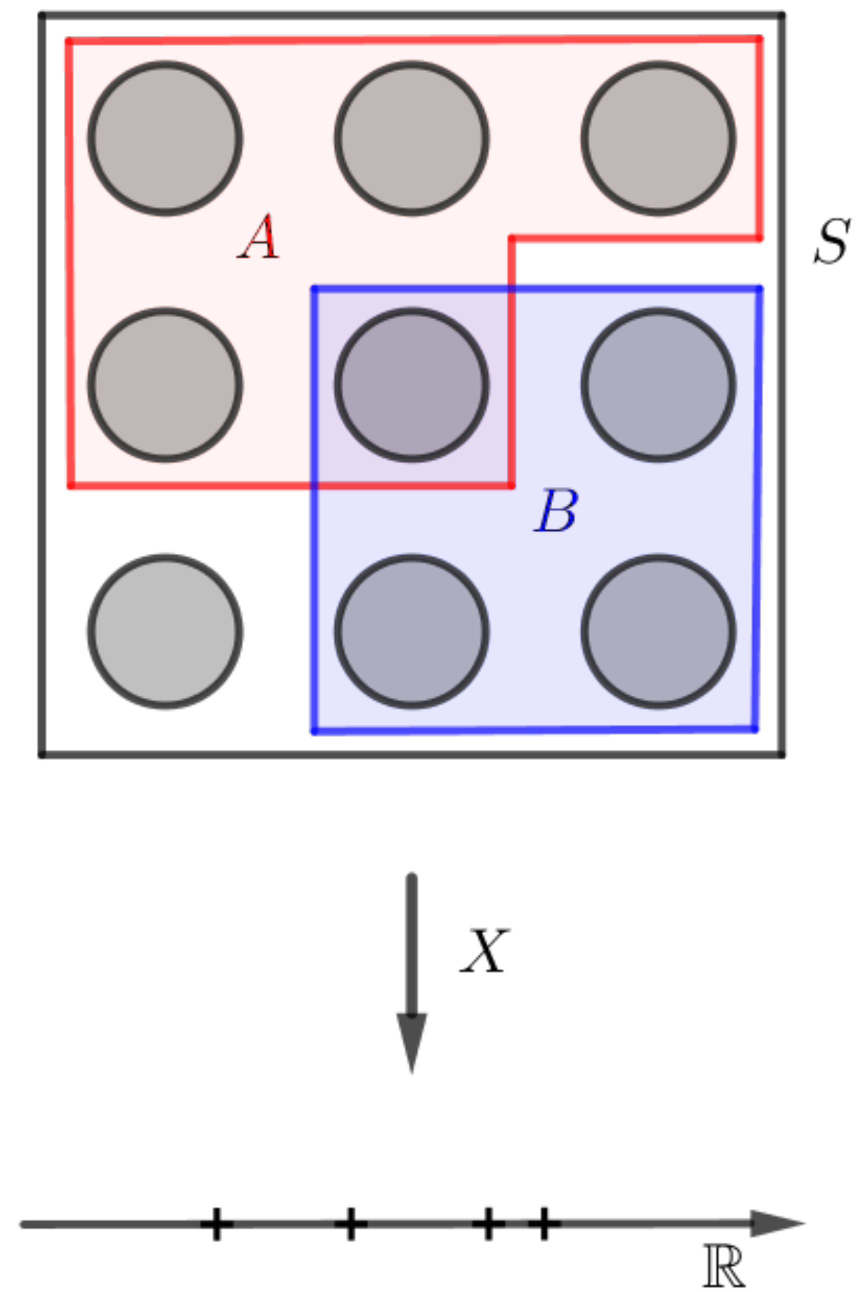
Lecture 4:

Random variables & their distributions

Random variables

Random variables

Definition 3.1.1 (Random variable). Given an experiment with sample space S, a **random variable** (r.v.) is a function from the sample space S to the real numbers $\mathbb{R}$. It is common to denote random variables by capital letters, like X.



