36 - 226 Introduction to Statistical Inference

Homework assignment 2

Due: Wednesday, January 30, 2013

- Write your full name, the course number, and the homework number at the top of each page.
- **STAPLE** your entire assignment together with a staple.
- Write clearly. Electronic submission of homework assignments is not accepted.

1. Wackerly 4.77. The SAT and ACT college entrance exams are taken by thousands of students each year. The mathematics portions of each of these exams produce scores that are approximately normally distributed. In recent years, SAT mathematics exam scores have averaged 480 with standard deviation 100. The average and standard deviation for ACT mathematics scores are 18 and 6, respectively. *Hint: Review Section 4.5.*

   (a) An engineering school sets 550 as the minimum SAT math score for new students. What percentage of studens will score below 550 in a typical year?

   (b) What score should the engineering school set as a comparable standard on the ACT math test?

2. Using Wackerly 5.154 (the questions are different): In a clinical study of a new drug formulated to reduce the effects of rheumatoid arthritis, researchers found that the proportion p of patients who respond favorably to the drug is a random variable that varies from batch to batch of the drug. Assume that p has a probability density function given by

$$f(p) = \begin{cases} 12p^2(1-p), & 0 \leq p \leq 1, \\ 0, & \text{elsewhere} \end{cases}$$

   Suppose that n patients are injected with portions of the drug taken from the same batch. Let Y denote the number showing a favorable response. Note that the conditional distribution of Y given p, i.e., $P(Y = y \mid p)$ is Binomial(n, p). In the problem you are given the marginal distribution of p, i.e., $f(p)$. Please do the following: *Hint: see lecture notes corresponding to the Chapter 5 review.*

   (a) Find the joint probability density function of Y and p, that is $f(y, p)$.

   (b) Next, please find the marginal probability distribution of Y, $f(y)$ (i.e. the unconditional probability distribution of Y).

   (c) Using Bayes' theorem, find the conditional distribution of $f(p \mid y)$ (this is know as the *posterior distribution*). Without doing a derivation, state the mean and variance of the posterior distribution of p (i.e. use the fact that $f(p \mid y)$ belongs to a recognizble family).

3. 5.97 (Wackerly). The random variables Y_1 and Y_2 are such that $E(Y_1) = 4$, $E(Y_2) = -1$, $V(Y_1) = 2$, and $V(Y_2) = 8$.

   (a) What is $Cov(Y_1, Y_1)$?

   (b) Assuming that the means and variances are correct, as given, is it possible that $Cov(Y_1, Y_2) = 7$? *Hint: If $Cov(Y_1, Y_2) = 7$, what is the value of ρ, the coefficient of correlation?*

(c) Assuming that the means and variances are correct, what is the largest possible value for $Cov(Y_1, Y_2)$? If $Cov(Y_1, Y_2)$ achieves this largest value, what does that imply about the relationship between Y_1 and Y_2?

(d) Assuming that the means and variances are correct, what is the smallest possible value for $Cov(Y_1, Y_2)$? If $Cov(Y_1, Y_2)$ achieves this largest value, what does that imply about the relationship between Y_1 and Y_2?

4. 6.7 Wackerly. Suppose Z has a standard normal distribution.

   (a) Find the density function of $U = Z^2$.

   (b) Does U have a gamma distribution? What are the values of α and β?

   (c) What is another name for the distribution of U?

5. Let $Y_1, \ldots, Y_{10}$ be independent and identically distributed (iid) random variables such that for $0 < p < 1$, $P(Y_i = 0) = 1 - p$ and $P(Y_1 = 1) = p$ (i.e. Bernoulli random variables). Let $W = Y_1 + Y_2 + \cdots + Y_{10}$.

   (a) Find the moment-generating function of W, $m_W(t)$.

   (b) What is the distribution of W? (That is, what is the name of the distribution?)

   (c) Using the moment-generating funtion of W, find $V(W)$.

6. 6.74 Wackerly. Let $Y_1, \ldots, Y_n$ be independent, uniformly distributed random variables on the interval $[0, \theta]$. Find the

   (a) probability distribution function of $Y_{(n)} = \max(Y_1, Y_2, \ldots, Y_n)$.

   (b) density function of $Y_{(n)}$.

   (c) mean and variance of $Y_{(n)}$.

7. 8.8 Wackerly (Yes, you already know how to do problems in Chapter 8! Read Sections 8.1 and 8.2 to introduce some definitions). Suppose Y_1, Y_2, and Y_3 denote a random sample from an exponential distribution with density function

$$f(y) = \begin{cases} \left(\frac{1}{\theta}\right) e^{-y/\theta}, & y > 0, \\ 0, & \text{elsewhere} \end{cases}$$

   Consider the following five estimators (i.e. functions of the random variables) of θ:

$$\hat{\theta}_1 = Y_1, \quad \hat{\theta}_2 = \frac{Y_1 + Y_2}{2}, \quad \hat{\theta}_3 = \frac{Y_1 + 2Y_2}{2}, \quad \hat{\theta}_4 = \min(Y_1, Y_2, Y_3), \quad \hat{\theta}_5 = \bar{Y}$$

   (a) Which estimators are unbiased? (That is, for which estimators is it the case that $E(\hat{\theta}_i) = \theta$? If the expected value of the estimator is equal to θ, we call it an unbiased estimator.)

   (b) Among the unbiased estimators, which has the smallest variance? (Find the variance of the *unbiased* estimators above, and determine which is the smallest.)