STAT 400 Fall 2017

Two Sample and Chi-Square Testing Practice

NINTH (EIGHTH) (SEVENTH)

1. a) **7.2-2 (6.3-2) (6.5-2)** "Hint": equal variances

b) Use Welch's T to construct a 90% confidence interval for $\mu_X - \mu_Y$ (without the assumption that the variances are equal).

2. **7.3-12 (6.5-18) (6.7-18)**

3. **8.3-14 (7.1-20) (8.1-20)**

4. **8.2-2 (7.3-2) (8.3-2 (a),(b),(d))**

5. **8.2-11 (7.3-15) (8.3-15)**

6. **9.1-1 (8.1-1) (8.5-1)**

7. **9.1-2 (8.1-2) (8.5-2)**

8. **9.1-4 (8.1-4) (8.5-6)**

9. **9.2-1 (8.2-1) (8.6-1)**

10. **9.2-2 (8.2-2) (8.6-2)**

11. **9.2-8 (8.2-8) (8.6-8)**

12. **9.2-10 (8.2-10) (8.6-10)**

Answers:

1. a) 7.2-2 (6.3-2) (6.5-2) "Hint": equal variances

	x	$(x-\bar{x})^2$	y	$(y-\bar{y})^2$
	644	10983.04	623	6142.640625
	493	2134.44	472	5274.390625
	532	51.84	492	2769.390625
	462	5959.84	661	13543.140625
	565	665.64	540	21.390625
			502	1816.890625
			549	19.140625
			518	708.890625
Total:	2696	19794.80	4357	30295.875000

$$\bar{x} = \frac{2696}{5} = 539.2, \qquad s_x^2 = \frac{19794.8}{4} = 4948.7,$$

$$\bar{y} = \frac{4357}{8} = 544.625, \qquad s_y^2 = \frac{30295.875}{7} \approx 4327.982.$$

$$s_{\text{pooled}}^2 = \frac{(5-1)\cdot 4948.7 + (8-1)\cdot 4327.982}{5+8-2} \approx 4553.7, \qquad s_{\text{pooled}} \approx 67.481.$$

$$t_{0.05}(11) = 1.796, \qquad (539.2 - 544.625) \pm 1.796 \cdot 67.481 \cdot \sqrt{\frac{1}{5} + \frac{1}{8}}$$

$$\mathbf{-5.425 \pm 69.092} \quad \text{or} \quad \mathbf{(-74.517,\ 63.667)}.$$

b) Use Welch's T to construct a 90% confidence interval for $\mu_X - \mu_Y$ (without the assumption that the variances are equal).

$$\left\lfloor \frac{\left(\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}\right)^2}{\frac{1}{n_1-1}\cdot\left(\frac{s_1^2}{n_1}\right)^2+\frac{1}{n_2-1}\cdot\left(\frac{s_2^2}{n_2}\right)^2} \right\rfloor = \left\lfloor \frac{\left(\frac{4948.7}{5}+\frac{4327.982}{8}\right)^2}{\frac{1}{5-1}\cdot\left(\frac{4948.7}{5}\right)^2+\frac{1}{8-1}\cdot\left(\frac{4327.982}{8}\right)^2} \right\rfloor$$

$$= \lfloor 8.172642 \rfloor = \mathbf{8} \text{ degrees of freedom}$$

$$t_{0.05}(8) = 1.860, \qquad (539.2 - 544.625) \pm 1.860 \cdot \sqrt{\frac{4948.7}{5}+\frac{4327.982}{8}}$$

$$\mathbf{-5.425 \pm 72.772} \quad \text{or} \quad \mathbf{(-78.197,\ 67.347)}.$$

2. **7.3-12** **(6.5-18)** **(6.7-18)**

$n_A = 460,\quad y_A = 170.\qquad \hat{p}_A = \frac{y_A}{n_A} = \frac{170}{460} = 0.37.$

