ESTADÍSTICA ESPAÑOLA
Vol. 37, Núm. 140, 1995, págs. 473 a 495

Convergencia de estimadores de la tasa de azar acumulada basados en procesos βeta

por
ALFREDO MÉNDEZ y NIEVES GARCÍA

RESUMEN

El análisis de la función de distribución de una variable aleatoria de tiempo de fallo, con observaciones censuradas por la derecha, puede realizarse mediante procesos de recuento; y utilizando los procesos βeta, obtener estimadores Bayes no paramétricos de la tasa de azar acumulativa.

En este trabajo, mediante métodos martingala, probamos que el comportamiento asintótico de sucesiones de estimadores basados en procesos βeta converge a la verdadera tasa de azar acumulativa. También presentamos diversos procedimientos para construir intervalos y bandas de confianza de dicha tasa de azar.

Palabras clave: Datos censurados por la derecha, distribución asintótica, martingalas, procesos βeta, procesos de recuento, tasa de azar acumulativa.

Clasificación AMS: 62C10, 62G20

1. INTRODUCCIÓN

Para abordar el problema de la estimación de una función de supervivencia desde un punto de vista Bayesiano no paramétrico N.L.Hjort [10] utiliza como espacio paramétrico el conjunto de las funciones tasa de azar acumulativa y como inductores de medidas de probabilidad en dicho espacio, con la topología de los conjuntos de Borel cilíndricos, a los procesos βeta, demostrando que tienen un amplio soporte, que permiten interpretar fácilmente el conocimiento previo en términos matemáticos, que la distribución posterior dada una muestra posiblemente censurada por la derecha, también es βeta y, que las estimaciones dadas indican la repercusión de la información prior y de la proporcionada por la muestra. Posteriormente, utilizando el producto integral, se consiguen estimaciones de la función de supervivencia en base a las estimaciones de la tasa de azar acumulativa obtenidas.

En este artículo se prueban algunos resultados sobre los estimadores de la tasa de azar acumulativa basados en los procesos βeta y se comprueba que el comportamiento asintótico de dichos estimadores, bajo ciertas condiciones de regularidad, coincide con el comportamiento asintótico del estimador no paramétrico de Nelson-Aalen ([1], [2], [14] y [15]). Para establecer dicho resultado se utilizan técnicas de martingalas, esta elegante aproximación reemplaza los largos y complicados métodos algebraicos que se utilizaban para establecer propiedades de ciertos estadísticos. La mayor parte de los resultados teóricos sobre martingalas y martingalas locales que se utilizarán pueden encontrarse en Fleming y Harrington [7] y Andersen et al [3]. Estos trabajos realizan un estudio minucioso sobre el cálculo de martingalas asociadas a procesos de recuento y, en particular, analizan propiedades del estimador de Nelson-Aalen.

También se proponen algunos estimadores de la precisión para los estimadores propuestos basados en los procesos de variación predecible y de variación opcional. Finalizando con la construcción de intervalos y bandas de confianza para la función tasa de azar acumulativa.

2. MOMENTOS DEL ESTIMADOR DE LA TASA DE AZAR ACUMULATIVA

Consideraremos el modelo de datos censurados por la derecha en el cual $(T_i,U_i)_{i=1,\dots,n}$ es una m.a.s. proviniente de la v.a. bidimensional (T,U), siendo T una v.a. de tiempo de fallo y U una v.a. de censura, con S(t), F(t) y Λ(t) las funciones de supervivencia, distribución y tasa de azar acumulativa, respectivamente, asociadas a la v. a. T y donde las observaciones son $X_i = min(T_i,U_i)$ y $\delta_i = I(X_i = T_i)$, con i=1,... ,n. Para su estudio sean N(t), el nº de items que fallan verdaderamente hasta t, dN(t), el nº de items que fallan verdaderamente en t, $N^U(t)$, el nº de items que se censuran hasta t, Y(t), el nº de items que sobreviven hasta t, y el proceso estocástico M(t), definidos por:

$$N(t)=\sum_{i=1}^{n} N_i(t)=\sum_{i=1}^{n} I(X_i \le t, \delta_i = 1);\quad dN(t)=N(t)-N(t^-)=\sum_{i=1}^{n} N_i(t)-N_i(t^-);$$

$$Y(t)=\sum_{i=1}^{n} Y_i(t)=\sum_{i=1}^{n} I(X_i \ge t);$$

$$N^U(t)=\sum_{i=1}^{n} N_i^U(t)=\sum_{i=1}^{n} I(Xi \le t, \delta i=0);$$

$$M(t)=\sum_{i=1}^{n} M_i(t)=\sum_{i=1}^{n}(N_i(t)-\int_0^t Y_i(s)d\Lambda(s))=\sum_{i=1}^{n}(N_i(t)-A_i(t))=N(t)-\int_0^t Y(s)d\Lambda(s)$$

siendo $N_i(t), dN_i(t), N_i^U(t), Y_i(t)$, y $M_i(t)$ los correspondientes procesos individuales. Además, supondremos que las v.a. $T_i(t)$ y $U_i(t)$, i=1,... ,n, cumplen

$$\frac{dF(z)}{1-F(z^-)}=-\frac{dP(T_i \ge z, U_i \ge T_i)}{(T_i \ge z, U_i \ge z)} \quad \forall z \ge 0 \text{ tales que } P(T_i \ge z, U_i \ge z)>0 \qquad [1]$$

y que $\forall t \ge 0$ las v.a. $\{dN_i(t)/ F_{t^-}(t)\}_{i=1}^{n}$ son independientes y toman los valores {0,1}, donde F_{t^-} es la σ-álgebra historia de los procesos de recuento, y $F=\{ F_i(t)/ t>0\}_{i=1}^{n}$.

Los estimadores de la tasa de azar acumulativa que vamos a utilizar están basados en los procesos βeta (ver Hjort [10], Definición y Teorema 4.3).

Definición 2.1 Sea Λ_0 una tasa de azar acumulativa con un número finito de saltos en los puntos $t_1,\ldots,t_n$, y sea c(·) una función continua a trozos, no negativa y definida sobre [0, ∞). Diremos que un proceso de Levy, Λ, es un proceso βeta de parámetros c(·) y $\Lambda_0(\cdot)$, notándolo también como Λ ~ βeta (c, Λ_0), si el proceso Λ tiene una representación de Levy de la forma

$$\mathcal{E}[e^{-\theta\Lambda(t)}]=(\prod_{j/t_j \le t}\mathcal{E}[e^{-\theta s_j}])\exp(\int_0^\infty (1-e^{-\theta s})dL_t(s)), \qquad \text{para } t \ge 0,\ \theta \ge 0$$

siendo: $s_j=\Lambda\{t_j\} \sim \beta\text{eta}(c(t_j)\Lambda_0\{t_j\},\ c(t_j)(1-\Lambda_0\{t_j\}))$ y

$$dL_t(s) = (\int_0^t c(z)s^{-1}(1-s)^{c(z)-1}d\Lambda_{0,c}(z))ds, \qquad \forall t \geq 0 \quad y \quad 0 < s < 1,$$

en donde $\Lambda_{0,c}(t) = \Lambda_0(t) - \sum_{t_j \leq t} \Lambda_0\{t_j\}$, es el proceso Λ_0 sin los saltos dados.

Proposición 2.1. Sea Λ un proceso βeta con parámetros c(t) y, según la Def. 2.1, y sean las observaciones $(T_1,\delta_1),\dots,(T_n,\delta_n)$, posiblemente censuradas. Entonces, el estimador Bayes de Λ, bajo función de pérdida cuadrática, es

$$\hat{\Lambda}(t) = \int_0^t \frac{c(s)d\Lambda_0(s) + dN(s)}{c(s) + Y(s)} \qquad [2]$$

En primer lugar estableceremos que el proceso M(t) es una martingala, que resulta fundamental, ya que se puede utilizar como proceso integrador para procesos suficientemente regulares.

Proposición 2.2. Para el modelo de censura considerado obtenemos que el proceso definido por $M(t) = N(t) - \int_0^t Y(s)d\Lambda(s) = N(t) - A(t)$ es una martingala continua por la derecha con respecto a la filtración F y el proceso A(t), compensador de N(t) según la Descomposición de Doob-Meyer, es no decreciente, F_t -predecible y en el cero vale cero c.s. Además, el proceso M(t) es martingala local continua por la derecha y de cuadrado integrable.

Demostración. La demostración es consecuencia directa del Teorema 1.3.2 y del Teorema 2.3.1 de Fleming y Harrington [7].

