Discussion 14 Solutions

1. Province Ranch is an insurance company that provides homeowners policies to tenants and non-tenants in Neverland. A random sample of 16 policies for tenants yielded the sample mean basic premium of \$250 and the sample standard deviation of \$34. Assume that the basic premiums for tenant policies at Province Ranch are normally distributed.

a) Construct a 90% confidence interval for the overall mean basic premium for tenant policies at Province Ranch.

$\overline{X} = 250.$ $s = 34.$ $n = 16.$ $\alpha = 0.10.$

σ is unknown. The confidence interval : $\overline{X} \pm t_{\alpha/2} \dfrac{s}{\sqrt{n}}.$

$\alpha/2 = 0.05.$ number of degrees of freedom $= n - 1 = 16 - 1 = 15.$

$t_{\alpha/2} = 1.753.$ $250 \pm 1.753 \cdot \dfrac{34}{\sqrt{16}}$ **250 ± 14.9** **(235.1 , 264.9)**

b) Construct a 95% confidence interval for the overall standard deviation of the basic premiums for tenant policies at Province Ranch.

The confidence interval : $$\left(\sqrt{\frac{(n-1)\cdot s^2}{\chi^2_{\alpha/2}}}, \sqrt{\frac{(n-1)\cdot s^2}{\chi^2_{1-\alpha/2}}} \right).$$

95% confidence level $\alpha = 0.05$ $\alpha/2 = 0.025.$

number of degrees of freedom $= n - 1 = 16 - 1 = 15.$

$\chi^2_{\alpha/2} = \chi^2_{0.025} = 27.49.$ $\chi^2_{1-\alpha/2} = \chi^2_{0.975} = 6.262.$

$$\left(\sqrt{\frac{(16-1)\cdot 34^2}{27.49}}, \sqrt{\frac{(16-1)\cdot 34^2}{6.262}} \right)$$ **(25.115 , 52.622)**

c) Suppose also that a random sample of 9 non-tenant policies yielded the sample mean basic premium of \$375 and the sample standard deviation of \$40. Assume that the basic premiums for non-tenant policies at Province Ranch are also normally distributed, and the two overall standard deviations are equal. Construct a 90% confidence interval for the difference between the overall mean basic premiums for tenants and non-tenants at Province Ranch.

$$s_{\text{pooled}}^2 = \frac{(16-1)\cdot 34^2 + (9-1)\cdot 40^2}{16+9-2} \approx 1310.4348 \qquad s_{\text{pooled}} \approx 36.2$$

$$\left(\overline{X}-\overline{Y}\right) \pm t_{\alpha/2}\cdot s_{\text{pooled}}\cdot\sqrt{\frac{1}{n}+\frac{1}{m}} \qquad 16+9-2 = 23 \text{ degrees of freedom}$$

$$\alpha/2 = 0.05. \qquad t_{0.05}(23) = 1.714.$$

$$(375-250) \pm 1.714\cdot 36.2\cdot\sqrt{\frac{1}{16}+\frac{1}{9}}$$

125 ± 25.85 **(99.15, 150.85)**

d) Repeat part (c) without the assumption that the two overall standard deviations are equal. Use Welch's T.

$$\left\lfloor \frac{\left(\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}\right)^2}{\frac{1}{n_1-1}\cdot\left(\frac{s_1^2}{n_1}\right)^2+\frac{1}{n_2-1}\cdot\left(\frac{s_2^2}{n_2}\right)^2} \right\rfloor = \left\lfloor \frac{\left(\frac{34^2}{16}+\frac{40^2}{9}\right)^2}{\frac{1}{16-1}\cdot\left(\frac{34^2}{16}\right)^2+\frac{1}{9-1}\cdot\left(\frac{40^2}{9}\right)^2} \right\rfloor$$

$$= \lfloor 14.54278 \rfloor = \mathbf{14} \text{ degrees of freedom}$$

$$t_{0.05}(14) = 1.761 \qquad (375-250) \pm 1.761\cdot\sqrt{\frac{34^2}{16}+\frac{40^2}{9}}$$

125 ± 27.8454 **(97.1546, 152.8454)**

2. Over a long period of time at a suburban restaurant, 60% of the customers ordered coffee, 20% ordered tea, 15% ordered soda, and 5% ordered milk. The owners opened a new restaurant in the business section of town, where they think that the drinking preferences of the customers may be different. In a random sample of 300 customers, 150 ordered coffee, 78 ordered tea, 51 ordered soda, and 21 ordered milk. Determine at a 1% level of significance whether drinking patterns are different in town than at the suburban restaurant.

$H_0: \quad p_C = 0.60, \quad p_T = 0.20, \quad p_S = 0.15, \quad p_M = 0.05.$

	C	T	S	M
O	150	78	51	21
E	$300 \cdot 0.60 = 180$	$300 \cdot 0.20 = 60$	$300 \cdot 0.15 = 45$	$300 \cdot 0.05 = 15$
$\frac{(O-E)^2}{E}$	$\frac{(150-180)^2}{180}$	$\frac{(78-60)^2}{60}$	$\frac{(51-45)^2}{45}$	$\frac{(21-15)^2}{15}$
	5.0	5.4	0.8	2.4

$$Q = \sum_{cells} \frac{(O-E)^2}{E} = 5.0 + 5.4 + 0.8 + 2.4 = \mathbf{13.6}.$$

$k - 1 = 4 - 1 = 3$ d.f.

