Chapter 3 - Probabilites

Joe Rovalino

Chapter 3 - Probability

1) Dice rolls. (3.6, p. 92) If you roll a pair of fair dice, what is the probability of

(a) getting a sum of 1?

JR Answer: 1/6 or 0.1666667

```
x<- 1/6
x
```

```
## [1] 0.1666667
```

(b) getting a sum of 5? JR Answer: Same as above. 1/6 or 0.1666667
(c) getting a sum of 12?

JR Answer: Same as above 1/6 or 0.1666667

2) Poverty and language. (3.8, p. 93) The American Community Survey is an ongoing survey that provides data every year to give communities the current information they need to plan investments and services. The 2010 American Community Survey estimates that 14.6% of Americans live below the poverty line, 20.7% speak a language other than English (foreign language) at home, and 4.2% fall into both categories.

(a) Are living below the poverty line and speaking a foreign language at home disjoint? JR Answer: They are not disjoint or mutually exclusive as 4.2% are both poverty true and foreign language at home as true.

(b) Draw a Venn diagram summarizing the variables and their associated probabilities. JR Answer: Please see Venn diagram in the submission of HMWK "PovertyandLangVenn.jpg"

(c) What percent of Americans live below the poverty line and only speak English at home? JR Answer: Below poverty line and only speak english = P(Poverty) - P(OtherLan) = 0.146-0.042 = 0.104

```
y<- 0.146 - 0.042
y
```

```
## [1] 0.104
```

(d) What percent of Americans live below the poverty line or speak a foreign language at home? JR Answer: Use the Addition rule. Probablity(Poverty)+Probability(OtherLan) - P(both poverty and Lang) = 0.146 + .207 - 0.042 = 0.311

```
z <- 0.146 + .207 - 0.042
z
```

```
## [1] 0.311
```

(e) What percent of Americans live above the poverty line and only speak English at home? JR Answer:

```
Three_e <- 1 - z
Three_e
```

```
## [1] 0.689
```

(f)Is the event that someone lives below the poverty line independent of the event that the person speaks a foreign language at home? JR Answer:

```
PbelowP <- 0.146
POthLang <- 0.207
PbelowPnOthLang <- 0.042
Three_f <- PbelowP * POthLang
Three_f
```

```
## [1] 0.030222
```

```
typeof(Three_f)
```

```
## [1] "double"
```

```
is.logical(Three_f != PbelowPnOthLang)
```

```
## [1] TRUE
```

Since the mutliplication rule of PbelowP * POthLang = .311 and this doesn't equal PbelowPnOthLang = 0.042, then they are not independent.

3) Assortative mating. (3.18, p. 111) Assortative mating is a nonrandom mating pattern where individuals with similar genotypes and/or phenotypes mate with one another more frequently than what would be expected under a random mating pattern. Researchers studying this topic collected data on eye colors of 204 Scandinavian men and their female partners. The table below summarizes the results. For simplicity, we only include heterosexual relationships in this exercise.

```
F_Bl_Blue <- 78
F_Br_Blue <- 19
F_Gr_Blue <- 11
F_Bl_Green <- 13
F_Bl_Brown <- 23
F_Br_Brown <- 23
F_G_Brown <- 9
F_Bl_Green <- 13
F_Br_Green <- 12
F_G_Green <- 16
Total_M_Blue <- 114
Total_M_Brown <- 54
Total_M_Green <-36
Total_F_Blue <- 108
Total_F_Brown <-55
Total_F_Green <- 41
Grand_Total <- 204
```

(a) What is the probability that a randomly chosen male respondent or his partner has blue eyes?

```
a_answer <- (Total_F_Blue/Grand_Total) + (Total_M_Blue/Grand_Total) - (F_Bl_Blue/Grand_Total)
a_answer
```

```
## [1] 0.7058824
```

The answer is 0.706

(b) What is the probability that a randomly chosen male respondent with blue eyes has a partner with blue eyes?

```
b_answer <-F_Bl_Blue/Total_M_Blue
b_answer
```

```
## [1] 0.6842105
```

The answer is 0.684 of a randomly chosen male respondent with blue eyes has a partner with blue eyes

(c) What is the probability that a randomly chosen male respondent with brown eyes has a partner with blue eyes?

```
c_answer <- F_Br_Blue/Total_M_Brown
c_answer
```

```
## [1] 0.3518519
```

