Soc 3811 Basic Social Statistics
Second Midterm Exam Spring 2010

Your Name [50 points]: ________________________ ID #:___________

ANSWERS

INSTRUCTIONS:

(A) Write your name on the line at top front of every sheet.

(B) If you use a page of notes in taking this exam, sign & insert it inside this booklet before turning in your exam.

(C) *Show your calculations for numerical problems in the space provided!*

- -

1. Fill in the blanks. **[1 point each]:**

   a. Which alphabet is used to designate sample statistics?

   Latin, Roman, English

   b. The minimum value(s) of Z necessary to designate an alpha area is called

   critical value(s)

   c. The average dispersion of a sampling distribution is measured by its

   standard error

   d. The type of hypothesis that puts the entire alpha area into half of a sampling distribution called

   one-tailed

   e. The branch of statistics concerned about using sample statistics to make conclusions about population parameters is called

   inferential statistics

2. Use the edited version of the Appendix C Areas Under Normal Curve table on the next page to find either the alpha area(s) (α) or the Z score(s) corresponding to the alpha areas in the specified tail(s) of a standardized normal distribution. **[5 points]:**

A. Area between the mean and *Z* = +2.80:

0.49744

B. Area from Z = -3.60 to $-\infty$:

0.000159

C. *Z* score(s) for an α area = 0.0082 in the right tail:

+2.40

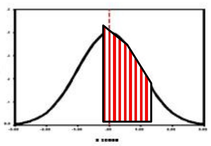
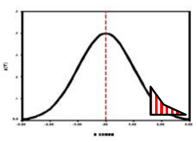
D. What is the sum of the α areas for *Z* $\pm$3.05?

0.00228

E. If the sum of α areas in both tails = 0.0100, what are the *Z* score(s)?

±2.58

Appendix C
Areas Under the Normal Curve

Z-score	Area from 0 to Z	Area from Z to ∞
		
0.00	0.0000	0.5000
0.50	0.1915	0.3085
1.00	0.3413	0.1587
1.50	0.4332	0.0668
1.60	0.4452	0.0548
1.65	0.4505	0.0495
1.70	0.4554	0.0446
1.75	0.4599	0.0401
1.80	0.4641	0.0359
1.85	0.4678	0.0322
1.90	0.4713	0.0287
1.95	0.4744	0.0256
1.96	0.4750	0.0250
2.00	0.4772	0.0228
2.05	0.4798	0.0202
2.10	0.4821	0.0179
2.15	0.4842	0.0158
2.20	0.4861	0.0139
2.25	0.4878	0.0122
2.30	0.4893	0.0107
2.33	0.4901	0.0099
2.35	0.4906	0.0094
2.40	0.4918	0.0082
2.45	0.4929	0.0071
2.50	0.4938	0.0062
2.55	0.4946	0.0054
2.58	0.4951	0.0049
2.60	0.4953	0.0047
2.65	0.4960	0.0040
2.75	0.4970	0.0030
2.80	0.49744	0.00256
2.85	0.49781	0.00219
2.90	0.49813	0.00187
2.95	0.49841	0.00159
3.00	0.49865	0.00135
3.05	0.49886	0.00114
3.10	0.49903	0.00097
3.15	0.49918	0.00082
3.20	0.49931	0.00069
3.25	0.49942	0.00058
3.29	0.49950	0.00050
3.30	0.49952	0.00048
3.35	0.49960	0.00040
3.40	0.49966	0.00034
3.45	0.49972	0.00028
3.50	0.499767	0.000233
3.60	0.499841	0.000159
3.70	0.499892	0.000108
3.80	0.499928	0.000072
3.90	0.499952	0.000048
4.00	0.499968	0.000032

Your name: ____

3. In a population of church members, the parameter values for number of times they attend church each year are known to be $\mu_Y = 38.7$ and $\sigma_Y^2 = 16.0$. Calculate the expected means and standard errors of two sampling distributions with the sample sizes shown below. (Show your work to receive partial credit for wrong answers)**[5 points]:**

(A) For samples of N = 4,900 respondents

$$\mu_{\bar{Y}} = 38.7$$

$$\sigma_{\bar{Y}} = \sqrt{\sigma_Y^2 / N} = \sigma_Y / \sqrt{N}$$
$$= \sqrt{16.0 / 4900} = 4.0 / \sqrt{4900} = 4 / 70 = 0.057$$

