Final - Winter 2015 - Math 462

NAME: ____________________

INSTRUCTIONS: You must show all your work to obtain credit for any answer. If your handwriting is not legible, your work may not be possible to grade. No calculators, smartphones, or other electronic devices may be used during this exam, nor are written notes of any kind.

Problem	Points Possible	Points Earned
1	15	
2	15	
3	10	
4	10	
5	10	
6	10	
TOTAL	70	

Problem 1 (15 points)**.** Suppose that $X_1, X_2, \dots, X_n$ are i.i.d. from the Normal$(0, \theta)$ pdf

$$f(x;\theta) = \frac{1}{\sqrt{2\pi\theta}} e^{-\frac{1}{2\theta}x^2}$$

Note: $E[X_1^4] = 3\theta^2$.

(a) Find a sufficient statistic for θ.

(b) Find the MLE of θ.

(c) Find the method of moments estimator of θ.

(d) Is the MLE unbiased? (Justify.) If not, modify it to obtain an unbiased estimator.

(e) Compute the Cramer-Rao Lower bound for the variance of unbiased estimators of θ. Calculate the variance for the unbiased estimator found above and compare to this bound.

Solution. (a) Let $\ell(\theta) = \ln L(\theta)$. Then

$$\begin{aligned}
f(x_1, \dots, x_n; \theta) &= (2\pi)^{-n/2}\theta^{-n/2} \exp\left(-\frac{1}{2\theta}\sum_i x_i^2\right) \\
\ell(\theta) &= \frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\theta) - \frac{n}{2\theta}\bar{x^2} \\
\ell'(\theta) &= -\frac{n}{2}\frac{1}{\theta} + \frac{n}{2\theta^2}\bar{x^2}
\end{aligned}$$

Thus the equation $\ell'(\theta) = 0$ is

$$\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n} x_i^2$$

(b) Since

$$E(X_i^2) = \theta,$$

the method of moments estimator is $\tilde{\theta} = \bar{x^2}$.

(c) Since $E(X_i^2) = \theta$, the MLE is unbiased.

(d) We have

$$\begin{aligned}
\ell'(\theta) &= -\frac{n}{2}\frac{1}{\theta} + \frac{n}{2\theta^2}\bar{x^2} \\
-\ell''(\theta) &= -\frac{n}{2\theta^2} + \frac{n}{\theta^3}\bar{x^2} \\
-E[\ell''(\theta)] &= -\frac{n}{2\theta^2} + \frac{n}{\theta^2} \\
&= \frac{n}{2\theta^2},
\end{aligned}$$

Thus the CL lower bound is $\frac{2\theta^2}{n}$.

(e)

$$\mathrm{Var}(X_i^2) = E(X_i^4) - E(X_i^2)^2 = 3\theta^2 - \theta^2 = 2\theta^2.$$

Thus the MLE meets the CR lower bound

□

Problem 2 (15 points)**.** Let $X_1, X_2, \ldots, X_n$ be i.i.d. with pdf

$$f(x;\theta) = \begin{cases} \frac{1}{\theta} & \text{if } 0 \le x \le \theta, \\ 0 & \text{otherwise}. \end{cases}$$

(a) Find a sufficient statistic for θ. Justify your answer.

(b) Find the pdf for $T_n = \max\{X_1, \ldots, X_n\}$.

(c) Calculate $P(\theta - \epsilon < T_n \le \theta)$, and then show that T_n is a consistent estimator of θ

(d) Suppose that the prior distribution on θ (with known parameters α and b) has pdf

$$f(\theta) = \begin{cases} \frac{\alpha b^\alpha}{\theta^{\alpha+1}} & \text{if } \theta \ge b \\ 0 & \text{otherwise}. \end{cases}$$

Note: This pdf has mean $\frac{\alpha b}{\alpha - 1}$.

Find the Bayes Estimator for θ assuming squared-error loss.

