

Week 6 - Combinatronics / Conditional Probability

Christian Thieme
9/29/2020

1. A box contains 54 red marbles, 9 white marbles, and 75 blue marbles. If a marble is randomly selected from the box, what is the probability that it is red or blue? Express your answer as a fraction or a decimal number rounded to four decimal places.

We can solve this question using the probability of an event rule:

$$P(\text{event}) = \frac{\text{outcomes that meet our criteria}}{\text{all possible outcomes}}$$

Here the outcomes that meet our criteria are that the marble is either blue or red. We'll add the count of red marbles to the count of blue marbles, and then divide by the total amount of marbles.

```
red <- 54
white <- 9
blue <- 75

total <- red + white + blue

prob <- round((red + blue) / total,4)
```

The probability is 0.9348.

2. You are going to play mini golf. A ball machine that contains 19 green golf balls, 20 red golf balls, 24 blue golf balls, and 17 yellow golf balls, randomly gives you your ball. What is the probability that you end up with a red golf ball? Express your answer as a simplified fraction or a decimal rounded to four decimal places.

We will use the same formula as mentioned in #1, however, here the outcome that meets our criteria is a red golf ball which we'll divide by the total number of golf balls.

```
green <- 19
red <- 20
blue <- 24
yellow <- 17

total <- green + red + blue + yellow

prob2 <- red / total
```

The probability is 0.25.

3. A pizza delivery company classifies its customers by gender and location of residence. The research department has gathered data from a random sample of 1,399 customers. The data is summarized in the table below. What is the probability that a customer is not male or does not live with parents? Write your answer as a fraction or a decimal number rounded to four decimal places.

Gender and Residence of Customers	
	Males/Females
Apartment	81/228
Dorm	116/79
With Parent(s)	215/252
Seniorly Parent(s) Home	130/97
Other	129/72

```
total <- 81 + 228 + 116 + 79 + 215 + 252 + 130 + 97 + 129 + 72

males <- 81 + 116 + 215+ 130 + 129

females_with_parents <- 252

prob3 <- round((total - males - females_with_parents) / total,4)
```

The probability is 0.3402.

4. Determine if the following events are independent: Going to the gym and losing weight

A. Going to the gym and losing weight are dependent events. A person's weight is often positively affected by going to the gym. Gym attendance also often leads to other healthy habits such as eating healthier, which in turn can lead to weight loss as well.

5. A veggie wrap at City Subs is composed of 3 different vegetables and 3 different condiments wrapped up in a tortilla. If there are 8 vegetables, 7 condiments, and 3 types of tortilla available, how many different veggie wraps can be made?

The question above does not state that order is important, so I'll assume that the question is asking how many combination can be made with the different ingredients (not permutations). We can determine the number of combinations by using the following formula for each ingredient and then multiplying the answers together (product rule):

$${}_nC_k = \binom{n}{k} = \frac{n!}{k!(n-k)!}$$

We will use the `choose` function in R to generate these combinations as opposed to working them by hand.

```
library(arrangements) #library to see combinations
```

```
veggies <- choose(8,3)

condiments <- choose(7,3)

combinations <- veggies * condiments * 3
```

There are 5880 combinations possible.

Sometimes looking at the above number seems unbelievable. We can use the `arrangements` library to view the combinations, which makes it easier for us to wrap our head around what is happening. Let's look at the combinations we can create if we had 7 condiments and could only pick 3.

```
arrangements::combinations(7,3)
```

```
##      [,1] [,2] [,3]
## [1,] 1 2 3
## [2,] 1 2 4
## [3,] 1 2 5
## [4,] 1 2 6
## [5,] 1 2 7
## [6,] 1 3 4
## [7,] 1 3 5
## [8,] 1 3 6
## [9,] 1 3 7
## [10,] 1 4 5
## [11,] 1 4 6
## [12,] 1 4 7
## [13,] 1 5 6
## [14,] 1 5 7
## [15,] 1 6 7
## [16,] 2 3 4
## [17,] 2 3 5
## [18,] 2 3 6
## [19,] 2 3 7
## [20,] 2 4 5
## [21,] 2 4 6
## [22,] 2 4 7
## [23,] 2 5 6
## [24,] 2 5 7
## [25,] 2 6 7
## [26,] 3 4 5
## [27,] 3 4 6
## [28,] 3 4 7
## [29,] 3 5 6
## [30,] 3 5 7
## [31,] 3 6 7
## [32,] 4 5 6
## [33,] 4 5 7
## [34,] 4 6 7
## [35,] 5 6 7
```

You can see that this alone generates 35 unique combinations without replacement. This multiplied by the 56 combinations from the veggies and the 3 from the tortillas gets us to our answer. Visualizing it in this way helps us to understand why the final combinations are so large for a seemingly small set of choices.

