

COMPARACIÓN DE CURVAS DE SUPERVIVENCIA GAMMA ESTOCÁSTICAMENTE ORDENADAS

JOSÉ D. BERMÚDEZ EDO*
EDUARDO BEAMONTE CÓRDOBA**
Universitat de València

En este trabajo se propone un análisis de supervivencia basado en un modelo Gamma. Se obtienen las condiciones teóricas bajo las cuales dos funciones de supervivencia Gamma están estocásticamente ordenadas. Estos resultados se utilizan para proponer un método sencillo que permite comparar dos poblaciones cuando, a priori, se conoce que sus curvas de supervivencia están estocásticamente ordenadas. Los resultados se ejemplifican con el análisis de un banco de datos reales sobre tiempos de desempleo.

Comparison of stochastically ordered gamma survival curves

Palabras clave: supervivencia bayesiana, información inicial, métodos de Monte Carlo, predicción

Clasificación AMS (MSC 2000): 62E25, 62F15, 62N05

* Departamento de Estadística e Investigación Operativa. Universitat de València.

** Departamento de Economía Aplicada. Universitat de València.

Dirección para correspondencia: Departamento de Economía Aplicada. Bloque departamental oriental.
Av. Los Naranjos s/n. 46022 Valencia. Tfno.: 963828618/963828428. E-mail: beamonte@uv.es.

– Recibido en julio de 1999.

– Aceptado en abril de 2000.

1. INTRODUCCIÓN

En el análisis de datos de supervivencia provenientes de dos poblaciones es habitual que se conozca, a priori, que una de ellas debe presentar tiempos de vida estocásticamente más pequeños que los de la otra población (grupo control frente a grupo tratamiento). En este trabajo modelizamos el tiempo de vida de cada individuo mediante una distribución Gamma, cuyos parámetros —que dependen de ciertas covariables entre las que incluimos un indicador de la población de pertenencia— son modelizados como una realización de una superpoblación; así, nuestro modelo tiene en cuenta que no existen dos individuos iguales, pero todos ellos se «parecen». La información inicial de tiempos de vida estocásticamente ordenados la traducimos a restricciones sobre los hiperparámetros del modelo jerárquico, restricciones que incluimos en la distribución inicial por lo que también se cumplen en la distribución final.

La mayor parte de las investigaciones en este sentido se dirigen a la obtención de estimadores máximo verosímiles restringidos a la condición de orden para modelos no paramétricos. Utilizando la metodología bayesiana, un importante estudio es el llevado a cabo por Arjas y Gasbarra (1996) que también proporcionan una completa bibliografía. En este trabajo realizamos el análisis de la distribución final mediante una muestra obtenida a partir de la misma, utilizando métodos de Monte Carlo basados en cadenas de Markov.

2. EL MODELO Y LA INFORMACIÓN INICIAL

El modelo que se considera en este trabajo para el análisis de tiempos de supervivencia progresivamente censurados por la derecha es:

$$t \sim \text{Ga}(t|\alpha, \beta) \\ (\log \alpha, \log \beta)' \sim \text{N}_2((\log \alpha, \log \beta)'|Bx, H),$$

modelo jerárquico en el que cada individuo tiene un tiempo de supervivencia Gamma de parámetros α y β que dependen a su vez de los hiperparámetros B y H y del vector x de covariables del individuo. De este modo, podemos incluir modelos de poblaciones en los que a pesar de la ordenación estocástica de las curvas de supervivencia es posible el solapamiento de las mismas para individuos concretos. Este modelo supone una generalización de uno anterior (Bermúdez y Beamonte, 1993) para permitir la inclusión de covariables en el estudio y también está relacionado con un modelo semiparamétrico aditivo propuesto en Beamonte (1998).

Utilizamos las siguientes definiciones de orden estocástico para funciones de riesgo y de supervivencia genéricas (Arjas y Gasbarra, 1996).

Definición 1

Dos funciones de riesgo, $h_1(t)$ y $h_2(t)$, se dicen estocásticamente ordenadas si $h_1(t) < h_2(t) \forall t > 0$, ó $h_1(t) > h_2(t) \forall t > 0$.

Definición 2

Dos funciones de supervivencia, $S_1(t)$ y $S_2(t)$, se dicen estocásticamente ordenadas cuando $S_1(t) < S_2(t) \forall t > 0$, ó $S_1(t) > S_2(t), \forall t > 0$.

Dada la relación que liga a las funciones de supervivencia y riesgo:

$$S(t) = \exp \left\{ - \int_0^t h(x) dx \right\},$$

se tiene que si $h_1(t) < h_2(t) \forall t > 0$, entonces $S_1(t) > S_2(t) \forall t > 0$, no siendo cierto, en general, el recíproco. La equivalencia de ordenaciones estocásticas sí que se da entre funciones de supervivencia y funciones de riesgo acumulado, $H(t) = \int_0^t h(x) dx$.

