes tiene comportamiento espejo y el otro no, por ejemplo,

$$P(A_i \mid B_j) = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \quad i = 1, 2, \ldots, n$$

$$P(B_j \mid A_i) = b_{ij}, \quad i, j = 1, 2, \ldots, n;$$

donde alguna de las probabilidades anteriores verifica $b_{ij}(1 - b_{ij}) \neq 0$ tendríamos un modelo incompatible.

En caso de que el comportamiento sea determinista, aunque sobre acciones distintas, bastará permutar adecuadamente la numeración de las acciones para obtener los mismos resultados

***iv*.- Modelo de reacción 2-uniforme.**

$$P(A_i \mid B_j) = \frac{1}{n}, \quad i = 1, 2, ..., n; \quad j = 1, 2, ..., m$$

$$P(B_j \mid A_i) = \frac{1}{m}, \quad j = 1, 2, ..., m; \quad i = 1, 2, ..., n$$

Veamos bajo qué condiciones son compatibles:

$$A = \begin{bmatrix} \frac{1}{n} & \frac{1}{n} & \cdots & \frac{1}{n} \\ \frac{1}{n} & \frac{1}{n} & \cdots & \frac{1}{n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{n} & \frac{1}{n} & \cdots & \frac{1}{n} \end{bmatrix} \qquad B = \begin{bmatrix} \frac{1}{m} & \frac{1}{m} & \cdots & \frac{1}{m} \\ \frac{1}{m} & \frac{1}{m} & \cdots & \frac{1}{m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{m} & \frac{1}{m} & \cdots & \frac{1}{m} \end{bmatrix}$$

$$C = \begin{bmatrix} \frac{m}{n} & \frac{m}{n} & \cdots & \frac{m}{n} \\ \frac{m}{n} & \frac{m}{n} & \cdots & \frac{m}{n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{m}{n} & \frac{m}{n} & \cdots & \frac{m}{n} \end{bmatrix}$$

Puesto que *rang(C)=1* para todo *n* y *m*, el modelo 2-uniforme es compatible.

Calculamos las probabilidades de acción:

$$P(A_i B_j) = P(A_i \mid B_j) P(B_j) = \frac{1}{n} P(B_j) \quad \forall i, j = 1, 2, ..., n$$

Sumando en *j*, $P(A_i) = \frac{1}{n} \quad \forall i = 1, 2, ..., n$

Por otra parte, $P(A_i B_j) = P(B_j \mid A_i) P(A_i) = \frac{1}{m} P(A_i) \quad \forall i, j = 1, 2, ..., n$

Sumando en *i*, $P(B_j) = \frac{1}{m} \quad \forall j = 1, 2, ..., n$

Note que $P(A_i B_j) = \frac{1}{n} \frac{1}{m} \quad \forall i, j = 1, 2, ..., n$ lo que constituye un comportamiento independiente

Este modelo representa el caso en el que cada agente no está influenciado por las acciones del otro, reaccionando al azar.

v.- Modelo de reacción 1-uniforme.

$$P(A_i \mid B_j) = \frac{1}{n}, \quad i = 1,2,...,n; \quad j = 1,2,...,m$$

$$P(B_j \mid A_i) = b_{ij}, \quad j = 1,2,...,m; \quad i = 1,2,...,n$$

Note que una condición necesaria para que estas distribuciones sean compatibles es que $b_{ij} \neq 0, \quad j = 1,2,...,m; \quad i = 1,2,...,n$

Por otra parte, las matrices *A*, *B* y *C* quedan:

$$A = \begin{bmatrix} \frac{1}{n} & \frac{1}{n} & \cdots & \frac{1}{n} \\ \frac{1}{n} & \frac{1}{n} & \cdots & \frac{1}{n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{n} & \frac{1}{n} & \cdots & \frac{1}{n} \end{bmatrix} \qquad B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1m} \\ b_{21} & b_{22} & \cdots & b_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nm} \end{bmatrix}$$

$$C = \begin{bmatrix} \frac{1}{nb_{11}} & \frac{1}{nb_{12}} & \cdots & \frac{1}{nb_{1m}} \\ \frac{1}{nb_{21}} & \frac{1}{nb_{22}} & \cdots & \frac{1}{nb_{2m}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{nb_{n1}} & \frac{1}{nb_{n2}} & \cdots & \frac{1}{nb_{nm}} \end{bmatrix}$$

Observemos que $rang(C) = 1 \leftrightarrow rang(B) = 1$, por tanto esa última es la condición de compatibilidad.

Suponiendo que esa condición se verifica, calculemos las probabilidades de las acciones:

$$P(A_i B_j) = P(A_i \mid B_j) P(B_j) = \frac{1}{n} P(B_j) \quad \forall i, j = 1, 2, \ldots, n$$

Sumando en j, $P(A_i) = \frac{1}{n} \quad \forall i = 1, 2, \ldots, n$

Por otra parte,

$$P(B_j) = \sum_{i=1}^{n} P(B_j \mid A_i) P(A_i) = \sum_{i=1}^{n} P(B_j \mid A_i) \frac{1}{n} = \frac{1}{n} \sum_{i=1}^{n} b_{ij} \quad \forall j = 1, 2, \ldots, m$$

Que resulta ser la media de las filas de B.

Este modelo representa el caso en el que un agente no está influenciado por las acciones del otro, reaccionando al azar, mientras que el otro si lo está.

