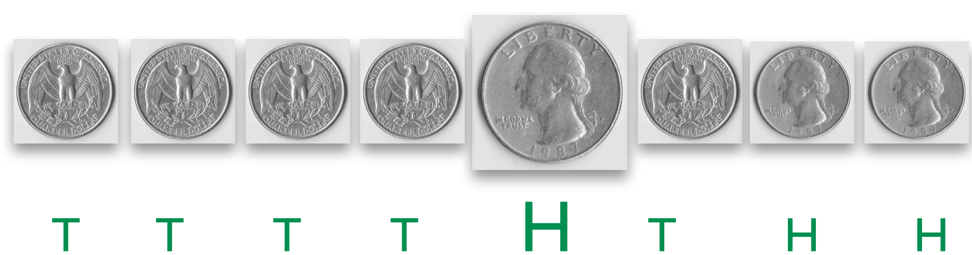
random variables



let X = index of

A *random variable* X assigns a real number to each outcome in a probability space.

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

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

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

$$\vdots$$

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

Alternatively:

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

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) = $(1-p)^2p$

...

Check that it is a valid probability distribution:

$$P\left(\bigcup_{i\geq 1}\{X=i\}\right) = \sum_{i\geq 1}(1-p)^{i-1}p = p\sum_{i\geq 0}(1-p)^i = p\frac{1}{1-(1-p)} = 1$$

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

Ex:

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

Y = index of 1st head in seq of coin flips, $1 \le Y, Y \in \mathbb{N}$

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

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

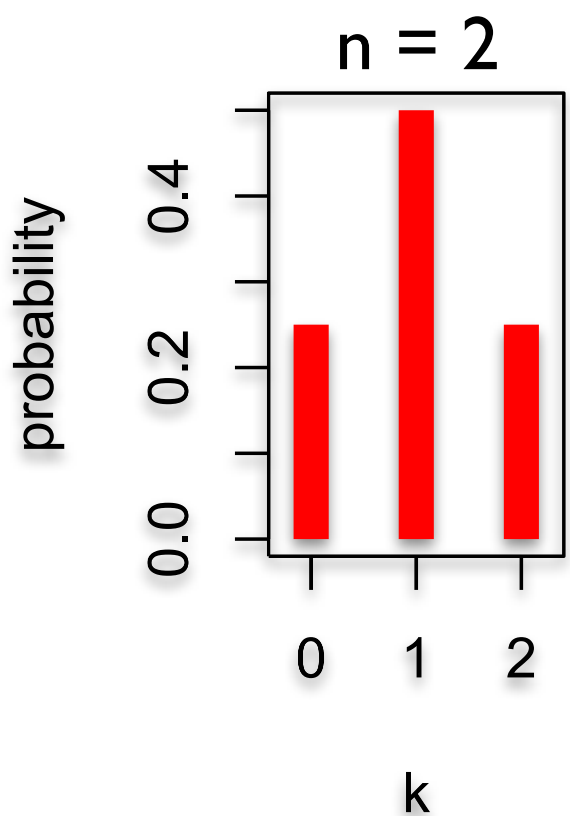
$$p(a) = \begin{cases} P(X = a) & \text{for } a \in T \\ 0 & \text{otherwise} \end{cases}$$

