Welcome to St. Petersburg!

- Game set-up
  - We have a fair coin (come up "heads" with \( p = 0.5 \))
  - Let \( n \) = number of coin flips before first "tails"
  - You win \$2^n
- How much would you pay to play?
- Solution
  - Let \( X \) = your winnings
  - \( E[X] = \sum_{i=0}^{\infty} \left( \frac{1}{2} \right)^i 2^i = \sum_{i=0}^{\infty} \frac{1}{2} = \infty \)
  - I’ll let you play for \$1 million... but just once! Takers?

Vegas Breaks You

- Why doesn’t everyone do this?
  - Real games have maximum bet amounts
  - You have finite money
    - Not be able to keep doubling bet beyond certain point
  - Casinos can kick you out
- But, if you had:
  - No betting limits, and
  - Infinite money, and
  - Could play as often as you want...
- Then, go for it!
  - And tell me which planet you are living on

Breaking Vegas

- Consider even money bet (e.g., bet "Red" in roulette)
  - \( p = 18/38 \) you win \$Y, otherwise \((1 - p)\) you lose \$Y
- Consider this algorithm for one series of bets:
  1. \( Y = \$1 \)
  2. Bet \( Y \)
  3. If Win, stop
  4. If Loss, \( Y = 2 \times Y \), goto 2
- Let \( Z = \) winnings upon stopping
  - \( E[Z] = \left( \frac{18}{38} \right) [20 \left( \frac{18}{38} \right) 2 - 1] + \left( \frac{20}{38} \right) \left[ \left( \frac{18}{38} \right) 4 - 2 - 1 \right] + \ldots \)
  - \( = \sum_{j=0}^{\infty} \left( \frac{18}{38} \right) \left( 2^j - \sum_{i=1}^{j} 2^j \right) = \left( \frac{18}{38} \right) \left( \frac{20}{38} \right) = \left( \frac{18}{38} \right) \left( \frac{1}{1 - \frac{20}{38}} \right) = 1 \)
- Expected winnings \( \geq 0 \). Use algorithm infinitely often!

Variance

- Consider the following 3 distributions (PMFs)
  - All have the same expected value, \( E[X] = 3 \)
  - But “spread” in distributions is different
- Variance = a formal quantification of “spread

Computing Variance

\[
\text{Var}(X) = E[(X - \mu)^2] = \sum_{x} (x - \mu)^2 p(x) = \sum_{x} (x^2 - 2\mu x + \mu^2) p(x) = \sum_{x} x^2 p(x) - 2\mu \sum_{x} x p(x) + \mu^2 \sum_{x} p(x) = E[X^2] - 2\mu E[X] + \mu^2 = E[X^2] - 2\mu^2 + \mu^2 = E[X^2] - \mu^2 = E[X^2] - (E[X])^2
\]

Ladies and gentlemen, please welcome the 2nd moment!
### Variance of 6 Sided Die

- Let $X =$ value on roll of 6 sided die
- Recall that $E[X] = 7/2$
- Compute $E[X^2]$

\[
E[X^2] = \left( \left( \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{array} \right) \right)^2 = \left( \begin{array}{c} \frac{1}{6} \\ \frac{1}{6} \\ \frac{1}{6} \\ \frac{1}{6} \\ \frac{1}{6} \\ \frac{1}{6} \end{array} \right) = \frac{91}{6}
\]

\[
\text{Var}(X) = E[X^2] - (E[X])^2 = \frac{91}{6} - \left( \frac{7}{2} \right)^2 = \frac{25}{12}
\]

### Properties of Variance

- $\text{Var}(aX + b) = a^2 \text{Var}(X)$
- Proof:

\[
\text{Var}(aX + b) = E[(aX + b)^2] - (E[aX + b])^2 = a^2 E[X^2] + 2abE[X] + b^2 - (aE[X] + b)^2
\]

\[
= a^2 E[X^2] - a^2 (E[X])^2 = a^2 (E[X^2] - (E[X])^2) = a^2 \text{Var}(X)
\]

- Standard Deviation of $X$, denoted $SD(X)$, is:

\[
SD(X) = \sqrt{\text{Var}(X)}
\]

### Jacob Bernoulli

- Jacob Bernoulli (1654-1705), also known as "James", was a Swiss mathematician

- One of many mathematicians in Bernoulli family
- The Bernoulli Random Variable is named for him
- He is my academic 11th-grandfather
- Resemblance to Charlie Sheen weak at best

### Bernoulli Random Variable

- Experiment results in “Success” or “Failure”
- $X$ is random indicator variable ($1 = \text{success}, 0 = \text{failure}$)
- $P(X = 1) = p(1) = p$
- $P(X = 0) = p(0) = 1 - p$
- $X$ is a Bernoulli Random Variable: $X \sim \text{Ber}(p)$
- $E[X] = p$
- $\text{Var}(X) = p(1 - p)$
- Examples
  - coin flip
  - random binary digit
  - whether a disk drive crashed

