

Problem Set 4

Felipe Riveroll Aguirre*

November 13, 2012

Titanic Survival. The following table describes the survival status of passengers on the Titanic, stratified by Passenger Class (First, Second, or Third), Sex/Age (Child, Women, or Man), and Survival Status. The Frequency column indicates the number of passengers in each stratum. (For example there were 4 1st class women passengers who did not survive and 140 1st class women passengers who did survive). These data were obtained from the website anesi.com and refers to British Parliamentary Papers, Shipping. Casualties (Loss of the Steamship “Titanic”), 1912. cmd 6352 Report of a Formal Investigation into the circumstances attending the foundering on the 15 th April 1912 of the British Steamship “Titanic” of Liverpool after striking ice in or near Latitude 41 46 N., Longitude 50 14 W., North Atlantic Ocean, whereby loss of life ensued (London; His Majesty’s Stationary Office, 1912) page 42.

Passenger Class	Age/Sex	Survival Status	Frequency
First	Child	Survived	6
First	Child	Did not survive	0
First	Women	Survived	140
First	Women	Did not survive	4
First	Man	Survived	57
First	Man	Did not survive	118
Second	Child	Survived	24
Second	Child	Did not survive	0
Second	Women	Survived	80
Second	Women	Did not survive	13
Second	Man	Survived	14
Second	Man	Did not survive	154
Third	Child	Survived	27
Third	Child	Did not survive	52
Third	Women	Survived	76
Third	Women	Did not survive	89
Third	Man	Survived	75
Third	Man	Did not survive	387

*friveroll@gmail.com

1. Use these data to calculate the Risk Ratio for surviving comparing “women or children” as the exposed group and “all other passengers” as the non-exposed group.

```
> titanic <- read.csv("https://dl.dropbox.com/u/4828275/titanic.csv")

> # All Women or Children
> Child_Women <- subset(titanic, Age.Sex == "Women" | Age.Sex == "Child")
> Child_Women_R1 <- sum(subset(Child_Women,
+                           Survival.Status == "Survived"),4))/
+                           sum(Child_Women[,4])
> # All man
> Man <- subset(titanic, Age.Sex == "Man")
> Man_R0 <- sum(subset(Man,
+                     Survival.Status == "Survived"),4))/
+                     sum(Man[,4])
> # Risk Ratio R1 / R0
> Child_Women_R1/Man_R0

[1] 3.808876
```

2. Repeat this calculation for each passenger class.

(a) First Class Risk Ratio

```
> Child_Women_1st <- subset(Child_Women,
+                           Passenger.Class == "First")
> Child_Women_1st_R1 <- sum(subset(Child_Women_1st,
+                                 Survival.Status == "Survived"),4))/
+                                 sum(Child_Women_1st[,4])
> Man_1st <- subset(Man,
+                  Passenger.Class == "First")
> Man_1st_R0 <- sum(subset(Man_1st,
+                          Survival.Status == "Survived"),4))/
+                          sum(Man_1st[,4])
> Child_Women_1st_R1/Man_1st_R0

[1] 2.988304
```

(b) Second Class Risk Ratio

```
> Child_Women_2nd <- subset(Child_Women,
+                           Passenger.Class == "Second")
> Child_Women_2nd_R1 <- sum(subset(Child_Women_2nd,
+                                 Survival.Status == "Survived"),4))/
+                                 sum(Child_Women_2nd[,4])
> Man_2nd <- subset(Man,
+                  Passenger.Class == "Second")
> Man_2nd_R0 <- sum(subset(Man_2nd,
+                          Survival.Status == "Survived"),4))/
+                          sum(Man_2nd[,4])
> Child_Women_2nd_R1/Man_2nd_R0

[1] 10.66667
```

(c) Third Class Risk Ratio

```
> Child_Women_3th <- subset(Child_Women,
+                             Passenger.Class == "Third")
> Child_Women_3th_R1 <- sum(subset(Child_Women_3th,
+                                 Survival.Status == "Survived"),4)/
+                             sum(Child_Women_3th[,4])
> Man_3th <- subset(Man,
+                   Passenger.Class == "Third")
> Man_3th_R0 <- sum(subset(Man_3th,
+                           Survival.Status == "Survived"),4)/
+                   sum(Man_3th[,4])
> Child_Women_3th_R1/Man_3th_R0
[1] 2.600328
```

Incidence Rate Ratio Blood Pressure and CHD. The following table uses data from the NHLBI teaching data set and displays the blood pressure distribution for 4,434 participants in the Framingham Heart Study attending an examination in 1956. For each blood pressure category, the tables displays the number of subjects with existing Coronary Heart Disease (CHD) at that exam (Prevalent Cases of CHD) and also the number of new cases of CHD and the total amount of person-years of follow-up that was observed during a 24 year follow-up period **for those subjects who did not have CHD at the 1956 exam.** Follow-up for each subject began in 1956 and ended with the development CHD (fatal or non-fatal), death from another cause, loss to follow-up, or the end of the follow-up period (whichever came first).

