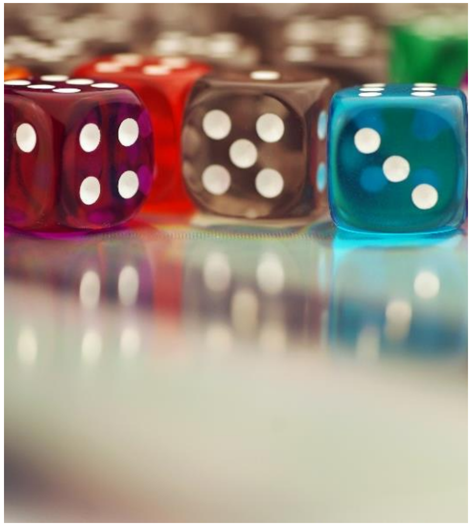
USC Viterbi School of Engineering

CSCI 104
Discrete Probability

Aaron Cote
Mark Redekopp'

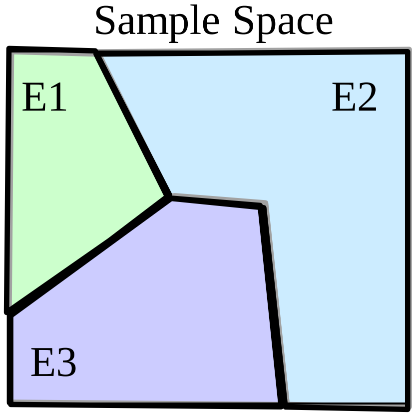


Key Ideas In this Unit

- Basic probability calculation
- Conditional probability and its definition
  - Law of total probability
  - Definition of (mutual and pairwise) independence
- Bernoulli trials and binomial distribution
- Bayes Theorem
- Random Variables and Expected Value
- Linearity of Expectation
- Geometric Distribution

Definitions

- A trial is a procedure that yields one of a set of possible outcomes. A sequence of trials is an experiment.
  - Flipping a coin is a trial.
- The sample space is the set of possible outcomes
  - {heads, tails, on edge}
- An event is a subset of the sample space
  - {heads, on edge}
- You roll two dice and are wondering if you'll get snake eyes (two 1's).
- What is the trial?
- What is the sample space?
- What is the event?
- What is the probability of the event?



Definitions

- A trial is a procedure that yields one of a set of possible outcomes. A sequence of trials is an experiment.
  - Flipping a coin is a trial.
- The sample space is the set of possible outcomes
  - {heads, tails, on edge}
- An event is a subset of the sample space
  - {heads, on edge}
- You roll two dice and are wondering if you'll get snake eyes (two 1's).
- What is the trial?
  - **Rolling two dice**
- What is the sample space?
  - **{1,2,3,4,5,6} x {1,2,3,4,5,6}**
- What is the event?
  - **Snake eyes (1,1)**
- What is the probability of the event?
  - **1/36**

The Uniform Distribution

- In the previous problem, we assumed something.
- The uniform distribution assigns the probability $\frac{1}{n}$ to each of the n elements in the sample space.
  - An unbiased coin assigns $^1/_2$ to both heads and tails
  - An unbiased die assigns $^1/_6$ to each side.
  - Two unbiased dice assigns $^1/_{36}$ to each result.
- If S is a sample space of *equally likely outcomes*, and E is an event of S, then the probability of E is
  - $p(E) = \frac{|E|}{|S|}$
  - To compute $p(E)$ use counting techniques to find |E| and |S|

Practice 1

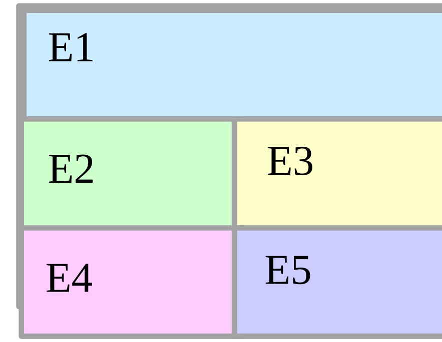
- What is the probability that the sum of two dice rolls is 7?
- How many 5-card Poker hands are there?
- How many 5-card Poker hands contain a 4-of-a-kind?
- What is the probability of a 4-of-a-kind?
- What is the probability of a full house?

Practice 1 Solution

- What is the probability that the sum of two dice rolls is 7?
  - $\frac{6}{36} = \frac{1}{6}$
- How many 5-card Poker hands are there?
  - $\binom{52}{5}$
- How many 5-card Poker hands contain a 4-of-a-kind?
  - $13 \cdot 48$
- What is the probability of a 4-of-a-kind?
  - $\frac{13 \cdot 48}{\binom{52}{5}} = 0.024\%$
- What is the probability of a full house?
  - $\frac{{}_{13}P_2 \cdot \binom{4}{3}\binom{4}{2}}{\binom{52}{5}}$

Rules for Probabilities

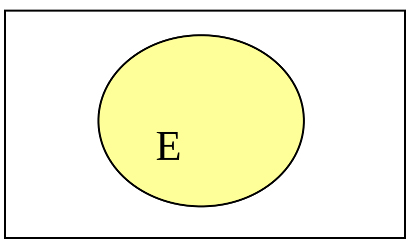
- **Rule 1**: If $S = \{E_1, \ldots, E_n\}$ is the sample space, then $1 = \sum_{i=1}^{n} p(E_i)$ and $\forall i: 0 \le p(E_i) \le 1$

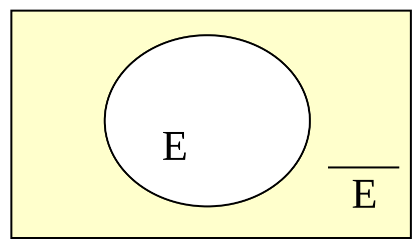


Sample Space

- **Rule 2**: Given event E in sample space S, the probability of the complement of E is:

$$p(\bar{E}) = 1 - p(E)$$



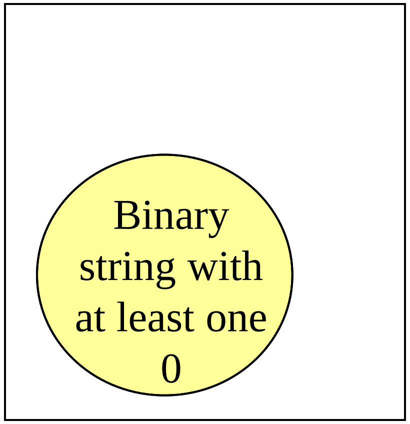


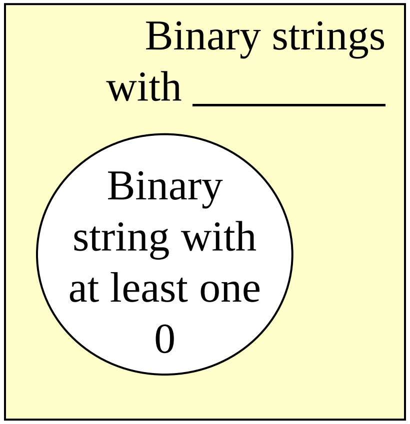
Summary:

- If the sample space is broken into disjoint events, then the probabilities of those events **must sum to 1**
- If the probability of an event happening is p(E), then the probability that the event does NOT happen is **1-p(E)**

Inverting Probabilities

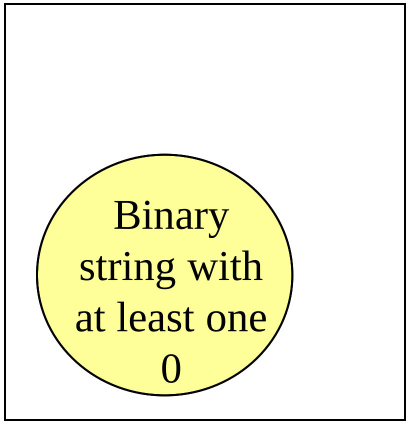
- You have a sequence of 10 bits, generated by the uniform distribution. What is the probability that at least one bit is a 0?