Random variables

Example 3.1.2 (Coin tosses). We toss a fair coin twice. The sample space is $S = \{HH, HT, TH, TT\}$. Here are some r.v.-s on this space:

- $X =$ # of Heads:
  $X(HH) = 2,\ X(HT) = X(TH) = 1,\ X(TT) = 0$
- $Y =$ # of Tails: $Y = 2 - X$
- $I = \begin{cases} 1, \text{ if 1st toss = Heads} \\ 0, \text{ otherwise} \end{cases}$ – **indicator random variable**

Distributions and probability mass functions (PMFs)

Distributions and PMFs

There are two main types of r.v.-s: **discrete** and **continuous**. In this lecture, we'll focus on discrete r.v.-s

Definition 3.2.1 (Discrete random variable). A random variable X is said to be **discrete** if there is a finite list of values $a_1, a_2, \ldots, a_n$ or an infinite list $a_1, a_2, \ldots$ such that $P(X = a_j \text{ for some } j) = 1$. If X is a discrete r.v., then this finite or countably infinite set of values it takes and such that $P(X = x) > 0$ is called the **support** of X.

Continuous r.v.-s can take any real value in an interval, we'll discuss these in more detail in the next lecture.

Distributions and PMFs

The **distribution** of an r.v. specifies the probabilities of all events associated with the r.v. For a discrete r.v., the most natural way to do this is:

Definition 3.2.2 (Probability mass function). The **probability mass function (PMF)** of a discrete r.v. X is the function p_X given by $p_X(x) = P(X = x)$. It's > 0 if $x \in$ (support of X), and 0 otherwise.

In writing $P(X = x)$, $X = x$ denotes an **event**. (Sometimes also written as $\{X = x\}$ – formally, $\{s \in S : X(s) = x\}$

