STAT 200

Group Exercise 6

Pond's Age-Defying Complex is a cream with alpha-hydroxy acid

- **Pond's Age-Defying Complex is a creme with alpha-hydroxy acid,** a product that is advertised to improve the skin. In a study, 33 women over age 40 used a cream with alpha-hydroxy acid, for 22 weeks. At the end of the study period the women were examined by dermatologists and 23 exhibited skin improvement.
- Conduct a test of a proportion to see if the ingredient improved the skin of more than 60% of the subjects using $\alpha = .05$.

What do I know?

- p = 23/ 33 = .697 q = .303
- n = 33
- **Null Hypothesis p = .6**
- **What is the Standard Error?**

Standard Error = $[(.6)(.4)/33]^{.5}$ = .0853

Pond's Age-Defying Complex

Null hypothesis	**H_0: p = .6**
Alternative	**H_a: p > .6** one-tailed test,upper
Assumptions	**Large sample, binomial = normal**
Test Statistic	**z* = (.697 – .6)/.0853**
Rejection Region	**$z_{\alpha=.05.}$ = 1.645**
Calculation	**z* = 1.14**
Conclusion	**$z^* < z_{\alpha=.05}$** **1.14 < 1.645** **Cannot reject H_0: p = .6**

What is the p-value for the test statistic?

- **What is the p-value for your test?**
- **Look up the probability up to the test statistic**
  - **1.14 in the standard normal table shows a probability of .3729**
- **p = .5 - .3729 = .1271**

Current technology uses X-rays and lasers for inspection of solder-joints

- **Current technology uses X-rays and lasers for inspection of** solder-joints on printed circuit boards. A current manufacturer claims its product can inspect on average at least 10 solder joints per second. A potential buyer tested the equipment on 27 different printed circuit boards (PCB). The sample mean and standard deviation are given below. **We will assume that this distribution is approximately normal.**
- $\bar{x}$ = 10.54 s = 1.61 n = 27

Calculate a 95% confidence interval for this sample estimate.

- **Standard Error = 1.61/(27)$^{.5}$ = .3098**
- **$t_{.05/2,\ 26\ d.f.}$ = 2.056**
- **10.54 ± 2.056(.3098)**
- **10.54 ± .637**
- **9.903 to 11.177**

Inspection of solder-joints

- Test to see if the sample data is different from the manufacturer's claim. You will need to determine the Null Hypothesis and your Alternative Hypothesis.
- Use α = .05.
- **Note: use the t-distribution for the rejection region.**

Set up the Hypothesis Test

Null hypothesis	**H_0: μ = 10**
Alternative	**H_a: $\mu \neq$ 10 two-tailed test**
Assumptions	**Small sample, normal**
Test Statistic	**$t^* = (10.54 - 10)/(1.61/(27)^{.5})$**
Rejection Region	**$t_{\alpha=.05/2,\ 26\ d.f.} = \pm 2.056$**
Calculation	**$t^* = 1.743$**
Conclusion	**$t^* < t_{\alpha=.05/2,\ 26\ d.f.}$** **1.743 < 2.056** **Cannot reject H_0: μ = 10**

Differences in Cholesterol Levels Between Males and Females

- Briefly review the descriptive statistics and describe differences between the two groups.
- Then construct a 95% Confidence Interval around each mean. Be sure and show the standard error of the mean for each group.

Describe Differences

DESCRIPTIVE STATISTICS FOR CHOLESTEROL LEVELS OF MALES AND FEMALES

Cholesterol	*Females*	*Males*
Mean	200.31757	196.085366
Standard Error	0.8812586	0.96610995
Median	201	196
Mode	194	196
Standard Deviation	10.720974	12.372244
Sample Variance	114.93928	153.072423
Kurtosis	-0.4930988	0.01540046
Skewness	-0.1089159	0.08632739
Range	47	61
Minimum	176	166
Maximum	223	227
Sum	29647	32158
Count	148	164
Confidence Level(95.0%)	1.7415714	1.90770209

Description

- **The mean level of cholesterol for females is slightly higher (200.32 versus 196.09)**
- **The median for each group is very close to the mean, indicating no extreme outliers and possibly a symmetrical distribution**
- **The spread of the data for males is larger than that of females.**
  - **Males have the larger range, from 166 to 227**
  - **Std Dev for males is 12.37 versus 10.72 for females**

95% Confidence Interval

- **95% C.I. for females**
  - 200.32 ± 1.96(.8813) =
  - 200.32 ± 1.73
- **95% C. I. for males**
  - 196.09 ± 1.96(.9661) =
  - 196.09 ± 1.89

Hypothesis Test

- Conduct the test that females cholesterol levels are different from males.
- The null hypothesis should be that the two groups are equal.
- **Do this by hand.**
- Use $\alpha = .05$.

