TMA 4275 V2019

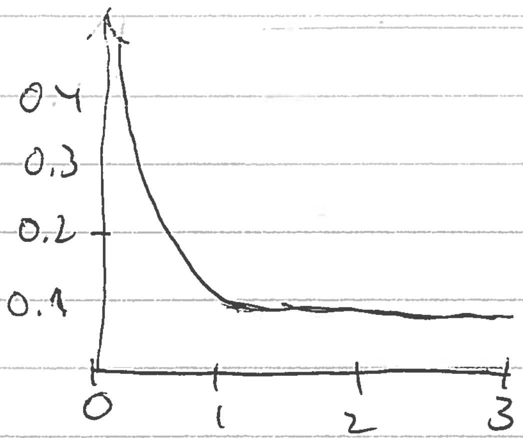
Problem 1

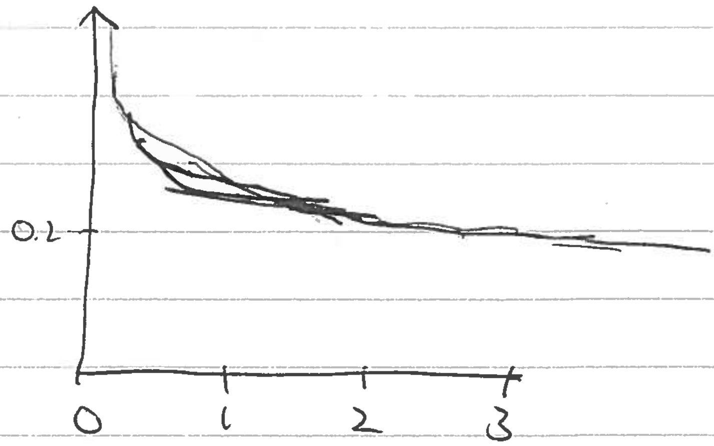
- $Z(t) = \int_0^t z(u)\,du$
- $f(t) = z(t)\, e^{-\int_0^t z(u)\,du}$
- $R(t) = e^{-\int_0^t z(u)\,du}$

$Z(t)$ is estimated by NA-estimator

$$\hat{Z}_{NA}(t) = \sum_{t_{(i)} \le t} \frac{d_i}{n_i}$$ where $d_i = \#$ failing at $t_{(i)}$, $n_i = \#$ at risk at $t_{(i)}$

Sketches: Take the derivatives to plot.





No-AMI: Hazard large in beginning, stabilizes at level ≈ 0.1 at 1 year, slightly down after this

At least one-AMI: Hazard very large in beginning, constantly high from approx 0.5 year, when it stabilizes. Lower after 1.5 years

Problem 2.

a) See Slides 13

Arrhenius model:

T is lifetime
s is temperature in °C
β_0, β_1 are regression parameters
σ is scale parameter
$\mathcal{U}$ is a standardized r.v. with values in $\mathbb{R}$

Use: Put units on test at different temperatures, usually (much) higher than natural temperature for the unit. Estimate parameters by M.L; predict lifetime characteristics under normal temperature.

b) Estimated model:

$$\ln T = -13.4693 + 0{,}6279 \cdot \frac{11604.83}{s + 273{,}16} + 0.977823 \cdot \mathcal{U}$$

$U \sim N(0,1)$ in lognormal model.

$P(T > 30000) = P(\ln T > \ln 30\,000)$

$= P(\ln T > 10.30895)$

$= 1 - \Phi(\ln T \leq 10.30895)$

Putting $s = 10$ in the estimated model we get

$$\ln T = 12.26411 + 0.977823\, U$$

so we estimate:

$$\widehat{P(T>30000)} = 1 - \Phi\left(\frac{10.30895 - 12.26411}{0.977823}\right)$$

$$= 1 - \Phi(-2.00) = \Phi(2.00) = \underline{0.97719}$$

The requirement was $P(T > 30000) = 0.95$, so (modulo statistical error) the requirement is satisfied.

