Tarea 4

Funciones en análisis de supervivencia

Modelos de series de Tiempo y Supervivencia

Profesor: Naranjo Albarrán Lizbeth

Adjuntos: Reyes González Belén

Rivas Godoy Yadira

Integrantes: Cuéllar Chávez Eduardo de Jesús

García Tapia Jesús Eduardo

Miranda Meraz Areli Gissell

Ramírez Maciel José Antonio

Saldaña Morales Ricardo

Grupo: 9249

Fecha: 02/DIC/2021

1. Resuelva lo siguiente:

a) Suponga que la función de riesgo de asociada a un tiempo de supervivencia es una función lineal $h(t) = a + bt$ donde \$ a > 0\$ y $b \in \mathbb{R}$. Obtenga: $S(t)$, $F(t)$, $f(t)$, $h(t)$, y la media, mediana y moda de la distribución. Grafique las funciones.

b) Supongo que una función de supervivencia está definida por $S(t) = exp(-t^{\gamma})$ para $0 \leq t$. Obtenga su función de densidad $f(t)$ y su función de riesgo $h(t)$.

Solución

a) Sea $h(t) = a + bt$

Como

$$
\begin{aligned}
H(t) &= \int_0^t h(u)du \\
\implies H(t) &= \int_0^t a + budu = au|_0^t + \frac{bu^2}{2}|_0^t = at + \frac{bt^2}{2}
\end{aligned}
$$

Veamos $S(t)$,

$$S(t) = e^{\int_0^t h(u)du} = e^{-H(t)} = e^{-(at+\frac{bt^2}{2})}$$

Ahora,veamos $F(t)$, como

$$
\begin{aligned}
S(t) &= 1 - F(t) \\
\implies F(t) &= 1 - S(t) = 1 - e^{-(at+\frac{bt^2}{2})}
\end{aligned}
$$

Por otro lado, como

$$
\begin{aligned}
&F'(t) = f(t) \\
&\implies \frac{d\left(1 - e^{-\left(at+\frac{bt^2}{2}\right)}\right)}{dt} = f(t) = -\left(-\left(a + \frac{2bt}{2}\right)e^{\left(at+\frac{bt^2}{2}\right)}\right) = (a+bt)e^{-\left(at+\frac{bt^2}{2}\right)}
\end{aligned}
$$

De manera alternativa,

$$f(t) = h(t)e^{\int_0^t h(u)du} = h(t)S(t) = (a+bt)e^{-\left(at+\frac{bt^2}{2}\right)}$$

Veamos la media:

$$\int_0^{\infty} t(a+bt)e^{-\left(at+\frac{bt^2}{2}\right)dt}$$

Haciendo cambio de variable, sea $u = \frac{bt^2}{2} + at \implies du = (a+bt)dt, \quad$ despejando t

$$
\begin{aligned}
\frac{bt^2}{2} + at - u = 0 \implies & \frac{a + \sqrt{a^2 - 4\left(\frac{b}{2}\right)(-u)}}{2\left(\frac{b}{2}\right)} = x \\
& \text{(Nos quedamos con el signo positivo por el soporte)} \\
& \implies \frac{a + \sqrt{a^2 + 2bu}}{b} = x \\
& \implies \mathbb{E}(X) = \int_0^\infty \left[\frac{a + \sqrt{a^2 + 2bu}}{b}\right] e^{-u} du \\
& = \frac{a}{b}\int_0^\infty e^{-u} du + \frac{1}{b}\int_0^\infty \sqrt{a^2 + 2bu} e^{-u} du \\
& = -\frac{a}{b} e^{-u}|_0^\infty + \frac{1}{b}\int_0^\infty \sqrt{a^2 + 2bu} e^{-u} du \\
& = -\frac{a}{b}\left[\lim_{u\to\infty} e^{-u} - e^0\right] + \frac{1}{b}\int_0^\infty \sqrt{a^2 + 2bu} e^{-u} du \\
& = -\frac{a}{b}[0 - 1] + \frac{1}{b}\int_0^\infty \sqrt{a^2 + 2bu} e^{-u} du \\
& = \frac{a}{b} + \frac{1}{b}\int_0^\infty \sqrt{a^2 + 2bu} e^{-u} du \quad (1)
\end{aligned}
$$

Sea

$$
\begin{aligned}
y &= \frac{(a^2 + 2bu)^{\frac{3}{2}}}{3b} \\
\implies dy &= \sqrt{a^2 + 2bu} du \\
\implies 3by &= (a^2 + 2bu)^{\frac{3}{2}} \\
\implies (3by)^{\frac{3}{2}} &= a^2 + 2bu \\
\implies \frac{(3by)^{\frac{3}{2}} - a^2}{2b} &= u
\end{aligned}
$$

