BIOS 760: Solution to Final 2010

1. (a) Note that since the exponential distribution is an exponential family, $\bar{T}_n \equiv \frac{1}{n}\sum_{i=1}^n T_i$ is complete and sufficient statistic. Since $E(\bar{T}_n) = \lambda$, $\bar{T}_n$ is the UMVUE. Since the variance of an exponential variable with rate λ is λ^2, we have

$$\mathrm{Var}(\bar{T}_n) = \frac{1}{n^2}\sum_{i=1}^n \mathrm{Var}(T_i) = \frac{\lambda^2}{n}\,.$$

We compute the information for one observation. Since $f(t) = \lambda^{-1}e^{-t/\lambda}$, $t > 0$, we have

$$\begin{aligned} l(t) &= -\log\lambda - t/\lambda \\ \dot{l}(t) &= -\lambda^{-1} + t/\lambda^2 = \lambda^{-2}(t-\lambda) \\ I(\lambda) &= E[\dot{l}^2] = \lambda^{-4}E[(T-\lambda)^2] = \lambda^{-4}\mathrm{Var}(T) = \lambda^{-2} \end{aligned}$$

We conclude that $I_n(\lambda) = n/\lambda^{-2}$, and hence the information bound $I^{-1}(\lambda) = \lambda^2/n$, i.e., the estimator $\bar{T}_n$ attains the information bound.

(b) The estimator $nT_{(1)}$ where $T_{(1)} = \min\{T_1, \ldots, T_n\}$ is unbiased estimator for λ. Note that the variance of $nT_{(1)}$ is $n^2\mathrm{Var}(T_{(1)}) = \lambda^2$. Hence, the variance of $nT_{(1)}$ is n times larger than the variance of $\bar{T}_n$, and thus $\bar{T}_n$ is better.

(c) Since T_n is complete and sufficient statistic, we need to find a function $g_k(\bar{T}_n)$ such that $E[g_k(\bar{T}_n)] = \lambda^k$. In other words, we need to find a function $g_k(\bar{T}_n)$ such that

$$\begin{aligned} \frac{1}{(n-1)!}\int_0^\infty g_k(t)t^{n-1}\lambda^{-n}e^{-t/\lambda}dt &= \lambda^k \\ \Leftrightarrow \frac{1}{(n-1)!}\int_0^\infty g_k(t)t^{n-1}\lambda^{-n-k}e^{-t/\lambda}dt &= 1 \end{aligned}$$

Substituting $g_k(t) = t^k(n-1)!/(n+k-1)!$ we have

$$\begin{aligned} \frac{1}{(n-1)!}\int_0^\infty g_k(t)t^{n-1}\lambda^{-n-k}e^{-t/\lambda}dt &= \frac{1}{(n-1)!}\int_0^\infty t^k\frac{(n-1)!}{(n+k-1)!}t^{n-1}\lambda^{-n-k}e^{-t/\lambda}dt \\ &= \frac{1}{(n+k-1)!}\int_0^\infty t^{n+k-1}\lambda^{-n-k}e^{-t/\lambda}dt = 1\,, \end{aligned}$$

where the last equality follows since the left hand side is an integral of the gamma density with parameters $n+k$ and λ.

2. (a) We start with the likelihood of (z_i, d_i) given λ. Assume first that $d_i = 1$, in that case $T_i = z_i$ and the likelihood of that $T_i = t_i$ and that $C_i \geq t_i$, by the independency is

$$f(t_i = z_i|\lambda)\cdot P(C > z_i) = \lambda^{-1}e^{z_i/\lambda}(1 - G(z_i)\,.$$

Assume now that $d_i = 0$, in that case $C_i = z_i$, and $T_i > z_i$. By the independency, the likelihood of $T_i > z_i$ and $C_i = z_i$ is given by

$$\lambda^{-1} e^{z_i/\lambda} \cdot g(z_i)\,.$$

Summarizing, the likelihood of $(Z_i = z_i, \Delta_i = d_i)$ is given by

$$\left(\lambda^{-1} e^{-z_i/\lambda}(1 - G(z_i))\right)^{d_i} \left(e^{-z_i/\lambda} g(z_i)\right)^{1-d_i}$$

