Math 263, Section 6 **Name**____
Third Test Spring 2010
Total points: 125

For Questions 1-10, circle *one* answer. (4 points each)

1. An agricultural researcher plants 25 plots with a new variety of yellow corn. The yield per acre for the new variety of yellow corn follows a normal distribution with unknown mean, μ, and standard deviation $\sigma = 10$ bushels per acre. If the average yield for these 25 plots is $\bar{x} = 150$ bushels per acre, what is a 90% confidence interval for μ?
   (a) 150 ± 0.784
   (b) 150 ± 2.00
   (c) 150 ± 3.29
   (d) 150 ± 3.92
   (e) 150 ± 5.15

2. (Continuation of Problem 1): Which of the following would produce a confidence interval with a smaller margin of error than the 90% confidence interval? (*You can do this problem even if you can't do Problem* 1)
   (a) Plant only 5 plots rather than 25, because 5 are easier to manage and control.
   (b) Plant 10 plots rather than 25, because a smaller sample size will result in a smaller margin of error.
   (c) Plant 100 plots rather than 25, because a larger sample size will result in a smaller margin of error.
   (d) Compute a 99% confidence interval rather than a 90% confidence interval, because a higher confidence level will result in a smaller margin of error.

3. In the last mayoral election in a large city, 47% of the adults over the age of 65 voted Republican. A researcher wishes to determine if the proportion of adults over the age of 65 in the city who plan to vote Republican in the next mayoral election has changed. Let p represent the proportion of the population of all adults over the age of 65 in the city who plan to vote Republican in the next mayoral election. In terms of p, the researcher should test which of the following null and alternative hypotheses?
   (a) $H_0: p = 0.47$ versus $H_a: p < 0.47$
   (b) $H_0: p = 0.47$ versus $H_a: p \neq 0.47$
   (c) $H_0: p = 0.47$ versus $H_a: p > 0.47$
   (d) $H_0: p \neq 0.47$ versus $H_a: p < 0.47$
   (e) $H_0: p \neq 0.47$ versus $H_a: p = 0.47$
   (f) $H_0: p \neq 0.47$ versus $H_a: p > 0.47$

4. A simple random sample of 20 third-grade children is selected from a school district and each child is given a test to measure his/her reading ability. We want to calculate a 95% confidence interval for the population mean score. In the sample, the mean score is 64 points and the standard deviation is 12 points. What is the margin of error of the confidence interval?
   (a) 2.68 points
   (b) 4.64 points
   (c) 5.26 points
   (d) 5.62 points
   (e) 6.84 points

5. A report indicated that automobiles manufactured in North America are less fuel-efficient than automobiles manufactured in Asia. It is known that automobiles from Asia have a mean fuel efficiency of 22 miles per gallon. To determine if there is evidence to support this claim, a random sample of automobiles manufactured in North American is to be selected and their fuel efficiency determined. If μ is the mean fuel efficiency of North American automobiles, the appropriate hypotheses to be tested are

(a) H_0: $\mu = 22$ against H_a: $\mu > 22$.
(b) H_0: $\mu = 22$ against H_a: $\mu \neq 22$
(c) H_0: $\mu = 22$ against H_a: $\mu < 22$.
(d) H_0: $\bar{x} = 22$ against H_a: $\bar{x} > 22$.
(e) H_0: $\bar{x} = 22$ against H_a: $\bar{x} \neq 22$.
(f) H_0: $\bar{x} = 22$ against H_a: $\bar{x} < 22$.

For Questions 6-10, use the following information, and circle *one* answer. (4 points each)

Recent revenue shortfalls in a southwestern state led to a reduction in the state budget for higher education. To offset the reduction, the largest state university proposed a 25% tuition increase. It was determined that such an increase was needed simply to compensate for the lost support from the state. Random samples of 50 freshmen, 50 sophomores, 50 juniors, and 50 seniors from the university were asked whether or not they were strongly opposed to the increase, given that it was the minimum increase necessary to maintain the university's budget at current levels. The results are given in the following table:

	Year in school				
Opinion	Freshman	Sophomore	Junior	Senior	Total
Strongly opposed	39	36	29	18	122
Not strongly opposed	11	14	21	32	78
Total	50	50	50	50	200

6. To compare the four classes (the four years in school) with respect to their opinion regarding the tuition increase, which distribution should we calculate?
(a) The joint distribution of year in school and opinion.
(b) The marginal distribution of opinion.
(c) The conditional distribution of year in school, given opinion.
(d) The conditional distribution of opinion, given year in school.

7. Which hypotheses would be tested by a chi-square test based on the data in this table?

(a) H_0: The proportion of students strongly opposed is the same in all four classes.
H_a: The proportion of students strongly opposed is different in all four classes.

(b) H_0: The proportion of students strongly opposed is the same in all four classes.
H_a: The proportion of students strongly opposed is different in at least one class.

(c) H_0: The number of students strongly opposed is the same in all four classes.
H_a: The number of students strongly opposed is different in all four classes.

(d) H_0: The number of students strongly opposed is different in all four classes.
H_a: The number of students strongly opposed is the same in all four classes.

