MATH 667 EXAM 2 (November 9, 2017: 2:30pm–3:45pm) 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) Let $X_1, \ldots, X_n$ be independent identically distributed random variables each with probability mass function

$$P(X = x|p) = \begin{cases} p(1-p)^x & \text{if } x \text{ is a nonnegative integer} \\ 0 & \text{otherwise} \end{cases},$$

and let p be a random variable with a beta prior distribution which has probability density function

$$\pi(p) = \begin{cases} \dfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha-1}(1-p)^{\beta-1} & \text{if } 0 < p < 1 \\ 0 & \text{otherwise} \end{cases}$$

and mean $E[p] = \dfrac{\alpha}{\alpha+\beta}$. Assume α and β are known and fixed.

(a - 7 pts) Let $Y = \sum_{i=1}^{n} X_i$. Find the posterior distribution of p given that $Y = y$.

(b - 3 pts) Find the Bayes estimator of p.

(a) Since $X_1, \ldots, X_n$ are independent geometric random variables representing the number of failures before a success, $\sum_{i=1}^{n} X_i$ is a negative binomial random variable with pmf $f(y|p) \propto p^n(1-p)^y$.

$$\pi(p|y) \propto f(y|p)\pi(p)$$

$$\propto p^n(1-p)^{\Sigma x_i} p^{\alpha-1}(1-p)^{\beta-1}, \quad 0<p<1$$

$$= p^{n+\alpha-1}(1-p)^{\Sigma x_i+\beta-1}$$

So $\pi(p|y) = C(\alpha,\beta)\, p^{(n+\alpha)-1}(1-p)^{(\Sigma x_i+\beta)-1}$

$\Rightarrow p|Y \sim \text{Beta}(n+\alpha, \Sigma x_i+\beta)$

(b) $E[p|X] = \boxed{\dfrac{n+\alpha}{n+\sum_{i=1} X_i+\alpha+\beta}}$

Problem 2. (10 points) Suppose that $X_1, \ldots, X_n$ are independent identically distributed (iid) random variables each with probability density function $f(x) = \lambda x^{-\lambda-1} I_{(1,\infty)}(x)$ where $\lambda > 0$.

(a - 3 pts) Show that $\frac{1}{n}\sum_{i=1}^{n} \ln X_i$ is an unbiased estimator of $\frac{1}{\lambda}$.

(b - 4 pts) Calculate the Cramér-Rao Lower Bound for the variance of an unbiased estimator of $\frac{1}{\lambda}$.

(c - 3 pts) Does $\frac{1}{n}\sum_{i=1}^{n} \ln X_i$ attain the Cramér-Rao Lower Bound? Show work to justify your answer.

(Hint: If X is a random variable with pdf or pmf of the form $f(x|\theta) = h(x)c(\theta)\exp(w(\theta)t(x))$, then $\mathrm{E}[w'(\theta)t(X)] = -\frac{d}{d\theta}\{\ln c(\theta)\}$ and $\mathrm{Var}[w'(\theta)t(X)] = -\frac{d^2}{d\theta^2}\{\ln c(\theta)\} - \mathrm{E}[w''(\theta)t(X)]$.)

(a) X_i has a pdf in the form $f(x) = \left(\frac{1}{x} I_{(1,\infty)}(x)\right)\cdot \lambda \cdot e^{(-\lambda)\cdot \ln x}$

$= h(x)\cdot c(\lambda)\cdot e^{w(\lambda)\cdot t(x)}$

so $E\left[\frac{d}{d\lambda}[-\lambda]\cdot \ln X_i\right] = -\frac{d}{d\lambda}[\ln \lambda] \Rightarrow E[-\ln X_i] = -\frac{1}{\lambda} \Rightarrow E[\ln X_i] = \frac{1}{\lambda}$.

$\Rightarrow E\left[\frac{1}{n}\sum_{i=1}^{n} \ln X_i\right] = \frac{1}{n}\sum_{i=1}^{n} E[\ln X_i] = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{\lambda} = \frac{1}{n}\cdot\frac{n}{\lambda} = \frac{1}{\lambda}$.

