
Will Monroe
July 5, 2017
with materials by
Mehran Sahami
and Chris Piech
image: Therightclicks
Independence

Announcements: Problem Set 1 due!

Will Monroe
CS 109

Problem Set #1
June 28, 2017

Problem Set #1

Due: 12:30pm on Wednesday, July 5th

With problems by Mehran Sahami and Chris Piech

For each problem, briefly explain/justify how you obtained your answer. Brief explanations of your answer are necessary to get full credit for a problem even if you have the correct numerical answer. The explanations help us determine your understanding of the problem whether or not you got the correct answer. Moreover, in the event of an incorrect answer, we can still try to give you partial credit based on the explanation you provide. It is fine for your answers to include summations, products, factorials, exponentials, or combinations; you don't need to calculate those all out to get a single numeric answer.

Note: all assignment submissions will be made online through Gradescope. You can find information on signing up to submit assignments though Gradescope on the class webpage. If you handwrite your solutions, you are responsible for making sure that you can produce **clearly legible** scans of them for submission. You may use any word processing software you like for writing up your solutions. On the CS109 webpage we provide a template file and tutorial for the LaTeX system, if you'd like to use it.

This problem set includes one question where we ask you to write some code. You'll need to include a printout of your code in PDF or image form in your Gradescope submission. Double-check that indentation is preserved and the code isn't cut off (at the end of the line or at the end of the page). For LaTeX, we recommend the `minted` package (`https://www.sharelatex.com/learn/Code_Highlighting_with_minted`) with the `breaklines` option.

1. Introduce yourself! Fill out this Google form to tell me a bit about you:

   `https://goo.gl/forms/DuJ8v0UMpsTKDDlB2`

   (No need to copy the answers into your Gradescope submission; you can select an arbitrary page or write "done" so there is something to select.)

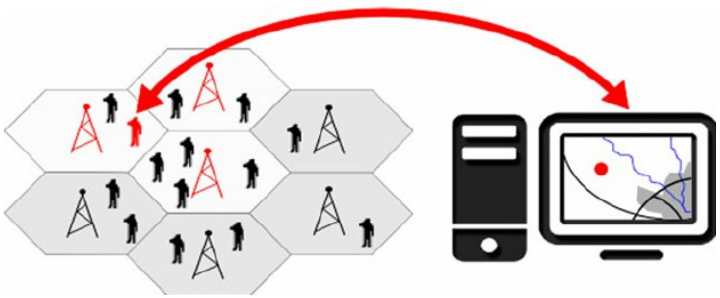
2. 10 computers are brought in for servicing (and machines are serviced one at a time). Of the 10 computers, 3 are PCs, 4 are Macs, 2 are Linux machines, and 1 is an Amiga. Assume that all computers of the same type are indistinguishable (i.e., all the PCs are indistinguishable, all the Macs are the indistinguishable, etc.).
   a. In how many distinguishable ways can the computers be ordered for servicing?
   b. In how many distinguishable ways can the computers be ordered if the first 5 machines serviced must include all 4 Macs?
   c. In how many distinguishable ways can the computers be ordered if 1 PC must be in the first three and 2 PCs must be in the last three computers serviced?

gradescope

Solutions to be posted next Wednesday

(no submissions allowed >1 week late, even if you're willing to take lots of late penalties!)

Announcements: Problem Set 2 out



(Cell phone location sensing)

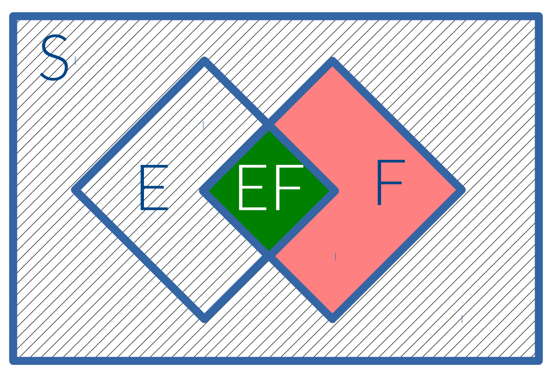
Due next Wednesday, 7/12, at 12:30pm (before class).

13 problems
1 coding problem (this time with a bit of data to work with!)

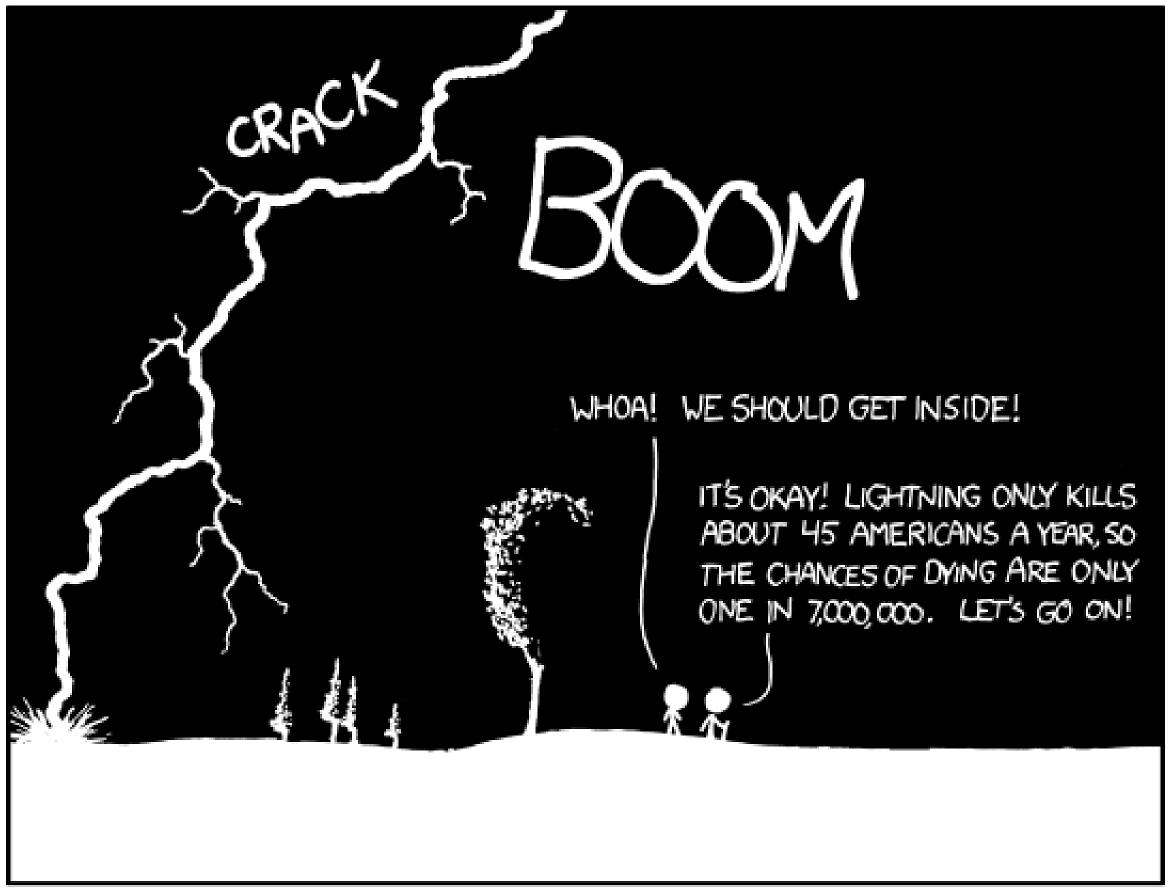
Review: Conditional probability

The conditional probability P(E | F) is the probability that E happens, **given** that F has happened. F is the new sample space.

