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:

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

Intuition through proof

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

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

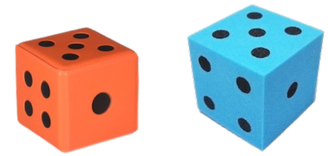
$$= P(E)$$ Taking the bus to cancellation city

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

Dice, our misunderstood friends

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

- 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

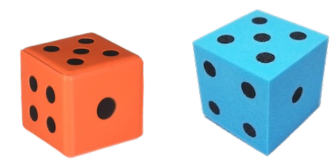
Three events E, F, and G are independent if:

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

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

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

Dice, increasingly misunderstood (still our friends)

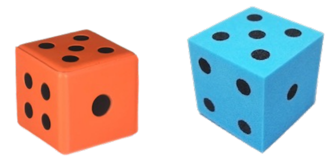


- Each roll of a 6-sided die is an independent trial.
- Two rolls: D_1 and D_2.
- Let event E: $D_1 = 1$
  event F: $D_2 = 6$
  event G: $D_1 + D_2 = 7$ $G = \{(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)\}$

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:

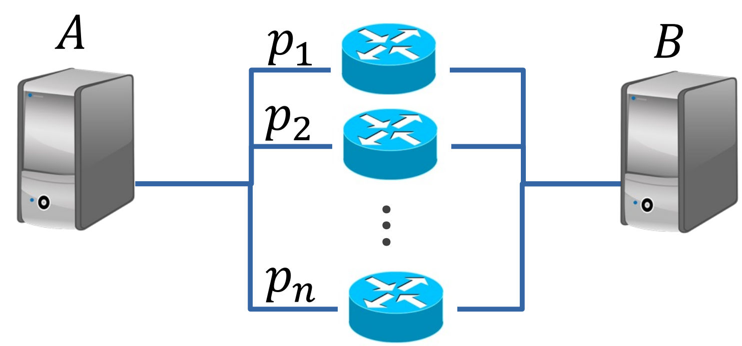
- Flip a coin n times
- Roll a die n times
- Send a multiple-choice survey to n people
- Send n web requests to k different servers

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 $P(E)$?

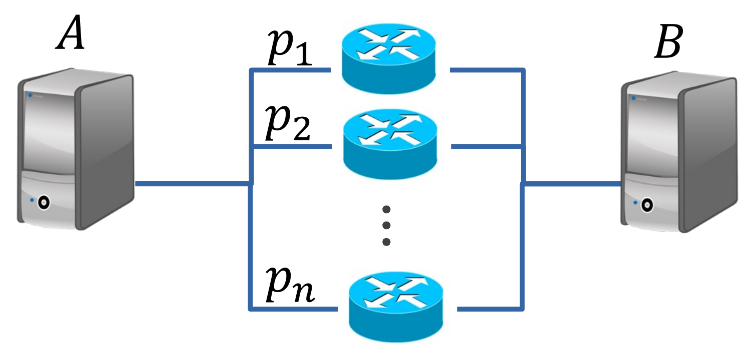


Network reliability

Consider the following parallel network:

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

What is $P(E)$?



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

≥ 1 with independent trials: take complement

Exercises

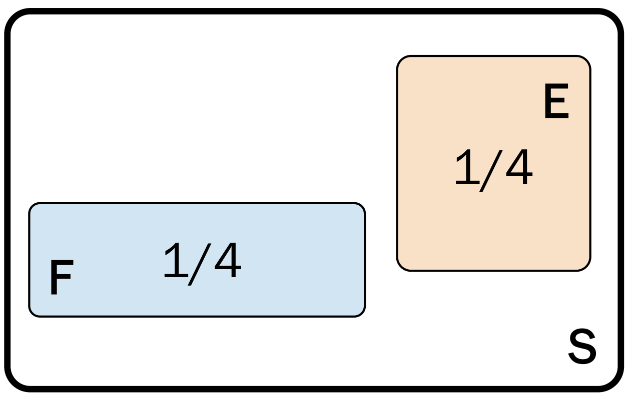
Independence?

