

Sobre los Modelos Biparamétricos en Auditoría de Cuentas

María del Carmen Martel Escobar

Universidad de Las Palmas de G.C.

Agustin Hernández Bastida

Universidad de Granada

Resumen

Los modelos estadísticos bayesianos desarrollados en la literatura para estimar el error de una contabilidad en auditoría de cuentas se han basado habitualmente en un planteamiento biparamétrico. En todos esos modelos ha sido fundamental la hipótesis de independencia entre los parámetros.

Tal hipótesis tiene un escaso fundamento en la práctica y sólo razones de conveniencia matemática la sustentan. En este trabajo se analiza el comportamiento de dichos modelos frente a separaciones o alejamientos de la hipótesis de independencia.

1 Introducción y Planteamiento.

Es bien conocida la utilidad de los procedimientos estadísticos en auditoría contable y en particular la utilidad del análisis estadístico bayesiano.

Entre los modelos bayesianos para la estimación del error total en una contabilidad se sitúan los modelos biparamétricos, consistentes en disponer de información muestras sobre dos magnitudes observables (el número de errores encontrados en la muestra y su tamaño), cuyas verosimilitudes dependen de dos parámetros.

Estos parámetros corresponden a dos magnitudes sobre las que es posible emitir un juicio a priori y son la probabilidad de aparición de error, la tasa de error, θ , y el tamaño medio de error en presencia de error, denotado habitualmente por μ .

La cantidad de interés en el problema, sobre la que se desea realizar inferencias, es la cantidad total de error en la población, T_y , que es posible expresar de forma proporcional al producto de los parámetros ϕ y μ , siendo el valor de libro conocido, T_x , la constante de proporcionalidad.

Es decir, es posible expresar:

$$T_y = T_x \cdot \phi \cdot m$$

(para un desarrollo detallado de esta cuestión ver Martel (1996)).

La literatura publicada en esta línea consiste en

- suponer especificadas sendas distribuciones a priori para los parámetros ϕ y μ .
- disponer de una información muestras dada por el tamaño de la muestra, n , el número de errores detectados en ella, y su tamaño medio, \bar{z} (\bar{z} es la media muestral de la fracción muestras de error en los ítems con error, i.e., $\bar{z} = \frac{1}{m} \sum_{i=1}^m z_i : z_i = \frac{y_i}{x_i}$, siendo y_i la cantidad de error del ítem (unidad muestras) i y x_i el valor total registrado de ese ítem).
- suponer que los parámetros ϕ y m son independientes, y
- calcular la distribución a posteriori para el parámetro de interés $\omega = \phi \cdot m$

Según sean las elecciones de las distribuciones a priori de los parámetros y según se formalice la información muestras, se plantean distintos modelos: Beta-Normal de *Felix* y *Grimlund* (1977), Gamma - Gamma inversa de *Cox* y *Snell* (1979) y Gamma - Uniforme y Beta Uniforme de *Godfrey* y *Neter* (1984), entre los más utilizados.

Pero surge la duda sobre si es realista tal suposición de independencia, es decir, ¿son de verdad independientes la fracción media de error (en ítems con error) y la tasa de error de la población? A este respecto, *Menzefricke* (1984), considera que es una suposición importante pero que hay escasa evidencia empírica disponible para comprobarla. Para *Godfrey* y *Neter* (1984), hay relativamente poco conocimiento del proceso por el que se generan los errores, pero podría ser que la magnitud de la cantidad de error está relacionada con su plausibilidad. Discuten así la independencia entre el proceso que determina si un ítem tiene o no error (tasa de error) y la cantidad de ese error (fracción de error) y cobra interés estudiar qué efecto tiene

sobre el modelo suponer independencia aún cuando no la hubiera. Un estudio de este tipo es el que planteamos aquí: analizar la sensibilidad de alguno de estos modelos a desviaciones de la hipótesis de independencia.

Como hemos indicado antes, esta hipótesis de independencia es habitual e imprescindible en todos los modelos publicados en la literatura. No obstante, en la práctica, es una hipótesis de difícil justificación sobre la que se han hecho diversas consideraciones a lo largo del tiempo para intentar su justificación aunque la verdadera razón de fondo para esta hipótesis siempre ha sido técnica. Si no se supone independencia entre ϕ y μ , la obtención de la distribución a posteriori de $\omega = \phi \cdot \mu$ es técnicamente de una complejidad extrema.