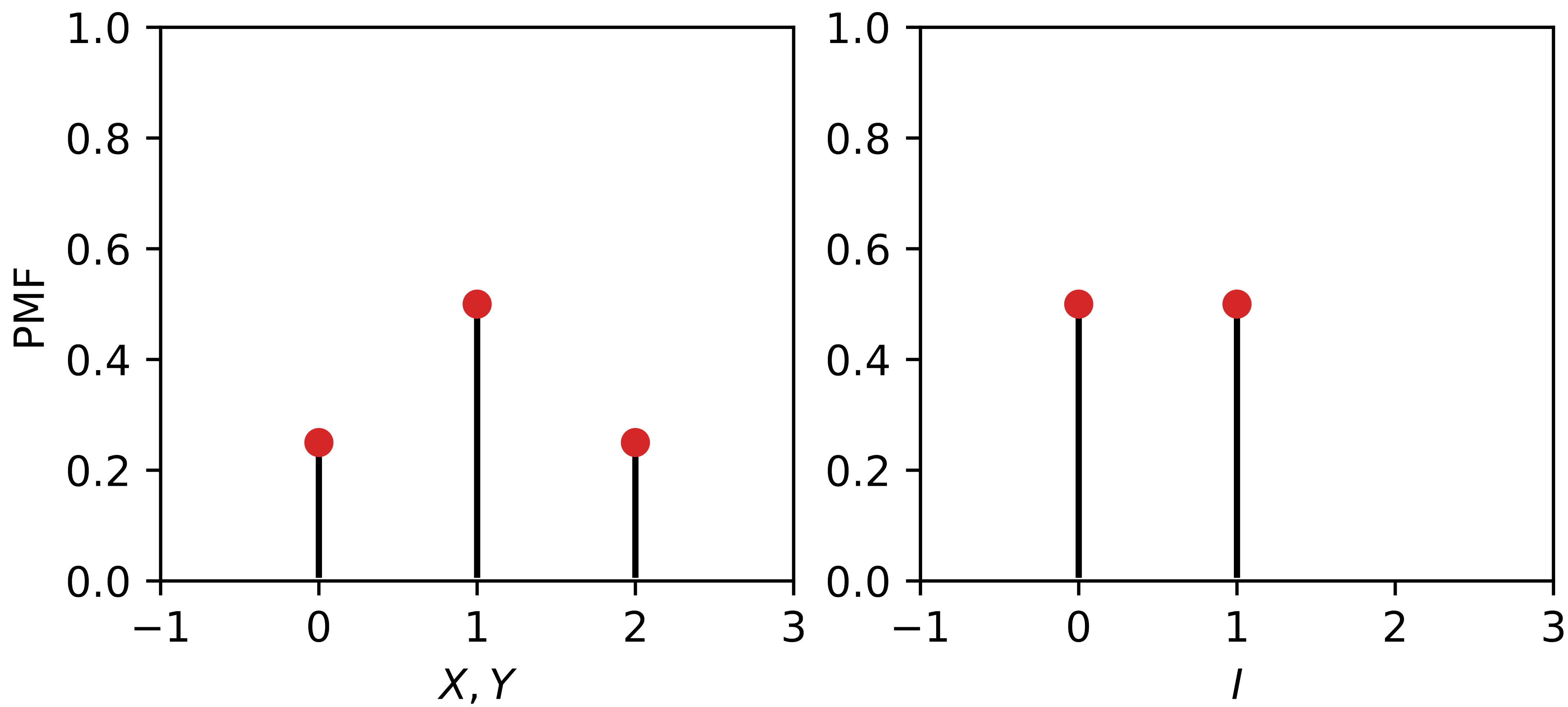
Distributions and PMFs

Example 3.2.4 (Two coin tosses again). $X = \#$ of Heads,

$Y = \#$ of Tails, $I = \begin{cases} 1, \text{ if 1st toss = Heads} \\ 0, \text{ otherwise} \end{cases}$ – indicator variable

