# PROPUESTA DE ESTIMADORES PARA ENFRENTAR SITUACIONES DE RIESGOS COMPETITIVOS DEPENDIENTES

Rolando Uranga Piña[1]*, Mayelin Mirabal Sosa[2]**, Armando Seuc Jo[3]***
*Centro Nacional Coordinador de Ensayos Clínicos
**Instituto Finlay. Investigación y Producción de Vacunas
***Instituto Nacional de Angiología y Cirugía Vascular

**ABSTRACT**
In a classic problem of competing risks, estimates of diverse rates are carried out starting from suppositions about the causes that compete. These suppositions have included essentially the independence among the competing causes. The present work intends to relax these suppositions to more general situations. It is demonstrated that the Net Rate can be estimated from observed data in the presence of Competing Risks with certain type of dependence. Generalizations of the Actuarial estimator for the Net Rate are suggested for the two classes of dependencies studied. By means of simulations, the proposed estimators and the one reported in the literature for independent competing risks are compared, and the results confirm the fact that the new estimators can be used in the circumstances defined to estimate the probabilities of survival.

**KEY WORDS:** Dependent Competing Risks, Joint Survival Function, Net Rate, Crude Rate.

**MSC**: 62F10

**RESUMEN**
En un problema clásico de riesgos competitivos se realizan estimaciones de diversas tasas a partir de supuestos hechos sobre las causas que compiten. Los supuestos han incluido fundamentalmente la independencia entre las causas competitivas. El presente trabajo se propone relajar estos supuestos a situaciones más generales. Se demuestra que la Tasa Neta puede ser estimada a partir de datos observados en presencia de Riesgos Competitivos con determinado tipo de dependencia. Se proponen, para los dos tipos de dependencia estudiados, generalizaciones del estimador tipo Tabla de Vida para la Tasa Neta. Se comparan, mediante simulaciones, los estimadores propuestos y el reportado en la literatura para riesgos competitivos independientes lo que corrobora que los nuevos estimadores pueden ser utilizados en las circunstancias que se han definido para estimar la probabilidad de supervivencia.

## 1. INTRODUCCIÓN

Es universalmente conocido el Análisis de Supervivencia como rama de la Estadística dedicada al estudio de variables aleatorias no negativas, interpretadas como tiempo hasta la ocurrencia de un determinado evento (Cox [1984], Ibrahim [2005]). El evento posible puede ser único – muerte, progresión de un tumor, rotura de una componente mecánica – o múltiple – muerte por padecimiento cardíaco, cáncer, u otra causa; rotura por vibración, corrosión, etc. –. Cuando el evento potencial es múltiple y sólo puede observarse uno para cada sujeto, se habla de una situación de Riesgos Competitivos.

La teoría de riesgos competitivos se remonta a los intentos, en 1760, de Daniel Bernoulli de separar el riesgo de muerte por viruela del riesgo de muerte por otras causas (Lindqvist [2007], Blower [2004]). Usualmente, el evento múltiple queda especificado por las múltiples causas de ocurrencia de un único evento. De aquí el nombre de causas o riesgos competitivos. Chiang [1961] estructura la situación de riesgos competitivos mediante la definición de tres tasas de interés diferentes: Tasa Global, definida como la probabilidad global de ocurrencia del evento por cualquiera de las causas; Tasa Neta, definida como la probabilidad de ocurrencia

[1] rolando@cencec.sld.cu
[2] mmirabal@finlay.edu.cu
[3] metodoli@infomed.sld.cu

del evento por una determinada causa en una población hipotética en la cual sólo estuviera activa esa causa; Tasa Cruda, definida como la probabilidad de ocurrencia del evento por una determinada causa en presencia de las causas restantes.

De ordinario, los estudios de riesgos competitivos parten de estimadores empíricos de las funciones de supervivencia crudas y se centran en intentar estimar las probabilidades de supervivencia netas, con el objetivo de predecir patrones de mortalidad esperados en condiciones hipotéticas en que ciertas causas de muerte han sido eliminadas (Tsiatis [1975]). Sin embargo, este mismo autor demuestra que es imposible estimar las tasas netas a partir solamente de las tasas crudas, sin asumir supuestos adicionales acerca del tipo de dependencia entre las causas competitivas. A este resultado se le conoce como "Problema de Identificación" – Identifiability Problem – (Lindqvist [2007]).

