Topic 6 –
Variance / Standard Deviation

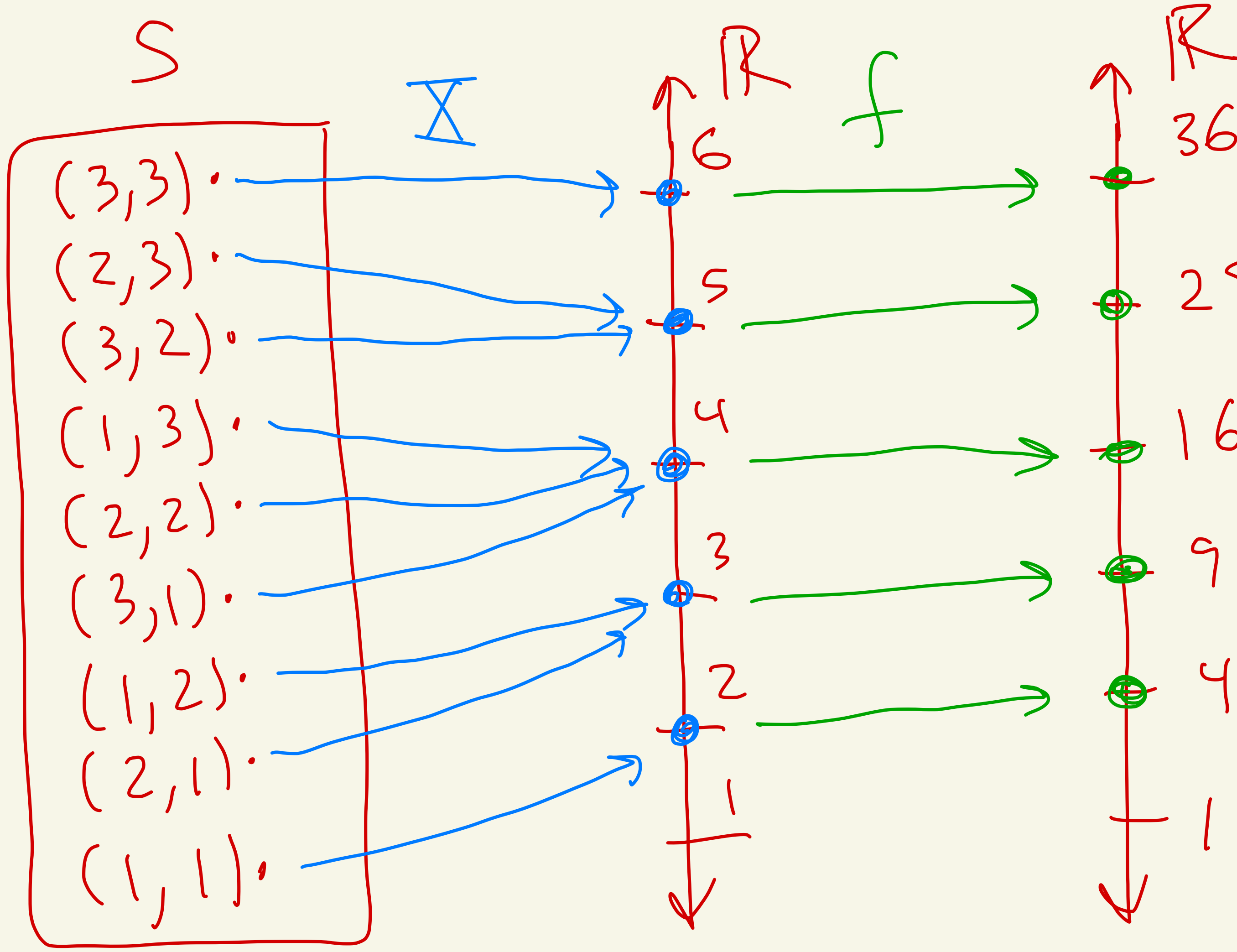
Topic 6 – More on Expected Value, Variance, Standard Deviation

