05: Independence

David Varodayan

January 15, 2020

Adapted from slides by Lisa Yan

Monty Hall Problem from Let's Make a Deal

Behind one door is a car (equally likely to be any door).

Behind the other two doors are goats

1. You choose a door
2. Host opens 1 of other 2 doors, revealing a goat
3. You are given an option to change to the other door.

Should you switch?

Doors A, B, C

What happens if you switch

You picked:

1/3 → A = prize, 1/3 → B = prize, 1/3 → C = prize

A = prize

- Host opens
- You switch to
- Result

P(win | A prize, you picked ___, switch) =

B = prize

- Host opens
- You switch to
- Result

P(win | B prize, you picked ___, switch) =

C = prize

- Host opens
- You switch to
- Result

P(win | C prize, you picked ___, switch) =

P(win | you picked ___, switched) =

Monty Hall, 1000 envelope version

Start with 1000 envelopes (of which 1 is the prize).

1. You choose 1 envelope.

   $\frac{1}{1000}$ = P(envelope is prize)

   $\frac{999}{1000}$ = P(other 999 envelopes have prize)

2. I open 998 of remaining 999 (showing they are empty).

   $\frac{999}{1000}$ = P(998 empty envelopes had prize) + P(last other envelope has prize)

   = P(last other envelope has prize)

3. Should you switch?

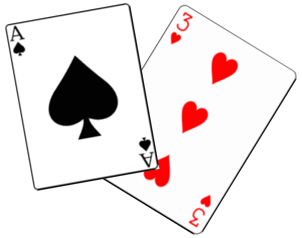
   P(you win without switching) = $\frac{1}{\text{original \# envelopes}}$

   P(you win with switching) = $\frac{\text{original \# envelopes - 1}}{\text{original \# envelopes}}$

This class going forward

<u>Last week</u>

Equally likely events



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

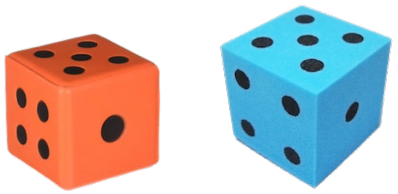
(counting, combinatorics)

<u>For most of this course</u>

Not equally likely events

$P\left(\begin{array}{c} E \text{ given some evidence} \\ \text{has been observed} \end{array}\right)$

