05: Independence

Lisa Yan
April 15, 2020

Quick slide reference

Generalized Chain Rule

Chain Rule

Review

Definition of conditional probability:

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

The Chain Rule:

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

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

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

Generalized Chain Rule

$$P(E_1 E_2 E_3 \ldots E_n) \\ = P(E_1)P(E_2|E_1)P(E_3|E_1 E_2) \ldots P(E_n|E_1 E_2 \ldots E_{n-1})$$

Lisa Yan, CS109, 2020

Quick check

$P(E_1E_2E_3 \ldots E_n) = P(E_1)P(E_2|E_1) \ldots P(E_n|E_1E_2 \ldots E_{n-1})$ Chain Rule

You are going to a friend's ~~Halloween~~ virtual costume party.

Let C = there is candy
M = there is music
E = no one wears your costume
W = you wear a costume

An awesome party means that all of these events must occur.

What is $P(\text{awesome party}) = P(CMEW)$?

A. $P(C)P(M|C)P(E|CM)P(W|CME)$
B. $P(M)P(C|M)P(E|MC)P(W|MCE)$
C. $P(W)P(E|W)P(CM|EW)$
D. A, B, and C
E. None/other

Quick check

$$P(E_1 E_2 E_3 \dots E_n) = P(E_1)P(E_2|E_1) \dots P(E_n|E_1 E_2 \dots E_{n-1})$$ Chain Rule

virtual costume

You are going to a friend's ~~Halloween~~ party.

Let C = there is candy
M = there is music
E = no one wears your costume
W = you wear a costume

An awesome party means that all of these events must occur.

What is $P(\text{awesome party}) = P(CMEW)$?

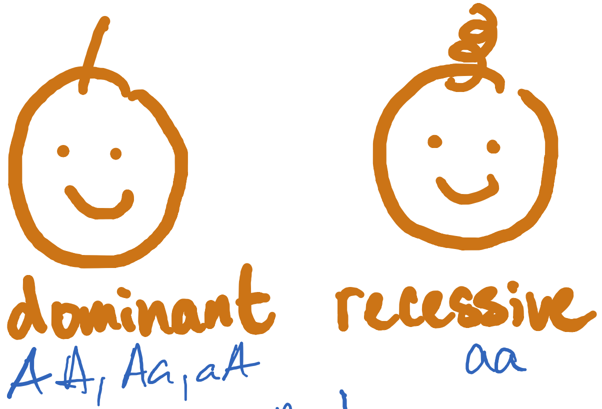
A. $P(C)P(M|C)P(E|CM)P(W|CME)$ $= P(CMEW)$
B. $P(M)P(C|M)P(E|MC)P(W|MCE)$ $= P(MCEW)$
→ C. $P(W)P(E|W)P(CM|EW)$ $= P(WECM)$
[✓ D.] A, B, and C
E. None/other

Chain Rule is a way of introducing "order" and "procedure" into probability.

Think of the children

Two parents both have an (A, a) gene pair.

- Each parent will pass on one of their genes (each gene equally likely) to their child.
- The probability of a single child having curly hair (recessive trait) is 0.25.
- There are three children.



dad \ mom	A	a
A	AA	Aa
a	aA	aa

What is the probability that all three children have curly hair?

Let E_1, E_2, E_3 be the events that child 1, 2, and 3 have curly hair, respectively.

$$P(E_1 E_2 E_3) = \underset{\underset{0.25}{\uparrow}}{P(E_1)} P(E_2 | E_1) P(E_3 | E_1 E_2)$$

up next: simplify chain rule!

$P(E_i) = 0.25$

05b_independence_i

Independence I

1. independence
2. conditional probability
3.

⋮

10.

Independence

Two events E and F are defined as independent if:

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

Otherwise E and F are called dependent events.

If E and F are independent, then:

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

Intuition through proof

Independent events E and F $\Leftrightarrow$ $P(EF) = P(E)P(F)$

Statement:

If E and F are independent, then $P(E|F) = P(E)$.

Proof:

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

Definition of conditional probability

$$= \frac{P(E)\cancel{P(F)}}{\cancel{P(F)}}$$

Independence of E and F

$$= \underline{P(E)}$$

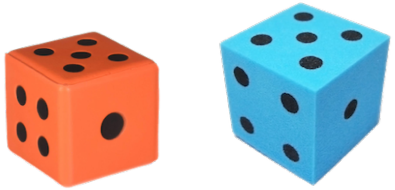
Taking the bus to cancellation city

Knowing that F happened does not change our belief that E happened.

Dice, our misunderstood friends

Independent events E and F ⟺ $P(EF) = P(E)P(F)$
$P(E|F) = P(E)$

- Roll two 6-sided dice, yielding values D_1 and D_2.
- Let event E: $D_1 = 1$
  event F: $D_2 = 6$
  event G: $D_1 + D_2 = 5$ $G = \{(1,4),(2,3),(3,2),(4,1)\}$ $|G| = 4$



1. Are E and F independent?

$P(E) = 1/6$ $EF = \{(1,6)\}$
$P(F) = 1/6$
$P(EF) = 1/36$ $\frac{1}{36} = \frac{1}{6} \cdot \frac{1}{6}$

independent

2. Are E and G independent?

$P(E) = 1/6$ $EG = \{(1,4)\}$
$P(G) = 4/36 = 1/9$
$P(EG) = 1/36 \neq P(E)P(G)$ $\neq 1/6 \cdot 1/9$