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

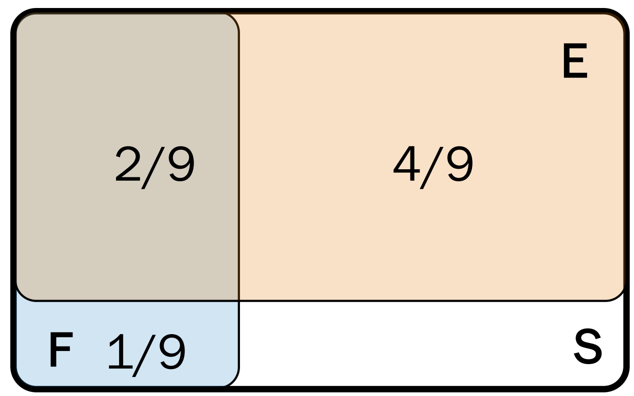
1. True or False? Two events E and F are independent if:
   A. Knowing that F happens means that E can't happen.
   B. Knowing that F happens doesn't change probability that E happened.

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

A.



B.



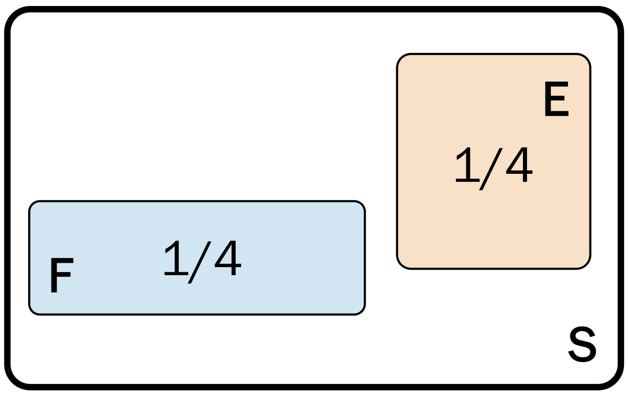
Independence?

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

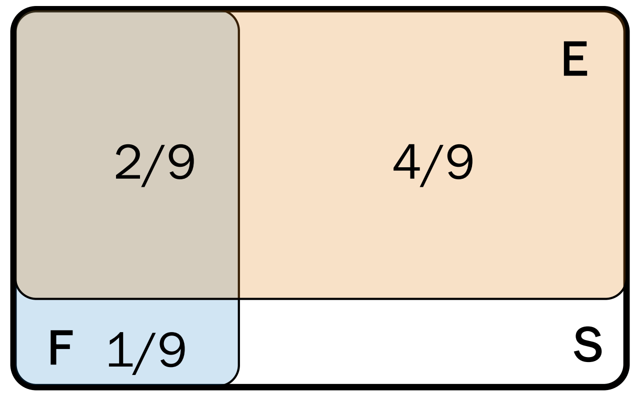
1. True or False? Two events E and F are independent if:
   A. Knowing that F happens means that E can't happen.
   B. Knowing that F happens doesn't change probability that E happened.

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

A.



B.



