Math 164 HW 5

Andrew Zastovnik

March 24, 2016

Problem 1

Let $X_1, ..., X_n$ be a random sample from a distribution with probability mass function

$$f(a;\theta) = \theta(1-\theta)^a \mathbb{1}_{\{0,1,...\}}(a),$$

where $0 < \theta < 1$.

(a) Find the maximum likelihood estimator of θ.
For X_i in $\{0, 1, ...\}$

$$\frac{d}{d\theta} log[f(X_1, ..., X_n;\theta)] = \frac{d}{d\theta} log[\prod_{i=1}^{n} \theta(1-\theta)^{X_i}] = \frac{d}{d\theta} \sum_{i=1}^{n} [log(\theta) + X_i log(1-\theta)]$$

$$= \sum_{i=1}^{n} [\frac{1}{\theta} - \frac{X_i}{1-\theta}] = [\frac{n}{\theta} - \frac{\sum_{i=1}^{n} X_i}{1-\theta}]$$

Set equal to 0

$$\frac{n}{\hat{\theta}} - \frac{\sum_{i=1}^{n} X_i}{1-\hat{\theta}} = 0 \Rightarrow \frac{n}{\hat{\theta}} = \frac{\sum_{i=1}^{n} X_i}{1-\hat{\theta}} \Rightarrow (1-\hat{\theta})n = \hat{\theta}\sum_{i=1}^{n} X_i \Rightarrow 1-\hat{\theta} = \frac{1}{n}\hat{\theta}\sum_{i=1}^{n} X_i \Rightarrow$$

$$\frac{1}{\hat{\theta}} - 1 = \frac{1}{n}\sum_{i=1}^{n} X_i \Rightarrow \frac{1}{\hat{\theta}} = \frac{1}{n}\sum_{i=1}^{n} X_i + 1$$

$$\Rightarrow \hat{\theta} = \frac{1}{\frac{1}{n}\sum_{i=1}^{n} X_i + 1}$$

(b) Find the maximum likelihood estimator of

$$\tau(\theta) = E(X_1) = \frac{1-\theta}{\theta}$$

From Proposition 1 we know that

$$\widehat{\tau(\theta)} = \tau(\hat{\theta}) = \frac{1 - \frac{1}{\frac{1}{n}\sum_{i=1}^{n} X_i + 1}}{\frac{1}{\frac{1}{n}\sum_{i=1}^{n} X_i + 1}} = \frac{\frac{\frac{1}{n}\sum_{i=1}^{n} X_i}{\cancel{\frac{1}{n}\sum_{i=1}^{n} X_i + 1}}}{\frac{1}{\cancel{\frac{1}{n}\sum_{i=1}^{n} X_i + 1}}}$$

$$= \frac{1}{n}\sum_{i=1}^{n} X_i$$

(c) Find the Cramr-Rao lower bound for variances of unbiased estimators of $\tau(\theta)$.
Show $f(a;\theta)$ is one parameter exponential family

$$\theta(1-\theta)^a \mathbb{1}_{\{0,1,...\}}(a) = \mathbb{1}_{\{0,1,...\}}(a)\theta e^{a log(1-\theta)} = h(a)c(\theta)e^{t(a)q(\theta)}$$

and that $\frac{d}{d\theta}q(\theta)$ is continuous for $0 < \theta < 1$

$$\frac{d}{d\theta}q(\theta) = \frac{d}{d\theta} log(1-\theta) = \frac{-1}{1-\theta}$$

Since this is continuous and $f(a;\theta)$ is one parameter exponential family R1-R4 hold Therefore

$$I_n(\theta) = -E\Big[\frac{d^2}{d\theta^2} log[f(X_1, ..., X_n;\theta)]\Big] = -E\Big[\frac{d}{d\theta}\Big(\frac{n}{\theta} - \frac{\sum_{i=1}^{n} X_i}{1-\theta}\Big)\Big] = -E\Big[\frac{-n}{\theta^2} - \frac{\sum_{i=1}^{n} X_i}{(1-\theta)^2}\Big]$$