En la literatura ha sido indiscriminadamente usado el estimador de la tasa neta bajo el supuesto de independencia. Si las causas que compiten no son independientes, este estimador proporciona resultados sesgados. Como ejemplo, el incremento observado en los últimos 5 a 10 años en las tasas crudas de mortalidad por cáncer en algunos países europeos puede deberse a que las tasas crudas de mortalidad por enfermedades cardíacas han disminuido, lo que posibilita que los sobrevivientes de estas últimas tengan ahora la posibilidad – y de hecho lo hagan – de morir por cáncer (Llorca [2001]). Sería errado asumir independencia en este caso y si se desea ofrecer un estimador de la tasa neta de mortalidad por cáncer, es necesario plantear supuestos adicionales.

En el presente trabajo se proponen dos escenarios de dependencia de las causas competitivas en los cuales es posible estimar la tasa neta. Se ofrecen además expresiones para los estimadores respectivos que constituyen generalizaciones del estimador usual de Tabla de Vida. Al primer escenario se le llama *escenario de independencia condicional* y al segundo, *escenario de independencia parcial*. Se ilustra la teoría expuesta mediante un ejemplo simulado.

El desarrollo teórico se basa en la formulación del problema de riesgos competitivos en términos de tiempos latentes o conceptuales, en el uso apropiado del estimador clásico de Tabla de Vida y en las facilidades de programación que brinda el lenguaje S, implementación R, versión 2.4.1 (R Development Core Team [2006]).

## 2. RIESGOS COMPETITIVOS

Se estudia una situación en la cual cada individuo está expuesto a una serie de $k$ posibles eventos de interés (o alternativamente, a un evento que tiene $k$ posibles causas de interés).

Una forma intuitiva de describir el problema de riesgos competitivos, es en términos de tiempos latentes o conceptuales. Ver Lindqvist [2007], Moeschberger ML, David HA [1971], Cox DR [1972], Elwert F [2006].

Sean $T_1, T_2, ..., T_k$, variables aleatorias no negativas. Se considera que la variable $T_i$ representa el tiempo latente de ocurrencia de un evento del tipo $i$ $(1 \leq i \leq k)$. Es decir, $T_i$ es el tiempo hipotético de ocurrencia del evento por la causa $i$ si el resto de las causas no estuvieran presentes. Sean $T = \min\{T_1, T_2, ..., T_k\}$ y $J = j$ si $T = T_j$ ($j \in \{1, 2, ..., k\}$).

La información básica disponible en un problema clásico de Riesgos Competitivos es entonces la distribución conjunta de las variables aleatorias ($T$, $J$).

Las causas del evento o los tipos de eventos que intervienen en situaciones como las anteriores se conocen como riesgos o causas competitivas. Las causas competitivas son excluyentes porque se asume que sólo se observa una de ellas en cada sujeto.

### 2.1. Riesgos Competitivos independientes

Un aspecto muy discutido teóricamente y relevante en la literatura concerniente a riesgos competitivos es la posible interrelación entre éstos, expresada en términos de la independencia o no entre las causas competitivas. Muy pocos artículos aplicados hacen referencia a la suposición de independencia en el marco de los estudios de Riesgos Competitivos; sin embargo, esta suposición es restrictiva en este escenario porque es muy probable que las causas competitivas estén interrelacionadas. Ver Elwert F [2006], Lindqvist BH, Ruggeri F, Kenett R, Faltin FW [2007], Lindqvist BH, Stove B, Langseth H [2006].

De acuerdo a la formulación de una situación de riesgos competitivos en términos de tiempos latentes expuesta arriba, puede formalizarse la definición de riesgos competitivos independientes de la siguiente forma:

Se dirá que los riesgos competitivos son independientes, si las variables aleatorias $T_1, T_2, ..., T_k$ son mutuamente independientes.

Shen y Thall comentan que siempre asumirán independencia entre los riesgos competitivos, porque a partir de los datos no se puede confirmar/contrastar ese supuesto (Shen Y, Thall PF [1998]). El mismo argumento serviría para justificar el supuesto de dependencia (en general mucho más probable que el de independencia en el sentido práctico).

En el marco de un estudio de riesgos competitivos ha sido de interés la estimación de las funciones de sub-distribución:

$$F_j(t) = P(T < t, J = j); \quad t > 0,\ j = 1,...,k$$

Es decir, ha sido de interés estimar la probabilidad de ocurrencia del evento por una causa en presencia de las restantes causas (Tasa Cruda).

También en este marco, ha sido común la estimación de la probabilidad de ocurrencia del evento en una situación hipotética en que la causa $j$ es la única que está presente y el resto de las causas han sido eliminadas, $P(T_j < t)$ (Tasa Neta). A continuación se proponen dos variantes que permiten estimar esta probabilidad cuando existe dependencia entre dos causas competitivas.

