STAT 151: Lecture U01, Final Exam: V1; Instructor: P. Cartledge; April 19, 2007

Name: _____ Student ID: _____

Instructions: This is a closed book exam. You may use the formula sheet and the tables provided and a non-programmable calculator. The exam has twleve pages and has 45 multiple choice questions. In each multiple choice question, choose the answer you think is closest to being correct. There are no deductions for incorrect guesses. Mark your choices clearly in the answer section below by writing the letter corresponding to your chosen answer. Make sure your answers are correctly located and clearly marked. Each correct answer is worth 1 mark.

Hand in ALL pages (including the formula sheet/tables handout). Print your name and student ID at the top of this page.

1. _____
2. _____
3. _____
4. _____
5. _____

6. _____
7. _____
8. _____
9. _____
10. _____

11. _____
12. _____
13. _____
14. _____
15. _____

16. _____
17. _____
18. _____
19. _____
20. _____

21. _____
22. _____
23. _____
24. _____
25. _____

26. _____
27. _____
28. _____
29. _____
30. _____

31. _____
32. _____
33. _____
34. _____
35. _____

36. _____
37. _____
38. _____
39. _____
40. _____

41. _____
42. _____
43. _____
44. _____
45. _____

Note: You MUST place an answer on THIS page to get a mark for a question.

STAT 151: Lecture U01, Final Exam: V1; Instructor: P. Cartledge; April 19, 2007

1. Superman is having a dinner party at the Fortress of Solitude, but Superman's lazy distant cousin, Gary, is inviting the same guests to his get-together at the Fortress of Hebetude. Each guest's decision to go to a party is independent of the decisions of other guests. If a single guest has a probability of 0.52 of going to Superman's party, what is the approximate probability that at most 4 of 5 guests go to Gary's party?
A) 0.025 **B)** 0.975 C) 0.038 D) 0.962

3. An online resource to evaluate professors, known as www.ratemyprofessors.com, provides ratings between the range of 0.0 to 5.0 for each professor. A sample of 5 average ratings of statistics professors reveals a sample mean of 3.68 and a sample *variance* of 0.697. Assume the population of average ratings is normal. A 99% confidence interval for the mean rating of ALL statistics professors is
A) 3.68 ± 1.719 B) 3.68 ± 1.257 C) 3.68 ± 1.435 D) 3.68 ± 1.168

4. Suppose you now wish to estimate the mean rating so that the margin of error is within 0.11 of the true population mean. Most average ratings are between 1.8 and 4.7. How large should n be to achieve 90% confidence in the mentioned level of accuracy?
A) 52 B) 90 C) 64 **D)** 53

5. A student wanted to measure the monthly amount of alcohol they purchased (by individual receipts) during one academic term. From three different months, they randomly sampled the same number of receipts. Assume all populations are normal with some common variance. Unfortunately, some excess alcohol washed away some of their work, leaving only the information in the table below.

Source	**df**	**SS**	**MS**	***F***
Treatments				
Error			50	
Total	**53**	**10000**		

From the table above, which of the following is TRUE?
A) There are 3 observations per group and the student will likely reject H_0.
B) There are 18 observations per group and the student will likely fail to reject H_0.
C) There are 3 observations per group and the student will likely fail to reject H_0.
D) There are 18 observations per group and the student will likely reject H_0.

STAT 151: Lecture U01, Final Exam: V1; Instructor: P. Cartledge; April 19, 2007

6. In a film from the summer of 2006, an undercover cop was heard to remark, "Probability is a lot like gravity…it doesn't *negotiate*." Suppose 1 out of every 50 people who saw this film believed this "opinion" and a sample consisted of 815 people. What is the approximate probability that the sample proportion is greater than 3 out of 100?
A) 0.0212 B) 0.0475 **C)** 0.9525 D) 0.9788

7. A colour commentator (e.g. Ray Ferraro) provides exceptional insight to coincide with the play-by-play of hockey games. As a former player/professional doing his job, he first assumes that NHL players are talented and deserve to play at this level, unless there is enough evidence refuting this assumption. On Jan. 4, 2007 (http://www.youtube.com/watch?v=uI0CFWPxdss), Mr. Ferraro commented the following: "Patrik Stefan, you should be embarrassed for what you just did!" Mr. Stefan also has 7 seasons of NHL experience, indicating he is talented and deserves to play at the NHL level. This situation is analogous to
A) committing Type I and Type II errors simultaneously.
B) committing no error at all.
C) committing a Type I error.
D) committing a Type II error.

