Homework 6 Solutions

Timothy Daley

1. Chapter 18, Exercise 1. Explain your logic in finding your answer.

Consider raffle ticket example in Section 18.1.1. Suppose 500 tickets are sold, and you have data on 8 of them. Continue to assume sampling with replacement. Consider the Maximum Likelihood and Methods of Moments estimators.

a. Find the probability that the MLE is exactly equal to the true value of C.

The MLE in this situation is $\max_{1 \leq i \leq n} x_i$. We want to calculate $\Pr(\max_{1 \leq i \leq n} x_i = C)$. By the complement rule this equals $1 - \Pr(\max_{1 \leq i \leq n} x_i \leq C - 1)$. If the maximum is at most $C - 1$, then all of the tickets must be at most $C - 1$. Therefore

$$\Pr(\max_{1 \leq i \leq n} x_i = C) = 1 - \Pr(\max_{1 \leq i \leq n} x_i \leq C - 1) = 1 - \cap_{1 leq i \leq n} \Pr(x_i \leq C - 1).$$

Sampling with replacement means independence of the samples, and therefore the intersection can be turned into a product

$$\Pr(\max_{1 \leq i \leq n} x_i = C) = 1 - \cap_{1 leq i \leq n} \Pr(x_i \leq C - 1) = 1 - \prod_{i=1}^{n} \Pr(x_i \leq C - 1) = 1 - \left(\frac{C-1}{C}\right)^n$$

Since $C = 500$ and $n = 8$, this equals

```
1 - (499/500)^8
```

```
## [1] 0.01588845
```

b. Find the exact probability that the MLE is within 50 of the true value.

Similar to above, the probability the MLE is within 50 is equal to

$$\Pr(\max_{1 \leq i \leq n} x_i \geq 450) = 1 - \Pr(\max_{1 \leq i \leq n} x_i \leq 449) = 1 - \prod_{i=1}^{n} \left(\frac{449}{500}\right)^8.$$

I think an argument can be made for the value 450 above rather than 449, so please accept both. This is equal to

```
1 - (449/500)^8
```

```
## [1] 0.5771263
```

c. Find the approximate probability that the Method of Moments estimator is within 50 of the true value.

The method of moments estimator is $2\bar{x} - 1$. To find the approximate probability we will apply the CLT, even though the sample size (8) is small. The expectation of x is $(500 + 1)/2 = 250.5$ and the variance is given by $((500 - 1 + 1)^2 - 1)/12 = 20833.25$. Therefore the probability $2\bar{x} - 1$ is within 50 of the true value 500 is approximately equal to

$$\Pr(450 \leq 2\bar{x} - 1 \leq 550) = \Pr(\frac{451}{2} \leq \bar{x} \leq \frac{551}{2}) = \Pr(\frac{225.5 - 250.5}{\sqrt{20833.25}/\sqrt{8}} \leq z \leq \frac{275.5 - 250.5}{\sqrt{20833.25}/\sqrt{8}})$$

1

```
pnorm((275.5 - 250.5)/(sqrt(20833.25)/sqrt(8))) - pnorm((225.5 - 250.5)/(sqrt(20833.25)/sqrt(8)))
```

[1] 0.3757946

2. Chapter 18, Exercise 2a.

Suppose $I = 1$ or 0, with probability p and $1 - p$, respectively. Given I, X has a Poisson distribution with mean λ_I. Suppose we have $X_1, \ldots, X_n$, a random sample of size n from the (unconditional) distribution of X. (We do not know the associated values of I, i.e. $I_1, \ldots, I_n$.) This kind of situation occurs in various applications. The key point is the effect of the unseen variable. In terms of estimation, note that there are three parameters to be estimated.

   a. Set up the likelihood function, which if maximized with respect to the three parameters would yield the MLEs for them.

