**Assignment 2**

Question 5

1. A = = 0.45

B = = 0.55

Posterior mean = A \* θ + B \*

For the **high** dose group:  
0.45\*4.04 + 0.55\*4.25

Posterior mean is 4.15

For the **low** dose group:  
0.45\*3.83 + 0.55\*3.75

Posterior mean is 3.79

Posterior standard deviation = = 0.22

So, the posterior distributions are:

µhigh dose | Xhigh dose ~ N(mean=4.15, standard deviation = 0.22)

µlow dose | Xlow dose ~ N(mean=3.79, standard deviation = 0.22)

1. Mean = difference of means between two groups (4.15-3.79) = 0.36units/mL

Variance for the high dose group = = = 0.9

Variance for the low dose group = = = 0.9

Sum of the variances =

Standard deviation = = 0.42

So, the posterior distribution of the difference between two groups:

µhigh dose -µlow dose | data ~ N(mean=0.36, standard deviation=0.42)

1. 2.5% quantile = θ – 1.96\*τ = 0.36 – 1.96\*0.42

97.5% quantile = θ + 1.96\*τ = 0.36 + 1.96\*0.42

CrI 95% (-0.46, 1.18)

The credibility interval is large, and as it encompasses negative and positive values, it suggests that the true mean difference between groups may be outside of the zone of clinical significance (which would be an increase in titers > 1, as stated in problem 3 of assignment 2).

Posterior probability of the difference in titers having clinical significance:  
P(µ>1): 1-pnorm(1,0.36,0.42) = 0.064

By calculating the posterior probability of clinically meaningful significance between two groups, we have 6.4% of chance of having a titer greater than 1.

1. Considering the following non-informative prior distribution:

µ ~ N(mean=100, standard deviation = 50)

A = ≈ 0

B = ≈ 1

B > A, meaning that the posterior mean will be weighted more towards the observed mean than the prior non-informative mean.

Posterior mean = A \* θ + B \*

For the **high** dose group:  
0\*4.04 + 1\*4.25

Posterior mean is 4.25

For the **low** dose group:  
0\*3.83 + 1\*3.75

Posterior mean is 3.75

Posterior standard deviation = = 0.3

So, the posterior distributions with non informative means would be:

µhigh dose | Xhigh dose ~ N(mean=4.25, standard deviation = 0.3)

µlow dose | Xlow dose ~ N(mean=3.75, standard deviation = 0.3)

1. Mean = (4.25-3.75) = 0.5units/mL

Variance for the high dose and low dose groups are the same, as is unchanged = = = 0.9

Sum of the variances = 0.9+0.9= 0.18

Standard deviation = = 0.42

So, the posterior distribution of the difference between two groups:

µhigh dose -µlow dose | data ~ N(mean=0.5, standard deviation=0.42)

1. 2.5% quantile = θ – 1.96\*τ = 0.5 – 1.96\*0.42

97.5% quantile = θ + 1.96\*τ = 0.5 + 1.96\*0.42

CrI 95% (-0.32,1.32)

The credible interval is wider in both directions, providing less certainty of where the value of true difference in titers lays.

**Assignment 3**

Question 1

1. Epinephrine’s safety for out-of-hospital cardiac arrest in adult patients  
   Intervention arm: parenteral epinephrine

Control arm: saline placebo

1. Rate of survival at 30 days post event
2. N = 8000, risk ratio 1.25

|  |  |  |  |
| --- | --- | --- | --- |
|  | Survival at 30 days | Death at 30 days | Total |
| Epinephrine | 7.5% - 300 | 92.5% - 3700 | 4000 |
| Placebo | 6% - 240 | 94% - 3760 | 4000 |
| Total | 540 | 7460 | 8000 |

Odds ratio = 1.27

Standard deviation for the risk ratio: = 0.08366

CI 95% = In(relative risk) – 1.96 x 0.08366, In(relative risk) + 1.96 x 0.08366

CI 95% = 1.25 – 1.96\*0.08366, 1.25 + 1.96\*0.08366

CI 95% = (1.086, 1.413)

1. power.prop.test(4000,p1=0.075,p2=0.06,sig.level=1-0.95,alternative="one.sided")

The power for the planned sample was 84.8%

1. Recruited patients and results

|  |  |  |  |
| --- | --- | --- | --- |
|  | Survival at 30 days | Death at 30 days | Total |
| Epinephrine | 3.2% - 130 | 3882 | 4012 |
| Placebo | 2.4% - 94 | 3901 | 3995 |
| Total | 214 | 7793 | 8007 |

Although the sample sizes are slightly different, they are close to the original estimate. As such I chose to still use the power.prop.test(), that requires equal sample sizes:

power.prop.test(n=4000, p1=0.032, p2=0.024, sig.level=1-0.95, alternative="one.sided")

The power based on the real frequencies observed is ≈ 70%.