(b) $\ln f(x|\lambda) = \ln \lambda - (\lambda+1)\ln x$

$\frac{d}{d\lambda}\{\ln f(x|\lambda)\} = \frac{1}{\lambda} - \ln x$

$E\left[\left(\frac{d}{d\lambda}\{\ln f(x|\lambda)\}\right)^2\right] = E\left[\left(\ln X - \frac{1}{\lambda}\right)\right]^2 = \mathrm{Var}[\ln X] = -\frac{d}{d^2\lambda}[\ln \lambda] = -\left(-\frac{1}{\lambda^2}\right) = \frac{1}{\lambda^2}$

$\Rightarrow \mathrm{CRLB} = \frac{\left(\frac{d}{d\lambda}\left[\frac{1}{\lambda}\right]\right)^2}{n E\left[\left(\frac{d}{d\lambda}\ln f(X|\lambda)\right)^2\right]} = \frac{\left(-\frac{1}{\lambda^2}\right)^2}{n E\left[\left(\ln X - \frac{1}{\lambda}\right)^2\right]} = \frac{\frac{1}{\lambda^4}}{n\cdot\frac{1}{\lambda^2}} = \boxed{\frac{1}{n\lambda^2}}$.

(c) Yes, $\mathrm{Var}\left[\frac{1}{n}\sum_{i=1}^{n}\ln X_i\right] = \frac{1}{n^2}\sum_{i=1}^{n}\mathrm{Var}[\ln X_i] = \frac{1}{n^2}\left(n\cdot\frac{1}{\lambda^2}\right) = \frac{1}{n\lambda^2}$.

Problem 3. (10 points) Let $X_1, \ldots, X_n$ be a random sample from a population with parameter θ. Suppose that W is an unbiased estimator of θ and T is a sufficient statistic for θ, where both W and T are functions of $X_1, \ldots, X_n$.

(a - 1 pt) Let $\phi(T) = \mathrm{E}[W|T]$. Is $\phi(T)$ a statistic?

(b - 4 pts) Prove that $\phi(T)$ is an unbiased estimator of θ.

(c - 5 pts) Prove that $\mathrm{Var}[\phi(T)] \leq \mathrm{Var}[W]$ for all θ.

(a) Yes, since T is sufficient, $\phi(T)$ does not depend on θ.

(b) $E[\phi(T)] = E[E[W|T]] = E[W] = \theta$

since W is unbiased for θ

(c) $\mathrm{Var}[W] = \mathrm{Var}[E[W|T]] + \underbrace{E[\mathrm{Var}[W|T]]}_{\geq 0 \text{ since } \mathrm{Var}[W|T] \geq 0}$

$\Rightarrow \mathrm{Var}[W] \geq \mathrm{Var}[E[W|T]] = \mathrm{Var}[\phi(T)]$.

Problem 4. (10 points) Suppose $X_1, \ldots, X_n$ are independent identically distributed 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 $0 < p < 1$.

(a - 5 pts) Find a complete statistic which is sufficient for p. Justify your answer.

(b - 5 pts) Assume $n \geq 2$. Find the UMVUE of p^2. Justify your answer.

(a) X_i has the form of an exponential family since

$$f(x) = I_{\{0,1\}}(x)\,(1-p)\left(\frac{p}{1-p}\right)^x = \underbrace{I_{\{0,1\}}(x)}_{h(x)}\,\underbrace{(1-p)}_{c(p)}\,e^{\overbrace{x}^{t(x)}\overbrace{\ln\left(\frac{p}{1-p}\right)}^{w(p)}}.$$

Thus, $\sum_{j=1}^n t(X_j) = \sum_{j=1}^n X_j$ is sufficient for p.

Since $p \in (0,1)$ contains an open subset of $\mathbb{R}^1$, $\sum_{j=1}^n X_j$ is complete.

(b) First, we need an unbiased estimator for p^2.

Since X_1 and X_2 are independent and $E[X_1] = E[X_2] = p$,

$E[X_1 X_2] = E[X_1]E[X_2] = p^2$.

