Name: Key **Math 127 Exam 3 Fall 2014**

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

Sign Name: Key

Show all work when appropriate. Points are in parentheses. This test is graded out of 100 points and counts for 20% of your Math 127 grade.

Thanks for a nice semester. Best wishes. Don't forget to sign up for Math 128 in Spring 2015.

The graded exams are kept on file for at least one year in my office and students are welcome to come see them whenever I'm available in my office.

An answer key will be posted to Blackboard shortly after the testing is completed.

A

1a. (2) ISFJ (The Protector) is the most common personality type making up 14% of the general population. If we take repeated random samples of 200 people, determine the mean and standard deviation of the model for $\hat{p}$. Show your work and round to three decimal places on everything in this problem.

$\mu_{\hat{p}} = p = 0.14$ $\sigma_{\hat{p}} = \sqrt{\frac{0.14(0.86)}{200}} \approx 0.025$

1b. (2) The correct model for $\hat{p}$ is the Normal model. Presuming we have unbiased samples, the 10% condition is met, and personality type being categorical, what is the final condition that needs to be met for Normality to kick in? Show that it is met.

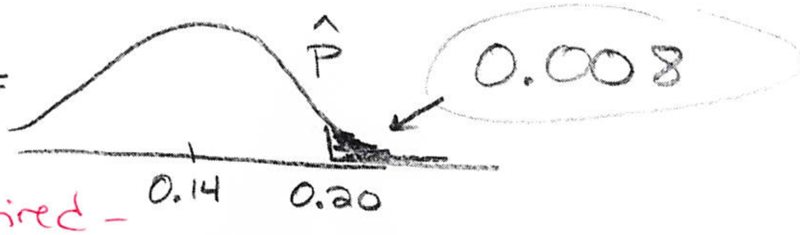
10 Successes Expect $200(0.14) = 28$ ISFJ ✓

10 Failures Expect $200(0.86) = 172$ NONS ✓

1c. (2) What's the probability of getting a sample of 200 people with 40 or more ISJFs? Show shaded Normal model.

$\hat{p} = \frac{40}{200} = 0.20$ ISFJ

$P(\hat{p} > 0.20) =$



ZZZ Retired -

2. Open up the "**Fall 2014 Personality Types**" dataset. Construct a 98% confidence interval for the proportion of Cecil students that are Extraverted.

2a. (2) Confidence Interval: (34.13%, 53.37%)

2b. (2) We would like to test the claim that the majority of Cecil students are Extraverted.

Hypotheses: $H_0: p = 0.50$ vs. $H_A: p > 0.50$

2c. (2) Based on your confidence interval, make a decision and concluding remark in context. Justify your answer with complete sentences in the context of the problem.

Fail to Reject H_0. No evidence a majority of all CC students are extraverted.

2d. (2) What is the margin of error in your interval? 9.62%

2e. (2) What two things could we do to reduce the margin of error?

Thing 1: Bigger Sample Size Thing 2: Less Confidence

ZZZ Retired

3. Professor Sheppard is an ESFJ (The Provider). ESFJ is the second most common type in the population. ESFJs make up: 12% of the general population, **17% of women**, and 8 % of men. Continue with the "**Fall 2014 Personality Types**" dataset.

Run a hypothesis test to determine if the proportion of **women** at Cecil College that are ESFJ differs from the proportion in the general population of **women**.

3a. (2) Hypotheses: $H_0: p = 0.17$ vs. $H_A: p \neq 0.17$

3b. (1) Summarized data from the sample: $\hat{p}_F = \frac{12}{92} = 0.1304$

3c. (2) Test statistic: $Z = -1.01$ P-value: 0.3124

3d. (1) Decision at the 5% significance level: Fail to Reject H_0

3e. (2) Concluding remark in the context of the problem: No evidence to say the percentage of ESFJ females at CC differs from 17%.

Now run a two-sample hypothesis test to determine if the proportion of **women** at Cecil College that are ESFJ is higher than the proportion of **men** at Cecil College that are ESFJ.

3f. (2) Hypotheses: $H_0: p_F = p_M$ vs. $H_A: p_F > p_M$

3g. (1) Summarized data including the difference in sample proportions:

$\hat{p}_F = \frac{12}{92} = 0.1304$ $\hat{p}_M = \frac{1}{52} = 0.0192$

Difference $= 0.1304 - 0.0192$
$= 0.1112$

If the conditions are met, run the test:

3h. (2) Test statistic: $Z = 2.24$ P-value: 0.0127

3i. (1) Decision at the 5% significance level: Reject H_0 -3

3j. (2) Concluding remark in the context of the problem: There is evidence a higher % females are ESFJ at CC.