Hint: Use indicator functions to keep track of where the pdf's are positive. The posterior distribution for θ has the same form as the prior distribution (but with different values replacing parameters α and b).

(e) Is the Bayes estimator consistent? Justify your answer.

Solution. (a) We have

$$f(x_1, \ldots, x_n) = \prod_{i=1}^{n} I(0 \le x_i \le \theta)\theta^{-1} = \theta^{-n} I(0 \le \max_i x_i \le \theta)$$

whence $\max_i x_i$ is a sufficient statistic.

(b) For $0 \le t \le \theta$,

$$P(T_n \le t) = P(X_1 \le t)^n = (t/\theta)^n.$$

Thus

$$f_{T_n}(t) = \frac{n t^{n-1}}{\theta^n}, \quad 0 \le t \le \theta.$$

(c) We have, writing throughout that $\max\{b, x_i\} = \max\{b, x_1, \ldots, x_n\}$,

$$\begin{aligned} f(\theta \mid x_1, \ldots, x_n) &= c f(x_1, \ldots, x_n \mid \theta) f_\Theta(\theta) \\ &= c\theta^{-n} I(\theta > \max_i x_i) \frac{\alpha b^\alpha}{\theta^{\alpha+1}} I(\theta > b) \\ &= c' \frac{1}{\theta^{n+\alpha+1}} I(\theta > \max\{b, x_i\}) \\ &= \frac{(n+\alpha)(\max\{b, x_i\})^{n+\alpha}}{\theta^{n+\alpha+1}} I(\theta > \max\{b, x_i\}) \,. \end{aligned}$$

Thus, the posterior is the same form as the prior, with $\max\{b, x_i\}$ replacing m and $n+\alpha$ replacing α. We conclude that its expectation, the Bayes estimator is

$$\frac{(n+\alpha)\max\{b, x_i\}}{n+\alpha+1} \,.$$

Since

$$P(\theta \geq \max X_i > \theta - \varepsilon) = \int_{\theta-\varepsilon}^{\theta} \theta^{-n} t^{n-1} dt = 1 - \left(1 - \frac{\varepsilon}{\theta}\right)^n \to 1 \,,$$

we have $\max X_i \xrightarrow{\text{Pr}} \theta$. The function $w \mapsto \max\{w, b\}$ is continuous, it must be that $\max\{b, x_i\} \xrightarrow{\text{Pr}} \max\{\theta, b\} = \theta$.

Thus, since $(n+\alpha)/(n+\alpha+1) \to 1$, we have

$$\frac{(n+\alpha)\max\{b, x_i\}}{n+\alpha+1} \xrightarrow{\text{Pr}} \theta \,,$$

and the estimator is consistent.

$\square$

Problem 3 (10 points)**.** Suppose that you have two samples from Normal distributions, $N(\mu_X, \sigma_X^2)$ and $N(\mu_Y, \sigma_Y^2)$, respectively. The first sample is of size 4, and the second is of size 8. The sample variances are 2 and $\frac{4}{7}$, respectively.

(a) If the sample means are 3.1 and 4.3, test at level 0.05 the hypothesis that $\mu_X = \mu_Y$ against the two-sided alternative. Assume that the variances of the distributions are the same when carrying out this test.

Use the approximation $\sqrt{\frac{3}{8}} \approx 0.6$.

(b) Test at level 0.01 the hypothesis that $\sigma_X^2 = \sigma_Y^2$ against the alternative that $\sigma_X^2 > \sigma_Y^2$.

(c) Prior knowledge leads you to believe that $\sigma_X^2 = 3$. Test this hypothesis, at level 0.01, against the alternative that $\sigma_X^2 \neq 3$.

Proof. We have $S_p^2 = (3 \cdot 2 + 7 \cdot (4/7))/10 = 1$, so $S_p = 1$. Thus

$$T = \frac{3.1 - 4.3}{1\sqrt{\frac{1}{4} + \frac{1}{8}}} = -1.95$$

Since $t_{0.025} = 2.23$, we do not reject H_0.

