Sandy Chan
PA 706
Homework #4
Due April 24, 2008

One Sample

1.

6	AVG (x bar) = 5.4
4	STDEV (s) = 2.07
7	
7	
1	
8	Ho: $\mu \leq 5$
4	H1: $\mu > 5$
5	one-tail upper t-test
5	ά = 0.05
7	dF = n - 1 = 9
n=10	T-crit = 2.26

T-statistic = $(5.4-5) / (2.07/\sqrt{10}) = 0.61$

Conclusion

Statistical language: With 95% confidence, we fail to reject the null hypothesis that the small town's average number of citations is less than or equal to that of the neighboring towns' average of 5 per month.

Plain English: We are 95% confident that the small town's average number of citations is close enough to that of other towns. It has not exceeded the average of neighboring towns.

2. Ho: $\mu \geq 2.5$
H1: $\mu < 2.5$ one-tail lower t-test, ά = 0.05, dF = 409, T-crit = -1.645, T-stat = -0.712

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
rating for environmental protection	410	2.47	.763	.038

One-Sample Test

	Test Value = 2.5					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
rating for environmental protection	-.712	409	.477	-.027	-.10	.05

Conclusion
Statistical language: With 95% confidence, we fail to reject the null hypothesis that the population rating for protecting the environment is at least 2.5.
Plain English: We are 95% confident that the mean rating for protecting the environment is at least 2.5.

The 95% confidence interval for the mean of JOBENV:

Descriptives

			Statistic	Std. Error
rating for environmental protection	Mean		2.47	.038
	95% Confidence Interval for Mean	Lower Bound	2.40	
		Upper Bound	2.55	
	5% Trimmed Mean		2.47	
	Median		2.00	
	Variance		.582	
	Std. Deviation		.763	
	Minimum		1	
	Maximum		4	
	Range		3	
	Interquartile Range		1	
	Skewness		.323	.121
	Kurtosis		-.309	.240

Lower bound: 2.40 Upper bound: 2.55
This means that we are 95% confident that the mean rating for protecting the environment falls between 2.4 and 2.55.

3. Ho: $\mu = 2.5$
H1: $\mu \neq 2.5$
Two-tail t-test, $\alpha = 0.05/2 = 0.025$
T-crit = 1.96, T-stat = -0.712 (± 0.712)

Conclusion
Statistical language: With 95% confidence, we fail to reject the null hypothesis that the population rating for protecting the environment is equal to 2.5.
Plain English: We are 95% confident that the mean rating for protecting the environment is neutral.

The 95% confidence interval for the mean of JOBENV-Lower bound: 2.40, Upper bound: 2.55
This means that we are 95% confident that the mean rating for protecting the environment is between 2.4 and 2.55.

Two samples

4.

Town 1		Town 2	
6	AVG (x1 bar) = 5.4	3	AVG (x2 bar) = 5.1
4	STDEV (s1) = 2.07	5	STDEV (s2) = 1.85
7		6	
7		5	
1		7	Ho: $\mu 1 \leq \mu 2$
8		4	H1: $\mu 1 > \mu 2$
4		9	One-tail upper
5		4	ά = 0.10
5		5	dF = n1+n2 – 2 = 18
7		3	T-crit = 1.33
n1 = 10		n2 = 10	

sp-squared) = [9(2.07squared) + 9(1.85squared)] / (9+9) = 3.85

T-stat = 0.3 /√[3.85 (1/10+1/10)] = 0.34

Conclusion

Statistical language: With 90% confidence, we fail to reject the null hypothesis that the average number of citations is less than or equal to that of the neighboring town's average.
Plain English: We are 90% confident that there is no statistical difference between the average number of citations between the two towns.

5. ***Part 1 - Excel***

Ho: μ-men = μ-women
H1: μ-men ≠ μ-women
two-tail
ά = 0.05/2 = 0.025, T-crit = 1.96, T-stat = 0.619

t-Test: Two-Sample Assuming Unequal Variances

	Men	*Women*
Mean	2.495327103	2.44898
Variance	0.683076653	0.47431
Observations	214	196
Hypothesized Mean Difference		
df	404	
t Stat	0.618688826	
P(T<=t) one-tail	0.268235022	
t Critical one-tail	1.64863405	
P(T<=t) two-tail	0.536470044	
t Critical two-tail	1.965853199	

Conclusion

Statistical language: With 95% confidence, we fail to reject the null hypothesis that mean rating by men and by women for the county's performance in protecting the environment are equal.
Plain English: We are 95% confident that the men and women rate the county's performance in protecting the environment similarly.

Part 2 –Excel - # of data points in each range were different – those with missing values for either jobenv or jobserv were omitted from the pool of responses used for the test. This is what SPSS ended up doing also, as shown by the number of observations used in problem #6.

Ho: μ-env - μ-serv = 0

H1: μ-env - μ-serv ≠ 0

Two-tail, ά = 0.05/2 = 0.025, T-crit = 1.96, T-stat = -5.6 (± 5.6)

t-Test: Paired Two Sample for Means

	JOBENV	*JOBSERV*
Mean	2.197468354	2.465822785
Variance	0.544663625	0.569260425
Observations	395	395
Pearson Correlation	0.185358992	
Hypothesized Mean Difference		
df	394	
t Stat	-5.598663254	
P(T<=t) one-tail	2.02788E-08	
t Critical one-tail	1.648730227	
P(T<=t) two-tail	4.05577E-08	
t Critical two-tail	1.966003126	

Conclusion

Statistical language: With 95% confidence, we reject the null hypothesis that the mean rating for the county's performance in protecting the environment and for the county's provision of public services is the same.