***vi*.- Modelo de comportamiento probabilístico idéntico.**

Supongamos que $n=m$ y además que los A_i y B_j representan las mismas opciones para los dos agentes. Diremos que dos agentes siguen un modelo de comportamiento probabilístico idéntico si las reacciones de cada agente ocurren con igual probabilidad ante las acciones del otro.

$$P(A_i \mid B_j) = P(B_j \mid A_i) = b_{ij}, \quad i, j = 1, 2, ..., n;$$

En este caso la condición necesaria de compatibilidad se verifica trivialmente. Además, $A=B$ y por tanto, estas matrices son biestocásticas. Los elementos de la matriz C que se pueden calcular son la unidad. Las distribuciones son compatibles pues, C admite, al menos, una extensión de rango uno que coincide con la matriz formada por unos. Este caso nos proporciona las siguientes probabilidades de las acciones:

$$P(A_i) = P(B_i) = \frac{1}{n}, \quad i = 1, 2, ..., n$$

Sin más que observar que la matriz de constantes de proporcionalidad por filas K_f está formada por unos.

La unicidad deberá ser estudiada en cada caso pues, como sabemos, dependerá de la posición de los ceros si los hay, ya que es obvio que este modelo proporciona un único par de conjuntos de probabilidades de acciones si no hay ceros.

Ejemplo 7:

Consideremos las matrices de reacciones:

$$A = B = \begin{bmatrix} 0.1 & 0.6 & 0.3 & 0 & 0 \\ 0.5 & 0.1 & 0.4 & 0 & 0 \\ 0.4 & 0.3 & 0.3 & 0 & 0 \\ 0 & 0 & 0 & 0.2 & 0.8 \\ 0 & 0 & 0 & 0.8 & 0.2 \end{bmatrix}$$

Nótese que es biestocástica.

La matriz (incompleta) de cocientes C es:

$$C=\begin{bmatrix}1&1&1&*&*\\1&1&1&*&*\\1&1&1&*&*\\*&*&*&1&1\\*&*&*&1&1\end{bmatrix}$$

Que puede ser completada de manera que tenga rango unidad como

$$C=\begin{bmatrix}1&1&1&\lambda&\lambda\\1&1&1&\lambda&\lambda\\1&1&1&\lambda&\lambda\\\frac{1}{\lambda}&\frac{1}{\lambda}&\frac{1}{\lambda}&1&1\\\frac{1}{\lambda}&\frac{1}{\lambda}&\frac{1}{\lambda}&1&1\end{bmatrix}$$

para cada $\lambda>0$. Por lo tanto, no es única y tampoco lo son las marginales. El cálculo de las marginales proporciona que las probabilidades de las acciones de β y α vienen dadas por los vectores:

$$\vec{\beta}=\begin{bmatrix}\frac{1}{3+2\lambda}\\\frac{1}{3+2\lambda}\\\frac{1}{3+2\lambda}\\\frac{\lambda}{3+2\lambda}\\\frac{\lambda}{3+2\lambda}\end{bmatrix}\qquad\qquad\vec{\alpha}=\begin{bmatrix}\frac{\lambda}{2+3\lambda}\\\frac{\lambda}{2+3\lambda}\\\frac{\lambda}{2+3\lambda}\\\frac{1}{2+3\lambda}\\\frac{1}{2+3\lambda}\end{bmatrix}$$

Para $\lambda = 1$ se obtienen las distribuciones uniformes en cada agente.

Observemos que corresponde al caso general del comportamiento espejo.

3.- Algunas aplicaciones.

El modelo propuesto, puede aplicarse a gran variedad de situaciones prácticas.

a.- Publicidad de respuesta directa.

Recordemos que la publicidad de respuesta es un tipo de campaña publicitaria en el que se requiere cierto tipo de respuesta. En él, las acciones son, por un lado el visionado del spot, y por otro la reacción inmediata del cliente. Se trata por tanto de un modelo dos por dos.

b.- Fijación de estrategia de precios.

Véase los ejemplos de la sección uno.

c.- Formación de equipos.

Se trata de estudiar la inclusión en un equipo de dos individuos determinados *A* y *B.*

d.- Pruebas diagnósticas generalizadas.

El análisis de pruebas diagnósticas estudia las probabilidades de los modelos dos por dos cuando las acciones son, por un lado enfermo, sano y por otro test positivo, test negativo. Esto es susceptible de ser generalizado para el caso en el test puede presentar más niveles.

Conclusiones y Futuras Líneas de Investigación.

Conclusiones.

En este trabajo, se ha abordado el problema de la compatibilidad de dos distribuciones de probabilidad para el caso de vectores bidimensionales, discretos y finitos. Encontramos una fuerte relación existente con el rango de cierta matriz (completa o no) de cocientes. Esta relación nos proporciona, no sólo información sobre la compatibilidad de las distribuciones condicionadas si no, además, métodos de cálculo de tal distribución, bien directamente o mediante otras matrices, que hemos denominado matrices de constantes de proporcionalidad. Este método puede aplicarse al estudio de la estructura de las matrices condicionadas que indican ciertas propiedades probabilísticas de las distribuciones marginales. Se han revisado, también, algunos métodos alternativos.

Respecto a la unicidad, se ha mostrado que el conjunto de soluciones de un problema de compatibilidad, es convexo. A continuación se mostró que la unicidad está relacionada con la situación de los ceros en las matrices de probabilidad condicionadas. Resultando que tanto teoría de grafos, como la estructura de la matriz de cocientes son poderosas herramientas para estudiar esta cuestión.

La descomposición en valores singulares de la matriz de cocientes, puede ser utilizada para para encontrar de mínima incompatibilidad, en el sentido de ciertas normas matriciales.