(b) By the previous question,

$$\begin{aligned} l(\{z_i, \Delta_i\}|\lambda) &= \sum_{i=1}^{n} (\Delta_i(-\log\lambda - z_i/\lambda + \log(1 - G(z_i))) + (1 - \Delta_i) - z_i/\lambda + \log g(z_i)) \\ \dot{l}(\{z_i, \Delta_i\}|\lambda) &= -\frac{1}{\lambda}\sum_{i=1}^{n} \Delta_i + \frac{1}{\lambda^2}\sum_{i=1}^{n} z_i \end{aligned}$$

Hence the MLE is given by

$$\hat{\lambda} = \frac{\sum_{i=1}^{n} z_i}{\sum_{i=1}^{n} \Delta_i}\,.$$

(c) By question 2a, after substituting $\theta = \lambda^{-1}$, for one observation, we obtain

$$\begin{aligned} f(z, \Delta|\theta) &= \left(\theta e^{-z\theta}(1 - G(z))\right)^{\Delta} \left(e^{-z\theta} g(z)\right)^{1-\Delta} \\ \log f &= -\theta z + \log\theta\Delta + \Delta\log(1 - G(z)) + (1 - \Delta)\log g(z) \\ \dot{l} &= -z + \theta^{-1}\Delta \\ \ddot{l} &= -\theta^{-2}\Delta \\ I(\theta) &= -E[\ddot{l}] = \frac{E(\Delta)}{\theta^2} = \frac{P(T \le C)}{\theta^2}\,. \end{aligned}$$

Since the information bound for λ is $(\dot{\lambda}(\theta))^2 I(\theta)^{-1}$, we obtain that the information bound for λ is

$$I(\lambda)^{-1} = \left(\frac{-1}{\theta^2}\right)^2 \frac{\theta^2}{P(T \le C)} = \frac{\lambda^2}{P(T \le C)}\,.$$

(d) Note that $\hat{\theta} = \hat{\lambda}^{-1}$ is the MLE for θ. By the MLE theory,

$$\sqrt{n}(\hat{\theta} - \theta) \to_d N\left(0, \frac{\theta^2}{P(T \le C)}\right)\,.$$

Denote $\phi(\theta) = 1/\theta$. Hence, $\phi'(\theta) = -1/\theta^{-2}$. Let V be r.v. distributed $N\left(0, \frac{\theta^2}{P(T\le C)}\right)$. Using the delta-method, we obtain

$$\sqrt{n}(\hat{\lambda} - \lambda) = \sqrt{n}(\phi(\hat{\theta}) - \phi(\theta)) \to_d \phi'(\theta)V \sim N\left(0, \frac{\lambda^2}{P(T \le C)}\right)$$

3. (a) Note that when $d_i = 1$, this means that t_i equals z_i. However, when $d_i = 0$, $t_i > z_i$ and given z_i, using the memoryless property of exponential we have that $f(t_i|z_i, d_i = 0, \lambda) = \lambda^{-1}e^{-(t_i-z_i)/\lambda}$. Summarizing we have

$$f(t_i|z_i, d_i, \lambda) = d_i 1_{\{t_i=z_i\}} + (1-d_i)1_{\{t_i>z_i\}}\lambda^{-1}e^{-(t_i-z_i)/\lambda}.$$

(b) First note that

$$f(t_i, z_i, d_i|\lambda) = f(z_i, d_i|\lambda)f(t_i|z_i, d_i\lambda)\,.$$

$f(z_i, d_i|\lambda)$ is given by question 2a. $f(t_i|z_i, d_i\lambda)$ is given by previous question. Multiplying, we obtain

$$\begin{aligned} f(t_i, z_i, d_i|\lambda) &= \left(\lambda^{-1}e^{-t_i/\lambda}(1-G(t_i))\right)^{d_i 1_{\{z_i=t_i\}}}\left(e^{-z_i/\lambda}\lambda^{-1}e^{(t_i-z_i)/\lambda}g(z_i)\right)^{(1-d_i)1_{\{t_i>z_i\}}} \\ &= (\lambda^{-1}e^{-t_i/\lambda})^{1_{\{t_i\ge z_i\}}}(1-G(t_i))^{d_i 1_{\{z_i=t_i\}}}g(z_i)^{(1-d_i)1_{\{t_i>z_i\}}} \end{aligned}$$

The result follows by taking log.

