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] $= \left(\frac{1}{2}\right)^1 2^0 + \left(\frac{1}{2}\right)^2 2^1 + \left(\frac{1}{2}\right)^3 2^2 + \left(\frac{1}{2}\right)^4 2^3 + \ldots = \sum_{i=0}^{\infty} \left(\frac{1}{2}\right)^{i+1} 2^i$

    $= \sum_{i=0}^{\infty} \frac{1}{2} = \infty$
  - I'll let you play for $1 million... but just once! Takers?

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 * Y, goto 2
  - Let Z = winnings upon stopping
  - E[Z] $= \left(\frac{18}{38}\right)1 + \left(\frac{20}{38}\right)\left(\frac{18}{38}\right)(2-1) + \left(\frac{20}{38}\right)^2\left(\frac{18}{38}\right)(4-2-1) + \ldots$

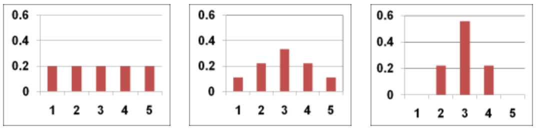
    $= \sum_{i=0}^{\infty}\left(\frac{20}{38}\right)^i\left(\frac{18}{38}\right)\left(2^i - \sum_{j=1}^{i} 2^{j-1}\right) = \left(\frac{18}{38}\right)\sum_{i=0}^{\infty}\left(\frac{20}{38}\right)^i = \left(\frac{18}{38}\right)\frac{1}{1-\frac{20}{38}} = 1$
  - Expected winnings ≥ 0. Use algorithm infinitely often!

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

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"

Variance

- If X is a random variable with mean μ then the **variance** of X, denoted Var(X), is:

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

- Note: Var(X) ≥ 0
- Also known as the 2nd Central Moment, or square of the Standard Deviation

Computing Variance

$$\begin{aligned}
\text{Var}(X) &= E[(X-\mu)^2] \\
&= \sum_x (x-\mu)^2 p(x) \\
&= \sum_x (x^2 - 2\mu x + \mu^2) p(x) \\
&= \sum_x x^2 p(x) - 2\mu \sum_x x p(x) + \mu^2 \sum_x p(x) \\
&= \boxed{E[X^2]} - 2\mu E[X] + \mu^2 \\
&= E[X^2] - 2\mu^2 + \mu^2 \\
&= E[X^2] - \mu^2 \\
&= \boxed{E[X^2] - (E[X])^2}
\end{aligned}$$

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[X²]

$$E[X^2] = (1^2)\frac{1}{6} + (2^2)\frac{1}{6} + (3^2)\frac{1}{6} + (4^2)\frac{1}{6} + (5^2)\frac{1}{6} + (6^2)\frac{1}{6} = \frac{91}{6}$$

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

$$= \frac{91}{6} - \left(\frac{7}{2}\right)^2 = \frac{35}{12}$$

Properties of Variance

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

    Var(aX + b) $= E[(aX + b)^2] – (E[aX + b])^2$

    $= E[a^2X^2 + 2abX + b^2] – (aE[X] + b)^2$

    $= a^2E[X^2] + 2abE[X] + b^2 – (a^2(E[X])^2 + 2abE[X] + b^2)$

    $= a^2E[X^2] – a^2(E[X])^2 = a^2(E[X^2] – (E[X])^2)$

    $= a^2\text{Var}(X)$
- Standard Deviation of X, denoted SD(X), is:

$$\text{SD}(X) = \sqrt{\text{Var}(X)}$$

  - Var(X) is in units of X^2
  - SD(X) is in same units as 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* great[11]-grandfather
- Resemblance to Charlie Sheen weak at best

Bernoulli Random Variable

- Experiment results in "Success" or "Failure"
  - X is random indicator variable (1 = success, 0 = failure)
  - P(X = 1) = p(1) = p   P(X = 0) = p(0) = 1 – p
  - X is a **<u>Bernoulli</u>** Random Variable: X ~ Ber(p)
  - E[X] = p
  - Var(X) = p(1 – p)
- Examples
  - coin flip
  - random binary digit
  - whether a disk drive crashed

Binomial Random Variable

- Consider *n* independent trials of Ber(p) rand. var.
  - X is number of successes in *n* trials
  - X is a **<u>Binomial</u>** Random Variable: X ~ Bin(n, p)

$$P(X = i) = p(i) = \binom{n}{i} p^i (1-p)^{n-i} \quad i = 0,1,\ldots,n$$

  - By Binomial Theorem, we know that $\sum_{i=0}^{\infty} P(X = i) = 1$
- 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 ~ Bin(3, 0.5)

$$P(X = 0) = \binom{3}{0} p^0 (1-p)^3 = \frac{1}{8}$$

$$P(X = 1) = \binom{3}{1} p^1 (1-p)^2 = \frac{3}{8}$$

$$P(X = 2) = \binom{3}{2} p^2 (1-p)^1 = \frac{3}{8}$$

$$P(X = 3) = \binom{3}{3} p^3 (1-p)^0 = \frac{1}{8}$$

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 ~ Bin(7, 0.1)
  - But, parity bits allow us to correct at most 1 bit error
- P(a correctable message is received)?
  - P(X = 0) + P(X = 1)