✗ dependent

Generalizing independence

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

$$\left\{\begin{array}{l} P(EFG) = P(E)P(F)P(G), \text{ and} \quad \leftarrow \\ \left.\begin{array}{l} P(EF) = P(E)P(F), \text{ and} \\ P(EG) = P(E)P(G), \text{ and} \\ P(FG) = P(F)P(G) \end{array}\right\} \text{pairwise independence} \end{array}\right.$$

n events $E_1, E_2, \ldots, E_n$ are independent if:

$$\left\{\begin{array}{l} \text{for } r = 1, \ldots, n: \\ \quad \text{for every subset } E_1, E_2, \ldots, E_r: \\ \quad\quad P(E_1, E_2, \ldots, E_r) = P(E_1)P(E_2) \cdots P(E_r) \end{array}\right.$$

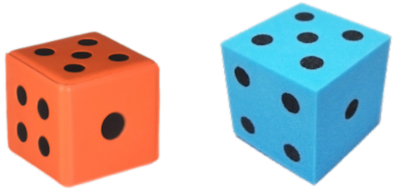
Independent trials:

$\underline{H}\quad \underline{\quad}\quad \underline{\quad}\quad \underline{H}\ \underline{H}\ \underline{H}$

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.

Dice, increasingly misunderstood (still our friends)



- Each roll of a 6-sided die is an independent trial.
- Two rolls: D_1 and D_2.
- Let event E: $D_1 = 1$

  event F: $D_2 = 6$

  event G: $D_1 + D_2 = 7$

$G = \{(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)\}$

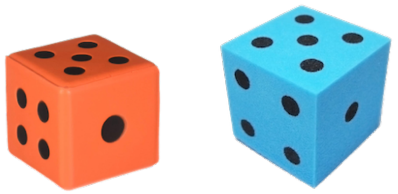
1. Are E and F independent? ✅
2. Are E and G independent?
3. Are F and G independent?
4. Are E, F, G independent?

$EF = \{(1,6)\}$

$P(EF) = 1/36$

$= 1/6 \cdot 1/6$

Dice, increasingly misunderstood (still our friends)



- Each roll of a 6-sided die is an independent trial.
- Two rolls: D_1 and D_2.
- Let event E: $D_1 = 1$
  event F: $D_2 = 6$
  event G: $D_1 + D_2 = 7$

$P(G) = \frac{|G|}{|S|} = \frac{6}{36} = \frac{1}{6}$

$G = \{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}$

1. Are E and F ✅ independent?

   $P(E) = 1/6$

   $P(F) = 1/6$

   $P(EF) = 1/36$

2. Are E and G independent?

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

   $EG = \{(1,6)\}$

   $P(EG) = 1/36 = 1/6 \cdot 1/6$

3. Are F and G independent?

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

   $FG = \{(1,6)\}$

   $= P(FG) = 1/36 = 1/6 \cdot 1/6$

4. Are E, F, G ❌ independent?

   $EFG = \{(1,6)\}$

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

Pairwise independence is not sufficient to prove independence of >2 events!

05b_independence_ii

Independence II

Independent trials

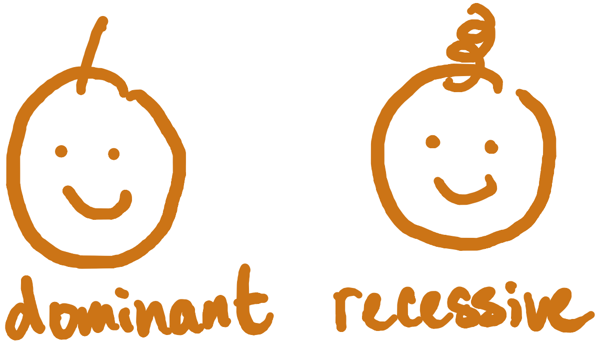
We often are interested in experiments consisting of n independent trials.

- n trials, each with the same set of possible outcomes
- n-way independence: an event in one subset of trials is independent of events in other subsets of trials

Examples:

- Flip a coin n times
- Roll a die n times
- Send a multiple choice survey to n people
- Send n web requests to k different servers

Think of the children as independent trials



Two parents both have an (A, a) gene pair.

- Each parent will pass on one of their genes (each gene equally likely) to their child.
- The probability of a single child having curly hair (recessive trait) is 0.25.
- There are three children. **Each child is an independent trial.** ←

What is the probability that all three children have curly hair?

$P(E_i) = 0.25$

E_1, E_2, E_3 are independent.

Let E_1, E_2, E_3 be the events that child 1, 2, and 3 have curly hair, respectively.

$$P(E_1E_2E_3) = P(E_1)\underbrace{P(E_2|E_1)}\underbrace{P(E_3|E_1E_2)}$$

$$= P(E_1)\,P(E_2)\,P(E_3)$$

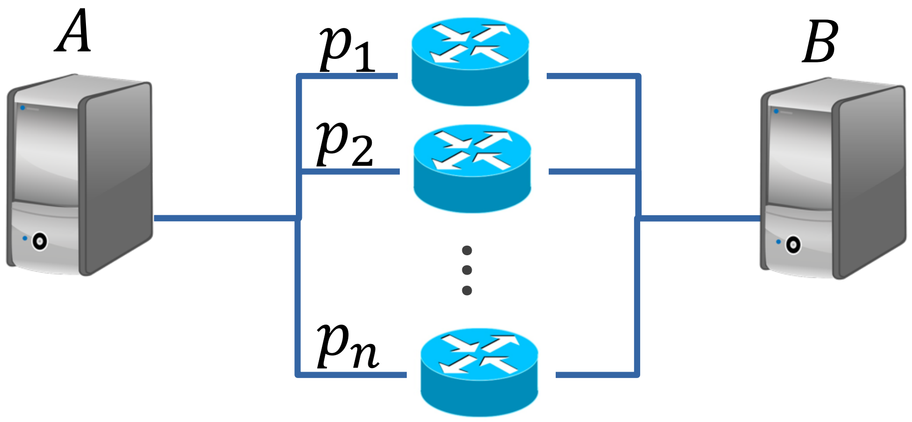
$$= (0.25)^3$$

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.

