COMPARACIÓN DE DIVERSOS ESTIMADORES DEL VECTOR DE PRIORIDADES EN AHP*

Alfredo Altuzarra Casas

Grupo Decisión Multicriterio Zaragoza
Departamento de Métodos Estadísticos
Universidad de Zaragoza
e-mail: altuzarr@unizar.es

José María Moreno Jiménez

Grupo Decisión Multicriterio Zaragoza
Departamento de Métodos Estadísticos
Universidad de Zaragoza
e-mail: moreno@unizar.es

Manuel Salvador Figueras

Grupo Decisión Multicriterio Zaragoza
Departamento de Métodos Estadísticos
Universidad de Zaragoza
e-mail: salvador@unizar.es

Resumen

La obtención del vector de prioridades en el Proceso Analítico Jerárquico (AHP) se ha realizado tradicionalmente a partir del Método del Autovector principal por la derecha de Saaty (Saaty, 1977) o mediante la Media Geométrica por Filas (Crawford y Williams, 1985). En este trabajo, utilizando técnicas bayesianas en un modelo multiplicativo con errores lognormales, se proponen diversos estimadores para el vector de prioridades. Estos estimadores son comparados en términos del Error Cuadrático Medio, lo que lleva a decantarse por la mediana de la distribución a posteriori del vector de prioridades como procedimiento de priorización.

Palabras clave: Proceso Analítico Jerárquico (AHP), Análisis Bayesiano, Estimación.

Área temática: Métodos cuantitativos.

* Este trabajo está parcialmente sufragado por el proyecto multidisciplinar del Gobierno de Aragón "*Gobierno electrónico. Toma de decisiones complejas basadas en Internet: e-democracia y e-cognocracia*" (ref. PM2004-052).

1. Introducción.

El Proceso Analítico Jerárquico (AHP) es una técnica de decisión multicriterio discreta que permite obtener, en una escala de razón, las prioridades asociadas al conjunto de alternativas comparadas. Desarrollada por Saaty (1977, 1980) aborda la resolución de problemas multicriterio, multientorno y multiactores, incorporando en el modelo los aspectos subjetivos y la incertidumbre inherente a todo Proceso de Toma de Decisiones (PTD).

Esta técnica consta de cuatro etapas: modelización, valoración, priorización y síntesis. En la tercera etapa de AHP, se calculan, a partir de las matrices de comparaciones pareadas emitidas por el decisor, los vectores de prioridades locales que determinan la importancia relativa de los elementos de un nivel respecto a un nodo del nivel superior. Para calcular el vector de prioridades se han propuesto en la literatura distintos métodos. El método inicialmente propuesto por Saaty (1980) es el del Autovector Principal por la derecha, *Eigenvector Method* (EVM), que fue el más utilizado en un principio. En los últimos años ha tenido un gran auge la utilización del Método de la Media Geométrica por Filas, *Row Geometric Mea*n (RGM), (Saaty, 1980; Crawford y Williams, 1985; Aguarón y Moreno-Jiménez, 2000).

En lo que sigue se presenta un nuevo procedimiento de priorización, basado en técnicas estadísticas bayesianas. Para ello se supone un modelo multiplicativo con errores lognormales en el que las estimaciones del vector de prioridades se obtienen como las distribuciones a posteriori de los parámetros correspondientes. En los casos que sea posible, estos valores serán calculados de forma exacta, y cuando no lo sea, de forma aproximada mediante métodos MCMC.

El trabajo se ha estructurado de la siguiente forma: en la siguiente sección se presentan los métodos de priorización habituales en AHP; en la Sección 3 se muestra el modelo bayesiano utilizado, mientras que en la Sección 4 se recogen sus correspondientes estimaciones. La Sección 5 se dedica a la comparación entre las estimaciones obtenidas mediante el Método RGM y las obtenidas mediante la utilización de técnicas bayesianas. Finalmente en la Sección 6 se comentan las conclusiones y se indican posibles extensiones.

2. Métodos de priorización.

Como ya se ha comentado anteriormente, AHP es una técnica multicriterio, propuesta por Saaty a mediados de los 70, que permite abordar la resolución de problemas de decisión que se plantean en situaciones complejas incluyendo, en la misma, aquellos aspectos cualitativos y cuantitativos que recojan la subjetividad existente en la mayoría de los procesos de toma de decisiones.

Siguiendo una modelización jerárquica del problema, en la que se incluyen todos los aspectos que se consideren relevantes para la comprensión de dicho problema (criterios, alternativas, escenarios, decisores, etc), se incorporan las preferencias del decisor mediante una serie de comparaciones pareadas que dan lugar a las denominadas matrices de juicio (recíprocas y positivas): A partir de estas matrices se obtienen, mediante cualquiera de los procedimientos de priorización existentes, las prioridades relativas de los elementos comparados evaluadas en una escala de razón. En lo que sigue se revisan, brevemente los más utilizados (EGV y RGM).

2.1. Método EGV.

Dada una matriz recíproca de comparaciones pareadas, $\mathbf{A} = (a_{ij})$ con $a_{ij} > 0$, el vector de prioridades obtenido según el método del *autovector principal por la derecha* (EGV), viene dado por $\mathbf{w}$, que verifica la siguiente expresión:

$$\mathbf{Aw} = \lambda_{max}\mathbf{w}, \qquad \sum_{i=1}^{n} w_i = 1 \text{ (condición de normalización)} \qquad (2.1)$$

donde λ_{max} es el valor propio principal de la matriz **A.**

Una justificación intuitiva del uso de este método se basa en suponer conocidas las prioridades de las alternativas, $\mathbf{w} = (w_1,\ldots,w_n)$, y a partir de ellas construir la matriz de comparaciones pareadas, $\mathbf{W} = (w_{ij}) = \frac{w_i}{w_j}$, verificando $\mathbf{Ww} = n\mathbf{w}$.