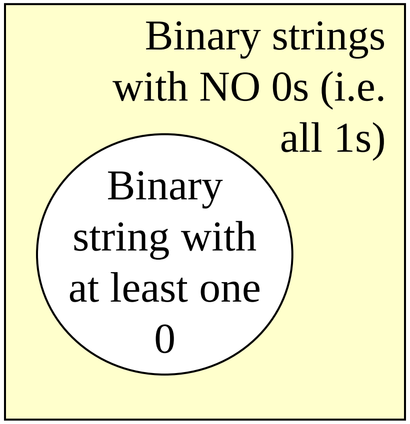




Inverting Probabilities

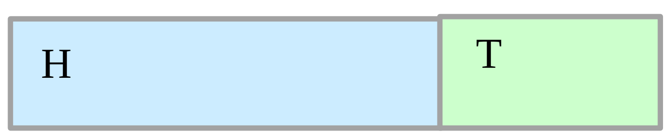
- You have a sequence of 10 bits, generated by the uniform distribution. What is the probability that at least one bit is a 0?
  - $1 - P(\text{all bits are } 1) = 1 - {}^{1}/_{2^{10}}$





Non-Uniform Distributions

- A biased coin is twice as likely to come up heads as tails. What is the probability it comes up heads?
  - $^{2}/_{3}$



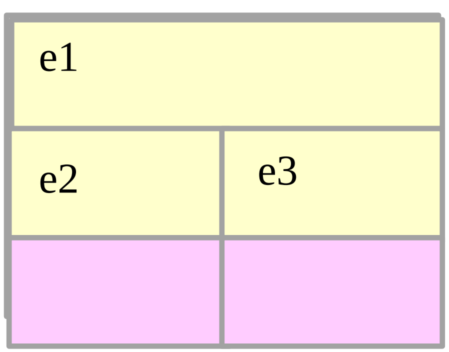
Sample Space

The Sum Rule for Probability

- A biased die has 3 appear twice as often as any other individual number. What is the probability an odd number is rolled?
  - $P(1) + P(3) + P(5) = {}^{1}/_{7} + {}^{2}/_{7} + {}^{1}/_{7} = {}^{4}/_{7}$

- $p(E) = \sum_{e_i \in E} p(e_i)$
  - provided all e_i are disjoint (mutually exclusive)

Sample Space



The Subtraction Rule for Probability

- There is a line of 100 people.
- Everyone at an even position is a pirate.
- Everyone at a position divisible by 5 is a ninja
- You choose a person in the line according to the uniform distribution.
- What is the probability they are a pirate or ninja (or both)?

  - ______________________________

- $p(E1 \cup E2) = p(E1) + p(E2) - p(E1 \cap E2)$

- This is also known as the principle of inclusion-exclusion (aka PIE)

The Subtraction Rule for Discrete Probabilities

- There is a line of 100 people.
- Everyone at an even position is a pirate.
- Everyone at a position divisible by 5 is a ninja
- You choose a person in the line according to the uniform distribution.
- What is the probability they are a pirate or ninja (or both)?
  - $\frac{1}{2}+\frac{1}{5}-\frac{1}{10}=\frac{3}{5}$
- $p(E1 \cup E2) = p(E1) + p(E2) - p(E1 \cap E2)$
- This is also known as the principle of inclusion-exclusion (aka PIE)

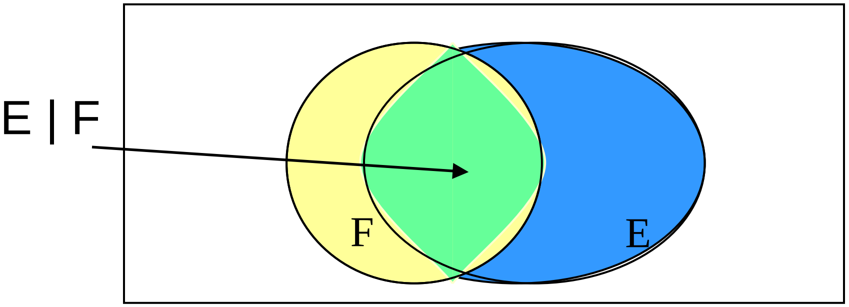
CONDITIONAL PROBABILITY AND INDEPENDENCE

Conditional Probability

- The **conditional probability** of E, given that F occurred, is:

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

- For the biased die, we know that an odd number was rolled. What is the probability a 3 was rolled?
  - $P(3|odd) = \frac{2/7}{4/7} = 1/2$



Practice 2

There are 4 people in line. Each person is either a pirate or a ninja (but not both), assigned by the uniform distribution. What is the probability there are two consecutive ninjas, given the first person in line is a ninja?

$$p(\mathrm{NN}|1N) = \frac{p(\mathrm{NN} \cap 1\mathrm{N})}{P(1\mathrm{N})}$$

Practice 2 Solution

There are 4 people in line. Each person is either a pirate or a ninja (but not both), assigned by the uniform distribution. What is the probability there are two consecutive ninja, given the first person in line is a ninja?

$$p(\mathrm{NN}|1N) = \frac{p(\mathrm{NN} \cap 1\mathrm{N})}{P(1\mathrm{N})}$$

$$= \frac{p(\mathrm{NN} \cap 1\mathrm{N})}{\frac{1}{2}} = \frac{5}{8}$$

<u>Compute $p(\mathrm{NN} \cap 1\mathrm{N})$:</u>

NNPP
NNPN
NNNP
NNNN
NPNN

$$p(\mathrm{NN} \cap 1\mathrm{N}) = {}^{5}/_{16}$$

Practice 3

- A family has two children, the gender is generated u.a.r. (uniformly at random). What is the probability of two boys, given the youngest is a boy?
  - $^{1}/_{2}$

- A family has two children, the gender is generated u.a.r. (uniformly at random). What is the probability of two boys, given at least one boy?
  - Intuition A: This is just the same problem, right? So it is $^{1}/_{2}$?
  - Intuition B: There are 3 possible events for "at least one boy": BG, GB, BB. So it is $^{1}/_{3}$?
- Intuition B is correct.

For the Unconvinced

- Gather data from all U.S. families with exactly 2 children (lets say there are 10 million).
  - 2.5 million families have two girls.
  - 2.5 million families have two boys.
  - 2.5 million families have BG.
  - 2.5 million families have GB.
- Given there is at least one boy, there are 7.5 million families we might have selected from. Of them, $^{1}/_{3}$ of them have two boys.

The Prosecutor's Fallacy (1)

- You are accused of speeding.
  - Define S as the event you are speeding (vs. not speeding)
- The prosecutor notes that you have a radar detector (it detects when police are using their radar system which identifies speeders).
  - Define D as the event you have a radar detector
- Write the probability of your guilt as a conditional probability:
- The prosecutor also notes that 80% of speeders have radar detectors, so you are probably guilty. Is there any flaw in his reasoning?
  - What is the conditional probability described in this paragraph?

The Prosecutor's Fallacy (2)

- You are accused of speeding.
- The prosecutor notes that you have a radar detector (it detects when police are using their radar system which identifies speeders).
- The prosecutor also notes that 80% of speeders have radar detectors, so you are probably guilty.
- Is there any flaw in his reasoning?

- The question at hand is P(S|D) = probability you were a speeder given you had a detector
- The prosecutor is using: P(D|S) = probability you had a detector given you are a speeder
- It could be that 1% of the population speeds, but 80% of all drivers have radar detectors, for example. Now the same argument indicates you probably aren't guilty!
- The fallacy is mistaking p(E|F) for p(F|E)

The Law of Total Probability

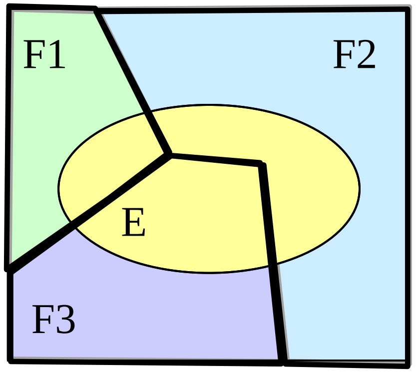
- For any partition of the sample space into disjoint events $F1, \ldots, Fk$, where $F1 \cup \ \ldots \cup Fk = S$ and $\forall_{i,j} i \neq j, F_i \cap F_j = \emptyset$:

$$p(E) = p(E|F1) \cdot p(F1) + \ldots + p(E|Fk) \cdot p(Fk)$$

- Because:

$$p(E) = \frac{p(E \cap F1)}{p(F1)} \cdot p(F1) + \ \ldots \ + \frac{p(E \cap Fk)}{p(Fk)} \cdot p(Fk)$$
$$= p(E \cap F1) + \ \ldots \ + p(E \cap Fk)$$

It may be easier to break the sample space into cases and find the probability of an event in each case!
Go back to the previous example of "at least one boy".



INDEPENDENCE OF EVENTS

Independence

- Two events E and F are defined to be **independent** exactly when:
  - $P(E|F) = P(E)$ (which says knowing that F occurred didn't affect P(E)) or equivalently when
  - $P(E \cap F) = P(E) \cdot P(F)$
- To see why the two are equivalent…
  - Recall: $P(E|F) = \frac{P(E \cap F)}{P(F)}$, but for that to be equal to $\mathrm{P}(E)$ as stated in the first expression for independence we would need the numerator to be $P(E \cap F) = P(\mathrm{E}) \cdot \mathrm{P}(F)$ to make

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

Practice 4

- A family has 2 children, the gender is determined u.a.r. The event E is when there are two girls. F is when there is at least one girl. Are E and F independent?

- A family has 3 children. E is when there is a boy and a girl. F is when there is at most 1 boy. Are E and F independent?

- There are 4 people in line. Each person is a pirate or ninja (not both), generated u.a.r. E is when the first person is a ninja. F is when there are an even number of ninjas. Are E and F independent?

