Announcements

PS1 due today
  ◦ Pain poll
  ◦ Last day to hand in + solutions released next Friday 7/13

PS2 out today! Due next Friday 7/13
Goals for today

- Properties of probability
- Independence
- Mutual exclusivity, independence, and DeMorgan’s
- Conditional Independence
Summary from last time

**CRACK**

**BOOM**

WHOA! WE SHOULD GET INSIDE!

IT'S OKAY! LIGHTNING ONLY KILLS ABOUT 45 AMERICANS A YEAR, SO THE CHANCES OF DYING ARE ONLY ONE IN 7,000,000. LET'S GO ON!

THE ANNUAL DEATH RATE AMONG PEOPLE WHO KNOW THAT STATISTIC IS ONE IN SIX.

xkcd by Randall Munroe
Summary (conditional probability)

\[
P(F \mid E) = \frac{P(EF)}{P(E)}
\]

**Definition of Conditional Probability**

\[
P(F \mid E) = \frac{P(E \mid F)P(F)}{P(E)} = \frac{P(E \mid F)P(F)}{P(E \mid F)P(F) + P(E \mid F^c)P(F^c)} = \sum \frac{P(E \mid F_j)P(F_j)}{P(F_j)P(E \mid F_j)}
\]

**Chain Rule**

\[P(EF) = P(E)P(F \mid E) = P(F)P(E \mid F)\]

**Law of Total Probability**

\[P(E) = P(F)P(E \mid F) + P(F^c)P(E \mid F^c) = \sum \ P(F_j)P(E \mid F_j)\]

**Bayes’ Theorem**
Bayesian Interpretations

\[
P(H|E) = \frac{P(E|H)P(H)}{P(E)}
\]

Prior: Probability of H before you observe E
Likelihood: Probability of E given that H holds
Posterior: Probability of H after you observe E
DNA Paternity testing

Child is born with (A, a) gene pair (event $B_{A,a}$)

- Mother has (A, A) gene pair
- Two possible fathers: $M_1$: (a, a)  $M_2$: (a, A)
- $P(M_1) = p$  $P(M_2) = 1 - p$

What is $P(M_1 | B_{A,a})$?

Solution:

\[
P(M_1 | B_{A,a}) = \frac{P(M_1 B_{A,a})}{P(B_{A,a})}
\]

\[
= \frac{P(B_{A,a} | M_1) P(M_1)}{P(B_{A,a} | M_1) P(M_1) + P(B_{A,a} | M_2) P(M_2)}
\]

\[
= \frac{1 (p)}{1 (p) + (1/2) (1 - p)} = \frac{2p}{1+p} > p
\]

$M_1$ more likely to be father than he was before, since $P(M_1 | B_{A,a}) > P(M_1)$
Properties of probability

For an event A:

\[ P(A) = 1 - P(A^c) \]  \hspace{1cm} \text{(Complement)}

For any events A and B:

\[ P( A \cap B ) = P( B \cap A ) \]  \hspace{1cm} \text{(Commutativity)}

\[ P( A \cap B^c ) = P( A ) - P( A \cap B ) \]  \hspace{1cm} \text{(Intersection)}

\[ P( A ) = P( A \cap B ) + P( A \cap B^c ) \]  \hspace{1cm} \text{(Law of Total Probability)}

\[ = P(A \mid B) P(B) + P(A \mid B^c) P(B^c) \]

\[ P(A \cap B) = P(A) P(A \mid B) \]  \hspace{1cm} \text{(Chain Rule)}

\[ P( A \cap B ) \geq P( A ) + P( B ) - 1 \]  \hspace{1cm} \text{(Bonferroni)}
P(•|G) is also a probability space!

Formally, P(•|G) satisfies 3 axioms of probability. For example:

Axiom 1 of probability 0 ≤ P(E|G) ≤ 1

Corollary 1 (complement) P(E^c|G) = 1 − P(E|G)

More generally:

Chain Rule: \[ P(EF|G) = P(F|G) \cdot P(E|FG) \]

Law of Total Probability: \[ P(E|G) = P(F|G) \cdot P(E|FG) + P(F^c|G) \cdot P(E|F^cG) \]

Bayes Theorem: \[ P(F|EG) = \frac{P(E|FG) \cdot P(F|G)}{P(E|G)} \]
And now...
Two Dice

Roll two 6-sided dice, yielding values $D_1$ and $D_2$.

