Final Studyguide Problems Solution Set

1. Let X be a random variable whose pmf under $H_0 : \theta = \theta_0$ and $H_1 : \theta = \theta_1$ is given by

x	1	2	3	4	5	6	7
$f(x\|\theta_0)$	0.01	0.01	0.01	0.01	0.01	0.01	0.94
$f(x\|\theta_1)$	0.06	0.05	0.04	0.03	0.02	0.01	0.79

   (a) Find the UMP test for H_0 versus H_1 with size $\alpha = 0.05$.

   Solution: By Neyman-Pearson Lemma, the rejection region of the UMP test is $\{x : f(x|\theta_1) > kf(x|\theta_0)\}$. From the table we can see that such rejection region of size .05 is determined by any value of k between 1 and 2, which yields
   $x \in R = \{1, 2, 3, 4, 5\}$ with $P_{\theta_0}(X \in R) = 0.01 + 0.01 + 0.01 + 0.01 + 0.01 = 0.05$.

   (b) Compute the probability of Type II Error for this test.

   Solution: The probability if Type II Error is
   $P_{\theta_1}(X \in R^c) = P_{\theta_1}(X \in \{6, 7\}) = 0.80$.

2. Let $X_1, \ldots, X_n$ be *i.i.d.* random variables from Uniform$(0, \theta)$ distribution with pdf

$$f_X(x|\theta) = \frac{1}{\theta} I(0 \leq x \leq \theta), \qquad \theta > 0$$

   (a) Show that the $T = X_{(n)} = \max(X_1, \cdots, X_n)$ has Monotonic Likelihood Ratio (MLR) property.

   Solution: The joint likelihood function is $L(\theta|\mathbf{x}) = \frac{1}{\theta^n} I(\theta \geq x_{(n)})$, so the MLE is $\hat{\theta} = x_{(n)}$. Because $X_{(n)}/\theta$ follows a Beta$(n, 1)$ distribution, the pdf of $X_{(n)}$ is

$$g(t|\theta) = f_{X_{(n)}}(t|\theta) \quad = \quad \frac{n}{\theta}\left(\frac{t}{\theta}\right)^{n-1} I(0 < t < \theta).$$

   Let $\theta_2 > \theta_1$, then when $0 \leq t \leq \theta_1 < \theta_2$

$$\begin{aligned} \frac{g(t|\theta_2)}{g(t|\theta_1)} &= \frac{\frac{n}{\theta_2}\left(\frac{t}{\theta_2}\right)^{n-1}}{\frac{n}{\theta_1}\left(\frac{t}{\theta_1}\right)^{n-1}} \\ &= \left(\frac{\theta_1}{\theta_2}\right)^n \end{aligned}$$

   When $\theta_1 < t \leq \theta_2$, the denominator $g(t|\theta_1) = 0$, thus $\frac{g(t|\theta_2)}{g(t|\theta_1)} = \infty$. Hence $g(t|\theta_2)/g(t|\theta_1)$ is a non-decreasing function of t in $0 \leq t \leq \theta_2$, and it has an MLR.

(b) Show that the hypothesis testing procedure specified by rejection region

$$R = \{\mathbf{X} : X_{(n)} = \max(X_1, \cdots, X_n) > k\theta_0\}$$

is the UMP level α test for testing $H_0 : \theta \le \theta_0$ against $H_1 : \theta > \theta_0$ for its size. Represent k in terms of α in a closed form.

Solution: It is known that $T = X_{(n)}$ is a sufficient statistic for θ. Combining with the MLR property in part (a), by Karlin-Rubin Theorem, the UMP level α test rejects H_0 if $T > t_0$. The size of test when $t_0 \le \theta_0$ is

$$\begin{aligned}
\sup_{\theta \le \theta_0} \Pr(T > t_0|\theta) &= 1 - \inf_{\theta \le \theta_0} \Pr(T \le t_0|\theta) \\
&= 1 - \inf_{\theta \le \theta_0} \left[\prod_{i=1}^{n} \Pr(X_i \le t_0)\right] \\
&= 1 - \left(\frac{t_0}{\theta_0}\right)^n = \alpha
\end{aligned}$$

Therefore, $t_0 = (1-\alpha)^{\frac{1}{n}}\theta_0$ and $k = (1-\alpha)^{\frac{1}{n}}$.

(c) Find the $(1-\alpha)$ confidence interval of θ obtained from the hypothesis testing procedure above. (If you have not finished part (a), you may represent the confidence interval using k).

Solution: The acceptance region of the test above is

$$A(\theta_0) = \{\mathbf{x} : x_{(n)} \le (1-\alpha)^{\frac{1}{n}}\theta_0\}$$

Thus the one-sided $(1-\alpha)$ confidence interval is

$$C(\mathbf{x}) = \{\theta : x_{(n)} \le (1-\alpha)^{\frac{1}{n}}\theta\} = \{\theta : \theta \ge (1-\alpha)^{-\frac{1}{n}} x_{(n)}\}$$

3. Let $X_1, \ldots, X_n$ be *i.i.d.* random variables from Logistic$(\theta, 1)$ distribution with pdf

$$f_X(x|\theta) = \frac{e^{-(x-\theta)}}{1+e^{-(x-\theta)}}, \qquad -\infty < x < \infty, -\infty < \theta < \infty$$

(a) Show that

$$V_n = \frac{1}{1+\exp(\overline{X})}$$

are consistent estimator for $\Pr(X \le 0) = \frac{1}{1+e^{\theta}}$.

