Survival Analysis HW 1

Mario Ibanez

January 18, 2016

Problem 2.5

A model for lifetimes, with a bathtub-shaped hazard rate, is the exponential power distribution with survival function $S(x) = exp(1 - exp((\lambda x)^\alpha))$.
a) If $\alpha = 0.5$, show that the hazard rate has a bathtub shape and find the time at which the hazard rate changes from decreasing to increasing.
b) If $\alpha = 2$, show that the hazard rate of x is monotone increasing.

Answer

First we find the hazard function:

$$\begin{aligned} h(x) &= \frac{-S'(x)}{S(x)} \\ &= \frac{exp(1 - exp((\lambda x)^\alpha)) * e^{(\lambda x)^\alpha} \lambda^\alpha \alpha x^{\alpha-1}}{exp(1 - exp((\lambda x)^\alpha))} \\ &= \alpha \lambda^\alpha x^{\alpha-1} e^{(\lambda x)^\alpha} \end{aligned}$$

and

$$\begin{aligned} h'(x) &= \alpha\lambda^\alpha \left[(\alpha-1)x^{\alpha-2} exp((\lambda x)^\alpha) + x^{\alpha-1} exp((\lambda x)^\alpha)\alpha(\lambda x)^{\alpha-1}\lambda\right] \\ &= \alpha\lambda^\alpha exp((\lambda x)^\alpha)\left[(\alpha-1)x^{\alpha-2} + x^{\alpha-1}\alpha(\lambda x)^{\alpha-1}\lambda\right] \end{aligned}$$

It's worth noting that in this distribution, $\lambda > 0$ is the scale parameter, $\alpha > 0$ is the shape parameter, and the support is $x > 0$.

a) Let $\alpha = 0.5$. Then:

$$\begin{aligned} h'(x) &= \frac{1}{2}\sqrt{\lambda} e^{\sqrt{\lambda x}} \left[\frac{-1}{2} x^{-3/2} + x^{-1/2}\frac{1}{2}(\lambda x)^{-1/2}\lambda\right] \\ &= \frac{e^{\sqrt{\lambda x}}\sqrt{\lambda}}{2}\left[\frac{\lambda}{2\sqrt{x}\sqrt{\lambda}\sqrt{x}} - \frac{1}{2x\sqrt{x}}\right] \\ &= \frac{e^{\sqrt{\lambda x}}\sqrt{\lambda}}{2}\left[\frac{\sqrt{\lambda}}{2x} - \frac{1}{2x\sqrt{x}}\right] \\ &= \frac{e^{\sqrt{\lambda x}}\sqrt{\lambda}}{2}\left[\frac{\sqrt{\lambda x} - 1}{2x\sqrt{x}}\right] \end{aligned}$$

This is the product of two terms. The term

$$\frac{e^{\sqrt{\lambda x}}\sqrt{\lambda}}{2}$$

is always positive. The other term

$$\frac{\sqrt{\lambda x}-1}{2x\sqrt{x}}$$

is zero when $x=\frac{1}{\lambda}$, is negative when $x<\frac{1}{\lambda}$ and is positive when $x>\frac{1}{\lambda}$. (The denominator is always positive, only the numerator affects the sign of this term) This shows that the hazard rate function has a bathtub shape and changes from decreasing to increasing at $x=\frac{1}{\lambda}$.

b) Let $\alpha=2$. Then:

$$h'(x)=2\lambda^2 e^{(\lambda x)^2}\left[1+x(2)(\lambda x)\lambda\right]$$

Since $x>0$ and $\lambda>0$, it is easy to see that the derivative is always positive since the only operations are addition and multiplication of postive terms. Thus, the hazard rate of x is monotone increasing.

Problem 2.8

The battery life of an internal pacemaker, in years, follows a Pareto distribution with $\theta=4$ and $\lambda=5$.
a) What is the probability the battery will survive for at least 10 years?
b) What is the mean time to battery failure?
c) If the battery is scheduled to be replaced at the time t_0, at which 99% of all batteries have yet to fail (that is, at t_0 so that $P(X>t_0)=0.99$), find t_0.

Answer

The pdf of the Pareto distribution is

$$f(x)=\begin{cases}\frac{\theta\lambda^\theta}{x^{\theta+1}} & \text{for } x\geq\lambda\\ 0 & \text{otherwise}\end{cases}$$

and with $\theta=4$ and $\lambda=5$

$$f(x)=\begin{cases}\frac{2500}{x^5} & \text{for } x\geq 5\\ 0 & \text{otherwise}\end{cases}$$

a) The probability that the battery will last for at least 10 years is

$$\begin{aligned}P(X\geq 10)&=\int_{10}^{\infty}\frac{2500}{x^5}dx\\ &=\left[\frac{-625}{x^4}\right]_{10}^{\infty}\\ &=0+\frac{625}{10000}\\ &=0.0625\\ &=6.25\%\end{aligned}$$

There is a 6.25% chance that the battery lasts longer than 10 years.

b) The mean time to battery failure is

$$
\begin{aligned}
E(X) &= \int_5^\infty x\frac{2500}{x^5}dx \\
&= \int_5^\infty \frac{2500}{x^4}dx \\
&= \left[\frac{2500}{(-3)x^3}\right]_5^\infty \\
&= 0 + \frac{2500}{(3)(125)} \\
&= \frac{20}{3} \\
&\approx 6.667
\end{aligned}
$$

The mean time to battery failure is $\frac{20}{3} \approx 6.667$ years.

c) In order to find t_0 such that $P(X > t_0) = 0.99$, we solve the equation:

$$\int_5^{t_0} \frac{2500}{x^5}dx = 0.01$$

for t_0. Then doing so,

$$
\begin{aligned}
\int_5^{t_0} \frac{2500}{x^5}dx &= 0.01 \\
\left[\frac{-625}{x^4}\right]_5^{t_0} &= 0.01 \\
1 - \frac{625}{t_0^4} &= 0.01 \\
\frac{625}{0.99} &= t_0^4 \\
t_0 &\approx 5.012579
\end{aligned}
$$

So, $t_0 \approx 5.01$ years, or 5 years and roughly 4 days.

