9-11 Conditional Probability

Wednesday, September 11, 2019 11:28 AM

9-11 Conditional Probability

9-11: Conditional Probability and Counting
CSCI 3022 Fall 19

Opening **Example**: Suppose we draw a card from a traditional 52-card playing deck. What is the probability that the card is the $A\diamondsuit$? What is the probability that the card is either an A or a $\diamondsuit$?

Opening Example Sol'n

Suppose we draw a card from a traditional 52-card playing deck. What is the probability that the card is the $A\diamondsuit$? What is the probability that the card is either an A or a $\diamondsuit$?

1) $\dfrac{|\{A\diamondsuit\}|}{|\text{whole deck}|} = \dfrac{1}{52}$.

2) $P(A \text{ or } \diamondsuit) = P(A) + P(\diamondsuit) - P(A \cap \diamondsuit)$

$= \dfrac{4}{52} + \dfrac{13}{52} - \dfrac{1}{52} = \dfrac{16}{52} = \dfrac{4}{13}$.

Opening Example Sol'n

Suppose we draw a card from a traditional 52-card playing deck. What is the probability that the card is the $A\diamondsuit$? What is the probability that the card is either an A or a $\diamondsuit$?

1. $P(\{A\diamondsuit\}) = \frac{1}{52}$
2. $P(\{A \cup \diamondsuit\}) = P(\{A\}) + P(\{\diamondsuit\}) - P(\{A \cap \diamondsuit\}) = \frac{4}{52} + \frac{13}{52} - \frac{1}{52} = \frac{16}{52} = \frac{4}{13}$

Announcements and Reminders

- Homework due Friday the 13th.
- Another notebook day Friday, with real probabilities!.
- Last time in lecture: sets! events! probabilities!

Last Time...

A few big takeaways from our first lecture on probability.

- A *sample space* (denoted Ω) of a probabilistic process is the set of all possible outcomes of that process.
- An *event* is any collection (subset) of outcomes from the sample space.
- *Probability* is a function that takes in events and random variables and outputs numbers in $[0, 1]$.
- **Idea:** two or more trials are *independent* if they don't affect each other.

Last Time...

A few big takeaways from our first lecture on probability.

- A *sample space* (denoted Ω) of a probabilistic process is the set of all possible outcomes of that process.
- An *event* is any collection (subset) of outcomes from the sample space.
- *Probability* is a function that takes in events and random variables and outputs numbers in $[0, 1]$.
    1. If A and B are disjoint (mutually exclusive) sets, $P(A \cup B) = P(A) + P(B)$.
    2. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
- **Idea:** two or more trials are *independent* if they don't affect each other.

Formal Probability

Suppose we know $P(\omega)$ for each outcome ω in Ω.

We can compute the probability of an event A which may include one or more outcomes as the sum of all of the probabilities of the outcomes in A:

$$P(A) = \sum_{\omega \in A} P(\omega)$$

Example: Suppose we flip a biased coin with a probability function given by $P(\{H, T\}) = \{p, 1 - p\}$ three times. What is the probability we get two or more tails?

Adding Outcomes

Example: Suppose we flip a biased coin with a probability function given by $P(\{H,T\}) = \{p, 1-p\}$ three times. What is the probability we get two or more tails?

HHH
HHT
⋮

1) what are the outcomes?

$A = \{\{TTT\}, \{TTH\}, \{THT\}, \{HTT\}\}$

2) what are the probs?

$P(A) = P(\{TTT\}) + P(HTT) + P(THT) + P(TTH)$

$\overset{\text{indep.}}{=} P(T)\cdot P(T)\cdot P(T) + \underbrace{P(H)\cdot P(T)\cdot P(T) + P(T)\cdot P(H)\cdot P(T) + \ldots}$

$= (1-p)^3 + p(1-p)(1-p) + (1-p)p(1-p) + \ldots$

$= (1-p)^3 + \underline{\underline{3}}\, p(1-p)^2$

Adding Outcomes

Example: Suppose we flip a biased coin with a probability function given by $P(\{H, T\}) = \{p, 1-p\}$ three times. What is the probability we get two or more tails?

- A is the event that that we see two or more tails. It includes the following elements of Ω: $\{\{TTH\}, \{THT\}, \{HTT\}, \{TTT\}\}$.
- $P(A) = \sum_{\omega \in A} P(\omega) =$ $P(\{TTH\}) + P(\{THT\}) + P(\{HTT\}) + P(\{TTT\})$ because of these outcomes are *disjoint*.
- These probabilities are the products of probabilities of the 3 flips within each, because each flip is *independent* and the probabilities are identical. As a result:

$$
\begin{aligned}
P(A) &= (1-p)\cdot(1-p)\cdot p + (1-p)\cdot p\cdot(1-p) + p\cdot(1-p)\cdot(1-p) \\
&\quad + (1-p)\cdot(1-p)\cdot(1-p) \\
&= 3p(1-p)^2 + (1-p)^3
\end{aligned}
$$

Conditional Probability

Example: If you stop a random person on the street and ask them what month they were born, what is the probability they were born in a long month? (i.e., with 31 days in it)

$\{J, M, M, J, A, O, D\} = \frac{7}{12}$

Example: What is the probability they were born in a month with an r in the name?