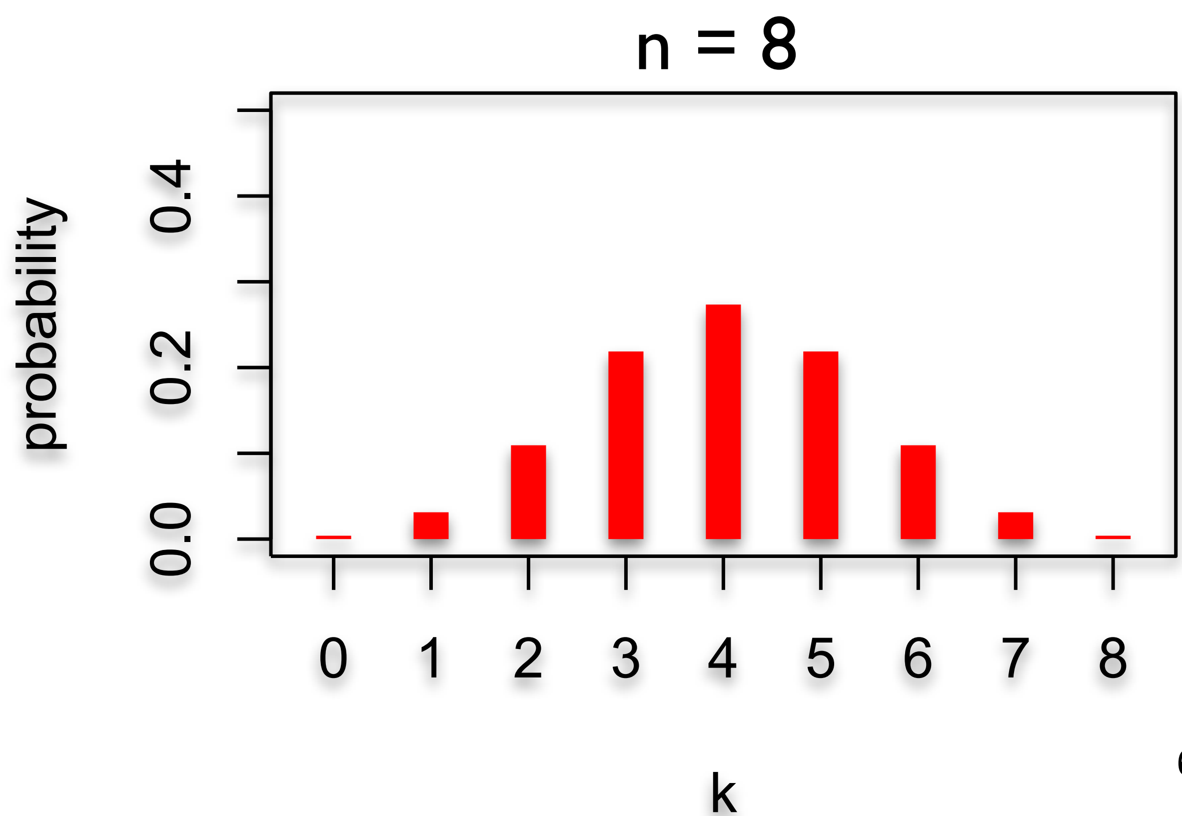
is called the *probability mass function*. Note: $\sum_{a \in T} p(a) = 1$

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:





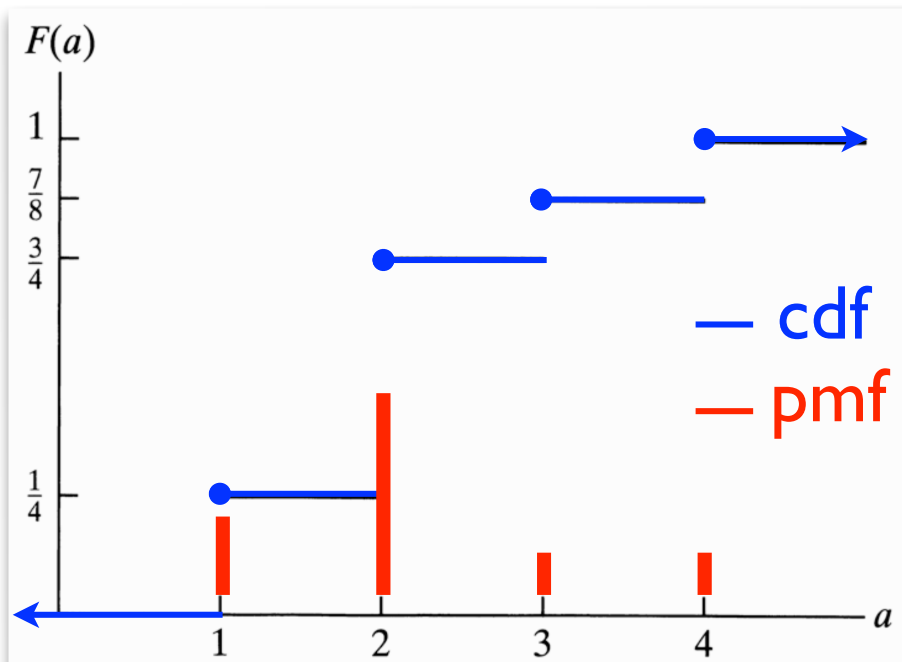
The *cumulative distribution function* for a random variable X is the function $F: \mathbb{R} \to [0,1]$ defined by