$$P(E|F) = \frac{P(EF)}{P(F)}$$



Conditional probability: A cautionary tale



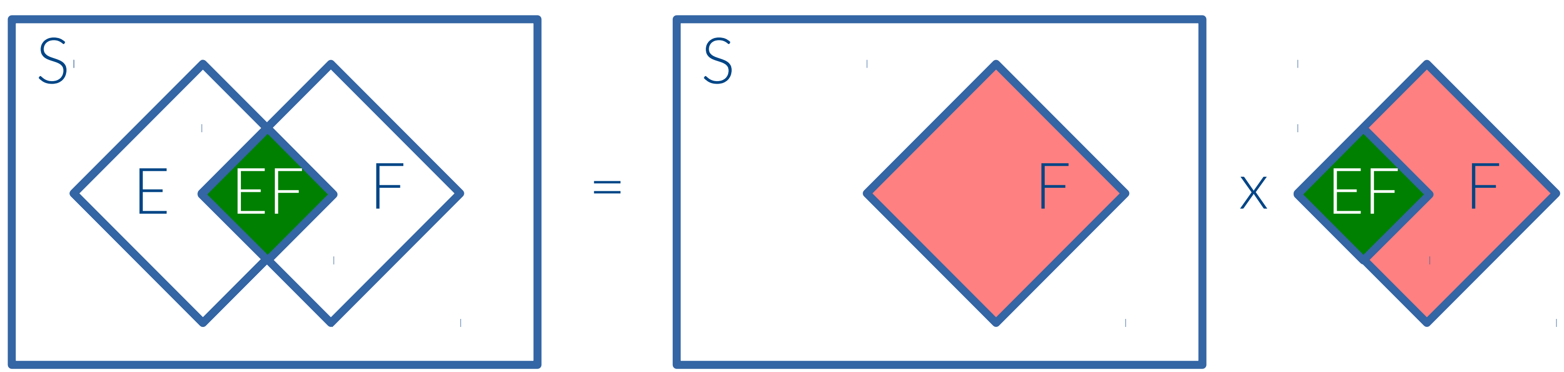
THE ANNUAL DEATH RATE AMONG PEOPLE WHO KNOW THAT STATISTIC IS ONE IN SIX.

xkcd by Randall Munroe

Review: Chain rule of probability

The probability of **all** events happening
is the probability of **the first** happening
times the prob. of **the second given the first**
times the prob. of **the third given the first two**
...etc.

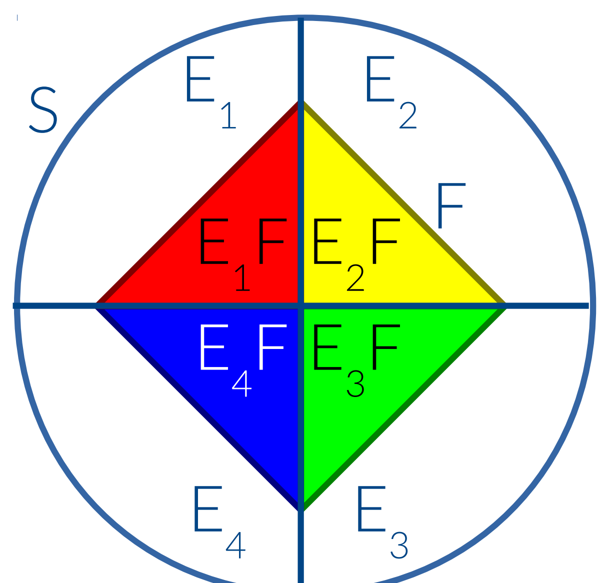
$$P(EFG\ldots)=P(E)P(F|E)P(G|EF)\ldots$$



Review: Law of total probability

You can compute an overall probability by summing over **mutually exclusive** and **exhaustive** sub-cases.

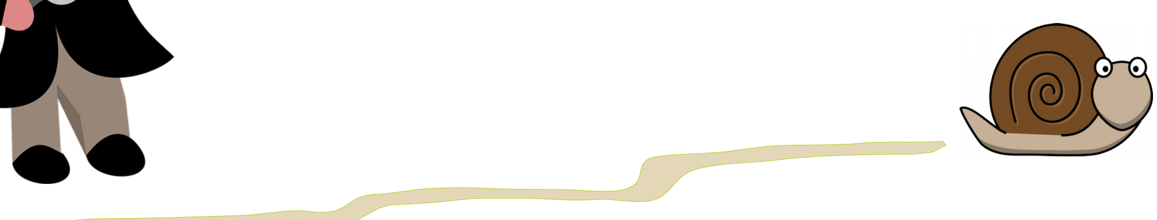
$$P(F) = \sum_i P(E_i F)$$
$$= \sum_i P(E_i) P(F|E_i)$$



Review: Bayes' theorem

You can "flip" a conditional probability if you multiply by the probability of the **hypothesis** and divide by the probability of the **observation**.

$$P(E|F) = \frac{P(F|E)\,P(E)}{P(F)}$$



Independence

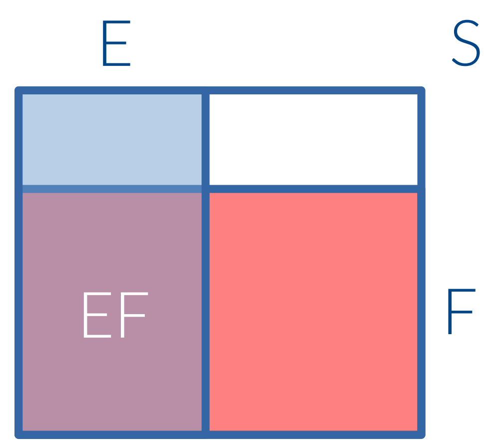
Two events are **independent** if you can **multiply** their probabilities to get the probability of **both** happening.

