MATH 667 EXAM 3 (December 6, 2017: 11:30am–2:00pm) NAME KEY

The exam is closed book; students are permitted to prepare one 8.5×11 page of formulas, notes, etc. that can be used during the exam. A calculator is permitted but not necessary for the exam. Do 4 out of the 5 problems (10 points each, 40 points total). Clearly indicate the problem that you are omitting; if it is not clear, then the first 4 problems will be graded.

Problem 1. (10 points) Suppose the probability mass function of a random variable is $f(x|\theta)$ where θ equals either 1, 2, or 3 and the values of the $f(x|\theta)$ are given in the following table for each θ.

x	1	2	3	4
$f(x\|\theta=1)$	.05	.05	.05	.85
$f(x\|\theta=2)$	.15	.25	.35	.25
$f(x\|\theta=3)$	.40	.30	.20	.10

(a - 2 pts) What is the rejection region for the UMP test with size .10 for testing $H_0 : \theta = 1$ versus $H_1 : \theta = 2$?
(b - 2 pts) What is the power for the UMP test in part (a) when $\theta = 2$?
(c - 2 pts) What is the rejection region for the UMP test with size .10 for testing $H_0 : \theta = 1$ versus $H_1 : \theta = 3$?
(d - 2 pts) What is the power for the UMP test in part (c) when $\theta = 3$?
(e - 2 pts) Is there a UMP test for testing $H_0 : \theta = 1$ versus $H_1 : \theta \neq 1$? If so, find it. If not, explain why not.

(a)

x	1	2	3	4
$\frac{f(x\|\theta=2)}{f(x\|\theta=1)}$	3	5	7	$\frac{.25}{.85}$

If we only reject when $x=3$, $P_{\theta=1}(\text{reject } H_0) = .05$. Then $P_{\theta=1}(X\in\{2,3\}) = .05+.05 = .10$. Next, we add $x=2$ to the rejection region.
By the Neyman-Pearson Lemma, $R=\{2,3\}$ is the rejection region of the UMP level .10 test of $H_0: \theta=1$ vs. $H_1: \theta=2$.

(b) $P_{\theta=2}(X\in\{2,3\}) = P_{\theta=2}(X=2) + P_{\theta=2}(X=3) = .25 + .35 = .60$

(c) Similar to (a), we find $R=\{1,2\}$ since

x	1	2	3	4
$\frac{f(x\|\theta=3)}{f(x\|\theta=1)}$	8	6	4	$\frac{.10}{.85}$

(d) $P_{\theta=3}(X\in\{1,2\}) = P_{\theta=3}(X=1) + P_{\theta=3}(X=2) = .40+.30 = .70$.

(e) The intended meaning of the problem is to ask if there is a UMP level .10 test.
There is not a UMP level .10 test since the test with $R=\{2,3\}$ is most powerful when $\theta=2$, but $P_{\theta=3}(X\in\{2,3\}) = .30+.20 = .50 < .70 = P_{\theta=3}(X\in\{1,2\})$.
Likewise, the test with $R=\{1,2\}$ is most powerful when $\theta=3$, but $P_{\theta=2}(X\in\{1,2\}) = .15+.25 = .40 < .60 = P_{\theta=2}(X\in\{2,3\})$.

Alternate answer: There is a UMP level .15 test since $R=\{1,2,3\}$ is the rejection region of the test with the most power when $\theta=2$ and $\theta=3$.

Problem 2. (10 points) Suppose $\mathcal{X}$ is the set of all possible values of a continuous random variable/vector $\boldsymbol{X}$ which has pdf $f(\boldsymbol{x}|\theta)$ where $\theta \in \{\theta_0, \theta_1\}$. Let R be a set such that $\boldsymbol{x}$ is in R if $f(\boldsymbol{x}|\theta_1) > kf(\boldsymbol{x}|\theta_0)$ and $\boldsymbol{x}$ is in R^c if $f(\boldsymbol{x}|\theta_1) \leq kf(\boldsymbol{x}|\theta_0)$ and let

$$\phi(\boldsymbol{x}) = \begin{cases} 1 & \text{if } \boldsymbol{x} \in R \\ 0 & \text{if } \boldsymbol{x} \in R^c \end{cases}.$$

For some $\alpha \in (0,1)$ and $\beta \in (\alpha, 1)$, assume that k is selected so that $\int \phi(\boldsymbol{x})f(\boldsymbol{x}|\theta_0)\, d\boldsymbol{x} = \alpha$ and $\int \phi(\boldsymbol{x})f(\boldsymbol{x}|\theta_1)\, d\boldsymbol{x} = \beta$.