La utilización de procesos que verifican cierta propiedad en sentido local se debe a la dificultad de establecer dichas propiedades en sentido global y a la sencillez de comprobarlas mediante sucesiones de localización.

Proposición 2.3. Sean c(t) una función real con valores en [0, ∞) continua por la izquierda con límites finitos por la derecha, y el proceso Y(t) de items en riesgo. Entonces, $W(t) = \dfrac{J(t)}{Y(t)}$,

$$G(t) = \frac{c(t)J(t)}{(c(t)+Y(t))Y(t)}, \quad y \quad L(t) = \frac{I(Y(t)>0)}{c(t)+Y(t)} = \frac{J(t)}{c(t)+Y(t)} = \begin{cases} \dfrac{1}{c(t)+Y(t)} & \text{si } Y(t) > 0 \\ & \text{si } Y(t) = 0 \end{cases}$$

son procesos acotados, continuos por la izquierda y F_t – predecibles.

Demostración. Por ser los procesos N(t), $N^U(t)$ e $Y(t) = n - (N(t^-) + N^U(t^-))$ continuos por la izquierda y F_t − predecibles, concluimos que W, G y L también son continuos por la izquierda y F_t − predecibles. Y por la propia definición de W, G y L concluimos que son acotados.

Proposición 2.4. Sea Λ la tasa de azar acumulativa de la v.a. T y supongamos que Λ es un proceso βeta de parámetros (c, Λ_0) según la Def. 2.1, siendo c(t) continua por la izquierda. Sea el estimador de Λ, dado en la Ec. (2) y el proceso definido por

$$\Lambda^* = \int_0^t \frac{c(s)d\Lambda_0(s) + Y(s)dN(s)}{c(s) + Y(s)} = \int_0^t \frac{c(s)d\Lambda_0(s) + Y(s)J(s)dN(s)}{c(s) + Y(s)}$$

Entonces, el proceso es una martingala respecto de la filtración F. Además, $\hat{\Lambda}(t) - \Lambda^*(t)$ es martingala local de cuadrado integrable.

Demostración. Puesto que $M(t) = N(t) - \int_0^t Y(s)d\Lambda(s) \Rightarrow Y(t)d\Lambda(t) = dN(t) - dM(t)$. Entonces, $\forall t \geq 0$, $\hat{\Lambda}(t) - \Lambda^*(t) = \int_0^t \frac{1}{c+Y}(dN - YJd\Lambda) = \int_0^t \frac{1-J}{c+Y}dN + \int_0^t \frac{J}{c+Y}dM = \int_0^t LdM$, donde la última igualdad es consecuencia de que $\forall t \geq 0$, con Y(t)=0 se cumple que dN(t)=0.

Aplicando el Teorema 2.4.5 de Fleming y Harrington [7] para martingalas locales de la forma $\int LdM$ se tiene que $\hat{\Lambda}(t) - \Lambda^*(t)$ es martingala respecto de la filtración F. Además, por la proposición anterior L(t) es F_t − predecible y localmente acotado, luego $\int LdM$ es martingala local de cuadrado integrable.

Por tanto, Λ^* es el compensador de $\hat{\Lambda}$.

Proposición 2.5. En las mismas hipótesis de la proposición anterior $\forall t \geq 0$ se verifica que:

i) $E[\hat{\Lambda}(t) - \Lambda^*(t)] = 0$

ii) $E[\hat{\Lambda}(t)] \leq \int_0^t E\left[\frac{c}{c+Y}\right]d\Lambda_0 + \int_0^t P(Y>0)d\Lambda = \int_0^t E\left[\frac{c}{c+Y}\right]d\Lambda_0 + \Lambda(t) - \int_0^t P(Y=0)d\Lambda$

iii) $E[\hat{\Lambda}(t)] = \Lambda(t) - \int_0^t E\left[\frac{c}{c+Y}\right]dK$, con $K(t)=\Lambda(t)-\Lambda_0(t)$

Demostración. El apartado i) es consecuencia de que $\hat{\Lambda} - \Lambda^*$ es martingala que vale cero c.s. en t=0.

$$E[\hat{\Lambda}(t)] = E[\Lambda^*(t)] = E\left[\int_0^t \frac{c}{c+Y} d\Lambda_0\right] + E\left[\int_0^t \frac{YJ}{Y} d\Lambda\right] \leq \int_0^t E\left[\frac{c}{c+Y}\right] d\Lambda_0 + \int_0^t E\left[\frac{YJ}{Y}\right] d\Lambda =$$

$$= \int_0^t E\left[\frac{c}{c+Y}\right] d\Lambda_0 + \int_0^t E[I(Y>0)]\, d\Lambda = \int_0^t E\left[\frac{c}{c+Y}\right] d\Lambda_0 + \int_0^t P(Y>0)\, d\Lambda =$$

$$= \int_0^t E\left[\frac{c}{c+Y}\right] d\Lambda_0 + \Lambda(t) - \int_0^t P(Y=0)\, d\Lambda$$, por lo que obtenemos (ii).

Por último, $\hat{\Lambda}(t) - \Lambda^*(t) = \int_0^t \frac{c\,d\Lambda_0 + dN - c\,d\Lambda - Y\,d\Lambda}{c+Y} = \int_0^t \frac{dN - Y\,d\Lambda - c(d\Lambda - d\Lambda_0)}{c+Y} =$

$= \int_0^t \frac{dM - c\,dK}{c+Y}$. Entonces, puesto que Λ(t) es la verdadera tasa de azar y tomando esperanzas

$E[\hat{\Lambda}(t)] = \Lambda(t) + E\left[\int_0^t \frac{dM - c\,dK}{c+Y}\right] = \Lambda(t) - \int_0^t E\left[\frac{c}{c+Y}\right] dK$, siendo $K(t)=\Lambda(t)-\Lambda_0(t)$, y donde la última igualdad se deduce del punto i) de esta proposición y del teorema de Fubini.

Observación 2.1. Para el análisis del momento de segundo orden del estimador $\hat{\Lambda}$ se consideran los procesos de variación predecible y de variación opcional. Como se menciona en la sección II.3 de Andersen et al [3], si M y M' son martingalas locales de cuadrado integrable existen dos procesos $\langle M, M \rangle$ y $\langle M, M' \rangle$ llamados de variación y de covariación predecibles, respectivamente, y que son c.s. los únicos procesos predecibles, de variación acotada y tal que $M^2 - \langle M, M \rangle$ y $M\,M' - \langle M, M' \rangle$ son martingalas locales que valen cero en el instante cero.

También se definen los procesos, [M, M] y [M, M'], de variación y covariación opcional, respectivamente. Estos procesos coinciden con los procesos $\langle M, M \rangle$ y $\langle M, M' \rangle$, respectivamente, en caso de que M y M' sean continuos, en caso contrario verifican que $\langle M, M' \rangle$ es el compensador de [M, M'] y, en particular, $\langle M, M \rangle$ es el compensador de [M, M].

Proposición 2.6. Sea $(N_1,\dots,N_k)$ un proceso de recuento multivariante y consideremos los estadísticos $\hat{\Lambda}_i$ y Λ^*_i, con i=1,...,k, definidos de forma análoga a los dados en la Prop. 2.1, entonces $\forall t \geq 0$ y $\forall i, j = 1,\dots,n$

$$\left\langle \hat{\Lambda}_i - \Lambda^*_i,\ \hat{\Lambda}_j - \Lambda^*_j \right\rangle(t) = \delta_{ij} \int_0^t \frac{Y_i(s)}{(c(s)+Y_i(s))^2}(1-\Delta\Lambda(s))\,d\Lambda(s)$$

$$\left\langle \hat{\Lambda} - \Lambda^*,\ \hat{\Lambda} - \Lambda^* \right\rangle(t) = \int_0^t \frac{Y(s)}{(c(s)+Y(s))^2}(1-\Delta\Lambda(s))\,d\Lambda(s)$$

Demostración. Sean $L_i = \dfrac{J_i}{c+Y_i}$, con i=1,...,k. Entonces, por la Prop. 2.3, L_i está acotado, por lo que $\forall t \geq 0$ existirá $\Gamma_i > 0$ tal que $|L_i(t)| < \Gamma_i$, y $E\left|\int_0^t L_i^2\, d\langle M_i, M_i\rangle\right| \leq \Gamma_i^2\, E\left|\langle M_i, M_i\rangle(t)\right|$.

