Statistics 252 "Practice" Midterm Solutions – Winter 2005

1. (a) Since

$$\log f(y|\theta) = 2\log(\theta) - \theta^2 y,$$

we find

$$\frac{\partial^2}{\partial\theta^2}\log f(y|\theta) = \frac{-2}{\theta^2} - 2y.$$

Thus,

$$I(\theta) = -\mathbb{E}\left(\frac{\partial^2}{\partial\theta^2}\log f(Y|\theta)\right) = \frac{2}{\theta^2} + 2\mathbb{E}(Y) = \frac{4}{\theta^2}$$

since $\mathbb{E}(Y) = \theta^{-2}$. (This is because $Y \sim \text{Exp}(\theta^{-2})$.)

(b) Let $Z = \theta^2 Y$ so that

$$P(Z \le z) = P(Y \le \theta^{-2}z) = \int_0^{\theta^{-2}z} \theta^2 e^{-\theta^2 y}\,dy = 1 - e^{-z}.$$

Thus, we must find a and b so that

$$\int_0^a e^{-z}\,dz = \alpha_1 \quad \text{and} \quad \int_b^\infty e^{-z}\,dz = \alpha_2.$$

Computing the integrals we find $a = -\log(1-\alpha_1)$ and $b = -\log(\alpha_2)$. Hence,

$$\begin{aligned} 1 - (\alpha_1 + \alpha_2) = P(a \le Z \le b) &= P(-\log(1-\alpha_1) \le \theta^2 Y \le -\log(\alpha_2)) \\ &= P\left(\frac{-\log(1-\alpha_1)}{Y} \le \theta^2 \le \frac{-\log(\alpha_2)}{Y}\right). \end{aligned}$$

In other words,

$$\left(\frac{-\log(1-\alpha_1)}{Y}, \frac{-\log(\alpha_2)}{Y}\right)$$

is a confidence interval for θ^2 with coverage probability $1 - (\alpha_1 + \alpha_2)$.

2. (a) The likelihood function is

$$L(\theta) = \prod_{i=1}^n f(y_i|\theta) = (\theta - 1)^n \left(\prod_{i=1}^n y_i\right)^{-\theta}$$

so that the log-likelihood function is

$$\ell(\theta) = n\log(\theta - 1) - \theta\sum_{i=1}^n \log(y_i).$$

Hence $\ell'(\theta) = 0$ implies

$$0 = \frac{n}{\theta - 1} - \sum_{i=1}^n \log(y_i).$$

Since

$$\ell''(\theta) = -\frac{n}{(\theta-1)^2} < 0,$$

we conclude that

$$\hat{\theta}_{\text{MLE}} = 1 + \frac{n}{\sum_{i=1}^n \log(Y_i)}.$$

(b) Since

$$\log f(y|\theta) = \log(\theta - 1) - \theta \log(y),$$

we find

$$\frac{\partial^2}{\partial\theta^2} \log f(y|\theta) = \frac{-1}{(\theta - 1)^2}.$$

Thus,

$$I(\theta) = -\mathbb{E}\left(\frac{\partial^2}{\partial\theta^2} \log f(Y|\theta)\right) = \frac{1}{(\theta - 1)^2}.$$

(c) An approximate 90% confidence interval for θ is given by

$$\hat{\theta}_{\text{MLE}} \pm 1.64 \frac{1}{\sqrt{n\, I(\hat{\theta}_{\text{MLE}})}}.$$

Since $n = 25$ and $\sum \log y_i = 5$, we conclude that

$$\hat{\theta}_{\text{MLE}} = 1 + \frac{25}{5} = 6$$

and

$$I(\hat{\theta}_{\text{MLE}}) = \frac{1}{(\hat{\theta}_{\text{MLE}} - 1)^2} = \frac{1}{25}.$$

Hence, an approximate 95% confidence interval for θ is

$$6 \pm 1.64.$$

3. To find the method of moments estimators for λ and θ, we must solve the system of equations

$$\mathbb{E}(Y) = \overline{Y} \quad \text{and} \quad \mathbb{E}(Y^2) = \frac{1}{n}\sum_{i=1}^{n} Y_i^2$$

Thus, some trivial algebra gives

$$\hat{\theta}_{\text{MOM}} = \overline{Y} \quad \text{and} \quad \hat{\lambda}_{\text{MOM}} = \sqrt{\frac{2n}{\sum_{i=1}^{n} Y_i^2}}.$$

4. (a) Consider a population. A parameter is a number which summarizes the population. In general, this number is unknown. In Stat 252, the most common example is when the population is summarized in terms of a density function $f(y|\theta)$ which is parametrized by an unknown parameter θ. An estimator, therefore, is a rule for calculating an estimate of the parameter based on a random sample from the population. That is, if $Y_1, \ldots, Y_n$ constitute an iid collection of random variables each with $Y_i \sim f(y|\theta)$, then *any* random variable $g(Y_1, ..., Y_n)$ is an estimator of θ. (Of course, we prefer to work with minimum variance unbiased estimators, but that is not a requirement of the definition.)

