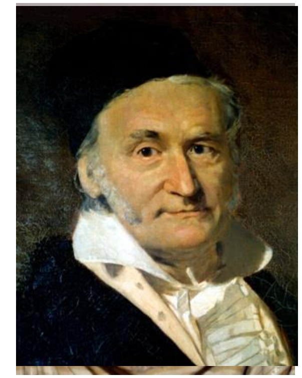
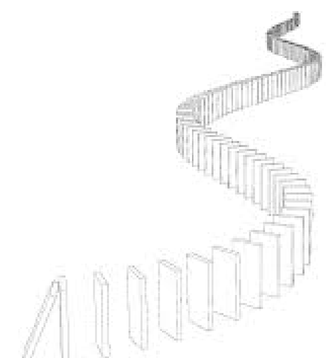


# COMP 1002

# Logic for Computer Scientists

## Lecture 29

B 5 2 J



# Expectations

- Often we are interested in what outcome we would see "on average".
  - How fast does this program run "on average"?
- Suppose that possible outcomes of an experiment are numbers $a_1, \ldots, a_n$
  - E.g., time a program takes to sort n elements
- Its **expected value (mean)** is $\Sigma_{k=1}^{n} a_k \mathrm{Pr}(a_k)$
  - Often phrased in terms of a "random variable" X, where X is a *function* from outcomes to numbers.
  - Write E(X) to mean the expectation of X.

# Random variables

- "Random variable" X is a *function* from outcomes to numbers.
  - In Computer Science applications, usually X counts something.
    - Number of heads out of n coin tosses.
    - Number of steps a program takes on an input
  - An *indicator* random variable X takes value 0 or 1 depending on whether an event occurred or not.

- Expectation of a random variable X is

$$\mathrm{E}(\mathrm{X}) = \sum_{i \in Outcomes} X(i) * \Pr(i)$$

# Expectation example

Suppose we roll two fair dice. What is the expected sum of their values?

- X can take values from 2 to 12.
  - Pr(X=2) = Pr(X=12) = 1/36
  - Pr(X =3) = Pr(X =11) =2/36=1/18,
  - Pr(X =4) = Pr(X =10) =3/36=1/12,
  - Pr(X =5) = Pr(X =9) =4/36=1/9,
  - Pr(X =6) = Pr(X =8) =5/36,
  - Pr(X =7) =6/36=1/6

$$E(X) = 2 \cdot \frac{1}{36} + 3 \cdot \frac{1}{18} + 4 \cdot \frac{1}{12} + 5 \cdot \frac{1}{9} + 6 \cdot \frac{5}{36} + 7 \cdot \frac{1}{6} + 8 \cdot \frac{5}{36} + 9 \cdot \frac{1}{9} + 10 \cdot \frac{1}{12} + 11 \cdot \frac{1}{18} + 12 \cdot \frac{1}{36} = 7$$

# Expected win in a lottery

- Rules of Lotto 6/49:
  - A player chooses 6 numbers, 1 to 49.
  - During a draw, 6 randomly generated numbers are revealed.
  - If all 6 numbers chosen by the player match 6 numbers in the draw, the player gets the jackpot of $5,000,000 or more.
    - There are also smaller prizes; let's ignore them for simplicity.
  - A ticket costs $3.
  - According to Atlantic Lotto Corporation, chances of winning the jackpot are roughly 1/14,000,000.

# Expected win in a lottery

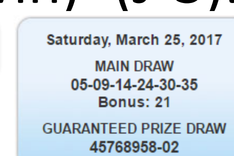
- Rules of Lotto 6/49:
  - A player chooses 6 numbers, 1 to 49.
  - During a draw, 6 randomly generated numbers are revealed.
  - If all 6 numbers chosen by the player match 6 numbers in the draw, the player gets the jackpot of $5,000,000 or more.
    - There are also smaller prizes; let's ignore them for simplicity.
  - A ticket costs $3.
  - According to Atlantic Lotto Corporation, chances of winning the jackpot are roughly 1/14,000,000.

- What is the expected amount a player would win if the jackpot is 5,000,000?
  - Pr(win) = 1/14,000,000. Pr(loss)=1-Pr(win) = 13,999,999/14,000,000.
  - Let the random variable X encode the amount a player wins.
    - For all but one player, that amount is -3. So Pr(X=-3)=Pr(loss)
    - For the lucky one, the amount is the jackpot minus ticket price. Pr(X=4,999,997)=Pr(win)
  - Expected amount to win is E(X) = Pr(loss)*(-3) +Pr(win)*(5,000,000-3) = *-2.64*
    - If counting smaller prizes , just add their amount*odds to the sum, and adjust Pr(loss)
    - E(X)=Pr(loss)*(-3)+Pr(jackpot)*(4,999,997)+Pr(5/6+bonus)*374,997+Pr(5/6)*312,497…

# Expected win in a lottery

- Rules of Lotto 6/49:
  - A player chooses 6 numbers, 1 to 49.
  - During a draw, 6 randomly generated numbers are revealed.
  - If all 6 numbers chosen by the player match 6 numbers in the draw, the player gets the jackpot of $5,000,000 or more.
  - A ticket costs $3.
  - According to Atlantic Lotto Corporation, chances of winning the jackpot are roughly 1/14,000,000.

