

Chapter 2 – Probability Homework

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

- (a) getting a sum of 1?
- (b) getting a sum of 5?
- (c) getting a sum of 12?

Ans:

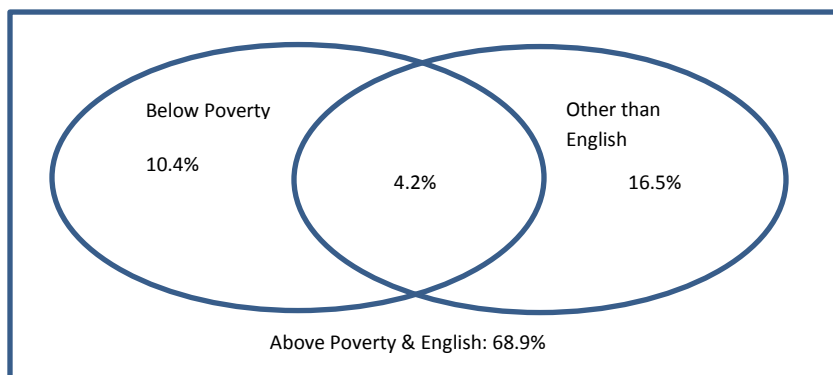
- a) $P(1) = 0$ (The minimum sum that we get with a fair dice is 2. So getting a 1 is not possible)
- b) $P(5) = 4 / 36 = 1 / 9 = 0.111$
- c) $P(12) = 1 / 36 = 0.0278$

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?
- (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?
- (d) What percent of Americans live below the poverty line or speak a foreign language at home?
- (e) What percent of Americans live above the poverty line and only speak English at home?
- (f) Is the event that someone lives below the poverty line independent of the event that the person speaks a foreign language at home?

Ans:

- a) No, they are not disjoint. There is a 4.2% that belong to both groups.
- b) Below is the diagram



- c) 10.4% of the Americans live below poverty line and only speak English.
- d) 31.1% (10.4 + 4.2 + 16.5) of the Americans live below poverty line and only a foreign language.
- e) 68.9% of the Americans live above the poverty line and speak English at home.
- f) If the events are independent then $P(\text{BPL} \mid \text{FL}) = P(\text{BPL})$
 Now $P(\text{BPL} \mid \text{FL}) = P(\text{BPL and FL}) / P(\text{FL}) = 4.2\% / 20.7\% = 20.3\%$ which is $\neq 14.6\%$.
 Hence the 2 events are dependent.

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.⁶⁵

		<i>Partner (female)</i>			Total
		Blue	Brown	Green	
<i>Self (male)</i>	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?
- (b) What is the probability that 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? What about the probability 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.

Ans:

- a) $P(\text{Male-Blue or Female-Blue}) = P(\text{Male-Blue}) + P(\text{Female-Blue}) - P(\text{Male-Blue and Female-Blue})$
 $= (114 / 204) + (108 / 204) - (78 / 204) = (114 + 108 - 78) / 204 = \mathbf{0.70588 = 70.59\%}$
- b) $P(\text{Female-Blue} \mid \text{Male-Blue}) = P(\text{Male-Blue and Female-Blue}) / P(\text{Male-Blue})$
 $= (78 / 204) / (114 / 204) = \mathbf{0.6842 = 68.42\%}$
- c) $P(\text{Female-Blue} \mid \text{Male-Brown}) = P(\text{Male-Brown and Female-Blue}) / P(\text{Male-Brown})$
 $= (19 / 204) / (54 / 204) = \mathbf{0.3519 = 35.19\%}$

 $P(\text{Female-Blue} \mid \text{Male-Green}) = P(\text{Male-Green and Female-Blue}) / P(\text{Male-Green})$
 $= (11 / 204) / (36 / 204) = \mathbf{0.3056 = 30.56\%}$
- d)

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.

	<i>Format</i>		<i>Total</i>
	Hardcover	Paperback	
<i>Type</i>	Fiction	13	59
	Nonfiction	15	8
	Total	28	67

- Find the probability of drawing a hardcover book first then a paperback fiction book second when drawing without replacement.
- Determine the probability of drawing a fiction book first and then a hardcover book second, when drawing without replacement.
- 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.
- The final answers to parts (b) and (c) are very similar. Explain why this is the case.

