STAT 200 Week 6 Homework Problems

9.1.2

Many high school students take the AP tests in different subject areas. In 2007, of the 144,796 students who took the biology exam 84,199 of them were female. In that same year, of the 211,693 students who took the calculus AB exam 102,598 of them were female ("AP exam scores," 2013). Estimate the difference in the proportion of female students taking the biology exam and female students taking the calculus AB exam using a 90% confidence level.

n1=144,796

n2=211,693

p1=84199/144796 = 0.582

p2=102598/211693 = 0.485

q1=1-0.582 = 0.418

q1=1-0.485 = 0.515

zc = 1.645

E=1.645*sqrt[{(0.582*0.418)/144796}+{90.485*0.515)/211693}]

E= 0.0028

$(0.582-0.485)-0.0028<p1-p2<(0.582-0.485)+0.0028$

$0.0942<p1-p2<0.998$

There is a 90% chance that $0.0942<p1-p2<0.998$ contains true difference in proportions.

9.1.5

Are there more children diagnosed with Autism Spectrum Disorder (ASD) in states that have larger urban areas over states that are mostly rural? In the state of Pennsylvania, a fairly urban state, there are 245 eight year olds diagnosed with ASD out of 18,440 eight year olds evaluated. In the state of Utah, a fairly rural state, there are 45 eight year olds diagnosed with ASD out of 2,123 eight year olds evaluated ("Autism and developmental," 2008). Is there enough evidence to show that the proportion of children diagnosed with ASD in Pennsylvania is more than the proportion in Utah? Test at the 1% level.

n1=18,440

n2=2,123

x1=245

x2=45

p1=245/18440 = 0.013

p2=45/2123 = 0.021

q1=1-0.013 = 0.987

q1=1-0.021 = 0.979

p-= (245+45)/(18440+2123) = 0.014

q-=1-0.014 = 0.986

z={(0.013+0.021)-0}/sqrt[{(0.014*0.986)/18440}+{(0.014*0.986)/2123}]

z=12.760

9.2.3

All Fresh Seafood is a wholesale fish company based on the east coast of the U.S. Catalina Offshore Products is a wholesale fish company based on the west coast of the U.S. Table #9.2.5 contains prices from both companies for specific fish types ("Seafood online," 2013) ("Buy sushi grade," 2013). Do the data provide enough evidence to show that a west coast fish wholesaler is more expensive than an east coast wholesaler? Test at the 5% level.

Table #9.2.5: Wholesale Prices of Fish in Dollars

Fish	All Fresh Seafood Prices	Catalina Offshore Products Prices
Cod	19.99	17.99
Tilapi	6.00	13.99
Farmed Salmon	19.99	22.99
Organic Salmon	24.99	24.99
Grouper Fillet	29.99	19.99
Tuna	28.99	31.99
Swordfish	23.99	23.99
Sea Bass	32.99	23.99
Striped Bass	29.99	14.99

Fish	All Fresh Seafood Prices	Catalina Offshore Products Prices	d=x1-x2
Cod	19.99	17.99	2
Tilapi	6	13.99	-7.99
Farmed Salmon	19.99	22.99	-3
Organic Salmon	24.99	24.99	0
Grouper Fillet	29.99	19.99	10
Tuna	28.99	31.99	-3
Swordfish	23.99	23.99	0
Sea Bass	32.99	23.99	9

Striped Bass	29.99	14.99	15
		Mean=	2.446
		Standard Dev=	7.399
		Test Stat=	0.110
		df=	8
		p-value=	0.542

Since p-value>0.05, accept Ho.

9.2.6

The British Department of Transportation studied to see if people avoid driving on Friday the 13th. They did a traffic count on a Friday and then again on a Friday the 13th at the same two locations ("Friday the 13th," 2013). The data for each location on the two different dates is in table #9.2.6. Estimate the mean difference in traffic count between the 6th and the 13th using a 90% level.

Table #9.2.6: Traffic Count

Dates	6th	13th
1990, July	139246	138548
1990, July	134012	132908
1991, September	137055	136018
1991, September	133732	131843
1991, December	123552	121641
1991, December	121139	118723
1992, March	128293	125532
1992, March	124631	120249
1992, November	124609	122770
1992, November	117584	117263

Dates	6th	13th	d=x1-x2
1990, July	139246	138548	698
1990, July	134012	132908	1104
1991, September	137055	136018	1037
1991, September	133732	131843	1889
1991, December	123552	121641	1911
1991, December	121139	118723	2416
1992, March	128293	125532	2761
1992, March	124631	120249	4382

1992, November	124609	122770	1839
1992, November	117584	117263	321
		Mean=	1835.8
		Standard Dev=	1173.065
		df=	9
		tc=	1.833
		E=	679.962

1835.8-679.962<ud<1835.8+679.962

There is a 90% chance that 2515.762<ud<2515.762 contains true mean difference betwen the 6th and the 13th.

9.3.1

The income of males in each state of the United States, including the District of Columbia and Puerto Rico, are given in table #9.3.3, and the income of females is given in table #9.3.4 ("Median income of," 2013). Is there enough evidence to show that the mean income of males is more than of females? Test at the 1% level.