Practice 4 Solution

- A family has 2 children, the gender is determined u.a.r. The event E is when there are two girls. F is when there is at least one girl. Are E and F independent?
  - No. $\mathrm{p(E)} = {}^{1}/_{4}$ $p(F) = {}^{3}/_{4}$ $p(E) \cdot p(F) = {}^{3}/_{16}$ $p(E \cap F) = \frac{1}{4}$
- A family has 3 children. E is when there is a boy and a girl. F is when there is at most 1 boy. Are E and F independent?
  - Yes. $\mathrm{p(E)} = {}^{3}/_{4}$ $p(F) = {}^{1}/_{2}$ $p(E) \cdot p(F) = {}^{3}/_{8}$ $p(E \cap F) = \frac{3}{8}$
- There are 4 people in line. Each person is a pirate or ninja (not both), generated u.a.r. E is when the first person is a ninja. F is when there are an even number of ninjas. Are E and F independent?
  - Yes. $\mathrm{p(E)} = {}^{1}/_{2}$ $p(F) = {}^{1}/_{2}$ $p(E) \cdot p(F) = {}^{1}/_{4}$ $p(E \cap F) = \frac{1}{4}$

Practice 5

- You have a jar with 6 red marbles and 4 blue marbles. You draw 3 of them from the jar, u.a.r.
  - Let E be the event where not all marbles are the same color.
  - Let F be the event where the first marble drawn is red.
- Are E and F independent?
  - $p(F) =$
  - $p(\bar{E}) =$ $p(E) =$
  - $p(E) \cdot p(F) =$
  - $p(E \cap F) =$
- ___ independent!

Practice 5 Solution

- You have a jar with 6 red marbles and 4 blue marbles. You draw 3 of them from the jar, u.a.r.
  - Let E be the event where not all marbles are the same color.
  - Let F be the event where the first marble drawn is red.
- Are E and F independent?
  - $p(F) = {}^{6}/_{10}$
  - $p(\bar{E}) = \frac{\binom{6}{3}+\binom{4}{3}}{\binom{10}{3}} = {}^{1}/_{5}$ $\quad p(E) = {}^{4}/_{5}$
  - $p(E) \cdot p(F) = \frac{4}{5} \cdot \frac{6}{10} = \frac{12}{25}$
  - $p(E \cap F) = {}^{3}/_{5} \cdot (1 - {}^{5}/_{9} \cdot {}^{4}/_{8}) = {}^{13}/_{30}$
- Not independent!

Indendence of many events

- Events E_1, … E_n are **pairwise** independent if every pair of events are independent.
  - For E1,E2,E3 we must show:
  - E1 and E2 are independent, E1 and E3 are independent, and E2 and E3 are independent

- They are **mutually** independent if, for **any subset** of events, they are independent.
  - For E1, E2, E3 we must show:
  - (E1,E2) are independent
  - (E1,E3) are independent
  - (E2,E3) are independent
  - (E1, E2, E3) are independent [i.e. $p(E_1 \cap E_2 \cap E_3) = p(E_1) \cdot p(E_2) \cdot p(E_3)$]

Practice 6

- Consider 2 fair coin flips.
  - E1 is the event where the first flip is heads.
  - E2 is the event where the second flip is heads.
  - E3 is the event where exactly one flip is heads.
- Are these pairwise independent? Mutually independent?
  - $p(E1) = p(E2) = p(E3) =$
  - $p(E1 \cap E2) = p(E1 \cap E3) = p(E2 \cap E3) =$
  - $p(E1 \cap E2 \cap E3) =$
- Results:

Practice 6 Solution

- Consider 2 fair coin flips.
  - E1 is the event where the first flip is heads.
  - E2 is the event where the second flip is heads.
  - E3 is the event where exactly one flip is heads.
- Are these pairwise independent? Mutually independent?
  - p(E1) = p(E2) = p(E3) = $^1/_2$
  - p(E1 ∩ E2) = p(E1 ∩ E3) = p(E2 ∩ E3) = $^1/_4$
  - p(E1 ∩ E2 ∩ E3) = 0
- Results: Pairwise independent, but not mutual.

Practice 7

- An integer between 1 and 8 inclusive is chosen u.a.r.
  - E1 is the event that the number is ≤ 4.
  - E2 is the event that the number is odd.
  - E3 is the event that the number is prime.
- Pairwise independent? Mutually independent?
  - $p(E1) = p(E2) = p(E3) =$
  - $p(E1 \cap E2 \cap E3) =$
  - $p(E2 \cap E3) =$
- Neither!

Practice 7 Solution

- An integer between 1 and 8 inclusive is chosen u.a.r.
  - E1 is the event that the number is ≤ 4.
  - E2 is the event that the number is odd.
  - E3 is the event that the number is prime.
- Pairwise independent? Mutually independent?
  - p(E1) = p(E2) = p(E3) = $^{1}/_{2}$
  - p(E1 ∩ E2 ∩ E3) = $^{1}/_{8}$
  - p(E2 ∩ E3) = $^{3}/_{8}$
- Neither!

The Birthday Paradox

- What are the odds that two or more people watching this lecture share the same birthday? Is it above or below 50%?
- Each person's birthday is independent of the others' and you may assume that any of the 365 possible birthdays are equally likely. Because of independence, we can use a sequence of trials and multiply their probabilities.
- The probability that among n people, at least two have the same birthday, is:

1	2	3		...		365

The Birthday Paradox

- What are the odds that two or more people watching this lecture share the same birthday? Is it above or below 50%?
- Each person's birthday is independent of the others' and you may assume that any of the 365 possible birthdays are equally likely.
- The probability that among n people, at least two have the same birthday, is:
- $1 - \mathrm{p}(each\ of\ the\ n\ birthdays\ is\ unique)$
- $1 - (\frac{365}{365} \cdot \frac{364}{365} \cdot \frac{363}{365} \cdot \ldots \cdot \frac{366-n}{365})$

- n = 10: 11.7%
- n = 23: 50.7%
- n = 30: 70.6%
- n = 50: 97%
- n = 70: 99.9%
- n=366: 100%

BERNOULLI TRIALS AND THE BINOMIAL DISTRIBUTION

Bernoulli Trials Motivation

- A biased coin has probability of heads ${}^{2}/_{3}$. Given 3 flips, what is the probability of exactly 2 heads?
  - Because the flips are independent, we should be able to just multiply the probability of each event in sequence
  - So, it is (2/3)*(2/3)*(1/3)?

Bernoulli Trials

- A Bernoulli Trial is a trial with 2 outcomes that may or may not have equal probability. One outcome is the success with probability p, the other is the failure with probability q = 1-p.
- If you have an experiment of n identical Bernoulli trials, the probability of getting exactly k successes is:
  - $\binom{n}{k} p^k q^{n-k} = \binom{n}{k} p^k (1-p)^{n-k}$
- This looks like the Binomial Theorem!

Practice 8

- A biased coin has probability of heads $^2/_3$. Given 3 flips, what is the probability of exactly 2 heads?
  - ➢
- There are 10 people in line. Each person has a 90% chance of being a pirate and is otherwise a ninja. What is the probability of exactly 8 pirates?
  - ➢

Practice 8 Solution

- A biased coin has probability of heads $^2/_3$. Given 3 flips, what is the probability of exactly 2 heads?
  - $\binom{3}{2} \cdot (^2/_3)^2 \cdot (^1/_3)^1 = 44.4\%$
- There are 10 people in line. Each person has a 90% chance of being a pirate, and is otherwise a ninja. What is the probability of exactly 8 pirates?
  - $\binom{10}{2} \cdot (^1/_{10})^2 \cdot (^9/_{10})^8 = 19\%$

Randomized Algorithms

- Suppose a group of size **n** got together yesterday, and they're worried that they contracted coronavirus during this gathering.
  - Let's say that enough social mingling occurred that if even one person had it, then each person now has a 10% chance of having it now.
- You could test EVERYONE, but suppose testing is limited.
- Instead, you can test a subset, **k** of the members (chosen u.a.r.), and assert that if none of the **k** members are positive, none of the entire group of **n** is positive.
- ***<u>If no one actually contracted the virus</u>***, what is the probability that your testing of only **k** members will indicate no one got COVID (i.e. what is $p(0\ positive\ out\ of\ k\ tests)$)?
  - Define your Bernoulli trial success as positive test: P(success)=_____
  - $p(0\ positive\ out\ of\ k\ tests) =$
- ***<u>If someone did contract the virus</u>***, what is the probability that your testing of only **k** members will (incorrectly) indicate no one got COVID
  - Now P(success)=
  - $p(0\ positive\ out\ of\ k\ tests) =$
- If k = 66, this is < 0.1%. If k = 132, this is < 0.0001%

Randomized Algorithms Solution

- Suppose a group of size **n** got together yesterday, and they're worried that they contracted coronavirus during this gathering.
  - Let's say that enough social mingling occurred that if even one person had it, then each person now has a 10% chance of having it now.
- You could test EVERYONE
  - But suppose testing is limited.
- Instead, you can test **k** of the members (chosen u.a.r.), and assert that if none of the **k** members are positive, none of the entire group of **n** is positive.
- ***If no one actually contracted the virus***, what is the probability that your testing of only **k** members will indicate no one got COVID (i.e. what is $p(0\ positive\ out\ of\ k\ tests)$)?
  - Define your Bernoulli trial success as positive test: P(success)=0
  - $p(0\ positive\ out\ of\ k\ tests) = \binom{k}{0}0^0(1)^k = 1$ (i.e. 100%)
- ***If someone did contract the virus***, what is the probability that your testing of only **k** members will (incorrectly) indicate no one got COVID
  - Now P(success)=$\frac{1}{10}$
  - $p(0\ positive\ out\ of\ k\ tests) = \binom{k}{0}(\frac{1}{10})^0(\frac{9}{10})^k$
- If k = 66, this is < 0.1%. If k = 132, this is < 0.0001%

Randomized Algorithms

- This was an example of a **Monte Carlo** randomized algorithm: such an algorithm is always fast, but sometimes returns the wrong answer.
  - In a few lectures we will see that a **Bloom Filter** is a Monte Carlo data structure.