## 3. TRATAMIENTO DE RIESGOS COMPETITIVOS DEPENDIENTES

Sea $K(t_1,...,t_k) = P(T_1 \geq t_1,...,T_k \geq t_k)$ la función de supervivencia conjunta de $T_1, T_2, ..., T_k$. Se considerará el caso $k = 2$.

El problema consiste en identificar condiciones que determinen la existencia de una única $K(t_1, t_2)$, asumiendo que se han dado como prefijadas las tasas crudas
$F_1(t) = P(T_1 < t, J = 1) = P(T_1 < t, T_1 < T_2)$ y $F_2(t) = P(T_2 < t, J = 2) = P(T_2 < t, T_2 < T_1)$ - estas son las probabilidades de supervivencia a cada una de las causas de interés en presencia de la otra causa, de manera que sea posible estimar las probabilidades de supervivencia a cada una de las causas de interés cuando la otra causa ha sido eliminada; es decir, se desea estimar las tasas netas $P(T_1 < t)$ y $P(T_2 < t)$. Los resultados que se proponen se resumen en las dos proposiciones siguientes.

**Proposición 1 (Escenario de Independencia Condicional).** Sea *A* un evento y $A^c$ el evento complementario. Dadas las sub-distribuciones:

$$G_1(t) = P(T_1 < t, T_1 < T_2, A),\ G_2(t) = P(T_2 < t, T_2 < T_1, A)$$

$$H_1(t) = P(T_1 < t, T_1 < T_2, A^c)\,,\ H_2(t) = P(T_2 < t, T_2 < T_1, A^c)$$

y bajo el supuesto de independencia condicional:

$$P(T_1 \geq t_1, T_2 \geq t_2 / A) = P(T_1 \geq t_1 / A) P(T_2 \geq t_2 / A)$$

$$P(T_1 \geq t_1, T_2 \geq t_2 / A^c) = P(T_1 \geq t_1 / A^c) P(T_2 \geq t_2 / A^c)$$

se verifica que:

$$K(t_1, t_2) = \overline{G}(0) \exp\left\{ -\int_0^{t1} \frac{g_1(u)}{\overline{G}(u)} du - \int_0^{t2} \frac{g_2(u)}{\overline{G}(u)} du \right\} + \overline{H}(0) \exp\left\{ -\int_0^{t1} \frac{h_1(u)}{\overline{H}(u)} du - \int_0^{t2} \frac{h_2(u)}{\overline{H}(u)} du \right\}$$

donde:

$$g_1(t) = G_1'(t)\,,\ g_2(t) = G_2'(t)$$

$$h_1(t) = H_1'(t)\,,\ h_2(t) = H_2'(t)$$

$$\overline{G}(t) = G(\infty) - G(t)\,,\ \overline{H}(t) = H(\infty) - H(t)$$

$$G(t) = G_1(t) + G_2(t)$$

$$H(t) = H_1(t) + H_2(t)$$

**Demostración:**

$$K(t_1, t_2) = P(T_1 \geq t_1, T_2 \geq t_2)$$

$$= P(A) P(T_1 \geq t_1, T_2 \geq t_2 / A) + P(A^c) P(T_1 \geq t_1, T_2 \geq t_2 / A^c)$$

Sustituyendo:

$$P(A) = \overline{G}(0)\,,\quad P(A^c) = \overline{H}(0)$$

Se obtiene,

$$K(t_1, t_2) = \overline{G}(0) P(T_1 \geq t_1, T_2 \geq t_2 / A) + \overline{H}(0) P(T_1 \geq t_1, T_2 \geq t_2 / A^c)$$

Note que bajo el supuesto de independencia condicional:

$$P(T_1 \geq t_1, T_2 \geq t_2 / A) = P(T_1 \geq t_1 / A) . P(T_2 \geq t_2 / A)$$

Además: $P(T_1 < t, T_1 < T_2 / A) = \dfrac{P(T_1 < t, T_1 < T_2, A)}{P(A)} = \dfrac{G_1(t)}{\overline{G}(0)}$ y

$$P(T_2 < t, T_2 < T_1 / A) = \frac{P(T_2 < t, T_2 < T_1, A)}{P(A)} = \frac{G_2(t)}{\overline{G}(0)}$$

