Week 2 Assignment Math - Puneet Auluck

7/12/15

Short Answers:

- **2.16** 0.9750
- **2.18**
    (a) 0.2839
    (b) 0.2795
    (c) 0.3212
    (d) Yes, independent
- **2.20**
    (a) 0.7059
    (b) 0.6842
    (c) 0.3519, 0.3056
    (d) No, not independent
- **2.23** 0.4615

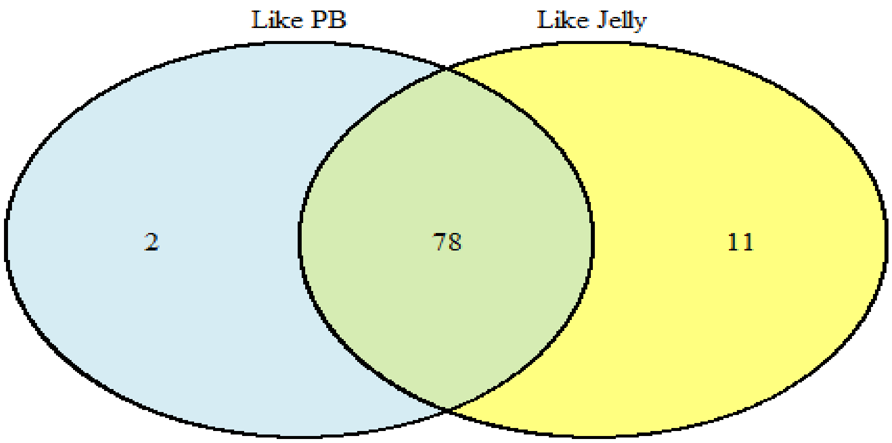
Detailed Answers:

2.16 PB&J.
Suppose 80% of people like peanut butter, 89% like jelly, and 78% like both. Given that a randomly sampled person likes peanut butter, what's the probability that he also likes jelly?

I have created a Venn diagram that displays the proportions of PB&J preferences.

P(like PB) = 80%, P(like Jelly) = 89% and P(like Jelly and like PB) = 78%

```
suppressMessages(library(VennDiagram))
draw.pairwise.venn(80, 89, 78, category = c("Like PB", "Like Jelly"), fill =
c("light blue", "yellow"), alpha = rep(0.5, 2), cat.pos = c(0,0), cat.dist =
rep(0.025, 2), scaled = FALSE, rotation.degree = 180)
```




```
## (polygon[GRID.polygon.1], polygon[GRID.polygon.2],
polygon[GRID.polygon.3], polygon[GRID.polygon.4], text[GRID.text.5],
text[GRID.text.6], text[GRID.text.7], text[GRID.text.8], text[GRID.text.9])
```

```
p_like_PB <- 0.80
p_like_Jelly <- 0.89
p_like_PB_and_Jelly <- 0.78
p_like_Jelly_given_PB <- p_like_PB_and_Jelly/p_like_PB
sprintf("The probability that the random person selected jelly is %1.4f.",
p_like_Jelly_given_PB)
```

```
## [1] "The probability that the random person selected jelly is 0.9750."
```

2.18 Weight and health coverage, Part II.
Exercise 2.14 introduced a contingency table summarizing the relationship between weight status, which is determined based on body mass index (BMI), and health coverage for a sample of 428,638Americans. In the table below, the counts have been replaced by relative frequencies (probability estimates).

		Weight Status			
		Neither overweight nor obese (BMI < 25)	Overweight ($25 \leq$ BMI < 30)	Obese (BMI ≥ 30)	Total
Health Coverage	Yes	0.3145	0.3306	0.2503	0.8954
	No	0.0352	0.0358	0.0336	0.1046
	Total	0.3497	0.3664	0.2839	1.0000

(a) What is the probability that a randomly chosen individual is obese?

```
P_Obese <- 0.2839
sprintf("The probability a randomly chosen individual is %1.4f.", P_Obese)
```

```
## [1] "The probability a randomly chosen individual is 0.2839."
```

(b) What is the probability that a randomly chosen individual is obese given that he has health coverage?

```
P_HC_Yes <- 0.8954
P_Obese_and_HC_Yes <- 0.2503
P_Obese_given_HC_Yes <- P_Obese_and_HC_Yes/P_HC_Yes
sprintf("The probability is %1.4f.", P_Obese_given_HC_Yes)

## [1] "The probability is 0.2795."
```

(c) What is the probability that a randomly chosen individual is obese given that he doesn't have health coverage?

```
P_HC_No <- 0.1046
P_Obese_and_HC_No <- 0.0336
P_Obese_given_HC_No <- P_Obese_and_HC_No/P_HC_No
sprintf("The probability is %1.4f.", P_Obese_given_HC_No)

## [1] "The probability is 0.3212."
```

(d) Do being overweight and having health coverage appear to be independent?

