Kris Harper
STAT 24400
November 18, 2010

Homework 7

Problem 1. *Suppose X follows a geometric distribution,*

$$P(X = k) = p(1-p)^{k-1}$$

and assume an i.i.d sample of size n.
(b) Find the mle of p.
(c) Find the asymptotic variance of the mle.

(b) The log likelihood is given by

$$l(p) = \sum_{i=1}^{n} \log\left(p(1-p)^{k_i-1}\right) = \sum_{i=1}^{n} \left(\log(p) + (k_i - 1)\log(1-p)\right) = n\log(p) + \log(1-p)\sum_{i=1}^{n}(k_i - 1).$$

Then

$$l'(p) = \frac{n}{p} - \frac{1}{1-p}\sum_{i=1}^{n}(k_i - 1).$$

Setting this equal to 0 gives

$$\hat{p} = n\left(\sum_{i=1}^{n} k_i\right)^{-1}.$$

This is indeed a maximum because

$$l''(p) = -\frac{n}{p^2} - \frac{1}{(1-p)^2}\sum_{i=1}^{n}(k_i - 1)$$

and

$$l''(\hat{p}) = -\left(\sum_{i=1}^{n} k_i\right)^2 \left(\frac{1}{n} + \frac{1}{\sum_{i=1}^{n}(k_i - 1)}\right) < 0.$$

(c) Since $\hat{p}$ is an average, the central limit theorem will tell us that $\mathrm{Var}(\hat{p}) \approx \mathrm{Var}(X)/n = p^2(1-p)/n$.

Problem 2. *Consider an i.i.d. sample of random variables with density function*

$$f(x \mid \sigma) = \frac{1}{2\sigma}\exp\left(-\frac{|x|}{\sigma}\right).$$

(b) Find the maximum likelihood estimate of σ.
(c) Find the asymptotic variance of the mle.

(b) The log likelihood is given by

$$l(\sigma) = \sum_{i=1}^{n} \log\left(\frac{1}{2\sigma}\exp\left(-\frac{|x_i|}{\sigma}\right)\right) = -\frac{1}{\sigma}\sum_{i=1}^{n}|x_i| - n\log(2\sigma).$$

Then

$$l'(\sigma) = \frac{1}{\sigma^2}\sum_{i=1}^{n}|x_i| - \frac{n}{\sigma}.$$

Setting this equal to 0 gives

$$\hat{\sigma} = \frac{1}{n}\sum_{i=1}^{n} |x_i|.$$

This is indeed a maximum because

$$l''(\sigma) = \frac{n}{\sigma^2} - \frac{2}{\sigma^3}\sum_{i=1}^{n} |x_i|$$

and

$$l''(\hat{\sigma}) = \left(\sum_{i=1}^{n} |x_i|\right)^{-2}\left(\frac{1}{n} - 2\right) < 0.$$

(c) We have

$$\begin{aligned}
I(\sigma) &= -E\left(\frac{\partial^2}{\partial\sigma^2}\log\left(\frac{1}{2\sigma}\exp\left(-\frac{|x|}{\sigma}\right)\right)\right) \\
&= -E\left(\frac{\partial^2}{\partial\sigma^2}\left(-\frac{|x|}{\sigma} - \log(2\sigma)\right)\right) \\
&= -E\left(\frac{\partial}{\partial\sigma}\left(\frac{|x|}{\sigma^2} - \frac{1}{\sigma}\right)\right) \\
&= -E\left(\frac{1}{\sigma^2} - \frac{2|x|}{\sigma^3}\right) \\
&= -\left(\frac{1}{\sigma^2} - \frac{2E(|x|)}{\sigma^3}\right).
\end{aligned}$$

Note that for this distribution, $|X|$ follows an exponential distribution with parameter $1/\sigma$. Thus we have

$$E(|X|) = \int_{-\infty}^{\infty} \frac{x}{2\sigma} e^{-|x|/\sigma} dx = \sigma.$$

Plugging this in gives $I(\sigma) = -(1/\sigma^2 - 2\sigma/\sigma^3) = 1/\sigma^2$. Thus $\mathrm{Var}(\hat{\sigma}) = \sigma^2/n$

Problem 3. *Suppose that $X_1, X_2, \ldots, X_n$ are i.i.d. random variables on the interval $[0, 1]$ with the density function*

$$f(x \mid \alpha) = \frac{\Gamma(2\alpha)}{\Gamma(\alpha)^2}[x(1-x)]^{\alpha-1}$$

where $\alpha > 0$ is a parameter to be estimated from the sample. It can be shown that

$$\begin{aligned}
E(X) &= \frac{1}{2} \\
\mathrm{Var}(X) &= \frac{1}{4(2\alpha+1)}.
\end{aligned}$$

(c) What equation does the mle of α satisfy?
(d) What is the asymptotic variance of the mle?