1. Using the table from the previous item, we have that:

Odds ratio:

SD for the odds ratio: = 0.137

CI 95% = (1.39\*exp(-1.96\*0.137), 1.39\*exp(+1.96\*0.137))

CI 95% = (1.06, 1.79)

1. Reporting using odds ratio was appropriate. As the number of patients who survived is much smaller than those who died, the odds ratio is a reliable approximation of the relative risk.
2. An odds ratio of 1 indicates that there is no difference to the placebo and the treatment groups, and likely the number of survivals and death in each group is the same. If we carry out a Chi square test, the value of p will be 1. An approximate test is adequate due to the large number of subjects (a, b, c and d on the 2x2 table are greater than 5).

**Question 2**

1. Observed values

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | MD | RN | Conventional health care provider | Complementary therapy | Total |
| Rural area | 57 | 7 | 3 | 52 | 119 |
| Small city | 44 | 23 | 12 | 74 | 153 |
| Large city | 35 | 18 | 10 | 63 | 126 |
| Total | 136 | 48 | 25 | 189 | 398 |

To calculate the chi-square, we have to obtain the expected value for each cell, using the formula expected value = (column total) x (row total) / overall total

Expected values

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | MD | RN | CHCP | CT |
| Rural area | 40.663 | 14.351 | 7.474 | 56.510 |
| Small city | 52.281 | 18.452 | 9.610 | 72.655 |
| Large city | 43.055 | 15.195 | 7.914 | 59.834 |

To obtain chi-square:

1. Observed values

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | MD | RN | CHCP | CT | Total |
| Never | 29 | 9 | 1 | 14 | 53 |
| Sometimes | 84 | 33 | 6 | 29 | 152 |
| Often/always | 25 | 15 | 18 | 137 | 195 |
| Total | 138 | 57 | 25 | 180 | 400 |

Expected values

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | MD | RN | CHCP | CT |
| Never | 18.285 | 7.5525 | 3.3125 | 23.85 |
| Sometimes | 52.44 | 21.66 | 9.5 | 68.4 |
| Often/always | 67.275 | 27.7875 | 12.1875 | 87.75 |

Given that the expected and observed values for the row “Never” and the column “Conventional health care provider” are smaller than 5, it would be more appropriate to use the Fisher’s exact test.

1. a=matrix(c(29,9,1,14,84,33,6,29,25,15,18,137),ncol=4,byrow=T)

fisher.test(a,simulate.p.value = TRUE,alternative = "two.sided")

Fisher's Exact Test for Count Data with simulated p-value (based on 2000 replicates)

data: a

p-value = 0.0004998

alternative hypothesis: two.sided

We have enough evidence to reject the null hypothesis, which would be that there is no difference between patient-provider communication among health care providers.

**Question 3**

1. Analyze the pseudovirus-neutralizing activity of an mRNA vaccine by comparing the immunologic response to different doses in mice, administered at 0 and 3 weeks.
2. H0: μlow dose = μhigh dose vs. HA μlow dose ≠ μhigh dose

(assigned “L” for mice that received lower dose and “H” for higher dose)

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Titer | 40 | 40 | 40 | 40 | 40 | 40 | 116.4 | 138.7 | 173.4 | 550.9 | 631.2 | 637.3 | 643.6 | 730.2 | 1152.7 | 1441.2 | 1483.9 | 2556.4 | 4534.6 | 12453 |
| Group | L | L | L | L | L | L | L | L | L | H | H | H | H | H | H | H | H | H | H | L |
| Rank | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |

Tc = 1+2+3+4+5+6+7+8+9+20 = 65

Z statistic = = = -3.023716

P-value = 0.002

We have enough evidence to **reject** the null hypothesis.

1. As done in the methods, I used the log() of the titers. The log(highdose) and log(lowdose), when in a Q-Qplot, have a distribution closer to normal.

logb=log(highdose)

loga=log(lowdose)

t.test(loga,logb,alternative = 'two.sided', paired = TRUE, var.equal = FALSE)

t = -4.4652, df = 9, **p-value = 0.001566**

Based on this test, we can **reject** the null hypothesis and accept the alternative hypothesis. In this case, it is appropriate to use a t-test given that the sample size is small.

1. BOOTSTRAP ON LECTURE 10  
   High dose group:

mb = median(logb) = 3.68

sb = sd(logb) = 0.7

qt(1-0.2/2, 9) = 1.38

CI 80% (mb-1.38\*(sb/sqrt(10)), mb+1.38\*(sb/sqrt(10))

CI 80% (6.82-1.38\*0.22), (6.82+1.38\*0.22)

CI 80% (6.51, 7.12)

Low dose group:

ma = median(loga) = 3.68

sa = sd(loga) = 1.78

qt(1-0.2/2, 9) = 1.38

CI 80% (ma-1.38\*(sa/sqrt(10)), ma+1.38\*(sa/sqrt(10))

CI 80% (3.68-1.38\*0.56), (3.68+1.38\*0.56)

CI 80% (2.91, 4.46)