Variance

- The variance measures how spread out a random variable is around its mean.

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

→ Note that $\mathbb{E}[X]$ is a constant while X is random. It may be easier to understand this formula using notation that captures this:

$$\mu_X = \mathbb{E}[X] \qquad \mathrm{Var}[X] = \underbrace{\mathbb{E}\left[(X - \mu_X)^2\right]}$$

Can think of this as the expected value of the function $Y = (X - \mu_X)^2$.

- The standard deviation σ_X of X is the square root of its variance: $\boxed{\sigma_X = \sqrt{\mathrm{Var}[X]}}$

Sometimes denote variance by σ_X^2.

- Simple formula for calculating the variance of a linear function:

$$\boxed{\text{If } Y = aX + b, \text{ then } \mathrm{Var}[Y] = a^2 \mathrm{Var}[X].}$$

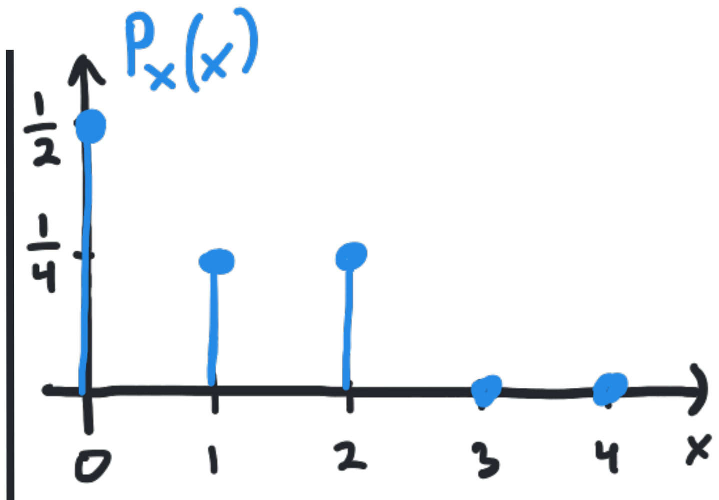
→ Why?