Two Dice

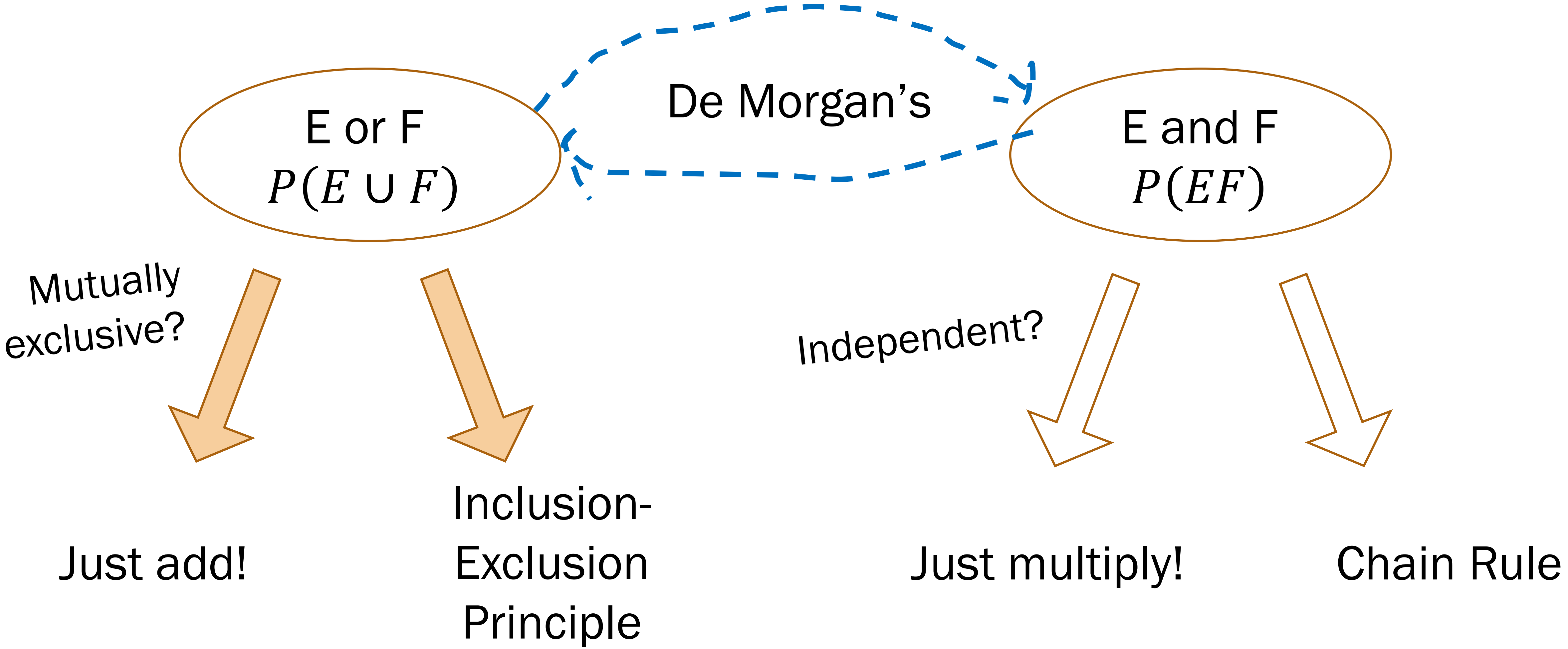


- Roll two 6-sided dice, yielding values D_1 and D_2.
- Let event E: $D_1 = 5$
  event F: $D_2 = 5$

1. Roll a 5 on one of the rolls or both
2. Roll a 5 on both rolls
3. Neither roll is 5
4. Roll a 5 on roll 2
5. Do not roll a 5 on one of the rolls or both

A. $P(F)$
B. $P(E \cup F)$
C. $P(E^C \cup F^C)$
D. $P(EF)$
E. $P(E^C F^C)$

Probability of events



Inclusion-Exclusion

- P(student programs in Java) = 0.28 $P(E)$
- P(student programs in Python) = 0.07 $P(F)$
- P(student programs in Java and Python) = 0.05. $P(E \cap F) = P(EF)$

What is P(student does not program in (Java or Python))?

1. Define events & state goal
2. Identify known probabilities
3. Solve

Let: E: Student programs in Java
F: Student programs in Python

Want: $P\left((E \cup F)^C\right)$

$$
\begin{aligned}
P\left((E \cup F)^C\right) &= 1 - P(E \cup F) \\
&= 1 - [P(E) + P(F) - P(E \cap F)] \\
&= 1 - [0.28 + 0.07 - 0.05] \\
&= 0.70
\end{aligned}
$$

Chain Rule

Definition of conditional probability:

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

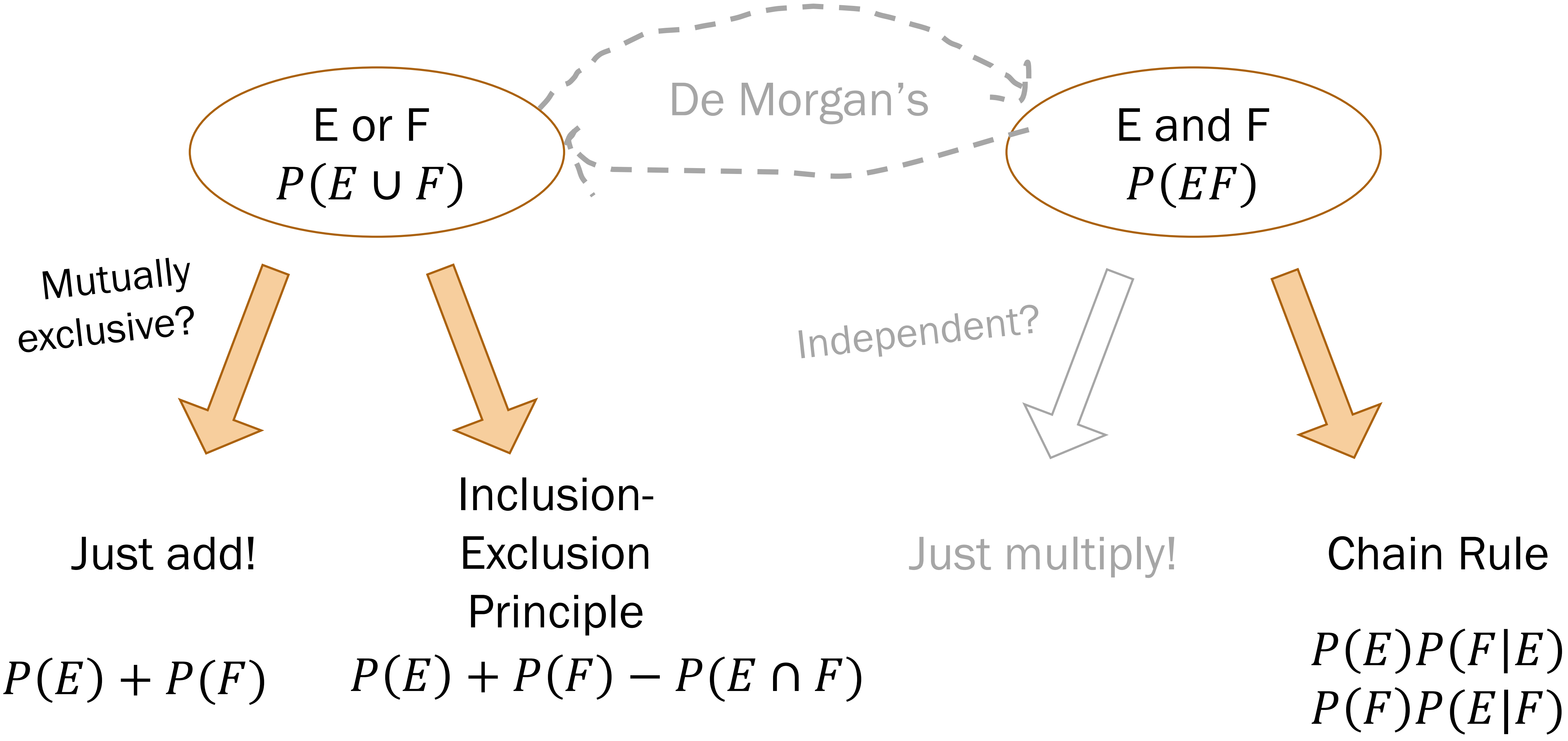
The Chain Rule:

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

Generalized Chain Rule

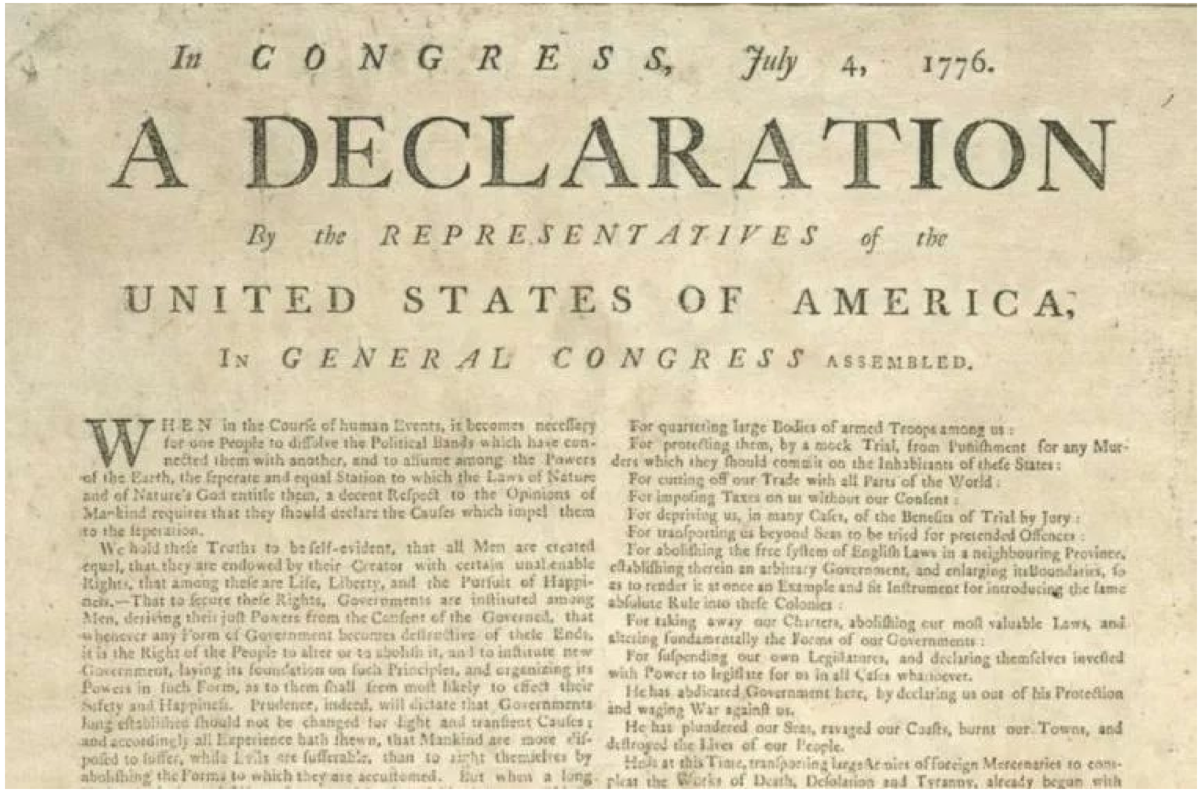
$$P(E_1 E_2 E_3 \dots E_n) = P(E_1)P(E_2|E_1)P(E_3|E_1 E_2) \dots P(E_n|E_1 E_2 \dots E_{n-1})$$

