HW 6 Solutions

Stat 61

Problem 1

(a)

Let $n = 4$ be the sample size, and let X_1, X_2, X_3 be independent identically distributed Bernoulli random variables (they represent the outcome of George's spins) with the parameter θ (probability of heads). Next, let X_4 be a negative binomial random variable (it represents the outcome of Hilary's spins) with the parameter $r = 1$ (desired number of favorable outcomes) and θ (probability of heads).

The likelihood function is defined as:

$$\operatorname{lik}(\theta) = f(x_1, x_2, x_3, x_4 \mid \theta) = f(x_1 \mid \theta) \cdot f(x_2 \mid \theta) \cdot f(x_3 \mid \theta) \cdot f(x_4 \mid \theta)$$

$$= P(X_1 = x_1 \mid \theta) \cdot P(X_2 = x_2 \mid \theta) \cdot P(X_3 = x_3 \mid \theta) \cdot P(X_4 = x_4 \mid \theta) \,.$$

Substituting the definitions of the probability distribution functions of X_i's, we have that the likelihood function is

$$\operatorname{lik}(\theta) = (1-\theta)^3 \cdot \binom{4-1}{1-1} \cdot \theta^1 \cdot (1-\theta)^3 = \boxed{\theta \cdot (1-\theta)^6} \,,$$

since the probability of tails, due to complement rule, is $1 - \theta$ (and George observed 3 tails), and since the probability distribution function of a negative binomial random variable Y with the parameters r (desired number of favorable outcomes) and θ (probability of success) is

$$P(Y = k) = \binom{k-1}{r-1} \cdot \theta^r \cdot (1-\theta)^{k-r} \,.$$

Solution source: QuizletPlus

Problem 1

(b)

To find the MLE of θ, consider the log-likelihood, which is the natural logarithm of the above likelihood:

$$l(\theta) = \ln(\text{lik}(\theta)) = \ln(\theta) + 6 \cdot \ln(1-\theta)\,.$$

Now we need to find its global maximum on the interval $\langle 0, 1\rangle$ (where θ can take on values).

The derivative of l is

$$l'(\theta) = \frac{1}{\theta} - \frac{6}{1-\theta}\,.$$

Stationary points are the null points of the above derivative, so

$$l'(\theta) = 0 \iff \frac{1}{\theta} - \frac{6}{1-\theta} = 0 \iff 1-\theta = 6\cdot\theta \iff 7\theta = 1 \iff \theta = \frac{1}{7}\,.$$

The second derivative of l is

$$l''(\theta) = -\frac{1}{\theta^2} - \frac{6}{(1-\theta)^2} < 0,$$

which is strictly negative for all $0 < \theta < 1$. This means that the log-likelihood is strictly concave, which shows that at the point $\hat{\theta} = \frac{1}{7}$ the log-likelihood (and the likelihood as well) reaches its global maximum.

Therefore, the MLE of θ is

$$\boxed{\hat{\theta} = \frac{1}{7}}\,.$$

Solution source: QuizletPlus

Problem 2

(a)

For the method of moments, we first need to find the expected value of X, which is a continuous random variable with the density function as defined in the exercise (uniform distribution).

Since X is uniform on $[0, \theta]$, then its density function is

$$f(x \mid \theta) = \frac{1}{\theta}, \quad x \in [0, \theta].$$

So, using the definition of the expected value of a continuous random variable, we have

$$E(X) = \int_{-\infty}^{+\infty} x \cdot f(x \mid \theta)\, dx = \int_{0}^{\theta} \frac{x}{\theta}\, dx = \frac{1}{\theta} \cdot \left(\frac{x^2}{2}\right)\Bigg|_0^\theta = \frac{1}{\theta} \cdot \frac{\theta^2}{2} = \frac{\theta}{2}.$$

From the above expression, we can express θ as

$$\theta = 2 \cdot E(X).$$

Now, the method of moments simply suggests writing the **sample mean** $\overline{X}$ in place of $E(X)$, and that would be the method of moments estimate of θ. So, the desired estimate is

$$\boxed{\hat{\theta} = 2 \cdot \overline{X}}.$$

Using the linearity of the expectation, the mean value of $\hat{\theta}$ can easily be found as

$$E(\hat{\theta}) = 2 \cdot E(\overline{X}) = 2 \cdot E\left(\frac{1}{n} \cdot \sum_{i=1}^{n} X_i\right) = \frac{2}{n} \cdot \sum_{i=1}^{n} E(X_i) = \frac{2}{n} \cdot n \cdot \frac{\theta}{2} = \boxed{\theta}.$$

Now let's find the variance of $\hat{\theta}$.