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

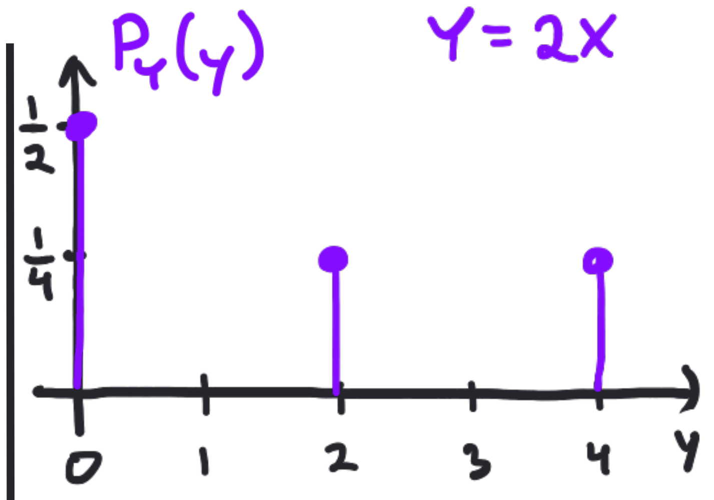
Linearity of Expectation

- **Example:**

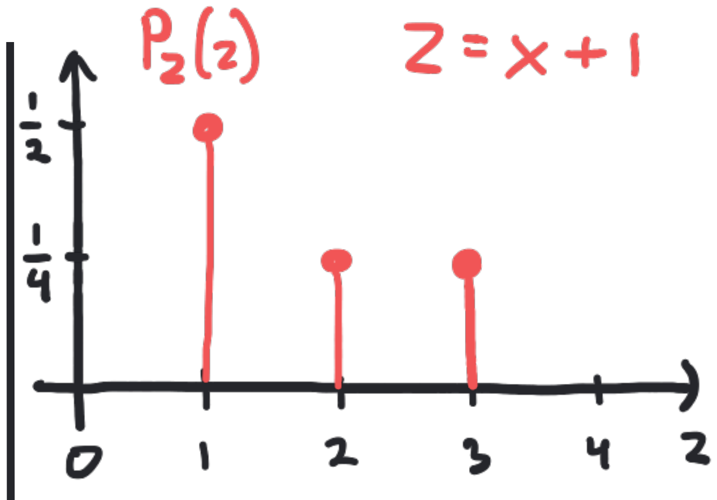
PMF



$Y = 2X$



$Z = X + 1$



$$\mathbb{E}[X] = \sum_{x \in R_X} x\, P_X(x)$$

$$0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{4} + 2 \cdot \frac{1}{4} = \frac{3}{4}$$

$$0 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} + 4 \cdot \frac{1}{4} = \frac{3}{2} = 2 \cdot \mathbb{E}[X]$$

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

$$\mathrm{Var}[X] = \mathbb{E}\left[(X - \mathbb{E}[X])^2\right] = \sum_{x \in R_X} (x - \mathbb{E}[X])^2 P_X(x)$$

$$\left(0 - \frac{3}{4}\right)^2 \cdot \frac{1}{2} + \left(1 - \frac{3}{4}\right)^2 \cdot \frac{1}{4} + \left(2 - \frac{3}{4}\right)^2 \cdot \frac{1}{4}$$
$$= \frac{9}{16} \cdot \frac{1}{2} + \frac{1}{16} \cdot \frac{1}{4} + \frac{25}{16} \cdot \frac{1}{4} = \frac{18 + 1 + 25}{64} = \frac{44}{64} = \frac{11}{16}$$

$$\left(0 - \frac{3}{2}\right)^2 \cdot \frac{1}{2} + \left(2 - \frac{3}{2}\right)^2 \cdot \frac{1}{4} + \left(4 - \frac{3}{2}\right)^2 \cdot \frac{1}{4}$$
$$= \frac{9}{4} \cdot \frac{1}{2} + \frac{1}{4} \cdot \frac{1}{4} + \frac{25}{4} \cdot \frac{1}{4} = \frac{18 + 1 + 25}{16} = \frac{44}{16} = \frac{11}{4} = 2^2\, \mathrm{Var}[X]$$

$$\left(1 - \frac{7}{4}\right)^2 \cdot \frac{1}{2} + \left(2 - \frac{7}{4}\right)^2 \cdot \frac{1}{4} + \left(3 - \frac{7}{4}\right)^2 \cdot \frac{1}{4}$$
$$= \frac{9}{16} \cdot \frac{1}{2} + \frac{1}{16} \cdot \frac{1}{4} + \frac{25}{16} \cdot \frac{1}{4} = \frac{18 + 1 + 25}{64} = \frac{44}{64} = \frac{11}{16} = \mathrm{Var}[X]$$

- Another formula for calculating the variance:

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

→ Why?

$$Var[X] = \sum_{x \in R_X} (x - \mu_X)^2 P_X(x) \quad \text{where} \quad \mu_X = \mathbb{E}[X]$$

$$= \sum_{x \in R_X} (x^2 - 2\mu_X x + \mu_X^2) P_X(x)$$

$$= \sum_{x \in R_X} x^2 P_X(x) - 2\mu_X \sum_{x \in R_X} x P_X(x) + \mu_X^2 \underbrace{\sum_{x \in R_X} P_X(x)}_{\text{Normalization}}$$

$$= \mathbb{E}[X^2] - 2\mu_X \mathbb{E}[X] + \mu_X^2 \cdot 1$$

$$= \mathbb{E}[X^2] - 2(\mathbb{E}[X])^2 + (\mathbb{E}[X])^2$$

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

- a bit more terminology: n^{th} moment $\mathbb{E}[X^n]$
  n^{th} central moment $\mathbb{E}[(X - \mathbb{E}[X])^n]$