Probability of events



Today's plan

- ➡ Independence
- Independent trials
- De Morgan's Laws
- Conditional independence (if time)



In CONGRESS, July 4, 1776.

A DECLARATION

By the REPRESENTATIVES of the

UNITED STATES OF AMERICA,

In GENERAL CONGRESS ASSEMBLED.

WHEN in the Courſe of human Events, it becomes neceſſary for one People to diſſolve the Political Bands which have connected them with another, and to aſſume among the Powers of the Earth, the ſeperate and equal Station to which the Laws of Nature and of Nature's God entitle them, a decent Reſpect to the Opinions of Mankind requires that they ſhould declare the Cauſes which impel them to the ſeparation.

We hold theſe Truths to be ſelf-evident, that all Men are created equal, that they are endowed by their Creator with certain unalienable Rights, that among theſe are Life, Liberty, and the Purſuit of Happineſs.—That to ſecure theſe Rights, Governments are inſtituted among Men, deriving their juſt Powers from the Conſent of the Governed, that whenever any Form of Government becomes deſtructive of theſe Ends, it is the Right of the People to alter or to aboliſh it, and to inſtitute new Government, laying its Foundation on ſuch Principles, and organizing its Powers in ſuch Form, as to them ſhall ſeem moſt likely to effect their Safety and Happineſs. Prudence, indeed, will dictate that Governments long eſtabliſhed ſhould not be changed for light and tranſient Cauſes; and accordingly all Experience hath ſhewn, that Mankind are more diſpoſed to ſuffer, while Evils are ſufferable, than to right themſelves by aboliſhing the Forms to which they are accuſtomed. But when a long

For quartering large Bodies of armed Troops among us:

For protecting them, by a mock Trial, from Puniſhment for any Murders which they ſhould commit on the Inhabitants of theſe States:

For cutting off our Trade with all Parts of the World:

For impoſing Taxes on us without our Conſent:

For depriving us, in many Caſes, of the Benefits of Trial by Jury:

For tranſporting us beyond Seas to be tried for pretended Offences:

For aboliſhing the free Syſtem of Engliſh Laws in a neighbouring Province, eſtabliſhing therein an arbitrary Government, and enlarging its Boundaries, ſo as to render it at once an Example and fit Inſtrument for introducing the ſame abſolute Rule into theſe Colonies:

For taking away our Charters, aboliſhing our moſt valuable Laws, and altering fundamentally the Forms of our Governments:

For ſuſpending our own Legiſlatures, and declaring themſelves inveſted with Power to legiſlate for us in all Caſes whatſoever.

He has abdicated Government here, by declaring us out of his Protection and waging War againſt us.

He has plundered our Seas, ravaged our Coaſts, burnt our Towns, and deſtroyed the Lives of our People.

He is at this Time, tranſporting large Armies of foreign Mercenaries to compleat the Works of Death, Deſolation and Tyranny, already begun with

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.