First note that

$$\mathrm{Var}(\hat{\theta}) = \mathrm{Var}(2 \cdot \overline{X}) = 4 \cdot \mathrm{Var}(\overline{X}) = 4 \cdot \mathrm{Var}\left(\frac{1}{n} \cdot \sum_{i=1}^{n} X_i\right) = \frac{4}{n^2} \cdot \sum_{i=1}^{n} \mathrm{Var}(X_i)$$

$$= \frac{4}{n^2} \cdot n \cdot \mathrm{Var}(X_1) = \frac{4}{n} \cdot \mathrm{Var}(X_1),$$

since all X_i's are independent, so the variance of their sum is actually the sum of their variances, and they all have the same variance $\mathrm{Var}(X_1)$.

Solution source: QuizletPlus

Now, the variance of X_1 can be found as

$$\operatorname{Var}(X_1) = E(X_1^2) - [E(X_1)]^2.$$

Therefore, using the definition of the expected value of a (function of a) continuous random variable, we have

$$E(X_1^2) = \int_{-\infty}^{+\infty} x^2 \cdot f(x \mid \theta)\, dx = \int_0^\theta \frac{x^2}{\theta}\, dx = \frac{1}{\theta} \cdot \left(\frac{x^3}{3}\right)\Big|_0^\theta = \frac{1}{\theta} \cdot \frac{\theta^3}{3} = \frac{\theta^2}{3}.$$

Finally, the variance of X_1 is

$$\operatorname{Var}(X_1) = \frac{\theta^2}{3} - \left(\frac{\theta}{2}\right)^2 = \frac{\theta^2}{12},$$

which means that the variance of $\hat{\theta}$ is

$$\operatorname{Var}(\hat{\theta}) = \frac{4}{n} \cdot \frac{\theta^2}{12} = \boxed{\frac{\theta^2}{3n}}.$$

(b)

Let n be the sample size, and let $X_1, \ldots, X_n$ be independent identically distributed random variables with the same probability distribution function (the one described in the exercise).

To find the MLE of θ, we first define the likelihood function:

$$\operatorname{lik}(\theta) = f(x_1, \ldots, x_n \mid \theta) = f(x_1 \mid \theta) \cdots f(x_n \mid \theta).$$

Substituting the definition of the density function of X yields

$$\operatorname{lik}(\theta) = \underbrace{\frac{1}{\theta} \cdot \ldots \cdot \frac{1}{\theta}}_{n \text{ times}} = \frac{1}{\theta^n}.$$

Notice that this is a strictly decreasing function. But on what interval? Remember that the density function equals 1 when $x_i \in [0, \theta]$, i.e. when $0 \le x_i \le \theta$. This means that each x_i, $i \in \{1, 2, \ldots, n\}$ has to be less than or equal to θ, so their maximum has to be less than or equal to θ, i.e. $\max\limits_{1 \le i \le n} x_i \le \theta$.

Since the likelihood is a strictly decreasing function, we have to choose θ the least possible to maximize the likelihood. An obvious choice is $\theta = \max\limits_{1 \le i \le n} x_i$, since that's the lower bound of the interval where the likelihood function is defined, so we can conclude that at that point, the likelihood reaches its global maximum.

Therefore, the MLE of θ is

$$\boxed{\tilde{\theta} = \max_{1 \le i \le n} X_i}.$$

Solution source: QuizletPlus

Let's denote the MLE by $\widetilde{\theta} = X_{(n)}$, and let's first find its cumulative distribution function.

For $x \in [0, \theta]$, we have

$$F_{X_{(n)}}(x) = P(X_{(n)} \leq x) \overset{(1)}{=} P(X_1 \leq x, \ldots, X_n \leq x)$$

$$\overset{(2)}{=} P(X_1 \leq x) \cdot \ldots \cdot P(X_n \leq x) \overset{(3)}{=} [F_{X_1}(x)]^n$$

$$\overset{(4)}{=} \left(\frac{x}{\theta}\right).$$

In (1) we used a fact that if the maximum on n values is less than a certain value, then all those n values have to be less than that value.
In (2) we used that all X_i's are independent.
In (3) we used that all X_i's are identically distributed (they are all uniform on $[0, \theta]$), so they all have the same cumulative distribution function F_{X_1}.
In (4) we used that the cumulative distribution function of a uniform random variable on $[0, \theta]$ is $F_{X_1}(x) = \frac{x}{\theta}$, for $x \in [0, \theta]$.

