Emma Erdtmann
Daniel Detore
I pledge my honor that I have abided by the Stevens Honor System.

Section 1
Problem 1
Null distribution: Standard normal (z)
Observed test statistic (z): 7.369312
p-value: 1.715113e-13
As the p-value is less than the significance level of 0.05, reject the null hypothesis. This means there is a difference in SRP participation with male and female students.

Problem 2
Null distribution: Chi-square with df = 1
Observed X^2 statistic: 54.30675
p-value: 1.715113e-13
H0 Null hypothesis would be that gender and SRP participation are independent, Ha alternative hypothesis would be that they are associated; the X^2 test shows that we should reject the null hypothesis meaning there is significant association.

Problem 3
Null distribution: Chi-square with df = 1
Observed X^2 statistic: 9.725788
p-value: 0.001817004

H0 would be P(male) = 0.4, P(female) = 0.6, Ha would be that at least one of the probabilities differs. The p-value is less than the significance level of 0.05 so we reject the null hypothesis; there is some significant difference in the gender distribution.

Section 2
Problem 4
The two involved categorical random values are
- "Version Viewed" with values "1st", "2nd", "3rd", and "4th", and
- "Likely to Buy" with values "Yes" or "No"

Problem 5
Observed table for Version Viewed and Likely to Buy based on the data:

	Likely to Buy		
Version Viewed	Yes	No	Total
1st	25	35	60
2nd	18	17	35

3rd	30	15	45
4th	24	26	50
Total	97	93	190

Problem 6

The value of each inner cell $E_{i,j} = \frac{O_{i,+}O_{+,j}}{O_{+,+}}$ where $O_{+,+} = 190.$

Expected table for Version Viewed and Likely to Buy based on H_0:

	Likely to Buy		
Version Viewed	Yes	No	Observed Total
1st	$\frac{97*60}{190} = 30.63$	$\frac{93*60}{190} = 29.37$	60
2nd	$\frac{97*35}{190} = 17.87$	$\frac{93*35}{190} = 17.13$	35
3rd	$\frac{97*45}{190} = 22.97$	$\frac{93*45}{190} = 22.03$	45
4th	$\frac{97*50}{190} = 25.53$	$\frac{93*50}{190} = 24.47$	50
Observed Total	97	93	190

Problem 7

```
> e = c(30.63, 17.87, 22.97, 25.53,  29.37, 17.13, 22.03, 24.47)
> o = c(25, 18, 30, 24,  35, 17, 15, 26)
> xx = sum((o-e)^2/e); xx
[1] 6.698234
```

thus $x^2 = 6.70.$

Problem 8

i

H_0: there is no association between Version Viewed and Likely to Buy.

H_a: there is some association between Version Viewed and Likely to Buy.

ii

$\chi^2 \sim \chi^2_k$ where df $k = (r - 1)(c - 1) = (4 - 1)(2 - 1) = 3$

Problem 9

```
> 1-pchisq(6.698234, 3)
```

```
[1] 0.08216406
```

thus $P_{H_0}(\chi^2 > 6.698234) = 1 - pchisq(6.698234, 3) = 0.08$.

Problem 10

If our $\alpha > 0.08$, we can reject H_0 and assume H_a. If $\alpha \le 0.08$, we cannot reject H_0 and must assume H_0 is true. In other words, we can assume H_a with 92% confidence.

Problem 11

With a somewhat high 92% confidence (i.e. at the 0.08 significance level), we can say that there is some association between which version of the product the customer viewed and whether they were likely to buy it or not. If the confidence level were higher (or the significance level were lower) this would not be possible.

Problem 12

```
> chisq.test(matrix(o, nrow=4), correct=FALSE)

	Pearson's Chi-squared test

data:  matrix(o, nrow = 4)
X-squared = 6.694, df = 3, p-value = 0.08232
```

This means our test statistic $\chi^2 = 6.694$, our null distribution is χ^2_3, and the resulting p-value is 0.08232. We are given $\alpha = 0.01 < 0.08232$ therefore we cannot reject H_0, which means there is no association between which version of the product the customer viewed and whether they were likely to buy it or not.