- The other type of randomized algorithm is a **Las Vegas** algorithm: it is always right, but sometimes takes a long time.
  - QuickSort, and HashTables, are examples of this type.

BAYE'S THEOREM

A motivating problem

- There are two boxes.
  - Box 1 contains 2 gold coins and 7 cardinal coins
  - Box 2 contains 4 gold coins and 3 cardinal coins
- You choose a box u.a.r.
- Then from that box, you choose a coin u.a.r.
- You draw a cardinal coin
- What is the probability the coin came from the first box?

A motivating problem

- CC is the event that you draw a cardinal coin.
- B1 is the event that you draw from box 1.
- Calculating p(CC|B1) is…
  - EASY! ($^{7}/_{9}$)
- Calculating p(B1|CC) is…
  - Seemingly much harder

Bayes Theorem

- Consider:
  - $p(F|E) = \frac{p(E \cap F)}{p(E)}$ but also $p(E|F) = \frac{p(F \cap E)}{p(F)}$
  - $p(F|E) \cdot p(E) = p(E \cap F) = p(E|F) \cdot p(F)$
  - $p(F|E) = \frac{p(E|F) \cdot p(F)}{p(E)}$
- Bayes' Theorem:
  - $p(F|E) = \frac{p(E|F) \cdot p(F)}{p(E)}$

USC Viterbi School of Engineering

Calculating Baye's

- Bayes' Theorem:
  - $p(F|E) = \frac{p(E|F) \cdot p(F)}{p(E)}$

- Often calculating p(E) is more difficult and can more easily be accomplished by splitting its calculation over 2 or more disjoint sets to get (recall the Law of Total Probability)
  - $p(F|E) = \frac{p(E|F) \cdot p(F)}{p(E|F) \cdot p(F) + p(E|\bar{F}) \cdot p(\bar{F})}$ (breaking E over F and NOT F)

  or

  - $p(F|E) = \frac{p(E|F) \cdot p(F)}{p(E|F1) \cdot p(F1) + \ldots + p(E|Fk) \cdot p(Fk)}$ (breaking E over F1, F2, … Fk)

The Cardinal Coin Problem

- Let CC is the event that you draw a cardinal coin.
- Let B1 is the event that you draw from box 1.
- Find (B1|CC)…

- Box 1 contains 2 gold coins and 7 cardinal coins
- Box 2 contains 4 gold coins and 3 cardinal coins

- $p(\mathrm{B1}|\mathrm{CC}) =$

- $p(\mathrm{B1}) = p(B2) =$
- $p(\mathrm{CC}|\mathrm{B1}) =$
- $p(\mathrm{CC}|B2) =$
- $p(\mathrm{B1}|\mathrm{CC}) =$

The Cardinal Coin Problem (Solution)

- Let CC is the event that you draw a cardinal coin.
- Let B1 is the event that you draw from box 1.
- Find (B1|CC)…

> - Box 1 contains 2 gold coins and 7 cardinal coins
> - Box 2 contains 4 gold coins and 3 cardinal coins

- $p(\mathrm{B1}|\mathrm{CC}) = \frac{p(CC|B1)p(B1)}{p(CC|B1)p(B1)+p(CC|B2)p(B2)}$
- $p(\mathrm{B1}) = p(B2) = {}^{1}/_{2}$
- $p(\mathrm{CC}|\mathrm{B1}) = {}^{7}/_{9}$
- $p(\mathrm{CC}|B2) = {}^{3}/_{7}$
- $p(\mathrm{B1}|\mathrm{CC}) = \frac{p(CC|B1)p(B1)}{p(CC|B1)p(B1)+p(CC|B2)p(B2)} = \frac{{}^{7}/_{9}\cdot{}^{1}/_{2}}{{}^{7}/_{9}\cdot{}^{1}/_{2}+{}^{3}/_{7}\cdot{}^{1}/_{2}} = 64\%$

Practice 9

- Say that 10% of the population has coronavirus.
- 85% of those infected display a fever.
- 5% of those who are not infected display a fever (from other causes)
- You display a fever. What is the probability you have coronavirus?
  - Let F = you display a fever. Let C = you have coronavirus.
  - Describe the probability we are solving for?

- $\mathrm{p}(\mathrm{F}|\mathrm{C}) = \quad , p\left(F\middle|\overline{C}\right) = \quad , p(C) =$
- $p(\mathrm{C}|\mathrm{F}) =$

Practice 9 Solution

- Say that 10% of the population has coronavirus.
- 85% of those infected display a fever.
- 5% of those who are not infected display a fever (from other causes)
- You display a fever. What is the probability you have coronavirus?
  - Let F = you display a fever. Let C = you have coronavirus.
  - Describe the probability we are solving for? $\mathbf{p}(\boldsymbol{C}|\boldsymbol{F})$
  - And by Baye's Theorem: $p(C|F) = \frac{p(F|C) \cdot p(C)}{p(F|C) \cdot p(C) + p(F|\overline{C}) \cdot p(\overline{C})}$
- $\mathrm{p(F|C)} = 85\%, p(\mathrm{F}|\overline{\mathrm{C}}) = 5\%, p(\mathrm{C}) = 10\%$
- So $p(\mathrm{C|F}) = \frac{85\% \cdot 10\%}{85\% \cdot 10\% + 5\% \cdot 90\%} = 65\%$

USC Viterbi School of Engineering

Oddities from Bayes' Theorem

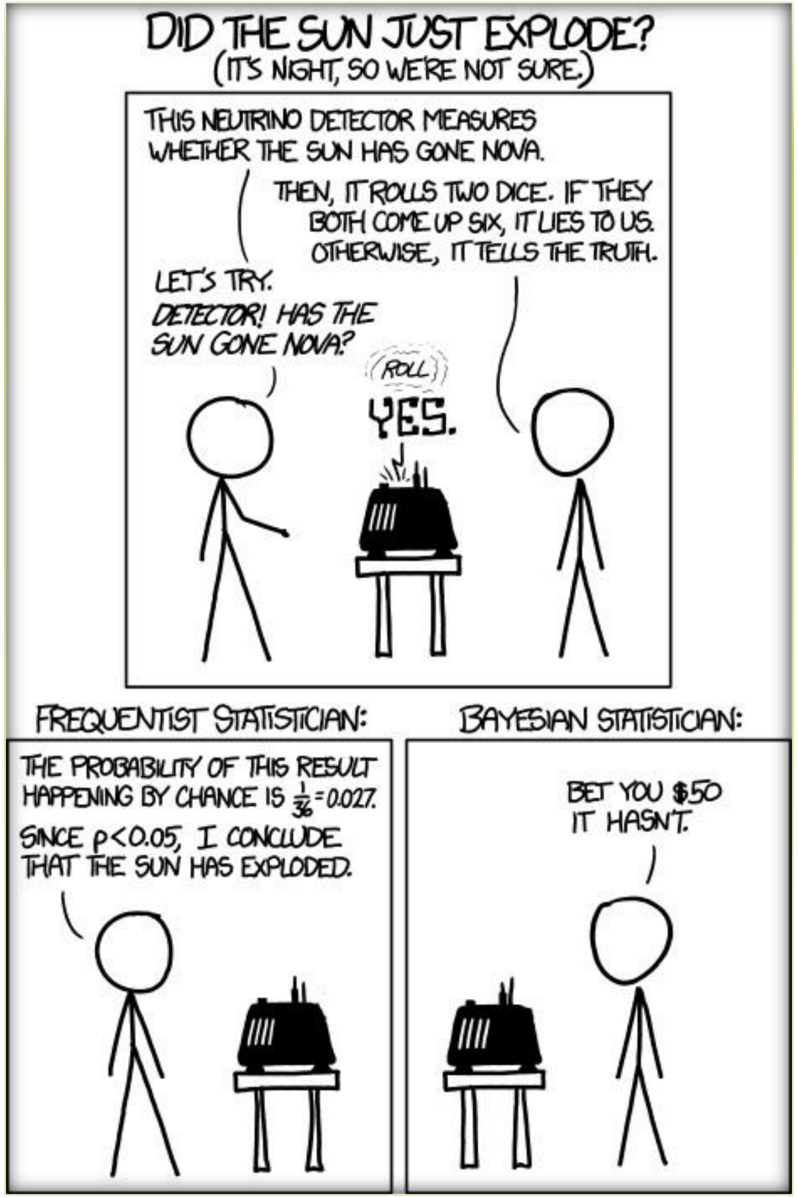
- Suppose there is a much rarer disease that infects 1 out of 100,000 people.
- You have a very accurate test for the disease:
  - If you have the disease, it correctly says so 99% of the time (i.e. true positive)
  - If you don't, it correctly says so 99.5% of the time. (i.e. true negative)
- You take the test, and it comes back positive. What is the probability you have the disease?
  - Let TP = test comes back positive, D = you have the disease.
  - What are we trying to solve? ________________
- $p(TP|\mathrm{D}) =$ _____, $p(TP|\overline{\mathrm{D}}) =$ ______, $p(\mathrm{D}) =$ ________

