Problem Set 6 Solutions

Problem 2 in Lecture 6

(a) Pmf is

$$f(x) = \begin{cases} p & \text{if} \quad x = 1 \\ 1-p & \text{if} \quad x = 0 \end{cases}$$

Thus, it is easy to check $f(x) = p^x(1-p)^{1-x}$.

(b)

$$l_n(p) = \sum_{i=1}^{n} \ln f(X_i) = \sum_{i=1}^{n} X_i \ln p + (1 - X_i)\ln(1-p).$$

(c) FOC implies

$$\sum_{i=1}^{n} X_i \frac{1}{\hat{p}} - (1 - X_i)\frac{1}{1-\hat{p}} = 0.$$

Thus,

$$\hat{p} = \bar{X}.$$

Problem 5 in Lecture 6

(a)

$$l_n(\alpha) = n \ln \alpha - (1+\alpha)\sum_{i=1}^{n} \ln X_i$$

(b) FOC is

$$\frac{n}{\hat{\alpha}} - \sum_{i=1}^{n} \ln X_i = 0.$$

Thus, $\hat{\alpha} = \frac{n}{\sum_{i=1}^{n} \ln X_i}$.

Problem 6 in Lecture 6

(a)

$$l_n(\theta) = -n \ln \pi - \sum_{i=1}^{n} \ln(1 + (X_i - \theta)^2)$$

(b) Taking derivative wrt θ gives the FOC

$$2\sum_{i=1}^{n} \frac{(X_i - \theta)}{(1 + (X_i - \theta)^2)} = 0.$$

Finding analytical solution may not be possible.

Problem 7 in Lecture 6

(a)
$$l_n(\theta) = -n \ln 2 - \sum_{i=1}^{n} |X_i - \theta|$$

(b) MLE maximizes $l_n(\theta)$. Since the first term of l_n is independent from the parameter, we have
$$\hat{\theta} \in \arg\min \sum_{i=1}^{n} |X_i - \theta|$$

The solution for this minimum absolute loss is a median of $\mathbf{X} = \{X_1, ..., X_n\}$. The basic intuition here is that as long as θ is not the median, by moving θ toward the median value with amount of ϵ one can decrease the value of the objective function by $n\epsilon$ and increase by $m\epsilon$ where $n > m$.

Think about the case that we only have $\{X_1, X_2, X_3\}$. Now suppose that $\theta > X_2$. Now by moving θ toward X_2 by ϵ the distance between θ and X_1, X_2 decreases by ϵ, and the distance between θ and X_3 increases by ϵ, so the change in total distance is $-2\epsilon + \epsilon = \epsilon$. The same argument holds when $\theta < X_2$. So as long as $\theta \neq X_2$, we can always decrease the value of the objective function.

Here, I will suggest a rigorous proof for odd n case. Let $X_{[i]}$ be an order statistic (i.e. $X_{[i]} = i$ th largest element of $\mathbf{X}$). When n is odd $\hat{\theta} = med(\mathbf{X}) = X_{[(n+1)/2]}$. Take any $\theta^* > \hat{\theta}$. Let i^* be such that $X_{[i^*]} \leq \theta^* \leq X_{[i^*+1]}$. By definition of θ^*, $(n+1)/2 \leq i^*$. Then

$$\begin{aligned}
\sum_{i=1}^{n} |X_i - \theta| - \sum_{i=1}^{n} |X_i - \hat{\theta}| &= \sum_{i>i^*} (X_{[i]} - \theta^*) + \sum_{i \leq i^*} (\theta^* - X_{[i]}) \\
&- \sum_{i \geq (n+1)/2} (X_{[i]} - \hat{\theta}) + \sum_{i<(n+1)/2} (\hat{\theta} - X_{[i]}) \\
&= \sum_{i>i^*} (\hat{\theta} - \theta^*) + \sum_{i<(n+1)/2} (\theta^* - \hat{\theta}) \\
&+ \sum_{i^* \geq i \geq (n+1)/2} (\hat{\theta} + \theta^* - 2X_{[i]}) \\
&= \sum_{i>i^*} (\hat{\theta} - \theta^*) + \sum_{i<(n+1)/2} (\theta^* - \hat{\theta}) \\
&+ \sum_{i^* \geq i \geq (n+1)/2} (\hat{\theta} - \theta^* + 2(\theta^* - X_{[i]})) \\
&= \sum_{i \geq (n+1)/2} (\hat{\theta} - \theta^*) + \sum_{i<(n+1)/2} (\theta^* - \hat{\theta}) \\
&+ \sum_{i^* \geq i \geq (n+1)/2} 2(\theta^* - X_{[i]}) \\
&= (\hat{\theta} - \theta^*) + \sum_{i^* \geq i \geq (n+1)/2} 2(\theta^* - X_{[i]}) \\
&= (\theta^* - \hat{\theta}) + \sum_{i^* \geq i > (n+1)/2} 2(\theta^* - X_{[i]}) \\
&> 0.
\end{aligned}$$

The similar argument holds for $\theta^* < \hat{\theta}$ case. For the even n case, essentially the similar argument holds but with a bit more complication because of the tie breaking issue.