La modelización que hemos adoptado para medir la separación de la independencia está basada en las clases de contaminación.

Sea $\phi = (\phi, \mu)$ donde $\phi, \mu \in [0, 1]$ suponemos que la distribución a priori de ϕ pertenece a la clase

$$T_e = \{(1 - e)p_0 + ep, \text{ tal que } p \in Q\},$$

donde, $p_0 = p_{01}p_{02}$, siendo p_{01} y p_{02} dos distribuciones marginales para f y μ , $e \in [0,1]$, Q es la clase contaminante que está formada por todas las distribuciones bidimensionales para $f = (f, \mu)$ con marginales dadas, p_{01}, p_{02} .

Obsérvese que,

$e = 0$ supone independencia entre los parámetros y distribución a priori bidimensional completamente especificada,

$e = 1$, supone conocer únicamente que la distribución a priori para $q = (f, \mu)$ tiene como marginales p_{01} y p_{02} .

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

Se trata de obtener el rango de variación sobre la clase T_e , de las esperanzas a posteriori de la función paramétrica de interés (en este caso, la dada por su producto: $f \cdot \mu$, para ello habrá que calcular los extremos que la esperanza a posteriori alcanza sobre la clase referida.

Según sea el valor de la diferencia de tales extremos, podrá hablarse de robustez o de su ausencia. Nótese que un valor pequeño para este rango, implicará que cualquier elección de distribución a priori sobre esa clase conduce a resultados parecidos, por lo que habrá justificación para escoger la distribución π_0 y por tanto para la hipótesis de independencia. Sin embargo, un valor elevado para esa diferencia entre supremo e ínfimo, implicará una gran variabilidad: que la hipótesis de independencia no es robusta.

Pero esto es una labor muy compleja que no puede ser abordada directamente, ya que conduce a complejos problemas de cálculo de extremos, a problemas de optimización de una función no lineal, como es la esperanza a posteriori de $\mathbf{f.m}$, sobre la clase T_ϵ ,

Para hacer frente a esta situación se recurre a la técnica de linealización que se propone en Lavine et al. (1991) y Wasserman et al. (1993), que 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 (es la función que consigue la linealización).

En el epígrafe siguiente se describe esta técnica así como la aproximación del problema al caso discreto, que va a permitir su tratamiento y resolución. A continuación se ofrece una aplicación al modelo Ganuna - Uniforme de Godfrey y Neter.

2 Linealización del problema.

Como se acaba de indicar, hay que calcular las cotas:

$$\left\{ \begin{array}{l} p_e = \sup_{p_e \in r} E_p [\mathbf{f.m} | \text{datos}]_p \\ p_e = \inf_{p_e \in r_e} E_p [\mathbf{f.m} | \text{datos}]_p \end{array} \right\}$$

y el rango $\overline{p_e} - \underline{p_e}$ a diferentes valores de $\epsilon \in [0, 1]$, donde $E\pi, [\mathbf{f.m} | \text{datos}]$ representa la esperanza a posteriori de $\mathbf{f.m}$, para una distribución $\pi_\epsilon \in T_\epsilon$

En el epígrafe anterior ya vimos que este cálculo conduce a un problema de optimización no lineal. Como una vía de resolución se utilizará la técnica de linealización y posterior

discretización propuesta por Lavine, Wasserman y Wolpert (1991); Wasserman et al. (1993).

