Name: Solutions

Math 30 – Mathematical Statistics

Midterm 1 Practice Exam 1

Instructions:

1. Show all work. You may receive partial credit for partially completed problems.
2. You may use calculators and a one-sided sheet of reference notes. You may not use any other references or any texts, apart from the provided tables (if any), and distribution sheet.
3. You may not discuss the exam with anyone but me.
4. Suggestion: Read all questions before beginning and complete the ones you know best first. Point values per problem are displayed below if that helps you allocate your time among problems, as well as shown after each part of the problem.
5. Good luck!

Problem	1	2	3	Total
Points Earned				
Possible Points	15	10	10	35

1. An orchard is trying to estimate the proportion of bad apples in their large shipments to local markets. They plan to take a random sample of apples from the next outgoing shipment. For each apple in the random sample, they will record a 0 or 1 (0 = good, 1 = bad) as the observation for that apple. Let θ be the proportion of bad apples in the shipment. Hence, the pdf associated with a single observation is $f(x \mid \theta) = \theta^x (1-\theta)^{1-x}, x = 0,1, \theta \in (0,1)$.

a. For this example, the random sample is a sample from a Bernoulli distribution with parameter θ. (1)

b. Find the MLE for θ based on a random sample of size n. (4)

$$f_n(\underset{\sim}{x} \mid \theta) = \theta^{\sum x_i}(1-\theta)^{n-\sum x_i} = L(\theta)$$

$$\ell(\theta) = \sum X_i \ln \theta + (n - \sum X_i) \ln (1-\theta)$$

$$\ell'(\theta) = \frac{\sum X_i}{\theta} + \frac{n-\sum X_i}{1-\theta}(-1) = 0$$

$$\frac{\sum X_i}{\theta} = \frac{n-\sum X_i}{1-\theta} \qquad (1-\theta)\sum X_i = n\theta - \theta \sum X_i$$

$$\sum X_i = n\theta \qquad \hat{\theta} = \frac{\sum X_i}{n} = \bar{X}$$

$\bar{X}$ is MLE for θ

c. Find the MLE for θ(1-θ) based on a random sample of size n. (2)

By the invariance property, the MLE is $\bar{X}(1-\bar{X})$.

d. Sufficient statistics can be identified using the factorization criterion. (1)

e. Identify a sufficient statistic for θ. (3)

$$L(\theta) = \theta^{\sum x_i}(1-\theta)^{n-\sum x_i}$$

$$= \underbrace{\theta^T (1-\theta)^{n-T}}_{g(T,\theta)} \cdot \underbrace{1}_{h(\underset{\sim}{x})}$$

Let $T = \sum X_i$

$T = \sum X_i$ is suff for θ by FC.

f. Is the MLE you found in b. minimal sufficient for θ? Explain in one sentence. (2)

yes. $\bar{X}$ is a 1-1 fn of $T = \sum X_i$, so $\bar{X}$ is suff and if an MLE is suff, it is min. suff.

g. Is the MLE you found in b. an MVUE for θ? Explain in one sentence. (2)

$$E(\bar{X}) = \frac{1}{n} E(\sum X_i) = \frac{1}{n} \sum E(X_i) = \frac{1}{n} \sum \theta = \theta$$

Yes b/c MLE is unbiased and suff ⇒ MVUE.

2. You are sampling a single observation from a distribution with pdf: $f(x \mid \theta) = \theta^2 x \exp(-\theta x)$, x>0, and θ>0.

$X \sim \text{Gamma}\left(\alpha = 2, \beta = \frac{1}{\theta}\right)$

a. Find an unbiased estimator for θ from your single observation. (6)

$E(X) = \alpha\beta = \frac{2}{\theta}$ so see inverse relationship.

$E\left(\frac{1}{X}\right) = \int_0^\infty \theta^2 x \left(\frac{1}{x}\right) e^{-\theta x} dx = \theta^2 \int_0^\infty e^{-\theta x} dx$

improper use limits

$= \theta^2 \left(-\frac{1}{\theta} e^{-\theta x}\right)\Big|_0^\infty = \theta^2\left(\frac{1}{\theta}\right) = \theta$

$\Rightarrow \frac{1}{X}$ is unbiased

b. Is your estimator in a. consistent for θ? Explain how to show consistency and then show it or argue that your estimator isn't consistent. (4)

Probably NOT b/c only 1 observation

B/c unbiased, just need lim of variance to be 0.

Need $\lim_{n \to \infty} Var\left(\frac{1}{X}\right) = 0$

No dependence on n b/c only 1 obs, so need $Var(1/X) = 0$ and it's not going to be

$Var\left(\frac{1}{X}\right) = E\left(\left(\frac{1}{X}\right)^2\right) - \left[E\left(\frac{1}{X}\right)\right]^2$

$= E\left(\frac{1}{X^2}\right) - \theta^2$

$E\left(\frac{1}{X^2}\right) = \int_0^\infty \theta^2 x \left(\frac{1}{x^2}\right) e^{-\theta x} dx = \theta^2 \int_0^\infty \frac{e^{-\theta x}}{x} dx \Rightarrow DNE$

$\Rightarrow$ seems to indicate we don't have consistency with one obs.

$\frac{1}{X}$ actually has what is known as an inverse gamma dist, and the variance doesn't exist unless $\alpha > 2$, and here $\alpha = 2$.

3. (Data from Devore and Peck) Recent studies have looked at the duration of waiting time between a diagnosis and recommendation of surgery until the actual surgery. In the case of potentially life-threatening conditions, the waiting time on average is hopefully of short duration. For a (assume random) sample of 539 patients recommended for a heart bypass procedure, the sample mean waiting time was 19 days, and the sample standard deviation was 10 days (i.e. $\bar{x} = 19, s = 10, n = 539$). We will assume the distribution of waiting times is normal with unknown population mean and variance.

Z and t Critical values for 90%: z*=1.645, t*(200)=1.653, t*(369.5)=1.649, t*(538)=1.648, t*(739)=1.647 (You will not need all these). Chi-square table (provided).

a. Compute a 90% confidence interval for the population mean waiting time. (3)

$\bar{x} \pm z^* s/\sqrt{n}$ b/c $n = 539$ is large

$19 \pm 1.645 \frac{10}{\sqrt{539}} \Rightarrow 19 \pm .7086$

$(18.2914, 19.7086)$

b. Compute a 95% confidence interval for the population standard deviation of waiting time, assuming s=10, but changing the sample size to be 26. (3)

$n = 26 \quad n-1 = 25 \quad s = 10$

$$\left(\sqrt{\frac{(n-1)s^2}{\chi^2_u}}, \sqrt{\frac{(n-1)s^2}{\chi^2_L}}\right)$$

$$\Rightarrow \left(\frac{\sqrt{2500}}{\sqrt{40.6465}}, \frac{\sqrt{2500}}{\sqrt{13.1197}}\right) \Rightarrow (7.84257, 13.80410)$$

c. Compute a 95% lower confidence bound for the population standard deviation of waiting time, again assuming s=10 for a sample size of 26. (3)

$\chi^2_{.05} = \chi^2_u = 37.6525$

$$\left(\sqrt{\frac{(n-1)s^2}{\chi^2_u}}, \infty\right) \qquad (8.148414, \infty)$$

d. A pivot must have a distribution that ***does*** ***does not*** depend on the parameter of interest. (Circle one) (1)