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

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 "≤" vs "<"

For a discrete r.v. X with p.m.f. p(•), 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

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

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

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

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

For a discrete r.v. X with p.m.f. p(•), 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 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), ..., p(6) = 1/6

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

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

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

E[X] = (+1)•p(+1) + (-1)•p(-1) = 1•(1/2) +(-1)•(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 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?

$$\begin{aligned} P(H) &= p; \quad P(T) = 1 - p = q \\ p(i) &= pq^{i-1} \\ E(x) &= \textstyle\sum_{i \geq 1} ip(i) = \sum_{i \geq 1} ipq^{i-1} = p \sum_{i \geq 1} iq^{i-1} \quad (*) \end{aligned}$$

A calculus trick:

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

$dy^0/dy = 0$

So $(*)$ becomes:

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

How much would you pay to play?

E.g.:

p=1/2; on average head every 2nd flip

p=1/10; on average, head every 10th flip.

Calculating E[g(X)]:

Y=g(X) is a new r.v. Calc P[Y=j], then apply defn:

X = sum of 2 dice rolls

i	p(i) = P[X=i]	i•p(i)
2	1/36	2/36
3	2/36	6/36
4	3/36	12/36
5	4/36	20/36
6	5/36	30/36
7	6/36	42/36
8	5/36	40/36
9	4/36	36/36
10	3/36	30/36
11	2/36	22/36
12	1/36	12/36

$E[X] = \Sigma_i \, ip(i) =$ 252/36 = 7

Y = g(X) = X mod 5

j	q(j) = P[Y = j]	j•q(j)
0	4/36+3/36 =7/36	0/36
1	5/36+2/36 =7/36	7/36
2	1/36+6/36+1/36 =8/36	16/36
3	2/36+5/36 =7/36	21/36
4	3/36+4/36 =7/36	28/36

$E[Y] = \Sigma_j \, jq(j) =$ 72/36 = 2

expectation of a *function* of a random variable

Calculating E[g(X)]: Another way – add in a different order, using P[X=...] instead of calculating P[Y=...]

X = sum of 2 dice rolls

i	p(i) = P[X=i]	g(i)•p(i)
2	1/36	2/36
3	2/36	6/36
4	3/36	12/36
5	4/36	0/36
6	5/36	5/36
7	6/36	12/36
8	5/36	15/36
9	4/36	16/36
10	3/36	0/36
11	2/36	2/36
12	1/36	2/36

$E[g(X)] = \Sigma_i\, g(i)p(i) =$ 72/36 $= 2$

Y = g(X) = X mod 5

j	q(j) = P[Y = j]	j•q(j)
0	4/36+3/36 =7/36	0/36
1	5/36+2/36 =7/36	7/36
2	1/36+6/36+1/36 =8/36	16/36
3	2/36+5/36 =7/36	21/36
4	3/36+4/36 =7/36	28/36

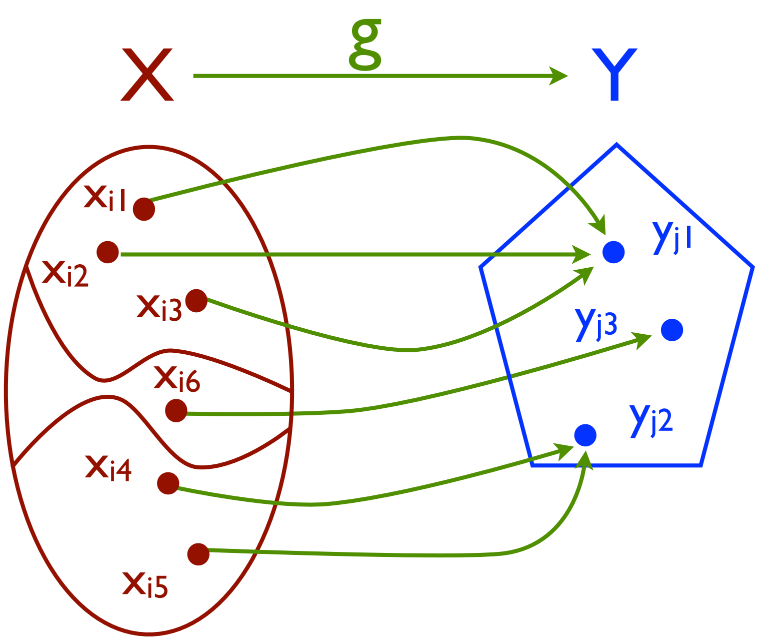
$E[Y] = \Sigma_j\, jq(j) =$ 72/36 $= 2$

expectation of a *function* of a random variable

Above example is not a fluke.

