HW#1 Survival Analysis I

NAME: JIA-HAN SHIH

Problem 1.

Please derive the mean and variance under the Weilbull distribution.

Solution 1.

Let random variable T follow the Weilbull distribution with survival function

$$S_T(t) = \exp\{-(\lambda t)^\nu\}, \quad \lambda > 0,\ \nu > 0,\ t \ge 0.$$

Its probability density function is

$$f_T(t) = -\frac{d}{dt} S_T(t) = \nu\lambda^\nu t^{\nu-1} \exp\{-(\lambda t)^\nu\}.$$

By straightforward calculations, the *k*-th moment of T is

$$E(T^k) = \int_0^\infty t^k \nu\lambda^\nu t^{\nu-1} \exp\{-(\lambda t)^\nu\} dt = \int_0^\infty \nu\lambda^\nu t^{k+\nu-1} \exp\{-(\lambda t)^\nu\} dt.$$

We consider a change of variable $x = t^\nu$ ($t = x^{1/\nu}$) and then $dx = \nu t^{\nu-1} dt$. Thus, the integral becomes

$$\int_0^\infty \lambda^\nu x^{k/\nu} \exp\{-\lambda^\nu x\} dx = \lambda^\nu \int_0^\infty x^{k/\nu+1-1} \exp\{-\lambda^\nu x\} dx = \lambda^\nu \frac{\Gamma(k/\nu+1)}{(\lambda^\nu)^{k/\nu+1}}$$
$$= \frac{\Gamma(k/\nu+1)}{\lambda^k}.$$

Then second equality follows from the gamma integral. Then, the mean and variance are

$$E(T) = \frac{\Gamma(1/\nu+1)}{\lambda} \quad \text{and}$$

$$\text{var}(T) = E(T^2) - \{E(T)\}^2 = \frac{\Gamma(2/\nu+1) - \{\Gamma(1/\nu+1)\}^2}{\lambda^2},$$

respectively. On the other hand, the median of the Weibull distribution is obtained by solving

$$\exp\{-(\lambda t_{0.5})^\nu\} = 1/2 \quad \Rightarrow \quad t_{0.5} = \frac{(\log 2)^{1/\nu}}{\lambda}.$$

Problem 2.

Suppose the median survival times of male and female are 70 and 80, respectively. Please derive the mean survival times of male and female (including parameters λ and β) under the Weibull model by specifying a reasonable ν.

Solution 2. Under the Weilbull model, human survival time follows

$$S_T(t\,|\,x) = \exp\{-(\lambda t)^{\nu} e^{\beta x}\}, \qquad \lambda > 0,\ \nu > 0,\ \beta \in \mathrm{R},\ t \geq 0,$$

where the covariate $x=1$ indicates male and $x=0$ indicates female. The formulas in Problem 1 can be applied with λ replaced by $\lambda e^{\beta x/\nu}$. Here, we set the shape parameter $\nu = 2$ for convenience (see NOTE).

If the median survival times for male and female are 70 and 80, respectively. We have

$$80 = \frac{(\log 2)^{1/2}}{\lambda} \qquad \text{and} \qquad 70 = \frac{(\log 2)^{1/2}}{\lambda e^{\beta/2}}.$$

Solving these two equations, we obtain

$$\lambda = \frac{(\log 2)^{1/2}}{80} = 0.0104 \qquad \text{and} \qquad \beta = \log(\log 2) - 2\log(70\lambda) = 0.2671.$$

Then, the mean survival times for male and female are

$$\frac{\Gamma(1/2+1)}{\lambda e^{\beta/2}} = 74.51 \qquad \text{and} \qquad \frac{\Gamma(1/2+1)}{\lambda} = 85.16,$$

respectively. These results seem more reasonable than the exponential model.

NOTE: The hazard function of the Weibull distribution is

$$h_T(t) = \frac{f_T(t)}{S_T(t)} = \nu \lambda^{\nu} t^{\nu-1}.$$

Thus, if $\nu > 1$, the hazard function is increasing in t. This may represent that the probability of a person dies in old age is higher than he (or she) dies in young age.