Perrón (1907) demostró que si $\mathbf{A} = (a_{ij})$, $a_{ij} > 0$, i, j = 1,…, n, entonces **A** tiene un valor propio positivo λ_{max} (valor propio principal de **A**), cumpliendo además que:

$\lambda_{max} > |\lambda_k|\ \forall\ k$, con λ_k valor propio de **A**. Más aún, el autovector principal por la derecha de la matriz **A**, solución de $\mathbf{Aw} = \lambda_{max}\mathbf{w}$, cumple que $w_i > 0$, $\forall$ i = 1, ..., n.

2.2. Método RGM.

Dada una matriz recíproca de comparaciones pareadas, $\mathbf{A} = (a_{ij})$ con $a_{ij} > 0$, el vector de prioridades obtenido según el método de la *media geométrica por filas* (RGM) viene dado por:

$$w_i = \left(\prod_{j=1,..,n} a_{ij}\right)^{1/n} \quad i = 1, ..., n \qquad \sum_{i=1}^{n} w_i = 1 \text{ (condición de normalización)} \tag{2.2}$$

Saaty considera este método como una aproximación al EGV, al que denomina método exacto, proporcionando buenos resultados para situaciones de poca inconsistencia. Por el contrario, es defendido por otros autores, como Crawford y Williams (1985) y Barzilai (1997), que lo prefieren por sus propiedades estadísticas y psicológicas. En particular, De Jong (1984) y Crawford y Williams (1985) demuestran que dicho estimador coincide con el de mínimos cuadrados logarítmicos cuando las comparaciones pareadas vienen descritas por el siguiente modelo multiplicativo, donde e_{ij} es el término de error aleatorio:

$$a_{ij} = \frac{w_i}{w_j} e_{ij} \tag{2.3}$$

Además, el RGM es el estimador máximo verosímil (MLE) de w si los errores e_{ij} se suponen independientes e idénticamente distribuidos, con distribución logarítmico-normal. El RGM proporciona el mismo vector de prioridades que el EGV si la matriz de comparaciones pareadas es consistente y también cuando el tamaño de la matriz de comparaciones pareadas es $n \leq 3$.

3. Modelo.

Sea un decisor **D** que debe ordenar/seleccionar entre un conjunto compuesto por n alternativas $\mathbf{A} = \{A_1, A_2, \ldots, A_n\}$ tomando como base un criterio de comparación C.

Para ello, y siguiendo la metodología AHP (Saaty, 1977, 1980) emite un conjunto de juicios que dan como resultado una matriz $\mathbf{A} = (a_{ij})$ de comparaciones pareadas en la que a_{ij} representa la comparación relativa entre la alternativa i-ésima y la j-ésima, con $i, j \in \{1,\dots,n\}$ $i < j$. Dicha matriz es recíproca verificando las restricciones $a_{ii} = 1 \ \forall\ i$; $a_{ij} \cdot a_{ji} = 1 \ \forall\ i < j$.

El objetivo del problema consiste en determinar, a partir de dicha información, un vector de prioridades $\mathbf{w} = (w_1,..., w_n)$ con $w_i \geq 0$ $i = 1, \dots, n$, donde w_i represente la prioridad relativa de la alternativa A_i respecto al criterio C.

La aproximación seguida en este trabajo se encuadra dentro del denominado AHP estocástico y toma como punto de partida el modelo multiplicativo dado por la expresión (2.3), donde e_{ij} es el error cometido al comparar las alternativas A_i y A_j.

Si en este modelo se toman logaritmos y se denota:

$$Y_{ij} = \log a_{ij} \qquad \mu_i = \log w_i \quad \varepsilon_{ij} = \log e_{ij} \ \ i, j \in \{1,\dots,n\}\ i < j$$

la expresión (2.3) queda de la forma:

$$Y_{ij} = \mu_i - \mu_j + \varepsilon_{ij} \tag{3.1}$$

En lo que sigue se supondrá que los errores ε_{ij} siguen distribuciones $N(0, \sigma^2)$ independientes (equivalentemente, $e_{ij} \sim LN(0,\sigma^2)$) y, por lo tanto, los juicios a_{ij} siguen una distribución:

$$a_{ij} \mid \mathbf{w} \sim LN\left(\frac{w_i}{w_j}, \sigma^2\right) \Leftrightarrow Y_{ij} \mid \boldsymbol{\mu} \sim N\left(\mu_i - \mu_j, \sigma^2\right)$$

Este modelo ha sido muy utilizado en la literatura (De Jong, 1984; Crawford y Williams, 1985; Fichtner, 1986; Alho y Kangas, 1997), aunque, en general, desde un punto de vista clásico.

Desde un punto de vista bayesiano su tratamiento ha sido menor y generalmente centrado en el caso de la decisión en grupo (Alho y otros, 1996; Alho y Kangas, 1997; Leskinen y Kangas, 1998 o Basak, 1998).

Con el fin de evitar problemas de identificabilidad, simplificar la notación y los cálculos se toma $\mu_n = 0$ (equivalentemente $w_n = 1$), esto es, se fija la última alternativa (A_n) como la de referencia. Se podría seleccionar cualquiera sin imponer ninguna restricción adicional pues suponer $w_n = 1$ no es más que una de las posibles normalizaciones en AHP. También por simplicidad se supone que σ^2 es conocida.

3.1. Distribución a priori.

La distribución a priori de las prioridades μ_i i = 1, …, n-1 viene dada por:

$$\mu_i \sim N(0, \lambda\sigma^2) \ \forall \ i = 1,..., n\text{-}1 \text{ independientes con } \lambda > 0$$

Se supone por lo tanto, que, a priori, nuestra creencia acerca de las alternativas es que todas tienen la misma prioridad $w_1 = \ldots = w_n = 1$. Además, cuanto mayor es λ, más difusa es la distribución a priori y, por tanto, mayor es la ignorancia que a priori se tiene acerca del valor de las prioridades μ_i. En el límite, si $\lambda \to \infty$ la distribución a priori para cada μ_i se hace impropia y es la distribución de referencia no informativa para $\theta = (\mu_1,\ldots, \mu_{n-1})'$.