Dado lo anterior, continuamos desarrollando donde nos quedamos en (1)

$$
\begin{aligned}
&= \frac{1}{b}\int_{\frac{a^3}{b}}^\infty e^{-\left[\frac{(3by)^{\frac{3}{2}} - a^2}{2b}\right]} dy \\
&= \frac{1}{b}\int_{\frac{a^3}{b}}^\infty e^{-\frac{(3by)^{\frac{3}{2}}}{2b}} e^{\frac{a^2}{2b}} dy \\
&= \frac{1}{b} e^{\frac{a^2}{2b}} \int_{\frac{a^3}{b}}^\infty e^{-\frac{(3by)^{\frac{3}{2}}}{2b}} dy \\
\therefore \mathbb{E}[X] &= \frac{a}{b} + \frac{1}{b} e^{\frac{a^2}{2b}} \int_{\frac{a^3}{b}}^\infty e^{-\frac{(3by)^{\frac{3}{2}}}{2b}} dy
\end{aligned}
$$

Lo dejamos expresado solamente, debido a que no se puede integrar.

Ahora veamos la moda:

**

$$
\begin{aligned}
&(a+bt)e^{-\left(at+\frac{bt^2}{2}\right)} \\
&= ae^{-\left(at+\frac{bt^2}{2}\right)} + bte^{-\left(at+\frac{bt^2}{2}\right)} \\
\frac{df(t)}{df} &= a(-a-bt)e^{-(at+\frac{bt^2}{2})} + b\left(e^{-at+\frac{bt^2}{2}} + t\left[(-a-bt)e^{\left(at+\frac{bt^2}{2}\right)}\right]\right) \\
&= -(a^2+abt)e^{-(at+\frac{bt^2}{2})} + be^{-at+\frac{bt^2}{2}} - bt(a+bt)e^{-(at+\frac{bt^2}{2})} \\
&= e^{-(at+\frac{bt^2}{2})}(-a^2-abt+b-bat-b^2t^2) \\
&= -e^{-(at+\frac{bt^2}{2})}(a^2+abt-b+bat+b^2t^2) \\
&= -e^{-(at+\frac{bt^2}{2})}(a^2+2abt+b^2t^2-b) \quad (2)
\end{aligned}
$$

Igualando a cero, como $e^{-x} > 0 \quad \forall x$;

$$
\begin{aligned}
(2) = 0 \quad \text{si y solo si} \quad & a^2+2abt+b^2t^2-b = 0 \\
\implies x = & \frac{-2ab \pm \sqrt{(2ab)^2 - 4(b^2)(a^2-b)}}{2b^2} \\
= & -\frac{a}{b} \pm \frac{1}{2b}\sqrt{4a^2b^2 - 4a^2b^2 + 4b} \\
= & -\frac{a}{b} + \frac{1}{2b}\sqrt{4b} \\
t = & -\frac{a}{b} + \frac{2}{2b}\sqrt{b} \quad \text{si b} > 0 \\
t = & -\frac{a}{b} + \frac{1}{\sqrt{b}} > 0 \\
& \text{si y solo si} \quad \frac{1}{\sqrt{b}} > \frac{a}{b} \\
& \text{si y solo si} \quad \sqrt{b} > a \\
& \text{si y solo si} \quad b > a^2 \\
& a^2+2abt+b^2t^2-b > 0
\end{aligned}
$$

**

$$
\begin{aligned}
&- e^{-\left(at+\frac{bt^2}{2}\right)}(a^2+2abt+b^2t^2-b)\\
&= a^2(a+bt)e^{-\left(at+\frac{bt^2}{2}\right)} + 2ab\left(t(a+bt)e^{-(at+\frac{bt^2}{2})} - e^{-\left(at+\frac{bt^2}{2}\right)}\right)\\
&+ b^2\left(t^2(a+bt)e^{-\left(at+\frac{bt^2}{2}\right)} - 2te^{-\left(at+\frac{bt^2}{2}\right)}\right) - b(a+bt)e^{-\left(at+\frac{bt^2}{2}\right)}\\
&= e^{-\left(at+\frac{bt^2}{2}\right)}\left(a^3+a^2bt+2a^2bt+2ab^2t^2-2ab+ab^2t^2+b^3t^3-t-ab-b^2t\right)
\end{aligned}
$$