USC Viterbi School of Engineering

Oddities from Bayes (Solution)

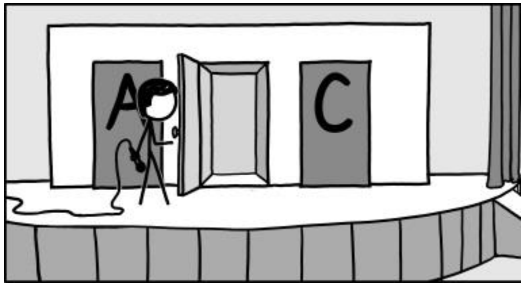
- Suppose there is a much rarer disease that infects 1 out of 100,000 people.
- You have a very accurate test for the disease:
  - If you have the disease, it correctly says so 99% of the time (i.e. true positive)
  - If you don't, it correctly says so 99.5% of the time. (i.e. true negative)
- You take the test, and it comes back positive. What is the probability you have the disease?
  - Let TP = test comes back positive, D = you have the disease.
  - What are we trying to solve? p(D|TP)
- $p(TP|\mathrm{D}) = 99\%, p(TP|\overline{\mathrm{D}}) = 0.5\%, p(\mathrm{D}) = 0.001\%$
- $p(\mathrm{D}|\mathrm{TP}) = \frac{99\% \cdot 0.001\%}{99\% \cdot 0.001\% + 0.5\% \cdot 99.999\%} = \frac{0.0000099}{.0000099 + 0.00499995} = 0.2\%$

XKCD #1132 Frequentists vs. Bayesians



The Monty Hall Three-Door Puzzle

- There is a famous problem from a game show, "Let's Make a Deal" with host Monty Hall. It goes as follows:
- There are three doors.
  - Behind one is a car.
  - Behind each of the other two doors is a goat.
  - The gameshow host knows where the car is.



- You are asked to choose one door, but don't open it.
- The gameshow host intentionally chooses a door which you did not choose, which does not have the car, and opens it, revealing a goat.
- You can now open the door you originally chose, or the other closed door. What should you do?

Monty Hall

- **Intuition A:** there are two doors left, and each is equally likely, so it doesn't matter which door you open.

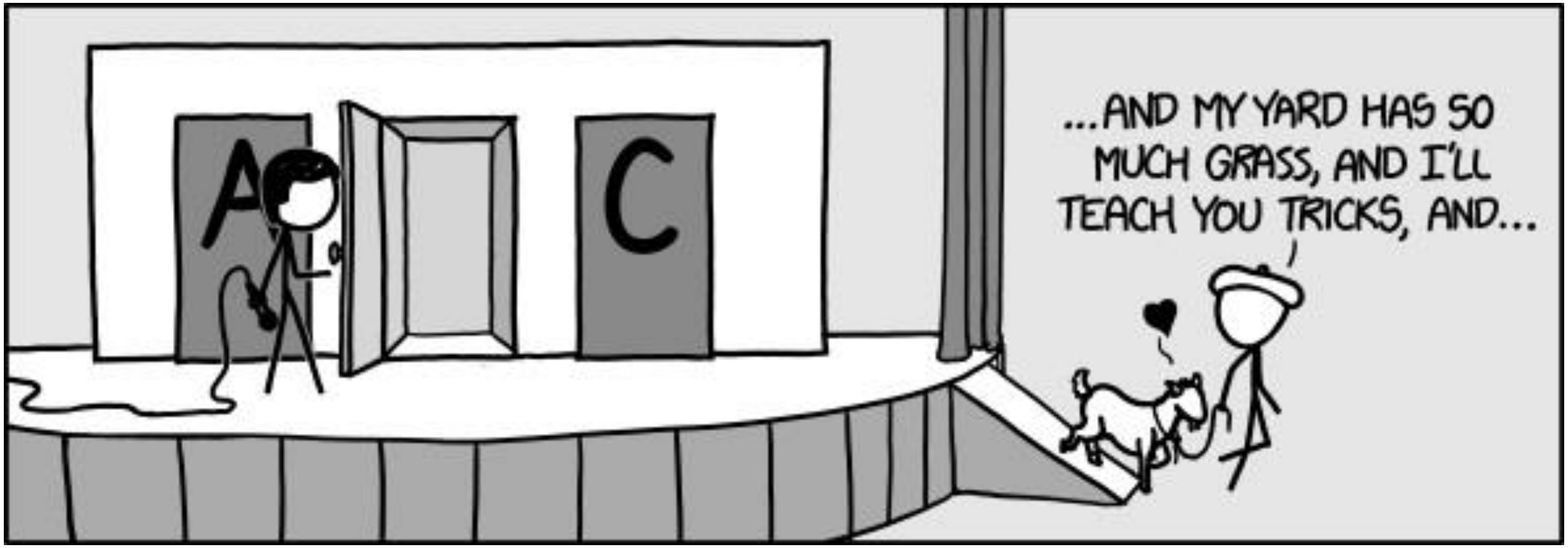
- **Intuition B:** when you guessed which door to open, you had a $^1/_3$ chance of it having the car. That is still the case, so the other doors have a $^2/_3$ probability of a car. So, you should switch doors.

- Intuition B is correct. Why?

Monty Hall for the unconvinced

- There are 100 doors.
  - Behind one is a car.
  - Behind each of the other 99 doors is a goat.
  - The gameshow host knows where the car is.
- You choose one door (door 1), but don't open it.
- The gameshow host opens 98 other doors (which does not include your door) that do NOT have the car behind it
- Does the probability you originally chose the correct door magically increase from $^1/_{100}$ to $^1/_2$? (No)
  - The 99/100 chance that the car was behind some OTHER door that you did not choose, is simply "transferred" to the remaining unopened door.

XKCD #1282: Monty Hall



Monty Hall and Baye's

- Let's formulate this problem using Baye's theorem.
- There are 3 doors. Without loss of generality: You choose door 1. The host then opens door 2 to reveal a goat. What is the probability you should switch (i.e. the car is behind door 3)?
  - Let **Ci** = the car is behind door i, **E**=You chose door 1 and the host opens door 2.
  - We are asking you to solve for P(C3 | E)
- By Baye's Theorem:

  $$p(\mathrm{C}3|E) =$$

- Find the component probabilities:
  - $p(\mathrm{C}i) =$
  - $p(E|\mathrm{C}3) =$
  - $p(E|\mathrm{C}2) =$
  - $p(E|\mathrm{C}1) =$
- Thus, $p(\mathrm{C}3|E) =$

Monty Hall and Baye's (Solution)

- Let's formulate this problem using Baye's theorem.
- There are 3 doors. Without loss of generality: You choose door 1. The host then opens door 2 to reveal a goat. What is the probability you should switch (i.e. the car is behind door 3)?
  - Let **Ci** = the car is behind door i, **E**=You chose door 1 and the host opens door 2.
  - We are asking you to solve for P(C3 | E)
- By Baye's Theorem:

$$p(\mathrm{C3}|E) = \frac{p(E|\mathrm{C3})\cdot p(\mathrm{C3})}{p(E)} = \frac{p(E|\mathrm{C3})\cdot p(\mathrm{C3})}{p(E|\mathrm{C1})p(\mathrm{C1})+p(E|\mathrm{C2})p(\mathrm{C2})+p(E|\mathrm{C3})p(\mathrm{C3})}$$

- Find the component probabilities:
  - $p(\mathrm{F}i) = {}^1/_3$
  - $p(E|\mathrm{C3}) = 1$
  - $p(E|\mathrm{C2}) = 0$, because the host never opens the door with the car.
  - $p(E|\mathrm{C1}) = {}^1/_2$, because the host can open door 2 or 3.
- Thus

$$p(F3|E) = \frac{{}^1/_3}{{}^1/_2 \cdot {}^1/_3 + 0 \cdot {}^1/_3 + 1 \cdot {}^1/_3} = {}^2/_3$$

RANDOM VARIABLES, EXPECTED VALUE AND VARIANCE

Definitions

- A **Random Variable** is a **mapping/function** [NOT a VARIABLE] from the sample space, S, to the set of real numbers, $\mathbb{R}$.
  - That is, every single outcome is assigned a number.
- If your trial is to flip 3 coins, there are 8 possible outcomes in the sample space. You could have a random variable X that calculates the number of heads for each outcome: X(HHT) = 2, for example
- The distribution of a random variable is the probability of each possible number.
- So for the above trial:

- $p(X = 0) = {}^{1}/_{8}$
- $p(\mathrm{X} = 1) = {}^{3}/_{8}$
- $p(\mathrm{X} = 2) = {}^{3}/_{8}$
- $p(\mathrm{X} = 3) = {}^{1}/_{8}$

Outcomes of 3 flips:

HHH
HHT
HTH
HTT
THH
THT
TTH
TTT

A Dice Problem

- You roll two fair dice. X(t) is a random variable that outputs the sum of the dice. What is the distribution of X?

