University of Colorado **Boulder**

Department of Computer Science
CSCI 2824: Discrete Structures
Chris Ketelsen

Discrete Probability
Conditional Probability and Independence

Conditional Probability

Conditional Probabilities

Example: Suppose that we flip a fair coin three times

Now, suppose that the first coin flip is Tails

Given this information, what is the probability that an *odd* number of coins come up Tails?

Conditional Probability

Conditional Probabilities

Example: Suppose that we flip a fair coin three times

Now, suppose that the first coin flip is Tails

Given this information, what is the probability that an *odd* number of coins come up Tails?

Of the possible three coin flips, the ones that start with a T are

$$\{THH, THT, TTH, TTT\}$$

Since 2 of the 4 possible outcomes have an odd number of Tails, we say that the probability of an odd number of Tails **given** that the first coin flip was Tails is $\frac{2}{4} = \frac{1}{2}$

Conditional Probability

Let F be the event that the first coin flip is tails and let E be the event that an odd number of tails appear

The probability that E occurs **given** that F occurred is called the **conditional probability** of E given F and denoted $p(E \mid F)$

In general, to compute the conditional probability of E given F we use F as the sample space. For E to then occur it must be the case that E belongs to $E \cap F$

Def: Let E and F be events with $p(F) > 0$. The conditional probability of E given F is defined as

$$p(E \mid F) = \frac{p(E \cap F)}{p(F)}$$

Conditional Probability

Let's check this definition with the coin flip example

$F = \{THH, THT, TTH, TTT\}$

$E = \{HHT, HTH, THH, TTT, \}$

$E \cap F = \{THH, TTT\}$

Recalling that the original sample space S has size 8, we have

$$P(E \cap F) = \frac{2}{8} = \frac{1}{4} \quad \text{and} \quad p(F) = \frac{4}{8} = \frac{1}{2}$$

Thus, by the definition of conditional probability

$$p(E \mid F) = \frac{1/4}{1/2} = \frac{1}{2} \quad \checkmark$$

Conditional Probability

Example: A bit string of length four is generated at random so that each of the 16 bit strings of length four is equally likely. What is the probability that it contains at least two consecutive 0s, given that its first bit is a 0?

Conditional Probability

Example: A bit string of length four is generated at random so that each of the 16 bit strings of length four is equally likely. What is the probability that it contains at least two consecutive 0s, given that its first bit is a 0?

Let F be the event that the first bit is 0 and E be the event that the bit string contains at least two consecutive zeros.

Note that there are $2^4 = 16$ length 4 bit strings

$E \cap F = \{0000, 0001, 0011, 0010, 0100\} \Rightarrow p(E \cap F) = \frac{5}{16}$

There are $2^3 = 8$ bit strings that start with $0 \Rightarrow p(F) = \frac{8}{16} = \frac{1}{2}$

Thus $p(E \mid F) = \frac{5/16}{8/16} = \frac{5}{8}$

Conditional Probability

Example: Suppose you flip a coin two times. If the first flip is Heads, what is the probability that the second flip is Heads as well?

In other words, what is the probability that the the second flip is Heads *given* that the first flip was heads?

Conditional Probability

Example: Suppose you flip a coin two times. If the first flip is Heads, what is the probability that the second flip is Heads as well?

In other words, what is the probability that the the second flip is Heads *given* that the first flip was heads?

Let F be the event that the first flip is Heads and E be the event that the second flip is Heads as well.

$$F = \{HH, HT\} \quad \text{and} \quad E \cap F = \{HH\}$$

There are 4 equally likely outcomes when you flip two coins. Thus

$$p(E \mid F) = \frac{1/4}{2/4} = \frac{1}{2}$$

Is this surprising?

Conditional Probability

No way! Because the result of the first flip has nothing to do with the result of the second flip

Mathematically, we have shown that for this particular example

$$p(E \mid F) = p(E)$$

when this happens we say that events E and F are **Independent.**

Example: Suppose you draw two cards from a deck. Are the events that the first card is Red and that the second card is Red *independent*?

Conditional Probability

Example: Suppose you draw two cards from a deck. Are the events that the first card is Red and that the second card is Red *independent*?

