Social Studies 201
Winter 2004

Answers to Problem Set 5

1. **Shift work and health**

   (a) Let p be the true proportion of all married men working evening shift who have problems with partner. The sample in Table 1 of the problem set indicates that 36% or $\hat{p} = 0.36$ of the married men in the sample who work evening shift report problems with their partner. Since $n = 93$ is large ($n = 93 > 5/0.36 = 13.9$), the sample proportion $\hat{p}$ is normally distributed with mean p and standard deviation $\sqrt{pq/n}$. For the 90% interval estimate, $Z = \pm 1.645$. By using $p = q = 0.5$, the widest possible interval for any of the proportions is obtained. Alternatively, $\hat{p} = 0.36$ and $\hat{q} = 1 - \hat{p} = 1 - 0.36 = 0.64$ could be used for p and q in the estimate of the standard deviation of $\hat{p}$. The interval estimates are:

$$\begin{aligned}
\hat{p} \pm Z\sqrt{\frac{pq}{n}} &= \hat{p} \pm 1.645\sqrt{\frac{0.5 \times 0.5}{93}} \\
&= \hat{p} \pm 1.645\sqrt{0.002688} \\
&= \hat{p} \pm (1.645 \times 0.0518) \\
&= \hat{p} \pm 0.0853 \\
&= 0.36 \pm 0.0853
\end{aligned}$$

   or from 0.27 to 0.45. The 90% interval estimate for married men is from 27% to 45%.

   For the married women, the method is the same and with a sample size of $n = 99$, again $\hat{p}$ is normally distributed and $Z = \pm 1.645$ for 90% confidence. $\hat{p} = 0.29$ and $\hat{q} = 1 - \hat{p} = 1 - 0.29 = 0.71$, and these could be used to construct the interval, although $p = q = 0.5$ is used here. The interval is

$$\hat{p} \pm Z\sqrt{\frac{pq}{n}} = \hat{p} \pm 1.645\sqrt{\frac{0.5 \times 0.5}{99}}$$

$$\begin{aligned} &= \hat{p} \pm 1.645\sqrt{0.002525} \\ &= \hat{p} \pm (1.645 \times 0.0503) \\ &= \hat{p} \pm 0.08266 \\ &= 0.29 \pm 0.08266 \end{aligned}$$

or from 0.21 to 0.37. The 90% interval estimate for the proportion of all married women who work evening shift and have problems with partner is from 21% to 37%.

(b) For estimation of any proportion, the sample size required to achieve the specified accuracy E is

$$n = \left(\frac{Z}{E}\right)^2 pq.$$

For 95% confidence, $Z = 1.96$. From the question, the precision required is ± 4 percentage points so $E = 0.04$, and the largest possible sample size for this Z and E are given by $p = q = 0.5$. The required sample size is

$$n = \left(\frac{Z}{E}\right)^2 pq = \left(\frac{1.96}{0.04}\right)^2 \times 0.5 \times 0.5 = 49^2 \times 0.25 = 600.25$$

The required sample size is $n = 601$.

(c) i. Let p be the true proportion of all men working evening shift who have high stress. The question is to test whether this proportion exceeds 0.33, or 33%, the per cent of men who report high stress but work regular shift. Since the question uses "exceeds," this is a one-tailed test in the positive direction. The null hypothesis is an equality so the hypotheses are:

$$H_0 \ : \ p = 0.33$$

$$H_1 \ : \ p > 0.33$$

The sample size is $n = 137$ and the test statistic is $\hat{p}$, where $\hat{p} = 0.44$, $\hat{q} = 1 - \hat{p} = 1 - 0.44 = 0.56$. To ensure this is a large sample size, n divided by the smaller of p or q gives $5/0.33 = 15.1$ and $n = 137 > 15.1$ so this is a large sample size and $\hat{p}$ has a normal distribution with mean p and standard deviation $\sqrt{pq/n}$. At $\alpha = 0.05$ significance, with a one-tailed test, the critical $Z = 1.645$. When $Z > +1.645$, the null hypothesis can be rejected but when $Z <= 1.645$, the null hypothesis is not rejected.

$$\begin{aligned} Z &= \frac{\hat{p} - p}{\sqrt{pq/n}} \\ &= \frac{0.44 - 0.33}{\sqrt{(0.33 \times 0.67)/137}} \\ &= \frac{0.11}{0.0402} \\ &= 2.738 \end{aligned}$$

