Homework 2

ESE 402/542

Due October 9, 2020 at 11:59pm

Type or scan your answers as a single PDF file and submit on Canvas.

Problem 1. Suppose that $X_1, X_2, ..., X_n$ are i.i.d. random variables in a sample with the density function

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

(a) Use method of moments to estimate σ?

(b) Find the mle estimate of σ?

(c) What is the asymptotic variance of the mle?

Solution

(a)

$$\begin{aligned} E(X) &= \int_{-\infty}^{\infty} x f(x|\sigma) dx \\ &= \frac{1}{2\sigma} \int_{-\infty}^{\infty} x \exp\left(-\frac{|x|}{\sigma}\right) dx = 0 \end{aligned}$$

Since x is an odd function and $\exp -\frac{|x|}{\sigma}$ is an even function, integrating over $(-\infty, \infty)$ gives zero

First moment is independent of σ.
Second moment gives

$$\begin{aligned} E\left(X^2\right) &= \frac{1}{2\sigma} \int_{-\infty}^{\infty} x^2 \exp\left(-\frac{|x|}{\sigma}\right) dx \\ &= \frac{1}{\sigma} \int_{0}^{\infty} x^2 \exp\left(-\frac{x}{\sigma}\right) dx \\ &= \sigma^2 \int_{0}^{\infty} y^2 e^{-y} dy \\ &= 2\sigma^2 \end{aligned}$$

Hence,

$$\sigma = \sqrt{E\left(X^2\right)/2}$$
$$\hat{\sigma} = \sqrt{\frac{\hat{\mu}_2}{2}} = \sqrt{\frac{\sum_{i=1}^{n} X_i^2}{2n}}$$

(b) log likelihood is given by

$$\begin{aligned} l(\sigma) &= \log\left[\prod_{i=1}^{n} \frac{1}{2\sigma} \exp\left(-\frac{|X_i|}{\sigma}\right)\right] \\ &= -n\log(2\sigma) - \frac{1}{\sigma}\sum_{i=1}^{n} |X_i| \end{aligned}$$

Differentiating $l(\sigma)$ wrt σ, we get

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

Equating it to zero we get mle of σ as

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

(c) The asymptotic variance of the mle $\hat{\sigma}$ is given by 1/[nI(σ)] where

$$\begin{aligned} I(\sigma) &= -E\left[\frac{\partial^2}{\partial\sigma^2} \log f\left(X_1|\sigma\right)\right] \\ &= -E\left[\frac{1}{\sigma^2} - \frac{2\,|X_1|}{\sigma^3}\right] \\ &= -\frac{1}{\sigma^2} + \frac{2}{\sigma^2} \\ &= \frac{1}{\sigma^2} \end{aligned}$$

Variance of $\hat{\sigma}$ is σ^2/n

Problem 2. Given

$$I(\theta) = E\left[\frac{\partial}{\partial\theta}\log f(X|\theta)\right]^2$$

Under appropriate smoothness conditions, it can be proved that the probability distribution $\sqrt{nI(\theta_0)}\left(\hat{\theta}-\theta_0\right)$ tends to standard normal distribution.

(a) Show that $I(\theta) = -E\left[\frac{\partial^2}{\partial\theta^2}\log f(X|\theta)\right]$

(b) For the distribution in **Problem 1** find the confidence interval for the estimate $\hat{\sigma}$ (Hint: use standard normal property of $\sqrt{nI(\theta_0)}\left(\hat{\theta}-\theta_0\right)$ to find the confidence bounds)

(c) Suppose the distribution changes to a uniform distribution defined on on $[0,\ \theta]$ for $X_1, X_2, ..., X_n$ i.i.d random variables in a sample

- find the mle estimate $\hat{\theta}$ for the uniform distribution?
- find the asymptotic variance for mle estimate $\hat{\theta}$?
- find the confidence interval w.r.t $\hat{\theta}$?

(d) Suppose the distribution changes to a uniform distribution defined on on $[a,\ b]$ for $X_1, X_2, ..., X_n$ i.i.d random variables in a sample,

- find the MLE estimate for parameters a and b?

