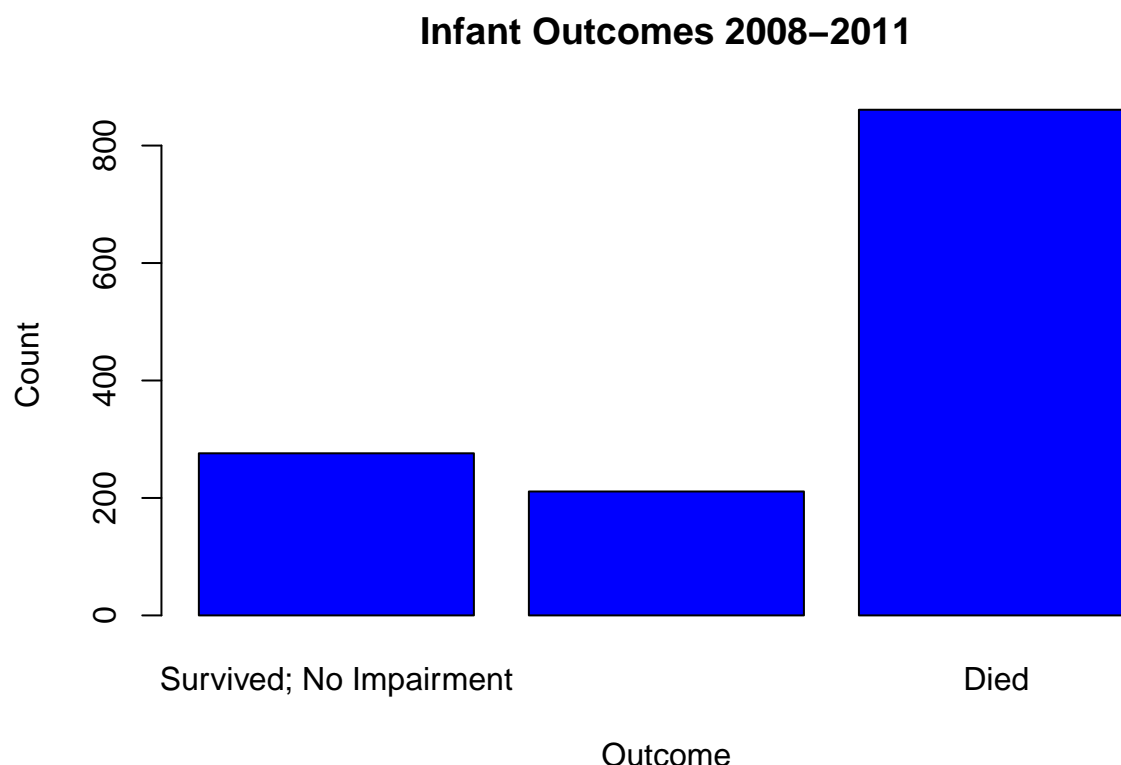


Sample Solution for Homework 1

Xu Chen

1. Analysis of periviable neonates (8 points)

In a 2017 New England Journal of Medicine paper, Duke researchers examined outcomes of infants born at 22-24 weeks of gestation. Of 1348 infants in the study born at these gestational ages in 2008-2011, 861 died, 211 survived with neurodevelopmental impairment, and 276 survived without neurodevelopmental impairment.



Suppose historical rates of all three outcomes (from the early 2000's) are 15% survival without impairment, 15% survival with impairment, and 70% death.

(a) Estimate the proportions of infants surviving without impairment, surviving with impairment, and dying, providing point and interval estimates.

Here we use MLE as point estimate. Let π_1, π_2 and π_3 denote the proportion of infants surviving without impairment, surviving with impairment, and dying respectively. The MLEs are given by

$$\hat{\pi}_i = \frac{n_i}{n} \quad \text{for } i = 1, 2, 3$$

```
n <- 1348
ni <- c(276, 211, 861)
```

```
pi.hat <- ni / n
print(pi.hat)
```

```
## [1] 0.2047478 0.1565282 0.6387240
```

The marginal distributions of the number of infants surviving without impairment, surviving with impairment, and dying are binomial distributions. One may construct confidence intervals for these three proportions *individually* using Wald, score or likelihood ratio test statistics. We provide Wald intervals for π_1, π_2 and π_3 . Exact confidence interval (CI) and CIs based on score and likelihood ratio test statistics for π_1 are included to illustrate how to compute them in practice.

A 95% Wald confidence interval (CI) of π_i is given by

$$\left[\hat{\pi}_i - z_{0.025} \sqrt{\frac{\hat{\pi}_i(1 - \hat{\pi}_i)}{n}}, \hat{\pi}_i + z_{0.025} \sqrt{\frac{\hat{\pi}_i(1 - \hat{\pi}_i)}{n}} \right]$$

```
wald.ci <- pi.hat + (qnorm(0.975) * sqrt(pi.hat * (1 - pi.hat)/n))%o%c(-1,1)
wald.ci
```

```
##           [,1]      [,2]
## [1,] 0.1832068 0.2262887
## [2,] 0.1371312 0.1759252
## [3,] 0.6130804 0.6643677
```

Exact CI

A 95% Clopper-Pearson CI for π_1 is constructed by

$$\left\{ \pi_1 : \sum_{k=0}^{n_1} \binom{n}{k} \pi_1^k (1 - \pi_1)^{n-k} > \frac{0.05}{2} \quad \text{and} \quad \sum_{k=n_1}^n \binom{n}{k} \pi_1^k (1 - \pi_1)^{n-k} > \frac{0.05}{2} \right\}$$

Using the formula in section 16.6.1 in @agresti2013, this interval is

```
1/c((1+(n-ni[1]+1)/ni[1]/qf(1-0.975,df1 = 2*ni[1],df2=2*(n-ni[1]+1))), (1+(n-ni[1])/(ni[1]+1)/qf(1-0.025,
```

```
## [1] 0.1834981 0.2272927
```

binom.test can also be used.

```
binom.test(x=ni[1],n=n)$conf.int
```

```
## [1] 0.1834981 0.2272927
## attr(,"conf.level")
## [1] 0.95
```

