Name: _Key_

Version 24

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

Sign Name: _Key_

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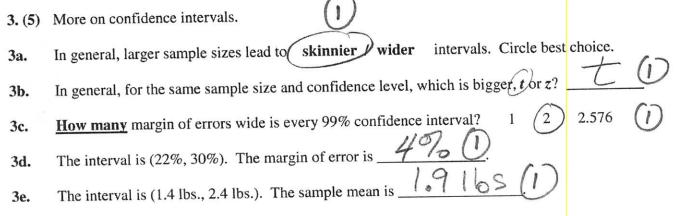No cell phone calculators.
You must staple your sheet of notes to the exam.

- **<u>Show all work when appropriate.</u> StatCrunch provided numbers are OK always, but on sample sizes and minor algebraic calculations, support your answers.**

- 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 pick them up whenever I am available in my office.

- Exam grades will be posted to Blackboard by Friday, December 11th (but possibly sooner). Final grade announcements will be posted to Blackboard by Friday, December 11th (but possibly sooner). Your numerical "Course Grade" on Blackboard is your final grade in Math 127 and you will know your letter grade based on my announcement.

- An answer key will be posted on Blackboard shortly after all Math 127 testing is completed for the Fall 2015 semester. Most likely, the answer key will be posted on Friday, December 11th.

- Letter grades will be posted to MyCecil, but students may see WIP for a few days.

- Good luck on this exam, good luck in the future. It's been my pleasure to work with you this semester.

- Stat II runs in the Spring semester on Tuesday / Friday from 12:00 noon until 1:50 pm. Cheap course, just my notes and a StatCrunch subscription. Go sign up.

1. Use the "**Calendar Year 2015 Dream Home**" dataset. Run the complete two-sample test to determine if, on average, houses that have a "**Garage**" have a newer "**Year Built**" when compared to houses without a "**Garage**".

1a. (2) Hypotheses: H_0: $\mu_{YES} = \mu_{NO}$ (1) H_A: $\mu_{YES} > \mu_{NO}$ (1)

1b. (2) Summarized Data For This Test:

Yes Garage: $n = 104$, $\bar{Y} = 1992.06$, $S = 25.51$ (1)

No Garage: $n = 32$, $\bar{Y} = 1982.41$, $S = 30.73$

Difference: 9.65 YEARS (1)

1c. (3) What three conditions must be met to proceed with the hypothesis test? Explain if each is met. The variable is quantitative, so no need to go over that one.

Condition 1: Presume House Data is Unbiased. (1)

Condition 2: Both $n = 32$, $n = 104$ less than 10% of all homes (1)

Condition 3: Both n exceed 30 ✓ (Needed, both histo. (1) skewed).

1d. (1) Test Statistic: $t = 1.614$ (1) which has a P-value = 0.0568

1e. (2) Decision using $\alpha = 0.05$: Fail to Reject H_0 (2)

1f. (2) Concluding remark in context: No evidence to say that homes with Garages are, on average, newer than homes without garages. (2)

1g. (2) Interpret your standard error of 5.98 with a sentence in the context of the problem:

With repeated samples, we'd expect the difference in the $\bar{Y}$'s to vary by about 5.98 years. (2)

1h. (2) Interpret the test statistic with a sentence in context.

Our difference of 9.65 years was 1.614 standard errors above a 0 difference. (2)

2. (2) You're going to cook up a couple of confidence intervals.

2a. Two-sample CI for the difference in two proportions. 92% confidence. Give the z. ___1.75 ①___

2b. One-sample CI for a mean. 95% confidence. Sample size is 17. Give the t. ___2.12 ①___

3. (5) More on confidence intervals. ①

3a. In general, larger sample sizes lead to (skinnier)/ wider intervals. Circle best choice. ①

3b. In general, for the same sample size and confidence level, which is bigger, t or z? ___t ①___

3c. **How many** margin of errors wide is every 99% confidence interval? 1 (2) 2.576 ①

3d. The interval is (22%, 30%). The margin of error is ___4% ①___.