$n_B = 440,\quad y_B = 141.\qquad \hat{p}_B = \frac{y_B}{n_B} = \frac{141}{440} = 0.32.$

a) $$(\hat{p}_A - \hat{p}_B) \pm z_{\alpha/2} \cdot \sqrt{\frac{\hat{p}_A\cdot(1-\hat{p}_A)}{n_A}+\frac{\hat{p}_B\cdot(1-\hat{p}_B)}{n_B}}$$

95% confidence level $\qquad \alpha = 0.05 \qquad \alpha/2 = 0.025. \qquad z_{\alpha/2} = 1.96.$

$$(0.37 - 0.32) \pm 1.96 \cdot \sqrt{\frac{0.37\cdot 0.63}{460}+\frac{0.32\cdot 0.68}{440}} \quad \text{or} \quad \mathbf{[-0.012,\ 0.112]}.$$

$$\mathbf{0.05 \pm 0.062}$$

b) Yes, the interval in part (a) includes zero.

3. **8.3-14** **(7.1-20)** **(8.1-20)**

$H_0: p_1 = p_2$ vs. $H_1: p_1 > p_2$. Right – tailed.

or $H_0: p_1 \le p_2$ vs. $H_1: p_1 > p_2$

a) The test statistic is

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}\cdot(1-\hat{p})\cdot\left(1/n_1 + 1/n_2\right)}}, \quad \text{where} \quad \hat{p} = \frac{y_1 + y_2}{n_1 + n_2} = \frac{n_1 \cdot \hat{p}_1 + n_2 \cdot \hat{p}_2}{n_1 + n_2}.$$

$\alpha = 0.05$. The critical (rejection) region is $z > z_{\alpha} = 1.645$.

b) $n_1 = 900$, $y_1 = 135$. $\hat{p}_1 = \dfrac{y_1}{n_1} = \dfrac{135}{900} = 0.15.$

$n_2 = 700$, $y_2 = 77$. $\hat{p}_2 = \dfrac{y_2}{n_2} = \dfrac{77}{700} = 0.11.$

$$\hat{p} = \frac{y_1 + y_2}{n_1 + n_2} = \frac{135 + 77}{900 + 700} = \frac{212}{1600} = 0.1325.$$

The observed value of z (test statistic)

$$z = \frac{0.15 - 0.11}{\sqrt{0.1325 \cdot 0.8675 \cdot \left(1/900 + 1/700\right)}} = 2.341$$

is greater than 1.645 (the test statistic does fall into the rejection region)
so Reject H_0.

c) $\alpha = 0.01$. The critical (rejection) region is $z > z_{\alpha} = 2.326$.
$z = 2.341 > 2.326$ (the test statistic does fall into the rejection region)
so Reject H_0.

d) p-value = $P(Z > 2.341) = 0.0096$.

4. **8.2-2** **(7.3-2)** **(8.3-2 (a),(b),(d))**

a) $H_0 : \mu_X = \mu_Y$ vs. $H_1 : \mu_X > \mu_Y$ Right – tailed.

$n = 16$, $m = 13$.

If we assume that population variances are equal, then we can use the test statistic

$$t = \frac{(\bar{x} - \bar{y})}{s_{pooled} \cdot \sqrt{\frac{1}{16} + \frac{1}{13}}}$$

where

$$s^2_{pooled} = \frac{15 \cdot s_x^2 + 12 \cdot s_y^2}{27}.$$

Then the number of degrees of freedom = $n + m - 2 = 16 + 13 - 2 = 27$.

$\alpha = 0.01$. The critical (rejection) region is $t > t_\alpha =$ **2.473**.

b) $\bar{x} = 415.16$, $s_x^2 = 1356.75$, $\bar{y} = 347.40$, $s_y^2 = 692.21$.

$$s^2_{pooled} = \frac{(n-1) \cdot s_x^2 + (m-1) \cdot s_y^2}{n+m-2} = \frac{15 \cdot 1356.75 + 12 \cdot 692.21}{27} \approx 1061.398889.$$

$s_{pooled} \approx 32.579$.