Since

$$F = \frac{2}{(4/7)} = 3.5$$

and $F_{3,7,0.0.05} = 8.45$, we do not reject H_0.

Since

$$\frac{3 \cdot 2}{3} = 2,$$

and the critical values for the chi-squared with 3 degrees of freedom are 0.07 and 12.8, we do not reject H_0. $\square$

Problem 4 (10 points)**.** Suppose that a stock can either go up, stay the same, or go down each day. Denote the probabilities of these three events as p_u, p_s and p_d, respectively.

You have a theory that the probabilities of these three events are $0.4, 0.2, 0.4$, respectively.

You observe the stock on 100 days and find that it goes up, stays the same, and goes down $45, 15$ and 40 days, respectively.

(a) Find the expected number of days out of 100 for which the stock goes up, stays the same, and goes down. Write down the sum, but no need to complete the arithmetic, which equals the chi-squared statistic for testing the hypothesis $H_0 : (p_u, p_s, p_d) = (0.4, 0.2, 0.4)$.

(b) The value of the chi-squared statistic is 1.875. Carry out, at level 0.025, the hypothesis test of $H_0 : (p_u, p_s, p_d) = (0.4, 0.2, 0.4)$.

Solution. The expected counts are $40, 20, 40$ respectively.

$$D = (40-45)^2/40 + (20-15)^2/20 + (40-40)^2/40\,.$$

The critical value for D based on a chi-squared with 2 degrees of freedom equals 7.38, since $D < 7.38$ we do not reject. $\square$

Problem 5 (10 points)**.** Suppose that the random variable X has the pdf

$$f(x;\theta) = \frac{1}{2}e^{-|x-\theta|}$$

(a) Consider the test $H_0 : \theta = 0$ vs $H_0 : \theta \neq 0$, based on the single observation X, which rejects H_0 if and only if $|X| > 2$. Find the significance level of this test.

(b) Find the power of the above test when $\theta = 1$.

Solution. We have

$$P(|X| > 2) = 2\int_2^\infty \frac{1}{2}e^{-x}dx = e^{-2}$$

so the signficance level is e^{-2}.

Note that $X - \theta$ has pdf $f(x;0)$ for all θ. Also

$$\begin{aligned}
P(|X| > 2 \mid \theta = 1) &= P(X > 2) + P(X < -2) \\
&= P(X - 1 > 1) + P(X - 1 < -3) \\
&= \int_1^\infty \frac{1}{2}e^{-x}dx + \int_{-\infty}^{3} \frac{1}{2}e^{x}dx \\
&= \frac{e^{-1} + e^{-3}}{2}.
\end{aligned}$$

□

Problem 6 (10 points)**.** Suppose you want to evaluate a new drug for the flu. The manufacturer claims that the drug increases the probability p that a patient recovers in under three days. You give the drug to 100 flu patients and record the number X of patients who recover in under three days. Note that, if you give no treatment, the chance a person will recover from the flu in under three days is 0.5.

(a) Write down, in terms of X, an approximate level 0.05 test of the hypothesis that the treatment is no better than doing nothing (i.e. that $p = 1/2$) versus the alternative that it is better ($p > 1/2$).

(b) Carry out this test if you observe $X = 55$.

(c) Find the approximate power of the test above when $p = 0.6$.
(You can use the approximation $\sqrt{24} \approx 5$.)

Solution. Reject if

$$Z = \frac{X - 50}{\sqrt{100(1/2)(1/2)}} = \frac{X - 50}{5} > 1.64$$

For $X = 55$ we have $Z = 1$ so we do not reject.

The power when $p = 0.6$ is

$$P((X - 50)/5 > 1.64) = P((X - 60)/5 > -0.36) \approx 0.64\,.$$

□