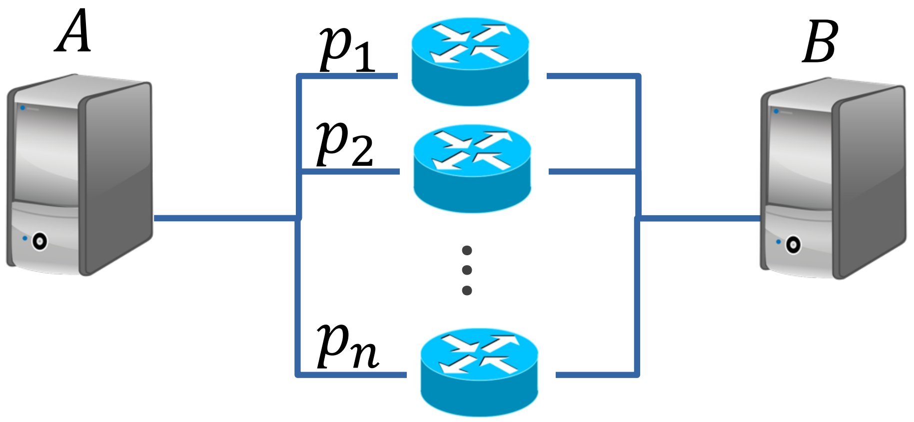
What is $P(E)$?



Network reliability

Consider the following parallel network:

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



router i: functioning p_i

failure mode $(1-p_i)$

What is $P(E)$?

$$P(E) = P(\geq 1 \text{ one router works})$$

$$= 1 - P(\text{all routers fail}) = 1 - P(R_1 \text{fail} \cap R_2 \text{fail} \cap \cdots \cap R_n \text{fail})$$

$$= 1 - (1-p_1)(1-p_2)\cdots(1-p_n)$$

$$= 1 - \prod_{i=1}^{n}(1-p_i)$$

≥ 1 with independent trials: take complement

05: Independence (live)

Lisa Yan
April 15, 2020

Independence

Review

Two events E and F are defined as independent if:

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

For independent events E and F,

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

Think

Slide 24 has two questions to think over by yourself. We’ll go over it together afterwards.

Post any clarifications here!

https://us.edstem.org/courses/109/discussion/27279

Think by yourself: 2 min

(by yourself)

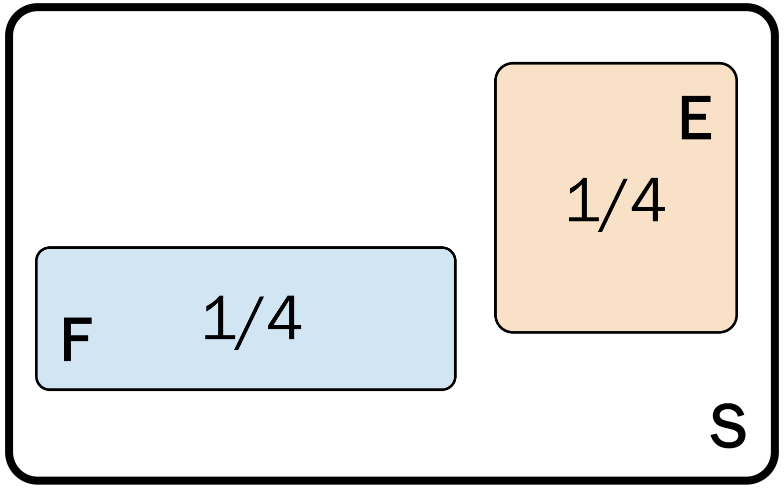
Independence?

Choose A or B or both

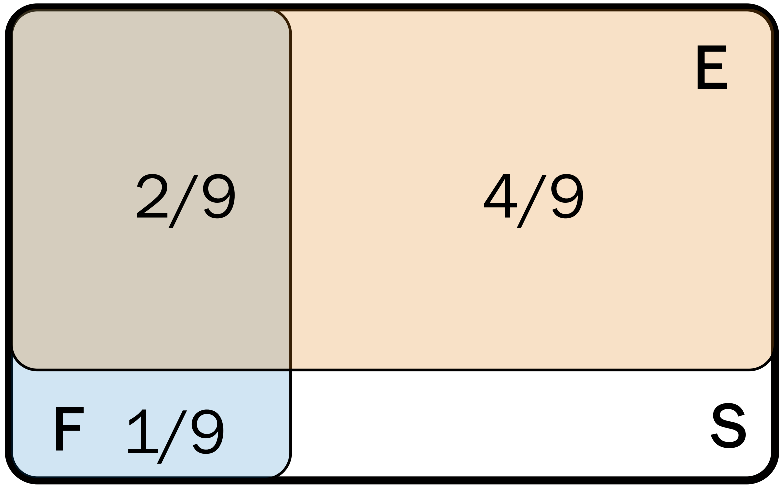
Independent events E and F $\Leftrightarrow$ $P(EF) = P(E)P(F)$
$P(E|F) = P(E)$

1. Two events E and F are independent if:
   A. Knowing that F happens means that E can't happen.
   B. Knowing that F happens doesn't change probability that E happened.
2. Are E and F independent in the following pictures?

A.



B.



(by yourself)

Independent events E and F ⟺ $P(EF) = P(E)P(F)$

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

Independence?

