Illapani_Homework_CH2

Srini Illapani

September 12, 2015

Chapter 2 - Probability

2.6

Dice rolls. If you roll a pair of fair dice, what is the probability of

(a) getting a sum of 1?
- The minimum number for a six face dice is 1. So for a pair of dice, the least or minimum total possible is 2. So the probability of getting a sum of 1 is zero.

(b) getting a sum of 5?
- There are 4 possible ways of rolling a pair of dice that would result in a sum of 5. Dice 1 - 1, 2, 3, 4 Dice 2 - 4, 3, 2, 1

The probability of getting P(1, 4) = P(1) x P(2) = 1/6 x 1/6 = 1/36

The probability of getting P(2, 3) = P(2) x P(3) = 1/6 x 1/6 = 1/36

The probability of getting P(3, 2) = P(3) x P(2) = 1/6 x 1/6 = 1/36

The probability of getting P(4, 1) = P(4) x P(1) = 1/6 x 1/6 = 1/36

Therefore the probability of getting a sum of 5 is 1/36 + 1/36 + 1/36 + 1/36 = 4/36 = 1/9

(c) getting a sum of 12?
- There is only one possible way of rolling a pair of dice that would result in a sum of 12 Dice 1 - 6 Dice 2 - 6

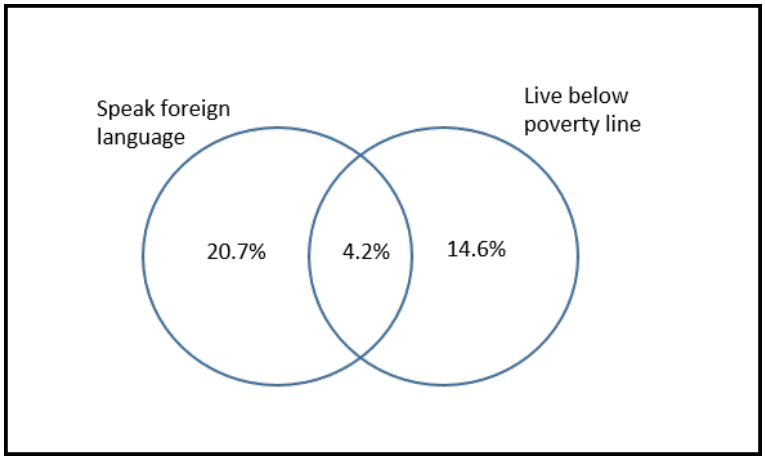
Therefore the probability of getting a sum of 12 is P(6) x P(6) = 1/6 x 1/6 = 1/36

2.8

Poverty and language. 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.59

(a) Are living below the poverty line and speaking a foreign language at home disjoint?
- No, they are non-disjoint. They can speak a foreign language and at the same time some of them can be below poverty line.

(b) Draw a Venn diagram summarizing the variables and their associated probabilities.



(c) What percent of Americans live below the poverty line and only speak English at home?

- 14.6 - 4.2 = 10.4%

(d) What percent of Americans live below the poverty line or speak a foreign language at home?

- (20.7 + 14.6) - 4.2 = 31%

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

- We know that the percentage of Americans that speak English are 79.3% and the percentage of Americans that are above poverty line are 85.4%. But we do not have the data that speak English and also above the poverty line.

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

- Yes, they are 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() = (114+108-78)/ 204 = 0.705 or 70.5%

(b) What is the probability that a randomly chosen male respondent with blue eyes has a partner with blue eyes?
   - P() = 78/114 = 0.68 or 68%

(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(Br, Bl) = 19/54 = 0.35 or 35%
   - P(Gr, Bl) = 11/36 = 0.30 or 30%

(d) Does it appear that the eye colors of male respondents and their partners are independent? Explain your reasoning.
   - The data shows that there is a greater probability of the partners having same color eyes. There is a corelation that is enough to say the colors are dependent.

2.30

Books on a bookshelf. The table below shows the distribution of books on a bookcase based on whether they are nonfiction or fiction and hardcover or paperback.