8. Deep within the inner reaches of Mexico lies a sleepy village harassed by a gang of outlaws, led by the infamous El Guapo. The saviours brought in to resist this devious señor will overthrow him with a probability of 0.58. The probability for them to just sit and drink tequila is 0.40. Lastly, the probability of overthrowing El Guapo, given that they drink tequila, is 0.86. The probability of the saviours drinking tequila or beating the beastly bandito (and his band of banditos) is approximately
A) 0.748 **B)** 0.636 C) 0.980 D) 0.916

9. The probability distribution for the number of heads in four tosses of a coin is given by

# of heads	0	1	2	3	4
Probability	0.0625	0.25	0.375	0.25	0.0625

Let X represent the number of heads. The probability of at least one tail is given by
A) $P(X > 3)$ B) $P(X \geq 3)$ C) $P(X < 3)$ **D)** $P(X \leq 3)$

10. A street vendor is concerned with buying the best beverage to compliment his advertised product. To determine this, he first takes two independent surveys of random customers and records whether they prefer Mountain Dew or crab juice. The first sample produces 22 of 40 customers who like crab juice whereas the second sample produces 43 of 66. To judge consistency, the vendor then tests to see if the proportions are significantly different. Consequently, he obtains a test statistic of $z_0 = -1.04$. Which of the following pieces of information correspond to his hypothesis test?
A) H_a: $p_1 > p_2$, p-value = 0.1492
B) H_a: $p_1 > p_2$, p-value = 0.2984
C) H_a: $p_1 \neq p_2$, p-value = 0.2984
D) H_a: $p_1 \neq p_2$, p-value = 0.1492

STAT 151: Lecture U01, Final Exam: V1; Instructor: P. Cartledge; April 19, 2007

Use the following information to answer questions 11 – 14:

A researcher was interested in establishing a relationship between the salary of an NCAA basketball coach (x) and the team's winning percentage (y). A random sample of 10 teams gives the following statistical output (salary measured in millions, winning percentage in decimal form):

	df	*SS*	*MS*	*F*	*P*-value
Regression	1	0.0410	0.0410	23.2441	0.001319
Residual	8	0.0141	0.0018		
Total	9	0.0552			

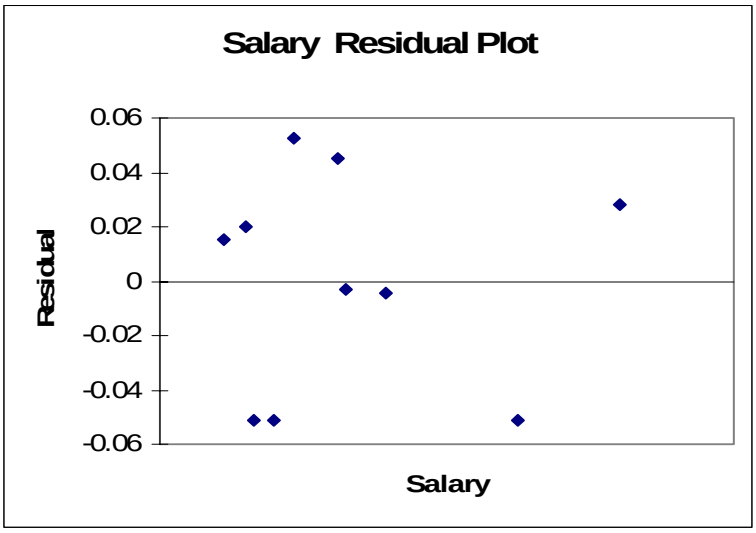
	Coefficients	*Std Error*	*t Stat*	*P-value*
Intercept	0.5083	0.0248	20.5350	3.31E-08
Salary	0.2024	0.0420	4.8212	0.001319

11. A 95% confidence interval for the estimated slope is approximately
A) 0.202 ± 0.042 **B)** 0.202 ± 0.097 C) 0.202 ± 0.078 D) 0.202 ± 0.094

12. To test the validity of regression of team winning percentage on coach salary such that the test is two-sided, the respective test statistic and critical value are
A) 23.24 and $F_{\alpha, 1, n-2}$.
B) 4.82 and $t_{\alpha/2, n-2}$.
C) Both A) and B) are acceptable.
D) None of the above.