1. Two events E and F are independent if:

Assuming $P(E) > 0$

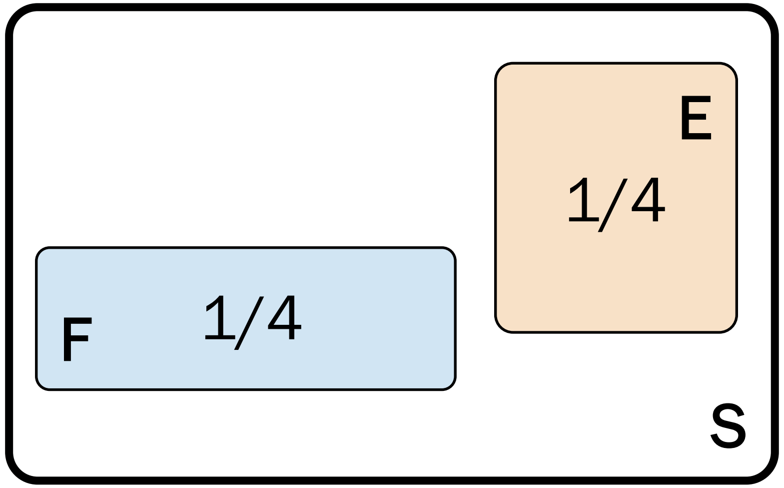
~~A.~~ Knowing that F happens means that E can't happen. $P(E|F) = 0 \neq P(E)$

B. Knowing that F happens doesn't change probability that E happened.

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

2. Are E and F independent in the following pictures?

~~A.~~

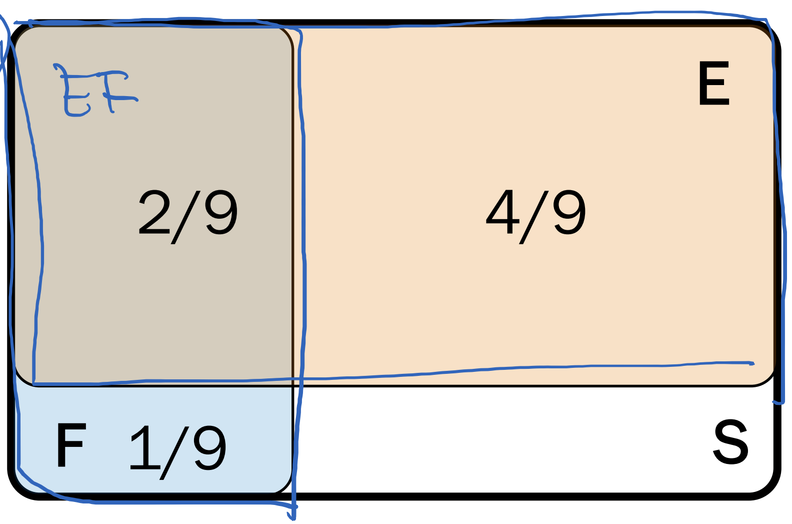


$P(E) = 1/4$ $P(F) = 1/4$

$P(EF) = 0 \neq P(E)P(F)$

$EF = \emptyset$ MUTUALLY EXCLUSIVE

B.



$P(E) = 2/9 + 4/9$
$= 2/3$

$P(F) = 2/9 + 1/9 = 1/3$

$P(EF) = 2/9$
$= P(E)P(F)$
$= \frac{2}{3} \cdot \frac{1}{3}$

Be careful:

- Independence is NOT mutual exclusion.
- Independence is difficult to visualize graphically.

Independence

Two events E and F are defined as independent if:

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

For independent events E and F,

- $P(E|F) = P(E)$
- E and F^C are independent.

new

Independence of complements

Statement:

If E and F are independent, then E and F^C are independent.

Proof:

$P(\underline{EF^C}) = P(E) - P(EF)$	Intersection
$= P(E) - P(E)P(F)$	Independence of E and F
$= P(E)[1 - P(F)]$	Factoring
$= P(\underline{E})P(\underline{F^C})$	Complement
E and F^C are independent	Definition of independence

$P(E \mid F^C) = P(E)$

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

Knowing that F **did or didn't happen** does not change our belief that E happened.

Independence

Review

Two events E and F are defined as independent if:

$$\boxed{P(EF) = P(E)P(F)}$$

For independent events E and F,

- $P(E|F) = P(E)$
- E and F^C are independent

Independent trials are when we observe independent sub-experiments, each of which has the same set of possible outcomes.

Breakout Rooms

Check out the questions on the next slide (Slide 30). Post any clarifications here!

https://us.edstem.org/courses/109/discussion/27279

Breakout rooms: 5 min. Introduce yourself!

🤔

(biased) Coin Flips

Suppose we flip a coin n times. Each coin flip is an **independent trial** with probability p of coming up heads. Write an expression for the following:

1. $P(n$ heads on n coin flips$)$
2. $P(n$ tails on n coin flips$)$
3. $P($first k heads, then $n-k$ tails$)$
4. $P($exactly k heads on n coin flips$)$

$\underbrace{H}_{p}\ \underbrace{H}_{p}\ \underbrace{H}_{p} \cdots \underbrace{H}_{p}$ $\quad p^n$

$\underbrace{T}_{1-p}\ T \cdots T$ $\quad (1-p)^n$

$\underbrace{H\ H \cdots H}_{k}\ \underbrace{T \cdots T}_{n-k}$ $\quad p^k (1-p)^{n-k}$

HHHT…T ⟸ $p^k (1-p)^{n-k}$

THHTH‥T ⟸ same prob

THTHTTTT ⟸ "

… HHH ⟸ …

$\binom{n}{k}$ 1. count w/ places to put heads

2. compute specific prob of this outcome

(biased) Coin Flips

Suppose we flip a coin n times. Each coin flip is an **independent trial** with probability p of coming up heads. Write an expression for the following:

1. $P(n$ heads on n coin flips$)$
2. $P(n$ tails on n coin flips$)$
3. $P($first k heads, then $n-k$ tails$)$
4. $P($exactly k heads on n coin flips$)$

p^n

$(1-p)^n$

$p^k(1-p)^{n-k}$

TTT··H··H

H··HT-··T

HTHT···HT

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

\# of mutually exclusive outcomes

$P($a particular outcome's k heads on n coin flips$)$

Make sure you understand #4! It will come up again.