Por ser M_i martingala local de cuadrado integrable, utilizando el Teorema 2.6.1 de Fleming y Harrington [7] y por suponer $\Lambda(t) < \infty$ $\forall t \geq 0$ c.s., podemos concluir que

$$E\left[\int_0^t L_i^2\, d\langle M_i, M_i\rangle\right] \leq \Gamma_i^2\, E\left[\int_0^t (1-\Delta A_i)\,dA_i\right] \leq \Gamma_i^2 \int_0^t E[1-\Delta A_i]\ dA_i \leq \Gamma_i^2 A_i(t) < \infty$$

por el Lema 2.6.1 de Fleming y Harrington [7] y por la Prop. II.4.1 de Andersen et al [3]

i) $\left\langle \hat{\Lambda}_i - \Lambda^*_i,\ \hat{\Lambda}_j - \Lambda^*_j \right\rangle(t) = \left\langle \int_0^{(\cdot)} L_i\, dM_i, L_j\, dM_j \right\rangle(t) = \int_0^t L_i L_j\, d\langle M_i, M_j\rangle = 0$ c.s. si $i \neq j$

ii) $\left\langle \hat{\Lambda}_i - \Lambda^*_i,\ \hat{\Lambda}_i - \Lambda^*_i \right\rangle(t) = \int_0^t L_i^2\, d\langle M_i, M_i\rangle = \int_0^t L_i^2\, Y_i(1-\Delta\Lambda)\,d\Lambda = \int_0^t \frac{Y_i(1-\Delta\Lambda)}{(c+Y_i)^2}\, d\Lambda$

Por otra parte, $M^2 = \left(\sum_{i=1}^k M_i\right)^2 = \sum_{i=1}^k M_i^2 + 2\sum_{i<j}^k M_i M_j$, y por $M_i M_j = 0$ *c.s. si* $i \neq j$, el proceso compensador de M^2 será $\sum_{i=1}^k \langle M_i, M_i\rangle = \int \sum_{i=1}^k Y_i(1-\Delta\Lambda)\,d\Lambda = \int Y(1-\Delta\Lambda)\,d\Lambda$,

por tanto, $\left\langle \hat{\Lambda} - \Lambda^*,\ \hat{\Lambda} - \Lambda^* \right\rangle(t) = \int_0^t \frac{Y}{(c+Y)^2}(1-\Delta\Lambda)\,d\Lambda.$

Observación 2.2. A partir de los resultados obtenidos en las dos últimas proposiciones podemos concluir que el sesgo del estimador $\hat{\Lambda}$ está acotado por una cantidad que, en el caso de que c(t) sea pequeño y P(Y(s)=0), con s en (0,t], próximo a cero, tiene poca importancia. El sesgo también tiene poca importancia si la tasa de azar acumulativa verdadera, Λ, está suficientemente aproximada por la tasa de azar acumulativa elegida como prior, Λ_0.

La primera condición de la Prop. 2.5 se sigue verificando al sustituir t por τ, siendo τ un tiempo de parada (es una aplicación directa del Teorema de Parada Opcional).

Por tener los procesos de recuento saltos de tamaño uno se deduce que las martingalas $\{M_i\}_{i=1}^{k}$ son ortogonales, por lo que las martingalas $\{\hat{\Lambda}_i - \Lambda_i^*\}_{i=1}^{k}$ son también ortogonales, y el compensador de $(\hat{\Lambda} - \Lambda^*)^2$ es la suma de los compensadores $\sum_{i=1}^{k}\langle \hat{\Lambda}_i - \Lambda_i^*, \hat{\Lambda}_i - \Lambda_i^* \rangle$.

Para el estudio de la precisión vamos a considerar el error cuadrático medio de $\hat{\Lambda}$ propuesto por Aalen ([1], [2]).

Definición 2.2. Sean $(N_i, \ldots, N_k)$ un proceso de recuento multivariante, y los estadísticos $\hat{\Lambda}$ y Λ^* considerados anteriormente. Llamaremos error cuadrático medio de $\hat{\Lambda}$, y lo denotaremos por $\tilde{\sigma}^2$, al estadístico definido por $\tilde{\sigma}^2(t) = E[(\hat{\Lambda}(t) - \Lambda^*(t))^2]$.

Observación 2.3. Para la estimación de dicha cantidad vamos a proponer dos estimadores $\hat{\sigma}^2$ y $\bar{\sigma}^2$. El primero de ellos es el proceso de variación predecible asociado a la martingala local $(\hat{\Lambda} - \Lambda^*)$, es decir, es $\langle \hat{\Lambda} - \Lambda^*, \hat{\Lambda} - \Lambda^* \rangle$, mientras que el segundo es el proceso de variación opcional $[\hat{\Lambda} - \Lambda^*, \hat{\Lambda} - \Lambda^*]$.

Proposición 2.7. Con los mismos elementos de la definición anterior sean, $\forall t \geq 0$,

$$\hat{\sigma}^2(t) = \langle \hat{\Lambda} - \Lambda^*, \hat{\Lambda} - \Lambda^* \rangle(t) \quad \text{y} \quad \bar{\sigma}^2(t) = [\hat{\Lambda} - \Lambda^*, \hat{\Lambda} - \Lambda^*](t)$$

Entonces, $\hat{\sigma}^2(t)$ y $\bar{\sigma}^2(t)$ son estimadores insesgados para $\tilde{\sigma}^2(t)$, y se pueden expresar como

$$\hat{\sigma}^2(t) = \int_0^t \frac{Y}{(c+Y)^2}(1 - d\Lambda)\,d\Lambda \quad \text{y} \quad \bar{\sigma}^2(t) = \int_0^t \frac{J(1-2Y\Delta\Lambda)}{(c+Y)^2}\,dN + \int_0^t \frac{Y^2\Delta\Lambda}{(c+Y)^2}\,d\Lambda$$

Demostración. Por ser $(\hat{\Lambda}-\Lambda^*)^2-\left\langle\hat{\Lambda}-\Lambda^*,\hat{\Lambda}-\Lambda^*\right\rangle$ una martingala de media 0, el estadístico $\hat{\sigma}^2$$ es insesgado para $\tilde{\sigma}^2$.

Puesto que el proceso $\left\langle\hat{\Lambda}-\Lambda^*,\hat{\Lambda}-\Lambda^*\right\rangle$ es el compensador de $[\hat{\Lambda}-\Lambda^*,\hat{\Lambda}-\Lambda^*]$, su diferencia es una martingala con valor cero c.s. en t=0, y podemos concluir que $\bar{\sigma}^2$ es insesgado para $\tilde{\sigma}^2$, ya que $E[\tilde{\sigma}-\bar{\sigma}]=E[\tilde{\sigma}-\hat{\sigma}]+E[\hat{\sigma}-\bar{\sigma}]=0$ c.s.

La expresión de $\hat{\sigma}^2$ es un resultado obtenido en la proposición previa. Por la propia definición del proceso de variación opcional y aplicando la fórmula de integración por partes para la integral de Lebesgue-Stieljes se obtiene que, si M es la martingala considerada,

$$[M,M](\mathrm{t})=\int_0^{\mathrm{t}}\Delta\mathrm{M}\,\mathrm{dM}=\sum_{s\le t}\Delta\mathrm{N}(\Delta\mathrm{N}-\Delta\mathrm{A})-\sum_{s\le t}\Delta A(\Delta N-\Delta A)=$$

$$=\sum_{s\le t}(\Delta\mathrm{N})^2-2\sum_{s\le t}\Delta\mathrm{A}\Delta\mathrm{N}+\sum_{s\le t}(\Delta\mathrm{A})^2=\mathrm{N(t)}-2\int_0^{\mathrm{t}}\mathrm{Y}\,\Delta\Lambda\,\mathrm{dN}+\int_0^{\mathrm{t}}\mathrm{Y}^2\,\Delta\Lambda\,\mathrm{d}\Lambda.$$

Por la Prop. II.4.1 de Andersen et al [3] y teniendo en cuenta la expresión de $d[M,M](\mathrm{t})$,

$$[\hat{\Lambda}-\Lambda^*,\hat{\Lambda}-\Lambda^*](\mathrm{t})=\int_0^{\mathrm{t}}\mathrm{L}^2\,d[M,M]=\int_0^{\mathrm{t}}\mathrm{L}^2\,(dN-2Y\Delta\Lambda\,dN+Y^2\Delta\Lambda)d\Lambda=$$

$$=\int_0^{\mathrm{t}}\mathrm{L}^2\,(1-2Y\Delta\Lambda)dN+\int_0^{\mathrm{t}}\mathrm{L}^2\,Y^2\Delta\Lambda\,d\Lambda=\int_0^t\frac{J(1-2Y\Delta\Lambda)}{(c+Y)^2}dN+\int_0^{\mathrm{t}}\frac{Y^2\Delta\Lambda}{(c+Y)^2}d\Lambda$$

con lo que se tiene probada la proposición.