Then since $\sum_{j=1}^n X_j$ is complete and sufficient,

$E\left[X_1 X_2 \,\middle|\, \sum_{j=1}^n X_j\right]$ is the UMVUE of p^2.

To compute it, let $\phi(t) = E\left[X_1 X_2 \,\middle|\, \sum_{j=1}^n X_j = t\right] = P\left(X_1 X_2 = 1 \,\middle|\, \sum_{j=1}^n X_j = t\right)$

$$= P\left(X_1 = 1 \text{ and } X_2 = 1 \,\middle|\, \sum_{j=1}^n X_j = t\right)$$

$$= \frac{P\left(X_1 = 1 \text{ and } X_2 = 1 \text{ and } \sum_{j=1}^n X_j = t\right)}{P\left(\sum_{j=1}^n X_j = t\right)}$$

$$= \frac{P(X_1 = 1)\,P(X_2 = 1)\,P\left(\sum_{j=3}^n X_j = t-2\right)}{P\left(\sum_{j=1}^n X_j = t\right)}$$

$$= \frac{\cancel{p}\cdot\cancel{p}\cdot\binom{n-2}{t-2}\cancel{p^{t-2}}\,\cancel{(1-p)^{n-t}}}{\binom{n}{t}\cancel{p^t}\,\cancel{(1-p)^{n-t}}}$$

$$= \frac{\binom{n-2}{t-2}}{\binom{n}{t}} = \frac{\frac{(n-2)!}{(t-2)!\,(n-t)!}}{\frac{n!}{t!\,(n-t)!}} = \frac{t(t-1)}{n(n-1)}$$

So $\phi(\Sigma X_i) = \boxed{\frac{\Sigma X_i}{n} \cdot \frac{(\Sigma X_i - 1)}{(n-1)}}$

is the UMVUE for p^2.

Problem 5. (10 points) Suppose X_1 and X_2 are independent identically distributed normal random variables each with probability density function $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$ where μ and σ^2 are unknown, and suppose that the experimenter is interested in testing $H_0 : \mu = 0$ versus $H_1 : \mu \neq 0$ where $\Theta = \{(\mu, \sigma^2) : \mu \in (-\infty, \infty) \text{ and } \sigma^2 \in (0, \infty)\}$.

(a - 3 pts) Compute the likelihood ratio $\lambda(x_1, x_2) = \dfrac{\sup_{(\mu,\sigma^2)\in\Theta_0} L(\mu, \sigma^2; x_1, x_2)}{\sup_{(\mu,\sigma^2)\in\Theta} L(\mu, \sigma^2; x_1, x_2)}$.

(b - 3 pts) Show that the likelihood ratio test has a critical region of the form $\left\{(x_1, x_2) : \frac{|\bar{x}|}{s} \geq K\right\}$ where $\bar{x} = \frac{x_1 + x_2}{2}$ and $s^2 = (x_1 - \bar{x})^2 + (x_2 - \bar{x})^2$.

(c - 3 pts) Find the value of K (to at least 2 decimal places) such that the test in part (a) has size 0.05.

(d - 1 pt) If the observed data is $x_1 = 50$ and $x_2 = 60$, do we reject H_0 in the likelihood ratio test of $H_0 : \mu = 0$ versus $H_1 : \mu \neq 0$ with size 0.05?

Use the standard normal and/or t tables attached to this exam as needed.

(a) $L(\mu, \sigma^2) = \prod_{i=1}^{2} f(x_i) = \frac{1}{2\pi\sigma^2} e^{-\frac{1}{2\sigma^2}[(x_1-\mu)^2 + (x_2-\mu)^2]}$

Assuming $H_0 : \mu = 0$ is true, $\max_{\sigma^2} L(0, \sigma^2) = \frac{1}{2\pi\tilde{\sigma}^2} e^{-\frac{1}{2\tilde{\sigma}^2} 2\tilde{\sigma}^2}$

where $\tilde{\sigma}^2 = \frac{x_1^2 + x_2^2}{2}$ since $\tilde{\ell}(\sigma^2) = \ell(0, \sigma^2) = -\ln(2\pi) - \ln(\sigma^2) - \frac{1}{2\sigma^2}(x_1^2 + x^2)$ is minimized at $\tilde{\sigma}^2$.

