**541B Midterm 1 Solutions**[1]

## 1. Question 1

Let $X$ be a Gaussian random variable with mean $\mu \in \mathbf{R}$ and variance 1, so that $X$ has PDF

$$f(x) := \frac{1}{\sqrt{2\pi}} e^{-(x-\mu)^2/2}, \qquad \forall\, x \in \mathbf{R}.$$

Provide a 99% confidence interval for $\mu$.

Your answer should use the function $\Psi(t) := \int_{-t}^{t} e^{-x^2/2} dx/\sqrt{2\pi}$, $\Psi\colon (0,\infty) \to (0,1)$, and/or the function $\Psi^{-1}\colon (0,1) \to (0,\infty)$. (Recall that $\Psi(\Psi^{-1}(s)) = s$ for all $s \in (0,1)$ and $\Psi^{-1}(\Psi(t)) = t$ for all $t > 0$.)

*Solution.* Since $X - \mu$ is a standard Gaussian random variable (and $X - \mu$ is a pivotal quantity) we have

$$\mathbf{P}(-a \leq X - \mu \leq a) = \int_{-a}^{a} e^{-x^2/2} dx/\sqrt{2\pi} = \Psi(a), \qquad \forall\, a > 0.$$

Setting $\Psi(a) = .99$, we have $\Psi^{-1}(.99) =: a$, so that

$$\mathbf{P}(X - \Psi^{-1}(.99) \leq \mu \leq X + \Psi^{-1}(.99)) = \Psi(\Psi^{-1}(.99)) = .99.$$

That is, the confidence interval for $\mu$ is

$$[X - \Psi^{-1}(.99), \quad X + \Psi^{-1}(.99)].$$

## 2. Question 2

Suppose $X$ is a binomial distributed random variable with parameters 2 and $\theta \in \{1/2, 3/4\}$. (That is, $X$ is the number of heads that result from flipping two coins, where each coin has probability $\theta$ of landing heads.)

We want to test the hypothesis $H_0$ that $\theta = 1/2$ versus the hypothesis $H_1$ that $\theta = 3/4$.

Let $\mathcal{T}$ be the set of hypothesis tests with significance level at most $1/20$.

(Recall that the significance level of a hypothesis test $\phi\colon \mathbf{R} \to [0,1]$ is $\sup_{\theta \in \Theta_0} \mathbf{E}_\theta \phi(X)$.)

Find a uniformly most powerful (UMP) class $\mathcal{T}$ hypothesis test $\phi\colon \mathbf{R} \to [0,1]$.

Justify your answer.

Hint: you can freely use the following facts for the PMF $f_\theta$ of $X$

$$\frac{f_{3/4}(0)}{f_{1/2}(0)} = \frac{1}{4}, \qquad \frac{f_{3/4}(1)}{f_{1/2}(1)} = \frac{3}{4}, \qquad \frac{f_{3/4}(2)}{f_{1/2}(2)} = \frac{9}{4}.$$

*Solution.* The Neyman-Pearson Lemma says that a UMP test for the class of tests with an upper bound on the significance level must be a likelihood ratio test with significance level equal to $1/20$. That is, there is some $k > 0$ and $\gamma \in [0,1]$ such that the following hypothesis test is UMP class $\mathcal{T}$.

$$\phi(x) := \begin{cases} 1 & , \text{ if } f_{\theta_1}(x) > k f_{\theta_0}(x) \\ 0 & , \text{ if } f_{\theta_1}(x) < k f_{\theta_0}(x) \\ \gamma & , \text{ if } f_{\theta_1}(x) = k f_{\theta_0}(x). \end{cases}$$

[1]September 18, 2023, 

After examining the likelihood ratios, we decide to choose $k = 9/4$, so that

$$\phi(x) := \begin{cases} 1 & \text{, if } f_{\theta_1}(x) > (9/4) f_{\theta_0}(x) \\ 0 & \text{, if } f_{\theta_1}(x) < (9/4) f_{\theta_0}(x) \\ \gamma & \text{, if } f_{\theta_1}(x) = (9/4) f_{\theta_0}(x) \end{cases} \qquad = \begin{cases} 0 & \text{, if } x \neq 2 \\ \gamma & \text{, if } x = 2. \end{cases}$$

Then $\mathbf{E}_{\theta_0}\phi(X) = \mathbf{P}_{\theta_0}(X = 2)\gamma = \mathbf{P}_{1/2}(X = 2)\gamma = (1/4)\gamma$. Since this quantity is equal to $1/20$ by assumption, we choose $\gamma := 1/5$. That is, our UMP test is

$$\phi(x) := \begin{cases} 0 & \text{, if } x \neq 2 \\ 1/5 & \text{, if } x = 2. \end{cases}$$

## 3. Question 3

Let $X_1, \ldots, X_n$ be a real-valued random sample of size $n$ so that $X_1$ has PDF given by

$$f(x) = \lambda e^{-\lambda x} 1_{x>0}, \qquad \forall\, x \in \mathbf{R},$$

where $\lambda > 0$ is an unknown parameter.