Let F be the event that the first card is Red and E be the event that the second card is Red

Question: Suppose you draw two cards and know nothing about the first card. What is the probability that the second card is Red?

$$p(E) = \frac{26}{52} = \frac{1}{2}$$

Conditional Probability

Example: Suppose you draw two cards from a deck. Are the events that the first card is Red and that the second card is Red *independent*?

Let F be the event that the first card is Red and E be the event that the second card is Red

OK, now suppose that you know that the first card is Red, what is the probability that the second card is also Red?

$$p(E \mid F) = \frac{25}{51} \neq \frac{1}{2} = p(E)$$

Thus (as you probably already intuitively knew) the two events are **NOT** independent

Conditional Probability

When two events are independent, something nice happens

Assume that E and F are independent, then

$$p(E \mid F) = \frac{p(E \cap F)}{p(F)} = p(E)$$

If we solve the second equality for $p(E \cap F)$ we find that

$$p(E \cap F) = p(E)p(F)$$

So if two events are independent, we can compute the probability of their intersection (i.e. that they both occur) by simply multiplying their individual probabilities!

Conditional Probability

Example: Suppose you roll a fair die and flip a coin. What is the probability that you roll a 4 and the coin lands Tails?

Conditional Probability

Example: Suppose you roll a fair die and flip a coin. What is the probability that you roll a 4 and the coin lands Tails?

Clearly the two events are independent, so we have

$$p(4 \text{ AND } T) = p(4) \cdot p(T) = \frac{1}{6} \cdot \frac{1}{2} = \frac{1}{12}$$

Even in the case that the two events are dependent, we can still multiply probabilities, we just have to be careful ...

Conditional Probability

Example: Suppose you draw two cards from a deck. What is the probability that the 1^{st} card is a club and the 2^{nd} card is a heart?

Conditional Probability

Example: Suppose you draw two cards from a deck. What is the probability that the 1^{st} card is a club and the 2^{nd} card is a heart?

Note: Not independent b/c more likely to draw a heart after one of the clubs is removed

Conditional Probability

Example: Suppose you draw two cards from a deck. What is the probability that the 1^{st} card is a club and the 2^{nd} card is a heart?

Note: Not independent b/c more likely to draw a heart after one of the clubs is removed

Claim: $p(1^{\text{st}} - \clubsuit \cap 2^{\text{nd}} - \heartsuit) = \dfrac{13}{52} \cdot \dfrac{13}{51} = \dfrac{13}{204}$

Conditional Probability

Example: Suppose you draw two cards from a deck. What is the probability that the 1^{st} card is a club and the 2^{nd} card is a heart?

Note: Not independent b/c more likely to draw a heart after one of the clubs is removed

Claim: $p(1^{\text{st}} - \clubsuit \cap 2^{\text{nd}} - \heartsuit) = \dfrac{13}{52} \cdot \dfrac{13}{51} = \dfrac{13}{204}$

Proof: From definition of conditional probability

$$p(2^{\text{nd}} - \heartsuit \mid 1^{\text{st}} - \clubsuit) = \frac{p(1^{\text{st}} - \clubsuit \cap 2^{\text{nd}} - \heartsuit)}{p(1^{\text{st}} - \clubsuit)}$$

Conditional Probability

Example: Suppose you draw two cards from a deck. What is the probability that the 1^{st} card is a club and the 2^{nd} card is a heart?

Note: Not independent b/c more likely to draw a heart after one of the clubs is removed

Claim: $p(1^{\text{st}} - \clubsuit \cap 2^{\text{nd}} - \heartsuit) = \dfrac{13}{52} \cdot \dfrac{13}{51} = \dfrac{13}{204}$

Proof: From definition of conditional probability

$$p(1^{\text{st}} - \clubsuit \cap 2^{\text{nd}} - \heartsuit) = p(1^{\text{st}} - \clubsuit)\, p(2^{\text{nd}} - \heartsuit \mid 1^{\text{st}} - \clubsuit)$$

Conditional Probability

Example: Suppose you draw two cards from a deck. What is the probability that the 1^{st} card is a club and the 2^{nd} card is a heart?