Solution

(a)

$$0 = \frac{\partial}{\partial\theta}\int f(x|\theta)dx = \int\left[\frac{\partial}{\partial\theta}\log f(x|\theta)\right]f(x|\theta)dx$$

$$\begin{aligned}
0 &= \frac{\partial}{\partial\theta}\int\left[\frac{\partial}{\partial\theta}\log f(x|\theta)\right]f(x|\theta)dx \\
&= \int\left[\frac{\partial^2}{\partial\theta^2}\log f(x|\theta)\right]f(x|\theta)dx + \int\left[\frac{\partial}{\partial\theta}\log f(x|\theta)\right]^2 f(x|\theta)dx \\
I(\theta) &= -E\left[\frac{\partial^2}{\partial\theta^2}\log f(X|\theta)\right]
\end{aligned}$$

(b) We know that $I(\theta_0)$ from problem 1 is $1/\sigma^2$.
Hence the confidence bounds can be written as

$$P\left(-z(\alpha/2) \le \sqrt{nI(\hat{\theta})}\left(\hat{\theta}-\theta_0\right) \le z(\alpha/2)\right) \approx 1-\alpha$$

of the inequalities yields

$$\hat{\theta} \pm z(\alpha/2)\frac{1}{\sqrt{nI(\hat{\theta})}}$$

So the confidence bound can be written as

$$\hat{\sigma} \pm z(\alpha/2)\frac{\sigma}{\sqrt{n}}$$

(c) **EXTRA CREDIT**
The log-likelihood is given by:

$$l(\theta) = \begin{cases} -n\log(\theta) & \text{if } \theta > \max_i x_i \\ -\infty & \text{otherwise} \end{cases}$$

which is strictly decreasing for $\theta > \max_i x_i$. Therefore, the MLE corresponds to the maximum of all the X_i:

$$\hat{\theta}_{\text{MLE}} = \max_{i=1,\ldots,n} X_i$$

- Asymptotic Variance cannot be found as Uniform Distribution is not a regular function. The question should have beenFind the Variance of mle Estimate.
  CDF of the mle estimate is given by

$$\begin{aligned} F_{\theta_2}(y) &= P\left(x_{(n)} < y\right) \\ &= \prod_{i=1}^{n} F_x(y) = F_x(y)^n \\ &= \left(\tfrac{y}{\theta}\right)^n \end{aligned}$$

The pdf of the MLE is the derivative form, as

$$f_{\theta_2}(y) = \tfrac{n}{\theta^n} y^{n-1}$$

$$\begin{aligned} E\left(\hat{\theta}_2\right) &= \int_0^\theta f_{\theta_2}(y) y\, dy \\ &= \int_0^\theta \frac{n}{\theta^n} y^n dy \\ &= \frac{n}{n+1}\theta \end{aligned}$$

Therefore, the mean of the MLE is

$$E\left(\hat{\theta}_2\right) = \int_0^\theta f_{\theta_2}(y) y dy$$

$$= \int_0^\theta \frac{n}{\theta^n} y^n dy$$

$$= \frac{n}{n+1}\theta$$

And the variance of the MLE is

$$\text{Var}\left(\hat{\theta}_2\right) = E\left(\hat{\theta_2^2}\right) - E\left(\hat{\theta}_2\right)^2$$

$$= \int_0^\theta f_{\theta_2}(y) y^2 dy - \left(\frac{n}{n+1}\theta\right)^2$$

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

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

Note that the above derivation for MLE for part 2(c) is not expected to be known or used for midterm. The above if for your knowledge.

- Confidence Interval

A function $Q(X_1, ..., X_n,)$ is a pivot if the distribution of Q does not depend on θ
Let Q $= max(X_i)/\theta$

$$\mathbb{P}(Q \leq t) = \mathbb{P}\left(max(X_{(i)}) \leq t\theta\right) = t^n$$

so Q is a pivot. Let $c_n = \alpha^{1/n}$. Then

$$\mathbb{P}\left(Q \leq c_n\right) = \alpha$$

Also, P(Q $\leq 1) = 1$. Therefore

$$1 - \alpha = \mathbb{P}(c_n \leq Q \leq 1) = \mathbb{P}\left(c_n \leq \frac{max(X_{(i)})}{\theta} \leq 1\right)$$

$$= \mathbb{P}\left(\frac{1}{c_n} \geq \frac{\theta}{max(X_{(i)})} \geq 1\right)$$

$$= \mathbb{P}\left(max(X_{(i)}) \leq \theta \leq \frac{max(X_{(i)})}{c_n}\right)$$

so a $1 - \alpha$ confidence interval is

$$\left(max(X_{(i)}), \frac{max(X_{(i)})}{\alpha^{1/n}}\right)$$

(d) For the uniform distribution given as:

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{if } x \in [a, b] \\ 0 & \text{otherwise} \end{cases}$$