(e) H_0: The proportion of students strongly opposed is different in all four classes.
H_a: The proportion of students strongly opposed is the same in all four classes.

8. We test the null hypothesis that there is no association between year in school and opinion. Under the null hypothesis, what is the expected number of strongly opposed seniors?

(a) 18
(b) 19.5
(c) 25
(d) 30.5
(e) 50

9. What is the contribution to the chi-square statistic from the cell of strongly opposed seniors? (The contribution is the term that comes from this cell in the expression for chi-squared statistic.)

(a) 5.12
(b) 8.01
(c) 8.68
(d) 12.5
(e) 30.5

10. The chi-square statistic for these data equals 21.9. Are the data statistically significant?

(a) Yes at both the 5% and 1% significance levels.
(b) Yes at the 5% significance level but not at the 1% significance level.
(c) Yes at the 1% significance level but not at the 5% significance level.
(d) No at both the 5% and 1% significance levels.
(e) This cannot be determined from the information given.

11. (**22 points**) Medical researchers want to decide if a new drug has the side affect of raising patients' blood pressure above the usual level of 120 mmHg.

(a) What is the null hypothesis?
$H_{o:}$ Mean blood pressure of patients who take drug = 120 mm.

(b) What is the alternative hypothesis?
$H_{a:}$ Mean blood pressure of patients who take drug > 120 mm.

(c) What conclusion will researchers draw from a sufficiently high average blood pressure in a large random sample of patients taking the drug (*check one*):

__X___Reject the null hypothesis

_____ Reject the alternative hypothesis

_____ Do not reject the null hypothesis

_____ Do not reject the alternative hypothesis

(d) What conclusion will researchers draw from an average blood pressure close to 120 mmHg in a large random sample of patients taking the drug (*check one*):

_____ Reject the null hypothesis

_____ Reject the alternative hypothesis

__X __Do not reject the null hypothesis

_____ Do not reject the alternative hypothesis

(e) The P-value of the average blood pressure of patients in the sample is 0.015. This means (*check one or more*):

_____ The probability that the null hypothesis is true is 1.5%.

_____ There is a 1.5% chance that the alternative hypothesis is true

_____ The probability that the patients' average blood pressure is unaffected is 1.5%.

__X __If the drug does not raise blood pressure, there's a 1.5% chance of the random sample having an average as high as, or higher than, that observed.

_____ If the drug raises blood pressure, there's a 1.5% chance of the random sample having an average above 120 mmHg.

12. (**24 points**) A 2005 article reported on an anti-seizure drug, Valproate, given to bi-polar alcoholics.[1] Some participants were given the drug, some were given a placebo, and all were questioned six months later to see if they were drinking heavily; the results are in the table. Use the following steps to decide whether Valproate seems to have a significant effect on drinking in bi-polar alcoholics:

	Heavy drinking	No heavy drinking	Total
Valproate	14	18	32
Placebo	15	7	22
Total	29	25	54

This can be done as a 2-sample Z-test or using a chi-squared test.
Here's the solution for the Z-test, which can be one or two sided.
The chi-square solution is on the next page.

(a) What is the null hypothesis?

Let group1 be those who took Valproate and group 2 be those who took the placebo. We consider the proportions who remained heavy drinkers.
The null hypothesis is that the proportion that remain heavy drinkers are the same in both groups: $H_0: p_1 = p_2$

(b) What is the alternate hypothesis?

The alternate hypothesis is that the proportion that remains heavy drinkers is different in the group that took Valproate: $H_a: p_1 \neq p_2$
If we test that Valproate reduces drinking, then we have $H_a: p_1 < p_2$

(c) What is the test statistic?

The sample statistics are

$$\hat{p}_1 = \frac{14}{32} = 0.438 = 43.8\% \text{ and } \hat{p}_2 = \frac{15}{22} = 0.682 = 68.2\%$$

The standard error is

$$SE = \sqrt{\frac{0.438(1-0.438)}{32} + \frac{0.682(1-0.682)}{22}} = 0.1325$$

and the z-score is

$$z = \frac{0.438 - 0.682}{0.1325} = -1.844.$$

If we use the pooled proportion we have $\hat{p} = \frac{14+15}{32+22} = 0.5370 = 53.7\%$. Then

$$SE = \sqrt{0.537(1-0.537)\left(\frac{1}{32} + \frac{1}{22}\right)} = 0.1381,$$

and the z-score is

$$z = \frac{0.438 - 0.682}{0.1381} = -1.769.$$

(d) What is the distribution of the test statistic?

Standard normal distribution.

(e) What is the P-value?

One sided: The P-value is $P = 0.03843 = 3.843\%$. From this P-value, we can conclude that Valproate has a significant effect at the 5% level.
Two sided: The P-value is $P = 2 \cdot 0.0384 = 0.0769 = 7.69\%$. This P-value is larger and does not allow us to conclude significance.

(f) From this study, can you conclude that Valproate was reduces heavy drinking in bi-polar alcoholics? What assumption is necessary to enable you to draw a conclusion?

