Accelerated failure time models

HY, 2019

Accelerated failure time models

- Denote $\theta = \exp(\beta' Z)$ where β are the (log) regression parameters and Z a vector of covariates
- $S_0(u) = S(u; \theta = 1)$ denotes the survival function when $\theta = 1$, i.e. when $Z = 0$ (the covariates are at their baseline levels/values)
- So far we have considered proportional hazards models, in which the survival function with any θ is $S(t; \theta) = [S_0(t)]^\theta$
- In the accelerated failure time (AFT) models, the survival function is taken to depend on θ in a different way:

$$S(t; \theta) = S_0(\theta t)$$

- The covariates can be interpreted to effect directly on the survival times, in contrast to proportional hazards models in which they affect on the hazard; this means that in general θ (and β) parameters are *not* proportional hazards

Density and hazards function in AFT models

- Denote the density and hazard functions with $\theta = 1$ as $f_0(u)$ and $\lambda_0(u)$
  - Note that there are still parameters other than $\beta = \log(\theta)$ that define the baseline density (and hazard) function
- It follows immediately from the AFT form of the survival function that the density function can as well be expressed as follows:

$$f(t;\theta) = -\frac{dS(t;\theta)}{dt} = -\frac{dS_0(\theta t)}{dt} = \theta f_0(\theta t)$$

- The hazard function is

$$\lambda(t;\theta) = \frac{f(t;\theta)}{S(t;\theta)} = \frac{\theta f_0(\theta t)}{S_0(\theta t)} = \theta \lambda_0(\theta t)$$

Exponential regression

- Exponential regression models can be viewed either as proportional hazard models or AFT models
- This follows from the observation that the dependence of the survival function on θ can be written in two alternative ways:

$$
\begin{aligned}
S(t;\theta) = e^{-\theta\lambda t} \quad &= \overbrace{[e^{-\lambda t}]^{\theta}}^{\text{prop. hazards}} = \overbrace{e^{-\lambda(\theta t)}}^{\text{AFT}} \\
&= \overbrace{S_0(t)^{\theta}}^{\text{prop. hazards}} = \overbrace{S_0(\theta t)}^{\text{AFT}}
\end{aligned}
$$

- Weibull regression, an extension of exponential regression, is the only class of regressions models that has both interpretations (see page the next page)
- Otherwise, the two classes of regression models are different and their parameters have different interpretations

Weibull regression

- Let $\tilde{\theta}$ denote the proportional hazard under the Weibull model

$$S(t;\theta) = e^{-\tilde{\theta}(\alpha t)^p} \quad = \overbrace{[e^{-(\alpha t)^p}]^{\tilde{\theta}}}^{\text{prop. hazards}} = \overbrace{e^{-(\alpha\tilde{\theta}^{1/p}t)^p}}^{\text{AFT}}$$

$$= \overbrace{S_0(t)^{\tilde{\theta}}}^{\text{prop. hazards}} = \overbrace{S_0(\tilde{\theta}^{1/p}t)}^{\text{AFT}}$$

- This means that the proportional hazards Weibull regression models with hazard ratio $\tilde{\theta}$ is also an AFT model with regression parameter $\theta = \tilde{\theta}^{1/p}$
- Weibull is the only class of models that have both the AFT and proportional hazards interpretation
- To define more general AFT models, we next turn to characterising them as log-linear models for the survival time

AFT models as log-linear models

- It turns out that AFT models can be presented as log-linear models for the survival time
- To see this, we first show that the distribution of $\log(\theta T)$ does not depend on θ (it may still depend on other "baseline" parameters for $f_0(t)$)
- $U = \log(\theta T)$ has a density function as follows:

$$f_U(u) = f_T(t)|\frac{dt}{du}| = \frac{e^u}{\theta}\theta f_0(\theta t) = e^u f_0(e^u)$$

- The above result means that $\log(T)$ can be written as a sum of $\log(\theta)$ and a random variable whose distribution does not depend on θ:

$$\log(T) = \log(\theta T/\theta) = -\log(\theta) + \log(\theta T) = -\beta' Z + U$$

Example: exponential regression

- Under exponential regression $T \sim \text{Exp}(\lambda e^{\beta' Z})$
  - Indexing individuals is omitted for convenience
- Let θ include the baseline hazard, i.e. write $\theta := \lambda \exp(\beta' Z)$
- We have already seen (task 3/exercise 2) that under the exponential model

$$\log(\theta T) = W, \quad \text{where } f(w) = \exp(w - e^w), \quad -\infty < w < \infty$$

- It follows that

$$\log(T) = \log(\theta T/\theta)) = -\log(\theta) + W = -\log(\lambda) - \beta' Z + W$$

- In this case, the distribution of the "error" term W does not depend on any parameter

Example: Weibull regression

- The Weibull hazard is $p\alpha^p t^{p-1} \exp(\beta' Z)$
- Write $\theta = \alpha \exp(\beta' Z/p)$ so that the hazard is $p\theta^p t^{p-1}$
- We have already shown (in exercise 3) that under the Weibull model

$$\log(\theta T) = p^{-1} W \equiv \sigma W, \text{ where } f(w) = \exp(w - e^w), \ -\infty < w < \infty$$

- It follows that

$$\log(T) = \log(\theta T/\theta) = -\log(\theta) + \sigma W = -\log(\alpha) - \sigma\beta' Z + \sigma W$$

Remarks

- We have found that in AFT models $\log(T) = -\beta' Z + \sigma U$
- If we write the (individual) covariate as $Z_i = (1, Z_{i1}, \ldots, Z_{i,p})$, we have

$$\log(T) = -\beta_0 - \beta'(Z_{i1}, ..., Z_{ip})' + \sigma U$$

- This is the formulation we used for exponential and Weibull regression above ($\beta_0 = -\log(\lambda)$ or $\beta_0 = -\log(\alpha)$)

- For the Weibull model, the distribution of U was shown to be the so called extreme value distribution with density $f(w) = \exp(w - e^w)$
- In general, U does **not** need to has the extreme value distribution; in fact, different AFT models are formulated most easily by choosing an appropriate error distribution
- Often the AFT model is written as $\log(T) = \beta_0 + \beta' Z + \sigma U$, in which case values $\beta > 0$ mean increasing life times (smaller hazard)

Log-logistic regression

- In general, AFT models can be defined by choosing the distribution of the error term appropriately
- As an example of an AFT model which is *not* a proportional hazards model, consider so called log-logistic distribution
- We choose

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

- The regression model is now

$$\log(T) = -\log(\theta) + \sigma U = -\beta_0 - \beta' Z + \sigma U$$