

CS70 @ UC Berkeley, Spring 2026

Lecture 21

Random Variables II: Variance and Covariance

April 9, 2026

Get the Largest Number (Lecture 20)



Consider a deck of N cards each with a number written on one side, facing down. **Assume:**

- N is large and all numbers are distinct.
- The deck is well shuffled.

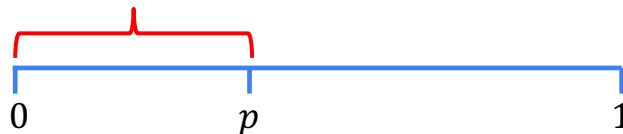
Goal: Get the largest number.

Rules:

1. Reveal one card at a time starting from top.
2. **STOP** at the current card or reveal the next card.
3. If you pass on a card that has been revealed, you can't choose to it later.

Strategy: Reveal a certain proportion, say p , of the cards and record the largest number, denoted M , you have seen. Then, **STOP** when you see a number larger than M .

$M = \text{largest number here}$



Question: What should p be to maximize your chance of winning?

Get the Largest Number (Lecture 20)

Success = get the largest number. $\mathbb{P}(\text{success})?$

Denote the sorted numbers as $X_1 > X_2 > \dots > X_N$

Case 1: $M = X_1 \Rightarrow$ fail

$$\mathbb{P}(M = X_1) = p$$



$$\mathbb{P}(\text{success} | M = X_1) = 0$$

Case 3: $M = X_3 \Rightarrow X_1, X_2$ are in $(p, 1]$

$$\mathbb{P}(M = X_3) \approx p(1-p)^2$$



$$\mathbb{P}(\text{success} | M = X_3) = \mathbb{P}(X_1 \text{ appears before } X_2 \text{ in } (p, 1]) = \frac{1}{2}$$

Case 2: $M = X_2 \Rightarrow X_1$ is in $(p, 1]$

$$\mathbb{P}(M = X_2) \approx p(1-p)$$



$$\mathbb{P}(\text{success} | M = X_2) = 1$$

General Case: $\mathbb{P}(M = X_{k+1}) \approx p(1-p)^k$ Good approximation for $k \ll N$