Plain English: We are 95% confident that the mean rating for the county's performance in protecting the environment is statistically different from the mean rating for the county's provision of public services.

6. ***Part 1 – SPSS*** Same set-up and conclusion as #5.

Group Statistics

	sex	N	Mean	Std. Deviation	Std. Error Mean
rating for environmental protection	male	214	2.50	.826	.056
	female	196	2.45	.689	.049

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
rating for environmental protection	Equal variances assumed	9.076	.003	.614	408	.540	.046	.076	-.102	.195
	Equal variances not assumed			.619	404.461	.536	.046	.075	-.101	.194

Because the t-statistics are similar for when equal variances are assumed and when not assumed, 0.614 and 0.619 respectively, we can run the t-test assuming equal variances.

Part 2 – SPSS

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	rating for environmental protection	2.47	395	.754	.038
	rating for overall public service	2.20	395	.738	.037

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	rating for environmental protection - rating for overall public service	.268	.953	.048	.174	.363	5.599	394	.000

More than two samples

7. Excel

Ho: μnever = μseldom = μsometimes = μoften

H1: not all μk are equal

ά = 0.05, F-crit = 2.62, F-stat = 4.52

Anova: Single Factor

SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
Never	78	189	2.42308	0.79271
Seldom	96	214	2.22917	0.51535
Sometimes	156	348	2.23077	0.43672
Often	145	298	2.05517	0.48305

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	7.11859	3	2.37286	4.52024	0.00388	2.62384
Within Groups	247.248	471	0.52494			
Total	254.366	474				

Conclusion
Statistical language: With 95% confidence, we reject the null hypothesis that the mean rating for the county's provision of public services among different use groups is equal.
Plain English: We're 95% confident that at least one of the use groups rates the county on its public services differently from the other use groups.

8. SPSS Same set-up and conclusion as #7 but with post hoc test.

ANOVA

rating for overall public service

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	7.119	3	2.373	4.520	.004
Within Groups	247.248	471	.525		
Total	254.366	474			

Post hoc test

Multiple Comparisons

rating for overall public service

LSD

(I) use services	(J) use services	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
never	seldom	.194	.110	.080	-.02	.41
	sometimes	.192	.100	.056	.00	.39
	often	.368*	.102	.000	.17	.57
seldom	never	-.194	.110	.080	-.41	.02
	sometimes	-.002	.094	.986	-.19	.18
	often	.174	.095	.069	-.01	.36
sometimes	never	-.192	.100	.056	-.39	.01
	seldom	.002	.094	.986	-.18	.19
	often	.176*	.084	.036	.01	.34
often	never	-.368*	.102	.000	-.57	-.17
	seldom	-.174	.095	.069	-.36	.01
	sometimes	-.176*	.084	.036	-.34	-.01

*. The mean difference is significant at the 0.05 level.

Often-users are statistically different from **Never**-users. **Often**-users are also statistically different from **Sometimes**-users.

9. Excel

Ho: Sex and educational level are independent

H1: Sex and educational level are not independent

ά = 0.05, dF = 5, chi-sq-crit = 11.07

observed X-tab	8th grade or less	Some high school	High school graduate	Some college/2-yr degree	College degree	Post-grad work/grad degree	**Total**
Male	4	12	47	52	77	58	250
Female	2	17	52	70	72	41	254
Total	6	29	99	122	149	99	504

expected X-tab	8th grade or less	Some high school	High school graduate	Some college/2-yr degree	College degree	Post-grad work/grad degree	**Total**
Male	2.98	14.38	49.11	60.52	73.91	49.11	250
Female	3.02	14.62	49.89	61.48	75.09	49.89	254
Total	6	29	99	122	149	99	504

chi-sq terms	0.35	0.39	0.09	1.2	0.13	1.61
	0.34	0.39	0.09	1.18	0.13	1.58
chi-sq stat (sum of terms =	**7.48**					

Conclusion

Statistical language: With 95% confidence, we fail to reject the null hypothesis that sex and educational level are independent.

Plain English: We are 95% confident that sex and educational level are independent.

10. SPSS – Same set-up and conclusion as #9.

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
sex * highest level of education	504	99.6%	2	.4%	506	100.0%

sex * highest level of education Crosstabulation

			highest level of education						
			8th grade or less	some high school	high school graduate	some college/2-year degree	college degree	post graduate work/graduate degree	Total
sex	male	Count	4	12	47	52	77	58	250
		Expected Count	3.0	14.4	49.1	60.5	73.9	49.1	250.0
	female	Count	2	17	52	70	72	41	254
		Expected Count	3.0	14.6	49.9	61.5	75.1	49.9	254.0
	Total	Count	6	29	99	122	149	99	504
		Expected Count	6.0	29.0	99.0	122.0	149.0	99.0	504.0

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	7.493[a]	5	.186
Likelihood Ratio	7.534	5	.184
Linear-by-Linear Association	3.220	1	.073
N of Valid Cases	504		

a. 2 cells (16.7%) have expected count less than 5. The minimum expected count is 2.98.