

05: Independence

Jerry Cain

April 6, 2022

Quick slide reference

- 3 Independence
- 9 Independent Trials
- 16 Exercises and deMorgan's Laws



Independence I

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

Statement:

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

Proof:

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

Definition of conditional probability

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

Independence of E and F

$$= P(E)$$

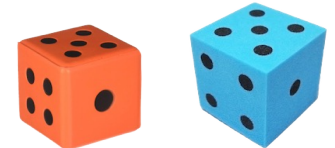
Taking the bus to cancellation city

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

Dice, our misunderstood friends

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

- 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 : $D_1 + D_2 = 5$



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

1. Are E and F independent?

$$P(E) = 1/6$$

$$P(F) = 1/6$$

$$P(EF) = 1/36$$

✓ independent

2. Are E and G independent?

$$P(E) = 1/6$$

$$P(G) = 4/36 = 1/9$$

$$P(EG) = 1/36 \neq P(E)P(G)$$

✗ dependent

Generalizing independence

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

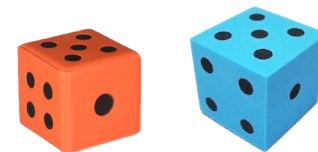
$$\left\{ \begin{array}{l} P(EFG) = P(E)P(F)P(G), \text{ and} \\ P(EF) = P(E)P(F), \text{ and} \\ P(EG) = P(E)P(G), \text{ and} \\ P(FG) = P(F)P(G) \end{array} \right.$$

n events E_1, E_2, \dots, E_n are independent if:

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

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 : $D_1 + D_2 = 7$



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

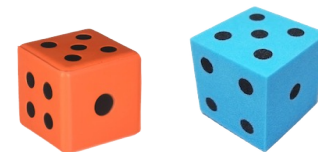
1. Are E and F independent?
2. Are E and G independent?
3. Are F and G independent?
4. Are E, F, G independent?

$$P(EF) = 1/36$$



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?

$$P(EF) = 1/36$$

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



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:

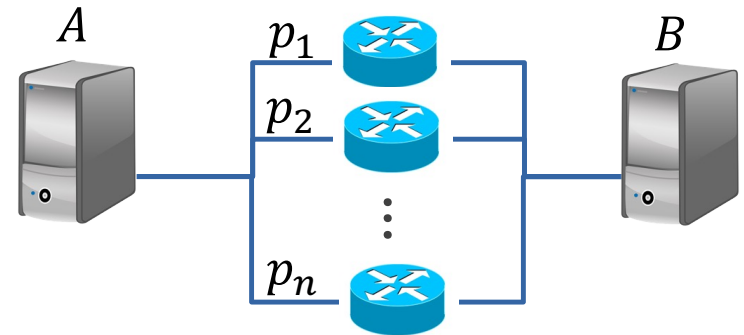
- Flip a coin n times
- Roll a die n times
- Send a multiple-choice survey to n people
- Send n web requests to k different servers

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.

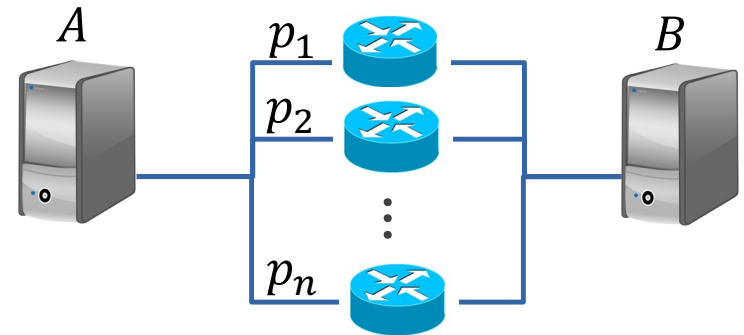
What is $P(E)$?



Network reliability

Consider the following parallel network:

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



What is $P(E)$?

$$\begin{aligned} P(E) &= P(\geq 1 \text{ one router works}) \\ &= 1 - P(\text{all routers fail}) \\ &= 1 - (1 - p_1)(1 - p_2) \cdots (1 - p_n) \\ &= 1 - \prod_{i=1}^n (1 - p_i) \end{aligned}$$

