TMA 4275 Lifetime analysis
Exercise 6 - solution

Problem 1

a) The likelihood is given by

$$L(\theta,\gamma|(y_1,d_1)\dots,(y_n,\delta_n)) = \prod_{i:\delta_i=1} f(y_i,\theta,\gamma) \prod_{j:\delta_j=0} R(y_j,\theta,\gamma)$$
$$= \frac{1}{\theta^r}\exp\left(-\frac{1}{\theta}\sum_{i=1}^{n}(y_i-\gamma)\right)$$

where $r = \sum_{i=1}^{n}\delta_i$ denotes the number of noncensored observations.

The loglikelihood is then given by

$$l(\theta,\gamma|(y_1,d_1)\dots,(y_n,\delta_n)) = -r\log(\theta) - \frac{1}{\theta}\sum_{i=1}^{n}(y_i-\gamma)$$
$$= -r\log(\theta) - \frac{1}{\theta}\sum_{i=1}^{n}y_i + \frac{n\gamma}{\theta}$$

b) The estimator of θ can be then found by solving $\frac{\partial l}{\partial \theta} = 0$ which gives

$$\hat{\theta} = \frac{1}{r}\sum_{i=1}^{n}(y_i-\gamma)$$

which is valid for $\gamma \leq \min\{y_1,\dots,y_n\}$.

It can be seen, that the loglikelihood is maximized by choosing γ as large as possible. At the same time $y_i \geq \gamma$, therefore $\hat{\gamma} = \min\{y_1,\dots,y_n\}$.

c) It is given that

$$2(l(\hat{\theta},\hat{\gamma}) - l(\theta,\hat{\gamma}(\theta)) \sim \chi_1^2$$

This implies that

$$P\left(2(l(\hat{\theta},\hat{\gamma}) - l(\theta,\hat{\gamma}(\theta))) \leq \chi^2_{1,\alpha}\right) = 1-\alpha$$

i.e. solve for θ

$$l(\theta,\hat{\gamma}(\theta)) \geq l(\hat{\theta},\hat{\gamma}) - \frac{\chi^2_{1,\alpha}}{2}$$

d) MINITAB output:

```
Distribution:   2-Parameter Exponential
Parameter Estimates
                        Standard   95.0% Normal CI
Parameter  Estimate        Error     Lower    Upper
Scale       86.1240      20.8880   53.5400  138.538
Threshold   138.467            0   138.467  138.467
Log-Likelihood = -92.749
```

Note, that the fit is not good.

e) If the censoring is also possible before time γ then the estimator of γ is the first observed failure because in this case the censored observations before the first failure does not change the value of the loglikelihood.

Problem 2

a) The likelihood is given by

$$\begin{aligned} L(\theta) &= \prod_{i=1}^{r} f(t_i, \theta) \prod_{j=1}^{n-r} R(c, \theta) \\ &= \frac{1}{\theta^r} \exp\left(-\frac{1}{\theta} \sum_{i=1}^{r} t_i + (n-r)c \right) \end{aligned}$$

where r denotes the number of failures.

Taking logarithm of the likelihood and solving the score function gives

$$\begin{aligned} \hat{\theta} &= \frac{1}{r} \left(\sum_{i=1}^{r} t_i + (n-r)c \right) \\ &= 19.25 \end{aligned}$$

b) Since this case is right truncation, we have to modify the likelihood function. This may be also seen as a nonrandom sampling process, nonrandom in the sense that we do not sample randomly from the whole population, but only from the population that fails before time $c = 10$. When we compute the likelihood, we thus need to compensate for the sampling procedure.

The density function is given by

$$f(t|t < c) = \frac{f(t \cap t < c)}{P(t < c)} = \frac{f(t)}{F(c)}$$

giving

$$L(\theta) = \prod_{i=1}^{r} f(t_i | t_i < c)$$

c) The score equation in this case is a highly nonlinear equation which can be solved only numerically. This gives $\hat{\theta} = 10.96$