[1] "Study: Anti-Seizure Drug Reduces Drinking in Bi-Polar Alcoholics" in *Bipolar Central* (online), Jan 6, 2005. Reported by N. Pfenning, *Elementary Statistics*, Brookes/Cole, 2011.

One sided: Yes, if the drug was given randomly.
Two sided: No, we cannot conclude this.

(g) What weaknesses do you find in the study?
The sample size is small; 54 people with only 22 in the control group. However, increasing the sample size may not be easy as it involves finding enough bipolar alcoholics.

Using the chi-squared test:

(a) Null hypothesis: There is no association between Valproate and drinking.
(b) Alternate hypothesis: There is an association between Valproate and drinking.
(c) We find the table of expected values:

	Heavy drinking	No heavy drinking	Total
Valproate	17.19	14.81	32
Placebo	11.81	10.19	22
Total	29	25	54

Then $\chi^2 = 3.13$ and the $df = (2-1)(2-1) = 1$.

(d) The statistics has the chi-squared distribution with 1 degree of freedom.
(e) The P-value is 0.0769.
(f) We cannot reject the null hypothesis; there is not significant evidence of an association.

13. (**39 points**) A sports writer wants to see if a football filled with helium travels farther, on average, than a football filled with air. To test this, the writer used 18 adult male volunteers and two identical footballs. One football was filled with helium to the recommended pressure; the other football was filled with air to the recommended pressure. The volunteers were randomly divided into two groups of nine people each. Group 1 kicked the football filled with helium. Group 2 kicked the football filled with air. The mean distance kicked in Group 1 was 30 yards, with a standard deviation of 8 yards. The mean distance kicked in Group 2 was 26 yards, with a standard deviation 6 yards. Assume the two groups of kicks are independent and that the distance kicked is normally distributed. Let μ_1 and μ_2 represent the mean distance kicked we would observe if all members of this population kicked, respectively, a helium-filled and an air-filled football.

(a) Find the 98% confidence interval for the difference in means. (*Show all your work. Keep three decimal places throughout.*)

Let population 1 be the distances kicked of the helium balloons. Let population 2 be the distances kicked of the air balloons.

We use a T-statistic with $df = 8$. Then $t = 2.896$. The standard deviation SE_D of the difference in means, $D = \bar{x}_1 - \bar{x}_2$, is given by

$$SE_D = \sqrt{\frac{8^2}{9} + \frac{6^2}{9}} = 3.333.$$

The confidence interval is

$$(\bar{x}_1 - \bar{x}_2 - t \cdot SE_D, \quad \bar{x}_1 - \bar{x}_2 + t \cdot SE_D)$$

$$(30 - 26 - 2.896\,(3.333), 30 - 26 + 2.896\,(3.333)) = (-5.652, 13.652).$$

For parts (b)-(d), assume the 98% confidence interval in part (a) is $(-6.123, 14.123)$. *This is not the correct answer, but it will enable you to do parts (b) –(d) even if you have not done part (a).*

(b) Mark the following statements T (true) or F (false). *No reasons needed.*

__F___The confidence interval tells us that if the two footballs are kicked by two different people, there's a 98% chance that the difference in the distances traveled will be between −6.123 yards and 14.123 yards.

__ F ___The confidence interval tells us that if the two footballs are kicked by the same person, there's a 98% chance that the difference in the distances traveled will be between −6.123 yards and 14.123 yards.

___ T __The confidence interval tells us that if everyone in the population kicked both footballs, there is a 98% chance that the difference in the mean distances traveled will be between −6.123 yards and 14.123 yards.

<u>#13 continued</u>. For parts (c)-(d), assume the 98% confidence interval in part (a) is $(-\mathbf{6.123}, \mathbf{14.123})$.

(c) What does the confidence interval tell you about whether there is a significant difference in the distances traveled by helium-filled footballs and air-filled footballs? In answering this question, state:

(i) Both hypotheses.

H_0: Mean distance traveled by helium-filled balls = Mean distance by air-filled balls.
H_a: Mean distance traveled by helium-filled balls ≠ Mean distance by air-filled balls.

(ii) The significance level. (That is, the value of α.)

With a 98% confidence interval, we can draw conclusions at the 2% level.

(iii) Your conclusion.

There is no significant difference in the mean distance traveled at the 2% significance level.

(iv) The reason for your conclusion in terms of the confidence interval.

Since 0 is in the interval (one end is negative and the other end is positive), the difference in means could be 0. Thus the confidence interval does not give us evidence that the means are different.

(d) What effect would each of the following changes have on the length of the confidence interval calculated in part (a)? In each case, everything else remains constant.
(*Check one answer in each line. No reasons needed*.)

Decreasing the confidence level:
Lengthens the interval____Shortens the interval__X_Does not change the length of the interval____

Decreasing the sample sizes:
Lengthens the interval_ X _Shortens the interval____Does not change the length of the interval____

Decreasing the standard deviations:
Lengthens the interval____Shortens the interval_ X _Does not change the length of the interval____

Decreasing the mean distance kicked by the helium-filled football by 2 yards:
Lengthens the interval____Shortens the interval____Does not change the length of the interval__ X