Theorem: if $Y = g(X)$, then $E[Y] = \Sigma_i\, g(x_i)p(x_i)$, where x_i, $i = 1, 2, \ldots$ are all possible values of X.

Proof: Let y_j, $j = 1, 2, \ldots$ be all possible values of Y.



Note that $S_j = \{ x_i \mid g(x_i)=y_j \}$ is a partition of the domain of g.

$$
\begin{aligned}
\sum_i g(x_i)p(x_i) &= \sum_j \sum_{i:g(x_i)=y_j} g(x_i)p(x_i) \\
&= \sum_j \sum_{i:g(x_i)=y_j} y_j p(x_i) \\
&= \sum_j y_j \sum_{i:g(x_i)=y_j} p(x_i) \\
&= \sum_j y_j P\{g(X) = y_j\} \\
&= E[g(X)]
\end{aligned}
$$

A & B each bet \$1, then flip 2 coins:

HH	A wins \$2
HT	Each takes back \$1
TH	Each takes back \$1
TT	B wins \$2

Let X be A's net gain: +1, 0, -1, resp.:

$P(X = +1) = 1/4$
$P(X = 0) = 1/2$
$P(X = -1) = 1/4$

What is E[X]?

$$E[X] = 1 \cdot 1/4 + 0 \cdot 1/2 + (-1) \cdot 1/4 = 0$$

What is $E[X^2]$?

$$E[X^2] = 1^2 \cdot 1/4 + 0^2 \cdot 1/2 + (-1)^2 \cdot 1/4 = 1/2$$

Note:
$E[X^2] \neq E[X]^2$

Linearity of expectation, I

For any constants a, b: $E[aX + b] = aE[X] + b$

Proof:

$$\begin{aligned} E[aX + b] &= \sum_x (ax + b) \cdot p(x) \\ &= a \sum_x xp(x) + b \sum_x p(x) \\ &= aE[X] + b \end{aligned}$$

Example:

Q: In the 2-person coin game above, what is E[2X+1]?

A: E[2X+1] = 2E[X]+1 = 2•0 + 1 = 1

Linearity, II

Let X and Y be two random variables derived from outcomes of a single experiment. Then

$$E[X+Y] = E[X] + E[Y]$$ True even if X,Y *dependent*

Proof: Assume the sample space S is countable. (The result is true without this assumption, but I won't prove it.) Let X(s), Y(s) be the values of these r.v.'s for outcome s∈S.

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

Proof: similar to that for "expectation of a function of an r.v.," i.e., the events "X=x" partition S, so sum above can be rearranged to match the definition of $E[X] = \sum_x x \cdot P(X = x)$

Then:

$$E[X+Y] = \Sigma_{s \in S}(X[s] + Y[s])\, p(s)$$
$$= \Sigma_{s \in S} X[s]\, p(s) + \Sigma_{s \in S} Y[s]\, p(s) = E[X] + E[Y]$$

Example

X = # of heads in *one* coin flip, where P(X=1) = p.

What is $E(X)$?

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

Let X_i, $1 \le i \le n$, be # of H in flip of coin with $P(X_i=1) = p_i$

What is the expected number of heads when all are flipped?

$E[\Sigma_i X_i] = \Sigma_i E[X_i] = \Sigma_i p_i$

Special case: $p_1 = p_2 = ... = p$:

$E[\text{\# of heads in } n \text{ flips}] = pn$

Note:

Linearity is special!

It is *not* true in general that

$$E[X \cdot Y] = E[X] \cdot E[Y]$$

$$E[X^2] = E[X]^2 \quad \leftarrow \text{counterexample above}$$

$$E[X/Y] = E[X] / E[Y]$$

$$E[\text{asinh}(X)] = \text{asinh}(E[X])$$

$$\vdots$$

Application: The Probabilistic Method

Bunch of prisoners in a jail.