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

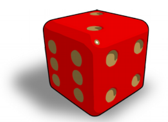
$$\Leftrightarrow$$

$$E \perp F$$

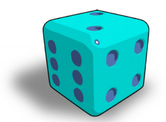
("independent of")



Rolling two dice



D_1



D_2

$$P(\overset{E}{D_1=1}, \overset{F}{D_2=1}) = \frac{1}{36}$$

P(EF) = 1/6 · 1/6

P(E) = 1/6

P(F) = 1/6

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),

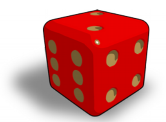
(3, 1), (3, 2), (3, 3), (3, 4), (3, 5), (3, 6),

(4, 1), (4, 2), (4, 3), (4, 4), (4, 5), (4, 6),

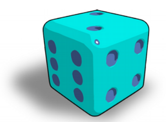
(5, 1), (5, 2), (5, 3), (5, 4), (5, 5), (5, 6),

(6, 1), (6, 2), (6, 3), (6, 4), (6, 5), (6, 6)}

Rolling two dice



D_1



D_2

E G

$$P(D_1 = 1, D_1 + D_2 = 4) = ?$$

P(EG) = 1/36

≠ 1/6 · 3/36

P(E) = 1/6

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),

P(G) = 3/36

(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)}

Independence

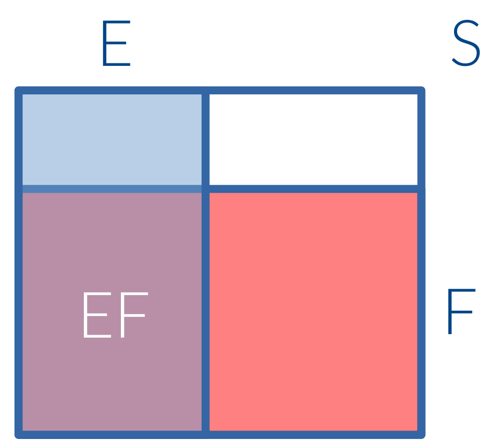
Two events are **independent** if you can **multiply** their probabilities to get the probability of **both** happening.

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

$$\Leftrightarrow$$

$$E \perp F$$

(“independent of”)



Independence as conditional

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

(if *and only if* E, F are independent)

Conditioning on the complement

If E, F are independent:

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

Three events

E, F, G are independent if:

$$P(EFG) = P(E)P(F)P(G)$$

and

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

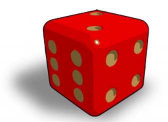
and

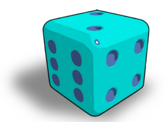
$$P(EG) = P(E)P(G)$$

and

$$P(FG) = P(F)P(G)$$

Rolling two dice again





D_1 D_2

E: event that $D_1 = 1$

F: event that $D_2 = 6$

G: event that $D_1 + D_2 = 7$

P(E) = 1/6

P(F) = 1/6

P(EF) = 1/36 ✓

= 1/6 · 1/6

Are **E** and **F** independent?

A) Yes

https://bit.ly/1a2ki4G → https://b.socrative.com/login/student/

Room: CS109SUMMER17

Independence and causation

If two events **don't affect each other**, and also have **no unknown factors in common** that affect both, then they're likely to be independent.

However:

If two events are independent, it **doesn't necessarily mean** they don't affect each other!

What it means is that knowing one doesn't give you **information** about the other.

Rolling two dice again

D_1 D_2

E: event that $D_1 = 1$

F: event that $D_2 = 6$

G: event that $D_1 + D_2 = 7$

$P(E) = 1/6$

$P(F) = 1/6$

$P(EF) = 1/36$ ✓

$= 1/6 \cdot 1/6$

$P(G) = 6/36 = 1/6$

$P(EG) = 1/36$ ✓

$= 1/6 \cdot 1/6$

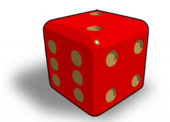
Are **E** and **G** independent?

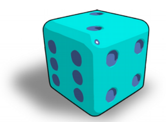
A) Yes

https://bit.ly/1a2ki4G → https://b.socrative.com/login/student/

Room: CS109SUMMER17

Rolling two dice again


D_1


D_2

E: event that $D_1 = 1$

F: event that $D_2 = 6$

G: event that $D_1 + D_2 = 7$

$P(E) = 1/6$

$P(F) = 1/6$

$P(EF) = 1/36 = 1/6 \cdot 1/6$ ✓

$P(G) = 6/36 = 1/6$

$P(EG) = 1/36 = 1/6 \cdot 1/6$ ✓

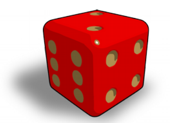
$P(FG) = 1/36 = 1/6 \cdot 1/6$ ✓

Are F and G independent?

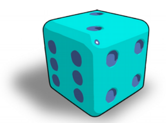
A) Yes

https://bit.ly/1a2ki4G → https://b.socrative.com/login/student/
Room: CS109SUMMER17

Rolling two dice again



D_1



D_2

E: event that $D_1 = 1$

F: event that $D_2 = 6$

G: event that $D_1 + D_2 = 7$

$P(E) = 1/6$

$P(F) = 1/6$

$P(EF) = 1/36 = 1/6 \cdot 1/6$ ✓

$P(G) = 6/36 = 1/6$

$P(EG) = 1/36 = 1/6 \cdot 1/6$ ✓

$P(FG) = 1/36 = 1/6 \cdot 1/6$ ✓

$P(EFG) = 1/36 \neq 1/6 \cdot 1/6 \cdot 1/6$ ✗

Are E, F and G independent?

B) No (!)

https://bit.ly/1a2ki4G → https://b.socrative.com/login/student/

Room: CS109SUMMER17

Many events

$$E_1, E_2, \ldots, E_n$$

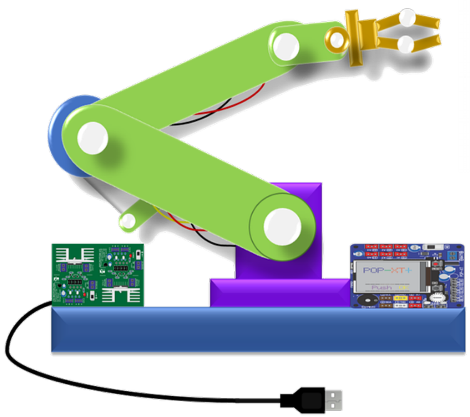
are independent if

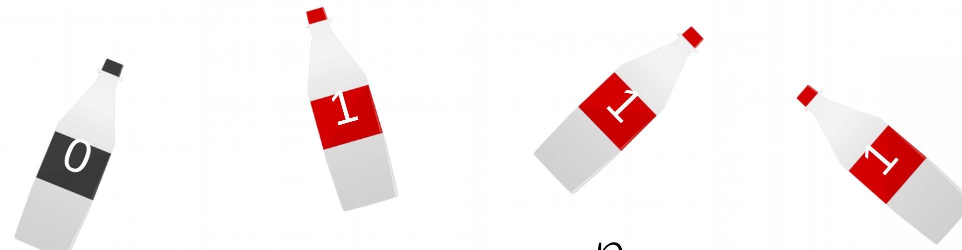
for every subset

$$E_{i_1}, E_{i_2}, \ldots, E_{i_r}:$$

$$P\left(E_{i_1} E_{i_2} \ldots E_{i_r}\right) = P\left(E_{i_1}\right) P\left(E_{i_2}\right) \ldots P\left(E_{i_r}\right)$$

Generating random bits





$(1 - p)$ $\quad p \quad p \quad p$

$\underbrace{\qquad\qquad\qquad}_{n}$

P(first n are 1's) =
$p \cdot p \cdot \ldots = p^n$

Each bit is a 1 with probability p, 0 with probability $(1 - p)$.

