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?
What happens if you switch

You picked: C

A = prize
- Host opens B
- You switch to A
- Result Win

P(win | A prize, you picked C, switch) = 1

B = prize
- Host opens A
- You switch to B
- Result Win

P(win | B prize, you picked C, switch) = 1

C = prize
- Host opens A or B
- You switch to B or A
- Result Lose

P(win | C prize, you picked C, switch) = 0

P(win | you picked C, switched) = \frac{1}{3} \cdot 1 + \frac{1}{3} \cdot 1 + \frac{1}{3} \cdot 0 = \frac{2}{3}
Monty Hall, 1000 envelope version

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

1. You choose 1 envelope.
   \[
   \frac{1}{1000} = \text{P(envelope is prize)} \\
   \frac{999}{1000} = \text{P(other 999 envelopes have prize)}
   \]

2. I open 998 of remaining 999 (showing they are empty).
   \[
   \frac{999}{1000} = \text{P(998 empty envelopes had prize)} \\
   + \text{P(last other envelope has prize)}
   \]
   \[
   = \text{P(last other envelope has prize)}
   \]

3. Should you switch?
   \[
   \text{P(you win without switching)} = \frac{1}{\text{original # envelopes}}
   \]
   \[
   \text{P(you win with switching)} = \frac{1}{\text{original # envelopes} - 1}
   \]
This class going forward

Last week
Equally likely events

\[ P(E \cap F) \quad P(E \cup F) \]
(counting, combinatorics)

For most of this course
Not equally likely events

\[ P(E \text{ given some evidence has been observed}) \]
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

- **E or F**
  - $P(E \cup F)$
  - Mutually exclusive?
  - Just add!

- **De Morgan’s**
  - Independent?

- **E and F**
  - $P(EF)$
  - Just multiply!
  - Chain Rule
Inclusion-Exclusion

- \( P(\text{student programs in Java}) = 0.28 \)
- \( P(\text{student programs in Python}) = 0.07 \)
- \( P(\text{student programs in Java and Python}) = 0.05 \).

What is \( P(\text{student does not program in (Java or Python)})? \)

1. Define events & state goal
   
   Let: 
   
   \( E: \) Student programs in Java 
   
   \( F: \) Student programs in Python 
   
   Want: \( P\left( (E \cup F)^C \right) \)

2. Identify known probabilities
   
   \[
   P\left( (E \cup F)^C \right) = 1 - P(E \cup F) \\
   = 1 - [P(E) + P(F) - P(E \cap F)] \\
   = 1 - [0.28 + 0.07 - 0.05] \\
   = 0.70
   \]

3. Solve
**Chain Rule**

**Definition** of conditional probability:

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

The **Chain Rule**:

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

\[ = P(F|\epsilon)P(\epsilon) \]
Generalized Chain Rule

\[ P(E_1 E_2 E_3 \ldots E_n) = P(E_1)P(E_2|E_1)P(E_3|E_1E_2) \ldots P(E_n|E_1E_2 \ldots E_{n-1}) \]
Probability of events

- **E or F**
  \[ P(E \cup F) \]
  - Mutually exclusive?
  - Just add!
  - \[ P(E) + P(F) \]

- **E and F**
  \[ P(EF) \]
  - Independent?
  - Just multiply!
  - Chain Rule
  \[ P(E)P(F|E) \]
  \[ P(F)P(E|F) \]

**Inclusion-Exclusion Principle**
\[ P(E) + P(F) - P(E \cap F) \]
Today’s plan

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

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

- Roll two 6-sided dice, yielding values $D_1$ and $D_2$.
- Let event $E$: $D_1 = 1$ 
  event $F$: $D_2 = 6$
  event $G$: $D_1 + D_2 = 5$

\[ EF = \{(1,6)\} \]
\[ G = \{(1,4), (2,3), (3,2), (4,1)\} \]
\[ \subseteq G = \{(1,4)\} \]

1. Are $E$ and $F$ independent?

\[ P(E) = \frac{1}{6} \quad \text{YES} \]
\[ P(F) = \frac{1}{6} \]
\[ P(EF) = \frac{1}{36} = P(E)P(F) \]

2. Are $E$ and $G$ independent?

\[ P(G) = \frac{4}{36} = \frac{1}{9} \]
\[ P(E \cap G) = \frac{1}{36} \neq P(E)P(G) \]
Independence?

\[ P(A \cap B) = 0 \neq P(A)P(B) \]

Mutually Exclusive

\[ P(A) = \frac{2}{3} = P(A|B) \]

