DNA Paternity Testing

- Child is born with (A, a) gene pair (event $B_{A,a}$)
  - Mother has (A, A) gene pair
  - Two possible fathers: M_1: (a, a) M_2: (a, A)
  - $P(M_1) = p$ $P(M_2) = 1 - p$
  - What is $P(M_1 \mid B_{A,a})$?
- Solution
  - $P(M_1 \mid B_{A,a}) = P(M_1\ B_{A,a}) / P(B_{A,a})$

$$= \frac{P(B_{A,a} \mid M_1)P(M_1)}{P(B_{A,a} \mid M_1)P(M_1) + P(B_{A,a} \mid M_2)P(M_2)}$$

$$= \frac{1 \cdot p}{1 \cdot p + \frac{1}{2}(1 - p)} = \frac{2p}{1 + p} > p$$

M_1 more likely to be father than he was before, since $P(M_1 \mid B_{A,a}) > P(M_1)$

Reminder of Geometric Series

- Geometric series: $x^0 + x^1 + x^2 + x^3 + ... + x^n = \sum_{i=0}^{n} x^i$

- From your "Calculation Reference" handout:

$$\sum_{i=0}^{n} x^i = \frac{1-x^{n+1}}{1-x}$$

- As $n \to \infty$, and $|x| < 1$, then

$$\sum_{i=0}^{n} x^i = \frac{1-x^{n+1}}{1-x} \to \frac{1}{1-x}$$

Simplified Craps

- Two 6-sided dice repeatedly rolled (roll = ind. trial)
  - E = 5 is rolled before a 7 is rolled
  - What is P(E)?
- Solution
  - F_n = no 5 or 7 rolled in first n – 1 trials, 5 rolled on n^{th} trial
  - P(E) = $P\left(\bigcup_{n=1}^{\infty} F_n\right) = \sum_{n=1}^{\infty} P(F_n)$
  - P(5 on any trial) = 4/36 P(7 on any trial) = 6/36
  - $P(F_n) = (1 - (10/36))^{n-1} (4/36) = (26/36)^{n-1} (4/36)$
  - P(E) = $\frac{4}{36}\sum_{n=1}^{\infty}\left(\frac{26}{36}\right)^{n-1} = \frac{4}{36}\sum_{n=0}^{\infty}\left(\frac{26}{36}\right)^{n} = \frac{4}{36}\frac{1}{\left(1-\frac{26}{36}\right)} = \frac{2}{5}$

From Urns to Coupons

- "Coupon Collecting" is classic probability problem
  - There exist N different types of coupons
  - Each is collected with some probability p_i ($1 \le i \le N$)
- Ask questions like:
  - After you collect m coupons, what is probability you have k different kinds?
  - What is probability that you have ≥ 1 of each N coupon types after you collect m coupons?
- You've seen concept (in a more practical way)
  - N coupon types = N buckets in hash table
  - collecting a coupon = hashing a string to a bucket

Digging Deeper on Independence

- Recall, two events E and F are called independent if

  $$P(EF) = P(E)\,P(F)$$

- If E and F are independent, does that tell us whether the following is true or not:

  $$P(EF \mid G) = P(E \mid G)\,P(F \mid G),$$

  where G is an arbitrary event?

- In general, No!

Not-so Independent Dice

- Roll two 6-sided dice, yielding values D_1 and D_2
  - Let E be event: $D_1 = 1$
  - Let F be event: $D_2 = 6$
  - Let G be event: $D_1 + D_2 = 7$
- E and F are independent
  - $P(E) = 1/6$, $P(F) = 1/6$, $P(EF) = 1/36$
- Now condition both E and F on G:
  - $P(E|G) = 1/6$, $P(F|G) = 1/6$, $P(EF|G) = 1/6$
  - $P(EF|G) \neq P(E|G)\, P(F|G)$ → E|G and F|G ***dependent***
- Independent events can become dependent by conditioning on additional information

Do CS Majors Get Fewer A's?

