# Stat 513
# Homework key 4

February 3, 2016

### REGULAR PROBLEMS

1. Suppose $Y_1, \dots, Y_n$ are i.i.d. lognormal random variables with $Y_i \sim \mathcal{LN}(\mu, \sigma^2)$ Denote the unknown parameter vector by $\theta = (\mu, \sigma^2)$.

   (a) Derive the asymptotic distribution for the maximum likelihood estimate $\hat{\eta}$ of $\eta = EY_i$.

   (b) Derive the asymptotic distribution for the nonparametric estimate $\tilde{\eta}$of $\eta$.

   (c) Find the asymptotic relative efficiency

$$e(\tilde{\eta}, \hat{\eta}) = \frac{Var(\hat{\eta})}{Var(\tilde{\eta})}.$$

**Ans:**

(a) Define

$$X_i = \log Y_i \sim \mathcal{N}(\mu, \sigma^2).$$

Hence, the MLE of $\hat{\mu}$ and $\hat{\sigma^2}$ are going to be $\bar{X}$ and $S_x = \dfrac{\sum_{i=1}^n (x_i - \bar{X})^2}{n}$ respectively. Now,

$$\hat{\eta} = e^{\hat{\mu} + \hat{\sigma^2}/2} = g(\mu, \hat{\sigma^2}/2)$$

where $g(x, y) = \exp(x + \frac{y}{2})$. We will now use Delta method to get the asymptotic distribution of $\hat{\eta}$. Recall from HW 1 that

$$\sqrt{n} \begin{bmatrix} \bar{X} - \mu \\ \hat{\sigma^2} - \sigma^2 \end{bmatrix} \xrightarrow{\mathcal{D}} \mathcal{N}_2\left(0, \begin{bmatrix} \sigma^2 & \gamma_3 \\ \gamma_3 & \gamma_4 - \sigma^4 \end{bmatrix}\right).$$

When $X_i-$s are normal, $\gamma_3 = 0$ and $\gamma_4 = \sigma^4$. Substituting the values we get that

$$\sqrt{n} \begin{bmatrix} \bar{X} - \mu \\ \hat{\sigma^2} - \sigma^2 \end{bmatrix} \xrightarrow{\mathcal{D}} \mathcal{N}_2\left(0, \begin{bmatrix} \sigma^2 & 0 \\ 0 & 2\sigma^4 \end{bmatrix}\right).$$

Now we will be able to use Delta method with $g(x, y) = \exp(x + \frac{y}{2})$ which gives $g(\mu, \sigma^2) = \eta$. Hence, we obtain

$$\sqrt{n}(\hat{\eta} - \eta) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \nabla g^T \begin{bmatrix} \sigma^2 & 0 \\ 0 & 2\sigma^4 \end{bmatrix} \nabla g\Big|_{\mu,\sigma^2}\right)$$

. Now $\nabla g\Big|_{\mu,\sigma^2} = \exp(\mu + \sigma^2/2)\begin{bmatrix}1, & 1/2\end{bmatrix}^T$. Therefore, we get

$$\sqrt{n}(\hat{\eta} - \eta) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, e^{2\mu+\sigma^2}(\sigma^2 + \sigma^4/2)\right).$$

(b) The non-parametric estimate of $\eta$ is $\bar{Y}$. Hence $\tilde{\eta} = \hat{Y}$. By CLT, we get that

$$\sqrt{n}(\tilde{\eta} - \eta) \xrightarrow{\mathcal{D}} \mathcal{N}(0, Var(Y_i)).$$

Since $Var(Y_i) = (e^{\sigma^2} - 1)e^{2\mu+\sigma^2}$,

$$\sqrt{n}(\tilde{\eta} - \eta) \xrightarrow{\mathcal{D}} \mathcal{N}(0, (e^{\sigma^2} - 1)e^{2\mu+\sigma^2})$$

(c)

$$e(\tilde{\lambda}, \hat{\lambda}) = \frac{Var(\hat{\lambda})}{Var(\tilde{\lambda})} = \frac{e^{2\mu+\sigma^2}(\sigma^2 + \sigma^4/2)}{(e^{\sigma^2} - 1)e^{2\mu+\sigma^2}} = \frac{\sigma^2 + \sigma^4/2}{e^{\sigma^2} - 1}.$$

Notice that

$$e^{\sigma^2} - 1 = (\sigma^2 + \frac{\sigma^4}{2!}) + \frac{\sigma^6}{3!} + \dots$$

Hence, $e(\tilde{\lambda}, \hat{\lambda}) < 1$ for $\sigma^2 > 0$.

