Tarea 4

Leobardo Enriquez

2024-04-28

4.5 Ejercicios

1. Se tiene una variable aleatoria X con función de densidad $f(x) = (1 + 2x^2)e^{-2x}$ $x > 0$

a) Determine la función de supervivencia $S(x)$

La función de supervivencia $S(t)$, tanto en el caso continuo como en el discreto, se define como la probabilidad de que un individuo sobreviva más allá del tiempo t. Para el caso continuo:

$$S(t) = \mathbb{P}(T > t) = 1 - F_T(t) = \int_t^\infty f_T(u)du$$

Para la variable aleatoria X dada, planteamos:

$$S(x) = \int_x^\infty f_X(u)du = \int_x^\infty (1 + 2u^2)e^{-2u}du = \int_x^\infty e^{-2u}du + \int_x^\infty 2u^2e^{-2u}du =$$

$$= -\frac{e^{-2u}}{2}|_x^\infty + 2[-\frac{1}{2}e^{-2u}u^2|_x^\infty - \int_x^\infty -e^{-2u}udu] = -\frac{e^{-2u}}{2}|_x^\infty - e^{-2u}u^2|_x^\infty - 2\int_x^\infty -e^{-2u}udu =$$

$$= -\frac{e^{-2u}}{2}|_x^\infty - e^{-2u}u^2|_x^\infty - 2[-\frac{1}{4}(-2e^{-2u}u|_x^\infty - \int_x^\infty -2e^{-2u}(1))]$$

$$= -\frac{e^{-2u}}{2}|_x^\infty - e^{-2u}u^2|_x^\infty - 2[-\frac{1}{4}(-2e^{-2u}u|_x^\infty - (e^{-2u}|_x^\infty)] =$$

$$= -\frac{e^{-2u}}{2}|_x^\infty - e^{-2u}u^2|_x^\infty - e^{-2u}u|_x^\infty - \frac{1}{2}e^{-2u}|_x^\infty = -0 + \frac{e^{-2x}}{2} - 0 + e^{-2x}x^2 - 0 + e^{-2x}x - 0 + \frac{1}{2}e^{-2x} =$$

$$= \frac{e^{-2x}}{2} + e^{-2x}x^2 + e^{-2x}x + \frac{1}{2}e^{-2x} = e^{-2x}(x^2 + x + 1)$$

b) Determine la función de riesgo $h(x)$

La función de riesgo $h(t)$ (hazard function), también llamada **tasa de falla condicional** (en el análisis de confiabilidad) o *tasa de mortalidad* (en demografía), se define como la probabilidad de falla durante un intervalo de tiempo muy pequeño suponiendo que el individuo ha sobrevivido hasta el inicio del intervalo (aunque en la definición de $h(t)$ se tenga explícitamente la palabra “probabilidad”, hay que tener en claro que esta función no es una función de probabilidad, si no tal cual una **tasa**, ya que la acumulación de esta puede dar valores superiores a 1).

$$h(t) = \frac{f(t)}{S(t)} = \frac{(1 + 2x^2)e^{-2x}}{e^{-2x}(x^2 + x + 1)} = \frac{1 + 2x^2}{x^2 + x + 1}$$

c) Determine la función de media residual $mr(x)$

Para individuos de edad x este parámetro, denotado por $mr(x)$, mide la esperanza de vida residual; esto es, "la esperanza de vida que les queda a partir de la edad x".

Se define el caso contínuo como:

$$mr(x) = \frac{\int_x^\infty S(t)dt}{S(x)} = \frac{\int_x^\infty e^{-2t}(t^2+t+1)dt}{e^{-2x}(x^2+x+1)} = \frac{\int_x^\infty t^2e^{-2t}dt + \int_x^\infty te^{-2t}dt + \int_x^\infty e^{-2t}dt}{e^{-2x}(x^2+x+1)} =$$

$$\frac{(-\frac{1}{2}e^{-2t}t^2 - \frac{1}{2}e^{-2t}t - \frac{1}{4}e^{-2t})|_x^\infty + (\frac{1}{4}(-2e^{-2t}t - e^{-2t})|_x^\infty) - (\frac{1}{2}e^{-2t}|_x^\infty)}{e^{-2x}(x^2+x+1)} =$$

