Name: ______________________________

Practice Second Midterm Exam

Statistics 1000
Spring 2007
Dr. Nancy Pfenning

This is a closed book exam worth 200 points. You are allowed to bring needed tables, a calculator, and a two-sided sheet of notes. There are 12 problems, with point values as shown. If you want to receive partial credit for wrong answers, show your work. Don't spend too much time on any one problem.

1. (30 pts.) In a group of about 400 students, the proportion with pierced ears is .65. Pick 30 students at random; we are interested in the probability that the number X with pierced ears is less than 5.

   (a) Which of these numbers tells the sample size n? (i) 400 (ii) .65 (iii) 30 (iv) 5

   (b) We verify that the population is at least ten times the sample size so we can say that

      i. X is approximately normal
      ii. X is approximately binomial
      iii. X has only two possible categories, success or failure

   (c) We verify that $np \geq 10$ and $n(1-p) \geq 10$ so we can say that

      i. X is approximately normal
      ii. X is approximately binomial
      iii. X has only two possible categories, success or failure

   (d) Find the mean ______and standard deviation ______of X.

   (e) Use a normal approximation to find the probability that fewer than 5 of those sampled students have pierced ears.

   (f) Finding fewer than 5 with pierced ears in the sample is
   (i) virtually impossible (ii) very unusual (iii) fairly common (iv) almost guaranteed

   (g) Find the mean ______and standard deviation ______of sample proportion $\hat{p}$.

2. (15 pts.) In the year 2000, combined SAT scores in the U.S. were approximately normal with mean 1020 and standard deviation 210.

   (a) Roughly 95% of students scored between ______and ______.

   (b) The NCAA required Division I athletes to score at least 820 on the combined SAT in order to compete in their first college year. What proportion of all students scored 820 or better?

   (c) The top 3% scored above how many points?

3. (15 pts.) The number rolled on a die has mean 3.5 and standard deviation 1.7.

   (a) What are the mean ______and standard deviation ______for sample mean roll of 25 dice?

   (b) 50 students each roll 25 dice and calculate the sample mean roll. If each student uses sample mean roll to construct a 90% confidence interval for the mean roll of all dice, about how many of their intervals should contain the true mean?
   (i) 0 (ii) 5 (iii) 10 (iv) 15 (v) 20 (vi) 25 (vii) 30 (viii) 35 (ix) 40 (x) 45 (xi) 50

   (c) If each student in a class of 50 uses sample mean roll to test the true null hypothesis that the mean roll of all dice is 3.5, about how many should reject at the $\alpha = .10$ level?
   (i) 0 (ii) 5 (iii) 10 (iv) 15 (v) 20 (vi) 25 (vii) 30 (viii) 35 (ix) 40 (x) 45 (xi) 50

4. (30 pts.) Suppose a personnel manager in charge of hiring has been going under the assumption that 20% of all applicants for executive positions lie about having a degree. The company ran a background check and found that of 89 applicants for such positions, 16 lied about having a degree.

   (a) What is the sample proportion who lied? (Round to two decimal places.)

   (b) Formulate null and alternative hypotheses to test if the data are consistent with the personnel manager's assumption.

   (c) Calculate the test statistic.

   (d) Find the p-value.

   (e) Is there convincing evidence against the manager's claim?
   (i) yes (ii) no (iii) results are borderline

   (f) Construct a 95% confidence interval for the proportion of all applicants for executive positions who lie about having a degree.

   (g) Suppose the manager wants to estimate the proportion of all applicants who lie about having a degree with a conservative margin of error of .10. How many applicants should be sampled?

5. (20 pts.) In a sample of 16 college students, the mean hours of sleep reported for the night before was 7.1 and the standard deviation was 1.56.

   (a) Give a 99% confidence interval for mean hours slept by all college students.

   (b) Presumably, adults in the U.S. average 7.0 hours of sleep a night. Does our confidence interval suggest that college students' overall mean may also be 7.0?

      i. yes
      ii. no; it's almost surely lower than 7.0
      iii. no; it's almost surely higher than 7.0

   (c) To interpret your interval in (a), circle one of the following:

      i. The sample mean has a 99% probability of falling in this interval.
      ii. The hours slept by 99% of all students was somewhere in this interval.
      iii. The hours slept by 99% of sampled students was somewhere in this interval.
      iv. The population mean has a 99% probability of falling in this interval.
      v. The interval has a 99% probability of containing the population mean.

   (d) Circle any of the following which would result in a more precise interval (you may circle anywhere from none to all three):

      i. reduce our level of confidence
      ii. the population standard deviation is larger
      iii. the sample size is larger

6. (5 pts.) I tested $H_0 : \mu = 64.5$ vs. $H_a : \mu > 64.5$ about the mean height μ of all female college students, based on my sample of 230 female students' heights. Their sample mean height turned out to be 64.9 inches, and the P-value was .004. To give a correct interpretation of this P-value, circle one of the following:

   (a) If the mean height of all female college students is **equal to** 64.5, then .004 is the probability that a random sample of 230 female students would have a mean height **equal to** 64.9.

   (b) If the mean height of all female college students is **equal to** 64.5, then .004 is the probability that a random sample of 230 female students would have a mean height **greater than or equal to** 64.9.