$$= \frac{n}{\theta^2} + \frac{E(\sum_{i=1}^n X_i)}{(1-\theta)^2} = \frac{n}{\theta^2} + \frac{\sum_{i=1}^n E(X_i)}{(1-\theta)^2} = \frac{n}{\theta^2} + \frac{\sum_{i=1}^n \frac{\cancel{1-\theta}}{\theta}}{(1-\theta)^{\cancel{2}}} = \frac{n}{\theta^2} + \frac{n}{\theta(1-\theta)}$$

$$= \frac{n}{\theta^2} 1 + \frac{\theta}{(1-\theta)} = \frac{n}{\theta^2} \frac{(1-\theta)}{(1-\theta)} + \frac{\theta}{(1-\theta)} = \frac{n}{\theta^2} \frac{(1-\theta)+\theta}{(1-\theta)}$$

$$I_n(\theta) = \frac{n}{\theta^2(1-\theta)}$$

Therefore the Cramer-Rao lower bound is

$$var(T) \geq \frac{[\frac{d}{d\theta}\frac{1-\theta}{\theta}]^2}{I_n(\theta)} = \frac{[\frac{-1}{\theta} - \frac{1-\theta}{\theta^2}]^2}{I_n(\theta)} = \frac{[\frac{1}{\theta^2} + 2\frac{1}{\theta}\frac{1-\theta}{\theta^2} + \frac{(1-\theta)^2}{\theta^4}]}{I_n(\theta)} = \frac{\frac{\cancel{1}}{\cancel{\theta^2}}[1 + 2\frac{1-\theta}{\theta} + \frac{(1-\theta)^2}{\theta^2}]}{\frac{n}{\cancel{\theta^2}(1-\theta)}}$$

$$= \frac{\left(1 + \frac{1-\theta}{\theta}\right)^2}{\frac{n}{1-\theta}} = \frac{\left(\frac{1}{\theta}\right)^2}{\frac{n}{1-\theta}} = \frac{\frac{1}{\theta^2}}{\frac{n}{1-\theta}} \frac{\theta^2(1-\theta)}{\theta^2(1-\theta)} = \frac{1-\theta}{n\theta^2}$$

$$var(T) \geq \frac{1-\theta}{n\theta^2}$$

(d) Is the maximum likelihood estimator of $\tau(\theta)$ a uniformly minimum variance unbiased estimator? Justify your answer.
First check if T is unbiased

$$E(T) = E[\frac{1}{n}\sum_{i=1}^n X_i] = \frac{1}{n}\sum_{i=1}^n E[X_i] = \frac{1}{n}\sum_{i=1}^n [\frac{1-\theta}{\theta}] = \frac{1-\theta}{\theta}$$

Therefore our MLE of $\tau(\theta)$ is unbiased. In part a we found that

$$\frac{d}{d\theta} log[f(X_1, ..., X_n; \theta)] = \frac{n}{\theta} - \frac{\sum_{i=1}^n X_i}{1-\theta} = \frac{-n}{1-\theta}[-\frac{1-\theta}{\theta} + \frac{1}{n}\sum_{i=1}^n X_i] = a(\theta)[T - \tau(\theta)]$$

Therefore, T is and UMVUE that obtains the Cramer-Rao Lower Bound.

Problem 2

Let $X_1, ..., X_n$ be a random sample from a distribution with probability mass function

$$f(a; \mu) = \frac{1}{\sqrt{18\pi}} exp\Big\{ - \frac{(a-\mu)^2}{18} \Big\},$$

where $-\infty < \mu < \infty$.

(a) Find the Cramr-Rao lower bound for variances of unbiased estimators of μ.
First prove this is exponential family

$$\frac{1}{\sqrt{18\pi}} exp\Big\{ -\frac{(a-\mu)^2}{18} \Big\} = \frac{1}{\sqrt{18\pi}} exp\Big\{ -\frac{(a^2 - 2a\mu + \mu^2}{18} \Big\} = \frac{1}{\sqrt{18\pi}} e^{-a^2} e^{-\mu^2} e^{2a\mu} = h(a)c(\mu)e^{t(a)q(\mu)}$$