Table #9.3.3: Data of Income for Males

$42,951	$52,379	$42,544	$37,488	$49,281	$50,987	$60,705
$50,411	$66,760	$40,951	$43,902	$45,494	$41,528	$50,746
$45,183	$43,624	$43,993	$41,612	$46,313	$43,944	$56,708
$60,264	$50,053	$50,580	$40,202	$43,146	$41,635	$42,182
$41,803	$53,033	$60,568	$41,037	$50,388	$41,950	$44,660
$46,176	$41,420	$45,976	$47,956	$22,529	$48,842	$41,464
$40,285	$41,309	$43,160	$47,573	$44,057	$52,805	$53,046
$42,125	$46,214	$51,630				

Table #9.3.4: Data of Income for Females

$31,862	$40,550	$36,048	$30,752	$41,817	$40,236	$47,476	$40,500
$60,332	$33,823	$35,438	$37,242	$31,238	$39,150	$34,023	$33,745
$33,269	$32,684	$31,844	$34,599	$48,748	$46,185	$36,931	$40,416
$29,548	$33,865	$31,067	$33,424	$35,484	$41,021	$47,155	$32,316
$42,113	$33,459	$32,462	$35,746	$31,274	$36,027	$37,089	$22,117
$41,412	$31,330	$31,329	$33,184	$35,301	$32,843	$38,177	$40,969
$40,993	$29,688	$35,890	$34,381				

Male Earnings
N1: 52

df1 = N - 1 = 52 - 1 = 51
M1: $46,453.31
SS1: 2520976841.08

s21 = 2520976841.08/(52-1) = 49430918.45

Female Earnings
N2: 52
df2 = N - 1 = 52 - 1 = 51
M2: $35,376.38
SS2: 4047089204.31
s22 = 4047089204.31/(52-1) = 79354690.28

T-value Calculation

s2p = ((51/102) * 49430918.45) + ((51/102) * 79354690.28) = 64392804.37

s2M1 = 64392804.37/52 = 1238323.16

s2M2 = 64392804.37/52 = 1238323.16

t = 11076.92/sqrt2476646.32 = 7.04

The t-value is 7.04. The p-value is < .00001. The result is therefore significant at p < 0.01. Therefore there is enough evidence to show that the mean income of males is more than females at a 1% level

9.3.3

A study was conducted that measured the total brain volume (TBV) (in mm^3) of patients that had schizophrenia and patients that are considered normal. Table #9.3.5 contains the TBV of the normal patients and table #9.3.6 contains the TBV of schizophrenia patients ("SOCR data oct2009," 2013). Is there enough evidence to show that the patients with schizophrenia have less TBV on average than a patient that is considered normal? Test at the 10% level.

Table #9.3.5: Total Brain Volume (in mm^3) of Normal Patients

1663407	1583940	1299470	1535137	1431890	1578698
1453510	1650348	1288971	1366346	1326402	1503005
1474790	1317156	1441045	1463498	1650207	1523045
1441636	1432033	1420416	1480171	1360810	1410213
1574808	1502702	1203344	1319737	1688990	1292641
1512571	1635918				

Table #9.3.6: Total Brain Volume (in mm^3) of Schizophrenia Patients

1331777	1487886	1066075	1297327	1499983	1861991

1368378	1476891	1443775	1337827	1658258	1588132
1690182	1569413	1177002	1387893	1483763	1688950
1563593	1317885	1420249	1363859	1238979	1286638
1325525	1588573	1476254	1648209	1354054	1354649
1636119					

Mean1	1463339.219	Mean 2	1451293.194
Standard Dev1	125458.276	Standard Dev2	171932.2264
DF=	61		

t=(1451293.194-1463339.219)/sqrt({171932.226^2/31}+{125458.276^2/32}) = -0.32

t critical value = 1.296

Since the absolute value of the t statistic (0.32) is less than the t critical value, plus the p value associated with t = -0.32 at 61 degrees freedom is 0.7501. Since the p-value is greater than the level of significance (0.1), we fail to reject the null hypothesis. This means there isn't enough evidence to suggest that patients with schizophrenia have less TBV on average than a patient that is considered normal.

P-value computed from online calculator.

9.3.4

A study was conducted that measured the total brain volume (TBV) (in mm^3) of patients that had schizophrenia and patients that are considered normal. Table #9.3.5 contains the TBV of the normal patients and table #9.3.6 contains the TBV of schizophrenia patients ("SOCR data oct2009," 2013). Compute a 90% confidence interval for the difference in TBV of normal patients and patients with Schizophrenia.

9.3.8

The number of cell phones per 100 residents in countries in Europe is given in table #9.3.9 for the year 2010. The number of cell phones per 100 residents in countries of the Americas is given in table #9.3.10 also for the year 2010 ("Population reference bureau," 2013). Find the 98% confidence interval for the different in mean number of cell phones per 100 residents in Europe and the Americas.