Se ha propuesto una aplicación de las cuestiones que nos ocupan al estudio de la coherencia entre las decisiones de dos agentes que reacciones ante acciones del otro.

Futuras Líneas de Investigación.

1.- Aplicación de las diversas propiedades de los conjuntos convexos al estudio de las propiedades probabilísticas de las distribuciones marginales y conjuntas asociadas al problema que nos ocupa, supuesta la compatibilidad.

2.- Extensión de resultados a problemas de mayor dimensión.

3.- Cálculo de distribuciones de mínima compatibilidad, utilizando la descomposición en valores singulares de la matriz de cocientes al caso general.

5.- Aplicaciones reales del modelo teórico de acción y reacción.

6.- Extensión de los métodos estudiados al caso de variables discretas de rango infinito.

Anexo I

Otros métodos.

Este anexo contiene material complementario. Concretamente tres enfoques para el análisis de los problemas de compatibilidad, unicidad e incompatibilidad que no han sido tratados en este trabajos.

1.- Uso de programación lineal en problemas de compatibilidad.

El uso de sistemas lineales y métodos de programación lineal para estudiar la compatibilidad y unicidad de matrices, ha sido estudiado en el capítulo 3, completamos aquí ese estudio mostrando el trabajo de Ghosh y Nadarajah (2014).

Cálculo de distribuciones marginales:

Supongamos que $n \le m$ (si $n>m$ entonces reemplaza A y B por sus transpuestas).

Supongamos que A y B son compatibles, entonces, con la notación habitual, la probabilidad conjunta p_{ij} verifica, $p_{ij} = b_{ij}\alpha_i$, $i = 1,2,...,n$; $j = 1,2,...,m$ y además,

$$a_{ij} = \frac{p_{ij}}{\sum_{s=1}^{n} p_{sj}} = \frac{b_{ij}\alpha_i}{\sum_{s=}^{n} b_{sj}\alpha_s}, \quad i = 1,2,...,n; \; j = 1,2,...,m$$

$$\beta_j = \sum_{s=}^{n} b_{sj}\alpha_s \quad j = 1,2,...,m$$

Por tanto la marginal $\vec{\beta}$ puede calcularse a partir de $\vec{\alpha}$

$$a_{ij}\sum_{s=}^{n} b_{sj}\alpha_s - b_{ij}\alpha_i = 0, \quad i = 1,2,...,n; \; j = 1,2,...,m$$

El sistema lineal anterior puede escribirse en forma matricial como $D\vec{\alpha} = \vec{0}_{nm}$, donde *D* es una matriz *nm*×*n*. $\vec{0}_{nm}$ denota, como en otras ocasiones, el vector columna nulo de dimensión *nm*.

Por otra parte, $D\vec{\alpha} = \vec{0}_{nm}$ es un sistema homogéneo de *nm* ecuaciones con *n* incógnitas, que se puede reducir a un sistema de *r* ecuaciones linealmente independientes, utilizando, por ejemplo su forma escalonada reducida por filas (ver, por ejemplo, Lay, 2007) o por cualquier otro método, obteniendo $D_r\vec{\alpha} = \vec{0}_{nm}$

Observación: en rigor, según el método utilizado para la reducción de *D* a D_r es posible que las coordenadas de $\vec{\alpha}$ cambien respecto a su posición inicial, en lo que sigue, supondremos que no.

Consideramos el sistema de ecuaciones lineales $D_r\vec{\gamma} = \vec{0}_{nm}$ donde $\vec{\gamma}$ es un vector columna de *n* incognitas.

Cuando A y B son compatibles, existe una solución no negativa con, al menos un elemento no nulo $\vec{\gamma}^*$ que puede normalizarse para obtener un vector de probabilidad que coincidirá con la marginal $\vec{\alpha}$.

Cuando A y B son incompatibles, la solución de $D_r\vec{\gamma} = \vec{0}_{nm}$ es el vector nulo.

Para estudiar si existe la solución $\vec{\gamma}^*$ podemos usar métodos de programación lineal. Consideremos el siguiente programa:

$$\begin{aligned}
&\text{máx } \sum_{i=1}^{n} \alpha_i \\
\text{S.A.} \quad & D_r\vec{\alpha} = \vec{0}_{nm} \\
& \alpha_i \geq 0, i = 1,2,\ldots,n \\
& \sum_{i=1}^{n} \alpha_i \leq 1
\end{aligned}$$

Si el máximo de la función objetivo es estrictamente positivo entonces, por normalización del vector solución se obtiene la marginal $\vec{\alpha}$, y por tanto, A y B son compatibles.
Si el máximo es cero, A y B son incompatibles

Teorema I.1.1. (Ghosh y Nadarajah, 2014). Las matrices A y B son compatibles si $rang(D) \leq n-1$, se tiene la igualdad cuando existe una única solución.

Cálculo de la distribución conjunta.

Supongamos que A y B son compatibles y P la distribución condicionada compatible con A y B. El sistema a utilizar se deduce de:

$$a_{ij} p_{\bullet j} = b_{ij} p_{i\bullet},\ i = 1,2,...,n;\ j = 1,2,...,m$$

$$a_{ij} \sum_{s=1}^{n} p_{sj} - b_{ij} \sum_{k=1}^{m} p_{ik} = 0,\ i = 1,2,...,n;\ j = 1,2,...,m$$

Y supongamos que, en forma matricial adopta la forma $H\vec{p} = \vec{0}_{nm}$ donde H es una matriz $nm \times nm$ y $\vec{p} = \left(p_{11}, p_{12}, ..., p_{nm} \right)^t$ es un vector, de dimensión nm, que contiene las probabilidades conjuntas, escritas en orden lexicográfico de sus subíndices.
Ghosh y Nadarajah (2014) prueban,

Teorema I.1.2 (Ghosh y Nadarajah, 2014). El espacio de soluciones Ω del sistema $H\vec{p} = \vec{0}_{nm}$ es $(I - M)\vec{z}$, donde M es idempotente ($M^2 = M$) y $\vec{z}$ es un vector columna de dimensión nm.

