STA 261s2006 Assignment 3

Do this assignment in preparation for Test One, which will be in class on Wednesday Jan 25th. The questions are practice for the test, and are not to be handed in.

1. Read Section 8.1; do Exercises 8.2 and 8.3.

2. Let $X_1, \ldots, X_n$ be a random sample from a distribution with expected value μ and variance σ^2. The Law of Large Numbers says $\overline{X}_n \stackrel{P}{\to} \mu$. Prove that $\overline{X}_n + a \stackrel{P}{\to} \mu + a$, where a is a constant.

3. As before, let $X_1, \ldots, X_n$ be a random sample from a distribution with expected value μ and variance σ^2. Prove that $3\overline{X}_n \stackrel{P}{\to} 3\mu$. Hint: Let $W_n = 3\overline{X}_n$, and use Chebyshev's inequality (see the new formula sheet).

4. Let $X_1, \ldots, X_n$ be a random sample from a distribution with expected value μ_x and variance σ_x^2. Independently of $X_1, \ldots, X_n$, let $Y_1, \ldots, Y_n$ be a random sample from a distribution with expected value μ_y and variance σ_y^2. This is a model for taking independent random samples from two different populations (like Males and Females). For this setup, the *parameters* are μ_x, σ_x^2, μ_y and σ_y^2. When you are asked for expected values and variances below, please express your answer in terms of the parameters. It is always possible.

   (a) How do you know $\overline{X}_n$ and $\overline{Y}_n$ are independent?

   (b) i. What is $E(\overline{X}_n + \overline{Y}_n)$?

   ii. What is $Var(\overline{X}_n + \overline{Y}_n)$?

   iii. Prove that $\overline{X}_n + \overline{Y}_n \stackrel{P}{\to} \mu_x + \mu_y$. Hint: Let $W_n = \overline{X}_n + \overline{Y}_n$, and use Chebyshev's inequality.

   (c) Let $W_n = \overline{X}_n \overline{Y}_n$.

   i. What is $E(W_n)$?

   ii. What is $E(W_n^2)$?

   iii. What is $Var(W_n)$?

   iv. Prove that $\overline{X}_n \overline{Y}_n \stackrel{P}{\to} \mu_x \mu_y$.

5. Let $f_{X_n}(x) = \frac{1}{4}I(x=0) + \frac{1}{2}I(x=1) + \frac{1}{4}I(x = \frac{n+1}{n})$.

(a) Is X_n discrete, or is it continuous?

(b) What is $F_{X_n}(x)$? You may write the answer as a case function, or you may write it using indicators.

(c) Let $g(x) = \lim_{n\to\infty} f_{X_n}(x)$. Consider the cases $x = 0$, $x = 1$ and x equals something else separately. Is $g(x)$ a probability distribution?

(d) Let $G(x) = \lim_{n\to\infty} F_{X_n}(x)$. Your answer should apply to all x. Is $G(x)$ a cumulative distribution function?

(e) Let X be a Bernoulli random variable with $\theta = \frac{3}{4}$. Denote the cumulative distribution function of X by $F_X(x)$. At what points is $F_X(x)$ discontinuous? Does $G(x)$ equal $F_X(x)$ except possibly at those points? Does this mean $\overline{X}_n \stackrel{d}{\to} X$? (Check the definition.)

(f) Do we have $\lim_{n\to\infty} E(X_n) = E(X)$? Answer Yes or No. Show your work.

(g) Do we have $\lim_{n\to\infty} Var(X_n) = Var(X)$? Answer Yes or No. Show your work.

6. Let $f_{X_n}(x) = \frac{1}{3}I(x=0) + \frac{2}{3}\left(\frac{n-1}{n}\right)I(x=1) + \frac{1}{3n}I(x=n)$.

(a) Why would you *never* integrate $f_{X_n}(x)$?

(b) What is $F_{X_n}(x)$? You may write the answer as a case function, or you may write it using indicators.

(c) Let $f(x) = \lim_{n\to\infty} f_{X_n}(x)$. Consider the cases $x = 0$, $x = 1$ and x equals something else separately. Is $f(x)$ a probability distribution?

