CSCA67 Tutorial, Week 5

Oct. 19th-Oct. 23rd, 2015

Compiled by *G. Singh Cadieux*

Adapted from
A. Bretscher, *CSCA67 Week 5 Lecture Notes* &
Evans and Rosenthal. *Probability and Statistics: The Science of Uncertainty.* W. H. Freeman and Company, 2010.

1 Review of week 5's lecture

1.1 *The Birthday Problem*

Q: What is the probability that, in a group of n people, at least 2 have the same birthday?

Let us represent the days of the year by the integers 1, 2, ..., 365. Then we choose our sample space S to be {all possible combinations of n birthdays}. That is, we include all possible combinations of n days, with repetition (up to n repetitions of the same day, where all n birthdays fall on the same day).

For example, if $n = 3$, we include:

- all the single days of the year (eg. (1, 1, 1), (2, 2, 2), (3, 3, 3), ...), in the case that all 3 birthdays fall on the same day,
- all combinations of 2 different days of the year (eg. (1, 1, 2), (1, 1, 3), (1, 1, 4), ...), in the case that 2 of the birthdays fall on the same day, and
- all combinations of 3 different days of the year (eg. (1, 2, 3), (1, 2, 4), (1, 2, 5), ...), in the case that all 3 birthdays fall on different days

Suppose that all birthdays are equally likely. Then, by the classical definition of probability,

$$P(\{\text{at least 2 people share a birthday}\}) = \frac{|\{\text{at least } 2\ldots\}|}{|S|}$$

We know from counting principles that

$$\begin{aligned}|S| &= \#\text{ of ways to select the first birthday} \times \#\text{ of ways to select the second birthday} \\ &\quad \times \ldots \times \#\text{ of ways to select the } n\text{th birthday} \\ &= 365 \times 365 \times \ldots \times 365 \\ &= 365^n\end{aligned}$$

and that

$$\begin{aligned}|\{\text{at least } 2\ldots\}| &= \#\text{ of arrangements of } n \text{ birthdays where 2 people share a birthday} \\ &\quad + \#\text{ of arrangements of } n \text{ birthdays where 3 people share a birthday} \\ &\quad + \ldots + \#\text{ of arrangements of } n \text{ birthdays where } n \text{ people share a birthday}\end{aligned}$$

So

$$P(\{\text{at least 2 people share a birthday}\}) = \frac{|\{\text{2 people share a birthday}\}| + \ldots + |\{n \text{ people share a birthday}\}|}{365^n}$$

HOWEVER, this seems very laborious to compute, particularly if n is large.
We can instead use the complement rule to determine that

$$\begin{aligned} P(\{\text{at least 2 people share a birthday}\}) &= 1 - P(\overline{\{\text{at least 2 people share a birthday}\}}) \\ &= 1 - P(\{\text{no shared birthdays}\}) \\ &= 1 - \frac{\#\text{ of ways to select } n \text{ unshared birthdays}}{365^n} \\ &= 1 - \frac{365 \times 364 \times \ldots \times (365 - n + 1)}{365^n} \end{aligned}$$

NOTICE that, if $n > 365$, the above calculation produces

$$\begin{aligned} P(\{\text{at least } 2 \ldots\}) &= 1 - \frac{365 \times (365 - 1) \times \ldots \times (365 - 364) \times (365 - 365) \times \ldots \times (365 - n + 1)}{365^n} \\ &= 1 - \frac{365 \times \ldots \times 0 \times \ldots \times (365 - n + 1)}{365^n} \\ &= 1 - \frac{0}{365^n} \\ &= 1 - 0 = 1 \end{aligned}$$

Why does this make sense?

By the pigeonhole principle, if we have n objects to place in fewer than n pigeonholes, at least 1 pigeonhole will contain multiple objects. In this case, if there are more than 365 birthdays to distribute over 365 days, at least 2 birthdays will fall on the same day. Thus, the probability of at least 2 people sharing a birthday is 1, or absolutely certain.

NOTE that this counting method counts *ordered* n-tuples: for example, where $n = 3$, we consider (1, 1, 2) and (1, 2, 1) to be different combinations of birthdays.
If we were to instead consider unordered n-tuples, we could not use the classical definition of probability, since not all outcomes would be equally likely. For example, where $n = 3$, the unordered combination (1, 1, 2) is more likely than the unordered combination (1, 2, 3), since there are more ways in which the former can occur.