La metodología consiste en lo siguiente:

Se define la siguiente función real:

$$c(q, (\mathbf{f}, \mathbf{m})) = L(\mathbf{f}, \mathbf{m})(\mathbf{f} \cdot \mathbf{m} - q), (\mathbf{f}, \mathbf{m}) \in \Theta; q \in R,$$

que es claramente una función lineal en q , y decreciente (recordemos que la cantidad de interés es el producto $\mathbf{y} = \mathbf{f} \cdot \mathbf{m}$). Consideremos las funciones de variable $q \in R$

$$g_e(q) = \sup_{p_e \in \Gamma} E_{p_e} [c(q, (\mathbf{f}, \mathbf{m}))] \text{ y } h_e(q) = \inf_{p_e \in \Gamma} E_{p_e} [c(q, (\mathbf{f}, \mathbf{m}))]$$

$$g(q) = \sup_{p_e \in Q} E_{p_e} [c(q, (\mathbf{f}, \mathbf{m}))] \text{ y } h(q) = \inf_{p_e \in Q} E_{p_e} [c(q, (\mathbf{f}, \mathbf{m}))]$$

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

$$\begin{aligned} g_e(q) &= (1 - e) E_{p_0} [c(q, (\mathbf{f}, \mathbf{m}))] + e g(q) \\ &= (1 - e) E_{p_0} [L(\mathbf{f}, \mathbf{m})(\mathbf{f} \cdot \mathbf{m} - q)] + e g(q) \\ &= (1 - e) E_{p_0} [L(\mathbf{f}, \mathbf{m}) \mathbf{f} \cdot \mathbf{m}] - q \cdot (1 - e) E_{p_0} [L(\mathbf{f}, \mathbf{m})] + e g(q) \\ &= (1 - e) \cdot p \cdot (E_{p_0} [\mathbf{f} \cdot \mathbf{m} \mid \text{datos}] - q) + e g(q) \end{aligned}$$

donde $p = \int_{\Theta} L(\mathbf{f}, \mathbf{m}) p_0(\mathbf{f}, \mathbf{m}) d\mathbf{f} d\mathbf{m} = E_{p_0} [L(\mathbf{f}, \mathbf{m})]$, es la distribución predictiva de los datos

bajo π_0 . Es decir:

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

y análogamente,

$$h_e(q) = (1 - e) \cdot p \cdot (E_{p_0} [\mathbf{f} \cdot \mathbf{m} \mid \text{datos}] - q) + e h(q)$$

lavine et al (1991) demuestra el siguiente resultado.

Teorema.

$$\overline{p_e} = \inf \{q \in R : g_e(q) \leq 0\}$$

y, el paso a extremo inferior es claro que

$$\underline{p_e} = \inf \{q \in R : h_e(q) \leq 0\}$$

y en particular

$$\left\{ \begin{array}{l} \text{Si } \exists q \in R : g_{\epsilon}(q) = 0 \Rightarrow \bar{p}_{\epsilon} = \inf \{q : g_{\epsilon}(q) = 0\} \\ \text{Si } \exists q \in R : h(q) = 0 \Rightarrow \underline{p}_{\epsilon} = \inf \{q : h_{\epsilon}(q) = 0\} \end{array} \right\},$$

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 de funciones lineales). Por ello, $g_{\epsilon}(q)$ y $h_{\epsilon}(q)$ heredan estas propiedades como funciones de q ; y al ser $\Theta = [0, 1] \times [0, 1]$, $0 \leq \phi, \mu \leq 1$, entonces $g_{\epsilon}(O) > 0$, $g_{\epsilon}(I) < 0$ ($h_{\epsilon}(O) > 0$, $h_{\epsilon}(I) < 0$), con lo que está asegurada la existencia de una única raíz para g_{ϵ} y h_{ϵ} en $[0, 1]$, estas raíces serán respectivamente \bar{p}_{ϵ} y \underline{p}_{ϵ} . Es decir: la única raíz en $[0, 1]$ de $h_{\epsilon}(q)$ es \underline{p}_{ϵ} y la única raíz de $g_{\epsilon}(q)$ es \bar{p}_{ϵ} .

En esto consiste la técnica de linealización, en el uso de la función lineal en q , $c(q, (0, i_{\epsilon}))$ 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 $q \in R$, valores de $g(q)$ y $h(q)$ y que consiste en discretizar el espacio paramétrico, y pasar de la familia C 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) buscando la distribución discreta que optimice $E[c(q, (f, m))]$.

A continuación, dado $\epsilon \in [0, 1]$, y sustituyendo en (5) y (6), habremos calculado valores de $g_{\epsilon}(q)$, $h_{\epsilon}(q)$. 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). En este artículo presentamos algunos de los resultados obtenidos correspondientes al modelo Gamma - Uniforme (GU).