Las funciones de riesgo y de supervivencia del modelo Gamma pueden ser expresadas en función de los parámetros α y β , de modo que las condiciones acerca de la ordenación estocástica de las mismas se traducen a restricciones sobre una columna de la matriz de hiperparámetros B. A continuación, establecemos dichos resultados cuyas demostraciones pueden ser consultadas en el apéndice.

Proposición 1

Si $(a, b)'$ es la columna de B asociada a la covariable que indica la población de pertenencia, entonces las respectivas funciones de supervivencia del modelo jerárquico Gamma están estocásticamente ordenadas si, y sólo si, $\text{sig}(a) \neq \text{sig}(b)$.

En el caso particular del modelo considerado en este trabajo resultan coincidentes las restricciones sobre los hiperparámetros en las ordenaciones de las funciones de riesgo y en las de las supervivencias.

Proposición 2

En el modelo jerárquico Gamma se tiene la equivalencia de ordenaciones estocásticas entre funciones de supervivencia y funciones de riesgo.

Para el análisis bayesiano del modelo utilizamos una distribución inicial sobre los hiperparámetros Normal-Wishart truncada para cumplir las condiciones de ordenación

estocástica, esto es, $a < 0$ y $b > 0$ (ó $a > 0$ y $b < 0$, según el orden), para $(a, b)'$ la columna de B asociada a la covariable que indica la población de pertenencia. De este modo, al cumplirse las restricciones de signo sobre las componentes de la correspondiente columna de hiperparámetros de la distribución inicial, también se cumplirán las mismas restricciones en la distribución final con probabilidad uno.

Obtenemos una muestra de la distribución final utilizando el muestreo de Gibbs (ver, por ejemplo, Casella, 1996 y referencias allí citadas). Para ello, es necesario realizar una adecuada descomposición del vector paramétrico completo de modo que sea sencillo simular a partir de las distribuciones condicionales completas. Para una consulta detallada del procedimiento ver Beamonte y Bermúdez (1995).

El cálculo de cualquier parámetro de interés —supervivencias medias o distribuciones predictivas, por ejemplo— lo aproximamos por Monte Carlo utilizando la muestra de la distribución final obtenida con anterioridad.

Un análisis alternativo, sin utilizar restricciones de orden en las funciones de riesgo o supervivencia, permite calcular, también por Monte Carlo, la probabilidad de que exista ordenación, lo que puede ser un indicador útil de la adecuación de los datos a esa afirmación inicial.

3. ANÁLISIS DE UNOS DATOS DE DESEMPLEO

Los datos analizados en este trabajo proceden de una encuesta dirigida a los licenciados en Ciencias Matemáticas por la Universitat de València. La labor de campo con los cuestionarios cumplimentados se cerró en octubre de 1994 (Encuesta sobre la valoración de la adecuación de los estudios a la actividad profesional. Convenio de la Universitat de València con la Conselleria d'Educació i Ciència).

El banco de datos finalmente considerado consta de 559 individuos, para todos y cada uno de los cuales aparece medida una covariable indicando su tiempo de supervivencia (meses transcurridos desde la obtención de la licenciatura hasta la consecución del primer empleo), una variable indicadora de la censura (1, si ha encontrado primer empleo y 0, en caso contrario) y cinco covariables: x_1 , constante e igual a uno, x_2 , indicador de la población (1, licenciados hasta 1991, inclusive y 0, resto), x_3 , sexo (1, si es hombre y 0, si es mujer), x_4 , nota media de la licenciatura (1, aprobado, 2, notable y 3, sobresaliente) y x_5 , actitud ante el hipotético hecho de volver a cursar la licenciatura (1, sí y 0, no). La covariable x_2 es una dicotomización de una covariable original igual al año de licenciatura. Realizamos la misma en ese sentido dada la información inicial disponible: en 1992 se produjo la congelación de plazas de oposición en enseñanza secundaria y siendo ésta la principal salida laboral de estos licenciados son de prever supervivencias mayores para el grupo en el que $x_2 = 0$.

Analizamos estos datos utilizando el modelo comentado en el apartado anterior, incorporando la información inicial acerca de la ordenación de las supervivencias en ambas poblaciones vía la correspondiente restricción paramétrica en la distribución inicial y obteniendo una muestra de la distribución final mediante el algoritmo de Gibbs.

