Rincón de la Bioestadística

Tópicos de análisis de sobrevida: cuarta parte

Gabriel Cavada Ch.[1,2]
[1]Facultad de Medicina, Universidad de Los Andes.
[2]División de Bioestadística, Escuela de Salud Pública, Universidad de Chile.

Survival analysis topics: part four

Recordando los conceptos de función de riesgo (h(t)) y función de riesgo acumulados, los que fueron expuestos en "Tópicos de Análisis de Sobrevida: Primera parte" es posible introducir los Modelos Riesgos Proporcionales:

Modelo de riesgos proporcionales

Con el objetivo de poder identificar los potenciales factores de riesgo, se ilustra la metodología de los riesgos proporcionales, que se puede formular como sigue:

Si se considera que cada individuo tiene un particular perfil o conjunto de variables explicativas que son independientes del tiempo y se denotan por el vector $\boldsymbol{X}=(X_1,X_2,\ldots,X_p)$, así la función de riesgo condicional es proporcional a la función de riesgo no condicionada:

$$h(t \mid X) = h(t)\cdot e^{X\beta}$$

se llama función de razón de riesgos (Hazard Risk) a la expresión:

$$HR = e^{X\beta} = \frac{h(t \mid X)}{h(t)}$$

La función HR permite comparar dos sujetos con perfiles distintos, en efecto, si se consideran los perfiles $\boldsymbol{X}$ y $\boldsymbol{X}^*$, se tiene que:

$$\frac{h(t \mid X)}{h(t \mid X^*)} = \frac{h(t)\cdot e^{X\beta}}{h(t)\cdot e^{X^*\beta}} = e^{(X-X^*)\beta}$$

Expresión que representa la razón de riesgos del perfil $\boldsymbol{X}$ sobre el perfil $\boldsymbol{X}^*$. Se deduce de inmediato que:

e^{β}, representa la razón de riesgos si los perfiles contienen componentes indicatrices y representa el cambio de la razón de riesgos por unidad de componente si ésta es una variable continua.

También es posible plantear el modelo de riesgos proporcionales a partir de la función de riesgo acumulado, H(t), que se define como:

$$H(t) = \int_0^t h(t)\cdot dt$$

En efecto, al integrar la ecuación:

$$h(t \mid X) = h(t)\cdot e^{X\beta}$$

el modelo puede plantearse como sigue:

$$H(t \mid X) = H(t)\cdot e^{X\beta}$$

Y de la relación

$$S(t) = e^{-H(t)}$$

Se deduce:

$$S(t \mid X) = e^{-H(t)\cdot e^{X\beta}} = \left[S(t)\right]^{e^{X\beta}}$$

de donde se concluye que la función de razón de riesgos se puede escribir en términos de la función de sobrevida y de las funciones de sobrevida condicionales como:

$$e^{X\beta} = \frac{\ln(S(t \mid X))}{\ln(S(t))}$$

La estimación máximo-verosímil de los parámetros del modelo se hace maximizando la función de verosimilitud:

$$\ln L = \sum_{i \text{ no censurados}} \ln(h(t_i)\cdot e^{X\beta}) - \sum_{i=1}^{n} H(t_i)\cdot e^{X\beta}$$

En general, no es posible encontrar las estimaciones de los parámetros algebraicamente y se debe usar el método de Newton-Raphson.

La teoría asintótica asegura que para n suficientemente grande los parámetros estandarizados siguen una distribución normal de media cero y varianza 1:

$$\frac{\hat{\beta}_j - \beta_j}{\sqrt{Var(\hat{\beta}_j)}} \sim \mathrm{N}(0,1)$$

Lo que permite realizar test de hipótesis sobre el HR asociado a variables de naturaleza categórica o numérica.