Solution: By WLLN,

$$\overline{X} \xrightarrow{P} \theta$$

By Continuous Mapping Theorem, let $g(y) = \frac{1}{1+e^y}$, then

$$V_n = g(\overline{X}) \xrightarrow{P} g(\theta) = \frac{1}{1+e^\theta} = \Pr(X \le 0)$$

Therefore, V_n is also a consistent estimator for $\Pr(X \le 0)$.

(b) We know that

$$W_n = \frac{1}{n}\sum_{i=1}^{n} I(X_i \le 0)$$

follows an asymptotic distribution

$$W_n \sim \mathcal{AN}\left(\frac{1}{1+e^\theta}, \frac{e^\theta}{n(1+e^\theta)^2}\right)$$

Compute $\mathrm{ARE}(V_n, W_n)$ and determine whether V_n is asymptotically more efficient than W_n. Justify your answer.

Solution: By CLT, $\overline{X}$ follows the asymptotic normal distribution

$$\overline{X} \sim \mathcal{AN}\left(\mathrm{E}X, \frac{\mathrm{Var}X}{n}\right) = \mathcal{AN}\left(\theta, \frac{\pi^2}{3n}\right)$$

Therefore, by the Delta Method, $V_n = g(\overline{X})$ follows an asymptotic normal distribution

$$V_n \sim \mathcal{AN}\left(g(\theta), \frac{\pi^2[g'(\theta)]^2}{3n}\right) = \mathcal{AN}\left(\frac{1}{1+e^\theta}, \frac{\pi^2 e^{2\theta}}{3n(1+e^\theta)^4}\right)$$

$$\begin{aligned} ARE(V_n, W_n) &= \frac{\frac{e^\theta}{n(1+e^\theta)^2}}{\frac{\pi^2 e^{2\theta}}{3n(1+e^\theta)^4}} \\ &= \frac{3(1+2e^\theta+e^{2\theta})}{\pi^2 e^\theta} = \frac{3}{\pi^2}\left(\frac{1}{e^\theta}+2+e^\theta\right) \ge \frac{12}{\pi^2} > 1 \end{aligned}$$

Therefore, V_n is asymptotically more efficient than W_n.

(c) When $n = 1$, construct a UMP level α test for testing $H_0 : \theta = 0$ against $H_1 : \theta = 1$.

Solution: By Neyman-Pearson Lemma, the UMP level α test reject H_0 if

$$\frac{f(x|\theta=1)}{f(x|\theta=0)} = \frac{\frac{e^{-x+1}}{1+e^{-x+1}}}{\frac{e^{-x}}{1+e^{-x}}} = \frac{e+e^{-x+1}}{1+e^{-x+1}} > k$$

Because $\frac{e+e^{-x+1}}{1+e^{-x+1}}$ is an increasing function of x, the rejection region is $X > k^*$ such that

$$\Pr(X > k^*|\theta = 0) = 1 - \frac{1}{1+e^{-k^*}} = \alpha$$

Therefore $e^{-k^*} = \frac{1}{1-\alpha} - 1 = \frac{\alpha}{1-\alpha}$, so $k^* = \log\left(\frac{1-\alpha}{\alpha}\right)$.

4. Let $X_1, \ldots, X_n$ be *i.i.d.* random variables from the following pdf

$$f_X(x|\theta) = \frac{1}{x\log\theta} I(1 \le x \le \theta), \qquad \theta > 1$$

(a) Construct a size α likelihood ratio test (LRT) for testing $H_0 : \theta \ge \theta_0$ against $H_1 : \theta < \theta_0$ for arbitrary $\theta_0 > 1$.

Solution: The likelihood function

$$L(\theta|\mathbf{x}) = \frac{I(x_{(1)} \ge 1)I(x_{(n)} \le \theta)}{(\log\theta)^n \prod_{i=1}^n x_i}$$

is a decreasing function of θ, so $\hat{\theta}_0 = \max(x_{(n)}, \theta_0)$ and $\hat{\theta} = x_{(n)}$. The LRT test statistic is

$$\lambda(\mathbf{x}) = \frac{L(\hat{\theta}_0|\mathbf{x})}{L(\hat{\theta}|\mathbf{x})} = \begin{cases} 1, & x_{(n)} \ge \theta_0 \\ \frac{(\log x_{(n)})^n}{(\log\theta_0)^n}, & x_{(n)} < \theta_0 \end{cases}$$

The rejection region $\lambda(\mathbf{x}) \le c$ can be reduced to $\log x_{(n)} \le k\log\theta_0$, or $x_{(n)} \le \theta_0^k$. The size of the test is

$$\begin{aligned} \sup_{\theta\ge\theta_0} \Pr(X_{(n)} \le \theta_0^k|\theta) &= \Pr(X_{(n)} \le \theta_0^k|\theta_0) = \prod_{i=1}^n \Pr(X_i \le \theta_0^k) \\ &= \left(\frac{\log\theta_0^k}{\log\theta_0}\right)^n = k^n = \alpha \end{aligned}$$

This, $k = \alpha^{\frac{1}{n}}$, and the size α test rejects H_0 if and only if $X_{(n)} \le \theta_0^{\alpha^{\frac{1}{n}}}$.

