STAT 613 Final Solutions (2011)

Please note my solutions are not necessarily the only correct solution

1) (i) $T \sim \frac{1}{\lambda}\exp\left(-\frac{t}{\lambda}\right)$ $C \sim \frac{1}{\mu}\exp\left(-\frac{c}{\mu}\right)$

hence $P(T > x) = \exp(-t/\lambda)$ and $P(C > c) = \exp(-c/\mu)$.

Thus

$$P(T < C + x) = \int P(T < c + x)\underbrace{f_c(c)}_{\text{density of } C}dx$$

$$= \int_0^\infty \left[1 - \exp\left(-\frac{(c+x)}{\lambda}\right)\right]\frac{1}{\mu}\exp\left(-\frac{c}{\mu}\right)dc$$

$$= 1 - e^{-x/\lambda}\frac{1}{\mu}\int_0^\infty e^{-c\left(\frac{1}{\lambda}+\frac{1}{\mu}\right)}dc$$

$$= 1 - e^{-x/\lambda}\cdot\frac{1}{\mu}\int e^{-c\left(\frac{\mu+\lambda}{\mu\lambda}\right)}dc$$

$$= 1 - e^{-x/\lambda}\cdot\frac{\lambda}{\lambda+\mu}.$$

(ii) $$\frac{\partial \mathcal{L}_n(\lambda)}{\partial\lambda} = \sum_{i=1}^{T}\delta_i\cdot\frac{1}{f}\frac{\partial f}{\partial\lambda} + \sum_{i=1}^{T}(1-\delta_i)\frac{1}{\mathcal{F}(t;\lambda)}\frac{\partial\mathcal{F}}{\partial\lambda}$$

$f(x;\lambda) = \frac{1}{\lambda}e^{-t/\lambda}$ and $\mathcal{F}(x;\lambda) = e^{-t/\lambda}$. hence

$$\hat{\lambda}_n = \frac{\sum_i \delta_i T_i + \sum(1-\delta_i)C_i}{\sum_i \delta_i} \quad \text{where} \quad \delta_i = \begin{cases} 1 & \text{if } T_i < C_i \\ 0 & \text{if } T_i > C_i \end{cases}$$

$$\mathbb{E}\{\delta_i T_i\} = \mathbb{E}\{T_i\, I(T_i < C_i)\} = \mathbb{E}\{T_i\,\mathbb{E}[I(C_i > T_i | T_i)]\}$$

$$= \mathbb{E}\left\{T_i e^{-T_i/\mu}\right\} = \int t e^{-t/\mu}\frac{1}{\lambda}e^{-t/\lambda}d\lambda$$

$$= \frac{1}{\lambda}\int t e^{-t\left(\frac{\mu+\lambda}{\mu\lambda}\right)}dt =$$

$$= \frac{(\mu+\lambda)}{\mu\lambda^2}\int\left(\frac{\mu\lambda}{\mu+\lambda}\right)t e^{-t\left(\frac{\mu+\lambda}{\mu\lambda}\right)}dt = \frac{(\mu+\lambda)^2}{\mu^2\lambda^3}.$$

Similarly we have

$$\mathbb{E}\{(1-\delta_i)C_i\} = \frac{(\mu+\lambda)^2}{\lambda^2 \mu^3}.$$

Finally $\mathbb{E}(\delta_i) = P(T<C) = 1 - \dfrac{\lambda}{\lambda+\mu} = \left(\dfrac{\mu}{\lambda+\mu}\right)$.

Therefore (by Slutsky's lemma).

$$\hat{\lambda}_n \xrightarrow{P} \frac{\frac{(\mu+\lambda)^2}{\mu^2\lambda^3} + \frac{(\mu+\lambda)^2}{\mu^3\lambda^2}}{\frac{\mu}{(\lambda+\mu)}} = \frac{(\lambda+\mu)^3}{\mu}\left\{\frac{1}{\mu^2\lambda^3} + \frac{1}{\lambda^2\mu^3}\right\}$$

$$= \frac{(\lambda+\mu)^3}{\mu^3\lambda^3}\left\{\lambda+\mu\right\} = \frac{(\lambda+\mu)^4}{\mu^3\lambda^3}.$$

Clearly when $\mu \neq \infty$, this is a biased estimator of λ.