We know this problem well as a mixture problem with the latent variable I indicating class membership. If I is known then $X|I$ is Poisson(λ_I), i.e. $\Pr(X = x|I) = \lambda_I^x e^{-\lambda}/x!$. Note that

$$\Pr(X = x) = \Pr(X = x \cap I = 0) + \Pr(X = x \cap I = 1) = \Pr(X = x|I = 0)\Pr(I = 0) + \Pr(X = x|I = 1)\Pr(I = 1) = (1-p)\lambda_0^x e^{-\lambda_0}$$

Therefore the likelihood equals

$$\prod_{i=1}^{n}\left((1 - p)\lambda_0^{x_i} e^{-\lambda_0}/x_i! + p\lambda_1^{x_i} e^{-\lambda_1}/x_i!\right)$$

3. Chapter 18, Exercise 3.

Find the Method of Moments and Maximum Likelihood estimators of the following parameters in famous distribution families:

   a. p in the binomial family (n known)

MLE: Suppose we have observations x_i for $i = 1, \ldots, m$. The log likelihood is $\sum_{i=1}^{m} \log \binom{n}{x_i} + x_i \log p + (n - x_i) \log(1 - p)$ and the derivative in p is

$$\frac{\partial}{\partial p} \log L(p|x_1, \ldots, x_n) = \sum_{i=1}^{m} \frac{x_i}{p} - \frac{n - x_i}{1 - p} = 0$$

$$0 = \sum_{i=1}^{m} x_i(1 - p) - (n - x_i)p = \sum_{i=1}^{m} x_i - np$$

therefore $\hat{p}_{MLE} = \frac{1}{nm} \sum_{i=1}^{m} x_i$.

MoM: $\mathrm{E}(X) = np$ and therefore the MoM equation is $\bar{x} = np$ and therefore $\hat{p}_{MoM} = \bar{x}/n = \frac{1}{nm} \sum_{i=1}^{m} x_i$.

   b. p in the geometric family.

MLE: If X is geometrically distributed then the pdf of X is $f(x) = (1 - p)^x p$. The log likelihood is $\log L(p|x_1, \ldots x_n) = \sum_{i=1}^{n} \log(p) + x_i \log(1 - p)$. The derivative in p is

$$\frac{\partial}{\partial p} \log L(p|x_1, \ldots, x_n) = \sum_{i=1}^{n} \frac{1}{p} - \frac{x_i}{1 - p} = 0$$

and this implies that $np - n = -p \sum_{i=1}^{n} x_i$ and therefore $\hat{p}_{MLE} = n/\left(n + \sum_{i=1}^{n} x_i\right)$.

MoM: The expectation of X is equal to $\frac{1-p}{p}$, therefore the method of moments equation is $\bar{x} = \frac{1-p}{p}$, or $\hat{p} = 1/(1 + \bar{x})$

Note that the book has the opposite pdf, $f(x) = (1-p)p^x$, so the estimators will be slightly different.

    c. μ in the normal family (σ known)

MLE: If X is normally distributed then $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$. For observations $x_1, \ldots, x_n$ the log likelihood is $\sum_{i=1}^{n} -\frac{1}{2}(\log 2 + \log \pi + \log \sigma^2) - \frac{(x-\mu)^2}{2\sigma^2}$. The derivative in μ is equal to

$$\frac{\partial}{\partial \mu} \log L = \sum_{i=1}^{n} 2(x_i - \mu)/2\sigma^2 = 0$$

and therefore $\hat{\mu}_{MLE} = \bar{x}$.

MoM: The expectation is μ and therefore $\hat{\mu}_{MoM} = \bar{x}$.

    d. λ in the Poisson family.

MLE: If X is Poisson distributed then $f(x) = \lambda^x e^{-\lambda}/x!$. For observations $x_1, \ldots, x_n$, the log likelihood is $\sum_{i=1}^{n} x_i \log \lambda - \lambda - \log x!$. The derivative in λ is equal to $\frac{\partial}{\partial \lambda} \log L = \sum_{i=1}^{n} x_i/\lambda - 1 = 0$. Therefore $\hat{\lambda}_{MLE} = \bar{x}$.

MoM: The expecation is λ and therefore $\hat{\lambda}_{MoM} = \bar{x}$

    4. Chapter 18, Exercise 4 a, b, and c. By this I mean the first 3 bullet points.

In part a and b, assume that $x_1, x_2, \ldots, x_n$ are iid observations from a Bernoulli(p).

For each of the following quantities, state whether the given estimator is unbiased in the given context:

    a. $\hat{p}$ as an estimator of p.

$x_1, \ldots, x_n$ are iid samples from a Bernoulli(p) distribution. $\hat{p} = \sum_{i=1}^{n} x_i/n$. The expectation of $\hat{p}$ equals

$$E\left(\sum_{i=1}^{n} x_i/n\right) = \frac{1}{n} \sum_{i=1}^{n} E x_i = \frac{1}{n} \sum_{i=1}^{n} p = p.$$

Therefore $\hat{p}$ is unbiased.

    b. $\hat{p}(1 - \hat{p})$ as an estimator for $p(1 - p)$.