The answer is 0.352 of a randomly chosen male respondent with brown eyes has a partner with blue eyes

(c2) What about the probability of a randomly chosen male respondent with green eyes having a partner with blue eyes?

```
c2_answer <- F_Gr_Blue/Total_M_Green
c2_answer
```

```
## [1] 0.3055556
```

c2_ answer is 0.306 of a randomly chosen male respondent with green eyes having a partner with blue eyes

(d) Does it appear that the eye colors of male respondents and their partners are independent? Explain your reasoning.

```
P_FB <- Total_F_Blue/Grand_Total
P_FB
```

```
## [1] 0.5294118
```

d_answer: Since we already calculated b answer to be 0.684 of a randomly chosen male respondent with blue eyes has a femail partner with blue eyes. Whish is 0.684. If we take the probability of females = Blue as 108/204 should be equal to probability of when the female is blue is 0.529. They would have been equal if they were independent

4) Books on a bookshelf. (3.26, p. 114) The table below shows the distribution of books on a bookcase based on whether they are nonfiction or fiction and hardcover or paperback.

```
FicHard <- 13
NonFicHard <- 15
FicPaper <- 59
NocFicPaper <- 8
Tot_H <- 28
Tot_P <- 67
Tot_F <- 72
Tot_NonF <-23
BookGTotal <- 95
```

(a) Find the probability of drawing a hardcover book first then a paperback fiction book second when drawing without replacement.

```
ans_4_a <- (Tot_H/BookGTotal) * (FicPaper/(BookGTotal -1))
ans_4_a
```

```
## [1] 0.1849944
```

Answer 4a is 0.185 probability of drawing a hardcover book first then a paperback fiction book second when drawing without replacement

(b) Determine the probability of drawing a fiction book first and then a hardcover book second, when drawing without replacement.

```
ans_4_b <- (Tot_F/BookGTotal) * (Tot_H/(BookGTotal-1))
ans_4_b
```

```
## [1] 0.2257559
```

```
ans_4_b2 <- (Tot_F/BookGTotal) * ((Tot_H -1)/(BookGTotal-1))
ans_4_b2
```

```
## [1] 0.2176932
```

The probability will differ if the first fiction book was a hardcover or not. As the Total hardcover may have been reduced by 1. So the answer is 0.218 if the first book was a hardcover. If the first selection was a paperback then the probabilty shitfs to 0.226.

(c) Calculate the probability of the scenario in part (b), except this time complete the calculations under the scenario where the first book is placed back on the bookcase before randomly drawing the second book.

```
ans_4_c <- (Tot_F/BookGTotal) * (Tot_H/(BookGTotal))
ans_4_c
```

```
## [1] 0.2233795
```

Answer c is 0.223 as the book is put back into the shelf.

(d) The final answers to parts (b) and (c) are very similar. Explain why this is the case.

Since the fiction book that may or may not have been hardcoover but ultimately is returned to the shelf - it returns the numbers for total hardcover to the original number as well as the total to the oringal number. This makes the grand total on the second drawing (denominator) continue to be 95 count.

5 Baggage fees. (3.34, p. 124) An airline charges the following baggage fees: $25 for the first bag and $35 for the second. Suppose 54% of passengers have no checked luggage, 34% have one piece of checked luggage and 12% have two pieces. We suppose a negligible portion of people check more than two bags. (a) Build a probability model, compute the average revenue per passenger, and compute the corresponding standard deviation.

```
no_bag <- 0 * .54
onebag <- 25* .34
twobag <- 60 * .12
avgrev <-(no_bag + onebag + twobag)
avgrev
```

```
## [1] 15.7
```

```
s1 <- (0-avgrev)^2 * .54
s1
```

```
## [1] 133.1046
```

```
s2 <- (25-avgrev)^2 * 0.34
s2
```

```
## [1] 29.4066
```

```
s3 <- (60-avgrev)^2 * 0.12
s3
```

```
## [1] 235.4988
```

```
v_bag <- s1+s2+s3
v_bag
```

```
## [1] 398.01
```

```
sd = sqrt(v_bag)
sd
```

```
## [1] 19.95019
```

Answer: Average Revenue is $15.70 and the standard deviation is 19.95

(b) About how much revenue should the airline expect for a flight of 120 passengers? With what standard deviation? Note any assumptions you make and if you think they are justified.

```
p_no <-120
rev <- p_no * avgrev
rev
```

```
## [1] 1884
```

```
rev_a <-rev - sd
rev_a
```

```
## [1] 1864.05
```

