SAMPLE FINAL B
INSTRUCTIONS

1. The only items allowed within your reach during an exam are calculators, the exam paper, pencils, erasers, a highlighter pen and your tool cards. Take off your hat and put it under your seat.

2. In the upper left corner of your NCS answer sheet PRINT YOUR LAST NAME. Skip a square and print your first name. Bubble in the corresponding spaces beneath those letters.

3. DO NOT FILL IN YOUR STUDENT ID NUMBER.

4. In spaces KLMNO, PRINT THE FIVE DIGITS OF YOUR TEAM NUMBER. Bubble in the corresponding spaces beneath those numbers.

5. **WHOEVER PERFORMS #2, 3 & 4 ABOVE CORRECTLY WILL RECEIVE *ONE* POINT EXTRA CREDIT!!!!!**

6. **Remember, a student is to avoid even the appearance of cheating. Keep your eyes on your exam or on the ceiling. If any member of the teaching team observes questionable behavior on your part, he or she has the right to confiscate your exam and ask you to leave the room.**

7. **ANY talking between students will be interpreted as cheating and all parties will fail the course.**

8. **Absolutely ALL cell phones must be turned off and out of reach. ANY cell phone usage for any purpose will be interpreted as cheating and the user will fail the course.**

9. **ANY TOOL CARDS WHICH HAVE PHOTOCOPIED, COMPUTER GENERATED OR TYPE-WRITTEN INFORMATION ATTACHED TO THEM WILL BE CONFISCATED.**

10. There are 40 multiple choice questions on this exam. Each question is worth 5 points. NOTE: EVEN THE HARD QUESTIONS ARE ONLY WORTH 5 POINTS!

11. I believe that there is only one completely correct answer to each question on this test. Look for it and select it as your best choice.

12. It is possible, though unlikely, that there is a mistake on the exam. **ALWAYS ASK!**

13. Don't let yourself get stuck on one question. Get all the answers you are sure of, then go back to the ones you are not sure of.

14. If you find a question particularly difficult, remember to draw a picture.

15. Stay calm and do your best!

1. Which of the following is/are ethical hypothesis testing behaviors?

   A. Establish hypotheses to correspond with your sample.
   B. Establish hypotheses based on your beliefs.
   C. Delete observations to improve your results.
   D. All of these are ethical.

2. Which of these parameters associated with a binomial experiment will produce a probability distribution with the smallest standard deviation, given *n* is the same for each situation?

   A. $\pi = 0.5$
   B. $\pi = 0.4$
   C. $\pi = 0.3$
   D. $(1-\pi) = 0.1$
   E. $(1-\pi) = 0.8$

3. If, when constructing an interval estimate of the mean number of videos college students watch per month, an E370 student wrote down $\overline{X} \pm T.INV(0.05,35) * s.e.$, which α value was she using?

   A. 0.10 B. 0.05. C. 0.025. D. 0.01.

According to Gallup's U.S. Employee Engagement Index, just over half of surveyed U.S. employees (51%) told Gallup last month that they agree with the statement that the people who run most companies are "honest and ethical." **Use this information to answer the following *SIX* questions. *Note: the questions continue onto the next page.***

4. If 25 employees were randomly drawn, what number would you expect to agree that those who run most companies are “honest and ethical”?

   A. 0.51 B. 12.25 C. 12.75 D. 13.51

5. If 25 employees were randomly drawn, what is the standard deviation of the number who agree that those who run most companies are “honest and ethical”?

   A. 0.10 B. 2.50 C. 6.25 D. 10

6. What is the probability that of 25 employees, fewer than 13 would agree that those who run most companies are “honest and ethical”?