**

$$= a^3+2a^2bt+2a^2bt+3ab^2t^2+b^3t^3-3b^2t*3ab$$

**

$$
\begin{aligned}
&= (a+bt)(a^2+2abt+b(bt^2-3))\\
t &= -\frac{a}{b}+\frac{1}{\sqrt{b}} \implies bt = -a+\sqrt{b} \implies 2abt = -a^2+2a\sqrt{b}\\
b^2t^2 &= b^2\left(\frac{a^2}{b^2}-\frac{2a}{b^{\frac{3}{2}}}+\frac{1}{b}\right) = a^2-2a\sqrt{b}+b\\
&= (a+\sqrt{b}-a)(a^2-2a^2+2a\sqrt{b}+a^2-2a\sqrt{b}+b-3b)\\
&= (\sqrt{b})(-2b) = -2b^{\frac{3}{2}} < 0\\
&\therefore \text{Es un máximo como *}
\end{aligned}
$$

$$
\begin{aligned}
&a^2+2abt+b^2t^2-b < a^2+2a\\
&- a^2-2abt-b(bt^2-1)\\
&< -2abt-b^2t^2+b\\
&< -2abt+b\\
&< -2abt\\
&< 0
\end{aligned}
$$

Entonces, si $b < 0$ siempre es decreciente, por lo que su máximo es en $t = 0$ (donde empieza). Lo mismo si $b < a^2$; ya que alcanza su máximo en un t < 0; Por lo que de ahí decrece, entonces su máximo sería en $t = 0$.

$$
\begin{gathered}
(2b^{\frac{3}{2}}-a)(a^2-4ab^{\frac{3}{2}}+4b^3-3b)\\
2a^2b^{\frac{3}{2}}-8ab^3+8b^{\frac{a}{2}}-6b^{\frac{5}{2}}-a^3+4a^2b^{\frac{3}{2}}-4ab^3+3ab\\
ba^2b^{\frac{3}{2}}-12ab^3+8b^{\frac{a}{2}}-6b^{\frac{5}{2}}+3ab-a^3
\end{gathered}
$$

$$
\begin{gathered}
ln(0.5) = -at\\
-a^{-1}ln(0.5) = t
\end{gathered}
$$

$$
\begin{aligned}
1-e^{-(at+\frac{bt^2}{2})}=.05 &\Leftrightarrow 0.5=e^{-}at+\frac{bt^2}{2}\\
&\Leftrightarrow ln(.05)=-at-\frac{bt^2}{2}\\
&\Leftrightarrow -ln(0.5)=at+\frac{bt^2}{2}\\
&\Leftrightarrow \frac{b}{2}\left(t^2+\frac{2at}{b}\right)=-ln(0.5)\\
&\Leftrightarrow \frac{b}{2}\left(t^2+\frac{2at}{b}+\left(\frac{a}{b}\right)^2-\left(\frac{a}{b}\right)^2\right)=ln(2)\\
&\Leftrightarrow \frac{b}{2}\left(\left(t+\frac{a}{b}\right)^2-\left(\frac{a}{b}\right)^2\right)=ln(2)\\
&\Leftrightarrow \left(t+\frac{a}{b}\right)^2-\left(\frac{a}{b}\right)^2=\frac{2ln(2)}{b}\\
&\Leftrightarrow \left(t+\frac{a}{b}\right)^2=\frac{2ln(2)}{b}+\frac{a^2}{b^2}\\
&\Leftrightarrow t+\frac{a}{b}=\sqrt{\frac{2bln(2)+a^2}{b^2}}\\
&\Leftrightarrow t=\frac{\sqrt{a^2+2bln(2)}b}{b}-\frac{a}{b}\\
&\Leftrightarrow t=\frac{-a+\sqrt{a^2+2bln(2)}}{b}\\
\therefore \quad \text{La media es} \quad t=\frac{-a+\sqrt{a^2+2bln(2)}}{b}
\end{aligned}
$$

Ahora, como $t>0$, solo está definido si:

$$
-a+\sqrt{a^2+2bln(2)}>0 \Leftrightarrow \sqrt{a^2+2bln(2)}>a
$$

Pero, $ln(2)>0$; $2>0$, entonces necesitamos que $b>0$!, ya que, de lo contrario $\sqrt{a^2+2bln(2)}<a$. Ahora, si $b=0$, se reduce a una exponencial con $\lambda=a$. Si $b<0$, el soporte debería ser $(0,\frac{a}{-b})$; ya que si $b<0$ y $t>\frac{a}{-b}$; $at+\frac{bx^2}{2}$ es decreciente.