Bajo el supuesto de independencia y de acuerdo al resultado de Tsiatis [1975] la función de supervivencia conjunta, dado el evento *A*, es identificable y puede ser escrita como:

$$P(T_1 \geq t_1, T_2 \geq t_2 / A) = \exp\left\{ -\int_0^{t1} \frac{g_1(u) / \overline{G}(0)}{(\overline{G}(0) - G(u)) / \overline{G}(0)} du - \int_0^{t2} \frac{g_2(u) / \overline{G}(0)}{(\overline{G}(0) - G(u)) / \overline{G}(0)} du \right\}$$

$$= \exp\left\{ -\int_0^{t1} \frac{g_1(u)}{\overline{G}(u)} du - \int_0^{t2} \frac{g_2(u)}{\overline{G}(u)} du \right\}$$

Por un procedimiento similar queda:

$$P(T_1 \geq t_1, T_2 \geq t_2 / A^c) = \exp\left\{-\int_0^{t1} \frac{h_1(u)}{\overline{H}(u)} du - \int_0^{t2} \frac{h_2(u)}{\overline{H}(u)} du\right\}$$

De este modo se obtiene la expresión deseada para $K(t_1, t_2)$.

**Proposición 2 (Escenario de Independencia Parcial).** Dadas las sub-distribuciones:

$$F_1(t) = P(T_1 < t, T_1 < T_2),\ F_2(t) = P(T_2 < t, T_2 < T_1)$$

y bajo el supuesto de independencia parcial:

$K(t_1, t_2) = P(T_1 \geq t_1, T_2 \geq t_2) = A(t_1)B(t_2)C(t_1, t_2)$, $C(t_1, t_2)$ conocida, se verifica:

$$K(t_1, t_2) = \frac{C(t_1, t_2)}{C(0,0)} \exp\left\{-\int_0^{t1} \frac{f_1(u)}{\overline{F}(u)} du - \int_0^{t2} \frac{f_2(u)}{\overline{F}(u)} du - \int_0^{t1} \frac{C_1(u,u)}{C(u,u)} du - \int_0^{t2} \frac{C_2(u,u)}{C(u,u)} du\right\}$$

donde:

$$f_1(t) = F_1'(t),\ f_2(t) = F_2'(t)$$

$$\overline{F}(t) = 1 - F(t) = 1 - F_1(t) - F_2(t)$$

$$C_1(t_1, t_2) = \frac{\partial}{\partial t_1} C(t_1, t_2),\ C_2(t_1, t_2) = \frac{\partial}{\partial t_2} C(t_1, t_2)$$

**Demostración:**

Tsiatis [1975] demuestra que dadas las funciones de sub-distribución, y cualquiera sea la función de supervivencia conjunta *K*, se tiene que:

$$K_1(t,t) = -f_1(t) \qquad (1)$$

$$K_2(t,t) = -f_2(t) \qquad (2)$$

donde $K_1(t_1, t_2) = \frac{\partial}{\partial t_1} K(t_1, t_2)$, $K_2(t_1, t_2) = \frac{\partial}{\partial t_2} K(t_1, t_2)$.

Sumando (1) y (2) e integrando se obtiene: $K(t,t) = \overline{F}(t)$.

Sustituyendo ahora en (1) y en (2) $K(t,t) = A(t)B(t)C(t,t)$ queda:

$$A'(t)B(t)C(t,t) + A(t)B(t)C_1(t,t) = -f_1(t) \qquad (3)$$

$$A(t)B'(t)C(t,t) + A(t)B(t)C_2(t,t) = -f_2(t) \qquad (4)$$

Dividiendo (3) por $K(t,t) = A(t)B(t)C(t,t)$ y reagrupando:

$$\frac{A'(t)}{A(t)} = \frac{-f_1(t)}{\overline{F}(t,t)} - \frac{C_1(t,t)}{C(t,t)} \qquad (5)$$

Análogamente de (4) se obtiene:

$$\frac{B'(t)}{B(t)} = \frac{-f_2(t)}{\overline{F}(t,t)} - \frac{C_2(t,t)}{C(t,t)} \qquad (6)$$

Integrando en ambos miembros de (5) y (6) y despejando $A(t)$ de (5) y $B(t)$ de (6):

$$A(t)=D_1\exp\left\{-\int_0^t \frac{f_1(u)}{\overline{F}(u)}du-\int_0^t \frac{C_1(u,u)}{C(u,u)}du\right\} \tag{7}$$

$$B(t)=D_2\exp\left\{-\int_0^t \frac{f_2(u)}{\overline{F}(u)}du-\int_0^t \frac{C_2(u,u)}{C(u,u)}du\right\} \tag{8}$$

Multiplicando (7) y (8) y evaluando en $t=0$ se llega a que $D_1D_2=\dfrac{1}{C(0,0)}$.

