STA4322 Exam2 (closed book, 9:30 - 10:45, 7/24/2009) Name_____

1. Let $Y \sim exp(\theta)$, i.e., the density function of Y is $f(y) = \frac{1}{\theta}e^{-y/\theta}$, $y \geq 0$.

   Suppose that θ is unknown. (i) Show that $X = Y/\theta$ has a distribution independent of θ. (5%)

   Sol: Let $x = y/\theta$. Then $dx = dy/\theta$ and the density for X is $f(x) = e^{-x}$ for $x > 0$ independent of θ.

   (ii) Suppose we have only 1 observation Y. Construction a 95% confidence interval for θ based on Y. (10%)

   Sol: We wish to find a and b such that $P(a < \frac{Y}{\theta} < b) = 0.95$. By (i) we know that $\frac{Y}{\theta}$ is an exponential distribution with mean 1. Thus we need

   $$p(X > b) = \int_b^\infty e^{-t}dt = e^{-b} = 0.025.$$

   Or $b = -\ln(0.025) = 3.69$. Similarly, $a = -\ln(0.975) = 0.025$. Thus a 95% C.I. is

   $$\frac{Y}{3.69} \leq \theta \leq \frac{Y}{0.025}$$

   (iii) Suppose we have n observations y_1, y_2, ..., y_n. Find the maximum likelihood estimate (MLE) for θ. (10%)

   Sol: The likelihood is $l(y_1, ..., y_n) = \frac{1}{\theta^n}e^{-\sum y_i/\theta}$. Taking log, we have

   $$\ln L = -n \ln \theta - \sum y_i/\theta.$$

   By differentiating it with respect to θ and set the derivative to 0, we have the MLE $\hat{\theta} = \bar{y}$.

   (iv) Is the MLE the MVUE for θ? Justify your answer. (5%)

   Sol: Yes, because from the likelihood equation, we see that $\sum y_i$ is sufficient for θ. Moreover, from the gamma distribution, $E[y_1] = \theta \Rightarrow \bar{y}$ is unbiased.

2. (i) Two random samples both of 200 persons were taken from cities A and B. The question was whether they were happy with their living conditions. The number of yes answers from city A was 132 and that from city B was 120. Use 95% confidence interval to judge whether we can say that the residents in city A is more satisfied with their city. (15%)

Sol: $\hat{p}_1 = 132/200 = 0.66$ and $\hat{p}_2 = 120/200 = 0.60$. From table 8.1, we know the 95% C.I. for $p_1 - p_2$ is

$$0.66 - 0.60 \pm 1.96\sqrt{\frac{0.66 \times 0.34}{200} + \frac{0.6 \times 0.4}{200}}$$

It becomes 0.06 ± 0.094. Since the interval covers 0, we better not make the conclusion that the residents in city A is happier.

(ii) In (i), we are actually estimating the difference between two proportions p_1 and p_2. Suppose we wish to estimate this difference $p_1 - p_2$ with standard error 0.01 with equal sample size from both populations. What should be the sample size. (Note: standard error is defined in Table 8.1.) (10%)

Sol: We wish

$$\sqrt{\frac{p_1 q_1}{n} + \frac{p_2 q_2}{n}} = 0.01.$$

Since at the design stage the p's are unknown, we take the conservative approach by letting them to be 0.5. That means

$$\sqrt{\frac{0.25}{n} + \frac{0.25}{n}} = 0.01 \;\Rightarrow\; n = 5000.$$

3. The IQ score is normally distributed. A psychologist is interested in estimating the average IQ from children with malnutrition. Twenty randomly picked children were tested. The average score was 94 with sample standard deviation 20.4. Find a 95% confidence interval for the true mean IQ of this type of children. We consider 20 as a small sample. (15%)

Sol: from the t formula, a 95% C.I. for μ is

$$94 \pm t_{19,0.025}\frac{20.4}{\sqrt{20}} \;\Rightarrow\; 94 \pm 9.55.$$

by $t_{19,0.025} = 2.093$.

4. Let Y_1, Y_2, ..., Y_n be a random sample from a normal distribution with unknown mean μ and known variance 1. Show that the MVUE of $\theta = \mu^2$ is $(\bar{Y})^2 - 1/n$. (10%)

Sol:

$$\begin{aligned} l(y_1, ..., y_n) &= (\frac{1}{2\pi})^{n/2} exp(-\frac{1}{2}\sum(y_i - \mu)^2) \\ &= (\frac{1}{2\pi})^{n/2} exp(-\frac{1}{2}[\sum(y_i - \bar{y})^2 + n(\bar{y} - \mu)^2]) \\ &\Rightarrow \bar{y} \text{ is sufficient for } \mu. \end{aligned}$$

We need only to show that $\mathrm{E}[(\bar{Y})^2 - 1/n] = \mu^2$. This can be seen from

$$E[(\bar{Y})^2] = [mean(\bar{y})]^2 + Var(\bar{y}) = \mu + \frac{1}{n}.$$

5. Let Y_1, Y_2, ..., Y_n be a random sample from the uniform distribution on the interval $(\theta, \theta+1)$. We wish to use $Y_{(n)} = \max(Y_1, Y_2, ..., Y_n)$ to construct an estimator for θ. (i) Find $\mathrm{E}[Y_{(n)}]$ and from here to construct an unbiased estimator for θ. (15%)

   Sol: Note that $P(Y_1 < x) = \int_\theta^x dt = x - \theta$. Let $X = Y_{(n)}$. Thus,

$$P(X \le x) = P(\max(Y_1, Y_2, ..., Y_n) \le x) = [p(Y_1 < x)]^n = (x-\theta)^n.$$

   The density function of X is then

$$f(x) = n(x-\theta)^{n-1},$$

   and

$$E[X] = \int_\theta^{\theta+1} nx(x-\theta)^{n-1}dx = \int_0^1 n(z+\theta)z^{n-1}dz = \theta + \frac{n}{n+1}.$$

   Thus an unbiased estimator based on $Y_{(n)}$ is

$$\hat{\theta} = Y_{(n)} - \frac{n}{n+1}.$$

   (ii) Is $Y_{(n)}$ sufficient for θ? Justify your answer. (5%)

   Sol: No. Apparently $Y_{(1)} = \min(Y_1, Y_2, ..., Y_n)$ also contains information of θ.