random variables

let $X_i$ = index of
A random variable is a numeric function of the outcome of an experiment, not the outcome itself.

Ex.

- Let $H$ be the number of Heads when 20 coins are tossed
- Let $T$ be the total of 2 dice rolls
- Let $X$ be the number of coin tosses needed to see 1st head

**Note:** even if the underlying experiment has “equally likely outcomes,” the associated random variable may not

<table>
<thead>
<tr>
<th>Outcome</th>
<th>$X = #H$</th>
<th>$P(X)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>TT</td>
<td>0</td>
<td>$P(X=0) = 1/4$</td>
</tr>
<tr>
<td>TH</td>
<td>1</td>
<td>${P(X=1) = 1/2}$</td>
</tr>
<tr>
<td>HT</td>
<td>1</td>
<td>$}$</td>
</tr>
<tr>
<td>HH</td>
<td>2</td>
<td>$P(X=2) = 1/4$</td>
</tr>
</tbody>
</table>
20 balls numbered 1, 2, ..., 20

Draw a subset of size 3 at random.

Let $X$ = the maximum of the numbers on those 3 balls

What is $P(X \geq 17)$

$$P(X = 20) = \frac{\binom{19}{2}}{\binom{20}{3}}$$

$$P(X = 18) = \frac{\binom{17}{2}}{\binom{20}{3}}$$
20 balls numbered 1, 2, ..., 20

Draw a subset of size 3 at random.

Let $X =$ the maximum of the numbers on those 3 balls

What is $P(X \geq 17)$

\[
P(X = 20) = \frac{19}{20} \times \frac{19}{20} = \frac{3}{20} = 0.150
\]

\[
P(X = 19) = \frac{18}{20} \times \frac{18}{20} \times \frac{18}{20} \times \frac{18}{20} = \frac{18 \times 17}{20 \times 19 \times 18} = 0.134
\]

\[
\vdots
\]

\[
\sum_{i=17}^{20} P(X = i) \approx 0.508
\]

\[
P(X \geq 17) = 1 - P(X < 17) = 1 - \frac{16}{20} = 0.508
\]
Flip a (biased) coin (probability $p$ of Heads) repeatedly until 1-st head observed.

What is the sample space?

$\Omega = \{H, TH, TTH, \ldots \}$

How many flips? Let $X$ be that number.

$P(X=1) = p$
$P(X=2) = (1-p) \cdot p$
$P(X=3) = (1-p)^2 \cdot p$
$\ldots$
$P(X=i) = (1-p)^{i-1} \cdot p$. 
Flip a (biased) coin repeatedly until 1st head observed

How many flips? Let $X$ be that number.

$P(X=1) = P(H) = p$

$P(X=2) = P(TH) = (1-p)p$

$P(X=3) = P(TTH) = (1-p)^2p$

...

$P(X=i) = P(T^{i-1}H) = (1-p)^{i-1}p$

\[
\sum_{i \geq 0} x^i = \frac{1}{1-x},
\]

when $|x| < 1$

memorize me!
A discrete random variable is one taking on a countable number of possible values.

Ex:

- $X =$ sum of 3 dice, $3 \leq X \leq 18$, $X \in \mathbb{N}$
- $Y =$ position of 1st head in seq of coin flips, $1 \leq Y$, $Y \in \mathbb{N}$
- $Z =$ largest prime factor of $(1+Y)$, $Z \in \{2, 3, 5, 7, 11, \ldots\}$
A discrete random variable is one taking on a countable number of possible values.

Ex:

\[ X = \text{sum of 3 dice}, \quad 3 \leq X \leq 18, \quad X \in \mathbb{N} \]

**Definition:** If \( X \) is a discrete random variable taking on values from a countable set \( T \subseteq \mathbb{R} \), then

\[
p_x(a) = \begin{cases} 
P(X = a) & \text{for } a \in T \\ 0 & \text{otherwise} \end{cases}
\]

is called the *probability mass function*. Note:

\[
\sum_{a\in T} p_x(a) = 1
\]
Let $X$ be the number of heads in $n$ independent coin tosses, each with probability $p$ of heads.

$$p_X(k) = Pr(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$
Let $X$ be the number of heads observed in $n$ coin flips

$$P(X = k) = \binom{n}{k}p^k(1 - p)^{n-k}, \text{ where } p = P(H)$$

Probability mass function ($p = \frac{1}{2}$):
The cumulative distribution function for a random variable $X$ is the function $F: \mathbb{R} \rightarrow [0,1]$ defined by

$$F(a) = P[X \leq a]$$

Ex: 3 students; homework returned according to random permutation.