   A. =1-BINOM.DIST(13,25,0.51,1)
   B. =1-BINOM.DIST(12,25,0.51,1)
   C. =BINOM.DIST(13,25,0.51,1)
   D. =BINOM.DIST(12,25,0.51,1)

7. A 96% confidence interval estimate of the percentage of employees who agree that those who run most companies are "honest and ethical" was calculated to be [0.48 , 0.54]. How many employees did the Gallup organization contact? A useful number is one of the following:

=NORM.S.INV(0.96) = 1.75	=T.INV(0.04,1) = 2.28
=NORM.S.INV(0.98) = 2.05	=T.INV(0.02,1) = 2.65
=NORM.S.INV(0.02) = -2.05	=T.INV(0.08,1) = 1.90

A. 851 B. 1003 C. 1167 D. 1444

8. Which of the following is the correct interpretation of the confidence interval estimate in the previous question, [0.48 , 0.54]?

A. We are 96% confident that the interval 48% to 54% includes the true proportion of US employees who agree that those who run most companies are "honest and ethical".
B. We are confident that 96% of US employees agree that those who run most companies are "honest and ethical" between 48% and 54% of the time.
C. Of all US employees, we are confident that 48% to 54% of them agree that those who run most companies are "honest and ethical" 96% of the time.
D. We are 96% confident that the interval 48% to 54% includes the sample proportion of US employees who agree that those who run most companies are "honest and ethical".

9. Using the 96% interval estimate mentioned above, [0.48 , 0.54], what conclusion would we reach if we test H_0: $\pi = 0.56$ versus H_1: $\pi \neq 0.56$ at $\alpha = 0.04$?

A. Accept the null and reject the alternative.
B. Fail to reject the null and conclude the alternative.
C. Reject the null and conclude the alternative.
D. We cannot tell what our decision will be from the information given.

10. In classical hypothesis testing, the test statistic is to the critical value what the _____.

A. p-value is to alpha.
B. critical value is to alpha.
C. test statistic is to the p-value.
D. level of significance is to the test statistic.

11. When we perform linear regression we are

    A. minimizing the sum of the squared predicted value of Y calculated at the mean values of X.
    B. minimizing the sum of the squared difference between the predicted values of Y and the mean of Y.
    C. minimizing the sum of the squared difference between the observed values of Y and the predicted values of Y.
    D. minimizing the squared difference between the mean of the observed values of Y and the mean of the predicted values of Y.

12. The weight of the 2 ounce size of Hershey's chocolate bars is a normally distributed random variable with a mean of 2.1 ounces and a standard deviation of 0.1 ounces. Assuming the production process is running correctly, which is more likely, the random selection of a single bar that weighs no more than 1.9 ounces or the random selection of a sample of 4 bars with a mean weight of no more than 2 ounces?

    A. The likelihood of the single bar weight is greater.
    B. The likelihood of the sample mean weight is greater.
    C. The likelihood of the single bar weight is the same as that of the sample mean weight.
    D. The question requires a computer or other computation aid in order to answer it.

13. You performed a right tailed hypothesis test and your sample gave you a test statistic in the left tail. Which of the following is considered the appropriate next step in this hypothesis test?

    A. Fail to reject the null.
    B. Change the direction of the hypotheses.
    C. Change the level of significance of your test.
    D. Change the sample and do the test again.
    E. All of the above.

14. Net Exports are calculated by subtracting Imports from Exports. Assume Exports and Imports are independent of one another. If mean exports are \$25M with a standard deviation of \$3.5M and mean imports are \$30M with a standard deviation of \$5M, what is the expected value and variance of Net Exports?

    A. $E(X_n)$ = -\$5M $V(X_n)$ = 37.25
    B. $E(X_n)$ = -\$5M $V(X_n)$ = -12.75
    C. $E(X_n)$ = -\$55M $V(X_n)$ = -12.75
    D. $E(X_n)$ = -\$55M $V(X_n)$ = 162.56

In an attempt to find which factors may be related to a car's fuel consumption, we came across a data set of 392 car models. In the data we have four variables, MPG (miles per gallon), Origin (1 if imported, 0 otherwise), Accelerate (time to accelerate from 0 to 60 mph in seconds), and Cylinders (number of cylinders). Partial regression output is as follows. **Use this information to answer the next *FOUR* questions.**

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	36.848	2.425		0.000	32.081	
Origin	2.387		3.740	0.000		3.642
Accelerate	0.145	0.103	1.412	0.159	-0.057	0.348
Cylinders	-3.025					-2.626

15. Researchers are thinking about a more detailed analysis of the role number of cylinders plays in gas mileage, but the analysis will not begin unless the researchers conclude that number of cylinders is important to predicting gas mileage. According to the output above, at the 5% level, what should the researchers do?

