Medical Statistics – Answers lab 4

Part 1: Inference for categorical data

Question 1

Using the data provided in the table, calculate an approximate 95% confidence interval for the difference in proportions of post-surgical complications between smokers and non-smokers.

Answer question 1

- Step 1: Extract the data
  - Smokers with complications: $x_1 = 8$, Total smokers: $n_1 = 20$
  - Non-smokers with complications: $x_2 = 10$, Total non-smokers: $n_2 = 60$
- Step 2: Calculate the sample proportions
  - $p_1 = \frac{x_1}{n_1} = \frac{8}{20} = 0.4$
  - $p_2 = \frac{x_2}{n_2} = \frac{10}{60} \approx 0.167$
- Step 3: Compute the difference in proportions
  - Difference $= p_1 - p_2 = 0.4 - 0.167 \approx 0.233$
- Step 4: Calculate the standard error (SE) of the difference
  - $\text{SE} = \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$
  - $\text{SE} = \sqrt{\frac{0.4(1-0.4)}{20} + \frac{0.167(1-0.167)}{60}} \approx 0.120$
- Step 5: Determine the 95% confidence interval
  - $Z_{critical} = 1.96$
  - Lower bound $=$ Difference $- Z_{critical} \times \text{SE} = 0.233 - 1.96 \times 0.0.120 \approx -0.001$
  - Upper bound $=$ Difference $+ Z_{critical} \times \text{SE} = 0.233 + 1.96 \times 0.120 \approx 0.468$

- Step 6: Conclusion
  - The 95% confidence interval is approximately $(-0.001, 0.468)$

Question 2

Based on the 95% confidence interval, can we conclude that there is a statistically significant difference in the proportion of post-surgical complications between smokers and non-smokers?

Answer question 2

The 95% confidence interval for the difference in proportions is $(-0.001, 0.468)$. Since this interval includes zero, we cannot reject the null hypothesis that the two proportions are equal. Therefore, we do not have sufficient evidence to conclude that there is a statistically significant difference in the proportion of post-surgical complications between smokers and non-smokers.

Results of the two-sample test of proportions based on the normal approximation with continuity correction:

```r
# Create a contingency table
complications <- matrix(c(8, 12, 10, 50), nrow = 2, byrow = TRUE)
colnames(complications) <- c("Complication", "No Complication")
rownames(complications) <- c("Smokers", "Non-smokers")
complications <- as.table(complications)

prop.test(complications)
```

```
    2-sample test for equality of proportions with continuity correction

data:  complications
X-squared = 3.4409, df = 1, p-value = 0.0636
alternative hypothesis: two.sided
95 percent confidence interval:
 -0.03449898  0.50116565
sample estimates:
   prop 1    prop 2
0.4000000 0.1666667
```

Question 3

Based on the results of the test, can we conclude that there is a statistically significant difference in the proportion of post-surgical complications between smokers and non-smokers?

Answer question 3

The p-value from the two-sample test of proportions is 0.064. Since this p-value is greater than the significance level of 0.05, we do not have sufficient evidence to reject the null hypothesis. Therefore, we cannot conclude that there is a statistically significant difference in the proportion of post-surgical complications between smokers and non-smokers.

Question 4

In addition to the p-value, output of the `prop.test()` function also provides an approximate 95% confidence interval for the difference in proportions. How does this confidence interval compare to the one you calculated manually?

Answer question 4

The 95% confidence interval for the difference in proportions was $(-0.001, 0.468)$. This is an approximate interval without continuity correction. The confidence interval provided in the output above is slightly wider due to the continuity correction. Without continuity correction (`adjust = FALSE` in R), the confidence interval provided by the `prop.test()` function is the same as the one that was calculated manually. SPSS gave the same result as R.

Checking of assumptions

Exercise

Check this assumption by calculating the expected counts for each cell in the contingency table.

Question 5

Is it reasonable to use the normal approximation in this case?

Answer question 5

The expected counts for each cell in the contingency t able are as follows:

```
chisq_test_overall <- chisq.test(complications)
chisq_test_overall$expected
```

```
            Complication No Complication
Smokers              4.5            15.5
Non-smokers         13.5            46.5
```

One of the cells in the contingency table has an expected count just below 5, which indicates that the two sample Z-test may not be fully accurate in this case.