≥ 1 with independent trials:
take complement

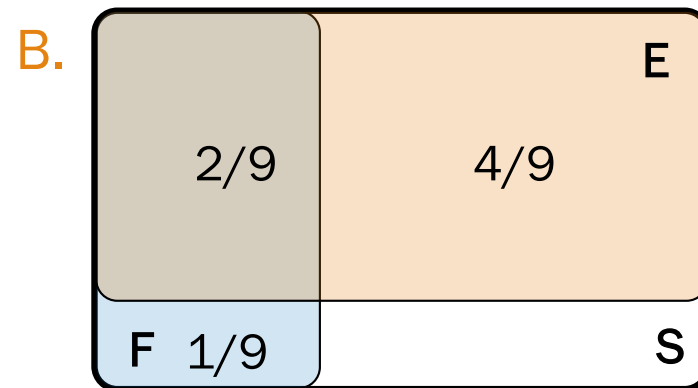
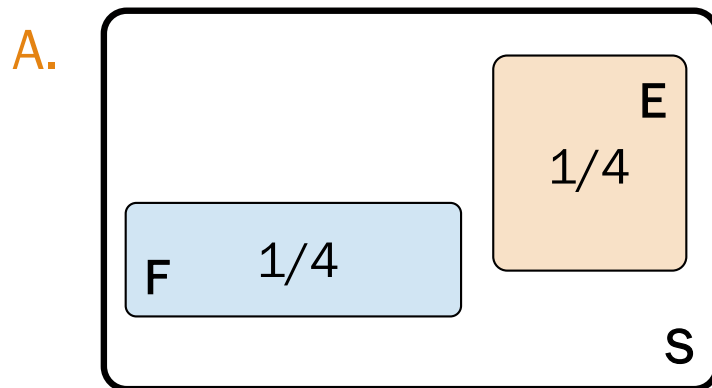


Exercises

Independence?

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

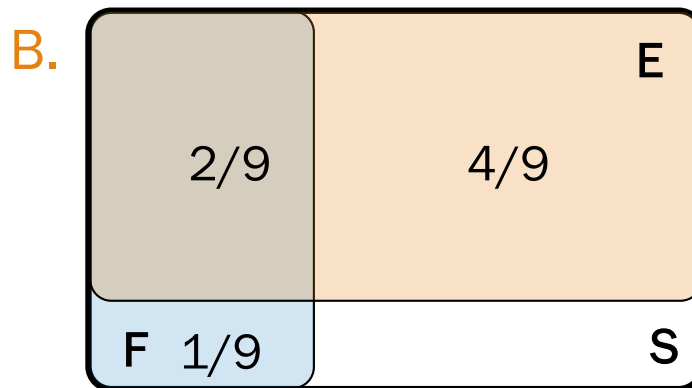
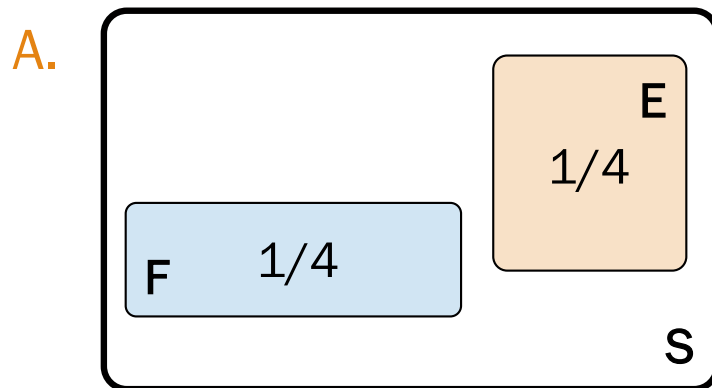
1. True or False? Two events E and F are independent if:
 - A. Knowing that F happens means that E can't happen.
 - B. Knowing that F happens doesn't change probability that E happened.
2. Are E and F independent in the following pictures?



Independence?

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

1. True or False? 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?



Be careful:

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

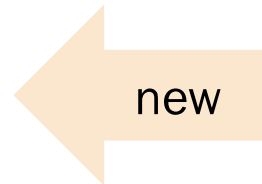
Independence

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.



Independence of complements

Statement:

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

Proof:

$$\begin{aligned}P(EF^C) &= P(E) - P(EF) \\ &= P(E) - P(E)P(F) \\ &= P(E)[1 - P(F)] \\ &= P(E)P(F^C)\end{aligned}$$