- $p(X = 2) = {}^{1}/_{36} = p(X = 12)$
- $p(X = 3) = {}^{2}/_{36} = p(X = 11)$
- $p(X = 4) = {}^{3}/_{36} = p(X = 10)$
- $p(X = 5) = {}^{4}/_{36} = p(X = 9)$
- $p(X = 6) = {}^{5}/_{36} = p(X = 8)$
- $p(X = 7) = {}^{6}/_{36}$

Random Variables: Origins

- In 1940, two mathematicians were planning to publish papers about random variables, and they were trying to decide what to call them (random variable or chance variable, each preferred by one of the mathematicians).
- Given the subject matter, how do you suppose they decided?
- They flipped a coin, and we've been stuck with random variable ever since.

Expectation

- The Expected Value E(X) of a random variable X is the average value outputted by X. That is:

$$\mathbf{E}(\boldsymbol{X}) = \sum_{s \in SampleSpace} \mathbf{X}(\mathbf{s}) \cdot \mathbf{p}(\mathbf{s})$$

$$\mathbf{E}(\boldsymbol{X}) = \sum_{x \in Range(X)} \boldsymbol{x} \cdot \mathbf{p}(\boldsymbol{X} = \boldsymbol{x})$$

- What is the expected value of a fair die roll?
  - $1 \cdot {}^{1}/_{6} + 2 \cdot {}^{1}/_{6} + 3 \cdot {}^{1}/_{6} + 4 \cdot {}^{1}/_{6} + 5 \cdot {}^{1}/_{6} + 6 \cdot {}^{1}/_{6} = 3.5$

Expectation

- What is the expected number of heads for a single coin flip?
  - $E[\text{Heads from 1 coin flip}] =$
- What is the expected number of heads in 3 coin flips?
  - $E[\text{Heads from 3 coin flips}] =$
  - Isn't that just ___ * E[heads from 1 coin flip]...
  - In general, **given n Bernoulli trials**, each with **probability of success p**, the **expected number of successes is _____**
- Why can we do that? ____________ of expectations...

Expectation

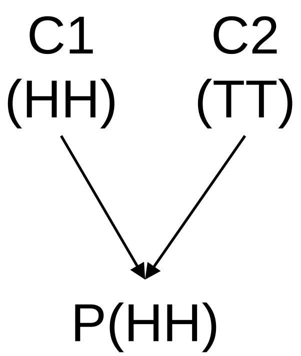
- What is the expected number of heads for a single coin flip?
  - $E[\text{Heads from 1 coin flip}] = {}^1/_2 \cdot 0 + {}^1/_2 \cdot 1 = 0.5$
- What is the expected number of heads in 3 coin flips?
  - $E[\text{Heads from 3 coin flips}]$
    $= {}^1/_8 \cdot 0 + {}^3/_8 \cdot 1 + {}^3/_8 \cdot 2 + {}^1/_8 \cdot 3 = {}^{12}/_8 = 1.5$
  - Isn't that just 3 * E[heads from 1 coin flip]…
  - In general, **given n Bernoulli trials**, each with **probability of success p**, the **expected number of successes is $n \cdot p$**
- Why can we do that? Linearity of expectations…

Linearity of Expectations

- $E(X1 + \cdots + Xn) = E(X1) + \cdots + E(Xn)$
  - The **expectation of the sum** of random variables is the **sum of the expectations** of each random variable…this is what we'll use the most
  - That is, if you want to calculate the expected number of successes over n trials, you can calculate the expected number of successes for the 1st trial, and the 2nd trial, and the 3rd trial, etc, and add them all up.
- $E(a \cdot X + b) = a \cdot E(X) + b$
  - We'll use this identify less often
- What is the expected value of the sum of 3 fair dice?
  - $3 \cdot 3.5 = 10.5$

Expectation on Non-Independent Random Variables

- You have two coins. One has heads on both sides, the other has tails on both sides.
- You choose a coin u.a.r., and then you flip it twice.
  - Let **C1** = choose HH coin, **C2** = choose TT coin, and **HH** you get two heads.
- Are HH and C1 independent (or HH and C2)?
  - Recall independence test: P(HH | C1) = P(HH)
  - P(HH | C1) = 1, P(HH) = 1/2 = P(HH | C1) + P(HH | C2), so they are NOT independent
- Define the following R.V.s…
  - X = R.V. of how many heads you get in your 2 flips
  - X1 = R.V. of how many heads you get with C1
  - X2 = R.V. of how many heads you get with C2
- What is the expected number of heads?
  E[X] = E[0.5*X1+0.5*X2] = 0.5*E[X1]+0.5*E[X2] = (0.5*2) + (0.5*0) = 1
- Linearity of Expectations holds, even if the variables aren't independent!



Expected Chain Length

- In a hash table that uses chaining, recall that loading factor, α, defined as:
  - (n=number of items in the table) / (m=tableSize) => α = n / m
  - It is just the fraction of locations currently occupied
- For chaining, α, can be greater than 1
  - This is because n > m
  - For given values of n and m, let L = the length of a chain at some location = number of items that hashed to that location
  - **What is E[L]**? (Hint: Consider an item hashing to location x as a Bernoulli trial
    - P(success) = P(1 key hashes to some location x)
  - **E[L]** = _______________
- Best to keep the loading factor, α, below 1
  - Resize and rehash contents if load factor too large (using new hash function)

Hirings

- n applicants are interviewed for a single job opening, 1 applicant per day.
- The first applicant is automatically hired.
- For every day thereafter, if the current applicant is better than the current employee, then the applicant is hired and the employee is fired.
- What is the expected number of hirings?

Hirings Solution

- Let I_k return 1 if the kth arriving person is hired, and 0 otherwise. We want:
- $E(\sum_{k=1}^{n} I_k) = \sum_{k=1}^{n} E(I_k)$
- E(I_1) = ___ (the first person is ______________).
- E(I_2) = ___ (there are __ Ways to (relatively) order the first 2 arrivers and only ____ of them has the best arriving second)
- E(I_3) $=$ __ (there are ___ Ways to (relatively) order the first 3 arrivers and only __ of them have the best one last)
- …
- $\sum_{k=1}^{n} E(I_k) = \sum_{k=1}^{n} \frac{1}{k} = \Theta(\log n)$

1234
1243
1324
1342
1423
1432
2134
2143
2314
2341
2413
2431
3124
3142
3214
3241
3412
3421
4123
4132
4213
4231
4312
4321

Hirings Solution

- Let I_k return 1 if the kth arriving person is hired, and 0 otherwise. We want:
- $E(\sum_{k=1}^{n} I_k) = \sum_{k=1}^{n} E(I_k)$
- E(I_1) = 1 (the first person is always hired).
- E(I_2) = $\frac{1}{2}$ (there are 2! Ways to (relatively) order the first 2 arrivers and only one of them has the best arriving second)
- E(I_3) $= \frac{1}{3}$ (there are 3! Ways to (relatively) order the first 3 arrivers and only 2! of them have the best one last)
- …
- $\sum_{k=1}^{n} E(I_k) = \sum_{k=1}^{n} \frac{1}{k} = \theta(\log n)$

Hash Tables

- A hash table has k buckets and n items.
- Each item is distributed to buckets u.a.r., and independently from each other.
- What is the expected number of empty buckets?
- We can figure this out by determining the probability a specific bucket is empty.
- Let Bi be the event where bucket i is empty.
- p(Bi) = E(Bi) = $\left(\frac{k-1}{k}\right)^n$
- We want E(B1 + … + Bk)
- By Linearity of Expectations, this is $k \cdot \left(\frac{k-1}{k}\right)^n$
- If n = k = 1000, this is about 368. If n =1000 and k = 10000, this is about 9048.

Hash Table Analysis

Assume a universal hash function, and $\alpha=1$

- When finding the item I'm looking for, how many other items, on average, will be in the same bucket?
- (m-1)/m (Linearity of Expectation!)

On average, every item has 1 "roommate".

- Note that this is different than asking what the average bucket size is.

The Geometric Distribution

- A coin comes up heads with probability p. We will flip the coin until it comes up heads. What is the expected number of flips?
- pr(X=1) = p
- pr(X=2) = $(1-p) \cdot p$
- pr(X=3) = $(1-p)^2 \cdot p$
- $E(X) = \sum_{i=1} i \cdot (1-p)^{i-1} \cdot p = \frac{1}{p}$
- Because for |x| < 1:
- $\sum_{k=1}^{\infty} kx^{k-1} = \frac{1}{(1-x)^2}$