		Format		
		Hardcover	Paperback	Total
Type	Fiction	13	59	72
	Nonfiction	15	8	23
	Total	28	67	95

(a) Find the probability of drawing a hardcover book first then a paperback fiction book second when drawing without replacement.
   - P() = (28/95) * (59/94) = 0.184

(b) Determine the probability of drawing a fiction book first and then a hardcover book second, when drawing without replacement.
   - P1(Hardcover Fiction & Hardcover) = (13/95) * (27/94) = 0.03
   - P2(Paperback Fiction & Hardcover) = (59/95) * (28/94) = 0.18
   - P(Fiction first & Hardcover second) = P(1) + P(2)= 0.224 or 22.4%

(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.
   - P1(Fiction) = 72/95 = .75
   - P2(Hardcover) = 28/95 = .29
   - P() = P1*P2 = 0.22 or 22%

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

- The differenc eis small because we are only taking out 1 book out of a total of 95. This is not significant. Also, the hardcover books are less compared to paperback.

2.38

Baggage fees. 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.

Bags	Cost, A	P(A)	A * P(A)	$(A - E(A))^2$	$(A - E(A))^2 * P(A)$
0	$0.00	0.54	0	$246.49	$133.10
1	$25.00	0.34	8.5	$86.49	$29.41
2	$60.00	0.12	7.2	$1,962.49	$235.50
		E(A) =	**15.7**	**Variance =**	**$398.01**
				SD=	**19.95**

- The average revenue per passenger is $15.7 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.

- The revenue for 120 passengers would be 120 * $15.7 = $1,884
- The stadard deviation will be 19.95

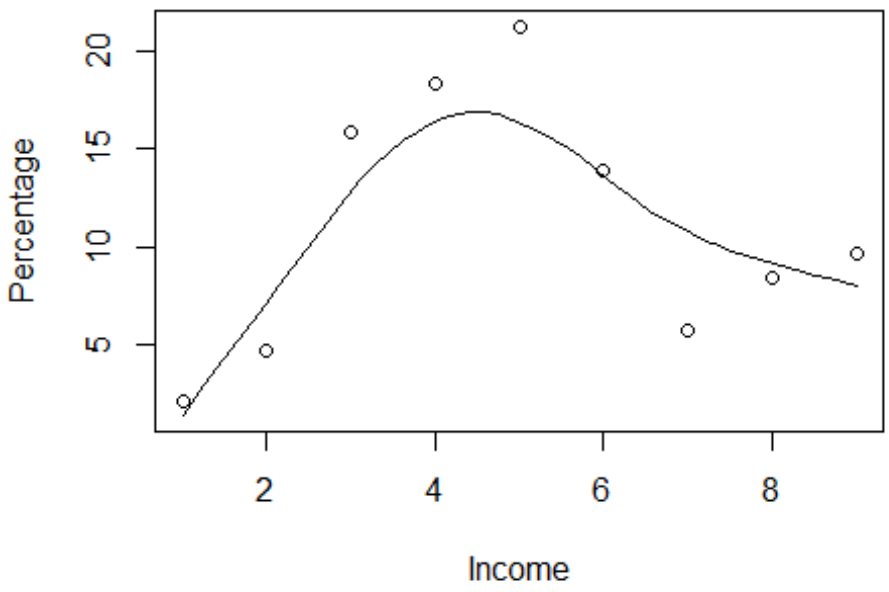
2.44

Income and gender. 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	*Total*
$1 to $9,999 or loss	2.2%
$10,000 to $14,999	4.7%
$15,000 to $24,999	15.8%
$25,000 to $34,999	18.3%
$35,000 to $49,999	21.2%
$50,000 to $64,999	13.9%
$65,000 to $74,999	5.8%
$75,000 to $99,999	8.4%
$100,000 or more	9.7%

(a) Describe the distribution of total personal income.

```
income <- read.csv("Incomes.csv", sep = ",")
scatter.smooth(income)
```




```
- The distribution of the personal income looks normal. Most of them are at t
he mid range of $35000 to $49999.
```

(b) What is the probability that a randomly chosen US resident makes less than $50,000 per year?

- The P(<$50k) = 62.2%

(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.

- Females = 41% and assuming that the income ranges represent men and women making equal income at all the ranges. P(F<$50k) = 62.2% * 41% = 25.5%

(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.

- My assumption is based on the men and women making equal income in all the ranges. The 71.8% means more women make less money compared to men. So my assumption is not valid.