Assignment 6

Problem 1

UCBAdmissions is a built-in R dataset. For the year 1973, applicants to the graduate programs of the six largest academic departments at the University of California Berkeley are cross-classified by admission status, gender, and the department to which they applied.

a. Calculate and display the marginal contingency table cross-classifying admission by gender (ignoring department). When displaying this table, include the row, column, and table totals on the bottom, right, and bottom-right corner.

```
ucb_table <- margin.table(UCBAdmissions, c(1, 2))
ucb_table_with_totals <- addmargins(ucb_table)
ucb_table_with_totals
```

```
##           Gender
## Admit      Male Female  Sum
##   Admitted 1198    557 1755
##   Rejected 1493   1278 2771
##   Sum      2691   1835 4526
```

b. Compute the marginal table giving the proportions admitted and rejected by gender (ignoring department). Then display your results as percentages, rounded to one decimal place. For whom is the admission rate higher overall, males or females?

```
ucb_proportions <- round(prop.table(ucb_table, margin = 2) * 100, 1)
ucb_proportions
```

```
##           Gender
## Admit      Male Female
##   Admitted 44.5   30.4
##   Rejected 55.5   69.6
```

```
# As shown in table, overall admissions rate is higher for males than for females
```

c. Compute the table giving the proportions admitted and rejected for each gender within each department. Display your results as percentages, rounded to one decimal place, in a flat format using ftable(). How do these results compare to what you found in (1b)?

```
ucb_dept_proportions <- round(prop.table(UCBAdmissions, margin = c(2, 3)) * 100, 1)
ftable(ucb_dept_proportions)
```

```
##                 Dept    A    B    C    D    E    F
## Admit    Gender
## Admitted Male        62.1 63.0 36.9 33.1 27.7  5.9
##          Female      82.4 68.0 34.1 34.9 23.9  7.0
## Rejected Male        37.9 37.0 63.1 66.9 72.3 94.1
##          Female      17.6 32.0 65.9 65.1 76.1 93.0
```

```
# Even though the overall admission rate was higher for males from 1(b)
# The admissions rate for females is higher in 4 out of the 6 departments
```

d. Carry out the chi-square test of the null hypothesis of independence between the genders of applicants and the departments to which they applied. (This will be a test for marginal independence, ignoring the applicants' admission status.) What do you conclude?

```
gender_dept_table <- margin.table(UCBAdmissions, margin = c(2, 3))
chi_square_test <- chisq.test(gender_dept_table)
chi_square_test
```

```
##
##  Pearson's Chi-squared test
##
## data:  gender_dept_table
## X-squared = 1068.4, df = 5, p-value < 2.2e-16
```

```
# Since p-value is less than 0.05, we reject the null hypothesis at the 5% significance level
# Hence, there is a statistically significant relationship b/w gender & admission rates across departme
```

Problem 2

How many roots does the function have? What are they? Describe the behavior of Newton's method (see Lecture 12) applied to this function if the initial guess is each of the following:

3 3.2 2.99 3.01

a. Determine the number of rows and columns.

```
# This equation has one root at x=3. exp(-x) is always greater than zero.


x0 <- c(3, 3.2, 2.99, 3.01)  # Initial guesses
tolerance <- 0.000001

for (initial_x in x0) {
  x <- initial_x
  f <- (x ^ 2 - 6 * x + 9) * exp(-x)
  iteration <- 0
  while (abs(f) > tolerance) {
    f.prime <- (-x ^ 2 + 8 * x - 15) * exp(-x)
    x <- x - f / f.prime
    f <- (x ^ 2 - 6 * x + 9) * exp(-x)
    iteration <- iteration + 1
```

```
    cat("Iteration:", iteration, "x:", x, "\n")
  }
  cat("Root found for initial guess", initial_x, "is", x, "\n")
}
```

```
## Root found for initial guess 3 is 3
## Iteration: 1 x: 3.088889
## Iteration: 2 x: 3.042377
## Iteration: 3 x: 3.02073
## Iteration: 4 x: 3.010256
## Iteration: 5 x: 3.005102
## Iteration: 6 x: 3.002544
## Root found for initial guess 3.2 is 3.002544
## Iteration: 1 x: 2.994975
## Iteration: 2 x: 2.997481
## Root found for initial guess 2.99 is 2.997481
## Iteration: 1 x: 3.004975
## Iteration: 2 x: 3.002481
## Root found for initial guess 3.01 is 3.002481
```

```
# Recall that x=3 is the root
# For all initial guesses, the iteration stops as the value becomes closer to 3
# Specifically, if the guess is within the specified tolerance level 0.000001
```