Conditional Probability

Example: If you stop a random person on the street and ask them what month they were born, what is the probability they were born in a long month? (i.e., with 31 days in it)
Lazy answer: let L be the event that their birth month has 31 days in it. $\{Jan, Mar, May, Jul, Aug, Oct, Dec\}$ are the elements in L out of 12 months total, so $P(L) = \frac{7}{12}$ if all months are equally likely.

Example: What is the probability they were born in a month with an r in the name?

Conditional Probability

Example: If you stop a random person on the street and ask them what month they were born, what is the probability they were born in a long month? (i.e., with 31 days in it)

Slightly less lazy answer: let L be the event that their birth *day* in a month with 31 days in it. The months in L, now span $7 \cdot 31 = 217$ days out of 365 (.2422) total, so $P(L) = \frac{217}{365}$ if all days are equally likely.

Example: What is the probability they were born in a month with an r in the name?

J, F, M, A, S, O, N, D $= \frac{8}{12}$

Conditional Probability

Example: If you stop a random person on the street and ask them what month they were born, what is the probability they were born in a long month? (i.e., with 31 days in it)

Example: What is the probability they were born in a month with an r in the name?

(Only the lazy answer): Let R be the event that their birth month has an 'r' in the name. $\{Jan, Feb, Mar, Apr, Sep, Oct, Nov, Dec\}$ are the elements in R, so $P(R) = \frac{8}{12}$ if all months are equally likely.

Conditional Probability

Example: Now suppose this person tells you that they were born in a 31 day month. What, then, is the probability that they were born in a month with an 'r' in it?

Conditional Probability

Example: Now suppose this person tells you that they were born in a 31 day month. What, then, is the probability that they were born in a month with an 'r' in it?

Recall $\{Jan, Mar, May, Jul, Aug, Oct, Dec\} \in L$ and $\{Jan, Feb, Mar, Apr, Sep, Oct, Nov, Dec\} \in R$.

New $\Omega = L$.

$$P(R|L) = \frac{P(R \text{ that's also in } L)}{P(L)}$$

$$= \frac{P(R \cap L)}{P(L)}$$

Conditional Probability

Example: Now suppose this person tells you that they were born in a 31 day month. What, then, is the probability that they were born in a month with an 'r' in it?

Recall $\{Jan, Mar, May, Jul, Aug, Oct, Dec\} \in L$ and $\{Jan, Feb, Mar, Apr, Sep, Oct, Nov, Dec\} \in R$.
Our *given* knowledge has reduced the sample space to just those elements in L! Now that $\Omega = L$, we are only interested in the elements in R that are also in L.

$$P(R \textbf{ given } L) = \frac{\#event\,in\,both}{\#events\,in\,L}$$

$$P(R|L) = \frac{P(R \cap L)}{P(L)}$$

$$= \frac{4/12}{7/12}$$

$$= 4/7$$

Conditional Probability

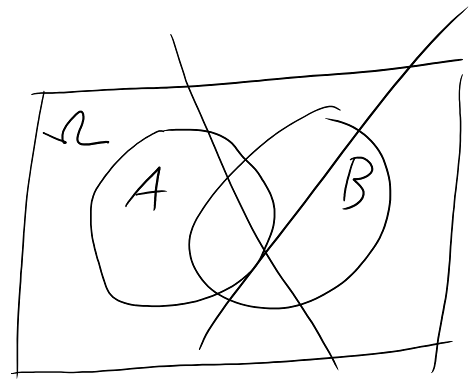
Notation:
We will use the notation $P(A|B)$ to represent the conditional probability of event A *given* that the event B has occurred. B is the "conditioning event."

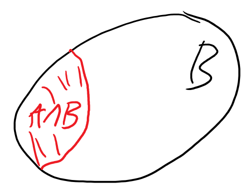
Definition: *Conditional Probability* is defined as:

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

provided that $P(B) > 0$.

Example: A bit string (1s and 0s) of length 4 is generated at random so that each of the 16 possible bit-strings is equally likely. What is the probability that it contains at least 2 consecutive 1s, given that the first bit is a 1?