1.2 *Conditional Probability*

Given two events A and B for an experiment, the CONDITIONAL PROBABILITY of A given B, written $P(A|B)$, represents the fraction of time that A occurs once we *know* that B occurs.

$P(E)$ represents the fraction of time that an event E occurs. For example, if $P(E) = 50\%$ and we conduct 10 trials of our experiment, 5 of those trials will result in E.
Recall that $A \cap B$ is the event that both A and B occur.
Thus, the conditional probability

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad \text{with } P(B) > 0$$

represents the ratio between the fraction of time that B occurs and the fraction of time that $A \cap B$ occurs.

Consider the experiment of tossing two fair coins. Let A be the event that the first coin is heads, and let B be the event that exactly two coins are heads.

If we let our sample space $\mathcal{S}$ be $\{HH, HT, TH, TT\}$, then this is an equally-likely sample space and we can compute that

$$P(A) = \frac{|A|}{|\mathcal{S}|} = \frac{|\{HH, HT\}|}{|\{HH, HT, TH, TT\}|} = \frac{2}{4} \qquad P(B) = \frac{|B|}{|\mathcal{S}|} = \frac{|\{HH\}|}{|\{HH, HT, TH, TT\}|} = \frac{1}{4}$$

$$P(A \cap B) = \frac{|A \cap B|}{|\mathcal{S}|} = \frac{|\{HH\}|}{|\{HH, HT, TH, TT\}|} = \frac{1}{4}$$

Using the definition above, we can also compute that

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{|A \cap B|/|\mathcal{S}|}{|B|/|\mathcal{S}|} = \frac{|A \cap B|}{|B|} = \frac{|\{HH\}|}{|\{HH\}|} = 1$$

Notice that $P(A \cap B) = \frac{|A \cap B|}{|\mathcal{S}|}$ is very similar to $P(A|B) = \frac{|A \cap B|}{|B|}$. The difference is that, in the former case, we consider the fraction of time that $A \cap B$ occurs out of *all* possible outcomes, while in the latter case, we consider the fraction of time that $A \cap B$ occurs out of the times that B occurs (ignoring the times that B does *not* occur).
Put another way: $P(B|A)$ answers the question, "if B occurs, what is the probability that A *also* occurs?"

Notice also that, in this example, $P(A|B) > P(A)$. This tells us that B occurring increases the likelihood that A will occur.
Intuitively, this makes sense: if both coins are heads (B occurs), then the first coin *must* be heads (event A occurs with probability 1); but if we do not know whether either coin is heads, it is possible for the first coin to be tails (event A occurs with probability < 1).

However, if C is the event that exactly two coins are tails, then

$$P(A|C) = \frac{|A \cap C|}{|C|} = \frac{|\emptyset|}{|\{TT\}|} = 0$$

Here, $P(A|C) < P(A)$ - that is, C occurring *decreases* the likelihood that A will occur.
Again, intuitively, this makes sense: if both coins are tails (C occurs), then the first coin *cannot* be heads (event A occurs with probability 0); but if we do not know whether either coin is heads, it is possible for the first coin to be heads (event A occurs with probability > 0).

And if D is the event that the second coin is heads, then

$$P(A|D) = \frac{|A \cap D|}{|D|} = \frac{|\{HH\}|}{|\{TH, TT\}|} = \frac{1}{2}$$

Here, $P(A|D) = P(A)$ - that is, D occurring *does not affect* the likelihood that A will occur.

Independence of Events

If $P(E|F) = P(E)$ or $P(F|E) = P(F)$ for two events E and F, where $P(F) > 0$ or $P(E) > 0$ respectively, we say that E and F are independent. Intuitively, this means that the occurrence of E does not affect the probability of F occurring, and vice versa.

Product Rule

We can reformulate the above definition of conditional probability for two events A and B as:

$$P(A \cap B) = P(A|B) \cdot P(B) = P(B|A) \cdot P(A)$$

This allows us to compute the probability of $A \cap B$ when we are given the probability of A and the conditional probability of B given A, or the probability of B and the conditional probability of A given B.