2. Suppose $Y_1, \dots, Y_n$ are i.i.d. lognormal random variables with $Y_i \sim \mathcal{E}(\lambda)$ with hazard $\lambda$. Further define random variables $W_i = 1_{[Y_i \le c]}$ for scalar $c$.

   (a) Derive the asymptotic distribution for the maximum likelihood estimate $\hat{\lambda}$ of $\lambda$.

   (b) Derive the asymptotic distribution for the maximum likelihood estimate $\tilde{\lambda}$ of $\lambda$ using only $W_i$-s.

   (c) Find the asymptotic relative efficiency

$$e(\tilde{\lambda}, \hat{\lambda}) = \frac{Var(\hat{\lambda})}{Var(\tilde{\lambda})}.$$

**Ans:**

(a) The log-likelihood

$$l(\lambda | y_1, \dots, y_n) = n \log \lambda - \lambda \sum_{i=1}^{n} y_i$$

$$\Rightarrow \frac{\partial}{\partial \lambda} l(\lambda) = \frac{n}{\lambda} - \sum_{i=1}^{n} y_i$$

Setting $\frac{\partial}{\partial \lambda} l(\lambda) = 0$ we get that the MLE $\hat{\lambda}$ is $\frac{1}{\bar{Y}}$. By CLT,

$$\sqrt{n}(\bar{Y} - \frac{1}{\lambda}) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \frac{1}{\lambda^2}).$$

We will now use Delta method with $g(x) = \frac{1}{x}$. Hence, $g(\bar{Y}) = \hat{\lambda}$. Also

$$g'(x)^2 \Big|_{x=1/\lambda} = \lambda^4.$$

This gives

$$\sqrt{n}(\hat{\lambda} - \lambda) \xrightarrow{\mathcal{D}} \mathcal{N}(0, g'(1/\lambda)^2 \lambda^2)$$

or

$$\sqrt{n}(\hat{\lambda} - \lambda) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \lambda^2).$$

(b) It is clear that $W_i \sim Ber(p)$ where $p = 1 - e^{-\lambda c}$. This gives

$$\lambda = -\frac{\log(1-p)}{c} \tag{1}$$

or equivalently

$$\tilde{\lambda} = -\frac{\log(1-\hat{p})}{c} \tag{2}$$

where $\hat{p}$ is the maximum likelihood of $p$, namely $\bar{W}$. Also by CLT,

$$\sqrt{n}(\bar{W} - p) \xrightarrow{\mathcal{D}} \mathcal{N}(0, p(1-p)).$$

Therefore we need to apply delta method again with $g(p) = -\frac{\log(1-p)}{c}$ to get the asymptotic distribution of $\tilde{\lambda}$.

$$g'(p) = \frac{1}{c(1-p)}.$$

$$\sqrt{n}(g(\bar{W}) - g(p)) \xrightarrow{\mathcal{D}} \mathcal{N}(0, g'(p)^2 p(1-p))$$

or

$$\sqrt{n}(g(\bar{W}) - g(p)) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \frac{p}{c^2(1-p)}\right).$$

By (2)