Note: Not independent b/c more likely to draw a heart after one of the clubs is removed

Claim: $p(1^{\text{st}} - \clubsuit \cap 2^{\text{nd}} - \heartsuit) = \frac{13}{52} \cdot \frac{13}{51} = \frac{13}{204}$

Proof: From definition of conditional probability

$$\begin{aligned} p(1^{\text{st}} - \clubsuit \cap 2^{\text{nd}} - \heartsuit) &= p(1^{\text{st}} - \clubsuit)\, p(2^{\text{nd}} - \heartsuit \mid 1^{\text{st}} - \clubsuit) \\ &= \frac{13}{52} \cdot \frac{13}{51} \end{aligned}$$

Conditional Probability

Note that in the previous problem we did some algebra and found that

$$p(1^{\text{st}} - \clubsuit \cap 2^{\text{nd}} - \heartsuit) = p(1^{\text{st}} - \clubsuit)\, p(2^{\text{nd}} - \heartsuit \mid 1^{\text{st}} - \clubsuit)$$

This type of rearrangement is so useful that we give it a name:

Product Rule: Let E and F be two events with $p(E) \neq 0$ and $p(F) \neq 0$, then

$$p(E \cap F) = p(E \mid F)p(F) = p(F \mid E)p(E)$$

Proof:

Conditional Probability

The Birthday Problem

What is the minimum number of people who need to be in a room so that the probability that at least two of them have the same birthday is greater than $\frac{1}{2}$?

Conditional Probability

The Birthday Problem

What is the minimum number of people who need to be in a room so that the probability that at least two of them have the same birthday is greater than $\frac{1}{2}$?

First we make some assumptions:

1. Birthdays of any two people are independent
2. Being born on any day of the year is equally likely
3. There are 366 days in the year

Conditional Probability

The Birthday Problem

What is the minimum number of people who need to be in a room so that the probability that at least two of them have the same birthday is greater than $\frac{1}{2}$?

First we make some assumptions:

1. Birthdays of any two people are independent
2. Being born on any day of the year is equally likely
3. There are 366 days in the year

Restatement: The probability that two people have the same birthday is 1 minus the probability that no two people have the same birthday

Conditional Probability

Now let's compute the probability that with n people in the room no two people have the same birthday

Think of asking people their birthdays 1 at a time

$1.$ When the first person enters the room, the probability is ___ that they do not have the same birthday as anyone else

$2.$ When the second person enters the room, the probability is _____ that they do not have the same birthday as the first person

$3.$ When the third person enters the room, the probability is _____ that they do not have the same birthday as the first two people

$n.$ When the n^{th} person enters the room, the probability is _____ that they do not have the same birthday as the first two people

Conditional Probability

Since birthdays are independent, the probability that no two people have the same birthday is

$$p_n = 1 \cdot \frac{365}{366} \cdot \frac{364}{366} \cdots \frac{367-n}{366}$$

And so the probability that two of the n people have the same birthday is

$$1 - p_n = 1 - 1 \cdot \frac{365}{366} \cdot \frac{364}{366} \cdots \frac{367-n}{366}$$

We want to know what is the first value of n such that $1 - p_n \geq \frac{1}{2}$

We could find n with Calculus, but let's just plug in some numbers

Conditional Probability

When $n = 3$ the probability is $1 - p_n \approx 0.027$

When $n = 10$ the probability is $1 - p_n \approx 0.117$

When $n = 22$ the probability is $1 - p_n \approx 0.475$

When $n = 23$ the probability is $1 - p_n \approx 0.506$

When $n = 50$ the probability is $1 - p_n \approx 0.970$

When $n = 100$ the probability is $1 - p_n \approx 0.9999996$

Conditional Probability

When $n = 3$ the probability is $1 - p_n \approx 0.027$

When $n = 10$ the probability is $1 - p_n \approx 0.117$

When $n = 22$ the probability is $1 - p_n \approx 0.475$

When $n = 23$ the probability is $1 - p_n \approx 0.506$

When $n = 50$ the probability is $1 - p_n \approx 0.970$

When $n = 100$ the probability is $1 - p_n \approx 0.9999996$