ÍNDICE

Estudios de Economía Aplicada
Nº 11, 1999. Págs. 101-120

Análisis de robustez de los modelos bayesianos para Auditoría de Cuentas: La independencia entre Tasa y Cantidad de Error[1]

MARTEL ESCOBAR, Mª C.
HERNÁNDEZ BASTIDA, A.
VÁZQUEZ POLO, F. J.
Universidad de Las Palmas de Gran Canaria
Universidad de Granada

Esta versión incluye todas las correcciones sugeridas por el evaluador, las cuales nos han parecido oportunas y por las que le quedamos muy agradecidos.

RESUMEN

Se tratan los modelos bayesianos biparamétricos para el cálculo de cotas del error total en una auditoría en el contexto de los diferentes modelos estadísticos desarrollados en la literatura contable. En particular, en este artículo se estudia la robustez de la (discutible) hipótesis de independencia entre los parámetros en la que se apoyan dichos modelos. El estudio se aplica al modelo Beta-Uniforme de Godfrey y Neter, obteniendo conclusiones de falta de robustez.

Palabras clave: Análisis Bayesiano, Modelos Biparamétricos, Auditoría de Cuentas, Análisis de Robustez.

ABSTRACT

We study biparametric Bayesian models in auditing among the wide range of Bayesian statistical models. Specifically, we try to measure the robustness of the hypothesis of the essential and polemic independence between the parameters of the models. We present an application of this study to the

---

1. Investigación financiada por el proyecto PB95-1194 de la Dirección General de Investigación Científica y Técnica (D.G.I.C.Y.T.). Correspondencia a: María Martel Escobar. Departamento de Métodos Cuantitativos en Economía y Gestión. Universidad de Las Palmas de G.C. F. de CC. Económicas y Empresariales. Módulo D. 35017 Las Palmas de Gran Canaria. E-mail: maria@empresariales.ulpgc.es

ÍNDICE

Beta-Uniform model (Godfrey and Neter, 1984) and we conclude that there exist reasons to assume a great lack of robustness.

Keywords: Bayesian analysis, Biparametric models, Auditing, Bayesian Robustness.

Clasificación A.M.S.: 62A15, 62C10

Artículo recibido en marzo de 1998. Revisado en agosto de 1998.

1. Introducción

En los modelos estadísticos desarrollados en la literatura contable para utilizar en una investigación de auditoría de cuentas hay una gran variedad de métodos entre los que destacan los métodos biparamétricos bayesianos (Cox y Snell, 1979 y Godfrey y Neter, 1984, entre otros). Esencialmente estos procedimientos parten de la consideración de dos magnitudes observables en la muestra: m que representa el número de fallos observados en la muestra y z el orden de magnitud de los fallos observados.

Los diversos procedimientos han surgido a partir de diferentes suposiciones sobre la forma de las verosimilitudes y sobre la modelización probabilística de las distribuciones a priori para los parámetros del problema. Muy resumidamente estas aproximaciones han sido: m sigue una distribución Binomial de parámetros n (tamaño muestral) y ϕ (tasa de error) o bien su aproximación por una distribución de Poisson de parámetros $n\phi$. Y z sigue una distribución Exponencial de parámetros m (Cox y Snell, 1979) o Exponencial truncada en (0,1) de media m (Godfrey y Neter, 1984).

Para los parámetros del problema las diferentes distribuciones a priori consideradas han sido:

- ϕ sigue una Gamma y m sigue una Gamma inversa (Cox y Snell, 1979).
- ϕ sigue una Gamma truncada y m una Uniforme (Godfrey y Neter, 1984).
- ϕ sigue una Beta y m una Uniforme (Godfrey y Neter, 1984).

El parámetro de interés en el problema de auditoría de cuentas es $\psi = \phi \cdot \mu$. Por tanto desde un punto de vista bayesiano el objetivo es la determinación de la distribución a posteriori de ψ, y a partir de ella, el problema de buscar una cota superior para el error total de una contabilidad (que no es más que el producto entre ψ y valor de libro registrado) se convierte sencillamente en determinar un cuantil de la distribución a posteriori.

En las tres situaciones descritas anteriormente para la modelización de la información a priori, los diversos autores consideran que los parámetros ϕ y μ son independientes.

Otra gran variedad de modelos estadísticos bayesianos desarrollados en la literatura para la auditoría de cuentas son los basados en la distribución multinomial. En todos ellos la modelización es semejante y consiste en suponer establecido un rango de categorías de fallo (cada categoría de fallo es un porcentaje de fallo), a partir de ello la información muestral proporcionada por una muestra DUS o MUS (dollar unit sampling o monetary unit sampling) consiste en contar el número de observaciones que pertenece a cada una de las categorías. Por tanto la verosimilitud es multinomial. Observemos que la multinomial no viene impuesta desde el principio como un modelo hipotético más o menos justificado, sino que surge del diseño, perfectamente natural, para la toma de datos que se ha establecido.

En esta variedad de modelos el problema tiene un parámetro multidimensional de la distribución multinomial, donde cada componente p_i es la probabilidad de un fallo de la categoría i-ésima (es decir, del *i*%). El parámetro de interés en el problema y sobre el que deseamos realizar inferencias es la cantidad de error en la contabilidad, que puede expresarse como una combinación lineal de las componentes p_i:

$$q = \frac{RBV}{100} \cdot \sum_{i=-100}^{100} i \cdot p_i ,$$

donde RBV es el valor registrado total del libro y es conocido.

La literatura publicada para esta variedad de modelos se puede clasificar en dos líneas. La primera línea (ver Tsui, Matsumura y Tsui, 1985; Matsumura, Tsui y Wong, 1987 o Matsumura et al., 1991) consiste en suponer especificada una distribución a priori para el parámetro multidimensional (de dimensión 201 ya que tenemos categorías desde -100 a 100, y por supuesto una distribución conjugada de Dirichlet), obtener la distribución a posteriori y a partir de ella calcular la a posteriori del parámetro de interés, θ. Dos aspectos destacamos de los trabajos de esta línea: El primero, es la elección normalizada de la distribución a priori, el usuario es incapaz de especificar su opinión sobre un parámetros de tan alta dimensión. El segundo es, la dificultad e incluso imposibilidad de obtener explícitamente la distribución a posteriori del parámetro de interés, lo que lleva a varias y diversas aproximaciones.