    A. Fail to reject the null and begin the detailed analysis.
    B. Fail to reject the null and don't begin the detailed analysis.
    C. Reject the null and begin the detailed analysis.
    D. Reject the null and don't begin the detailed analysis.

16. What is the underlying null hypothesis for the coefficient of Origin?

    A. $H_0 = 0$.
    B. $\mu = 0$.
    C. $\beta_0 = 0$.
    D. $\beta_{Origin} = 0$.

17. According to the output, we won't drop Accelerate as an independent variable because

    A. the p-value is less than the probability of Type II error.
    B. the confidence interval is wide enough.
    C. the t-statistic is greater than 1.
    D. the coefficient is not small.

18. If the p-value for Origin was 0.05, we would have

    A. rejected at $\alpha = .01$ but not at $\alpha = .05$.
    B. rejected at $\alpha = .05$ but not at $\alpha = .01$.
    C. rejected at both $\alpha = .01$ and $\alpha = .05$.
    D. not rejected at either $\alpha = .01$ or $\alpha = .05$.

The table below is an Excel generated correlation matrix for a set of variables from an 1854 survey conducted by the Massachusetts Commission on Lunacy. The variables are: NBR = Number of lunatics, by county, DIST = Distance to nearest mental health center, POP = County population , 1850 (thousands), PDEN = County population density per square mile, and PHOME = Percent of lunatics cared for at home.

	NBR	DIST	POP	PDEN	PHOME
NBR	1				
DIST	-0.04	1.00			
POP	0.98	-0.02	1.00		
PDEN	0.40	-0.18	0.77	1.00	
PHOME	-0.58	0.41	-0.51	-0.44	1.00

19. If an analysis of the factors contributing to the percentage of lunatics cared for at home was planned, for which, if any, independent variable pairs would there be a multicollinearity problem?

A. Between NBR and DIST; DIST and POP; possibly DIST and PDEN
B. Between POP and NBR; possibly POP and PDEN
C. Between DIST and PHOME; PDEN and PHOME; possibly POP and PHOME
D. No evidence of multicollinearity exists in this table.

20. Assuming a linear relationship between X and Y, if Pearson's Correlation Coefficient (r) is equal to **- 0.30,**

A. variable X is larger than variable Y.
B. the intercept (b_0) is negative.
C. the slope (b_1) is negative.
D. there is no correlation.

21. A student claims that she can correctly identify whether a person is a business major or an economics major by the way the person dresses. She is presented with one person and considers this to be a hypothesis test with the null hypothesis being that the person is a business major and the alternative that the person is an economics major. A business major presents herself; the student fails to reject the null. This is an example of

A. a Type I error.
B. a Type II error.
C. a correct decision.
D. Insufficient information.

22. You bought a 25 pack of blank CDR discs. The rumor is that the manufacturer has some quality control problem; some batches of their 25 pack CDR discs are real good, while others are just crap. Overall there is a 50-50 chance that a certain disc burns and plays flawlessly. You would like to know the chance that all of your 25 discs will work great. Your E370 team member said, "Well it's simple--just use BINOM.DIST(25,25,0.5,0)." You treat your team member to a cold brewski for such a brilliant idea and as you stumble home it hits you: Problem! What was it you remembered when you thought of a problem with the BINOM.DIST solution?

    A. The result of one trial is not independent of the results of the other trials.
    B. The number of trials should have been at least 30, but it's only 25.
    C. The outcome of each trial is not one of two mutually exclusive categories.
    D. The random variable is not discrete, as a binomial variable should be.

