

Lab #3

The focus of this lab will be on using R to provide you with summary statistics of different variables with a data set. We will look at both categorical and continuous variables. We will also go through calculating a weighted average and how to stratify contingency tables.

1 Tables in R

1.1 First import in the dataset `titanic.txt`

1.2 Use the command `attach()` to attach the dataset to R

```
attach(titanic)
```

Now you do not need to use the `$` to tell R where to look for the variables. When you are done working with the dataset, use `detach(titanic)`

Read in the table and notice the output

```
Table1 <- table(Class)
Table1
```

To view proportions/percentages of tables

```
Props1 <- prop.table(Class)
Props1
Perc1 <- 100*Props1
Perc1
```

Two and three way tables can also be created

```
Table2 <- table(Sex,Survived)
Table2
Table3 <- table(Sex,Survived,Age)
Table3
```

You can also use `prop.table` for these two and three way tables to see percentages across rows and columns.

```
prop.table(Table2) ## Overall using Table2 created previously
prop.table(Table2,1) ## Across row in Table2
prop.table(Table2,2) ## Across col in Table2
```

For three way tables things get a little trickier. The code below will calculate proportions for each combination of Sex and Age using our Table3 from before. (Note 1 and 3 are the positions of the variables Sex and Age in the construction of Table3).

```
prop.table(Table3,c(1,3))
```

2 Graphs in R

R is a great tool to view charts and graphs. We will introduce the graphical displays from the Lattice package, which was written as an alternative to the default displays that R provides. The code to load the Lattice package is:

```
require(lattice)
```

You should now see a message that the Lattice package was loaded. Now we can create bar charts to graph categorical data. Several examples are provided below:

```
barchart(Table1)
barchart(table(Survived))
## A simple barchart showing the counts of survivors and non-survivors,

barchart(table(Survived),horizontal=FALSE)
## Changing horizontal from its default to =FALSE will create a vertical barchart

barchart(table(Class,Survived))
## Using "Survived" as a grouping variable for "Class"
barchart(table(Class,Survived),auto.key=TRUE)
## Includes a key indentifying the groups by color

barchart(table(Class,Sex,Survived),auto.key=TRUE)
barchart(table(Class,Sex,Age,Survived),auto.key=TRUE)
barchart(table(Class,Sex,Age,Survived),auto.key=TRUE,scales="free")
#Allows the scale to change across graphs
```

3 Weighted Averages

Weighted averages have been a tough concept for students to grasp, so now we will spend some time going over how to calculate weighted averages and what they represent. Consider the next two tables

Class	Survived Count	N	Rate
1st	203	325	62.46
2nd	118	285	41.40
3rd	178	706	25.21
Crew	212	885	23.95

Class	Sex	Survived Count	N	Rate
1st	F	141	145	97.24
	M	62	180	34.44
2nd	F	93	106	87.74
	M	25	179	13.97
3rd	F	90	196	45.92
	M	88	510	17.25
Crew	F	20	23	86.96
	M	192	862	22.27

Based on the information above, what proportion of all passengers were female? How about male? These will now be our weights.

Weighted Averages

We will now calculate the weighted averages for the first, second, third, and crew classes, adjusting for sex. Try this on your own. To get you started, the adjusted survival rate for the first class is:

$$\begin{aligned} & \left(\frac{\# \text{ of Female Passengers}}{\text{Total \# of Passengers}} \right) \left(\frac{\# \text{ of Females Survive 1st class}}{\# \text{ Female Passengers in 1st class}} \right) \\ & + \left(\frac{\# \text{ of Male Passengers}}{\text{Total \# of Passengers}} \right) \left(\frac{\# \text{ of Males Survive 1st class}}{\# \text{ Male Passengers in 1st class}} \right) \\ & = \left(\frac{470}{2201} \right) \left(\frac{141}{145} \right) + \left(\frac{1731}{2201} \right) \left(\frac{62}{180} \right) = 0.479 \end{aligned}$$

Find the adjusted survival rates for the rest of the categories and compare them to the raw rates.

This concludes Lab 3. DON'T FORGET TO DETACH YOUR DATASET.

```
detach(titanic)
```

■