***Unit 6 REVIEW (Ch. 23 – 25)***

1. The distribution of scores of students taking the LSATs is claimed to have a mean of 521. We take a sample of 100 incoming Harvard Law School freshman LSAT scores and find a mean of 589 and a standard deviation of 37. Since Harvard is an Ivy League school, they think their freshmen are smarter than average law students. Test this theory (that Harvard students score higher than average on the LSATs) at the 0.05 significance level.

µ = 521 $\overbar{x}=589$ s = 37 α = 0.05 n = 100 df = 99

Ho: µ = 521

Ha: µ > 521

Conditions:

* Random - assumed representative
* 10% condition - there are more than 1000 Harvard law school freshman
* Normality - n = 100 > 30

Conditions met 🡪 Students t-distribution 🡪 1 sample t test

$t= \frac{589-521}{^{37}/\_{\sqrt{100}}}$ = 18.378

P(t > 18.378| df = 99) = 5.6007 x 10-34

We reject Ho b/c p-value of 5.6007 x 10-34 is < α = 0.05. We have sufficient evidence that the average score on the LSATs for Harvard freshman is greater than 521 points.

1. A teacher wants to test the effectiveness of a new textbook. She believes that this new textbook is easier to read, and that her students should have better grades on their tests this year than they have in the past. She took a random sample of test scores from last year’s classes, and then a random sample of test scores from this year’s classes. Assume normal populations for both years. Test her theory at α= 0.01.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Old book** |  |  |  | **New book** |  |  |
| 85 | 84 | 91 | 75 | 65 | 94 | 62 | 86 | 89 | 80 |
| 75 | 82 | 84 | 89 | 62 | 96 | 88 | 88 | 79 | 75 |
| 74 | 64 | 58 | 95 | 50 | 94 | 84 | 86 | 78 | 64 |

 $\overbar{x}\_{O}$ = 75.533 $\overbar{x}\_{N}$ = 82.867 α = 0.01

 sO = 13.266 sN = 10.11

 nO = 15 nN = 15

 Ho: µO = µN

 Ha: µO < µN

 Conditions

* Random - stated random
* Independence - assume new and old books are independent of each other
* 10% condition - there are more than 150 students using the new and old books each year.
* Normality - stated normal populations for both years (the normal probability plots for

both samples are linear too)

Conditions met 🡪 students t-distribution 🡪 2 sample t test

$t= \frac{75.533-82.867}{\sqrt{\frac{13.266^{2}}{15}+\frac{10.11^{2}}{15}}}$ = -1.7026

P(t < -1.7026|df = 26.164) = 0.0502

We fail to reject Ho b/c p-value of 0.0502 > α = 0.01. We have insufficient evidence that the new book’s score are greater than the old book’s average scores. The new book did not improve test scores.

1. A football coach is frustrated with his team’s lack of speed. He measures each player’s 40-yard dash speed and then sends all of them to a speed and agility camp. He then measures their times again after. The data is below. Is their sufficient evidence to say that the camp helped the players speed? Run a test.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Before** | 4.88 | 5.1 | 4.41 | 4.73 | 4.6 | 4.8 | 4.95 | 4.98 | 5.2 | 5.13 | 5.05 | 4.9 | 4.7 | 4.6 | 5.11 |
| **After** | 4.7 | 4.85 | 4.35 | 4.77 | 4.56 | 4.78 | 4.7 | 4.9 | 5 | 5.1 | 5.1 | 4.7 | 4.56 | 4.34 | 4.9 |

 Ho: µd = 0 µd = after - before

 Ha: µd < 0

 Conditions:

1. Random 1) assumed players are representative of all people that do the camp
2. Paired data 2) the 2 sets of data were recorded on the sample people
3. 10% condition 3) There are more than 150 football players that will take the camp
4. Normality 4) the normal probability plot of differences is approx. linear 🡪

 Therefore normal pop

 Conditions met 🡪 use Student’s t-distribution 🡪 1 sample **PAIRED** t-test

 t = $\frac{-0.122-0}{^{0.1077}/\_{\sqrt{15}}}$ = -4.387

 P(t < -4.387| df = 14) = 3.103 x 10-4

We reject Ho because p-value of 3.103 x 10-4 < alpha = 0.05. We have sufficient evidence that the average difference in the speeds before and after the camp is less less than 0. Therefore, the camp did work at improving the average speeds of the players.