The log-likelihood is:

$$l(a, b) = \begin{cases} -n \log(b - a) & \text{if } x_i \in [a, b] \\ -\infty & \text{otherwise} \end{cases}$$

Here, we have lower limit of the interval defines which is not equal to zero. If we have to maximize the log-likelihood given the data, we need to ensure that all the data values remain in the interval $[a, b]$. Hence, we can argue that the MLE estimates of a and b are:

$$\hat{a}_{\text{MLE}} = \min_{i=1,\ldots,n} X_i$$

$$\hat{b}_{\text{MLE}} = \max_{i=1,\ldots,n} X_i$$

Problem 3. Suppose $X_1, X_2, ..., X_n$ are i.i.d distributed in a sample with pdf

$$f(x|\theta) = \begin{cases} e^{-(x-\theta)}, & \text{if } x \geq \theta \\ 0, & \text{otherwise} \end{cases}$$

(a) Find the method of moments estimate of θ.

(b) Find the mle of θ. (Hint: Be careful, and do not differentiate before thinking. For what values of is the likelihood positive?)

Solution

(a)

$$\begin{aligned} \mu_1 = E[X] &= \int_\theta^\infty x e^{-(x-\theta)} dx \\ &= \theta + \int_0^\infty y e^{-y} dy \\ &= \theta + [(y)(-e^y)]|_{y=0}^{y=\infty} + \int_0^\infty e^{-y} dy \\ &= \theta + 1 \end{aligned}$$

Equating the sample first moment to the population first moment:

$$\begin{aligned} \mu_1 &= \hat{\mu}_1 \\ \theta + 1 &= \frac{1}{n} \sum_1^n X_i = \overline{X} \\ \Rightarrow \quad \hat{\theta} &= \overline{X} - 1 \end{aligned}$$

(b)

$$\begin{aligned} \text{lik}(\theta) &= f(X_1, \ldots, X_n|\theta) \\ &= \prod_{i=1}^n f(X_i|\theta) \\ &= \prod_{i=1}^n \left[e^{-(X_i-\theta)} \mathbf{1}_{[\theta,\infty)}(X_i) \right] \\ &= \left[e^{-\sum_{i=1}^n (X_i-\theta)} \right] \prod_{i=1}^n \left[\mathbf{1}_{[0,X_i]}(\theta) \right] \\ &= \left[e^{-\sum_{i=1}^n X_i} e^{n\theta} \right] \left[\mathbf{1}_{[0,\min(X_1,\ldots,X_n)]}(\theta) \right] \end{aligned}$$

lik (θ) is maximized by maximizing θ subject to $\theta \leq X_i$ for all $i = 1, \ldots, n$

i.e., $\quad \hat{\theta}_{MLE} = \min(X_1, \ldots, X_n)$

Problem 4. Suppose $X_1, X_2, ..., X_n \sim$ Poisson(λ). Given that the random variables are i.i.d, for $\theta = \exp(-\lambda)$,

1. Find an unbiased estimator of θ? (Note that it may not be the best estimator. Any unbiased estimator is fine).
2. Find the variance of the unbiased estimator you found and compare with the Cramer Rao lowerbound?

Solution

1. Consider $Pr(X_i = 0) = \frac{\lambda^0}{0!}e^{-\lambda} = e^{-\lambda}$

   For calculating the unbiased estimator, we can use this intuition and let $\hat{\theta} = \frac{1}{n}\Sigma_{i=1}^n \mathbf{1}(X_i = 0)$

$$\begin{aligned} E[\hat{\theta}] &= E[\frac{1}{n}\Sigma_{i=1}^n \mathbf{1}(X_i = 0)] \\ &= \frac{1}{n}E[\Sigma_{i=1}^n \mathbf{1}(X_i = 0)] \\ &= \frac{1}{n}ne^{-\lambda} \\ &= e^{-\lambda} \end{aligned}$$

   Hence, we can say that $\hat{\theta} = \frac{1}{n}\Sigma_{i=1}^n \mathbf{1}(X_i = 0)$ is an unbaised estimator.