Finalmente, sustituyendo las expresiones de $A(t_1)$ y $B(t_2)$ se obtiene la expresión deseada de $K(t_1,t_2)$.

Sea dado un problema de riesgos competitivos dependientes donde se tiene el supuesto de independencia entre los riesgos, condicionado a los valores de una covariable binaria que representa la ocurrencia o no de un evento $A$ - supuesto de independencia condicional -. El estimador Tabla de Vida para la supervivencia a la causa 1 definido para tal situación y utilizando el resultado de la proposición 1 quedaría como:

$$\hat{S}_1(t)=\hat{\alpha}\hat{S}_1^A(t)+\hat{\beta}\hat{S}_1^{A^c}(t)$$

donde:

$\hat{\alpha}$ y $\hat{\beta}$ son los estimaciones de $P(A)$ y $P(A^c)$ respectivamente.

$\hat{S}_1^A(t)$ y $\hat{S}_1^{A^c}(t)$ son los estimadores de la probabilidad de supervivencia a la causa 1, definidos según el esquema del estimador tipo Tabla de Vida para la Tasa Neta bajo el supuesto de independencia (ver Farley TMM [1986]), para los sujetos a los que le ocurre $A$ y $A^c$ respectivamente.

Lo anterior da lugar a un estimador tipo Tabla de Vida ajustado para la Tasa Neta que incorpora la interdependencia entre dos causas competitivas del tipo "independencia condicional". Es un estimador muy simple de calcular, pues sencillamente pondera el estimador de la supervivencia de los individuos que cumplen con la condición $A$ por la probabilidad de que ocurra el evento $A$ y lo suma al producto de la supervivencia de los individuos que no cumplen con $A$ y la probabilidad de que no ocurra $A$.

Por otra parte, sea $T_1=\min(U_1,V_1)$, $T_2=\min(U_2,V_2)$, $T=\min(T_1,T_2)$, se observa $(T,J,V_1,V_2)$, $V_1,V_2$ son dependientes, $U_1,U_2$ son independientes, $(U_1,U_2)$ es independiente de $(V_1,V_2)$. Entonces:

$$\begin{aligned}P(T_1\ge t_1,T_2\ge t_2)=K(t_1,t_2)&=P(U_1,V_1\ge t_1;U_2,V_2\ge t_2)\\&=P(U_1\ge t_1,U_2\ge t_2;V_1\ge t_1,V_2\ge t_2)\\&=P(U_1\ge t_1,U_2\ge t_2)P(V_1\ge t_1,V_2\ge t_2)\\&=P(U_1\ge t_1)P(U_2\ge t_2)P(V_1\ge t_1,V_2\ge t_2)\end{aligned}$$

La expresión anterior tiene la forma de la función de supervivencia definida en la Proposición 2, por lo que es "identificable". En esta misma expresión se observa que $T_1$ y $T_2$ son causas competitivas dependientes y que la dependencia entre estos tiempos está motivada por la dependencia entre $V_1$ y $V_2$. El estimador Tabla de Vida para la supervivencia a la causa 1 definido para tal situación y utilizando el resultado de la proposición 2 quedaría como:

$$\hat{S}_1(t)=\hat{C}(t,0)\frac{\hat{S}_1^T(t_1)}{\hat{S}_1^V(t_1)}$$

donde: $\hat{C}(t,0)$ se obtiene al estimar la probabilidad $P(V_1 \geq t, V_2 \geq 0) = P(V_1 \geq t)$, como el número de observaciones que cumplen la condición $V_1 \geq t$, dividido por el número total de observaciones.

$\hat{S}_1^T(t_1)$ es el estimador de la probabilidad de supervivencia a la causa 1, definido según el esquema del estimador tipo Tabla de Vida para la Tasa Neta bajo el supuesto de independencia, a partir de los datos $(T, J)$.

$\hat{S}_1^V(t_1)$ es el estimador de la probabilidad de supervivencia a la causa 1, definido según el esquema del estimador tipo Tabla de Vida para la Tasa Neta bajo el supuesto de independencia, a partir de los datos $(V, J')$ con $V = \min(V_1, V_2)$ y $J'$ igual a 1 ó 2 según $V = V_1$ ó $V = V_2$.

Lo anterior da lugar a un estimador tipo Tabla de Vida ajustado para la Tasa Neta que incorpora la interdependencia entre dos causas competitivas del tipo "independencia parcial".

