## 2.6

a) 0

b) 4/36

c) 1/36

## 2.8

a) No, living below the poverty line and speaking a foreign language are not disjoint. 4.2% fall into both categories.

b) Venn Diagram

Speak language other

English Live below poverty line

16.5% 4.2% 10.4%

c) 10.4% live below the poverty line and only speak English

d) 10.4% + 4.2% + 16.5% = 31.1%

e) 1-0.311 = 68.9%

f) P(LBP) x P(SLOE) =? P(LBP and SLOE)

0.146 x 0.207 = 0.0302 <> 4.2%

No, the events are not independent.

## 2.20

a) P(Mbl or Fbl) = P(Mbl) + P(Fbl) - P(Mbl and Fbl) = (114+108-78)/204 = 144/204 = 70.6%

b) P(Mbl | Fbl) = (78/204)/(108/204) = 72.21%

c) P(Mbr | Fbl) = (19/204)/(108/204) = 17.59%

P(Mg | Fbl) = (11/204)/(108/204) = 10.19%

d) Yes, it does appear that the eye colors of male respondents and their partners are independent.

72.21% + 17.59% + 10.19% = 100% (slightly off from rounding)

The probabilities are conditioned on the same information (Female partner having blue eyes) and they sum to 1. This indicates independence.

## 2.30

a) (28/95)\*(67/94) = 21%

b) (72/95)\*(28/94) = 22.6%

c) (72/95)\*(28/95) = 22.3%

d) When the sample size is only a small fraction of the population, observations are nearly independent even when sampling without replacement.

## 2.38

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *i* (Scenario) | 0 (no bag) | 1 (1 bag) | 2 (2 bags) | Total |
| *xi* | $0 | $25 | $60 |  |
| P(X= *xi*) | 0.54 | 0.34 | 0.12 |  |
| *xi*  x P(X= *xi*) | 0 | $8.50 | $7.20 | $15.70 |
| *xi* - µ | -15.20 | 9.30 | 44.30 |  |
| (*xi* - µ)2 | 231.04 | 86.49 | 1962.49 |  |
| (*xi* - µ)2 x P(X= *xi*) | 124.76 | 29.4 | 235.5 | 389.66 |

1. The average revenue per passenger is $15.70. The variance is 389.66, therefore the standard deviation is the sqrt of 389.66 = $19.74.
2. 120 \* $15.70 = $1,884 with a standard deviation of $19.74 assuming that the distribution is normally distributed. I am not sure this is justified since it seems to be right skewed.

## 2.44

1. The distribution is fairly symmetric and unimodal.
2. 2.2+4.7+15.8+18.3+21.2 = 62.2%
3. P(<50k and F) = P(<50k) x P(F) = .622\*.41 = 25.5% assuming these are independent
4. The assumption of independence is not valid seeing as the actual probability of P(<50k and F) is much higher than part c.