3 La hipótesis de independencia en el modelo GU.

Godfrey y Neter (1984) desarrollaron modelos para la búsqueda de cotas del error total cuando se utiliza el muestreo por unidades monetarias (MUS). Para ello consideran que z tiene distribución uniforme en $[0, 1]$, y para ϕ consideran que su distribución es Gamma, pero truncada en el intervalo $[0,1]$ (modelo Gamma - Uniforme, GU) o que sigue una distribución Beta (modelo Beta - Uniforme, BU). En concreto, el modelo GU, considera como verosimilitudes una distribución de Poisson para el número de errores en la muestra y una distribución exponencial truncada en 1 para su tamaño medio. La desventaja de estos modelos es que la distribución uniforme no es conjugada con la verosimilitud del problema, lo que lleva a la no obtención de una distribución conocida para $\Psi = f.m$, que Godfrey y Neter aproximan numéricamente. La consideración, en este trabajo, de los modelos como bidimensionales reduce la importancia de este inconveniente. En las tablas siguientes se presentan los resultados obtenidos para el caso en que la opinión del auditor sobre la tasa de error de la contabilidad en cuestión es que ésta es alta (en concreto, correspondientes a una distribución Gamma para ϕ con $E[\phi] = 0.25$ y $\sigma_\phi = 0.177$); para una gran variedad de resultados muestrales (correspondientes a una muestra de tamaño $n = 100$, en la que se encuentran $m = 1$ y 10 errores con tamaños $\bar{z} = 0.01, 0.05, 0.7, 1$) Para diferentes grados de contaminación ($\epsilon = 0.05, 0.1, 0.2, 0.25, 0.5, 1$) se presentan las cotas $\bar{r}_e, \underline{p}_e$ la medida de referencia bajo $p_0: E_{p_0}[f.m | \text{datos}]$ así como una medida tipificadora de la robustez dada por el factor de

sensibilidad relativa $RS = \frac{\bar{p}_e - \underline{p}_e}{2.E_{p_0}[f.m | \text{dato}]} \times 100$, útil para medir la cantidad de variación, en porcentaje, sobre Γ_E , de la cantidad $E_{p_0}[f.m | \text{datos}]$ (que representa la estimación de la cantidad de error bajo hipótesis de independencia). Los resultados se ofrecen multiplicados por un hipotético valor de libro $T_x = 10^6$ u.m.

Tabla 1: Modelo: GU Tasa de Error: Alta

ε	$m = 1, \bar{z} = 0.01$				ε	$m = 1, \bar{z} = 0.5$			
	$\overline{p_e}$	Media Post.	$\underline{p_e}$	RS		$\overline{p_e}$	Media Post.	$\underline{p_e}$	RS
0	7381.00	7381.00	7381.00	0.00	0	17112.00	17112.00	17112.00	0.00
0.05	9971.05	7381.00	4006.20	40.41	0.05	19470.79	17112.00	16519.98	8.62
0.1	12366.30	7381.00	2995.23	63.48	0.1	21512.41	17112.00	15950.40	16.25
0.2	16650.96	7381.00	2222.66	97.74	0.2	24859.81	17112.00	14860.30	29.22
0.25	18578.53	7381.00	2034.10	112.07	0.25	26251.79	17112.00	14343.60	34.79
0.5	26537.13	7381.00	1612.67	168.84	0.5	32155.04	17112.00	11804.78	59.46
1	37092.88	7381.00	1397.61	241.81	1	73394.11	17112.00	4272.94	201.97
$m = 1, \bar{z} = 0.7$					$m = 1, \bar{z} = 1$				
0	18273.00	18273.00	18273.00	0.00	0	19501.00	19501.00	19501.00	0.00
0.05	20686.53	18273.00	17808.18	7.88	0.05	22015.57	19501.00	19136.24	7.38
0.1	22729.30	18273.00	17346.75	14.73	0.1	24072.84	19501.00	18768.12	13.60
0.2	25988.96	18273.00	16449.31	26.10	0.2	27239.04	19501.00	18028.07	23.62
0.25	27314.57	18273.00	16449.31	29.73	0.25	28478.05	19601.00	17662.24	27.73
0.5	33875.85	18273.00	13720.92	55.15	0.5	35950.66	19501.00	15669.44	52.00
1	90012.65	18273.00	4275.23	234.60	1	98011.49	19501.00	0.00	251.30