[redacted]

Use this *additional* information for question 14:



STAT 151: Lecture U01, Final Exam: V1; Instructor: P. Cartledge; April 19, 2007

14. From the residual plot on the previous page, one may conclude that
A) a small sample size suggests further investigation of the model.
B) an increase in the coach's salary causes an increase in winning percentage.
C) the constant variance assumption of the errors is *completely* satisfied.
D) a nonlinear pattern is emerging from the data.

Use the following information to answer questions 15 – 17:

In a slightly naïve effort to prove that certain Canadian NHL teams have better offensive talent, a random sample of six players from each team from the 2006-07 NHL Season provided the following output regarding each player's individual point total.

SUMMARY

Groups	*Count*	*Average*	*Std Dev*
Cal	6	33.6667	24.4840
Edm	6	21.8333	15.7533
Mon	6	30.5000	18.8547
Ott	6	43.3333	22.5093
Tor	6	35.5000	13.3679
Van	6	30.5000	18.4689

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*
Between Groups	1496.8889	5	299.3778	0.8056	0.5547
Within Groups	11148.0000	30	371.6000		
Total	12644.8889	35			

15. In terms of the assumptions of ANOVA, one may conclude that
A) it is better to use *t*-tools with the above data structure.
B) the assumption of equal variances is not satisfied.
C) the small sample sizes may affect satisfaction of the normality assumption.
D) equality among the sample sizes worsens the violation of equal variances.

16. The appropriate set of hypotheses for this test would be
A) H_0: $\mu_{Cal} = \mu_{Edm} = \ldots = \mu_{Van}$ and H_a: at least one μ is different from the others.
B) H_0: $\mu_{Cal} = \mu_{Edm} = \ldots = \mu_{Van} = 0$ and H_a: each μ is different from the others.
C) H_0: $\mu_{Cal} = \mu_{Edm} = \ldots = \mu_{Van} = 0$ and H_a: at least one μ is different from the others.
D) H_0: $\mu_{Cal} = \mu_{Edm} = \ldots = \mu_{Van}$ and H_a: each μ is different from the others.

17. From the results of the analysis, one may conclude that
A) the null hypothesis is rejected and at least one group is different from the others.
B) the high variability within groups might be masking any significant differences.
C) the null hypothesis is accepted and there are no differences between the groups.
D) an estimate of the common population variance is 19.28.

STAT 151: Lecture U01, Final Exam: V1; Instructor: P. Cartledge; April 19, 2007

18. Which of the following statements about errors in hypothesis testing is TRUE?
A) The power of the test is the complement of the Type I error.
B) Increasing the sample size will decrease both errors.
C) Rejecting the null hypothesis when it is true is called the Type II error.
D) Lowering the Type II error will lower the Type I error.

Use the following information to answer questions 19 – 22:

A professor of archaeology has a secondary business as an "obtainer of rare antiquities". Unsure if he's making enough profit, he consults two separate appraisers and asks them to value the same 10 randomly selected items from his travels. The professor calculates the following summary statistics, regarding item value, from the two appraisers (units are in US dollars). Assume the samples come from normal populations.

Summary statistic	Appraiser 1	Appraiser 2	Difference
Average	1981	1936	65
Standard Deviation	44.00	44.51	54.77

19. Which two-sample test is appropriate?
A) Two-Sample Independent *t*-test: Equal Variances.
B) Paired *t*-test.
C) Two-Sample Independent *t*-test: Unequal Variances.
D) One purely performed by "top…men".

20. If testing whether the first appraiser value the items significantly higher, the appropriate set of hypotheses is
A) H_0: $\mu_1 - \mu_2 > 0$ and H_a: $\mu_1 - \mu_2 \leq 0$.
B) H_0: $\mu_1 - \mu_2 \leq 0$ and H_a: $\mu_1 - \mu_2 > 0$.
C) H_0: $\mu_1 - \mu_2 \geq 0$ and H_a: $\mu_1 - \mu_2 < 0$.
D) H_0: $\mu_1 - \mu_2 = 0$ and H_a: $\mu_1 - \mu_2 \neq 0$.

