Week 2 Assignment Data Science Math

Exercises 2.16, 2.18, 2.20 and 2.26 p. 108 – 114

Exercise 2.16 PB & J. Suppose 80% of people like peanut butter, 89% like jelly, and 78% like both. Given that a randomly selected person likes peanut butter, what’s the probability that he also likes jelly?

P( jelly | peanut butter ) = P( peanut butter and jelly ) = 0.78 = 0.975

P( peanut butter ) 0.80

Exercise 2.18 Weight and Health coverage, Part II. Exercise 2.14 introduced a contingency table summarizing the relationship between weight status, which is determined based on body mass index (BMI), and health coverage for a sample of 428,638 Americans. In the table below, the counts have been replaced by relative frequencies (probability estimates)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  | Weight Status | | |  |
|  |  | Neither overweight | Overweight | Obese |  |
|  |  | nor obese (BMI < 25) | (25 ≤ BMI ≤ 30) | (BMI ≥ 30) | Total |
| Health | Yes | 0.3145 | 0.3306 | 0.2503 | 0.8954 |
| Coverage | No | 0.0352 | 0.0358 | 0.0336 | 0.1046 |
|  | Total | 0.3497 | 0.3664 | 0.2839 | 1.0000 |

1. What is the probability that a randomly chosen individual is obese? 0.2839
2. What is the probability that a randomly chosen individual is obese given that he has health coverage?

P ( obese | health coverage ) = P ( obese and health coverage ) = 0.2503 = 0.2795

P ( health coverage ) 0.8954

1. What is the probabaility that a randomly chosen individual is obese given that he doesn’t have health coverage?

P ( obese | no health coverage ) = P ( obese and no health coverage ) = 0.0336 = 0.3212

P ( no health coverage ) 0.1046

1. Do being overweight and having health coverage appear to be independent?

If being overweight and having health coverage are independent then

P ( health coverage | overweight ) should be equal to ( or relatively close to, given that there is variability in any population involving actual data) P (health coverage)

P ( health coverage | overweight ) = P (health coverage and overweight ) = 0.3306 = 0.9022

P ( overweight ) 0.3664

P (health coverage) = 0.8954

Given natural variability, being overweight and having health coverage appear to be independent.

Exercise 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 |
|  | Blue | 78 | 23 | 13 | 114 |
| Self (Male) | Brown | 19 | 23 | 12 | 54 |
|  | Green | 11 | 9 | 16 | 36 |
|  | Total | 108 | 55 | 41 | 204 |

1. What is the probability that a randomly chosen male respondent or his partner has blue eyes?

30

Female blue eyes

78 Male and female blue

eyes

36

Male blue eyes

55 neither male nor female blue eyes

P ( Male = Blue or Female = Blue ) = (36 + 78 + 30) / 204 = 144 / 204 = 0.7059

1. What is the probability that a randomly chosen male respondent with blue eyes has a partner with blue eyes? 78 / 204 = 0.3824
2. What is the probability that a randomly chosen male respondent with brown eyes has a partner with blue eyes? 19 / 204 = 0.0931

What about the probability of a randomly chosen male respondent with green eyes having a partner with blue eyes? 11 / 204 = 0.0539

1. Does it appear that the eye colors of male respondents and their partners are independent? Explain your reasoning.

If the eye colors of the male respondents and their partners are independent then the color of the male partners eyes would give no indication of the color of the female partners eyes. If the two are independent then P( Female eye color | Male eye color ) should be equal to

P( Female eye color). Looking at the numbers, we can see this is not the case.

P(female blue | male blue) = P (female blue and male blue ) = (78/204) = 78 = 0.6842

P ( male blue ) (114/204) 114

Whereas P(female blue) = 108 / 204 = 0.5294

P(female brown | male brown) = P (female brown and male brown) = (23/204) = 23 = 0.4259

P ( male brown ) (54/204) 54

Whereas P(female brown) = 55 / 204 = 0.2696

and

P(female green | male green) = P (female green and male green) = (16/204) = 16 = 0.4444

P ( male green ) (36/204) 36

Whereas P(female green) = 41 / 204 = 0.2010

The eye color of the female partneris not independent of the eye color of the male partner, the relationship is dependent.

Exercise 2.26 Twins. About 30% of human twins are identical, and the rest are fraternal. Identical twins are necessarily the same sex – half are males and the other half are females. One-quarter of fraternal twins are both male, one-quarter both female, and one-half are mixes: one male, one female. You have just become a parent of twins and are told they are both girls. Given this information, what is the probability that they are identical?

0.7 x 0.5 = 0.35

0.7 x 0.25 = 0.175

0.7 x 0.25 = 0.175

0.3 x 0.5 = 0.15

0.3 x 0.5 = 0.15

0.5

0.25

0.25

0.5

0.5

0.7

0.3

P(Both Female) = 0.15 + 0.175 = 0.325

P(Identical) = 0.3

P(Fraternal) = 0.7

P(Both Female | Identical) = P (Both Female and Identical) / P(Identical) = 0.15 / 0.3 = 0.5

P(Both Female | Fraternal) = P(Both Female and Fraternal) / P(Fraternal) = 0.175 / 0.7 = 0.25

P(Identical | Both Female) =

= = = = 0.4615

Given that both are girls, the probability of identical twins increases from 30% to 46%.