Two lunch slots: A and B.

R pairs of prisoners are risky.

Is there a way to assign the prisoners to lunch slots so that at least 1/2 the risky pairs are broken up (assigned to different lunch slots)?

X: number of risky pairs that are broken up

E(X) = |R|/2.

==> there is an assignment of prisoners to lunch slots such that at least half of the risky pairs are broken up.

Cool!   We showed it exists without finding it, using a probabilistic argument.

Alice & Bob are gambling (again). X = Alice's gain per flip:

$$X = \begin{cases} +1 & \text{if Heads} \\ -1 & \text{if Tails} \end{cases}$$

E[X] = 0

… Time passes …

Alice (yawning) says "let's raise the stakes"

$$Y = \begin{cases} +1000 & \text{if Heads} \\ -1000 & \text{if Tails} \end{cases}$$

E[Y] = 0, as before.

Are you (Bob) equally happy to play the new game?

E[X] measures the "average" or "central tendency" of X.

What about its *variability*?

Definition

The *variance* of a random variable X with mean $E[X] = \mu$ is $Var[X] = E[(X-\mu)^2]$, often denoted σ^2.

Alice & Bob are gambling (again). X = Alice's gain per flip:

$$X = \begin{cases} +1 & \text{if Heads} \\ -1 & \text{if Tails} \end{cases}$$

E[X] = 0

Var[X] = 1

… Time passes …

Alice (yawning) says "let's raise the stakes"

$$Y = \begin{cases} +1000 & \text{if Heads} \\ -1000 & \text{if Tails} \end{cases}$$

E[Y] = 0, as before.

Var[Y] = 1,000,000

Are you (Bob) equally happy to play the new game?

E[X] measures the "average" or "central tendency" of X.

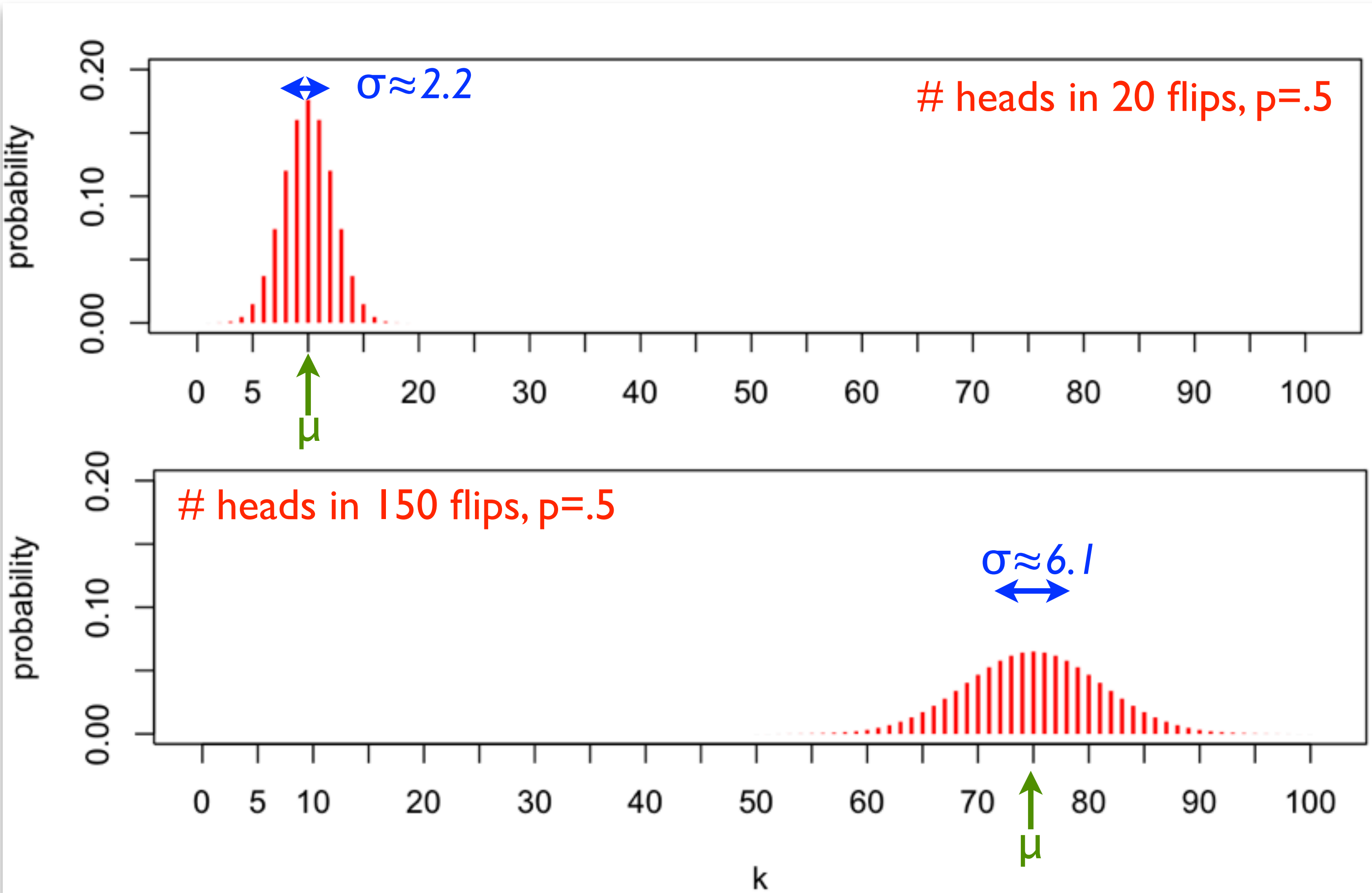
What about its *variability*?