(c) The E-step of an E-M algorithm consists on computing $E[\sum_{i=1}^n \log f(t_i, z_i, d_i|\lambda)|\lambda^{(k)}, \{z_i, d_i\}]$. Noting that for each expression $\log f(t_i, z_i, d_i|\lambda)$, the expectation is with respect to the density $f(t_i|\lambda^{(k)}, z_i, d_i)$. Hence we can write

$$E[\sum_{i=1}^n \log f(t_i, z_i, d_i|\lambda)|\lambda^{(k)}, \{z_i, d_i\}] = \sum_{i=1}^n \int \log f(t_i, z_i, d_i|\lambda)f(t_i|\lambda^{(k)}, z_i, d_i)dt\,.$$

Substituting we have

$$\begin{aligned} E[\log f(Y|\lambda)|\lambda^{(k)}, Y_{obs}] &= \sum_{i=1}^n \int \log f(t_i, z_i, d_i|\lambda)f(t_i|\lambda^{(k)}, z_i, d_i)dt \\ &= \sum_{i=1}^n (d_i\log(1-G(z_i)) + (1-d_i)g(z_i)) - n\log\lambda - \sum_{i=1}^n d_i\frac{z_i}{\lambda} \\ &\quad -\sum_{i=1}^n \frac{(1-d_i)}{\lambda}\int_{z_i}^{\infty}\frac{(t_i-z_i+z_i)}{\lambda^{(k)}}e^{(t_i-z_i)/\lambda^{(k)}}dt_i \\ &= \sum_{i=1}^n (d_i\log(1-G(z_i)) + (1-d_i)g(z_i)) - n\log\lambda - \sum_{i=1}^n d_i\frac{z_i}{\lambda} \\ &\quad -\frac{\lambda^{(k)}}{\lambda}\sum_{i=1}^n(1-d_i) - \sum_{i=1}^n(1-d_i)\frac{z_i}{\lambda}\,. \end{aligned}$$

(d) Taking derivative of $E[\log f(Y|\lambda)|\lambda^{(k)}, Y_{obs}]$, that was obtained in the previous question, with respect to λ, and equating to zero, we obtain

$$\begin{aligned} -\frac{n}{\lambda} + \frac{1}{\lambda^2}\sum_{i=1}^n z_i + \frac{\lambda^{(k)}}{\lambda^2}\sum_{i=1}^n(1-d_i) &= 0 \\ \Leftrightarrow \lambda^{(k+1)} &= \frac{1}{n}\sum_{i=1}^n z_i + \frac{\lambda^{(k)}}{n}\sum_{i=1}^n(1-d_i)\,. \end{aligned}$$

The algorithm consists on choosing an inial value $\lambda^{(1)}$ for λ. Then iterating between calculating the E-step as in the previous question based on the estimated $\lambda^{(k)}$, and then finding $\lambda^{(k+1)}$. The algorithm stops when the difference in the likelihood of $\lambda^{(k)}$ and $\lambda^{(k+1)}$ is less than a given criterion.

(e) Substituting $\lambda = \lambda^{(k)} = \lambda^{(k+1)}$, we obtain that

$$\tilde{\lambda} = \frac{\sum_{i=1}^{n} z_i}{\sum_{i=1}^{n} \Delta_i} = \hat{\lambda}, .$$

Note that for each $\lambda^{(k)}$ that is not a maximum of the likelihood, the iteration of the EM algorithm increases the observed likelihood function. Thus, since in a fixed point the likelihood does not increase, the fixed point is a maximum point of the likelihood.