Table #9.3.9: Number of Cell Phones per 100 Residents in Europe

100	76	100	130	75	84
112	84	138	133	118	134
126	188	129	93	64	128
124	122	109	121	127	152
96	63	99	95	151	147
123	95	67	67	118	125
110	115	140	115	141	77
98	102	102	112	118	118
54	23	121	126	47	

Table #9.3.10: Number of Cell Phones per 100 Residents in the Americas

158	117	106	159	53	50
78	66	88	92	42	3
150	72	86	113	50	58
70	109	37	32	85	101
75	69	55	115	95	73
86	157	100	119	81	113
87	105	96			

Mean1 = 108.15

Standard Dev1 = 897.90

n1 = 53

Mean2 = 87.205

Standard Dev2 = 1235.904

n2 = 39

df= 39+53-2 = 90

t=2.37

CI= (108.15-87.205)+-2.37*sqrt(897.90/53+1235.904/39) = 20.9450+-16.5275 = (4.4175;37.4725)

Since 0 is not within the interval, there is a significant difference in mean number of cell phones per 100 residents in Europe and the Americas

11.3.2

Levi-Strauss Co manufactures clothing. The quality control department measures weekly values of different suppliers for the percentage difference of waste between the layout on the computer and the actual waste when the clothing is made (called run-up). The data is in table #11.3.3, and there are some negative values because sometimes the supplier is able to layout the pattern better than the computer ("Waste run up," 2013). Do the data show that there is a difference between some of the suppliers? Test at the 1% level.

Table #11.3.3: Run-ups for Different Plants Making Levi Strauss Clothing

Plant 1	Plant 2	Plant 3	Plant 4	Plant 5
1.2	16.4	12.1	11.5	24
10.1	-6	9.7	10.2	-3.7
-2	-11.6	7.4	3.8	8.2
1.5	-1.3	-2.1	8.3	9.2
-3	4	10.1	6.6	-9.3
-0.7	17	4.7	10.2	8
3.2	3.8	4.6	8.8	15.8
2.7	4.3	3.9	2.7	22.3
-3.2	10.4	3.6	5.1	3.1
-1.7	4.2	9.6	11.2	16.8
2.4	8.5	9.8	5.9	11.3
0.3	6.3	6.5	13	12.3
3.5	9	5.7	6.8	16.9
-0.8	7.1	5.1	14.5	
19.4	4.3	3.4	5.2	
2.8	19.7	-0.8	7.3	
13	3	-3.9	7.1	
42.7	7.6	0.9	3.4	
1.4	70.2	1.5	0.7	
3	8.5			
2.4	6			
1.3	2.9			

Mean	4.5227	8.8318	4.8316	7.4895	10.3769		
Standard Dev	10.0320	15.3535	4.4032	3.6571	9.5550		
N	22	22	19	19	13		
	Anova: Single Factor						

	SUMMARY						
	Groups	*Count*	*Sum*	*Average*	*Variance*		
	1.2	21	98.3	4.680952381	105.095619		
	16.4	21	177.9	8.471428571	244.5151429		
	12.1	18	79.7	4.427777778	17.24800654		
	11.5	18	130.8	7.266666667	13.16235294		
	24	12	110.9	9.241666667	81.32083333		
	ANOVA						
	urce of Variat	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
	Between Groups	327.4154841	4	81.85387103	0.8279165192	0.5110488415	3.548191096
	Within Groups	8403.720516	85	98.86730019			
	Total	8731.136	89				

Since p-value is more than 0.01, we accept Ho, therefore there is no evidence between the suppliers.

11.3.4

A study was undertaken to see how accurate food labeling for calories on food that is considered reduced calorie. The group measured the amount of calories for each item of food and then found the percent difference between measured and labeled food, $\frac{(\text{measured} - \text{labeled})}{\text{labeled}} * 100\%$. The group also looked at food that was nationally advertised, regionally distributed, or locally prepared. The data is in table #11.3.5 ("Calories datafile," 2013). Do the data indicate that at least two of the mean percent differences between the three groups are different? Test at the 10% level.

Table #11.3.5: Percent Differences Between Measured and Labeled Food

National Advertised	Regionally Distributed	Locally Prepared
2	41	15
-28	46	60
-6	2	250
8	25	145
6	39	6
-1	16.5	80
10	17	95
13	28	3
15	-3	
-4	14	
-4	34	
-18	42	
10		
5		
3		
-7		
3		
-0.5		
-10		
6		

Mean	0.125	25.125	81.75				
Standard Dev	10.52050004	16.07387067	83.96895685				
N	20	12	8				
	Anova: Single						

Factor						
SUMMARY						
Groups	*Count*	*Sum*	*Average*	*Variance*		
National Advertised	20	2.5	0.125	110.6809211		
Regionally Distributed	12	301.5	25.125	258.3693182		
Locally Prepared	8	654	81.75	7050.785714		
ANOVA						
urce of Variat	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	38095.9	2	19047.95	12.9791466	0.00005361081359	2.45201433
Within Groups	54300.5	37	1467.581081			
Total	92396.4	39				

Since the p-value is less than 0.1, e reject Ho. therefore, there is evidence that at least two of the mean percent differences between the three groups are different.