STAT 630 Problem Set 1

Nikko Cleri

September 10, 2021

Question I.

1.2.2

Suppose $S = \{1, 2, 3, 4, 5, 6, 7, 8\}$, with $P(\{s\}) = 1/8$ for $1 \leq s \leq 8$

(a) What is $P(\{1, 2\})$?

Since $P(1) = P(2) = 1/8$, then $P(\{1, 2\}) = P(1) + P(2) = 2/8 = 1/4$.

(b) What is $P(\{1, 2, 3\})$?

By the same arguments as part (a), $P(\{1, 2, 3\}) = P(1) + P(2) + P(3) = 3/8$.

(c) How many events A are there such that $P(A) = 1/2$?

Any combination of four outcomes in an event will give a probability of $P(A) = 1/2$ (order does not matter), so $\binom{8}{4} = 70$ events.

1.2.9

Suppose $S = \{1, 2, 3, 4\}$ and $P(\{1\}) = 1/12$, and $P(\{1, 2\}) = 1/6$, and $P(\{1, 2, 3\}) = 1/3$. Compute $P(\{1\})$, $P(\{2\})$, $P(\{3\})$, $P(\{4\})$.

$P(\{1\}) = 1/12$ is given,

$P(\{1, 2\}) = 1/6 = P(\{1\}) + P(\{2\}) \implies P(\{2\}) = 1/12$

$P(\{1, 2, 3\}) = 1/3 = P(\{1\}) + P(\{2\}) + P(\{3\}) \implies P(\{3\}) = 1/6$

$P(\{4\}) = 1 - P(\{1, 2, 3\}) = 2/3$

Sanity check: $P(\{1\}) + P(\{2\}) + P(\{3\}) + P(\{4\}) = 1/12 + 1/12 + 1/6 + 2/3 = 1$

Question II.

Two 6-sided dice are thrown sequentially and the values they show are recorded.

(a) List the sample space.

$$
\begin{aligned}
S = \{&(1,1), (1,2), (1,3), (1,4), (1,5), (1,6),\\
&(2,1), (2,2), (2,3), (2,4), (2,5), (2,6),\\
&(3,1), (3,2), (3,3), (3,4), (3,5), (3,6),\\
&(4,1), (4,2), (4,3), (4,4), (4,5), (4,6),\\
&(5,1), (5,2), (5,3), (5,4), (5,5), (5,6),\\
&(6,1), (6,2), (6,3), (6,4), (6,5), (6,6)\}
\end{aligned}
$$

(b) List the outcomes that make up the following events: A = "the sum of the two values is at least 9", B = "the value of the first die is higher than the value of the second die", and C = "the second die has value 4".

$A = \{(3,6),(4,5),(4,6),(5,4),(5,5),(5,6),(6,3),(6,4),(6,5),(6,6)\}$

$B = \{(2,1),(3,1),(3,2),(4,1),(4,2),(4,3),(5,1),(5,2),(5,3),(5,4),(6,1),(6,2),(6,3),(6,4),(6,5)\}$

$C = \{(1,4),(2,4),(3,4),(4,4),(5,4),(6,4)\}$

(c) List the elements of the following events; $A \cap C$, $B \cup C$, $A \cap (B \cup C)$.

$A \cap C = \{(5,4),(6,4)\}$

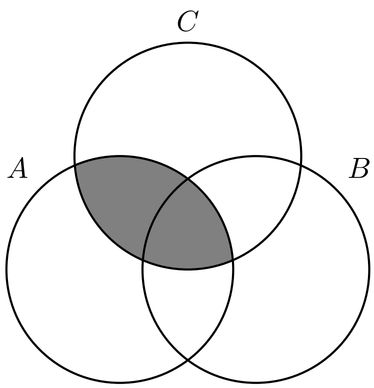
$B \cup C = \{(2,1),(3,1),(3,2),(4,1),(4,2),(4,3),(5,1),(5,2),(5,3),(5,4),$
$(6,1),(6,2),(6,3),(6,4),(6,5),(1,4),(2,4),(3,4),(4,4)\}$

$A \cap (B \cup C) = \{(5,4),(6,3),(6,4),(6,5)\}$

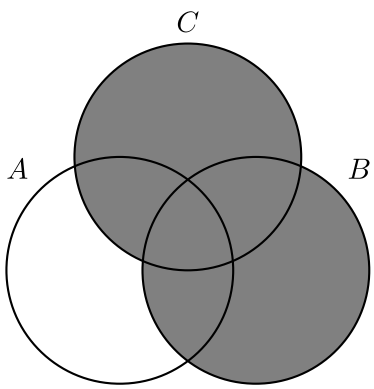
(d) Illustrate the events of part (c) in Venn diagrams with regions depicting A, B, and C.

