Name: Key _____ **Math 127 Exam 3 Spring 2017**

Version Applesauce

Oath: *"I will not discuss the exam contents with anyone on Earth until the answer key is posted to BB."*

Sign Name: _____

Permitted Materials: One-sheet of handwritten or typed notes. No copies of published materials.
The datasets are found on www.statcrunch.com. No other webpages.
Any calculator is permitted or use the calculator found on the computers.
No cell phones on the desk. No cell phone calculators.
You must staple your sheet of notes to the exam.

Sign Name: _____

- Show all work when appropriate. StatCrunch can be used for everything unless you are explicitly asked to show a calculation.
- Points are in parentheses for each problem.
- This test is graded out of 100 points and counts for 25% of your Math 127 grade.
- The graded exams are kept on file for at least one year and students are welcome to come collect them whenever I am available in my office.
- An answer key will be posted on Blackboard on Monday, May 8 after the testing is completed.
- Final letter grade cutoffs will be posted to Blackboard on Monday, May 8 around 5 pm. Your numerical "Course Grade" on Blackboard is your final grade in Math 127 and you will know your letter grade based on my announcement. 89.5% is a guaranteed A. 79.5% for a B. 69.5% for a C. 59.5% for a D.
- Letter grades will be posted to MyCecil, but students may see WIP for a few days.
- Good luck on this exam. It has been my pleasure to work with you this semester.

1. (2) 61277 We have a 97% confidence interval for the mean salary of all community college computer science professors, ($_____ to \$65,330)

Margin of error was \$4,053. Determine the sample mean salary.

$65,330 - 4053 = 61,277$

2. (2) _____ We have an interval for the proportion of Baltimore county households who subscribe to Verizon FIOS. The interval is (23.90% to 25.20%). Calculate the margin of error:

$$ME = \frac{25.20\% - 23.90\%}{2} = 0.65\%$$

3. (1) We have a test for a population proportion, and the P-value = 0.387, hypothesized value of 25%. If we're testing for "Greater Than", the sample proportion was:

(Greater Than 25%) Less Than 25% Cannot Tell

4. (2) $t =$ 3.4995 Give the t value for $\bar{y} \pm t\left(s/\sqrt{n}\right)$ if $n = 8$ data points, 99% confidence.

$z =$ 2.807 Give the z value for $\hat{p} \pm z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ for 99.5% confidence.

5. (3) Professor Kupe wants to estimate his true mean total cholesterol. He is willing to pay out-of-pocket for bloodwork multiple times. He'd like to estimate his mean cholesterol to within 5 points. He'd like to be 95% confident. He believes the standard deviation for his true cholesterol to be about 15 points.

Show the calculation to determine the sample size for the number of times he must go get bloodwork done.

$$n = \left[\frac{1.96(15)}{5}\right]^2 = 34.57$$

So $n = 35$

6a. (1) If the P-value is less than the significance level, we will reject the null hypothesis.

6b. (1) Statistically significant data will have small (big / small) P-values.

6c. (1) Type II errors can only occur when the P-value is big (big / small).

7a. (2) Presume we know with certainty that 22% of Cecil stat students use our stat tutors. We will run a large in-house advertising campaign to attempt to increase the proportion. After the campaign, we will take a random sample of $n = 100$ stat students.

To start, presuming $p = 0.22$ still, determine the mean and standard deviation of the model for $\hat{p}$. Show your work.

$\mu_{\hat{p}} = p = 0.22$ $\quad \sigma_{\hat{p}} = \sqrt{\frac{0.22(0.78)}{100}} \approx 0.0414$

7b. (2) How many students in the sample would need to use the stat tutors to convince you our in-house advertising campaign was successful? Show calculation.

Rule of Thumb $\hat{p} + 2\sigma_{\hat{p}}$

$0.22 + 2(0.0414) = 0.3028$

$\hat{p} = 0.3028 \approx \frac{30}{100}$ so $30+$

$31+$

7c. (2) Use your model to determine the probability that the sample has at most 25% who use the tutor.

Answer: 0.7657

8. (3) It is believe that 7% of people are walking around with cracked phone screens. We will run a survey here to see if Cecil students are in line with this figure. First, we need to determine the required sample size. Use a 2.5% margin of error and use 95% confidence. Show the calculation:

$n = \frac{1.96^2(0.07)(0.93)}{(0.025)^2} \approx 400.14$

so $n = 400$

$n = 401$

9a. (1) True or False: True P(Making a Type I Error) = P-value, if you reject the null hypothesis.

9b. (1) Which hypothesis is presumed true before you collect your data? Null

9c. (1) The hypothesized value for our two-sample difference tests (both proportions and means) is 0.

10a. (2) We know the mean age of first time mothers is 26 and let's presume the standard deviation is 5 years. We are going to get the sampling distribution for the sample mean here. Determine the mean and standard deviation of that model if the sample size is $n = 225$. Show work.