23. Which of the following is ***TRUE*** about the *Student's t* distribution?

    A. It is bell shaped and symmetrical.
    B. It has more area in the center and less in the tails than does the normal distribution.
    C. It is used to construct confidence intervals for the population mean when the population standard deviation is known.
    D. As the number of degrees of freedom increases, the *t* distribution approaches the binomial.

For train travelers, one of the biggest complaints is of the "off-track" waiting time when the passenger train is bumped off the track to allow a freight train to pass. This waiting time is known to have a normal distribution with a mean of 25 minutes and a standard deviation of 5 minutes. Suppose 100 train trips have been randomly sampled. **Use this information to answer the next *TWO* questions.** ***NOTE: the questions continue onto the next page*.**

24. Describe the sampling distribution of the mean "off-track" waiting time for these 100 trips.

    A. Distribution is normal with mean = 25 minutes, standard error = 5 minutes.
    B. Distribution is normal with mean = 25 minutes, standard error = 0.5 minutes.
    C. Distribution is approximately normal with mean = 25 minutes, standard error = 5 minutes.
    D. Distribution is approximately normal with mean = 25 minutes, standard error = 0.5 minutes.

25. What is the probability that the mean "off-track" waiting time is more than 20 minutes?

A. =NORM.DIST(20,25,0.5,1)
B. =NORM.DIST(20,25,5,1)
C. =1-NORM.DIST(20,25,0.5,1)
D. =1-NORM.DIST(20,25,5,1)

26. Which of the following statements is consistent with the Central Limit Theorem?

A. When μ and σ are known, the population will be approximately normally distributed.
B. If a population has μ and σ, a sample from that population will be normally distributed if the sample size is large enough.
C. When we know σ, the variation in the sample means will be equal to that of the population.
D. Means of samples for n=30 from a Left Handed Weibull distribution will be approximately normally distributed.

The Advertising Council is interested in tracking the size of television advertising budgets for firms with excellent customer product identification recall. The Ad Council would like to make a case that the mean TV advertising budget for such firms has increased from its traditional level of $49 M. The following is partial Descriptive Statistics output as generated by Excel. SPEND is TV advertising budget in millions of dollars. **Use this information to answer the next *FIVE* questions. *NOTE: the questions continue onto the next page*.**

SPEND ($M)	
Mean	50.4
Median	27
Mode	26.9
Sample Variance	2927.8
Range	180.9
Minimum	5
Maximum	185.9
Sum	1058.4
Count	21

27. Which of the following is the appropriate set of hypotheses for a test such as that described above?

A. H_0: $\mu = \$49$ M H_1: $\mu \neq \$49$ M
B. H_0: $\mu \leq \$49$ M H_1: $\mu > \$49$ M
C. H_0: $\mu \leq \$50.4$M H_1: $\mu > \$50.4$M
D. H_0: $\bar{x} \geq \$50.4$M H_1: $\bar{x} < \$50.4$M

28. Assuming that all conditions are appropriate, the p-value to use to test the null hypothesis associated with the Ad Council's research question is

    A. =1-T.DIST(50.4-49/11.8,20,1)
    B. =T.DIST(1.4/54.1,20,1)
    C. =1-NORM.S.DIST(49-50.4)/11.8)
    D. =2*NORM.DIST(1.4,50.4,54.1,1)

29. If the p-value for the hypothesis test described above is 0.005, the null hypothesis of the test would be

    A. rejected at α=.01 but not at α=.001.
    B. rejected at α=.001 but not at α=.01.
    C. rejected at both α=.01 and α=.001.
    D. not rejected at α=.01 and α=.001.