Este tipo de distribuciones a priori permite realizar un análisis conjugado del problema y es el que se suele adoptar habitualmente en la literatura (Lipscomb y otros (1998) o Alho y Kangas (1997), aunque estos últimos autores se centran en el problema de decisión en grupo).

En resumen, el modelo considerado viene dado por las siguientes condiciones:

MODELO

$$\boxed{\begin{array}{l} Y_{ij} = \mu_i - \mu_j + \varepsilon_{ij} \\ \mu_n = 0 \quad \mu_i \sim N(0, \lambda\sigma^2) \ \text{ independientes } \ \forall\, i = 1,..., n-1 \\ \varepsilon_{ij} \sim N(0, \sigma^2) \ \text{ independientes } \ \forall\, i, j = 1,..., n \quad\quad i \neq j \end{array}} \tag{3.2}$$

3.2. Forma matricial.

Sea $\mathbf{Y}_N = \left(Y_{i_1 j_1},...., Y_{i_N j_N}\right)'$ con $Y_{i_k j_k} = \log(a_{i_k j_k})$ el vector de juicios emitidos donde se supondrá, sin pérdida de generalidad, que $1 \leq i_k < j_k \leq n \ \forall \ k = 1, \ldots, N$, siendo N el número de comparaciones realizadas por el decisor **D**, cumpliendo $1 \leq N \leq \frac{n(n-1)}{2}$.

Por lo tanto, si $N = \frac{n(n-1)}{2}$ la información proporcionada por el decisor **D** es completa. En caso contrario no se han realizado todas las comparaciones posibles entre las alternativas y la información de la que se dispone es incompleta.

Sean $\theta = (\mu_1, ..., \mu_{n-1})'$ el vector del logaritmo de las prioridades y $\boldsymbol{\varepsilon}_N = (\varepsilon_{i_1,j_1}, ..., \varepsilon_{i_N,j_N})'$ el vector de los errores.

El Modelo (3.2) puede expresarse en forma matricial como:

$$\mathbf{Y}_N = \mathbf{X}_N\theta + \varepsilon_N \text{ con } \varepsilon_N \sim N_N\left(\mathbf{0}_N, \sigma^2\mathbf{I}_N\right) \tag{3.3}$$

$$\theta \sim N_{n-1}\left(\mathbf{0}_{n-1}, \lambda\sigma^2\mathbf{I}_{n-1}\right) \tag{3.4}$$

donde $\mathbf{X}_N = (x_{kh})$ es una matriz de dimensión (N x n-1) de forma que:

$$x_{k,h} = \begin{cases} 1 & h = i_k \neq n \\ -1 & h = j_k \neq n \\ 0 & \text{resto} \end{cases} \tag{3.5}$$

Cada comparación realizada se representa mediante una fila de la matriz $\mathbf{X}_N$. Dicha fila contiene, una sola vez, los valores 1 y –1 en las columnas correspondientes a las alternativas comparadas, excepto que una de ellas sea la de referencia. En ese caso, si una de las alternativas comparadas es la de referencia, el juicio emitido será de la forma Y_{in} con $i < n$ y su fila tan sólo contendrá un 1 en la columna i-ésima.

3.3. Distribución a posteriori.

Sabiendo que de (3.3) se sigue que $\mathbf{Y}_N \mid \theta \sim N_N\left(\mathbf{X}_N\theta, \sigma^2\mathbf{I}_N\right)$ y que la distribución a priori (3.4) es conjugada se obtiene, aplicando el Teorema de Bayes, la distribución a posteriori del vector de parámetros θ:

$$\theta \mid (\mathbf{X}_N, \mathbf{Y}_N) \sim N_{n-1}\left(\left(\mathbf{X}_N'\mathbf{X}_N + \frac{1}{\lambda}\mathbf{I}_{n-1}\right)^{-1}\mathbf{X}_N'\mathbf{Y}_N,\ \sigma^2\left(\mathbf{X}_N'\mathbf{X}_N + \frac{1}{\lambda}\mathbf{I}_{n-1}\right)^{-1}\right) \tag{3.6}$$

En particular si $\lambda \to \infty$ se tiene que:

$$\theta \mid (\mathbf{X}_N, \mathbf{Y}_N) \sim N_{n-1}\left((\mathbf{X}'_N \mathbf{X}_N)^{-1} \mathbf{X}'_N \mathbf{Y}_N,\ \sigma^2 (\mathbf{X}'_N \mathbf{X}_N)^{-1}\right) \qquad (3.7)$$

A partir de estas distribuciones es posible realizar inferencias acerca del vector de parámetros θ o de transformaciones suyas como, el vector de prioridades **w**.

4. Estimación del vector de prioridades.

Si se conoce la matriz completa de comparaciones pareadas, es decir, se han realizado todas las comparaciones pareadas posibles, el número total de juicios emitidos es $N = \frac{n(n-1)}{2}$, y en ese caso se tiene que:

$$\left(\mathbf{X}'_N \mathbf{X}_N\right)^{-1} = \frac{1}{n}\begin{pmatrix} 2 & 1 & \dots & 1 \\ \dots & \dots & \dots & \dots \\ 1 & 1 & \dots & 2 \end{pmatrix} = \frac{1}{n}(\mathbf{I}_{n-1} + \mathbf{J}_{n-1}) \qquad (4.1)$$

donde: $$\mathbf{J}_{n-1} = \begin{pmatrix} 1 & 1 & \dots & 1 \\ \dots & \dots & \dots & \dots \\ 1 & 1 & \dots & 1 \end{pmatrix} \qquad (4.2)$$

Teorema 4.1

Si se dispone de la matriz completa de comparaciones, la distribución a posteriori dada por la expresión (3.7), del vector de prioridades θ, viene caracterizada por:

$\mu_i \mid (\mathbf{X}, \mathbf{Y}) \sim N(E[\mu_i \mid (\mathbf{X}, \mathbf{Y})], \mathrm{Var}[\mu_i \mid (\mathbf{X}, \mathbf{Y})])$, con:

$$E[\mu_i \mid (\mathbf{X}, \mathbf{Y})] = \frac{1}{n}\left(\sum_{j=1}^{n} Y_{ij} + \sum_{j=1}^{n} Y_{jn}\right) \quad \forall\ i = 1, \dots, n-1 \qquad (4.3)$$

$$\mathrm{Var}[\mu_i \mid (\mathbf{X}, \mathbf{Y})] = \frac{2}{n}\sigma^2 \quad \forall\ i = 1, \dots, n-1 \qquad (4.4)$$

donde se ha tomado $Y_{ii} = 0$ e $Y_{ji} = -Y_{ij}\ \forall\ i < j = 1, \dots, n$.