Conditional Probability

Example: A bit string (1s and 0s) of length 4 is generated at random so that each of the 16 possible bit-strings is equally likely. What is the probability that it contains at least 2 consecutive 1s, given that the first bit is a 1?

$$P(2\times 1\text{'s} \mid S_1 = 1) = \frac{P(\text{both})}{P(S_1=1)} = \frac{P(\#2\text{'s a 1 or } 3\,\&\,4)}{1/2}$$

0000
0001
0010
0100
1000
0011
0101
1001
0110
1010
1100 ✓
0111
1011 ✓
1101 ✓
1110 ✓
1111 ✓

$5/8$

?? ask your computer.

Conditional Probability

Example: A bit string (1s and 0s) of length 4 is generated at random so that each of the 16 possible bit-strings is equally likely. What is the probability that it contains at least 2 consecutive 1s, given that the first bit is a 1?

List all 16 strings is an option... Or consider the event A : that there are consecutive 1's. Maybe it's easier to find A^C! The strings without consecutive 1's that start with a 1 are $\{1010, 1000, 1001\}$. Let C be the event that the first bit is a 1.

P(no consec. 1's | $b_1 = 1$)

$$P(A^C|C) = \frac{P(A^C \cup C)}{P(C)}$$
$$= \frac{3/16}{8/16} = 3/8$$

Conditional probability $P(\cdot|C)$ is a valid probability function, so the complementation property $P(A|C) = 1 - P(A^C|C) = \frac{5}{8}$ holds.

The Multiplication Rule

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

The definition of conditional probability yields the following result:

Multiplication Rule:

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

$p(\text{both}) = P(B|A) \cdot P(A)$

The multiplication rule is particularly useful when conditional probabilities are easier to calculate than intersections.

The Multiplication Rule

The definition of conditional probability yields the following result:
Multiplication Rule:

- $P(A \cup B) = P(A|B)P(B)$

The multiplication rule is particularly useful when conditional probabilities are easier to calculate than intersections.

The Multiplication Rule

The definition of conditional probability yields the following result:
Multiplication Rule:

- $P(A \cup B) = P(A|B)P(B)$
- $P(A \cup B) = P(B|A)P(A)$

The multiplication rule is particularly useful when conditional probabilities are easier to calculate than intersections.

Example: You draw two cards from a standard ~~playing deck~~. What is the probability that they are both red?

$$P(R_1 \cap R_2) = P(R_2 | R_1) \cdot P(R_1) = \boxed{\frac{25}{51} \cdot \frac{26}{52}}$$

$$= P(R_1 | R_2) \cdot P(R_2)$$

Sampling without replacement

Independence, formally

Definition: Two events A and B are said to be *independent* if $P(A|B) = P(A)$.

This definition, combined with the product rule give us three equivalent tests for independence:

1. $P(A|B) = P(A)$ $\Leftrightarrow$
2. $P(B|A) = P(B)$ $\Leftrightarrow$
3. $\boxed{P(A \cap B) = P(A)P(B)}$ $\Leftrightarrow$

$\Leftrightarrow$ A & B are independent.

$P(A) = 1/2$

$P(B) = 1/2$

$P(C) = 1/2$

$$P(A \cap B) = \frac{1}{2} \cdot 1/2$$
$$= P(A|B) \cdot P(B)$$
$$= 1/2 \cdot 1/2$$

$\hookrightarrow P(\text{flip}1 = H \mid \text{flip}2 = H) = 1/2$

$$P(A|C) = \frac{P(\{HH\})}{P(\{HH\} \cup \{TT\})} = 1/2 = P(B|C)$$

$$P(\{HH\}) = 1/4$$
$$\neq P(A) \cdot P(B) \cdot P(C)$$
$$= 1/2 \cdot 1/2 \cdot 1/2$$

$$P(C \mid A \cap B) = 1$$

Independence, in detail!

Why does this matter? Consider the following
Example: Flip a fair coin twice, and define

1. A : "heads on flip 1"
2. B : "heads on flip 2"
3. C : "same outcomes on both flips"

What are $P(A), P(B), P(C), P(A|B), P(A|C), P(B|C)$?
What about $P(A \cap B \cap C?)$
Any *pair* of A, B, C looks independent, since
$P(A) = P(B) = P(C) = P(A|B) = P(A|C) = P(B|C) = 1/2$.
However, $P(A \cap B \cap C?) = P(\{HH\}) = 1/4$ which is not the same as the triple product $P(A)P(B)P(C) = \frac{1}{8}$.
Ultimately, event C is determine by the combination of A and B.

The Law of Total Probability

Example: Suppose I have a couple of bags of marbles. The first bag contains 6 white marbles and 4 black marbles. The second bag contains 3 white marbles and 7 black marbles.

$P(1) = 1/2$

Now suppose I put the two bags in a box. If I close my eyes, grab a bag from the box, and then grab a marble from that bag, what is the probability that it is black?

