Extra exercises in STK4080 2019

Solution Exercise E5.1

Let the situation be as on pages 14-16 in Slides 15.

Assume now that $\alpha(t) = \lambda$ is a constant, and let $Z \sim \text{Gamma}(1/\delta, 1/\delta)$.

a) Write down the resulting expressions for $A(t), \mu(t), S(t)$ and $f(t)$.

Solution:

$$\begin{aligned}
A(t) &= \lambda t \\
\mu(t) &= \lambda(1 + \delta\lambda t)^{-1} \\
S(t) &= (1 + \delta\lambda t)^{-\frac{1}{\delta}} \\
f(t) &= \mu(t)S(t) = \lambda(1 + \delta\lambda t)^{-\frac{1}{\delta}-1}
\end{aligned}$$

Let $(\tilde{T}_i, D_i); i = 1, 2, \ldots, n$ be a (possibly right censored/independent censoring) sample from the distribution. Observe that this is the same situation as considered in Exercise 1 of Slides 16, but now assuming that each cluster consists of exactly one item (individual).

b) Use the cited Exercise 1 to write down the likelihood and log likelihood for the present problem. Show that the log likelihood can be written

$$\ell(\lambda, \delta) = D_{\bullet} \log \lambda - \sum_{i=1}^{n} \left(D_i + \frac{1}{\delta} \right) \log \left(1 + \delta\lambda\tilde{T}_i \right)$$

where $D_{\bullet} = \sum_{i=1}^{n} D_i$ is the number of events.

Solution: Use the expression for $\log L$ on p. 7 in Slides 16. Note that we now have $n_i = 1$ and $V_i = \lambda\tilde{T}_i$. Further, $D_{i\bullet} = D_i$ and $\mathcal{L}(c) = (1+\delta c)^{-1/\delta}$, $\mathcal{L}'(c) = -(1+\delta c)^{-1/\delta-1}$. Check that the above expression corresponds to the one of p. 7 in Slides 16 both for $D_i = 0$ and $D_i = 1$.

c) Derive the score functions and explain how you can find the maximum likelihood estimates for λ and δ.

Solution: It is straightforward to derive the equations $\frac{\partial\ell(\lambda,\delta)}{\partial\lambda} = 0$ and $\frac{\partial\ell(\lambda,\delta)}{\partial\delta} = 0$.

Exercise E5.2

We have intuitively explained the fact that the hazard rates are “dragged down” on p. 17 of Slides 15 by the fact that for large values of δ, there are many individuals with high hazard, and which fail first.

In this exercise we will show that the expression for $E(Z_i|H_i)$ on p. 11 of Slides 16 can be used to get some information on this effect of “failing first”.

The calculation in Slides 16 is for a cluster of n_i individuals. We can, however, put $n_i = 1$.

a) Consider the population hazard as calculated in Slides 15 on p. 16 and illustrated on p. 17. Explain why the formula on p. 11 of Slides 16 implies that

$$E(Z|T=t) = \frac{\delta + 1}{\delta A(t) + 1}$$

Describe in words what this equation tells us about the size of Z for an individual that fails at time t. Consider in particular $t = 0$.

Solution: The setting is now without data, so there is no censoring. Moreover, $n_i = 1$ so $D_{i\bullet} = 1$ and $V_i = A(t)$. Thus the formula on p. 11, Slides 16 gives the result of the exercise. This tells us the expected value of the frailty Z for an individual that fails at time t. It is a decreasing function of t, meaning intuitively that the most frail individuals fail first. Close to 0, the value is $\delta + 1$, which is the maximum value of the expectation.

b) Let now $\alpha(t) = t^2$ as on p. 17 of Slides 15. Calculate $E(Z|T=t)$ for some values of t and comment. For which value of t is $E(Z|T=t) = 1$?

Solution:

$$E(Z|T=t) = \frac{3\delta + 3}{t^3\delta + 3}$$

This equals 1 if $t = 3^{1/3} \approx 1.44$.