21. The test statistic and degrees of freedom for the test in question 20, respectively, are
A) $t = 3.75$ and $df = 9$.
B) $t = 11.87$ and $df = 9$.
C) $t = 2.27$ and $df = 18$.
D) $t = 2.27$ and $df = 9$.

22. Using your answer from question 21 and a significance level of $\alpha = 0.05$, then
A) fail to reject H_0 and conclude that the first appraiser values items higher.
B) fail to reject H_0 and conclude that the second appraiser values items equally or higher.
C) reject H_0 and conclude that the first appraiser values items higher.
D) reject H_0 and conclude that the second appraiser values items equally or higher.

STAT 151: Lecture U01, Final Exam: V1; Instructor: P. Cartledge; April 19, 2007

23. If a straight-line relationship exists between two variables, then
A) any pattern of error for a simple linear regression model cannot be verified.
B) the intercept may have no interpretive value.
C) only a simple linear regression model, with every component in the model being fixed, will be useful.
D) ANOVA analysis will not provide any meaningful interpretation.

24. A neurotic neurologist in Newfoundland spends an average of 45 minutes with a patient, whereas the standard deviation is 8.5 minutes. Suppose they take a random sample of 64 patients. The respective probabilities of a sample mean being less than 42.5 and an individual value being less than 40 are
A) 0.0094 and 0.2276.
B) 0.3859 and 0.2276.
C) 0.0094 and not enough information given.
D) 0.3859 and 0.5882.

25. Which of the following statements about statistical distributions is CORRECT?
A) The F distribution is symmetric.
B) The t distribution has two parameters.
C) The F and t distributions are related in some cases.
D) The t distribution is used for tests about proportions.

26. During one of his blackouts when he was infected by Venom, Spider-Man did something insanely insidious: he hired a corrupt accountant. This accountant tried to exploit Spidey's goodwill by claiming all the saved civilians as dependents. Through the accountant's analysis, he discovered that the proportion of saved civilians per month had a 90% confidence interval of 0.84 ± 0.08. From this interval, one may conclude
A) Spider-Man saves 84% of all civilians every month.
B) Spider-Man saves between 76% and 92% of all civilians every month.
C) that many more additional samples, with 90% confidence intervals of their own, will contain p such that exactly 90% of the intervals will contain p.
D) that many more additional samples, with 90% confidence intervals of their own, will contain p such that approximately 90% of all the intervals will contain p.

27. Which of the following statements about hypothesis testing is TRUE?
A) Using the given t-table, there is no need to multiply a calculated range of p-values if the test is one-tailed.
B) The judgment approach states that a p-value greater than 0.05 provides conclusive evidence against H_0.
C) A high p-value will allow you to accept the null hypothesis.
D) Assuming H_0 is true, the p-value measures the probability of obtaining a test statistic value at least as consistent with H_0 as what actually resulted.

STAT 151: Lecture U01, Final Exam: V1; Instructor: P. Cartledge; April 19, 2007

Use the following information to answer questions 28 – 31:

In a recent trend, the quality of movies has sadly decreased (exceptions being *Grindhouse* and 75% of *Blades of Glory*). Using the adage, "[Movies] get better with age", a researcher by the name of Alan Smithee was interested in predicting movie quality from the year a film was made. A random sample of 17 movies gives the following summary data (time measured in years, movie quality measured via the IMDB.com rating system):

$n = 17$, $\bar{x} = 1984.94$, $\bar{y} = 8.26$,
$s_x = 15.36$, $s_y = 0.32$, $r = -0.596$
(Assumptions for regression inference were satisfied.)

28. The estimated regression line, in terms of the variables, is
A) $\hat{\mu}(Rating \mid Year) = 1985.04 - 0.0124 Year$
B) $\hat{\mu}(Rating \mid Year) = 32.91 - 0.0124 Year$
C) $\hat{\mu}(Rating \mid Year) = 56793.42 - 28.61 Year$
D) $\mu(Rating \mid Year) = 32.91 - 0.0124 Year$

Use the following *additional* information for questions 29 – 31:

	df	*SS*	*MS*	*F*
Regression	1	0.5964	0.5964	8.2644
Residual	15	1.0824	0.0722	
Total	16	1.6788		

	Coefficients	*Std Error*	*t Stat*	*P-value*
Intercept	[redacted]	8.6811	3.8267	[redacted]
Year	[redacted]	0.0044	-2.8748	[redacted]