Realizamos 200.000 pasos en la cadena de Markov obtenida a partir de la distribución final y monitorizamos la evolución de las medias (calculadas cada 50 pasos) de los hiperparámetros de la matriz B. Obtuvimos una rápida convergencia, además de muy poca variabilidad en todos ellos. En la figura 1 se muestra la evolución de las medias de las columnas de hiperparámetros $(b_{12}, b_{22})'$ y $(b_{15}, b_{25})'$.

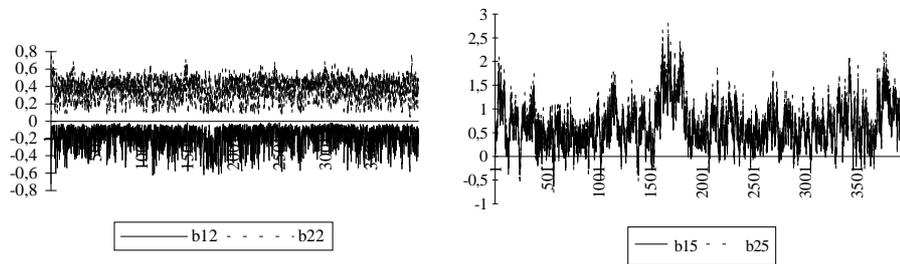


Figura 1. Evolución de las medias de los parámetros $(b_{12}, b_{22})'$ y $(b_{15}, b_{25})'$.

Obtuvimos una muestra de la distribución final desechando los 5.000 primeros pasos de la cadena de Markov, para así alcanzar la convergencia y, posteriormente, registrando uno de cada 50, a fin de tener independencia entre las observaciones muestrales, hasta un tamaño muestral igual a 500.

A continuación, realizamos una selección de variables influyentes en el modelo combinando la obtención de regiones de confianza de cada una de las columnas de hiperparámetros de B (Wei y Tanner, 1990), con el cálculo de las distancias de Mahalanobis de las muestras de cada una de ellas al vector origen.

En la figura 2 se representan las regiones de confianza 0.95 de algunas de dichas columnas y en la tabla 1 aparecen las comentadas distancias de Mahalanobis.

Tabla 1. Distancias de Mahalanobis del $(0, 0)'$ a las muestras de $(b_{1j}, b_{2j})'$.

	x_2	x_3	x_4	x_5
D^2	19.8631	1.0039	20.3579	2.2249

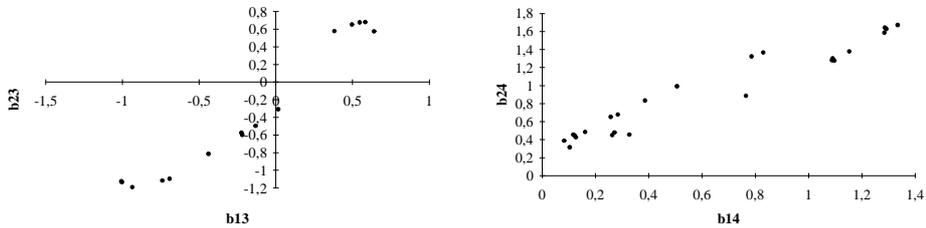


Figura 2. Región de confianza 0.95 de los hiperparámetros $(b_{13}, b_{23})'$ y $(b_{14}, b_{24})'$.

Las distancias de Mahalanobis correspondientes a las covariables x_3 y x_5 están asociadas a los cuantiles 0.6328 y 0.8907 de una distribución F con 2 y 498 grados de libertad. A su vez, las regiones de confianza 0.95 de los hiperparámetros correspondientes incluyen claramente al vector origen, por lo que ambas pueden ser eliminadas del estudio.

El cálculo de funciones de supervivencia predictivas lo realizamos considerando tan sólo las covariables x_1, x_2 y x_4 en el modelo.

En la figura 3 se comparan las supervivencias predictivas de ambas poblaciones para distintos valores de la covariable nota media.

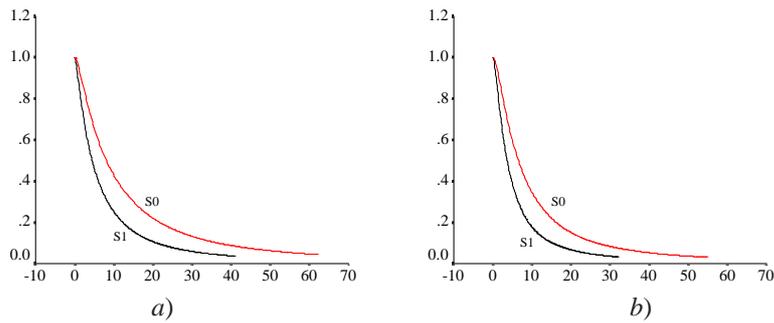


Figura 3. Funciones de supervivencia para a) $x_2 = 0, 1$ y $x_4 = 1$ y b) $x_2 = 0, 1$ y $x_4 = 2$.