E: generate n 1's, followed by a single 0

$$P(E) = p^n (1 - p)$$

(Biased) coin flipping

n flips, each flip heads with probability p, tails with probability $(1 - p)$

E: n heads

$$P(E) = p^n$$

F: n tails

$$P(F) = (1 - p)^n$$

G: k heads, then $n - k$ tails

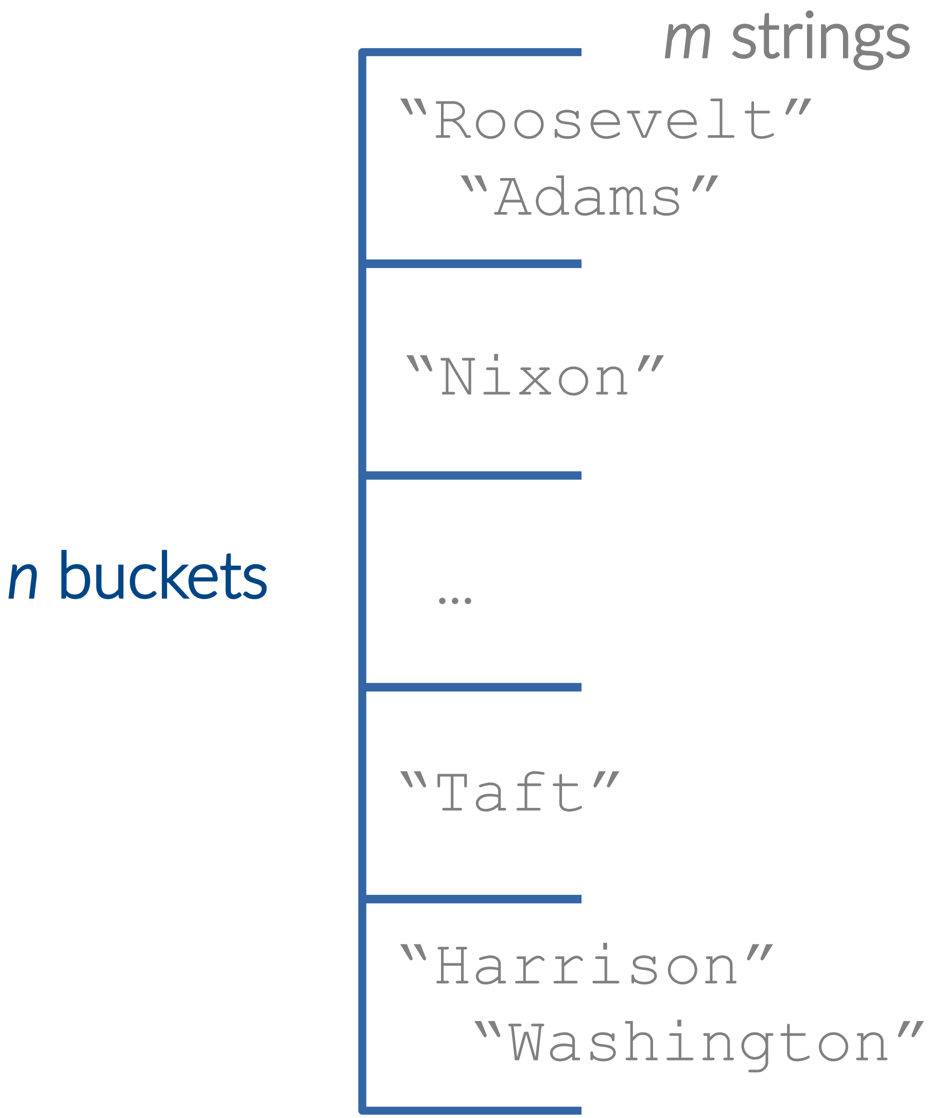
$$P(G) = p^k(1 - p)^{n-k}$$

H: exactly k heads

$$P(H) = \binom{n}{k} p^k (1-p)^{n-k}$$

Break time!

String hashing



All buckets equally likely, all independent

E: at least 1 hashed to first bucket

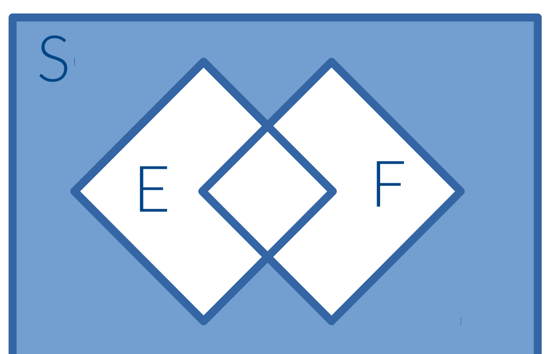
F_i: string *i* hashed to first bucket

$$P(E) = P(F_1 \cup F_2 \cup \ldots \cup F_m)$$

Review: Getting rid of ORs

Finding the probability of an OR of events can be nasty. Try **using De Morgan's laws** to turn it into an AND!

$$P(A \cup B \cup \cdots \cup Z) = 1 - P(A^c B^c \cdots Z^c)$$



String hashing

m strings

"Roosevelt"
"Adams"

"Nixon"

...

"Taft"

"Harrison"
"Washington"

n buckets

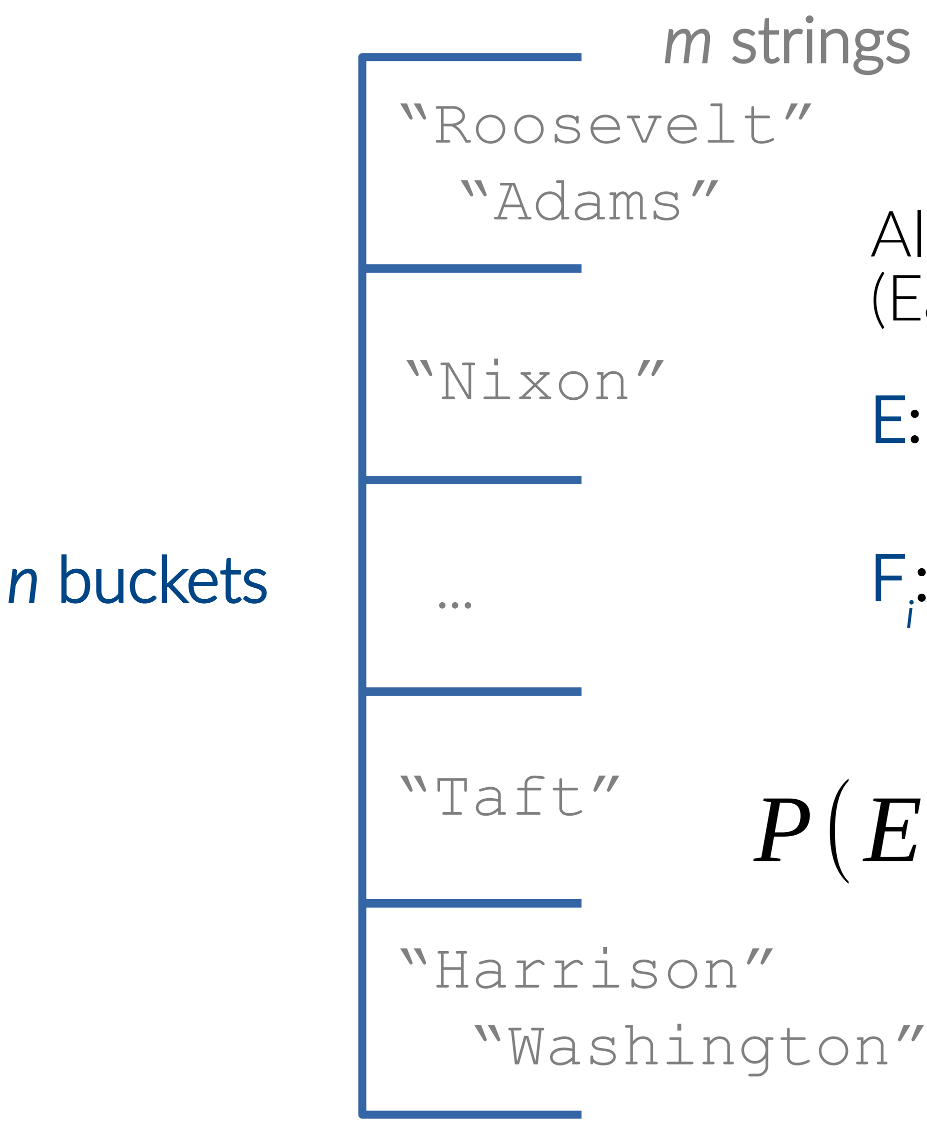
All buckets equally likely, all independent