- Say you are in a dorm with 100 students
  - 20 of the students are CS majors: $P(CS) = 0.2$
  - 30 of the students get straight A's: $P(A) = 0.3$
  - 6 students are CS majors who get straight A's
    - $P(CS, A) = 0.06$
    - $P(CS, A) = P(CS)P(A)$, so CS and A are independent
  - At faculty night, only CS majors and A students show up
    - So, 44 (= 20 + 30 – 6) students arrive
    - Of 44 students, 30 get A's $\Rightarrow$ $P(A \mid \text{faculty night}) = 30/44 \approx 0.68$
    - But, $P(A \mid \text{faculty night}, CS) = 6/20 = 0.3$
    - Appears that being CS major lowers probability of straight A's
    - But, weren't they supposed to be independent?
  - In fact, CS and A <u>conditionally dependent</u> at faculty night

Explaining Away

- Say you have a lawn
  - It gets watered by rain or sprinklers
  - P(rain) and P(sprinklers were on) are independent
  - Now, you come outside and see the grass is wet
    - You know that the sprinklers were on
    - Does that lower probability that rain was cause of wet grass?
  - This phenomena is called “explaining away”
    - One cause of an observation makes other causes less likely
  - Only CS majors and A students come to faculty night
    - Knowing you came because you’re a CS major makes it less likely you came because you get straight A’s

Conditioning Can Break Dependence

- Consider a randomly chosen day of the week
  - Let A be event: It is not Monday
  - Let B be event: It is Saturday
  - Let C be event: It is the weekend
- A and B are dependent
  - $P(A) = 6/7$, $P(B) = 1/7$, $P(AB) = 1/7 \neq (6/7)(1/7)$
- Now condition both A and B on C:
  - $P(A|C) = 1$, $P(B|C) = 1/2$, $P(AB|C) = 1/2$
  - $P(AB|C) = P(A|C)\, P(B|C) \rightarrow A|C$ and $B|C$ ***independent***
- Dependent events can become independent by conditioning on additional information

Conditional Independence

- Two events E and F are called ***conditionally independent given G***, if

  $$P(E\,F \mid G) = P(E \mid G)\,P(F \mid G)$$

  Or, equivalently: $P(E \mid F\,G) = P(E \mid G)$

- Exploiting conditional independence to generate fast probabilistic computations is one of the main contributions CS has made to probability theory
  - Judea Pearl wins 2011 Turing Award

    "For fundamental contributions to artificial intelligence through the development of a calculus for probabilistic and causal reasoning"

Random Variable

- A **Random Variable** is a real-valued function defined on a sample space
- Example:
  - 3 fair coins are flipped.
  - Y = number of “heads” on 3 coins
  - Y is a random variable
  - $P(Y = 0) = 1/8$ (T, T, T)
  - $P(Y = 1) = 3/8$ (H, T, T), (T, H, T), (T, T, H)
  - $P(Y = 2) = 3/8$ (H, H, T), (H, T, H), (T, H, H)
  - $P(Y = 3) = 1/8$ (H, H, H)
  - $P(Y \geq 4) = 0$

Binary Random Variables

- A binary random variable is a random variable with 2 possible outcomes (e.g., coin flip)
  - Now consider *n* coin flips, each which independently come up heads with probability *p*
  - Y = number of “heads” on *n* flips
  - P(Y = k) = $\binom{n}{k} p^k (1-p)^{n-k}$, where *k* = 0, 1, 2, ..., n
  - So, $\sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} = 1$
  - Proof: $\sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} = (p + (1-p))^n = 1^n = 1$

Simple Game

- Urn has 11 balls (3 green, 3 red, 5 black)
  - 3 balls drawn. +$1 for green, -$1 for red, $0 for black
  - Y = total winnings
  - P(Y = 0) = $\left[\binom{5}{3}+\binom{3}{1}\binom{3}{1}\binom{5}{1}\right]\Big/\binom{11}{3}=\frac{55}{165}$
  - P(Y = 1) = $\left[\binom{3}{1}\binom{5}{2}+\binom{3}{2}\binom{3}{1}\right]\Big/\binom{11}{3}=\frac{39}{165}$ = P(Y = -1)
  - P(Y = 2) = $\binom{3}{2}\binom{5}{1}\Big/\binom{11}{3}=\frac{15}{165}$ = P(Y = -2)
  - P(Y = 3) = $\binom{3}{3}\Big/\binom{11}{3}=\frac{1}{165}$ = P(Y = -3)

Probability Mass Functions

- A random variable *X* is <u>discrete</u> if it has countably many values (e.g., x_1, x_2, x_3, …)

- Probability Mass Function (PMF) of a discrete random variable is:

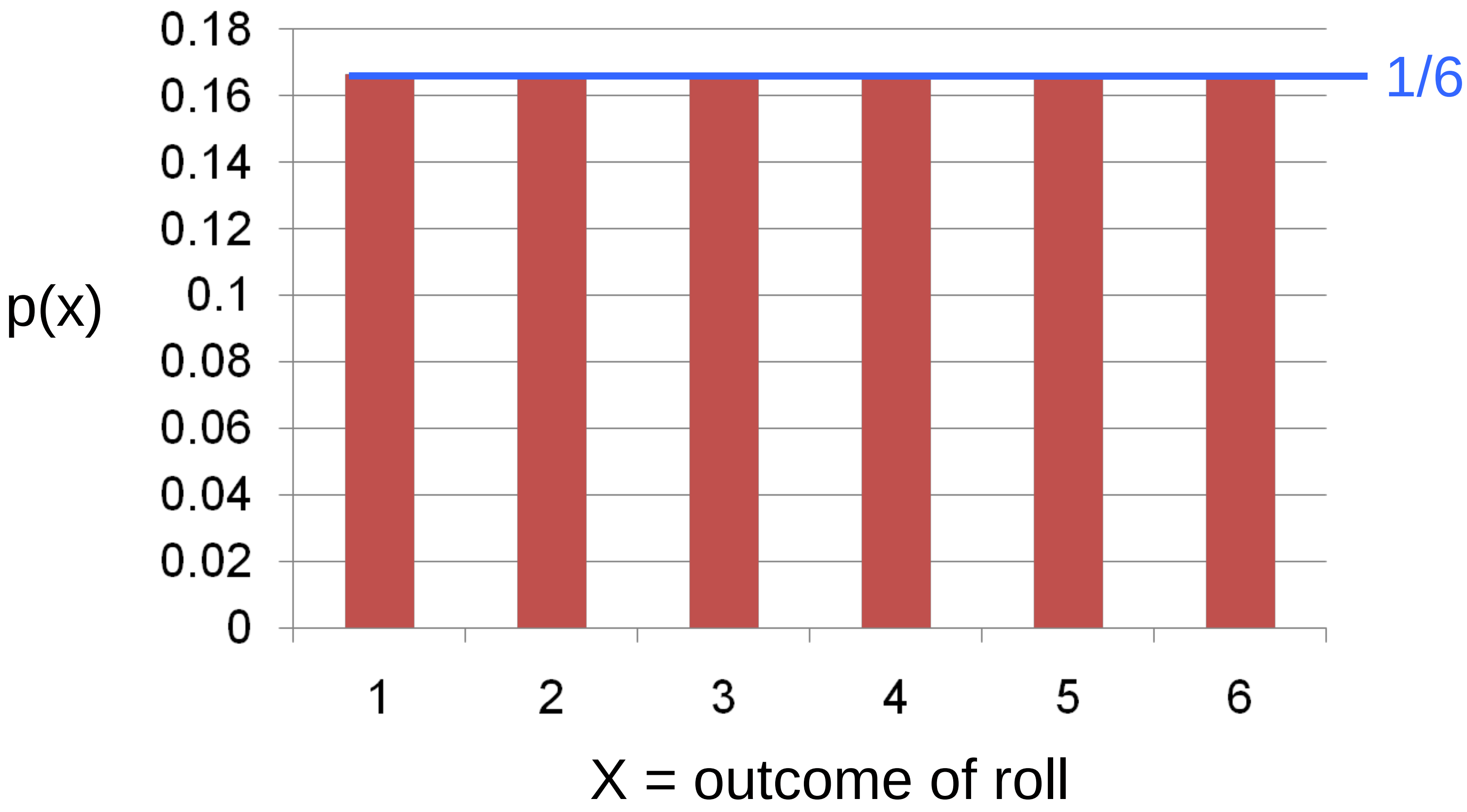
$$p(a) = P(X = a)$$

- Since $\sum_{i=1}^{\infty} p(x_i) = 1$, it follows that:

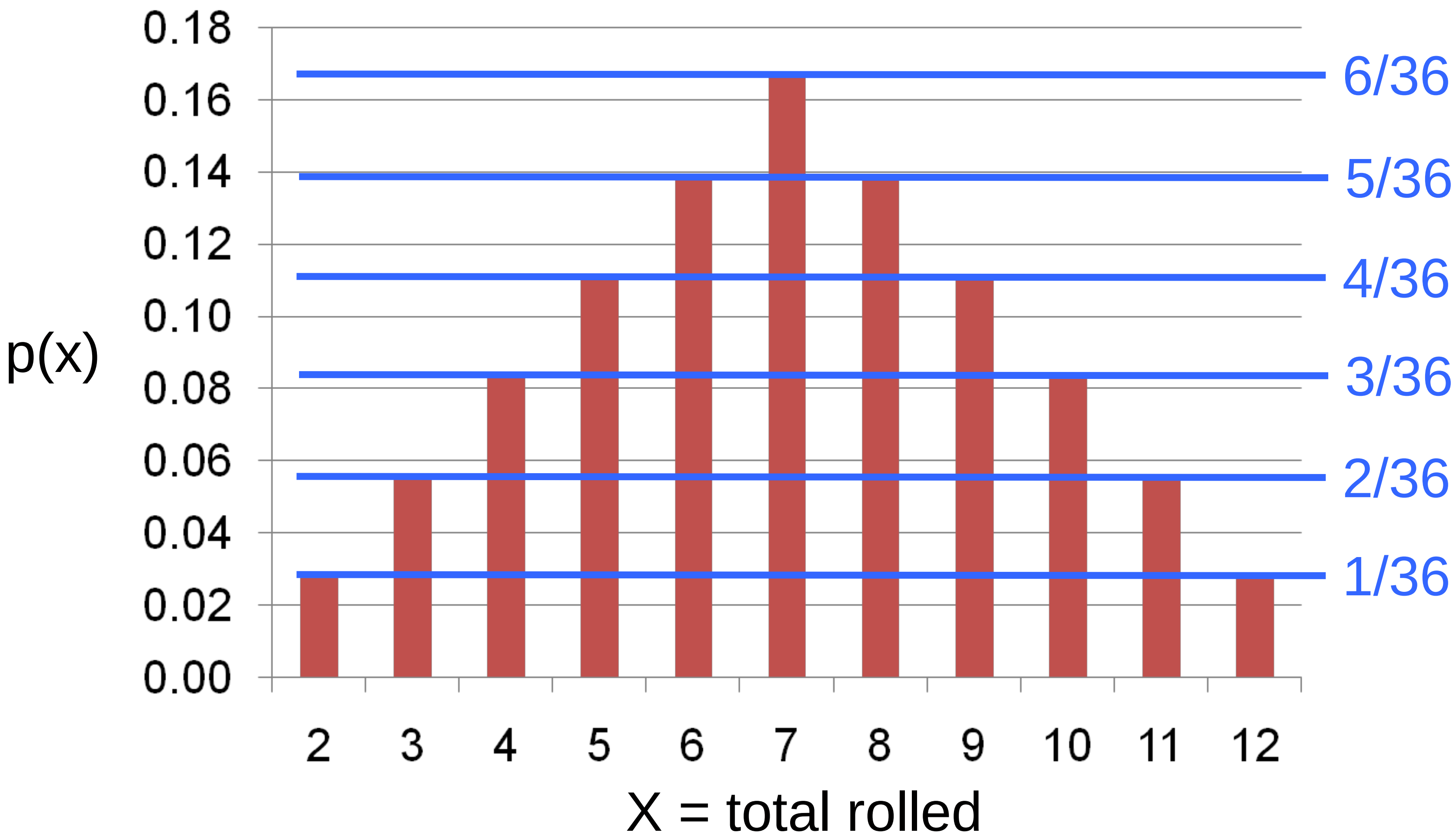
$$P(X = a) = \begin{cases} p(x_i) \geq 0 \text{ for } i = 1, 2, \ldots \\ p(x) = 0 \text{ otherwise} \end{cases}$$