Score CI

A 95% CI of π_1 based on score test statistic is given by inverting the test as follows

$$\left\{ \pi_1 : \left| \frac{\hat{\pi}_1 - \pi_1}{\sqrt{\pi_1(1 - \pi_1)/n}} \right| < z_{0.025} \right\}.$$

We need to solve a quadratic equation for π_1 to find the endpoints of the interval or directly use formula (1.14) in @agresti2013.

```
a <- (n + qchisq(0.95, 1)) / n
b <- -qchisq(0.95, 1) / n - 2 * ni[1] / n
c <- ni[1] ^ 2 / n ^ 2
(-b + c(-1, 1) * sqrt(b ^ 2 - 4 * a * c)) / 2 / a
```

```
## [1] 0.1840601 0.2271135
```

prop.test can be used for score CI.

```
prop.test(x=ni[1],n=n,correct = F)$conf.int
```

```
## [1] 0.1840601 0.2271135
## attr("conf.level")
## [1] 0.95
```

Likelihood ratio CI

A 95% CI based on likelihood ratio test is

$$\left\{ \pi_1 : 2 \left[n_1 \log \frac{\hat{\pi}_1}{\pi_1} + (n - n_1) \log \frac{1 - \hat{\pi}_1}{1 - \pi_1} \right] < \chi_1^2(0.05) \right\}$$

```
LR <- function(pi1, n1, n, alpha){
  pi1.hat <- n1 / n
  2 * (n1 * log(pi1.hat / pi1) + (n - n1) * log((1 - pi1.hat) / (1 - pi1))) - qchisq(1 - alpha, 1)
}
c(uniroot(f=LR, interval = c(0.1,0.2), n1=ni[1],n=n,alpha=0.05)$root, uniroot(f=LR, interval = c(0.2,0.3), n1=ni[1],n=n,alpha=0.05)$root)
```

```
## [1] 0.1837735 0.2268302
```

It is clear that different types of CIs are quite similar as counts in all categories are moderate and hence normal approximation is appropriate.

(b) Evaluate the hypothesis that the current rates are equal to the historical rates of 15, 15, and 70%, respectively, using the multinomial distribution.

Pearson χ^2 test

```
chisq.test(x=ni,p=c(0.15,0.15,0.7))
```

```
##
## Chi-squared test for given probabilities
##
## data:  ni
## X-squared = 34.549, df = 2, p-value = 3.145e-08
```

The p -value is 3.15×10^{-8} .

Likelihood ratio χ^2 test

$$G^2 = 2 \sum_{j=1}^3 n_j \log(n_j / \mu_j) = 31.9801$$

```
chisq.lr <- 2*sum(ni*log(ni/n/c(0.15,0.15,0.7)))
1-pchisq(chisq.lr,2)
```

```
## [1] 1.136612e-07
```

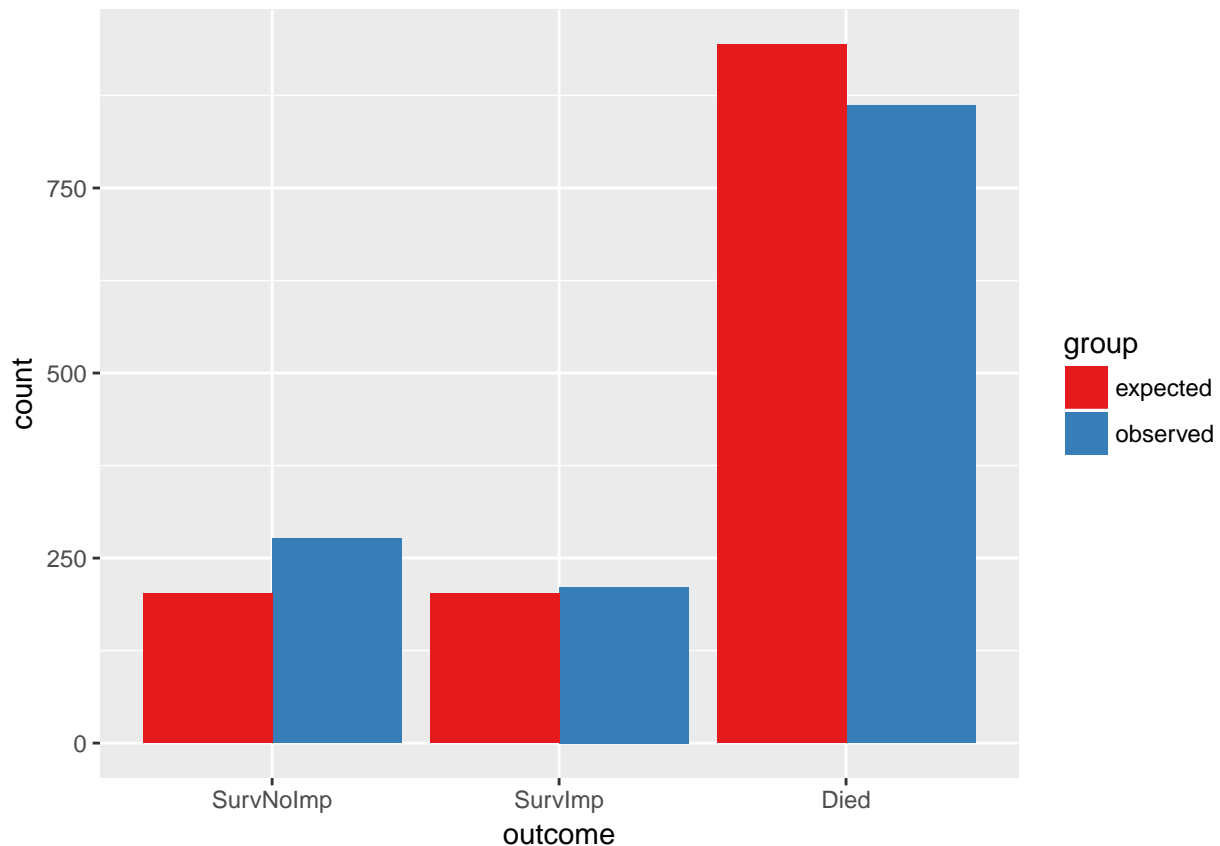
The p -value is 1.14×10^{-7} .