**4. SIMULACIÓN**

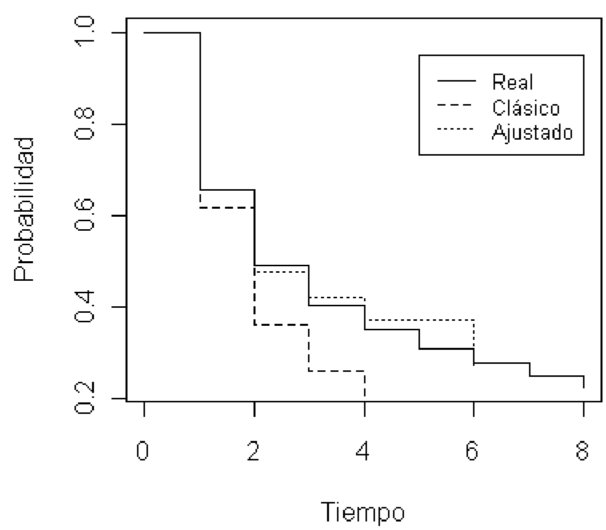


**Figura No.1**. Comparación de las estimaciones de la supervivencia usando el estimador clásico y el ajustado con respecto a la supervivencia real, para datos simulados con dos riesgos competitivos dependientes.

Para ilustrar los resultados de la proposición 1, se generó primeramente una muestra de tamaño 1000 proveniente de una variable aleatoria binaria Z ~ B($\pi$), con $\pi$=0.5, donde B($\pi$) representa la distribución de Bernoulli de parámetro $\pi$: P(Z = 1) = $\pi$, P(Z = 0) = 1 - $\pi$. A continuación, se generaron tiempos latentes de acuerdo a la ley: T1 ~ exp(p), T2 ~ exp(1-p) si Z = 1; T1 ~ exp(q), T2 ~ exp(1-q) si Z = 0; con p = 0.1, q = 0.9. Se representa por exp($\rho$) la distribución exponencial de parámetro $\rho$, cuya función de supervivencia es $S(t) = e^{-\rho t}$. De este modo, se obtuvo una muestra aleatoria de tamaño 1000 formada por tríos $(Z_i, T_{1i}, T_{2i})$, $i=1,\ldots,1000$, lo cual permitió construir la muestra simulada $(T_i, C_i)$ de tamaño 1000, donde $T_i = min(T_{1i}, T_{2i})$, $C_i = 1$ si $T_{1i} \leq T_{2i}$, $C_i = 2$ si $T_{1i} > T_{2i}$. A partir de la muestra simulada fue posible estimar la función de supervivencia neta asociada a la causa 1, o sea, $S_1(t) = P(T_1 \geq t)$, por dos métodos diferentes: mediante el estimador Tabla de Vida para la Tasa Neta bajo el supuesto incorrecto de independencia de las dos causas competitivas (ver Farley TMM [1986]); mediante el estimador Tabla de Vida ajustado para la Tasa Neta que incorpora la interdependencia entre las dos causas competitivas de tipo “independencia condicional”. Se dispone en este caso, además, de la expresión analítica de $S_1(t)$,

$$S_1(t) = K(t,0) = 0.5\exp\{-qt\} + 0.5\exp\{-pt\}$$

En la Figura No. 1 se muestran los resultados de la simulación. Con el estimador propuesto se obtienen estimaciones de la supervivencia más precisas que las obtenidas usando el método clásico. Si por el contrario se generan tiempos latentes independientes entonces las estimaciones obtenidas a partir del estimador clásico y el estimador ajustado definido bajo las condiciones de la Proposición 1 son indistinguibles. Ver Figura No. 2. Las simulaciones fueron generadas mediante el lenguaje de programación S, implementación R, versión 2.4.1 (R Development Core Team [2006]). Los códigos pueden consultarse en el apéndice 1.

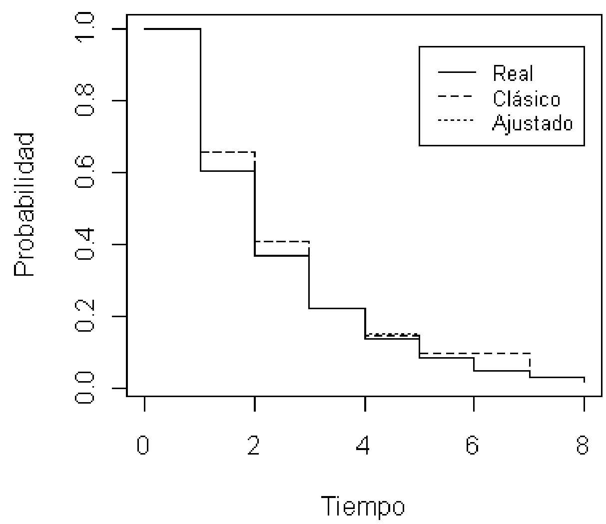


**Figura No.2**. Comparación de las estimaciones de la supervivencia usando el estimador clásico y el ajustado con respecto a la supervivencia real para datos simulados con dos riesgos competitivos independientes.