Since this Z-value is greater than 1.645, the the null hypothesis is rejected at the 0.05 level of significance. At the 0.05 level of significance the data provide evidence that the proportion of men reporting high stress is greater for those working evening shift than for those working regular shift.

ii. For women, the procedure is the same, with p being the proportion of all women working evening shift who have high stress. This is a two-tailed test since the question is whether the two proportions differ from each other – no direction is stated in the question.

$$\begin{aligned} H_0 &: p = 0.43 \\ H_1 &: p \neq 0.43 \end{aligned}$$

and $\hat{p} = 0.41$.

$$\begin{aligned} Z &= \frac{\hat{p} - p}{\sqrt{pq/n}} \\ &= \frac{0.41 - 0.43}{\sqrt{(0.43 \times 0.57)/149}} \\ &= \frac{-0.02}{0.0406} \\ &= 0.493 \end{aligned}$$

Since this Z-value is between the critical values of -1.96 and $+1.96$, the null hypothesis is not rejected at the 0.05 level of significance. At the 0.05 level of significance the data provide no evidence that the proportion of women reporting high stress differs for those working evening shift and for those working regular shift.

(d) From a., the interval estimates for percentage of each of men and womenn, working evening shift reporting problems with partner, do not overlap the corresponding percentages for those working regular shift. For women though, there is close to overlap. This leads to the conclusion that evening shift is more associated with having problems with partner, for both men and women. This evidence supports the first statement in the quote, that is, that men working evening shift report more problems.

From c., the test demonstrates that larger percentages of men working evening shift report high stress than for men working regular shift. Again, this supports the contention in the first part of the statement.

For women, the claim is that evening shift was not so much associated with these problems as it was for men. While the hypothesis test in c. supports this contention, the interval estimate of a. comes close to demonstrating that married women working evening shift have more problems with partner than do married women working regular shift. For both high stress and problems with partner, the samples show greater per cent of problems for women working evening shift than for women working regular shift. Even though the test and interval estimate do not prove greater problems, the second half of the quote is questionable – there is certainly no evidence that problems are less for women working evening shift.

2. **Television and internet use**

Part (i). Let μ be the true weekly hours watching television for all Saskatchewan residents aged 15-24. The sample size is $n = 180$ and the test statistic is $\bar{X}$. Given this large sample size, $\bar{X}$ has a normal distribution with mean μ and standard deviation $\sigma/\sqrt{n}$, where the sample standard deviation s is used as an estimate of σ. No significance level is stated in the question, so the $\alpha = 0.05$ level can be used. The test is a two-tailed test since the researcher gives no suggestion that the mean might be greater than or less than 11 hours per week. For a normal distribution, a two tailed test at the $\alpha = 0.01$ significance level has a critical value of $Z = 2.575$. H_0 is rejected for all $Z < -2.575$ or greater than $Z > +2.575$. The hypotheses are:

$$H_0 : \mu = 11$$

$$H_1 : \mu \neq 11$$

and the test gives a Z-value of

$$\begin{aligned} Z &= \frac{\bar{X} - \mu}{s/\sqrt{n}} \\ &= \frac{13.46 - 11}{10.18/\sqrt{180}} \\ &= \frac{2.46}{10.18/13.416} \\ &= \frac{2.46}{0.758} \\ &= 3.242 \end{aligned}$$

Since this Z-value is greater than +2.575, the the null hypothesis is rejected at the 0.01 level of significance. At the 0.01 level of significance this sample provides evidence that the mean weekly hours of watching television for Saskatchewan residents aged 15-24 is not equal to 11 hours.

Part (ii). For this part, let μ be the true weekly hours using the internet at work for all Saskatchewan residents aged 15-24. The hypotheses are

$$H_0 : \mu = 11$$

$$H_1 : \mu \neq 11$$

where the research hypothesis is again a two-directional hypothesis since no direction is indicated in the question. The statistic used to test the hypotheses is $\bar{X}$. Since the sample size is only $n = 15$ for this sample, the t-distribution is used. In order to use the t-distribution, the distribution of the population should be close to normally distributed – while there is no assurance of this, hours using internet at work is likely to have a reasonably symmetric distribution around the mean, so may not be too far from normally distributed. The sample mean $\bar{X}$ has a t-distribution with mean μ, standard deviation $s/\sqrt{n}$, and $n - 1 = 15 - 1 = 14$ degrees of freedom. For 14 d.f., $\alpha = 0.01$ significance, and a two-tailed test, the critical $t = 2.977$ and H_0 is rejected for all $t < -2.977$ or $t > +2.977$.