E and F^C are independent

Intersection

Independence of E and F

Factoring

Complement

Definition of independence

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

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



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

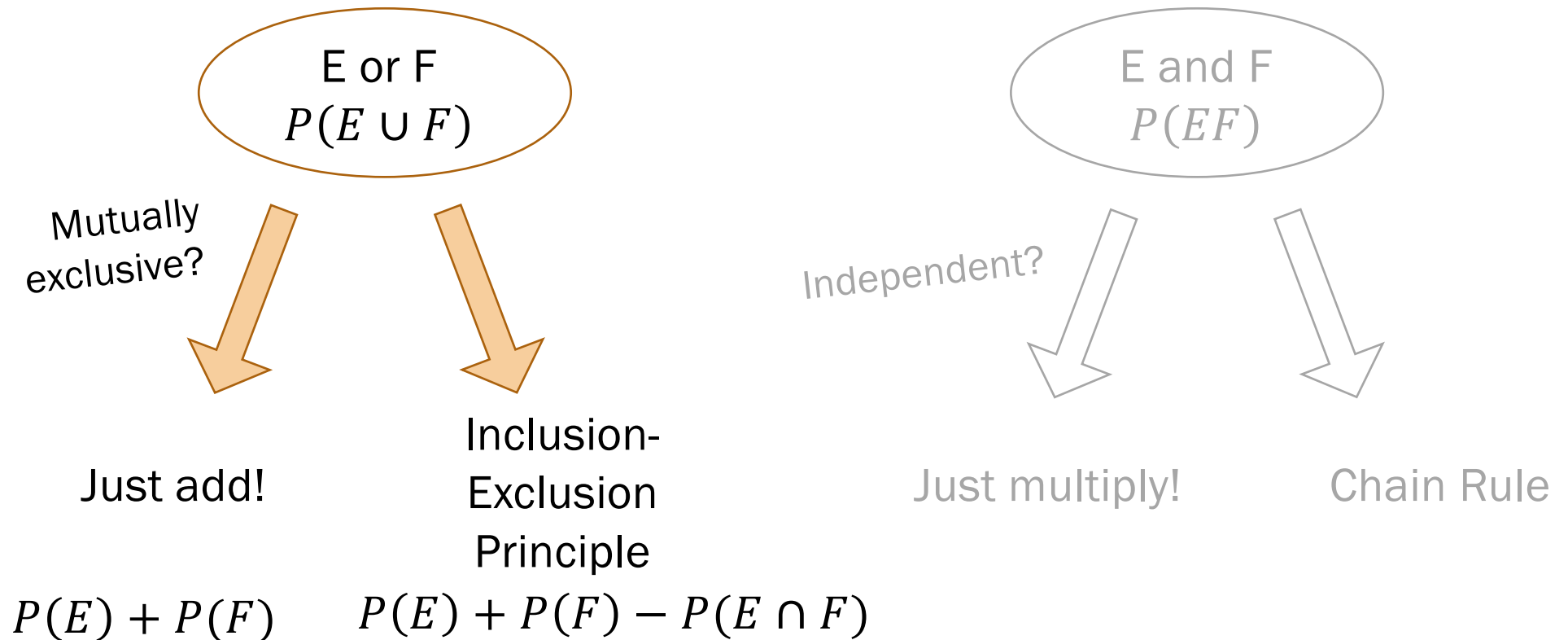
$$\binom{n}{k} p^k (1 - p)^{n-k}$$

of mutually
exclusive
outcomes

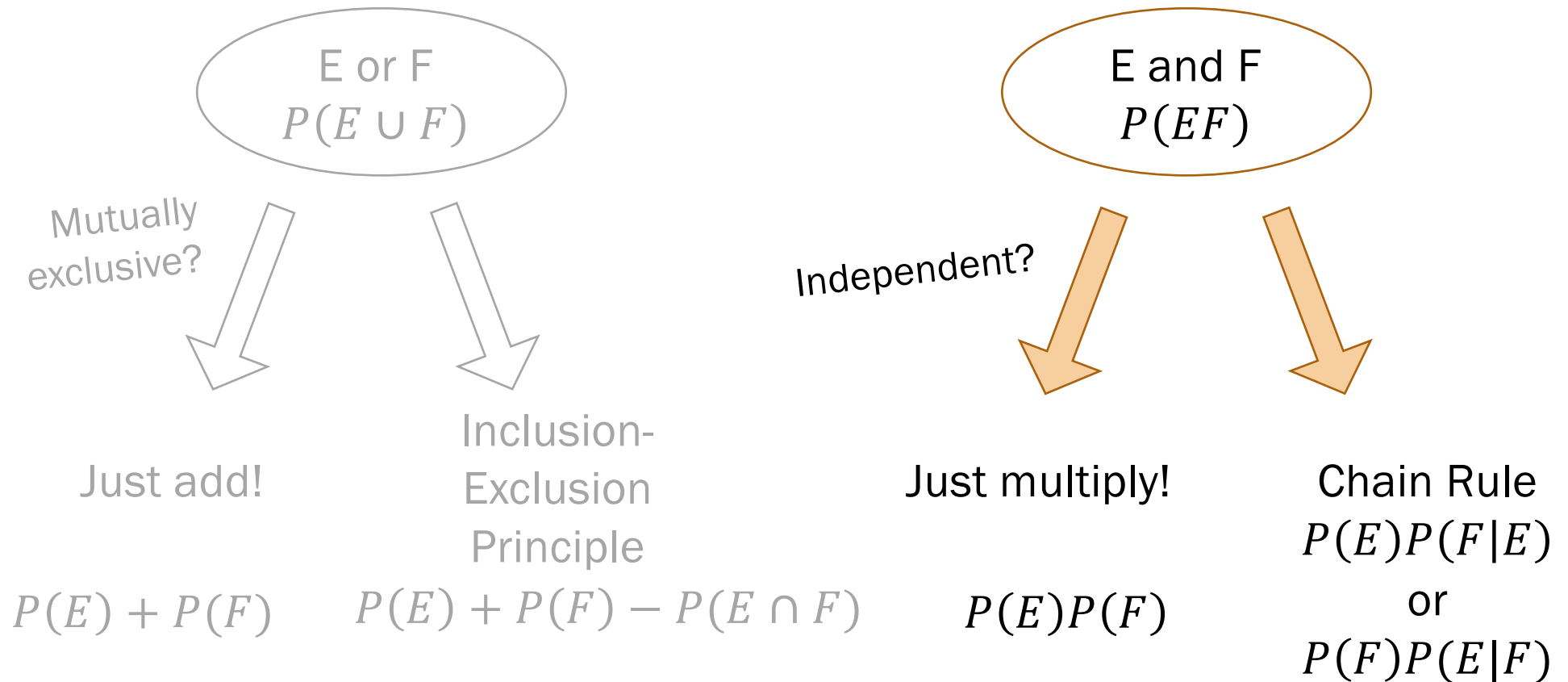
$P(\text{a particular outcome's}$
 $k \text{ heads on } n \text{ coin flips})$

