

Homework 4

Alizeh Azar

9/26/2021

Problem 1: OIS Exercise 3.8, page 93 (5 pts)

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.

- (a) Are living below the poverty line and speaking a foreign language disjoint?

Answer:

NO, there are Americans who both live below the poverty line and speak other language other than English. They fall in both categories.

- (b) *should not included it in R Markdown.*

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

Answer:

$$14.6 - 4.2 = 10.4 \\ = 10.4\%$$

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

Answer:

Using the general Addition Rule:

$$0.146 + 0.207 - 0.042 = 0.311 \\ = 31.1\%$$

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

Answer:

$$1 - 0.311 = 0.689 \\ = 68.9\%$$

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

Answer:

$P(\text{living below poverty line}) \times P(\text{speaking a foreign language at home}) = 0.146 \times 0.207 = 0.0302$ which does not equal $P(\text{both categories}) = 0.042$. Therefore, the events are not independent.

Problem 2: OIS Exercise 3.11, page 94 (4 points)

The table below shows the distribution of education level attained by US residents by gender based on data collected in the 2010 American Community Survey (note that there are only two genders recorded).

<i>Highest education attained</i>	Male	Female
Less than 9th grade	0.07	0.13
9th to 12th grade, no diploma	0.10	0.09
HS graduate (or equivalent)	0.30	0.20
Some college, no degree	0.22	0.24
Associate's degree	0.06	0.08
Bachelor's degree	0.16	0.17
Graduate or professional degree	0.09	0.09
Total	1.00	1.00

- (a) What is the probability that a randomly chosen man has at least a Bachelor's degree?

Answer:

$$0.16 + 0.09 = 0.25$$

- (b) What is the probability that a randomly chosen woman has at least a Bachelor's degree?

Answer:

$$0.17 + 0.09 = 0.26$$

- (c) What is the probability that a man and woman getting married both have at least a Bachelor's degree?

Note any assumptions you must make to answer this question.

Answer:

$$0.25 \times 0.26 = 0.065$$

Assumptions:

- The education level of husband and wife are independent
- The decision to get married is not related to education level

- (d) If you made an assumption in part (c), do you think it was reasonable? If you didn't make an assumption, double check your earlier answer and then return to this part.

Answer:

The education level of husband and wife independent assumption is not reasonable because often people marry another person with a comparable level of education.

The decision to get married is not related to education level assumption is also not reasonable because some people oftenly decide to marry after attaining a certain education level.

Problem 3: Conditional Probability in R (7 points)

Refer to the Titanic dataset we used in class.

```
titanic <- read.csv("http://vincentarelbundock.github.io/Rdatasets/csv/carData/TitanicSurvival.csv")
```

- (a) Using titanic, create a table comparing the survived and sex variables (again, there are only two sexes recorded). Print the table. (1 pt)

```
library(summarytools)
ctable(titanic$survived,titanic$sex,prop = "n")
```

Cross-Tabulation

survived * sex

Data Frame: titanic

	sex	female	male	Total
survived				
no		127	682	809
yes		339	161	500
Total		466	843	1309

- (b) According to the table, what is the probability that a passenger survived? (1 pt)

Answer:

$$500 \div 1309 = 0.381971$$

- (c) According to the table, what is the probability that a passenger was male? (1 pt)

Answer:

$$843 \div 1309 = 0.6440031$$

- (d) According to the table, what is the probability a passenger was male and survived? (1 pt)

Answer:

$$162 \div 1309 = 0.1237586$$

- (e) According to the table, what is the probability of a passenger surviving given that the passenger was male? (1 pt)

Answer:

$$162 \div 843 = 0.1921708$$

- (f) Are sex and survival independent? Give a mathematical justification! (2 pts)

Answer:

Sex and survival are not independent. Probabilities calculated above are different and show that sex and survival are not independent. That is, $P(\text{male and survived})$ is not equal to $P(\text{survived} | \text{male})$

Problem 4 (4 pts)

Of volunteers coming into a blood center, 1 in 3 have type O^+ blood, 1 in 15 have O^- , 1 in 3 have A^+ , and 1 in 16 have A^- (Note: These blood types are mutually exclusive– you cannot have both type O^- and A^+ at the same time). This distribution can be summarized by the following table:

Blood type	O^+	O^-	A^+	A^-
$p(\text{blood type})$	$\frac{1}{3}$	$\frac{1}{15}$	$\frac{1}{3}$	$\frac{1}{16}$

- (a) Is this distribution described in the table a valid probability distribution? Why or why not?

Answer:

The distribution described in the table is not a valid probability distribution. This is because the sum of the probabilities does not sum to 1.

- (b) What is the probability that the randomly selected person has type O blood?

Answer:

$$\frac{1}{3} + \frac{1}{15} = 0.4$$

- (c) What is the probability that the randomly selected person has type A blood?

Answer:

$$\frac{1}{3} + \frac{1}{16} = 0.3958333$$

- (d) What is the probability that the randomly selected person has neither type A or type O blood?

Answer:

$$1 - \left(\frac{1}{3} + \frac{1}{16} + \frac{1}{3} + \frac{1}{15}\right) = 0.2041667$$

Problem 5 (5 pts)

A local college has three types of rooms in their dormitories: singles, doubles, and triples. Thus, the distribution of the number of roommates a randomly chosen student has, M , is as follows:

M	0	1	2
$p(M)$	$\frac{1}{5}$	$\frac{1}{2}$	$\frac{3}{10}$

- (a) In R, calculate the expected value of the number of roommates a randomly chosen student has. Save it as `exp_roommates`. (2pts)

```
v1 <- c(0,1,2)
v2 <- c(0.2,0.5,0.3)
```

```
exp_roommates <- sum(v1*v2);exp_roommates
```

```
## [1] 1.1
```

- (b) In R, calculate the variance of the number of roommates a randomly chosen student has. Save it as `var_roommates`. (3 pts)

```
v3 <- v1^2
```

```
var_roommates <- sum((exp_roommates-v1)^2*v2);var_roommates
```

```
## [1] 0.49
```