La segunda línea (McCray, 1984, 1986, 1992a y 1992b) consiste en considerar que la magnitud intuitiva de claro significado y sobre la que el usuario tiene información a priori es la cantidad de error θ. Esta segunda línea tiene un inconveniente y una gran ventaja. El inconveniente es que la distribución a priori no es combinable de manera inmediata con la verosimilitud del problema, pero esta dificultad técnica tiene también una solución técnica (Hernández-Bastida y Vázquez-Polo, 1997). La gran ventaja consiste en que al usuario se le ha pedido que especifique la distribución a priori sobre la magnitud que realmente es familiar para él y sobre la que verdaderamente tiene información a priori.

Obsérvese que en la última línea de trabajo descrita, el problema es uniparamétrico y por tanto no existe el problema de especificar una distribución a priori bidimensional, como en la primera variedad de trabajos descrita, ni hay que recurrir a la alternativa allí empleada, a saber, la especificación de marginales bajo una hipótesis de independencia entre los parámetros.

Ahora bien la gran ventaja, desde el punto de vista de la especificación de distribuciones a priori, que se ha señalado en la línea de McCray podría verse mermada si los modelos biparamétricos del principio fuesen robustos. En otras palabras:

(i) La especificación de distribuciones a priori marginales precisa especificar distribuciones a priori sobre magnitudes menos intuitivas para el auditor que "la cantidad total de error θ", pero si las conclusiones de los modelos fuesen robustas frente a variaciones razonables en estas distribuciones a priori, no habría un problema especial en estas especificaciones.

(ii) Al especificar marginales es imprescindible suponer independencia entre los parámetros para obtener la distribución conjunta del parámetro bidimensional (ϕ ,μ). ¿Hay evidencia empírica en las poblaciones contables que justifique la hipótesis de independencia? Menzefricke (1984) o Godfrey y Neter (1984) sugieren que la respuesta es negativa. Entonces la única alternativa es plantearnos si hay robustez en las conclusiones de estos modelos respecto a variaciones razonables en la hipótesis de independencia.

Desde un punto de vista bayesiano de especificación de distribuciones a priori para los parámetros del problema, (i) y (ii) son los dos grandes criterios para comparar los modelos biparamétricos, indicados al principio, con el modelo uniparamétrico, indicado más tarde dentro de la segunda variedad de modelos. Además las conclusiones que se obtengan serán valiosas para el usuario a la hora de elegir los diversos modelos propuestos en la literatura.

En este trabajo se estudia el criterio señalado como (ii). El trabajo está organizado de la siguiente manera: En la sección 2 se expone la metodología que utilizaremos para analizar la robustez frente a la hipótesis de independencia. En la sección 3 se presenta el modelo biparamétrico para auditoría de cuentas que se analiza en este trabajo. El epígrafe 4 recoge los resultados obtenidos y por último en el epígrafe 5 se recogen las conclusiones y posibles extensiones.

2. Metodología

Como ya se ha indicado antes, el planteamiento del problema biparamétrico se basa en dos parámetros ϕ y μ, que se suponen independientes; especificadas dos distribuciones marginales π_{01} y π_{02} para ϕ y para μ, se puede determinar completamente la distribución de (ϕ,μ) y de aquí la distribución de $\psi = \phi \cdot \mu$. Nuestro objetivo es examinar la robustez frente a la independencia; para ello la primera dificultad es ¿qué vamos a entender por no independencia entre los parámetros ϕ y μ, para que sea operativa nuestra idea? En otras palabras, es necesario modelizar la separación de la hipótesis de independencia. La modelización que hemos adoptado está basada en las clases de contaminación y es la siguiente:

Suponemos que la distribución a priori de (ϕ, μ) pertenece a la clase,

$$\Gamma_e = \{(1-e)p_0 + eq : q \in Q\},$$

ÍNDICE

donde,

$\boldsymbol{p}_0 = \boldsymbol{p}_{01} \bullet \boldsymbol{p}_{02}$, siendo π_{01} y π_{02} las distribuciones marginales para ϕ y μ, respectivamente.

$\boldsymbol{e} \in [0,1]$, es el grado de contaminación, y

Q es la clase contaminante y está formada por todas las distribuciones bidimensionales para (f, m) con marginales dadas por y

Obsérvese que,

$\boldsymbol{e} = 0$, supone independencia entre los parámetros y distribución a priori bidimensional completamente especificada.

$\boldsymbol{e} = 1$, supone conocer únicamente que la distribución a priori para (ϕ, μ) tiene como marginales π_{01} y π_0

$\boldsymbol{e} \in (0,1)$, se puede interpretar como un alejamiento gradual de la hipótesis de independencia.

Por supuesto que se ha elegido una forma concreta de alejamiento de la hipótesis de independencia que no es la única pero que tiene dos cualidades definitivas, desde nuestro punto de vista: es técnicamente tratable y, si se observara falta de robustez en esta forma concreta, podríamos deducir falta general de robustez frente a la hipótesis de independencia.

Nuestro objetivo es obtener el rango de variación sobre la clase Γ_e de las esperanzas a posteriori del parámetro de interés, que ha sido utilizado habitualmente para inferir la cantidad total media de error. Según sea el valor de la diferencia de tales extremos podrá hablarse de robustez o de falta de ella. Concretando, nuestro objetivo es determinar las siguientes cantidades,

$$\begin{aligned} \overline{r_e} &= \sup_{p_e \in \Gamma_e} E_{p_e}\left[f \cdot m \middle| datos \right] \\ \underline{r_e} &= \inf_{p_e \in \Gamma_e} E_{p_e}\left[f \cdot m \middle| datos \right] \end{aligned} \qquad (1)$$

donde la verosimilitud bidimensional será el producto de las verosimilitudes unidimensionales (ver Moors, 1983).

Para facilitar la discusión sobre la amplitud del rango de variación indicado usaremos un factor de sensibilidad relativa, *RS*, debido a Sivaganesan (1991) y definido por,

$$RS = \frac{\overline{r_e} - \underline{r_e}}{2 \cdot E_{p_0}\left[f \cdot m \mid datos\right]} \times 100$$

RS es útil para medir la cantidad de variación en porcentaje sobre Γ_e, alrededor de $E_{p_0}\left[f \cdot m \middle| datos\right]$ siendo ésta utilizada como patrón de comparación, en el papel de "centro" del intervalo de variación.