## 5. DISCUSIÓN Y CONCLUSIONES

Si el interés de un estudio es la estimación de la probabilidad de ocurrencia del evento en presencia de las restantes causas, entonces puede utilizarse el estimador de la Tasa Cruda independientemente de si los riesgos competitivos son o no independientes. Si por el contrario, el objetivo es la estimación de la probabilidad de ocurrencia del evento en ausencia de las restantes causas el estimador clásico de las Tasas Netas es apropiado sólo si los tiempos latentes de ocurrencia de los eventos son independientes.

Con los resultados propuestos en este trabajo se puede concluir que la Tasa Neta para riesgos competitivos dependientes puede ser estimada en dos escenarios:

I) Bajo el supuesto de independencia condicional definido en la proposición 1.
II) Bajo el supuesto de independencia parcial definido en la proposición 2.

El uso del estimador Tabla de Vida ajustado a los tipos de dependencia entre las causas competitivas que se han denominado independencia condicional e independencia parcial, mejora al clásico, basado en el supuesto usual de independencia, como muestran las simulaciones.

El resultado que se ha presentado, a pesar de ser de corte teórico, tiene un impacto social significativo pues mediante su aplicación se pueden confeccionar escenarios plausibles relacionados con reducciones en las tasas de mortalidad y morbilidad que permitan evaluar el impacto de estas reducciones sobre, por ejemplo, la Esperanza de Vida al Nacer, y sobre la Esperanza de Vida Saludable.

Es conocida la importancia de una correcta evaluación de las tasas de mortalidad y morbilidad, así como de la Esperanza de Vida Saludable, con el fin de planificar recursos para las distintas esferas de un país. La herramienta que se propone es de muy fácil aplicación para los investigadores biomédicos después que se tiene un conocimiento profundo del problema que se está abordando y de las posibles interrelaciones entre las causas competitivas.

**RECEIVED APRIL 2010**
**REVISED OCTOBER 2010**

## REFERENCIAS

[1] COX DR, OAKES D (1984): **Analysis of Survival Data**. Chapman and Hall, London.

[2] IBRAHIM JG (2005): Applied Survival Analysis. **Presentation in the 21st Annual Summer Workshop of the American Statistical Association.**

[3] LINDQVIST BH, RUGGERI F, KENETT R, FALTIN FW (2007): **Competing Risks. Encyclopedia of Statistics in Quality and Reliability**. John Wiley & Sons Ltd, Chichester,

[4] BLOWER S (2004): An attempt at a new analysis of the mortality caused by smallpox and of the advantages of inoculation to prevent it. Daniel Bernoulli. **Reviews in Medical Virology,** 14, 275-288.

[5] CHIANG CL (1961): A stochastic study of the life table and its applications: III. The follow-up study with the consideration of competing risks. **Biometrics** 17, 57-78.

[6] TSIATIS A (1975): A Nonidentifiability Aspect of the Problem of Competing Risks. **Proc. Nat. Acad. Sci.** 72, 20-22.

[7] LLORCA J, DELGADO RM (2001): Competing risks analysis using Markov chains: impact of cerebrovascular and ischemic heart disease in cancer mortality. **International Epidemiological Association** 30, 99-101.

[8] R DEVELOPMENT CORE TEAM (2006): R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

[9] MOESCHBERGER ML, DAVID HA (1971): Life tests under competing causes of failure and the theory of competing risks. **Biometrics** 27, 909-933.

[10] COX DR (1972): Regression model and life tables. **J Roy Statist. Soc.** 34, 187–220.

[11] ELWERT F. (2006): Non-Independence in Competing Risk Models. Disponible en: URL: http://www.iq.harvard.edu/blog/sss/archives/2006/03/nonindependence.shtml, acceso en noviembre de 2006.

[12] LINDQVIST BH, STOVE B, LANGSETH H (2006): Modelling of dependence between critical failure and preventive maintenance: The repair alert model. **Journal of Statistical Planning and Inference *136*,** 1701-1717.