(a - 5 pts) For any function ψ such that $\psi(\boldsymbol{x}) \in [0,1]$ for all $\boldsymbol{x} \in \mathcal{X}$, prove that

$$(\phi(\boldsymbol{x}) - \psi(\boldsymbol{x}))\,(f(\boldsymbol{x}|\theta_1) - kf(\boldsymbol{x}|\theta_0)) \geq 0 \text{ for all } \boldsymbol{x} \in \mathcal{X}.$$

(b - 5 pts) If $\int \psi(\boldsymbol{x})f(\boldsymbol{x}|\theta_0)\, d\boldsymbol{x} \leq \alpha$, then prove that $\int \psi(\boldsymbol{x})f(\boldsymbol{x}|\theta_1)\, d\boldsymbol{x} \leq \beta$.

(a) If $x \in R$, then $\phi(x) - \psi(x) = 1 - \psi(x) \geq 0$ and $f(x|\theta_1) - kf(x|\theta_0) > 0$
so that $(\phi(x) - \psi(x))(f(x|\theta_1) - kf(x|\theta_0)) \geq 0$.

If $x \in R^c$, then $\phi(x) - \psi(x) = 0 - \psi(x) \leq 0$ and $f(x|\theta_1) - kf(x|\theta_0) \leq 0$
so that $(\phi(x) - \psi(x))(f(x|\theta_1) - kf(x|\theta_0)) \geq 0$.

So, the statement is true for any $x \in \mathcal{X}$.

(b) From part (a), it follows that
$$\int (\phi(x) - \psi(x))(f(x|\theta_1) - kf(x|\theta_0))\, dx \geq 0.$$

Now, $\int (\phi(x) - \psi(x))(f(x|\theta_1) - kf(x|\theta_0))\, dx =$
$$\int \phi(x) f(x|\theta_1)\, dx - \int \psi(x) f(x|\theta_1)\, dx - k\int \phi(x) f(x|\theta_0)\, dx + k\int \psi(x) f(x|\theta_0)\, dx$$
$$= \beta - \beta' - k\alpha + k\alpha'$$

where $\beta' = \int \psi(x) f(x|\theta_1)\, dx$ and $\alpha' = \int \psi(x) f(x|\theta_0)\, dx$.

Since $k > 0$ (follows from $\alpha > 0$) and $\alpha' \leq \alpha$, we have

$$\beta - \beta' \cancel{- k\alpha + k\alpha} \geq \beta - \beta' - k\alpha + k\alpha' \geq 0$$

So that $\beta \geq \beta'$.

Problem 3. (10 points) Suppose that $X_1, \ldots, X_9$ is independent, identically distributed random variables from a distribution with probability density function $f(x) = \frac{1}{\theta} I_{[0,\theta]}(x)$ where $\theta > 0$ is unknown, and suppose that the experimenter is interested in testing

$$H_0 : \theta \leq 1 \text{ versus } H_1 : \theta > 1.$$

(a - 2 pts) Find a sufficient statistic $T(X_1, \ldots, X_9)$ for θ.
(b - 5 pts) Find the probability density function $g(t|\theta)$ of T and show that the family of pdf's has a nondecreasing monotone likelihood ratio.
(c - 3 pts) Find the value t_0 such that the test which rejects H_0 if and only if $T > t_0$ is a UMP level .05 test.