Observación 2.4. Para la aproximación de los estimadores $\hat{\sigma}^2$ y $\bar{\sigma}^2$ se puede recurrir a los estadísticos correspondientes que se obtienen al sustituir la tasa de azar por su estimación, es decir, las cantidades:

$$\hat{\sigma}^{*2}(t)=\int_0^t\frac{JY(1-\Delta\hat{\Lambda})}{(c+Y)^2}d\hat{\Lambda}=\int_0^t\frac{JY(c+Y-1)}{(c+Y)^4}dN+\int_0^t\frac{JYc(c+Y-c\Delta\Lambda_0-2\Delta N)}{(c+Y)^4}d\Lambda_0$$

$$\bar{\sigma}^{*2}(t)=\int_0^t\frac{J(1-2Y\Delta\hat{\Lambda})}{(c+Y)^2}dN+\int_0^t\frac{Y^2\Delta\hat{\Lambda}}{(c+Y)^2}d\hat{\Lambda}=\int_0^t\frac{Jc^2(1-2Y\Delta\Lambda_0)}{(c+Y)^4}dN+\int_0^t\frac{Y^2c^2\Delta\Lambda_0}{(c+Y)^4}d\Lambda_0$$

son aproximaciones de $\hat{\sigma}^2$ y $\bar{\sigma}^2$, respectivamente.

Si consideramos el caso en que Λ es continua, $\forall t \geq 0$ $d\langle M,M\rangle(t) = dN(t) = d[M,M](t)$, por lo que coinciden las dos estimaciones de $\tilde{\sigma}^2$ y, además, dichos estadísticos son cantidades muestrales, obteniéndose que $\hat{\sigma}^2 = \bar{\sigma}^2 = \int_0^t \frac{J}{(c+Y)^2} dN = \bar{\sigma}^{*2}(t) = \hat{\sigma}^{*2}(t)$ es una cantidad que nos sirve directamente para la estimación.

La razón para preferir $\hat{\sigma}^2$ frente a $\bar{\sigma}^2$, como estimador de $\tilde{\sigma}^2$, es que resulta ser un medio más directo para calcular una aproximación de dicho estadístico, y en el caso particular de que Λ sea continua ambos estadísticos coinciden.

Estos estimadores siguen siendo válidos en el caso de que el proceso N sea un proceso de recuento multivariante, ya que en ese caso se sigue cumpliendo que el proceso $M = \sum_{i=1}^{k} M_i = = \sum_{i=1}^{k} N_i - A_i = \sum_{i=1}^{k} N_i - \int_0^{(\cdot)} Y_i \, d\Lambda$ verifica que es martingala y localmente de cuadrado integrable, que el compensador de $M^2(t)$) es $\langle M,M\rangle(t) = \int_0^t Y(1-\Delta\Lambda)\,d\Lambda$ y que $\hat{\Lambda}(t) - \Lambda^*(t)$ es martingala local de cuadrado integrable, con $\left\langle \hat{\Lambda} - \Lambda^*, \hat{\Lambda} - \Lambda^* \right\rangle(t) = \int_0^t \frac{Y}{(c+Y)^2}(1-\Delta\Lambda)\,d\Lambda$.

Otro posible estimador de la varianza de $\hat{\Lambda}$ se puede obtener en base a suponer que $\Lambda/datos$ es un proceso βeta, en cuyo caso podríamos considerar como estimador de la varianza, la varianza a posteriori, es decir,

$$\bar{\sigma}^2(t) = \int_0^t \frac{1-\Delta\hat{\Lambda}}{c+Y+1} d\hat{\Lambda}.$$

3. DISTRIBUCIÓN ASINTÓTICA

Para el análisis de propiedades asintóticas de los estimadores de la tasa de azar acumulativa basada en procesos βeta se incluye una notación en la que los superíndices representan un proceso individual. Más específicamente, sea una sucesión de procesos de recuento, $N^{(n)} = (N_1^{(n)}, \ldots, N_k^{(n)})$, con n=1,2,..., cada uno de ellos con tasa de azar acumulativa $(\Lambda_1, \ldots, \Lambda_k)$, independiente de n, y para cada i=1,...,k, Λ_i es un proceso βeta de parámetros (c_i, Λ_{0i}). Consideramos también los procesos $J_i^{(n)}(t) = I(Y_i^{(n)}(t) > 0)$; i=1,...,k; n=1,2,..., y supongamos que $\frac{J_i^{(n)}}{c_i + Y_i^{(n)}}$ es localmente acotado para cada i y para cualquier n, siendo $c_i(\cdot)$ las funciones que forman parte de los parámetros de los procesos βbeta, continuas por la izquierda y con límites por la derecha.

Comprobaremos que el comportamiento asintótico del estimador $\hat{\Lambda}$ coincide con el comportamiento asintótico del estimador de Nelson-Aalen, lo cual resulta razonable en el sentido de que si se dispone de gran cantidad de información muestral, el conocimiento a priori debe tener, en comparación, poco peso a la hora de realizar estimaciones, es decir, las estimaciones se realizan fundamentalmente con la información muestral, igual que el estimador de Nelson-Aalen. Aplicando que el estimador de Nelson-Aalen converge en probabilidad a la verdadera tasa de azar acumulativa, podemos concluir que el estimador basado en los procesos βeta también converge a la verdadera tasa de azar acumulativa.

Teorema 3.1 Sea $(N^{(n)})_{n\in N}$ una sucesión de procesos de recuento, y para cada n, con $n \in \mathbf{N}$, sea $M^{(n)} = N^{(n)} - \int Y^{(n)}\, d\Lambda$ la martingala local de cuadrado integrable asociada y los demás elementos relativos a dichos proceso de recuento y considerados anteriormente. Sean, los estimadores de Nelson-Aalen, $\forall t \geq 0$ $\overline{\Lambda}^{(n)}(t) = \int_0^t \frac{dN^{(n)}}{Y^{(n)}}$, y para cada t>0 fijo supongamos que, cuando $n \longrightarrow \infty$, se verifican las siguientes hipótesis:

$$\int_0^t \frac{c(s)\, J^{(n)}(s)}{c(s)+Y(s)}(d\Lambda_0(s) - d\Lambda(s)) \xrightarrow{P} 0 \qquad [3]$$

$$\int_0^t \frac{J^{(n)}(s)}{Y^{(n)}(s)} d\Lambda(s) \xrightarrow{P} 0 \qquad [4]$$

$$\int_0^t (1 - J^{(n)}(s))\, d\Lambda(s) \xrightarrow{P} 0 \qquad [5]$$

Entonces, cuando $n \longrightarrow \infty$, $\sup_{s\in[0,t]} \left|\hat{\Lambda}^{(n)}(s) - \Lambda(s)\right| \xrightarrow{P} 0.$

Demostración. Para cualquier t>0 fijo se tiene

$$\left|\hat{\Lambda}^{(n)}(t) - \Lambda(t)\right| \leq \left|\hat{\Lambda}^{(n)}(t) - \overline{\Lambda}^{(n)}(t)\right| + \left|\overline{\Lambda}^{(n)}(t) - \Lambda(t)\right| \qquad [6]$$

Obsérvese que $\hat{\Lambda}(t)-\overline{\Lambda}(t)=\int_0^t \frac{c\,d\Lambda_0+dN}{c+Y}-\int_0^t \frac{dN}{Y}=\int_0^t \frac{c(Yd\Lambda_0-dN)}{(c+Y)Y}=$

$$=\int_0^t \frac{c(Yd\Lambda_0-Yd\Lambda)}{(c+Y)Y}-\int_0^t \frac{c(dN-Yd\Lambda)}{(c+Y)Y}=\int_0^t \frac{cJ(d\Lambda_0-d\Lambda)}{c+Y}-\int_0^t \frac{c\,J}{(c+Y)Y}dM \Rightarrow$$

$$\Rightarrow \hat{\Lambda}^{(n)}(t)-\overline{\Lambda}^{(n)}(t)=\int_0^t \frac{c\,J^{(n)}}{c+Y^{(n)}}(d\Lambda_0-d\Lambda)-\int_0^t \frac{c\,J^{(n)}}{(c+Y^{(n)})\,Y^{(n)}}dM^{(n)}$$

Por la condición (3) se cumple que, cuando $n\to\infty$, el primer sumando del miembro de la derecha de la igualdad anterior converge en probabilidad a cero.

