Variance of a Sum
\[ \text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y) + 2 \text{Cov}(X,Y) \]
\[ \text{Var} \left( \frac{1}{n} \sum_{i=1}^{n} X_i \right) = \frac{1}{n^2} \sum_{i=1}^{n} \text{Var}(X_i) + \frac{2}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \text{Cov}(X_i, X_j) \]
\[ = \frac{1}{n} \text{Var}(X_1) + \frac{2}{n} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \text{Cov}(X_i, X_j) \]

Proof: Exercise (Apply (2) from last class)

Example

$k$ people put their hats in a box, then pick a hat at random. Find \( \text{Var}(X) \), where $X = \#$ of people who pick their own hat.

Define $X_i = 1$ if $i$th person picks their own hat, $0$ otherwise.

$X_i \sim \text{Ber}(p)$, with $p = P(X_i = 1) = \frac{1}{k}$

$E[X_i] = \frac{1}{k}$, $\text{Var}(X_i) = p(1-p) = \frac{1}{k} \left( 1 - \frac{1}{k} \right)$

For $i \neq j$,

$\text{Cov}(X_i, X_j) = E[X_i X_j] - E[X_i] E[X_j]

X_i X_j = \begin{cases} 1 & \text{if } X_i = 1 \text{ and } X_j = 1 \\ 0 & \text{otherwise} \end{cases}$

$E[X_i X_j] = P(X_i X_j = 1) = P(X_i = 1 \text{ and } X_j = 1)$
\[ P(X_i = 1 | X_j = 1) = \frac{1}{k-1} \cdot \frac{1}{k} \]

\[ \text{Cov}(X_i, X_j) = \mathbb{E}[X_i X_j] - \mathbb{E}[X_i] \mathbb{E}[X_j] \]

\[ = \frac{1}{k-1} \cdot \frac{1}{k} - \frac{1}{k} \cdot \frac{1}{k} \]

\[ = -\frac{1}{k^2 (k-1)} \]

\[ \text{Var}(X) = \text{Var} \left( \sum_{i=1}^{k} X_i \right) \]

\[ = \sum_{i=1}^{k} \text{Var}(X_i) + \sum_{i=1}^{k} \sum_{j=1}^{k} \text{Cov}(X_i, X_j) \]

\[ = k \cdot \frac{1}{k} \cdot \left( 1 - \frac{1}{k} \right) + k(k-1) \frac{1}{k^2 (k-1)} \]

\[ = \frac{k-1}{k} + \frac{1}{k} \]

\[ = 1 \]

\[ \text{Var} \left( \sum_{i=1}^{n} X_i \right) = \sum_{i=1}^{n} \text{Var}(X_i) + 2 \sum_{1 \leq i < j \leq n} \text{Cov}(X_i, X_j) \]

**Uncorrelated Variables**

\( X_1, \ldots, X_n \) are called **uncorrelated** if \( \text{Cov}(X_i, X_j) = 0 \) for all \( i, j \) with \( i \neq j \).

**If** \( X_1, \ldots, X_n \) are uncorrelated, **then**

\[ \text{Var} \left( \sum_{i=1}^{n} X_i \right) = \sum_{i=1}^{n} \text{Var}(X_i) \]