La matriz del M del teorema anterior viene dada por $M = H^{-}H$ donde H^{-} es la g-inversa de C.

Teorema I.1.3 (Ghosh y Nadarajah, 2014). Si existe un vector del espacio de soluciones Ω tal que todos sus elementos son no negativos y al menos uno es positivo, entonces A y B son compatibles.

Teorema I.1.4 (Ghosh y Nadarajah, 2014). Si cualquier vector del espacio de soluciones Ω es no positivo, entonces A y B son incompatibles.
Consideremos el siguiente programa lineal.

$$\text{máx} \sum_{i,j} p_{ij}$$

$$\begin{aligned} \text{S.A.} \quad & H\vec{p} = \vec{0}_{nm} \\ & 0 \le p_{ij} \le 1, \qquad i = 1,2,...,n;\ j = 1,2,...,m \end{aligned}$$

Entonces se tiene:

Teorema I.1.5 (Ghosh y Nadarajah, 2014). El valor máximo del programa lineal anterior es estrictamente positivo si y sólo si las matrices *A* y *B* son compatibles.

2.- Uso de Grafos Ponderados. El método gráfico.

El siguiente método, básicamente, una formalización del método basado en la fórmula de Bessag, que estudiamos en los capítulos 1, 2 y 3.

Consideremos un grafo (*V*,*E*) donde la arista de vértices *u* y *v* se denota $\{u,v\}$. A cada arista $\{u,v\}$ le asociamos un cociente o ratio $\frac{r(u,v)}{r(v,u)}$ donde $r(u,v)$ y $r(v,u)$ son números positivos. Sea *R=R(E)* el conjunto de ratios.

De esta forma *(V,E,R)* es un grafo ponderado que Yao, Chen y Wang (2014) llaman representación gráfica.

Sean *A* y *B* dos matrices en las condiciones habituales y que verifican la condición necesaria de compatibilidad $N_A = N_B = N$. El conjunto de vértices es $V = N$. El conjunto de aristas se define como sigue. Sean $(x,y),(x',y') \in V$ distintos $\{(x,y),(x',y')\} \in E$ si y sólo si $x = x'$ o $y = y'$

Los ratios se construyen de la siguiente forma

$$\frac{r((x,y),(x,y'))}{r((x,y'),(x,y))} = \frac{b_{xy}}{b_{xy'}}, \qquad y \neq y'$$

$$\frac{r((x,y),(x',y))}{r((x',y),(x,y))} = \frac{a_{xy}}{a_{x'y}}, \qquad x \neq x'$$

Al grafo ponderado *(V,E,R)* se le denomina representación gráfica del problema de compatibilidad.

Definición 1. Diremos que el conjunto de ratios *R* es compatible si y sólo si existe una distribución de probabilidad positiva sobre *V* tal que

$$\frac{p(x,y)}{p(x',y')}=\frac{r((x,y),(x',y'))}{r((x',y'),(x,y))}, \qquad \forall (x,y),(x',y')\in E$$

Diremos así mismo que *p* satisface *R*.

Teorema I.2.1. **(**Yao, Chen y Wang, 2014) Sean *A* y *B* dos matrices en las condiciones habituales y que verifican la condición necesaria de compatibilidad y sea *(V,E,R)* la representación gráfica asociada. Entonces, una distribución con soporte *V* tiene condicionadas *A* y *B* si y sólo su conjunto de ratios es *R*.

En consecuencia, *A* y *B* son compatibles si y sólo si *R* lo es.

Definición 2. Dos representaciones gráficas *(V,E,R)* y *(V,E',R')* con el mismo conjunto de vértices se dicen equivalentes sii

i.- *R* y *R'* coinciden sobre $E\cap E'$.

ii.- Para todo $\{u,v\}\in E-E'$ existe un camino de *E'*, $u=v_0v_1...v_{k-1}v_k=v$ con $\{v_l,v_{l+1}\}\in E',\ \ l=0,1,...,k-1$ tales que

$$\frac{r(u,v)}{r(v,u)}=\prod_{l=0}^{k-1}\frac{r'(v_{l+1},v_l)}{r'(v_l,v_{l+1})}$$

iii.- Para $\{u,v\}\in E'-E$ existe un camino de *E*, $u=v_0v_1...v_{k-1}v_k=v$ con $\{v_l,v_{l+1}\}\in E,\ \ l=0,1,...,k-1$ tales que

$$\frac{r'(u,v)}{r'(v,u)}=\prod_{l=0}^{k-1}\frac{r(v_{l+1},v_l)}{r(v_l,v_{l+1})}$$

Informalmente, *(V,E,R)* y *(V,E',R')* con el mismo conjunto de vértices se dicen equivalentes si los conjuntos de ratios coinciden en las aristas comunes y los ratios en un mismo conjunto de vértices *E* (*E'*) pueden obtenerse a partir de ratios en el otro conjunto. Puede utilizarse para simplificar representaciones.