Ans:

- $(28 / 95) * (59 / 94) = 0.2947 * 0.6277 = \mathbf{0.1849 = 18.49\%}$
- $(72 / 95) * (28 / 94) = 0.7579 * 0.2979 = \mathbf{0.2258 = 22.58\%}$
- $(72 / 95) * (28 / 95) = 0.7579 * 0.2947 = \mathbf{0.2234 = 22.34\%}$
- The sample size is small compared to the population. In such cases, the observations are nearly independent even when sampling without replacement.

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.

- Build a probability model, compute the average revenue per passenger, and compute the corresponding standard deviation.
- 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.

Ans:

Event	X	P(X)	Y = X * P(X)	Power(X - Y, 2)
No Bags	\$ -	54%	\$ -	\$ 133.10
1 Bag	\$ 25.00	34%	\$ 8.50	\$ 29.41
2 Bags	\$ 60.00	12%	\$ 7.20	\$ 235.50
		100%	\$ 15.70	\$ 398.01

Average Revenue Per Passenger =	\$ 15.70
Variance =	\$ 398.01
SD =	19.950188

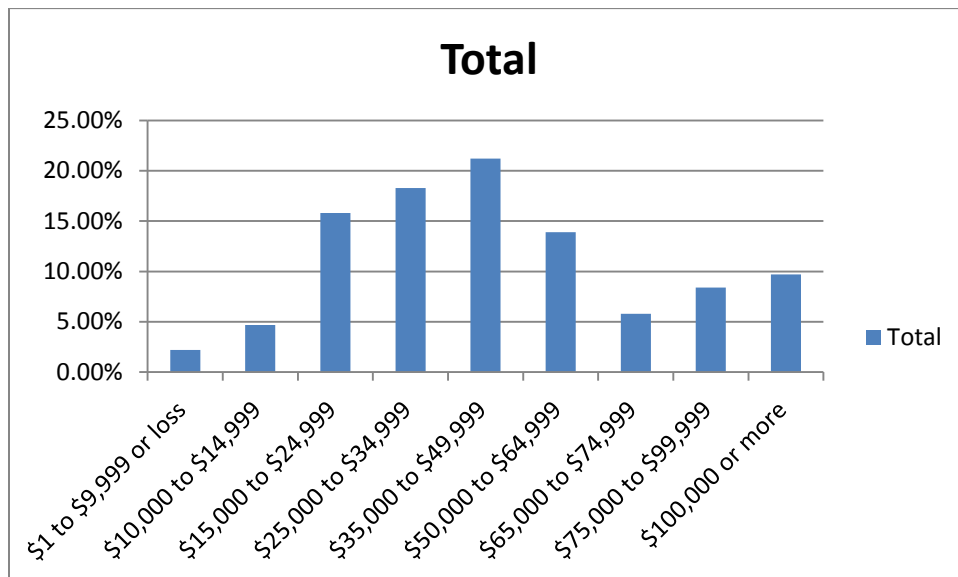
- Average Revenue per Passenger = \$ 15.7
Standard Deviation = \$ 19.95
- Revenue for 120 passengers = (\$15.7 * 120) = \$ 1884.0
The SD would be the same about \$ 19.95

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.⁶⁹

- Describe the distribution of total personal income.
- What is the probability that a randomly chosen US resident makes less than \$50,000 per year?
- 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.
- 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.

<i>Income</i>	<i>Total</i>
\$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%

Ans:



- The distribution looks to be unimodal with a slightly right skew with significant outliers.
- $P(< \$50K) = (2.2\% + 4.7\% + 15.8\% + 18.3\% + 21.2\%) = \mathbf{62.2\%}$
- If we assume that $P(\text{Female})$ and $P(< \$50K)$ are independent, then
 $P(< \$50K \text{ and Female}) = P(< \$50K) * P(\text{Female}) = 0.622 * 0.41 = \mathbf{0.2550 = 25.5\%}$
- With the new information we now have $P(< \$50K \text{ and Female}) = 71.8\%$. This is significantly different from what we had before. Hence, I would conclude that my assumption of independence is not valid.