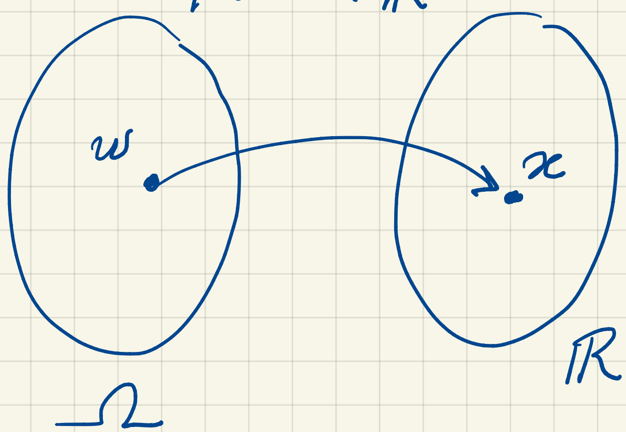
Lecture 5. Random Variables (Discrete case)

Definition: A random variable is a mapping from the sample space Ω to the real line.

$$\phi : \Omega \to \mathbb{R}$$

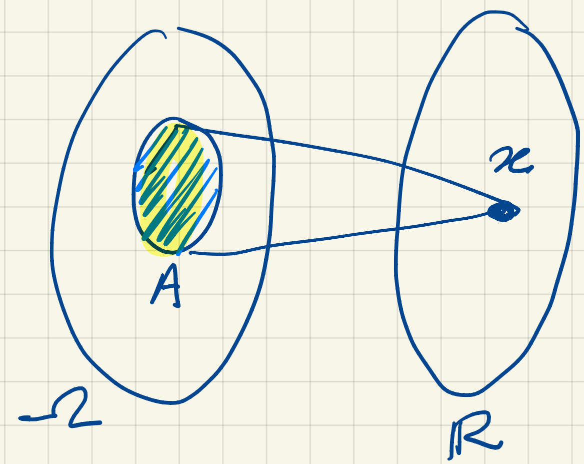


X takes on value x

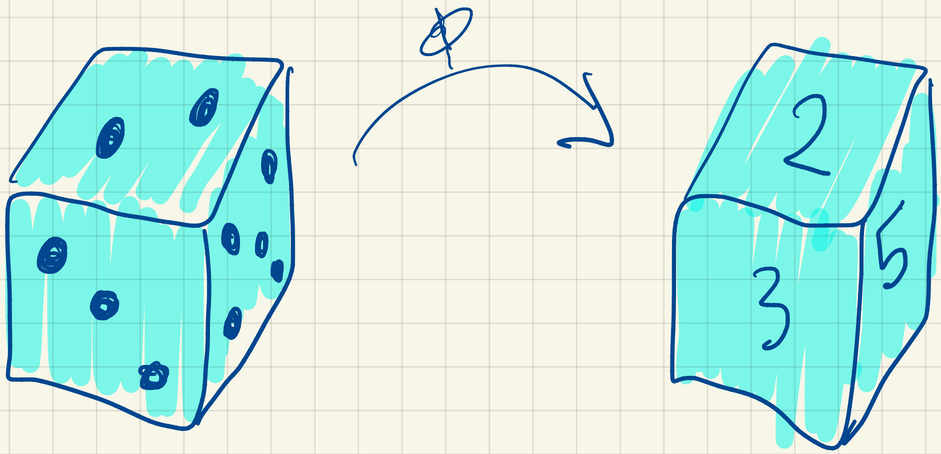
$$P(X=x) = \sum_{\substack{w \in \Omega \\ \phi(w)=x}} P(w)$$

Mapping is not necessarily one-to-one

We can work with outcomes analytically, define averages, deviations, etc...

