MATH 322 Probability & Statistics
Spring 2011
Midterm I

1. Suppose $X_1, X_2, \ldots, X_n$ is a realization of a random sample from a continuous r.v. with the following density function:

$$f(x; \theta) = (\theta + 1)x^\theta \quad 0 \leq x \leq 1$$

Find the moment estimator of $\theta$. [5]

$$E(x) = \frac{1}{n} \sum_{i=1}^{n} X_i$$

$$\Rightarrow \frac{\theta + 1}{\theta + 2} = \bar{x}$$

$$\Rightarrow \theta + 1 = \frac{\theta \bar{x} + 2 \bar{x}}{1}$$

$$\Rightarrow \theta - \theta \bar{x} = 2 \bar{x} - 1$$

$$\Rightarrow \frac{\theta (1 - \bar{x}) = 2 \bar{x} - 1}{1 - \bar{x}}$$

2. Suppose $X_1, X_2, \ldots, X_n$ is a realization of a random sample from a continuous r.v. with the following density function:

$$f_\theta (x) = \frac{x}{\theta^2} e^{-\frac{x^2}{2\theta^2}} \quad \text{for} \quad x \geq 0$$

What is the likelihood function $L(\theta)$? What is the loglikelihood function $l(\theta)$? [5]

$$L(x_1, x_2, \ldots, x_n; \theta) = f(x_1, \theta) \cdot f(x_2, \theta) \cdot \cdots \cdot f(x_n, \theta)$$

$$= \frac{x_1}{\theta^2} e^{-\frac{x_1^2}{2\theta^2}} \times \frac{x_2}{\theta^2} e^{-\frac{x_2^2}{2\theta^2}} \times \cdots \times \frac{x_n}{\theta^2} e^{-\frac{x_n^2}{2\theta^2}}$$

$$= \frac{\prod_{i=1}^{n} x_i}{\theta^{2n}} e^{-\frac{1}{2\theta^2} \sum_{i=1}^{n} x_i^2}$$

$$l(\theta, L) = \ln \left( \frac{\prod_{i=1}^{n} x_i}{\theta^{2n}} \right) + \ln \left( e^{-\frac{1}{2\theta^2} \sum_{i=1}^{n} x_i^2} \right)$$

$$= \sum_{i=1}^{n} \ln(x_i) - 2n \ln(\theta) - \frac{1}{2\theta^2} \sum_{i=1}^{n} x_i^2$$

$$= \sum_{i=1}^{n} \ln(x_i) - 2n \ln(\theta) - \frac{1}{2\theta^2} \sum_{i=1}^{n} x_i^2$$
For exponential distribution

\[ P(X \leq x) = 1 - e^{-\lambda x} \]

3. Suppose \(X_1, X_2, X_3\) are i.i.d. \(\text{Exp}(\lambda)\), and that we observe the realizations \(X_1 = 1.0, X_2 = 2.0,\) and \(X_3 = 3.0\). What is the maximum likelihood estimate of \(\Pr(X_1 > 2)\)? [10]

The likelihood function is

\[
L(x_1, x_2 \ldots x_n ; \theta) = f(x_1, \theta) \times f(x_2, \theta) \ldots \times f(x_n, \theta)
\]

\[
= \lambda e^{-\lambda x_1} \times \lambda e^{-\lambda x_2} \ldots \lambda e^{-\lambda x_n}
\]

\[
= \lambda^n e^{-\lambda \sum_{i=1}^{n} x_i}
\]

The log-likelihood function

\[\ln(L) = n \ln(\lambda) - \lambda \sum_{i=1}^{n} x_i\]

\[
\frac{d \ln(L)}{d \lambda} = 0 \Rightarrow \frac{n}{\lambda} = \sum_{i=1}^{n} x_i
\]

\[
\Rightarrow \lambda = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{3}{6} = \frac{1}{2}
\]

\[\Pr(X_1 > 2) = \int_{2}^{\infty} \lambda e^{-\lambda x} dx = e^{-\frac{1}{2} \times 2} = e^{-1}\]

4. A sample of 20 patients at a doctor's office reveals an average waiting time of 16 minutes, and a standard deviation of 5 minutes.