An equivalent definition:

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

Intuition through proof

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

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

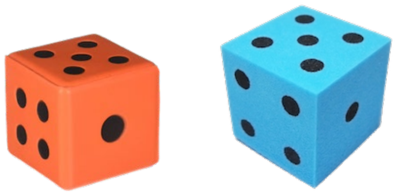
$$= P(E)$$

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: $D_1 + D_2 = 5$

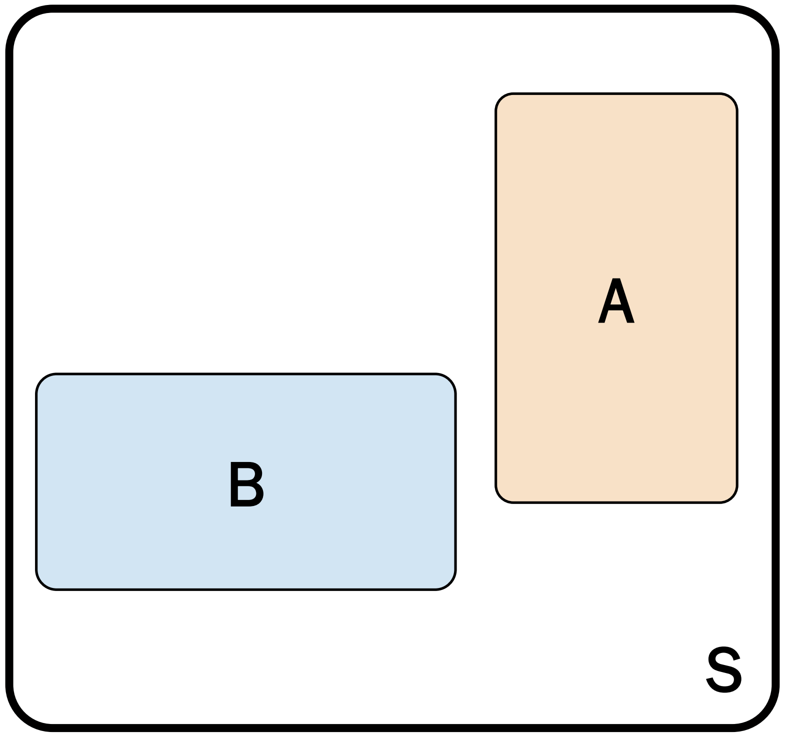


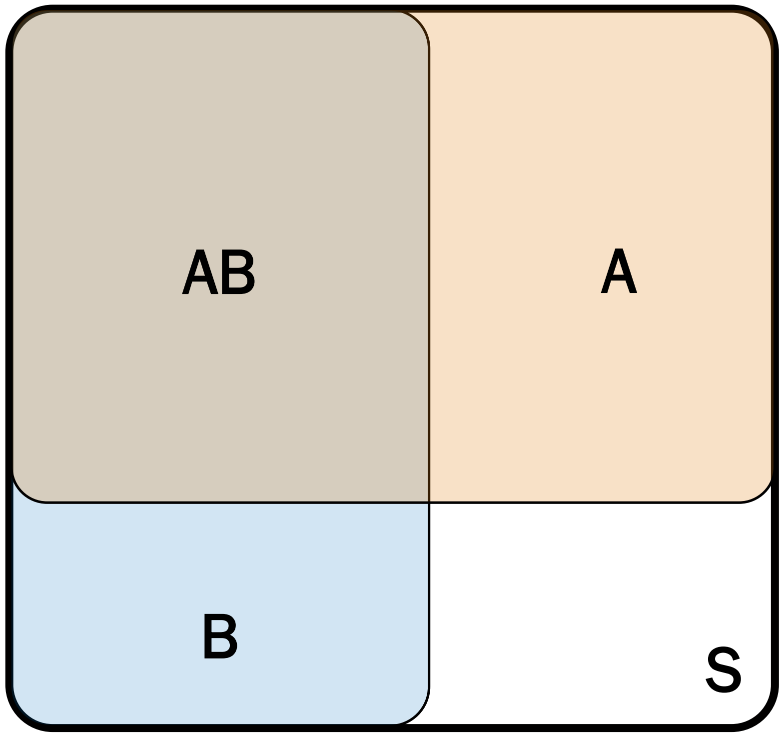
$G = \{(1,4), (2,3), (3,2), (4,1)\}$

1. Are E and F independent?
2. Are E and G independent?

Independence?