29. To test for clear evidence that increasing the year associates with a decreasing rating, the respective alternative hypothesis and appropriate range of *p*-values are
A) $H_a: \beta < 0$, $0.010 > p\text{-value} > 0.005$
B) $H_a: \beta \neq 0$, $0.020 > p\text{-value} > 0.010$
C) $H_a: \beta < 0$, $0.020 > p\text{-value} > 0.010$
D) $H_a: \beta \neq 0$, $0.010 > p\text{-value} > 0.005$

30. A 99% prediction interval for *Rating* for the year 1990 is
A) 8.23 ± 0.192 B) 8.23 ± 0.779 **C)** 8.23 ± 0.792 D) 8.23 ± 0.189

31. Given all the information above, we may conclude that
A) *Rating* and *Year* have a strong, positive relationship.
B) 2015 will be a terrible year for movie quality.
C) movies made during Roman times were the best movies ever made.
D) *Rating* and *Year* have a moderately strong, linear relationship.

STAT 151: Lecture U01, Final Exam: V1; Instructor: P. Cartledge; April 19, 2007

Use the following information to answer questions 32 – 33:

32. A popular online community known as "Facebook" allows individuals to maintain connections with a large variety of people they meet. A researcher interested in measuring popularity via this community takes a random sample of 11 individuals and notes the "number of friends" they have. The sample mean is 150.27 and the sample standard deviation is 61.97. Since 13 is an "unlucky" number, then he believes that 13^2 must be far worse, so he tests to see if the population "number of friends" is hopefully unequal to this number. Which of the following pieces of information correspond to his hypothesis test? Assume that the observations come from a normal population.
A) H_a: $\mu \neq \mu_0$, reject H_0.
B) H_a: $\mu > \mu_0$, fail to reject H_0.
C) H_a: $\mu > \mu_0$, reject H_0.
D) H_a: $\mu \neq \mu_0$, fail to reject H_0.

33. If you took a second sample of 40 people and miraculously got the exact same sample statistics, a test of the current sample will produce a *p*-value range of
A) (0.025, 0.050) B) (0.100, 1.000) C) (0.000, 0.001) **D)** (0.050, 0.100)

34. A linguist surveyed 1000 people to see if they understood the following phrase:

"Jeden i jeden równa się dwa."

A 95% confidence interval for the proportion who understood is (0.151, 0.235). Which of the following will DEFINITELY decrease the margin of error of a second interval?
A) Increase the confidence level to 99% and increase the sample size to 1100.
B) Decrease the confidence level to 90% and decrease the sample size to 900.
C) Decrease the confidence level to 90% and increase the sample size to 1100.
D) Increase the confidence level to 99% and decrease the sample size to 900.

35. For question 34, which of the following is a VALID assumption for the confidence interval to be appropriate?
A) The sample size is greater than or equal to 30.
B) The sample contains 1000 specifically-chosen people.
C) Both the product of the sample size and the true proportion and the product of the sample size and the complement of the true proportion have to be ≥ 15.
D) The sample contains 1000 randomly-selected people.

STAT 151: Lecture U01, Final Exam: V1; Instructor: P. Cartledge; April 19, 2007

Use the following information to answer questions 36 – 38:

Any fan of sports and resident of Canada will always have to face one of the ultimate decisions: American football or Canadian football? One of the most common arguments is that one league has more skill. Suppose the skill in question is the ability of the quarterback to complete passes to his receivers. A random sample of 10 football players from each league provides the following summary statistics (units are in 100%). Assume the samples come from normal populations.

Summary statistic	NFL	CFL	Difference
Average	57.9	59.9	-2.0
Standard Deviation	3.6	7.9	9.1

36. Which two-sample test is appropriate?
A) Two-Sample Independent *t*-test: Equal Variances.
B) Paired *t*-test.
C) Two-Sample Independent *t*-test: Unequal Variances.
D) Only a summit series à la Canada-Russia 1972 will ever solve this dispute.

37. If testing for any difference between the two leagues, the test statistic and degrees of freedom, respectively, are
A) $t = -0.73$ and $df = 9$.
B) $t = -0.70$ and $df = 9$.
C) $t = -0.73$ and $df = 18$.
D) $t = -0.26$ and $df = 9$.