Fisher's exact test

```
fisher.test(complications)
```

```
    Fisher's Exact Test for Count Data

data:  complications
p-value = 0.05967
alternative hypothesis: true odds ratio is not equal to 1
95 percent confidence interval:
  0.9154957 11.7051187
sample estimates:
odds ratio
  3.274581
```

Question 6

Based on the results of Fisher's exact test, can we conclude that there is a statistically significant difference in the proportion of post-surgical complications between smokers and non-smokers?

Answer question 6

The p-value from Fisher's exact test is 0.060. Since this p-value is greater than the significance level of 0.05, we do not have sufficient evidence to reject the null hypothesis. Therefore, we cannot conclude that there is a statistically significant difference in the proportion of post-surgical complications between smokers and non-smokers.

Vaccine side effects across age groups

Chi-square test of homogeneity

```
# Create a contingency table
side_effects <- matrix(c(50, 30, 10, 40, 40, 20, 30, 50, 40), nrow = 3, byrow = TRUE)
colnames(side_effects) <- c("None", "Mild", "Severe")
rownames(side_effects) <- c("18-39", "40-59", "60+")
side_effects <- as.table(side_effects)
side_effects
```

```
      None Mild Severe
18-39   50   30     10
40-59   40   40     20
60+     30   50     40
```

```
# Perform the chi-square test of homogeneity
chisq_test_overall <- chisq.test(side_effects)
print(chisq_test_overall)
```

```
    Pearson's Chi-squared test

data:  side_effects
X-squared = 25.136, df = 4, p-value = 4.723e-05
```

Question 7

Based on the results of the chi-square test, can we conclude that the distribution of vaccine side effects is consistent across the three age groups?

Answer question 7

The p-value from the chi-square test of homogeneity is <0.0001. Since this p-value is less than the significance level of 0.05, we have sufficient evidence to reject the null hypothesis. Therefore, we can conclude that the distribution of vaccine side effects is not consistent across the three age groups.

Checking of assumptions

Question 8

Are the expected cell counts greater than 5 for the different cells in the contingency table?

Expected cell counts:

```
# Retrieve the table of expected counts
chisq_test_overall$expected
```

```
          None     Mild   Severe
18-39 34.83871 34.83871 20.32258
40-59 38.70968 38.70968 22.58065
60+   46.45161 46.45161 27.09677
```

Answer question 8

The expected counts in all cells are well above 5, which indicates that the normal approximation is appropriate in this case.

Post-hoc pairwise comparisons

Comparison of age groups 18–39 and 40–59

```
table_12 <- side_effects[c("18-39", "40-59"), ]
table_12
```

```
      None Mild Severe
18-39   50   30     10
40-59   40   40     20
```

```r
# Perform chi-square test for the subset of data
chisq_test_12 <- chisq.test(table_12)
print(chisq_test_12)
```

```
    Pearson's Chi-squared test

data:  table_12
X-squared = 5.3616, df = 2, p-value = 0.06851
```

```r
# Adjust the p-value for multiple testing
p_adjusted_12 <- 3*chisq_test_12$p.value
```

The Bonferroni-corrected p-value for this comparison is 0.206.

> Exercise
>
> Perform the pairwise comparison between the other two pairs of age groups (40–59 and 60+, 18–39 and 60+) using the same approach.

Comparison of age groups 18–39 and 60+

```r
table_13 <- side_effects[c("18-39", "60+"), ]
table_13
```

```
      None Mild Severe
18-39   50   30     10
60+     30   50     40
```

```r
# Perform chi-square test for the subset of data
chisq_test_13 <- chisq.test(table_13)
print(chisq_test_13)
```

```
    Pearson's Chi-squared test

data:  table_13
X-squared = 24.208, df = 2, p-value = 5.536e-06
```

```r
# Adjust the p-value for multiple testing
p_adjusted_13 <- 3*chisq_test_13$p.value
```

The Bonferroni-corrected p-value for this comparison is <0.0001.