(b) It is desirable to find unbiased estimators because by having an unbiased estimator we know $\mathbb{E}(\hat{\theta}) = \theta$. Thus, the most likely "value" of $\hat{\theta}$ is θ. If we have the unbiased estimator with the smallest variance, then the distribution of $\hat{\theta}$ is clustered as tightly as possible about its mean, namely θ. Thus, the MVUE is the "most likely" of all unbiased estimators to "closest" to θ.

5. (a) If $Y \sim \text{Unif}(0, \theta)$, then $\mathbb{E}(Y) = \theta/2$ and $\text{Var}(Y) = \theta^2/12$. Thus,

$$\hat{\theta}_{\text{MOM}} = 2\overline{Y}.$$

Since

$$\mathbb{E}(\hat{\theta}_{\text{MOM}}) = 2\mathbb{E}(\overline{Y}) = 2\mathbb{E}(Y_1) = 2\frac{\theta}{2} = \theta$$

we conclude that $\hat{\theta}_{\text{MOM}}$ is an unbiased estimator of θ.

(b) In order to find $\mathbb{E}(\hat{\theta}_{\text{MLE}})$ we must find the density function of $\hat{\theta}_{\text{MLE}}$. Now,

$$P(\hat{\theta}_{\text{MLE}} \le t) = \left[\int_0^t \theta^{-1}\, dy\right]^{10} = \frac{t^{10}}{\theta^{10}}, \quad 0 \le t \le \theta$$

so that $f(t) = 10\,\theta^{-10}\, t^9$, $0 \le t \le \theta$. Thus,

$$\mathbb{E}(\hat{\theta}_{\text{MLE}}) = \int_0^\theta 10\,\theta^{-10}\, t^{10}\, dt = \frac{10}{11}\,\theta.$$

Thus, an unbiased estimator of θ which is a function of the MLE is given by

$$\hat{\theta}_B = \frac{11}{10}\,\max(Y_1, \ldots, Y_{10}).$$

Also, note that

$$\mathbb{E}(\hat{\theta}^2_{\text{MLE}}) = \int_0^\theta 10\,\theta^{-10}\, t^{11}\, dt = \frac{10}{12}\,\theta^2.$$

(c) From **(a)**, we conclude

$$\text{Var}(\hat{\theta}_{\text{MOM}}) = 4\,\text{Var}(\overline{Y}) = \frac{4}{10}\,\text{Var}(Y_1) = \frac{4\theta^2}{10 \cdot 12} = \frac{\theta^2}{30}.$$

¿From **(b)**, we conclude

$$\text{Var}(\hat{\theta}_B) = \frac{121}{100}\,\text{Var}(\max(Y_1, \ldots, Y_{10})) = \frac{121}{100}\left(\frac{10}{12} - \frac{100}{121}\right)\theta^2 = \frac{\theta^2}{120}.$$

Thus,

$$\text{eff}(\hat{\theta}_{\text{MOM}}, \hat{\theta}_B) = \frac{\text{Var}(\hat{\theta}_B)}{\text{Var}(\hat{\theta}_{\text{MOM}})} = \frac{1}{4}.$$

Since both $\hat{\theta}_{\text{MOM}}$ and $\hat{\theta}_B$ are unbiased, the one with the smaller variance is preferrable, namely $\hat{\theta}_B$.

(d) Since

$$\log f(y|\theta) = -\log(\theta),$$

we find

$$\frac{\partial^2}{\partial\theta^2}\log f(y|\theta) = \frac{1}{\theta^2}$$

Thus,

$$I(\theta) = -\mathbb{E}\left(\frac{\partial^2}{\partial\theta^2}\log f(Y|\theta)\right) = -\frac{1}{\theta^2}.$$

(e) The Cramer-Rao inequality tells us that that if $\hat{\theta}$ is any unbiased estimator of θ based on $(Y_1, \ldots, Y_{10})$, then

$$\mathrm{Var}(\hat{\theta}) \geq \frac{1}{10\, I(\theta)} = \frac{-\theta^2}{10}.$$

Of course, for any random variable X, $\mathrm{Var}(X) \geq 0$. Thus, having a negative lower bound in the C-R inequality is useless. It give us no new information.

The problem in this question arises from the fact that the density function of a uniform random variable is discontinuous. Therefore, technically, the computation of the Fisher inequality is invalid. My reason for asking you this question was to draw your attention to this important fact.