Therefore, the p -values given by both tests are very small and the null hypothesis is rejected under level $\alpha = 0.05$. The data provides significant evidence that the proportions in recent years are not equal to the historical rates.

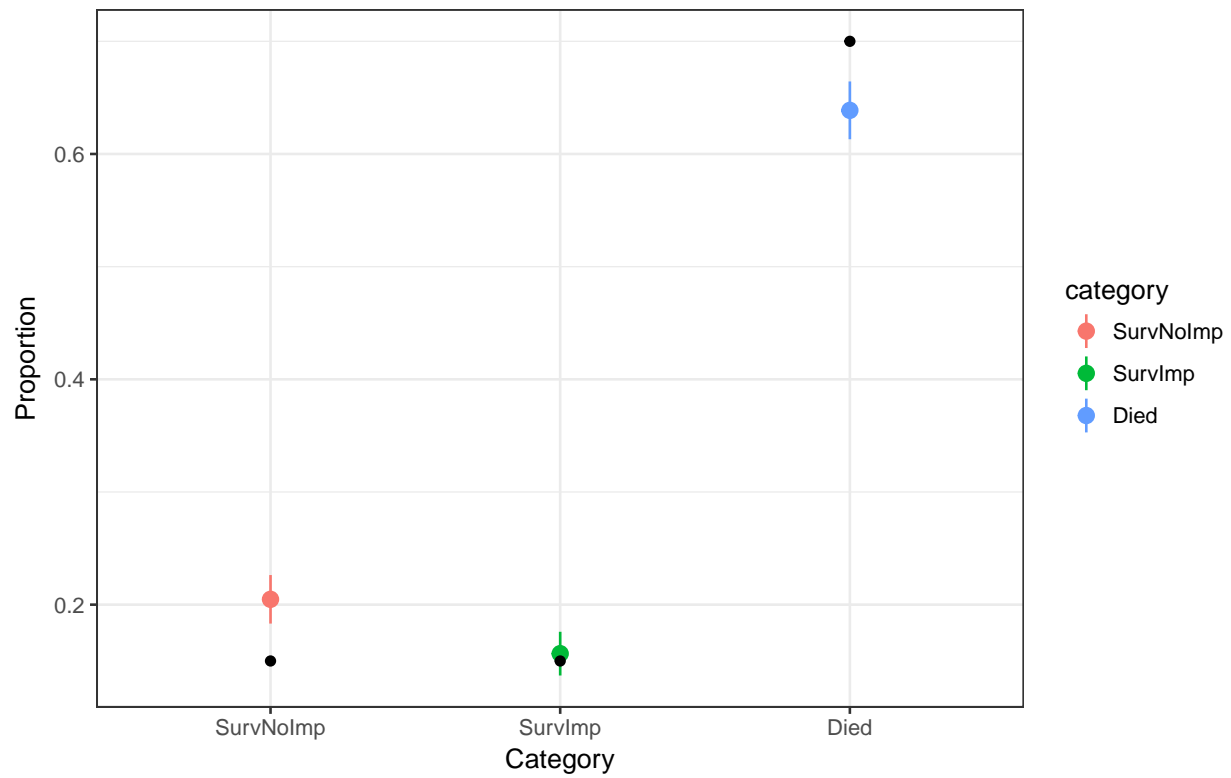
(c) Display results using informative graphics.

The small p -value indicates that the data provides evidence against the null hypothesis. Informative graphs should capture the departure from data to null hypothesis. A histogram comparing observed counts and expected counts under the null and a graph showing CIs of some multinomial parameters does not cover the null are typical choices.

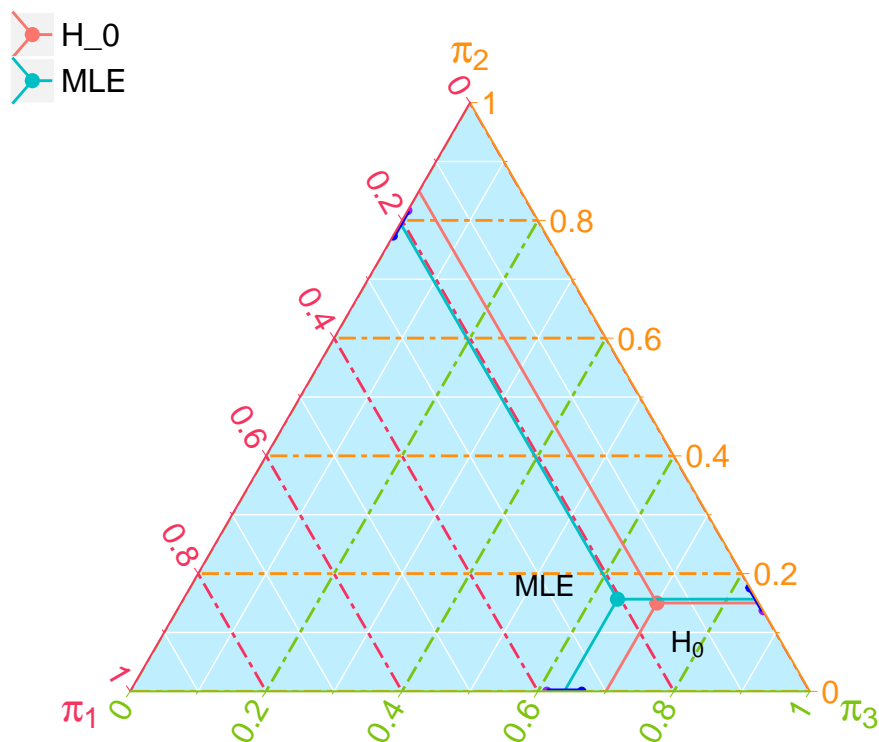
(1) The histogram below compares expected counts and observed counts for all categories.