Rejection Region: "Reject H_0 if $Q > \chi^2_\alpha(3)$"

$\chi^2_{0.01}(3) = \mathbf{11.34}$.

$13.6 = Q > \chi^2_\alpha(3) = 11.34.$ **Reject H_0.**

3. A stuff member of an emergency medical service wishes to determine whether the number of accidents is equally distributed during the week. A week was selected at random, and the following data were obtained. Perform the appropriate test at a 5% significance level. State the null hypothesis. Report the value of the test statistic, the critical value(s), the p-value (you may give a range), and state your decision.

Day	Mo	Tu	We	Th	Fr	Sa	Su
No. of accidents	22	16	11	13	22	31	25

$$\mathbf{H_0}: p_{\mathbf{Mo}} = p_{\mathbf{Tu}} = p_{\mathbf{We}} = p_{\mathbf{Th}} = p_{\mathbf{Fr}} = p_{\mathbf{Sa}} = p_{\mathbf{Su}} = \frac{\mathbf{1}}{\mathbf{7}}$$

H_1 : not H_0

$n = 22 + 16 + 11 + 13 + 22 + 31 + 25 = 140.$

Each expected frequency is $140 \times \frac{1}{7} = 20.$

$$Q = \frac{(22-20)^2}{20} + \frac{(16-20)^2}{20} + \frac{(11-20)^2}{20} + \frac{(13-20)^2}{20} + \frac{(22-20)^2}{20} + \frac{(31-20)^2}{20} + \frac{(25-20)^2}{20}$$

$$= 0.20 + 0.80 + 4.05 + 2.45 + 0.20 + 6.05 + 1.25 = \mathbf{15}.$$

Rejection Region: "Reject H_0 if $Q \ge \chi^2_\alpha(6)$" $\qquad \chi^2_{0.05}(6) = \mathbf{12.59}.$

$15 = Q > \chi^2_\alpha(6) = 12.59.$ $\qquad$ **Reject $\mathbf{H_0}$ at $\alpha = 0.05$.**

$16.81 > 15 > 14.45.$

$0.01 < \text{p-value} < 0.025.$

p-value $\approx 0.02.$

4. A group of 100 children (50 boys and 50 girls) were asked to select their favorite color out of Red, Green, and Blue. The following data was obtained:

Gender	Red	Green	Blue	
Boys	21	7	22	50
Girls	14	18	18	50
	35	25	40	100

A toy manufacturer wants to know if the color preferences of boys and girls differ. Perform χ^2 test of homogeneity using a 5% level of significance. State the null hypothesis, report the value of the test statistic, the critical value(s), and state your decision.

$$\mathbf{H_0:}\ p_{\mathbf{BR}} = p_{\mathbf{GR}},\ p_{\mathbf{BG}} = p_{\mathbf{GG}},\ p_{\mathbf{BB}} = p_{\mathbf{GB}} \quad \text{vs.} \quad H_1: \text{ not } H_0$$

$$E_{11} = \frac{50 \times 35}{100} = 17.5 \qquad E_{12} = \frac{50 \times 25}{100} = 12.5 \qquad E_{13} = \frac{50 \times 40}{100} = 20$$

$$E_{21} = \frac{50 \times 35}{100} = 17.5 \qquad E_{22} = \frac{50 \times 25}{100} = 12.5 \qquad E_{23} = \frac{50 \times 40}{100} = 20$$

$$Q = \frac{(21-17.5)^2}{17.5} + \frac{(7-12.5)^2}{12.5} + \frac{(22-20)^2}{20} + \frac{(14-17.5)^2}{17.5} + \frac{(18-12.5)^2}{12.5} + \frac{(18-20)^2}{20}$$

$$= 0.70 + 2.42 + 0.20 + 0.70 + 2.42 + 0.20 = \mathbf{6.64}.$$

$(3-1) \times (2-1) = 2$ degrees of freedom $\qquad \chi^2_{0.05}(2) = \mathbf{5.991}$.

$Q = 6.64 > 5.991 = \chi^2_{\alpha}$. $\qquad$ **Reject $\mathbf{H_0}$.**

P-value $\approx$ 0.036.

5. A survey of high school girls classified them by two attributes: whether or not they participated in sports and whether or not they had one or more older brothers. Use the following data to test the null hypothesis that these two attributes are independent:

	Participated in Sports		
Older Brother(s)	Yes	No	Totals
Yes	11	9	20
No	13	27	40
Totals	24	36	60

a) Use $\alpha = 0.05$.

b) Use $\alpha = 0.10$.

H_0 : sports and older brothers are independent

$$E_{11} = \frac{20 \times 24}{60} = 8 \qquad E_{12} = \frac{20 \times 36}{60} = 12$$

$$E_{21} = \frac{40 \times 24}{60} = 16 \qquad E_{22} = \frac{40 \times 36}{60} = 24$$

$$Q = \frac{(11-8)^2}{8} + \frac{(9-12)^2}{12} + \frac{(13-16)^2}{16} + \frac{(27-24)^2}{24}$$

$$= 1.125 + 0.75 + 0.5625 + 0.375 = \mathbf{2.8125}.$$

a) $(2-1) \times (2-1) = 1$ degree of freedom $\qquad \chi^2_{0.05}(1) = \mathbf{3.841}$.

$2.8125 < 3.841$. **Do NOT Reject H_0**

b) $\chi^2_{0.10}(1) = \mathbf{2.706}$. $\qquad$ p-value ≈ 0.0935.

$2.8125 > 2.706$. **Reject H_0**