

OBTENCIÓN DE GRÁFICAS ALTERNATIVAS PARA FACILITAR EL USO DE LOS RESULTADOS DE UNA REGRESIÓN DE COX

José D. Bermudez*

El objeto de este informe es mostrar cómo se pueden construir dos tipos de gráficas que pueden resultar muy útiles. La primera, es una gráfica que relaciona el índice pronóstico (resultado de la ecuación de las covariables seleccionadas en el modelo con sus correspondientes coeficientes) con la probabilidad de sobrevivir a un tiempo dado. La segunda, muestra la relación entre el índice pronóstico y la mediana de supervivencia (esto es, el valor del tiempo t para el que un individuo con índice pronóstico x presenta una supervivencia del 50%).

En él se muestra como se pueden construir esos dos tipos de gráficas, utilizando los datos Mieloma. Múltiple.

Como índice pronóstico, x , vamos a considerar la combinación lineal de covariables, utilizando los coeficientes propuestos por la regresión de Cox.

De esta forma simplificamos el cálculo del índice, pues no hará falta emplear la función exponencial.

Para la construcción de las dos gráficas es necesario conocer la función de supervivencia base, $S_0(t)$, o cualquier equivalente suyo: función de azar acumulado, función de densidad, función log menos log, etc.

El SPSS no nos muestra una tabla de la función de supervivencia (sí podríamos ver su gráfica y a partir de ella, construir la tabla nosotros mismos). pero sí nos muestra una tabla de la función de azar acumulado base, $h_0(t)$, así que vamos a hacer todos los cálculos a partir de esa función.

* Documento Internet Explorer. Hospital de Sagunto.

Una vez leídos los datos de **Mieloma Múltiple**:

Estadística

- Supervivencia
- Regresión de Cox

Tiempo: meses

Status: censura (0) Covariables : logbun y hb Opciones: Ver función de línea base.

Al ejecutarlo obtendremos, entre otros resultados, la siguiente tabla:

TIME	CUM HAZARD	SURVIVAL	SE	CUM HAZARD
1	0.0084	0.9742	0.0175	0.0261
2	0.0219	0.9339	0.0284	0.0684
3	0.0266	0.9202	0.0313	0.0831
5	0.0375	0.8893	0.0371	0.1174
6	0.0622	0.8234	0.0461	0.1944
7	0.0828	0.7721	0.0515	0.2587
9	0.0906	0.7534	0.0534	0.2831
11	0.1328	0.6602	0.0599	0.4152
13	0.1426	0.6404	0.0613	0.4456
14	0.1535	0.6190	0.061	0.4797
15	0.1647	0.5976	0.064	0.5149
16	0.1887	0.5543	0.0658	0.5900
17	0.2153	0.5102	0.0671	0.6730
18	0.2296	0.4878	0.0676	0.7178
19	0.2603	0.4432	0.0680	0.8137
24	0.2777	0.4198	0.0683	0.8680
25	0.2955	0.3171	0.0683	0.9232
26	0.3139	0.3749	0.0679	0.9811
32	0.3344	0.3516	0.0675	1.0452
35	0.3561	0.3285	0.0668	1.1133
37	0.3797	0.3052	0.0657	1.1868
41	0.4049	0.2821	0.0646	1.2656
42	0.4355	0.2563	0.0633	1.3613
51	0.4714	0.2291	0.0613	1.4774

TIME	CUM HAZARD	SURVIVAL	SE	CUM HAZARD
52	0.5125	0.2015	0.0595	1.6019
54	0.5613	0.1730	0.0568	1.7546
58	0.6229	0.1427	0.0536	1.9470
66	0.7025	0.1112	0.0476	2.1960
67	0.8131	0.0787	0.0417	2.5415
88	0.9859	0.0459	0.0321	3.0816
89	1.2451	0.0204	0.0196	3.8918
92	0	0	0	0

Todo lo que necesitamos son las dos primeras columnas: la primera es el tiempo, t , y la segunda la función de azar acumulado base, $h_0(t)$, las demás columnas tienen que ver con un individuo con valores medios de covariantes, y no las vamos a necesitar para nada.

La función $h_0(t)$ es escalonada (como todas las gráficas tipo Kaplan Meier), así que su valor para un tiempo t que no aparezca en la tabla, coincide con el del tiempo más cercano e inferior a t que aparezca en tablas. Por ejemplo, para $t = 10$ esa función vale 0.0906.

Vamos a guardar esta tabla (solo las líneas de número sin el encabezamiento), copiando y pegando, en un documento solo texto que llamaremos "tabla.dar".

Gráfica 1: Relación entre índice pronóstico y la supervivencia a un tiempo fijo.

Vamos a dibujar en la misma gráfica tres curvas: una para 10 meses, otra para 40 meses y la última para 60 meses. (De la misma forma se dibujarían cualquier número de gráficas).