(B) For samples of N = 256 respondents

$$\mu_{\bar{Y}} = 38.7$$

$$\sigma_{\bar{Y}} = \sqrt{\sigma_Y^2 / N} = \sigma_Y / \sqrt{N}$$
$$= \sqrt{16.0 / 256} = 4.0 / \sqrt{256} = 4 / 16 = 0.25$$

Your name: ____

4. A sample of 1,111 college students was asked how many credit cards they had. The sample mean = 3.25 and the standard deviation = 2.14. First, estimate the standard error of the sampling distribution. Then, calculate and report the lower and upper limits of the 99% confidence interval (for Z scores see table on page 3). Finally, state your inference about the population parameter in regard to that confidence interval. **[5 points]:**

$$s_{\bar{Y}} = s_Y / \sqrt{N} = 2.14 / \sqrt{1100} = 2.14 / 33.166 = 0.0645$$

Estimated standard error of the sampling distribution: 0.0645

$$CI_{(1-\alpha)(100\%)} = \bar{Y} \pm Z_{\alpha/2} \sigma_{\bar{Y}}$$

$$CI_{99\%} = 3.25 \pm (2.58)\,(0.0645) = 3.25 \pm 0.166$$

$$UCL_{99\%} = 3.25 + 0.166 = 3.58$$

$$LCL_{99\%} = 3.25 - 0.166 = 3.08$$

Lower limit of the 95% CI: 3.08

Upper limit of the 95% CI: 3.58

Your inference about the population parameter:

Because a population parameter falls between the lower and upper limits in 99 of 100 such confidence intervals, we can be 99% confident that the mean number of credit cards in the population lies between 3.08 and 3.58.

5. Last year, an operations manager found the error rate in her widget factory was 78 defects (per million widgets produced). She hypothesizes that the error rate this year will be lower, because of new procedures she put into operation. Write the manager's null and research hypotheses in both English-language and symbolic notation forms. State the critical value(s) of Z for testing the null hypothesis at $\alpha = .01$, assuming that a large random sample of widgets will be selected from the population. (For Z scores, see table on page 3) **[5 points]:**

English language form:

H_0: Error rate will be 78 or more (per million).

H_1: Error rate will be less than 78 (per million).

Symbolic notation form:

H_0: $\mu \geq 78$

H_1: $\mu < 78$

Critical value(s) of Z = -2.33 (for α = .01)

Your name: ____

6. A sports analyst hypothesizes that Minnesotans attend fewer than 10 athletic events per year. In a sample of 463 Minnesotans, he finds that mean = 9.5 events and standard deviation = 4.6. Write his null and research hypotheses in symbolic form. Set α = .001 and state the critical value(s) of *Z* (see page 3). Calculate the *t* test statistic and state your decision about the null hypothesis. If you decide to reject H_0, report the probability of making a Type I error (false rejection error). State a substantive conclusion of your decision. **[5 points]:**

H_0: $\mu \geq 10$

H_1: $\mu < 10$

Critical value(s) of Z = -3.30 for α = .001

$$s_Y = 4.6$$

$$t = \frac{\bar{Y} - \mu}{s_Y / \sqrt{N}} = \frac{9.5 - 10.0}{4.6 / \sqrt{463}} = \frac{-0.5}{4.6 / 21.5} = \frac{-0.5}{0.21} = +2.34$$

Decision about H_0: Do not reject H_0

Probability of Type I error: ----

State the substantive conclusion of your decision:

Minnesotans do not attend fewer than 10 athletic events per year.

7. A journalist hypothesizes that the frequency of newspaper reading differs by gender. He analyzes these data on newspaper reading (a 100-point scale from "never" to "every day"):

	Women	Men
Mean $\overline{Y}$	75	70
Variance s_Y^2	361	441
Sample size N	861	648

Write the researcher's null and research hypothesis pair in symbolic form. Calculate the *t* test statistic and state your decision about the null hypothesis. If you reject H_0, report the **lowest probability** of a Type I error (false rejection error). State the substantive conclusion of your decision. **[5 points]:**

$H_0: \mu_W = \mu_M$ OR: $H_0: \mu_W - \mu_M = 0$

$H_1: \mu_W \neq \mu_M$ OR: $H_1: \mu_W - \mu_M \neq 0$

$$t = \frac{(\overline{Y}_W - \overline{Y}_M) - (\mu_W - \mu_M)}{\sqrt{s_W^2 / N_W + s_M^2 / N_M}} = \frac{(75-70)-0}{\sqrt{\frac{361}{861} + \frac{441}{648}}}$$

$$= \frac{+5.0}{1.049} = +4.77$$

For α = .05, c.v. of Z = ±1.96
For α = .01, c.v. of Z = ±2.58
For α = .001, c.v. of Z = ±3.30

Decision about H_0: Reject H_0 at all three α's

Probability of Type I error: $p < .001$

State the substantive conclusion of your decision:

Women read newspapers more than men
OR Women and men differ in newspaper reading.