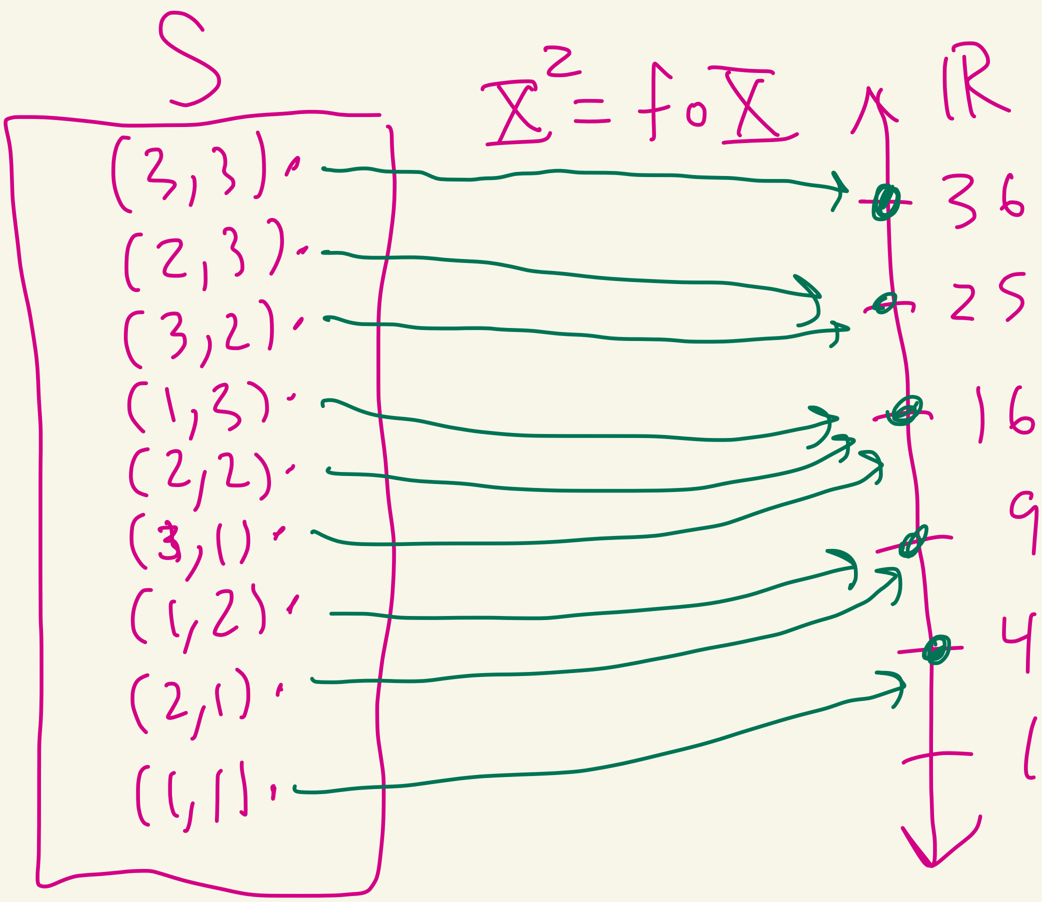
Given a discrete random variable $X : S \longrightarrow \mathbb{R}$, if you take a function $f : \mathbb{R} \to \mathbb{R}$ and compute the composition $f \circ X$ then you will get a new random variable

under appropriate conditions such as a finite sample space

<u>Ex:</u> Suppose we roll two 3-sided dice, each labeled 1, 2, 3 where each side is equally likely. Let X be the sum of the dice.

Let $f:\mathbb{R}\to\mathbb{R}$ be $f(t)=t^2$.