If the conditions are not met, explain why (5):

Only one success for males. Should not run test.

ZZZ Retired

4. Continue with the "**Fall 2014 Personality Types**" dataset. Run a hypothesis test to determine if people who are "***Judging***" are older on average than people who are "***Perceiving***".

4a. (2) Hypotheses: $H_0: \mu_J = \mu_P$ vs. $H_0: \mu_J > \mu_P$

4b. (1) Summarized data from the samples including the difference in the sample means:

$\bar{Y}_J = 24.53$ $\bar{Y}_J - \bar{Y}_P = 4.23$ years

$\bar{Y}_P = 20.30$

4c. (2) Test statistic: $t = 3.46$ P-value: 0.0004

4d. (1) Decision: Reject H_0

4e. (2) Concluding remark in the context of the problem: There is evidence that the mean age of Judgers exceeds the mean age of Perceivers for all CC students.

4f. (2) What is the probability you've made a Type I error? 0.0004

4g. (2) Explain with a sentence in the context of the problem the meaning of the test statistic:

Our difference of 4.23 years was 3.46 standard errors above 0.

5. **True or False.** In hypothesis testing…**(1 point each)**

I. T F The P-value is the strength of the evidence against the null hypothesis.

II. T F The P-value depends on the significance level.

III. T F The P-value depends on the data.

IV. T F The P-value is P(making a Type I error, if you decided to reject the null hypothesis).

V. T F The P-value is P(the null hypothesis is true).

VI. T F The P-value is a conditional probability.

6. Americans average \$861 on Christmas gifts each year with a standard deviation of \$212, based on one research institution. Make no assumptions about the shape of the distribution for the amount Americans spend on Christmas gifts.

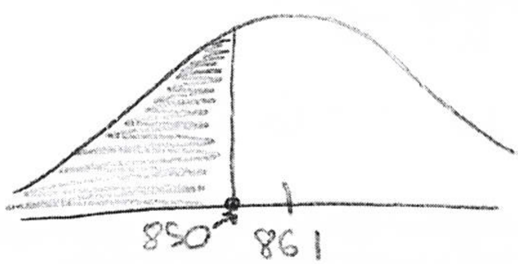
6a. (2) If we were to take repeated random samples of size 10, what condition would we need to confirm in order to model the sample mean with a Normal model?

Purchase amounts must be Normal (since $n < 30$)

6b. (2) With repeated samples of size 90, determine the mean and standard deviation for the sampling distribution for the sample mean. Show work, round the standard deviation to two decimals.

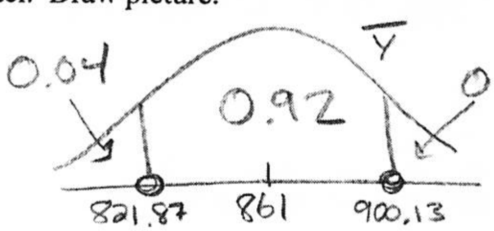
$\mu_{\bar{y}} = \mu_y = 861$ $\qquad \sigma_{\bar{y}} = \frac{\sigma_y}{\sqrt{n}} = \frac{212}{\sqrt{90}} \approx 22.35$

6c. (2) For $n = 90$, what percentage of samples will have a sample mean of at most \$850? Draw picture.



$P(\bar{y} \le 850) = 0.311$

6c. (2) For $n = 90$, give the two values for y-bar that will capture the central 92% of the sampling distribution model. Draw picture.



\$821.87 to \$900.13

6d. (2) Suppose we take a random sample of size 90 and obtain $\bar{y} = \$792$. Is that unusual? Justify.

Use Z-Score

$$Z = \frac{792 - 861}{22.35}$$

$= -3.09$

Yes, very unusual.

Use Probability

$P(\bar{y} < 792) = 0.001$

very unusual.

42%

7. (2) We are going to collect some data in the county – we'd like to know what proportion of households have dogs. There are 42,113 households total, far too many to take a census. A survey done by the ASPCA claims approximately ~~37%-47%~~ of all households in America have dogs. If we'd like to be 95% confident and require a 3% margin of error, determine the required number of households we will need to survey.

$$n = \frac{1.96^2(0.42)(0.58)}{(0.03)^2} \approx 1039.8$$

use $n = 1040$

8. (2) A doctor wants to determine her own true mean resting heart rate to within one beat with 99% confidence. If her standard deviation is approximately three beats, on how many randomly selected mornings should she take a measurement? Show calculation.

$$n = \left[\frac{2.576(3)}{1}\right]^2 = 59.7$$

So $n = 60$

9. Based on a random sample of 40 Cecil students, a 98% confidence interval for the average number of credits students are taking is (8.2 credits, ?????? credits) with a margin of error of 2.18 credits.