Error Correcting Codes (cont)

- Using error correcting codes: X ~ Bin(7, 0.1)

$$P(X=0)=\binom{7}{0}(0.1)^0(0.9)^7 \approx 0.4783$$

$$P(X=1)=\binom{7}{1}(0.1)^1(0.9)^6 \approx 0.3720$$

  - P(X = 0) + P(X = 1) = 0.8503
- What if we didn’t use error correcting codes?
  - X ~ Bin(4, 0.1)
  - P(correct message received) = P(X = 0)

$$P(X=0)=\binom{4}{0}(0.1)^0(0.9)^4 = 0.6561$$

- 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 children with brown eyes)?
  - Child has blue eyes: p = (½) (½) = ¼ (2 blue genes)
  - P(child has brown eyes) = 1 – (¼) = 0.75
  - X = # of children with brown eyes. X ~ Bin(4, 0.75)

$$P(X=3)=\binom{4}{3}(0.75)^3(0.25)^1 \approx 0.4219$$

Properties of Bin(n, p)

- We have X ~ Bin(n, p)

$$E[X^k]=\sum_{i=0}^{n} i^k \binom{n}{i} p^i (1-p)^{n-i} = \sum_{i=1}^{n} i^k \binom{n}{i} p^i (1-p)^{n-i}$$

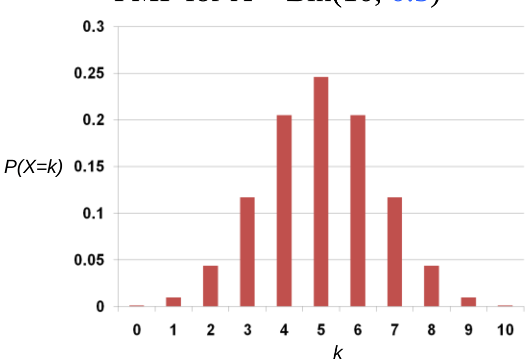
  - Noting that: $i\binom{n}{i} = \frac{i\,n!}{i!(n-i)!} = \frac{n(n-1)!}{(i-1)!((n-1)-(i-1))!} = n\binom{n-1}{i-1}$

$$E[X^k]=np\sum_{i=1}^{n} i^{k-1}\binom{n-1}{i-1}p^{i-1}(1-p)^{n-i} = np\sum_{j=0}^{n-1}(j+1)^{k-1}\binom{n-1}{j}p^j(1-p)^{n-(j+1)}, \quad \text{where } i=j+1$$

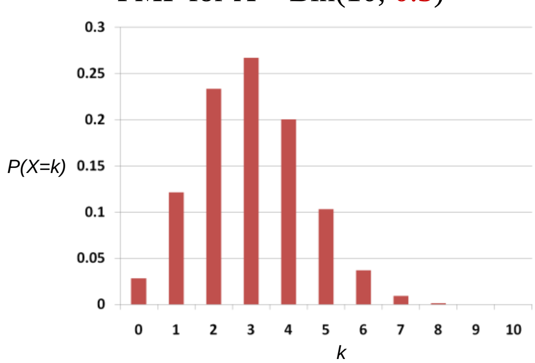
$$= npE[(Y+1)^{k-1}], \quad \text{where } Y \sim Bin(n-1, p)$$

  - Set $k = 1$ → E[X] = np
  - Set $k = 2$ → E[X^2] = np E[Y + 1] = np[(n – 1)p + 1]
  - Var(X) = np[(n – 1)p + 1] – $(np)^2$ = np(1 – p)
- Note: Ber(p) = Bin(1, p)

PMF for X ~ Bin(10, 0.5)



PMF for X ~ 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

$$P(2n \text{ voters tie}) = \binom{2n}{n}\left(\frac{1}{2}\right)^n\left(\frac{1}{2}\right)^n = \frac{(2n)!}{n!n!2^{2n}}$$

  - Use Stirling's Approximation: $n! \approx n^{n+1/2}e^{-n}\sqrt{2\pi}$

$$P(2n \text{ voters tie}) \approx \frac{(2n)^{2n+1/2}e^{-2n}\sqrt{2\pi}}{n^{2n+1}e^{-2n}2\pi 2^{2n}} = \frac{1}{\sqrt{n\pi}}$$

  - Power = P(tie) * Elec. Votes = $\frac{1}{\sqrt{(a/2)\pi}}(ac) = \frac{c\sqrt{2a}}{\sqrt{\pi}}$
  - Larger state = more power