Lema I.2.1 (Yao, Chen y Wang, 2014)**.** Suponga que *(V,E,R)* y *(V,E',R')* con el mismo conjunto de vértices son equivalentes. Entonces una distribución de probabilidad positiva sobre *V* satisface *R* si y sólo si satisface *R'*.

En consecuencia, *R* es compatible si y sólo si *R'* es compatible.

Denotamos $\coprod$ a la unión disjunta de grafos, esto es, los grafos que se unen no tienen ni vértices ni aristas comunes.

Teorema I.2.2 (Yao, Chen y Wang, 2014) **.** Se considera la representación gráfica $(V,E,R)=\coprod_{i=1}^{k}(V_i,E_i,R_i)$ donde cada (V_i,E_i) es un subgrafo conexo de (V,E) y R_i es la restricción de *R* a E_i. Entonces, *R* es compatibles si los R_i son compatibles *i*=1,2,…,*n*.

Observación. Las relaciones entre las distribuciones de probabilidad que hacen compatibles a *R* y R_i son las siguientes.

Si *R* es compatible y *p* es la distribución de probabilidad sobre *V* que satisface *R*, las distribuciones de probabilidad sobre V_i, $p_i(v)=\dfrac{p(v)}{\sum_{v_i\in V_i}p(v_i)}$, $i=1,2,...,k$ satisfacen R_i.

Si R_i son compatibles y las distribuciones de probabilidad sobre V_i, p_i satisfacen R_i, entonces, para cada $\lambda_1,\lambda_2,...,\lambda_k\geq 0$ con $\sum_{i=1}^{k}\lambda_i=1$, la distribución de probabilidad sobre *V*, $p_\lambda(v)=\sum_{i=1}^{k}\lambda_i p_i(v)1_{V_i}(v)$, verifica *R*.

Nótese que si k>1 la solución, si existe no es única.

Teorema I.2.3 (Yao, Chen y Wang, 2014)**.** Para cada representación *(V,E,R),* las siguientes condiciones son equivalentes.

i.- R es compatible.

ii.- Para cualquier par de caminos $v_0v_1v_2...v_l$ y $w_0w_1w_2...w_s$ con $v_0 = w_0$ y $v_l = w_s$ se verifica que

$$\prod_{i=0}^{l-1}\frac{r(v_{i+1},v_i)}{r(v_i,v_{i+1})}=\prod_{i=0}^{s-1}\frac{r(w_{i+1},w_i)}{r(w_i,w_{i+1})}$$

iii.- Para cualquier camino $v_0v_1v_2...v_lv_0$,

$$\prod_{i=0}^{l}\frac{r(v_{i+1},v_i)}{r(v_i,v_{i+1})}=1 \text{ donde } v_{l+1}=v_0$$

La condición *ii* indica que los productos de ratios a través de caminos que unan los mismos vértices coinciden. La condición *iii* indica que el producto de los ratios a través de cualquier camino cerrado (ciclo) vale uno.

Este método puede ser extendido sin dificultad a más de dos variables.

3.- Soluciones aproximadas. Muestreo de Gibbs.

Una de las técnicas de simulación más relacionadas con la especificación condicional es el denominado Muestreo de Gibbs.

El muestreo de Gibbs es un algoritmo que permite generar muestras de la variable aleatoria X, componente de un vector aleatorio, utilizando las distribuciones condicionadas univariantes de dicho vector.

En el caso de un vector bidimensional, caso que aquí nos interesa, obtenemos una muestra de X a partir de $f(x|y)$ y $f(y|x)$.

Estas líneas están basadas en el trabajo de Casella y George (1992).

Sea (X,Y) un vector aleatorio bidimensional, y supongamos conocidas $f(x|y)$ $f(y|x)$ las funciones de densidad o probabilidad de las variables $X|Y=y$ y $Y|X=x$. Generaremos un valor de la variable aleatoria X a partir de estas distribuciones condicionadas. Para ello, construimos la sucesión (sucesión de Gibbs)

$$Y_0^{'}\ X_0^{'}\ Y_1^{'}\ X_1^{'}\ Y_2^{'}\ X_2^{'}...Y_k^{'}X_k^{'}$$

de la siguiente forma:

0.- $j = 0$

1.- Considerar un valor cualquiera y'_j de la variable Y.

2.- Generar un valor x'_j de la variable aleatoria $X\,|\,Y = y'_j$ utilizando $f(x\,|\,y'_j)$.

3.- Generar un valor y'_{j+1} de la variable aleatoria $Y\,|\,X = x'_j$ utilizando $f(y\,|\,x'_j)$.

4.- Para $j: = j+1$ volver al paso 2.

Puede probarse que bajo condiciones razonables (Geman y Geman 1984), X'_k tiende en distribución a X. Así, para k suficientemente grande, podemos suponer que $X'_k = x'_k$ es una realización muestral de X. La elección de k , es decir, de la longitud de la sucesión de Gibbs que haga aceptable que X'_k y X son idénticamente distribuidas es un problema complicado que ha recibido varias propuestas de soluciones (Geffand y Smith (1990), Tanner (1991)).

Para la formación de una muestra de tamaño *m* usando este algoritmo, se han propuesto diversas soluciones. A nuestro entender, la más razonable, es generar *m* sucesiones de Gibbs, indepencdientes,

$$y_0^{(l)}, x_0^{(l)}, y_1^{(l)}, x_1^{(l)}, ..., y_k^{(l)} x_k^{(l)} \qquad\qquad l = 1...m$$

Es decir, realizar las elecciones de los valores iniciales de forma aleatoria e independiente $y_0^l,\ l = 1...m$, la muestra estaría formada por $x_k^{(1)}, x_k^{(2)}, ..., x_k^{(m)}$.

