Sept 27th Recall: expected value / expectation

$$\mu = E[X] = \sum_{x:p(x)>0} x \cdot p(x)$$

$$= \sum_{\text{values of } X} (\text{value of } x) \cdot P_X(\text{value of } x)$$

If g is a function, the $g(X)$ has an expectation of

$$E[g(X)] = \sum_{\text{all } x} g(x) \cdot P(X=x)$$

ex $\underset{g(X)=X^2}{E[X^2]} = \sum_{\text{all } x} \underset{g(x)}{x^2} \cdot P(X=x)$

stays as x

ex roll a die & $X = \text{value} = \{1,2,3,4,5,6\}$

find $E[X^2]$.

$P(X=1) = \frac{1}{6}$

$P(X=4) = \frac{1}{6}$

Let $X^2 = Y = \{1,4,9,16,25,36\}$

6 equally likely still

$$E[Y] = 1 \cdot \tfrac{1}{6} + 4 \cdot \tfrac{1}{6} + 9 \cdot \tfrac{1}{6} + 16 \cdot \tfrac{1}{6} + 25 \cdot \tfrac{1}{6} + 36 \cdot \tfrac{1}{6} = \frac{91}{6}$$

$$E[X^2] = 1^2 \cdot \tfrac{1}{6} + 2^2 \cdot \tfrac{1}{6} + 3^2 \cdot \tfrac{1}{6} + 4^2 \cdot \tfrac{1}{6} + 5^2 \cdot \tfrac{1}{6} + 6^2 \cdot \tfrac{1}{6} = 91/6$$

x squared

prob stay the same

<u>Variance</u>: if we have a r.v. X with $\mu = E[X]$ then the variance is defined as

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

$(X-\mu)^2 = g(X)$

↳ tells you how spread out the values of your r.v. are

<u>ex</u> r.v. X, Y, Z, W w/ pmf below. Compute $E[\cdot]$ & $\mathrm{Var}(\cdot)$

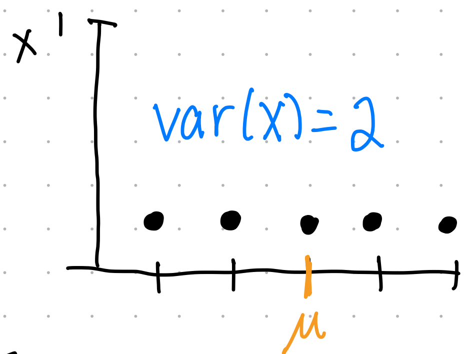
value of X,Y,Z,W	1	2	3	4	5
pmf $p(x)$	$1/5$	$1/5$	$1/5$	$1/5$	$1/5$
pmf $p(y)$	$1/10$	$2/10$	$4/10$	$2/10$	$1/10$
pmf $p(z)$	$5/10$	0	0	0	$5/10$
pmf $p(w)$	0	0	1	0	0

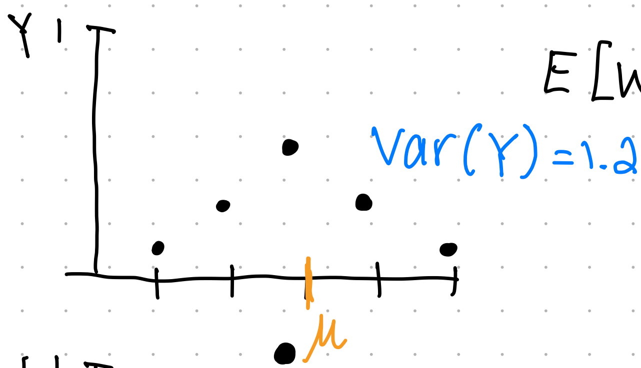
$E[X] = 1\cdot\frac{1}{5} + 2\cdot\frac{1}{5} + 3\cdot\frac{1}{5} \cdot 4\cdot 1/5 + 5\cdot 1/5 = 3$

$E[Y] = 1\cdot\frac{1}{10} + 2\cdot\frac{2}{10} + 3\cdot\frac{4}{10} + 4\cdot\frac{2}{10} + 1\cdot\frac{1}{10} = 3$

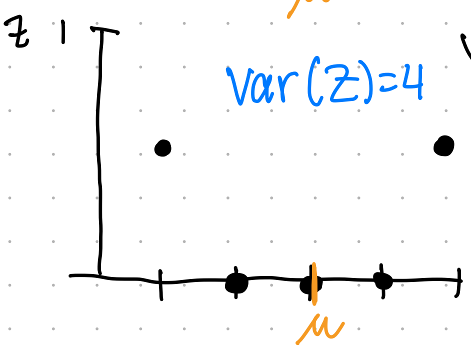
$E[Z] = 1\cdot\frac{5}{10} + 5\cdot\frac{5}{10} = 3$