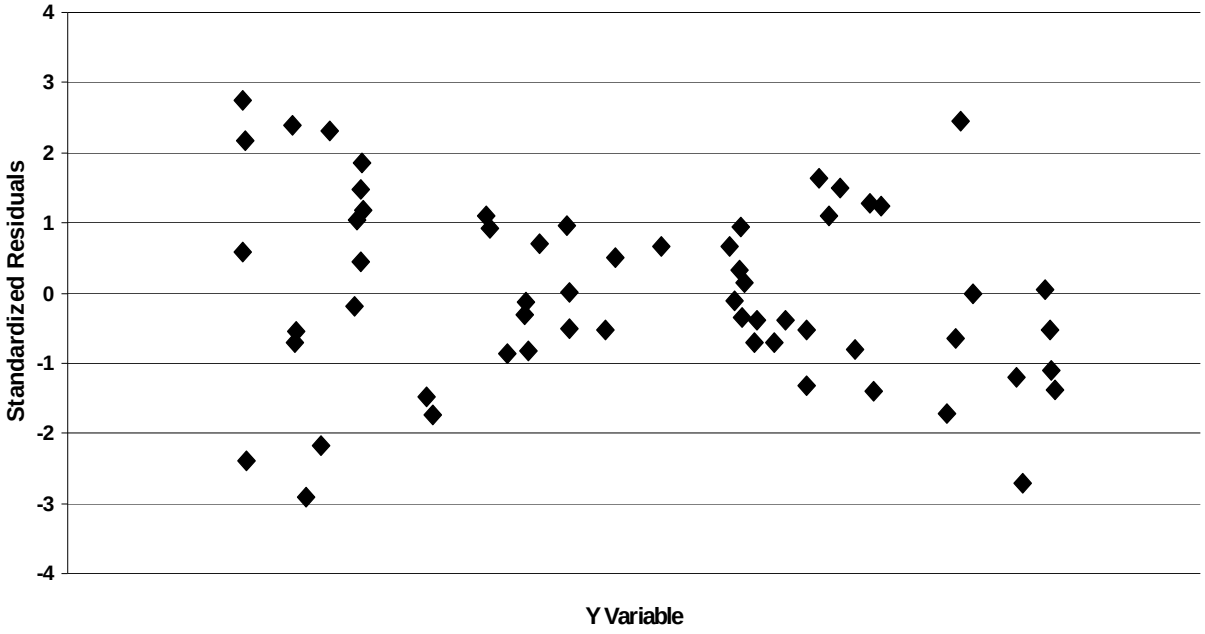
30. If the p-value calculated for the hypothesis test described on the previous page is 0.45, this would mean

    A. the probability of getting a sample mean at least as great as \$49M should be rejected at 0.45, as it is quite small.
    B. the probability of getting a sample mean at least as great as \$50.4M should be rejected at 0.45, as it is quite large.
    C. the probability of getting a sample mean at least as great as \$49M, assuming μ = \$49M is 0.45, that is, highly unlikely.
    D. the probability of getting a sample mean at least as great as \$50.4M , assuming μ = \$49M is 0.45, that is, highly likely.

31. With reference to the Excel output, can a hypothesis test legitimately be performed using this data set?

    A. Yes, because of the Central Limit Theorem.
    B. No, because the data set indicates the population is right skewed.
    C. No, because we don't know the population standard deviation.
    D. Yes, because we know the sample size and the sample standard deviation.

32. The graph below demonstrates a violation of which important linear regression assumption?



A. The mean of the distribution of ε is 0.
B. The standard deviation of ε is independent of the level of the dependent variable.
C. The probability distribution of ε is normal.
D. The value of ε associated with any value of X_i is independent of any other ε.

The Roman Senate has become concerned about the loyalty of the army in Gaul commanded by Julius Caesar. They claim that of the 80,000 men in the army, at least 28,000 are foreign barbarians. Caesar believes that there are fewer barbarians than that, so the Senate should not worry. He polls one legion of 1,000 men and finds that 340 of them are barbarians. **Use this information to answer the next *TWO* questions.**

33. What is the appropriate set of hypotheses to determine if the Senatorial concern is valid?

A. $H_0: \pi = 0.35$ $H_1: \pi \neq 0.35$
B. $H_0: \pi \leq 0.34$ $H_1: \pi > 0.34$
C. $H_0: \mu \leq 28{,}000$ $H_1: \mu > 28{,}000$
D. $H_0: \pi \geq 0.35$ $H_1: \pi < 0.35$