Problem 2.10

In some applications, a third parameter, called a guarantee time, is included in the models discussed in this chapter. This parameter G is the smallest time at which a failure could occur. The survival function of the three-parameter Weibull distribution is given by

$$S(x) = \begin{cases} 1 & \text{if } x < G \\ exp(-\lambda(x-G)^\alpha) & \text{if } x \geq G \end{cases}$$

a) Find the hazard rate and the density function of the three-parameter Weibull distribution.
b) Suppose that the survival time X follows a three-parameter Weibull distribution with $\alpha = 1$, $\lambda = 0.0075$ and $G = 100$. Find the mean and median lifetimes.

Answer

a) Using the fact that $f(x) = -S'(x)$, we have:

$$f(x) = \begin{cases} 0 & \text{if } x < G \\ exp(-\lambda(x-G)^{\alpha})(\lambda\alpha)(x-G)^{\alpha-1} & \text{if } x \geq G \end{cases}$$

And then using the fact that $h(x) = \frac{f(x)}{S(x)}$ we have:

$$h(x) = \begin{cases} 0 & \text{if } x < G \\ \lambda\alpha(x-G)^{\alpha-1} & \text{if } x \geq G \end{cases}$$

b) With those values, the survival function is then

$$S(x) = \begin{cases} 1 & \text{if } x < 100 \\ e^{(-0.0075(x-100))} & \text{if } x \geq 100 \end{cases}$$

or rewritten

Mean Life - The mean life is defined by:

$$\mu = mrl(0) = \int_0^{\infty} S(t)dt$$

So we have:

$$\begin{aligned} \mu &= \int_0^{infty} S(t)dt \\ &= 100 + \int_{100}^{\infty} S(t)dt \\ &= 100 + \int_{100}^{\infty} e^{(-0.0075(x-100))}dt \\ &= 100 + \int_{100}^{\infty} e^{(-0.0075(x-100))}dt \\ &= 100 + \frac{e^{0.75}}{-0.0075}\left[1/e^{0.0075x}\right]_{100}^{\infty} \\ &= 100 + \frac{e^{0.75}}{-0.0075} * [0 - 1/e^{0.75}] \\ &= 100 + \frac{1}{0.0075} \\ &\approx 233.3 \end{aligned}$$

The mean life is approximately 233.3.

Median Life - Since X is continuous, the median life is $x_{0.5}$ such that $S(x_{0.5}) = 0.5$. The median life occurs sometime after $x = 100$, and so we need to solve the equation

$$0.5 = e^{0.75-0.0075x}$$

Doing so

$$
\begin{aligned}
0.5 &= e^{0.75-0.0075x} \\
ln(0.5) &= 0.75 - 0.0075x \\
0.0075x &= 0.75 - ln(0.5) \\
x &= \frac{0.75 - ln(0.5)}{0.0075} \\
x &\approx 192.42
\end{aligned}
$$

The median life is approximately 192.42.

Problem 2.12

Let X have a uniform distribution on the interval 0 to θ with density function

$$
f(x) = \begin{cases} 1/\theta & \text{for } 0 \leq x \leq \theta \\ 0 & \text{otherwise} \end{cases}
$$

a) Find the survival function of X.
b) Find the hazard rate of X.
c) Find the mean residual-life function.

Answer

Problem 2.15

Based on data reported to the International Bone Marrow Transplant Registry, the survival function for a person given an HLA- identical sibling transplant for refractory multiple myeloma is given by

x	S(x)
Months Post Transplant	Survival Probability
$0 \leq x < 6$	1.00
$6 \leq x < 12$	0.55
$12 \leq x < 18$	0.43
$18 \leq x < 24$	0.34
$24 \leq x < 30$	0.30
$30 \leq x < 36$	0.25
$36 \leq x < 42$	0.18
$42 \leq x < 48$	0.10
$48 \leq x < 54$	0.06
$x \geq 54$	0

a) Find the probability mass function for the time to death for a refractory multiple myeloma bone marrow transplant patient.

b) Find the hazard rate of X.

c) Find the mean residual life at 12, 24, and 36 months post transplant.

d) Find the median residual life at 12, 24, and 36 months.

Answer

Problem 2.18

Given a covariate Z, suppose that the log survival time Y follows a linear model with a logistic error distribution, that is

$$Y = ln(X) = \mu + \beta Z + \sigma W$$

where the pdf of W is given by

$$f(w) = \frac{e^w}{(1+e^w)^2}, \quad -\infty < w < \infty$$

a) For an individual with covariate Z, find the conditional survival function of the survival time X, given Z, namely, $S(x|Z)$.

b) The odds that an individual will die prior to time x is expressed by $\frac{1-S(x|Z)}{S(x|Z)}$. Compute the odds of death prior to time x for this model.

c) Consider two individuals with different covariate values. Show that, for any time x, the ratio of their odds of deaths is independent of x. The log logistic regression model is the only model with this property.

Answer