Suppose we want to test the hypothesis $H_0$ that $0 < \lambda \leq 2$ versus the hypothesis $H_1$ that $\lambda > 2$.

Describe the uniformly most powerful hypothesis test among all hypothesis tests with significance level at most $1/3$. Justify your answer.

*Solution.* We have for any $x = (x_1, \ldots, x_n) \in \mathbf{R}^n$,

$$f_\lambda(x) = 1_{x_1,\ldots,x_n \geq 0} \prod_{i=1}^{n} \lambda e^{-\lambda x_i} = 1_{x_1,\ldots,x_n \geq 0} \lambda^n e^{-\lambda \sum_{i=1}^n x_i}.$$

So, if $\lambda_1 > \lambda_0 > 0$, we have for any $x = (x_1, \ldots, x_n) \in \mathbf{R}^n$,

$$\frac{f_{\lambda_1}(x)}{f_{\lambda_0}(x)} = 1_{x_1,\ldots,x_n \geq 0} (\lambda_1/\lambda_0)^n e^{-(\lambda_1 - \lambda_0)\sum_{i=1}^n x_i}.$$

Since $\lambda_1 - \lambda_0 > 0$, this likelihood ratio is a strictly increasing function of $-\sum_{i=1}^n x_i$. We conclude from the Karlin-Rubin Theorem that there is a UMP test is of the form

$$\phi(x) := \begin{cases} 1 & \text{, if } -\sum_{i=1}^n x_i > c \\ 0 & \text{, if } -\sum_{i=1}^n x_i < c \\ \gamma & \text{, if } -\sum_{i=1}^n x_i = c. \end{cases}$$

for some $c \in \mathbf{R}$. Rewriting this, and noting that $-\sum_{i=1}^n x_i = c$ has probability zero for any $\lambda > 0$,

$$\phi(x) := \begin{cases} 1 & \text{, if } \sum_{i=1}^n x_i < -c \\ 0 & \text{, if } \sum_{i=1}^n x_i \geq -c. \end{cases}$$

## 4. Question 4

Let $X_1, \ldots, X_n$ be a real-valued random sample of size $n$ from a family of distributions $\{f_\theta \colon \theta \in \Theta\}$. (That is, $X$ has distribution $f_\theta$, where $X = (X_1, \ldots, X_n)$.) Suppose $\Theta = \mathbf{R}$. Fix $\theta \in \mathbf{R}$. Suppose $\{f_\theta \colon \theta \in \Theta\}$ has the monotone likelihood ratio property with respect to a statistic $Y = t(X)$ that is a continuous random variable.

Consider the set of hypothesis tests with rejection region $\{x \in \mathbf{R}^n \colon t(x) > c\}$, where $c \in \mathbf{R}$ is a constant (so that different values of $c$ correspond to different hypothesis tests.) Fix $\theta_0 \in \Theta$. Suppose we are testing $H_0 = \{\theta \leq \theta_0\}$ versus $H_1 = \{\theta > \theta_0\}$. For any $0 < \alpha < 1$, let $c_\alpha \in \mathbf{R}$ such that the rejection region $\{x \in \mathbf{R}^n \colon t(x) > c_\alpha\}$ has significance level $\alpha$. Define the $p$-value quantity

$$p(x) := \inf\{\alpha \in [0,1] \colon t(x) > c_\alpha\}, \qquad \forall\, x \in \mathbf{R}^n.$$

(Here we define the infimum of the empty set to be 1.)

Show that, if $X = x$, then $p(x)$ satisfies

$$p(x) = \mathbf{P}_{\theta_0}(t(X) > t(x)).$$

*Solution.* Recall that significance level $\alpha$ means that

$$\alpha = \sup_{\theta \in \Theta_0} \mathbf{E}_\theta \phi(X) = \sup_{\theta \leq \theta_0} \mathbf{P}_\theta(t(X) > c_\alpha)$$

Since the Karlin-Rubin Theorem implies that the power function is nondecreasing in $\theta$, we have

$$\alpha = \mathbf{P}_{\theta_0}(t(X) > c_\alpha). \qquad (*)$$

We also have

$$p(x) = \inf\{\alpha \in [0,1] \colon \phi_\alpha(x) = 1\} = \inf\{\alpha \in [0,1] \colon t(x) > c_\alpha\}.$$

The nested property implies that $\{\alpha \in [0,1] \colon t(x) > c\}$ is an interval, so that the infimum of this set is the smaller endpoint of that interval. That is, there exists some $\alpha \in [0,1]$ such that $p(x) = \alpha$ and $t(x) = c_\alpha$. So, from $(*)$,

$$\alpha = p(x) = \mathbf{P}_{\theta_0}(t(X) > c_\alpha) = \mathbf{P}_{\theta_0}(t(X) > t(x)).$$