(2) Here we plot the point estimates and 95% Wald intervals against the values of parameters under the null hypothesis. Clearly, Wald intervals of π_1 and π_3 do not cover the values of parameters under the null hypothesis.



- (3) We provide a ternary graph of multinomial parameters given by MLE and null hypothesis. On each edge of this triangle, we plot the 95% Wald interval (purple line segment) of the parameter. Again, two of the CIs do not cover the values of parameters under the null hypothesis.



2. Agresti 1.7 (12 points total)

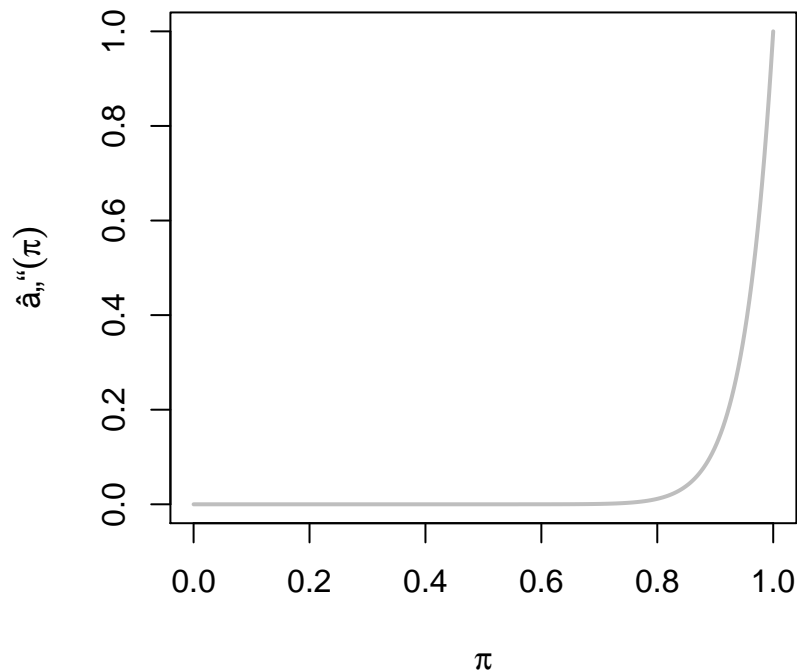
Complete the given parts and add part f, given below. In a crossover trial comparing a new drug to a standard, π denotes the probability that the new one is judged better. It is desired to estimate π and to test $H_0 : \pi = 0.50$ against $H_a : \pi \neq 0.50$. In 20 independent observations, the new drug is better each time.

2.a. (2 points)

Plot the likelihood function. Is it close to the quadratic shape that large-sample normal approximations utilize?

The likelihood function is given by

$$\ell(\pi) = \binom{20}{20} \pi^{20} (1 - \pi)^{20-20} = \pi^{20}$$



Clearly, the likelihood function does not exhibit a quadratic shape and hence large-sample normal approximation is not appropriate here.

2.b. (2 points)

Give the ML estimate of π . Conduct a Wald test and construct a 95% Wald CI for π . Are these sensible?

The ML estimate of π is 1. Since $\hat{\pi} = 1$, the Wald test statistic is $z_W = \frac{\hat{\pi} - \pi_0}{\sqrt{\hat{\pi}(1-\hat{\pi})/n}} = +\infty$ and hence p -value is 0. A 95% Wald confidence interval degenerates to a point at $\hat{\pi} = 1$.

Both of the results are not sensible because Wald test or interval are based on valid normal approximation. However, as shown in part (a), normal approximation is no longer valid in this example. Two results are overconfident and misleading as extreme situation occurs.

2.c. (2 points)

Conduct a score test, reporting the p -value. Construct a 95% score confidence interval. Interpret.

```
score.test <- prop.test(x=20,n=20,correct=F)
score.test$p.value
```

```
## [1] 7.744216e-06
```

```
score.test$conf.int
```

```
## [1] 0.8388748 1.0000000  
## attr(,"conf.level")  
## [1] 0.95
```

The p -value is 7.74×10^{-6} . Data exhibits significant evidence against the null indicating the proportion is not 0.5 and we reject the null hypothesis under level $\alpha = 0.05$. A 95% score CI is (0.8389, 1) which is more reasonable than the Wald CI.

2.d. (2 points)

Conduct a likelihood-ratio test and construct a likelihood-based 95% confidence interval. Interpret.

The likelihood-ratio test statistic is

$$-2(L_0 - L_1) = 2 \left[y \log \frac{y}{n\pi_0} + (n - y) \log \frac{n - y}{n - n\pi_0} \right] = 2 \left[20 \log \frac{20}{10} \right] = 27.7259$$

```
chisq.lr <- 40*log(2)  
pval.lr <- 1-pchisq(chisq.lr,df=1)
```

The p -value is 1.398×10^{-7} .

A 95% confidence interval based on likelihood-ratio test statistic is

$$\begin{aligned} 2 \left[y \log \frac{y}{n\pi} + (n - y) \log \frac{n - y}{n - n\pi} \right] &< \chi_1^2(0.05) \\ 2(-20 \log \pi) &< \chi_1^2(0.05) \\ \pi &> 0.9084 \end{aligned}$$

```
exp(-qchisq(0.95,1)/40)
```

```
## [1] 0.9084309
```

Therefore, a 95% likelihood-ratio confidence interval is (0.9084, 1). The p -value is smaller and CI is narrower than those given by score test.

2.e.(2 points)

Construct an exact binomial test. Interpret.