Therefore it is exponential Check if $q'(\mu)$ is continuous for $-\infty < \mu < \infty$

$$\frac{d}{d\mu} q(\mu) = \frac{d}{d\mu}\mu = 1$$

which is ok. Thus R1-R4 hold.
Therefore,

$$I_n(\mu) = nI(\mu) = -nE\Big[\frac{d^2}{d\mu^2} log[f(X_1; \mu)]\Big] = -nE\Big[\frac{d^2}{d\mu^2} log[\frac{1}{\sqrt{18\pi}} e^{-\frac{1}{18}a^2} e^{-\frac{1}{18}\mu^2} e^{\frac{2}{18}X_i\mu}]\Big] = -nE\Big[\frac{d^2}{d\theta^2}(log[\frac{1}{\sqrt{18\pi}}] - a^2$$

$$= -nE\Big[\frac{d}{d\theta}(-\frac{2}{18}\mu + \frac{2}{18})\Big] = -nE\Big[-\frac{2}{18}\Big] = \frac{n}{9}$$

Therefore by theorem 1

$$var(T) \geq \frac{[\frac{d}{d\mu}\tau(\mu)]^2}{I_n(\theta)} = \frac{[\frac{d}{d\mu}\mu)]^2}{\frac{n}{9}} = \frac{9}{n}$$

(b) Is the maximum likelihood estimator of μ a uniformly minimum variance unbiased estimator? Justify your answer. Find the MLE of μ

$$\frac{d}{d\mu}log[f(X_1,...,X_n;\mu)] = \frac{d}{d\mu}log[\prod_{i=1}^{n}\frac{1}{\sqrt{18\pi}}exp\Big\{-\frac{(X_i-\mu)^2}{18}\Big\}] = \frac{d}{d\mu}\sum_{i=1}^{n}log(\frac{1}{\sqrt{18\pi}})-\sum_{i=1}^{n}\frac{(X_i-\mu)^2}{18}$$

$$= 0 - \sum_{i=1}^{n}\frac{-2(X_i-\mu)}{18} = \sum_{i=1}^{n}\frac{X_i}{9} - \sum_{i=1}^{n}\frac{\mu}{9} = \sum_{i=1}^{n}\frac{X_i}{9} - \frac{n\mu}{9}$$

Set equal to 0

$$\sum_{i=1}^{n}\frac{X_i}{9} - \frac{n\hat{\mu}}{9} = 0 \Rightarrow \sum_{i=1}^{n}\frac{X_i}{9} = \frac{n\hat{\mu}}{9} \Rightarrow \sum_{i=1}^{n}X_i = n\hat{\mu} \Rightarrow \hat{\mu} = \frac{1}{n}\sum_{i=1}^{n}X_i$$

Check the second derivative

$$\frac{d}{d\mu}\sum_{i=1}^{n}\frac{X_i}{9} - \frac{n\mu}{9} = -\frac{n}{9} < 0$$

Since this is strictly less than zero $\hat{\mu}$ must be a maximum
Professor Lee asked us to check if the MLE is unbiased so

$$E\Big[\frac{1}{n}\sum_{i=1}^{n}X_i\Big] = \frac{1}{n}\sum_{i=1}^{n}E[X_i] = \frac{n}{n}E[X_i] = \mu$$

Let's try to rearrange the score function in the form $a(\mu)[\frac{1}{n}\sum_{i=1}^{n}X_i - \mu]$

$$\frac{d}{d\mu}log[f(X_1,...,X_n;\mu)] = \sum_{i=1}^{n}\frac{X_i}{9} - \frac{n\mu}{9} = \frac{n}{9}[\frac{1}{n}\sum_{i=1}^{n}X_i - \mu] = a(\mu)[\frac{1}{n}\sum_{i=1}^{n}X_i - \mu]$$

Therefore, $\frac{1}{n}\sum_{i=1}^{n}X_i$ is a UMVUE of μ which obtains the Cramer-Rao Lower Bound.