3e. The interval is (1.4 lbs., 2.4 lbs.). The sample mean is ___1.9 lbs ①___

4. (2) You'd like to estimate the mean starting salary for a Cecil College nursing graduate to within $2000 with 99% confidence. You'll need a reasonable estimate for the standard deviation, so use the $\dfrac{\text{Range}}{6}$ estimate as we did in class. How many recent graduates should be sampled? Show calculations.

$$n = \left[\frac{2.576 \cdot SD\ EST}{2000}\right]^2 =$$

Answers Vary

Will take anything

$n = 5$ to $n = 75$

②

5. (2) The most recent figures from 2014 has 12.8% of the population of Colorado being without health insurance. The state government will run a mini study to see if things have improved in 2015. The required sample size is to be found first. If 95% confidence and a 5% margin of error is required, how many Coloradans should be surveyed in this mini study? Show calculation.

$$n = \frac{1.96^2 (0.128)(0.872)}{(0.05)^2} = 171.51$$

So $n = 172$ ②

6. (12) Open up the "**Calendar Year 2015 Dream Home**" dataset. Test if more than three-quarters of all dream homes have "**Central Air**". Presume this is an unbiased sample of dream homes and that all conditions are met. Show all steps, but StatCrunch can be used for all calculations.

$H_0 : p = 0.75$ (2)

$H_A : p > 0.75$ (2)

Data: $\hat{p} = \frac{113}{129} = 0.876$ (2)

Test Stat: $z = 3.30$ (2)

P-Value $= 0.0005$ (2)

Decision: Reject H_0 (2)

Conclusion: There is tons of evidence that more than 3/4 of all Dream Homes have (2) Central Air.

7. Historically, 38.2% of all American citizens have been organ donors. Recently, a huge, nationwide advertising campaign took place that may have positively impacted the national figure.

7a. (2) The statistician working for www.organdonor.gov will be overseeing a data collection effort in which $n = 400$ adults are randomly surveyed by the government. Presuming the historical figure above still holds, give the mean and standard deviation of the sampling distribution model for the sample proportion. Show the calculation.

$\mu_{\hat{p}} = p = 0.382$ (1)

$\sigma_{\hat{p}} = \sqrt{\frac{0.382(0.618)}{400}} = 0.0243$ (1)

7b. (2) Use your model to answer this question. Do not run a test or compute a confidence interval, because no data has been collected. There may be a number of acceptable answers. How many people in the sample would need to be organ donors to convince you that the advertising campaign was effective? Justify your answer. Explain how you arrived at your number of people.

Answers VARY

Effective means more than 38.2% donors.

I say take:

$\mu_{\hat{p}} + 2\sigma_{\hat{p}} = 0.382 + 2(0.0243)$ (2)

$= 0.4306$

43.06% of 400 = 172 people / 173 people

8. Suppose the mean *"Credit Card Debt"* for <u>all</u> Cecil College students is $1,809 with a standard deviation of $1,880 and the shape of the distribution is very skewed right.

8a. (1) What is the minimum required sample size for the Normality of $\bar{y}$ to kick in? $n \geqslant 30$ ①

8b. (1) Give the mean of $\bar{y}$: $\mu_{\bar{y}} = 1809$ ①

8c. (1) Determine the standard deviation of $\bar{y}$ for sample sizes of $n = 100$. Show calculation.

$$\sigma_{\bar{y}} = \frac{\sigma_y}{\sqrt{n}} = \frac{1880}{\sqrt{100}} = 188 \text{ ①}$$

8d. (1) For $n = 100$, $P(\bar{y} < \$1,750) = 0.3768$ ①

8e. (1) For $n = 100$, give P_3 for the $\bar{y}$ model. $P_3 = \$1455.41$ ①. P_3 is the third percentile, PS.

8f. (1) For $n = 100$, convert a $\bar{y} = \$1,884.26$ to a z-score. Show calculation.

$$z = \frac{1884.26 - 1809}{188} = 0.400 \text{ ①}$$

9. A statistician created a confidence interval for the true proportion of all Democratic voters who would support Bernie Sanders. The interval was 28.242% to 35.698% based on responses from $n = 660$ Democrats.

9a. (2) How many in the sample said they'd vote for Sanders? 211 ②

9b. (2) Can we conclude statistically that Bernie Sanders has at least one-third of the Democratic vote? Why or why not?

① No. Whole① interval would need to exceed 33.33% (It doesn't)

9c. (2) Two things could be done by the statistician to make the interval more precise (skinnier). What are they?

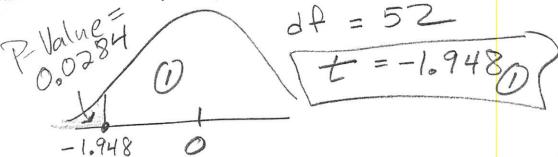1. Collect more data ①
2. Reduce the confidence level. ①

10. Open up the "**Calendar Year 2015 Library Data**" dataset. A few years back, library books used to have a mean "*Weight*" of 750 grams. We have data from $n = 315$ randomly selected books from our library. Is there evidence the mean "*Weight*" has changed?