Let E be event: $D_1 = 1$
Let F be event: $D_2 = 1$

1. What is $P(E)$, $P(F)$, and $P(EF)$?

$$P(E) = 1/6, \quad P(F) = 1/6, \quad P(EF) = 1/36$$

$P(EF) = P(E) \cdot P(F) \Rightarrow E$ and $F$ independent

2. Let $G$ be event: $D_1 + D_2 = 5$. What is $P(E)$, $P(G)$ and $P(EG)$?

$G = \{(1,4), (2,3), (3,2), (4,1)\}$

$$P(E) = 1/6, \quad P(G) = 4/36 = 1/9, \quad P(EG) = 1/36$$

$P(EG) \neq P(E) \cdot P(G) \Rightarrow E$ and $G$ dependent
Independence

Two events $E$ and $F$ are defined as independent if:

$$P(EF) = P(E) P(F)$$

Or, equivalently:

$$P(E | F) = P(E)$$

(otherwise $E$ and $F$ are called dependent events).

Three events $E$, $F$, and $G$ are independent if:

$$P(EFG) = P(E) P(F) P(G), \text{ and}$$
$$P(EF) = P(E) P(F), \text{ and}$$
$$P(EG) = P(E) P(G), \text{ and}$$
$$P(FG) = P(F) P(G)$$

Pairwise independence of 3 variables is not enough to prove independence!
Independence?

A \cup B = P(A) \cup P(B)
When $E$ and $F$ are independent...

Given independent events $E$ and $F$, prove:

$$P(E | F) = P(E | F^c)$$

Proof:

$$P(E F^c) = P(E) - P(EF)$$ \hspace{1cm} \text{Intersection}

$$= P(E) - P(E) P(F)$$ \hspace{1cm} \text{Independence}

$$= P(E) [1 - P(F)]$$ \hspace{1cm} \text{Factoring}

$$= P(E) P(F^c)$$ \hspace{1cm} \text{Complement}

So, $E$ and $F^c$ independent, implying that:

$$P(E | F^c) = P(E) = P(E | F)$$

Intuitively, if $E$ and $F$ are independent, knowing whether $F$ holds gives us no information about $E$. 
Generalized Independence

Events $E_1, E_2, ..., E_n$ are independent if for every subset $E_1', E_2', ..., E_r'$ (where $r \leq n$) it holds that:

$$P(E_1'E_2'...E_r') = P(E_1') P(E_2') P(E_3')... P(E_r')$$

Example:
Outcomes of $n$ separate flips of a coin are all independent of one another. Each flip in this case is a **trial** of the experiment.
Two!!!! Dice!!!!

Roll two 6-sided dice, yielding values $D_1$ and $D_2$.

Let $E$ be event: $D_1 = 1$

Let $F$ be event: $D_2 = 6$

Let $G$ be event: $D_1 + D_2 = 7$

1. Are $E$ and $F$ independent?
   
   $P(E) = 1/6$, $P(F) = 1/6$, $P(EF) = 1/36$  
   
   Yes!

2. Are $E$ and $G$ independent?
   
   $P(E) = 1/6$, $P(G) = 1/6$, $P(EG) = 1/36$  
   
   Yes!

3. Are $F$ and $G$ independent?
   
   $P(F) = 1/6$, $P(G) = 1/6$, $P(FG) = 1/36$  
   
   Yes!

4. Are $E$, $F$, and $G$ independent?
   
   $P(EFG) = 1/36 \neq 1/216 = (1/6)(1/6)(1/6)$  
   
   $EFG = \{ (1,6) \}$  
   
   No!
Generating Random Bits

A computer produces a series of random bits.
• with probability $p$ of producing a 1.
• Each bit generated is an independent trial.
• $E =$ first $n$ bits are 1’s, followed by a single 0

What is $P(E)$?

