**Stat 135, Fall 2006 A. Adhikari**

# HOMEWORK 4 SOLUTIONS

**1.** The log of the normal density is

$$\log f(x|\mu,\sigma) \quad = \quad \log \frac{1}{\sqrt{2\pi}} \; - \; \log\sigma \; - \; \frac{(x-\mu)^2}{2\sigma^2}$$

So the log likelihood function is

$$l(\mu,\sigma) \quad = \quad n\log \frac{1}{\sqrt{2\pi}} \; - \; n\log\sigma \; - \; \frac{\sum_{i=1}^{n}(X_i-\mu)^2}{2\sigma^2}$$

**a.** If $\mu$ is a known constant then all you have to do is differentiate the log likelihood with respect to $\sigma$, set equal to 0, and solve:

$$-\frac{n}{\hat{\sigma}} \; + \; \frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\hat{\sigma}^3} \quad = \quad 0$$

so $\hat{\sigma}^2 \; = \; \frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2$, no big surprise. Take the square root to get the MLE of $\sigma$.

**b.** This time treat $\sigma$ as the constant and differentiate the log likelihood with respect to $\mu$:

$$\frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i-\hat{\mu}) \quad = \quad 0$$

so $\hat{\mu} = \bar{X}$, again no big surprise.

**c.** We know that $\hat{\mu}$ is unbiased and has variance $\sigma^2/n$. Now

$$\frac{d}{d\mu}[\log \frac{1}{\sqrt{2\pi}} \; - \; \log\sigma \; - \; \frac{(x-\mu)^2}{2\sigma^2}] \quad = \quad \frac{x-\mu}{\sigma^2}$$

So the Fisher information is

$$I(\mu) \quad = \quad E[(\frac{X-\mu}{\sigma^2})^2] \quad = \quad \frac{\sigma^2}{\sigma^4} \quad = \quad \frac{1}{\sigma^2}$$

The Cramer-Rao bound says that no unbiased estimate has variance less than $1/nI(\mu) = \sigma^2/n = Var(\hat{\mu})$. So $\hat{\mu}$ has the smallest variance among all unbiased estimates.

**2a.** Notice that the density is that of $T+\theta$ where $T$ has the exponential density with parameter 1. Therefore the first moment of the density is $\mu_1 = E(T)+\theta = 1+\theta$, and therefore $\theta = \mu_1 - 1$. Therefore the MOM estimate is $\hat{\theta}_{MOM} = \bar{X} - 1$.

**b.** It is important to notice that with probability 1, $\theta \leq \min(X_1, X_2, \ldots, X_n)$. So the likelihood function is

$$e^{-\sum_{i=1}^{n}(X_i-\theta)} \quad = \quad e^{-n\bar{X}} \cdot e^{n\theta}$$

for $\theta \leq \min(X_1, X_2, \ldots, X_n)$. This is an increasing function of $\theta$ so there is no need to differentiate it to find its maximum. The function is maximized by the maximum possible value of $\theta$, which is $\min(X_1, X_2, \ldots, X_n)$ by our earlier observation. So $\hat{\theta}_{MLE} = \min(X_1, X_2, \ldots, X_n)$.

**3.** We have just one observation $X$ from the uniform distribution on $\{1, 2, \ldots, N\}$.

The first moment of the distribution is $\mu_1 = (N+1)/2$, so $N = 2\mu_1 - 1$, so the MOM estimate is $\hat{N}_{MOM} = 2X - 1$. The observed value of $X$ is 888, so the observed value of $\hat{N}_{MOM}$ is 1775.

The likelihood function is $1/N$ for $N \geq X$. This is a decreasing function of $N$ and is maximized when $N$ is at its minimum possible value. That's $X$. So $\hat{N}_{MLE} = X$ and its observed value is 888.

**4. a)** Let $T$ be the time until the first failure. Then $T$ is the minimum of five i.i.d. exponential variables. What we have is one observation of $T$. Hence, the likelihood function is just its density:

$$lik(\tau) = f_T(t) = \frac{5}{\tau} e^{-\frac{5}{\tau} t}$$

If you don't know where that came from, look at Example A of Section 3.7, page 104. The minimum of independent exponentials is itself an exponential.