$$\begin{aligned} t &= \frac{\bar{X} - \mu}{s/\sqrt{n}} \\ &= \frac{7.60 - 11}{10.80/\sqrt{15}} \\ &= \frac{-3.40}{10.80/3.873} \\ &= \frac{-3.40}{2.789} \\ &= -0.861 \end{aligned}$$

This t-value is between -2.977 and $+2.977$, so the the null hypothesis cannot be rejected at the 0.01 level of significance. At the 0.01 level of significance the sample does not provide sufficient evidence to conclude that the researcher's claim that mean hours of internet use at work equals 11 hours is incorrect.

For (i), the large sample size and the difference of 2.46 hours (13.46 - 11) leads to a strong rejection of the hypothesis that mean hours of television use equal 11. That is, the Z-value is 3.242, leading to a decisive rejection of the null hypothesis. for (ii), while the difference between the sample mean of 7.60 hours and the hypothesized mean of 11 hours is larger (3.40) than for (i), the smaller sample size makes it more difficult to reject the null hypothesis. The t-value here was only -0.861, not large enough to reject the hypothesis that the mean hours used internet at work differs from 11. The two samples differ a lot in sample size and it is this difference tht is the primary factor in leading to the different conclusion.

3. **Years of education**.

(a) Let μ be the true years of education for all Saskatchewan adults. The sample size is $n = 3,862$ and the test statistic is $\bar{X} = 12.44$. Given this large sample size, by the Central Limit Theorem, $\bar{X}$ has a normal distribution with mean μ and standard deviation $\sigma/\sqrt{n}$, where the sample standard deviation s is used as an estimate of σ. The significance level is $\alpha = 0.02$ and the test is a two-tailed test since no direction is indicated in the question. For a normal distribution, a two tailed test at the $\alpha = 0.02$ significance level has a critical value of $Z = 2.33$. and H_0 is rejected for all $Z < -2.33$ or greater than $Z > +2.33$. the hypotheses are:

$$H_0 \ : \ \mu = 12$$

$$H_1 \ : \ \mu \neq 12$$

and the test gives a Z-value of

$$\begin{aligned} Z &= \frac{\bar{X} - \mu}{s/\sqrt{n}} \\ &= \frac{12.44 - 12}{3.14/\sqrt{3,862}} \\ &= \frac{0.44}{3.14/62.145} \\ &= \frac{0.44}{0.0505} \\ &= 8.708 \end{aligned}$$

Since this Z-value is much greater than +2.33, the the null hypothesis is very strongly rejected at the 0.02 level of significance. At the 0.02 level of significance this sample provides strong evidence that the mean years of education for Saskatchewan adults differs from 12 years.

(b) For this part, let μ be the true mean education level for all Saskatchewan adults. The hypotheses are

$$H_0 \ : \ \mu = 12$$

$$H_1 \ : \ \mu > 12$$

where the research hypothesis is a one-directional hypothesis since the question suggests that the mean education level may exceed 12 years. The statistic used to test the hypotheses is $\bar{X}$. Since the sample size is only $n = 20$ for this sample, the t-distribution must be used. The sample mean $\bar{X}$ has a t-distribution with mean μ, standard deviation $s/\sqrt{n}$, and $n - 1 = 20 - 1 = 19$ degrees of freedom. For 19 d.f., $\alpha = 0.05$ significance, and a one-tailed t-test, the critical $t = 1.729$ and H_0 is rejected for all $t > +1.729$.

$$\begin{aligned} t &= \frac{\bar{X} - \mu}{s/\sqrt{n}} \\ &= \frac{13.73 - 12}{2.80/\sqrt{20}} \\ &= \frac{1.73}{2.80/4.472} \\ &= \frac{1.73}{0.626} \\ &= 2.763 \end{aligned}$$

Since this t-value of 2.763 exceeds +1.729, the the null hypothesis is rejected at the 0.05 level of significance. Even from this small sample, there is evidence that the mean years of education for all Saskatchewan adults exceeds 12 years.

(c) In a. the sample mean differed by only 0.44 years from the hypothesized mean of 12 years, yet this proved to be a large difference statistically, leading to rejection of the hypothesis that the overall mean is 12 years. For b., the result is the same – a larger difference of 1.73 years between the sample and hypothesized mean exists and, even though the sample size is small, this larger difference of 1.73 years is sufficient to reject the null hypothesis.