Además: $$\mathrm{Cov}[\mu_i, \mu_j \mid (\mathbf{X}, \mathbf{Y})] = \frac{1}{n}\sigma^2 \quad \forall\ i, j = 1, \dots, n-1;\ \text{con } i \neq j \qquad (4.5)$$

Demostración. Inmediata, basta aplicar las expresiones (3.7), (4.1) y (4.2). □

A partir de las expresiones anteriores y utilizando el cambio **w** = exp(θ) se deduce:

Teorema 4.2 (Estimaciones bayesianas del vector de prioridades)
La mediana, media y moda de la distribución $w_i \mid (\mathbf{X}, \mathbf{Y})$ vienen dadas por:

$$\text{- [Mediana]: } \mathrm{Med}[w_i \mid (\mathbf{X}, \mathbf{Y})] = \left(\prod_{j=1}^{n} a_{ij} \Big/ \prod_{j=1}^{n} a_{nj} \right)^{1/n} \tag{4.6}$$

$$\text{- [Media]: } \mathrm{E}[w_i \mid (\mathbf{X}, \mathbf{Y})] = \mathrm{Med}[w_i \mid (\mathbf{X}, \mathbf{Y})] \exp\left[\frac{1}{n} \sigma^2 \right] \tag{4.7}$$

$$\text{- [Moda]: } \mathrm{Moda}(w_i \mid (\mathbf{X}, \mathbf{Y})] = \mathrm{Med}[w_i \mid (\mathbf{X}, \mathbf{Y})] \exp\left[-\frac{2}{n} \sigma^2 \right] \tag{4.8}$$

Demostración. A partir del Teorema 4.1 se sigue que:

$$\mu_i \mid (\mathbf{X}, \mathbf{Y}) \sim \mathrm{N}\left(\mathrm{E}\left[\mu_i \mid (\mathbf{X}, \mathbf{Y})\right], \frac{2}{n}\sigma^2 \right) \text{ donde:}$$

$$\mathrm{E}[\mu_i \mid (\mathbf{X}, \mathbf{Y})] = \overline{Y}_{i.} - \overline{Y}_{n.} \qquad \text{con } \overline{Y}_{i.} = \frac{1}{n}\sum_{j=1}^{n} Y_{ij} \quad i = 1, \ldots, n$$

Dado que $\mu_i = \log(w_i)$ se tendrá que:

$$w_i \mid (\mathbf{X}, \mathbf{Y}) \sim \mathrm{LN}\left(\mathrm{E}\left[\mu_i \mid (\mathbf{X}, \mathbf{Y})\right], \frac{2}{n}\sigma^2 \right) \quad i = 1, \ldots, n$$

Por lo tanto, basta calcular la media, mediana y moda de una distribución lognormal, (Johnson y Kotz, 1970), para obtener las expresiones (4.6), (4.7) y (4.8). □

Teorema 4.3
Las estimaciones, normalizadas en modo distributivo (su suma vale 1), del vector de prioridades obtenidas a partir de la mediana a posteriori (4.6) coinciden con las obtenidas aplicando el método RGM y normalizadas.

Demostración. Sean **w** y **v** los vectores obtenidos a partir de la mediana a posteriori y el RGM, respectivamente; es decir:

$$w_i = \left(\prod_{j=1}^{n} \left(a_{ij} \big/ a_{nj} \right) \right)^{1/n} \qquad \forall\ i = i, ..., n\text{-}1 \qquad w_n = 1$$

$$v_i = \left(\prod_{j=1}^{n} a_{ij} \right)^{1/n} \qquad \forall\ i = i, ..., n$$

Si denominamos $\mathbf{w^D}$ y $\mathbf{v^D}$ a los mismos vectores anteriores, pero normalizados en modo distributivo (suma igual a la unidad), se obtiene:

a) Prioridades calculadas mediante el RGM:

$$v_i^D = \frac{\left(\prod_{j=1}^{n} a_{ij} \right)^{1/n}}{\sum_{i=1}^{n} \left(\prod_{j=1}^{n} a_{ij} \right)^{1/n}} \qquad \forall\ i = i, ..., n$$

b) Prioridades calculadas mediante priorización bayesiana, excepto la última:

$$w_i^D = \frac{w_i}{\sum_{i=1}^{n-1} w_i + 1} = \frac{\left(\prod_{j=1}^{n} a_{ij} \right)^{1/n}}{\left(\prod_{j=1}^{n} a_{nj} \right)^{1/n} + \sum_{i=1}^{n-1} \left[\left(\prod_{j=1}^{n} a_{ij} \right)^{1/n} \right]} = \frac{\left(\prod_{j=1}^{n} a_{ij} \right)^{1/n}}{\sum_{i=1}^{n} \left(\prod_{j=1}^{n} a_{ij} \right)^{1/n}},\ \ i=1,\ldots,n\text{-}1$$

c) Para la última prioridad w_n^D:

$$w_n^D = \frac{1}{1 + \sum_{i=1}^{n-1} \left[\prod_{j=1}^{n} \left(a_{ij} \big/ a_{nj} \right) \right]^{1/n}} = \frac{\left(\prod_{j=1}^{n} a_{nj} \right)^{1/n}}{\sum_{i=1}^{n} \left(\prod_{j=1}^{n} a_{ij} \right)^{1/n}}$$

Por lo tanto $w_i^D = v_i^D$ $\forall$ i = i, ..., n, con lo que la tesis queda demostrada. □-