$\tilde{\ell}'(\sigma^2) = \frac{1}{\sigma^2} - \frac{1}{2(\sigma^2)^2}(x_1^2 + x_2^2) = 0$

$\Downarrow$

$\tilde{\sigma}^2 = \frac{x_1^2 + x_2^2}{2}$

Also, $\max_{(\mu,\sigma^2)} L(\mu, \sigma^2) = \frac{1}{2\pi\hat{\sigma}^2} e^{-\frac{1}{2\hat{\sigma}^2} 2\hat{\sigma}^2}$

where $\hat{\sigma}^2 = \frac{(x_1-\hat{\mu})^2 + (x_2-\hat{\mu})^2}{2}$ since $\ell(\mu, \sigma^2) = -\ln(2\pi) - \ln(\sigma^2) - \frac{1}{2\sigma^2}[(x_1-\mu)^2 + (x_2-\mu)^2]$ is minimized at $\hat{\mu} = \frac{x_1 + x_2}{2}$ and $\hat{\sigma}^2 = \frac{(x_1-\bar{x})^2 + (x_2-\bar{x})^2}{2}$

$\frac{\partial \ell}{\partial \mu} = -\frac{1}{2\sigma^2}[-2(x_1-\mu) - 2(x_2-\mu)] = 0$

$\Downarrow$

$\hat{\mu} = \frac{x_1 + x_2}{2}$

$\frac{\partial \ell}{\partial \sigma^2} = -\frac{1}{\sigma^2} - \frac{1}{2(\sigma^2)^2}[(x_1-\mu)^2 + (x_2-\mu)^2]$

$\Downarrow$

$\hat{\sigma}^2 = \frac{(x_1-\hat{\mu})^2 + (x_2-\hat{\mu})^2}{2} = \frac{(x_1-\bar{x})^2 + (x_2-\bar{x})^2}{2}$.

(b)

So $\lambda(x_1, x_2) = \frac{\hat{\sigma}^2}{\tilde{\sigma}^2} = \frac{(x_1-\bar{x})^2 + (x_2-\bar{x})^2}{x_1^2 + x_2^2}$

$= \frac{(x_1-\bar{x})^2 + (x_2-\bar{x})^2}{[(x_1-\bar{x})^2 + 2\bar{x}^2] + [(x_2-\bar{x})^2 + 2\bar{x}^2]}$

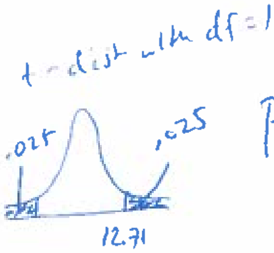
$= \frac{(x_1-\bar{x})^2 + (x_2-\bar{x})^2}{(x_1-\bar{x})^2 + (x_2-\bar{x})^2 + 4\bar{x}^2}$

$$= \frac{1}{1+\frac{4\bar{x}^2}{s^2}} = \frac{1}{1+4\left(\frac{|\bar{x}|}{s}\right)^2}$$

$\lambda(x_1,x_2)$ is small $\iff \frac{|\bar{x}|}{s}$ is large

So the rejection region has the form $\left\{(x_1,x_2) : \frac{|\bar{x}|}{s} \geq K\right\}$.

(c) If $\mu = 0$, then $\frac{\bar{X}}{\sqrt{S^2/2}} \sim t_{2-1=1}$ so



$$P\left(\frac{|\bar{X}|}{S/\sqrt{2}} \geq 12.71\right) = .05$$

$$\Downarrow$$

$$P\left(\frac{|\bar{X}|}{S} \geq \underbrace{\frac{12.71}{\sqrt{2}}}_{\boxed{8.98 = K}}\right) = .05$$

(d) $\frac{|\bar{X}|}{S} = \frac{55}{\sqrt{50}} \approx 7.78$

$\bar{X} = \frac{50+60}{2} = \frac{110}{2} = 55$

$S^2 = (50-55)^2 + (60-55)^2$

$= 25 + 25 = 50$

So we fail to reject H_0.