Example

$(f\circ X)(2,3)$

$= f(X(2,3))$

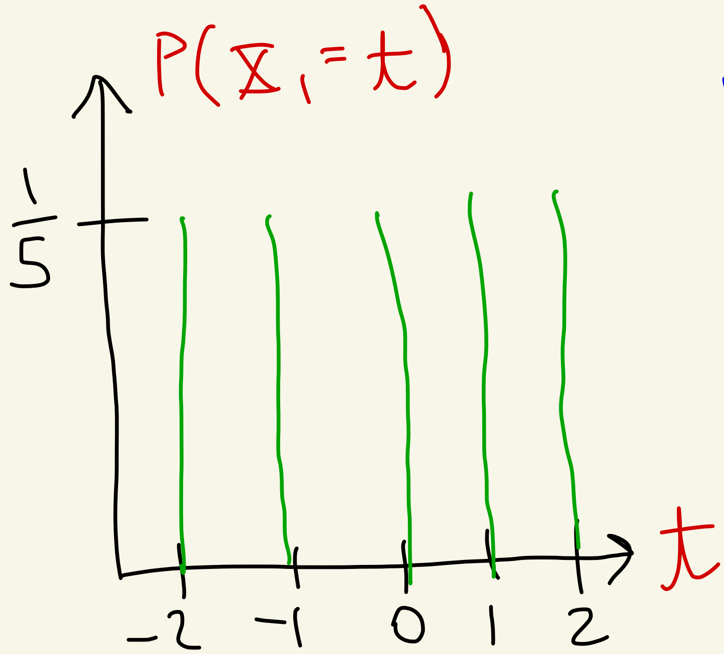
$= f(2+3)$

$= f(5)$

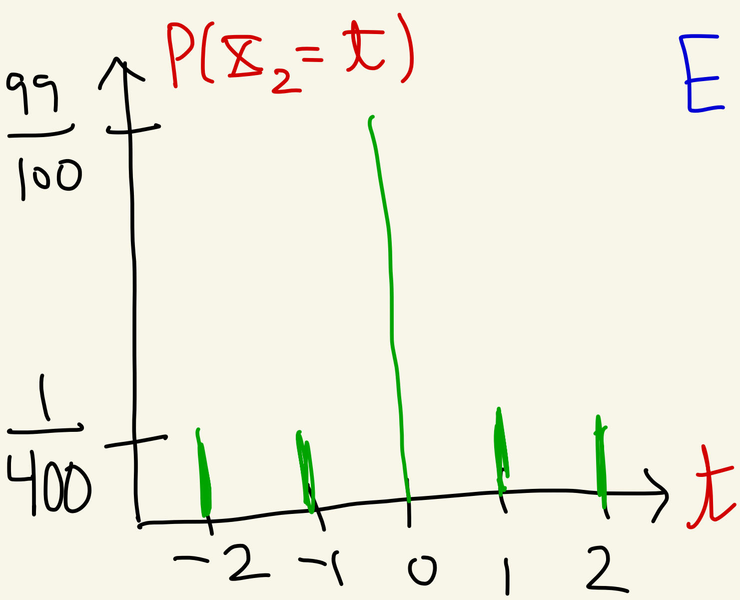
$= 5^2 = 25$

Expected value doesn't give all the info for a probability function. It can't detect how much the data is spread out or not spread out.

<u>Ex:</u> (Two probability functions w/ same expected value but data spread out differently)



$$E[X_1] = (-2)\left(\tfrac{1}{5}\right) + (-1)\left(\tfrac{1}{5}\right) + (0)\left(\tfrac{1}{5}\right) + (1)\left(\tfrac{1}{5}\right) + (2)\left(\tfrac{1}{5}\right) = 0$$



$$E[X_2] = (-2)\left(\tfrac{1}{400}\right) + (-1)\left(\tfrac{1}{400}\right) + (0)\left(\tfrac{99}{100}\right) + (1)\left(\tfrac{1}{400}\right) + (2)\left(\tfrac{1}{400}\right) = 0$$

We want a number that measures the average magnitude of the fluctuations of the random variable from it's expected value.

Let $\mu = E[X]$.

One might try to measure the expected value of $|X-\mu|$, ie the expected value of the distance between X's values and μ. This is too hard to use.

So instead we measure

$$E\Big[\underbrace{(X-\mu)^2}_{\text{square of distance between } X \text{ and } \mu}\Big]$$

$$\boxed{\begin{aligned} |X-\mu| &= \sqrt{(X-\mu)^2} \\ (X-\mu)^2 &= |X-\mu|^2 \end{aligned}}$$

Def: Let X be a discrete random variable. Define the variance of X to be $\mathrm{Var}(X) = E\left[(X-\mu)^2\right]$.

Define the standard deviation of X to be $\sigma_X = \sigma = \sqrt{\mathrm{Var}(X)}$

(where $\mu = E[X]$)

Note: One can prove that if $x_1, x_2, x_3, \ldots$ are the outputs of X, and $f : \mathbb{R} \to \mathbb{R}$, then

$$E[f(X)] = \sum_i f(x_i) \cdot P(X = x_i)$$

proof is below

Thus,

$$\mathrm{Var}(X) = \sum_i (x_i - \mu)^2 \cdot P(X = x_i)$$

(where $\mu = E[X]$)

proof of above formula:

Let $A = \{x_1, x_2, x_3, \dots\}$ be the range of X.
The range of $f \circ X$ is $f(A) = \{f(x_1), f(x_2), f(x_3), \dots\}$.

Thus,

$$E[f(X)] = \sum_{x \in A} f(x) \cdot \underbrace{P(f \circ X = f(x))}_{P(\{w \mid f(X(w)) = f(x)\})}$$

$= P(\{w \mid X(w) = y \text{ and } f(y) = f(x)\})$

$$= \sum_{x \in A} f(x) \cdot \sum_{\substack{y \in A \\ \text{where} \\ f(y) = f(x)}} P(X = y)$$