$X$ is number of homeworks returned to their correct homework.

<table>
<thead>
<tr>
<th>Prob</th>
<th>outcome</th>
<th>$X$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/6</td>
<td>123</td>
<td>3</td>
</tr>
<tr>
<td>1/6</td>
<td>132</td>
<td>1</td>
</tr>
<tr>
<td>1/6</td>
<td>213</td>
<td>1</td>
</tr>
<tr>
<td>1/6</td>
<td>231</td>
<td>0</td>
</tr>
<tr>
<td>1/6</td>
<td>312</td>
<td>0</td>
</tr>
<tr>
<td>1/6</td>
<td>321</td>
<td>1</td>
</tr>
</tbody>
</table>

What is probability mass function?

Cumulative distribution function?
The **cumulative distribution function** for a random variable $X$ is the function $F: \mathbb{R} \to [0, 1]$ defined by

$$F(a) = P[X \leq a]$$

Ex: if $X$ has **probability mass function** given by:

$$p(1) = \frac{1}{4} \quad p(2) = \frac{1}{2} \quad p(3) = \frac{1}{8} \quad p(4) = \frac{1}{8}$$

$$F(a) = \begin{cases} 0 & a < 1 \\ \frac{1}{4} & 1 \leq a < 2 \\ \frac{3}{4} & 2 \leq a < 3 \\ \frac{7}{8} & 3 \leq a < 4 \\ 1 & 4 \leq a \end{cases}$$

NB: for discrete random variables, be careful about “$\leq$” vs “$<$”
expectation
For a discrete r.v. $X$ with p.m.f. $p(\cdot)$, the **expectation of $X$**, aka **expected value** or **mean**, is

$$E[X] = \sum_x x p(x)$$

average of random values, weighted by their respective probabilities

For the equally-likely outcomes case, this is just the average of the possible random values of $X$

For unequally-likely outcomes, it is again the average of the possible random values of $X$, **weighted by their respective probabilities**

**Ex 1:** Let $X =$ value seen rolling a fair die  $p(1), p(2), \ldots, p(6) = 1/6$

$$E(X) = \sum_{i=1}^{6} i \cdot \frac{1}{6} = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + \ldots + 6 \cdot \frac{1}{6} = 3.5$$
For a discrete r.v. \( X \) with p.m.f. \( p(\cdot) \), the *expectation of \( X \)*, aka *expected value* or *mean*, is

\[
E[X] = \sum_x xp(x)
\]

For the equally-likely outcomes case, this is just the average of the possible random values of \( X \).

For *unequally-likely* outcomes, it is again the average of the possible random values of \( X \), *weighted* by their respective probabilities.

**Ex 1:** Let \( X = \) value seen rolling a fair die \( p(1), p(2), \ldots, p(6) = 1/6 \)

\[
E[X] = \sum_{i=1}^{6} ip(i) = \frac{1}{6} (1 + 2 + \cdots + 6) = \frac{21}{6} = 3.5
\]

**Ex 2:** Coin flip; \( X = +1 \) if H (win $1), -1 if T (lose $1)

\[
E(X) = \frac{1}{2} \cdot \frac{p_r(X=1)}{2} + \frac{(-1) \cdot p_r(X=-1)}{2} = 0
\]
For a discrete r.v. $X$ with p.m.f. $p(\cdot)$, the \textit{expectation of $X$}, aka \textit{expected value} or \textit{mean}, is

\[
E[X] = \sum_x x p(x)
\]

For the equally-likely outcomes case, this is just the average of the possible random values of $X$

For \textit{unequally}-likely outcomes, it is again the average of the possible random values of $X$, \textit{weighted by their respective probabilities}

\begin{itemize}
    \item Ex 1: Let $X =$ value seen rolling a fair die $p(1), p(2), \ldots, p(6) = 1/6$
    \[
    E[X] = \sum_{i=1}^{6} i p(i) = \frac{1}{6} (1 + 2 + \cdots + 6) = \frac{21}{6} = 3.5
    \]
    
    \item Ex 2: Coin flip; $X = +1$ if H (win $1$), -1 if T (lose $1$)
    \[
    E[X] = (+1)\cdot p(+1) + (-1)\cdot p(-1) = 1\cdot(1/2) +(-1)\cdot(1/2) = 0
    \]
\end{itemize}
For a discrete r.v. $X$ with p.m.f. $p(\cdot)$, the *expectation of $X$*, aka *expected value* or *mean*, is

$$E[X] = \sum_x x p(x)$$

Another view: A 2-person gambling game. If $X$ is how much you win playing the game once, how much would you expect to win, on average, per game, when repeatedly playing?
For a discrete r.v. \(X\) with p.m.f. \(p(\cdot)\), the expectation of \(X\), aka expected value or mean, is

\[
E[X] = \sum x p(x)
\]

Another view: A 2-person gambling game. If \(X\) is how much you win playing the game once, how much would you expect to win, on average, per game, when repeatedly playing?

