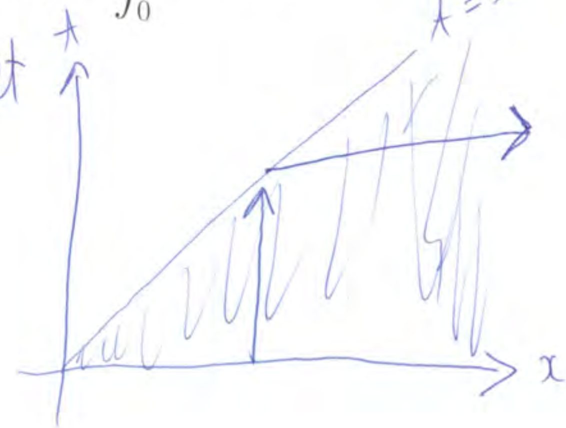
Sample Questions: Survival and Hazard Functions

STA312 Fall 2023. Copyright information is at the end of the last page.

For all these questions, T is a continuous random variable with $P(T > 0) = 1$, density $f(t)$ and cumulatve distribution function $F(t) = P(T \leq t)$.

1. The survival function is $S(t) = P(T > t)$. Prove $E(T) = \displaystyle\int_0^\infty S(t)\,dt$.

$$\int_0^\infty S(t)\,dt = \int_0^\infty \int_t^\infty f_T(x)\,dx\,dt$$



~~As t goes from 0 to ∞~~

As t goes from 0 to ∞, x goes from t to ∞

As $x \to \infty$, t goes from 0 to x

$$= \int_0^\infty \int_0^x f_T(x)\,dt\,dx$$

$$= \int_0^\infty f_T(x)\left(\int_0^x dt\right) dx = \int_0^\infty f_T(x)\,x\,dx$$

$$= \int_0^\infty x f_T(x)\,dx = E(T)$$

2. The hazard function is defined by $h(t) = \lim_{\Delta \to 0} \dfrac{P(t < T < t + \Delta | T > t)}{\Delta}$, where $\Delta > 0$. Prove $h(t) = \frac{f(t)}{S(t)}$.

$$h(t) = \lim_{\Delta \to 0} \frac{P(t < T < t+\Delta \cap \widetilde{T > t})}{\Delta \, P(T > t)}$$

$$= \lim_{\Delta \to 0} \frac{P(t < T < t+\Delta)}{\Delta \, S(t)}$$

$$= \frac{1}{S(t)} \lim_{\Delta \to 0} \frac{F(t+\Delta) - F(t)}{\Delta}$$

$$= \frac{1}{S(t)} f(t) = \frac{f(t)}{S(t)}$$

3. Prove $S(t) = e^{-\int_0^t h(x)\,dx}$.

$$\int_0^t h(x)\,dx = \int_0^t \frac{f(x)}{S(x)}\,dx$$

$$= -\int_0^t \frac{-f(x)}{1-F(x)}\,dx \qquad \begin{aligned} u &= 1-F(x) \\ du &= -f(x)\,dx \end{aligned} \qquad \begin{array}{c|c} x & u \\ \hline t & 1-F(t) \\ 0 & \underset{0}{\underset{\|}{1-F(0)}} \end{array}$$

$$= -\int_1^{1-F(t)} \frac{1}{u}\,du = \int_{1-F(t)}^{1} \frac{1}{u}\,du$$

$$= \log u \Big|_{1-F(t)}^{1} = \underset{0}{\underset{\|}{\log(1)}} - \log(1-F(t))$$

$$= -\log S(t) = \int_0^t h(x)\,dx$$

$$\Rightarrow S(t) = e^{-\int_0^t h(x)\,dx}$$

4. Let $T \sim \exp(\lambda)$. Find the hazard function $h(t)$ for $t > 0$.

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

Constant hazard rate.

5. Let T have the Pareto density $f(t|\theta) = \begin{cases} \frac{\theta}{t^{\theta+1}} & \text{for } t \geq 1 \\ 0 & \text{for } t < 1 \end{cases}$

(a) Find the hazard function $h(t)$ for $t > 1$.

$$h(t) = \frac{f(t)}{S(t)}$$

$$S(t) = \int_t^\infty \frac{\theta}{x^{\theta+1}}\,dx = \theta \int_t^\infty x^{-\theta-1}\,dx$$

$$= \cancel{\theta}\, \frac{x^{-\theta \cancel{-1+1}}}{-\cancel{\theta}} \Bigg|_t^\infty = \frac{-1}{x^\theta} \Bigg|_t^\infty$$

$$= -\left(0 - \frac{1}{t^\theta}\right) = \frac{1}{t^\theta}$$

$$h(t) = \frac{\frac{\theta}{t^{\theta+1}} (t^\theta)}{\frac{1}{\cancel{t^\theta}} \cancel{(t^\theta)}} = \frac{\theta}{t} \quad \downarrow$$

(b) Earlier, we found the MLE $\widehat{\theta}_n = \frac{n}{\sum_{i=1}^{n} \log t_i}$, and $\widehat{v}_n = \frac{\widehat{\theta}_n^2}{n}$.

i. Give $\widehat{h(t)}$, the maximum likelihood estimate of the hazard function evaluated at a particular time $t > 1$. Your answer is a formula involving t and $\widehat{\theta}_n$.

$$h(t) = \frac{\theta}{t} \qquad \widehat{h(t)} = \frac{\hat{\theta}}{t}$$

ii. We want a confidence interval for $h(t)$, the hazard function evaluated at a particular time $t > 1$. Give formulas for the lower and upper 95% confidence limits. Show your work.

$$se_{\hat{\theta}} = \sqrt{\frac{\hat{\theta}^2}{n}} = \frac{\hat{\theta}}{\sqrt{n}}$$

$$0.95 \approx P\left(\frac{\hat{\theta} - 1.96 \frac{\hat{\theta}}{\sqrt{n}}}{t} < \frac{\theta}{t} < \frac{\hat{\theta} + 1.96 \frac{\hat{\theta}}{\sqrt{n}}}{t} \right)$$