(a) What is the point estimate of the average waiting time for all patients in this particular doctor's office? [2]

\[\hat{\mu} = \bar{x} = 16\]

(b) Calculate the (estimated) standard error of this point estimate, and explain what this number means. [5]

\[\text{S.E.} = s \cdot d(\bar{x}) = \frac{s}{\sqrt{n}} = \frac{s}{\sqrt{n}} (\text{estimated}) = \frac{s}{\sqrt{20}} = 1.118\]

We can expect a sample average to differ from the population average by about 1.12 minutes.

(c) Explain why we cannot construct a confidence interval for the average waiting time for all patients using the \(Z\) distribution even if we could assume that the distribution of waiting time is normal. [2]

Sample size is small \((n < 30)\).
5. A sanitation department is interested in estimating the mean amount of garbage per bin for all bins in the city. In a random sample of 36 bins, the sample mean amount was 51.5 pounds and the population standard deviation was 4 pounds. Construct 99% confidence intervals for \( \mu \). [5]

\[
\bar{x} \pm z_{\alpha/2} \frac{s}{\sqrt{n}} \quad \alpha = 1\% \quad \Rightarrow z_{0.005} = 2.576
\]

\[
51.5 \pm 2.576 \times \frac{4}{\sqrt{36}}
\]

\[
(49.783, 53.217)
\]

6. As part of his class project, a Statistics student took a random sample of 50 College students and recorded how many hours a week they spent on the Internet. The sample reveals an average of 6.9 hrs. Calculate the 90% Confidence Interval for average Internet usage among college students. Assume that the standard deviation of Internet usage for college students is known to be \( \sigma = 4.5 \) hrs/week. [5]

\[
6.9 \pm 1.645 \times \frac{4.5}{\sqrt{50}}
\]

\[
(5.853, 7.947)
\]

7. What percentage of college students have made at least one online purchase in the last three months? To answer this question, a market researcher surveyed 200 college students. Of those surveyed, 76 said that they had made at least one online purchase. Calculate the corresponding point estimate and its standard error. [5]

\[
\hat{p} = \frac{x}{n} = \frac{76}{200} = 0.38
\]

\[
SE = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \sqrt{\frac{0.38(1-0.38)}{200}} = 0.0343
\]

8. A web based software company is interested in estimating the proportion of individuals who use the Firefox browser. In a sample of 200 of individuals, 31 users stated that they used Firefox.

(a). Using this data, construct a 95% confidence interval for the proportion of all individuals that use Firefox. [5]

\[
\hat{p} = \frac{x}{n} = \frac{31}{200} = 0.155
\]

\[
\alpha = 5\% \quad \Rightarrow z_{0.025} = 1.96
\]

\[
\frac{31}{200} \pm 1.96 \times \sqrt{\frac{0.155(1-0.155)}{200}}
\]

\[
(0.1048, 0.2052)
\]
(b). What sample size would be required so that the width of the 95% confidence interval would be at most 0.08 units wide? [5]

\[ n = 315 \]

\[ n = \left\lfloor \frac{2 \cdot Z_{\alpha/2} \times \sqrt{\frac{\beta (1-\beta)}{\bar{y}^2}}}{w} \right\rfloor \]

\[ = 4 \times Z_{\alpha/2} \times \frac{\beta (1-\beta)}{\bar{y}^2} = 4 \times (1.96)^2 \frac{(0.155)(0.845)}{(0.08)^2} = 315 \]

9. The amount of lateral expansion (mils) was determined for a sample of \( n = 9 \) pulsed-power gas metal arc welds used in LNG ship containment tanks. The resulting sample standard deviation was \( s = 2.80 \) mils. Assuming normality, derive a 95% CI for \( \sigma^2 \). [5]

\[ (\frac{(n-1)s^2}{\chi^2_{1-\alpha/2, n-1}}, \frac{(n-1)s^2}{\chi^2_{\alpha/2, n-1}}) \]

\[ (3.577, 28.7706) \]

10. A sample of 14 joint specimens of a particular type gave a sample mean proportional limit stress of 8.50 MPa and a sample standard deviation of 0.80 MPa. Calculate and interpret a 95% prediction interval for the proportional limit stress of a single joint of this type. [5]

\[ \bar{x} = t_{\alpha/2, n-1} \times s \sqrt{1 + \frac{1}{n}} \]

\[ 8.5 \pm 2.160 \times (0.8) \sqrt{1 + \frac{1}{14}} \]

\[ (6.711, 10.289) \]