Calculate the Standard Error

$$\sigma_{(\bar{x}_1 - \bar{x}_2)} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

$$\sigma_{(\bar{x}_1 - \bar{x}_2)} = \sqrt{\frac{114.939}{148} + \frac{153.072}{164}} = 1.308$$

Set up the Hypothesis Test

Null hypothesis	**H_0: $\mu_1 - \mu_2 = 0$**
Alternative	**H_a: $\mu_1 - \mu_2 \neq 0$ two-tailed test**
Assumptions	**Large, independent samples**
Test Statistic	**$z^* = (200.32-196.09-0)/1.308$**
Rejection Region	**$z_{\alpha=.05/2} = \pm 1.96$**
Calculation	**$z^* = 3.234$**
Conclusion	**$z^* > z_{\alpha=.05/2}$.** **3.234 >1.96** **Reject H_0: $\mu_1 - \mu_2 = 0$**

Now Look at Excel Output

- **Look at the output from Excel for the same problem.** Find the same elements for the Hypothesis Test you just calculated.
- What is the meaning of:
  - **P(T<=t)one-tail**
  - **t Critical one-tail**

Excel Output

-Test: Two-Sample Assuming Unequal Variances

	Female	*Male*
ean	200.31757	196.08537
ariance	114.93928	153.07242
bservations	148	164
Hypothesized Mean ifference	0	
f	310	
t Stat	3.2364596	
(T<=t) one-tail	0.0006706	
t Critical one-tail	1.6497825	
(T<=t) two-tail	0.0013411	
t Critical two-tail	1.9676463	

Excel Output

- What is the meaning of:
  - **P(T<=t)two-tail**
    - This is the p-value for the test statistic, under a two-tailed test p = .0013
  - **t Critical two-tail**
    - This the t-value at the rejection region for a two-tailed test
    - t = 1.968

SAS Output

The TTEST Procedure

Statistics

Variable	GENDER	N	Lower CL Mean	**Mean**	Upper CL Mean	Lower CL Std Dev	Std Dev	Upper CL Std Dev	Std Err
CHOLEST	0	164	194.18	196.09	197.99	11.163	12.372	13.878	0.9661
CHOLEST	1	148	198.58	200.32	202.06	9.623	10.721	12.104	0.8813
CHOLEST	Diff (1-2)		-6.824	**-4.232**	-1.64	10.772	11.619	12.611	1.3173

T-Tests

Variable	Method	Variances	DF	t Value	Pr > |t|
CHOLEST	Pooled	Equal	310	-3.21	0.0015
CHOLEST	**Satterthwaite**	**Unequal**	**310**	**-3.24**	**0.0013**

Equality of Variances

Variable	Method	Num DF	Den DF	F Value	Pr > F
CHOLEST	Folded F	163	147	1.33	0.0770

SAS Output

- **Look at the results from SAS and find the same elements.** What do you think is the meaning of the Equality of Variances test?
  - When n is large we don't have to assume equal variances, but we would gain information if we could.
  - The equality of variance test is the ratio of the variances - we test for a ratio = 1.
  - Since p = .077, this test did not meet the criteria for significance at the alpha = .05 level.

Geneticists have identified E2F1 transcription factor as an important component of cell proliferation control.

- The researchers induced DNA synthesis in two batches of serum-starved cells.
- In one group of 92 cells (treatment), cells were micro-injected with the E2F1 gene.
- A control group of (158 cells) was not exposed to E2F1.
- After 30 hours, researchers determined the number of altered growth cells in each batch. Test to see if the proportion for the treatment group is larger than that of the control.
- Conduct the hypothesis test using $\alpha = .01$.

This is a difference of proportions test, and here is the data

	Control	E2F1 Treated
Total Cells	**158**	**92**
Number of growth altered cells	**15**	**41**
	P =	**P =**

This is a difference of proportions test, and here is the data

	Control	E2F1 Treated
Total Cells	**158**	**92**
Number of growth altered cells	**15**	**41**
	P = .0949	**P = .4457**

Estimate a "pooled" proportion based on the Null Hypothesis that the groups are equal

	Control	E2F1 Treated
Total Cells	**158**	**92**
Number of growth altered cells	**15**	**41**
	P = .0949	**P = .4457**
Pooled estimate of p_p = (15 + 41)/(158 + 92)		
p_p = 56/250 = .224 q_p = .776		

Conduct a Hypothesis Test

Null Hypothesis	H_0: $p_T - p_C = 0$
Alternative	H_0: $p_T - p_C > 0$
Assumptions of	Proportions, large sample, equality of p under
Pooled Estimate of p	$p_p = (15 + 41)/(158 + 92) = 56/250 = .225$ $q_p = .776$
Test Statistic (z* or t*)	z* = $(.4457 - .0949 - 0)/[(.224)(.776)(1/158 + 1/92)]^{.5}$
Rejection Region	$z_{.01} = 2.33$
Calculation of Test	z* = 6.4159
Comparison of z* or t* with Rejection Region	$z^* > z_{.01}$ 6.4159 > 2.33 We can reject H_0: $p_T - p_C = 0$