Since the binomial distribution under the null hypothesis is symmetric, the exact p -value is given by

$$p = P_{H_0}(y \geq 20) + P_{H_0}(y \leq 0) = 2 \binom{20}{20} 0.5^{20} = 1.907 \times 10^{-6}$$

```
2*0.5^20
```

```
## [1] 1.907349e-06
```


One may also use `binom.test` directly.

```
binom.test(x=20,n=20,alternative = 'two.sided')

##
## Exact binomial test
##
## data: 20 and 20
## number of successes = 20, number of trials = 20, p-value =
## 1.907e-06
## alternative hypothesis: true probability of success is not equal to 0.5
## 95 percent confidence interval:
## 0.8315665 1.0000000
## sample estimates:
## probability of success
## 1
```

Therefore we reject the null hypothesis under level $\alpha = 0.05$. Exact binomial test is more reliable and preferred here since the sample size is small. The confidence interval and p -value are close to those given by the score test.

2.f. (2 points)

Estimate a 95% HPD interval for π when a $U[0,1]$ prior is used. Interpret.

When using a $U[0,1]$ prior, the posterior is given by

$$\begin{aligned} p(\pi | y) &\propto p(y | \pi)p(\pi) \\ &\propto \pi^{20} \end{aligned}$$

Hence $\pi | y \sim \text{Be}(21, 1)$. Since $p(\pi | y)$ is monotone increasing in $[0,1]$, a HPD interval must be of the form $[l, 1]$. The lower bound is given by

```
qbeta(0.05,21,1)
```

```
## [1] 0.8670541
```

Therefore, a 95% HPD interval is given by $(0.8671, 1)$ which is close to the CI given by exact test. Bayesian method is relatively robust when observed data is extreme.

3. Agresti 1.41 (10 points)

For the Dirichlet prior for multinomial probabilities, show the Bayesian estimator of the posterior expectation is a weighted average of the prior mean and sample proportions as in Agresti (1.19).

Proof. Suppose the multinomial probabilities $\boldsymbol{\pi} = (\pi_1, \dots, \pi_c)$ have Dirichlet prior $\text{Dir}(\alpha_1, \dots, \alpha_c)$. Let (n_1, \dots, n_c) denote cell counts from the multinomial distribution $\boldsymbol{\pi}$. Then the posterior density of $\boldsymbol{\pi}$ is

$$\begin{aligned}
p(\boldsymbol{\pi} \mid n_1, \dots, n_c) &\propto \prod_{i=1}^c \pi_i^{n_i} \prod_{i=1}^c \pi_i^{\alpha_i-1} \\
&= \prod_{i=1}^c \pi_i^{n_i + \alpha_i - 1}
\end{aligned}$$

Hence the posterior distribution of $\boldsymbol{\pi}$ is $\text{Dir}(n_1 + \alpha_1, \dots, n_c + \alpha_c)$, which has mean

$$E(\pi_i \mid n_1, \dots, n_c) = \frac{n_i + \alpha_i}{n + K}$$

where $n = \sum_{i=1}^c n_i$, $K = \sum_{i=1}^c \alpha_i$.

Let $p_i = n_i/n$ and $\gamma_i = \alpha_i/K$. We have

$$\begin{aligned}
E(\pi_i \mid n_1, \dots, n_c) &= \frac{n_i}{n + K} + \frac{\alpha_i}{n + K} \\
&= \frac{n}{n + K} \cdot p_i + \frac{K}{n + K} \cdot \gamma_i
\end{aligned}$$

Therefore the posterior mean of π_i is a weighted average of sample proportion p_i and prior mean γ_i .

4. Agresti 2.26 (10 points)

Show that the OR and RR need not be similar when π_i is close to 1.0 for both groups.

According to the definitions of RR and OR, we have

$$\text{RR} = \frac{1 - \pi_1}{1 - \pi_2} \text{OR}.$$

When π_1 and π_2 are close to 1.0, the range of $\frac{1 - \pi_1}{1 - \pi_2}$ is large. For example, when $\pi_1 = 1 - 10^{-2}$ and $\pi_2 = 1 - 10^{-5}$, then $\text{RR} \approx 1$ but $\text{OR} \approx 10^{-3}$. If $\pi_1 = 1 - 10^{-5}$ and $\pi_2 = 1 - 10^{-2}$, then RR is still approximately 1 but $\text{OR} \approx 10^3$.

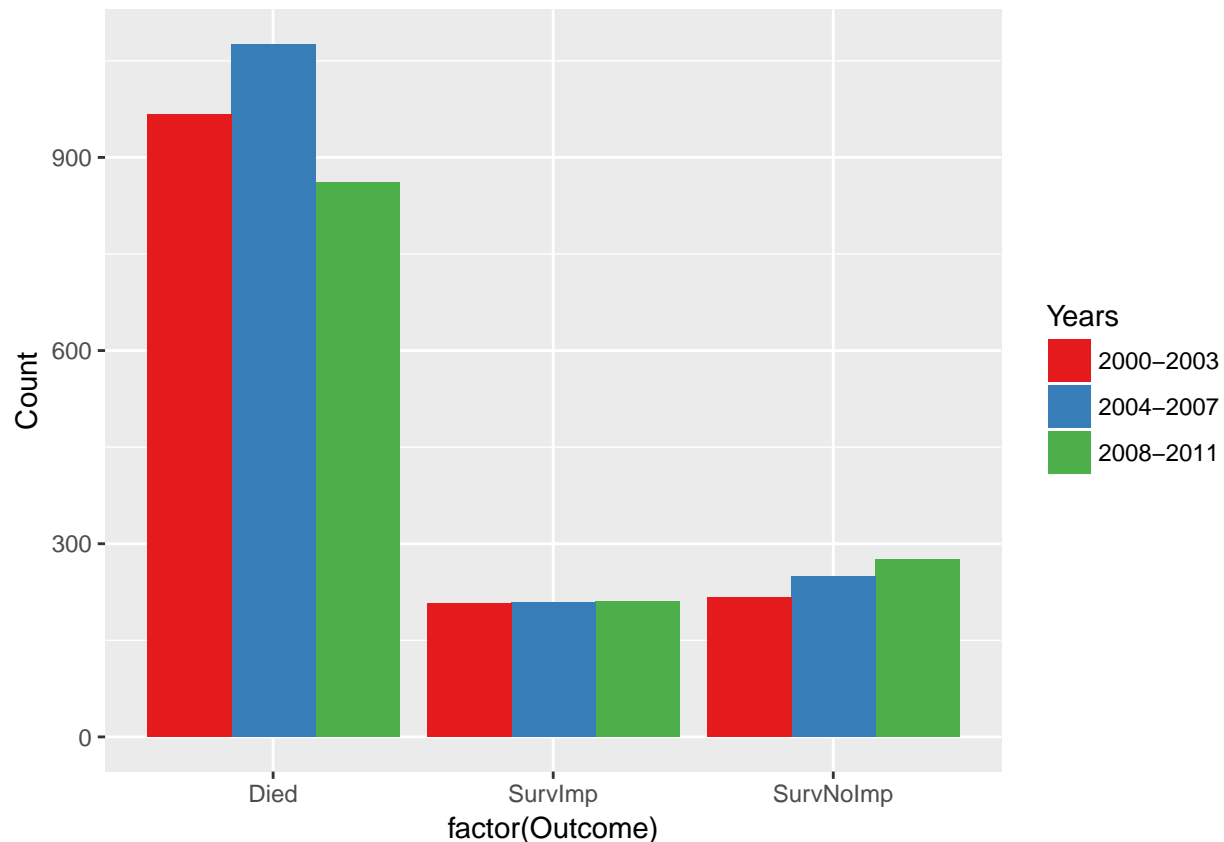
5. Agresti 2.28 (10 points)