6. Determine if the following events are independent: Jeff runs out of gas on the way to work and Liz watches the evening news.

B. Independent - these events are completely separate and have no effect on each other whatsoever UNLESS Jeff ran out of gas because Liz was watching the evening news instead of filling up Jeff's truck.

7. The newly elected president needs to decide the remaining 8 spots available in the cabinet she is appointing. If there are 14 eligible candidates for these positions (where rank matters), how many different ways can the members of the cabinet be appointed?

When rank matters, we use permutation. The following formula is used to calculate permutations:

$${}_nP_k = \frac{n!}{(n-k)!}$$

Instead of doing the calculation manually by hand, we can use the R function, `permutations` from the `gtools` package, which will output a row for every permutation. We'll then use `nrow` to summarize how many rows there are. From the formula above n is 14 and k is 8.

```
library(gtools)

x <- seq(1:14)

ways <- nrow(gtools::permutations(n = 14, r = 8, v = x))
```

There are 121080960 different ways that members of the cabinet can be appointed.

8. A bag contains 9 red, 4 orange, and 9 green jellybeans. What is the probability of reaching into the bag and randomly withdrawing 4 jelly beans such that the number of red ones is 0, the number of orange ones is 1, and the number of green ones is 3? Write your answer as a fraction or a decimal number rounded to four decimal places.

To get this set up well start with the following information:

```
red <- 9
orange <- 4
green <- 9

total <- red + orange + green
total
```

```
## [1] 22
```

We can see that there are a total of 22 jelly beans. Now we need to see how many combinations we can make from each of the colors. Conceptually, this is a little difficult. The best way to think about this is to imagine that you have a magic marker and you write the number 1-22 on each jelly bean and that as you go through and determine the possible combinations for each color, that not only does the color matter, but also the actual number on the bean you are selecting as well. In saying this, when looking at the combinations for green jelly beans, initially it seems like what order you select a green jelly bean doesn't matter because each time you come back with green - however, imagining there is a number on each green jelly bean helps to illustrate that each green jelly bean is its own choice and is separate from another green jelly bean. That being said, we calculate the combinations for each color below using the `choose` function.

```
red_combos <- choose(9,0)
orange_combos <- choose(4,1)
green_combos <- choose(9,3)
```

In our calculations above, if we choose 0 out of 9 red jelly beans, we come back with 1 combinations, which follows the rule of factorials that when we take the factorial of 0 we get 1. There are 4 different ways to select a single orange jelly bean from the 4, which makes sense intuitively. Lastly, there are 84 ways to select 3 different green jelly beans from the 9 green jelly beans.

We will multiply these three values together to get the numerator for our final calculation, which tells us how many total choices there are taking into consideration the color of each jelly bean:

```
red_combos * orange_combos * green_combos
```

```
## [1] 336
```

Now, we need to determine how many TOTAL combinations are possible for selecting 4 jelly beans out of the 22 jelly beans. Imagine again that each jelly bean has the number 1-22 written on it.

```
total_combos <- choose(22,4)
```

To find the probability, we'll divide the combinations of different colors we found above into the total combinations possible:

```
probability <- round((red_combos * orange_combos * green_combos) / total_combos,4)
```

The final probability of selecting 0 red, 1 orange, and 3 green jelly beans out of these 22 is 0.0459.

9. Evaluate the following expression:

```

$$\frac{11!}{7!}$$
factorial(9) / (factorial(3) * factorial(9-3))
```

```
## [1] 84
```

10. Describe the compliment of the given event: 67% of subscribers to a fitness magazine are over the age of 34.

When it comes to percents, our percentage is a total of something, meaning it should total to 100%. That being said, in our question we are told that 67% of subscribers to a fitness magazine are over the age of 34. Looking at the compliment, we would say, ok so what percentage doesn't meet that criteria? In this case, it would be the 33% of individuals who are 34 or younger.

11. If you throw exactly three heads in four tosses of a coin you win \$97. If not, you pay me \$30.

Step 1. Find the expected value of the proposition. Round your answer to two decimal places.