(iii) By part (i) we observe that

$$\hat{p} = \frac{1}{T}\sum_i \delta_i \xrightarrow{P} \left(\frac{\mu}{\lambda+\mu}\right)$$

iid random variables

and by part (ii) we have that

$$\hat{\lambda}_n = \frac{\sum_i \delta_i T_i + \sum(1-\delta_i)C_i}{\sum_i \delta_i} \xrightarrow{P} \frac{(\lambda+\mu)^4}{\mu^3\lambda^3}.$$

Thus using $\hat{p}$ and $\hat{\lambda}_n$ we can obtain estimators of μ and λ. (by solving these two equations).

<u>note</u> An alternative solution is to ~~const~~ construct the likelihood for μ based on left censoring. This together with the above likelihood will lead to 2 estimators ~~whic~~ which are different functions of $(\mu,\lambda) \to$ solve this

(2i) Since $\mathcal{F}_1$ and $\mathcal{F}_2$ are monotonically decreasing positive functions where $\mathcal{F}_1(0) = \mathcal{F}_2(0) = 1$ and $\mathcal{F}_1(\infty) = \mathcal{F}_2(\infty) = 0$, then it immediately follows that

$$\mathcal{F}(t,x) = p\,\mathcal{F}_1(t)^{e^{\beta_1 x}} + (1-p)\,\mathcal{F}_2(t)^{e^{\beta_2 x}}$$

use that $\frac{d\mathcal{F}_i(t)}{dt} = -f_i(t)$

is the same, thus $\mathcal{F}(t;x)$ is a survival function.

$$\frac{\partial \mathcal{F}(t,x)}{\partial t} = -p e^{\beta_1 x} f_1(t)\,\mathcal{F}_1(t)^{e^{\beta_1 x}-1} - (1-p)e^{\beta_2 x} f_2(t)\,\mathcal{F}_2(t)^{e^{\beta_2 x}-1}$$

$$\Rightarrow f(t;x) = +p e^{\beta_1 x} f_1(t)\,\mathcal{F}_1(t)^{e^{\beta_1 x}-1} + (1-p)\,e^{\beta_2 x} f_2(t)\,\mathcal{F}_2(t)^{e^{\beta_2 x}-1}$$

(ii) The censored loglikelihood is

$$\mathcal{L}_n(\beta_1,\beta_2,p) = \sum_i \Big[\delta_i \log f(Y_i;\beta_1,\beta_2,p) + (1-\delta_i)\log \mathcal{F}(Y_i;\beta_1,\beta_2,p)\Big]$$

clearly directly maximising the above is extremely difficult. Thus we look for an alternative method via the EM algorithm. Define the unobserved variable

$$I_i = \begin{cases} 1 & \text{with } P(I_i=1) = p = p_1 \\ 2 & \text{with } P(I_i=2) = (1-p) = p_2 \end{cases}$$

Then the joint density of (Y_i,δ_i,I_i) is

$$\delta_i\{\log p_{I_i} + \beta_{I_i} x + \log f_I(Y_i) + (e^{\beta_I x}-1)\log \mathcal{F}_I(Y_i)\}$$
$$+(1-\delta_i)\{\log p_I + e^{\beta_I x}\log \mathcal{F}_I(Y_i)\}$$

Thus the complete log-likelihood is

$$\mathcal{L}_T(\underline{Y}, \underline{\delta}, \underline{I}; \beta_1, \beta_2, p)$$

$$= \sum_{i=1}^{T} \Big\{ \delta_i \Big[\log p_{I_i} + \beta_{I_i} x_i + \log f_{I_i}(Y_i) + \big(e^{\beta_{I_i} x_i} - 1\big) \log \mathcal{F}_{I_i}(Y_i) \Big] + (1-\delta_i)\Big[\log p_{I_i} + e^{\beta_{I_i} x_i} \log \mathcal{F}_{I_i}(Y_i) \Big] \Big\}$$

Now we need to calculate $P(I_i \mid Y_i, \delta_i)$. We have

$$\omega_i^{\delta_i = 1} = P(I_i = 1 \mid Y_i, \delta_i = 1, p^*, \beta_1^*, \beta_2^*)$$

$$= \frac{p^* e^{\beta_1^* x} f_1(Y_i) \mathcal{F}_1(Y_i)^{[e^{\beta_1^* x} - 1]}}{p^* e^{\beta_1 x_i} f_1(Y_i) \mathcal{F}_1(Y_i)^{e^{\beta_1^* x_i} - 1} + (1-p^*) e^{\beta_1 x_i} f_2(Y_i) \mathcal{F}_2(Y_i)^{e^{\beta_2^* x} - 1}}$$

$$\omega_i^{\delta_i = 0} = P(I_i = 1 \mid Y_i, \delta_i = 0, p^*, \beta_1^*, \beta_2^*)$$

$$= \frac{p^* \mathcal{F}_1(Y_i)^{e^{\beta_i^* x_i}}}{p^* \mathcal{F}_1(Y_i)^{e^{\beta_1^* x_i}} + (1-p^*) \mathcal{F}_2(Y_i)^{e^{\beta_2^* x_i}}}$$