34. What is the standardized test statistic for this hypothesis test?

A. (0.35-0.34)/100
B. (0.34-0.35)/0.015
C. (0.35-0.34)/0.2275
D. (0.34-0.35)/0.063

You have a consulting agency and a client has developed a new ice cream manufactured of all the very best ingredients, which is costly to produce. The client has some questions about ice cream demand and you have gathered data for a regression. Ice cream consumption was measured over 30 four-week periods for the purpose of determining if ice cream consumption depends on the variables price, income, or temperature. Below is part of the Excel output of a regression of per person ice cream consumption (in pints) on price of ice cream ($/pint), family income ($/week) and temperature (Warm (above 80°) = 1, otherwise Cool = 0.) **Use this information to answer the next *FOUR* questions.** *Note: the questions continue onto the next page.*

SUMMARY OUTPUT

Regression Statistics	
Multiple R	
R Square	0.8653
Adjusted R Square	
Standard Error	0.0271
Observations	30

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	6		0.0181	24.6092	6.34e-09
Residual			0.0007		
Total	29	0.1255			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	0.6026	0.2511	2.3993	0.0249	0.0830	1.1221
price	-0.8668	0.7224	-1.1999	0.0924	-2.3611	0.6276
income		0.0019	-1.3063	0.2044	-0.0063	0.0014
temp	0.0019	0.0006	3.0017	0.0064	0.0006	0.0032

35. What is that minimized value that makes this line the line of best fit ?

A. 0.1347 B. 0.0698 C. 0.1086 D. 0.0169

36. What is the point estimate of the coefficient on Income?

A. -0.0025 B. -0.0013 C. 0 D. 0.0007

37. The R Square value 0.8653 means

A. the ratio of consumer and environment characteristics to ice cream consumption is 0.8653.
B. the ratio of ice cream consumption to consumer and environment characteristics is 0.8653.
C. the ratio of the variation in consumer and environment characteristics due to variation in ice cream consumption is 0.8653.
D. the ratio of the variation in ice cream consumption due to variation in consumer and environment characteristics is 0.8653.

38. The meaning of the coefficient on Temp is

A. ceteris paribus, for each additional degree of temperature above 80° ice cream consumption increases by 0.0019 pints per person.
B. ceteris paribus, ice cream consumed on cool days is 0.0019 pints greater than ice cream consumed on warm days.
C. ceteris paribus, warm days see 0.0019 pints more ice cream consumed on average than cool days.
D. ceteris paribus, for each additional pint of ice cream consumed, the temperature moves 0.0019 degrees closer to being warm.

A regression of the price of fine china on variables contributing to the price was calculated, with the following estimated equation a result. Use this equation to answer the next ***TWO*** questions.

$$PRICE' = -32.84 + 8.92\,(BOWL) + 4.39\,(DIAM) + 1.84\,(TIME)$$

The variables are
BOWL: if the piece is a bowl = 1, otherwise = 0
DIAM: the diameter of the piece in millimeters
TIME: the polishing time of the piece in minutes

39. What is the predicted price when the diameter of a bowl is 400 millimeters, and it was polished for 15 minutes?

A. $86.92 B. $260.76 C. $782.28 D. $1759.68

40. The fifth piece of china in the data set used to calculate this regression was a round platter with a diameter of 350 millimeters and a polishing time of 30 minutes. Its price was $1,500.00. How much did this observation contribute to the sum of squared error for the regression?

A. -91.7 B. -58.86 C. 3464.5 D. 8408.89

ANSWERS														
1	B	6	D	11	C	16	D	21	C	26	D	31	B	36	A
2	D	7	C	12	C	17	C	22	A	27	B	32	B	37	D
3	A	8	A	13	A	18	D	23	A	28	A	33	D	38	C
4	C	9	C	14	A	19	B	24	B	29	A	34	B	39	D
5	B	10	A	15	C	20	C	25	C	30	D	35	D	40	C

Sorry–typo on #3. Correct as of 7:00 PM 12/10/11