In comparing new and standard treatments with success probabilities π_1 and π_2 , the number needed to treat (NNT) is the number of patients that would need to be treated with the new treatment instead of the standard in order for one patient to benefit. Explain why a natural estimate of this is $\frac{1}{\pi_1 - \pi_2}$.

Let n denote the number needed to treat and $\hat{\pi}_1$ and $\hat{\pi}_2$ are estimates of π_1 and π_2 respectively. Then if standard treatments are used on n patients, we may expect $n\hat{\pi}_2$ patients recover. Similarly, we expect $n\hat{\pi}_1$ patients recover using the new treatment. Therefore, if supposing π_1 is larger than π_2 , in order to get one more recovery patient, we need to have $n\hat{\pi}_1 - n\hat{\pi}_2 = 1$ which is $n = \frac{1}{\hat{\pi}_1 - \hat{\pi}_2}$.

6. Periviable infants (12 points total)

Recall the data on outcomes of periviable infants. Of 1348 infants in the study born at these gestational ages in 2008-2011, 861 died, 211 survived with neurodevelopmental impairment, and 276 survived without neurodevelopmental impairment. In 2000-2003, of 1391 infants born at these gestational ages, 217 survived with no impairment, 207 survived with impairment, and 967 died. In 2004-2007, of 1535 total infants, 250 survived without impairment, 209 survived with impairment, and 1076 died.



6.a. Nominal variables (6 points)

Provide a formal test of the hypothesis that time period is not associated with outcome. Interpret results of your test in clear language that could be used in a scientific publication. If there is a change over time, clearly describe the nature of that change.

We first create a 2-way contingency table for data.

```
neonate.table <- xtabs(Count ~ Years + Outcome, data = Neonate)
print(neonate.table)
```

```
##           Outcome
## Years      Died SurvImp SurvNoImp
## 2000-2003   967     207     217
## 2004-2007 1076     209     250
## 2008-2011  861     211     276
```

We apply Pearson and likelihood ratio χ^2 tests to examine whether time period and outcome are independent.

```
chisq.test(neonate.table)
```

```
##  
## Pearson's Chi-squared test  
##  
## data:  neonate.table  
## X-squared = 17.954, df = 4, p-value = 0.00126  
expected.table <- apply(neonate.table, 1, sum) %*% t(apply(neonate.table, 2, sum)) / sum(neonate.table)  
G.sq <- 2 * sum(neonate.table * log(neonate.table / expected.table))  
1 - pchisq(G.sq, 4)  
  
## [1] 0.001394679
```

Therefore, we reject the null hypothesis under level $\alpha = 0.05$. The data provides evidence that time period is associated with outcome. However, we cannot tell how outcome is associated with time period from the test itself. According to the histogram, the number of surviving with impairment and surviving without impairment is increasing with respect to time periods. However, the counts of death do not have a clear trend. We extract a 2×2 table from the original table

Outcome

Died

SurvNoImp

Year

00-03

967

217

08-11

861

276

The odds ratio is given by

$$\hat{\theta}_{13} = \frac{n_{11}n_{33}}{n_{13}n_{31}} = \frac{967 \times 276}{217 \times 861} = 1.43$$
$$\hat{\sigma}(\log \hat{\theta}_{13}) = \sqrt{\frac{1}{n_{11}} + \frac{1}{n_{13}} + \frac{1}{n_{31}} + \frac{1}{n_{33}}} = 0.10$$

A 95% confidence interval of $\hat{\theta}_{13}$ is (1.17, 1.74). Both the point estimate and confidence interval of the odds ratio provide evidence that the odds ratio is not equal to 1 and hence the observed data conflicts with independence.

6.b. Ordinal variables (6 points)

Now assume that the outcomes are ordered so that death is the worst outcome and survival without impairment is the best outcome. Conduct a linear trend test using scores (1,2,3) for time (1=2000-2003) and outcome (3=dead, 2=impaired, 1=survived and not impaired) and interpret the results. Do the results change if we use the outcome weights 100=dead, 10=impaired, 1=survived not impaired? If so, explain how and why the results differ. Comment on the appropriateness of the scores for both variables.

```

neonate <- Neonate
levels(neonate$Years) <- c(1,2,3)
levels(neonate$Outcome) <- c(3,2,1)
res <- neonate[,1:2][rep(1:nrow(neonate), neonate$Count),]
M.sq <- (cor(as.numeric(as.character(res$Years)), as.numeric(as.character(res$Outcome)))^2) * (nrow(neonate)-1-pchisq(M.sq,1))

```

```
## [1] 0.0003991334
```

Therefore, we reject the null hypothesis under level $\alpha = 0.05$ and conclude that the linear component of the association is evident in data.

```

levels(neonate$Outcome) <- c(100,10,1)
res <- neonate[,1:2][rep(1:nrow(neonate), neonate$Count),]
M.sq <- (cor(as.numeric(as.character(res$Years)), as.numeric(as.character(res$Outcome)))^2) * (nrow(neonate)-1-pchisq(M.sq,1))

```

```
## [1] 0.001217493
```

After adopting the new scores for outcome, the p -value becomes larger but still very small and we reject the null under level $\alpha = 0.05$. According to the histogram, the data presents an obvious nonlinear trend in “died” category. We increase the score of “died” category leading to a smaller correlation.