Yes

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

\[ P(E|F) = P(E) \]
Independence of complements

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:

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

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

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

**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? YES
   - $P(E) = 1/6$
   - $P(F) = 1/6$
   - $P(EF) = 1/36$

2. Are $E$ and $G$ independent? YES
   - $P(E) = 1/6$
   - $P(G) = 1/6$
   - $P(EG) = 1/36$

3. Are $F$ and $G$ independent? YES
   - $P(F) = 1/6$
   - $P(G) = 1/6$
   - $P(FG) = 1/36$

4. Are $E, F, G$ independent? NO
   - $P(EGF) = 1/36$
   - $P(E) \cdot P(F) \cdot P(G) = 1/6 \cdot 1/6 \cdot 1/6$

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 $1 \leq i \leq n$)
- $E = \text{functional path from A to B exists.}$

What is $P(E)$?

\[
P(E) = P(\text{at least one router works})
= 1 - P(\text{all routers fail})
= 1 - (1-p_1)(1-p_2) \ldots (1-p_n)
= 1 - \prod_{i=1}^{n} (1-p_i)
\]
(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. $P(n \text{ heads on } n \text{ coin flips})$
2. $P(n \text{ tails on } n \text{ coin flips})$
3. $P(\text{first } k \text{ heads, then } n - k \text{ tails})$

4. $P(\text{exactly } k \text{ heads on } n \text{ coin flips})$
   
   # ways to choose $k$ tosses out of $n$ to be heads
   
   $\binom{n}{k} p^k (1-p)^{n-k}$
Announcements

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

This quarter
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 \(\leftrightarrow\) 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 \(i\).
- \(F =\) bucket 1 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 $i$.

1. $E = \text{bucket 1}$ has $\geq 1$ string hashed into it.

What is $P(E)$?

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

Define

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

$$P(S_i) = p_1$$
$$P(S_i^C) = 1 - p_1$$
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. \( E = \text{bucket 1} \) has \( \geq 1 \) string hashed into it.

What is \( P(E) \)?

\[
P(E) = P(S_1 \cup S_2 \cup \cdots \cup S_m)
= 1 - P((S_1 \cup S_2 \cup \cdots \cup S_m)^c)
= 1 - P(S_1^c S_2^c \cdots S_m^c)
= 1 - P(S_1^c)P(S_2^c)\cdots P(S_m^c)
= 1 - (1 - p_i)^m
\]
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 $P(E)$?

$$P(E) = P(S_1 \cup S_2 \cup \cdots \cup S_m)$$
$$= 1 - P\left( (S_1 \cup S_2 \cup \cdots \cup S_m)^c \right)$$
$$= 1 - P(S_1^c S_2^c \cdots S_m^c)$$
$$= 1 - P(S_1^c)P(S_2^c)\cdots P(S_m^c)$$
$$= 1 - (1 - p_1 - p_2 - \cdots - p_k)^m$$

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

$P(S_i^c) = 1 - p_1 - p_2 - \cdots - p_k$
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. \( E = \text{bucket 1} \) has \( \geq 1 \) string hashed into it.
2. \( E = \text{at least 1 of buckets 1 to } k \) has \( \geq 1 \) string hashed into it.
3. \( E = \text{each of buckets 1 to } k \) has \( \geq 1 \) string hashed into it.

What is \( P(E) \)?

Define \( F_i = \text{bucket } i \) 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. $E = \text{bucket 1}$ has $\geq 1$ string hashed into it.
2. $E = \text{at least 1 of buckets 1 to } k \text{ has } \geq 1 \text{ string hashed into it}$.
3. $E = \text{each of buckets 1 to } k \text{ has } \geq 1 \text{ string hashed into it}$.

What is $P(E)$?

\[
P(E) = P(F_1 F_2 \cdots F_k) \\
= 1 - P\left( (F_1 F_2 \cdots F_k)^C \right) \\
= 1 - P\left( F_1^C \cup F_2^C \cup \cdots \cup F_k^C \right) \\
= 1 - P\left( \bigcup_{i=1}^{k} F_i^c \right) = 1 - \sum_{r=1}^{k} (-1)^{r+1} \sum_{i_1 < \cdots < i_r} P(F_{i_1}^c F_{i_2}^c \cdots F_{i_r}^c) \\
\text{where } P(F_{i_1}^c F_{i_2}^c \cdots F_{i_r}^c) = (1 - p_{i_1} - p_{i_2} \cdots - p_{i_r})^m
\]