**Solution:**

Independent trials

$$P(\text{first } n \text{ 1’s}) = P(1^{\text{st}} \text{ bit}=1) \cdot P(2^{\text{nd}} \text{ bit}=1) \cdots P(n^{\text{th}} \text{ bit}=1)$$

$$= p^n$$

$$P(n+1 \text{ bit}=0) = (1 - p)$$

$$P(E) = P(\text{first } n \text{ 1’s}) \cdot P(n+1 \text{ bit}=0) = p^n (1 - p)$$
(biased) Coin Flips

Suppose we flip a coin $n$ times.

- A coin comes up heads with probability $p$.
- Each coin flip is an independent trial.

1. $P(n$ heads on $n$ coin flips$) = p^n$
2. $P(n$ tails on $n$ coin flips$) = (1 - p)^n$
3. $P(first$ $k$ heads, then $n - k$ tails$) = p^k(1 - p)^{n-k}$
4. $P(exactly$ $k$ heads on $n$ coin flips$) = {n \choose k} p^k(1 - p)^{n-k}$
Break

Attendance: tinyurl.com/cs109summer2018
Probability in a nutshell

Our goal today:

- Level 6
- Level 15
- Level 30
Hash Tables

$m$ strings are hashed (equally randomly) into a hash table with $n$ buckets.

• Each string hashed is an independent trial.
• Let event $E =$ at least one string hashed into first bucket.

What is $P(E)$?

Solution:
Define: $F_i =$ string $i$ not hashed into first bucket (where $1 \leq i \leq m$)
$F_1F_2\ldots F_m = $ no strings hashed into first bucket

WTF: $P(E) = 1 - P(F_1F_2\ldots F_m)$

$P(F_i) = 1 - 1/n = (n - 1)/n$ (for all $1 \leq i \leq m$)

$P(F_1F_2\ldots F_m) = P(F_1) P(F_2) \ldots P(F_m) = ((n - 1)/n)^m$

$P(E) = 1 - P(F_1F_2\ldots F_m) = 1 - ((n - 1)/n)^m$

Similar to $\geq 1$ of $m$ people having same birthday as you.
More Hash Table Fun

$m$ strings are hashed (unequally) into a hash table with $n$ buckets

- Each string hashed is an independent trial, with probability $p_i$ of getting hashed to bucket $i$
- $E = \text{At least 1 of}$ buckets 1 to $k$ has $\geq 1$ string hashed to it

Solution:

Define: $F_i = \text{at least one string hashed into } i\text{-th bucket (where } 1 \leq i \leq n)$

WTF: $P(E) = P(F_1 \cup F_2 \cup ... \cup F_k) = 1 - P((F_1 \cup F_2 \cup ... \cup F_k)^c)$

$= 1 - P(F_1^c F_2^c ... F_k^c)$ \hspace{1cm} (DeMorgan’s Law)

$P(F_1^c F_2^c ... F_k^c) = P(\text{no strings hashed to buckets 1 to } k)$

$= (1 - p_1 - p_2 - ... - p_k)^m$

$P(E) = 1 - (1 - p_1 - p_2 - ... - p_k)^m$
No seriously, it’s more Hash Table Fun

$m$ strings are hashed (unequally) into a hash table with $n$ buckets

- Each string hashed is an independent trial, with probability $p_i$ of getting hashed to bucket $i$
- $E = \text{Each of}$ buckets 1 to $k$ has $\geq 1$ string hashed to it

Solution:

Define: $F_i = \text{at least one string hashed into } i\text{-th bucket}$

$$P(E) = P(F_1 F_2 \ldots F_k) = 1 - P((F_1 F_2 \ldots F_k)^c)$$

$$= 1 - P(F_1^c \cup F_2^c \cup \ldots \cup F_k^c)$$  \hspace{1cm} \text{(DeMorgan’s Law)}

$$= 1 - P \left( \bigcup_{i=1}^{k} F_i^c \right) = 1 - \sum_{r=1}^{k} (-1)^{(r+1)} \sum_{i_1 < \ldots < i_r} P(F_{i_1}^c F_{i_2}^c \ldots F_{i_r}^c)$$

Where $P(F_{i_1}^c F_{i_2}^c \ldots F_{i_r}^c) = (1 - p_{i_1} - p_{i_2} \ldots - p_{i_r})^m$  \hspace{1cm} \text{General Inclusion-Exclusion Principle of Probability}
Network reliability

Consider the following parallel network:

- $n$ independent routers, each with probability $p_i$ of functioning (where $1 \leq i \leq n$)
- $E =$ functional path from $A$ to $B$ exists.