Therefore the complete likelihood conditioned on what we observe is

$$
\begin{aligned}
Q(\theta,\theta^{*}) = \sum_{i=1}^{T_1} \Big\{ & \delta_i w_i^{\delta_i} \big[\log p + \beta_1 x_i + \log f_1(Y_i) \\
& + (e^{\beta_1 x_i} - 1) \log F_1(Y_i) \big] + \\
& w_i^{1-\delta_i} (1-\delta_i) \big[\log p + e^{\beta_1 x_i} \log F_1(Y_i) \big] \Big\} \\
+ \sum_{i=1}^{T_1} \Big\{ & \delta_i (1 - w_i^{\delta_i}) \big[\log(1-p) + \beta_2 x_i + \log f_2(Y_i) \\
& + (e^{\beta_2 x_i} - 1) \log F_2(Y_i) \big] + \\
& (1 - w_i^{1-\delta_i})(1-\delta_i) \big[\log(1-p) + e^{\beta_2 x_i} \log F_2(Y_i) \big] \Big\}
\end{aligned}
$$

The conditional likelihood, above, look unwieldy, however the parameter estimates tend to be quite separate.

$$
\begin{aligned}
\frac{\partial Q}{\partial p} = & \sum_{i=1}^{T_1} \delta_i w_i^{\delta_i} \frac{1}{p} + \sum_{i=1}^{T_1} w_i^{1-\delta_i} (1-\delta_i) \frac{1}{p} \\
& - \sum_{i=1}^{T_1} \delta_i (1 - w_i^{\delta_i}) \frac{1}{(1-p)} - \sum_{i=1}^{T_1} (1 - w_i^{1-\delta_i})(1-\delta_i) \frac{1}{1-p}
\end{aligned}
$$

Now let
$$a = \sum_{i=1}^{T_1} \{ \delta_i w_i^{\delta_i} + w_i^{1-\delta_i} (1-\delta_i) \}$$
$$b = \sum_{i=1}^{T_1} \{ \delta_i (1 - w_i^{\delta_i}) + (1-\delta_i)(1 - w_i^{1-\delta_i}) \}$$

Thus
$$\frac{1}{p} a = \frac{1}{(1-p)} b \Rightarrow (1-p) a = p b$$

$$\Rightarrow \quad a = (a+b) p \quad \text{hence} \quad \boxed{\hat{p} = \frac{a_r}{a_r + b_r}}$$

estimate of $\hat{p}$ at rth iteration step

Now we consider the estimates of β_1 and β_2 at the lth iteration step.

$$\frac{\partial Q}{\partial \beta_1} = \sum_{l=1}^{T} \Big\{ \delta_l w_l^{\delta_l} \left[1 + (e^{\beta_1 x_l}) \log \mathcal{F}_1(Y_l) \right] + (1-\delta_l) w_l^{1-\delta_l} e^{\beta_1 x_l} \log \mathcal{F}_1(Y_l) \Big\} x_l = 0$$

$$\frac{\partial Q}{\partial \beta_2} = \sum_{l=1}^{T} \Big\{ \delta_l (1-w_l^{\delta_l}) \left[1 + (e^{\beta_2 x_l}) \log \mathcal{F}_2(Y_l) \right] + (1-\delta_l) w_l^{1-\delta_l} e^{\beta_2 x_l} \log \mathcal{F}_2(Y_l) \Big\} x_l = 0$$

$$\frac{\partial^2 Q}{\partial \beta_1^2} = \sum_{l=1}^{T} \Big\{ \delta_l w_l^{\delta_l} e^{\beta_1 x_l} \log \mathcal{F}_1(Y_l) + (1-\delta_l) w_l^{1-\delta_l} e^{\beta_1 x_l} \log \mathcal{F}_1(Y_l) \Big\} x_l^2$$

$$\frac{\partial^2 Q}{\partial \beta_2^2} = \sum_{l=1}^{T} \Big\{ \delta_l (1-w_l^{\delta_l}) e^{\beta_2 x_l} \log \mathcal{F}_2(Y_l) + (1-\delta_l)(1-w_l^{1-\delta_l}) e^{\beta_2 x_l} \log \mathcal{F}_2(Y_l) \Big\} x_l^2$$