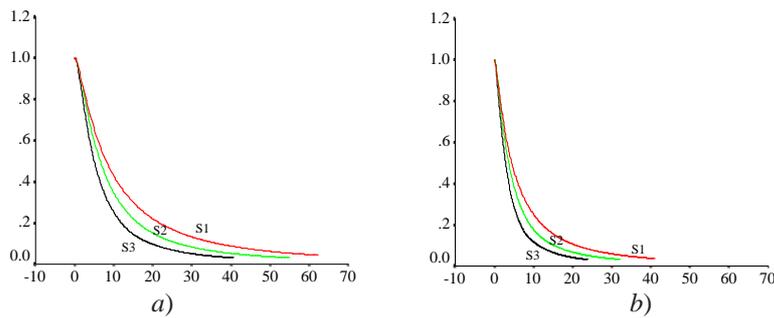


Figura 4. Funciones de supervivencia para a) $x_2 = 0$ y $x_4 = 1, 2, 3$ y b) $x_2 = 1$ y $x_4 = 1, 2, 3$.

La figura 4 recoge las diferencias existentes en las funciones de supervivencia según la nota media de la licenciatura para ambas poblaciones.

En la tabla 2 aparecen reflejadas la media y varianza de las densidades predictivas de nuevos individuos.

Tabla 2. Media y varianza de las densidades predictivas.

	$x_4 = 1$	$x_4 = 2$	$x_4 = 3$
$x_2 = 0$	$m = 16.9848$ $v = 1142.9119$	$m = 12.4307$ $v = 457.0186$	$m = 9.1554$ $v = 208.2967$
$x_2 = 1$	$m = 9.8468$ $v = 462.8604$	$m = 7.2001$ $v = 181.6949$	$m = 5.2976$ $v = 81.1224$

Observamos un decrecimiento en la media de la densidad predictiva conforme aumenta la nota media y, evidentemente, mayores medias en la población de licenciados después del año 1991. Por otra parte, en la figura 1 puede apreciarse claramente que $b_{12} < 0$ y $b_{22} > 0$, restricción paramétrica sobre la distribución final que indica la esperada ordenación de las supervivencias de las dos poblaciones.

4. CONSIDERACIONES FINALES

En primer lugar cabe destacar la aplicabilidad y adecuación de este modelo a diferentes estudios de supervivencia. También resalta la sencillez del método de análisis propuesto para la comparación de supervivencias estocásticamente ordenadas.

En el ejemplo analizado destaca la coherencia de los resultados obtenidos con los esperados. Parece lógico que las dos covariables finalmente consideradas sean las únicas influyentes en el modelo y que los individuos con mejor nota media tengan supervivencias menores.

Si se pretende comparar tres poblaciones —la generalización a cualquier número de poblaciones es inmediata—, pueden definirse dos variables indicadoras de la pertenencia a la población. Si, por ejemplo, codificamos con $(0, 0)$ a la población con mayor supervivencia, con $(0, 1)$ a la población con una supervivencia media y con $(1, 0)$ a la población con menor supervivencia, entonces las restricciones que deben cumplir las columnas de la matriz B asociadas a estas covariables, $(a, b)'$ y $(c, d)'$, son:

$$c < 0, \quad d > 0, \quad c - a > 0 \quad \text{y} \quad d - b < 0.$$

Obtenida una muestra aleatoria de la distribución final de B y H con inicial irrestricta, resulta sencillo transformarla en una muestra aleatoria proveniente de nuevas distribuciones finales utilizando diferentes iniciales. Para ello bastaría con utilizar el método de aceptación-rechazo con la primera distribución como función importante.

5. REFERENCIAS

- Arjas, E. y Gasbarra, D. (1996). «Bayesian inference of survival probabilities, under stochastic ordering constraints». *J. Amer. Statist. Assoc.*, 91, 1101-1109.
- Beamonte, E. y Bermúdez, J.D. (1995). «Comparación de curvas de supervivencia Gamma». *Qüestió*, 19, 171-186.
- Beamonte, E. (1998). *Aportaciones al análisis bayesiano semiparamétrico de datos de supervivencia*. Tesis doctoral. Universitat de València.
- Bermúdez, J.D. y Beamonte, E. (1993). «Análisis bayesiano de datos de supervivencia Gamma utilizando muestreo de Gibbs». *Estadística Española*, 35, 629-644.
- Casella, G. (1996). «Statistical inference and Monte Carlo algorithms (with comments)». *Test*, 5, 249-344.
- Wei, G.C.G. y Tanner, M.A. (1990). «Calculating the content and the boundary of the highest posterior density region via data augmentation». *Biometrika*, 77, 649-652.