Por la Prop. 2.3, $H^{(n)}(t)=\dfrac{c(t)\,J^{(n)}(t)}{(c(t)+Y^{(n)}(t))\,Y^{(n)}(t)}$ es acotado, continuo por la izquierda y predecible entonces, por el Teorema 2.4.1 de Fleming y Harrington [7], $\int_0^t \mathrm{H}^{(n)}(u)\,dM^{(n)}(u)$ es martingala local de cuadrado integrable, y por una versión de la desigualdad de Lenglart

$$P\left(\sup_{s\in[0,t]}\left|\int_0^s \mathrm{H}^{(n)}(u)\,dM^{(n)}(u)\right|>\eta\right)\leq \frac{\delta}{\eta^2}+P\left(\left\langle\int_0^{(\cdot)}\mathrm{H}^{(n)}\,dM^{(n)},\ \int_0^{(\cdot)}\mathrm{H}^{(n)}\,dM^{(n)}\right\rangle(t)>\delta\right)\quad [7]$$

por la Prop. II.4.1 [3], y que $d\left\langle M^{(n)},M^{(n)}\right\rangle=Y^{(n)}(1-\Delta\Lambda)\,d\Lambda$, ver Prop. 2.6, se obtiene

$$\left\langle\int_0^{(\cdot)}\mathrm{H}^{(n)}\,dM^{(n)},\ \int_0^{(\cdot)}\mathrm{H}^{(n)}\,dM^{(n)}\right\rangle(t)=\int_0^t \mathrm{H}^{(n)2}\,d\left\langle M^{(n)},M^{(n)}\right\rangle=$$

$$=\int_0^t \frac{c^2\,J^{(n)}}{(c+Y^{(n)})^2\,Y^{(n)2}}\,Y^{(n)}\,(1-\Delta\Lambda)\,d\Lambda=\int_0^t \frac{c^2\,J^{(n)}}{(c+Y^{(n)})^2\,Y^{(n)}}(1-\Delta\Lambda)\,d\Lambda$$

Por cumplirse que $Y(u)\,\Lambda(u)\leq 1\ \ \forall u\geq 0$, resultado cuya demostración puede encontrase en Meyer [12] y Gill [8], y puesto que "si $Y^{(n)}(s)\neq 1$, entonces $Y^{(n)}(s)\geq 1$, tenemos que $1-\Delta\Lambda(u)\leq 1$, y por tanto

$$\left\langle\int_0^{(\cdot)}\mathrm{H}^{(n)}\,dM^{(n)},\ \int_0^{(\cdot)}\mathrm{H}^{(n)}\,dM^{(n)}\right\rangle(t)\leq \int_0^t \frac{\mathrm{c}^2\,J^{(n)}}{\mathrm{c}^2\,Y^{(n)}}\mathrm{d}\Lambda=\int_0^t \frac{J^{(n)}}{Y^{(n)}}\mathrm{d}\Lambda.$$

Por lo que la desigualdad (7) queda

$P\left(\sup_{s\in[0,t]} \left| \int_0^s H^{(n)}(u)\, dM^{(n)}(u) \right| > \eta \right) \leq \frac{\delta}{\eta^2} + P\left(\int_0^t \frac{J^{(n)}}{Y^{(n)}} d\Lambda > \delta \right)$ y por la condición (4), cuando $n \to \infty$, $\sup_{s\in[0,t]} \left| \int_0^s H^{(n)}(u)\, dM^{(n)}(u) \right| \xrightarrow{P} 0$. Luego, queda probado que, cuando $n \to \infty$,

$$\left| \hat{\Lambda}^{(n)}(t) - \overline{\Lambda}^{(n)}(t) \right| \xrightarrow{P} 0$$

Dado que el estimador Nelson-Aalen converge a la verdadera tasa de azar acumulativa en la norma del supremo sobre intervalos finitos (ver Teorema IV.1.1 de Andersen et al [3], obtenemos que, cuando $n \to \infty$, $\left| \overline{\Lambda}^{(n)}(s) - \Lambda(s) \right| \xrightarrow{P} 0 \quad \forall s \geq 0$.

Combinando los resultados sobre convergencia en la desigualdad (6) queda demostrado el teorema.

Damos ahora una condición suficiente que debe cumplir el proceso Y(t) para que resulten ser ciertas las hipótesis (3), (4) y (5). Para la demostración de dicho resultado incluimos un proposición de tipo técnico, que es una consecuencia de un teorema debido a Gill [9].

Proposición 3.1. Supongamos que Λ es la función tasa de azar acumulativa de la v.a. T, y dicha variable verifica P(T ≥ t) ≥ m > 0, con m∈ **R** y t fijo. Sea una sucesión de procesos $\left\{ X^{(n)} \right\}_{n\in N}$, tal que $\forall t > 0$ fijo y para casi todo s ∈ [0,t], $X^{(n)}(s) \xrightarrow{P} f(s)$, cuando $n \to \infty$, siendo $f(\cdot)$ una función determinística que cumple las siguientes hipótesis:

$$\int_0^t \left| f(s) \right| d\Lambda(s) < \infty \tag{8}$$

y $\forall \delta > 0\ \exists K_\delta$ función no negativa tal que $\int_0^t K_\delta \; d\Lambda(s) < \infty$ [9]

$$\lim_{n\to\infty} \inf P(\left| X^{(n)}(s) \right| \leq K_\delta(s),\ \forall s \in [0,t]) \geq 1 - \delta \tag{10}$$

Entonces, cuando $n \to \infty$, $\sup_{s\in[0,t]} \left| \int_0^t X^{(n)}(u)\, d\Lambda(u) - \int_0^t f(u)\, d\Lambda(u) \right| \xrightarrow{P} 0.$

Demostración. Sin pérdida de generalidad podemos suponer que K_δ es función no creciente en δ, y que para cualquier δ>0, K_δ está acotada c.s. por una constante, Γ_δ, en el intervalo [0, t]. Para s fijo sean:

$$X_\delta^{(n)} = (sig\, X^{(n)})\; min\{|X^{(n)}|, K_\delta\}, \;\; \forall\delta > 0 \;\text{ y }\; \forall n \in \mathbf{N}$$

$$f_\delta = (sig\, f)\; min\{|f|, \mathrm{K}_\delta\}, \;\; \forall\delta > 0$$

Por lo que podemos realizar la siguiente descomposición, $\forall 0 < s \leq t$,

$$\left|\int_0^s \mathrm{X}^{(n)}\, d\Lambda - \int_0^s \mathrm{f}\, d\Lambda\right| \leq \left|\int_0^s \mathrm{X}^{(n)}\, d\Lambda - \int_0^s \mathrm{X}_\delta^{(n)}\, d\Lambda\right| + \left|\int_0^s \mathrm{X}_\delta^{(n)}\, d\Lambda - \int_0^s \mathrm{f}_\delta\, d\Lambda\right| + \left|\int_0^s \mathrm{f}_\delta\, d\Lambda - \int_0^s \mathrm{f}\, d\Lambda\right|$$

Entonces, por ser K_δ y Λ finitos c.s., aplicando el Teorema de la Convergencia Monótona y la Ec. (10), cuando $\delta \to 0$, se tiene que $\int_0^s f_\delta\, d\Lambda \to \int_0^s f\, d\Lambda$ y $P\left(\int_0^s X^{(n)}\, d\Lambda \neq \int_0^s X_\delta^{(n)}\, d\Lambda\right) \leq \delta$.

Sea F(t) la función de distribución asociada a la tasa de azar acumulativa Λ entonces, $\left|X_\delta^{(n)}\right|$ es uniformemente integrable, ya que para cualquier n ∈ **N**:

i) $E[\left|X_\delta^{(n)}\right|] \leq E[K_\delta] \leq \int_0^t \mathrm{K}_\delta\, dF \leq \int_0^t \mathrm{K}_\delta\, d\Lambda < \infty$, por hipótesis.

ii) Sea ε > 0 y $\mu(\varepsilon) = \dfrac{\varepsilon}{\Gamma_\delta}$, entonces para cualquier subconjunto A, con P(A) < μ, $\int_A |X_\delta^{(n)}|\, dP \leq \int_A K_\delta\, dP \leq \Gamma_\delta\, P(A) < \varepsilon$.

