Name: Drashti Mehta

Lab Assignment 11

(40 points)

DIRECTIONS

- ***This lab is due at the end of the next Friday.***
- ***Be sure to show all of your work by providing your hypotheses, R commands, solutions/output, and interpretation.***
- ***Please circle or highlight your final answers.***

1) In a study to test the effect of an acne cream, scientists found that 25 out of 106 randomly selected subjects who used the acne cream suffered from headaches. 19 of 133 randomly selected subjects who used a placebo cream suffered from headaches. Based on these data, can we conclude that the acne cream users had a different rate of headaches at α=0.02? (10 points).

A. 1) - Hypothesis:

$H_0: p_1 = p_2$

$H_a: p_1 \neq p_2$

- $n_1\hat{p}_1 \geq 5$, $n_1(1-\hat{p}_1) \geq 5$

$n_2\hat{p}_2 \geq 5$, $n_2(1-\hat{p}_2) \geq 5$

- $\hat{p} = \frac{(x_1+x_2)}{(n_1+n_2)} = \frac{25+19}{106+133} = \frac{44}{239} = 0.184$

- $z = \frac{(0.236-0.143)}{\left[0.184\left(\frac{1}{106}+\frac{1}{133}\right)\right]^{1/2}}$

$= 0.093 / 0.0505 = 1.84$

- $z_{\alpha/2} = 2.32$

→ In R:

```
> qnorm(0.99)
> 2.326348.
```

So z is not in the rejection region and we can't reject the null hypothesis. No enough evidence to conclude acne users had different rate of headaches.

In R:

> prop.test(x=c(25,19),n=c(106,133),correct=FALSE)

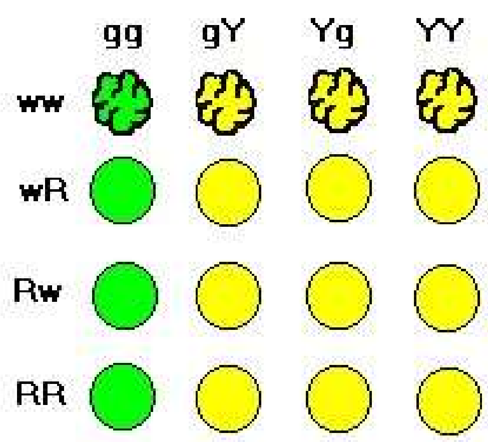
Result:

2-sample test for equality of proportions without
continuity correction

data: c(25, 19) out of c(106, 133)
X-squared = 3.3959, df = 1, p-value = 0.06536
alternative hypothesis: two.sided
95 percent confidence interval:
-0.007347869 0.193331696
sample estimates:
prop 1 prop 2
0.2358491 0.1428571

We see that the p-value = 0.06536 which is > 0.05. So, we cannot reject H0.

2) Mendel identified these pea types after crossing, determine if the observed numbers deviate from the expected frequency. Solve 'manually' by hand and also use an R test at alpha 0.05 (15 points).



Pea Type	Expected frequency	Observed number
yellow & round	9/16	301
yellow & wrinkled	3/16	111
green & round	3/16	118
green & wrinkled	1/16	35

H0: The distribution of the categories is 9/16, 3/16, 3/16 and 1/16
Ha: The distribution of the categories is not as claimed in H0
Degrees of freedom: 4-1 = 3 and $\alpha = 0.05$
table value: X2(3,0.05) = 7.815
Since X2= 2.507571 < 7.815

We fail to reject H0 and we conclude that the distribution of the categories is 9/16, 3/16, 3/16 and 1/16.

2. → observed table:

	Green	Yellow	Total
Round	118	301	419
wrinkled	35	111	146
Total	153	412	565

→ Expected table:

	Green	Yellow	Total
Round	565 x 3/16 = 105.9375	565 x 9/16 = 317.8125	423.75
Wrinkled	565 x 1/16 = 35.3125	565 x 3/16 = 105.9375	141.25
Total	141.25	423.75	565.

$$\chi^2 = \frac{(118-105.9375)^2}{105.9375} + \frac{(35-35.3125)^2}{35.3125} + \frac{(301-317.8125)^2}{317.8125} + \frac{(111-105.9375)^2}{105.9375}$$

$$= 2.507571$$

> pchisq(2.507571, 3, lower.tail = FALSE)
[1] 0.4739244
Since p value = 0.4739244 > 0.05, we do not reject the null hypothesis.

In R:
> x<- c(301, 118, 111, 35)
> p<- c(9/16, 3/16, 3/16, 1/16)
> chisq.test(x, p= p)

Chi-squared test for given probabilities

data: x
X-squared = 2.5076, df = 3, p-value = 0.4739

> qchisq(0.95,3)
[1] 7.814728

3) A plant scientist noticed that the seed of plant A (assume each plant only has one seed) can have three different colors (yellow, green, and red) and two different shapes (rounded and wrinkled). He wondered if there is any dependence between the shape and the color at alpha = 0.05. Solve 'manually' by hand and using an R test (15 points).

He randomly selected 250 seeds and noted the colors and shapes of the seeds as follows:

	Yellow	Green	Red
Rounded	70	30	50
Wrinkled	40	30	30

H0: the shape and the color variables for plant A are independent
Ha: the shape and the color variables for plant A are associated
Degrees of freedom: rows -1 * columns – 1 = 1 x 2 = 2
table value: X2(2,0.05) = 5.991
Since X2= 3.3144 < 5.991, we fail to reject H0 and we conclude that the shape and the color variables for plant A are independent.

3. → observed table :

	Yellow	Green	Red	Total
Round	70	30	50	150
Wrinkled	40	30	30	100
Total	110	60	80	250

→ Expected table:

	Yellow	Green	Red	Total
Round	110 x 150/250 = 66	60x150/250 = 36	80x150/250 = 48	150
wrinkled	110x100/250 44	60x100/250 = 24	80x100/250 = 32	100
Total	110	60	80	250

$$\chi^2 = \frac{(70-66)^2}{66} + \frac{(40-44)^2}{44} + \frac{(30-36)^2}{36} + \frac{(30-24)^2}{24} + \frac{(50-48)^2}{48} + \frac{(30-32)^2}{32}$$

$= 3.3144.$

```
> pchisq(3.3144, 2, lower.tail = F)
[1] 0.1906721
```

Since p value = 0.1906721 > 0.05, we do not reject the null hypothesis.

In R:

```
> x<- rbind(c(70,40), c(30,30), c(50,30))
> chisq.test(x)

        Pearson's Chi-squared test

data:  x
X-squared = 3.3144, df = 2, p-value = 0.1907

> qchisq(0.95, 2)
[1] 5.991465
```