$x_1, \ldots, x_n$ are iid samples from a Bernoulli(p) distribution. The expectation of $\hat{p}(1 - \hat{p})$ equals

$$E\left(\frac{1}{n} \sum_{i=1}^{n} x_i\left(1 - \frac{1}{n} \sum_{i=1}^{n} x_i\right)\right) = \frac{1}{n} \sum_{i=1}^{n} E(x_i) - \frac{1}{n^2} E\left(\left(\sum_{i=1}^{n} x_i\right)^2\right).$$

Note that $\left(\sum_{i=1}^{n} x_i\right)^2 = \sum_{i,j=1}^{n} x_i x_j = \sum_{i=1}^{n} x_i^2 + \sum_{i,j=1, i \neq j}^{n} x_i x_j$. Note that $E(x_i^2) = p$ for all $i = 1, \ldots, n$ since x_i is an indicator variable, and $E(x_i x_j) = p^2$ since x_i and x_j are independent. Therefore

$$E\left(\frac{1}{n} \sum_{i=1}^{n} x_i\left(1 - \frac{1}{n} \sum_{i=1}^{n} x_i\right)\right) = p - \frac{1}{n^2}(np + n(n-1)p^2) = p - p/n - p^2 + p^2/n = p(1-p) - p(1-p)/n$$

and we see that the estimate is biased downward by $p(1-p)/n$.

c. $\bar{X} - \bar{Y}$ as an estimate of $\mu_1 - \mu_2$.

Note that $\bar{X}$ is unbiased for μ_1 and $\bar{Y}$ is unbiased for μ_2. By the linearity of expectations, $\bar{X} - \bar{Y}$ is unbiased for $\mu_1 - \mu_2$.

5. Chapter 18, Exercise 6.

Suppose W has a uniform distribution on $(-c, c)$, and we draw a random sample of size n, $W_1, \ldots, W_n$. Find the Method of Moments and Maximum Likelihood estimators. (Note that in the Method of Moments case, the first moment won't work.)

MLE: The pdf of W is $f(w) = 1/2c$. Therefore the log likelihood of n independent observations is $\log L = \sum_{i=1}^{n} -\log 2 - \log c$. The derivative in terms of c is $\frac{\partial}{\partial c} \log L = \sum_{i=1}^{n} -1/c = 0$ This implies that $c = \infty$. Let's look at the second derivative $\frac{\partial^2}{\partial c^2} \log L = \sum_{i=1}^{n} 1/c^2 \geq 0$. Since the function has positive second derivative, $c = \infty$ is not a maximum, but is a minimum. Since the second derivative is positive at all points, the likelihood is decreasing at all possible values of c and therefore we should choose c to be as small as possible. The smallest c can be is the maximum of the observations, actually the maximum of the absolute value of the observations $\hat{c}_{MLE} = \max_{1 \leq i \leq n} |w_i|$.

MoM: As the problem states, $\mathrm{E}(W) = 0$ and therefore we can not use the first moment. In this case the variance is also equal to the second moment, which is equal to $(c - (-c))^2/12 = c^2/3$. Setting this equal to the empirical second moment, we obtain $\frac{1}{n} \sum_{i=1}^{n} x_i^2 = c^2/3$ and therefore $\hat{c}_{MoM} = \sqrt{3\bar{x^2}}$.

6. Chapter 18, Exercise 7.

An urn contains w marbles, one of which is black and the rest being white. We draw marbles from the urn one at a time, without replacement, until we draw the black one; let N denote the number of draws needed. Find the Method of Moments estimator of w based on N.

Consider the case when $n = 1$. Then the black ball comes out first, with probability $\frac{1}{w}$. Now consider the case $n = 2$. A white ball comes out first, with probability $\frac{w-1}{w}$, and then the black ball comes out with probability $\frac{1}{w-1}$, giving overall probability $\frac{1}{w}$. We can repeat this on and on and everytime the probability is $\frac{1}{w}$. Therefore the expectation is $\sum_{n=1}^{w} n \frac{1}{w} = \frac{1}{w} \frac{(w+1)w}{2} = (w+1)/2$. The method of moments estimator sets $\bar{N} = (w+1)/2$, or $\hat{w}_{MoM} = 2\bar{N} - 1$.