Y puesto que $X_\delta^{(n)} \xrightarrow{P} f_\delta$, cuando $n \to \infty \Rightarrow \int_0^s |\mathrm{X}_\delta^{(n)} - f_\delta|\, dF \to 0$ | c.s., cuando $n \to \infty$, luego

$$\left|\int_0^s (\mathrm{X}_\delta^{(n)} - f_\delta)\, dF\right| \leq \int_0^s |\mathrm{X}_\delta^{(n)} - f_\delta|\, dF \leq \int_0^s \left(|\mathrm{X}_\delta^{(n)}| + |f_\delta|\right) dF \leq 2K_\delta(s),$$

y como K_δ es integrable, podemos aplicar el Teorema de la Convergencia Dominada a la sucesión $\left(\int_0^s |\mathrm{X}_\delta^{(n)} - f_\delta|\, dF\right)_{n\in\mathrm{N}}$, obteniendo $\int_0^t \left(\int_0^s |\mathrm{X}_\delta^{(n)} - f_\delta|\, dF\right) dF \to 0$, por lo que $\int_0^s X_\delta^{(n)}\, dF$ converge en media, y por tanto en probabilidad, cuando $n \to \infty$, a $\int_0^s f_\delta\, dF$, y así queda establecido que $\int_0^s X_\delta^{(n)}\, d\Lambda \xrightarrow{P} \int_0^s f_\delta\, d\Lambda$.

Por lo que queda demostrada la proposición.

Teorema 3.2. Sea $\{N^{(n)}\}_{n\in N}$ una sucesión de procesos de recuento y todos los elementos asociados que venimos considerando. Supongamos que $\Lambda(s)<\infty$ y $\Lambda_0(s)<\infty$ $\forall s\in[0,t]$, con t>0 fijo. Si existe una función $y(\cdot)$ definida sobre [0, t], una sucesión $(a_n)_{n\in N}$ de números reales tal que $a_n\uparrow\infty$ y se verifican las siguientes hipótesis:

a.) $\inf_{s\in[0,t]} y(s)>0$ y b.) $\sup_{s\in[0,t]}\left|\frac{Y^{(n)}(s)}{a_n^2}-y(s)\right|\xrightarrow{P}0$, cuando $n\to\infty$.

Entonces, se cumplen las hipótesis (3), (4) y (5), y además

$$\forall s\in[0,t],\quad a_n^2\int_0^s\frac{J^{(n)}(u)}{Y^{(n)}(u)}\,d\Lambda(u)\xrightarrow{P}\int_0^s\frac{1}{y(u)}d\Lambda(u),\ \text{cuando } n\to\infty \qquad [11]$$

$$\forall\varepsilon>0\quad a_n^2\int_0^s\frac{J^{(n)}(u)}{Y^{(n)}(u)}I\left(\left|a_n\frac{J^{(n)}(u)}{Y^{(n)}(u)}\right|>\varepsilon\right)d\Lambda(u)\xrightarrow{P}0,\ \text{cuando } n\to\infty \qquad [12]$$

$$a_n\int_0^t(1-J^{(n)}(u))\,d\Lambda(u)\xrightarrow{P}0,\ \text{cuando } n\to\infty \qquad [13]$$

Demostración. Sea $0<m<\inf_{s\in[0,t]} y(s)$ entonces, por (a) y (b), $\exists n_0\in N$ tal que $\forall n\geq n_0$, se tiene que $\frac{Y^{(n)}(s)}{a_n^2}>m$ c.s.

Nótese que por $a_n\uparrow\infty$ y por cumplirse la hipótesis (b), $\inf_{s\in[0,t]} Y^{(n)}(s)\xrightarrow{P}\infty$, y por tanto:

i) Si $X_1^{(n)}(s)=\frac{c(s)\,J^{(n)}(s)}{c(s)+Y^{(n)}(s)}$, entonces $X_1^{(n)}(s)\xrightarrow{P}0$, por estar c(t) acotada. Además, para cada $n\in$ **N**, $X_1^{(n)}(s)\leq 1$ $\forall s\in[0,t]$, y $\int_0^t 1\cdot(d\Lambda_0(u)-d\Lambda(u))<\infty$, por las suposiciones realizadas. Entonces, $\forall\delta>0$, $P(|X_1^{(n)}(s)|\leq 1\ \forall s\in[0,t])=1>1-\delta$, y aplicando la proposición anterior, con $f=0$ y $K_\delta=1$ $\forall\delta>0$, podemos concluir que

$$\sup_{s\in[0,t]}\left|\int_0^t X_1^{(n)}(u)(d\Lambda_0(u)-d\Lambda(u))\right|\xrightarrow{P}0,\ \text{cuando } n\to\infty,$$

cumpliéndose (3).

ii) Si $X_2^{(n)}(s) = \frac{J^{(n)}(s)}{Y^{(n)}(s)}$, entonces $X_2^{(n)}(s) \xrightarrow{P} 0$, y $X_2^{(n)}(s)$ está acotado y a partir de un cierto n_1, con $n_1 \in \mathbf{N}$, se tiene que, $\forall\delta > 0$, $P(|X_2^{(n)}(s)| \leq \frac{1}{m},\ \forall s \in [0,t]) = 1 > 1-\delta$, y además $\int_0^t \frac{1}{m} d\Lambda(s) < \infty$. Luego, aplicando la proposición anterior, con $f = 0$ y $K_\delta = \frac{1}{m}$, se obtiene (4).

iii) Si $X_3^{(n)}(s) = 1 - J^{(n)}(s)$, y por $Y^{(n)}(s) > 0$ c.s., entonces $X_3^{(n)}(s) = 0$ c.s., por lo que se cumple la expresión (5).

Aplicando el apartado (b) se obtiene

$$\left| a_n^2 \int_0^s \frac{J^{(n)}}{Y^{(n)}} d\Lambda - \int_0^s \frac{1}{y} d\Lambda \right| \leq \int_0^s \frac{\left| \frac{Y^{(n)}}{a_n^2} - y \right|}{y \frac{Y^{(n)}}{a_n^2}} d\Lambda \leq$$

$$\leq \sup_{u \in [0,s]} \left| \frac{Y^{(n)}(u)}{a_n^2} - y(u) \right| \int_0^s \frac{1}{m^2} d\Lambda \leq \sup_{u \in [0,s]} \left| \frac{Y^{(n)}(u)}{a_n^2} - y(u) \right| \frac{1}{m^2} \Lambda(s) \xrightarrow{P} 0,$$

por estar $\Lambda(t)$ acotada. Por tanto se cumple (11).

Obsérvese que, $\left| I\left(a_n \frac{J^{(n)}}{Y^{(n)}} > \varepsilon \right) - I\left(\frac{1}{a_n y} > \varepsilon \right) \right| = \left| I\left(1 > \frac{Y^{(n)}}{a_n} \varepsilon \right) - I(1 > a_n y\varepsilon) \right| =$

$$= I\left(\left| \frac{Y^{(n)}}{a_n} \varepsilon - a_n y\varepsilon \right| > 0 \right) = I\left(\varepsilon a_n \left| \frac{Y^{(n)}}{a_n^2} - y \right| > 0 \right) = I\left(\left| \frac{Y^{(n)}}{a_n^2} - y \right| > 0 \right) \xrightarrow{P} 0,$$

por cumplirse la hipótesis (b). Entonces, para $\varepsilon > 0$, $\left| a_n^2 \int_0^t \frac{J^{(n)}}{Y^{(n)}} I\left(\left| a_n \frac{J^{(n)}}{Y^{(n)}} \right| > \varepsilon \right) d\Lambda \right| \leq$

$$\leq \left| a_n^2 \int_0^t \frac{J^{(n)}}{Y^{(n)}} I\left(\left| a_n \frac{J^{(n)}}{Y^{(n)}} \right| > \varepsilon \right) d\Lambda - \int_0^t \frac{1}{y} I\left(\left| a_n \frac{J^{(n)}}{Y^{(n)}} \right| > \varepsilon \right) d\Lambda \right| +$$

$$+\left|\int_0^t \frac{1}{y} I\left(\left|a_n \frac{J^{(n)}}{Y^{(n)}}\right| > \varepsilon\right) d\Lambda - \int_0^t \frac{1}{y} I\left(\frac{1}{a_n y} > \varepsilon\right) d\Lambda\right| + \left|\int_0^t \frac{1}{y} I\left(\frac{1}{a_n y} > \varepsilon\right) d\Lambda\right|$$

El primer sumando del miembro de la derecha está acotado por $\int_0^t \left|a_n^2 \frac{J^{(n)}}{Y^{(n)}} - \frac{1}{y}\right| d\Lambda$, y este valor converge, cuando $n \to \infty$, en probabilidad a cero como acabamos de probar.