2. Lets consider the variance of $\mathbf{1}(X_i = 0)$:

$$\begin{aligned} Var(\mathbf{1}(X_i = 0)) &= E[\mathbf{1}(X_i = 0)^2] - (E[\mathbf{1}(X_i = 0)])^2 \\ &= E[\mathbf{1}(X_i = 0)] - e^{-2*\lambda} \\ &= e^{-\lambda} - e^{-2*\lambda} \end{aligned}$$

   Hence we can calculate the variance of $\hat{\theta}$,

$$\begin{aligned} Var(\hat{\theta}) &= Var(\frac{1}{n}\Sigma_{i=1}^n \mathbf{1}(X_i = 0)) \\ &= \frac{1}{n^2}nVar(\mathbf{1}(X_i = 0)) \\ &= \frac{e^{-\lambda} - e^{-2*\lambda}}{n} \\ &= \frac{e^{-2*\lambda}(e^{\lambda} - 1)}{n} \end{aligned}$$

Now for n i.i.d Poisson distribution with parameter λ, we have the Cramer Rao bound as:

$$Var(\hat{\theta}) >= \frac{1}{\mathcal{I}(\theta)}$$

Here, we are estimating $\theta = e^{\lambda}$, hence we need to find the Fisher information for. So for finding that, we need to use the following identity:

$$\mathcal{I}(\theta) = \frac{\mathcal{I}(\lambda)}{(\frac{\partial\theta}{\partial\lambda})^2}$$

The Fisher information $\mathcal{I}(\lambda)$ for Poisson distribution is given as:

$$\begin{aligned}
\mathcal{I}(\lambda) &= E\left[\left(\frac{\partial f(X|\lambda)}{\partial\lambda}\right)^2\right] \\
&= E\left[\frac{\partial}{\partial\lambda}(-\lambda + X\log\lambda - \log X!)\right] \\
&= E\left[\left(\frac{X}{\lambda} - 1\right)^2\right] \\
&= Var_{\lambda}\left(\frac{X}{\lambda}\right) \\
&= \frac{Var(X)}{\lambda^2} = \frac{1}{\lambda}
\end{aligned}$$

Hence, for the sample with θ, Cramer Rao lower bound will be:

$$CRLB = \frac{(\frac{\partial\theta}{\partial\lambda})^2}{n\mathcal{I}(\lambda)} = \frac{e^{-2\lambda}\lambda}{n}$$

Now comparing the variance of the estimate and Cramer-Rao lower bound,

$$\begin{aligned}
\frac{e^{-2*\lambda}(e^{\lambda}-1)}{n} - \frac{e^{-2\lambda}\lambda}{n} ?0 \\
e^{\lambda} - 1 - \lambda ?0
\end{aligned}$$

For $\lambda > 0$, the above expression: $e^{\lambda} - 1 - \lambda > 0$. Hence, the variance is above lower bound.

Problem 5. We have access to a file consisting of $n = 10^4$ numbers. The numbers are either 1, or 2, or 3. Moreover, the value 1 appears $n_1 = 2600$ times in the file, the value 2 appears $n_2 = 5200$ times, and the value 3 appears $n_3 = 2200$ times. We know that these numbers are generated i.i.d. according to an unknown distribution.

(a) Let μ denote the mean of the distribution. Estimate the value of μ from sample data provided in the file and provide a 95% confidence interval.

(b) Assume now that the generating distribution of the data has the following form:

$$X = \begin{cases} 1, & \text{with probability } p_1, \\ 2, & \text{with probability } p_2, \\ 3, & \text{with probability } 1 - (p_1 + p_2) \end{cases}$$

We would like to estimate the value of the parameters p_1 and p_2. Consider the following estimator for the value of

$$p_1(X_1, \ldots, X_n) = \frac{1}{n} \Sigma_{i=1}^{n} 1\{X_i = 1\},$$

where $1\{A\}$ takes value 1 if A is true, and 0 otherwise.
Compute the estimate p_1 from the sample data provided in the file. Is this estimator and unbiased estimator for p_1? Justify your answer.

(c) Use the method of moments to estimate the value of p_1 and p_2. (you should compute the estimates from the sample data)

(d) Now, assume that the precise value of p_1 is given as $p_1 = \frac{1}{4}$. As a result, we now know that the distribution of the data has the form:

$$X = \begin{cases} 1, & \text{with probability } \frac{1}{4}, \\ 2, & \text{with probability } p_2, \\ 3, & \text{with probability } \frac{3}{4} - p_2 \end{cases}$$

We would like to estimate the value of the parameter p_2 from data. Find the maximum likelihood estimator for p_2 and provide a 95% confidence interval.