E: at least 1 hashed to first bucket

F_i: string *i* hashed to first bucket

$$P(E) = P(F_1 \cup F_2 \cup \ldots \cup F_m)$$
$$= 1 - P(F_1^C F_2^C \ldots F_m^C)$$
$$= 1 - P(F_1^C) P(F_2^C) \ldots P(F_m^C)$$
$$= 1 - \left(\frac{n-1}{n}\right)^m$$

String hashing



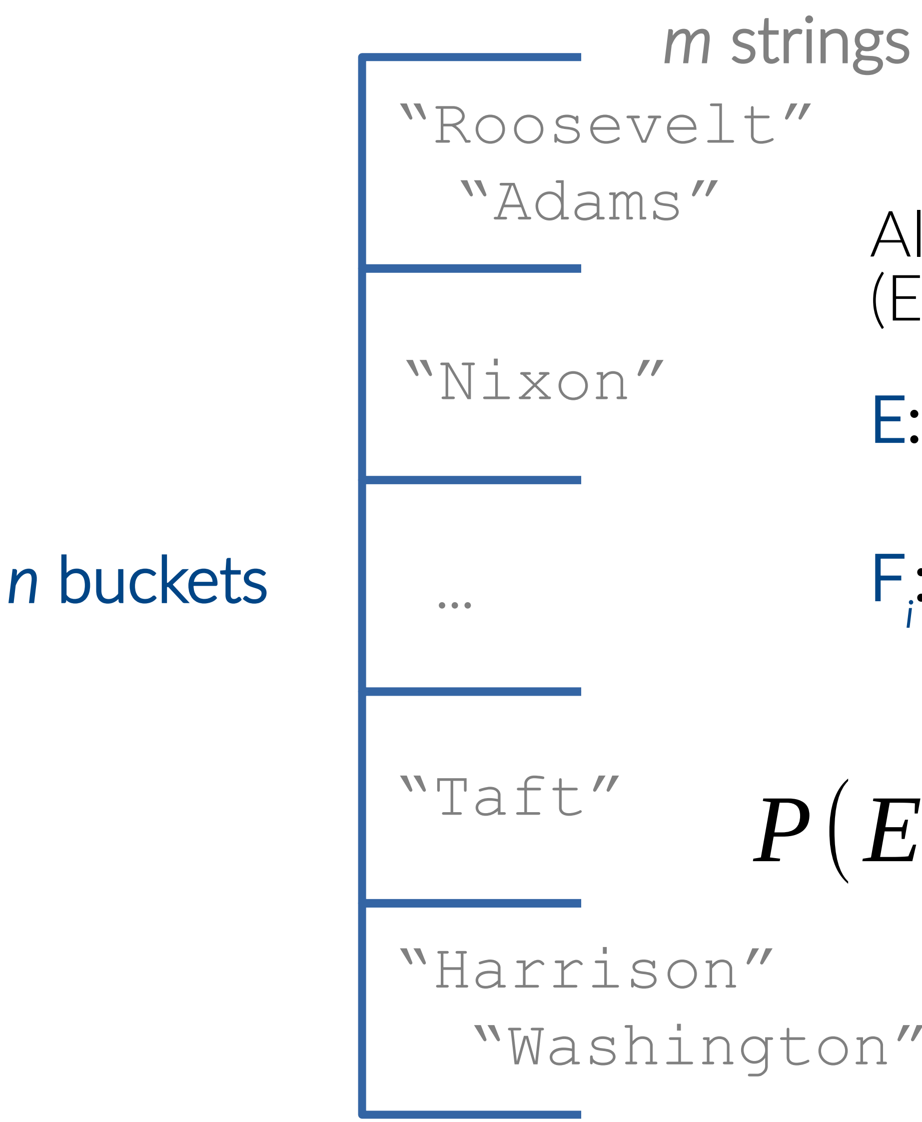
All independent, **not equally likely.**
(Each string → bucket i with prob. p_i)

E: at least 1 string hashed to **any** of the first k buckets

F_i: bucket i gets at least one string

$$P(E) = P(F_1 \cup F_2 \cup \ldots \cup F_k)$$
$$= 1 - P(F_1^C F_2^C \ldots F_k^C)$$
$$= 1 - (1 - p_1 - p_2 - \ldots - p_k)^m$$

String hashing



All independent, not equally likely.
(Each string → bucket i with prob. p_i)

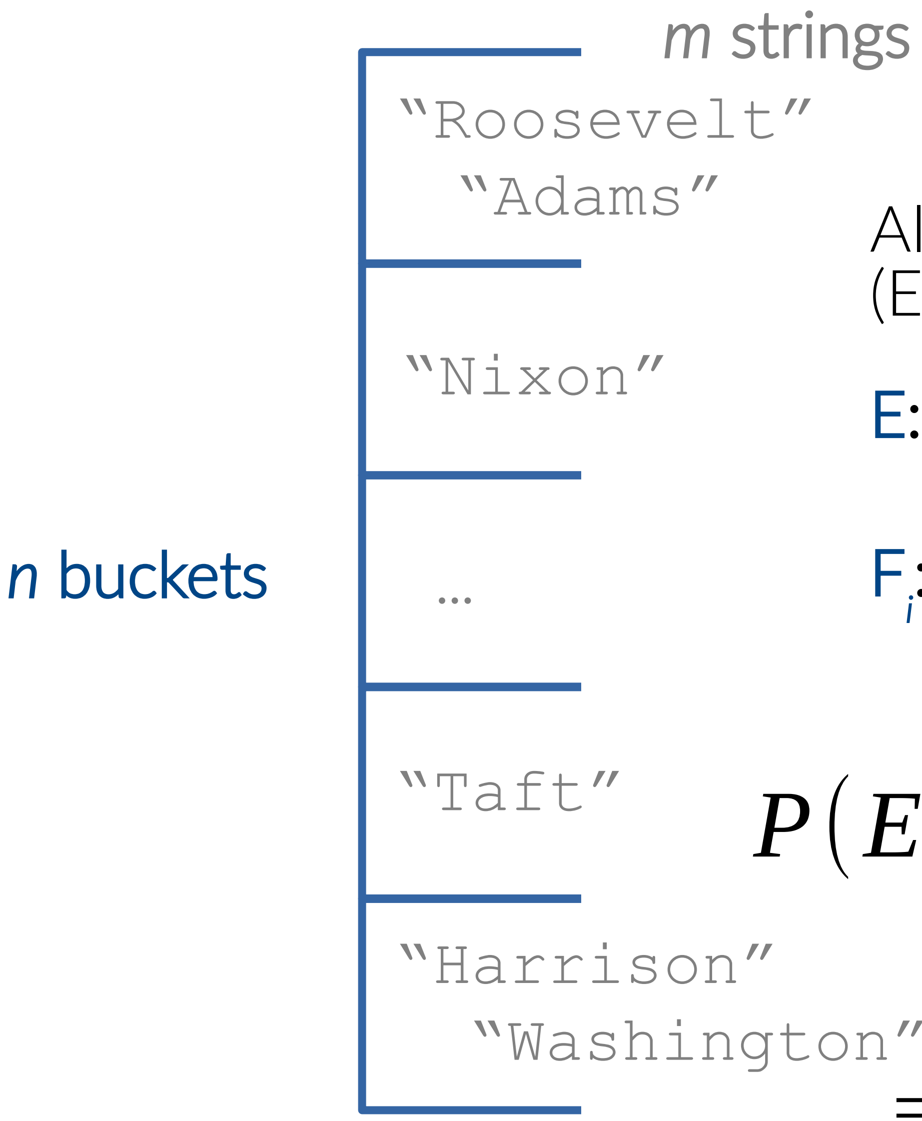
E: at least 1 string hashed to **each** of the first k buckets