   (c) If the mean height of all female college students is **greater than** 64.5, then .004 is the probability that a random sample of 230 female students would have a mean height **equal to** 64.9.

   (d) If the mean height of all female college students is **greater than** 64.5, then .004 is the probability that a random sample of 230 female students would have a mean height **greater than or equal to** 64.9.

7. (35 pts.) In a sample of 36 Stats students, the mean math SAT score was 611, with standard deviation 112. Is this significantly **different** from the mean for all Pitt students, which was 571?

   (a) State the null and alternative hypotheses.

   (b) Calculate the test statistic.

   (c) How many degrees of freedom should be used in Table A.2? (i) 30 (ii) 40 (iii) infinite

   (d) Give a range for the P-value.

   (e) Test at the $\alpha = .05$ level and draw your conclusions by circling one of the following:

      i. The P-value is small; the mean math SAT for Stat students may well be 571.

      ii. The P-value is small; we have evidence that mean math SAT for Stat students differs from 571.

      iii. The P-value is **not** small; the mean math SAT for Stat students may well be 571.

      iv. The P-value is **not** small; we have evidence that mean math SAT for Stat students differs from 571.

      v. Results are borderline.

   (f) Would your conclusions still be valid if the distribution of SAT scores were skewed? (Choose one:)

      i. Yes, because the distribution of sample mean is always normal.

      ii. Yes, because the sample size is large.

      iii. No, because the distribution of sample mean could also be skewed for a moderate sample size.

      iv. No, because distribution of sample mean would also be skewed, regardless of sample size.

   (g) What would be a range for the P-value if I had suspected in advance that Stats students' mean math SAT score is **higher** than 571? between ______and ______

8. (10 pts.) The mean and standard deviation for the number rolled on a die are 3.5 and 1.7. When I rolled 8 dice, I got a mean of 4.25 and a standard deviation of 1.56. Complete the following:

   (a) 3.5 is denoted (i) μ (ii) $\bar{x}$ (iii) σ (iv) s (v) p (vi) $\hat{p}$; it is a **statistic/parameter**

   (b) 1.7 is denoted (i) μ (ii) $\bar{x}$ (iii) σ (iv) s (v) p (vi) $\hat{p}$; it is a **statistic/parameter**

   (c) 4.25 is denoted (i) μ (ii) $\bar{x}$ (iii) σ (iv) s (v) p (vi) $\hat{p}$; it is a **statistic/parameter**

   (d) 1.56 is denoted (i) μ (ii) $\bar{x}$ (iii) σ (iv) s (v) p (vi) $\hat{p}$; it is a **statistic/parameter**

9. (15 pts.) Ten pilots performed tasks at a simulated altitude of 25,000 feet, in sober condition and again three days later after drinking alcohol. Seconds spent on useful performance was recorded for each pilot under both conditions.

	N	Mean	StDev	SE Mean
no alcohol	10	546.6	238.8	75.5
alcohol	10	351.0	210.9	66.7
Difference	10	195.6	230.5	72.9

```
95% lower bound for mean difference: 62.0
T-Test of mean difference = 0 (vs > 0): T-Value = 2.68  P-Value = 0.013
```

(a) Was this a matched pairs study? (i) yes (ii) no (iii) there is no way to tell

(b) Using $\alpha = .05$, was the mean difference in performance times significantly higher than zero? Circle the best answer below, and circle just **one** number in the output that lets you decide.

i. Yes, definitely.
ii. No, not at all.
iii. No, not quite.

(c) A possible flaw in the study design is that all pilots performed first without, then with alcohol. How could this be improved?

i. Flip a coin for each pilot to determine the order in which the conditions are tested.
ii. Have all the pilots test first with alcohol, rather than without.
iii. Ask the pilots who wants to be tested first with alcohol; the first five who agree will test in that order, and the other five will test in the opposite order.

10. (10 pts.) Tell what type of variables are involved in the following examples:

(a) A personnel manager gathers data about whether applicants lied about having a degree. (i) one quantitative variable (ii) one categorical variable (iii) one quantitative and one categorical variable (iv) two quantitative variables (v) two categorical variables

(b) College students are surveyed to see if they average more than 7 hours of sleep in a night. (i) one quantitative variable (ii) one categorical variable (iii) one quantitative and one categorical variable (iv) two quantitative variables (v) two categorical variables

11. (10 pts.) Note that a Type I error is rejecting H_0, even though it is true; a Type II error is failing to reject H_0, even though it is false.

   (a) Suppose H_0 states that a defendant is innocent. Since we wouldn't want an innocent man to be sent to jail, it's important to avoid a
   (i) Type I error (ii) Type II error

   (b) Suppose H_0 states that a person is free of HIV. Since we wouldn't want an infected person to be unaware and untreated, it's important to avoid a
   (i) Type I error (ii) Type II error

12. (5 pts.) Which one of the following is a two-sample design?

   (a) In order to compare number of yearly credits taken by freshmen vs. seniors, sample 100 each freshmen and seniors, and record how many credits they take this year.

   (b) In order to compare number of yearly credits taken by freshmen vs. seniors, sample 100 seniors and record how many credits they take this year, and how many they took in their freshman year.