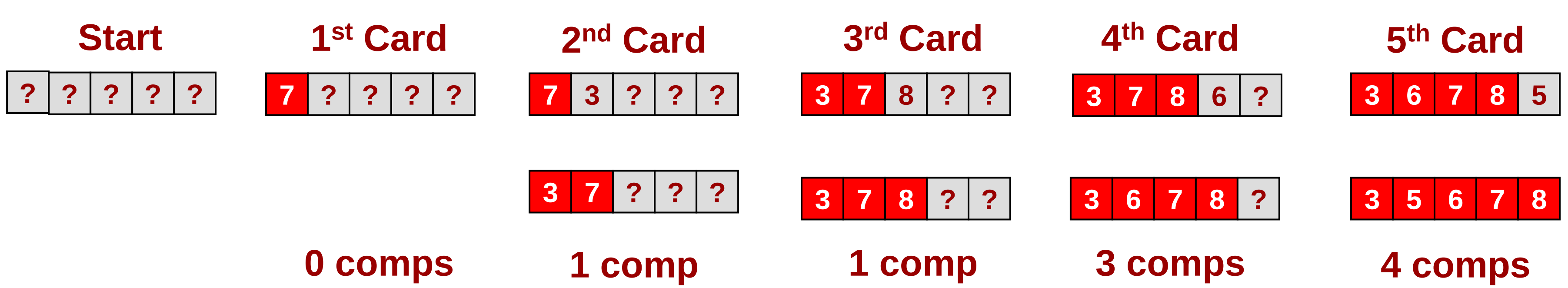
The Geometric Distribution

- A random variable has the geometric distribution with parameter $p : 0 \leq p \leq 1$ if $pr(X = k) = (1 - p)^{k-1} \cdot p$, for all positive integers k. It has expected value $1/p$.
- The probability that a randomly chosen 1000 digit number is prime is approximately $1/2302$. We choose a 1000 digit number u.a.r., and check if its prime. If it is, we're done, otherwise we randomly select another (possibly different) number and repeat. What is the expected number of times we select a number?
  - 2302
- A doctor is treating coronavirus victims. On each day, she has a $1/4$ probability of catching coronavirus, a $1/8$ probability of finishing treatment for the last victim, and otherwise she comes back to work the next day. What is the expected number of successive days she comes in to work?
  - $8/3$

MORE EXPECTED VALUE PROBLEMS

Insertion Sort Algorithm

- Recall how insertion sort works:
  - Imagine we pick up one element of the array at a time and then just insert it into the right position
  - Similar to how you sort a hand of cards in a card game
  - You pick up the first (it is by nature sorted), then pick up the second and insert it at the right position, etc.
- What is the expected number of comparisons made by Insertion Sort?

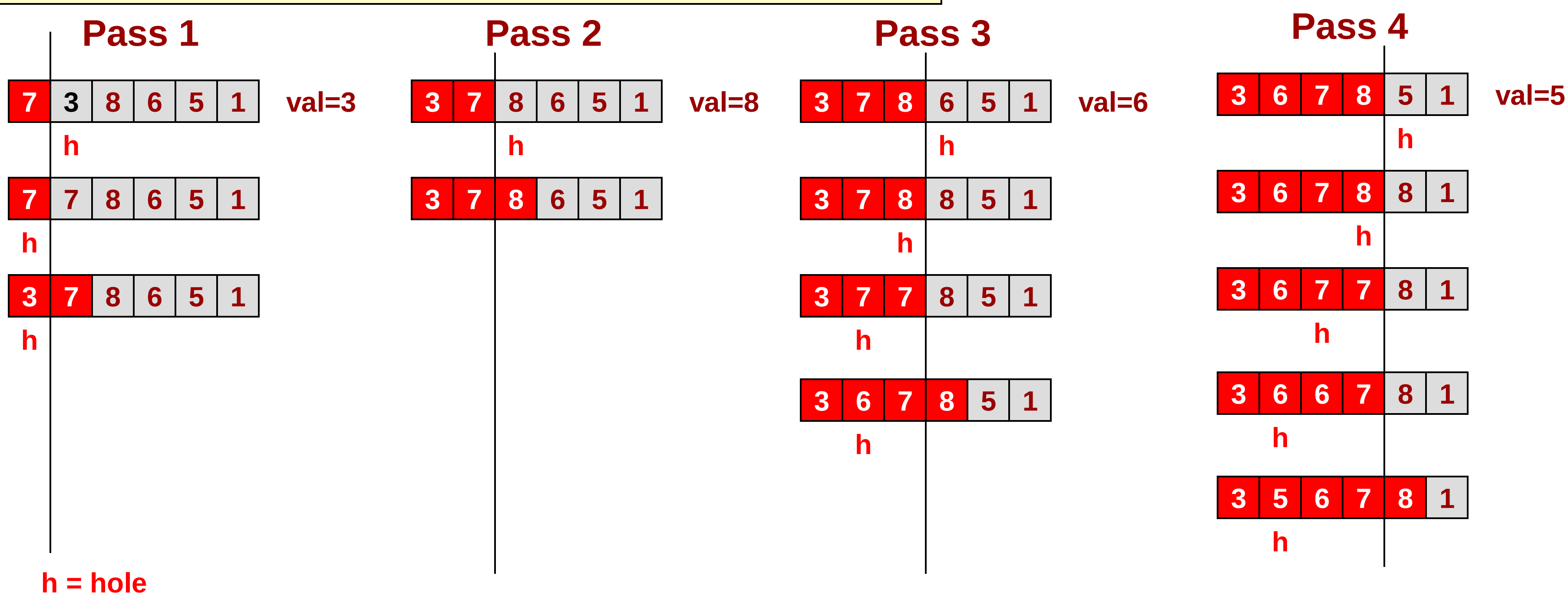


For this permutation, total = 9 comparisons

Insertion Sort Algorithm

```
void isort(vector<int>& mylist)
{  for(int i=1; i < mylist.size(); i++){
      int val = mylist[i];
      hole = i;
      while(hole != 0 && val < mylist[hole-1]){
         mylist[hole] = mylist[hole-1];
         hole--;
      }
      mylist[hole] = val;
}  }
```

Essentially, what is the expected number of times this comparison executes



Insertion Sort Analysis

- What is the expected number of comparisons made by insertion sort?
- Approach:
  - Define the RVs: **Xi** for $1 \ <= \ i \ <= \ n$ to be the number of comparisons needed to insert the i-th element.
  - We want to calculate E(X_1 + … + X_n)
  - Then by Linearity of Expectations, we can just calculate E(X_1), …, E(X_n)
- X_1 will always output 0 since the first item need not be compared, thus…
  - E(X_1) = 0
- X_2 will always output 1, because we need to compare a[1] and a[2], thus..
  - E(X_2) = 1
- X_3 will always output *at least* 1.
  - If it is the largest element so far, it will only output 1.
  - Otherwise, we will do another comparison with the smallest element, and it will output 2.
  - There are 2 cases where we do 2 comparisons (smallest and 2nd smallest)
  - E(X_3) $= {}^2/_3 \cdot 2 + {}^1/_3 \cdot 1 = {}^5/_3$

Insertion Sort, Continued

- $E(X_4) = {}^1/_4 \cdot 1 + {}^1/_4 \cdot 2 + {}^2/_4 \cdot 3 = {}^9/_4$
  - $E(X_4) = \frac{1+2+2\cdot3}{4}$
- $E(X_5) = {}^1/_5 \cdot 1 + {}^1/_5 \cdot 2 + {}^1/_5 \cdot 3 + {}^2/_5 \cdot 4 = {}^{14}/_5$
  - $E(X_5) = \frac{1+2+3+2\cdot4}{5}$
- $E(X_6) = {}^{20}/_6 = \frac{1+2+3+4+2\cdot5}{6}$

- At this point we can generalize the pattern:
- $E(X_k) = \frac{(\sum_{i=1}^{k-1} i))+(k-1)}{\text{k}} = \frac{\left(\sum_{i=1}^{k} i\right)-1}{\text{k}} = \frac{\frac{k(k+1)}{2}-1}{\text{k}} = \frac{\text{k}\cdot(\text{k}+1)-2}{2\text{k}}$

- We can find the expected value of the sum of the random variables by applying linearity of expectations to instead sum all the individual expected values
- $E[\sum_{\text{k}=1}^{n} X_k] = \sum_{\text{k}=1}^{n} E(X_k) = \sum_{k=1}^{n} \frac{\text{k}\cdot(\text{k}+1)-2}{2\text{k}} = \sum_{k=1}^{n} \frac{\text{k}+1}{2} - \sum_{k=1}^{n} \frac{1}{\text{k}} = \theta(n^2)$

Expected Value - Practice 10

- The ordered pair ⟨i,j⟩ is an inversion in a permutation of the first n positive integers if i < j, but j precedes i in the permutation (i.e. a **smaller** value comes AFTER a **larger** value in the permutation)
- There are 6 inversions in 3, 5, 1, 4, 2
  - ___________________________________

Idx:	**1**	**2**	**3**	**4**	**5**
Val:	3	5	1	4	2

- What is the max number of inversions in a list of n positive integers?

Worst case:

Idx:	**1**	**2**	**3**	**4**	**5**
Val:					

$(n-1)+(n-2)+\ldots+1+0$

- What is the expected number of inversions in a random permutation of the first n positive integers?
  - Let I_{ij} be the RV that assigns 1 if <i,j> is an inversion, and 0 otherwise
  - The probability that <i,j> is an inverted pair is __ so $E[I_{ij}]$ = ________________
  - If we then sum over all pairs, we'd expect:

Expected Value - Practice 10 (Solution)

- The ordered pair ⟨i,j⟩ is an inversion in a permutation of the first n positive integers if i < j, but j precedes i in the permutation.
- There are 6 inversions in 3, 5, 1, 4, 2
  - ⟨1,3⟩, ⟨1,5⟩, ⟨2,3⟩, ⟨2,4⟩, ⟨2,5⟩, ⟨4,5⟩
- What is the max number of inversions in a list of n positive integers?
  - $\binom{n}{2} = \frac{n\cdot(n-1)}{2}$
- What is the expected number of inversions in a random permutation of the first n positive integers?
  - The probability that <i,j> is an inverted pair is ½ so $E[I_{ij}] = (1/2)*1 + (1/2)*0=1/2$
  - If we then sum over all n(n-1)/2 pairs, we'd expect: $\frac{n\cdot(n-1)}{2} \cdot \frac{1}{2}$
  - $E[X] = \frac{n\cdot(n-1)}{4}$

VARIANCE AND STANDARD DEVIATION

Multiplying Expectations

- If X and Y are independent, then E(X · Y) = E(X) · E(Y).
- If we roll two fair dice and multiply their values, what is the expected result?
- 3.5^2 = 12.25

- If we roll one fair die, and square its value, what is the expected result?
- E(X · X) = E(X^2) $\neq$ E(X) · E(X), because X is not independent from X.

