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

\[
\text{Var}(X) = E[(X - \mu)^2] \\
= \sum_x (x - \mu)^2 p(x) \\
= \sum_x (x^2 - 2\mu x + \mu^2) p(x) \\
= \sum_x x^2 p(x) - 2\mu \sum_x x p(x) + \mu^2 \sum x p(x) \\
= E[X^2] - 2\mu^2 + \mu^2 \\
= E[X^2] - \mu^2
\]
Example:

What is $\text{Var}[X]$ 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

\[
\text{Var}(X) = \frac{91}{6} - \left( \frac{7}{2} \right)^2 = \frac{35}{12}
\]
 properties of variance

\[ \text{Var}[aX+b] = a^2 \text{Var}[X] \]

\[
\text{Var}(aX + b) = E[(aX + b - a\mu - b)^2] \\
= E[a^2(X - \mu)^2] \\
= a^2 E[(X - \mu)^2] \\
= a^2 \text{Var}(X)
\]

**Ex:**

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

\[ E[X] = 0 \quad \text{Var}[X] = 1 \]

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

\[ Y = 1000X \]

\[ E[Y] = E[1000X] = 1000E[X] = 0 \]
\[ \text{Var}[Y] = \text{Var}[1000X] = 10^6 \text{Var}[X] = 10^6 \]
In general: \( \text{Var}[X+Y] \neq \text{Var}[X] + \text{Var}[Y] \)

Ex 1:
Let \( X = \pm 1 \) based on 1 coin flip.
As shown above, \( \mathbb{E}[X] = 0, \text{Var}[X] = 1 \)
Let \( Y = -X \); then \( \text{Var}[Y] = (-1)^2\text{Var}[X] = 1 \)
But \( X+Y = 0 \), always, so \( \text{Var}[X+Y] = 0 \)

Ex 2:
As another example, is \( \text{Var}[X+X] = 2\text{Var}[X] \)?
a zoo of (discrete) random variables
A Bernoulli random variable $X$ is defined in an experiment where there are two possible outcomes: “Success” or “Failure.”

- $X$ is a random **indicator variable** ($1 = \text{success}, \ 0 = \text{failure}$)
- $P(X=1) = p$ and $P(X=0) = 1-p$

$X$ is called a **Bernoulli** random variable: $X \sim \text{Ber}(p)$

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

**Examples:**
- Coin flip
- Random binary digit
- Whether a disk drive crashed

---

Jacob (aka James, Jacques) Bernoulli, 1654 – 1705
Consider $n$ independent random variables $Y_i \sim \text{Ber}(p)$

$X = \sum_i Y_i$ is the number of successes in $n$ trials

$X$ is a \textit{Binomial} random variable: $X \sim \text{Bin}(n,p)$

$$P(X = i) = \binom{n}{i} p^i (1 - p)^{n-i} \quad i = 0, 1, \ldots, n$$

By Binomial theorem, $\sum_{i=0}^{n} P(X = i) = 1$

\textbf{Examples}

- \# of heads in $n$ coin flips
- \# of 1’s in a randomly generated length $n$ bit string
- \# of disk drive crashes in a 1000 computer cluster

$E[X] = pn$

$\text{Var}(X) = p(1-p)n$

\hspace{1cm} \leftrightarrow \text{(proof below, twice)}
PMF for $X \sim \text{Bin}(10, 0.5)$

PMF for $X \sim \text{Bin}(10, 0.25)$
PMF for $X \sim \text{Bin}(30, 0.5)$

PMF for $X \sim \text{Bin}(30, 0.1)$
mean and variance of the binomial

\[ E[X^k] = \sum_{i=0}^{n} i^k \binom{n}{i} p^i (1 - p)^{n-i} \]

\[ = \sum_{i=1}^{n} i^k \binom{n}{i} p^i (1 - p)^{n-i} \]

\[ E[X^k] = np \sum_{i=1}^{n} i^{k-1} \binom{n-1}{i-1} p^{i-1} (1 - p)^{n-i} \]

\[ = np \sum_{j=0}^{n-1} (j + 1)^{k-1} \binom{n-1}{j} p^j (1 - p)^{n-1-j} \]

\[ = np E[(Y + 1)^{k-1}] \]

where \( Y \) is a binomial random variable with parameters \( n - 1, p \).