Veamos que:

$$
\begin{aligned}
at+\frac{bt^2}{2}=0 &\Leftrightarrow t(a+\frac{bt}{2})=0\\
&\Leftrightarrow t=0 \quad \text{ó} \quad a+\frac{bt}{2}=0\\
&\Leftrightarrow \frac{bt}{2}=-a\\
&\Leftrightarrow t=\frac{-2a}{b}
\end{aligned}
$$

b) Sabemos que $S(t)=exp(-t^{\gamma})$, entonces obtendremos las funciones que se nos piden.

Función de densidad:

$$
\begin{aligned}
S(t) = exp(-t^{\gamma}) &\implies F(x) = 1 - e^{r^{\gamma}} \\
&\implies f(x) = -e^{-t^{\gamma}} * \left(-\gamma t^{\gamma-1}\right) \\
&= \gamma t^{\gamma-1} e^{-t^{\gamma}} \\
&\therefore \quad \text{la función de densidad es} = \gamma t^{\gamma-1} e^{-t^{\gamma}} \blacksquare \\
&\text{Ahora obtenemos la función de riesgo} \\
&\implies h(t) = \frac{\gamma t^{\gamma-1} e^{-t^{\gamma}}}{e^{-t^{\gamma}}} \\
&= \gamma t^{\gamma-1} \\
&\therefore \quad \text{la función de riesgo es} = \gamma t^{\gamma-1} \blacksquare
\end{aligned}
$$

2. Resume las siguientes distribuciones: Exponencial, Weibull, Log-Normal, Gamma, Gompertz, Log-Logística, Geométrica. (Hint: Ver capítulo 7 del libro: Kleinbaum, D. & Klein, M. (2005) Survival Analysis. A Self-Learning Text. Springer.)

a) ¿Qué distribuciones describen una tasa de riesgo constante?
b) ¿Qué distribuciones describen una tasa de riesgo creciente? Identifica sus diferencias.
c) ¿Qué distribuciones describen una tasa de riesgo decreciente? Identifica sus diferencias.
d) Con R grafique las funciones $S(t)$ y $h(t)$ dando valores fijos para los parámetros de las funciones. (Hint: Revisa la sección 2.4 del libro: Moore, D.F. (2016) Applied Survival Analysis Using R. Use R! Springer.)

Solución

De acuerdo con la bibliografía consultada y lo visto en clases sería:

a) Exponencial, Weibull con $\rho = 1$, la geométrica y la Gamma.

Exponencial; con

$$
\begin{aligned}
S(t) &= e^{\lambda x} \\
f(x) &= \lambda e^{-\lambda x} \\
h(t) = \frac{f(x)}{S(t)} &\implies h(t) = \frac{\lambda e^{\lambda x}}{e^{-\lambda x}} \\
&= \lambda
\end{aligned}
$$

Geométrica con:

$$
\begin{aligned}
S(t) &= (1-\rho)^{K-1} \quad \text{(para la que tiene el soporte en } \mathbb{N} \setminus \{0\}) \\
f(t) &= (1-\rho)^{K-1} \rho \\
\implies h(t) &= \frac{(1-\rho)^{K-1}\rho}{(1-\rho)^{K-1}} = \rho \to \text{constante} \\
S(t) &= (1-\rho)^{K} \quad \text{(para la que tiene el soporte en } \mathbb{N} \\
f(t) &= (1-\rho)^{K} \rho \\
\implies h(t) &= \frac{(1-\rho)^{K}\rho}{(1-\rho)^{K}} = \rho \to \text{constante}
\end{aligned}
$$

Por lo que la exponencial es buena para modelar riesgos constantes continuas, y la geométrica riesgos constantes discretos. Hay que aclarar que la Exponencial es un caso particular de la Weibull (Donde $\rho = 1$) y, como el libro nos comenta que la Weibull es un caso particular de la Gamma Generalizada, por ende, la exponencial lo es.

b) La Weibull, la Log-Normal, la Log-Logistica, La Gompertz y La Gamma.