- $\frac{1}{6} \cdot 1^2 + \frac{1}{6} \cdot 2^2 + \frac{1}{6} \cdot 3^2 + \frac{1}{6} \cdot 4^2 + \frac{1}{6} \cdot 5^2 + \frac{1}{6} \cdot 6^2 \approx 15.17$

Variance

- The expected value of a random variable tells us the average, but nothing about how widely its values are distributed.
- If the midterm has an average of 33, it's possible that everyone got 33, and it's also possible that half the class got 50 while the other half got 16. Variance is a metric that measures how widely values are distributed.
- The variance of X is

$$V(X) = \sum_{s \in S} \left(X(s) - E(X)\right)^2 \cdot p(s)$$

- $X(s) - E(X)$ measures how much individual values vary from the average.
- Squaring it heavily weights outliers: a large value will be produced if there are a lot of scores very far from the average.
- An easier way to calculate variance is $V(X) = E(X^2) - E(X)^2$

Practice 11

- A gambler's coin has heads on both sides. What is the variance on the number of heads after n flips?
- X is always n, X^2 is always n^2, so $E(X^2) – E(X)^2 = n^2 – n^2 = 0$
- That's what you'd expect.
- What is the variance on the number of successes for a single Bernoulli trial?
- $pr(X = 1) = p$
- $E(X) = p \cdot 1 + (1 - p) \cdot 0 = p$
- $E(X^2) = p \cdot 1^2 + (1 - p) \cdot 0^2 = p$
- $V(X) = E(X^2) – E(X)^2 = p - p^2 = p \cdot (1 - p) = p \cdot q$

Practice 12

- What is the variance on the result of a fair die roll?

$$E(X) = 3.5$$

$$E(X^2) = \frac{1}{6} \cdot (1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2) \approx 15.17$$

V(X) = 15.17 – 3.52 ≈ 2.92

- What is the variance of the number of heads for 1 coin flip?

$$pq = {}^{1}/_{4}$$

- What is the variance of the number of heads for 3 coin flips?

$$3pq = {}^{3}/_{4}$$

- Wait, why did that work?

Bienayme's Formula

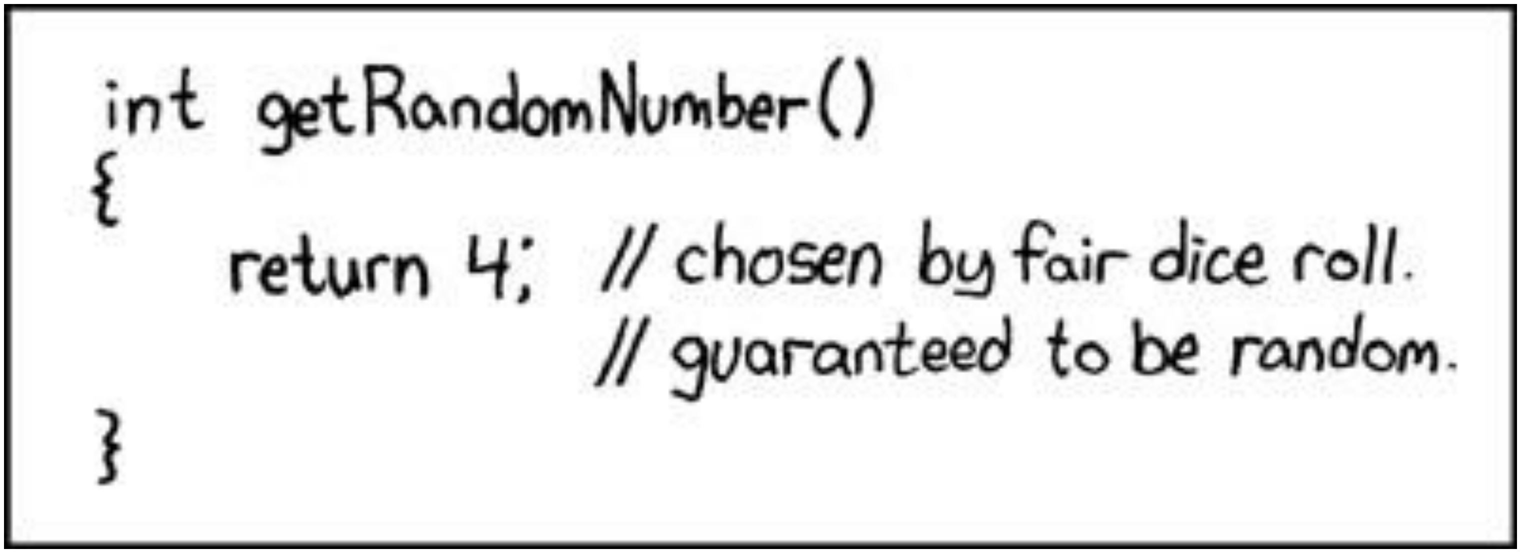
- If X1, ..., Xn are pairwise independent random variables, then:
- $$V(X1 + \ldots + Xn) = V(X1) + \ldots + V(Xn)$$
- What is the variance on the number of successes over n Bernoulli trials?
- npq

USC Viterbi School of Engineering

Standard Deviation

- The standard deviation of a random variable is the square root of its variance.
- This is used for test statistics quite frequently.
- If you assume that test scores follow a specific type of distribution known as the **bell curve** (which is often the case for well-constructed tests), then the standard deviation tells you how many students fall within a specific range of scores.
- 94% will fall within +/- 2 standard deviations
- 68% will fall within +/- 1 standard deviations

USC Viterbi School of Engineering



XKCD #221

Simpson's Paradox (1)

- Alice and Bob are studying for the CSCI 104 final. They each do 25 practice problems, split over 2 types (trees and probability).
- Bob does 5 tree problems and gets 3 of them right.
- Alice does 20 tree problems and gets 13 of them right.
- Bob does 20 probability problems and gets 15 of them right.
- Alice does 5 probability problems and gets 4 of them right.

- Bob got a 60% on trees, Alice got a 65% (Alice did better).
- Bob got a 75% on probability, Alice got an 80% (Alice did better).
- Bob got 18 out of 25 problems right, Alice got 17 out of 25 right (Bob did better overall).

Simpson's Paradox (2)

- A new study promotes the benefits of a specific diet.
- 3 out of 8 of the residents of Roshar survive to age 70 on diet 1.
- 40 out of 90 of Roshar residents survive on diet 2.
- 60 out of 92 of the residents of Tyrea survive to age 70 on diet 1.
- 7 out of 10 of Tyrea residents survive on diet 2.
- 63% of people on diet 1 survive, vs. 47% on diet 2.
- This is not enough information to extol the benefits of diet 1!

Simpson's Paradox (3)

- 38% of the residents of Roshar survive to age 70 on diet 1.
- 44% of Roshar residents survive on the new diet (diet 2).
- 65% of the residents of Tyrea survive to age 70 on diet 1.
- 70% of Tyrea residents survive on the new diet.
- It looks like diet 2 is the better diet! The natural longevity of Tyreans made diet 1 look better than it was.

BACKUP

USC Viterbi School of Engineering

Hirings

- Let I_k return 1 if the kth best employee is hired, and 0 otherwise. We want:
- $E(\sum_{k=1}^{n} Ik) = \sum_{k=1}^{n} E(I_k)$
- E(I_1) = ___ (the best employee is always hired).
- E(I_2) = ___ (there is a ___% chance the best employee shows up before the 2nd best employee)
- E(I_3) $=$ __ (there is a ___% chance the 3rd best employee shows up
before the best two employees)
- $\sum_{k=1}^{n} E(I_k) =$

Hirings Solution

- Let I_k return 1 if the kth best employee is hired, and 0 otherwise. We want:
- $E(\sum_{k=1}^{n} Ik) = \sum_{k=1}^{n} E(I_k)$
- E(I_1) = 1 (the best employee is always hired).
- E(I_2) = $\frac{1}{2}$ (there is a 50% chance the best employee shows up before the 2nd best employee)
- E(I_3) $= \frac{1}{3}$ (there is a 33% chance the 3rd best employee shows up
  before the best two employees)
- $\sum_{k=1}^{n} E(I_k) = \sum_{k=1}^{n} \frac{1}{k} = \theta(\log n)$

- Professor Slacker hates grading, and just assigned each student a grade (A, B, C, or D) u.a.r. If he had actually bothered to grade, he would have given 25% of students each grade. What is the expected percentage of students that got the correct grade?
  - 25%

Find the Expected Value

- What is the expected number of guesses when I play Wordle?