- How large should be the jackpot so that the players expect at least to break even?
  - Let's call the jackpot amount J.
  - Expected amount to win is E(X) = Pr(loss)*(-3) +Pr(win)*(J-3).
    - To break even, want E(X)=0.
  - J= 3+(E(X)-Pr(loss)*(-3))/Pr(win) = *42,000,000*



# Bernoulli trials and repeated experiments


- Suppose an experiment has two outcomes, 1 and 0 (success/failure), with Pr(1) = p.
  - Such experiment is called a **Bernoulli trial**.
- What happens if the experiment is repeated multiple times (independently from each other?)

  - A sample space after carrying out n Bernoulli trials is a set of all possible n-tuples of elements in {0,1} (or {success, fail}).
  - Number of n-tuples with k 1s is $\binom{n}{k}$
  - Probability of getting 1 in any given trial is p, of getting 0 is (1-p).
  - Probability of getting exactly k 1s (successes) out of n trials is $\binom{n}{k}p^k(1-p)^{n-k}$
  - Probability of getting the first success on exactly the $k^{th}$ trial is $p(1-p)^{k-1}$
- How many trials do we need, on average, to get a success?

# Expected number until…

- Suppose that Alan insists on buying lottery tickets until he wins. If probability of winning is $p$, how many tickets in expectation ("on average") would he have to buy?
  - Let X be a random variable for how many tickets he has to buy.
  - The probability of winning on exactly $i^{th}$ ticket is $p(1-p)^{i-1}$
  - $E(X) = \Sigma_{i \in \mathbb{N}}\, i * \Pr(X = i) = \frac{1}{p}$
    - So for Lotto 6/49 he'd have to buy 14,000,000 tickets (and spend $42,000,000 -- that's jackpot that would let him break even! )

# Expected number until…

- Suppose we have Bernoulli trials with success probability $p$. What is the expected number of trials to see success?
  – Let X be a random variable for the number of steps till success.
  – $E(X) = \Sigma_{i \in \mathbb{N}}\, i * \Pr(X = i) = \frac{1}{p}$
- Same reasoning applies to other processes, where there is a fixed probability of something happening at each experiment or time step.
  – A system has a 1% probability of hanging in any given hour. How long, on average, will it stay up?
    - 100 hours: a little over 4 days.

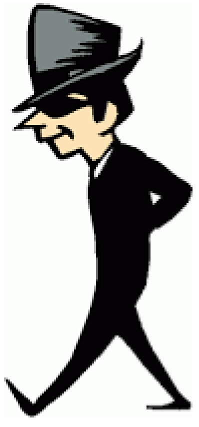
# Linearity of expectation

- Expectation is a very well-behaved operation:
  - $E(X_1 + X_2 + \cdots + X_n) = E(X_1) + \cdots + E(X_n)$
  - $E(aX + b) = a\,E(X) + b$
    - Where $X_1 \ldots X_n$ are random variables on some sample space S, and $a, b \in \mathbb{R}$

- Proof:
  - $E(X_1 + X_2) = \Sigma_{s\in S}\, p(s)(X_1(s) + X_2(s))$
    $= \Sigma_{s\in S}\, p(s)X_1(s) + \Sigma_{s\in S}\, p(s)X_2(s)$
    $= E(X_1) + E(X_2)$
  - Similar for $E(aX + b) = a\,E(X) + b$
    - Using the fact that $\Sigma_{s\in S}\, p(s)$ =1

# Hat-check problem

- Suppose n men came to an event, and checked in their hats at the door.
  - On the way out, in a hurry, they each picked up a random hat.
  - On average, how many men picked their own hat?



# Hat-check problem

- Suppose n men came to an event, and checked in their hats at the door.
  - On the way out, in a hurry, they each picked up a random hat.
  - How many men are expected to have picked their own hat?
- For each man, introduce a random variable $X_i$, where $X_i = 1$ iff he picked his own hat
  - Such random variables are called **indicator variables**.
  - The quantity we want is $E(X_1 + \cdots + X_n)$

  - Now, for each $X_i$, $E(X_i) = 1 \cdot \Pr(X_i = 1) = \frac{1}{n}$
  - By linearity of expectation, $E(X_1 + \cdots + X_n) = E(X_1) + \cdots + E(X_n) = \frac{1}{n} + \cdots + \frac{1}{n} = n \cdot \frac{1}{n} = 1$
- So on average, just one man will go home with his own hat!

# Tower of Hanoi game

- Rules of the game:
  - Start with all disks on the first peg.
  - At any step, can move a disk to another peg, as long as it is not placed on top of a smaller disk.
  - Goal: move the whole tower onto the second peg.
- Question: how many steps are needed to move the tower of 8 disks? How about n disks?