1. Poisoning by DDT causes tremors and convulsions and slows recovery times of muscles. In a study of DDT poisoning, researchers fed several lab rats a measured amount of DDT. They then made measurements of the rats’ refractory period (the time needed for a nerve to recover after a stimulus). In their sample they find the following times: 1.61, 1.9, 1.53, 1.4, 1.33, 1.81, 1.3, 1.25, 1.65.
	1. Estimate the average refractory period using 95% confidence.

Conditions:

1. Random 1) assumed representative
2. 10% condition 2) there are more than 90 rats that will be tested
3. Normality 3) the normal probability plot looks linear 🡪 normal population

Conditions met 🡪 use Student’s t-distribution 🡪 1 sample t-interval

$1.5311\pm 2.306\left(^{0.2303}/\_{\sqrt{9}}\right)$ = (1.3541, 1.7081)

We are 95% confident that the true average refractory period is between 1.3541 and 1.7081 miliseconds.

* 1. If we know that the mean time for unpoisoned rats is 1.3 milliseconds, does your interval give evidence that the average time is different for poisoned rats?

Yes, our interval gives us evidence that the mean time for poisoned rats is higher than unpoisoned because the entire 95% confidence interval is above 1.3 milliseconds. We are 95% confident that the average refractory time for poisoned rats is above 1.3 milliseconds.

1. The Chapin Social Insight Test is a psychological test designed to measure how accurately a person appraises other people. The possible scores on the test range from 0 to 41. During the development of the test, it was given to several groups of people. Here are the results for male and female college students at a liberal arts college:

|  |  |  |  |
| --- | --- | --- | --- |
|  | **n** | **avg.** | **std.dev** |
| **Male** | 133 | 25.34 | 5.05 |
| **Female** | 162 | 24.94 | 5.10 |

Does the data support the contention that female and male students differ in average social insight? Use 96% confidence to make your conclusion.

Conditions:

1. Random 1) assumed representative
2. Independent 2) Males and females are independent of each other
3. 10% condition 3) there are more than 1330 male and 1620 female college students
4. Normality 4) nM = 133 and nF = 162 >30

Conditions met 🡪 use student’s t-distribution 🡪 2 sample t-interval

$\left(25.34-24.94\right)\pm 2.063\sqrt{\frac{5.05^{2}}{133}+\frac{5.10^{2}}{162}}$ = (-0.8247, 1.6247) df = 283

We are 96% confident that the true difference between the average score of men and women is between -0.8247 and 1.6247 points. Since 0 is in the interval, we have evidence to say that there is no difference between the scores of the two genders.

1. Many drivers of cars that can run on regular gas actually buy premium in the belief that they will get better gas mileage. To test that belief, we use 10 cars in a company fleet in which all the cars run on regular gas. Each car is filled first with either regular or premium gasoline, decided by a coin toss, and the mileage for that tank-full is recorded. Then the mileage is recorded again for the same cars for a tank-full of the other kind of gasoline. We don’t let the drivers know about this experiment. Here are the results in miles per gallon:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Regular** | 16 | 20 | 21 | 22 | 23 | 22 | 27 | 25 | 27 | 28 |
| **Premium** | 19 | 22 | 24 | 24 | 25 | 25 | 26 | 26 | 28 | 32 |

Is there evidence that cars get significantly better fuel economy with premium gasoline? Use 0.01 level of significance and a test.

Ho: µd = 0 µd = regular - premium

 Ha: µd < 0

 Conditions:

1. Random 1) assumed cars are representative of all cars in the fleet
2. Paired data 2) the 2 sets of data were recorded on the sample cars
3. 10% condition 3) There are more than 100 cars in the fleet
4. Normality 4) the normal probability plot of the differences is approx. linear 🡪

 normal pop of differences.

 Conditions met 🡪 use Student’s t-distribution 🡪 1 sample **PAIRED** t-test

 t = $\frac{\overbar{x}-μ}{^{s}/\_{\sqrt{n}}}$ = -4.472

 P(t < -4.472| df = 9) = 7.749 x 10-4

We reject Ho because p-value of 7.749 x 10-4 < alpha = 0.01. We have sufficient evidence that the average difference in the mpg of regular and premium gas less than 0. Therefore, there is evidence that the cars do get better fuel economy with premium gas.