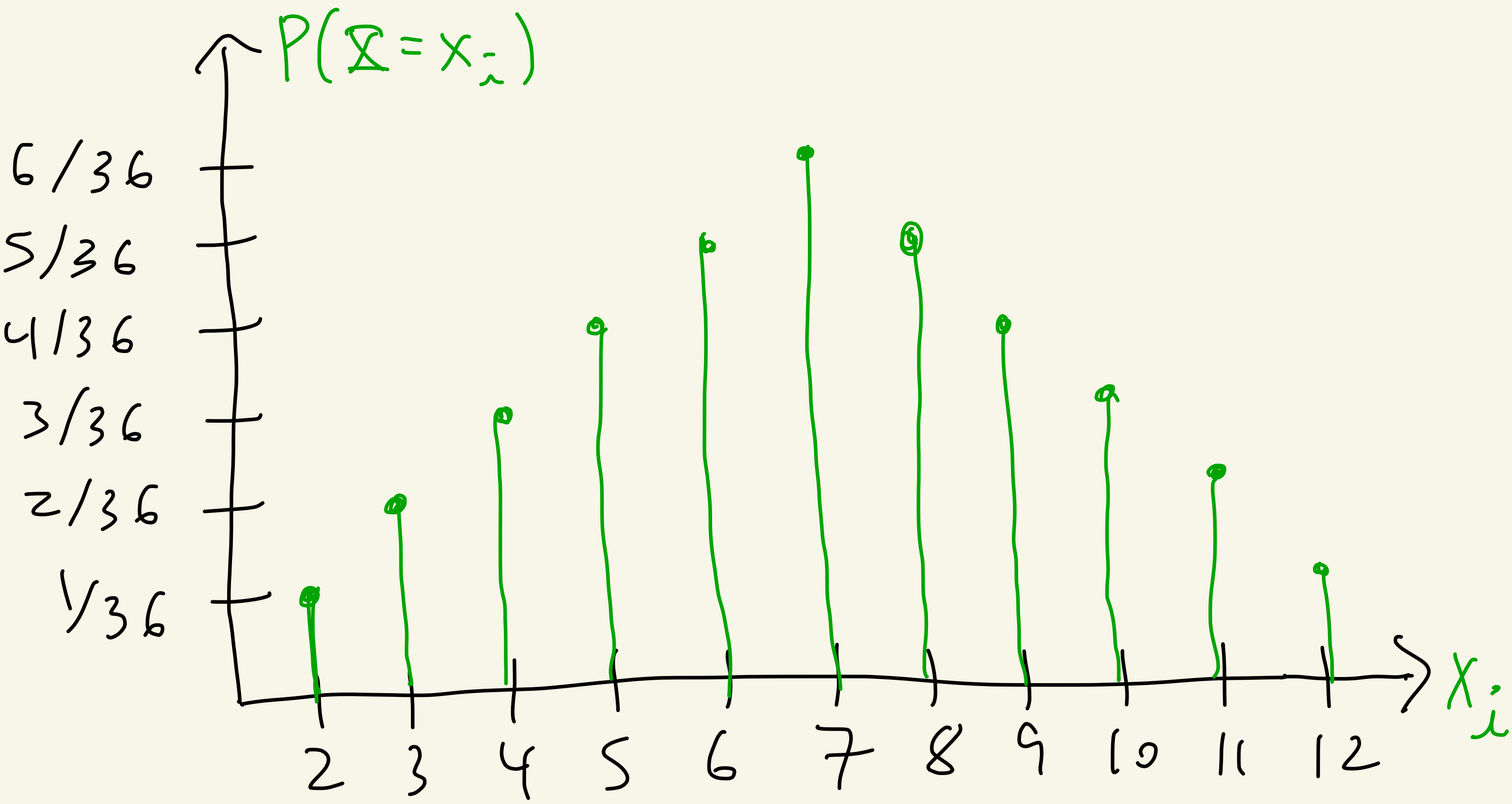
$$= \sum_{x \in A} \sum_{\substack{y \in A \\ \text{where} \\ f(y) = f(x)}} f(x) \cdot P(X = y)$$

$$= \sum_{x \in A} \sum_{\substack{y \in A \\ \text{where} \\ f(y) = f(x)}} f(y) \cdot P(X = y)$$

$$= \sum_{z \in A} f(z) P(X = z)$$

$\blacksquare$

Ex: Consider the experiment of rolling two 6-sided dice.
Let X be the sum of the dice



Recall that $\mu = E[X] = 7$.

Then,

$$Var(X) = \sum_{x_i} (x_i - 7)^2 \cdot P(X = x_i)$$

$$= (2-7)^2 \cdot \left(\frac{1}{36}\right) + (3-7)^2 \cdot \left(\frac{2}{36}\right)$$

$$+(4-7)^2\left(\frac{3}{36}\right)+(5-7)^2\left(\frac{4}{36}\right)$$

$$+(6-7)^2\cdot\left(\frac{5}{36}\right)+(7-7)^2\left(\frac{6}{36}\right)$$

$$+(8-7)^2\left(\frac{5}{36}\right)+(9-7)^2\cdot\left(\frac{4}{36}\right)$$

$$+(10-7)^2\left(\frac{3}{36}\right)+(11-7)^2\left(\frac{2}{36}\right)$$

$$+(12-7)^2\left(\frac{1}{36}\right)=\boxed{\frac{35}{6}}\approx\boxed{5.83}$$

$$\sigma_X=\sqrt{\mathrm{Var}(X)}=\sqrt{\frac{35}{6}}\approx\boxed{2.415}$$

Theorem: Let X be a discrete random variable. Let $\mu = E[X]$.