(a) The joint pdf of $X_1, \ldots, X_9$ is $f(\underline{x}) = \frac{1}{\theta^9} \prod_{i=1}^{9} I_{[0,\theta]}(x_i) = \frac{1}{\theta^9} I_{[0,\theta]}(\max_{1 \le i \le 9} x_i)$

so $\max_{1 \le i \le 9} X_i$ is sufficient for θ by the Factorization Theorem.

(b) First, find the cdf of $T = \max X_i$.

$$G(t) = P(T \le t) = P(X_1 \le t, \ldots, X_9 \le t) = \prod_{i=1}^{9} P(X_i \le t) = \begin{cases} 0 & \text{if } t < 0 \\ \left(\frac{t}{\theta}\right)^9 & \text{if } 0 \le t < \theta \\ 1 & \text{if } t \ge \theta \end{cases}.$$

The pdf of T is $g(t) = G'(t) = \frac{9t^8}{\theta^9} I_{[0,\theta]}(t)$.

For $\theta_1 < \theta_2$, $\frac{g(t|\theta_2)}{g(t|\theta_1)} = \begin{cases} \frac{\theta_1^9}{\theta_2^9} & \text{if } 0 \le t \le \theta_1 \\ \infty & \text{if } \theta_1 < t \le \theta_2 \end{cases}$ is nondecreasing in t,

so $\{g(t|\theta)\}$ has a nondecreasing monotone likelihood ratio.

(c) By the Karlin-Rubin Theorem, the test which rejects H_0 if and only if $T > t_0$ is a ~~UMP test~~ UMP level α test where $\alpha = P_{\theta=1}(T > t_0)$.

$$\underbrace{P_{\theta=1}(T > t_0) = 1 - P_{\theta=1}(T \le t_0) = 1 - \left(\frac{t_0}{1}\right)^9 = .05}_{t_0^9 = .95}$$

$$\boxed{t_0 = .95^{1/9}}$$

Problem 4. (10 points) Suppose X_1 and X_2 are independent identically distributed exponential random variables each with probability density function

$$f(x) = \begin{cases} \frac{1}{\beta}e^{-x/\beta} & \text{for } x > 0 \\ 0 & \text{otherwise} \end{cases}$$

where $\beta > 0$ is unknown.

(a - 5 pts) Show that $\dfrac{\min\{X_1, X_2\}}{\beta}$ is a pivot.

(b - 5 pts) Use the pivot in part (a) to find a $100(1-\alpha)\%$ confidence interval for β.

(a) $$P\left(\frac{X_{(1)}}{\beta} \le t\right) = P(X_{(1)} \le \beta t) = 1 - P(X_{(1)} > \beta t)$$

$$= 1 - P(X_1 > \beta t, X_2 > \beta t)$$

$$= 1 - P(X_1 > \beta t)\,P(X_2 > \beta t)$$

$$P(X > \beta t) = \int_{\beta t}^{\infty} \frac{1}{\beta} e^{-x/\beta}\,dx = \left[-e^{-x/\beta}\right]_{\beta t}^{\infty} = 0 + e^{-\beta t/\beta} = e^{-t}$$

$$= 1 - (e^{-t})(e^{-t})$$

$$= 1 - e^{-2t}$$

So the distribution of $\frac{X_{(1)}}{\beta}$ does not depend on β, and thus, $\frac{X_{(1)}}{\beta}$ is a pivot.

(b) $$P\left(\frac{X_{(1)}}{\beta} \le t\right) = \underbrace{1 - e^{-2t} = 1 - \alpha}_{}$$

$$\alpha = e^{-2t}$$

$$\ln \alpha = -2t$$

$$t = -\tfrac{1}{2}\ln\alpha$$

So $$P\left(\frac{X_{(1)}}{\beta} \le -\tfrac{1}{2}\ln\alpha\right) = 1 - \alpha$$

$$\Downarrow$$

$$P\left(\frac{-2X_{(1)}}{\ln\alpha} \le \beta\right) = 1 - \alpha$$

which shows that $\left[\frac{-2X_{(1)}}{\ln\alpha}, \infty\right)$ is a $100(1-\alpha)\%$ confidence interval for β.