$p_X(0) = P(X = 0) = 1/4, \;\; p_X(1) = 1/2, \;\; p_X(2) = 1/4$

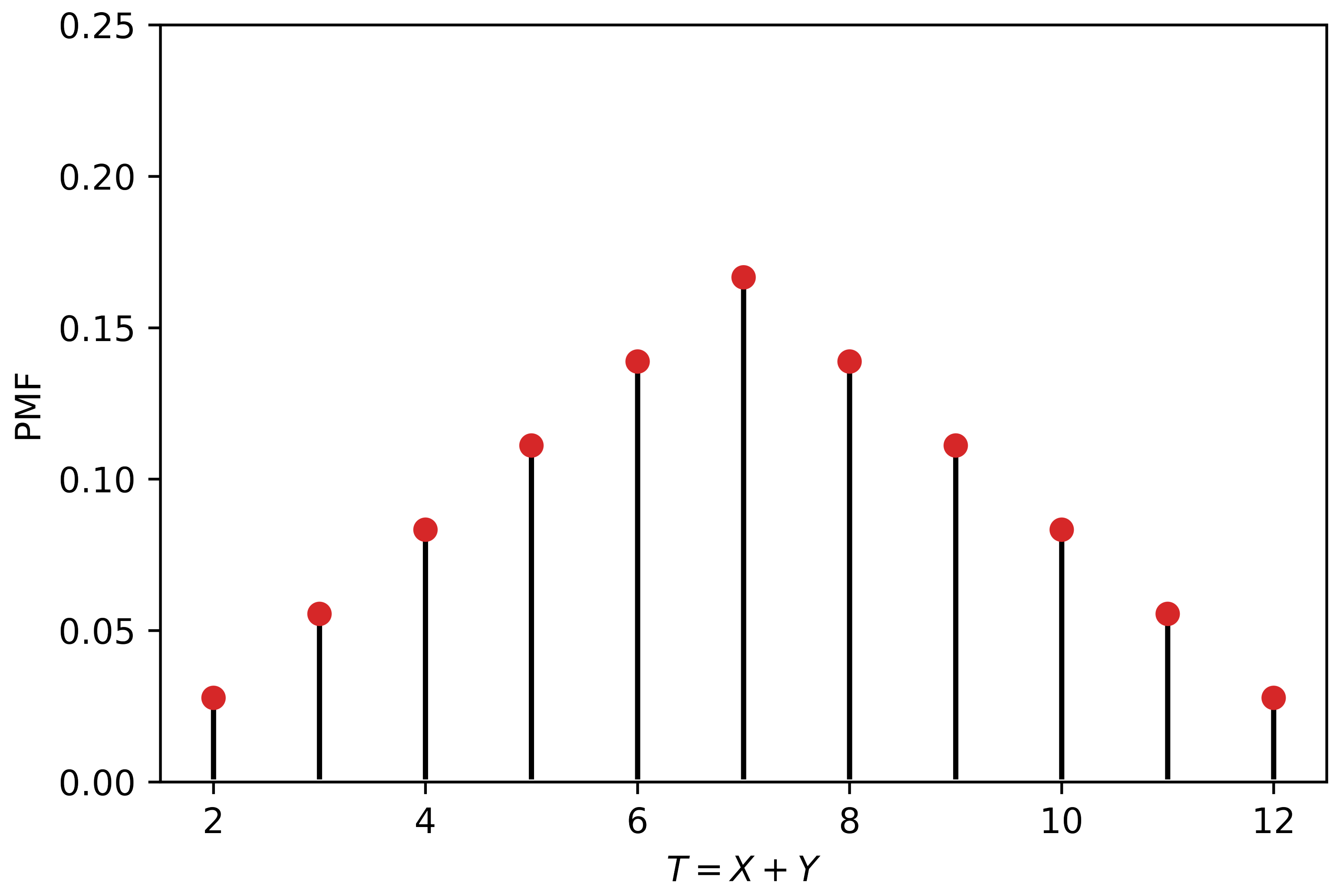
$Y = 2 - X$, so same PMF. $\;\; p_I(0) = 1/2, \;\; p_I(1) = 1/2$



Distributions and PMFs

Example 3.2.5 (Sum of die rolls). Roll two fair 6-sided dice. Let $T = X + Y$, where X, Y are individual rolls. The sample space is $S = \{(1,1), (1,2), \ldots, (6,5), (6,6)\}$.

$$p_T(2) = p_T(12) = \frac{1}{36}, \quad p_T(3) = p_T(11) = \frac{2}{36}, \quad \ldots, p_T(7) = \frac{6}{36}$$



Distributions and PMFs

Theorem 3.2.7 (Valid PMFs). Let X be a discrete r.v. with support $x_1, x_2, \ldots$ The PMF p_X of X must satisfy:

- Nonnegative: $p_X(x) > 0$ if $x = x_j$ for some j, $p_X(x) = 0$ otherwise
- Sums to 1: $\sum_{j=1}^{\infty} p_X(x_j) = 1$

Proof: First is true since probability is nonnegative. Second is true since X must take some value, and the events $\{X = x_j\}$ are disjoint, so

$$\sum_{j=1}^{\infty} P(X = x_j) = P\Big(\bigcup_{j=1}^{\infty} \{X = x_j\}\Big) = P(X = x_1 \text{ or } X = x_2 \ \ldots) = 1$$

Bernoulli and Binomial

Bernoulli and Binomial

Definition 3.3.1 (Bernoulli distribution). An r.v. X is said to have **Bernoulli distribution** with parameter p if $P(X = 1) = p$ and $P(X = 0) = 1 - p$, where $0 < p < 1$. We write $X \sim$ Bern(p) (X is Bernoulli-distributed). It's a **family** of distributions indexed by p.