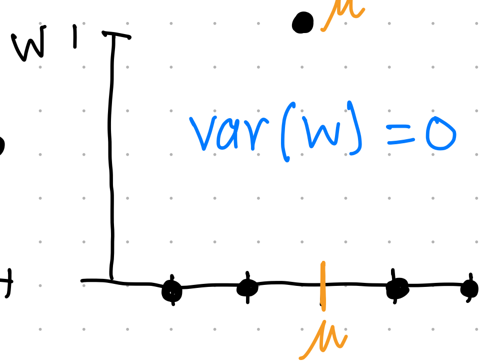
$E[W] = 3\cdot 1 = 3$





same $E[X] = \mu$!





$$\mathrm{Var}(X) = E[(X-\mu)^2] = E[(X-3)^2]$$

$$= (1-3)^2\cdot 1/5 + (2-3)^2\cdot 1/5 + (3-3)^2\cdot 1/5$$

$$+ (4-3)^2 \cdot 1/5 + (5-3)^2 \cdot 1/5 = 2$$

$$\mathrm{Var}(Y) = (1-3)^2 \cdot \frac{1}{10} + (2-3)^2 \cdot \frac{2}{10} + (3-3)^2 \frac{4}{10} + (4-3)^2 \frac{2}{10} + (5-3)^2 \frac{1}{10}$$

$$\mathrm{Var}(Z) = 4 \qquad\qquad = 1.2$$

$$\mathrm{Var}(W) = 0$$

---

alternatively, there is another way to compute variance

$$\mathrm{Var}(X) = E[(X-\mu)^2]$$ by def'n of $\mathrm{Var}(x)$

$$= \sum_x (x-\mu)^2 p(x)$$

$\mu = E[X]$ constant!

$$= \sum_x (x^2 - 2\mu x + \mu^2)\, p(x)$$

$$= \sum_x x^2 p(x) - 2\mu \underbrace{\sum_x x\, p(x)}_{\mu} + \mu^2 \underbrace{\sum_x p(x)}_{1}$$

$$= E[X^2] - 2\mu \cdot \mu + \mu^2$$

$$= E[X^2] - 2\mu^2 + \mu^2$$

$$= E[X^2] - \mu^2$$

$$= E[X^2] - (E[X])^2$$

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

ex roll a die, $X = \text{value} = \{1,2,3,4,5,6\}$

$$E[X] = 1\cdot\frac{1}{6} + 2\cdot\frac{1}{6} + 3\cdot\frac{1}{6} + 4\cdot\frac{1}{6} + 5\cdot\frac{1}{6} + 6\cdot\frac{1}{6} = \frac{7}{2} = \mu$$

$$Var(X) = E[(X-\mu)^2] = \left(1-\frac{7}{2}\right)^2\cdot\frac{1}{6} + \left(2-\frac{7}{2}\right)^2\cdot\frac{1}{6} + \dots + \left(6-\frac{7}{2}\right)^2\cdot\frac{1}{6} = \frac{35}{12}$$

OR $$E[X^2] = 1^2\cdot\frac{1}{6} + 2^2\frac{1}{6} + 3^2\cdot\frac{1}{6} + 4^2\cdot\frac{1}{6} + 5^2\cdot\frac{1}{6} + 6^2\cdot\frac{1}{6} = \frac{91}{6}$$

$$\Rightarrow Var(X) = E[X^2] - \mu^2 = \frac{91}{6} - \left(\frac{7}{2}\right)^2 = \frac{35}{12}$$

---

* the square root of the variance is called the <u>standard deviation</u>, $SD(X)$

$$SD(X) = \sqrt{Var(X)}$$

* Another Identity that is useful : $a, b \in \mathbb{R}$ (constants)

$$Var(aX+b) = E[((aX+b) - E[aX+b])^2]$$

$E[aX+b] = aE[X]+b = a\mu + b$

$$= E[(aX+b-(a\mu+b))^2]$$

$$= E[(aX - a\mu)^2] = E[a^2(X-\mu)^2]$$
$$= a^2 E[(X-\mu)^2]$$
$$= a^2 \, Var(X)$$

ex $Var(3X+5) = 3^2 Var(X)$

shift left/right $\longrightarrow$ does not change how spread out pmf is

affects how spread out pmf is

Let X be a r.v. The <u>cumulative distribution function</u> (CDF) of X is defined as