F_i: bucket i gets at least one string

$$P(E) = P(F_1 F_2 \ldots F_k)$$
$$= P(F_1) P(F_2) \ldots P(F_k)$$

[check your assumptions]

String hashing



All independent, not equally likely.
(Each string → bucket i with prob. p_i)

E: at least 1 string hashed to **each** of the first k buckets

F_i: bucket i gets at least one string

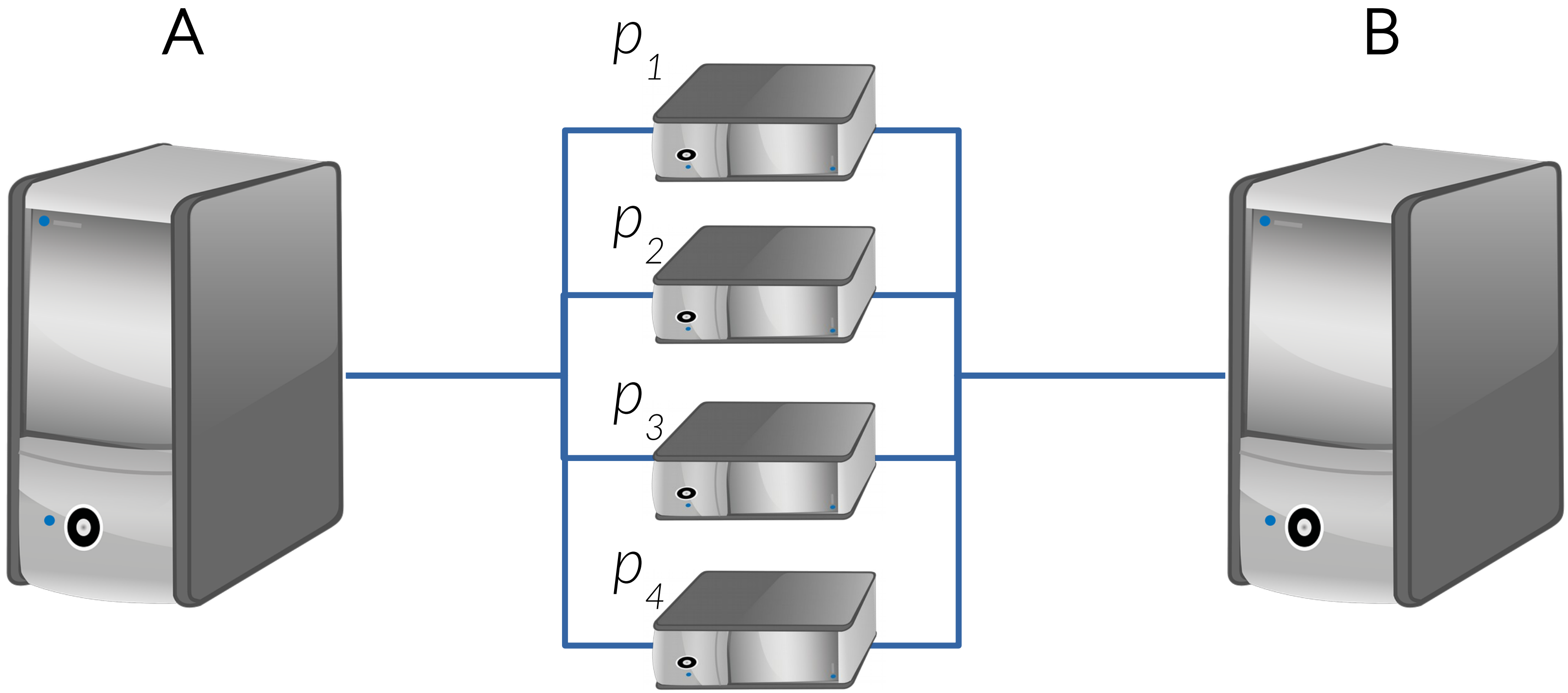
$$P(E) = P(F_1 F_2 \ldots F_k)$$

$$= 1 - P(F_1^C \cup F_2^C \cup \ldots \cup F_k^C)$$

$$= 1 - \sum_{r=1}^{k} (-1)^{(r+1)} \sum_{i_1 < \cdots < i_r} P\left(\bigcap_{j=1}^{r} F_{i_j}^C\right)$$

$$= 1 - \sum_{r=1}^{k} (-1)^{(r+1)} \sum_{i_1 < \cdots < i_r} (1 - p_{i_1} - p_{i_2} - \cdots - p_{i_r})^m$$

Network reliability



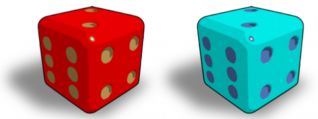
n independent routers, each works with prob. p_i

E: path exists from A to B

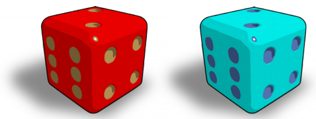
P(E) = ?

$$P(E) = 1 - P(\text{all routers fail})$$
$$= 1 - (1-p_1)(1-p_2)\ldots(1-p_n) = 1 - \prod_{i=1}^{n}(1-p_i)$$