[13] SHEN Y, THALL PF (1998): Parametric likelihoods for multiple non-fatal competing risk and death. **Stat Med** 17, 99-1015.

[14] FARLEY TMM (1986): Life-table methods for contraceptive research. **Stat Med** 5, 475-489.

**APÉNDICE 1. Códigos generados en R**

```
# Datos globales
n<-1000;p<-0.1;q<-0.9;r<-0.5*(p+q)
#  Generación de la covariable
z<-rep(0,n);for (i in 1:n) z[i]<-rbinom(1,1, 0.5)
# Estimador clásico de Tabla de Vida
survival<-function(T,C,n,nint,lint){
  # Número de eventos de interés y de censuras en cada intervalo
  d<-rep(0,nint);w<-rep(0,nint)
  for (j in 1:nint) for (i in 1:n){
  if ((j-1)*lint<=T[i] & T[i]<j*lint){
  if (C[i]==1) d[j]<-d[j]+1 else w[j]<-w[j]+1}}
  # Número en riesgo al inicio de cada intervalo
  nrisk<-rep(0,nint);nrisk[1]<-n
  for (j in 2:nint) nrisk[j]<-nrisk[j-1]-d[j-1]-w[j-1]
  # Cálculo la función de supervivencia clásica para cada
  # intervalo
  SC<-rep(1,nint);
  for (j in 1:nint) SC[j]<-1-d[j]/(nrisk[j]-(1/2)*w[j])
  SCA<-rep(1,nint);SCA[1]<-round(SC[1],3)
  for (j in 2:nint) for (i in 1:j) SCA[j]<-round(SCA[j]*SC[i],3)
  result<-SCA
}
########### Riesgos dependientes ###########
#  Generación de los tiempos latentes
T1<-rep(0,n);T2<-rep(0,n);for (i in 1:n){
if (z[i]==1) {T1[i]<-rexp(1,p);T2[i]<-rexp(1,1-p)}
        else {T1[i]<-rexp(1,q);T2[i]<-rexp(1,1-q)}}
#  Datos observados (tiempos y causas)
C<-rep(0,n);T<-rep(0,n);for (i in 1:n){
T[i]<-min(T1[i],T2[i]);C[i]<-which.min(c(T1[i],T2[i]))}
########## Riesgos independientes ##########
#  Generación de los tiempos latentes
T1ind<-rep(0,n);T2ind<-rep(0,n);
for (i in 1:n){T1ind[i]<-rexp(1,r);T2ind[i]<-rexp(1,1-r)}
#  Datos observados (tiempos y causas)
Cind<-rep(0,n);Tind<-rep(0,n);for (i in 1:n){
Tind[i]<-min(T1ind[i],T2ind[i]);
Cind[i]<-which.min(c(T1ind[i],T2ind[i]))}
###########################################################
# Número de intervalos
nint<-(trunc(max(T,Tind))+1);lint<-(trunc(max(T))+1)/nint
# Estimaciones
n0<-length(z[z==0]);n1<-length(z[z==1]);
TZ0<-T[z==0];TZ1<-T[z==1];CZ0<-C[z==0];CZ1<-C[z==1];
TZ0ind<-Tind[z==0];TZ1ind<-Tind[z==1];
CZ0ind<-Cind[z==0];CZ1ind<-Cind[z==1];
S<-survival(T,C,n,nint,lint)
S0<-survival(TZ0,CZ0,n0,nint,lint);
S1<-survival(TZ1,CZ1,n1,nint,lint)
Sind<-survival(Tind,Cind,n,nint,lint);
S0ind<-survival(TZ0ind,CZ0ind,n0,nint,lint)
S1ind<-survival(TZ1ind,CZ1ind,n1,nint,lint);
t<-seq(0,nint*lint,lint)
# Visualizando los resultados
par(mfrow=c(1,1));plot(t,0.5*exp(-q*t)+0.5*exp(-p*t),
type="s",xlab="Tiempo",ylab="Probabilidad")
lines(t,c(1,S),type="s",lty=2);
lines(t,c(1,(n0/n)*S0+(n1/n)*S1),type="s",lty=3)
legend(5,0.95,c("Real","Clásico","Ajustado"),lty=1:3,cex=0.8)
plot(t,exp(-r*t),type="s",xlab="Tiempo",ylab="Probabilidad")
lines(t,c(1,Sind),type="s",lty=2);
lines(t,c(1,(n0/n)*S0ind+(n1/n)*S1ind),type="s",lty=3)
legend(5,0.95,c("Real","Clásico","Ajustado"),lty=1:3,cex=0.8)
```