(b) Calculate the confidence coefficient of an interval estimator $[X_{(n)}, X_{(n)}^2]$ for θ.

Solution: The confidence coefficient is given by

$$\begin{aligned} \Pr[X_{(n)} \le \theta \le X_{(n)}^2] &= \Pr[\sqrt{\theta} \le X_{(n)} \le \theta] \\ &= \Pr[X_{(n)} \le \theta] - \Pr[X_{(n)} \le \sqrt{\theta}] \\ &= 1 - \left(\frac{\log\theta^{1/2}}{\log\theta}\right)^n \quad \text{(based on part (a) calculations)} \\ &= 1 - (1/2)^n \end{aligned}$$

(c) Consider a prior distribution of θ from the following pdf

$$f_Y(y|\alpha,\beta) = \frac{\beta\alpha^\beta}{y(\log y)^{\beta+1}}, \quad y > \alpha,\ \alpha > 0,\ \beta > 0$$

Calculate the posterior distribution of θ and justify whether the family of $f_Y(\theta|\alpha,\beta)$ is conjugate for $f_X(x|\theta)$ or not.

Solution: The prior distribution is $\pi(\theta;\alpha,\beta) = \frac{\beta\alpha^\beta I(\theta>\alpha)}{\theta(\log\theta)^{\beta+1}}$, and the posterior distribution becomes

$$\begin{aligned}
\pi(\theta|\mathbf{x};\alpha,\beta) &\propto \pi(\theta;\alpha,\beta)\prod_{i=1}^n f_X(x_i|\theta) \\
&= \frac{\beta\alpha^\beta I(x_{(n)} < \theta) I(\theta > \alpha)}{(\prod_{i=1}^n x_i)\theta(\log\theta)^{\beta+1}(\log\theta)^n} \\
&\propto \frac{\beta\alpha^\beta I(\theta > \max(x_{(n)},\alpha))}{\theta(\log\theta)^{\beta+n+1}} \propto \pi(\theta;\max(\alpha, x_{(n)}),\beta+n)
\end{aligned}$$

Therefore, the family of $\pi(\theta;\alpha)$ is a conjugate for $f_X(x|\alpha)$.

5. Let $X_1,\cdots,X_n$ be iid samples from the following geometric distribution

$$f(x|\theta) = \theta(1-\theta)^x \qquad x\in\{0,1,2,\cdots\},\ 0<\theta\le 1$$

whose expectation and variance is known to be $\mathrm{E}(X) = (1-\theta)/\theta$ and $\mathrm{Var}(X) = (1-\theta)/\theta^2$.

(a) Let $\pi(\theta) \sim Beta(\alpha,\beta)$ be the prior distribution, where α and β are known constants. Find the posterior distribution. Explain whether $\pi(\theta)$ is a conjugate family for $f(x|\theta)$ or not.

Solution: Using the usual calculation, the posterior distribution turns out to be

$$\pi(\theta|\mathbf{x}) = Beta\left(\alpha+n, \beta+\sum x_i\right)$$

Thus the prior is indeed a conjugate family for $f(x|\theta)$.

(b) Compute the Bayes' rule estimator of θ for loss function $L = (\theta-\hat{\theta})^2$

Solution Bayes estimator is the posterior mean, i.e.

$$\begin{aligned}
\hat{\theta} &= E[\theta|\mathbf{x}] = \frac{\alpha+n}{\alpha+n+\beta+\sum x_i} \\
&= \frac{\frac{\alpha}{n}+1}{\frac{\alpha+\beta}{n}+1+\overline{x}}
\end{aligned}$$

(c) Is the Bayes estimator in part (b) a consistent estimator? Justify your answer.

Solution: Here are the steps to demonstrate consistency:

(a) Using WLLN, $\overline{X} \xrightarrow{P} (1-\theta)/\theta$.

(b) Using Theorem 10.1.5, $\frac{\alpha+\beta}{n} + 1 + \overline{x} \xrightarrow{P} 1/\theta$.

(c) Using continuous mapping Theorem, $\frac{1}{\frac{\alpha+\beta}{n}+1+\overline{x}} \xrightarrow{P} \theta$.

(d) Using Theorem 10.1.5, $\frac{\frac{\alpha}{n}+1}{\frac{\alpha+\beta}{n}+1+\overline{x}} \xrightarrow{P} \theta$.

Therefore, the Bayes estimator is a consistent estimator.