Simplified craps

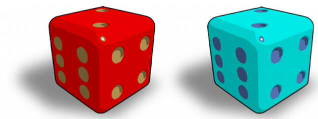


9



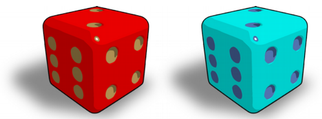
4

$

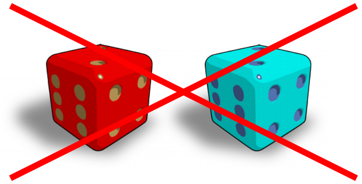


5

Simplified craps



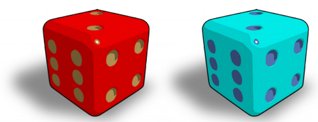
9



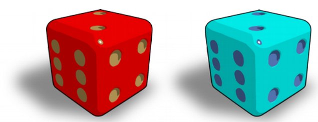
7

Simplified craps

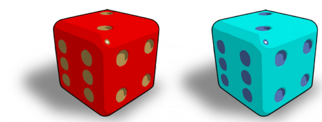
$



9



4



5

E: roll 5 before rolling 7

F_n: n – 1 rolls without a 5 or 7, then a 5

To Infinity! (and beyond?)

$$P(E) = P(F_1 \cup F_2 \cup \ldots)$$

$$= \sum_{n=1}^{\infty} P(F_n)$$

$$= \sum_{n=1}^{\infty} \left(1 - \frac{10}{36}\right)^{n-1} \left(\frac{4}{36}\right)$$

$$= \left(\frac{4}{36}\right) \sum_{n=1}^{\infty} \left(\frac{26}{36}\right)^{n-1}$$

$$P(5) = 4/36$$

$$P(7) = 6/36$$

Geometric series

$$x^0 + x^1 + x^2 + \cdots + x^n = \sum_{i=0}^{n} x^i$$

$$= \frac{1 - x^{n+1}}{1 - x}$$

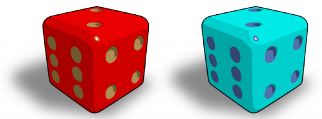
If $|x| < 1$, then as $n \to \infty$:

$$\sum_{i=0}^{n} x^i \to \frac{1}{1 - x}$$

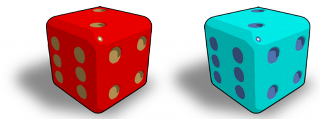
See "Calculation Reference"
for more super-useful sum
and product identities!

Simplified craps

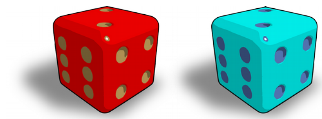
$



9



4



5

E: roll 5 before rolling 7

F_n: $n - 1$ rolls without a 5 or 7, then a 5

To Infinity! (and beyond?)

$$P(E) = P(F_1 \cup F_2 \cup \ldots)$$

$$= \sum_{n=1}^{\infty} P(F_n)$$

$$= \sum_{n=1}^{\infty} \left(1 - \frac{10}{36}\right)^{n-1} \left(\frac{4}{36}\right)$$

$$= \left(\frac{4}{36}\right) \sum_{n=0}^{\infty} \left(\frac{26}{36}\right)^{n}$$

$$P(5) = 4/36$$

$$P(7) = 6/36$$

$$= \left(\frac{4}{36}\right) \frac{1}{\left(1 - \frac{26}{36}\right)} = \frac{2}{5}$$

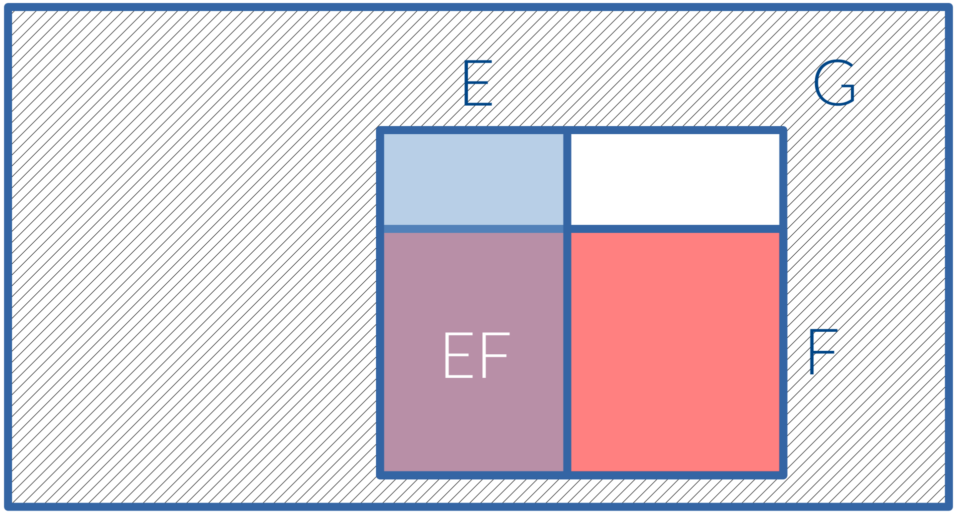
Conditional independence

Two events are **conditionally independent** if you can **multiply** their conditional probabilities to get the conditional probability of **both** happening.

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

$$\Leftrightarrow$$

$$(E \perp F) \big| G$$



Conditional independence

In general,

$$\left(E \perp F\right)$$

does **not** imply

$$\left(E \perp F\right) \Big| G$$

or vice versa!

Rolling two dice again

D_1 D_2

E: event that $D_1 = 1$

F: event that $D_2 = 6$

G: event that $D_1 + D_2 = 7$

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

$P(E \mid G) = 1/6$
$P(F \mid G) = 1/6$
$P(EF \mid G) = 1/6$
$\neq 1/6 \cdot 1/6$

Faculty night

At faculty night:

44 students

30 straight A's— $P(A \mid F) \approx 0.68$

20 CS majors

6 CS w/ straight A's— $P(A \mid CS, F) = 0.30$

Survey the whole dorm:

100 students

30 straight A's— $P(A) = 0.30$

20 CS majors

6 CS w/ straight A's— $P(A \mid CS) = 0.30$

Independent! What gives?

Faculty night

All the A students came to faculty night. So did all the CS majors.

A, CS **conditionally dependent**, conditioned on faculty night.

At faculty night:

44 students

30 straight A's— P(A | F) ≈ 0.68

20 CS majors

6 CS w/ straight A's— P(A | CS, F) = 0.30

Survey the whole dorm:

100 students

30 straight A's— P(A) = 0.30

20 CS majors

6 CS w/ straight A's— P(A | CS) = 0.30

Watering the lawn

E: event that it rained today
F: event that the sprinkler was on

Say E and F are independent.

G: event that the grass is wet

You observe the grass is wet. Probability of both rain and sprinkler goes up!

$$P(E \mid G) > P(E)$$
$$P(F \mid G) > P(F)$$

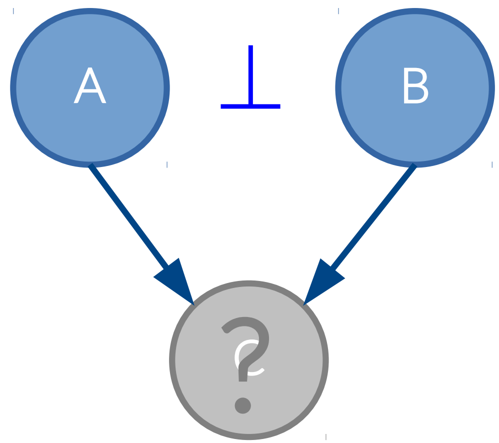
Now you find out the sprinklers were on.
Does your belief that it rained change?

