Chapter 2 - Probability

Binish Kurian Chandy

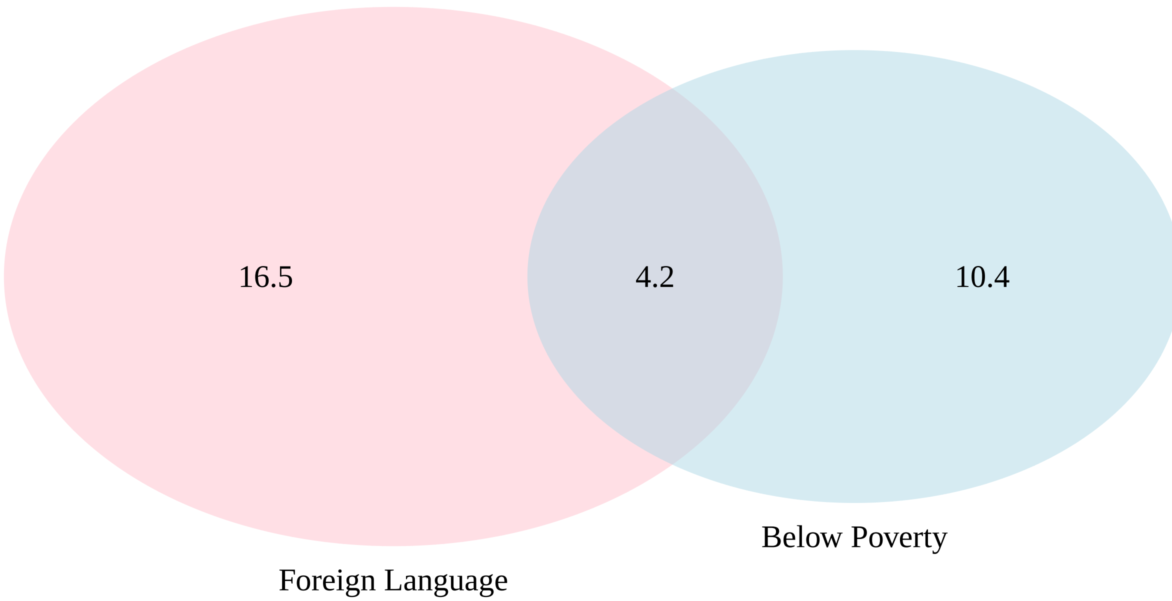
February 17, 2018

2.6

```
There are 36 ways the pair of fair dice can be thrown.
a. 0. The minimum sum is 2.
b. The sum 5 can appear 4 ways : (1,4), (2,3), (3,2), (4,1). So p = 4/36
c. The sum 12 can appear one way : (6,6). So p = 1/36
```

2.8

```
a. No,there are Americans who falls into two categories.
```



b.

```
## (polygon[GRID.polygon.1], polygon[GRID.polygon.2], polygon[GRID.polygon.3], polygon[GRID.polygon.4],
```

```
c. 10.4%
```

```
d. 31.1%
```

```
e. 68.9%
```

```
f. p(below poverty) * p(foreign language) = .104 * .165 = .01715 not equal to
p(below poverty and foreign language) ie .042. So the events are dependent.
```

2.20

```
a. A = Event of male partner having blue eyes
   B = Event of female partner having blue eyes
   P(AUB) = P(A) + P(B) - P(A AND B)
          = 114/204 + 108/204 - 78/204
          =.706
```

b. A = Event of male partner having blue eyes
B = Event of female partner having blue eyes
P(B/A) = P(A AND B) / P(A)
= 78/114
= .6842

c. A = Event of male partner having brown eyes
B = Event of female partner having blue eyes
P(B/A) = P(A AND B) / P(A)
= 19/54
= .352

A = Event of male partner having green eyes
B = Event of female partner having blue eyes
P(B/A) = P(A AND B) / P(A)
= 11/36
= .305

d. A = Event of male partner having blue eyes
B = Event of female partner having blue eyes
If the two events are independent then
P(A AND B) = P(A) * P(B)
P(A AND B) = 78/204 = .382
P(A) * P(B) = 114/204 * 108/204 = .2958
This shows that they are not independent.