The choice of scores depends on the context. In some situations, these three outcomes or time periods may be considered to have equal distances and then our first choice of scores appears to be more appropriate. However, in some medical experiments, death may be considered much more serious than other two outcomes. Hence assigning a high score to death is more sensible. Therefore, the choice of scores can only be justified under specific context. All conclusions based on the analysis are valid under the specific choice of scores and may change if adopting other scores. When the trend is significant, it may be relatively robust to the choice of scores.

7. Oral contraceptives and thromboembolism (8 points)

We consider a retrospective matched pair case-control study of thromboembolism (formation of a blood clot that breaks loose and is carried by the bloodstream to plug another vessel) and oral contraceptive use. The cases were 175 women of reproductive age, discharged alive from 43 hospitals in 5 cities after experiencing idiopathic thrombophlebitis, pulmonary embolism, or cerebral thrombosis or embolism. Controls were females matched with their cases for hospital, residence, time of hospitalization, race, age, marital status, parity, and insurance pay status. The question of interest is whether cases are more likely than controls to use oral contraceptives.

Of the 175 matched pairs, in 10 pairs both cases and controls used oral contraceptives (OC's), in 95 pairs neither used OC's, in 13 pairs only the control used OC's, and in 57 pairs only the case used OC's.

Using the appropriate test, evaluate the null hypothesis that the proportion of cases exposed to OC's is the same as the proportion of controls who are exposed, and provide the p -value for this test. Calculate the OR and exact 95% CI comparing the odds of thromboembolism in OC users to those in non-OC users. Interpret your results.

case

OC

non_OC

control

OC

10

13

non_OC

57

95

Based on the table, let π_{1+} and π_{+1} denote the proportions of controls and cases exposed to OC's respectively. Then the null hypothesis is $H_0 : \pi_{1+} = \pi_{+1}$ and McNemar test is used.

```
mcnemar.test(matrix(c(10,57,13,95),ncol=2), correct = F)
```

```
##
## McNemar's Chi-squared test
##
## data: matrix(c(10, 57, 13, 95), ncol = 2)
## McNemar's chi-squared = 27.657, df = 1, p-value = 1.448e-07
```

We may calculate OR and its exact 95% CI by inverting the exact McNemar test.

```
library(exact2x2)
mcnemar.exact(matrix(c(10,13,57,95),ncol=2))

##
## Exact McNemar test (with central confidence intervals)
##
## data: matrix(c(10, 13, 57, 95), ncol = 2)
## b = 57, c = 13, p-value = 1.029e-07
## alternative hypothesis: true odds ratio is not equal to 1
## 95 percent confidence interval:
##  2.371377 8.731311
## sample estimates:
## odds ratio
##  4.384615
```

p -value is very small and hence we reject the null hypothesis under level $\alpha = 0.05$. The odds ratio is 4.38 here implying that the odds of thromboembolism in OC users is 4.38 times that in non-OC users. An exact 95% CI for the odds ratio is (2.3714, 8.7313).

8. Passive smoking and cancer risk (15 points)

A landmark study examined the association between passive smoking and cancer in a group of 509 individuals with cancer and 489 cancer-free (unmatched) controls. Researchers defined passive smoking as exposure to the cigarette smoke of a partner who smoked at least one cigarette per day for at least 6 months. One potential confounding variable is smoking by the participants (i.e., personal smoking) because personal smoking is related to both cancer risk and partner smoking. Therefore, it is important to consider personal smoking when examining the relationship between passive smoking and cancer risk.

Among cases, 120 were exposed to passive smoke but did not smoke themselves; 161 were smokers also exposed to passive smoke; 111 had no exposure from passive or active smoking; and 117 were active smokers who were not exposed to passive smoke. Among controls, 80 were exposed to passive smoke but did not smoke themselves; 130 were smokers also exposed to passive smoke; 155 had no exposure from passive or active smoking; and 124 were active smokers who were not exposed to passive smoke.

*(a) Calculate the conditional odds ratios comparing the odds of being a case by passive smoking exposure status, conditional on active smoking status (active smoker or nonsmoker).

We summarize the data into the following $2 \times 2 \times 2$ table.

active smoker

not active

Passive

NotPass

Passive

NotPass

Case

161

117

120

111

Control

130

124

80

155

```
ctable <- array(c(161, 130, 117, 124, 120, 80, 111, 155), dim = c(2, 2, 2))
apply(ctable, 3, function(x) {x[1, 1] * x[2, 2] / x[1, 2] / x[2, 1]})
```

```
## [1] 1.312558 2.094595
```

(b) Evaluate whether case status and passive smoking are conditionally independent given active smoking status.

Cochran-Mantel-Haenszel test is used for testing conditional independence.

```
mantelhaen.test(ctable, correct = F)
```

```
##
## Mantel-Haenszel chi-squared test without continuity correction
##
## data: ctable
## Mantel-Haenszel X-squared = 14.422, df = 1, p-value = 0.0001461
## alternative hypothesis: true common odds ratio is not equal to 1
## 95 percent confidence interval:
## 1.263955 2.090024
## sample estimates:
## common odds ratio
## 1.625329
```

Therefore we reject the null hypothesis under level $\alpha = 0.05$ indicating the data provides strong evidence that case status and passive smoking are not conditionally independent given active smoking status.