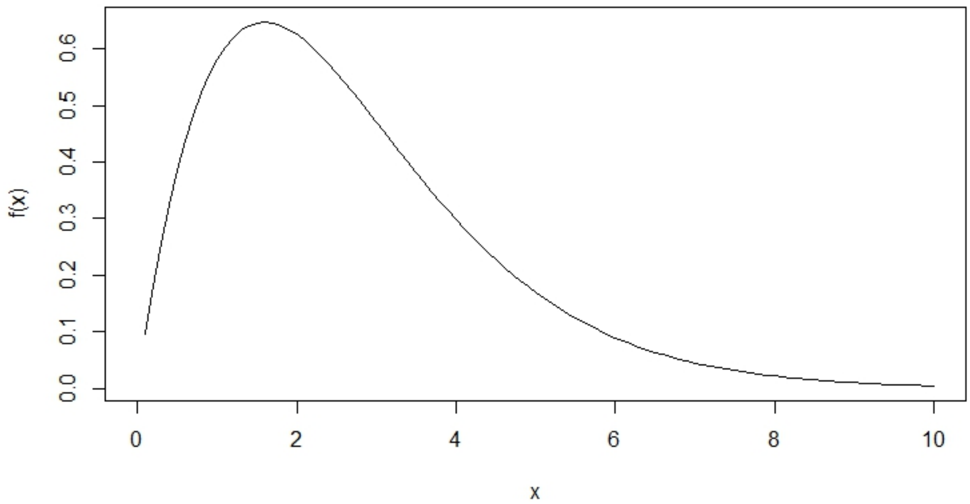
$$\sqrt{n}(\tilde{\lambda} - \lambda) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \frac{p}{c^2(1-p)}\right).$$

Since $p = 1 - e^{-\lambda c}$, $\frac{p}{1-p} = e^{\lambda c} - 1$.

(c)

$$e(\tilde{\lambda}, \hat{\lambda}) = \frac{Var(\hat{\lambda})}{Var(\tilde{\lambda})} = \frac{\lambda^2 c^2}{e^{\lambda c} - 1}.$$

Let $f(x) = \dfrac{x^2}{e^x - 1}$. Then $e(\tilde{\lambda}, \hat{\lambda}) = f(\lambda c)$. let us consider the behavior of $f(x)$ for positive x. For large value of $x$, $f(x) \to 0$. For smaller $x$, $f(x)$ behaves as follows:




Therefore we see that the efficiency is again less than 1. Also, the efficiency reaches its peak when $\lambda c$ is near 2. A better estimate of the location of the maximum can be found out by numerical methods. So for fixed $\lambda$, as $c$ increases the ARE first increases, then slowly decreases to 0.

## MORE INVOLVED PROBLEMS

3. Suppose $Y_1, Y_2, \ldots, Y_n$ are i.i.d. continuous random variables with cdf $F_Y(y|\theta)$ and density $f_Y(y|\theta)$. Let $\eta = F^{-1}(p)$ be the $p$-th quantile of the distribution of $Y$, and suppose that $f(\eta) > 0$. Then, the sample quantile $\tilde{\eta}$ computed using $\vec{Y}_n = (Y_1, \ldots, Y_n)$ can be shown to have asymptotic distribution

$$\sqrt{n}(\tilde{\eta} - \eta) \to_d \mathcal{N}\left(0, \frac{p(1-p)}{f^2(\eta|\theta)}\right).$$

Suppose we are interested in estimating medians of the distribution (so $p = 0.5$). For each of the following continuous distributions, find the asymptotic relative efficiency of the sample median $\tilde{\eta}$ compared to the MLE $\hat{\eta}$ of $\eta$.

(a) Normal distribution: $Y_i \sim \mathcal{N}(\mu, \sigma^2)$.

(b) Exponential distribution: $Y_i \sim \mathcal{E}(\lambda)$.

(c) $Y_i$ has density

$$f_Y(y|\sigma^2) = \frac{y}{\sigma^2} \exp\left\{-\frac{y^2}{2\sigma^2}\right\} \mathbf{1}_{(0,\infty)}(y).$$

(d) Uniform distribution: $Y_i \sim \mathcal{U}(0, \theta)$. (Find the ratio of variances for the approximate distribution based on the asymptotic distributions.)