For each part, there is a possibility of Type I error, rejecting the null hypothesis that $\mu = 12$, when this is not true. The chance of this is low that since the significance level is 0.02 in part a. and 0.05 in part b., so the probability of Type I error is less than 0.05 in either case. For part a., given the extremely large Z-value of 8.708, the probability of rejecting H_0 in error is much less than 0.02.

For part b. the probability of Type I error is larger, up to 0.05, although again the sample mean is well within the critical region, so the probability of rejecting the null hypothesis in error is less than 0.05.

Question 4. Computer problem for Problem Set 5

1. a. Religious and spiritual activities and household income

Part i.

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
RELHOURS RELIGIOUS ACTIVITIES HOURS	556	1.10	2.59	.11

One-Sample Test

	Test Value = 1.04					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
RELHOURS RELIGIOUS ACTIVITIES HOURS	.570	555	.569	6.25E-02	-.15	.28

Part ii.

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
INC income in thousands of dollars	617	65.46	43.89	1.77

One-Sample Test

	Test Value = 55.107					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
INC income in thousands of dollars	5.861	616	.000	10.35	6.89	13.82

b. Explanation.

For i. the true mean of weekly hours all undergraduates spend at religious and spiritual activities is hypothesized to be 1.04 hours, with the alternative hypothesis being that it is not 1.04, or that it differs from the mean for all Saskatchewan. Since n=556 is large, the distribution of the sample mean is normal (by Central Limit Theorem). If significance level 0.05 is selected, then for a two-tailed test, the region of rejection is all Z values less than –1.96 or greater than +1.96. From the SPSS table, the sample mean for undergraduates is 1.10 hours and the associated Z-value (t becomes Z once n>30) is 0.570. This Z is not in the region of rejection. As a result, the conclusion here is that, on average, undergraduates spend much the same amount of time at religious and spiritual activities as do all Saskatchewan adults, that is just over one hour per week.

For part ii, the test is that the true mean household income for all University of Regina undergraduates is $55,107, the same as for all Saskatchewan households. The alternative or research hypothesis is that the mean income of households of U of R undergraduates is greater than this Saskatchewan average. Again the sample is large and with a one-tailed test at the 0.01 level, the null hypothesis is rejected for all Z>+2.33. For this sample, Z=5.861>2.33, so the null hypothesis can be very decisively rejected. The sample mean income reported by undergraduates was $65,460, considerably greater than the Saskatchewan average, and statistically significantly greater than this average, at the 0.01 level or less.

We conclude that undergraduates come from better off economic background than the Saskatchewan average but spend similar time at religious and spiritual activities to that of all Saskatchewan adults.

2. Study hours

a. The SPSS tables for the three tests are as follows:

i. Test of mean study hours equal to 15.

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
STHOURS Study Hours	668	16.61	11.85	.46

One-Sample Test

	Test Value = 15					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference Lower	Upper
STHOURS Study Hours	3.502	667	.000	1.61	.71	2.51

ii. Test of mean study hours equal to 16

One-Sample Test

	Test Value = 16					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference Lower	Upper
STHOURS Study Hours	1.321	667	.187	.61	-.29	1.51

iii. Test of mean study hours equal to 17

One-Sample Test

	Test Value = 17					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference Lower	Upper
STHOURS Study Hours	-.860	667	.390	-.39	-1.29	.51

b. Description of tests on study hours

The parameter or population value being tested here is the true mean weekly hours spent studying by all University of Regina undergraduates in Fall 1998. The null hypothesis for each test is that this mean is equal to the specified value, that is, 15, 16, or 17. The sample size is n=667 for each test, a large sample size so that the sample mean is normally distributed (by the Central Limit Theorem). The large n also means that the t-value in the test is really a Z-value, since t-values approach Z-values as n becomes large. The sample mean is 16.61 hours and the standard deviation is 11.85 hours. This is the same for each test – what differs among the tests is the claim to be tested.

For (i), the alternative is that the true mean study hours exceeds 15 hours weekly. Since the suspicion is that the true mean exceeds 15 hours, this is a one-tailed test. At the 0.01 level of significance, the null hypothesis is rejected for all Z>2.33. From the SPSS table, Z=3.502 so the hypothesis that the mean study hours exceeds 15 is rejected at the 0.01 level of significance. This is a small level of significance, so rejection of the claim means that there is considerable certainty that mean study hours exceed 15 hours per week.