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

Proof: Let $x_1, x_2, x_3, \ldots$ be the values of X. Then

$$\mathrm{Var}(X) = \sum_i (x_i - \mu)^2 \cdot P(X = x_i)$$
$$= \sum_i x_i^2 \cdot P(X = x_i)$$
$$- 2\mu \boxed{\sum_i x_i \cdot P(X = x_i)}$$
$$+ \mu^2 \boxed{\sum_i P(X = x_i)}$$

equals $\mu = E[X]$

equals 1

$$= \boxed{\sum_i x_i^2 \cdot P(X = x_i)}$$

equals $E[X^2]$

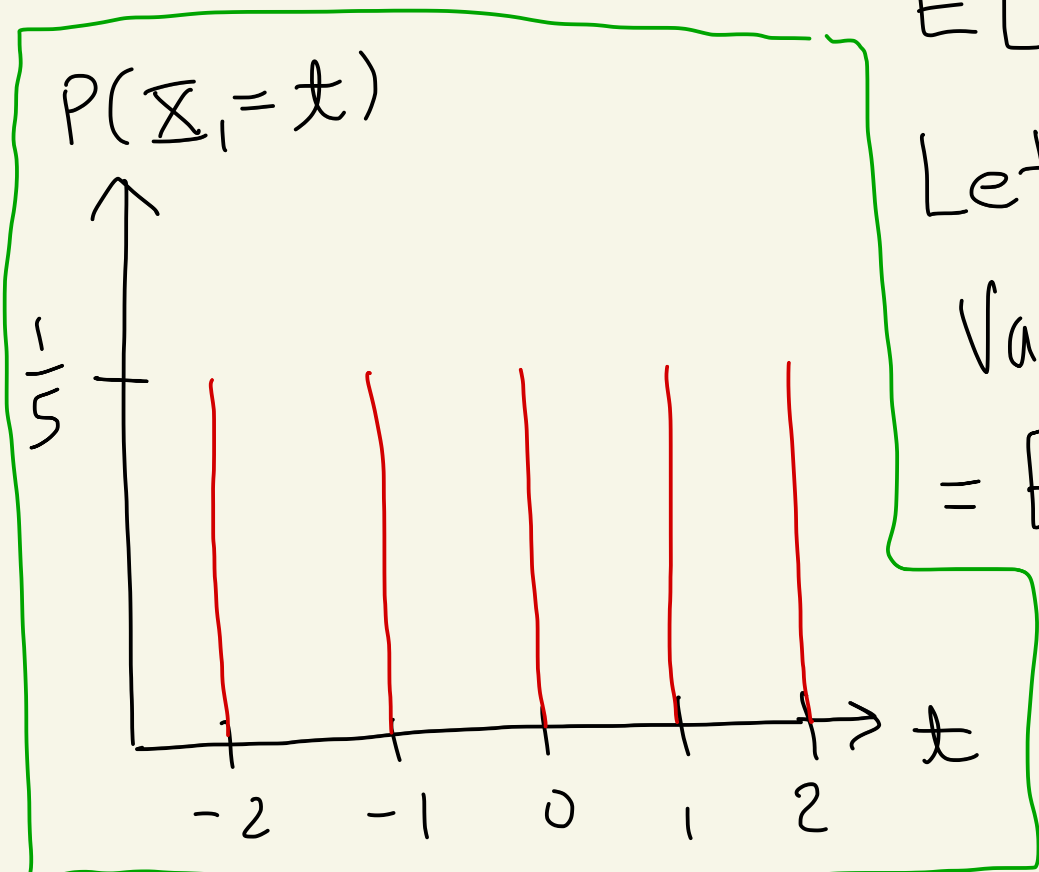
$$- 2\mu^2 + \mu^2$$

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

$\blacksquare$

Previously we had two examples with the same expected value of 0 but the data was spread out differently. Let's calculate the variance / standard deviation of those examples.

$\boxed{\text{Ex:}}$



We calculated $E[X_1] = 0$.