**Ans:**
To start, we specialize the asymptotic distribution of $\tilde{\eta}$ to the median ($p = 0.5$),

$$\sqrt{n}(\tilde{\eta} - \eta) \to_d \mathcal{N}\left(0, \frac{1}{4f^2(\eta|\theta)}\right).$$

Next, note that in parts (a-c) the distribution of $Y_i$ belongs to a regular parametric model, so that $\hat{\theta}$ (the MLE of $\theta$), has asymptotic distribution

$$\sqrt{n}(\hat{\theta} - \theta) \to_d \mathcal{N}(0, \mathbf{J}(\theta)^{-1}),$$

where $\mathbf{J}(\theta)$ is the information matrix (number) for $\theta$. Expressing the median $\eta = g(\theta)$ as a differentiable function of $\theta$, we can apply the delta method to obtain an asymptotic distribution for the MLE $\hat{\eta} = g(\hat{\theta})$ of the median (recalling the invariance of the MLE under transformations of the parameter $\theta$),

$$\sqrt{n}(\hat{\eta} - \eta) = \sqrt{n}(g(\hat{\theta}) - g(\theta)) \to_d \mathcal{N}(0, g'(\theta)^2 \mathbf{J}(\theta)^{-1}).$$

This means that we can express the ARE for (a-c) as

$$ARE(\hat{\eta}, \tilde{\eta}) = \frac{g'(\theta)^2 \mathbf{J}(\theta)^{-1}}{(4f^2(\eta|\theta))^{-1}} = 4f^2(\eta|\theta) g'(\theta)^2 \mathbf{J}(\theta)^{-1}.$$

(a) Since the $\mathcal{N}(\mu, \sigma^2)$ distribution is symmetric about $\mu$, we have that $\theta = \mu = \eta$. We have already seen that $\hat{\mu} = n^{-1} \sum_{i=1}^n Y_i$ is the MLE and has asymptotic distribution

$$\sqrt{n}(\hat{\mu} - \mu) = \sqrt{n}(\hat{\eta} - \eta) \to_d \mathcal{N}(0, \sigma^2).$$

Furthermore, we have

$$f^2(\eta|\theta) = f^2(\mu|\mu) = \left(\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(\mu-\mu)^2}{2\sigma^2}\right\}\right)^2 = \frac{1}{2\pi\sigma^2}.$$

Substituting into our expression for ARE, we obtain

$$ARE(\hat{\eta}, \tilde{\eta}) = 4f^2(\eta|\theta) g'(\theta)^2 \mathbf{J}(\theta)^{-1} = 4\frac{1}{2\pi\sigma^2}\sigma^2 = \frac{2}{\pi} \approx 0.6366 < 1.$$

(b) We consider the hazard parametrization for $\mathcal{E}(\lambda)$, so that

$$F_Y(y|\lambda) = \left(1 - e^{-\lambda y}\right) \mathbf{1}_{(0,\infty)}(y).$$

The median $\eta$ is found to be

$$\begin{aligned}\eta &= F^{-1}(1/2|\lambda)\\ &= \lambda^{-1}\log 2\\ &= g(\lambda),\end{aligned}$$

so that $g'(\lambda) = -\lambda^2 \log 2$ and

$$f^2(\eta|\theta) = f^2(\lambda^{-1}\log 2|\lambda) = \frac{\lambda^2}{4}.$$

Recall that the MLE for $\lambda$ in this model is $\hat{\lambda} = \bar{Y}_n^{-1}$, so that $\hat{\eta} = (\log 2)\bar{Y}_n$ (by invariance of the MLE). By the CLT,

$$\sqrt{n}(\hat{\eta} - \eta) = \log 2\sqrt{n}(\bar{Y}_n - \lambda^{-1}) \to_d \mathcal{N}(0, (\log 2)^2\lambda^{-2}) = \mathcal{N}(0, g'(\lambda)^2\mathbf{J}(\lambda)^{-1}).$$