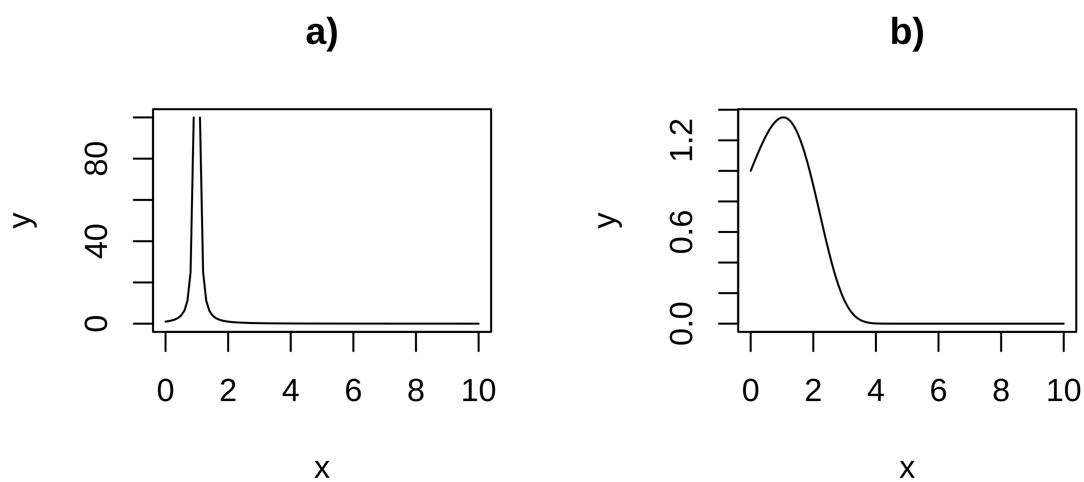
$$= \frac{(-\frac{1}{2}e^{-2t}t^2 - \frac{1}{2}e^{-2t}t - \frac{1}{4}e^{-2t} - \frac{1}{2}e^{-2t}t - \frac{1}{4}e^{-2t} - \frac{1}{2}e^{-2t})|_x^\infty}{e^{-2x}(x^2+x+1)} = \frac{(-\frac{1}{2}e^{-2t}t^2 - e^{-2t}t - e^{-2t})|_x^\infty}{e^{-2x}(x^2+x+1)} =$$

$$= \frac{(-0-0-0) - (-\frac{1}{2}e^{-2x}x^2 - e^{-2x}x - e^{-2x})}{e^{-2x}(x^2+x+1)} = \frac{\frac{1}{2}e^{-2x}x^2 + e^{-2x}x + e^{-2x}}{e^{-2x}(x^2+x+1)} = \frac{1}{2}$$

2. De los siguientes ¿cuál(es) no podría servir como un modelo de supervivencia?

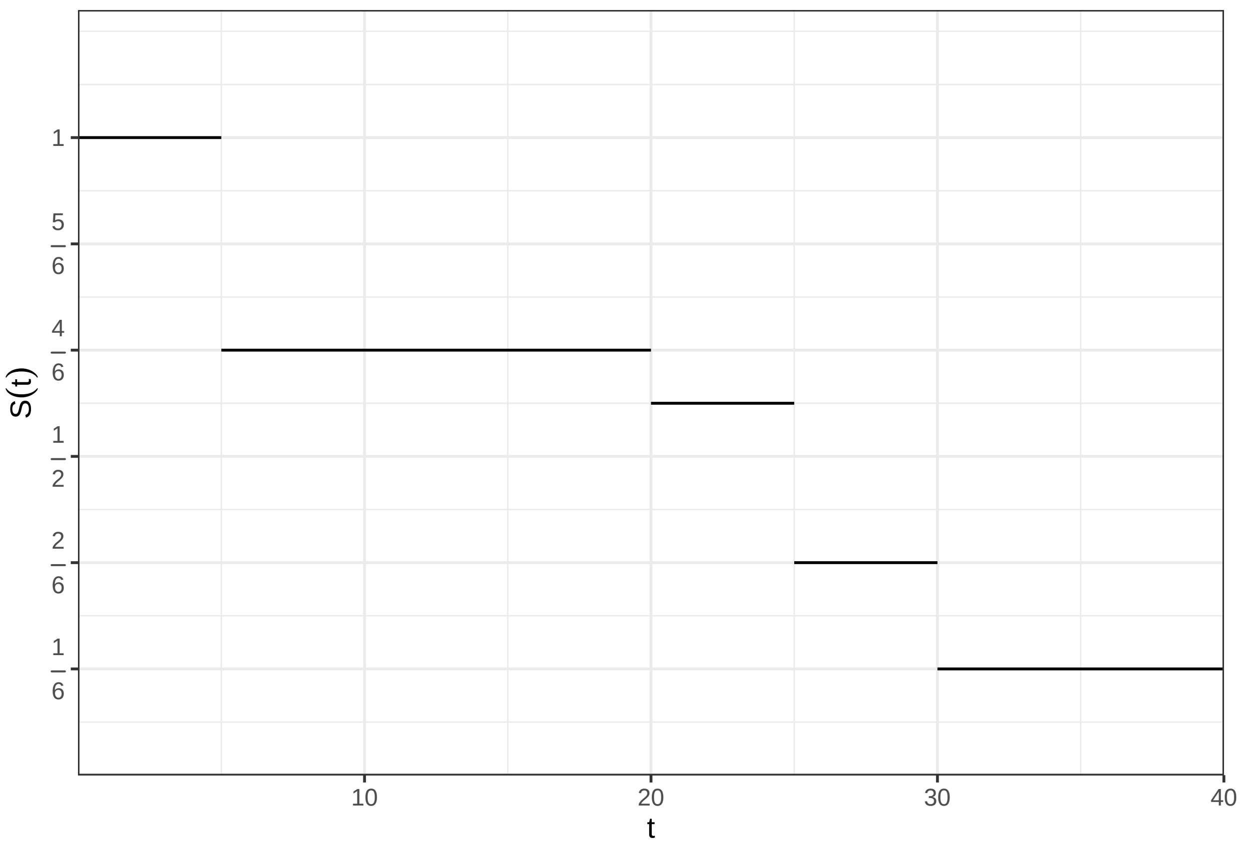
a) $S(t) = e^{\left[t-0.7\left(2^t-1\right)\right]}$