Definition 3.3.2 (Indicator random variable). **Indicator r.v.** of an event A = r.v. that equals 1 if A occurs and 0 otherwise. We'll denote it I_A or $I(A)$. Note that $I_A \sim$ Bern(p) with $p = P(A)$.

Story 3.3.3 (Bernoulli trial). An experiment that can result in a "success" or "failure" (but not both!) is called a **Bernoulli trial**. A Bernoulli r.v. thus = indicator r.v. of success in Bernoulli trial.

Bernoulli and Binomial

Story 3.3.4 (Binomial distribution). Suppose n **independent** Bernoulli trials are run, each with $P(\text{success}) = p$. Let $X =$ the number of successes. $X \sim$ Bin(n, p) – the **Binomial distribution** with parameters $n = 0,1,2,\ldots$ and $0 < p < 1$.

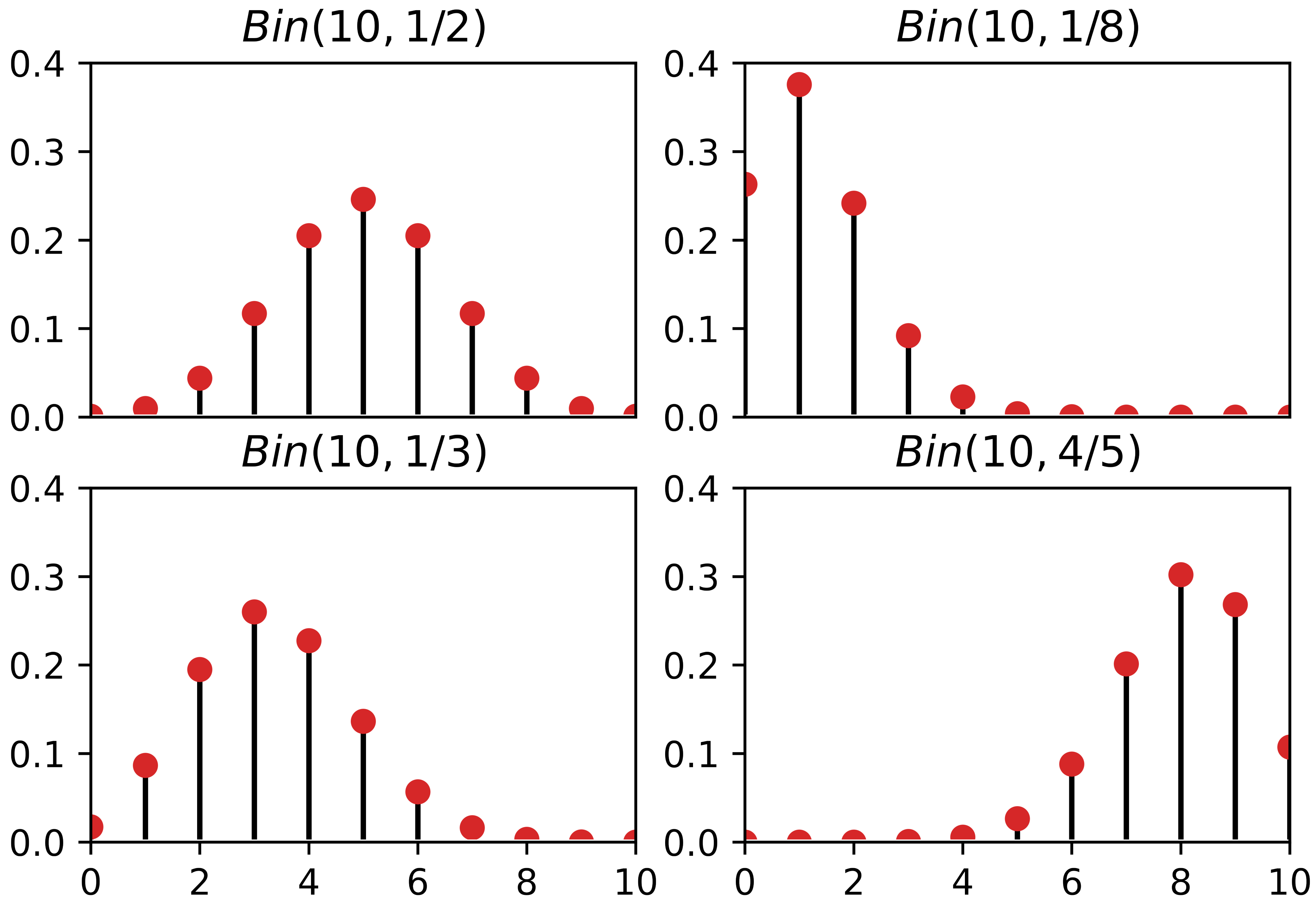
Note that Bern(p) is the same as Bin$(1,p)$.