Es claro, que toda inferencia basada en el muestreo de Gibbs, pasa por un estudio previo, de la convergencia del método y de la compatibilidad de las distribuciones condicionadas en las que se basa.

Chen, Ip y Wang (2011) utilizan el muestreo de Gibbs para estudiar la compatibilidad de las distribuciones condicionadas y obtener una solución aproximada a la distribución conjunta. Utilizaremos la misma notación que los autores aunque su construcción se realiza para vectores aleatorios discretos de cualquier dimensión.

Sean P_1 y P_2 las distribuciones condicionadas de las variables $X_1 | X_2$ y $X_2 | X_1$ respectivamente. Se llama conjunto de Gibbs al conjunto formado por las distribuciones $\hat{\pi}^{(1,2)}$ y $\hat{\pi}^{(2,1)}$ donde:

$\hat{\pi}^{(1,2)}$ es la distribución empírica de $\left\{(x_1^{2h\omega+1}, x_2^{2h\omega+1}) : 1 \le h \le l\right\}$ donde *l* es el número de ciclos completos generados según el siguiente esquema: generar $x_2^{(0)}$, y a partir de él obtener:

$$x_1^{(1)} \sim X_1 | X_2 = x_2^{(0)} \to x_2^{(1)} \sim X_2 | X_1 = x_1^{(1)} \to x_1^{(2)} \sim X_1 | X_2 = x_2^{(1)} \to x_2^{(2)} \sim X_2 | X_1 = x_1^{(2)} \to ...$$

$\hat{\pi}^{(2,1)}$ es la distribución empírica de $\left\{(x_1^{2h\omega+1}, x_2^{2h\omega+1}) : 1 \leq h \leq l\right\}$ donde *l* es el número de ciclos completos generados según el siguiente esquema: generar $x_1^{(0)}$, y a partir de él obtener:

$$x_2^{(1)} \sim X_2 \mid X_1 = x_1^{(0)} \rightarrow x_1^{(1)} \sim X_1 \mid X_2 = x_2^{(1)} \rightarrow x_2^{(2)} \sim X_2 \mid X_1 = x_1^{(1)} \rightarrow x_1^{(2)} \sim X_1 \mid X_2 = x_2^{(2)} \rightarrow \ldots$$

donde ω es un valor entero positivo, que nos permite saltar ciclos.

A partir de las distribuciones empíricas anteriores, obtenemos las condicionadas empíricas, $\hat{\pi}_{1|2}^{(1,2)}$ y $\hat{\pi}_{2|1}^{(1,2)}$ para $\hat{\pi}^{(1,2)}$ y $\hat{\pi}_{1|2}^{(2,1)}$ y $\hat{\pi}_{2|1}^{(2,1)}$ para $\hat{\pi}^{(2,1)}$ y el grado de separación mediante una medida de discrepancia *D* (ver más abajo), como

$e_D^{(1,2)} = D(\hat{\pi}_{1|2}^{(1,2)}, P_1) + D(\hat{\pi}_{2|1}^{(1,2)}, P_2)$, que se interpreta como el error (según *D*) entre $\hat{\pi}^{(1,2)}$ y la distribución conjunta cuyas condicionadas son P_1 y P_2.

$e_D^{(2,1)} = D(\hat{\pi}_{1|2}^{(2,1)}, P_1) + D(\hat{\pi}_{2|1}^{(2,1)}, P_2)$, que se interpreta como el error (según *D*) entre $\hat{\pi}^{(2,1)}$ y la distribución conjunta cuyas condicionadas son P_1 y P_2.

Definimos los pesos, para $\hat{\pi}^{(1,2)}$ y $\hat{\pi}^{(2,1)}$ vienen dados por

$$\varepsilon_D^{(1,2)} = \frac{1 / e_D^{(1,2)}}{E_D}, \qquad \varepsilon_D^{(2,1)} = \frac{1 / e_D^{(2,1)}}{E_D}$$

Respectivamente. Donde $E_D = \frac{1}{e_D^{(1,2)}} + \frac{1}{e_D^{(2,1)}}$,

Nótese que a mayor error, menor peso. Una aproximación a la distribución conjunta, que se conoce con el nombre de solución de Gibbs es:

$$\tilde{\pi}_D = e_D^{(1,2)} \hat{\pi}^{(1,2)} + e_D^{(2,1)} \hat{\pi}^{(2,1)}$$

Entre las medidas *D*, utilizadas habitualmente para medir la discrepancia entre dos distribuciones de probabilidad finitas $p=(p_1, p_2, ..., p_k)$ y $p'=(p'_1, p'_2, ..., p'_k)$ se encuentran:

$L^2=\sum(p'_i-p_i)^2$	Euclidea
$F^2=\sum(\sqrt{p'_i}-\sqrt{p_i})^2$	Freemann-Tuckey
$\chi^2=\sum\frac{(p'_i-p_i)^2}{p'_i}$	Pearson
$N^2=\sum\frac{(p'_i-p_i)^2}{p_i}$	Neyman
$I^2=\sum p'_i\log\left(\frac{p'_i}{p_i}\right)$	Información
$G^2=\sum p_i\log\left(\frac{p_i}{p'_i}\right)$	Kullback- Leibler.

Así mismo, si no se quiere tener en cuenta el error, puede utilizarse como solución de Gibbs viene dada por:

$$\tilde{\pi}_{EQ}=\frac{1}{2}\left(\hat{\pi}^{(1,2)}+\hat{\pi}^{(2,1)}\right)$$

Este método puede ser extendido sin dificultad a más de dos variables.

