

STAT 22000: Homework 3

Zixi Li

Problem 1 The North Carolina births data

(a)

Here I make a two-way contingency table between marital and lowbirthweight.

```
tally(~marital+lowbirthweight, data=nc, margins=FALSE)
```

```
##           lowbirthweight
## marital      low not low
##   married      49    564
##   not married    61    325
```

(b)

The proportion of babies who were born out of marriage is:

$$\frac{61 + 325}{49 + 564 + 61 + 325} \times 100\% = 38.64\%$$

The proportion of babies who were born with a low birth weight is:

$$\frac{49 + 61}{49 + 564 + 61 + 325} \times 100\% = 11.01\%$$

(c)

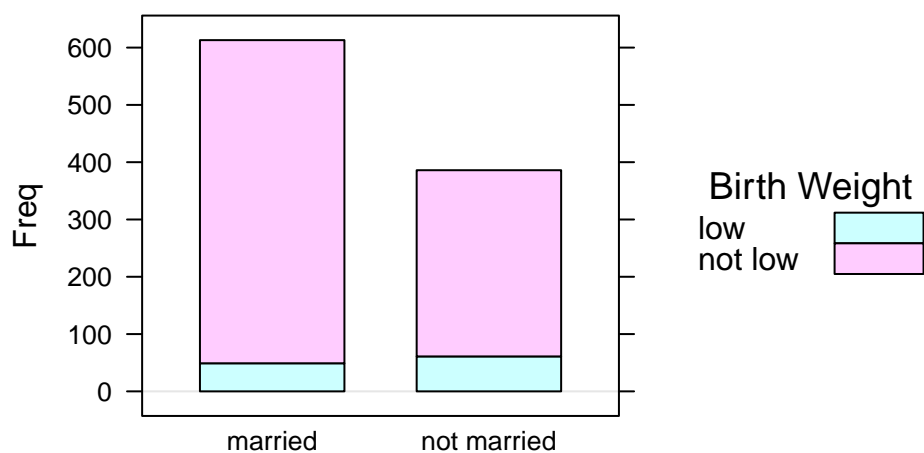
Among the babies born to unmarried moms, the proportion of having a low birth weight is:

$$\frac{61}{61 + 325} \times 100\% = 15.80\%$$

(d)

Here I make a segmented bar chart that displays the two-way table in part (a), where the bars represent the variable marital and the segments of the bars represent the levels of lowbirthweight.

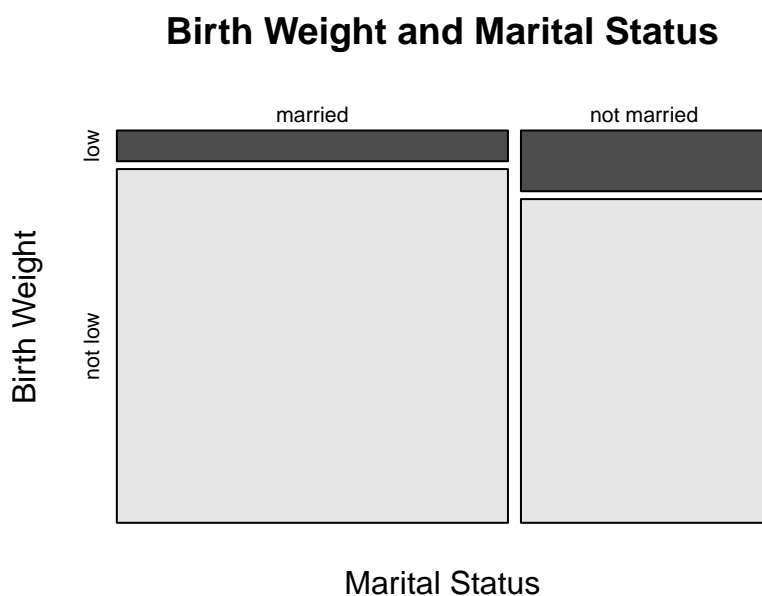
```
marilowbw = tally(~marital+lowbirthweight, data=nc, margins=FALSE)
barchart(marilowbw, horizontal=FALSE,
         auto.key=list(space="right", title="Birth Weight", cex.title=1.2))
```



(e)

Here I make a mosaic plot that compares the percentage of babies born with a low birth weight between those born to married women and those born to unmarried women.

```
mosaicplot(tally(~ marital+lowbirthweight, data=nc),
  main="Birth Weight and Marital Status",
  ylab= "Birth Weight", xlab="Marital Status", color=TRUE)
```



(f)

Based on the plot in the previous part, babies born outside of marriage were more likely to have a low birth weight.

(g)

The two variables marital and lowbirthweight are not independent.

Problem 2 The North Carolina births data continued.

(a)

Here I make a two-way contingency table between marital and habit.

```
tally(~marital+habit, data=nc, margins=FALSE)
```

```
##           habit
## marital      nonsmoker smoker
##   married          553    60
##   not married       320    66
```

Among married mothers, the percentage of smoking during pregnancy is:

$$\frac{60}{553 + 60} \times 100\% = 9.79\%$$

Among unmarried mothers, the percentage of smoking during pregnancy is:

$$\frac{66}{320 + 66} \times 100\% = 17.10\%$$

Unmarried mothers were more likely to smoke during pregnancy.

(b)

Here I make two tables for examining the relation of marital and lowbirthweight after adjusting for the mother's smoking habit.

```
tab = tally(~marital + lowbirthweight + habit, data=nc)
ftable(tab, col.vars = c("habit","marital"))
```

```
##           habit  nonsmoker          smoker
##           marital  married not married married not married
## lowbirthweight
## low                42         50         7         11
## not low            511        270        53        55
```

```
prop.table(ftable(tab, col.vars = c("habit","marital")), 2)
```

```
##           habit  nonsmoker          smoker
##           marital  married not married married not married
## lowbirthweight
## low                0.0759494  0.1562500 0.1166667  0.1666667
## not low            0.9240506  0.8437500 0.8833333  0.8333333
```

- i. Among the nonsmoking married mothers, the percentage of having babies with low birth weights is 0.0759494. Among the nonsmoking unmarried mothers, the percentage of having babies with low birth weights is 0.1562500. Thus considering babies with nonsmoking mothers only, those born to unmarried mothers were more likely to have low birth weights.

- ii. Among the smoking married mothers, the percentage of having babies with low birth weights is 0.1166667. Among the nonsmoking unmarried mothers, the percentage of having babies with low birth weights is 0.1666667. Thus considering babies with smoking mothers only, those born to unmarried mothers were more likely to have low birth weights.
- iii. Out-of-marriage birth is still associated with low birth weights after adjusting for mothers' smoking habit.

(c)

From the above analysis, it's highly possible that marriage reduces the chance of births having low birth weights, since we have observed the association even after adjusting for mothers' smoking habit. However, we cannot draw a safe conclusion because there might be other factors that confound the results besides smoking habit.

Problem 3 Smokers' relapse

(a)

If we assume there is no other possible factors, we can claim that bupropion is effective as a smoking cessation aid, because bupropion decreases relapse rate in both environments (living with a smoker or not).

(b)

Yes. The environment (living with a smoker or not) is the block.

(c)

We cannot. First, the amount of sample in the two groups are different (125 and 89) so it doesn't fulfill the requirement of randomization. Also, in order to claim that living with smokers makes it harder to quit smoking, we should take bupropion groups into account as well.