Problem 9 in Lecture 6

$$I = -E\left(\frac{\partial^2}{\partial\alpha^2}\log f(X_i|\alpha)\right) = -E\left(\frac{\partial}{\partial\alpha}\left(\frac{1}{\alpha} - \log X_i\right)\right) = -E\left(-\frac{1}{\alpha^2}\right) = \frac{1}{\alpha^2}.$$

Problem 12 in Lecture 6

(a) The information for θ is $I = -E\left(\frac{\partial^2}{\partial\theta^2}\log f(X_i|\theta)\right) = -E\left(-\frac{1}{\theta^2}\right) = \frac{1}{\theta^2}$. Thus, the Cramer-Rao lower bound for θ is $CRLB = (nI)^{-1} = \frac{\theta^2}{n}$.

(b) From problem 1, $\hat{\theta} = \frac{1}{\bar{X}_n}$. By CLT, $\sqrt{n}(\bar{X}_n - E(X_i)) \xrightarrow{d} N(0, Var(X_i)) = N(0, \frac{1}{\theta^2})$. The function $g(x) = \frac{1}{x}$ is continuously differentiable at $x = E(X_i) = \frac{1}{\theta} > 0$. Thus by the Delta Method, $\sqrt{n}(\hat{\theta} - \theta) = \sqrt{n}(g(\bar{X}_n) - g(E(X_i))) \xrightarrow{d} g'(E(X_i))N(0, \frac{1}{\theta^2})$. Since $g'(E(X_i)) = g'(\frac{1}{\theta}) = -\theta^2$, we have $\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{d} N(0, (-\theta^2)^2\frac{1}{\theta^2}) = N(0, \theta^2)$.

(c) $\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{d} N(0, I^{-1}) = N(0, \theta^2)$. The answer is the same as in part (b).

Problem 14 in Lecture 6

(a) From problem 2, $\sqrt{n}(\hat{p} - p) \xrightarrow{d} N(0, p(1-p))$. Thus, $V = p(1-p)$. Since $p = E(X_i)$, a natural estimator for it is $\hat{p} = \bar{X}_n$. Thus, an estimator of V is $\hat{V} = \bar{X}_n(1 - \bar{X}_n)$.

(b) By WLLN, $\bar{X}_n \xrightarrow{p} E(X_i) = p$. By CMT, $\hat{V} \xrightarrow{p} p(1-p) = V$. Thus, $\hat{V}$ is consistent for V as $n \to \infty$.

(c) $s(\hat{p}) = \sqrt{\frac{1}{n}\hat{V}} = \sqrt{\frac{1}{n}\bar{X}_n(1 - \bar{X}_n)}$.

Problem 15 in Lecture 6

(a)

$$F_X(c) = \begin{cases} 0, & \text{if } c < 0 \\ \frac{c}{\theta}, & \text{if } c \in [0,1] \\ 1, & \text{if } c > 1 \end{cases}$$

(b)

$$\begin{aligned} F_{n(\hat{\theta}_n - \theta)}(x) &= P\left(n(\hat{\theta}_n - \theta) \le x\right) \\ &= P\left(\hat{\theta}_n \le \theta + \frac{x}{n}\right) \end{aligned}$$

Recall that in class we derive $\hat{\theta}_n = \max_{i=1,\ldots,n}\{X_i\}$, so then the event

$$\hat{\theta}_n \le \theta + \frac{x}{n} \iff \max_{i=1,\ldots,n}\{X_i\} \le \theta + \frac{x}{n} \iff X_i \le \theta + \frac{x}{n}, \text{for } i = 1, \ldots, n$$

Thus,

$$\begin{aligned} F_{n(\hat{\theta}_n-\theta)}(x) = P\left(\hat{\theta}_n \le \theta + \frac{x}{n}\right) &= P\left(X_i \le \theta + \frac{x}{n}, \text{for } i = 1, ..., n\right) \\ &= \prod_{i=1}^{n} P\left(X_i \le \theta + \frac{x}{n}\right) \\ &= \prod_{i=1}^{n} F_X(\theta + \frac{x}{n}) = \left(F_X(\theta + \frac{x}{n})\right)^n \end{aligned}$$

where the middle two equalities comes from the iid property.

(c) If $x \ge 0$, then $F_X(\theta + \frac{x}{n}) = 1$ for all n, so then $F_{n(\hat{\theta}_n-\theta)}(x) = 1^n = 1$, so $\lim_{n\to\infty} F_{n(\hat{\theta}_n-\theta)}(x) = 1$.