\( k=1 \) gives: \( E[X] = np \); \( k=2 \) gives \( E[X^2] = np[(n-1)p+1] \)

hence:

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

\[ = np[(n - 1)p + 1] - (np)^2 \]

\[ = np(1 - p) \]
Theorem: If $X$ & $Y$ are *independent*, then $E[X\cdot Y] = E[X] \cdot E[Y]$

Proof:

Let $x_i, y_i, i = 1, 2, \ldots$ be the possible values of $X, Y$.

$$E[X \cdot Y] = \sum_i \sum_j x_i \cdot y_j \cdot P(X = x_i \land Y = y_j)$$

$$= \sum_i \sum_j x_i \cdot y_j \cdot P(X = x_i) \cdot P(Y = y_j)$$

$$= \sum_i x_i \cdot P(X = x_i) \cdot \left( \sum_j y_j \cdot P(Y = y_j) \right)$$

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

Note: *NOT* true in general; see earlier example $E[X^2] \neq E[X]^2$
Theorem: If $X$ & $Y$ are independent, then

$$\text{Var}[X+Y] = \text{Var}[X] + \text{Var}[Y]$$

Proof: Let 

$$\begin{align*}
\hat{X} & = X - E[X] \\
\hat{Y} & = Y - E[Y] \\
E[\hat{X}] & = 0 \\
E[\hat{Y}] & = 0 \\
\text{Var}[\hat{X}] & = \text{Var}[X] \\
\text{Var}[\hat{Y}] & = \text{Var}[Y]
\end{align*}$$

Then

$$\begin{align*}
\text{Var}[X + Y] & = \text{Var}[\hat{X} + \hat{Y}] \\
& = E[(\hat{X} + \hat{Y})^2] - (E[\hat{X} + \hat{Y}])^2 \\
& = E[\hat{X}^2 + 2\hat{X}\hat{Y} + \hat{Y}^2] - 0 \\
& = E[\hat{X}^2] + 2E[\hat{X}\hat{Y}] + E[\hat{Y}^2] \\
& = \text{Var}[\hat{X}] + 0 + \text{Var}[\hat{Y}] \\
& = \text{Var}[X] + \text{Var}[Y]
\end{align*}$$
mean, variance of binomial r.v.s

If $Y_1, Y_2, \ldots, Y_n \sim \text{Ber}(p)$ and independent,

then $X = \sum_{i=1}^{n} Y_i \sim \text{Bin}(n, p)$.

$$E[X] = E[\sum_{i=1}^{n} Y_i] = nE[Y_1] = np$$

$$\text{Var}[X] = \text{Var}[\sum_{i=1}^{n} Y_i] = n\text{Var}[Y_1] = np(1 - p)$$
A RAID-like disk array consists of $n$ drives, each of which will fail independently with probability $p$. Suppose it can operate effectively if at least one-half of its components function, e.g., by “majority vote.” For what values of $p$ is a 5-component system more likely to operate effectively than a 3-component system?

$$X_5 = \text{# failed in 5-component system} \sim \text{Bin}(5, p)$$

$$X_3 = \text{# failed in 3-component system} \sim \text{Bin}(3, p)$$
\( X_5 = \# \text{ failed in 5-component system} \sim \text{Bin}(5, p) \)

\( X_3 = \# \text{ failed in 3-component system} \sim \text{Bin}(3, p) \)

\[
P(\text{5 component system effective}) = P(X_5 < 5/2)
\]

\[
\binom{5}{0} p^0 (1-p)^5 + \binom{5}{1} p^1 (1-p)^4 + \binom{5}{2} p^2 (1-p)^3
\]

\[
P(\text{3 component system effective}) = P(X_3 < 3/2)
\]

\[
\binom{3}{0} p^0 (1-p)^3 + \binom{3}{1} p^1 (1-p)^2
\]

**Calculation:**

5-component system is better iff \( p < 1/2 \)