We need formulas for expected value; median; quantiles.
These are given in the lecture slides; Slide 10:

$$E(T) = e^{\mu + \frac{1}{2}\sigma^2}$$

$$\text{Median}(T) = e^{\mu}$$

$$t_p = e^{\mu + \sigma \Phi^{-1}(p)}$$

Here we estimate; ,

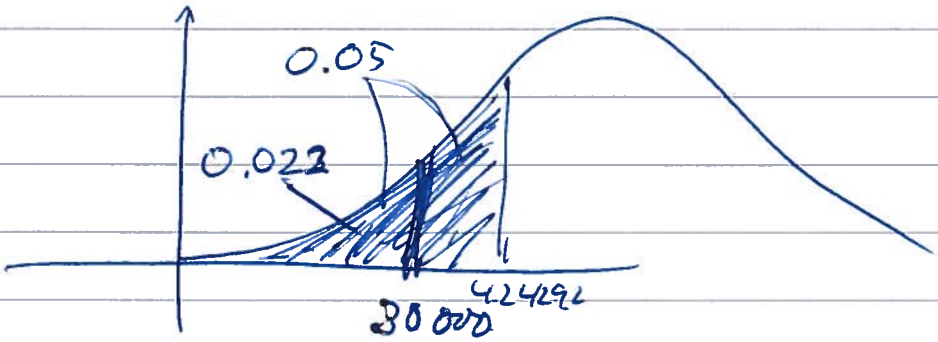
$$\widehat{E(T)} = e^{12.26411 + \frac{1}{2} \cdot 0.977823^2}$$

$$= 341867.6$$

$$\widehat{\text{Median}(T)} = e^{12.26411} = 211950.9$$

$$\widehat{t_{0.05}} = e^{12.26411 + 0.977823 \cdot (-1.645)}$$

$$= 42429.2$$



c) Let T be lognormal(μ, σ). Then

$$F(t) = P(T \le t) = P(\ln T \le \ln t)$$

$$= \Phi\left(\frac{\ln t - \mu}{\sigma}\right).$$

Thus

$$\frac{\ln t - \mu}{\sigma} = \Phi^{-1}(F(t))$$

$$\cancel{\ln t = \mu + \sigma \Phi^{-1}(F(t))}$$

or $\Phi^{-1}(F(t)) = \frac{1}{\sigma} \ln t - \frac{\mu}{\sigma}$

so

$(\ln t, \Phi^{-1}(F(t)))$ is on the line

$$y = \frac{1}{\sigma} x - \frac{\mu}{\sigma}$$

Figure 2 plots the points – for each temperature –

$$(\ln T_{(i)}, \Phi^{-1}(1 - \hat{R}(T_{(i)}))$$

where $\hat{R}$ is a (modified) KM-estimate ~~for each~~ temperature. The corresponding lines are

$$y = \frac{1}{\hat{\sigma}} x - \frac{\hat{\mu}}{\hat{\sigma}} = \frac{1}{\hat{\sigma}} x - \frac{\hat{\beta}_0 + \hat{\beta}_1 \cdot \frac{11604.83}{s + 273.16}}{\hat{\sigma}}$$

The lines are parallel because they have the same slope $1/\hat{\sigma}$ (but intercepts depend on temperature).

Model seems to fit to the data well. Can also read off expected behavior at $s = 10°C$ (but no points on this line since all observations at 10°C are censored).

In the model (1) we have

$$U = \frac{\ln T - \beta_0 - \beta_1 \frac{11604.83}{s + 273.16}}{\sigma}$$

Standardized residuals are the estimated version of this, calculated for each observation, censored or not. These are – if the model is correct – considered to be a censored sample from the distribution of U.

The plot in Figure 3 is obtained as on p. 4, now letting $\mu = 0, \sigma = 1$, and letting the data be the standardized residuals $\hat{u}_i$ (see Slides 11).

Residual plot indicates a satisfactory fit of the model.

Exercise 2

a) $W(t) = \int_0^t \alpha \beta e^{-\beta u}\, du = \alpha \left[-e^{-\beta u}\right]_0^t = \alpha(1-e^{-\beta t})$

$W(t) = E[N(t)] =$ ~~the~~ expected # errors found by time t.