Let's calculate

$$\mathrm{Var}(X_1) = E[X_1^2] - (\underbrace{E[X_1]}_{0})^2$$

$$= E[X_1^2]$$

We have

$$E[X_1^2] = (-2)^2 \cdot \underbrace{\left(\frac{1}{5}\right)}_{P(X_1=-2)} + (-1)^2 \cdot \underbrace{\left(\frac{1}{5}\right)}_{P(X_1=-1)}$$

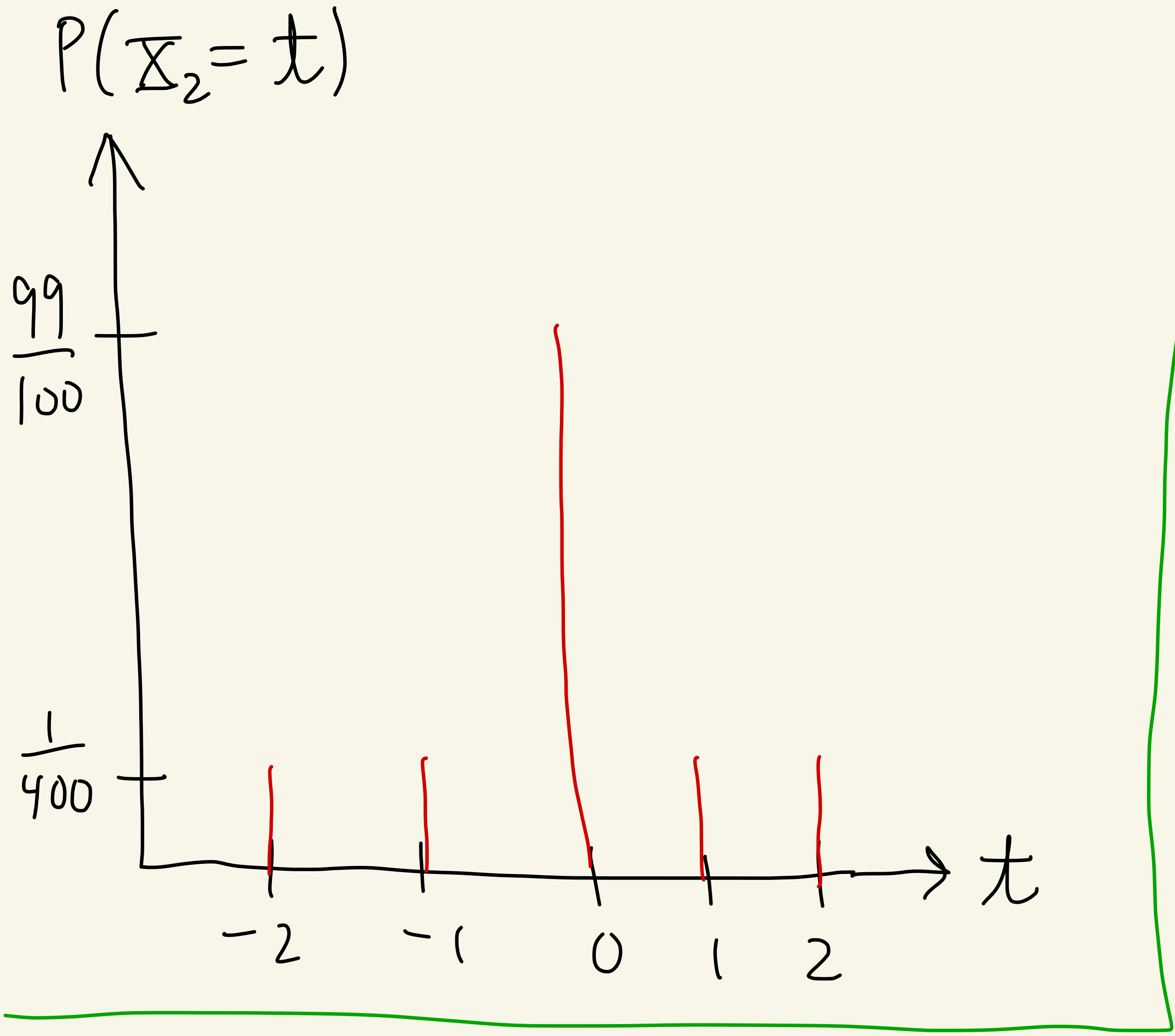
$$+ (0)^2\left(\frac{1}{5}\right) + (1)^2\left(\frac{1}{5}\right) + (2)^2\left(\frac{1}{5}\right)$$

$$= (4+1+0+1+4) \cdot \frac{1}{5} = 2$$

So, $\boxed{\mathrm{Var}(X_1^2) = 2}$

Then, $\boxed{\sigma_{X_1} = \sqrt{\mathrm{Var}(X_1^2)} = \sqrt{2} \approx 1.414}$

We also had the following example:



We saw that $E[X_2] = 0$ previously.