The observed value of t (test statistic)

$$t = \frac{(415.16 - 347.40)}{32.579 \cdot \sqrt{\frac{1}{16} + \frac{1}{13}}} = \mathbf{5.570}$$

is greater than 2.473 (the test statistic does fall into the rejection region)

so **Reject H_0**.

c) Welch's T:

$$\left\lfloor \frac{\left(\frac{s_x^2}{n}+\frac{s_y^2}{m}\right)^2}{\frac{1}{n-1}\cdot\left(\frac{s_x^2}{n}\right)^2+\frac{1}{m-1}\cdot\left(\frac{s_y^2}{m}\right)^2} \right\rfloor = \left\lfloor \frac{\left(\frac{1356.75}{16}+\frac{692.21}{13}\right)^2}{\frac{1}{16-1}\cdot\left(\frac{1356.75}{16}\right)^2+\frac{1}{13-1}\cdot\left(\frac{692.21}{13}\right)^2} \right\rfloor$$

$= \lfloor 26.628 \rfloor = 26$ degrees of freedom.

OR

$$c = \frac{\frac{s_x^2}{n}}{\frac{s_x^2}{n}+\frac{s_y^2}{m}} = \frac{\frac{1356.75}{16}}{\frac{1356.75}{16}+\frac{692.21}{13}} = 0.614275.$$

$$\frac{1}{r} = \frac{c^2}{n-1}+\frac{(1-c)^2}{m-1} = \frac{0.614275^2}{15}+\frac{0.385725^2}{12} = 0.037554.$$

$r = \frac{1}{0.037554} = 26.628.$ Round down. 26 degrees of freedom.

Test statistic: $$t = \frac{\left(\bar{x}-\bar{y}\right)}{\sqrt{\frac{s_x^2}{n}+\frac{s_y^2}{m}}} = \frac{(415.16-347.40)}{\sqrt{\frac{1356.75}{16}+\frac{692.21}{13}}} = \mathbf{5.7672}.$$

$\alpha = 0.01$. The critical (rejection) region is $t > t_\alpha = \mathbf{2.479}$.

Reject $\mathbf{H_0}$.

5. **8.2-11** **(7.3-15)** **(8.3-15)**

$H_0: \mu_X - \mu_Y = 0$ vs. $H_1: \mu_X - \mu_Y > 0$ Right – tailed.

or $H_0: \mu_X - \mu_Y \le 0$ vs. $H_1: \mu_X - \mu_Y > 0$

$n = 90$, $\bar{x} = 8.10$, $s_x = 0.117$, $m = 110$, $\bar{y} = 8.07$, $s_y = 0.054$.

a) $s_x^2 = 0.013689$, $s_y^2 = 0.002916$. Population variances do not seem to be equal.
n and m are both large. The observed value of z (test statistic) is

$$z = \frac{\left(\bar{x} - \bar{y}\right) - 0}{\sqrt{\frac{s_x^2}{n} + \frac{s_y^2}{m}}} = \frac{\left(8.10 - 8.07\right) - 0}{\sqrt{\frac{0.117^2}{90} + \frac{0.054^2}{110}}} = \mathbf{2.245}.$$

$\alpha = 0.05$. The critical (rejection) region is $z > z_\alpha = \mathbf{1.645}$.

The observed value of z (test statistic) does fall into the rejection region:
$2.245 > 1.645$. **Reject H_0**.

b) p-value $= P(Z > 2.245) \approx \mathbf{0.0124}$.

6. **9.1-1** **(8.1-1)** **(8.5-1)**

$$Q_3 = \frac{(42-56)^2}{56} + \frac{(64-56)^2}{56} + \frac{(53-56)^2}{56} + \frac{(65-56)^2}{56}$$

$$= 3.5 + 1.142857 + 0.160714 + 1.446429 = \mathbf{6.25} < 7.815 = \chi^2_{0.05}(3).$$

Thus we do not reject H_0 at $\alpha = 0.05$.