Blood Pressure Category		Number of Subjects at 1956 Exam	Prevalent Cases of CHD	Number Developing CHD During Follow up	Total Years of Followup
I	$SBP < 140$ and $DBP < 90$	2815	88	547	55384.42
II	$140 \leq SBP < 160$ or $90 \leq DBP < 95$	781	39	214	13191.79
III	$SBP \geq 160$ or $DBP \geq 95$	838	67	285	12348.94

1. What is the Incidence Rate Ratio for developing CHD for participants in Blood Pressure Groups II or III combined (exposed group) compared to participants in the Blood Pressure Group I (non-exposed group)

```
> BP <- read.csv("https://dl.dropbox.com/u/4828275/BP.csv", header = T)
> (sum(BP[2:3,5])/sum(BP[2:3,6]))/(BP[1,5]/BP[1,6])
[1] 1.978188
```

2. What is the Incidence Rate Ratio for developing CHD for participants in Blood Pressure Group III (exposed group) compared to participants in the Blood Pressure Group I (non-exposed group)?

```
> (BP[3,5]/BP[3,6])/(BP[1,5]/BP[1,6])
[1] 2.336767
```

3. What is the Incidence Rate Ratio for developing CHD for participants in Blood Pressure Group II (exposed group) compared to participants in the Blood Pressure Group I (non-exposed group)?

```
> (BP[2,5]/BP[2,6])/(BP[1,5]/BP[1,6])
[1] 1.642519
```

Risk Ratios, Odds Ratios, Rate Ratios for BMI, Death, and CHD. The following table displays categories of body mass index for the participants at the 1956 exam. (Note: 19 of the 4434 participants are excluded from this table because of missing data on bmi1.)

BMI Category	Range of BMI	Frequency
Underweight	$BMI < 18.5$	57
Normal weight	$18.5 \leq BMI < 25$	1936
Overweight	$25 \leq BMI < 30$	1845
Obese	$30 \leq BMI$	577

Use R to perform the following calculations. Hint: All of the following questions ask you to compare obese subjects to normal weight subjects. Create a new binary variable using bmi1 which equals 1 if the person is obese and 0 if the person is normal weight. Anyone who is underweight or overweight should be missing a value for the new binary variable you create.

```
> library("foreign") #needed for read.dta function
> ### Import Data from STATA ###
> data <- read.dta("https://dl.dropbox.com/u/4828275/fhs.dta",
+               convert.factors = TRUE,
+               missing.type = TRUE)
> attach(data)
>
> bmi1_obese <- NA
> bmi1_obese[bmi1>=30] <- 1
> bmi1_obese[bmi1>=18.5 & bmi1 < 25] <- 0
> bmi1_obese <- na.exclude(bmi1_obese)
```

1. Calculate the 24-year Risk Ratio for death comparing obese subjects (exposed group, n=577) to normal weight subjects (non-exposed group, n=1936).
2. Calculate the 24-year Odds Ratio for death comparing obese subjects (exposed group, n=577) to normal weight subjects (non-exposed group, n=1936).
3. Calculate the 24-year Rate Ratio for death comparing obese subjects (exposed group, n=577) to normal weight subjects (non-exposed group, n=1936).
4. Calculate the 24-year Rate Ratio for developing coronary heart disease comparing obese subjects (exposed group) to normal weight (non-exposed group), excluding subjects with prevalent CHD at the 1956 exam.
Hint: Use the if/in tab options to restrict the sample to those without prevalent CHD at the 1956 exam (prevchd1 = 0).
5. Calculate the 24-year Rate Ratio for developing coronary heart disease comparing (obese or overweight) subjects (exposed group) to normal weight (non-exposed group), excluding subjects with prevalent CHD at the 1956 exam.
6. Calculate the 24-year Rate Ratio for developing coronary heart disease comparing underweight subjects (exposed group) to normal weight (non-exposed group), excluding subjects with prevalent CHD at the 1956 exam.