El cálculo del superior e inferior que aparece en (1) es extraordinariamente complejo para abordarlo directamente, por ello utilizaremos la técnica de linealización de Lavine, Wasserman y Wolpert (1991) y Wasserman, Lavine y Wolpert (1993). Esta técnica permite el cálculo de extremos de las esperanzas a posteriori sobre una clase dada a través de la obtención de las raíces de unas funciones de variable real, definidas por los extremos de las esperanzas de una función de variable real, lineal en dicha variable.

Se define la siguiente función real:

$$c(q,(f,m)) = L(f,m) \cdot (f \cdot m - q), \quad (f,m) \in \Theta; \; q \in \Re, \qquad (2)$$

que es claramente una función lineal en *q*, y decreciente (recordemos que la cantidad de interés es el producto $\psi = \phi\,\mu$). Consideremos las funciones de variable $q \in \Re$

$$g_e(q) = \sup_{p_e \in \Gamma_e} E_{p_e}\left[c(q,(f,m))\right] \quad \text{y} \quad h_e(q) = \inf_{p_e \in \Gamma_e} E_{p_e}\left[c(q,(f,m))\right], \qquad (3)$$

$$g(q) = \sup_{p \in Q} E_p\left[c(q,(f,m))\right] \quad \text{y} \quad h(q) = \inf_{p \in Q} E_p\left[c(q,(f,m))\right], \qquad (4)$$

donde por $E_p\left[c(q,(f,m))\right]$ se nota la esperanza de la función *c* según la distribución π que sea. Es claro que

$$g(q) = (1-e)E_{p_0}\left[c(q,(f,m))\right] + eg(q) =$$

$$= (1-e)E_{p_0}\left[L(f,m)(f \cdot m - q)\right] + eg(q)$$

$$= (1-e)E_{p_0}\left[L(f,m) \cdot f \cdot m\right] - q \cdot (1-e)E_{p_0}\left[L(f,m)\right] + eg(q) =$$

$$= (1-e) \cdot p \cdot \left(E_{p_0}\left[f \cdot m \mid datos\right] - q\right) + eg(q),$$

donde $p = \int_{\Theta} L(f,m) p_0(f,m) \cdot d\, f dm = E_{p_0}[L(f,m)]$, es la distribución predictiva de los datos bajo π_0. Es decir:

$$g_e(q) = (1-e) \cdot p \cdot \left(E_{p_0}[f \cdot m \mid datos] - q\right) + e g(q),$$

y análogamente:

$$h_e(q) = (1-e) \cdot p \cdot \left(E_{p_0}[f \cdot m \mid datos] - q\right) + e h(q). \qquad (6)$$

Lavine et al. (1991) demuestran el siguiente resultado.

Teorema

$$\overline{r}_e = \inf\{q \in \Re : g_e(q) \leq 0\}$$

y, el paso a extremo inferior es claro que

$$\underline{r}_e = \inf\{q \in \Re : h_e(q) \leq 0\},$$

y en particular

$$\begin{cases} Si\ \exists q \in \Re : g_e(q) = 0 \Rightarrow \overline{r}_e = \inf\{q : g_e(q) = 0\}, \\ Si\ \exists q \in \Re : h_e(q) = 0 \Rightarrow \underline{r}_e = \inf\{q : h_e(q) = 0\}. \end{cases}$$

Y si la raíz es única, esta raíz será la cota buscada.

Pero, de la definición dada se deduce que *g*(*q*) es continua, decreciente en *q* y convexa, como supremo de funciones lineales (*h*(*q*) es continua, decreciente en *q* y cóncava como ínfimo defunciones lineales). Por ello $g_e(q)$ y $h_e(q)$, heredan estas propiedades como funciones de *q*; y al ser $\Theta = [0,1] \times [0,1]$, $0 \leq f, m \leq 1$, entonces $g_e(0) > 0$, $g_e(1) < 0$ ($h_e(0) > 0$, $h_e(1) < 0$), con lo que está asegurada la existencia de una única raíz para $g_e(q)$ y $h_e(q)$ en [0,1], estas raíces serán, respectivamente $\underline{r}_e$ y $\underline{r}_e$. Es decir: la única raíz en [0,1] de $h_e(q)$ es $\underline{r}_e$ y la única raíz de $g_e(q)$ es r_e.

En esto consiste la técnica de linealización, en el uso de la función lineal en $q, c(q,(f,m))$ como función objetivo, y en lugar de las esperanzas a posteriori.

Si observamos (5) y (6), lo complicado es conocer la expresión de las funciones *g*(*q*) y *h*(*q*). Esta complicación se salva aplicando la técnica de discretización que proponen *Lavine et al.* (1991), y que permitirá obtener para diferentes $q \in \Re$, valores de *g*(*q*) y *h*(*q*). A continuación, dado $e \in [0,1]$, y sustituyendo en (5) y (6), habremos calculado valores de $g_e(q)$ y $h_e(q)$. Se trata de utilizar estos valores para localizar la raíz en un cambio de signo (mediante cualquier método numérico, en este trabajo hemos utilizado el método de la bisección). Es conveniente, como se sugiere en *Wasserman et al.* (1993), calcular $\overline{r}_e$, $\underline{r}_e$ para diferentes valores de ε, y obtener gráficas de las cotas frente a ε.

Por tanto, necesitamos determinar *g*(*q*) (*h*(*q*)) para algunos valores de *q*. El método elegido para cubrir este objetivo consiste en **discretizar** el espacio paramétrico, y pasar de la familia Q de densidades conjuntas, a una familia de distribuciones discretas que será el conjunto de oportunidades de un problema de programación lineal (formulado como el problema del transporte, cuyos orígenes se encuentran en el problema de Monge - Kantorovich, o de transferencia de masa, con amplias aplicaciones en escenarios probabilísticos, ver para ello *Tchen* (1980) y *Rachev* (1985a,b); buscando la distribución discreta que optimice $E[c(q,(f,m))]$.