El segundo sumando está acotado por $\int_0^t \frac{1}{m} \left| I\left(\left|a_n \frac{J^{(n)}}{Y^{(n)}}\right| > \varepsilon\right) - I\left(\left|\frac{1}{a_n y}\right| > \varepsilon\right)\right| d\Lambda \le$

$\le \frac{1}{m} \sup_{s\in[0,t]} \left| I\left(\left|a_n \frac{J^{(n)}}{Y^{(n)}}\right| > \varepsilon\right) - I\left(\left|\frac{1}{a_n y}\right| > \varepsilon\right)\right| \Lambda(t) \xrightarrow{P} 0$, como hemos visto anteriormente y por estar $\Lambda(t)$ acotada. Nos queda por demostrar que $\left|\int_0^t \frac{1}{y} I\left(\frac{1}{a_n y} > \varepsilon\right) d\Lambda\right|$ converge en probabilidad a cero. Ahora bien, $\left|\int_0^t \frac{1}{y} I\left(\frac{1}{a_n y} > \varepsilon\right) d\Lambda\right| \le \frac{1}{m} \sup_{s\in[0,t]} I\left(\frac{1}{a_n y(s)} > \varepsilon\right) \Lambda(t)$, y por otra parte $a_n \uparrow \infty$ y $m > 0 \Rightarrow$ $\forall \varepsilon > 0$ $\exists n_0 \in \mathbf{N}$ tal que $\forall n \ge n_0$, $1 < \varepsilon m a_n \Rightarrow$ $\Rightarrow$(por $y(s) > m$ $\forall s \in [0,t]$) $\Rightarrow$ $\forall n \ge n_0$, $\frac{1}{a_n y(s)} < \frac{1}{a_n m} < \varepsilon$ $\Rightarrow I\left(\frac{1}{a_n y(s)} > \varepsilon\right) \xrightarrow{P} 0$. Por tanto, queda demostrado (12).

Es evidente, aplicando el punto (iii) de este mismo teorema, que se cumple (13).

Teorema 3.3 Supongamos que existen una sucesión de números reales positivos creciente a infinito, $(a_n)_{n\in N} \uparrow \infty$, y una función $y(\cdot)$ no negativa y tal que $\sigma^2(s) = \int_0^s \frac{1}{y(u)} d\Lambda(u) < \infty$ c.s. $\forall s \in [0,t]$. Si se verifican las hipótesis del teorema anterior entonces, $a_n (\hat{\Lambda}^{(n)} - \Lambda) \xrightarrow{\mathcal{D}} U$, cuando $n \to \infty$, sobre $D[0,t]$, donde $U(\cdot)$ es una martingala Gausiana, con $U(0)=0$ c.s. y que $\forall s_1, s_2 \in [0,t]$ $Cov[U(s_1),U(s_2)] = \sigma^2(\min(s_1,s_2))$,t]. Además, siendo $\sigma_*^2(s) = \int_0^s \frac{J(u)}{Y(u)} dN(u)$, se cumple

$$\sup_{s\in[0,t]} a_n^2 \left|\sigma_*^2(s) - \sigma^2(s)\right| \xrightarrow{P} 0, \text{ cuando } n \to \infty \qquad [14]$$

Por último, (14) sigue siendo cierta al remplazar $\sigma_*^2(s)$ por $\sigma_{**}^2(s) = \int_0^s \frac{J(u)(Y(u) - \Delta N(u))}{Y^3(u)} dN(u)$. Donde $D[0,t]$ es el espacio de las funciones continuas por la derecha y con límites por la izquierda en el intervalo [0,t].

Demostración. Por la demostración del Teorema 3.1 $\left|\hat{\Lambda}^{(n)}(t) - \overline{\Lambda}^{(n)}(t)\right| \xrightarrow{P} 0$, cuando $n \to \infty$, es decir, la diferencia $\left|\hat{\Lambda}^{(n)}(t) - \overline{\Lambda}^{(n)}(t)\right|$ es despreciable al crecer n, y por tanto ambos estimadores tienen las mismas propiedades asintóticas.

El Teorema IV.1.2 de Andersen et al [3] establece, para procesos de recuento multivariantes, la distribución asintótica del estimador de Nelson-Aalen, $\overline{\Lambda}^{(n)}$, sobre intervalos acotados.

Para concluir la demostración utilizamos el Teorema 4.1 de Billingsley [5] $(X_n \xrightarrow{\mathcal{D}} X$ y $X_n - Y_n \xrightarrow{P} 0$ *entonces*, $Y_n \xrightarrow{\mathcal{D}} X)$

Observación 3.1 La demostración del Teorema IV.1.2 de Andersen et al [3] se basa en el Teorema Central del Límite para martingalas locales de Rebolledo [17].

La hipótesis (a) de los Teoremas 3.2 y 3.3 es simplemente para evitar la inclusión de valores donde la proporción de individuos en riesgo sea cero c.s., mientras que la hipótesis (b) es una suposición estándar, que suele cumplirse frecuentemente, diciéndose que $\frac{Y^{(n)}}{a_n^2}$ converge en probabilidad a la función determinística $y(\cdot)$ uniformemente sobre conjuntos acotados.

Por cumplirse las hipótesis (a) y (b) se verifican las expresiones (3), (4), (5), (11), (12) y (13), puesto que la condición dada en (11) es equivalente a que $\forall s \in [0,t]$, y cuando $n \to \infty$, se verifique $a_n^2 \int_0^s \frac{J^{(n)}(u)}{Y^{(n)}(u)} d\Lambda(u) \xrightarrow{P} \sigma^2(s)$.

Si el intervalo considerado es $[t_1, t_2]$, con $0 < t_1 < t_2 < \infty$, sustituyendo los estimadores $\hat{\Lambda}^{(n)}(\cdot)$ por $\hat{\Lambda}^{(n)}(\cdot) - \hat{\Lambda}^{(n)}(t_1)$, los resultados dados en el teorema anterior siguen siendo ciertos, siempre que se den las mismas hipótesis pero en el intervalo $[t_1, t_2]$.

4. INTERVALOS Y BANDAS DE CONFIANZA

Acabamos de probar que bajo ciertas suposiciones se obtienen resultados sobre la convergencia asintótica del estimador propuesto. Más específicamente, si $a_n = \sqrt{n}$, con $n \in \mathbf{N}$, y suponemos que $\frac{Y^{(n)}(s)}{n}$ converge en probabilidad uniformemente sobre conjuntos acotados, entonces $\sqrt{n}(\hat{\Lambda}(s) - \Lambda(s))$ converge a un proceso Gausiano de media cero y varianza $\sigma_*^2(s)$, o $\sigma_{**}^2(s)$, además también contamos con los tres estimadores de la varianza propuestos. Estos resultados van a ser utilizados para obtener intervalos y bandas de confianza.

4.1. Intervalos de confianza

Si suponemos que $s \in [0,t]$ y es fijo, podremos obtener un intervalo de confianza, a nivel $(1-\alpha)$, para $\Lambda(s)$ mediante la normalidad asintótica, que resulta ser

$$\hat{\Lambda}^{(n)}(s) \pm c_{\alpha/2}\, \sigma_*^2(s) \qquad [15]$$

donde $c_{\alpha/2}$ es el valor que verifica $P(Z>c_{\alpha/2})=\frac{\alpha}{2}$, siendo $Z \sim \mathcal{N}(0,1)$. Como ya se ha comentado previamente, también puede considerarse la misma expresión (15), donde σ_*^2 puede ser sustituido por σ_{**}^2, $\bar{\sigma}^{*2}$, $\hat{\sigma}^{*2}$, o $\hat{\bar{\sigma}}^2$.

Para muestras pequeñas los intervalos aquí considerados no tiene porqué ser aceptables. Bie, Borgan, y Liestøl [4], mediante simulaciones, y considerando el estimador de Nelson-Aalen y el estimador $\sigma_*^2(s)$ no obtuvieron buenos resultados para muestras pequeñas. Dichos autores proponen realizar una transformación sobre los datos, g, que sea derivable con derivada continua en $[0,\Lambda(s)+\varepsilon)$. A continuación, aplicando el resultado del Teorema 3.3, se obtiene que, cuando $n \to \infty$, $\frac{g(\hat{\Lambda}^{(n)}(s)) - g(\Lambda(s))}{|g'(\hat{\Lambda}^{(n)}(s)|\sigma_*^2(s)} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1)$. Por este resultado se obtienen intervalos de confianza para $g(\Lambda(s))$ y, puesto que se suponen hipótesis para que g sea inversible, se considera como intervalo de confianza a nivel $(1-\alpha)$, para $\Lambda(s)$ a $g^{-1}(g(\hat{\Lambda}^{(n)}(s)) \pm c_{\alpha/2} |g'(\hat{\Lambda}^{(n)}(s)|\sigma_*^2(s))$.