### Binomial Random Variable

- Consider $n$ independent trials of $\text{Ber}(p)$ rand. var.
  - $X$ is number of successes in $n$ trials
  - $X$ is a Binomial Random Variable: $X \sim \text{Bin}(n, p)$

\[
P(X = i) = \binom{n}{i} p^i (1-p)^{n-i}, \quad i = 0, 1, ..., n
\]

- By Binomial Theorem, we know that $\sum_{i=0}^{n} P(X = i) = 1$
- Examples
  - # of heads in $n$ coin flips
  - # of 1’s in randomly generated length $n$ bit string
  - # of disk drives crashed in 1000 computer cluster
    - Assuming disks crash independently

### Three Coin Flips

- Three fair (“heads” with $p = 0.5$) coins are flipped
- $X$ is number of heads
- $X \sim \text{Bin}(3, 0.5)$

\[
P(X = 0) = \binom{3}{0} (0.5)^0 (1-0.5)^3 = \frac{1}{8}
\]

\[
P(X = 1) = \binom{3}{1} (0.5)^1 (1-0.5)^2 = \frac{3}{8}
\]

\[
P(X = 2) = \binom{3}{2} (0.5)^2 (1-0.5)^1 = \frac{3}{8}
\]

\[
P(X = 3) = \binom{3}{3} (0.5)^3 (1-0.5)^0 = \frac{1}{8}
\]
Error Correcting Codes

- Error correcting codes
  - Have original 4 bit string to send over network
  - Add 3 “parity” bits, and send 7 bits total
  - Each bit independently corrupted (flipped) in transition with probability 0.1
  - X = number of bits corrupted: \( X \sim \text{Bin}(7, 0.1) \)
  - P(a correctable message is received)?
    \[ P(X = 0) + P(X = 1) = 0.8503 \]

Using error correction improves reliability ~30%!

Genetic Inheritance

- Person has 2 genes for trait (eye color)
  - Child receives 1 gene (equally likely) from each parent
  - Child has brown eyes if either (or both) genes brown
  - Child only has blue eyes if both genes blue
  - Brown is “dominant” (d), Blue is recessive (r)
  - Parents each have 1 brown and 1 blue gene
  - 4 children, what is \( P(3 \text{ children with brown eyes})? \)
    \[ P(X = 3) = \binom{4}{3} \left( \frac{1}{2} \right)^3 \left( \frac{1}{2} \right)^1 = 0.75 \]

Properties of Bin(n, p)

- We have \( X \sim \text{Bin}(n, p) \)
  \[ E[X^i] = \sum_{j=i}^{n} \binom{n}{j} p^j (1-p)^{n-j} \]
  - Noting that: \( \binom{n}{i} = \frac{n!}{i!(n-i)!} \)
  \[ E[X^i] = np \sum_{j=i}^{n} \binom{n-i}{j} p^{j} (1-p)^{(n-j)} = np \sum_{j=i}^{n} \binom{n-i}{j-1} p^{j-1} (1-p)^{(n-j+1)} = npE[Y^{i-1}], \text{ where } Y \sim \text{Bin}(n-1, p) \]
  - Set \( k = 1 \rightarrow E[X] = np \)
  - Set \( k = 2 \rightarrow E[X^2] = npE[Y + 1] = np[(n-1)p + 1] \)
  - \( \text{Var}(X) = np[(n-1)p + 1] - (np)^2 = np(1-p) \)
  - Note: \( \text{Ber}(p) = \text{Bin}(1, p) \)

PMF for X \( \sim \text{Bin}(10, 0.5) \)

PMF for X \( \sim \text{Bin}(10, 0.3) \)
Power of Your Vote

- Is it better to vote in small or large state?
  - Small: more likely your vote changes outcome
  - Large: larger outcome (electoral votes) if state swings
  - a (= 2n) voters equally likely to vote for either candidate
- You are deciding (a + 1)st vote
  \[ P(2n \text{ voters tie}) = \binom{2n}{n} \frac{1}{2^n} \binom{1}{2} = \frac{(2n)!}{n!n!2^n} \]
- Use Stirling’s Approximation: \[ n! \approx n^{n+\frac{1}{2}}e^{-n}\sqrt{2\pi} \]
  \[ P(2n \text{ voters tie}) \approx \frac{(2n)^{2n+\frac{1}{2}}e^{-2n}\sqrt{2\pi}}{n^{2n}e^{-2n}\sqrt{2\pi}2^n} = \frac{1}{\sqrt{n\pi}} \]
- Power = P(tie) * Elec. Votes = \[ \frac{1}{\sqrt{(a/2)\pi}} (ac) = \frac{c\sqrt{2\pi}}{\sqrt{\pi}}. \]
- Larger state = more power