(c) The log likelihood is given by

$$l(\alpha) = \sum_{i=1}^{n}\log\left(\frac{\Gamma(2\alpha)}{\Gamma(\alpha)^2}[x_i(1-x_i)]^{\alpha-1}\right) = n\log(\Gamma(2\alpha)) - 2n\log(\Gamma(\alpha)) + (\alpha-1)\sum_{i=1}^{n}\log(x_i(1-x_i))$$

Differentiating we have

$$l'(\alpha) = \frac{2n\Gamma'(2\alpha)}{\Gamma(2\alpha)} - \frac{2n\Gamma'(\alpha)}{\Gamma(\alpha)} + \sum_{i=1}^{n}\log(x_i(1-x_i)).$$

Then $\hat{\alpha}$ must satisfy $l'(\alpha) = 0$ where $l'(\alpha)$ is as above.

(d) We have

$$\begin{aligned}
I(\alpha) &= -E\left(\frac{\partial^2}{\partial\alpha^2}\log\left(\frac{\Gamma(2\alpha)}{\Gamma(\alpha)^2}[x(1-x)]^{\alpha-1}\right)\right) \\
&= -E\left(\frac{\partial^2}{\partial\alpha^2}\left(\log(\Gamma(2\alpha)) - 2\log(\Gamma(\alpha)) + (\alpha-1)\log(x(1-x))\right)\right) \\
&= -E\left(\frac{\partial}{\partial\alpha}\left(\frac{2\Gamma'(2\alpha)}{\Gamma(2\alpha)} - \frac{2\Gamma'(\alpha)}{\Gamma(\alpha)} + \log(x(1-x))\right)\right) \\
&= -E\left(2\frac{\Gamma(2\alpha)\Gamma''(2\alpha) - \Gamma'(2\alpha)^2}{\Gamma(2\alpha)^2} - 2\frac{\Gamma(\alpha)\Gamma''(\alpha) - \Gamma'(\alpha)^2}{\Gamma(\alpha)^2}\right) \\
&= 2\left(\frac{\Gamma(\alpha)\Gamma''(\alpha) - \Gamma'(\alpha)^2}{\Gamma(\alpha)^2} - \frac{\Gamma(2\alpha)\Gamma''(2\alpha) - \Gamma'(2\alpha)^2}{\Gamma(2\alpha)^2}\right)
\end{aligned}$$

and so

$$\mathrm{Var}(\hat{\alpha}) = \frac{1}{2n}\left(\frac{\Gamma(\alpha)\Gamma''(\alpha) - 2\Gamma'(\alpha)}{\Gamma(\alpha)^2} - \frac{\Gamma(2\alpha)\Gamma''(2\alpha) - 2\Gamma'(2\alpha)}{\Gamma(2\alpha)^2}\right)^{-1}.$$

Problem 4. *Suppose that $X_1, X_2, \ldots, X_n$ are i.i.d. with density function*

$$f(x \mid \theta) = e^{-(x-\theta)}, \quad x \geq \theta$$

and $f(x \mid \theta) = 0$ otherwise.
(b) Find the mle of θ.

(b) The log likelihood is given by

$$l(\theta) = \sum_{i=1}^{n} \log\left(e^{-(x_i-\theta)}\right) = \sum_{i=1}^{n}(\theta - x_i).$$

Since we're subtracting x_i from θ in each term, we want θ to be as close to the x_i terms as possible. On the other hand, we want $\theta \leq x_i$ for each term, otherwise $f(x_i \mid \theta) = 0$ for some i and so the joint distribution function is 0. Therefore, we pick $\hat{\theta}$ to be the value of x_i which is less than or equal to all the other values. That is,

$$\hat{\theta} = \min\{x_i \mid 1 \leq i \leq n\}.$$

Problem 5. *The Pareto distribution has been used in economics as a model for a density function with a slowly decaying tail:*

$$f(x \mid x_0, \theta) = \theta x_0^\theta x^{-\theta-1}, \quad x \geq x_0, \quad \theta > 1.$$

Assume that $x_0 > 0$ is given and that $X_1, X_2, \ldots X_n$ is an i.i.d. sample.
(b) Find the mle of θ.
(c) Fin the asymptotic variance of the mle.

(b) The log likelihood is given by

$$l(\theta) = \sum_{i=1}^{n} \log\left(\theta x_0^\theta x_i^{-\theta-1}\right) = n\log(\theta) + n\theta\log(x_0) - (\theta+1)\sum_{i=1}^{n}\log(x_i).$$

Then

$$l'(\theta) = \frac{n}{\theta} + n\log(x_0) - \sum_{i=1}^{n}\log(x_i).$$

Setting this equal to 0 we have

$$\hat{\theta} = n\left(\sum_{i=1}^{n}\log(x_i) - n\log(x_0)\right)^{-1} = n\left(\sum_{i=1}^{n}\log\left(\frac{x_i}{x_0}\right)\right)^{-1}.$$

This is indeed a maximum because

$$l''(\theta) = -\frac{n}{\theta^2} < 0.$$

(c) We have

$$\begin{aligned}
I(\theta) &= -E\left(\frac{\partial^2}{\partial\theta^2}\log\left(\theta x_0^{\theta}x^{-\theta-1}\right)\right) \\
&= -E\left(\frac{\partial^2}{\partial\theta^2}(\log(\theta) + \theta\log(x_0) - (\theta+1)\log(x))\right) \\
&= -E\left(\frac{\partial}{\partial\theta}\left(\frac{1}{\theta} + \log(x_0) - \log(x)\right)\right) \\
&= -E\left(-\frac{1}{\theta^2}\right) \\
&= \frac{1}{\theta^2}
\end{aligned}$$

and so $\mathrm{Var}(\hat{\theta}) = \theta^2/n$.