Otra forma de obtener intervalos de confianza alternativos a los comentados con anterioridad, y en la línea presentada por Bie, Borgan, y Liestøl [4], consiste en variar los parámetros del proceso βeta elegido como prior. En particular, dependiendo de nuestro conocimiento a priori, podremos elegir una familia de funciones $c(\cdot)$ constantes a trozos, suponiendo que Λ es un proceso βeta de parámetros (c, Λ_0), que al variar proporcionaran dicho intervalo de confianza.

También es posible conseguir un intervalo de confianza para $\Lambda(s)$ basándonos en el supuesto de que Λ es un proceso βeta de parámetros (c, Λ_0) y que $\Lambda/_{\text{datos}}$ es un proceso βeta de parámetros $\left(c+Y, \int_0^{(\cdot)} \frac{c\,d\Lambda_0 + dN}{c+Y}\right)$. En este caso el intervalo de confianza a nivel (1-α), para $d\Lambda(s)$ es de la forma $(e_{\alpha/2}, e_{1-\alpha/2})$, siendo $P(X \leq e_{\alpha/2}) = \alpha$, con X una v.a. βeta de parámetros $cd\Lambda_0 + dN$ y $c(1-d\Lambda_0) + Y - dN$. Por tanto, un intervalo de confianza para $\Lambda(s)$ es $(s \cdot e_{\alpha/2}, s \cdot e_{1-\alpha/2})$.

4.2. Bandas de confianza

Analizamos a continuación las bandas de confianza para Λ sobre subconjuntos del intervalo [0,t]. Doksum y Yandell [6] estimaron bandas de confianza para el estimador de Nelson-Aalen de la tasa de azar acumulativa. Dichas estimaciones se basan en la elección de una función, q, verificando ciertas propiedades de regularidad (para una elección particular, $q(x) = \frac{1}{\sqrt{x(1-x)}}$, se obtienen bandas de confianza semejantes a los intervalos de confianza de la forma dada en (15)).

Otra posibilidad es utilizar la idea, expuesta previamente, de introducir una transformación en los datos, con las propiedades convenientes, realizar estimaciones de los datos transformados y obtener una banda de confianza para la tasa de azar acumulativa mediante la transformada inversa de la banda calculada.

Siguiendo de forma paralela con el análisis realizado para intervalos de confianza, y partiendo de procesos βeta, podemos obtener bandas de confianza haciendo variar los parámetros iniciales. Este método, al igual que antes, no nos asegura buenos resultados para muestras pequeñas, o muestras con pocas observaciones no censuradas, y disponiendo de poca información inicial. Las ventajas de este método son, que resulta sencillo de implementar, y que puede resultar suficiente para contrastar gráficamente el comportamiento de los tiempos de fallo estudiados.

Para los tres métodos propuestos anteriormente se puede compara la amplitud de las bandas obtenidas utilizando todos los estimadores de la varianza propuestos.

Con el mismo procedimiento utilizado para el cálculo de intervalos de confianza para v.a. βeta, se puede construir la banda de confianza para Λ basándonos en que

$\Lambda/_{\text{datos}}$ es un proceso βeta de parámetros $\left(c+Y, \int_0^{\cdot} \frac{c\,d\Lambda_0 + dN}{c+Y}\right)$. Para cada s, obtenemos un intervalo de confianza para $d\Lambda(s)$, de la forma $(e_{\alpha/2}, e_{1-\alpha/2})$, por lo que el intervalo de confianza para $\Lambda(s)$ vendría dada por $\left(\int_0^s e_{\alpha/2}(u)\,du, \int_0^s e_{1-\alpha/2}(u)\,du\right)$.

Al igual que el método anterior, este método proporciona un forma sencilla para analizar gráficamente el comportamiento del proceso considerado.

Observación 4.1 Algunas cuestiones aquí planteadas y que son susceptibles de próximos análisis puede ser el examen de la convergencia asintótica del estimador de la tasa de azar acumulativa propuesto, pero con los estimadores de la varianza dados por $\bar{\sigma}^{*2}$, $\hat{\sigma}^{*2}$, o $\hat{\bar{\sigma}}^2$.

También, y utilizando el producto integral, se puede estudiar la convergencia asintótica del estimador $\hat{F}(t) = 1 - \prod_{[0,t]}\left(1 - \frac{c\,d\Lambda_0 + dN}{c+Y}\right)$, comparándola con la del estimador no paramétrico de Kaplan-Meier $\left(\bar{F}(t) = 1 - \prod_{[0,t]}\left(1 - \frac{dN}{Y}\right)\right)$.

REFERENCIAS

AALEN, O.O.(1975). «Statistical inference for a family of counting processes». *PhD thesis, University of California,* Berkeley.

AALEN, O.O.(1978). «Nonparametric inference for a family of counting processes». *Ann. Statist.* **6**, 701-726.

ANDERSEN, P.K., BORGAN, Ø., GILL, R.D., y KEIDING, N.(1993). «Statistical Models Based on Counting Processes», *Springer-Verlag,* New York.

BIE, O., BORGAN, Ø., y LIESTØL, K.(1987). «Confidence intervals and confidence bands for the acumulative hazard rate function and their small sample properties». *Scand. J. Statist.* **14**, 221-233.

BILLINGSLEY, P.(1968). «Convergence of Probability Measures». *Wiley,* New York.

DOKSUM, K.A., y YANDELL, B.S.(1984). «Tests for exponentiality» . *In Krishnaiah P.R. and Sen, P.K., editors. Handbook of Statistics **4**,* 579-611. North-holland, Amsterdan.

FLEMING, T.R., y HARRINGTON, D.P.(1991). «Counting Processes and Survival Analysis». *Wiley,* New York.

GILL, R.D.(1980). «Censoring and Stochastic Integrals». *Mathematical Centre Tracts 124, Mathematisch Centrum,* Amsterdan.

GILL, R.D.(1983). «Discussion of the papers by Helland and Kurtz». *Bull. Internat. Statist. Inst.* **50**(3), 239-243.

HJORT, N.L.(1990). «Nonparametric Bayes estimators based on beta processes in models for life history data». *Ann. Statist.* **18**, 1259-1294.

LENGLART, E.(1977). «Relatión de domination entre deux processus». *Ann. Inst. Henri Poincaré* **13**, 171-179.

MEYER, P.A.(1976). «Un cours sur les intégrales stochastiques. Seminaire de Probabilités X. Lecture Notes in Mathematics» 511, 245-400. *Springer-Verlag, Berlín.*

MORALES, D., PARDO, L. y QUESADA, V.(1991). «Bayesian survival estimation for incomplete data when the life distribution is proportionally related to the censoring time distribution.» *Communications in Statistics. Theory Meth.* **20**(3), 831-850.

NELSON,W.(1969). «Hazard plotting for incomplete failure data». *J. Qual. Technol.* **1**, 27-52.

NELSON, W.(1972). «Theory and -0.24cm}applications of hazard plotting for censored failure data». *Technometrics* **14**, 945-965.

QUESADA, V. y VIVAR, A.(1984). «Estimación paramétrica en modelos no paramétricos Bayesianos de supervivencia». *Revista Estadística Española, Instituto Nacional de Estadística* **103**, 5-23.

REBOLLEDO, R.(1980). «Central limit theorems for local martingales». *Z. Wahrsch Verw Geb.* **51**, 269-286.

CONVERGENCE OF ESTIMATORS OF CUMULATIVE HAZARD RATE BASED ON BETA PROCESSES

SUMMARY

The analysis of the cumulative distribution function of a failure time random variable, based on right-censored observations, can be made by means of counting processes and using the martingale methods. On the other hand, the methodology of nonparametric Bayesian survival analysis can be applied to βeta processes for obtaining estimators of cumulative hazard rate that can be written as stochastic integrals of bounded processes verifing hypothesis of regularity about continuity, and where the integrators are counting processes, or associated martingales in the Doob-Meyer decomposition, or processes relative to the occurrence rates for those events.

In this paper, we use the martingale framework to establish limiting distribution results for sequences of nonparametric Bayesian estimators of cumulative hazard rate based on βeta processes. First, we prove that asymptotic distribution of the statistics considered coincides with asymptotic behavior of Nelson-Aalen nonparametric estimator. Next, we apply that the classical Nelson-Aalen estimator tends to true cumulative hazard rate. Finally, we present several procedures to construct confidence intervals and confidence bands for cumulative hazard rate.

Key words: Right censored date, asymptotic distribution, martingale, βeta processes , counting processes, cumulative hazard rate.