1. For inference on means, why is Student’s t-model used and not the Normal model?

Because we do not know both the mean and the standard deviation of the population. Therefore we cannot use Z-scores and cannot use the normal model.

Student’s t-model is used when the population standard deviation is not known and must be estimated using the sample standard deviation. Estimating the standard deviation creates more variability.

1. How is Student’s t-model the same as the Normal model?

They are both: unimodal, symmetric, centered at 0

1. How is Student’s t-model different than the Normal model?

The student’s t-model is wider than the normal model.

1. How is degrees of freedom calculated and what effect does it have on the Student’s t-model?

df = n – 1 for 1 sample procedures

df = on the calculator for 2 sample procedures, or we can use the smaller of n1 – 1 and n2 – 1

As the df increases, the t-distribution becomes narrower, and closer to the normal model.

1. A certain population is skewed to the right. We want to estimate the mean so we take a sample. What must we know about the sample if we wish to create a confidence interval to estimate the true mean?

We must know that the sample size is greater than or equal to 30, otherwise our third condition would not check out, and we could not do the confidence interval.

1. If we calculate a 95% confidence interval, how can we decrease the margin of error without losing confidence?

Increase sample size

1. What is the critical value (t\*) for a sample of 67 for 92% confidence?

T\* = 1.778 (use INVT program)

1. What is the critical value for a sample of size 153 for 97% confidence?

T\* = 2.191

1. What is the p-value if I have a test statistic of t = 2.145, sample size of 28, and am doing a 2-sided test?

p-value = 2 \* tcdf(2.145, E99, 27) = 0.0411

1. What is the p-value if I have a test statistic of t = -1.987, sample size of 193, and am doing a lower-tailed test?

p-value = tcdf(-E99, -1.987, 192) 0.0242

1. I have an interval that is (102, 105).
	1. What is the sample mean? 103.5 units
	2. What is the margin of error? 1.5 units
	3. Assuming that the standard deviation is 5 and the sample size is 40, what is the confidence level?

$1.5=t^{\*}\left(\frac{5}{\sqrt{40}}\right)$

t\* = 1.897

Conf level = tcdf(-1.897, 1.897, 39) = ***93.47% confidence***

We wish to see if the dial indicating the oven temperature for a certain model oven is properly calibrated. 12 ovens of this model are selected at random. The dial on each is set to 300° F; after one hour, the actual temperature of each is measured with a thermometer. The temperatures measured had a mean of 302.5° F with a standard deviation of 0.25° F. Assuming that the actual temperatures for this model when the dial is set to 300° are normally distributed*,* we test whether the dial is properly calibrated by testing the hypotheses *H*0: *µ* = 300, *H*A: *µ* ≠ 300

1. Assuming conditions are met, calculate the t-statistic and P-Value for this problem.

T = 34.641 p-value = 1.39 x 10-12

1. What does the P-Value mean in this context?

There is a less than 1% chance of getting a sample where the mean temperature is 302.5° F or higher, if the oven really is 300° F.

1. What would be a Type I error in this context?

We conclude that the oven temperature is not 300° F, when really it is 300° F.

1. What would be a Type II error in this context?

We conclude that the oven temperature is 300 ° F, when really it is not 300° F.

1. What is Power in this context?

We conclude that the temperature of the oven is not 300 ° F, and we are correct that it is not 300° F.

1. If a 96% confidence interval were calculated what would be the critical value?

N = 12 df = 11 Conf = 96%

T\* = 2.328

1. What does 96% confidence mean in this context?

In repeated samples of 12 ovens, the intervals created would catch the true average oven temperature 96% of the time.

1. If the test was redone with the same significance level on 24 ovens and the same mean and standard deviation were found:
	1. What would happen to the P-Value? Decrease (because your test stat would increase)
	2. What would happen to the Type I error? stay the same (because your alpha, or significance level would stay the same)
	3. What would happen to the Type II error? Decrease (opposite of power)
	4. What would happen to the Power? Increase (because n increases)
	5. How would the 96% confidence interval change? Decrease/smaller/skinnier (when the sample size increases, the margin of error decreases)

***NOTE: Extra Credit will be offered. Problems will be from Unit 5 material (proportions)***