(W/B for marbles, 1/2 for bags)

$$P(B) = P(B \text{ and } 1) + P(B \text{ and } 2)$$
$$= P(B|1)\cdot P(1) + P(B|2)\cdot P(2)$$
$$= \frac{4}{10} \quad \frac{1}{2} + \frac{7}{10} \quad \frac{1}{2} = \frac{11}{20}$$

The Law of Total Probability

Example: Suppose I have a couple of bags of marbles. The first bag contains 6 white marbles and 4 black marbles. The second bag contains 3 white marbles and 7 black marbles.
Now suppose I put the two bags in a box. If I close my eyes, grab a bag from the box, and then grab a marble from that bag, what is the probability that it is black?

There are two 'ways' we get a black marble: from bag 1 or from bag 2. We just have to add both up!

$$\begin{aligned} P(\mathbf{B}) &= P(\mathbf{B \text{ from } 1}) + P(\mathbf{B \text{ from } 2}) \\ &= P(\mathbf{B} \cap \mathbf{1}) + P(\mathbf{B} \cap \mathbf{2}) \\ &= P(\mathbf{B}|\mathbf{1})P(\mathbf{1}) + P(\mathbf{B}|\mathbf{2})P(\mathbf{2}) \\ &= \frac{4}{10} \cdot \frac{1}{2} + \frac{7}{10} \cdot \frac{1}{2} \\ &= \frac{11}{20} \end{aligned}$$

The Law of Total Probability

Example: As before, the first bag contains 6 white marbles and 4 black marbles. The second bag contains 3 white marbles and 7 black marbles.
But what if bag 1 is made of a much nicer material to touch, so I am twice as likely to randomly select bag 1 from between the 2 bags?

$$P(B) = P(B \text{ AND } 1) + P(B \text{ AND } 2)$$

$$\vdots$$

$$= \frac{4}{10} \cdot \left(\frac{2}{3}\right) + \frac{7}{10}\left(\frac{1}{3}\right)$$

The Law of Total Probability

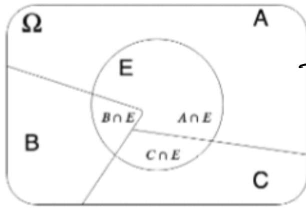
Example: As before, the first bag contains 6 white marbles and 4 black marbles. The second bag contains 3 white marbles and 7 black marbles.
But what if bag 1 is made of a much nicer material to touch, so I am twice as likely to randomly select bag 1 from between the 2 bags?

Same solution!:

$$\begin{aligned} P(\mathbf{B}) &= P(\mathbf{B\ from\ 1}) + P(\mathbf{B\ from\ 2}) \\ &= P(\mathbf{B} \cap \mathbf{1}) + P(\mathbf{B} \cap \mathbf{2}) \\ &= P(\mathbf{B}|\mathbf{1})P(\mathbf{1}) + P(\mathbf{B}|\mathbf{2})P(\mathbf{2}) \\ &= \frac{4}{10} \cdot \frac{2}{3} + \frac{7}{10} \cdot \frac{1}{3} \\ &= \frac{1}{2} \end{aligned}$$

The Law of Total Probability

Definition: A *Partition* of Ω is a set of disjoint events $E_1, E_2, \ldots E_k$ such that $E_1 \cup E_2 \cup \cdots \cup E_k = \Omega$. Given such a partition, any event A can be decomposed into:



$$P(A) = P(A \cap E_1) + P(A \cap E_2) + \cdots + P(A \cap E_k)$$

$$P(A) = P(A|E_1)P(E_1) + P(A|E_2)P(E_2) + \cdots + P(A|E_k)P(E_k)$$

P(black) Given bag

LTP, n00b

"Counting" Examples

Example 1: Suppose that your music app contains 100 songs, 10 of which are by the Beatles. Using the shuffle feature, what is the probability that the first Beatles song heard is the fifth song played?

$$P(\{NNNNNB\})$$

$$= P(\#5 = B \mid 1\text{-}4 \text{ are } N) \cdot P(1\text{-}4 \text{ are } N)$$

$$10/96 \cdot P(\#4 = N \mid \#1\text{-}3 = N) \cdot P(\#1\text{-}3 = N)$$

$$= \frac{87}{97} \qquad \frac{88}{98} \cdot \frac{89}{99} \quad \frac{90}{100}$$

$$= \frac{90}{100} \cdot \frac{89}{99} \quad \frac{88}{98} \quad \cdots \quad \frac{87}{97} \cdot \frac{10}{96}$$

$N_1 \qquad N_2 | N_1 \qquad N_3 | N_2 \text{ AND } N_1$

"Counting" Examples

Example 2: What is the probability of being dealt a flush in poker (five cards)?

"Counting" Examples

Example 3: What is the probability of being dealt all 4 kings in poker (five cards)?