Thus,

$$\mathrm{Var}(X_2) = E[X_2^2] - (\underbrace{E[X_2]}_{0})^2$$

$$= E[X_2^2]$$

And so,

$$E[\bar{X}_2^2] = (-2)^2 \cdot \underbrace{\left(\frac{1}{400}\right)}_{P(\bar{X}_2=-2)} + (-1)^2\left(\frac{1}{400}\right)$$

$$+ (0)^2\left(\frac{99}{100}\right) + (1)^2\left(\frac{1}{400}\right) + (2)^2\left(\frac{1}{400}\right)$$

$$= \frac{10}{400} = \frac{1}{40}$$

Thus,

$$\boxed{Var(\bar{X}_2) = \frac{1}{40}}$$

$$\boxed{\sigma_{\bar{X}_2} = \sqrt{Var(\bar{X}_2)} = \sqrt{\frac{1}{40}} \approx 0.158}$$

Theorem: Let X be a binomial random variable with parameters n and p. Then,

$$Var(X) = np(1-p)$$

$$\sigma_X = \sqrt{np(1-p)}$$

$\boxed{\text{proof:}}$ Recall that $E[X] = np$.

We have that

$$E[X^2] = \sum_{i=0}^{n} i^2 \overbrace{\binom{n}{i} p^i (1-p)^{n-i}}^{P(X=i)}$$

$$= \sum_{i=1}^{n} i^2 \frac{n!}{i!(n-i)!} \cdot p^i (1-p)^{n-i}$$

$$= np \sum_{i=1}^{n} i \frac{n!}{(i-1)!(n-i)!} \cdot p^{i-1} (1-p)^{n-i}$$

$k = i-1$

$$= np \sum_{k=0}^{n-1} (k+1) \frac{(n-1)!}{k!((n-1)-k)!} \cdot p^k (1-p)^{(n-1)-k}$$

$$= np \underbrace{\sum_{k=0}^{n-1} k \binom{n-1}{k} p^k (1-p)^{(n-1)-k}}_{\text{E[Y] where Y is a binomial random variable w/ parameters } n-1 \text{ \& } p.}$$

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

use binomial thm: $(a+b)^l = \sum_{j=0}^{l} \binom{l}{j} a^l b^{l-j}$

$$= (np) \cdot (n-1) \cdot p + (np) \cdot \big(p + (1-p)\big)^{n-1}$$

$$= n^2 p^2 - np^2 + np$$

Thus,

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

$$= n^2 p^2 - np^2 + np - (np)^2$$

$$= np - np^2 = np(1-p). \quad \blacksquare$$

Ex: Suppose we flip a coin 100 times. Let X be the number of heads that occur.

Then, X is a binomial random variable with $n = 100$ and $p = \frac{1}{2}$

probability of heads on a single flip

Then,

showed in topic 5

$$E[X] = np = 100\left(\frac{1}{2}\right) = 50$$

$$Var(X) = np(1-p) = 100\left(\frac{1}{2}\right)\left(1-\frac{1}{2}\right) = 25$$

$$\sigma_X = \sqrt{25} = 5$$

Theorem (Markov's Inequality)

Let X be a non-negative discrete random variable.

non-negative means: $X(\omega) \geq 0$ for all ω in the sample space

Let $\mu = E[X]$.

Then for any real number $t > 0$ we have that

$$P(X \geq t) \leq \frac{\mu}{t}$$

(since μ is fixed)

Note: $\frac{\mu}{t} \to 0$ as $t \to \infty$

add all these to get $P(X \geq t)$

proof:

Let A be the range of the function X.

Let

$B = \{x \mid x \in A \text{ and } x \geq t\}$.

Then,

$$
\begin{aligned}
E[X] &= \sum_{x \in A} x \cdot P(X = x) \\
&\geq \sum_{x \in B} x \cdot P(X = x) \\
&\geq \sum_{x \in B} t \cdot P(X = x) \\
&= t \sum_{x \in B} P(X = x) = t\, P(X \geq t)
\end{aligned}
$$

since $B \subseteq A$ and X is non-negative

since $x \geq t$ if $x \in B$

Thus, $P(X \geq t) \leq \dfrac{E[X]}{t}$. $\blacksquare$

Theorem: (Chebyshev's Inequality)