**b)**

$$l(\tau) = \log 5 - \log \tau - \frac{5t}{\tau}$$

$$l'(\tau) = -\frac{1}{\tau} + \frac{5t}{\tau^2} = 0 \quad \Rightarrow \quad \hat{\tau} = 5T$$

The observed value of $T$ is given as 100 days so the observed value of $\hat{\tau}$ is 500 days.

**c)** Since $T$ is exponential, the density of $5T$ is easy by the change of variable formula. But here's a calculation from first principles: Look at the cdf.

$$F_{\hat{\tau}}(t) = P[\hat{\tau} \leq t] = P[5T \leq t] = P[T \leq \frac{t}{5}] = F_T(\frac{t}{5}) = 1 - e^{-\frac{5}{\tau} \cdot \frac{t}{5}} = 1 - e^{-\frac{1}{\tau} t}$$

This is the cdf of the exponential distribution with parameter $1/\tau$, which is the same exponential density as that of the individual lifetimes.

**d)** The variance of the exponential with parameter $\frac{1}{\tau}$ is $\tau^2$. Hence the standard error of the estimate is simply $\tau$.

**5.** I'm just going to assume that you read it.

In what follows I've written the code so that it's more or less self-explanatory. I'm sure some of you will have done things more efficiently.

**6.** Here is the basic code. You can make the histogram have more bars, add labels to the axes, etc.

```
> alpha ← runif(1, min=2, max=4)
> lambda ← runif(1, min=1, max=2)
> firstsamp ← rgamma(200, shape = alpha, rate = lambda)
> hist(firstsamp, prob=TRUE)
> x ← seq(min(firstsamp), max(firstsamp), by=0.01)
> lines(x, dgamma(x, shape=alpha, rate=lambda))
```

**7.** I will use *firstsamp* from the previous problem. You really should be multiplying the variances by 199/200, but I'm prepared to ignore that.

```
> alpha.mm ← mean(firstsamp)ˆ2)/var(firstsamp)
> lambda.mm ← mean(firstsamp)/var(firstsamp)
> bigsample ← rgamma(200000, shape=alpha.mm, rate=lambda.mm)
> samples ← matrix(bigsample, nrow=200, ncol=1000)
> sampmeans ← colMeans(samples)
> sampvars ← apply(samples, 2, var)
> alphas ← sampmeansˆ2/sampvars
> lambdas ← sampmeans/sampvars
```

Now you've got your estimates; I'm sure you know how to draw the histograms. Both should be roughly normal, like Fig 8.4. You can use *points* or your pen to mark the true values (*alpha* and *lambda* from Problem 6) on the horizontal axes; you expect them to be somewhere around the middle (the estimates are biased, so the true values will be somewhat off center), but you'll have to see what happens in your particular replication of the experiment. The estimated standard error is simply the function *sd* applied separately to the two sets of estimates.

**8.** The only issue here is how to get the two values of $\hat{alpha}$. It goes without saying that we're using *alphas* from Problem 7.

```
> a.05 ← quantile(alphas, .05)
> a.95 ← quantile(alphas, .95)
```

Now it's just a question of drawing the graphs, which I'm sure you know how to do. Use the option *lty* with the command *lines* to get different line types, or use different colors and use a color printer. Just make sure that the three curves are distinguishable.

**9.** The only difference between this exercise and Problem 7 is in how you get the samples. After that the code is all the same. So:

```
> newbigsample ← sample(firstsamp, 200000, replace=TRUE)
```

and then you proceed as in Problem 7. The histograms should resemble those in Problem 7, perhaps a bit more skewed.

**10.** The worked example in *An Introduction to R* is for the logistic density. You just have to replace its log-likelihood function by that of the gamma, and remember to take the negative log-likelihood because you're using a minimizing command, not a maximizer. The vector of parameters is *p*; its first element is alpha and the second is lambda.

```
> fn ← function(p)
– 200*p[1]*log(p[2]) – (p[1]–1)*sum(log(firstsamp)) + p[2]*sum(firstsamp) + 200*log(gamma(p[1]))
> nlm(fn, p=c(alpha.mm, lambda.mm))
```

The estimates appear on the screen. Mine were near the center of the histograms in Problem 7, but this exercise does not give a sense of how the MLEs vary.