Independence check: P(Overweight and Health Coverage Yes) = 0.3306
P(OverWeight)*P(Health Coverage Yes) = 0.3664* 0.8954

```
P_Overweight_and_HC_Yes <- 0.3306
P_Overweight <- 0.3664
P_HC_Yes <- 0.8954
sprintf("Is P_Overweight_and_HC_Yes: %1.4f equal to
(P_Overweight)*(P_HC_Yes): %1.4f?", P_Overweight_and_HC_Yes,
P_Overweight*P_HC_Yes)

## [1] "Is P_Overweight_and_HC_Yes: 0.3306 equal to
(P_Overweight)*(P_HC_Yes): 0.3281?"
```

The probabilities calculated above are almost same so they appear to be independent.

2.20 Assortative mating. 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.

		Partner (female)			
		Blue	Brown	Green	Total
Self (male)	Blue	78	23	13	114
	Brown	19	23	12	54
	Green	11	9	16	36
	Total	108	55	41	204

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

```
P_MBL <- 114/204
P_FBL <- 108/204
P_MBL_and_FBL <- 78/204
P_MBL_or_FBL <- (P_MBL + P_FBL)-(P_MBL_and_FBL)
sprintf("The probability is %1.4f.", P_MBL_or_FBL)

## [1] "The probability is 0.7059."
```

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

```
P_FBL_given_MBL <- P_MBL_and_FBL/P_MBL
sprintf("The probability is %1.4f.", P_FBL_given_MBL)

## [1] "The probability is 0.6842."
```

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

```
P_MBR <- 54/204
P_MBR_and_FBL <- 19/204
P_FBL_given_MBR <- P_MBR_and_FBL/P_MBR
sprintf("The probability brown eyed male has a partner with blue eyes is
%1.4f.", P_FBL_given_MBR)

## [1] "The probability brown eyed male has a partner with blue eyes is
0.3519."

P_MGR <- 36/204
P_MGR_and_FBL <- 11/204
P_FBL_given_MGR <- P_MGR_and_FBL/P_MGR
sprintf("The probability green eyed male has a partner with blue eyes is
%1.4f.", P_FBL_given_MGR)

## [1] "The probability green eyed male has a partner with blue eyes is
0.3056."
```

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

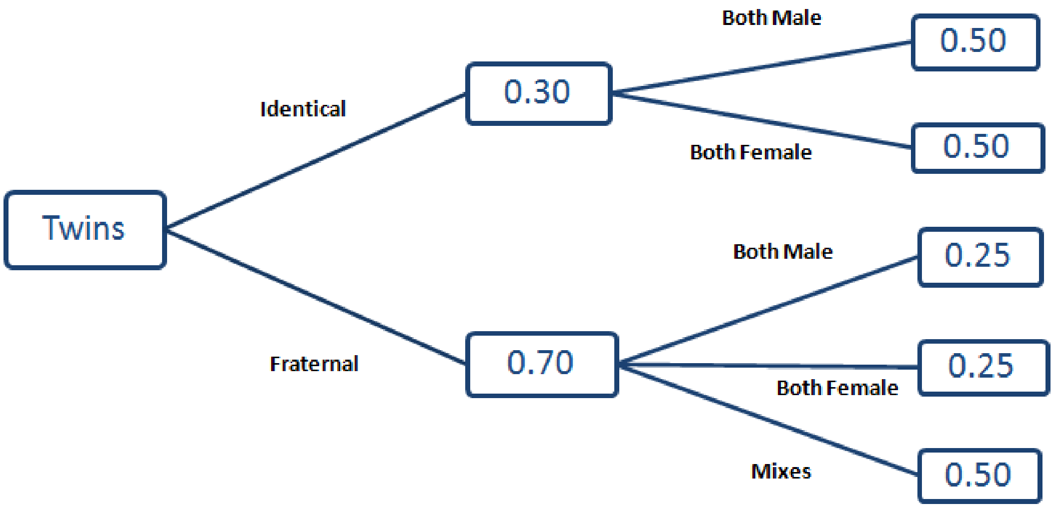
If eye colors of male respondents appear independent of their partners', then as an example, the following will be true: P(Male Blue Eyes and Partner Female Brown Eyes) = P(Male Blue Eyes) * P(Partner Female Brown Eyes)

```
P_MBL_and_FBR <- 23/204
P_MBL <- 114/204
P_FBR <- 55/204
sprintf("Is P_MBL_and_FBR: %1.4f equal to (P_MBL)*(P_FBR): %1.4f?",
P_MBL_and_FBR, P_MBL * P_FBR)
```

```
## [1] "Is P_MBL_and_FBR: 0.1127 equal to (P_MBL)*(P_FBR): 0.1507?"
```

As it can be seen, probabilities are different so it appears the eye colors of male respondents and their partners are independent.

2.26 Twins. About 30% of human twins are identical, and the rest are fraternal. Identical twins are necessarily the same sex - half are males and the other half are females. One-quarter of fraternal twins are both male, one-quarter both female, and one-half are mixes: one male, one female. You have just become a parent of twins and are told they are both girls. Given this information, what is the probability that they are identical?




```r
P_2G_given_Ident <- 0.3 * 0.5
P_2G_given_Frat <- 0.25 * 0.7
P_Ident_given_2G <- P_2G_given_Ident / (P_2G_given_Ident + P_2G_given_Frat)
sprintf("The probability is %1.4f.", P_Ident_given_2G)

## [1] "The probability is 0.4615."
```