10a. (2) Hypotheses: H_0: $\mu = 750\ g$ ① H_A: $\mu \neq 750\ g$ ①

10b. (2) Give the appropriate summary statistics for this test:

$$n = 315, \quad \bar{y} = 712.07 \,①\quad S = 449.56 \,①$$

10c. (2) Test Statistic: $t = -1.498$ ①   P-value: 0.1353 ①

10d. (2) Make an accurate, shaded, well-labeled drawing to show where the P-value came from:

$t, df = 314$

② ← P-Value

$-1.498 \quad 0 \quad 1.498$

10e. (2) **Decision:** Reject H_0 (Fail to Reject H_0) ②

10f. (2) Write a conclusion in context: There is no evidence ② to say that the mean weight of a library book has changed from 750 g.

10g. (1) What kind of error could you have made?   Type I   (Type II) ①   Type III   Type IV

10h. (2) If you did make this error, write a sentence that would explain the reality of the situation. ② In Reality, the mean weight of library books is different than 750 grams.

11. Since the economy is doing a little bit better, someone pondered if the mean age of all Harford Community College students has decreased from what it used to be:
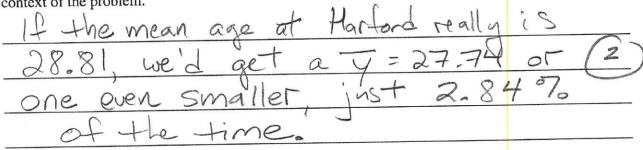
$$H_0: \mu = 28.81$$
$$H_A: \mu < 28.81$$

Suppose the statistics faculty at Harford ran the test. Sample size was $n = 53$, and the P-value was 0.0284. The null hypothesis was rejected because the faculty at Harford always run their hypothesis tests with a 5% significance level.

11a.(2) Determine the test statistic by shading under the appropriate Student's t model. Draw and label.

P-Value = 0.0284 $df = 52$ $t = -1.948$ ①

-1.948 0

11b.(2) We know that the sample mean was 27.74. Interpret the P-value of the test with a sentence in the context of the problem.

If the mean age at Harford really is 28.81, we'd get a $\bar{y} = 27.74$ or ② one even smaller, just 2.84% of the time.

12. (6) A test was run by Pfizer to determine if a new Xanax tablet was as effective as the existing tablet. The researchers failed to reject the null hypothesis at the $\alpha = 0.05$ significance level.

P-Value > 0.05

Circle the only correct choice for each row:

Are the data statistically significant at the 5% level?	Yes	(No) ①	Can't Tell
What kind of error could be made at the 5% level?	Type I	(Type II) ①	Type III
What would they do at the 1% level of significance?	Reject H₀	(Fail to Reject H₀) ①	Can't Tell
Are the data statistically significant at the 1% level?	Yes	(No) ①	Can't Tell
What would they do at the 10% level of significance?	Reject H₀	Fail to Reject H₀	(Can't Tell) ①
Are the data statistically significant at the 10% level?	Yes	No	(Can't Tell) ①

ZZZ Retired

13. (8) "Calendar Year 2015 Dream Home" dataset again. Run the following confidence intervals. StatCrunch is OK for all computations. Presume our sample is unbiased for all Cecil College students.

13a. 95% CI for the mean "*Square Footage*": (2179.39 to 2560.23) ②

13b. Interpret your interval from part **13a.** with a sentence in the context of the problem:

We're 95% confident that the interval ② 2179.39 SQ FT. to 2560.23 SQ.FT contains the true mean SQ FT of Dream Homes.

13c. 99% CI for the proportion of homes with exactly 3 "*Bedrooms*": $\hat{P} = \frac{55}{137}$, (29.36%, 50.93%) ②

13d. Is there statistical evidence that more than ½ of all homes have exactly 3 "*Bedrooms*"? Yes / **No** ②

ZZZ Retired

14. (12) Let's wrap this semester up by comparing females to males one last time. **"Calendar Year 2015 Large Survey"** dataset. Year of the Piano.

Run a complete hypothesis test to determine if a higher proportion of "*Females*" at Cecil College put more weight in "*Faith*". Compare the ladies to the men. Conditions are met. StatCrunch can be used for all calculations.

$H_0: P_F = P_M$ ②

$H_A: P_F > P_M$

Data: $\hat{P}_F = \frac{76}{230} = 33.04\%$ ①

$\hat{P}_M = \frac{29}{102} = 28.43\%$

Difference: 4.61% ①

Test Stat: $Z = 0.834$ ②

P-Value = 0.2022 ②

Decision: Fail to Reject H_0 ②

Conclusion

There is no evidence that a higher % of females ② (compared to the males) put the most weight in Faith.