If $x < 0$, then $\theta + \frac{x}{n} \le \theta$, and for n large enough ($n > \frac{-x}{\theta}$), $\theta + \frac{x}{n} \ge 0$; so this means, when $n \to \infty$, $\theta + \frac{x}{n} \in [0, \theta]$. Thus,

$$\begin{aligned} \lim_{n\to\infty} F_{n(\hat{\theta}_n-\theta)}(x) &= \lim_{n\to\infty} \left(F_X(\theta + \frac{x}{n})\right)^n \\ &= \lim_{n\to\infty} \left(1 + \frac{x/\theta}{n}\right)^n \\ &= e^{x/\theta} \end{aligned}$$

(d) Since Z is an exponential distribution with parameter θ,

$$F_Z(z) = \begin{cases} 0, & \text{if } z < 0 \\ 1 - e^{\frac{-z}{\theta}}, & \text{if } z \ge 0 \end{cases}$$

then the CDF of $-Z$ is

$$F_{-Z}(x) = \begin{cases} e^{\frac{x}{\theta}}, & \text{if } x < 0 \\ 1, & \text{if } x \ge 0 \end{cases}$$

By the discussion in (c), we see that when $x < 0$, $\lim_{n\to\infty} F_{n(\hat{\theta}_n-\theta)}(x) = e^{x/\theta} = F_{-Z}(x)$; and when $x \ge 0$, $\lim_{n\to\infty} F_{n(\hat{\theta}_n-\theta)}(x) = 1 = F_{-Z}(x)$. Thus, $\lim_{n\to\infty} F_{n(\hat{\theta}_n-\theta)}(x) = F_{-Z}(x)$ for all $x \in R$, so by definition of convergence in distribution, $n(\hat{\theta}_n - \theta) \xrightarrow{d} -Z$

Problem 1 in Lecture 7

A t test is used here. The test statistic is $T = |t| = \left|\frac{\bar{X}_n - 1}{\sqrt{s^2/n}}\right|$, where $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$, $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2$. Under H_0, $t \sim t_{n-1}$. For a pre-specified size α, pick the critical value c such that $\alpha = Pr(\text{Reject}|H_0) = Pr(T > c|H_0) = 2(1 - F(c))$, where $F(\cdot)$ is the CDF of t_{n-1}. This gives $c = F^{-1}(1 - \frac{\alpha}{2})$, the $1 - \frac{\alpha}{2}$ quantile of t_{n-1} distribution. The test is: reject H_0 if $T > c$, accept H_0 if $T \le c$.

Problem 3 in Lecture 7

Note that $P(T > c|\mu = 0) = P(T > c|\mu = 1) = \alpha$ leads to a size α test, since for any case of the parameter we always have the size of α. Now first let's show that $P(T > c|\mu = 0) = \alpha$.

$$\begin{aligned}
P(T > c|\mu = 0) &= P(\min\{|\sqrt{n}\bar{X}_n|, |\sqrt{n}(\bar{X}_n - 1)|\} > c|\mu = 0) \\
&= P(\min\{|Z|, |Z - \sqrt{n}|\} > c) \\
&= 1 - P(\min\{|Z|, |Z - \sqrt{n}|\} \leq c) \\
&= 1 - (1 - \alpha) = \alpha
\end{aligned}$$

Here, the second line comes from the fact that $\sqrt{n}\bar{X}_n$ follows $N(0, 1)$ under $\mu = 0$. Then the last line follows by the definition of our critical value.

We can do something similar using the fact that Z and $-Z$ all follows standard normal distribution in calculating $P(T > c|\mu = 1)$. Here, note that $\sqrt{n}(\bar{X}_n - 1)$ follows standard normal under $\mu = 1$. Therefore :

$$\begin{aligned}
P(T > c|\mu = 1) &= P(\min\{|\sqrt{n}\bar{X}_n|, |\sqrt{n}(\bar{X}_n - 1)|\} > c|\mu = 1) \\
&= P(\min\{|Z + \sqrt{n}|, |Z|\} > c|\mu = 1) \\
&= P(\min\{| - Z - \sqrt{n}|, | - Z|\} > c|\mu = 1) \\
&= P(\min\{|Z - \sqrt{n}|, |Z|\} > c|\mu = 1) \\
&= 1 - P(\min\{|Z - \sqrt{n}|, |Z|\} \leq c|\mu = 1) = 1 - (1 - \alpha) = \alpha
\end{aligned}$$

Here, the third line follows because taking minus does not change the absolute value, and the fourth line follows since Z and $-Z$ have the same distribution. Then the last line comes again from the definition of c. Therefore we conclude that this test is a size α test.