Midterm II

STA 131B
Winter 2016
University of California, Davis

Exam Rules: This exam is closed book and closed notes. Use of calculators, cell phones or other communication devices (including smart watch) is not allowed. You must show all of your work to receive credit.

Name : ______________________

ID : ______________________

Signature : ______________________

1. (48 points) Suppose $X_1, \dots, X_n$ form a random sample from a $N(\mu, \sigma^2)$ distribution with p.d.f.

$$\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad \text{for } -\infty < x < \infty.$$

In part (a)-(c) assume that σ is known.

a) Show that $\sum_{i=1}^n X_i$ is a sufficient statistic for μ. You must specify the u and v functions if you use the factorization theorem.
Answer. The joint p.d.f. of the distribution is

$$\begin{aligned} f_n(\mathbf{x}|\mu) &= \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}} = \frac{1}{(\sqrt{2\pi}\sigma)^n} e^{-\frac{\sum_{i=1}^n (x_i-\mu)^2}{2\sigma^2}} \\ &= \frac{1}{(\sqrt{2\pi}\sigma)^n} e^{-\frac{\sum_{i=1}^n x_i^2 - 2\mu \sum_{i=1}^n x_i + n\mu^2}{2\sigma^2}} \\ &= \frac{1}{(\sqrt{2\pi}\sigma)^n} e^{-\frac{\sum_{i=1}^n x_i^2}{2\sigma^2}} e^{-\frac{-2\mu \sum_{i=1}^n x_i + n\mu^2}{2\sigma^2}} \end{aligned}$$

By the factorization criterion, $u(\mathbf{x}) = \frac{1}{(\sqrt{2\pi}\sigma)^n} e^{-\frac{\sum_{i=1}^n x_i^2}{2\sigma^2}}$, $v(r(\mathbf{x}), \mu) = e^{-\frac{-2\mu \sum_{i=1}^n x_i + n\mu^2}{2\sigma^2}}$, and $r(\mathbf{x}) = \sum_{i=1}^n x_i$, so $\sum_{i=1}^n X_i$ is a sufficient statistic for μ.

b) Use the fact that the M.L.E. for μ is the sample mean. Is the M.L.E. minimal sufficient? Justify your answer.
Answer. The sample mean $\bar{X}_n$ is sufficient as it is a one-to-one function of $\sum_{i=1}^n X_i$. So it is minimal sufficient.

c) True or False (no details needed).

______ $e^{\sum_{i=1}^n X_i}$ is minimal sufficient.
Answer. True. Because $g(x) = e^x$ is a one-to-one function.
______ $(X_1, \sum_{i=2}^n X_i)$ is minimal sufficient.
Answer. False. Because $(X_1, \sum_{i=2}^n X_i)$ is not a function of the sufficient statistic $\sum_{i=1}^n X_i$.

d) Is the sample median inadmissible for μ (based on the square error loss function)? Justify your answer.

Answer. The sample median is inadmissible because it is not a function of $\sum_{i=1}^n X_i$.

e) Suppose that σ is also unknown, find a pair of joint sufficient statistics for $\theta = (\mu, \sigma)$. You must specify the u and v functions if you use the factorization theorem

Answer. Based on the joint p.d.f. and the factorization criterion, $u(\mathbf{x}) = 1$, $v(r(\mathbf{x}), \mu, \sigma) = \frac{1}{(\sqrt{2\pi}\sigma)^n} e^{-\frac{\sum_{i=1}^n x_i^2 - 2\mu \sum_{i=1}^n x_i + n\mu^2}{2\sigma^2}}$, and $r(\mathbf{x}) = (\sum_{i=1}^n x_i, \sum_{i=1}^n x_i^2)$, so the joint sufficient statistic is $\sum_{i=1}^n X_i$ and $\sum_{i=1}^n X_i^2$.

2. (52 points) Suppose that X is a random variable from a Bernoulli distribution with parameter $\theta = P(X = 1)$. That is, $f(x|\theta) = P(X = x) = \theta^x (1-\theta)^{1-x}$, for $x = 0$ and 1.
Note that $E(X) = \theta$ and $\text{var}(X) = \theta(1-\theta)$.