Teorema 4.4

En las condiciones del Teorema 4.1, si se realiza el cambio:

$$\psi_i = \mu_i - \bar{\theta} \quad \forall\ i = 1, ..., n; \text{ con } \bar{\theta} = \frac{1}{n}\sum_{i=1}^{n}\mu_i \tag{4.9}$$

$\psi \mid (\mathbf{X}, \mathbf{Y})$, sigue una distribución $\psi \mid (\mathbf{X}, \mathbf{Y}) \sim N_n(\mathbf{m}_\psi, \Sigma_\psi)$, con $\psi = (\psi_1, ..., \psi_n)$, con:

$$\mathbf{m}'_\psi = \left(\frac{1}{n}\sum_{j=1}^{n} Y_{1j}, ..., \frac{1}{n}\sum_{j=1}^{n} Y_{nj}\right) \qquad \Sigma_\psi = \frac{\sigma^2}{n^2}\begin{pmatrix} n-1 & -1 & ... & -1 \\ -1 & n-1 & ... & -1 \\ ... & ... & ... & ... \\ -1 & -1 & ... & n-1 \end{pmatrix}$$

Además, si se realiza el cambio de variable $\mathbf{v} = \exp(\psi)$, la estimación bayesiana de la prioridad i-ésima, dada por la mediana de la distribución a posteriori de $v_i \mid \mathbf{Y}$ coincide con la media geométrica por filas:

$$\text{Med}[v_i \mid (\mathbf{X}, \mathbf{Y})] = \left(\prod_{j=1}^{n} a_{ij}\right)^{1/n} \tag{4.10}$$

Demostración. Dada la distribución a posteriori (3.7), la expresión (4.9) implica un cambio de variable que no es más que la normalización geométrica utilizada por Crawford y Williams (1985), es decir, que el producto de las prioridades valga 1.

Como este cambio de variable (para cada componente, $\psi_i = \mu_i - \bar{\theta}$) es una transformación lineal, se sigue que $\psi \mid (\mathbf{X}, \mathbf{Y})$ tendrá una distribución Normal. Matricialmente, se puede expresar de la forma $\psi = \mathbf{T}\,\theta$. Por lo tanto:

$\mathbf{m}_\psi = \mathbf{T}\ E[\theta \mid (\mathbf{X}, \mathbf{Y})] = \mathbf{T}\ \mathbf{m}_\theta \qquad \Sigma_\psi = \mathbf{T}\ \text{Var}[\theta \mid (\mathbf{X}, \mathbf{Y})]\ \mathbf{T}'$

Realizando unos cálculos matriciales, se obtiene, para el vector de medias:

$$m_{\psi_i} = E(\psi_i \mid \mathbf{Y}) = \frac{1}{n}\left(-m_{\theta_1} - ... + (n-1)m_{\theta_i} - ... - m_{\theta_{n-1}}\right) = \frac{1}{n}\sum_j Y_{ij} \qquad \forall\ i = 1, ..., n$$

mientras que para la matriz de varianzas-covarianzas:

$$\Sigma_{\psi} = \frac{\sigma^2}{n^2}\begin{pmatrix} n-1 & -1 & \ldots & -1 \\ -1 & n-1 & \ldots & -1 \\ \ldots & \ldots & \ldots & \ldots \\ -1 & -1 & \ldots & n-1 \end{pmatrix}$$

Por lo tanto, cada prioridad $\psi_i \mid (\mathbf{X}, \mathbf{Y}) \sim N\left(m_{\psi_i}, \frac{\sigma^2}{n^2}(n-1)\right)$; y ya que $\psi_i = \log v_i$, se obtiene, de forma análoga al Teorema 4.2:

$$v_i \mid (\mathbf{X}, \mathbf{Y}) \sim LN\left(m_{\psi_i}, \frac{\sigma^2}{n^2}(n-1)\right)$$

y, a partir de aquí, se deduce la expresión (4.9) para la mediana. □

Salvo un factor de escala que refleja la asimetría de la distribución logarítmico-normal, los 3 estimadores obtenidos a partir del Teorema 4.2 se expresan en términos de la media geométrica de la fila i-ésima normalizada por la media geométrica de la fila n-ésima. Éste último término surge de la condición $w_n = 1$ adoptada en el modelo.

5. Comparación de las estimaciones del vector de prioridades.

A continuación se presenta un estudio comparativo de diferentes estimadores del vector de prioridades, tanto para el modelo original (2.3), como para el modelo en logaritmos (3.1). Por un lado, se estudian las tres medidas de posición calculadas en el Teorema 4.2 y, por otro, sus transformados, tal como se ha visto en el Teorema 4.4, donde también se hacía referencia a la media geométrica por filas. Se comprueba cuáles de ellos son insesgados y se comparan sus errores cuadráticos medios.

Teorema 5.1

i) El estimador de máxima verosimilitud y mínimos cuadrados propuesto por Crawford y Williams (1985), $\hat{\mu}_i = \frac{1}{n}\sum_j \log a_{ij}$, es un estimador insesgado del parámetro μ_i del modelo logarítmico (3.1).

ii) El transformado exponencial del estimador anterior, $\hat{w}_i = \exp(\hat{\mu}_i) = \prod_j a_{ij}^{1/n}$ es un estimador sesgado del parámetro w_i del modelo (2.3). Además este estimador coincide con el estimador bayesiano dado por la mediana de la distribución a posteriori $v_i = \exp(\Psi_i) \mid (\mathbf{X},\mathbf{Y})$, como resulta de la expresión (4.10) del Teorema 4.4.

iii) La media y la moda a posteriori de $v_i \mid (\mathbf{X}, \mathbf{Y})$ son, también, estimadores sesgados del parámetro w_i.

$$E[v_i \mid (\mathbf{X}, \mathbf{Y})] = \mathrm{Med}[v_i \mid (\mathbf{X}, \mathbf{Y})] \exp\left[\frac{(n-1)}{2n^2}\sigma^2\right] \tag{5.1}$$

$$\mathrm{Moda}[v_i \mid (\mathbf{X}, \mathbf{Y})] = \mathrm{Med}[v_i \mid (\mathbf{X}, \mathbf{Y})] \exp\left[-\frac{(n-1)}{n^2}\sigma^2\right] \tag{5.2}$$

Demostración. i) Basta con recordar que Crawford y Williamns (1985) utilizan el Modelo (3.1), con los errores ε_{ij} independientes con distribuciones $N(0, \sigma^2)$, imponiendo la condición de normalización $\prod_j w_i = 1 \Leftrightarrow \sum_j \mu_j = 0$.