Anexo II

Algoritmos y códigos.

Este anexo contiene las algunas referencias que contienen algoritmos, en algunos casos con el código correspondiente, que resuelven varios problemas tratados en este trabajo.

Compatibilidad.

Arnold, Castillo y Sarabia (1994) y con más detalles, Arnold, Castillo y Sarabia (1999) pg. 22, proporciona un algoritmo de cálculo de la representación marginal uniforme de una matriz.

Arnold, Castillo y Sarabia (2004), proporcionan un algoritmo para chequear la compatibilidad e vectores tridimensionales utilizando el método de extensión de rango unidad.

Song et al. (2010), proporciona algoritmo de compatibilidad usando la descomposición en bloques irreducibles de la matriz de cocientes *C* y, asumiendo la compatibilidad, obtener la distribución conjunta. También proporcionan un algoritmo para calcular la descomposición en bloques irreducibles y, a partir de ella obtener la extensión positiva de *C*.

Tian et al (2009) ofrecen el código de una función S-Plus que calcula la estimación de mínimos cuadrados restringidos al cubo unidad.

Wang y Kuo (2010) obtiene algoritmos para comprobar compatibilidad y, en su caso, calcula la distribución conjunta (basados en la fórmula de Besag). Así mismo proporciona código MATLAB.

Unicidad.

Song et al. (2010), proporciona algoritmo para comprobar la unicidad usando la descomposición en bloques irreducibles de la matriz de cocientes *C*.

Incompatibilidad.

Arnold, Castillo y Sarabia (1999) pg. 33 proporciona un algoritmo que proporciona la distribución de mínima incompatibilidad basada en la divergencia de Kullback-Leibler.

Castillo et al. (2002) contiene, entre otros, el algoritmo gamma para el cálculo del cono dual.

Wang y Kuo (2010) obtiene algoritmos para calcular una distribución próxima a la compatibilidad en el caso de incompatibilidad. Esta distribución se basa en la fórmula de Besag. Así mismo proporciona código MATLAB.

Yao, Chen, Wang (2014), ofrecen un algoritmo para estudiar la compatibilidad utilizando grafos ponderados.

REFERENCIAS Y BIBLIOGRAFÍA.

Abrahams J. y Thomas J. B. (1984) A note on the characterization of bivariate densities by conditional densities. *Comunications in Statistics: Theory and Methods,. 13(3), 395-400*

Amaya, J. (2013) Optimización. Disponible en http://escuela.mate304.org/NotasOptimizaci%C3%B3n_Convexa.pdf (última consulta: 9-8-2015)

Arnold B.C. Strauss D. (1991) Bivariate distribution with conditionals in prescribed exponential families. *Journal of the Royal Statistical Association.53, 365-375.*

Arnold B.C., Castillo E. Sarabia J.M. (1992) *Conditionally specified distributions.* Lectures Notes in Statistics, Volume 73, Springer Verlang. Berlin.

Arnold B.C., Castillo E. Sarabia J.M. . (1999). *Conditional Specification of Statistical Models* Springer

Arnold B.C., Castillo E. y Sarabia J.M.(2001) Conditionally Specified Distributions: An Introduction. *Statistical Science*, Vol. 16, No. 3, 249–274

Arnold B.C., Castillo E. y Sarabia J.M. (2002) Exact and near compatibility of discrete conditional distributions *Computational Statistics & Data Analysis* 40 231–252

Arnold, B. C., Castillo, E. and Sarabia, J. M. (2004). Compatibility of partial or complete conditional probability specifications. *Journal of Statistical Planning and Inference* 123, 133-159.

Arnold, B. C. Gokhale, D. V. (1994) On uniform marginal representation of contingency tables. *Statistics & Probability Letters* 21, 311-316.

Arnold, B. C. Gokhale, D. V. (1998). Distributions of the most nearly compatible with given families of conditional distributions. *Test* 7, 377-390.

Arnold B.C. Pourahmadi M. (1988) Conditional characterization of multivariate distributions. *Metrika. 35, 99-108.*

Arnold B.C. Press S. J. (1989) Compatible conditional distributions. *Journal of the American Statistical Association. 84, 152-156*

Ahsanullah M. (1985) Some characterizations of the bivariate normal distribution. *Metrika. 32, 215-218.*

Bapat, R. B. (2010) *Graphs and Matrices*, Springer

Besag, J. E. (1972). Nearest-neighboursystems and the autologistic model for binary data. *Journal of the Royal Statistical Society, Series B: Methodological*, 34:75–83.

Besag,. J. E. (1974) Spatial Interaction and the Statistical Analysis of Lattice Systems *Journal of the Royal Statistical Society*. Series B (Methodological), Vol. 36, No. 2: 192-236.

Bhattacharyya A. (1943) On some sets of sufficients conditions leading to the normal bivariate distribution. *Sankhyá, 6, Part 4, 499-406.*

Brucker J. (1979) A note on the bivariate normal distribution. *Comunications in Statistics: Theory and Methods. 8, 175-177.*

Casella G. y George E.I. (1992) Explaining the Gibbs Sampling. *The American Statistician 46(3), 167-174.*

Castillo, E., Cobo, A., Jubete, F., Pruneda, R. E. (1998). Orthogonal Sets and Polar Methods in Linear Algebra: Applications to Matrix Calculations, en *Systems of Equations and Inequalities, and Linear Programming*. Wiley, New York.