(d) Let $F(x) = \lim_{n\to\infty} F_{X_n}(x)$. Your answer should apply to all x.

i. What is $F(x)$? You may write the answer as a case function, or you may write it using indicators.

ii. Is $F(x)$ a cumulative distribution function? Does it correspond to $f(x)$?

iii. We seem to have $\overline{X}_n \stackrel{d}{\to} X$, where X is Bernoulli again. What is the parameter θ?

(e) Do we have $\lim_{n\to\infty} E(X_n) = E(X)$? Answer Yes or No. Show your work.

(f) Do we have $\lim_{n\to\infty} Var(X_n) = Var(X)$? Answer Yes or No. Show your work.

7. Read Section 8.2. The statement of the Law of Large Numbers in Theorem 8.2 is equivalent to what was given in class, by taking complements. You are responsible only for applications of the Central Limit Theorem (Theorem 8.3) and not for the proof. Do Exercises 8.63b, 8.66 (I get $P(Z > 1.875) \approx .0304$, interpolating in the table. You will never have to interpolate on a test or exam.), 8.67. 8.69, 8.73b.

8. Suppose the population mean number of rats in a variety store is a Poisson random variable with $\lambda = 5$.

   (a) For a single randomly selected store, what is the probability that the number of rats is less than 4?

   (b) For a random sample of eighty stores, what is the probability that the sample mean number of rats is less than 4?

9. A manufacturer of automobile batteries claims that its best battery has a mean life length of 54 months, with a standard deviation of 6 months. The exact shape of the distribution of battery life lengths is not reported. A consumer group decides to buy 49 of these batteries and test them. Assume that the 49 batteries represent a random sample from the entire population of car batteries of this type.

   (a) What is the probability that a *single* randomly chosen battery will last less than 52 months? Hint: This question is impossible to answer. Why?

   (b) What is the probability that mean life length of the 49 batteries is less than 52 months?

10. In the population of Toronto kitchen sinks, the mean concentration of lead in the tap water after running the water for one minute is .657 micrograms per liter, with a standard deviation of .22. The exact shape of the distribution of lead concentration measurements is unknown. Answer these questions if possible. If it is not possible to answer, say why.

    (a) What is the probability that the measurement from a *single* randomly chosen sink will be above .7 grams of lead per milliliter?

    (b) A random sample of nine sinks is tested. What is the probability that the sample mean is above .7 grams of lead per milliliter?

11. A standardized (multiple choice) test of academic achievement is normally distributed with mean μ=50 and standard deviation σ=10, for the population of North American university students. Answer these questions if possible. If it is not possible to answer, say why.

    (a) What is the probability that the score of a *single* randomly chosen student will be at least 45 (that is 45 or more)?

    (b) A random sample of sixteen North American university students were given the test. What is the probability that their mean score will be at least 45 (that is 45 or more)?

12. In a survey of smoking habits reported in the *Journal of the Canadian Medical Association*, a random sample of 1250 current smokers was asked "On the average, how many cigarettes do you now smoke a day?" The mean response was $\overline{x} = 21.8$, with a sample standard deviation of $s = 5.2$. A representative of the tobacco industry says that this figure is in keeping with a long-standing trend that average daily consumption is no more than one pack (20 cigarettes) a day, because the difference between 20 and 21.8 is tiny and could have arisen by chance.

    Evaluate this claim by answering the following question. If the true population mean is really $\mu = 20$, what is the probability of getting a sample mean as large as $\overline{x} = 21.8$ or larger just by chance? Hint: It is okay to use s in place of σ in the Central Limit Theorem.

    What do you conclude? Is the statement by the Industry representative a reasonable one, in light of your calculation?

13. A machine produces bolts, of which ten percent are defective. Find the probability that in a random sample of 400 bolts, 55 or more will be defective. Hint: Think of coding the data as $X_1, \ldots, X_{400}$, with $X_i = 1$ if bolt i is defective, and $X_i = 0$ if it is not defective. Translate the question into a question about the sample mean.

14. The amount of corn chips dispensed into a 24-ounce bag by a dispensing machine has been identified as possessing a normal distribution with a mean of 24.5 ounces and a standard deviation of 1 ounce. Suppose 16 bags of chips were randomly selected from this dispensing machine. Find the probability that the sample mean weight of these 400 bags exceeded 24.6 ounces. If the question is impossible to answer, say why.