Notice that, if A and B are independent, then

$$P(A \cap B) = P(A|B) \cdot P(B) = P(A) \cdot P(B) \quad \textit{and}$$
$$P(A \cap B) = P(B|A) \cdot P(A) = P(B) \cdot P(A)$$

which is an alternative definition of independence for two events A and B.

Bayes' theorem

We can also reformulate the above definition of conditional probability for two events A and B as:

$$\mathbf{P}(\mathbf{A}|\mathbf{B}) = \frac{P(A \cap B)}{P(B)} = \frac{P(A \cap B)}{P(B)} \cdot \frac{P(A)}{P(A)} = \frac{P(A \cap B)}{P(A)} \cdot \frac{P(A)}{P(B)} = \mathbf{P}(\mathbf{B}|\mathbf{A}) \cdot \frac{\mathbf{P}(\mathbf{A})}{\mathbf{P}(\mathbf{B})} \quad \text{with } P(A),\, P(B) > 0$$

This allows us to compute the conditional probability of B given A when we are given the probability of A, B, and the conditional probability of A given B.

For example, suppose that 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%.

Q: What is the conditional probability that it snows given that there is a traffic accident?

Using Bayes' theorem with A = snow and B = traffic accident, we can calculate that

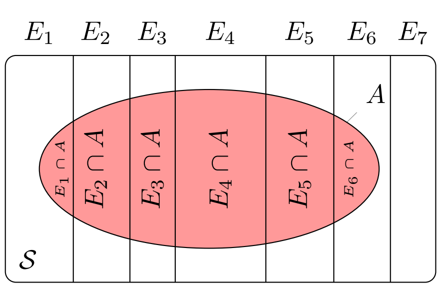
$$\begin{aligned} P(\text{snow}|\text{traffic accident}) &= \frac{P(\text{snow})}{P(\text{traffic accident})} \cdot P(\text{traffic accident}|\text{snow}) \\ &= \frac{0.2 \cdot 0.4}{0.1} \\ &= 0.8 = 80\% \end{aligned}$$

Law of Total (Conditioned) Probability

Let $E_1, E_2, \ldots, E_n$ be events that form a partition of the sample space $\mathcal{S}$, each with positive probability. Let $A \subseteq \mathcal{S}$ be any event. Then

$$\begin{aligned} P(A) &= P(E_1 \cap A) + P(E_2 \cap A) + \ldots + P(E_n \cap A) \\ &= P(E_1) \cdot P(A|E_1) + P(E_2) \cdot P(A|E_2) + \ldots + P(E_n) \cdot P(A|E_n) \qquad \text{[by the product rule]} \\ &= \sum_{i=1}^{n} P(E_i) \cdot P(A|E_i) \end{aligned}$$

Represented as a Venn diagram (with $n = 7$):



For example, suppose that a class contains 60% girls and 40% boys. Suppose that 30% of the girls have long hair and 20% of the boys have long hair.

Q: If a student is chosen randomly from the class, what is the probability that he or she has long hair?

Let $\mathcal{S}$ be the class of students.
If we let E_1 be the event that we choose a girl, and E_2 be the event that we choose a boy, then E_1 and E_2 partition $\mathcal{S}$ (since either E_1 or E_2 must occur, but E_1 and E_2 cannot both occur).

Using the law of total (conditioned) probability, we can calculate that

$$\begin{aligned} P(\text{long hair}) &= P(\text{girl}) \cdot P(\text{long hair}|\text{girl}) + P(\text{boy}) \cdot P(\text{long hair}|\text{boy}) \\ &= 60\% \cdot 30\% + 40\% \cdot 20\% \\ &= 26\% \end{aligned}$$

2 Probability problems

Suppose that we roll four fair six-sided dice.

Q: *What is the conditional probability that the first die shows 2, conditional on the event that exactly three dice show 2?*

Let our (equally-likely) sample space $\mathcal{S}$ be $\{(1,1,1,1),(1,1,1,2),\ldots,(6,6,6,6)\}$.
Let A be the event that exactly three dice show 2, and let B be the event that the first die shows 2.
Then we know from the definition of conditional probability that

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

$A \cap B$ is the event that exactly three dice show 2, including the first die.
So $A = \{(2,2,2,1),(2,2,2,3),\ldots,(6,2,2,2)\}$, and of this, $A \cap B = \{(2,2,2,1),(2,2,2,3),\ldots,(2,6,2,2)\}$.