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

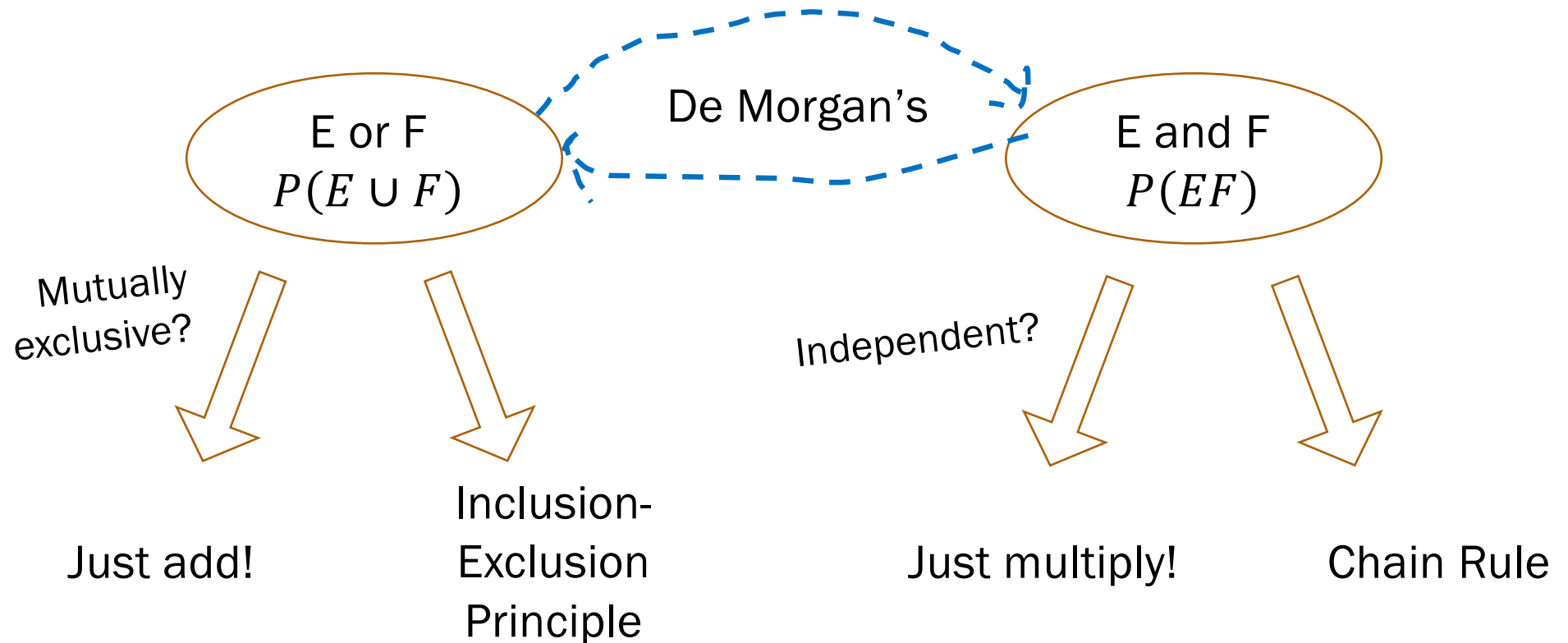
Probability of events



Probability of events

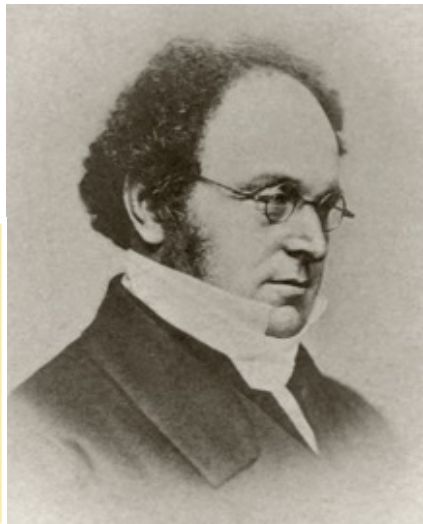
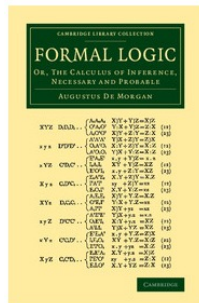


Probability of events



Augustus De Morgan

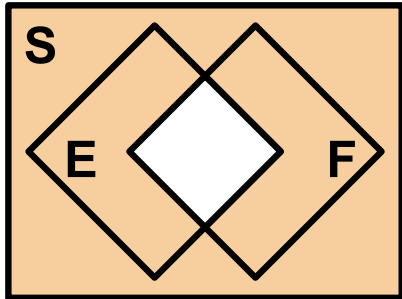
Augustus De Morgan (1806–1871):
British mathematician who wrote the book *Formal Logic* (1847).



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 between AND and OR.



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

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

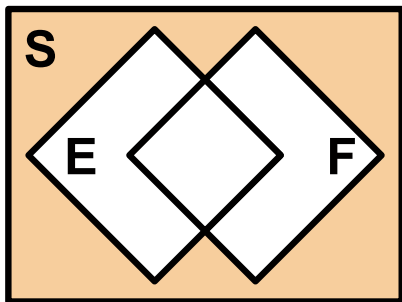
In probability:

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

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

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

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!

CS109 Mixer

Check out the questions on the next slide (Slide 27). **These are challenging problems.** Speak with your neighbor about how to best approach these problems.



Hash table fun

- m strings are hashed (not uniformly) 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 (not uniformly) 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?

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

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

Hash table fun

- m strings are hashed (not uniformly) 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?

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

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

$$= 1 - P\left((S_1 \cup S_2 \cup \dots \cup S_m)^C\right)$$

$$= 1 - P(S_1^C S_2^C \dots S_m^C)$$

