MATHEMATICS 201-510-LW **Name: ______________________**

Business Statistics

Martin Huard

Thursday December 4, 2008

Test #4

SOLUTIONS

Complete solutions are expected (formulas used, ect…)

Question 1 (11 points)

A systems specialist has studied the workflow of clerks all doing the same inventory work. Based on this study, she designed a new work flow layout for the inventory system. To compare average production for the old and new methods, a random sample of five clerks was used. The average production rate (number of inventory items processed per shift) for each clerk was measured both before and after the new system was introduced. The results are show below. Test the claim that the new system increases the mean number of items processed per shift. Use a 5 % level of significance. Use the *p*-value approach. Assume that the production rates are normally distributed.

Old	116	108	93	88	119
New	123	114	112	82	127
$d = N - O$	7	6	19	-6	8

Step 1 Assumptions: The sampled populations are normally distributed

Step 2 $H_0:\ \mu_d = 0$

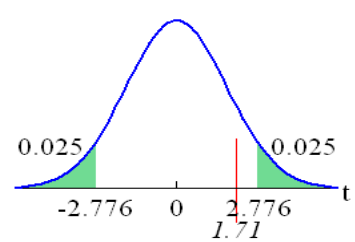
$H_A:\ \mu_d > 0$

Step 3 a) Test statistic: t with $df = n-1 = 4$

b) Right-tailed test with $\alpha = 0.05$

Step 4 a) $t = \dfrac{\bar{d} - \mu_d}{\frac{s_d}{\sqrt{n}}} = \dfrac{6.8 - 0}{\frac{8.871}{\sqrt{5}}} = 1.71$

b) $0.073 < p - \text{value} < 0.082$



Step 5 a) $p\text{-value} > 0.073 > \alpha = 0.05$

b) Fail to reject H_0.

$\therefore$ There is not sufficient evidence at the 5% level of significance to conclude that the system increases the mean number of items processed per shift.

Question 2 (11 points)

How long do batteries last on a camping trip? A random sample of 42 small camp flashlights were installed with brand I batteries and left on until the batteries failed. The sample mean lifetime was 9.8 hours with sample standard deviation 2.2 hours. Another random sample of 36 small flashlights of the same model were installed with brand II batteries and left on until the batteries failed. The sample mean of lifetimes was 8.1 hours with sample standard deviation 3.5 hours. Can you conclude, at the 5% level of significance, that there is a difference in the lifetime of the two batteries? Use the classical approach.

Step 1 Assumptions: $n_I = 42 \geq 30$, $n_{II} = 36 \geq 30$

The samples are independent

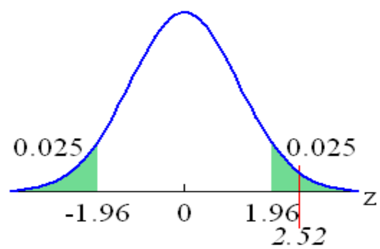
Step 2 $H_O: \mu_I - \mu_{II} = 0$

$H_A: \mu_I - \mu_{II} \neq 0$

Step 3 a) Test statistic: z

b) Two-tailed test with $\alpha = 0.05$

c) $z_{\frac{\alpha}{2}} = z_{0.025} = 1.96$



Step 4 $$z = \frac{(\bar{x}_B - \bar{x}_A) - (\mu_B - \mu_A)}{\sqrt{\frac{s_B^2}{n_B} + \frac{s_A^2}{n_A}}} = \frac{9.8 - 8.1}{\sqrt{\frac{2.2^2}{42} + \frac{3.5^2}{36}}} = 2.52$$

Step 5 a) z is in the critical region

b) Reject H_o.

∴ There is sufficient evidence at the 5% level of significance to conclude that there is a difference in the lifetime of the two batteries.

Question 3 (11 points)

A random sample of 100 Canadians was selected where each was asked whether or not they plan to spend less during the holidays this year due to the financial crisis. Can you conclude, at the 5% level of significance, that whether or not a person will spend less during the holidays is dependent of gender? Use the classical approach.

	Spend Less During the Holidays? Yes	No	*Total*
Men	32 (36)	13 (9)	45
Women	48 (44)	7 (11)	55
Total	80	20	100

Step 1 Assumptions: The classes are all inclusive and mutually exclusive

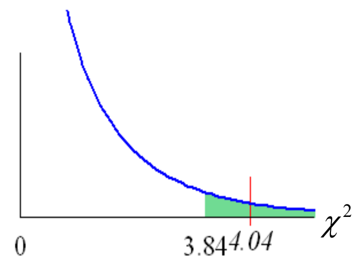
Step 2 H_o: Whether or not a person will spend less during the holidays is independent of gender.

H_A: Whether or not a person will spend less during the holidays is not independent of gender.

Step 3 a) Test statistic: χ^2 with $df = (1)(1) = 1$

b) Right-tailed test with $\alpha = 0.025$

c) $\chi^2_{(df,\alpha)} = \chi^2_{(1,0.05)} = 3.84$



Step 4

$$\chi^2 = \sum \frac{(O-E)^2}{E}$$

$$= \frac{(32-36)^2}{36} + \frac{(13-9)^2}{9} + \frac{(48-44)^2}{44} + \frac{(7-11)^2}{11}$$

$$= 4.04$$

Step 5 a) χ^2 is in the critical region

b) Reject H_o.

∴ There is sufficient evidence at the 5% level of significance to conclude that whether or not a person will spend less during the holidays is dependent of gender.

Question 4 (17 points)

To observe the relationship between the temperature in Quebec and in Montreal, a random sample of 10 summer days was taken, where the noon temperature was recorded both in Montreal and in Quebec. Here are the results (in degrees Celsius).