We calculate that $|A| = (4 \text{ ways to choose the die not showing } 2) \times (5 \text{ ways to choose its value}) = 20$ and $|A \cap B| = (3 \text{ ways to choose the die not showing } 2) \times (5 \text{ ways to choose its value}) = 15$. So

$$P(B|A) = \frac{15}{20} = \frac{3}{4}$$

Suppose we deal five cards from an ordinary 52-card deck.

Q: *What is the conditional probability that all five cards are spades, given that at least four of them are spades?*

Let our (equally-likely) sample space $\mathcal{S}$ be all possible combinations of 5 cards. Let A be the event that all five cards are spades, and let B be the event that at least four of the cards are spades.
Then we can use the definition of conditional probability to calculate

$$\begin{aligned} P(A|B) &= \frac{P(A \cap B)}{P(B)} = \frac{|A \cap B|/|\mathcal{S}|}{|B|/|\mathcal{S}|} = \frac{|A \cap B|}{|B|} \\ &= \frac{|\{5 \text{ spades and at least 4 spades}\}|}{|\{\text{at least 4 spades}\}|} = \frac{|\{5 \text{ spades}\}|}{|\{\text{at least 4 spades}\}|} \\ &= \binom{13 \text{ spades}}{5 \text{ spades}} \Big/ \left(\binom{13 \text{ spades}}{4 \text{ spades}} \cdot \binom{39 \text{ non-spades}}{1 \text{ non-spade}} + \binom{13 \text{ spades}}{5 \text{ spades}} \right) \\ &= \frac{1287}{29172} \approx 0.044 \end{aligned}$$

Suppose that we have one jar with 3 red and 2 blue balls, and a second jar with 4 red and 7 blue balls.

Q: *If we pick one of the jars at random, and then pick one of the balls in that jar at random, what is the probability that*

a) ***the second jar is picked and then a blue ball is picked?***

Let A be the event that the second jar is picked, and let B be the event that a blue ball is picked.
Then we can use the product rule to calculate

$$\begin{aligned} P(A \cap B) &= P(A) \cdot P(B|A) \\ &= \frac{1}{2} \cdot \frac{7 \text{ blue balls}}{11 \text{ total balls}} \\ &= \frac{7}{22} \end{aligned}$$

b) ***a blue ball is picked?***

Let B be the event that a blue ball is picked. Let A_1 be the event that the first jar is picked, and let A_2 be the event that the second jar is picked.
Then we can use the law of total (conditioned) probability to calculate

$$\begin{aligned} P(B) &= P(E_1) \cdot P(A|E_1) + P(E_2) \cdot P(A|E_2) P(A_1) \cdot P(B|A_1) + P(A_2) \cdot P(B|A_2) \\ &= \frac{1}{2} \cdot \frac{2 \text{ blue balls}}{5 \text{ total balls}} + \frac{1}{2} \cdot \frac{7 \text{ blue balls}}{11 \text{ total balls}} \\ &= 0.5\overline{18} \end{aligned}$$

c) ***the second jar was picked, given that we picked a blue ball?***

Again, let A be the event that the second jar is picked, and let B be the event that a blue ball is picked.
Since we know $P(A), P(B), P(B|A)$, we can use Bayes' theorem to calculate

$$\begin{aligned} P(A|B) &= \frac{P(A)}{P(B)} \cdot P(B|A) \\ &= \frac{1/2}{0.5\overline{18}} \cdot \frac{7 \text{ blue balls}}{11 \text{ total balls}} \\ &\approx 0.614 \end{aligned}$$

Suppose that we roll two fair six-sided dice.
Let A be the event that the two dice show the same value.
Let B be the event that the sum of the two dice is equal to 12.
Let C be the event that the first die shows 4.
Let D be the event that the second die shows 4.