$$= 1 - P(S_1^C)P(S_2^C) \dots P(S_m^C) = 1 - \left(P(S_1^C)\right)^m$$

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

Define

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

Complement

De Morgan's Law

S_i independent trials

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

More hash table fun: Possible approach?

- m strings are hashed (not uniformly) 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?

$$\begin{aligned}P(E) &= P(F_1 \cup F_2 \cup \dots \cup F_k) \\&= 1 - P\left((F_1 \cup F_2 \cup \dots \cup F_k)^c\right) \\&= 1 - P(F_1^c F_2^c \dots F_k^c) \\&? = 1 - P(F_1^c)P(F_2^c) \dots P(F_k^c)\end{aligned}$$

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

 F_i bucket events are *dependent*!

So we cannot approach with complement.

More hash table fun



- m strings are hashed (not uniformly) 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?

$$\begin{aligned} P(E) &= P(F_1 \cup F_2 \cup \dots \cup F_k) \\ &= 1 - P\left((F_1 \cup F_2 \cup \dots \cup F_k)^C\right) \\ &= 1 - P(F_1^C F_2^C \dots F_k^C) \end{aligned}$$

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

$$\begin{aligned} &= P(\text{buckets 1 to } k \text{ all denied strings}) \\ &= (P(\text{each string hashes to } k + 1 \text{ or higher}))^m \\ &= (1 - p_1 - p_2 \dots - p_k)^m \end{aligned}$$

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

The fun never stops with hash tables

- m strings are hashed (not uniformly) 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? 

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



Hint: Use Part 2's event definition:

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