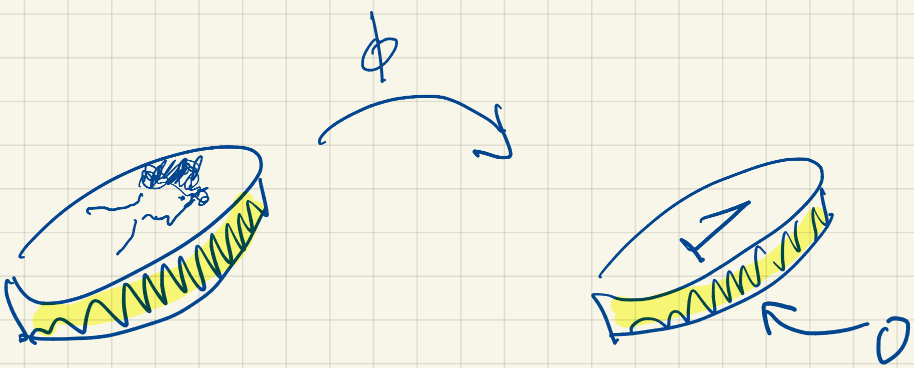


$$P(x) = P(A)$$



Fair die

$$P(X=x)=\frac{1}{6} \quad x\in\{1,2,\dots,6\}$$

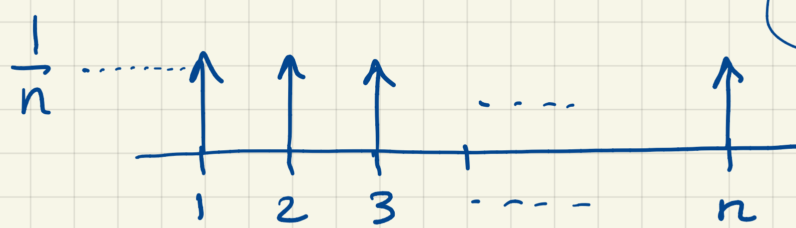


$$P(X=1)=1-P(X=0)=p$$

Probability Mass function: (Discrete case)

It's a function that assign a prob. for each value of X

Uniform PMF:



Uniform requires finite sample space

Binomial PMF:

Consider tossing a coin n times. (tosses are independent)

e.g. $n=3$ $S=\{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$

In general $|S| = 2^n$.

probability of an outcome with k heads and n-k tails is $p^k(1-p)^{n-k}$

p^3

$p^2(1-p)$

$p^2(1-p)$

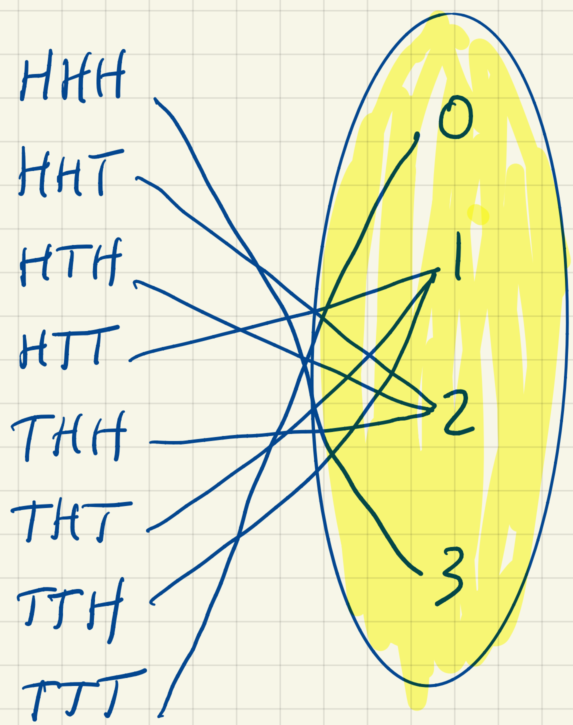
$p(1-p)^2$

$p^2(1-p)$

$p(1-p)^2$

$p(1-p)^2$

$(1-p)^3$



Random variable X = # heads in the outcome

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

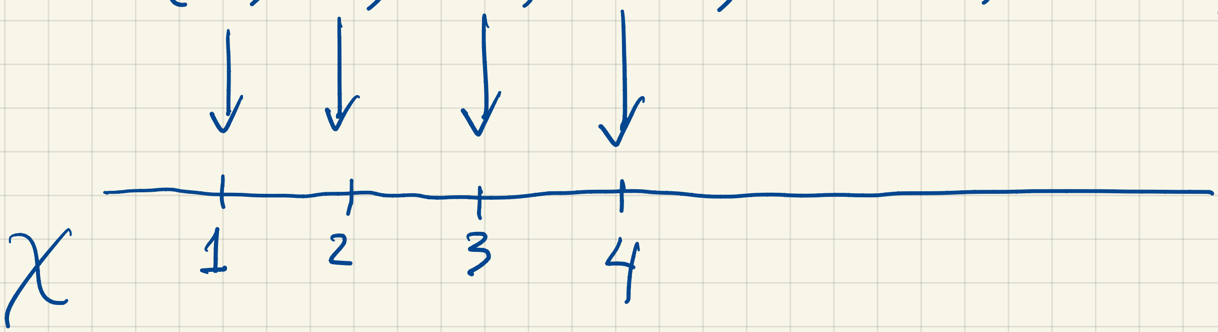
$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