Interlude for jokes/announcements

JOKE ON
SLIDE 40

Announcements

Free Online CTL Tutoring

CTL offers appointment tutoring for CS 109, in addition to tutoring for a number of other courses. For more information and to schedule an appointment, visit our tutoring appointments and drop-in schedule page. We also have a variety of remote learning resources and academic coaching available to assist with all of your learning needs!

Sections start today!

Late signups/change form: end of day

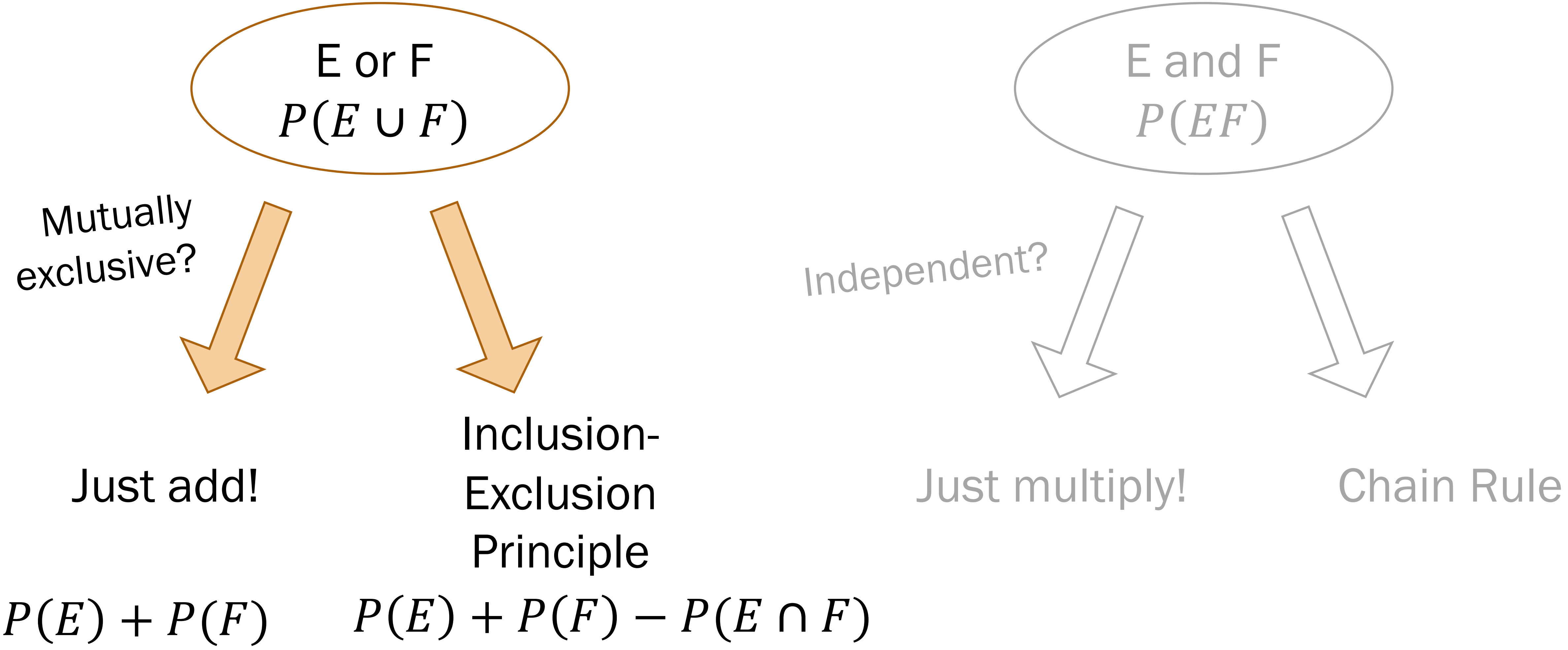
Problem Set 1

due: **10am** Pacific **Friday** (not 10:30am)

Still confused about Monty Hall? Check out the code!

https://us.edstem.org/courses/109/discussion/27277?comment=93040

Probability of events



E or F

$P(E \cup F)$

Mutually exclusive?

Just add!