First, we need to find all the possible permutations of a fair coin being flipped four times. Below we see that this will result in 16 different permutations. Next, we can count up how many of these permutation have exactly three heads, which we can see is 4.

```
x <- c('heads', 'tails')
gtools::permutations(n = 2, r = 4, v = x, repeats.allowed = TRUE)
```

```
##      [,1] [,2] [,3] [,4]
## [1,] "heads" "heads" "heads" "heads"
## [2,] "heads" "heads" "heads" "tails"
## [3,] "heads" "heads" "tails" "heads"
## [4,] "heads" "heads" "tails" "tails"
## [5,] "heads" "tails" "heads" "heads"
## [6,] "heads" "tails" "heads" "tails"
## [7,] "heads" "tails" "tails" "heads"
## [8,] "heads" "tails" "tails" "tails"
## [9,] "tails" "heads" "heads" "heads"
## [10,] "tails" "heads" "heads" "tails"
## [11,] "tails" "heads" "tails" "heads"
## [12,] "tails" "heads" "tails" "tails"
## [13,] "tails" "tails" "heads" "heads"
## [14,] "tails" "tails" "heads" "tails"
## [15,] "tails" "tails" "tails" "heads"
## [16,] "tails" "tails" "tails" "tails"
```

This value of 4 will be divided by the total number of permutations, 16, to get the probability of you winning the bet. Next, to get the expected value, we'll use the following formula where we multiply the expected value of you winning by the winnings, and then we multiply the probability of you losing by the amount you'd lose and add both of those together:

$$E(x) = P(4/16)(97) + P(1 - (4/16))(-30)$$

```
p <- 4/16

expected_value <- 97 * p + (-30) * (1-p)
```

Here we see the expected value is \$1.75, which is positive, meaning this is a favorable bet for you that you should take.

Step 2. If you played this game 559 times how much would you expect to win or lose? (Losses must be entered as negative.)

```
expected_winnings <- expected_value * 559
```

If you played this game 559 times you could expect to win: **\$978.25**.

12. Flip a coin 9 times. If you get 4 tails or less, I will pay you \$23. Otherwise you pay me \$26.

Step 1. Find the expected value of the proposition. Round your answer to two decimal places.

```
x <- c('heads', 'tails')

perms <- gtools::permutations(n = 2, r = 9, v = x, repeats.allowed = TRUE)

total <- nrow(perms)

tails_count <- expss::count_row_if("tails", perms)
less_than_4_tails <- length(tails_count[tails_count <= 4])

probability <- less_than_4_tails / total
```

```
expected_val <- 23 * probability + -26 * probability
expected_val
```

```
## [1] -1.5
```

Here we see the expected value is -1.5, which is negative, meaning this is an unfavorable bet for you that you should not take.

Step 2. If you played this game 994 times how much would you expect to win or lose? (Losses must be entered as negative.)

```
expected_loss <- expected_val * 994
```

If you played this game 994 times you could expect to lose: **-1491**.

13. The sensitivity and specificity of the polygraph has been a subject of study and debate for years. A 2001 study of the use of polygraph for screening purposes suggested that the probability of detecting a liar was .59 (sensitivity) and that the probability of detecting a "truth teller" was .90 (specificity). We estimate that about 20% of individuals selected for the screening polygraph will lie.

To solve the below problems, we will use Bayes' Theorem:

$$P(A_1|B) = \frac{P(B|A_1)P(A_1)}{\sum P(B|A_i)P(A_i)}$$

a. What is the probability that an individual is actually a liar given that the polygraph detected him/her as such?

- Let B = polygraph detected a liar
- Let A = individual is a liar
- $P(B|A1) = .59$ sensitivity
- $P(A1) = .20$ liars
- $P(B|A2) = 100 - .90 = .10$ will be called liars but are "truth tellers".
- $P(A2) = .80$ are not liars

$$P(\text{liar}|\text{polygraph detected liar}) = \frac{(.59)(.20)}{(.59)(.20) + (1-.59)(.80)} = \frac{0.118}{0.118 + 0.082} = \frac{0.118}{0.198} = 0.5959$$

b. What is the probability that an individual is actually a truth-teller given that the polygraph detected him/her as such?

- Let B = polygraph detected a truth teller
- Let A = individual is a truth-teller
- $P(B|A1) = .9$ sensitivity
- $P(A1) = .80$ tell the truth
- $P(B|A2) = 100 - .59 = .41$ will be called liars but are "truth tellers".
- $P(A2) = .20$ are not liars

$$P(\text{truth teller}|\text{polygraph detected truth teller}) = \frac{(.90)(.80)}{(.90)(.80) + (1-.59)(.20)} = \frac{0.72}{0.72 + 0.082} = \frac{0.72}{0.802} = 0.8976$$

c. What is the probability that a randomly selected individual is either a liar or was identified as a liar by the polygraph?

This question can be answered using the addition rule: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

- $P(A) = P(\text{liar}) = .20$
- $P(B) = P(\text{detected as liar}) = (.59(\text{liars detected}) * .20(\text{total liars})) + (.10(\text{truth tellers detected as liars}) * .80(\text{those telling the truth})) = 0.198$
- $P(A \cap B) = .20 * .59 = 0.118$ - liars who were detected as liars - removed so not double counting

The probability is:

$$P(A \cup B) = 0.20 + 0.198 - 0.118 = 0.28$$