Problem 6. *Let $X_1, \ldots, X_n$ be an i.i.d. sample from a Rayleigh distribution with parameter $\theta > 0$;*

$$f(x \mid \theta) = \frac{x}{\theta^2}e^{-x^2/(2\theta^2)}, \quad x \geq 0.$$

(b) Find the mle of θ.
(c) Find the asymptotic variance of the mle.

(b) The log likelihood is given by

$$l(\theta) = \sum_{i=1}^{n}\log\left(\frac{x_i}{\theta^2}e^{-x_i^2/(2\theta^2)}\right) = \sum_{i=1}^{n}\log(x_i) - \frac{1}{2\theta^2}\sum_{i=1}^{n}x_i^2 - 2n\log(\theta).$$

Then

$$l'(\theta) = \frac{1}{\theta^3}\sum_{i=1}^{n}x_i^2 - \frac{2n}{\theta}.$$

Setting this equal to 0 we have

$$\hat{\theta} = \pm\frac{1}{\sqrt{2n}}\sqrt{\sum_{i=1}^{n}x_i^2}.$$

Then we have

$$l''(\theta) = \frac{2n}{\theta^2} - \frac{3}{\theta^4}\sum_{i=1}^{n}x_i^2.$$

Putting in either the positive or negative square root gives

$$l''(\hat{\theta}) = -8n^2\left(\sum_{i=1}^{n}x_i^2\right)^{-1} < 0.$$

Thus both values of $\hat{\theta}$ are maximums.

(c) We have

$$
\begin{aligned}
I(\theta) &= -E\left(\frac{\partial^2}{\partial\theta^2}\log\left(\frac{x}{\theta^2}e^{-x^2/(2\theta^2)}\right)\right) \\
&= -E\left(\frac{\partial^2}{\partial\theta^2}(\log(x) - \frac{x^2}{2\theta^2} - 2\log(\theta))\right) \\
&= -E\left(\frac{\partial}{\partial\theta}\left(\frac{x^2}{\theta^3} - \frac{2}{\theta}\right)\right) \\
&= -E\left(\frac{2}{\theta^2} - \frac{3x^2}{\theta^4}\right) \\
&= \frac{3E(x^2)}{\theta^4} - \frac{2}{\theta^2}.
\end{aligned}
$$

We now have

$$E(X^2) = \int_0^\infty \frac{x^3}{\theta^2}e^{-x^2/(2\theta^2)}dx = e^{-x^2/(2\theta^2)}(2\theta^2 + x^2)\Big|_0^\infty = 2\theta^2.$$

Plugging this in gives us

$$I(\theta) = \frac{6\theta^2}{\theta^4} - \frac{2}{\theta^2} = \frac{4}{\theta^2}.$$

Thus $\mathrm{Var}(\hat{\theta}) = \theta^2/4n$.

Problem 7. *Let $X_1, \ldots, X_n$ be i.i.d. random variables with the density function*

$$f(x \mid \theta) = (\theta+1)x^\theta, \quad 0 \leq x \leq 1.$$

(b) Find the mle of θ.
(c) Find the asymptotic variance of the mle.

(b) The log likelihood is given by

$$l(\theta) = \sum_{i=1}^n \log\left((\theta+1)x_i^\theta\right) = n\log(\theta+1) + \theta\sum_{i=1}^n \log(x_i).$$

Then

$$l'(\theta) = \frac{n}{\theta+1} + \sum_{i=1}^n \log(x_i).$$

Setting this equal to 0 we have

$$\hat{\theta} = -1 - n\left(\sum_{i=1}^n \log(x_i)\right)^{-1}.$$

This is indeed a maximum because

$$l''(\theta) = -\frac{n}{(x+1)^2} < 0.$$

(c) We have

$$
\begin{aligned}
I(\sigma) &= -E\left(\frac{\partial^2}{\partial\theta^2}\log\left((\theta+1)x^\theta\right)\right) \\
&= -E\left(\frac{\partial^2}{\partial\theta^2}(\log(\theta+1)+\theta\log(x))\right) \\
&= -E\left(\frac{\partial}{\partial\theta}\left(\frac{1}{\theta+1}+\log(x)\right)\right) \\
&= -E\left(-\frac{1}{(\theta+1)^2}\right) \\
&= \frac{1}{(\theta+1)^2}
\end{aligned}
$$

so $\mathrm{Var}(\hat{\theta}) = (\theta+1)^2/n$.