$P(E) + P(F)$

Inclusion-Exclusion Principle

$P(E) + P(F) - P(E \cap F)$

E and F

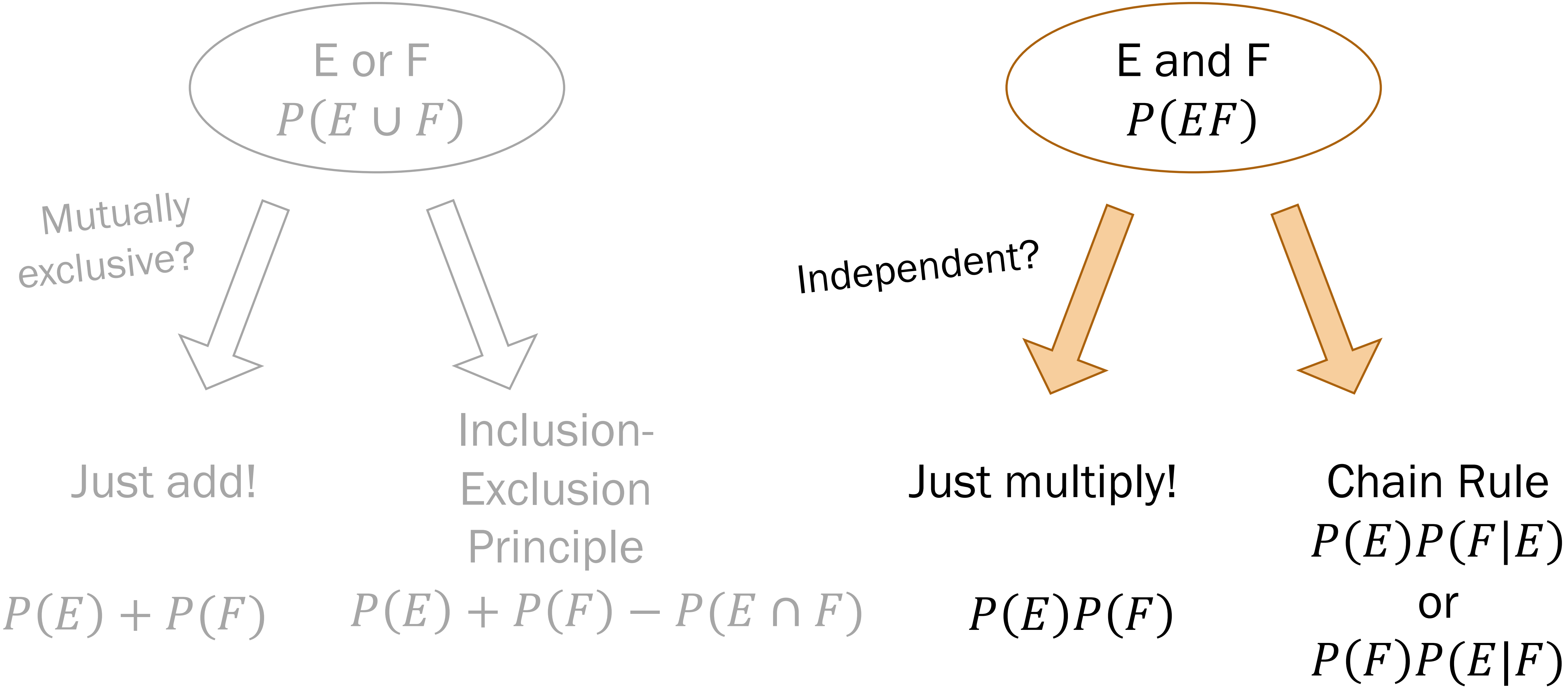
$P(EF)$

Independent?

Just multiply!

Chain Rule

Probability of events



E or F
$P(E \cup F)$

Mutually exclusive?

Just add!

$P(E) + P(F)$

Inclusion-Exclusion Principle

$P(E) + P(F) - P(E \cap F)$

E and F
$P(EF)$

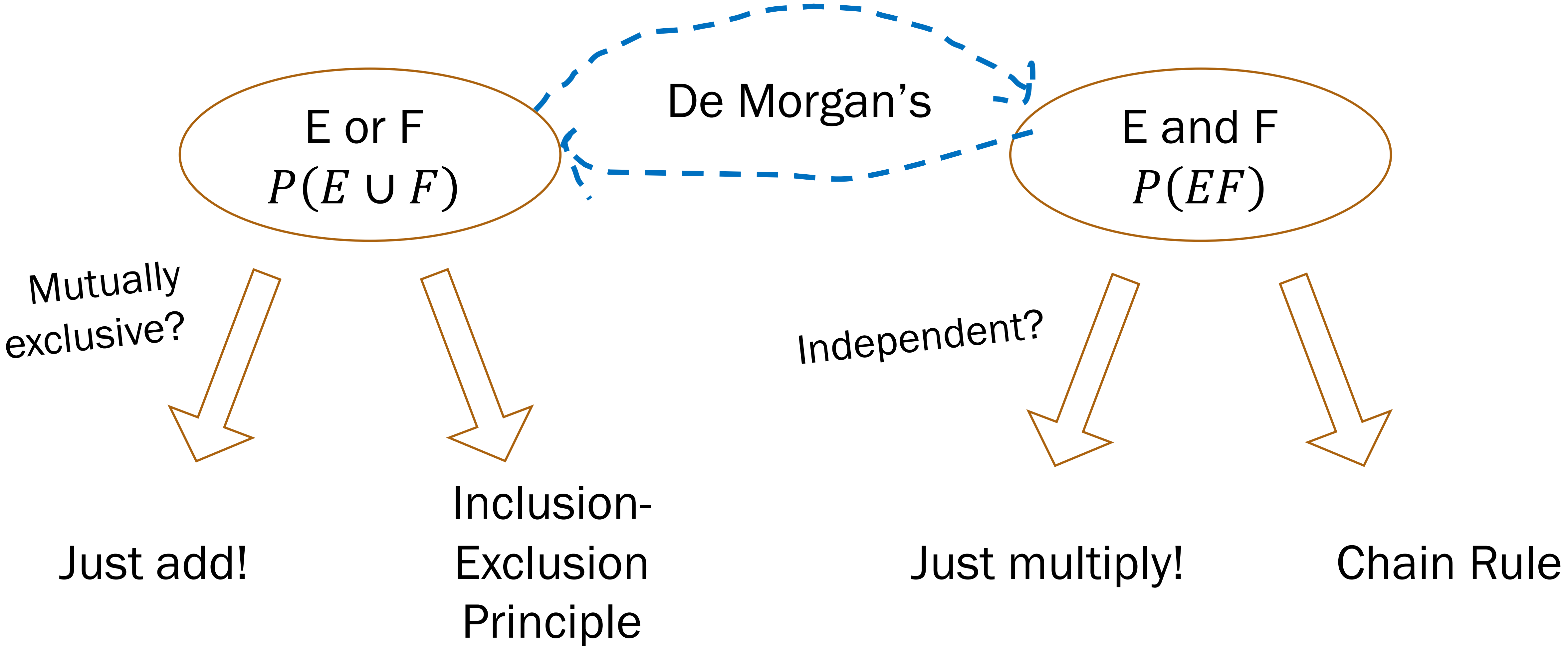
Independent?