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





Independence of complements

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

Statement:

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

Proof:

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

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

Today's plan

Independence

Independent trials

De Morgan's Laws

Conditional independence (if time)

Generalizing independence

Three events E, F, and 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
$P(FG) = P(F)P(G)$

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

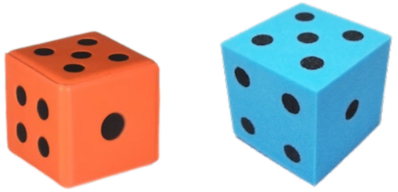
for $r = 1, \ldots, n$:
for every subset $E_1, E_2, \ldots, E_r$:
$P(E_1, E_2, \ldots, E_r) = P(E_1)P(E_2) \cdots P(E_r)$

Independent trials:

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?

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

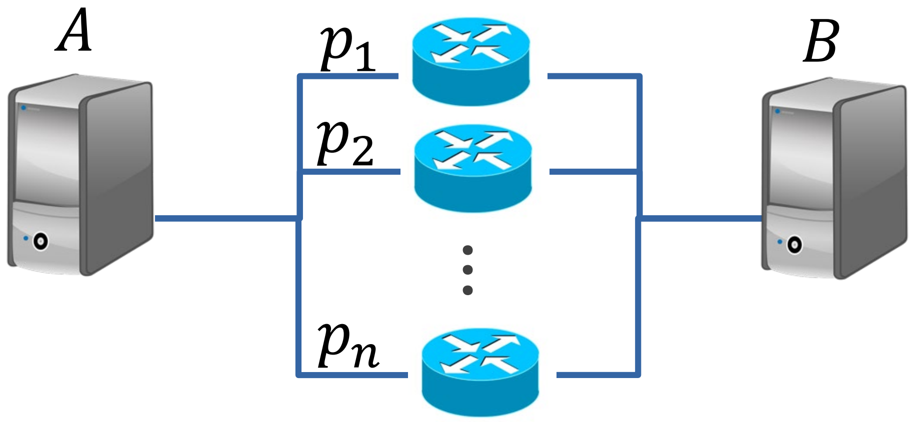
Pairwise independence is not sufficient to prove independence of >2 events!

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.

What is $P(E)$?



(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)
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)

Announcements

<u>Section</u>

Starts:	today
Late signups/changes:	by end of day
Solutions:	end of week

<u>This quarter</u>

Beginning:	fast-paced
Later:	deep into concepts
Counting:	the hardest part!

Today's plan

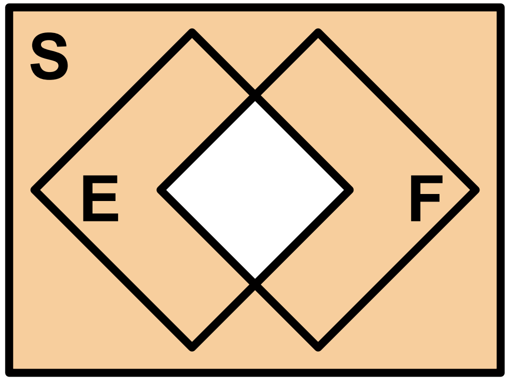
Independence

Independent trials

De Morgan's Laws

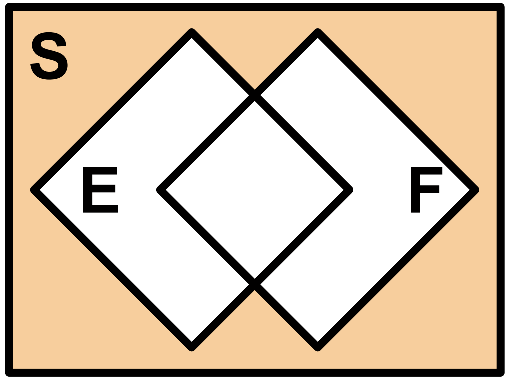
Conditional independence

De Morgan's Laws



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

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



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

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

In probability:

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

Great if E_i^C mutually exclusive!

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

Great if E_i independent!

De Morgan's: AND ⟷ OR

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.

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

What is $P(E)$?

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 (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.

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

What is $P(E)$?

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

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

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

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

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

$$= 1 - P\left(S_1^C\right)P\left(S_2^C\right) \cdots P\left(S_m^C\right)$$ S_i independent trials

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.

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.

What is $P(E)$?

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

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

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.

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 buckets 1 to k has $\geq$ 1 string hashed into it.

What is $P(E)$?

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.