Q: *Which events are independent of one another?*

Let us choose $\{(1,1),(1,2),\ldots,(6,6)\}$ as our (equally-likely) sample space $\mathcal{S}$. Then

$$P(A) = \frac{|A|}{|\mathcal{S}|} = \frac{|\{(1,1),(2,2),\ldots,(6,6)\}|}{6^2} = \frac{6}{6^2} = \frac{1}{6} \qquad P(C) = \frac{|C|}{|\mathcal{S}|} = \frac{|\{(4,1),(4,2),\ldots,(4,6)\}|}{6^2} = \frac{6}{6^2} = \frac{1}{6}$$

$$P(B) = \frac{|B|}{|\mathcal{S}|} = \frac{|\{(6,6)\}|}{6^2} = \frac{1}{6^2} \qquad P(D) = \frac{|D|}{|\mathcal{S}|} = \frac{|\{(1,4),(2,4),\ldots,(6,4)\}|}{6^2} = \frac{6}{6^2} = \frac{1}{6}$$

We know that two events E and F are independent if $P(E|F) = P(E)$ or if $P(F|E) = P(F)$, with $P(E), P(F) > 0$. And we know from the definition of conditional probability that $P(E|F) = \frac{P(E \cap F)}{P(F)}$. So

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{|A \cap B|/|\mathcal{S}|}{P(B)} = \frac{|\{(6,6)\}|/6^2}{1/6^2} = \frac{1/6^2}{1/6^2} = 1 \neq P(A) \qquad \Rightarrow A \text{ and } B \text{ are not independent}$$

$$P(A|C) = \frac{P(A \cap C)}{P(C)} = \frac{|A \cap C|/|\mathcal{S}|}{P(C)} = \frac{|\{(4,4)\}|/6^2}{1/6} = \frac{1/6^2}{1/6} = \frac{1}{6} = P(A) \qquad \Rightarrow A \text{ and } C \text{ are independent}$$

$$P(A|D) = \frac{P(A \cap D)}{P(D)} = \frac{|A \cap D|/|\mathcal{S}|}{P(D)} = \frac{|\{(4,4)\}|/6^2}{1/6} = \frac{1/6^2}{1/6} = \frac{1}{6} = P(A) \qquad \Rightarrow A \text{ and } D \text{ are independent}$$

$$P(B|C) = \frac{P(B \cap C)}{P(C)} = \frac{|B \cap C|/|\mathcal{S}|}{P(C)} = \frac{|\emptyset|/6^2}{1/6} = \frac{0}{1/6} = 0 \neq P(B) \qquad \Rightarrow B \text{ and } C \text{ are not independent}$$

$$P(B|D) = \frac{P(B \cap D)}{P(D)} = \frac{|B \cap D|/|\mathcal{S}|}{P(D)} = \frac{|\emptyset|/6^2}{1/6} = \frac{0}{1/6} = 0 \neq P(B) \qquad \Rightarrow B \text{ and } C \text{ are not independent}$$

$$P(C|D) = \frac{P(C \cap D)}{P(D)} = \frac{|C \cap D|/|\mathcal{S}|}{P(D)} = \frac{|\{(4,4)\}|/6^2}{1/6} = \frac{1/6^2}{1/6} = \frac{1}{6} = P(C) \qquad \Rightarrow C \text{ and } D \text{ are independent}$$

Monty Hall Problem

SEE *The Monty Hall Problem* OR *Monty Hall Problem for Dummies - Numberphile.*

3 Additional practice problems

Q: What is the probability of having the same birthday as your mother?

Q: Approximately how many people in the world have the same birthday as their mother?

Q: Approximately how many people in the world have the same birthday as their mother, father, *and* spouse?

Suppose we deal five cards from an ordinary 52-card deck.
Q: What is the conditional probability that the hand contains all four aces, given that the hand contains at least four aces?

Q: What is the conditional probability that the hand contains no pairs, given that it contains no spades?

Suppose we roll a fair six-sided die and then flip a number of fair coins equal to the number showing on the die (eg. if the die shows 4, then we flip 4 coins).
Q: What is the probability that the number of heads equals 3?

Q: Conditional on knowing that the number of heads equals 3, what is the probability that the die showed the number 5?

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.
Q: What is the probability that this batter will hit a home run on this pitcher's next pitch?