where X can assume values x_1, x_2, x_3, …

PMF For a Single 6-Sided Die



PMF For a Roll of Two 6-Sided Dice



Cumulative Distribution Functions

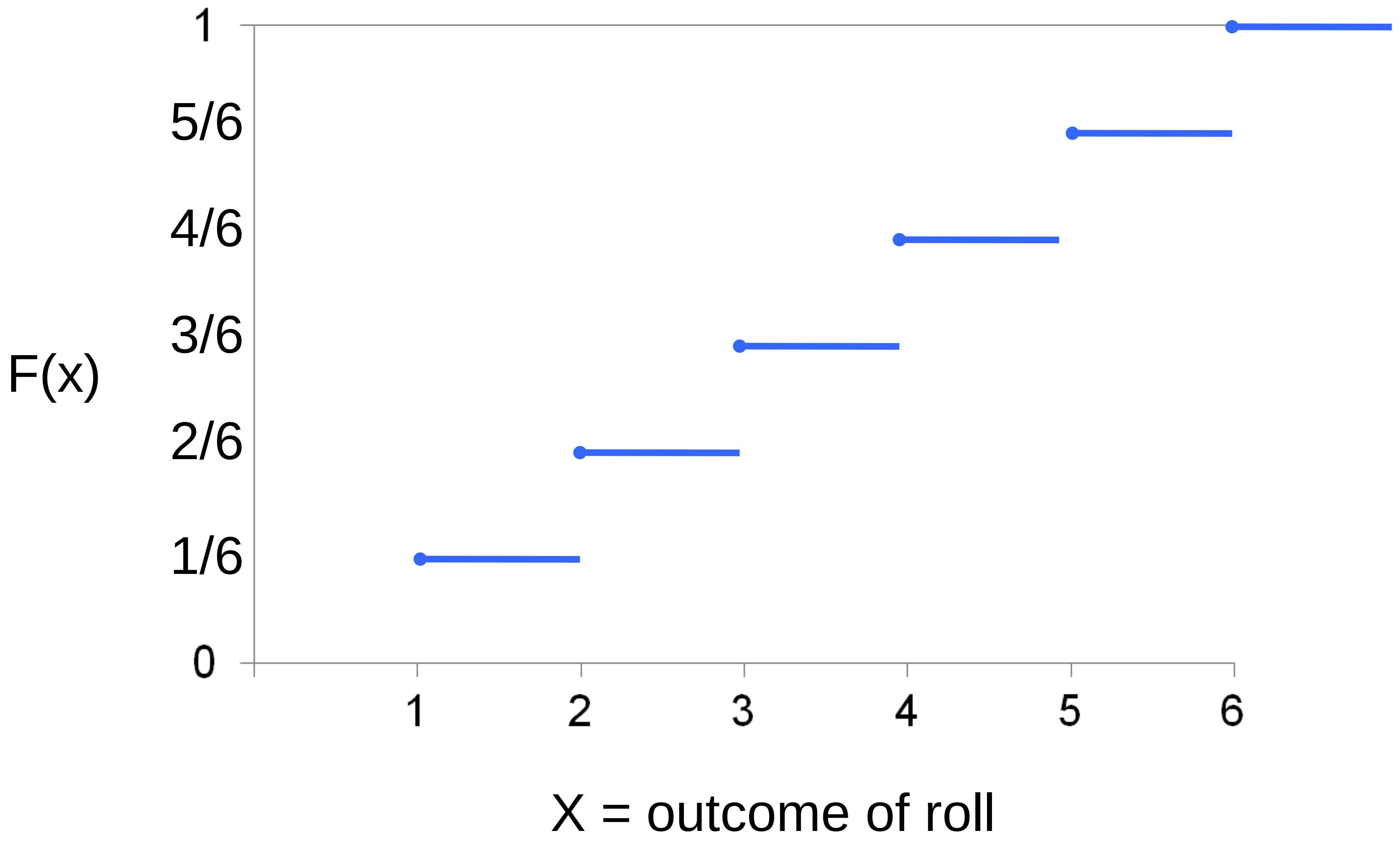
- For a random variable X, the Cumulative Distribution Function (CDF) is defined as:

$$F(a) = P(X \leq a) \quad \text{where } -\infty < a < \infty$$

- The CDF of a <u>discrete</u> random variable is:

$$F(a) = P(X \leq a) = \sum_{\text{all } x \leq a} p(x)$$

CDF For a Single 6-Sided Die



Expected Value

- The Expected Values for a discrete random variable X is defined as:

$$E[X] = \sum_{x:p(x)>0} x\, p(x)$$

- Note: sum over all values of x that have p(x) > 0.

- Expected value also called: *Mean*, *Expectation*, *Weighted Average*, *Center of Mass*, *1st Moment*

Expected Value Examples

- Roll a 6-Sided Die. X is outcome of roll
  - p(1) = p(2) = p(3) = p(4) = p(5) = p(6) = 1/6
- E[X] = $1\left(\frac{1}{6}\right)+2\left(\frac{1}{6}\right)+3\left(\frac{1}{6}\right)+4\left(\frac{1}{6}\right)+5\left(\frac{1}{6}\right)+6\left(\frac{1}{6}\right)=\frac{7}{2}$

- Y is random variable
  - P(Y = 1) = 1/3, P(Y = 2) = 1/6, P(Y = 3) = 1/2
- E[Y] = 1 (1/3) + 2 (1/6) + 3 (1/2) = 13/6

Indicator Variables

- A variable I is called an indicator variable for event *A* if

$$I = \begin{cases} 1 & \text{if } A \text{ occurs} \\ 0 & \text{if } A^c \text{ occurs} \end{cases}$$

- What is E[I]?
  - p(I =1) = P(A), p(I = 0) = 1 – P(A)
  - E[I] = 1 P(A) + 0 (1 – P(A)) = P(A)

We'll use this property frequently!