Just multiply!

$P(E)P(F)$

Chain Rule
$P(E)P(F|E)$
or
$P(F)P(E|F)$

Probability of events



E or F

$P(E \cup F)$

De Morgan's

E and F

$P(EF)$

Mutually exclusive?

Just add!

Inclusion-Exclusion Principle

Independent?

Just multiply!

Chain Rule

Augustus De Morgan

Augustus De Morgan (1806–1871):

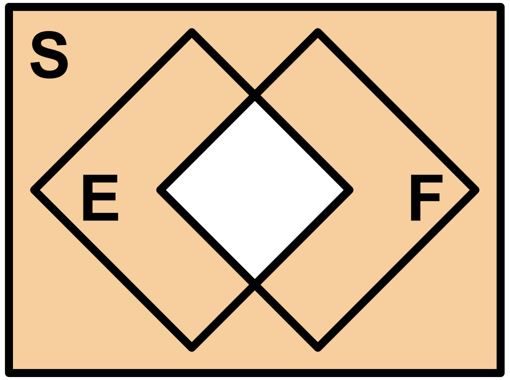
British mathematician who wrote the book *Formal Logic* (1847).

He looked remarkably similar to Jason Alexander (George from Seinfeld)

(but that's not important right now)

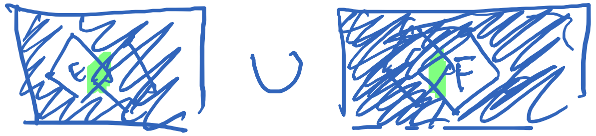
De Morgan’s Laws

DeMorgan’s lets you switch from AND to OR.



$$(E \cap F)^C = E^C \cup F^C$$

$$\left(\bigcap_{i=1}^{n} E_i\right)^C = \bigcup_{i=1}^{n} E_i^C$$



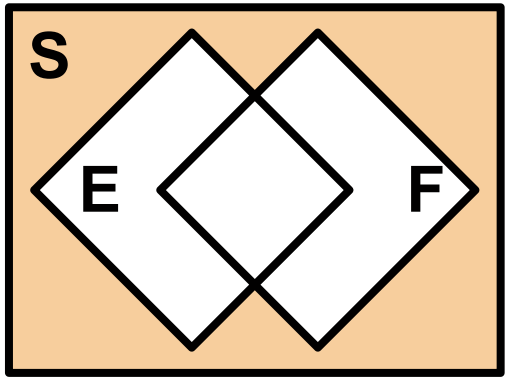
In probability:

$$P(E_1 E_2 \cdots E_n)$$

$$= 1 - P\left((E_1 E_2 \cdots E_n)^C\right)$$

$$= 1 - P(E_1^C \cup E_2^c \cup \cdots \cup E_n^c)$$

Great if E_i^C mutually exclusive!



$$(E \cup F)^C = E^C \cap F^C$$

$$\left(\bigcup_{i=1}^{n} E_i\right)^C = \bigcap_{i=1}^{n} E_i^C$$

In probability:

$$P(E_1 \cup E_2 \cup \cdots \cup E_n)$$

$$= 1 - P\left((E_1 \cup E_2 \cup \cdots \cup E_n)^C\right)$$

$$= 1 - P(E_1^c E_2^c \cdots E_n^c)$$

Great if E_i independent!

Think, then Breakout Rooms

Check out the questions on the next slide (Slide 40). **These are challenging problems.**

Post any clarifications here!

https://us.edstem.org/courses/109/discussion/27279

Think by yourself: 2 min

Breakout rooms: 5 min

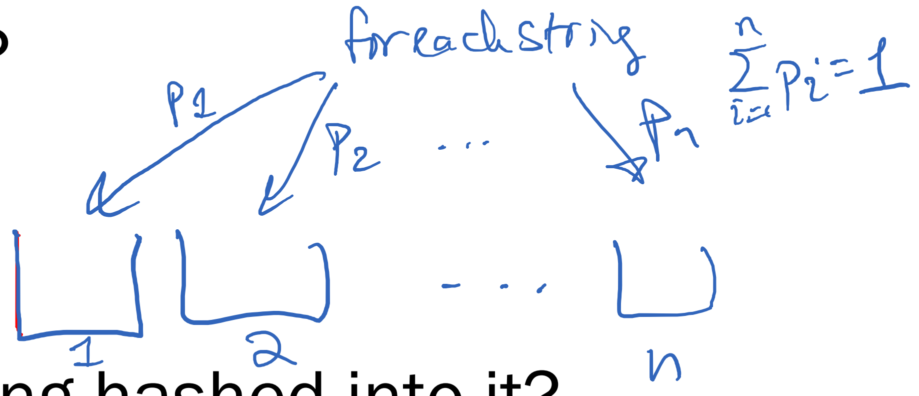
🤔

Hash table fun

- m strings are hashed (unequally) into a hash table with n buckets.
- Each string hashed is an independent trial w.p. p_i of getting hashed into bucket i.

What is $P(E)$ if

1. E = bucket 1 has ≥ 1 string hashed into it?



2. E = at least 1 of buckets 1 to k has ≥ 1 string hashed into it?

Hash table fun

- m strings are hashed (unequally) into a hash table with n buckets.
- Each string hashed is an independent trial w.p. p_i of getting hashed into bucket i.

