Homework 6

Kevin Ha

October 29, 2021

Problem 9.37

Problem 9.37a

```
##        Number allowed Number not allowed Total
## Small              51                  6    57
## Medium             12                  5    17
## Large               4                  1     5
## Total              67                 12    79
```

Problem 9.37b

Not allowed in Small Stratum: $\frac{6}{57} = 10.5263158\%$

Not allowed in Medium Stratum: $\frac{5}{17} = 29.4117647\%$

Not allowed in Large Stratum: $\frac{1}{5} = 20\%$

Problem 9.37c

Medium and large need to be combined as the expected count for the large stratum/not allowed is too small (< 5) and cannot be used for the χ^2 significance test.

Problem 9.37d

H_0 : There is no association between the size of the claim and the whether the claim is fraudulent or not.

Problem 9.37e

Calculating chi-squared manually:

```
##       Number allowed Number not allowed
## Small             51                  6
## Large             16                  6
```

```
## [1] "Expected counts"
```

```
##       Number allowed Number not allowed
## Small      48.341772           8.658228
## Large      18.658228           3.341772
```

```
##       Number allowed Number not allowed
## Small      0.1461712          0.8161227
## Large      0.3787164          2.1144998
```

$\chi^2 = 3.4555101$

$df = (r-1)(c-1) = (2-1)(2-1) = 1$

P-value:

```
pchisq(3.456, df = 1, lower.tail = FALSE)
```

```
## [1] 0.0630226
```

Problem 9.38

Problem 9.38a

Total number of claims not allowed:

Small: $\frac{6}{57} * 3342 = 351.7894737$

Medium: $\frac{5}{17} * 246 = 72.3529412$

Large: $\frac{1}{5} * 58 = 11.6$

Problem 9.38b

$SE_{estimate} = SE_{sample-proportion} * (2n)$

Margin of error for Small Claims:

$SE\hat{p}_{small} = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} * 2(n) = \sqrt{\frac{0.11(0.89)}{3342}} * 2(3342) = 36.1763348$

$MoE_{small} = 351.7894737 \pm 36.1763348$

Margin of error for Medium Claims:

$SE\hat{p}_{medium} = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} * 2(n) = \sqrt{\frac{0.29(1-0.29)}{246}} * 2(246) = 14.2339594$

$MoE_{small} = 72.3529412 \pm 14.2339594$

Margin of error for Large Claims:

$SE\hat{p}_{large} = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} * 2(n) = \sqrt{\frac{0.20(1-0.20)}{58}} * 2(58) = 6.0926185$

$MoE_{small} = 11.6 \pm 6.0926185$

Problem 9.50

Probability for Interval (1):

```
## [1] 0.2742531
```

Probability for Interval (2):

```
## [1] 0.185919
```

Probability for Interval (3):

```
## [1] 0.07965567
```

Probability for Interval (4):

```
## [1] 0.185919
```

Probability for Interval (5):

```
## [1] 0.2742531
```

Expected Counts:

Interval (1): $0.274 * 500 = 137$

Interval (2): $0.186 * 500 = 93$

Interval (3): $0.080 * 500 = 40$

Interval (4): $0.186 * 500 = 93$

Interval (5): $0.274 * 500 = 137$

Goodness of Fit Test:

$\chi^2 = \Sigma \frac{(obs-exp)^2}{exp}$

$\frac{(139-137)^2}{137} + \frac{(102-93)^2}{93} + \frac{(41-40)^2}{40} + \frac{(78-93)^2}{93} + \frac{(140-137)^2}{137} = 3.4102131$

Degrees of Freedom = $5 - 1 = 4$

P-value:

```
pchisq(3.41, df = 4, lower.tail = FALSE)
```

```
## [1] 0.4916943
```

Summary: Since the p-value is high, we fail to reject the null hypothesis and thus the random numbers follow a standard normal distribution.

Problem 9.51

Intervals used:

1. Less than or equal to 0.32
2. Less than or equal to 0.60
3. Greater than 0.40 and less than or equal to 0.60
4. Greater than -0.70 and less than or equal to -0.10
5. Greater than 1.10

```
x = rnorm(500)

obs1 = length(which(x <= 0.32))
exp1 = pnorm(0.32, mean(x), sd(x)) * 500

obs2 = length(which(x <= 0.60))
exp2 = pnorm(0.60, mean(x), sd(x)) * 500

obs3 = length(which(x <= 0.60 & x > 0.40))
exp3 = (pnorm(0.60, mean(x), sd(x)) - pnorm(0.40, mean(x), sd(x)))*500

obs4 = length(which(x <= -0.10 & x > -0.70))
exp4 = (pnorm(-0.10, mean(x), sd(x)) - pnorm(-0.70, mean(x), sd(x)))*500

obs5 = length(which(x > 1.10))
exp5 = pnorm(1.10, mean(x), sd(x), lower.tail = FALSE) * 500
```

Goodness of Fit Test:

$\chi^2 = \Sigma \frac{(obs-exp)^2}{exp}$

$\chi^2 = 2.5806017$

Degrees of Freedom = $5 - 1 = 4$

P-value:

```
pchisq(k, df = 4, lower.tail = FALSE)
```

```
## [1] 0.6302633
```