7. Chapter 18, Exercise 8.

Suppose $x_1, \ldots, x_n$ are uniformly distributed on $(0, c)$. Find the Method of Moments and Maximum Likelihood estimators of c, and compare their mean squared error.

The MLE is $\max_{1 \leq i \leq n} x_i$ and the MoM is $2\bar{x} - 1$. The mean square error of the MLE is $\mathrm{E}(\max_{1 \leq i \leq n} x_i - c)^2$. Note that $\Pr(\max_{1 \leq i \leq n} x_i \leq x) = F(x)^n = \left(\frac{x}{c}\right)^n$. Since the pdf is the deriviative of the cdf, the pdf for the MLE is $f_{MLE}(x) = \frac{\partial}{\partial x} \left(\frac{x}{c}\right)^n = \frac{n}{c} \left(\frac{x}{c}\right)^{n-1}$. Therefore the MSE is

$$\int_0^c (x - c)^2 \frac{n}{c} \left(\frac{x}{c}\right)^{n-1} dx = \int_0^c \frac{n}{c^n} x^{n+1} - \frac{2n}{c^{n-1}} x^n + \frac{n}{c^{n-2}} x^{n-1} dx$$

$$= \left(\frac{n}{c^n(n+2)} x^{n+2} - \frac{2n}{c^{n-1}(n+1)} x^{n+1} + \frac{1}{c^{n-2}} x^n\right)\Big|_0^c = \frac{3c}{(n+1)(n+2)}$$

The MoM is $2\bar{x}$. The distribution of $\bar{x}$ is approximately normal with mean $c/2$ and variance $c^2/12n$. Note that the MoM is unbiased. Therefore the MSE is equal to the variance.

8. Suppose that we want to model a species sampling experiment with a Negative Binomial distribution. In the experiment, species that are not captured are unobserved. Therefore the number of observed species follows a zero-truncated Negative Binomial distribution with pmf f(x) = P(X = x) / (1 - P(X = 0)). Suppose we observe D species with counts x1, x2, ..., xD.

a. Write out the Likelihood function for the observed data.

Note that this is a zero-truncated distribution, therefore $f(x) = \Pr(X = x | X > 0)$ with X negative binomially distributed. By the conditional probability law, this is equal to $\Pr(X = x)/(1 - \Pr(X = 0))$. Note that $1 - \Pr(X = 0) = 1 - (1 - p)^r$. Therefore $f(x) = \binom{x+r-1}{x}(1 - p)^r p^x/(1 - (1 - p)^r)$ and the likelihood is

$$\prod_{i=1}^{n} \binom{x_i + r - 1}{x_i}(1 - p)^r p^{x_i}/(1 - (1 - p)^r)$$

and log likelihood

$$\log L = \sum_{i=1}^{D} \log \binom{x_i + r - 1}{x_i} + r \log(1 - p) + x_i \log p - \log(1 - (1 - p)^r)$$

b. Write out the equations for Maximum Likelihood, but do not solve.

This is a two parameter distribution, so the MLE will require two derivatives.

$$\frac{\partial}{\partial r} \log L = \sum_{i=1}^{D} \binom{x_i + r - 1}{x_i}^{-1} \left(\frac{\partial}{\partial r} \binom{x_i + r - 1}{x_i}\right) + \log(1 - p) - (1 - (1 - p)^r)^{-1}(1 - p)^r \log(1 - p) = 0$$

$$\frac{\partial}{\partial p} \log L = \sum_{i=1}^{D} -r/(1 - p) + x_i/p - (1 - (1 - p)^r)(-r(1 - p)^{r-1}) = 0$$

c. Write out the equations for the Method of Moments.

The expectation is $\frac{pr}{1-p} \frac{1}{1-(1-p)^r}$. We can calculate the second moment as

$$\sum_{x=1}^{\infty} x^2 \binom{x + r - 1}{x}(1-p)^r p^x/(1-(1-p)^r) = \frac{1}{1-(1-p)^r} \sum_{x=0}^{\infty} \binom{x + r - 1}{x}(1-p)^r p^x = \frac{1}{1-(1-p)^r} \frac{pr}{(1-p)^2}$$

Therefore the MoM estimator satisfies

$$\bar{x} = \frac{pr}{1-p} \frac{1}{1-(1-p)^r}$$

and

$$\bar{x^2} = \frac{1}{1-(1-p)^r} \frac{pr}{(1-p)^2}$$