Problem 5. (10 points) Suppose $X_1, \ldots, X_n$ are iid Bernoulli random variables each with probability mass function

$$f(x) = \begin{cases} p^x(1-p)^{1-x} & \text{for } x = 0, 1 \\ 0 & \text{otherwise} \end{cases}$$

where $p \in (0,1)$ is unknown. Let $\hat{p}_n = \frac{\sum_{i=1}^n X_i}{n}$.

(a - 2 pts) Show that $\hat{p}_n$ is a consistent estimator of p.

(b - 4 pts) Show that $\frac{\sqrt{n}(\hat{p}_n - p)}{\sqrt{\hat{p}_n(1-\hat{p}_n)}}$ converges in distribution to a standard normal random variable.

(c - 4 pts) Use your answer to part (b) to construct an approximate 99% confidence interval for p.

(a) $E[\hat{p}_n] = \frac{1}{n} E[\sum_{i=1}^n X_i] = \frac{1}{n}\sum_{i=1}^n E[X_i] = \frac{1}{n}(np) = p \Rightarrow \text{Bias}[\hat{p}_n] = 0$ for all n

$$Var[\hat{p}_n] = \frac{1}{n^2} Var[\sum_{i=1}^n X_i] \underset{\text{by independence}}{=} \frac{1}{n^2}\sum_{i=1}^n Var[X_i] = \frac{1}{n^2} n\{p(1-p)\} = \frac{p(1-p)}{n} \to 0 \text{ as } n \to \infty.$$

So $\hat{p}_n$ is a consistent sequence of estimators of p.

(b) The Central Limit Theorem implies that

$$\frac{\sqrt{n}(\hat{p}_n - p)}{\sqrt{p(1-p)}} \to N(0,1) \text{ in distribution.}$$

Part (a) shows that $\hat{p}_n \to p$ in probability. Since $h(\hat{p}_n) = \sqrt{\frac{p(1-p)}{\hat{p}_n(1-\hat{p}_n)}}$ is continuous function of $\hat{p}_n \in (0,1)$, $h(\hat{p}_n) \to h(p) = 1$ in probability.

So

$$\frac{\sqrt{n}(\hat{p}_n - p)}{\sqrt{\hat{p}_n(1-\hat{p}_n)}} = \sqrt{\frac{p(1-p)}{\hat{p}_n(1-\hat{p}_n)}} \frac{\sqrt{n}(\hat{p}_n - p)}{\sqrt{p(1-p)}} \to N(0,1) \text{ in distribution}$$

by Slutsky's Theorem since $\sqrt{\frac{p(1-p)}{\hat{p}_n(1-\hat{p}_n)}} \to 1$ in prob. and $\frac{\sqrt{n}(\hat{p}_n - p)}{\sqrt{p(1-p)}} \to N(0,1)$ in dist.

(c) $$P\left(-z_{.005} \le \frac{\sqrt{n}(\hat{p}_n - p)}{\sqrt{\hat{p}_n(1-\hat{p}_n)}} \le \overset{2.576}{z_{.005}}\right) \to .99$$

$-2.576\sqrt{\frac{\hat{p}_n(1-\hat{p}_n)}{n}} \le \hat{p}_n - p$ $\qquad$ $\hat{p}_n - p \le 2.576\sqrt{\frac{\hat{p}_n(1-\hat{p}_n)}{n}}$

$p \le \hat{p}_n + 2.576\sqrt{\frac{\hat{p}_n(1-\hat{p}_n)}{n}}$ $\qquad$ $\hat{p}_n - 2.576\sqrt{\frac{\hat{p}_n(1-\hat{p}_n)}{n}} \le p$

So $$P\left(\hat{p}_n - 2.576\sqrt{\frac{\hat{p}_n(1-\hat{p}_n)}{n}} \le p \le \hat{p}_n + 2.576\sqrt{\frac{\hat{p}_n(1-\hat{p}_n)}{n}}\right) \to .99$$

shows that $\hat{p}_n \pm 2.576\sqrt{\frac{\hat{p}_n(1-\hat{p}_n)}{n}}$ is a 99% confidence interval for p.