Montreal	22	26	33	27	25	35	24	29	21	27
Quebec	21	28	30	25	24	32	20	27	19	26

The following information was calculated from this.

$$SS_x = 178.9 \quad SS_y = 165.6 \quad SS_{xy} = 160.2 \quad \bar{x} = 26.9 \quad \bar{y} = 25.2$$

The equation of the regression line was found to be

$$y = 1.112 + 0.895x$$

and standard error of estimate S_e

$$S_e = 1.664$$

a) Find a 95% confidence interval for the slope β of the regression line.

Step 1 Assumptions: $y_i \sim N\left(\alpha + \beta x_i, \sigma_e^2\right)$

Step 2 a) Test statistic: t with $df = n - 2 = 10 - 2 = 8$

b) Level of confidence: $1 - \alpha = 0.95$ or $\alpha = 0.05$

Step 3 Point estimate: $b = 0.895$

Step 4 a) $t_{\left(df, \frac{\alpha}{2}\right)} = t_{(8, 0.025)} = 2.306$

b) $E = t_{\left(df, \frac{\alpha}{2}\right)} \dfrac{s_e}{\sqrt{SS_x}} = 2.306 \dfrac{1.664}{\sqrt{178.9}} = 0.287$

c) $b - E < \beta < b + E$

$0.895 - 0.287 < \beta < 0.895 - 0.287$

$0.609 < \beta < 1.182$

Step 5 The 95% confidence interval for the regression coefficient β is 0.609 to 1.182.

b) The coefficient of correlation r was found to be 0.931. At the 5% level of significance, is there sufficient evidence to conclude that there is a correlation between the temperature in Montreal and in Quebec? Use the p-value approach.

Step 1 Assumptions: Bivariate normal population

Step 2 $H_o: \rho = 0$

$H_a: \rho \neq 0$

Step 3 a) Test statistic: t with $df = 8$

b) Two-tailed test with $\alpha = 0.05$

Step 4

a) $t = \dfrac{r\sqrt{n-2}}{\sqrt{1-r^2}} = \dfrac{0.931\sqrt{8}}{\sqrt{1-0.931^2}} = 7.2$

b) p-value $< 2(0.002) = 0.004$

Step 5 a) p-value $< 0.004 < \alpha = 0.05$

b) Reject H_o.

$\therefore$ There is sufficient evidence at the 5% level of significance to conclude that there is a correlation between the temperature in Montreal and in Quebec.

Formulas

$y = a + bx$ where $b = \frac{SS_{xy}}{SS_x}$ and $a = \overline{y} - b\overline{x}$

$$SS_{xy} = \sum xy - \frac{\left(\sum x\right)\left(\sum y\right)}{n}$$

$$r = \frac{SS_{xy}}{\sqrt{SS_x SS_y}}$$

$$SS_x = \sum x^2 - \frac{\left(\sum x\right)^2}{n}$$

$$t = \frac{\overline{d} - \mu_d}{\frac{s_d}{\sqrt{n}}}$$

$$E = t_{\left(df, \frac{\alpha}{2}\right)} \frac{s_d}{\sqrt{n}}$$

$$z = \frac{\left(\overline{x}_1 - \overline{x}_2\right) - \left(\mu_1 - \mu_2\right)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

$$E = z_{\frac{\alpha}{2}} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

$$t = \frac{\left(\overline{x}_1 - \overline{x}_2\right) - \left(\mu_1 - \mu_2\right)}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$E = t_{\left(df, \frac{\alpha}{2}\right)} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

$$s_p = \sqrt{\frac{\left(n_1 - 1\right) s_1^2 + \left(n_2 - 1\right) s_2^2}{n_1 + n_2 - 2}}$$

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}_p \hat{q}_p \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$E = z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}$$

$$\chi^2 = \sum \frac{\left(O - E\right)^2}{E}$$

$$\chi^2 = \frac{\left(n - 1\right) s^2}{\sigma^2}$$

$$\frac{\left(n - 1\right) s^2}{\chi^2_{\left(df, \frac{\alpha}{2}\right)}} < \sigma^2 < \frac{\left(n - 1\right) s^2}{\chi^2_{\left(df, 1 - \frac{\alpha}{2}\right)}}$$

$t = \frac{a - \alpha}{S_e} \sqrt{\frac{nSS_x}{SS_x + n\overline{x}^2}}$ with $df = n - 2$

$$S_e = \sqrt{\frac{SS_y - \frac{\left(SS_{xy}\right)^2}{SS_x}}{n - 2}}$$

$$E = t_{\left(df, \frac{\alpha}{2}\right)} S_e \sqrt{\frac{1}{n} + \frac{\overline{x}^2}{SS_x}}$$

$t = \frac{b - \beta}{S_e} \sqrt{SS_x}$ with $df = n - 2$

$$E = t_{\left(df, \frac{\alpha}{2}\right)} \frac{S_e}{\sqrt{SS_x}}$$

$$E = t_{\left(df, \frac{\alpha}{2}\right)} S_e \sqrt{1 + \frac{1}{n} + \frac{\left(x - \overline{x}\right)^2}{SS_x}}$$

$$t = \frac{r\sqrt{n - 2}}{\sqrt{1 - r^2}}$$

$$Z = \frac{1}{2} \ln \frac{1 + r}{1 - r}$$

$$Z - \frac{z_{\frac{\alpha}{2}}}{\sqrt{n - 3}} < \mu_z < Z + \frac{z_{\frac{\alpha}{2}}}{\sqrt{n - 3}}$$

$$r = \frac{e^{2Z} - 1}{e^{2Z} + 1}$$