What is $P(E)$ if

1. E = bucket 1 has $\geq$ 1 string hashed into it?

$i=1,\ldots,m$

Define S_i = string i is hashed into bucket 1
S_i^C = string i is not hashed into bucket 1

$$P(S_i) = p_1 \quad ✓$$
$$P\left(S_i^C\right) = 1 - p_1$$

Hash table fun

- m strings are hashed (unequally) into a hash table with n buckets.
- Each string hashed is an independent trial w.p. p_i of getting hashed into bucket i.

What is $P(E)$ if

1. E = bucket 1 has ≥ 1 string hashed into it?

$i = 1, \ldots, m$

Define S_i = string i is hashed into bucket 1
S_i^C = string i is not hashed into bucket 1

S_i are independent

$$P(S_i) = p_1$$
$$P(S_i^C) = 1 - p_1$$

WTF (not-real acronym for Want To Find):

$$P(E) = P(S_1 \cup S_2 \cup \cdots \cup S_m)$$

$$= 1 - P\left((S_1 \cup S_2 \cup \cdots \cup S_m)^C\right) \qquad \text{Complement}$$

$$= 1 - P\left(S_1^C S_2^C \cdots S_m^C\right) \qquad \text{De Morgan's Law}$$

$$= 1 - \overbrace{P(S_1^C)P(S_2^C)\cdots P(S_m^C)} = 1 - \left(P(S_1^C)\right)^m \qquad S_i \text{ independent trials}$$

$$= 1 - (1 - p_1)^m$$

More hash table fun: Possible approach?

- m strings are hashed (unequally) into a hash table with n buckets.
- Each string hashed is an independent trial w.p. p_i of getting hashed into bucket i.

What is $P(E)$ if

1. E = bucket 1 has $\geq$ 1 string hashed into it?
2. E = at least 1 of buckets 1 to k has $\geq$ 1 string hashed into it?

$i = 1, \dots, n$ buckets

Define F_i = bucket i has at least one string in it

$$
\begin{aligned}
P(E) &= P(F_1 \cup F_2 \cup \cdots \cup F_k) \\
&= 1 - P\left((F_1 \cup F_2 \cup \cdots \cup F_k)^C\right) \quad \text{Complement} \\
&= 1 - P\left(F_1^C F_2^C \cdots F_k^C\right) \quad \text{DeMorgan's} \\
&? = 1 - P\left(F_1^C\right) P\left(F_2^C\right) \cdots P\left(F_k^C\right)
\end{aligned}
$$

⚠ F_i bucket events are *dependent*!
So we cannot approach with complement.

$P(F_n \mid F_1^C F_2^C \cdots F_{n-1}^C) = 1$

More hash table fun

- m strings are hashed (unequally) into a hash table with n buckets.
- Each string hashed is an independent trial w.p. p_i of getting hashed into bucket i.

What is $P(E)$ if

1. E = bucket 1 has $\geq$ 1 string hashed into it?
2. E = at least 1 of buckets 1 to k has $\geq$ 1 string hashed into it?

Define F_i = bucket i has at least one string in it

$$P(E) = P(F_1 \cup F_2 \cup \cdots \cup F_k)$$

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

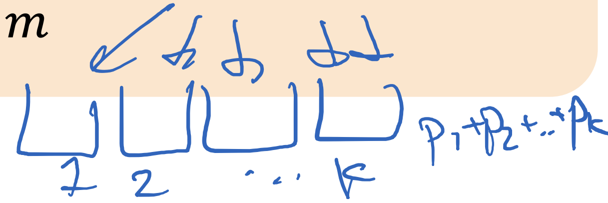
$$= 1 - P\big(F_1^C F_2^C \cdots F_k^C\big)$$

buckets 1 to k have no strings

$$= P(\text{no strings hashed to buckets 1 to } k)$$

$$= \big(P(\text{string hashed outside bkts 1 to } k)\big)^m$$

$$= (1 - p_1 - p_2 \ldots - p_k)^m$$



$$= 1 - (1 - p_1 - p_2 \ldots - p_k)^m$$

The fun never stops with hash tables

- m strings are hashed (unequally) into a hash table with n buckets.
- Each string hashed is an independent trial w.p. p_i of getting hashed into bucket i.

What is $P(E)$ if

1. E = bucket 1 has $\geq$ 1 string hashed into it?
2. E = at least 1 of buckets 1 to k has $\geq$ 1 string hashed into it?

Looking for a challenge? ☺

The fun never stops with hash tables

- m strings are hashed (unequally) into a hash table with n buckets.
- Each string hashed is an independent trial w.p. p_i of getting hashed into bucket i.

What is $P(E)$ if

1. E = bucket 1 has $\geq$ 1 string hashed into it?
2. E = at least 1 of buckets 1 to k has $\geq$ 1 string hashed into it?

3. E = each of ~~of~~ buckets 1 to k has $\geq$ 1 string hashed into it?

Hint: Use Part 2's event definition:

Define F_i = bucket i has at least one string in it

The fun never stops with hash tables

- m strings are hashed (unequally) into a hash table with n buckets.
- Each string hashed is an independent trial w.p. p_i of getting hashed into bucket i.

What is $P(E)$ if

3. E = each of ~~of~~ buckets 1 to k has ≥ 1 string hashed into it?

Define F_i = bucket i has at least one string in it

WTF: $$P(E) = P(F_1 F_2 \cdots F_k)$$

$$= 1 - P\left((F_1 F_2 \cdots F_k)^C\right)$$ Complement

$$= 1 - P\left(\underbrace{F_1^C \cup F_2^C \cup \cdots \cup}_{} F_k^C\right)$$ De Morgan's Law

general principle of inclusion-exclusion

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

where $P\left(F_{i_1}^c F_{i_2}^c \dots F_{i_r}^c\right) = (1 - p_{i_1} - p_{i_2} \dots - p_{i_r})^m$