38. Using your answer from question 37 and a significance level of $\alpha = 0.01$, then
A) fail to reject H_0 and conclude that there may be no difference between the 2 leagues.
B) fail to reject H_0 and conclude that there is a difference between the 2 leagues.
C) reject H_0 and conclude that there is a difference between the 2 leagues.
D) reject H_0 and conclude that there may be no difference between the 2 leagues.

39. Which of the following is a VALID set of hypotheses?
A) H_0: $\mu \neq 15$ vs. H_a: $\mu \neq 15$
B) H_0: $\hat{p} = 0.5$ vs. H_a: $\hat{p} \neq 0.5$
C) H_0: $\mu = 0.65$ vs. H_a: $\mu < 0.65$
D) H_0: $p < 0.76$ vs. H_a: $p \geq 0.76$

40. A survey of 90 University students asked them if they were going to see *Hot Fuzz* within 48 hours of taking this final exam. In this simple random sample, 70 students said they would not. A 90% confidence interval for the population proportion of University students who will be going to the exceptional film is
A) (0.219, 0.225) B) (0.136, 0.308) C) (0.125, 0.320) **D)** (0.150, 0.294)

STAT 151: Lecture U01, Final Exam: V1; Instructor: P. Cartledge; April 19, 2007

41. Using the sample from question 40, how many people should a second sample include in order for the population proportion of students who cannot wait to see *Hot Fuzz* to be estimated within 0.09 with 95% confidence?
A) 119 **B)** 82 C) 58 D) 43

42. Once upon a time, there was an employee at the Primatech Paper Company who got a puppy and cherished the animal's companionship every day. In fact, the employee always found it cute whenever the miniature dog would fascinate itself with the files brought home from work. Years quickly past and the employee got a husband and a son. Soon, the son was taking an undergraduate statistics course and, on the morning of handing in his last assignment, he encountered the inevitability he had hoped he would avoid all his life…his dog ate his homework. Thankfully, he caught the dog in time that he could save some of his work, which is in the table below. Can you help out a fellow student with this ANOVA table? Will you be his HERO?

Source	**df**	**SS**	**MS**	***F***
Treatments	4		206	
Error				
Total	**27**	**2040**		

From the information above, one may state that
A) there are equal observations per group and rejecting H_0 at $\alpha = 0.05$ will occur.
B) there are equal observations per group and failing to reject H_0 at $\alpha = 0.05$ will occur.
C) there are unequal observations per group and failing to reject H_0 at $\alpha = 0.05$ will occur.
D) there are unequal observations per group and rejecting H_0 at $\alpha = 0.05$ will occur.

43. If the coefficient of determination for two variables is 0.952, then
A) only a linear relationship is possible between the two variables.
B) most of the points are falling very close to the estimated regression line.
C) the residual sum of squares is approximately equal to the total sum of squares.
D) the relationship between the two variables is positive and strong.

STAT 151: Lecture U01, Final Exam: V1; Instructor: P. Cartledge; April 19, 2007

Use the following information to answer questions 44 – 45:

As you read this pre-amble, you become *highly* aware that you are sitting in the Pavilion, only minutes away from finishing your STAT 151 final exam! Desperately trying to contain your excitement, you take a deep breath and realize that there must be over ONE THOUSAND people taking an exam in the very same place! Your curiosity gets the better of you and you wonder what proportion of students get to experience the cinematic awe of the entire situation. Your keenly-aware professor telepathically senses your heightened alertness and cerebrally projects the numbers you need to use for the following questions…wait a moment…did it work?! No? Well, just in case, assume a random sample of 400 students, of which 90 are attaining this wonderfully higher level of consciousness. Also, let us claim that the true proportion is equal to 27%. Is the claimed proportion too high? Test the claim using $\alpha = 0.01$.

44. The test statistic for the current hypothesis test is
A) -2.03 B) -2.16 C) 2.03 D) 2.16

45. The *p*-value and conclusion for the test, respectively, are
A) 0.0154 and the claimed proportion is not significantly higher.
B) 0.0212 and the claimed proportion is not significantly higher.
C) 0.0424 and the claimed proportion is not significantly higher.
D) 0.0212 and the claimed proportion is significantly higher.

Now, since you are highly aware, you may want to check over your questions if you have time! Otherwise, make a silent cheer (I will try my best to "hear" it) and please hand in your final exam.

HAVE A GREAT SUMMER!