(c) Is the association between passive smoking and case status homogeneous across categories of active smoking status? Provide statistical evidence to support your answers.

Breslow-Day test is used for testing homogeneous association.

```
library(DescTools)
BreslowDayTest(ctable, correct = T)

##
## Breslow-Day Test on Homogeneity of Odds Ratios (with Tarone
## correction)
##
## data: ctable
## X-squared = 3.2736, df = 1, p-value = 0.0704
```

Since p -value is relatively large, we cannot reject the null under level $\alpha = 0.05$. There is no significant evidence that the association between passive smoking and case status is not homogeneous across categories of active smoking status.

9. Exercise without pain! (15 points)

An athletics company produces a new model of running shoe that includes a harder material in the insert to correct for overpronation. The company is worried that the material will cause heel tenderness due to some loss of cushioning with the strike of each step. The company enrolled 87 runners and asked them about occasional heel tenderness at study entry and then again after using the new shoe for one month. The company would like to know whether the proportion of runners with occasional heel tenderness is the same before the study and after using the study shoe for one month.

Analyze the data and produce a professional report (as part of this document) providing statistical evidence that addresses the question of interest to the company. Data are in the file dontbeaheel.xlsx.

The company collected the data based on the feedbacks from runners. We first summarize the data using the following 2×2 table. The company is interested in whether the proportion of runners with occasional heel tenderness is the same before and after the study. To respond to this question, we appeal to formal hypothesis testings. Since the records for each runner are dependent, the data is of the form of matched pairs. Therefore, we adopt McNemar test for this problem. The null hypothesis is the proportions of runners with occasional heel tenderness before study and after study are same. The alternative hypothesis is that the two proportions are different.

```
library(readxl)
shoe <- as.data.frame(read_xlsx('dontbeaheel.xlsx'))
shoe.reshape <- reshape(shoe, direction = 'wide', idvar = "ID", timevar = "Occasion")
shoe.table <- xtabs(~ Pain.Before + Pain.After, shoe.reshape)
shoe.table

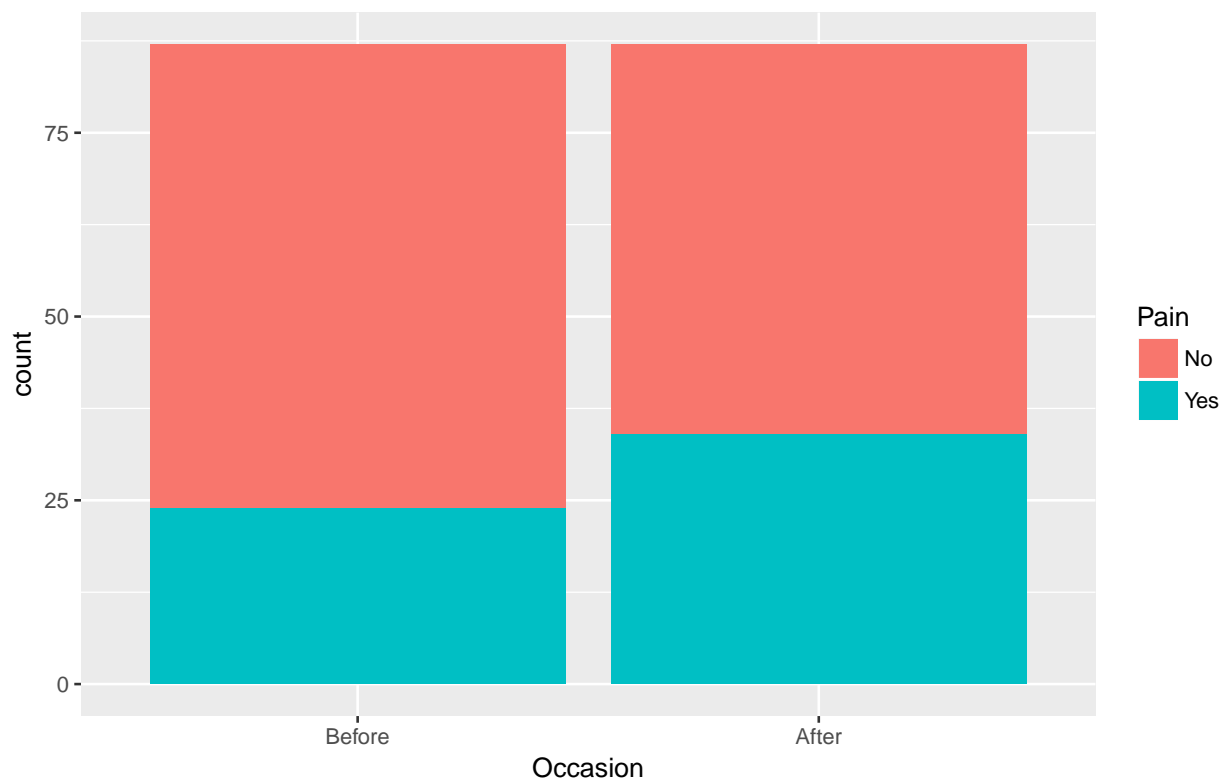
##              Pain.After
## Pain.Before No Yes
##           No  48  15
##           Yes   5  19

mcnemar.exact(shoe.table)
```



```
##
## Exact McNemar test (with central confidence intervals)
##
## data: shoe.table
## b = 15, c = 5, p-value = 0.04139
## alternative hypothesis: true odds ratio is not equal to 1
## 95 percent confidence interval:
##  1.03647 10.55115
## sample estimates:
## odds ratio
##      3
```

The p -value is 0.04139 and is smaller than the significance level $\alpha = 0.05$. We reject the null hypothesis and conclude that the data exhibits evidence that the proportion of occasional tenderness before and after the study are different. A point estimation of the odds ratio is 3 indicating the odds of the runners experiencing heel tenderness after the study is 3 times of that before the study. The following graph shows that more runners experience heel tenderness after the study than before the study.



According to the p -value of McNemar test and the estimate of the odds ratio, the data suggests that the new shoe may actually cause heel tenderness to more runners.