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$)$

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$)$

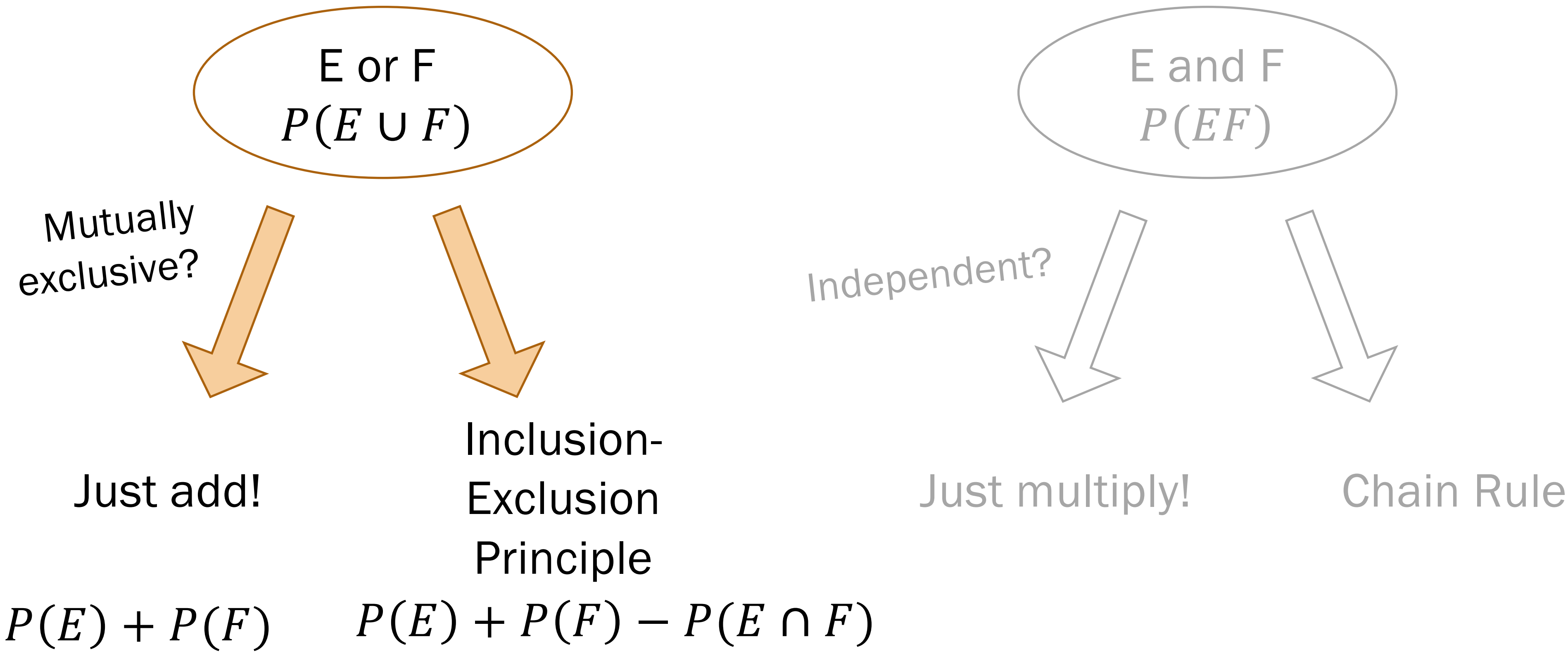
$$\binom{n}{k} p^k (1-p)^{n-k}$$

# 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.

Probability of events



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
$P(EF)$

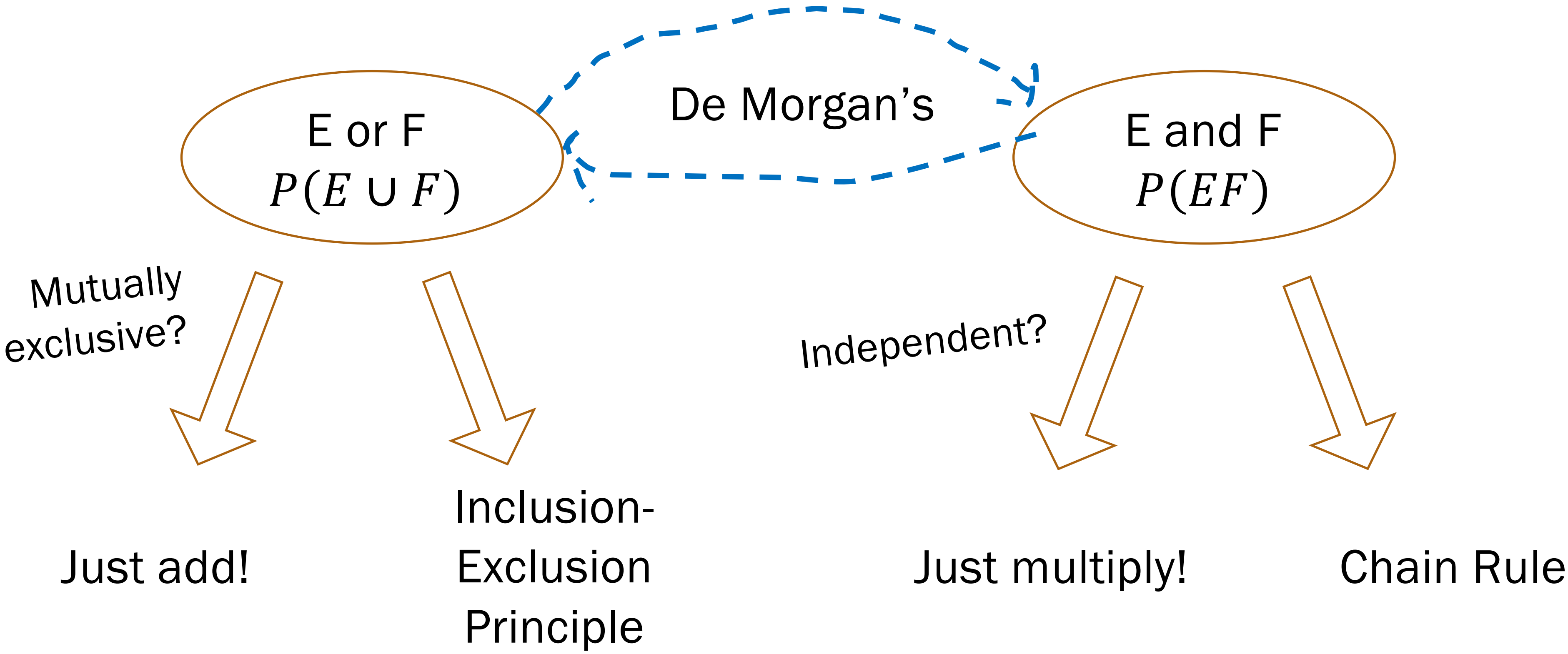
Independent?

Just multiply!

$P(E)P(F)$

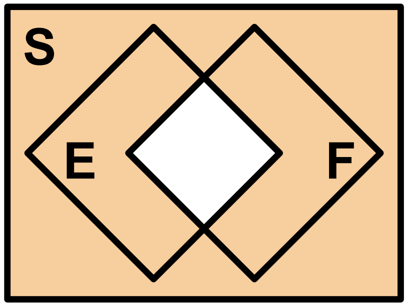
Chain Rule
$P(E)P(F|E)$
or
$P(F)P(E|F)$

Probability of events



De Morgan's Laws

De Morgan's lets you switch between AND and OR.



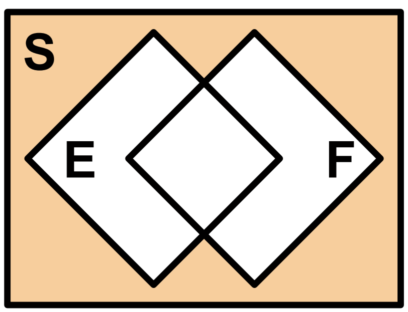
$$(E \cap F)^C = E^C \cup F^C$$

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

In probability:

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

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!

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?

2. E = at least 1 of buckets 1 to k has $\geq$ 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 $\geq$ 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 $\geq$ 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

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

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

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

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

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

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

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

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?

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

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

⚠️ F_i bucket events are *dependent*!

So we cannot approach with complement.

More hash table fun

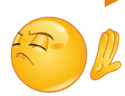
- m strings are hashed (not uniformly) into a hash table with n buckets.
- Each string hashed is an independent trial w.p. p_i of getting hashed into bucket i.

What is $P(E)$ if

1. E = 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

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



$= P(\text{buckets 1 to } k \textbf{ all denied strings})$

$= \left(P(\text{each string hashes to } k+1 \text{ or higher})\right)^m$

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

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