For (ii) and (iii), there are two claims that cannot be rejected at the 0.10 level of significance. These are both claims of difference, so are two-tailed tests. At 0.10 significance, the critical Z-value for a two-tailed test is at plus or minus 1.645. Since the Z-values are 1.321 and –0.860, neither the claim that study hours are 16, nor the claim they are 17, can be rejected. This leaves the researcher in some uncertainty, and both of these conclusions are examples of Type II error. That is, both claims cannot be true, but neither can be rejected. It is very likely that the true mean study hours are not exactly 16 or 17 hours, but what these results demonstrate is that mean study hours may not be all that different than 16 or 17 hours per week.

3. Interval estimate for mean weekly study hours

Descriptives

			Statistic	Std. Error
STHOURS Study Hours	Mean		16.61	.46
	95% Confidence Interval for Mean	Lower Bound	15.71	
		Upper Bound	17.51	
	5% Trimmed Mean		15.57	
	Median		15.00	
	Variance		140.409	
	Std. Deviation		11.85	
	Minimum		0	
	Maximum		100	
	Range		100	
	Interquartile Range		12.00	
	Skewness		1.689	.095
	Kurtosis		5.212	.189

b. The 95% interval estimate from the *Explore* procedure is from 15.71 to 17.51 hours per week, while the sample mean and standard deviation again are 16.61 and 11.85 hours. This interval is consistent with the findings from question 2 in two respects. First, the interval does not cross 15 hours (the lower end of the interval is 15.71 hours), so just as the hypothesis test led us to conclude that mean study hours exceed 15, so does the interval estimate. Second, the interval (from 15.71 to 17.51 hours) contains both 16 and 16 hours, so the interval provides evidence that the true mean may be somewhere near 16 or 17 hours. This was the same conclusion reached in parts (ii) and (iii) of question 2.

4. Cross-classification of V4 by sex and chi-square test

V4 Gay and Lesbians Married * SEX SEX OF RESPONDENT Crosstabulation

			SEX SEX OF RESPONDENT		
			1 MALE	2 FEMALE	Total
V4 Gay and Lesbians Married	1 Strongly Disagree	Count	71	70	141
		Expected Count	52.7	88.3	141.0
	2	Count	28	58	86
		Expected Count	32.2	53.8	86.0
	3	Count	70	113	183
		Expected Count	68.5	114.5	183.0
	4	Count	53	108	161
		Expected Count	60.2	100.8	161.0
	5 Strongly Agree	Count	38	86	124
		Expected Count	46.4	77.6	124.0
Total		Count	260	435	695
		Expected Count	260.0	435.0	695.0

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	14.820[a]	4	.005
Likelihood Ratio	14.599	4	.006
Linear-by-Linear Association	10.376	1	.001
N of Valid Cases	695		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 32.17.

b. Findings from table and chi-square test

The hypothesis being examined here is that there is no relationship between sex and opinion to statement concerning treating gays and lesbians as married. That is, this null hypothesis is that males and females respond much the same on this issue. The alternative hypothesis is that there is a relationship between sex and response on V4, that is, males and females have different views on this issue. This is a chi-square test, and the assumption that all cells have 5 or more expected cases is met. There are 4 degrees of freedom [2 columns and 5 rows so df = (5-1)x(2-1)=4]. At the 0.01 level of significance, and 4 degrees of freedom, the critical chi-square value is   . Since the chi-square value from the sample, 14.820 >   , the null hypothesis is rejected and the alternative hypothesis is accepted. That is, there is reasonably strong evidence that undergraduate males and females have different views on this issue.

IN terms of how males and females differ, there is little difference at the middle, or neutral response, where there are about the same number of each of males and females as expected. But at the strongly disagree and strongly agree end, there are some major differences. Only 53 males were expected to strongly disagree (under the assumption of no relation between sex and V4) but 71 strongly disagree. In contrast, there are 18 fewer females who strongly disagree (70-88) than expected. So males are more likely to strongly disagree and females less likely to strongly disagree. At the strongly agree end of opinions, the situation is reversed, with 8 more females and 8 fewer males than expected.

What these results demonstrate is that males and females have different views on the issue of accepting gays and lesbians as married, with males being less accepting of this and females being more accepting.

Last edited April 6, 2004