(c) Is the MLE of the 95^{th} percentile of f a UMVUE?

$$\int_{-\infty}^{\tau(\mu)} f(a;\mu)da = 0.95$$

Instead of solving this let's use the fact that $f(a;\mu) \sim N(\mu, 9)$

$$\tau(\mu) = \mu + 4.9344$$

By prop 1

$$T = \tau(\hat{\mu}) = \frac{1}{n}\sum_{i=1}^{n}X_i + 4.9344$$

Check if T is unbiased

$$E(T) = E[\frac{1}{n}\sum_{i=1}^{n}X_i + 4.9344] = \frac{1}{n}\sum_{i=1}^{n}E[X_i] + 4.9344 = \mu + 4.9344 = \tau$$

Let's try to put the score function in the form $a(\mu)[(\frac{1}{n}\sum_{i=1}^{n}X_i + 4.9344) - (\mu + 4.9344)$
Recall

$$\frac{d}{d\mu}log[f(X_1,...,X_n;\mu)] = \sum_{i=1}^{n}\frac{X_i}{9} - \frac{n\mu}{9} = \frac{n}{9}[\frac{1}{n}\sum_{i=1}^{n}X_i - \mu] = \frac{n}{9}[\frac{1}{n}\sum_{i=1}^{n}X_i + 4.9344 - \mu - 4.9344]$$

$$= a(\mu)[(\frac{1}{n}\sum_{i=1}^{n}X_i + 4.9344) - (\mu + 4.9344)]$$

Therefore T is a UMVUE.

Problem 3

Let $X_1, ..., X_n$ be a random sample from a distribution with probability density function $f(a;\theta) = \theta a^{\theta-1}\mathbb{1}_{(0,1)}(a)$, where $\theta > 0$. Find a function of θ, denoted $\tau(\theta)$, for which there exists an unbiased estimator whose variance attains the Cramer-Rao lower bound, and determine the uniformly minimum variance unbiased estimator of $\tau(\theta)$. First let's show that $f(a;\theta)$ belongs to the exponential family.

$$\theta a^{\theta-1}\mathbb{1}_{(0,1)}(a) = \mathbb{1}_{(0,1)}(a)\theta \exp\{(\theta-1)log(a)\} = h(a)c(\theta)exp\{t(a)q(\theta)\}$$

Therefore $f(a;\theta)$ belongs to the exponential family and R1-R3 hold.
Next find the score function.
For $X_1, ..., X_n$ in (0,1)

$$\frac{d}{d\theta}log(f(X_1, ..., X_n;\theta) = \frac{d}{d\theta}log(\prod_{i=1}^{n}\theta X_i^{\theta 1}) = \frac{d}{d\theta}\sum_{i=1}^{n}log(\theta) + (\theta-1)\sum_{i=1}^{n}log(X_i)$$

$$= \frac{d}{d\theta}nlog(\theta) + (\theta\sum_{i=1}^{n}log(X_i) - \sum_{i=1}^{n}log(X_i)) = \frac{n}{\theta} + \sum_{i=1}^{n}log(X_i)$$

$$= -n[-\frac{1}{n}\sum_{i=1}^{n}log(X_i) - \frac{1}{\theta}] = a(\theta)[T - \tau(\theta)]$$

Where $T = -\frac{1}{n}\sum_{i=1}^{n}log(X_i)$ and $\tau(\theta) = \frac{1}{\theta}$
By theorem 2 T is an UMVUE of $\tau(\theta)$ if T is unbiased
Let's check if T is unbiased

$$E(T) = E[-\frac{1}{n}\sum_{i=1}^{n}log(X_i)] = -\frac{1}{n}\sum_{i=1}^{n}E[log(X_i)]$$

Find $E[log(X_i)]$

$$E[log(X_i)] = \int_0^1 log(a)\theta a^{\theta-1}da = \theta\int_0^1 log(a)a^{\theta-1}da$$

Let $u = log(a)du = a^{-1}dadv = a^{\theta-1}v = \theta^{-1}a^{\theta}$ Then

$$\theta\int_0^1 log(a)a^{\theta-1}da = \theta[log(a)\theta^{-1}a^{\theta}\Big|_0^1 - \int_0^1\theta^{-1}a^{\theta}a^{-1}da = 0 - \theta\theta^{-1}\int_0^1 a^{\theta-1}da = -\frac{1}{\theta}a^{\theta}\Big|_0^1 = \frac{-1}{\theta}$$

Thus

$$E(T) = -\frac{1}{n}\sum_{i=1}^{n}E[log(X_i)] = -\frac{1}{n}\sum_{i=1}^{n}(-\frac{1}{\theta}) = \frac{1}{\theta}$$

Therefore, T is an unbiased UMVUE for $\frac{1}{\theta}$