b) $S(t) = (1-t)^{-2}$



Observamos que ninguno de los dos representa un modelo de supervivencia.

3. Suponga que la curva de supervivencia estimada de un conjunto de datos es la siguiente:



a) Calcule la esperanza de vida residual para $t = 15$.

$$mr(15) = \frac{\sum_{k=15}^{\infty} S(u_k)}{S(15)} = \frac{4/6 + 7/12 + 2/6 + 1/6}{4/6} = \frac{21/12}{4/6} = \frac{21}{8} = 2.625$$

b) Calcule el cuantíl correspondiente a una supervivencia del 80%

$$t_p = inf\{t : S(t) \leq 1 - p\}$$
$$t_{0.8} = inf\{t : S(t) \leq 0.2\} = 30$$

4. Suponga que Z tiene la siguiente función de riesgo $h(z) = \frac{1}{100-z}$, $0 < z < 100$

a) Determine la función de supervivencia $S(x)$

La función de riesgo acumulado:

$$H(t) = \int_0^t h(u)du = \int_0^t \frac{1}{100-u}du = (-ln|u-100|)|_0^t = -ln|t-100| + ln|-100| = ln(100) - ln|t-100|$$

$$S(t) = e^{-H(t)} = e^{ln(100)-ln|t-100|} = e^{ln(100)}e^{-ln|t-100|} = 100(-|t-100|) = -100|t-100|$$

b) Determine la función de distribución $F(x)$

$$F(t) = 1 - S(t) = 1 - (-100|t-100|) = 1 + 100|t-100| = 1 - 100(t-100) = -100t + 1001$$

c) Determine la función de probabilidad $f(x)$

$$f(t) = h(t)S(t) = \frac{1}{100-t}(-100|t-100|) = \frac{-100|t-100|}{100-t} = \frac{-100|100-t|}{100-t} = -100$$

$$f(t) = -\frac{dS(t)}{dt} = -[\frac{d}{dt}(-100(-(t-100)))] = -[\frac{d}{dt}(100t-10000)] = -100$$

6.8 Ejercicios

1. El tiempo de vida en años de una batería de marca paso sigue una distribución *Pareto* centrada con parámetros $\theta = 4$, $\lambda = 5$

a) ¿Cuál es la probabilidad de que la batería funcione por lo menos 10 años?

b) ¿Cuál es el tiempo de vida medio de la batería?

c) Encontrar el tiempo de reemplazo t_r para el cual 90% de las baterías no hayan fallado.

2. El tiempo de vida (en años) X de una bacteria que se encuentra en el lago "Gran Ojo" sigue una distribución "mezcla". Calcule la probabilidad de que el tiempo de vida de la bacteria sea menor a 200 años.

a) Con probabilidad 0.8, X tiene una distribución *Pareto* centrada con parámetros $\alpha = 2$ y $\theta = 100$.

b) Con probabilidad 0.2, X tiene una distribución *Pareto* centrada con parámetros $\alpha = 4$ y $\theta = 3000$.

3. Considere una distribución Weibull cuya función de de supervivencia está dada por $S(x) = exp(-\lambda x^{\alpha})$ con $\lambda > 0$ y $\alpha > 0$. Encuentre la función de verosimilitud bajo los siguientes casos:

a) Datos truncados por la izquierda (y_{l_i}, x_i), $y_{l_i} \leq x_i$, $i = 1, \ldots, n$ donde y_{l_i} es un tiempo de truncamiento por la izquierda (simplifique la expresión).

b) Datos censurados por intervalos (L_i, R_i), $i = 1, \ldots, n$.

c) Datos doblemente truncados (y_{l_i}, x_i, y_{r_i}), donde $y_{l_i} \leq x_i \leq y_{r_i}$, $i = 1, \ldots, n$ (simplifique la expresión).

d) n = 4 pacientes cuya edad de ocurrencia del evento son los siguientes intervalos $(90, 120]$, $(110, 115]$, $(80, 100]$, $(70, 75]$, sujetos a la condición de entrada $edad \geq 50$.