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 buckets 1 to k has $\geq$ 1 string hashed into it.

What is $P(E)$?

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

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

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

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

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

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

Today's plan

Independence

Independent trials

De Morgan's Laws

Conditional independence (if time)

Conditional Paradigm

For any events A, B, and E, you can condition consistently on E, and all formulas still hold:

Axiom 1	$0 \leq P(A\|E) \leq 1$
Corollary 1 (complement)	$P(A\|E) = 1 - P(A^C\|E)$
Commutativity	$P(AB\|E) = P(BA\|E)$
Chain Rule	$P(AB\|E) = P(B\|E)P(A\|BE)$
Bayes' Theorem	$P(A\|BE) = \frac{P(B\|AE)P(A\|E)}{P(B\|E)}$

Independence relationships can change with conditioning.

A and B independent does NOT always mean A and B independent given E.

Conditional Independence

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

Two events A and B are defined as conditionally independent given E if:

$$P(AB|E) = P(A|E)P(B|E)$$

An equivalent definition:

A. $P(A|B) = P(A)$
B. $P(A|BE) = P(A)$
C. $P(A|BE) = P(A|E)$

Netflix and Condition

Let E = a user watches Life is Beautiful.

Let F = a user watches Amelie.

What is $P(E)$?

$$P(E) \approx \frac{\text{\# people who have watched movie}}{\text{\# people on Netflix}} = \frac{10{,}234{,}231}{50{,}923{,}123} \approx 0.20$$

What is the probability that a user watches Life is Beautiful, given they watched Amelie?

$$P(E|F) = \frac{P(EF)}{P(F)} = \frac{\text{\# people who have watched both}}{\text{\# people who have watched Amelie}} \approx 0.42$$

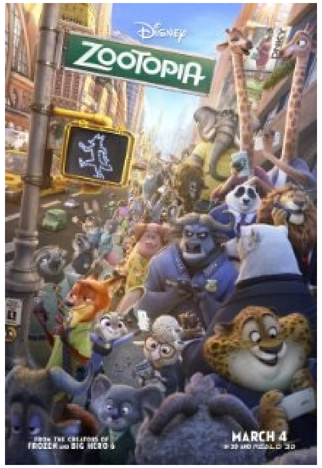
Netflix and Condition

Let E be the event that a user watches the given movie.

Let F be the event that the same user watches Amelie.

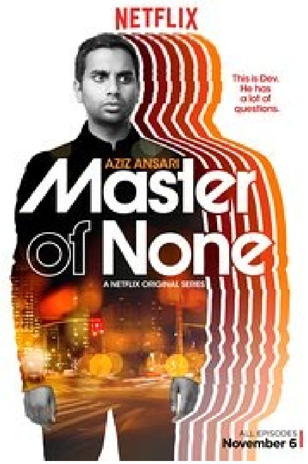
$P(E) = 0.19$

$P(E|F) = 0.14$



$P(E) = 0.32$

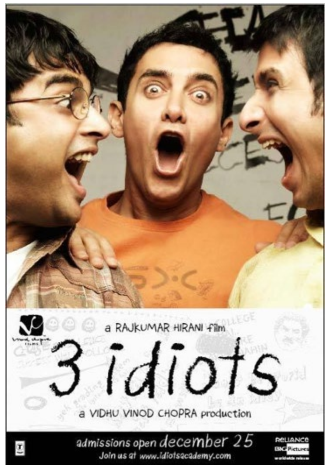
$P(E|F) = 0.35$



$P(E) = 0.20$

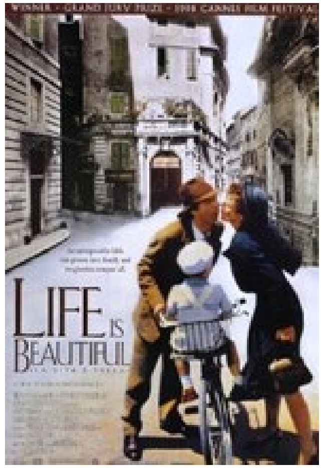
$P(E|F) = 0.20$

Independent!