Castillo E. Galambos J. (1987) Bivariate distributions with normal conditionals. *Procedings of the Internationals Association of Science for Developmen,* International Symposium on Simulation, Modeling an development, Cairo, Egypt, Acta Press, Anaheim, CA,59-62.

Castillo, E. Galambos, J.(1989) Conditional distributions and the bivariate normal distribution, *Metrika, 36 , 209-214.*

Castillo, E., Julete, F. Pruneda, R.E. Solares, C. (2002) Obtaining somultaneous solutions of linear subsystems of inequalities and duals. *Linear Algebra and its Applications 346, 131-154.*

Chen, S-H, Ip, E.H., Wang, Y.J. (2011) Gibbs ensembles for nearly compatible and incompatible conditional. *Computational Statistics and Data Analysis*.55, 1760-1769.

Diestel, R, (2010) *Graph Theory* Cuarta Edición Springer

Fraser D.A.S. y Streit F. (1980) A futher note on the bivariate normal distribution. *Comunications in Statistics: Theory and Methods.9 , 1097-1099.*

Gelfand A.E. y Smith A.F.M. (1990) Sampling based approaches to calculating marginal densities. *Journal of the American Statistical Association.,85,398-409.*

Gelman A. y Meng X.L. (1991) A note on bivariate distributions that are conditionally normal. *The American Statistician, 45(2), 125-126.*

Geman S. y Geman D. (1984) Stochastic relaxation, Gibbs distribution and the bayesian restoration of images. *IEEE Transaction on Pattern Analysis and Machine Intelligence 6, 721-741.*

Ghosh, I. Balakrishnan N. (2013) Study of incompatibility or near compatibility of bivariate discrete conditional probability distributions through divergence measures. *Journal of Statistical Computation and Simulation* 84 (1), 1-14.

Ghosh, I. Nadarajah, S. (2014) An alternative approach for compatibility of two discrete conditional distributions. *Comunications in Statistics: Theory and Methods* (por aparecer).

Gourieroux Ch. y Monfort A. (1979) On the characterization of a joint probability distribution by conditional distributions. *Journal of Econometrics 10 115-118.*

Gupta A.K y. Varga T. (1990) Characterization of joint density by conditional densities. *Comunications in Statistics: Theory and Methods. 19(12), 4643-4659.*

Hamedani G.G. (1991) On a recent characterization of the bivariate normal distribution. *Metrika. 38, 255-258.*

Hamedani G.G. (1992) Bivariate and multivariate normal characterization. A brief survey. *Comunications in Statistics: Theory and Methods.21(9), 2665-2688.*

Hürlimann, W. (1993) Bivariate distributions with diatomic conditionals and stop-loss transforms of random sums. *Statistics & Probability letters 17, 329-335*.

Isaacson D.L. y Madsen R.W. (1976) *Markov Chain: Theory and aplications.* John Wiley and Sons. New York.

Ip, E.H. Wang, Y.J. (2009) Canonical representation of specified multivariate discrete distributions. *Journal of M ultivariate Analysis*, 100, 1282-1290.

Kuo, K-L Wang, Y.J. (2011) A simple algorithm for cheking compatibility among discrete conditional distributions. *Computational Statistics and Data Analysis*.55, 2457-2462.

Lay, D.C. (2007) *Álgebra lineal y sus Aplicaciones* Tercera Edición. Pearson Educación

Nerlove, M. and Press, S. J. (1986). Multivariate log-linear probability models in econometrics. In *Advances in Statistical Analysis and Statistical Computing* (Edited by R. S. Mariano), 117-171. JAI Press, Greenwich, CT.

Ng, K. W. (1997). Inversion of Bayes Formula: Explicit Formulae for Unconditional pdf, Advances en *The Theory and Practice of Statistics* (N. L. Johnson and N. Balakrishnan, eds.) John Wiley and Sons.

Noble, B, Daniel J.W. (1989) *Álgebra Lineal Aplicada*. Tercera Edición. Prentice-Hall Interamerica S.A.

Mosteller, F. (1968) Association and estimation in contingency tables. *Journal of the American Statistical Association* 63, 1-28.

Pérez-Villalta, R (1997) *Distribuciones Bivariantes Especificadas con Condicionadas o Marginales* Trabajo de Investigación. Departamento de Economía Aplicada I Universidad de Sevilla.

Pérez-Villalta, R. (2000). Variables finitas condicionalmente especificadas. Questioo 24, 425-448

Rodrigues Correia, E.M. (2007) O papel de Distribuições Condicionalmente Especificadas em Estatística Bayesiana. Tesis Doctoral Universidad de Lisboa.

Schott J.R. (1997) *Matrix Analysis for Statistics* John Wiley and sons New York.

Song Ch.Ch., Li L.A., Chen Ch.H., Jiang T.J. Kuo K.L. (2010) Compatibility of finite discrete conditional distributions. *Statistica Sinica 20 (2010), 423-440*

Tanner M.A. (1991) *Tools for statistical inference, observed data and data argumentation methods.* Lectures Notes in Statistics, Volume 67, Springer Verlang. Berlin.

Tian, G-L Tan, M. Ng, K.W. Tang, M-L. (2009) A unified method for Cheking Compatibility and uniqueness for finite discrete conditional distributions. *Comunications in Statistics: Theory and Methods. 38, 115-129.*

Wang, Y.J. (2012) Comparisons of Three Approaches for Discrete Conditional Models *Communications in Statistics - Simulation and Computation* 41 (1) 32-43

Wang, Y.J. Kuo, K-L. (2010) Compatibility of discrete conditional distributions with structural zeros. *Journal of Multivariate Analysis* 101, 191-199.