La lectura de estos resultados evidencia una notable falta de robustez. Si se observa la columna correspondiente al factor RS, para casos con $m = 1$ error en la muestra, y dado un pequeño grado de desconfianza ($\varepsilon = 0.05, 0.1$), las fluctuaciones alrededor de la hipótesis de independencia superan el 40% (para $\bar{z} = 0.01$), estos resultados se moderan para valores mayores de \bar{z} , pero aumentan espectacularmente para $m = 10$ (obsérvese el caso $m = 10, \bar{z} = 0.01$), donde las fluctuaciones no bajan del 59%.

Esto nos permite afirmar que existen situaciones correspondientes a resultados muestrales plausibles en el escenario de población supuesto (y computacionalmente tratables) en los que

se aprecia una gran fluctuación de las conclusiones, es decir son situaciones de dependencia

Tabla 2: Modelo: GU Tasa de Error: Alta

ϵ	$m = 1, \bar{z} = 0.01$				ϵ	$m = 1, \bar{z} = 0.5$			
	$\overline{p_e}$	Media Post.	$\underline{p_e}$	RS		$\overline{p_e}$	Media Post.	$\underline{p_e}$	RS
0	1389.00	1389.00	1389.00	0.00	0	84972.00	84972.00	84972.00	0.00
0.05	102491.28	1389.00	976.75	3654.23	0.05	173002.34	84972.00	32966.04	82.40
0.1	102498.15	1389.00	976.75	3654.48	0.1	173194.03	84972.00	32880.21	82.56
0.2	102501.58	1389.00	976.75	3654.60	0.2	173254.11	84972.00	32837.10	82.63
0.25	102502.16	1389.00	976.75	3654.62	0.25	173278.71	84972.00	32828.71	82.64
0.5	102503.87	1389.00	976.75	36.54.68	0.5	173302.76	84972.00	32828.71	82.66
1	102504.44	1389.00	976.47	3654.71	1	173372.55	84972.00	32783.03	82.73
	$m = 1, \bar{z} = 0.7$					$m = 1, \bar{z} = 1$			
	$\overline{p_e}$	Media Post.	$\underline{p_e}$	RS		$\overline{p_e}$	Media Post.	$\underline{p_e}$	RS
0	93632.00	93632.00	93632.00	0.00	0	99372.00	99372.00	99372.00	0.00
0.05	209113.60	93632.00	33885.38	93.57	0.05	176963.71	99372.00	59639.17	59.03
0.1	211083.13	93632.00	33885.38	94.62	0.1	176963.71	99372.00	59639.17	59.03
0.2	211083.13	93632.00	33885.38	94.62	0.2	176963.71	99372.00	59639.17	59.03
0.25	211083.13	93632.00	33885.38	94.62	0.25	176963.71	99372.00	59639.17	59.03
0.5	211083.13	93632.00	33885.38	94.62	0.5	176963.71	99372.00	59639.17	59.03
1	215875.34	93632.00	28414.87	100.10	1	176963.76	99372.00	0.00	89.04

Entre ϕ y μ , lo que nos permite concluir que el modelo analizado es muy sensible a desviaciones de la hipótesis de independencia. En otras palabras, si no se está muy seguro de la independencia entre la tasa y la cantidad de error es extraordinariamente peligrosa su utilización.

Referencias Bibliográficas.

- [1] COX, D.R. y SNELL, E.J. (1979). On Sampling and the *Estimation of Rare Errors*. *Blometrika*; 66-1; 125-132.
- [2] FELIX, W.L. (Jr) y GRIMLUND, R.A. (1977). A Sampling *Model for Audit Tests of Composite Accounts*. *Journal of Accounting Research* (Spring); 23-40.

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

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

- [5] MARTEL ESCOBAR, M. (1996) *Aportaciones al Estudio de Técnicas Bayesianas en Auditoría*. Tesis Doctoral. Departamento de **Economía** Aplicada. Universidad de Las Palmas de G.C.

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

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