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

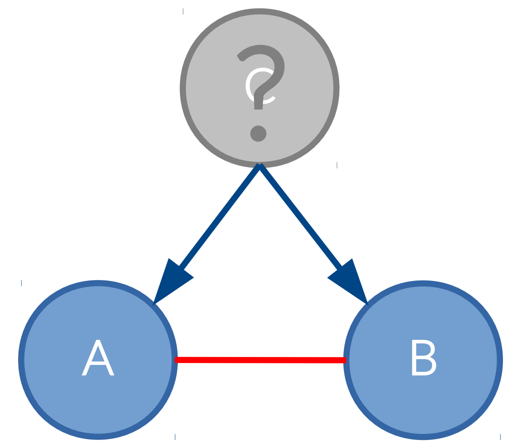
A graphical representation

C is unknown

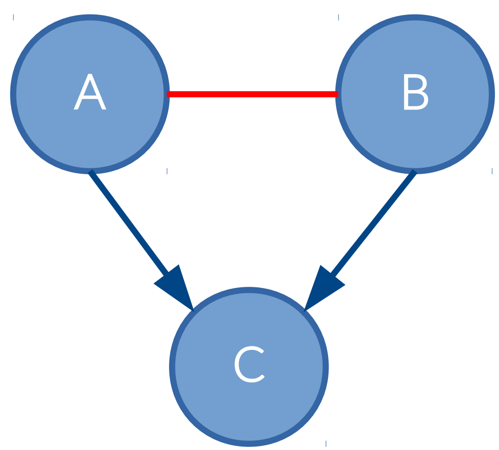
A, B independent!



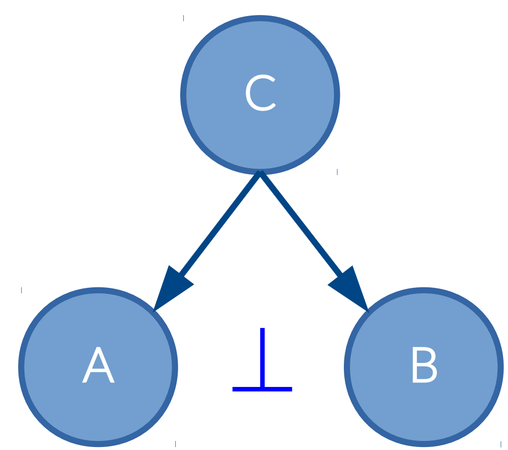
A, B **not** independent



Condition on C



A, B **not** conditionally independent

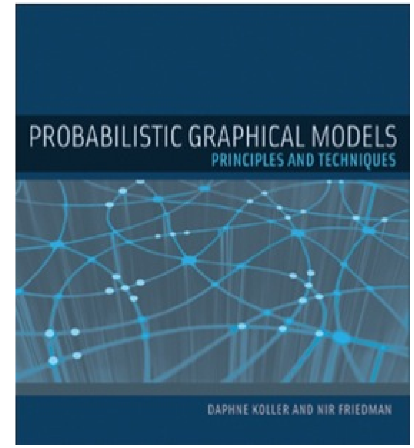


A, B conditionally independent!

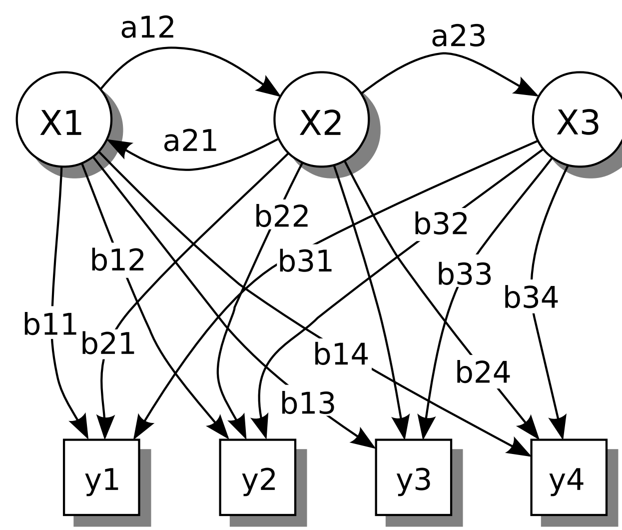
For more on this:

Daphne Koller—
Designed & taught earlier versions, "wrote the book" (literally & figuratively)

"the book"



Stefano Ermon—
Taught Winter 2017



CS228: Probabilistic Graphical Models

Tdunning/Razorbliss

Independence

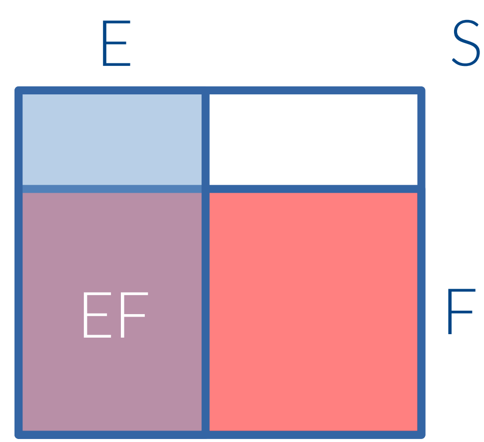
Two events are **independent** if you can **multiply** their probabilities to get the probability of **both** happening.

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

$$\Leftrightarrow$$

$$E \perp F$$

("independent of")



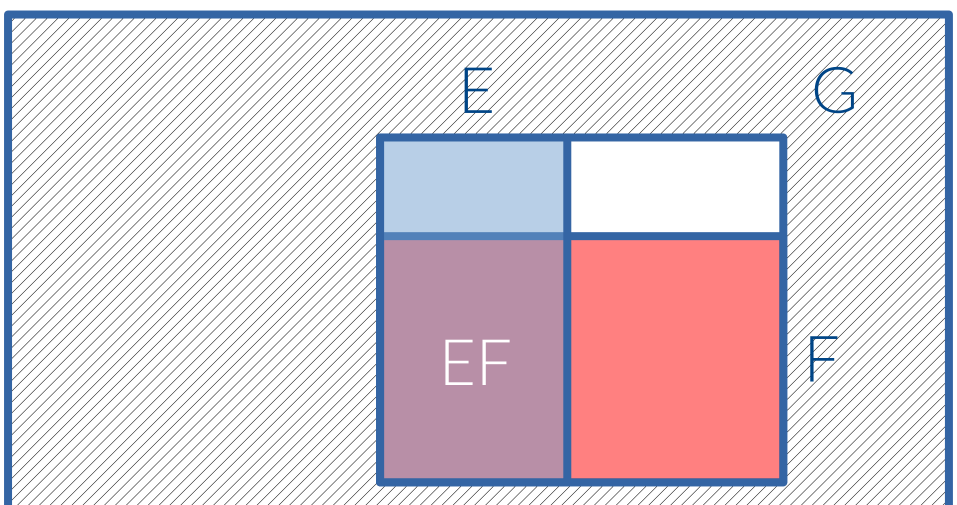
Conditional independence

Two events are **conditionally independent** if you can **multiply** their conditional probabilities to get the conditional probability of **both** happening.

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

$$\Leftrightarrow$$

$$(E \perp F)\,|\,G$$



Reminder: Python tutorial

Right now!

Follow Yuling to Gates B03