

Sobrevivencia de pacientes con cáncer según información genética

La base de datos contiene información sobre pacientes con cáncer de mama y ciertas características asociadas al paciente, tumor y características del perfil celular de éste. Algunos científicos mantienen la fe de que la clave para el cáncer se encuentra en el material genético.

Los datos se encuentran en el siguiente sitio web:

[METABRIC \(Molecular Taxonomy of Breast Cancer International Consortium\)](#)

Interesa estudiar el tiempo de sobrevivencia (en meses) de las pacientes, una vez que se le aplicó cirugía para tratar el cáncer de mama. Y considere como evento de interés el fallecimiento con motivo el cáncer.

a) Cargue la base de datos, asegúrese de no tener observaciones faltantes. La variable `Overall Survival (Months)` corresponde a y_i en meses y `Patient's Vital Status` indica el estado vital de la persona finalizado su seguimiento. Defina la censura como corresponda en este estudio. ¿Cuántos pacientes (considerando el período de observación/seguimiento) fallecieron con motivo de cáncer?

b) Las células de cáncer de mama con receptor hormonal positivo (o con hormonas positivas) tienen receptores de estrógeno (ER) o progesterona (PR) o ambos. Es posible detectar cierta proteína denominada receptor del factor de crecimiento epidérmico 2 (HER2). Los perfiles o combinaciones son ER+/HER2- (baja o alta proliferación), HER2+ y ER-/HER2-.

La variable `3-Gene classifier subtype` indica este perfil. Obtenga una tabla que permita observar cómo se distribuyen los fallos y no fallos dependiendo de los subtipos de cáncer de mama. Comente. Plantee un modelo Weibull con el subtipo celular como predictor. Utilice la función `WeibullReg` de la librería `SurvRegCensCov`. Muestre las estimaciones de la hazard acumulada logarítmica para cada subtipo de cáncer de mama. Comente.

c) Dados los resultados anteriores, se decide utilizar otras variables predictoras. Utilice la edad (`Age of diagnosis`) y el tamaño del tumor (`Size Tumor`) como variables predictoras. ¿Cómo varía la curva de sobrevivencia estimada bajo un modelo Weibull, considerando un tamaño de tumor fijo de 15 para las edades 20, 40, 60 y 80? Realice lo mismo pero fijando ahora la edad y probando diferentes tamaños de tumor. Comente.