Observemos que en este punto y con la metodología descrita, el problema se ha transformado en un problema lineal cuyo objetivo es determinar una distribución discreta $\{p_{ij}\}$ que, de acuerdo con las densidades marginales a priori p_{01} y p_{02}, minimice $\sum_{i,j} c_{ij} \cdot p_{ij}$, y planteado como sigue:

$$\min_{\{p_{ij}\}} \sum_{i=1}^{k} \sum_{j=1}^{l} c_{ij} \cdot p_{ij}$$

sujeto a:

$$\sum_{i=1}^{k} p_{ij} = \int_{\frac{j-1}{l}}^{\frac{j}{l}} p_{02}(m) \cdot dm, \;\; \mathrm{j} = 1,\dots,\mathrm{l}$$

$$\sum_{j=1}^{l} p_{ij} = \int_{\frac{i-1}{k}}^{\frac{i}{k}} p_{01}(f) \cdot df, \;\; \mathrm{i} = 1,\dots,\mathrm{k}$$

$$\sum_{i=1}^{k} \sum_{j=1}^{l} p_{ij} = 1$$

$$p_{ij} \geq 0, \;\; \forall \mathrm{i}, \mathrm{j}.$$

donde los coeficientes c_{ij} resultan de aplicar la función $c(q,(f,m))$ a los puntos de la partición escogida para discretizar el espacio paramétrico.

El valor óptimo de este problema para cada $q \in \Re$ dará el valor de la función –*g*(*q*), resolviendo el problema de dirección opuesta, es decir, el de maximización, obtendremos para cada $q \in \Re$ el valor de $-h(q)$. Sustituyendo los valores de $h(q)$ y $g(q)$ obtenidos en las expresiones (5) y (6), junto con el resto de las cantidades (que no dependen de q) se realiza un procedimiento iterativo de búsqueda de la raíz de $g_e(q)$ y $h_e(q)$ que culminará con la obtención de las cotas $\overline{r}_e$, $\underline{r}_e$, para ese valor de ε. La gran complicación viene dada por la discretización que se aplica al cálculo de cada valor de $g(q)$ y $h(q)$, complicación que crece con el tamaño de la partición.

Para el cálculo de los coeficientes de cada problema de programación lineal, se utilizó el programa MATLAB; mediante el paquete *Statistics "Continuous Distributions"* de MATHEMATICA se obtienen los coeficientes de los términos independientes de las restricciones de cada problema; conocidos los coeficientes y escritos en forma ecuacional (para lo que fue necesario el diseño de un programa apropiado en PASCAL), cada problema de programación lineal es resuelto con el paquete LINDO. Todo esto se realiza iterativamente, conectando a través de MATLAB los distintos programas. Una copia de ellos puede obtenerse solicitándola a cualquiera de los autores.

3. El Modelo Beta - Uniforme

De entre los modelos bayesianos biparamétricos mencionados en el epígrafe primero desarrollaremos nuestra metodología para el caso Beta - Uniforme (Godfrey y Neter, 1984), como caso representativo y más cercano a la naturaleza de los parámetros. Este modelo considera que la información a priori del auditor es tal que:

- $\phi \sim Beta(f_0 c, (1-f_0)c)$, por tanto

$$p_{01}(f) = \frac{\Gamma(c)}{\Gamma(f_0 c)\Gamma((1-f_0)c)} \cdot f^{f_0 c-1} \cdot (1-f)^{(1-f_0)c-1}, \quad 0 \le f \le 1,$$

de donde: $E[f] = f_0$, $V[f] = \dfrac{f_0(1-f_0)}{c+1}$.

- $\mu \sim U(0,1)$, por tanto

$$p_{02}(m) = 1, \; 0 \le m \le 1,$$

de donde, $E[m] = \dfrac{1}{2}$, $V[m] = \dfrac{1}{12}$.

La verosimilitud viene dada por:

- La v.a. número errores, M, que en una muestra de tamaño n se distribuye (aproximadamente) según una binomial de parámetros *n* y ϕ, por tanto

$$L_1(f \mid m) = P[M = m] = \binom{n}{m} f^m (1-f)^{n-m}.$$

Obsérvese que en este caso no se aproxima por distribución de Poisson, el mantener la distribución Binomial es útil ya que la distribución a priori para ϕ es beta, que es conjugada para la distribución binomial.

- La distribución de la fracción de error *Z* se considera exponencial truncada y para $m, \bar{z}$ es minimal suficiente, con lo que

$$L_2(m \mid \bar{z}) = \frac{\exp(-m\bar{z}/m)}{[m(1-\exp(-1/m))]^m}, \quad 0 \le \bar{z} \le 1.$$

Luego,

$$L(f,m) = L_1(f) \cdot L_2(m) = \binom{n}{m} f^m (1-f)^{n-m} \cdot \frac{\exp(-m\bar{z}/m)}{[m(1-\exp(-1/m))]^m}.$$

En cuanto a las **distribuciones finales** marginales se tiene que

- $f \mid m \sim \mathrm{Beta}(f_0 c + m, (1-f_0)c + n - m).$

- $m \mid \bar{z}$ se distribuye según la densidad a posteriori

$$p_{02}(m \mid \bar{z}) \propto L_2(m \mid \bar{z}) \cdot p_{02}(m) = \frac{\exp(-m\bar{z}/m)}{[m(1-\exp(-1/m))]^m},$$

con todo ello, y para aplicar (5) y (6), precisamos obtener

- $E_{p_0}[f \cdot m \mid datos] = E_{p_{01}}[f \mid datos] \cdot E_{p_{02}}[m \mid datos]$

$$= \frac{f_0 c + m}{c} \times k \times \int_0^1 \frac{\exp(-m\bar{z}/m)}{m^{m-1}(1-\exp(-1/m))^m} \cdot dm,$$

siendo $k = \left\{ \int_0^1 \frac{\exp(-m\bar{z}/m)}{[m(1-\exp(-1/m))]^m} \cdot dm \right\}^{-1}.$

ÍNDICE

- $p = E_{p_0}\left[L(f,m)\right]$ que obtenemos a partir de

$$\int_0^1 L_1(f)p_{01}(f)df = \binom{n}{m}\cdot\frac{\Gamma(c)\Gamma(f_0c+m)\Gamma((1-f_0)c+n-m)}{\Gamma(f_0c)\Gamma((1-f_0)c)\Gamma(c+n)}$$

$$\int_0^1 L_2(m)p_{02}(m)dm = \int_0^1\frac{\exp\left(-m\bar{z}/m\right)}{\left[m(1-\exp(-1/m))\right]^m}\cdot dm,$$

se tiene por tanto,

$$p = \binom{n}{m}\cdot\frac{\Gamma(c)\Gamma(f_0c+m)\Gamma((1-f_0)c+n-m)}{\Gamma(f_0c)\Gamma((1-f_0)c)\Gamma(c+n)}\cdot\int_0^1\frac{\exp\left(-m\bar{z}/m\right)}{\left[m(1-\exp(-1/m))\right]^m}\cdot dm,$$