The sets $A \cap C$, $B \cup C$, $A \cap (B \cup C)$ are denoted by the gray shaded regions in the following Venn diagrams.

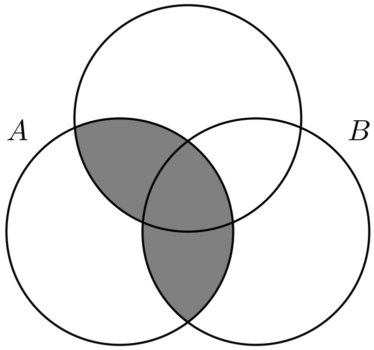
$A \cap C$



$B \cup C$



$A \cap (B \cup C)$



(e) Assume the outcomes are equally likely (they each have the same probability) and find the probabilities of the events in part (c).

$P(A \cap C) = 2/36 = 1/18$

$B \cup C = 19/36$

$P(A \cap (B \cup C)) = 4/36 = 1/9$

(f) Can $P(A \cap C)$ be computed by multiplying the probabilities of A and C?

No. $P(A \cap C) = 1/18$ and $P(A)P(C) = (10/36)(6/36) = 60/1296 = 5/108$. These are not equal, meaning A and C are not independent events.

(g) Imagine this experiment being repeated many times. What would be the longterm proportion of all of the experiments for which the sum of the two dice is 7?

There are 6 outcomes resulting in the sum of the two dice equaling 7, so the longterm proportion goes to $6/36 = 1/6$.

Question III.

1.3.2

Suppose that Al watches the six o'clock news 2/3 of the time, watches the eleven o'clock news 1/2 of the time, and watches both the six o'clock and eleven o'clock news 1/3 of the time. For a randomly selected day, what is the probability that Al watches only the six o'clock news? For a randomly selected day, what is the probability that Al watches neither news?

Notation for this problem: $P(6)$ is the probability of watching the 6 o'clock news. $P(11)$ is the probability of watching the 11 o'clock news.

The probability that Al watches *only* the 6 o'clock news is $P(6) - P(6 \cap 11) = 2/3 - 1/3 = 1/3$.

The probability that Al watches neither news is

$$\begin{aligned} P(none) &= 1 - P(6 \cup 11) \\ &= 1 - (P(6) + P(11) - P(6 \cap 11)) \\ &= 1 - (2/3 + 1/2 - 1/3) \\ &= 1/6 \end{aligned}$$

1.3.4

Suppose your right knee is sore 15% of the time, and your left knee is sore 10% of the time. What is the largest possible percentage of time that at least one of your knees is sore? What is the smallest possible percentage of time that at least one of your knees is sore?

The largest possible time will be the the sum of the times (no overlap in knee pain), so 25%. The smallest possible percentage will be when there is maximum overlap in knee pain, which means that all left knee pain occurs while the right knee is also in pain, so the minimum is 15%.

1.3.8

8 Suppose 55% of students are female, of which 4/5 (44%) have long hair, and 45% are male, of which 1/3 (15% of all students) have long hair. What is the probability that a student chosen at random will either be female or have long hair (or both)?

Notation for this problem: $P(F)$ is the probability of being female, $P(M)$ is the probability of being male, $P(L)$ is the probability of having long hair.

To find: $P(F \cup L)$

$$\begin{aligned} P(F \cup L) &= P(F) + P(L) - P(F \cap L) \\ &= .55 + (.44 + .15) - .44 \\ &= .70 \end{aligned}$$

1.3.10(a)

Suppose there are three events, A, B, and C. Prove that

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$$

To start this proof:

$$\begin{aligned} P(A \cup B \cup C) &= P(A \cup (B \cup C)) \\ &= P(A) + P(B \cup C) - P(A \cap (B \cup C)) \\ &= P(A) + P(B) + P(C) - P(A \cap (B \cup C)) - P(B \cap C) \\ P(A \cap (B \cup C)) &= P(A \cap B) + P(A \cap C) - P(A \cap B \cap C) \end{aligned}$$

The solution follows immediately

$$\begin{aligned} P(A \cup B \cup C) &= P(A) + P(B) + P(C) - P(B \cap C) - (P(A \cap B) + P(A \cap C) - P(A \cap B \cap C)) \\ &= P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C) \quad \square \end{aligned}$$

Question IV.

1.4.1

Suppose we roll eight fair six-sided dice.

