random variables

T T T T H T H H

let X = 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*

Outcome	*X = #H*	*P(X)*
TT	0	P(X=0) = 1/4
TH	1	} P(X=1) = 1/2
HT	1	
HH	2	P(X=2) = 1/4

---

20 balls numbered 1, 2, ..., 20

Draw 3 without replacement

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

What is P(X ≥ 17)

$$P(X=20) = \binom{19}{2}/\binom{20}{3} = \frac{3}{20} = 0.150$$

$$P(X=19) = \binom{18}{2}/\binom{20}{3} = \frac{18\cdot 17/2!}{20\cdot 19\cdot 18/3!} \approx 0.134$$

$$\vdots$$

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

$$P(X \geq 17) = 1 - P(X<17) = 1 - \binom{16}{3}\Big/\binom{20}{3} \approx 0.508$$

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!

probability mass functions

A *discrete* random variable is one taking on a *countable* number of possible values.

Ex:

X = sum of 3 dice, $3 \le X \le 18$, $X \in N$

Y = position of 1st head in seq of coin flips, $1 \le Y$, $Y \in N$

Z = largest prime factor of (1+Y), $Z \in \{2, 3, 5, 7, 11, \ldots\}$

probability mass functions

A *discrete* random variable is one taking on a *countable* number of possible values.

Ex:

X = sum of 3 dice, $3 \le X \le 18$, $X \in N$

Definition: If X is a discrete random variable taking on values from a countable set $T \subseteq \mathcal{R}$, then

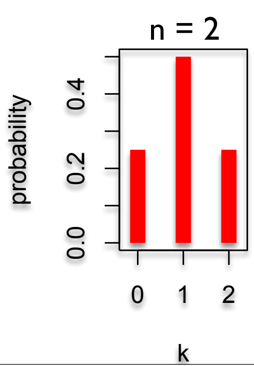
$$p_X(a) = \begin{cases} P(X = a) & \text{for } a \in T \\ 0 & \text{otherwise} \end{cases}$$

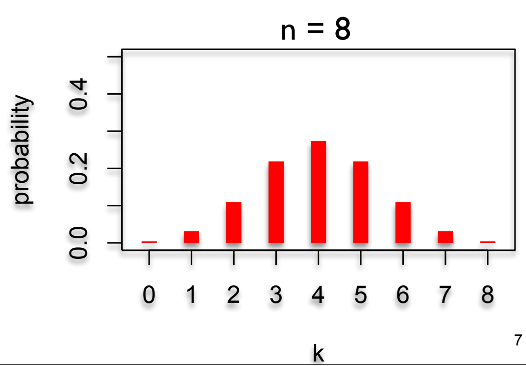
is called the *probability mass function*. Note: $\sum_{a \in T} p_X(a) = 1$

Let X be the number of heads 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 = ½):





cumulative distribution function

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

$$\boxed{F(a) = P[X \le a]}$$

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

X is number of homeworks returned to their correct owner.

Prob	outcome	X
1/6	123	3
1/6	132	1
1/6	213	1
1/6	231	0
1/6	312	0
1/6	321	1

What is probability mass function?

Cumulative distribution function?

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

$$\boxed{F(a) = P[X \le a]}$$

$$F(a) = \sum_{j \le a} p_X(a) = \sum_{j \le a} Pr(X = a)$$

summing over $j \in Range(X)$

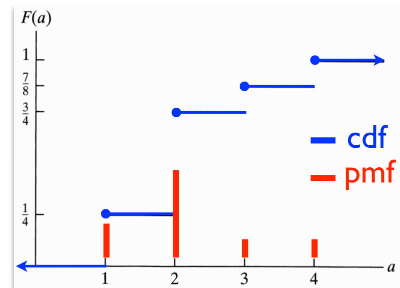
The *cumulative distribution function* for a random variable X is the function F: $\mathcal{R} \to [0,1]$ defined by

$$\boxed{F(a) = P[X \le a]}$$

Ex: if X has probability mass function given by:

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

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



NB: for discrete random variables, be careful about "≤" vs "<"

expectation

For a discrete r.v. X with p.m.f. p(•), the *expectation of X*, aka *expected value* or *mean*, is

$$\boxed{E[X] = \Sigma_x \, xp(x)}$$

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 *un*equally-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 $p(1), p(2), ..., p(6) = 1/6$

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

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

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

For a discrete r.v. X with p.m.f. $p(\bullet)$, the *expectation of X*, aka *expected value* or *mean*, is

$E[X] = \Sigma_x \, xp(x)$ average of random values, weighted by their respective probabilities

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. $p(\bullet)$, the *expectation of X*, aka *expected value* or *mean*, is

$E[X] = \Sigma_x \, xp(x)$ average of random values, weighted by their respective probabilities

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

If you win X dollars for that roll, how much do you expect to win?

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

For a discrete r.v. X with p.m.f. $p(\bullet)$, the *expectation of X*, aka *expected value* or *mean*, is

$E[X] = \Sigma_x \, xp(x)$ average of random values, weighted by their respective probabilities

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

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

"a fair game": in repeated play you expect to win as much as you lose. Long term net gain/loss = 0.

For a discrete r.v. X with p.m.f. $p(\bullet)$, the *expectation of X*, aka *expected value* or *mean*, is

$E[X] = \Sigma_x \, xp(x)$ Another view: $E[X] = \sum_{s \in S} X(s) \cdot p(s)$

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

X is number of homeworks returned to their correct owner.

Prob	outcome	X
1/6	123	3
1/6	132	1
1/6	213	1
1/6	231	0
1/6	312	0
1/6	321	1

E (X) = ?

$$E(X) = 0 \cdot \frac{2}{6} + 1 \cdot \frac{3}{6} + 3 \cdot \frac{1}{6}$$

$$= X(123)\cdot\frac{1}{6} + X(132)\cdot\frac{1}{6} + X(213)\cdot\frac{1}{6} + X(231)\cdot\frac{1}{6} + X(312)\cdot\frac{1}{6} + X(321)\cdot\frac{1}{6}$$

For a discrete r.v. X with p.m.f. $p(\bullet)$, the *expectation of X*, aka *expected value* or *mean*, is

$E[X] = \Sigma_x \, xp(x)$ Another view: $E[X] = \sum_{s \in S} X(s) \cdot p(s)$

20 balls numbered 1, 2, ..., 20

Draw 3 without replacement

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

E(X) = ?

$$E(X) = \sum_{k=3}^{20} k \frac{\binom{k-1}{2}}{\binom{20}{3}}$$