$P(E) = 0.09$

$P(E|F) = 0.72$



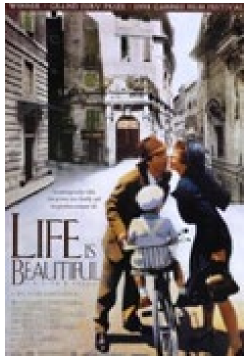
$P(E) = 0.20$

$P(E|F) = 0.42$

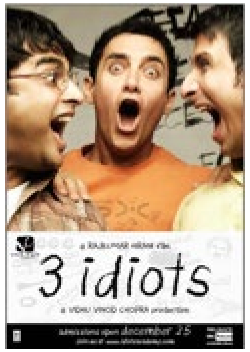
Netflix and Condition

Watched:

E_1

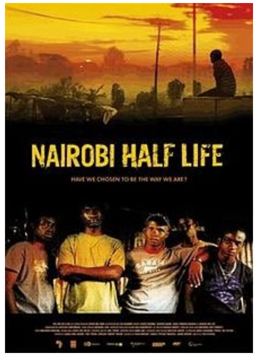


E_2



E_3

Will they watch?



E_4

What if $E_1 E_2 E_3 E_4$ are not independent? (e.g., all international emotional comedies)

$$P(E_4 | E_1 E_2 E_3) = \frac{P(E_1 E_2 E_3 E_4)}{P(E_1 E_2 E_3)}$$

$P(E_1 E_2 E_3 E_4)$ → $\frac{\text{\# people who have watched all 4}}{\text{\# people on Netflix}}$

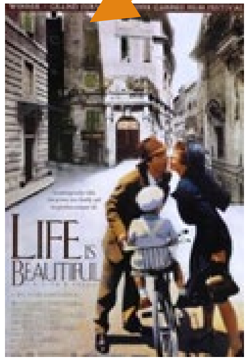
$P(E_1 E_2 E_3)$ → $\frac{\text{\# people who have watched those 3}}{\text{\# people on Netflix}}$

Netflix and Condition

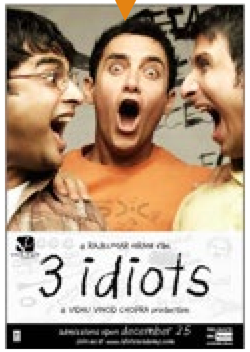
K: likes international emotional comedies

Watched:

E_1

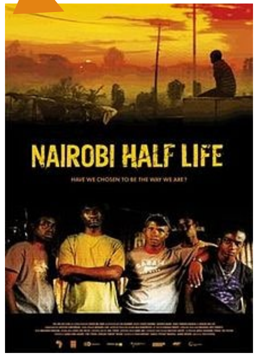


E_2



E_3

Will they watch?



E_4

What if $E_1E_2E_3E_4$ are *conditionally independent* K?

$$P(E_4|E_1E_2E_3) = \frac{P(E_1E_2E_3E_4)}{P(E_1E_2E_3)}$$

$$P(E_4|E_1E_2E_3K) = P(E_4|K)$$

Conditional independence is a Big Deal

Conditional independence is a practical, real-world way of decomposing hard probability questions.

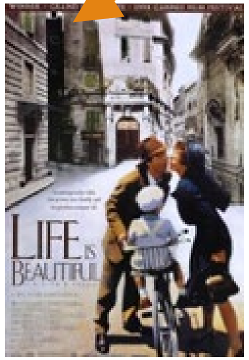
> "Exploiting conditional independence to generate fast probabilistic computations is one of the main contributions CS has made to probability theory."

–Judea Pearl wins 2011 Turing Award,
"For fundamental contributions to artificial intelligence through the development of a calculus for probabilistic and causal reasoning"

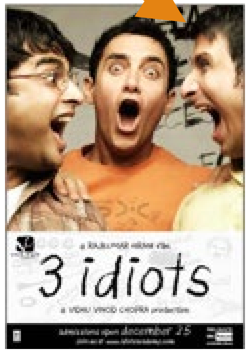
Netflix and Condition

K: likes international emotional comedies

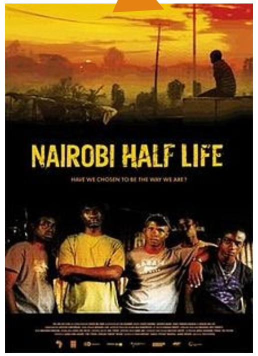
E_1



E_2



E_3



E_4

$E_1 E_2 E_3 E_4$ are dependent

$E_1 E_2 E_3 E_4$ are conditionally independent given K

Dependent events can become conditionally independent.