Los valores de la función de azar acumulado base para esos tres tiempos, obtenidos a partir de la tabla, son: 0.0906, 0.3797 y 0.6229 respectivamente.

Paso 1

Abrir un nuevo banco de datos en SPSS y crear una primera columna.

Esta columna debe tener tantos datos como puntos deseemos utilizar para el cálculo de la gráfica. Con 100 puntos suele ser ya bastante, pero si queremos más calidad en el dibujo puede que necesitemos 200 o incluso 500 puntos. Aquí vamos a utilizar 100 puntos.

No necesitamos los datos de esta columna para nada más, por lo que copiando y pegando podemos hacerla tan larga como queramos con poco esfuerzo.

Paso 2.

Crear la columna de valores de los índices pronóstico que vamos a representar.

Para ello hemos de decidir los índices mínimo y máximo a representar. Aquí vamos a representar desde $x=-2.5$ hasta $x=3.5$ (En general, vale la pena dibujar primero un rango de índices muy grande para, a la vista de gráfica resultante, decidir qué rango de valores es el más adecuado).

Transformar → Calcular

$$\text{Variable de destino: índice} = (\$casenum-1)/100 * 6-2.5$$

(100 es el número de puntos a utilizar, que coincide con el número de datos introducido en el Paso 1.6 es la distancia del x mínimo al máximo y -2.5 es el mínimo. \$casenum es una variable interna del SPSS que representa el número de individuos en el banco de datos).

Paso 3

Calcular los valores de las tres curvas a representar.

Transformar → Calcular

$$\text{Variable de destino: } t1 = \exp(-0.0906 * \exp(\text{índice}))$$

De forma similar se calculan las otras dos curvas, tan solo sustituir 0.2832 (valor de la función de azar acumulado para 10 meses) por los correspondientes valores para 40 y 60 meses:

Transformar → Calcular

$$\text{Variable de destino: } t2 = \exp(-0.3797 * (\text{índice}))$$

Transformar → Calcular

$$\text{Variable de destino : } t3 = \exp(-0.6229 * \exp(\text{índice}))$$

Paso 4

Dibujar las curvas

Gráficas → Dispersión

Superposición

Pares Y-X: t1-índice, t2-índice, t3-índice

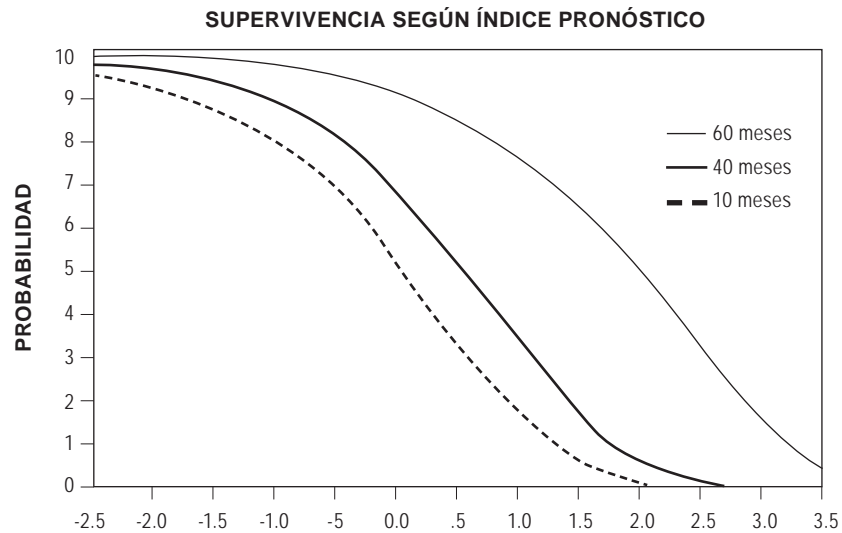
Editar la gráfica → Atributos

interpolación

Recto Marcadores: punto “.”

Luego, modificar los ejes para seleccionar la parte del plano a dibujar y modificar la leyenda.

El resultado final es la gráfica:



Paso 1

Leer tabla de azar acumulado base. Tabla que hablamos guardado en el fichero tabla.dat". con formato "solo texto"

Archivo → Leer datos ASCII

Formato de archivo : Libre

Nombre del archivo : tabla.dat

Variables definidas: t.h.v1.v2.v3.

Paso 2

Construir los valores de los índices pronósticos asociados a la función mediana.

Transformar → Calcular

Variable de destino: $x = \ln(\ln(2) / h)$

Paso 3

Dibujar la función.

Gráficas → Dispersión

Simple Eje Y:t Eje X:x

Editar la gráfica → Atributos

Interpolación: Salto a la derecha

Marcadores: punto “.”

Luego, modificar los ejes para seleccionar la parte del plano a dibujar, y modificar la leyenda.

El resultado final es la gráfica