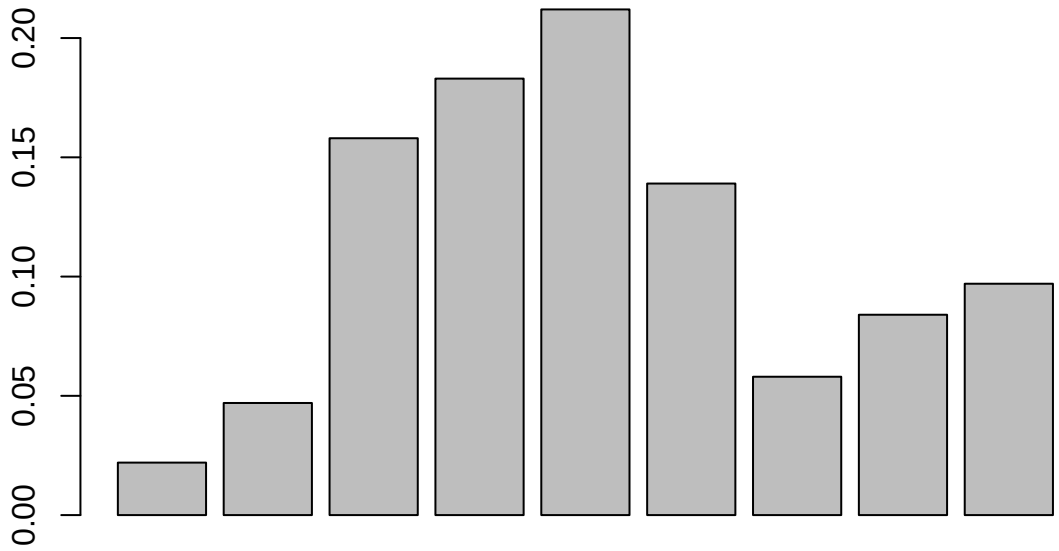
```
rev_b <- rev + sd
rev_b
```

```
## [1] 1903.95
```

The revenue would be $1884. Appling the sd -+ is 1864.05 or 1903.95

6 Income and gender. (3.38, p. 128) The relative frequency table below displays the distribution of annual total personal income (in 2009 inflation-adjusted dollars) for a representative sample of 96,420,486 Americans. These data come from the American Community Survey for 2005-2009. This sample is comprised of 59% males and 41% females. Income 1 $1 to $9,999 or loss 2 $10,000 to $14,999 3 $15,000 to $24,999 4 $25,000 to $34,999 5 $35,000 to $49,999 6 $50,000 to $64,999 7 $65,000 to $74,999 8 $75,000 to $99,999 9 $100,000 or more Total

```
buck1 <- 0.022
buck2 <- 0.047
buck3 <- 0.158
buck4 <- 0.183
buck5 <- 0.212
buck6 <- 0.139
buck7 <- 0.058
buck8 <- 0.084
buck9 <- 0.097
v <- c( 0.022, 0.047, 0.158, 0.183, 0.212, 0.139, 0.058, 0.084, 0.097)
barplot(v)
```




```
buck1to5 <- 0.022 + 0.047 + 0.158 +  0.183 +  0.212
buck1to5
```

```
## [1] 0.622
```

```
F_buck1to5 <- buck1to5 * .41
F_buck1to5
```

```
## [1] 0.25502
```

```
F2_buck1to5 <- buck1to5 * .718
F2_buck1to5
```

```
## [1] 0.446596
```

(a) Describe the distribution of total personal income.

Answer: The distribution is multimodal. The concentration in the 3-6 salary buckets. (b) What is the probability that a randomly chosen US resident makes less than $50,000 per year? Answer: Add the first five salary bucket % : 0.022 + 0.047 + 0.158 + 0.183 + 0.212: 0.622 or 62.2% chance of randomly choosing someone making less then $50K.

(c) What is the probability that a randomly chosen US resident makes less than $50,000 per year and is female? Note any assumptions you make.

Answer .6222* .41 or 0.25502 that she is femaile and makes less then 50K. Assuming that the male to female ratio is constant in each salary bucket.

(d) The same data source indicates that 71.8% of females make less than $50,000 per year. Use this value to determine whether or not the assumption you made in part (c) is valid.

AnswerL .6222 * .718 or 0.446596 that she is femail and makes less then 50K