- Y para aplicar la discretización necesitamos los coeficientes de la función objetivo del problema de programación lineal que son para cada $i=1,\ldots,k$ y $j=1,\ldots,l$:

$$c_{ij}(q) = \left(q-\frac{i\cdot j}{k\cdot l}\right)\cdot\binom{n}{m}\cdot\frac{i^m}{k^m}\left(1-\frac{i}{k}\right)^{n-m}\cdot\frac{\exp\left(\dfrac{-m\bar{z}}{j/l}\right)}{\left[\dfrac{j}{l}(1-\exp(-l/j))\right]^m}.$$

Como vemos, dependen del tamaño de la partición elegida y de la información muestral, de n, m y $\bar{z}$. Las restricciones dependerán del tamaño de la partición y de la información a priori, es decir, de ϕ_0 y c, su expresión será

$$\sum_{i=1}^{k} p_{ij} = \int_{\frac{j-1}{l}}^{\frac{j}{l}} p_{02}(m)\cdot dm = \frac{1}{l},\ \ j=1,\ldots,l,$$

$$\sum_{j=1}^{l} p_{ij} = \int_{\frac{i-1}{k}}^{\frac{i}{k}} p_{01}(f)\cdot df,\ \ i=1,\ldots,k\quad \left(p_{01}(f)\sim \mathrm{Beta}(f_0c,(1-f_0)c)\right).$$

4. Resultados obtenidos

Los resultados que se presentan de este modelo corresponden a cinco casos de información a priori diferentes:

1. Tasa de error baja.
   (a) Con Moda: $f_0 = 0.05,\ c = 25$ (varianza baja).
   (b) Sin Moda: $f_0 = 0.05,\ c = 3$ (varianza alta).

2. Tasa de error media.
   (a) Con Moda: $f_0 = 0.1,\ c = 15.$
   (b) Sin Moda: $f_0 = 0.1,\ c = 4.$

3. Tasa de error alta: $f_0 = 0.25,\ c = 5.$

En las tablas que se insertan a continuación se recogen estas situaciones correspondientes a diferentes valores de *m* y $\bar{z}$, centrando además nuestra atención en el factor RS.

4.1. *Caso de Tasa de Error Baja*

Los resultados obtenidos en este modelo para los casos de tasa de error a priori baja ($f_0 = 0.05$) y con varianza alta o baja ($c = 25$ ó 3) indican una fuerte sensibilidad a pequeñas contaminaciones, siendo mayor esta sensibilidad en el caso de varianza alta.

Tabla 1: Modelo: BU, Tasa de Error: Baja, Varianza alta

	m=10, $\bar{z}$ = 0.01				m=10, $\bar{z}$ =0.5			
ε	$\overline{r_e}$	Media Post.	$\underline{r_e}$	*RS*	$\overline{r_e}$	Media Post.	$\underline{r_e}$	RS
0	1231.80	1231.80	1231.80	0.00	75362.77	75362.77	75362.77	0.00
0.05	19493.01	1231.80	976.75	751.60	132541.18	75362.77	22492.07	73.01
0.1	19493.01	1231.80	976.75	751.60	132629.30	75362.77	22350.50	73.15
0.2	19493.01	1231.80	976.75	751.60	132629.30	75362.77	22284.89	73.21
0.25	19493.01	1231.80	976.75	751.60	132629.30	75362.77	22284.89	73.21
0.5	19493.01	1231.80	976.75	751.60	132629.30	75362.77	22284.89	73.21
1	19493.01	1231.80	976. 75	751.60	132671.00	75362.77	22284.80	73.24
	m = 10, $\bar{z}$ = 0.7				m = 10, $\bar{z}$ = 1			
0	83043.65	83043.65	83043.65	0.00	88134.64	88134.64	88134.64	0.00
0.05	146652.89	83043.65	33994.10	67.83	171742.34	88134.64	59802.82	63.50
0.1	146978.47	83043.65	33994.10	68.03	171742.34	88134.64	59802.82	63.50
0.2	146978.47	83043.65	33994.10	68.03	171742.34	88134.64	59802.82	63.50
0.25	146978.47	83043.65	33994.10	68.03	171742.34	88134.64	59802.82	63.50
0.5	146978.47	83043.65	33994.10	68.03	171742.34	88134.64	59802.82	63.50
1	146978.47	83043.65	23388.68	74.41	171746.09	88134.64	0.00	97.43

Tabla 2. Modelo: BU, Tasa de Error: Baja, Varianza baja

	m=10, $\bar{z}$ = 0.01				m=10, $\bar{z}$ =0.5			
ε	$\overline{r_e}$	Media Post.	$\underline{r_e}$	*RS*	$\overline{r_e}$	Media Post.	$\underline{r_e}$	RS
0	1125.00	1125.00	1125.00	0.00	76476.50	76476.50	76476.50	0.00
0.05	6052.49	1125.00	981.33	225.39	119340.42	76476.50	39894.68	51.94
0.1	6052.49	1125.00	981.33	225.39	119348.43	76476.50	39856.53	51.97
0.2	6052.49	1125.00	981.33	225.39	119524.30	76476.50	39853.10	52.09
0.25	6052.49	1125.00	981.33	225.39	119532.43	76476.50	39853.10	52.10
0.5	6052.49	1125.00	981.33	225.39	119632.16	76476.50	39843.56	52.16
1	6052.50	1125.00	981.24	225.40	119720.68	76476.50	39834.65	52.23
	m = 10, $\bar{z}$ = 0.7				m = 10, $\bar{z}$ = 1			
0	84270.90	84270.90	83043.65	0.00	80493.41	80493.41	80493.41	0.00
0.05	144007.02	84270.90	33994.10	60.81	171742.34	80493.41	59802.82	69.53
0.1	144050.50	84270.90	33994.10	60.83	171742.34	80493.41	59802.82	69.53
0.2	144050.50	84270.90	33994.10	60.98	171742.34	80493.41	59802.82	69.53
0.25	144050.50	84270.90	33994.10	60.98	171742.34	80493.41	59802.82	69.53
0.5	144050.50	84270.90	33994.10	60.98	171742.34	80493.41	59802.82	69.53
1	144050.51	84270.90	23388.68	60.98	171742.68	80493.41	0.00	106.68