APÉNDICE

Incluimos en este apéndice las demostraciones de las proposiciones del punto 2. Para ello, utilizamos unos resultados previos de modo que a partir de ellos ambas demostraciones resultan inmediatas.

Lema 1

Si $(a, b)'$ es la columna de B asociada a la covariable que indica la población de pertenencia, entonces las respectivas funciones de riesgo del modelo Gamma están estocásticamente ordenadas si, y sólo si, $\text{sig}(a) \neq \text{sig}(b)$.

Demostración

Dado $(\alpha_0, \beta_0) \in \mathbb{R}^+ \times \mathbb{R}^+$, definimos $R(\alpha_0, \beta_0) = \{(\alpha, \beta) \in \mathbb{R}^+ \times \mathbb{R}^+ \text{ tales que } h(t|\alpha_0, \beta_0) < h(t|\alpha, \beta) \forall t > 0\}$, donde $h(t|\alpha, \beta) = \frac{t^{\alpha-1} \exp(-\beta t)}{\int_t^\infty x^{\alpha-1} \exp(-\beta x) dx}$, es la función

de riesgo del modelo Gamma.

$$\begin{aligned} \log h(t|\alpha, \beta) &= (\alpha - 1) \log t - \beta t - \log \int_t^\infty x^{\alpha-1} \exp(-\beta x) dx \\ \frac{d}{d\alpha} \log h(t|\alpha, \beta) &= \log t - \frac{\int_t^\infty x^{\alpha-1} \log x \exp(-\beta x) dx}{\int_t^\infty x^{\alpha-1} \exp(-\beta x) dx} < 0, \end{aligned}$$

pues $\log x > \log t \forall x > t$, luego $h(t|\alpha, \beta)$ es estrictamente decreciente en $\alpha, \forall \beta > 0$ y $\forall t > 0$.

$$\frac{d}{d\beta} \log h(t|\alpha, \beta) = -t + \frac{\int_t^\infty x^{\alpha-1} x \exp(-\beta x) dx}{\int_t^\infty x^{\alpha-1} \exp(-\beta x) dx} > 0,$$

por lo tanto $h(t|\alpha, \beta)$ es estrictamente creciente en $\beta, \forall \alpha > 0$ y $\forall t > 0$.

Definimos $R_1 = \{(\alpha, \beta) \in \mathbb{R}^+ \times \mathbb{R}^+ \text{ tales que } \alpha < \alpha_0 \text{ y } \beta > \beta_0\}$ y $R_2 = \{(\alpha, \beta) \in \mathbb{R}^+ \times \mathbb{R}^+ \text{ tales que } \alpha > \alpha_0 \text{ y } \beta < \beta_0\}$.

$$(\alpha, \beta) \in R_1 \Rightarrow h(t|\alpha_0, \beta_0) < h(t|\alpha, \beta_0) < h(t|\alpha, \beta), \forall t > 0 \Rightarrow (\alpha, \beta) \in \mathbb{R}(\alpha_0, \beta_0)$$

$$(\alpha, \beta) \in R_2 \Rightarrow h(t|\alpha_0, \beta_0) > h(t|\alpha, \beta_0) > h(t|\alpha, \beta), \forall t > 0 \Rightarrow (\alpha, \beta) \in \mathbb{R}^C(\alpha_0, \beta_0).$$

Estudiamos ahora lo que ocurre fuera de esos dos rectángulos.

$$\lim_{t \rightarrow +\infty} h(t|\alpha, \beta) = \lim_{t \rightarrow +\infty} \frac{(\alpha - 1)t^{\alpha-2} \exp(-\beta t) - \beta t^{\alpha-1} \exp(-\beta t)}{-t^{\alpha-1} \exp(-\beta t)} = \beta, \forall \alpha > 0,$$

luego si $(\alpha, \beta) \in \mathbb{R}^+ \times \mathbb{R}^+$ con $\alpha < \alpha_0$ y $\beta < \beta_0$, $\lim_{t \rightarrow +\infty} h(t|\alpha, \beta) = \beta < \beta_0 = \lim_{t \rightarrow +\infty} h(t|\alpha_0, \beta_0)$ y, por continuidad, $\exists t^* > 0$ tal que $h(t|\alpha, \beta) < h(t|\alpha_0, \beta_0) \forall t > t^*$.