9a. (2) Determine the sample mean.

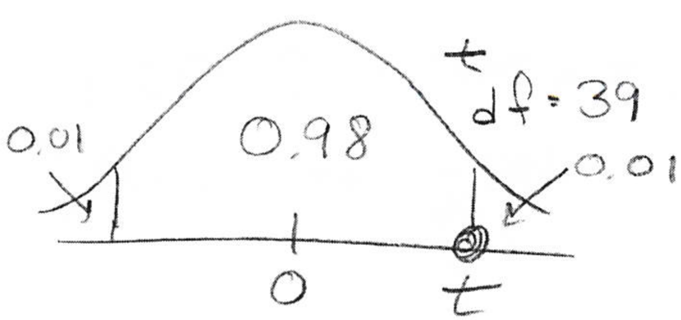
$\bar{y} = 8.2 + 2.18 = 10.38$ credits

9b. (2) Determine the upper boundary of the interval.

upper boundary $= 8.2 + 2(2.18)$

$= 12.56$ credits

9c. (2) Determine the critical t value that was used in the confidence interval formula.



$t = 2.43$

10. (2) In Professor Sheppard's class the average score on Exam 2 was a 66.52 and the standard deviation was 14.43. Pretending you are in her class, you have forgotten your score on Exam 2 but recall that your z-score is 1.38. Find your score on Exam 2. Show calculation.

$$z = \frac{y - \bar{y}}{s} \qquad 1.38 = \frac{y - 66.52}{14.43}$$

$$y = 66.52 + 1.38(14.43) = 86.4334$$

11. Scientists ran a hypothesis test to determine if the proportion of lab rats that found the cheese in a maze increased if the lab rats were given multiple injections of growth hormones. The team rejected the null at the 1% level of significance. **(1 point each)**

Circle the only correct choice for each row:

What would they do at the 5% level of significance?	Reject H_0	Fail to Reject H_0	Can't Tell
What would they do at the 10% level of significance?	Reject H_0	Fail to Reject H_0	Can't Tell
Are the data statistically significant at the 1% level?	Yes	No	Can't Tell
Are the data statistically significant at the 5% level?	Yes	No	Can't Tell
Are the data statistically significant at the 10% level?	Yes	No	Can't Tell
What kind of error could they have made at the 1% level?		Type I	Type II

ZZZ Retired - Calender Year - 2014 Large Survey

12. (2) All classes, use the "~~**Large Survey Math 127**~~" dataset. What condition is not met if we'd like to test if more than 5% of all Cecil students are Buddhists? The variable is "***Religion***". Be clear and show it is not met.

Don't have 10 successes.

Only have 4 Buddhists in sample.

13. (2) Cecil County ran the following test to determine if the mean income per household has decreased from $70,000. The sample mean from $n = 191$ households was $\bar{y} = \$64,550$ and the P-value of the test was 0.014.

$$H_0 : \mu = \$70,000 \quad \text{vs.} \quad H_A : \mu < \$70,000$$

Interpret the P-value with a sentence in the context of the problem: If the true mean in the county is still $70,000 we'd get a $\bar{y}$ = $64,550 or even smaller only 1.4% of the time.

14a. Use the "**Tuesday Men's Handicap**" dataset. Each week three games are bowled. Is there evidence that the mean score for **Game 1** exceeds 1000? This would indicate that for Game 1, individual bowlers are averaging over 200, since there are five dudes per team. Show all steps of the hypothesis test. Run the test using $\alpha = 0.01$. Treat this sample dataset as an unbiased representation of the bowlers in this league. You can assume the conditions are met. **(12 points)**

$H_0 : \mu = 1000$
$H_A : \mu > 1000$

Summarized data

$\bar{y} = 1005.0677$
$n = 768$

Test Stat: $t = 1.48$

P-Value $= 0.0696$

At $\alpha = 0.01$, we fail to reject H_0

There is no evidence that the mean score in Game 1 exceeds 1000.

14b. (2) If you made a mistake above, what type? Circle: Type I Type II Type III

14c. (2) Explain with a sentence in context what it would mean if you did make this mistake.

In reality, the mean score in Game 1 does exceed 1000, our sample just was unlucky and didn't detect it.

14d. (2) The standard error was 3.42. Explain what number means with a sentence in context.

With repeated samples we expect $\bar{y}$ to vary by about 3.42 pins.