Now, the density function of $X_{(n)}$ is simply the derivative of the above cumulative distribution function, which means that

$$f_{X_{(n)}}(x) = \left[F_{X_{(n)}}(x)\right]' = \left[\left(\frac{x}{\theta}\right)^n\right]' = \boxed{\frac{n \cdot x^{n-1}}{\theta^n}}, \quad x \in [0, \theta].$$

Problem 3

(a)

Next, using the definition of the expected value of a continuous random variable, we have that

$$E(X_{(n)}) = \int_{-\infty}^{+\infty} x \cdot f_{X_{(n)}}(x \mid \theta)\, dx = \int_{0}^{\theta} \frac{n \cdot x^n}{\theta^n}\, dx = \frac{n}{\theta^n} \cdot \left(\frac{x^{n+1}}{n+1} \right) \Bigg|_0^\theta = \boxed{\frac{n}{n+1} \cdot \theta}.$$

Furthermore, the variance of $X_{(n)}$ can be found as

$$\operatorname{Var}(X_{(n)}) = E(X_{(n)}^2) - [E(X_{(n)})]^2.$$

Therefore, using the definition of the expected value of a (function of a) continuous random variable, we have

$$E(X_{(n)}^2) = \int_{-\infty}^{+\infty} x^2 \cdot f_{X_{(n)}}(x \mid \theta)\, dx = \int_{0}^{\theta} \frac{n \cdot x^{n+1}}{\theta^n}\, dx = \frac{n}{\theta^n} \cdot \left(\frac{x^{n+2}}{n+2} \right) \Bigg|_0^\theta = \frac{n}{n+2} \cdot \theta^2 .$$

Finally, the variance of $X_{(n)}$ is

$$\operatorname{Var}(X_{(n)}) = \frac{n}{n+2} \cdot \theta^2 - \left(\frac{n}{n+1} \cdot \theta \right)^2 = \boxed{\frac{n\theta^2}{(n+1)^2 \cdot (n+2)}}.$$

We can notice that the variance of the MLE is significantly lower than the variance of the method of moments estimate, especially for larger values of n. This is because the variance of the MLE is of order n^{-2}, while the variance of the method of moments estimate is of order n^{-1}, and they both depend on θ the same way (both involve only θ^2).

Notice that the method of moments estimate is unbiased, since $E(\hat{\theta}) = \theta$. The MLE is biased, however, and the bias is

$$\text{Bias} = E(\tilde{\theta}) - \theta = \frac{n}{n+1} \cdot \theta - \theta = -\frac{\theta}{n+1} .$$

So, the bias is of order n^{-1}, but it can be lowered by forming a jackknife estimate $\tilde{\theta}_J$, as described in one of the exercises from Chapter 7.

Finally, remember that the mean squared error of an estimate $\hat{\theta}$ can be found as

$$\operatorname{MSE}(\hat{\theta}) = \operatorname{Var}(\hat{\theta}) + \text{Bias}^2 = \frac{\theta^2}{3n} + 0^2 = \frac{\theta^2}{3n} .$$

The mean squared error of $\tilde{\theta}$ is

$$\operatorname{MSE}(\tilde{\theta}) = \operatorname{Var}(\tilde{\theta}) + \text{Bias}^2 = \frac{n\theta^2}{(n+1)^2 \cdot (n+2)} + \left(-\frac{\theta}{n+1} \right)^2 = \frac{2\theta^2}{(n+1) \cdot (n+2)} .$$

Much like the variance, the mean squared error of the MLE is also significantly lower than the mean squared error of the method of moments estimate, especially for larger values of n.

Solution source: QuizletPlus

The method of moments estimate is unbiased, unlike the MLE, but the variance and the mean squared error of the MLE are significantly lower than with the method of moments estimate;

(b)

We have found that the expected value of the MLE $\tilde{\theta}$ of θ is

$$E(\tilde{\theta}) = \frac{n}{n+1} \cdot \theta \,.$$

Dividing the entire equality by $\frac{n}{n+1}$ yields that

$$\frac{n+1}{n} \cdot E(\tilde{\theta}) = \theta \,.$$

Using the linearity of the expectation, we have that

$$E\left(\frac{n+1}{n} \cdot \tilde{\theta}\right) = \theta \,.$$

So, if we define

$$\boxed{\tilde{\theta}_2 = \frac{n+1}{n} \cdot \tilde{\theta}} \,,$$

then we see that $E(\tilde{\theta}_2) = \theta$, which means that $\tilde{\theta}_2$ is a modification of the MLE that is unbiased for θ, so this is the desired estimate.

Solution source: QuizletPlus