Por otra parte, utilizando el teorema del valor medio, $\exists(\alpha^*, \beta^*)$ perteneciente al segmento que une (α, β) con (α_0, β_0) , de modo que:

$$\begin{aligned} h(t|\alpha_0, \beta_0) - h(t|\alpha, \beta) &= (\alpha_0 - \alpha) \frac{d}{d\alpha} h(t|\alpha, \beta)|_{(\alpha^*, \beta^*)} + (\beta_0 - \beta) \frac{d}{d\beta} h(t|\alpha, \beta)|_{(\alpha^*, \beta^*)} = \\ &= h(t|\alpha^*, \beta^*) \left[(\alpha_0 - \alpha) \frac{d}{d\alpha} \log h(t|\alpha, \beta)|_{(\alpha^*, \beta^*)} + (\beta_0 - \beta) \frac{d}{d\beta} \log h(t|\alpha, \beta)|_{(\alpha^*, \beta^*)} \right]. \end{aligned}$$

Y teniendo en cuenta que:

$$\begin{aligned} \lim_{t \rightarrow 0} \frac{d}{d\alpha} \log h(t|\alpha, \beta) &= \lim_{t \rightarrow 0} \log t - \frac{\int_0^\infty x^{\alpha-1} \log x \exp(-\beta x) dx}{\int_0^\infty x^{\alpha-1} \exp(-\beta x) dx} = \\ &= \lim_{t \rightarrow 0} \log t - \Psi(\alpha) + \log \beta, \end{aligned}$$

donde $\Psi(x) = \frac{d}{dx} \log \Gamma(x)$ y $\Gamma(x) = \int_0^\infty t^{x-1} \exp(-t) dt$, son las funciones digamma y gamma de Euler, respectivamente, y que:

$$\lim_{t \rightarrow 0} \frac{d}{d\beta} \log h(t|\alpha, \beta) = \lim_{t \rightarrow 0} -t + \frac{\int_0^\infty x^{\alpha-1} x \exp(-\beta x) dx}{\int_0^\infty x^{\alpha-1} \exp(-\beta x) dx} = \frac{\alpha}{\beta},$$

se tiene que:

$$\begin{aligned} \lim_{t \rightarrow 0} [h(t|\alpha_0, \beta_0) - h(t|\alpha, \beta)] &= \\ &= \lim_{t \rightarrow 0} h(t|\alpha^*, \beta^*) \left[(\alpha_0 - \alpha) \frac{d}{d\alpha} \log h(t|\alpha, \beta)|_{(\alpha^*, \beta^*)} + (\beta_0 - \beta) \frac{\alpha^*}{\beta^*} \right] = \\ &= \lim_{t \rightarrow 0} t^{\alpha-1} \log t < 0, \end{aligned}$$

luego $\exists t^{**} > 0$ tal que $h(t|\alpha_0, \beta_0) < h(t|\alpha, \beta) \forall t < t^{**}$.

Por consiguiente, si $\alpha < \alpha_0$ y $\beta < \beta_0$, los riesgos $h(t|\alpha, \beta)$ y $h(t|\alpha_0, \beta_0)$ se cruzan y por lo tanto no están estocásticamente ordenados. Idéntica conclusión se tiene si $\alpha > \alpha_0$ y $\beta > \beta_0$, por lo que $R(\alpha_0, \beta_0) = R_1$.

Si la ordenación de las funciones de riesgo es la recíproca, razonando de un modo totalmente análogo concluiríamos que $R^*(\alpha_0, \beta_0) = \{(\alpha, \beta) \in \mathbb{R}^+ \times \mathbb{R}^+ \text{ tales que } h(t|\alpha_0, \beta_0) > h(t|\alpha, \beta) \forall t > 0\} = R_2$.

Consecuentemente, las restricciones para los hiperparámetros son, respectivamente, $a < 0, b > 0$ y $a > 0, b < 0$, o equivalentemente, $\text{sig}(a) \neq \text{sig}(b)$. □

Lema 2

Dadas dos funciones de densidad, $f_1(t)$ y $f_2(t)$, si ambas intersectan en un único punto, entonces las respectivas funciones de supervivencia, $S_1(t)$ y $S_2(t)$, están estocásticamente ordenadas. Si $f_1(t)$ y $f_2(t)$ intersectan en dos puntos, entonces $S_1(t)$ y $S_2(t)$ no están estocásticamente ordenadas.

Demostración

Sea t_0 el único punto de intersección de las funciones $f_1(t)$ y $f_2(t)$. Supongamos que $f_1(t) > f_2(t) \forall t < t_0$ y que $f_1(t) < f_2(t) \forall t > t_0$, entonces $S_1(t) = 1 - F_1(t) < S_2(t) = 1 - F_2(t) \forall t < t_0$ y $S_1(t) < S_2(t) \forall t > t_0$, luego $S_1(t)$ y $S_2(t)$ están estocásticamente ordenadas.