Theorem 3.3.5 (Binomial PMF). If $X \sim$ Bin(n, p), then the PMF of X

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k} \quad \text{for } k = 0,1,\ldots,n.$$

Bernoulli and Binomial

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$ – for $p \neq 1/2$ it's **skewed**:



Bernoulli and Binomial

Theorem 3.3.7 Let $X \sim \text{Bin}(n, p)$ and $q = 1 - p$ (we often use q = failure probability in Bernoulli trial). Then $n - X \sim \text{Bin}(n, q)$.

Proof: Based on the property that $\binom{n}{n-k} = \binom{n}{k}$

Corollary 3.3.8 Let $X \sim \text{Bin}(n, p)$ with $p = 1/2$ and n – even. Then the distribution of X is symmetric about $n/2$ – that is,

$$P(X = n/2 + j) = P(X = n/2 - j)$$

Hypergeometric

Hypergeometric

Story 3.4.1 Urn with w white & b black balls, drawing n balls **with replacement** yields $\text{Bin}(n, w/(w+b))$ for X – number of white balls in n trials. If we instead sample **without replacement**, then X – # of white balls in n trials – follows a **Hypergeometric distribution:** $X \sim \text{HGeom}(w, b, n)$. In Bernoulli, trials are independent, in Hypergeometric trials are **dependent** (because without replacement)

Theorem 3.4.2 (Hypergeometric PMF). If $X \sim \text{HGeom}(w, b, n)$, then

$$P(X = k) = \frac{\binom{w}{k}\binom{b}{n-k}}{\binom{w+b}{n}} \quad \textit{(think of a proof!)}$$

Hypergeometric

Example 3.4.4 (Aces in a poker hand). In a 5-card hand from a well-shuffled deck, the # of aces $\sim$ HGeom$(4,48,5)$. Then,

$$P(3 \text{ aces}) = \frac{\binom{4}{3}\binom{48}{2}}{\binom{52}{5}} \approx 0.0017$$

Theorem 3.4.5 HGeom(w, b, n) and HGeom$(n, w + b - n, w)$ are identical.

Idea: $X \sim$ HGeom(w, b, n) – X = # of white in a sample of size n.
$Y \sim$ HGeom$(n, w + b - n, w)$ – Y = # of sampled balls among the white balls. (white/black -> sampled/not sampled).
(**Proof** follows from properties of binomial coefficients).

Discrete Uniform

Discrete Uniform

Story 3.5.1 (Discrete Uniform distribution). Let C be a finite, nonempty set of numbers. Choose one of these uniformly at random (i.e., all values are equally likely). Call the chosen number X.
Then X is said to have the **Discrete Uniform distribution** with parameter C, we denote this $X \sim$ DUnif(C).

The PMF is $P(X = x) = \dfrac{1}{|C|}$ for $x \in C$ (and 0 otherwise).

For any $A \subset C$, $P(X \in A) = \dfrac{|A|}{|C|}$