05: Independence

Jerry Cain
April 10th, 2024

Lecture Discussion on Ed
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:

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

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

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

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

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

   Yes! independent

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

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

   No! dependent

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

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

\[
\begin{align*}
P(EFG) &= P(E)P(F)P(G), \\
P(EF) &= P(E)P(F), \\
P(EG) &= P(E)P(G), \\
P(FG) &= P(F)P(G) 
\end{align*}
\]

for $r = 1, \ldots, n$:

for every subset $E_1, E_2, \ldots, E_r$:

\[
P(E_1E_2\ldots E_r) = P(E_1)P(E_2)\ldots P(E_r)
\]
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 3 or more 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 \( P(E) \)?

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

\[
= 1 - P(\text{all routers fail})
= 1 - (1 - p_1)(1 - p_2) \cdots (1 - p_n)
= 1 - \prod_{i=1}^{n}(1 - p_i)
\]
Exercises
Independence?

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?

   A. ![Diagram A]
   B. ![Diagram B]
Independence?

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?
   A. 
      \[
      \begin{array}{c|c}
      & E & \\
      F & 1/4 & 1/4 \\
      \hline
      S & 1/4 & \\
      \end{array}
      \]
      \( \text{EF} = \emptyset \)
      \( P(E) = 1/4 \)
      \( P(F) = 1/4 \)
      \( P(E)P(F) = 1/16 \neq 0 \)
   B. 
      \[
      \begin{array}{c|c|c}
      & E & \\
      F & \frac{2}{9} & \frac{4}{9} \\
      \hline
      S & \frac{2}{9} & \frac{4}{9} \\
      \end{array}
      \]
      \( \text{EF} = \emptyset \)
      \( P(E) = \frac{2}{9} + \frac{4}{9} = \frac{6}{9} = \frac{2}{3} \)
      \( P(F) = \frac{2}{9} + \frac{1}{9} = \frac{3}{9} = \frac{1}{3} \)
      \( P(E)P(F) = \frac{2}{3} \cdot \frac{1}{3} = \frac{2}{9} \)
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. $P(\text{exactly } k \text{ heads on } n \text{ coin flips})$
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

E or F
\[ P(E \cup F) \]

Just add!
\[ P(E) + P(F) \]

Mutually exclusive?

Inclusion-Exclusion Principle
\[ P(E) + P(F) - P(E \cap F) \]

E and F
\[ P(EF) \]

Just multiply!

Independent?

Just add!

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!
\[ P(E)P(F) \]

Just add!
\[ P(E) + P(F) \]

Inclusion-Exclusion Principle
\[ P(E) + P(F) - P(E \cap F) \]

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**
  - $P(E \cup F)$
- **E and F**
  - $P(EF)$

**Mutually exclusive?**
- Just add!
**Independent?**
- Just multiply!

- **Inclusion-Exclusion Principle**
- **De Morgan’s**
- **Chain Rule**
De Morgan’s Laws

\[(E \cap F)^c = E^c \cup F^c\]

In probability:
\[P(E_1E_2 \cdots E_n) = 1 - P((E_1E_2 \cdots E_n)^c)\]
Great if \(E_i^c\) mutually exclusive!

\[(E \cup F)^c = E^c \cap F^c\]

In probability:
\[P(E_1 \cup E_2 \cup \cdots \cup E_n) = 1 - P((E_1 \cup E_2 \cup \cdots \cup E_n)^c)\]
Great if \(E_i\) independent!
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. $\sum_{i=1}^{n} p_i = 1$
- Each string hashed is an independent trial w.p. $p_i$ of getting hashed into bucket $i$.

What is $P(E)$ if

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

Define

- $S_i = \text{string } i \text{ is hashed into bucket 1}$
- $S_i^C = \text{string } i \text{ is not hashed into bucket 1}$

$$P(S_i) = p_1$$
$$P(S_i^C) = 1 - p_1$$
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 \( \geq 1 \) string hashed into it?

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

\[
P(E) = P(S_1 \cup S_2 \cup \cdots \cup S_m) \\
= 1 - P\left((S_1 \cup S_2 \cup \cdots \cup S_m)^c\right) \\
= 1 - P\left(S_1^c S_2^c \cdots S_m^c\right) \\
= 1 - P(S_1^c)P(S_2^c)\cdots P(S_m^c) = 1 - \left(P(S_1^c)\right)^m \\
= 1 - (1 - p_1)^m
\]

Define \( S_i = \) string \( i \) is hashed into bucket 1
\( S_i^c = \) string \( i \) is not hashed into bucket 1

\[
P(S_i) = p_1 \quad P(S_i^c) = 1 - p_1
\]

\( S_i \) independent trials
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. $E = \text{bucket 1 has } \geq 1 \text{ string hashed into it}$?
2. $E = \text{at least 1 of buckets 1 to } k \text{ has } \geq 1 \text{ string hashed into it}$?

$$P(E) = P(F_1 \cup F_2 \cup \cdots \cup F_k)$$
$$= 1 - P\left((F_1 \cup F_2 \cup \cdots \cup F_k)^c\right)$$
$$= 1 - P(F_1^c F_2^c \cdots F_k^c)$$

$?? = 1 - P(F_1^c)P(F_2^c)\cdots P(F_k^c)$

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

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

Define $F_i = \text{bucket } i \text{ has at least one string in it}$

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

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