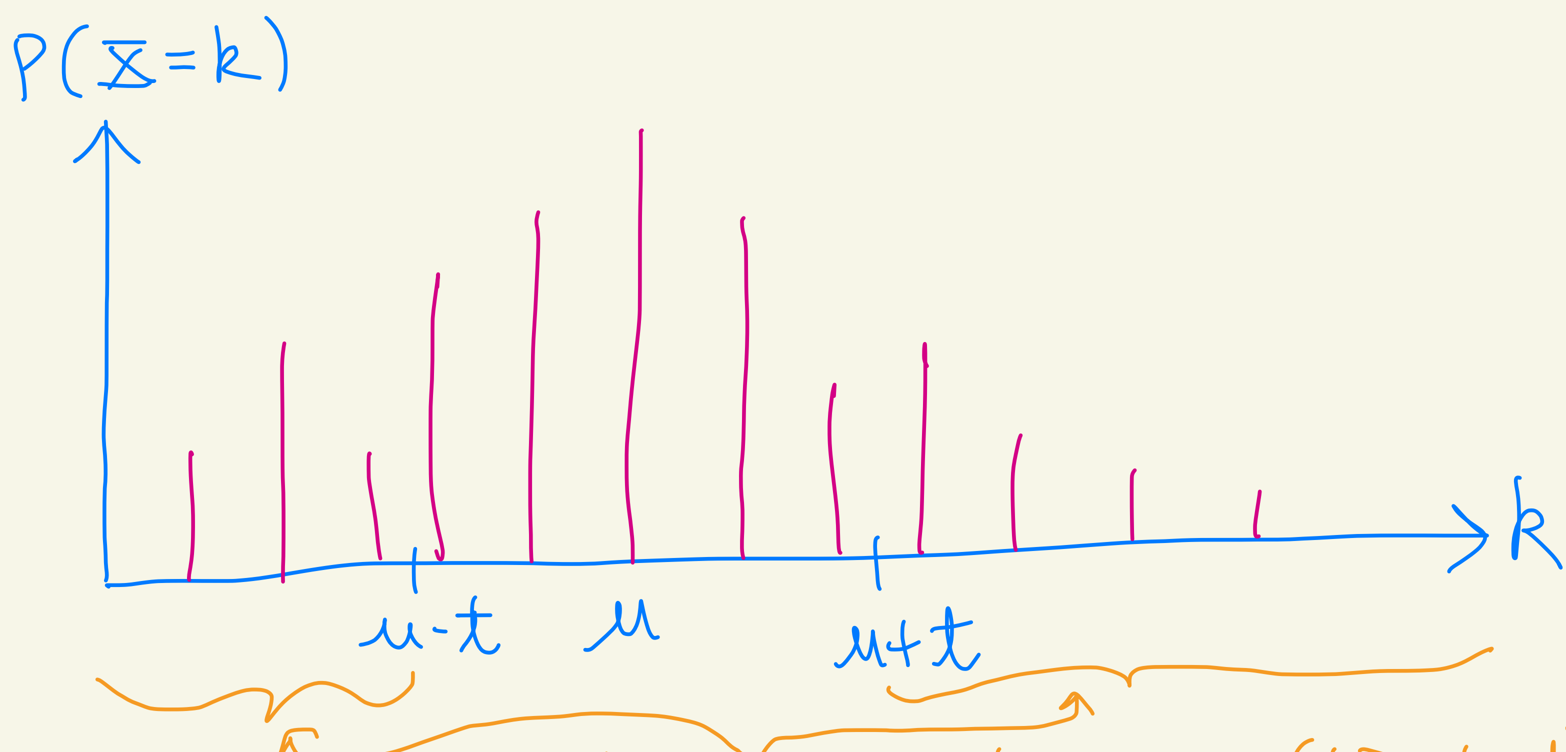
Let X be a discrete random variable. Let $\mu = E[X]$.

Let $\sigma = \sqrt{Var(X)}$.

Then for any $t > 0$, we have

$$\underbrace{P(|X-\mu| \geq t)} \leq \frac{\sigma^2}{t^2}$$

means: $P(\{\omega \mid \omega \in S \text{ with } |X(\omega)-\mu| \geq t\})$



add these up to get $P(|X-\mu| \geq t)$

<u>proof:</u> The random variable

$(\bar{X}-\mu)^2$ is non-negative.

So, Markov's inequality gives:

$$\underbrace{P\left((\bar{X}-\mu)^2 \geq t^2\right)}_{\text{same as } P(|\bar{X}-\mu| \geq t)} \leq \frac{E\left[(\bar{X}-\mu)^2\right]}{t^2}$$

$$= \frac{Var(\bar{X})}{t^2}$$

$$= \frac{\sigma^2}{t^2}$$

$\blacksquare$

Ex: (HW 6 #5(b))

Let X be a discrete random variable with $\mu = E[X]$ and $\sigma = \sqrt{Var(X)}$.

Show that $P(|X-\mu| \geq 2\sigma) \leq \frac{1}{4}$

pf: By Chebyshev:

$$P(|X-\mu| \geq 2\sigma) \leq \frac{\sigma^2}{(2\sigma)^2} = \frac{1}{4}$$

$\blacksquare$