Your name: ____

8. An economist believes that the proportion of minority women with high-paying jobs ($75,000 or more per year) is now greater than 0.20. In a sample of N = 225 women, she finds the proportion = 0.23, and estimates a standard error = 0.012. Write the null and research hypothesis pair in symbolic form. **Choose your own α-level** and state the critical value(s) of Z. Calculate the *t* test statistic and state your decision about the null hypothesis. State the substantive conclusion. **[5 points]:**

H_0: $\rho \leq 0.20$

H_1: $\rho > 0.20$

Your α-level: ________________ **Critical value(s) of *Z*:**

For α = .05, c.v. of Z = +1.65
For α = .01, c.v. of Z = +2.33
For α = .001, c.v. of Z = +3.10

$$t = \frac{p - \rho}{s_{\overline{P}}} = \frac{0.23 - 0.20}{0.012} = \frac{+0.03}{0.012} = +2.50$$

Decision about H_0: Reject H_0 at α = .05 or .01, don't reject at α = .001

Probability of Type I error: p < .05/.01 or -----

State the substantive conclusion of your decision:
The proportion of professional women with high-paying jobs is (or is not) greater than 0.20.

9. A criminologist hypothesizes that gun ownerships differs for middle-class and working-class people. Here are data on the proportions owning guns in a national survey:

	Middle class	**Working class**
Proportion that own guns	**0.35**	**0.43**
Proportion that do not own guns	**0.65**	**0.57**
N of respondents	**280**	**420**

Write the null and research hypothesis pair in symbolic form. **Choose your own α-level** and state the critical value(s) of Z. Calculate the *t* test statistic and state your decision about the null hypothesis. If you reject H_0, report the probability of a Type I error (false rejection error). State the substantive conclusion of your decision. **[5 points]:**

H_0: $\rho_M = \rho_W$ **OR:** **H_0: $\rho_M - \rho_W = 0$**

H_1: $\rho_M \neq \rho_W$ **OR:** **H_0: $\rho_M - \rho_W \neq 0$**

$$t = \frac{(p_M - p_W) - (\rho_M - \rho_W)}{\sqrt{p_M q_M / N_M + p_W q_W / N_W}} = \frac{(0.35 - 0.43) - 0}{\sqrt{\frac{(0.35)(0.65)}{280} + \frac{(0.43)(0.57)}{420}}}$$

$$= \frac{-0.08}{0.038} = -2.14 \quad (or + 2.14)$$

For α = .05, critical value of t = ±1.96
For α = .01, critical value of t = ±2.58
For α = .001, critical value of t = ±3.30

Decision about H_0: Reject H_0 at α < .05

Probability of Type I error: p < .05

State the substantive conclusion of your decision:

Working-class people are more likely to own guns.

10. A health researcher hypothesizes that grown daughters are more likely than their mothers to exercise regularly. Here are statistics from a paired sample of mothers and daughters:

How many days per week do you exercise?	Sample means $\bar{Y}$	Sample standard deviation s_D	Sample size (N)
Mothers	**3.2**	**9.0**	**196**
Daughters	**4.5**		

Write the null and research hypothesis pair in symbolic form. Show your *t* test calculations and state your decision about the null hypothesis. If you reject H_0, report the **lowest probability** of a Type I error (false rejection error). State the substantive conclusion of your decision. **[5 points]:**

H_0: $\mu_D = 0$

H_1: $\mu_D \neq 0$

$$t = \frac{\bar{Y}_D - \mu_D}{s_D / \sqrt{N}}$$

$$= \frac{(3.2 - 4.5) - 0}{9.0 / \sqrt{196}} = \frac{-1.3}{0.64} = -2.02 \quad (or + 2.02)$$

For α = .05, critical value of t = ±1.96

For α = .01, critical value of t = ±2.58

For α = .001, critical value of t = ±3.30

Decision about H_0: Reject H_0 at α < .05

Probability of Type I error: p < .05

State the substantive conclusion of your decision:

Daughters exercise more than their mothers.