Parametric regression models for survival data

Rasmus Waagepetersen

December 8, 2022

Parametric distributions for duration data

- Weibull
- log logistic
- log normal
- gamma
- ...

Essentially distributions for non-negative random variables.

Hazard function is useful for identifying essential differences between distributions (monotone/non-monotone,decreasing/increasing,bath-tub,hump-shaped,...)

Regression models

Let $S_0(\cdot;\psi)$ and $h_0(\cdot;\psi)$ be survival function and hazard function of a parametric duration data distribution.

Let z covariate vector of an individual (without a 1 for intercept).

Proportional hazards model:

$$h(t;z,\psi,\beta) = h_0(t;\psi)\exp(z^\mathsf{T}\beta) \quad S(t;z,\psi,\beta) = S_0(t;\psi)^{\exp(z^\mathsf{T}\beta)}$$

Accelerated failure time:

$$S(t;z,\psi,\gamma) = S_0(t\exp(-z^\mathsf{T}\gamma);\psi)$$
$$h(t;z,\psi,\beta) = h_0(t\exp(-z^\mathsf{T}\gamma);\psi)\exp(-z^\mathsf{T}\gamma)$$

Weibull

Weibull: $\psi = (\lambda, \alpha)$

$$H_0(t) = \lambda t^{\alpha} \quad h_0(t) = \lambda \alpha t^{\alpha - 1}$$

Proportional hazards:

$$h(t; z, \psi, \beta) = \lambda \alpha t^{\alpha-1} \exp(z^{\mathsf{T}} \beta)$$

Accelerated failure time:

$$\begin{aligned} h(t; z, \psi, \beta) &= \lambda \alpha t^{\alpha-1} \exp(-z^{\mathsf{T}} \gamma(\alpha - 1)) \exp(-z^{\mathsf{T}} \gamma) \\ &= \lambda \alpha t^{\alpha-1} \exp(-z^{\mathsf{T}} \gamma \alpha) \end{aligned}$$

Models coincide with $\beta = -\gamma\alpha$!

Weibull distribution is the only distribution where the proportional hazards model and the accelerated failure time are equivalent (see last slide) !

Log linear model

Model T as $\exp(Y)$ where

$$Y = \mu + z^{\mathsf{T}}\gamma + \sigma W$$

where W is a random variable (extreme-value, standard normal, logistic).

Then T is guaranteed to be non-negative.

Suppose $S_0(\cdot;\psi)$ is survival function of $\exp(\mu + \sigma W)$ where ψ is parameter for W. Then T follows the accelerated failure-time model:

$$S(t; z, \psi, \gamma) = S_0(t \exp(-z^{\mathsf{T}}\gamma); \psi)$$

Weibull

Density function and CDF of standard extreme-value distribution:

$$f(w) = \exp(w - \exp(w)) \quad 1 - \exp(-\exp(w))$$

Survival function of $\exp(\mu + \sigma W)$:

$$P(\exp(\mu + \sigma W) > t) = P(W > (\log(t) - \mu)/\sigma)$$
$$= \exp(-\exp((\log(t) - \mu)/\sigma)) = \exp(-\exp(-\mu/\sigma)t^{1/\sigma})$$

This is Weibull distribution with

$$\lambda = \exp(-\mu/\sigma) \quad \alpha = 1/\sigma$$

`survreg`

Procedure `survreg` is based on log linear model so provides estimate of σ (scale), μ (intercept) and γ (remaining parameters).

If we want λ, α, β for Weibull proportional hazards model:

$$\lambda = \exp(-\mu/\sigma) \quad \alpha = 1/\sigma \quad \beta = -\gamma/\sigma$$

PH and AFT implies Weibull

Suppose for *any* (t, z):

$$h(t) = h_1(t)\exp(\beta z) = h_2(t\exp(-\gamma z))\exp(-\gamma z)$$

Choose $z = 0$. Then $h_1(t) = h_2(t) := h_0(t)$.

Choose $z = \log(t)/\gamma$. Then

$$h_0(t)t^{\beta/\gamma} = h_0(tt^{-1})t^{-1} = h_0(1)t^{-1}$$

Thereby, with $\lambda\alpha = h_0(1)$ and $\alpha = -\beta/\gamma$,

$$h_0(t) = \lambda\alpha t^{\alpha-1}$$

That is, h_0 is the hazard function of a Weibull distribution.

(Note: in practice an individual will have a fixed z and t varying. The argument works the other way around: for any t find an individual with $z = \log(t)/\gamma$)