(a) What is the probability that all eight dice show a 6?

By the multiplication rule, $P(6,6,6,6,6,6,6,6) = (1/6)^8 = 1/1679616$. This is also one over the total number of possible rolls.

(b) What is the probability that all eight dice show the same number? The first number is irrelevant, so this will be $(1/6)^7 = 1/279936$

(c) What is the probability that the sum of the eight dice is equal to 9?

This is only possible if seven of the dice roll a 1 and 1 rolls a 2, which will be $8/1679616 = 1/209952$.

1.4.6

Suppose we pick two cards at random from an ordinary 52-card deck. What is the probability that the sum of the values of the two cards (where we count jacks, queens, and kings as 10, and count aces as 1) is at least 4?

The only cases where we will not get at least 4 is if we have an ace + ace or ace + 2. *Assuming no replacement*, the probability of getting one of these hands will be the probability of ace on first card then ace on second, plus ace on first then 2 on second, plus 2 on first plus ace on second.

The probability of drawing ace then ace is (4/52)(3/51).

The probability of drawing ace then 2 is (4/52)(4/51).

The probability of drawing 2 then ace is the same as ace then 2 (4/52)(4/51).

So the total probability of drawing a hand with a total of three or less is is the sum of these three probabilities = 0.017. The desired quantity is the probability of drawing a hand worth four or more, which will be 1-0.017 = 0.983.

1.4.11

Consider two urns, labelled urn #1 and urn #2. Suppose urn #1 has 5 red and 7 blue balls. Suppose urn #2 has 6 red and 12 blue balls. Suppose we pick three balls uniformly at random from each of the two urns. What is the probability that all six chosen balls are the same color?

The two cases where all balls are the same color are (1) pick 3 reds from each or (2) pick 3 blues from each. The sum of the probabilities of these two events will give the probability that all six are the same color. The probability of picking 3 reds from the first urn (assuming no replacement) is $\frac{\binom{5}{3}}{\binom{12}{3}}$, and 3 reds from the second is $\frac{\binom{6}{3}}{\binom{18}{3}}$. Similarly, the probabilities of choosing 3 blues from the first is $\frac{\binom{7}{3}}{\binom{12}{3}}$ and 3 blues from the second is $\frac{\binom{12}{3}}{\binom{18}{3}}$. The probability of choosing all six balls to be the same color is then:

$$\left(\frac{\binom{5}{3}}{\binom{12}{3}}\right)\left(\frac{\binom{6}{3}}{\binom{18}{3}}\right) + \left(\frac{\binom{7}{3}}{\binom{12}{3}}\right)\left(\frac{\binom{12}{3}}{\binom{18}{3}}\right) \approx 0.044$$

1.4.12

Suppose we roll a fair six-sided die and flip three fair coins. What is the probability that the total number of heads is equal to the number showing on the die?

There are six possible outcomes of the die roll, each with probability 1/6. For three coin flips, we can have all heads (1/8), one head (3/8) two heads (3/8) or all tails (1/8).

The three outcomes where the desired event occurs are: 1 on die and 1 head, 2 on die and 2 heads, 3 on die and 3 heads. The probability of any of these three occurring is then:

$$\frac{1}{6}\frac{3}{8} + \frac{1}{6}\frac{3}{8} + \frac{1}{6}\frac{1}{8} = \frac{7}{48}$$

Question V.

Long ago we posted grades identified by the last 4 digits of a student's social security number. (It is illegal now!) Assume each of the 10,000 configurations 0000 to 9999 are equally likely.

(a) Use R to estimate the probability that at least two students in a class of 100 share the same 4 digits. Do this by simulating 100,000 samples of size 100, with replacement, from the population $\{0, 1, 2 \ldots, 9999\}$ and then determining the proportion of times that a sample has at least one duplicate. (Note: if x is the sample then `(length(unique(x))!=100)` returns value TRUE, which computes to 1, when there is at least one duplicate. This is a bit simpler than the method shown in the birthday problem example.)

My result from R: 0.39221

(b) Find the actual probability and compare it to your estimate. (Recall the birthday problem discussed in class.)

This is done the same as the birthday problem, except with 10000 possible outcomes instead of 365.

The probability that no two share the same last four digits would be

$$\frac{P_{k,10000}}{10000^k}$$

So the probability that at least 2 share the same last four digits is:

$$P(k) = 1 - \frac{P_{k,10000}}{10000^k}$$

When calculated in R this gives 0.39143

(c) What is the smallest class enrollment for which the probability that at least two students have the same 4 digits is at least 0.50?