Por lo tanto: $\quad E(\hat{\mu}_i) = E\left(\frac{1}{n}\sum_j \log a_{ij}\right) = \frac{1}{n}\sum_j E(Y_{ij}) = \mu_i - \frac{1}{n}\sum_j \mu_j = \mu_i$

con lo que, tras aplicar la condición de normalización, este estimador es insesgado.

ii) Para la segunda parte de la propiedad se va a recurrir a la función generatriz de momentos de una distribución Normal.

Además, $\hat{w}_i = \prod_j a_{ij}^{1/n}$, se puede expresar como $\hat{w}_i = \exp\left[\frac{1}{n}\sum_j Y_{ij}\right]$, por lo que:

$$E(\hat{w}_i) = E\left(\exp\left[\frac{1}{n}\sum_j Y_{ij}\right]\right) = E\left(\exp[tX]\right) \tag{5.3}$$

con $t = \frac{1}{n}$; $X = \sum_j Y_{ij}$. De esta forma, resulta:

$$E(\hat{w}_i) = \exp\left[\frac{1}{n}(n\mu_i)\right]\exp\left[\frac{1}{2}\frac{1}{n^2}(n-1)\sigma^2\right] = w_i\exp\left[\frac{(n-1)\sigma^2}{2n^2}\right] \qquad (5.4)$$

El transformado exponencial dado por $\hat{w}_i = \prod_j a_{ij}^{1/n}$ es un estimador sesgado, que coincide con la expresión (4.10), mediana de la distribución a posteriori $v_i \mid (\mathbf{X}, \mathbf{Y})$.

iii) Sean: $\hat{w}_i^E = E[v_i \mid (\mathbf{X}, \mathbf{Y})]$ y $\hat{w}_i^{Mo}$ = Moda $[v_i \mid (\mathbf{X}, \mathbf{Y})]$ los estimadores dados por (5.1) y (5.2). Ambos se pueden poner en función de $\hat{w}_i$, y sus esperanzas resultan:

$$E(\hat{w}_i^E) = E\left(\hat{w}_i\exp\left[\frac{(n-1)}{2n^2}\sigma^2\right]\right) = w_i\exp\left[\frac{(n-1)}{n^2}\sigma^2\right] \qquad (5.5)$$

$$E(\hat{w}_i^{Mo}) = E\left(\hat{w}_i\exp\left[-\frac{(n-1)}{n^2}\sigma^2\right]\right) = w_i\exp\left[-\frac{(n-1)}{2n^2}\sigma^2\right] \qquad (5.6)$$

Por lo tanto, estos dos estimadores de w_i también son sesgados. □

Teorema 5.2

Los Errores Cuadráticos Medios (ECM) de los estimadores de w_i utilizados en el Teorema 5.1 (mediana, media y moda de $v_i \mid (\mathbf{X}, \mathbf{Y})$), son los siguientes:

$$ECM(\hat{w}_i) = w_i^2(g^2 - 2g^{1/2} + 1) \qquad (5.7)$$

$$ECM(\hat{w}_i^E) = w_i^2(g^3 - 2g + 1) \qquad (5.8)$$

$$ECM(\hat{w}_i^{Mo}) = w_i^2\left(\frac{1}{g^2} - \frac{3}{g} + 2\right),\ \text{donde } g = \exp\left[\frac{(n-1)}{n^2}\sigma^2\right] \qquad (5.9)$$

Demostración. Recordando que $ECM(\hat{w}_i) = Var(\hat{w}_i) + SESGO^2(\hat{w}_i)$ y utilizando, igual que en la demostración del Teorema anterior, la función generatriz de momentos de una distribución Normal, se deduce que:

$$(\hat{w}_i)^2 = \exp\left[\frac{2}{n}\sum_j Y_{ij}\right] = \exp[tX] \ \text{ con } t = \frac{2}{n};\ \ X = \sum_j Y_{ij} \sim N\,(n\,\mu_i, (n-1)\sigma^2)$$

$$E\left[(\hat{w}_i)^2\right] = \exp\left[\frac{2}{n} n\mu_i + \frac{1}{2}\frac{4}{n^2}(n-1)\sigma^2\right] = w_i^2 \; g^2 \tag{5.10}$$

$$E\left[(\hat{w}_i^E)^2\right] = E\left[(\hat{w}_i)^2 \exp\left[\frac{(n-1)}{n^2}\sigma^2\right]\right] = w_i^2 \; g^3 \tag{5.11}$$

$$E\left[(\hat{w}_i^{Mo})^2\right] = E\left[(\hat{w}_i)^2 \exp\left[\frac{-2(n-1)}{n^2}\sigma^2\right]\right] = w_i^2 \tag{5.12}$$

Y uniendo las expresiones (5.4) a (5.6) con (5.10) a (5.12), resulta:

$$\mathrm{Var}(\hat{w}_i) = w_i^2\left(g^2 - g\right) \tag{5.13}$$

$$\mathrm{Var}(\hat{w}_i^E) = w_i^2\left(g^3 - g^2\right) \tag{5.14}$$

$$\mathrm{Var}(\hat{w}_i^{Mo}) = w_i^2\left(1 - \tfrac{1}{g}\right) \tag{5.15}$$

A continuación se calculan los sesgos de estos estimadores:

$$\mathrm{SESGO}^2(\hat{w}_i) = \left(w_i - E[\hat{w}_i]\right)^2 = w_i^2\left(g^{1/2} - 1\right)^2 \tag{5.16}$$

$$\mathrm{SESGO}^2(\hat{w}_i^E) = w_i^2\left(g-1\right)^2 \tag{5.17}$$

$$\mathrm{SESGO}^2(\hat{w}_i^{Mo}) = w_i^2\left(\tfrac{1}{g} - 1\right)^2 \tag{5.18}$$