$P(\text{first error before } s) = P(N(0,s) \geq 1)$

$= 1 - P(N(0,s)=0) = 1-e^{-W(s)}$

$P(\text{first error between } s \text{ and } t)$

$= P(N(0,s)=0 \cap N(s,t) \geq 1)$

$= P(N(0,s)=0) \cdot (1 - P(N(s,t)=0))$

$= e^{-W(s)} \cdot (1 - e^{-(W(t)-W(s))})$

$= e^{-W(s)} - e^{-W(t)}$

$\lim_{t \to \infty} W(t) = \alpha$

i.e. α is the expected # errors found by time $t=\infty$, i.e., can be thought of as the expected # initial errors.

b)

$$R(t|s) = P(\text{no failures in } (s, s+t))$$

$$= e^{-(W(s+t) - W(s))}$$

$$= e^{-\alpha[(1-e^{-\beta(s+t)}) - (1-e^{-\beta s})]}$$

$$= e^{-\alpha[e^{-\beta s} - e^{-\beta(s+t)}]}$$

$$= e^{-\alpha e^{-\beta s}[1-e^{-\beta t}]}$$

$$= e^{-e^{-\beta s} \cdot W(t)}$$

Q.E.D.

Require: $R(t_0|s) \geq r$, r and t_0 given.

Find optimal s_0

$$R(t_0|s) \geq r$$

$$e^{-e^{-\beta s} W(t_0)} \geq r$$

$$-e^{-\beta s} W(t_0) \geq \ln r$$

$$e^{-\beta s} W(t_0) \leq -\ln r$$

$$-\beta s + \ln W(t_0) \leq \ln(-\ln r)$$

$$s \geq \frac{-\ln(-\ln r) + \ln W(t_0)}{\beta}$$

So optimal s_0 is

$$s_0 = \frac{-\ln(-\ln r) + \ln W(t_0)}{\beta}$$

$$= \frac{-\ln(-\ln r) + \ln \alpha + \ln(1-e^{-\beta t_0})}{\beta}$$

$t_0 = \infty$: $$s_0 = \frac{-\ln(-\ln r) + \ln \alpha}{\beta}$$

This s_0 is the time when the probability of finding no more errors is $\geq r$.

c) ~~d)~~ After time S_a there are no failures in the system, as all the a initial failures are found (by the assumptions of the model).

$$\underline{\underline{E[S_a]}} = E[T_1 + \cdots + T_a]$$

$$= \frac{1}{ba} + \frac{1}{b(a-1)} + \frac{1}{b(a-2)} + \cdots + \frac{1}{b}$$

$$\approx \frac{1}{b}\int_1^a \frac{1}{x}\,dx = \underline{\underline{\frac{\ln a}{b}}}$$

When $t_0 = \infty$, the optimal test time in (d) is

$$s_0 = \frac{-\ln(-\ln r) + \ln \alpha}{\beta}$$

Here $-\ln(-\ln r)$ is a slowly varying function,

$$-\ln(-\ln r) = \begin{cases} r = 0.9 : 2.25 \\ r = 0.99 : 4.60 \\ r = 0.999 : 6.91 \end{cases}$$

Thus for large α, the optimal time s_0 and $E(S_a)$ (assuming $a \approx \alpha$) are approximately the same.

d) $$L(a,b) = \prod_{i=1}^{m} \lambda_i e^{-\lambda_i (s_i - s_{i-1})}$$

$$= \prod_{i=1}^{m} (a-i+1) b\, e^{-(a-i+1) b (s_i - s_{i-1})}$$

Ok answer $\rightarrow$ $$= b^m \cdot \prod_{i=1}^{m} (a-i+1) e^{-(a-i+1) b (s_i - s_{i-1})}$$

Can show: $\rightarrow$ $$= b^m \left\{ \prod_{i=1}^{m} (a-i+1) \right\} \cdot e^{-b \left[\sum_{i=1}^{m-1} s_i + (a-m+1) s_m \right]}$$

where we used that

$$\sum_{i=1}^{m} (a-i+1)(s_i - s_{i-1})$$

$$= \sum_{i=1}^{m} (a-i+1) s_i - \sum_{i=1}^{m} (a-i+1) s_{i-1}$$

$$= \sum_{i=1}^{m} (a-i+1) s_i - \sum_{i=0}^{m-1} (a-i) s_i$$

can put $i=1$ since $s_0 = 0$

$$= \sum_{i=1}^{m-1} s_i + (a-m+1) s_m$$

Thus, sufficient statistic is $\left(\sum_{i=1}^{m-1} S_i, S_m \right)$

Maximizing: Note that a is a positive integer. Must be $a \geq m$. Derive profile likelihood for a by maximizing w.r.t. b for each $a = m, m+1, \dots$