La Weibull cuando $\gamma > 1$, con

$$\begin{aligned} S(t) &= e^{\lambda t^{\gamma}} \\ f(t) &= \gamma\lambda t^{\gamma-1}e^{-\lambda t^{\gamma}} \\ \implies h(t) &= \frac{\gamma\lambda t^{\gamma-1}e^{-\lambda t^{\gamma}}}{e^{-\lambda t^{\gamma}}} = \gamma\lambda t^{\gamma-1} \end{aligned}$$

Como $\gamma > 1 \implies \gamma - 1 > 0 \implies t^{\gamma-1}$ es creciente.\

La Log-Normal

$$S(t) = 1 - \phi\left(\frac{ln(t)-\mu}{\theta}\right)$$

donde $\phi(k)$ es la función de distribución de una $Normal(0,1)$ evaluada en $x = k$

$$h(t) = \frac{\frac{1}{\sqrt{2\pi\sigma^2}t}e^{-\left(\frac{(ln(t)-\mu)^2}{2\sigma^2}\right)}}{1-\phi\left(\frac{ln(t)-\mu}{\theta}\right)}$$

Aquí es "parcialmente" ya que, en clase vimos que cuando t es pequeña, comienza muy cerca de cero y va creciendo hasta alcanzar un máximo, para después decrecer.

La Log-logística con $\gamma > 1$; de manera parcial. Ya que crece hasta alcanzar un máximo, pero luego decae

$$\begin{aligned} S(t) &= \frac{1}{1+\lambda t^{\gamma}} \\ h(t) &= \frac{\gamma\lambda t^{\gamma-1}}{1+\lambda t^{\gamma}} \end{aligned}$$

La Gompertz cuando $\varphi > 1$; inicia cerca de cero y eventualmente crece hasta infinito

$$\begin{aligned} S(t) &= e^{-\left(\frac{\lambda}{ln(\varphi)}(\varphi^{t}-1)\right)} \\ h(t) &= \lambda\varphi^{t} \end{aligned}$$

Gamma, cuando $K > 1$

$$S(t) = 1 - GI(K, \lambda t)$$

Donde $GI(K,\lambda t) = \frac{1}{\Gamma(K)}\int_0^{\lambda t} u^{K-1}e^{-u}du$

$$h(t) = \frac{\frac{\lambda^K t^{K-1}e^{-\lambda t}}{\Gamma(k)}}{1-GI(K,\lambda t)}$$

Los riesgos son nulos en un inciso, para positivamente crecer hasta hacerse constantes (Ya que en elm límite en el límite converge a α)

c) La log-logística, la Weibull, la log-normal, la Gompertz y la Gamma.

La log-logística, con $\gamma \leq 1$; Aunque si $\gamma > 1$ (primero crece muy rápido hasta alcanzar un máximo para después decrecer).

$$S(t) = \frac{1}{1 + \lambda t^{\gamma}}$$
$$h(t) = \frac{\gamma \lambda t^{\gamma - 1}}{1 + \lambda t^{\gamma}}$$

La Weibull cuando $\gamma < 1$ con

$$S(t) = e^{-\lambda t^{\gamma}}$$
$$f(t) = \gamma \lambda t^{\gamma - 1} e^{-\lambda t^{\gamma}}$$
$$\implies h(t) = \frac{\gamma \lambda t^{\gamma - 1} e^{-\lambda t^{\gamma}}}{e^{-\lambda t^{\gamma}}} = \gamma \lambda t^{\lambda - 1}$$

como $\gamma < 1 \implies \gamma - 1 < 0 \implies t^{\gamma - 1}$ es creciente.

La Log-Normal

$$S(t) = 1 - \phi\left(\frac{ln(t) - \mu}{\theta}\right)$$

donde $\phi(k)$ es la función de distribución de una $Normal(0,1)$ evaluada en $x = K$

$$h(t) = \frac{\frac{1}{\sqrt{2\pi\sigma^2}t} e^{-\left(\frac{(ln(t) - \mu)^2}{2\sigma^2}\right)}}{1 - \phi\left(\frac{ln(t) - \mu}{\theta}\right)}$$

Cuando $\gamma \leq 1$ es monótona decreciente;\ Cuando $\gamma > 1$; es "parcialmente" ya que, en clase vimos que:

Cuando t es pequeña, comienza muy cerca de cero y va decreciendo hasta alcanzar un máximo, para después decrecer

La Gompertz cuando $\varphi > 1$; modela riesgos que

$$S(t) = e^{-\left(\frac{\lambda}{ln(\varphi)}(\varphi^t - 1)\right)}$$
$$h(t) = \lambda \varphi^t$$

La Gamma, cuando $K < 1$

$$S(t) = 1 - GI)(K, \lambda t)$$

Donde $GI(K, \lambda t) = \frac{1}{\Gamma(K)} \int_0^{\lambda t} u^{K-1} e^{-u} du$

$$h(t) = \frac{\frac{\lambda^K t^{K-1} e^{-\lambda t}}{\Gamma(k)}}{1 - GI(K, \lambda t)}$$

Los riesgos son grandes en un inicio, para posteriormente decrecer hasta hacerse constantes (Ya que el límite converge a α).

d)