Solution

(a) $\hat{\mu} = \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i = 1.96$

$\hat{\sigma} = Var(X) = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2 = 0.48$

95% confidence interval:

$[\hat{\mu} - Z(\frac{\alpha}{2})\frac{\hat{\sigma}}{\sqrt{n}}, \hat{\mu} + Z(\frac{\alpha}{2})\frac{\hat{\sigma}}{\sqrt{n}}]$

$= [1.95, 1.97]$

(b) $\hat{p}_1 = \frac{1}{n}\sum_{i=1}^{n} 1(X_i = 1) = 0.26$

$\hat{p}_2 = \frac{1}{n}\sum_{i=1}^{n} 1(X_i = 2) = 0.52$

$E[\hat{p}_1] = E[\frac{1}{n}\sum_{i=1}^{n} 1(X_i = 1)]$

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

$= Pr(X_i = 1)$

$= p_1$

Unbiased Same for $\hat{p}_2$

(c) $\mu_1 = E[X] = p_1 + 2p_2 + 3(1 - p_1 - p_2) = 3 - 2p_1 - p_2$

$\mu_2 = E[X^2] = p_1 + 4p_2 + 9(1 - p_1 - p_2) = 9 - 8p_1 - 5p_2$

$$\begin{cases} \mu_1 = 3 - 2p_1 - p_2 \\ \mu_2 = 9 - 8p_1 - 5p_2 \end{cases} \qquad \begin{cases} p_1 = 3 - \frac{5}{2}\mu_1 + \frac{1}{2}\mu_2 \\ p_2 = 4\mu_1 - \mu_2 - 3 \end{cases}$$

$\hat{\mu}_1 = \frac{1}{n} \sum_{i=1}^{n} X_i = 1.96$

$\hat{\mu}_2 = \frac{1}{n} \sum_{i=1}^{n} X_i^2 = 4.32$

$\hat{p}_1 = 0.26$

$\hat{p}_2 = 0.52$

(d)

$$f(x) = \begin{cases} \frac{1}{4} & , X = 1 \\ p_2 & , X = 2 \\ \frac{3}{4} - p_2 & , X = 3 \end{cases}$$

$$f(x) = \frac{1}{4} \mathbb{1}(X=1) + p_2 \mathbb{1}(X=2) + (\tfrac{3}{4} - p_2) \mathbb{1}(X=3)$$

$$\mathcal{L}(p_2) = \prod_{i=1}^{n} f(x)$$

$$\ell(p_2) = \sum_{i=1}^{n} \log(f(x))$$

$$= \sum_{x=1} \log \frac{1}{4} + \sum_{x=2} \log p_2 + \sum_{x=3} \log(\tfrac{3}{4} - p_2)$$

$$= n_1 \log \frac{1}{4} + n_2 \log p_2 + n_3 \log(\tfrac{3}{4} - p_2)$$

$$\frac{\partial \ell(p_2)}{\partial p_2} = \frac{n_2}{p_2} - \frac{n_3}{\frac{3}{4} - p_2} = 0, \quad \hat{p}_2 = \frac{3n_2}{4(n_2+n_3)}$$

$$\frac{\partial^2 \ell(p_2)}{\partial p_2^2} = -\frac{n_2}{p_2^2} + \frac{n_3}{(\frac{3}{4} - p_2)^2}, \quad \left.\frac{\partial^2 \ell(p_2)}{\partial p_2^2}\right|_{\hat{p}_2} < 0$$

$$\hat{p}_{2,mle} = \frac{3n_2}{4(n_2+n_3)} = \frac{39}{74} = 0.527$$

$$I(p_2) = E\left[\left(\frac{\partial}{\partial p_2} \log f(x)\right)^2\right]$$

$$= E\left[\left(\frac{1}{p_2} 1(X=2) - \frac{1}{\frac{3}{4}-p_2} 1(X=3)\right)^2\right]$$

$$= E\left[\frac{1}{p_2^2} 1(X=2) + \frac{1}{(\frac{3}{4}-p_2)^2} 1(X=3)\right]$$

$$= \frac{1}{p_2} + \frac{1}{\frac{3}{4}-p_2}$$

$$= 6.38$$

$$(\hat{p}_2 - p_2) \sim N\left(0, \frac{1}{nI(p_2)}\right)$$

95% Confidence Interval:

$[0.519, 0.535]$

Problem 6. Download data_HW2.csv and load it into Python. The numbers are observations drawn i.i.d. from an exponential distribution with unknown parameter λ. Include your code in your homework write-up.

(a) Compute estimates for the sample mean and sample variance without using inbuilt functions. Compare your answers with inbuilt numpy functions.

(b) Suppose now that the standard deviation is known to be 0.25. Compute a 90% confidence interval for the population mean.(python libraries can be used to calculate the confidence interval).