4.2. *Caso de Tasa de Error Media*

Tal y como vemos para este modelo aparecen respuestas superiores al 50% para grados de contaminación de ε =0.05. Los casos m=10 y $\bar{z}$ =0.01 son especialmente sensibles. En algunas situaciones las cotas superiores e inferiores son extremas para cualquier grado de contaminación ε, debido a que la predictiva, para estos casos, se hace prácticamente nula y la función $g_e(q)$ (respectivamente, $h_e(q)$), quede reducida a $g(q)$ (respectivamente, $h(q)$) dadas en las expresiones (5) y (6). Nuevamente obtenemos unos valores muy elevados para el coeficiente *RS* (obsérvese, por ejemplo, el caso de tasa media con varianza alta para m=10 y $\bar{z}$ = 0.01, para la misma situación muestral el caso de tasa media y varianza baja). En la Figura 1 se representan el mayor y menor valor de RS para m=1 y tasa media con varianza baja. La variabilidad de la sensibilidad relativa es elevada. Los casos con m=1 no presentan una respuesta tan contundente para ε =0.05,aunque para m=1, $\bar{z}$ =0.01, RS (ε=0.05)=29.85, que es ya un valor considerable.

Tabla 3: Modelo: BU, Tasa de Error: Media, Varianza alta

	m=10, $\bar{z}$ = 0.01				m=10, $\bar{z}$ =0.5			
ε	$\overline{r}_e$	Media Post.	$\underline{r}_e$	*RS*	$\overline{r}_e$	Media Post.	$\underline{r}_e$	RS
0	1250.00	1250.00	1250.00	0.00	76476.50	76476.50	76476.50	0.00
0.05	41398.72	1250.00	976.75	1616.88	110968.49	76476.50	54165.84	37.14
0.1	41398.72	1250.00	976.75	1616.88	118284.13	76476.50	47732.35	46.13
0.2	41398.72	1250.00	976.75	1616.88	128446.48	76476.50	41487.31	56.85
0.25	41398.72	1250.00	976.75	1616.88	132310.58	76476.50	39450.26	60.71
0.5	41398.72	1250.00	976.75	1616.88	141993.43	76476.50	32840.54	71.36
1	41398.72	1250.00	976.75	1616.88	148224.06	76476.50	22831.63	81.98

	m = 10, $\bar{z}$ = 0.7				m = 10, $\bar{z}$ = 1			
0	84270.90	84270.90	84270.90	0.00	89437.12	89437.12	89437.12	0.00
0.05	164196.68	84270.90	33994.10	77.25	171742.34	89437.12	59802.82	62.58
0.1	164196.68	84270.90	33994.10	77.25	171742.34	89437.12	59802.82	62.58
0.2	164196.68	84270.90	33994.10	77.25	171742.34	89437.12	59802.82	62.58
0.25	164748.86	84270.90	33994.10	77.58	171742.34	89437.12	59802.82	62.58
0.5	164748.86	84270.90	33994.10	77.58	171742.34	89437.12	59802.82	62.58
1	164748.90	84270.90	17802.31	87.19	171742.40	89437.12	0.00	96.01

Tabla 4: Modelo: BU, Tasa de Error: Media, Varianza baja

	m=10, $\bar{z}$ = 0.01				m=10, $\bar{z}$ =0.5			
ε	$\overline{r}_e$	Media Post.	$\underline{r}_e$	*RS*	$\overline{r}_e$	Media Post.	$\underline{r}_e$	RS
0	5776.74	5776.74	5776.74	0.00	13392.04	13392.04	13392.04	0.00
0.05	8066.08	5776.74	4617.12	29.85	15452.67	13392.04	13090.71	8.82
0.1	10163.78	5776.74	4078.87	52.67	17233.94	13392.04	12800.03	16.55
0.2	13877.39	5776.74	3566.17	89.25	20156.76	13392.04	12248.80	29.52
0.25	15527.63	5776.74	3423.88	104.76	21372.13	13392.04	11987.88	35.04
0.5	22233.30	5776.74	3076.74	165.81	25867.94	13392.04	10740.09	56.48
1	30856.99	5776.74	2861.4	242.31	34275.91	13392.04	8343.6	96.82

	m = 10, $\bar{z}$ = 0.7				m = 10, $\bar{z}$ = 1			
0	14301.33	14301.33	14301.33	0.00	15261.70	15261.70	15261.70	0.00
0.05	16385.94	14301.33	14069.94	8.10	17387.29	15261.70	15095.71	7.51
0.1	18156.91	14301.33	13840.29	15.09	19150.26	15261.70	14267.20	16.00
0.2	20997.33	14301.33	13386.73	26.61	21895.26	15261.70	14578.06	23.97
0.25	22156.62	14301.33	13158.99	31.46	22990.89	15261.70	14394.19	28.16
0.5	26395.51	14301.33	12023.74	50.25	27000.33	15261.70	13380.24	44.62
1	43040.94	14301.33	9244.80	118.16	53741.93	15261.70	10173.37	142.74

Tabla 5: Modelo: BU, Tasa de Error: Media, Varianza baja (continuación)

	m=10, $\bar{z}$ = 0.01				m=10, $\bar{z}$ =0.5			
ε	$\bar{r}_e$	Media Post.	$\underline{r}_e$	*RS*	$\bar{r}_e$	Media Post.	$\underline{r}_e$	RS
0	1250.00	1250.00	1250.00	0.00	76476.50	76476.50	76476.50	0.00
0.05	15322.78	1250.00	976.75	573.84	119340.42	76476.50	39894.68	51.94
0.1	15322.78	1250.00	976.75	573.84	119348.43	76476.50	39856.53	51.97
0.2	15322.78	1250.00	976.75	573.84	119524.30	76476.50	39853.10	52.09
0.25	15322.78	1250.00	976.75	573.84	119532.43	76476.50	39853.10	52.10
0.5	15322.78	1250.00	976.75	573.84	119632.43	76476.50	39843.56	52.17
1	15322.78	1250.00	976.72	573.84	119632.43	76476.50	39843.56	52.17
	m = 10, $\bar{z}$ = 0.7				m = 10, $\bar{z}$ = 1			
0	84270.90	84270.90	84270.90	0.00	89437.12	89437.12	89437.12	0.00
0.05	144007.02	84270.90	41518.21	60.81	171742.00	89437.12	59802.82	62.58
0.1	144050.50	84270.90	41518.21	60.83	171742.00	89437.12	59802.82	62.58
0.2	144050.50	84270.90	41275.98	60.98	171742.00	89437.12	59802.82	62.58
0.25	144050.50	84270.90	41275.98	60.98	171742.00	89437.12	59802.82	62.58
0.5	144050.50	84270.90	41275.98	60.98	171742.00	89437.12	59802.82	62.58
1	144050.51	84270.90	41275.88	60.98	171742.29	89437.12	0.00	96.01