What is $P(E)$?

Solution:

\[
P(E) = 1 - P(\text{all routers fail})
= 1 - (1 - p_1)(1 - p_2)\ldots(1 - p_n)
= 1 - \prod_{i=1}^{n} (1 - p_i)
\]
Two events $E$ and $F$ are defined as independent if:

$$P(EF) = P(E) P(F)$$

If $E$ and $F$ are independent, does that tell us whether the following is true?

$$P(EF \mid G) = P(E \mid G) P(F \mid G),$$

where $G$ is an arbitrary event?

In general, No!
Conditional Independence

Two events $E$ and $F$ are defined as conditionally independent given $G$ if:

$$P(EF \mid G) = P(E \mid G) \cdot P(F \mid G)$$

Or, equivalently:

$$P(E \mid FG) = P(E \mid G)$$

(otherwise $E$ and $F$ are called conditionally dependent given $G$).

**Warning**: After conditioning on additional information,

1. Independent events can become conditionally dependent.
2. Dependent events can become conditionally independent.
Roll two 6-sided dice, yielding values $D_1$ and $D_2$.

Let $E$ be event: $D_1 = 1$
Let $F$ be event: $D_2 = 6$
Let $G$ be event: $D_1 + D_2 = 7$

1. Are $E$ and $F$ independent? \text{Yes!}

\[ P(E) = 1/6, \quad P(F) = 1/6, \quad P(EF) = 1/36 \]

2. Are $E$ and $F$ independent given $G$? \text{No!}

\[ P(E|G) = 1/6, \quad P(F|G) = 1/6, \quad P(EF|G) = 1/6 \]

\[ P(EF|G) \neq P(E|G) \cdot P(F|G) \implies E|G \text{ and } F|G \text{ dependent} \]

Independent events can become conditionally dependent.
Faculty Night

In a dorm with 100 students:

- 30 students have straight A’s  \( P(A) = 0.30 \)
- 20 students are in CS  \( P(CS) = 0.20 \)
- 6 students are in CS with straight A’s  \( P(A,CS) = 0.06 \)

At faculty night F, only CS students and straight A students show up (44 total).

1. Are A and CS independent?  \( P(A|CS) = \frac{6}{20} = 0.30 = P(A) \)  Yes!

2. Are A and CS independent given F? \( P(A|CS,F) = \frac{30}{44} \) \( \neq \) \( P(A|F) \)  No!

Which is correct?

- Being a CS major lowers your probability of getting A’s
- Knowing you went to faculty night because you are a CS major makes it less likely that you went because you have straight A’s
Two events E and F are defined as independent if:

\[ P(\text{EF}) = P(E) \cdot P(F) \]

Or, equivalently:

\[ P(E|F) = P(E) \]

If E and F are independent, then E and Fc are also independent, and

\[ P(E|F) = P(E) = P(E|Fc) \]

Two events E and F are defined as conditionally independent given G if:

\[ P(\text{EF} | G) = P(E | G) \cdot P(F | G) \]

Or, equivalently:

\[ P(E|FG) = P(E | G) \]

Important notes:

1. Independence does not imply causality; it’s just math.
2. Conditioning on an event can break independence, or it can make dependent variables conditionally independent.
Summary

DeMorgan’s

\[ P(E \cup F) \] \quad \text{Mutually Exclusive?} \quad \text{Just Add!}

\[ P(E \cap F) \] \quad \text{Independent?} \quad \text{Just Multiply!}

\[ P(E) \quad \text{Inclusion Exclusion} \]

\[ P(F) \quad \text{Chain Rule} \]