Definition

The *variance* of a random variable X with mean $E[X] = \mu$ is $Var[X] = E[(X-\mu)^2]$, often denoted σ^2.

The *standard deviation* of X is $\sigma = \sqrt{Var[X]}$

$\mu = E[X]$ is about *location*; $\sigma = \sqrt{Var(X)}$ is about *spread*



Two games:

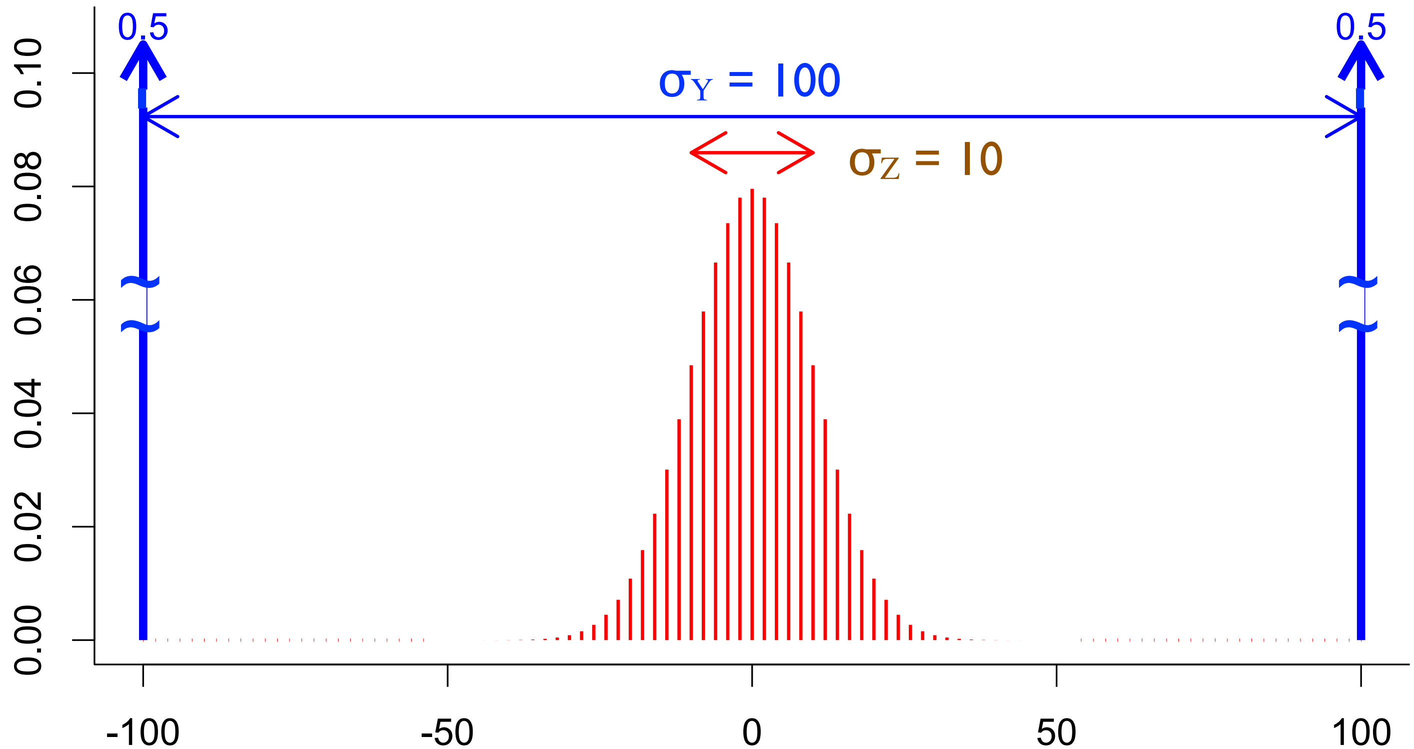
a) flip 1 coin, win Y = \$100 if heads, \$-100 if tails