4.3. *Caso de Tasa de Error Alta*

Para este último caso presentado se observa una sensibilidad moderada para los casos de una información muestral de $m=0,1$ errores detectados (menor para $m=0$), pero para los casos en los que aparezcan $m=10$ errores en la muestra, el factor de sensibilidad relativa aumenta (espectacularmente para $\bar{z}$ =0.01). Nuevamente nos encontramos con que la respuesta a una pequeña contaminación es elevada. Se trata de una situación en la que la oscilación de RS es elevada.

Tabla 6: Modelo: BU, Tasa de Error: Alta

	m=10, $\bar{z}$ = 0.01				m=10, $\bar{z}$ =0.5			
ε	$\bar{r}_e$	Media Post.	$\underline{r}_e$	*RS*	$\bar{r}_e$	Media Post.	$\underline{r}_e$	RS
0	1339.29	1339.29	1339.29	0.00	81393.11	81393.11	81393.11	0.00
0.05	98907.85	1339.29	976.75	3656.09	173166.28	81393.11	39873.70	81.34
0.1	98915.48	1339.29	976.75	3656.38	173469.77	81393.11	39854.24	81.53
0.2	98919.30	1339.29	976.75	3656.52	173591.23	81393.11	39853.10	81.61
0.25	98920.06	1339.29	976.75	3656.55	173591.23	81393.11	39853.10	81.61
0.5	98922.35	1339.29	976.75	3656.64	173659.90	81393.11	39843.56	81.66
1	98922.92	1339.29	976.66	3656.66	173660.88	81393.11	26059.91	90.07

Tabla 6: Modelo: BU, Tasa de Error: Alta (Cont.)

	m = 10, $\bar{z}$ = 0.7				m = 10, $\bar{z}$ = 1			
0	90290.25	90290.25	90290.25	0.00	95825.49	95825.49	95825.49	0.00
0.05	205636.22	90290.25	41518.21	90.88	171742.63	95825.49	59802.82	58.41
0.1	213019.94	90290.25	41275.98	95.11	171742.63	95825.49	59802.82	58.41
0.2	213019.94	90290.25	41275.98	95.11	171742.63	95825.49	59802.82	58.41
0.25	213019.94	90290.25	41275.98	95.11	171742.63	95825.49	59802.82	58.41
0.5	213019.94	90290.25	41275.98	95.11	171742.63	95825.49	59802.82	58.41
1	213020.09	90290.25	0.00	117.96	171742.63	95825.49	0.00	89.61

5. Conclusiones y posibles extensiones

Resulta conveniente, llegado este punto, recapitular y hacer una valoración de los resultados obtenidos para este modelo. En general, podemos decir que se observa una fuerte sensibilidad incluso para pequeños grados de desconfianza ($\varepsilon = 0.05$), espectacular en muchos casos. Es decir, hemos encontrado situaciones, plausibles con el escenario de población considerado, en las que es notable la ausencia de robustez para los modelos considerados. Esto indica que es peligrosa la utilización de tales modelos en aquellos casos en los que no hay seguridad sobre la hipótesis de independencia. Puede pensarse que la clase elegida para el estudio es excesivamente amplia, y que por ello el rango de variación obtenido es tan grande. De hecho es así, lo que ocurre es que esta clase se ha elegido como clase contaminante en un modelo de contaminaciones y es de esperar que para $\varepsilon = 1$ el valor de *RS* pueda ser elevado, pero es que en muchos casos se obtienen resultados espectaculares para RS con $\varepsilon = 0.05$. Recalcamos, una vez más, la fuerte respuesta de los modelos a un mínimo grado de desconfianza en la hipótesis de independencia.

La mayoría de los trabajos referidos al estudio de la robustez multiparamétrica o a medir la sensibilidad de la hipótesis de independencia de los parámetros, abundan en la gran complejidad del manejo de la clase de bidimensionales a priori con marginales dadas, debido a su gran variabilidad, lo que puede provocar conclusiones vacías, pero sí que coinciden en su utilidad como clase contaminante, como una buena herramienta para el diagnóstico de la independencia. Nótense los resultados que se obtienen en los apartados anteriores en muchos casos ya para $\varepsilon = 0.05$.

Es muy interesante determinar si, ante un modelo multiparamétrico, y antes de asignar distribuciones a priori marginales, es más sostenible la asunción de determinadas hipótesis a priori (independencia, unimodalidad,...) que la total determinación de las marginales a priori (punto de partida de los modelos utilizados en auditoría).

Berger y Moreno (1994) recomiendan partir de la hipótesis de independencia, con la excusa de que son imprescindibles fuertes suposiciones de este tipo en problemas multiparamétricos, y así consideran la clase

$$\Gamma_C = \{(1-\varepsilon)p_0 + \varepsilon q : q \in Q\},$$

siendo $e \in [0,1]$, p_0 la bidimensional bajo hipótesis de independencia (con marginales p_{01}, p_{02}) y para

$$Q = \{p(q_1, q_2) = [(1-e_1)p_{01}(q_1) + e_1 q_1(q_1)] \times [(1-e_2)p_{02}(q_2) + e_2 q_2(q_2)] : \mathrm{q_i} \in Q_i, \mathrm{i} = 1,2\},$$

$e_1, e_2 \in [0,1]$, es decir, construyen la clase contaminante como una clase que incorpora la suposición de independencia, formada por bidimensionales que son producto de marginales; esto permite introducir diferentes grados de confianza en las marginales p_{01}, p_{02}, a través de e_1, e_2, y proponen como clases contaminantes marginales, Q_1, Q_2, clases dadas por cuantiles conocidos y clases de distribuciones unimodales con la misma moda que las distribuciones p_{01}, p_{02}, (incorporando así, a la suposición de independencia, la suposición de unimodalidad). En este escenario, es posible aplicar resultados anteriores del caso univariante, lo que permite la obtención de expresiones elegantes para los extremos de la media a posteriori sobre estas clases.