Si $f_1(t)$ y $f_2(t)$ intersectan en dos puntos, t_1 y t_2 , supongamos que $f_1(t) > f_2(t) \forall t < t_1$ y que $f_1(t) > f_2(t) \forall t > t_2$, entonces $S_1(t) < S_2(t) \forall t < t_1$ y $S_1(t) > S_2(t) \forall t > t_2$, luego $S_1(t)$ y $S_2(t)$ no están estocásticamente ordenadas. □

Lema 3

Si $f(t|\alpha_0, \beta_0)$ y $f(t|\alpha_1, \beta_1)$ son funciones de densidad Gamma distintas, entonces intersectan en un único punto o en dos puntos.

Demostración

Al menos ha de existir un punto de intersección de las dos densidades para que ambas integren uno.

Consideramos el cociente de las funciones de densidad Gamma:

$$\frac{f(t|\alpha_0, \beta_0)}{f(t|\alpha_1, \beta_1)} = \frac{\frac{\beta_0^{\alpha_0}}{\Gamma(\alpha_0)} t^{\alpha_0-1} \exp(-\beta_0 t)}{\frac{\beta_1^{\alpha_1}}{\Gamma(\alpha_1)} t^{\alpha_1-1} \exp(-\beta_1 t)} = C t^{\alpha_0-\alpha_1} \exp\{-(\beta_0 - \beta_1)t\},$$

$$\text{con } C = \frac{\beta_0^{\alpha_0} \Gamma(\alpha_1)}{\beta_1^{\alpha_1} \Gamma(\alpha_0)}.$$

El comportamiento de esta función en ambas colas es:

$$\lim_{t \rightarrow 0} \frac{f(t|\alpha_0, \beta_0)}{f(t|\alpha_1, \beta_1)} = C \lim_{t \rightarrow 0} t^{\alpha_0-\alpha_1} = \begin{cases} 0, & \text{si } \alpha_0 > \alpha_1 \\ \infty, & \text{si } \alpha_0 < \alpha_1 \end{cases}$$

$$\lim_{t \rightarrow \infty} \frac{f(t|\alpha_0, \beta_0)}{f(t|\alpha_1, \beta_1)} = C \lim_{t \rightarrow \infty} t^{\alpha_0-\alpha_1} \exp\{-(\beta_0 - \beta_1)t\} = \begin{cases} 0, & \text{si } \beta_0 > \beta_1 \\ \infty, & \text{si } \beta_0 < \beta_1 \end{cases}$$

De este modo se tiene que:

- (i) si $\alpha_0 > \alpha_1$ y $\beta_0 > \beta_1$, entonces $\frac{f(t|\alpha_0, \beta_0)}{f(t|\alpha_1, \beta_1)}$ tiene un único máximo en el punto $\frac{\alpha_0 - \alpha_1}{\beta_0 - \beta_1}$ y, por lo tanto, interseca a uno en dos puntos.
- (ii) si $\alpha_0 < \alpha_1$ y $\beta_0 < \beta_1$, entonces $\frac{f(t|\alpha_0, \beta_0)}{f(t|\alpha_1, \beta_1)}$ tiene un único mínimo en el punto $\frac{\alpha_0 - \alpha_1}{\beta_0 - \beta_1}$ e interseca a uno en dos puntos.
- (iii) si $\alpha_0 \geq \alpha_1$ y $\beta_0 \leq \beta_1$, entonces $\frac{f(t|\alpha_0, \beta_0)}{f(t|\alpha_1, \beta_1)}$ es estrictamente creciente, luego interseca a uno en un único punto.
- (iv) si $\alpha_0 \leq \alpha_1$ y $\beta_0 \geq \beta_1$, entonces $\frac{f(t|\alpha_0, \beta_0)}{f(t|\alpha_1, \beta_1)}$ es estrictamente decreciente y también interseca a uno en un único punto.

□

Proposición 1

Si $(a, b)'$ es la columna de B asociada a la covariable que indica la población de pertenencia, entonces las respectivas funciones de supervivencia del modelo jerárquico Gamma están estocásticamente ordenadas si, y sólo si, $\text{sig}(a) \neq \text{sig}(b)$.

Demostración

Por el lema 3, las respectivas funciones de densidad Gamma, $f(t|\alpha_0, \beta_0)$ y $f(t|\alpha_1, \beta_1)$, intersectan en uno o dos puntos. Si las supervivencias están estocásticamente ordenadas entonces, por el lema 2, las densidades intersectan en un único punto y se cumple, por el lema 3, que $\alpha_0 \geq \alpha_1$ y $\beta_0 \leq \beta_1$, ó $\alpha_0 \leq \alpha_1$ y $\beta_0 \geq \beta_1$, es decir $\text{sig}(a) \neq \text{sig}(b)$.