$$F_X(x) := P(X \leq x) \quad \text{for any } x \in \mathbb{R}$$

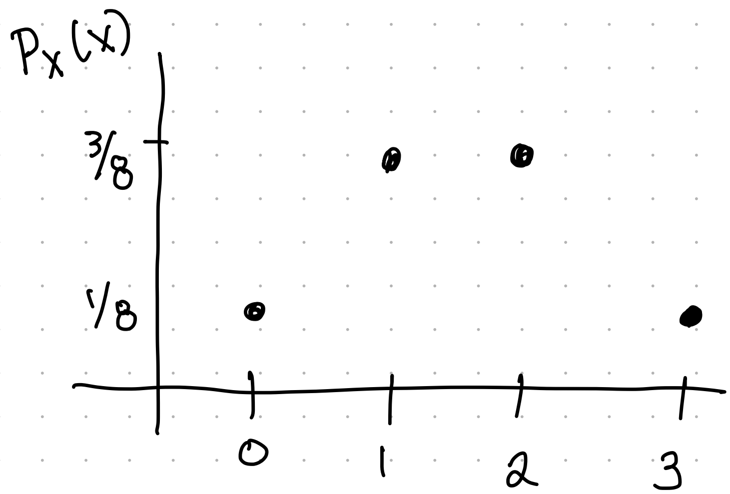
ex Let X be the # of heads shown after 3 coin flips.

$P(X=0) = \frac{1}{8}$

$P(X=1) = \frac{3}{8}$

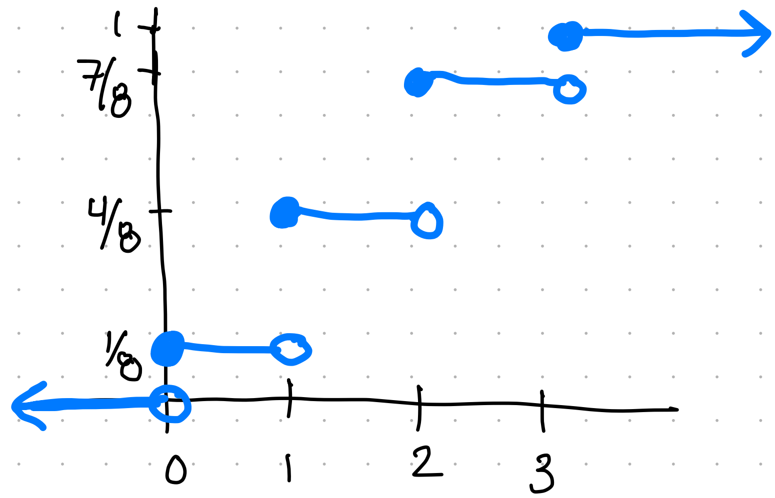
$P(X=2) = \frac{3}{8}$

$P(X=3) = \frac{1}{8}$



CDF adds up all the probs in order from left to right

$$F_X(x) = \begin{cases} 0 & x < 0 \\ 1/8 & 0 \leq x < 1 \\ 4/8 & 1 \leq x < 2 \\ 7/8 & 2 \leq x < 3 \\ 1 & x \geq 3 \end{cases}$$



Properties of CDF

① F is non-decreasing → always add pos probs

② $\lim_{x \to \infty} F(x) = 1$ ← should approach 1 on right hand side

③ $\lim_{x \to -\infty} F(x) = 0$ ← start off w/ zero prob

Specific examples of R.V. (some we have already seen! But we just didn't give it a name)

Bernoulli RV.: one trial / experiment where the outcome can be classified as either a success OR a failure, the prob of success is p.

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

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

Binomial RV: n # of independent trials where each success has a prob of p

ex: repeated coin flip

→ n trials of Bernoulli

If X represents the # of successes that occur in n trials, then X is called a Binomial R.V. with parameters (n,p).

(Bernoulli is just one trial of Binomial w/ $(1,p)$)

The pmf of Binomial R.V.

$$p(k) = \binom{n}{k} p^k (1-p)^{n-k} \qquad k=1,\ldots,n$$

↓

diff ways to list sequences of n outcomes w/ k successes

↳ order of successes does not matter

ex 3 fair coins are flipped. find pmf of # of heads shown. (Independence)