(a) Show that the Fisher information for θ based on a single observation X is $I(\theta) = \frac{1}{\theta(1-\theta)}$.

Answer.

$$\begin{aligned}
\log f(x|\theta) &= x \log\theta + (1-x)\log(1-\theta) \\
\frac{d \log f(x|\theta)}{d\theta} &= \frac{x}{\theta} - \frac{1-x}{1-\theta} \\
\frac{d^2 \log f(x|\theta)}{d\theta^2} &= -\frac{x}{\theta^2} - \frac{1-x}{(1-\theta)^2}
\end{aligned}$$

Thus,

$$I(\theta) = -E_\theta \left[\frac{d^2 \log f(X|\theta)}{d\theta^2} \right] = \frac{E_\theta(X)}{\theta^2} + \frac{1 - E_\theta(X)}{(1-\theta)^2} = \frac{\theta}{\theta^2} + \frac{1-\theta}{(1-\theta)^2} = \frac{1}{\theta} + \frac{1}{1-\theta} = \frac{1}{\theta(1-\theta)}.$$

(b) Find the Fisher information in a random sample $X_1, \ldots, X_{100}$ of size 100 from this Bernoulli distribution. You can just provide the answer without any derivation.

Answer. $I_n(\theta) = nI(\theta) = n/(\theta(1-\theta)) = 100/(\theta(1-\theta))$.

(c) Show that the sample mean $\bar{X} = \frac{1}{100}\sum_{i=1}^{100} X_i$ based on $X_1, \ldots, X_{100}$ is an unbiased estimator of θ.

Answer. $E(\bar{X}) = \frac{1}{100}\sum_{i=1}^{100} E(X_i) = \frac{1}{100} 100\theta = \theta$.

(d) Is the sample mean $\bar{X}$ in (c) an efficient estimator of θ? Justify your answer.

Answer. Yes. Because $m(\theta) = \theta$, and $\text{var}(\bar{X}) = \theta(1-\theta)/n = [m'(\theta)]^2 / I_n(\theta)$, which implies that $\bar{X}$ achieves the Cramer-Rao lower bound.

(e) Let $\hat{\theta}_1 = \frac{\sum_{i=1}^{100} X_i + 2}{105}$ be another estimator of θ. Find the bias and variance of $\hat{\theta}_1$.

Answer. $\text{bias}(\hat{\theta}_1) = E(\hat{\theta}_1) - \theta = \frac{\sum_{i=1}^{100} E(X_i) + 2}{105} - \theta = \frac{100\theta + 2}{105} - \theta = \frac{2 - 5\theta}{105}$.
$\text{var}(\hat{\theta}_1) = \frac{\sum_{i=1}^{100} \text{var}(X_i)}{(105)^2} = \frac{100\theta(1-\theta)}{(105)^2}$.

(f) Does $\bar{X}$ always have smaller MSE than $\hat{\theta}_1$? Provide a brief explanation.

Answer. No. $\text{MSE}(\hat{\theta}_1) = (\text{bias}(\hat{\theta}_1))^2 + \text{var}(\hat{\theta}_1) = \frac{(2-5\theta)^2}{(105)^2} + \frac{100\theta(1-\theta)}{(105)^2}$ and $\text{MSE}(\bar{X}) = \frac{\theta(1-\theta)}{100}$.

$$\text{MSE}(\bar{X}) - \text{MSE}(\hat{\theta}_1) = \left[\frac{1}{100} - \frac{100}{(105)^2} \right] \theta(1-\theta) - \frac{(2-5\theta)^2}{(105)^2} = \frac{1025}{100(105)^2}\theta(1-\theta) - \frac{(2-5\theta)^2}{(105)^2}$$

It is not always negative for any θ. For example, if $\theta = 2/5$, then $\text{MSE}(\bar{X}) - \text{MSE}(\hat{\theta}_1) > 0$.