Recíprocamente, si $\text{sig}(a) \neq \text{sig}(b)$ se tiene, por el lema 1, la ordenación estocástica de las funciones de riesgo y ésta siempre implica la ordenación de las funciones de supervivencia.

□

Proposición 2

En el modelo jerárquico Gamma se tiene la equivalencia de ordenaciones estocásticas entre funciones de supervivencia y funciones de riesgo.

Demostración

Es inmediata a partir de la proposición anterior y el lema 1.

□

ENGLISH SUMMARY

COMPARISON OF STOCHASTICALLY ORDERED GAMMA SURVIVAL CURVES

JOSÉ D. BERMÚDEZ EDO*
EDUARDO BEAMONTE CÓRDOBA**
Universitat de València

In this paper a survival analysis using a Gamma model is proposed. The theoretical conditions for two survival Gamma functions to be stochastically ordered are obtained. These results are used to propose a simple method that allows the comparison of two populations when it is known that they are stochastically ordered. The paper concludes with the analysis of a unemployment data bank.

Keywords: bayesian survival, prior information, Monte Carlo method, prediction

AMS Classification (MSC 2000): 62E25, 62F15, 62N05

* Departamento de Estadística e Investigación Operativa. Universitat de València.

** Departamento de Economía Aplicada. Universitat de València.

Dirección para correspondencia: Departamento de Economía Aplicada. Bloque departamental oriental.
Av. Los Naranjos s/n. 46022 Valencia. Tfno.: 963828618/963828428. E-mail: beamonte@uv.es.

– Received July 1999.

– Accepted April 2000.

In this paper we propose a Gamma hierarchical model for the Bayesian analysis of right-censored survival data,

$$t \sim \text{Ga}(t|\alpha, \beta)$$

$$(\log \alpha, \log \beta)' \sim N_2(\log \alpha, \log \beta)'|B\mathbf{x}, H).$$

The survival time of each individual is modelled through a Gamma distribution with parameter α and β . These parameters are specific for each individual in the population, but related to the vector of covariates \mathbf{x} through a second level of the hierarchical model, allowing certain heterogeneity in the population.

In the survival data analysis of two populations is habitual to know that the waiting time of one of them are uniformly lower than those of the other one. This initial information of stochastically ordered survival times is translated to restrictions on the hyperparameters of the hierarchical model. The theoretical conditions for two survival Gamma functions to be stochastically ordered are obtained. These results are used to propose a simple method that allows for comparison of two populations when it is known that they are stochastically ordered.

Following the Bayesian paradigm, we propose a truncated Normal-Wishart prior distribution for the hyperparameters B and H that includes the stochastic order restrictions. In this way, these restrictions will be verified by the posterior distribution with probability one.

The posterior distribution becomes so complicated that no analytical study seems possible, although the prior has been chosen to be as simple as possible. This problem is common to almost every model in populations. However, a simulation study using Gibbs sampling or another Markov Chain Monte Carlo technique is possible. In fact, having in mind that any conditional distribution has a simpler form than the joint distribution and is easier to simulate from it, we propose the Gibbs algorithm to obtain a sample from the posterior distribution.

The paper concludes with the analysis of an unemployment data bank. A survey on the adequacy of undergraduate studies in Mathematics to professional activity was carried out in 1994, through an agreement between the Valencia Education and Science Council and Valencia University. One thousand questionnaires were sent to graduates in Mathematics from Valencia University during the years 1978 to 1993. Finally, 559 answers were considered.

One of the items of the surveys was the time, in months, from graduation until the first employment. That is the variable we are going to analyse in this paper, using a few more items from the survey as covariates, (possibly) related to the unemployment time: population indicator, sex, mean grade and attitude to the hypothetical possibility of whether they would have studied the same subject.

The population indicator is related to the year of graduation and we have considered two populations: graduates in and before 1991, and graduates in and after 1992. The year 1992 should be a change-point in unemployment rates of mathematicians in the state of Valencia because no new teaching position in public high schools was created since 1991 to the end of our study, and these teaching positions were the common labor exit for graduates in Mathematics.

We analyse the data using the hierarchical Gamma model proposed and we incorporate the initial information about the ordered survival times through the adequate restrictions on the hyperparameters of the prior distribution. We obtain the estimates of some characteristics of interest of the posterior distribution using a sample from it, obtained with a Gibbs simulation process.