Similarly to the birthday problem, we do this computationally: a sample of 119 people gives a probability of .50584 that at least two will have the same last four digits.

Question VI.

1.5.1

Suppose that we roll four fair six-sided dice.

(a) What is the conditional probability that the first die shows 2, conditional on the event that exactly three dice show 2?

The simplest way to find the probability we are after is to find 1 minus the probability that the first die is not a 2, given that three of the dice are showing 2. By inspection, the probability that the first die is not a 2 given that 3 dice show 2 is $1/4$, so the probability that the first die does show 2 is simply $3/4$.

(b) What is the conditional probability that the first die shows 2, conditional on the event that at least three dice show 2?

From the definition of conditional probability we have

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Where A is the event that the first die shows 2 and B is the event that at least three dice show 2.

The probability that at least three dice show 2 is the probability that all four show 2 ($1/6^4$) plus the probability that three show 2 which is $5/6(1/6)^3$, so $P(B) = 21/1296$.

The intersect can be found using the multiplication definition $P(A \cap B) = P(A)P(B|A)$. We have

$$\begin{aligned} P(A) &= \frac{1}{6} \\ P(B|A) &= \frac{5}{6}\frac{1}{6^2} + \frac{1}{6^3} \\ &= \frac{16}{1296} \end{aligned}$$

So our final probability is $P(A|B) = 16/21$.

1.5.8

Suppose the probability of snow is 20%, and the probability of a traffic accident is 10%. Suppose further that the conditional probability of an accident, given that it snows, is 40%. What is the conditional probability that it snows, given that there is an accident?

From the definition of conditional probability we have

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Where A is the event that it snows and B is the event that at there is an accident. This calculation is trivial after knowing $P(A \cap B) = P(A)P(B|A) = .2(.4) = .08$. It follows that $P(A|B) = .08/.1 = .80$.

1.5.10

Consider two urns, labelled urn #1 and urn #2. Suppose, as in Exercise 1.4.11, that urn #1 has 5 red and 7 blue balls, that urn #2 has 6 red and 12 blue balls, and that we pick three balls uniformly at random from each of the two urns. Conditional on the fact that all six chosen balls are the same color, what is the conditional probability that this color is red?

Once again we can define the problem as

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Where A is the event that all six balls are red and B is the event that at all six balls are the same color. As calculated in Exercise 1.4.11, the probability that all six balls are the same color is

$$P(B) = \left(\frac{\binom{5}{3}}{\binom{12}{3}}\right)\left(\frac{\binom{6}{3}}{\binom{18}{3}}\right) + \left(\frac{\binom{7}{3}}{\binom{12}{3}}\right)\left(\frac{\binom{12}{3}}{\binom{18}{3}}\right)$$

And the probability that all balls are red, $P(A)$, is simply the first term:

$$P(A) = \left(\frac{\binom{5}{3}}{\binom{12}{3}}\right)\left(\frac{\binom{6}{3}}{\binom{18}{3}}\right)$$

And the probability $P(B|A)$ is trivially 1. The probability we are after is then

$$\begin{aligned} P(A|B) &= \frac{P(A)P(B|A)}{P(B)} \\ &= \frac{P(A)}{P(B)} \\ &= 0.025 \end{aligned}$$

Question VII.

1.5.7 Suppose a baseball pitcher throws fastballs 80% of the time and curveballs 20% of the time. Suppose a batter hits a home run on 8% of all fastball pitches, and on 5% of all curveball pitches.

(a) What is the probability that this batter will hit a home run on this pitcher's next pitch?

This is calculable by (law of total probability is implied):

$$\begin{aligned} P(\text{Home Run}) &= .8(.08) + .2(.05) \\ &= 0.074 \end{aligned}$$

(b) Suppose we are given (or know or assume) that the batter hits a home run. What is the (conditional) probability that he was thrown a curve ball?

Where A is the event that the batter was thrown a curveball, and B is that the batter hit a home run, we have:

$$\begin{aligned} P(A|B) &= \frac{P(A \cap B)}{P(B)} \\ &= \frac{P(A)P(B|A)}{P(B)} \\ &= .14 \end{aligned}$$

(c) As in (b), suppose the batter does not hit a home run. What is the probability that he was thrown a curve ball? If instead now A is the event that the batter was thrown a curveball, and B is that the batter did not hit a home run, we have:

$$
\begin{aligned}
P(A|B) &= \frac{P(A \cap B)}{P(B)} \\
&= \frac{P(A)P(B|A)}{P(B)} \\
&= .01
\end{aligned}
$$