Por último, uniendo las expresiones (5.13) a (5.18) se calculan los ECM:

$$\mathrm{ECM}(\hat{w}_i) = w_i^2\left(g^2 - g\right) + w_i^2\left(g^{1/2} - 1\right)^2 = w_i^2(g^2 - 2g^{1/2} + 1)$$

$$\mathrm{ECM}(\hat{w}_i^E) = w_i^2\left(g^3 - g^2\right) + w_i^2\left(g-1\right)^2 = w_i^2(g^3 - 2g + 1)$$

$$\mathrm{ECM}(\hat{w}_i^{Mo}) = w_i^2\left(1 - \frac{1}{g}\right) + w_i^2\left(1 - \frac{1}{g}\right)^2 = w_i^2\left(\frac{1}{g^2} - \frac{3}{g} + 2\right)$$

Estos resultados coinciden con las expresiones (5.7) a (5.9). □

Teorema 5.3

Los ECM de los estimadores del parámetro w_i, (mediana, media y moda de la distribución a posteriori $v_i \mid (\mathbf{X}, \mathbf{Y})$), dados en el Teorema 5.2, cumplen:

$$\mathrm{ECM}\,(\hat{w}_i^{E}) \geq \mathrm{ECM}\,(\hat{w}_i) \geq \mathrm{ECM}\,(\hat{w}_i^{Mo}) \tag{5.19}$$

Demostración. Para la primera desigualdad se calcula:

$$\mathrm{Var}(\hat{w}_i^{E}) - \mathrm{Var}(\hat{w}_i) = w_i^2\left(g^3 - g^2\right) - w_i^2\left(g^2 - g\right) = w_i^2\, g(g-1)^2$$

$$\mathrm{SESGO}^2\,(\hat{w}_i) - \mathrm{SESGO}^2\,(\hat{w}_i^{E}) = w_i^2\left(g^{1/2} - 1\right)^2 - w_i^2\,(g-1)^2$$

Sumando ambas expresiones, se obtiene:

$$\mathrm{ECM}(\hat{w}_i^{E}) - \mathrm{ECM}(\hat{w}_i) = w_i^2\left[(g-1)^3 + \left(g^{1/2} - 1\right)^2\right] \geq 0$$

Lo cual es cierto, puesto que $g = \exp\left[\dfrac{(n-1)}{n^2}\sigma^2\right] \geq 1$.

Para la segunda desigualdad, se descompone el ECM en sus dos términos, comprobando que:

$$\mathrm{Var}(\hat{w}_i) \geq \mathrm{Var}(\hat{w}_i^{Mo}) \qquad \mathrm{SESGO}^2\,(\hat{w}_i) \geq \mathrm{SESGO}^2\,(\hat{w}_i^{Mo})$$

Esto es equivalente, respectivamente, y utilizando que $g \geq 1$, a que:

$$w_i^2\left(g^3 - g^2\right) \geq w_i^2\,\frac{(g-1)}{g} \Leftrightarrow g^2(g-1) \geq \frac{(g-1)}{g} \Leftrightarrow g^3 \geq 1$$

$$w_i^2(g-1)^2 \geq w_i^2\left(\frac{1}{g} - 1\right)^2 \Leftrightarrow (g-1)^2 \geq \frac{(g-1)^2}{g^2} \Leftrightarrow g^2 \geq 1$$

Uniendo estas dos expresiones se sigue la segunda desigualdad. □

Teorema 5.4

El estimador bayesiano $\hat{\mu}_i^{B} = \dfrac{1}{n}\left(\sum_{j=1}^{n} Y_{ij} + \sum_{j=1}^{n} Y_{jn}\right)$ dado por la expresión (4.3) es un estimador insesgado del parámetro μ_i del modelo (3.2), mientras que la mediana, la media y la moda de la distribución a posteriori (w_i | (**X, Y**)), dados por las expresiones (4.6), (4.7) y (4.8) son estimadores sesgados de w_i.

Demostración. Para el estimador $\hat{\mu}_i^B$:

$$E\left(\hat{\mu}_i^B\right) = E\left(\frac{1}{n}\left(\sum_{j=1}^{n} Y_{jn} + \sum_{j=1}^{n} Y_{ij}\right)\right) = \frac{1}{n}\left(\sum_{j=1}^{n}(\mu_j - \mu_n) + \sum_{j=1}^{n}(\mu_i - \mu_j)\right) = \mu_i - \mu_n = \mu_i$$

Puesto que se ha aplicado la condición de identificabilidad ($\mu_n = 0$) utilizada en el Modelo (3.2), y por lo tanto, el estimador es insesgado.

Para los estimadores de w_i se hace uso de que su distribución a posteriori (w_i | (**X, Y**) es lognormal y además se utiliza de nuevo la función generatriz de momentos. Sean $\hat{w}_i^1$, $\hat{w}_i^2$ y $\hat{w}_i^3$ las estimaciones bayesianas del vector de prioridades dadas por las expresiones (4.6) a (4.8). Estos tres estimadores se pueden expresar también, como:

$$\hat{w}_i^1 = \exp\left[\frac{1}{n}\sum_j (Y_{ij} - Y_{nj})\right] = \exp[tX]$$

$$\hat{w}_i^2 = \hat{w}_i^1 \exp\left[\frac{\sigma^2}{n}\right] \qquad \hat{w}_i^3 = \hat{w}_i^1 \exp\left[\frac{-2\sigma^2}{n}\right]$$

donde $t = \frac{1}{n}$; $X = \sum_j (Y_{ij} - Y_{nj}) \sim N\ (n\ \mu_i,\ 2n\sigma^2)$.