Comparison of age groups 40-59 and 60+

```r
table_23 <- side_effects[c("40-59", "60+"), ]
table_23
```

```
      None Mild Severe
40-59   40   40     20
60+     30   50     40
```

```r
# Perform chi-square test for the subset of data
chisq_test_23 <- chisq.test(table_23)
print(chisq_test_23)
```

```
    Pearson's Chi-squared test

data:  table_23
X-squared = 7.4497, df = 2, p-value = 0.02412
```

```r
# Adjust the p-value for multiple testing
p_adjusted_23 <- 3*chisq_test_23$p.value
```

The Bonferroni-corrected p-value for this comparison is 0.072.

Question 9

Based on the results of the pairwise comparisons, which age groups have significantly different distributions of side effects?

Answer question 9

Based on the adjusted p-values from the pairwise comparisons, we can conclude that the distributions of side effects are significantly different between the age groups 18–39 and 60+ (Bonferroni-corrected p-value < 0.001). However, there were no significant differences between the age groups 40–59 and 60+ (Bonferroni-corrected p-value $= 0.072$) or between

the age groups 18–39 and 40–59 (Bonferroni-corrected p-value = 0.206).

Part 2: Analysis of paired continous data

Results of the paired t-test:

```r
# Perform paired t-test
t_test <- t.test(pockets$pocket_depth_before,
                 pockets$pocket_depth_after,
                 paired = TRUE)
print(t_test)
```

```
    Paired t-test

data:  pockets$pocket_depth_before and pockets$pocket_depth_after
t = 11.133, df = 95, p-value < 2.2e-16
alternative hypothesis: true mean difference is not equal to 0
95 percent confidence interval:
 0.09140772 0.13108091
sample estimates:
mean difference
      0.1112443
```

Question 10

Based on the results of the paired t-test, can we conclude that the intervention significantly reduces pocket depth?

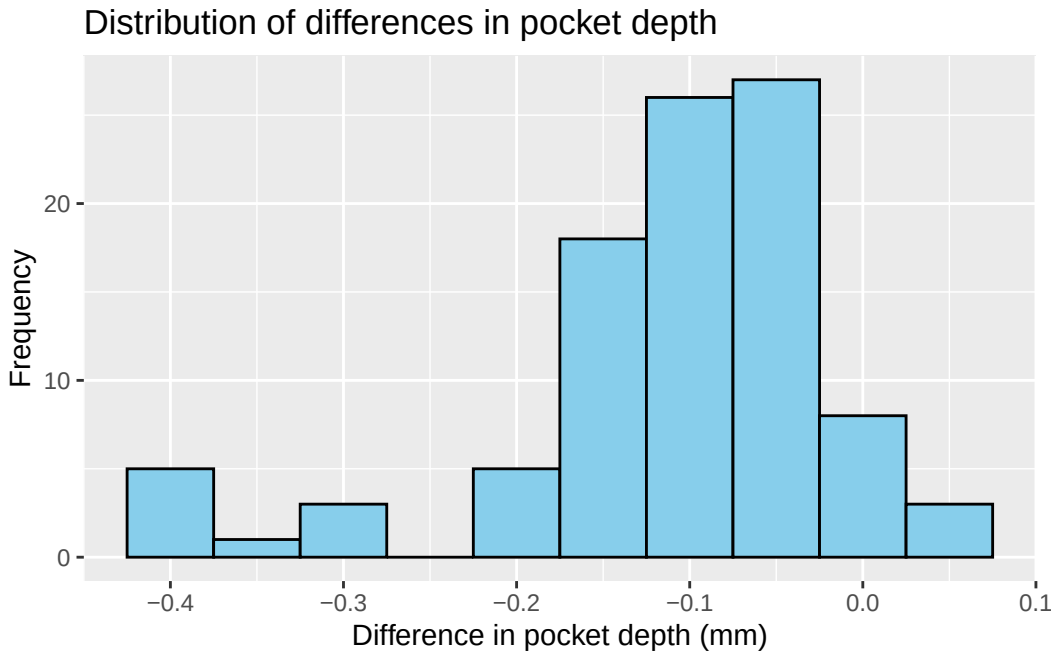
Answer question 10

The p-value from the paired t-test is <0.0001, which is less than the significance level of 0.05. Therefore, we have sufficient evidence to reject the null hypothesis and conclude that the intervention significantly reduces pocket depth.