Since $\chi^2_{0.10}(3) = 6.251$, p-value ≈ 0.10.

7. **9.1-2** **(8.1-2)** **(8.5-2)**

$$Q_4 = \frac{(224-232)^2}{232} + \frac{(119-116)^2}{116} + \frac{(130-116)^2}{116} + \frac{(48-58)^2}{58} + \frac{(59-58)^2}{58}$$

$$\approx \mathbf{3.7845}.$$

The null hypothesis will not be rejected at any reasonable significance level. Note that $E(Q_4) = 4$ when H_0 is true.

8. **9.1-4** **(8.1-4)** **(8.5-6)**

124 + 30 + 43 + 11 = 208

$208 \times \frac{9}{16} = 117$, $\quad 208 \times \frac{3}{16} = 39$, $\quad 208 \times \frac{3}{16} = 39$, $\quad 208 \times \frac{1}{16} = 13$.

$$Q_3 = \frac{(124-117)^2}{117} + \frac{(30-39)^2}{39} + \frac{(43-56)^2}{39} + \frac{(11-13)^2}{13}$$

$$\approx \mathbf{3.213675} < 7.815 = \chi^2_{0.05}(3).$$

Thus we do not reject H_0: $p_1 = 9/16$, $p_2 = 3/16$, $p_3 = 3/16$, $p_4 = 1/16$ at $\alpha = 0.05$.

9. 9.2-1 (8.2-1) (8.6-1)

O / **E** / $\frac{(O-E)^2}{E}$							
	95 **88.8** *0.432883*	36 **37.2** *0.03871*	71 **68.4** *0.09883*	21 **23.4** *0.246154*	45 **46.2** *0.031169*	32 **36** *0.444444*	300
	53 **59.2** *0.649324*	26 **24.8** *0.058065*	43 **45.6** *0.148246*	18 **15.6** *0.369231*	32 **30.8** *0.046753*	28 **24** *0.666667*	200
	148	62	114	39	77	60	500

$Q_5 = 3.230475 < 11.07 = \chi^2_{0.05}(5)$. Thus we do NOT reject H_0 at $\alpha = 0.05$.
Note that $E(Q_5) = 5$ when H_0 is true.

10. 9.2-2 (8.2-2) (8.6-2)

O / **E** / $\frac{(O-E)^2}{E}$							
	95 **83.4** *1.613429*	36 **41.1** *0.632847*	71 **75** *0.213333*	21 **21.6** *0.016667*	45 **41.4** *0.313043*	32 **37.5** *0.806667*	300
	53 **55.6** *0.121583*	26 **27.4** *0.071533*	43 **50** *0.98*	18 **14.4** *0.9*	32 **27.6** *0.701449*	28 **25** *0.36*	200
	130 **139** *0.582734*	75 **68.5** *0.616788*	136 **125** *0.968*	33 **36** *0.25*	61 **69** *0.927536*	65 **62.5** *0.1*	500
	278	137	250	72	138	125	1000

$Q_{10} = 10.1756 < 20.48 = \chi^2_{0.025}(10)$. Thus we do NOT reject H_0 at $\alpha = 0.025$.

11. 9.2-8 (8.2-8) (8.6-8)

51 **44.26744** *1.023943*	30 **36.73256** *1.233983*	81
43 **49.73256** *0.911422*	48 **41.26744** *1.098380*	91
94	78	172

$Q_1 = 4.26773 > 3.841 = \chi^2_{0.05}(1)$, reject H_0 at $\alpha = 0.05$.

12. 9.2-10 (8.2-10) (8.6-10)

O **E** $\frac{(O-E)^2}{E}$

5 **4.5** *0.055556*	11 **7** *2.285714*	9 **13.5** *1.5*	25
4 **4.5** *0.055556*	3 **7** *2.285714*	18 **13.5** *1.5*	25
9	14	27	50

$Q_2 = 7.68254 < 9.210 = \chi^2_{0.01}(2)$, do NOT reject H_0 at $\alpha = 0.01$.