Ex 1: Let \(X = \) value seen rolling a fair die \(p(1), p(2), \ldots, p(6) = 1/6\)
If you win \(X\) dollars for that roll, how much do you expect to win?

\[
E[X] = \sum_{i=1}^{6} i p(i) = \frac{1}{6} (1 + 2 + \cdots + 6) = \frac{21}{6} = 3.5
\]

Ex 2: Coin flip; \(X = +1\) if H (win $1), -1 if T (lose $1)

\[
E[X] = (+1)\cdot p(+1) + (-1)\cdot p(-1) = 1\cdot(1/2) +(-1)\cdot(1/2) = 0
\]

“a fair game”: in repeated play you expect to win as much as you lose. Long term net gain/loss = 0.
Let $X$ be the number of flips up to & including 1$^{\text{st}}$ head observed in repeated flips of a biased coin (with probability $p$ of coming up heads). If I pay you $1 per flip, how much money would you expect to make?

$E(X) = \sum_{i=1}^{\infty} i \cdot Pr(X=i)$

$= \sum_{i=1}^{\infty} i \cdot (1-p)^{i-1} \cdot p = p \sum_{i=1}^{\infty} i \cdot (1-p)^{i-1}$

$= p \left( \frac{1}{(1-(1-p))^2} \right) = \frac{1}{p}$

$\sum_{i \geq 0} x^i = \frac{1}{1-x}$, when $|x| < 1$

memorize me!
Let $X$ be the number of flips up to & including 1st head observed in repeated flips of a biased coin. If I pay you $1 per flip, how much money would you expect to make?

$$P(H) = p; \quad P(T) = 1 - p = q$$
$$p(i) = pq^{i-1} \leftarrow \text{PMF}$$
$$E[X] = \sum_{i \geq 1} ip(i) = \sum_{i \geq 1} ipq^{i-1} = p \sum_{i \geq 1} iq^{i-1} \quad (*)$$

A calculus trick:

$$\sum_{i \geq 1} iy^{i-1} = \sum_{i \geq 1} \frac{d}{dy} y^i = \sum_{i \geq 0} \frac{d}{dy} y^i = \frac{d}{dy} \sum_{i \geq 0} y^i = \frac{d}{dy} \frac{1}{1 - y} = \frac{1}{(1 - y)^2}$$

So $(*)$ becomes:

$$E[X] = p \sum_{i \geq 1} iq^{i-1} = \frac{p}{(1 - q)^2} = \frac{p}{p^2} = \frac{1}{p}$$

E.g.:

$p=1/2$; on average head every 2nd flip
$p=1/10$; on average, head every 10th flip.
Let $X$ be the number of heads observed in $n$ repeated flips of a biased coin. If I pay you $1 per head, how much money would you expect to make?

E.g.:

- $p=1/2$ \quad 50
- $p=1/10$ \quad 10
Let $X$ be the number of heads observed in $n$ repeated flips of a biased coin. If I pay you $1 per head, how much money would you expect to make?

E.g.:

$p = 1/2$; on average, $n/2$ heads

$p = 1/10$; on average, $n/10$ heads

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

$$= \sum_{i=1}^{n} i \binom{n}{i} p^i (1 - p)^{n-i}$$

$$= \sum_{i=1}^{n} n \binom{n-1}{i-1} p^i (1 - p)^{n-i}$$

$$= np \sum_{i=1}^{n} \binom{n-1}{i-1} p^{i-1} (1 - p)^{n-i}$$

$$= np \sum_{j=0}^{n-1} \binom{n-1}{j} p^j (1 - p)^{n-1-j}$$

$$= np(p + (1 - p))^{n-1} = np$$
For a discrete r.v. $X$ with p.m.f. $p(\cdot)$, the expectation of $X$, aka expected value or mean, is

$$E[X] = \sum_x x p(x) = \sum_x \Pr(X = x)$$

Another view:

$$E[X] = \sum_{s \in S} X(s) \cdot p(s)$$

$$E(X) = \sum_{\omega \in \Omega} X(\omega) \Pr(\omega)$$