Thus to estimate (β_1, β_2) at the jth iteration we use

$$\begin{bmatrix} \beta_1^{(j)} \\ \beta_2^{(j)} \end{bmatrix} = \begin{bmatrix} \beta_1^{(j-1)} \\ \beta_2^{(j-1)} \end{bmatrix} + \begin{bmatrix} \frac{\partial^2 Q}{\partial \beta_1^2} & 0 \\ 0 & \frac{\partial^2 Q}{\partial \beta_2^2} \end{bmatrix}^{-1}_{\beta^{(j-1)}} \begin{bmatrix} \frac{\partial Q}{\partial \beta_1} \\ \frac{\partial Q}{\partial \beta_2} \end{bmatrix}_{\beta^{(j-1)}}$$

Thus

$$\beta_1^{(j)} = \beta_1^{(j-1)} + \left(\frac{\partial^2 Q}{\partial \beta_1^2} \right)^{-1} \frac{\partial Q}{\partial \beta_1} \Big|_{\beta^{(j-1)}}$$

and similarly for $\beta_2^{(j)}$. Now we can rewrite $\left(\frac{\partial^2 Q}{\partial \beta_1^2} \right)^{-1} \frac{\partial Q}{\partial \beta_1} \Big|_{\beta^{(j-1)}}$

as $\quad (\underline{X}' \omega_1^{(j-1)} \underline{X})^{-1} x' \underline{s}_1^{(j-1)}$

where $\quad \underline{X}' = (x_1, x_2, \ldots, x_T)$ ~~and~~ ,

$$\omega_1^{(j-1)} = \text{diag}\left[\omega_{11}^{(j-1)}, \ldots, \omega_{1T}^{(j-1)}\right], \cancel{\text{with}},$$

$$\underline{s}_1^{(j-1)} = \begin{bmatrix} s_{11}^{(j-1)} \\ \vdots \\ s_{1T}^{(j-1)} \end{bmatrix}, \text{ with}$$

$$\omega_{1i}^{(j-1)} = \delta_i \, \omega_i^{\delta_i} \cancel{(1+\ldots)} e^{\beta_i^{(j-1)}} \log F_1(Y_i) + (1-\delta_i)\, \omega_i^{1-\delta_i} e^{\beta_i^{(j-1)} x_i} \times \log F_1(Y_i)$$

$$s_{1i}^{(j-1)} = \delta_i \omega_i^{\delta_i}\left[1 + e^{\beta_1^{(j-1)} x_i} \log F_1(Y_i)\right] + (1-\delta_i)\,\omega_i^{1-\delta_i} e^{\beta_1^{(j-1)} x_i} \log F_1(Y_i)$$

Thus altogether in the EM-algorithm we have:

Start with initial value $\beta_1^0, \beta_2^0, p^0$

<u>Step 1</u> Set $(\beta_{1,r-1}, \beta_{2,r-1}, p_r) = (\beta_1^*, \beta_2^*, p^*)$

evaluate $\omega_i^{\delta_i}$ and $\omega_i^{1-\delta_i}$ (these probabilities/weights stay the same through the iterative least squares.

<u>Step 2</u> maximise $Q(\theta, \theta^*)$ by using the algorithm

$$p_r = \frac{a_r}{a_r + b_r} \quad \text{where } a_r, b_r \text{ are defined on page 5}$$

Now recurse $\quad \beta_1^{(j)} = \beta_1^{(j-1)} + (\underline{X}' \omega_1^{(j-1)} \underline{X})^{-1} \underline{X}' \underline{s}_1^{(j-1)}$ same for $\beta_2^{(j)}$

~~Step 3 Go back to step 1 until converge~~

$$\beta_2^{(j)} = \beta_2^{(j)} + (\underline{X}' W_2^{(j-1)} \underline{X})^{-1} \underline{X}' \underline{s}_2^{(j-1)}$$

iterate until convergence.

<u>Step 9</u> Let $\beta_{1r}, \beta_{2r}, p_r$ be the limit of the iterative least squares, go back to step 1 until convergence.