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

Geometric PMF

Toss the coin until you get H.

$$S = \{H, TH, TTH, TTTH, TTTTH, \ldots\}$$



Infinite Sample space

$P(X=k) = ?$

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

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

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

$$\vdots$$

$$P(X=k) = p(1-p)^{k-1} \quad k \geq 1$$

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

$$p\left[1 + (1-p) + (1-p)^2 + (1-p)^3 + \ldots\right]$$

$$p \frac{1}{1-(1-p)} = p \frac{1}{p} = 1.$$

- Uniform $\quad P(X=k) = \frac{1}{n}$ where n is the # values (n is param)
- Binomial $\quad P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$ (p & n are parameters)
- Geometric $\quad P(X=k) = p(1-p)^{k-1}$ (p is parameter)

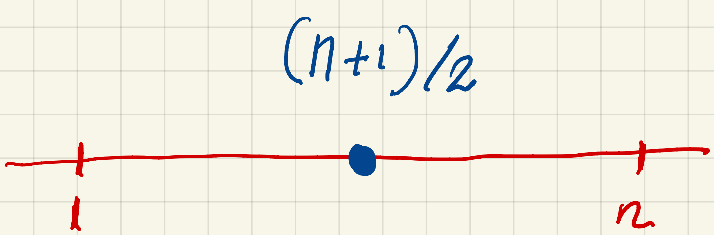
The Expected value or <u>mean</u> of a random Variable X.

$$E[X] = \sum_x x\, P(X=x)$$

Example: Consider uniform R.V. in $\{1, 2, 3, \ldots, n\}$

$$E[X] = \sum_x x\, P(X=x) = 1 \cdot \frac{1}{n} + 2 \cdot \frac{1}{n} + 3 \cdot \frac{1}{n} + \cdots n \cdot \frac{1}{n}$$

$$= \frac{1}{n} \underbrace{\left[1 + 2 + 3 + \cdots + n\right]}_{\frac{n(n+1)}{2}} = \frac{n+1}{2}$$



One can compute the expected value of any function $f(x)$

$$E[f(x)] = \sum_x f(x) P(X=x)$$

$E[X]$ is a special case when $f(x) = x$

Expectation has the following properties

- The expected value of a constant is the constant itself.
- $E[aX] = aE[X]$ where a is a constant.
- Linearity (very useful): $E[X+Y] = E[X] + E[Y]$ regardless of whether X and Y are independent
- $E[XY] = E[X]E[Y]$ if X and Y are independent

X and Y independent $\Longrightarrow E[XY] = E[X]E[Y]$

- Conditional Expectation: (Nested expectation)

$$E[X] = E\left[E[X \mid Y=y]\right]$$

- The inner expectation is expectation of X using $P(X \mid Y=y)$ (and not $P(x)$)

  This gives an expression in y.

- The outer expectation is over Y using $P(Y)$

$$\sum_{y} \underbrace{E[X \mid Y=y]}_{\Downarrow} P(Y=y)$$

$$f(y) = \sum_{x} x \, P(X=x \mid Y=y)$$

Proof:

$$\sum_y \left[\sum_x x\, P(X=x \mid Y=y) \right] P(Y=y)$$

$$= \sum_x \sum_y x\, P(X=x \mid Y=y)\, P(Y=y)$$

$$= \sum_x x \sum_y P(X=x \mid Y=y)\, P(Y=y)$$

$$= \sum_x x\, P(X=x)$$

It's also easy to prove linearity and the product in case of independence.

How Linear property can be useful.

Say X is binomial $P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$

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

Suppose $X_i = \begin{cases} 1 & p \\ 0 & 1-p \end{cases}$ Bernoulli Trial / Bernoulli Random Variable

$E[X_i] = p$ $\left(1 \cdot p + 0 \cdot (1-p)\right)$

If $X_1, X_2, \dots, X_n$ independent Bernoulli RVs

Then $X = X_1 + X_2 + \dots + X_n$

$$E[X] = E[X_1 + X_2 + \cdots X_n] = E[X_1] + E[X_2] + \cdots E[X_n] = np$$

So another way to look at binomial random variable is that it's the sum of independent Bernoullis

$$X = X_1 + X_2 + \cdots + X_n$$

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