2.30

a. P(Hardcover book) * P(Paperback fiction)
= 28/95 * 59/94 = .185

b. P(Fiction book) * P(Hardcover book)
= 72/95 * 28/94 = .2257

c. P(Fiction book) * P(Hardcover book)
= 72/95 * 28/95 = .2233

d. Since the sample space is relatively large, adding or removing a sample doesn't make much difference in the outcome.

2.38

a.

i	1	2	3	Total
xi	$0	$25	$60	
P(X=xi)	.54	.34	.12	1.00

Average revenue = E[X] = 0*.54 + 25*.54 + 60*.12
= 15.7
var = (0-15.7)^2*.54 + (25-15.7)^2*.34 + (60-15.7)^2*.12
= 398.01
SD = sqrt(398.01) = 19.95

b. Expected revenue for 120 passengers = 120 * E[X] = 1884
SD = sqrt((0-15.7)^2*.54*120 + (25-15.7)^2*.34*120 + (60-15.7)^2*.12*120)
= 218.54

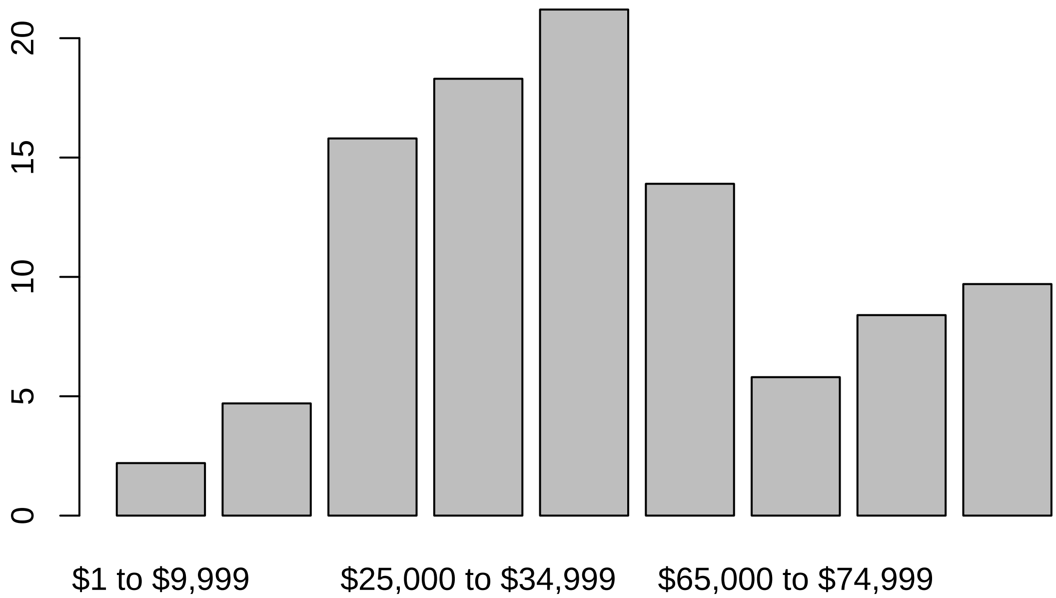
2.44

a.

```
sample_income <- c("$1 to $9,999","$10,000 to $14,999","$15,000 to $24,999",
                   "$25,000 to $34,999","$35,000 to $49,999",
                   "$50,000 to $64,999","$65,000 to $74,999",
                   "$75,000 to $99,999","$100,000 or more")

sample_percent <- c(2.2,4.7,15.8,18.3,21.2,13.9,5.8,8.4,9.7)

df <- data.frame(sample_income, sample_percent)
barplot(df$sample_percent, names.arg=df$'sample_income')
```




```
Fairly symmetric distribution with multiple modes

b. P(US resident makes < 50,000) = 1 - (.139+.058+.084+.097)
                                 = 62.2%

c. Assumption is income and gender are independent.
   P(US resident makes < 50,000 AND resident is female) = .622 * .41
                                                        = 26%

d. The new data shows the relationship ie females make less than $50k per year
   is 71.58%.
   This is different from what we calculated in step c.
   So the gender and income are not independent.
```