$\mu_{\bar{y}} = 26$, $\sigma_{\bar{y}} = \frac{\sigma_y}{\sqrt{n}} = \frac{5}{\sqrt{225}} = 0.3333$

10b. (2) What is the probability that in a random sample of 225 first time mothers, the average age is under 25?

Answer: 0.6013

10c. (2) What is the 99th percentile of the model for $\bar{y}$?

Answer: 26.78

11. (8) Confidence interval questions. Answer only is fine. Use the "**ZZZ Retired - Calendar Year 2017 Large Survey**" dataset.

11a. 95% confidence interval for the true mean "***Credit Card Debt***" for all Cecil College students.

Interval: ($595.72, $1021.30)

11b. 95% confidence interval for the true proportion of all Cecil students who "***Smoke***" cigarettes.

Interval: (12.92%, 19.30%)

11c. 95% confidence interval for the difference in proportions, "***Males***" and "***Females***", who drink "***Soda***" "***Daily or nearly daily***".

Interval: (−0.79%, +13.78%)

11d. 95% confidence interval for the difference in mean "***TV Time***", "***Males***" vs. "***Females***".

Interval: (1.98 hours, 6.71 hours)

12. Use the "**ZZZ Retired - Calendar Year 2017 Personality Type**" dataset for this hypothesis test. Test if more than half of Cecil College students are "***E***" = "***Extraverted***".

12a.(2) Hypotheses: $H_0: p = .50$ vs. $H_A: p > 0.50$

12b.(2) Actually check the four conditions.

Condition 1: I / E is categorical ✓

Condition 2: Presume sample unbiased ✓

Condition 3: Treat n=405 as < 10% of all Cecil students ✓

Condition 4: E = 198 Success, I = 207 Fails ✓

12c.(2) Summarize the data: $\hat{p} = \frac{198}{405} = 0.4889$

12d. (2) Test statistic: $Z = -0.447$

12e. (2) P-value: 0.6724

12f. (2) Decision: Fail to Reject H_0

12g.(2) Conclusion: No evidence at all to say more than 50% of all Cecil students are extraverted

12h.(2) Interpret your test statistic with a sentence in context: Our $\hat{p} = 48.89\%$ was 0.447 Standard errors below the hypothesized value of $p_0 = 50\%$

13. Use the "**ZZZ Retired - Calendar Year 2107 Food Bank**" dataset to test if the average "***Calories***" for grocery store food items is under 125 calories. Conditions are met, you do not have to check.

13a.(2) Hypotheses: $H_0: \mu = 125$ cal vs. $H_A: \mu < 125$ cal

13b.(2) Summarize the data: $n = 644$, $\bar{y} = 102.15$, $s = 75.52$

13c. (2) Test statistic: $t = -7.68$ **13d. (2)** P-value: < 0.0001

13e. (2) Decision: Reject H_0

13f.(2) Conclusion: We have tons of evidence that the mean calories for all packaged foods is under 125 / serving

13g.(2) Interpret your P-value with a sentence in context: If the true mean is 125 calories, we'd get a $\bar{y} = 102.15$, or one even smaller, less than 0.01% of the time

14. Use the "**ZZZ Retired - Calendar Year 2107 Cell Phone Addiction**" dataset to test if a higher proportion of "***Females***" are on "***Social Media***" when compared to the "***Males***".

Use the variable "***App Category***"!

Conditions are met, you do not have to check.

14a.(2) Hypotheses: $H_0: p_F = p_M$ vs. $H_A: p_F > p_M$

14b.(3) Summarize the data: $\hat{p}_F = \frac{140}{281} = 0.4982$, $\hat{p}_M = \frac{63}{232} = 0.2716$

Diff: 0.2266

14c. (2) Test statistic: $Z = 5.23$ **14d. (2)** P-value: < 0.0001

14e. (2) Decision: Reject H_0

14f.(2) Conclusion: We have tons of evidence that a higher % of females are on social media when compared to the males.

14g.(1) If you made a mistake, what kind? Type I

14h. (2) Explain the reality of the situation if you did make that mistake: In reality, the proportions on social media are equal (by gender)

15. Use the "**ZZZ Retired - Calendar Year 2107 Library Data**" dataset to test if books written by "***Males***" are, on average, longer than books written by "***Females***".

~~The variable is "*Pages*".~~ **~~Uncheck "Pool Variances" so we get the same answers.~~** Pages

Conditions are met, you do not have to check.

15a.(2) Hypotheses: $H_0: \mu_M = \mu_F$ vs. $H_A: \mu_M > \mu_F$

15b.(3) Summarize the data:

	n	$\bar{y}$	s
Males	336	355.09	231.45
Females	85	316.61	159.49
Diff:		38.48	

15c. (2) Test statistic: $t = 1.797$ **15d. (2)** P-value: 0.037

15e. (2) Decision: Reject H_0

15f.(2) Conclusion: We are convinced that, on average, books written by Males have more pages than books written by females.

15g.(2) Interpret your standard error with a sentence in context: With repeated samples, we'd expect the diff. in means to vary by about 26.42 pages.