Checking of assumptions

```r
# Calculate the differences in pocket depth
pockets$diff <- pockets$pocket_depth_after -
                pockets$pocket_depth_before

# Create a histogram of the differences
library(ggplot2)
ggplot(pockets, aes(x = diff)) +
  geom_histogram(binwidth = 0.05,
                 fill = "skyblue",
                 color = "black") +
  labs(title = "Distribution of differences in pocket depth",
       x = "Difference in pocket depth (mm)",
       y = "Frequency")
```



Question 11

Based on the histogram, do the differences in pocket depth appear to be approximately normally distributed?

Answer question 11

The histogram of the differences in pocket depth is left-skewed, indicating that the distribution is not normal and that we may therefore be better of using a sign test.

Sign test

```
# Perform the sign test
# There are 7 positive signs out of 96 pairs with either a positive or negative sign
sign_test <- binom.test(7, 96, p = 0.5)
print(sign_test)
```

```
    Exact binomial test

data:  7 and 96
number of successes = 7, number of trials = 96, p-value < 2.2e-16
alternative hypothesis: true probability of success is not equal to 0.5
95 percent confidence interval:
 0.02981784 0.14447971
sample estimates:
probability of success
            0.07291667
```

Question 12

Based on the results of the sign test, can we conclude that the intervention significantly reduces pocket depth?

Answer question 12

The p-value from the sign test is <0.0001, which is less than the significance level of 0.05. Therefore, we have sufficient evidence to reject the null hypothesis and conclude that the intervention significantly reduces pocket depth.

Wilcoxon signed-rank test

```r
# Perform the Wilcoxon signed-rank test
wilcox_test <- wilcox.test(pockets$pocket_depth_before,
                           pockets$pocket_depth_after,
                           paired = TRUE)
print(wilcox_test)
```

```
    Wilcoxon signed rank test with continuity correction

data:  pockets$pocket_depth_before and pockets$pocket_depth_after
V = 4572, p-value = 2.435e-16
alternative hypothesis: true location shift is not equal to 0
```

Question 13

Based on the results of the Wilcoxon signed-rank test, can we conclude that the intervention significantly reduces pocket depth?

Answer question 13

The p-value from the Wilcoxon signed-rank test is <0.0001, which is less than the significance level of 0.05. Therefore, we have sufficient evidence to reject the null hypothesis and conclude that the intervention significantly reduces pocket depth.

Part 3: Analysis of paired dichotomous data

```r
# Create the table for DNCB and Croton Oil responses
skin_response_table <- matrix(c(81, 23, 48, 21), nrow = 2, byrow = TRUE)
colnames(skin_response_table) <- c("DNCB +ve", "DNCB -ve")
rownames(skin_response_table) <- c("Croton Oil +ve", "Croton Oil -ve")
skin_response_table <- as.table(skin_response_table)

# Print the table
skin_response_table
```

```
               DNCB +ve DNCB -ve
Croton Oil +ve       81       23
Croton Oil -ve       48       21
```

Results of the McNemar test:

```r
# Perform the McNemar test
mcnemar_test <- mcnemar.test(skin_response_table)
print(mcnemar_test)
```

```
    McNemar's Chi-squared test with continuity correction

data:  skin_response_table
McNemar's chi-squared = 8.1127, df = 1, p-value = 0.004396
```

Question 14

Based on the results of the McNemar test, can we conclude that there is a significant difference in the proportions of patients with a negative response to DNCB and croton oil? If so, can you determine which substance is associated with a higher proportion of negative responses?

Answer question 14

The p-value from the McNemar test is 0.0044, which is less than the significance level of 0.05. Therefore, we have sufficient evidence to reject the null hypothesis and conclude that there is a significant difference in the proportions of patients with a negative response to DNCB and croton oil. The test does not provide information on which substance is associated with a higher proportion of negative responses. To address this question, we can calculate the estimated proportion of negative responses for each substance, which is $44/173 = 0.25$ for DNCB and $69/173 = 0.40$ for croton oil. Therefore, croton oil is associated with a higher proportion of negative responses.