Calculando sus esperanzas, resulta:

$$E(\hat{w}_i^1) = \exp\left[\frac{1}{n} n\mu_i + \frac{1}{2}\frac{1}{n^2} 2n\sigma^2\right] = w_i \exp\left[\frac{\sigma^2}{n}\right] \tag{5.20}$$

$$E(\hat{w}_i^2) = E\left(\hat{w}_i^1 \exp\left[\frac{\sigma^2}{n}\right]\right) = w_i \exp\left[\frac{2\sigma^2}{n}\right] \tag{5.21}$$

$$E(\hat{w}_i^3) = E\left(\hat{w}_i^1 \exp\left[\frac{-2\sigma^2}{n}\right]\right) = w_i \exp\left[\frac{-\sigma^2}{n}\right] \tag{5.22}$$

y, por lo tanto, los tres son estimadores sesgados de w_i. □

Teorema 5.5

Los ECM de los estimadores de w_i, dados por las expresiones (4.6) a (4.8), son:

i) Para la mediana de la distribución a posteriori (w_i | (**X, Y**))

$$\text{ECM}\,(\hat{w}_i^1) = w_i^2(h^4 - 2h + 1) \tag{5.23}$$

ii) Para la media de la distribución a posteriori (w_i | (**X, Y**))

$$\text{ECM}\,(\hat{w}_i^2) = w_i^2(h^6 - 2h^2 + 1) \tag{5.24}$$

iii) Para la moda de la distribución a posteriori (w_i | (**X, Y**))

$$\text{ECM}\,(\hat{w}_i^3) = \frac{2w_i^2}{h}(h-1) \text{ donde } h = \exp\left[\frac{\sigma^2}{n}\right] \tag{5.25}$$

Demostración. Similar a la realizada para el Teorema 5.2. □

Teorema 5.6

Los ECM de los estimadores mediana, media y moda de la distribución a posteriori (w_i | (**X, Y**)), calculados en el Teorema 5.5, y dados por las expresiones (5.23) a (5.25), todos ellos estimadores de w_i, cumplen la siguiente propiedad:

$$\text{ECM}\,(\hat{w}_i^2) \geq \text{ECM}\,(\hat{w}_i^1) \geq \text{ECM}\,(\hat{w}_i^3)$$

Demostración. Se hace de forma similar a lo expuesto para el Teorema 5.3. □

6. Conclusiones.

Suponiendo un modelo multiplicativo para los errores se ha propuesto un nuevo procedimiento de priorización, bayesiano, que se puede aplicar en situaciones con incertidumbre. Este nuevo método permite obtener el vector de prioridades para matrices de juicios completas (como en este trabajo), pudiendo extenderse al caso de aquellas matrices en las que algunos de los juicios han sido emitidos de forma imprecisa o, incluso, no han sido emitidos.

Respecto a los resultados obtenidos tras comparar estos estimadores, cabe indicar que la mediana de la distribución a posteriori (w_i | (**X, Y**)) dada por la expresión (4.6), coincide, salvo la condición de identificabilidad, con la media geométrica por filas, que es el estimador propuesto habitualmente en la literatura referente a AHP. Uno de sus inconvenientes es que es un estimador sesgado, aunque eso les ocurre a todos los estimadores que se han mostrado en los Teoremas 5.1 y 5.4. Por otra parte posee una gran ventaja respecto a la media y la moda a posteriori, y es que en su cálculo no interviene la varianza σ^2, que habitualmente es desconocida; es decir, que para dar una estimación basta únicamente con los valores de los juicios emitidos.

En términos del ECM, el mejor estimador de las prioridades w_i, cuando la varianza es conocida, es la moda a posteriori. Sin embargo σ^2 es habitualmente desconocida y debe ser estimada, lo cual resta interés a la utilización de la moda a posteriori como estimador, pues depende explícitamente de dicho parámetro.

Bibliografía:

1. Aguarón, J.; Moreno-Jiménez, J. M. (2000): Local Stability Intervals In The Analytic Hierarchy Process. *European Journal of Operational Research*, **125**, 113–132.

2. Alho, J.; Kangas, J. (1997): Analyzing Uncertainties in Expert´s Opinions of Forest Plan Performance. *Forest Science*, **43** (**4**), 521-528.

3. Alho, J.; Kangas, J.; Kolehmainen, O. (1996): Uncertainty in Expert Predictions of the Ecological Consequences of Forest Plans. Applied Statistics **45** (**1**), 1-14.

4. Barzilai, J. (1997): Deriving Weights from Pairwise Comparison Matrices. *Journal of the Operational Research Society*, **48**, 1226-1232.

5. Basak, I. (1998): Probabilistic Judgments Specified Partially in the Analytic Hierarchy Process. *European Journal of Operational Research*, **108**, 152-164.

6. Crawford G.; Williams C. (1985): A Note on the Analysis of Subjective Judgment Matrices, *Journal of Mathematical Psychology*, **29**, 387-405.

7. De Jong, P. (1984) : A Statistical Approach to Saaty's Scaling Method for Priorities. *Journal of Mathematical Psychology*, **28**, 467–478.

8. Fichtner, J. (1986): On Deriving Priority Vectors from Matrices of Pairwise Comparations. *Socio-Economic Planning Sciences,* **20**, 399-405.

9. Johnson, N. L.; Kotz, S. (1970): *Distributions in Statistics. Continuous Univariate Distribution-1.* Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, New York.

10. Leskinen, P.; Kangas, J. (1998): Analysing Uncertainties of Interval Judgment Data in Multi-Criteria Evaluation of Forest Plans. *Silva Fennica,* **32** (**4**), 363-372.

11. Lipscomb, J.; Parmigiani, G.; Hasselblad, V. (1998): Combinig Expert Judment by Hierarchical Modeling: an Aplication to Physician Staffing. *Management Science*, **44** (**2**), 149-161.

12. Perrón, O. (1907): Zur Theorie Der Matrices. *Math. Ann.*, **64**, 248-263.

13. Saaty T.L. (1977): A Scaling Method For Priorities In Hierarchical Structures, Journal of Mathematical Psychology, **15** (**3**), 234-281.

14. Saaty T.L. (1980): *Multicriteria Decision Making: The Analytic Hierarchy Process*, Mc Graw-Hill, New York. (2ª Impresión 1990, Rsw Pub. Pittsburg).