Substituting into our expression for ARE, we obtain

$$ARE(\hat{\eta}, \tilde{\eta}) = 4f^2(\eta|\theta)g'(\theta)^2\mathbf{J}(\theta)^{-1} = 4\frac{\lambda^2}{4}(\log 2)^2\lambda^{-2} = (\log 2)^2 \approx 0.4805 < 1.$$

(c) We have here that $F_Y(y) = \left(1 - \exp\left\{-\frac{y^2}{2\sigma^2}\right\}\right)\mathbf{1}_{(0,\infty)}(y)$, so that

$$\eta = \sigma\sqrt{2\log 2} = g(\sigma), \quad g'(\sigma) = \sqrt{2\log 2}, \quad \text{and}$$

$$f(\eta|\theta) = f(\sigma\sqrt{2\log 2}|\sigma) = \frac{\sqrt{2\log 2}}{2\sigma}$$

To find the information for $\sigma$, we derive the variance of the efficient score. First, we find that the score is

$$\dot{\ell}(\sigma|y) = \frac{\partial}{\partial\sigma}\left(\log y - 2\log\sigma - \frac{y^2}{2}\sigma^{-2}\right) = \frac{2}{\sigma}\left(\frac{y^2}{2\sigma^2} - 1\right).$$

Next, we find the variance of the score to be

$$\begin{aligned}\mathbf{J}(\sigma) &= Var(\dot{\ell}(\sigma|Y))\\ &= E(\dot{\ell}(\sigma|Y)^2)\\ &= \frac{4}{\sigma^2}E\left(\frac{Y^2}{2\sigma^2} - 1\right)^2\\ &= \frac{4}{\sigma^2}\int_0^\infty \left(\frac{y^2}{2\sigma^2} - 1\right)^2 \frac{y}{\sigma^2}\exp\left\{-\frac{y^2}{2\sigma^2}\right\}dy\\ &= \frac{4}{\sigma^2}\int_0^\infty (t-1)^2 e^{-t}dt\\ &= \frac{4}{\sigma^2}.\end{aligned}$$

The integration is made easy upon transforming to $T = Y^2/(2\sigma^2)$, recognizing the resultant integral as the variance of an $\mathcal{E}(1)$ random variable. Substituting into our expression for ARE, we obtain

$$ARE(\hat{\eta}, \tilde{\eta}) = 4f^2(\eta|\theta)g'(\theta)^2\mathbf{J}(\theta)^{-1} = 4\frac{2\log 2}{4\sigma^2}(2\log 2)\frac{\sigma^2}{4} = \frac{(2\log 2)^2}{4} = (\log 2)^2 \approx 0.4805 < 1.$$

(d) Now, we have that $\eta = \theta/2$ and $f(\eta|\theta) = \frac{1}{\theta}$. Since this family of distributions lacks common support in terms of $\theta$, our argument for the MLE $\hat{\theta}$ differs slightly. Recall that $\hat{\theta} = \max_{1\le i\le n} Y_i = Y_{(n)}$ and that

$$n(\theta - Y_{(n)}) \to_d \mathcal{E}(\theta),$$

with the mean parametrization for the exponential distribution. It follows that

$$n(\eta - \hat{\eta}) \to_d \mathcal{E}(\theta/2),$$

with $\hat{\eta} = \hat{\theta}/2$. We should account for the rates at which $\tilde{\eta}$ and $\hat{\eta}$ converge ($n^{1/2}$ and $n$, respectively) when comparing their asymptotic variances. Hence, we consider that for large enough $n$,

$$\tilde{\eta}\dot{\sim}(\eta, \frac{\theta^2}{4n}), \quad \hat{\eta}\dot{\sim}(\eta, \frac{\theta^2}{n^2}).$$

Taking the ratio of the above approximate variances, we have

$$ARE(\hat{\eta}, \tilde{\eta}) = \frac{\theta^2/n^2}{\theta^2/(4n)} = \frac{4}{n} \to 0.$$

Since this approximation is only valid for large values of $n$, we see that the MLE $\hat{\eta}$ is again more efficient than the non-parametric estimator $\tilde{\eta}$.