Se observan así dos tendencias a la hora de abordar el problema multiparamétrico: investigar el problema de desviaciones de la hipótesis de independencia, pero con las marginales totalmente determinadas (*Lavine et al.* (1991) , *Wasserman et al.* (1993)) o investigar el efecto de desviaciones de las marginales a priori, pero sin debilitar la hipótesis de independencia (*Berger y Moreno* (1994)). Uno de los objetivos de este artículo era el estudio de la robustez de la hipótesis de independencia, luego no parece conveniente contaminar con clases que preservan esta independencia.

Una posibilidad es considerar otra clase como clase contaminante en un modelo de contaminaciones. En vez de considerar la clase de distribuciones bidimensionales con marginales fijas, sería posible afinar más, y utilizar clases incluidas en ésta, pero que permitan medir el tipo de dependencia que existe entre las variables marginales. Este tipo de clases vienen dadas por distribuciones bivariantes de la forma $h(x,y) = f(x) \cdot g(y) \cdot s(x,y)$; donde *f* y *g* son las marginales y *s*(*x*, *y*) es una función que mide la dependencia entre ambas. Según las asignaciones para s(x, y) se obtienen diferentes clases (la de Farlie - Gumbel - Morgenstern (FGM), Gumbel tipo I, Gaussiana, Placket y Pareto).

Un estudio reciente para una asignación particular de s(x, y) que permite un tipo poligonal de dependencia puede verse en Long y Krzysztofowicz (1995).

En De la Horra y Fernández (1995) y Fernández (1994), se considera la clase FGM (es una clase con una estructura de dependencia algo rígida, con coeficiente de correlación en

$$\left(-\frac{1}{3}, \frac{1}{3}\right),$$

permitiendo desviaciones moderadas de la independencia a priori; es una clase adecuada cuando hay cierta confianza en la independencia a priori) como clase contaminante en un modelo de contaminaciones, y obtienen expresiones para los extremos de las esperanzas a posteriori.

Bibliografía

BERGER, J.O. y MORENO, E. (1994) *Bayesian Robustness in Bidimensional Models: Prior Independence.* Journal of Statistical Planning and Inference, 40, 161-176.

COX, D.R. y SNELL, E.J. (1979). *On Sampling and the Estimation of Rare Errors.* Biometrika; 66-1; 125-132.

DE LA HORRA, J. y FERNANDEZ, C. (1995). *Sensitivity to Prior Independence via Farlie-Gumbel-Morgenstern Model*. Communications in Statistics. Theory and Methods, Vol. 24, 4, 987-996.

FERNANDEZ LLANA, M.C.(1994). *Estudios sobre Robustez Bayesiana Global.* Tesis Doctoral. Univ. Autónoma de Madrid.

GODFREY, J. y NETER, J. (1984). *Bayesian bounds for monetary unit sampling in accounting and auditing.* Journal of Accounting Research; 22-2; 497-525.

HERNANDEZ-BASTIDA, A., y VAZQUEZ-POLO, F.J. (1997). *A Note on the Quasi-Bayesian Audit Risk Model for Dollar Unit Sampling.* The European Accounting Review, 6:3, 501-507.

LAVINE, M., WASSERMAN, L. y WOLPERT, R.(1991). *Bayesian Inference with Specified Prior Marginals.* Journal of American Statistical Associaton, 86, 964-981.

LONG, D. y KRZYSZTOFOWIC, R. (1995). *A Family of Bivariate Densities Constructed from Marginals.* Journal of American Statistical Associaton, Vol.90, 430, 739-746.

MATSUMURA, E.M., PLANTE, R., TSUI, K. y KANNAN, P. (1991). *Comparative Performance of two Multinomial-Based Methods for Obtaining Lower Bounds on the Total Overstatement Error in Accounting Populations.* Journal of Business and Economics Statistics, 9, 4, 423-429.

MATSUMURA, E.M., TSUI, K. y WONG, W. (1987). *An Extended Multinomial- Dirichlet Model for Error Bounds for Dollar-Unit Sampling.* Contemporary Accounting Research, 6, 2, 485-500.

McCRAY, J.H. (1984). *A quasi-bayesian audit risk model for dollar unit sampling.* The Accounting Review; LIX-1; 35-51.

McCRAY, J.H. (1986). *A General Bayesian Risk Model for Dollar-Unit Sampling and Multiple Populations.* The College of William and Mary. School of Business Administration.

McCRAY, J.H.(1992a). MLPC. *Most Likely Probability Curves. Description of Software.* The College of William and Mary. School of Business Administration.

McCRAY, J.H.(1992b). *MLPC. Most Likely Posterior Curves. Working Paper.* The College of William and Mary. School of Business Administration.

MENZEFRICKE, U. (1984). *Using Decision Theory for Planning Audit Sample Size with Dollar Unit Sampling.* Journal of Accounting Research, 22-2, 570-587.

MOORS, J.J.A. (1983). *Bayes' Estimation in Sampling for Auditing.* The Statistician, 32, 281-288.

RACHEV, S.T. (1985a). *On a Class of Minimal Functionals on a Space of Probability Measures*. Theory of Probability and its Applications, Vol.XXIX, 1, 41-49.

RACHEV, S.T. (1985b)*. The Monge-Kantorovich Mass Transference Problem and its Stochastics Applications.* Theory of Probability and its Applications, Vol.XXIX, 4, 647-676.

SIVAGANESAN, S. (1991) *Sensitivity of Some Posterior Summaries when the Prior is Unimodal with Specified Quantiles.* The Canadian Journal of Statistics. Vol.19, 1, 57-65.

TCHEN, A. H. (1980). *Inequalities for Distributions with Given Marginals.* The Annals of Probability. Vol.8, 4, 814-827.

TSUI, K., MATSUMURA, E.M. y TSUI, K. (1985). *Multinomial-Dirichlet Bounds for Dollar-Unit Sampling in Auditing.* The Accounting Review, 40, 1, 76-96.

WASSERMAN, L., LAVINE, M. y WOLPERT, R.L. (1993). *Linearization of Bayesian Robustness Problems.* Journal of Statistical Planning and Inference, 37, 307-316.