$X = \#$ of heads $\quad n=3 \quad p=\frac{1}{2}$

$$P(X=\underset{\hookrightarrow k}{0}) = \binom{3}{0}\cdot\left(\frac{1}{2}\right)^0\cdot\left(1-\frac{1}{2}\right)^{3-0}$$

$$= \frac{\cancel{3!}}{0!\,\cancel{3!}}\cdot 1\cdot\left(\frac{1}{2}\right)^3 = \frac{1}{8}$$

$$P(X=\underset{\hookrightarrow k}{1}) = \binom{3}{1}\cdot\left(\frac{1}{2}\right)^1\left(1-\frac{1}{2}\right)^{3-1}$$

$$= 3\cdot\frac{1}{2}\cdot\frac{1}{2^2} = \frac{3}{8}$$

$$P(X=2) = \binom{3}{2}\cdot\left(\frac{1}{2}\right)^2\left(1-\frac{1}{2}\right)^{3-2} = \frac{3}{8}$$

$$P(X=3) = \binom{3}{3}\cdot\left(\frac{1}{2}\right)^3\left(1-\frac{1}{2}\right)^{3-3} = \frac{1}{8}$$

expect this!

ex Bill has a 0.6 prob of making a free throw (basketball). Suppose each free throw is independent of the other. If he attempts 10 free throws, what is the prob he makes (a) 2 of them? (b) at least 2?

$P(K) = \binom{n}{K} p^K (1-p)^{n-K} \quad K = 1, \ldots, n$

Binomial R.V. $n = 10$ $p = .6$

a) $K = 2 \rightarrow P(2) = \binom{10}{2} (.6)^2 (.4)^{10-2}$

$$= \binom{10}{2} (.6)^2 (.4)^8$$

b) $P(K=2) + P(K=3) + P(K=4) + \ldots + P(K=10)$

OR $= 1 - \left[P(K=0) + P(K=1) \right]$

$$= 1 - \binom{10}{0} (.6)^0 (.4)^{10} - \binom{10}{1} (.6)^1 (.4)^9$$

$= 0.998$ 99.8%

Back to Bernoulli (one trial) $X=1$ success $X=0$ fail

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

Variance (2nd method) : $E[X^2] - (E[X])^2 = Var(x)$

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

$$\Rightarrow Var(x) = E[X^2] - (E[X])^2 = p - p^2 = p(1-p)$$

	Bernoulli	Binomial (n-trials)
$E[X]$:	p	np
$Var(X)$:	$p(1-p)$	$np(1-p)$

ex free throws above $n=10$ $p=.6$

expected # of free throw successes after 10 trials/throws $= np = 10 \cdot .6 = \underline{6}$

Negative Binomial R.V. Describes the # of trials needed to achieve a fixed # of successes in a series of independent Bernoulli trials, prob of success is p (failure is 1-p)

r = # of successes we want to achieve/reach

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

$n = r, r+1, r+2, \ldots$

↳ binomial coffecient (combination)

↳ # of ways to arrange r-1 successes in the first n-1 trials

normal Binomial: you count # of successes in a fixed # of Bern trials

"negative" Binomial, counts the # of trials (or failures) needed to achieve a fixed # of successes

"negative" ⟷ "opposite"

$$E[X] = \frac{r}{p} \qquad Var(X) = \frac{r(1-p)}{p^2}$$

Negative Binomial R.V.

ex Bill has 0.7 chance of making free throw. We want to find # of free throws n needed to make $r=5$ successful shots.

ex Let's calculate the prob that Bill needs to attempt exactly $n=8$ free throws to make 5 successful shots.

$n=8$

$r=5$

$p=.7$

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

$$P(X=8) = \binom{8-1}{5-1} (.7)^5 (.3)^{8-5}$$

$$= \binom{7}{4} (.7)^5 (.3)^3 \approx 0.159$$

the prob that Bill needs to attempt exactly 8 free throws to make 5 successful shots is 15.9 %

$\Rightarrow$ also $E[X] = \frac{r}{p} = \frac{5}{.7} \approx 7.14$

Bill is expected to attempt ~7.14 free throws to make 5 successful shots.

Hypergeometric R.V. models the # of successes in a fixed # of draws, without replacement, from a finite population that contains a known # of successes & failures. $\rightarrow$ No independence since prob changes after each draw

N : total population size
K : " successes in pop
n : # of draws or trials

# of ways to get success

$$P(X=k) = \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}$$

# of ways to get failures

$$\Rightarrow E[X] = \frac{nK}{N} \qquad Var(X) = n \cdot \frac{K}{N} \cdot \frac{N-K}{N} \cdot \frac{N-n}{N-1}$$

ex Imagine 10 balls. 6 red
4 blue

$N = 10 \quad K = 6$

You randomly select 4 w/o replacement.

Want prob of exactly 2 red balls out of 4 draws.

$n = 4$

$k = 2$

$$P(X=k) = \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}$$

$$P(X=2) = \frac{\binom{6}{2}\binom{4}{2}}{\binom{10}{4}}$$