b) flip 100 coins, win Z = (#(heads) - #(tails)) dollars

Same expectation in both: E[Y] = E[Z] = 0

Same extremes in both: max gain = \$100; max loss = \$100

But *variability* is very different:



$$\mathrm{Var}(X) = E[X^2] - (E[X])^2$$

$$\begin{aligned}
\mathrm{Var}(X) &= E[(X - \mu)^2] \\
&= \sum_x (x - \mu)^2 p(x) \\
&= \sum_x (x^2 - 2\mu x + \mu^2) p(x) \\
&= \sum_x x^2 p(x) - 2\mu \sum_x x p(x) + \mu^2 \sum_x p(x) \\
&= E[X^2] - 2\mu^2 + \mu^2 \\
&= E[X^2] - \mu^2
\end{aligned}$$

Example:

What is Var[X] when X is outcome of one fair die?

$$E[X^2] = 1^2\left(\frac{1}{6}\right) + 2^2\left(\frac{1}{6}\right) + 3^2\left(\frac{1}{6}\right) + 4^2\left(\frac{1}{6}\right) + 5^2\left(\frac{1}{6}\right) + 6^2\left(\frac{1}{6}\right)$$
$$= \left(\frac{1}{6}\right)(91)$$

E[X] = 7/2, so

$$\mathrm{Var}(X) = \frac{91}{6} - \left(\frac{7}{2}\right)^2 = \frac{35}{12}$$

$$\mathrm{Var}[aX+b] = a^2\,\mathrm{Var}[X]$$

$$\begin{aligned}\mathrm{Var}(aX + b) &= E[(aX + b - a\mu - b)^2] \\ &= E[a^2(X - \mu)^2] \\ &= a^2 E[(X - \mu)^2] \\ &= a^2 \mathrm{Var}(X)\end{aligned}$$

Ex:

$$X = \begin{cases} +1 & \text{if Heads} \\ -1 & \text{if Tails} \end{cases}$$

E[X] = 0

Var[X] = 1

$$Y = \begin{cases} +1000 & \text{if Heads} \\ -1000 & \text{if Tails} \end{cases}$$

Y = 1000 X

E[Y] = E[1000 X] = 1000 E[x] = 0

Var[Y] = Var[1000 X]

$= 10^6 \mathrm{Var}[X] = 10^6$

In general: $\mathrm{Var}[X+Y] \neq \mathrm{Var}[X] + \mathrm{Var}[Y]$

Ex 1:

Let $X = \pm 1$ based on 1 coin flip

As shown above, $E[X] = 0$, $\mathrm{Var}[X] = 1$

Let $Y = -X$; then $\mathrm{Var}[Y] = (-1)^2\mathrm{Var}[X] = 1$

But $X+Y = 0$, always, so $\mathrm{Var}[X+Y] = 0$

Ex 2:

As another example, is $\mathrm{Var}[X+X] = 2\mathrm{Var}[X]$?