$$\mathbb{P}(\text{success} | M = X_{k+1}) = \frac{1}{k}$$

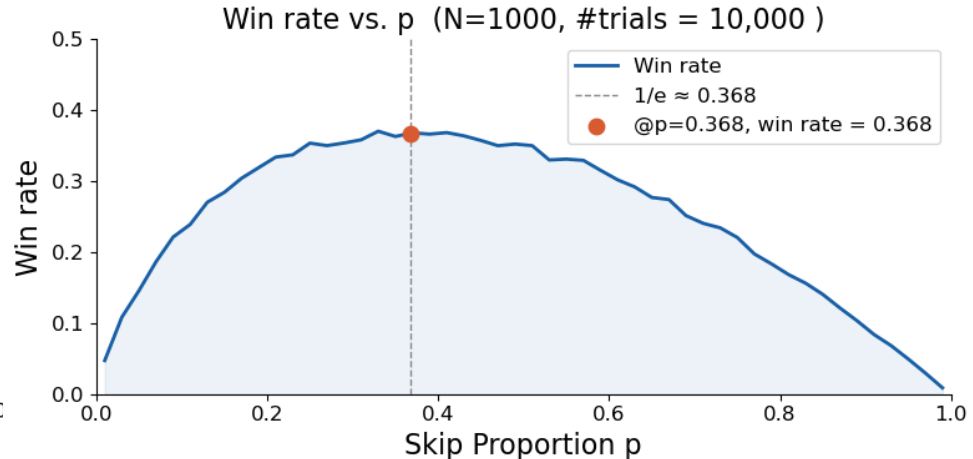
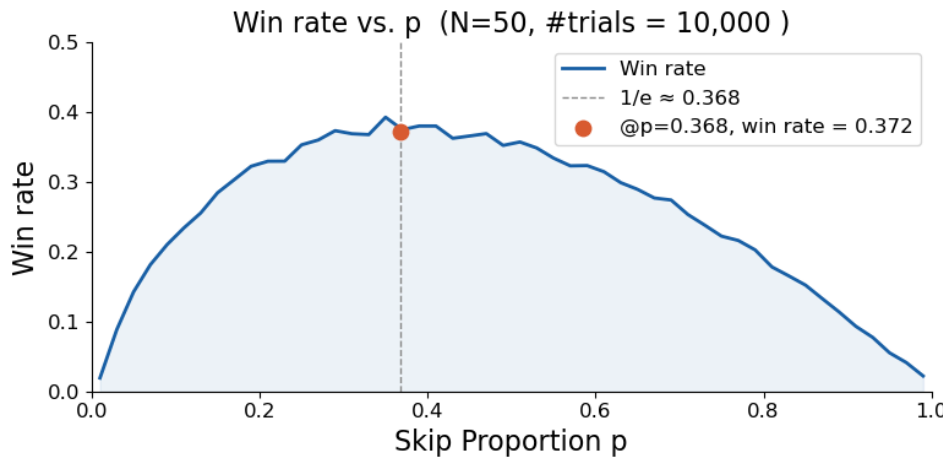
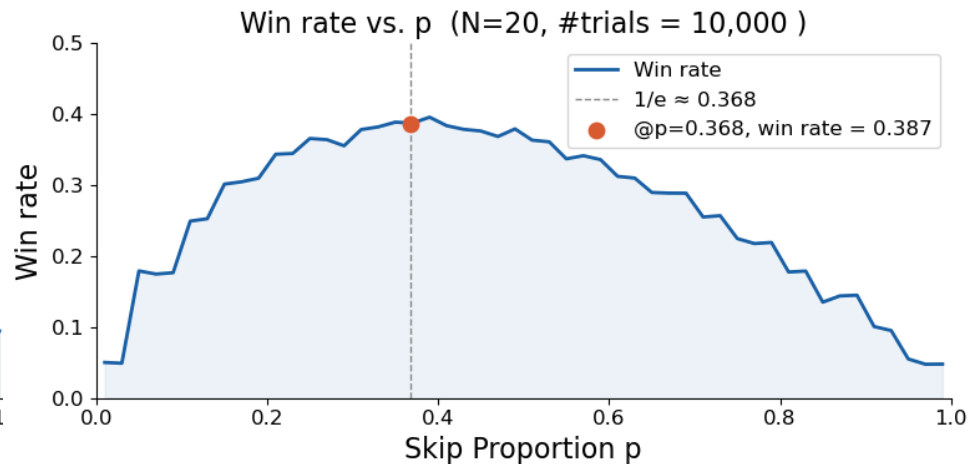
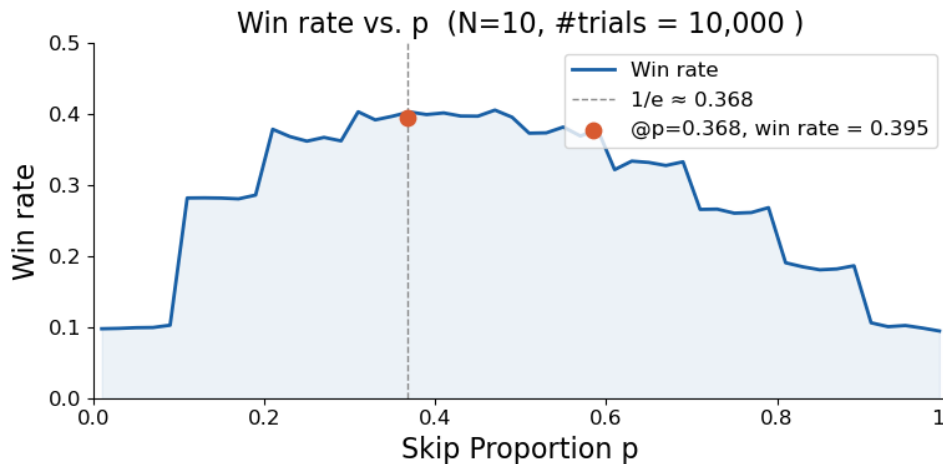
$$\mathbb{P}(\text{success}) = \sum_{i=1}^{(1-p)N} \mathbb{P}(\text{success} | M = X_i) \mathbb{P}(M = X_i)$$

decays exponentially fast as k increases

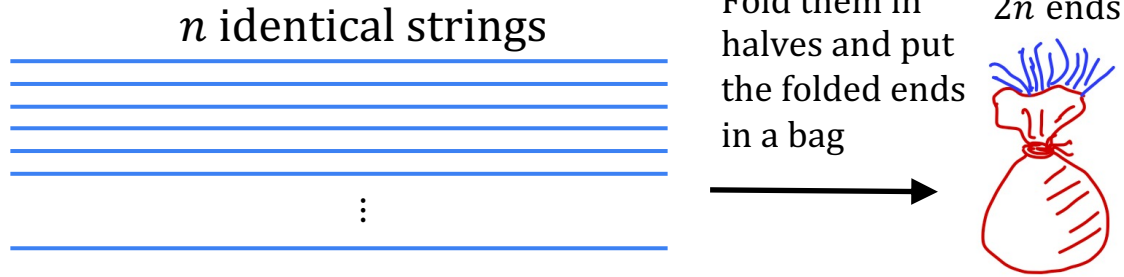
$$\approx \sum_{k=1}^{(1-p)N} \frac{1}{k} p(1-p)^k \approx -p \ln p$$

this is maximized when $p = \frac{1}{e} \approx 0.368^3$

Get the Largest Number (Simulation vs. Theory)



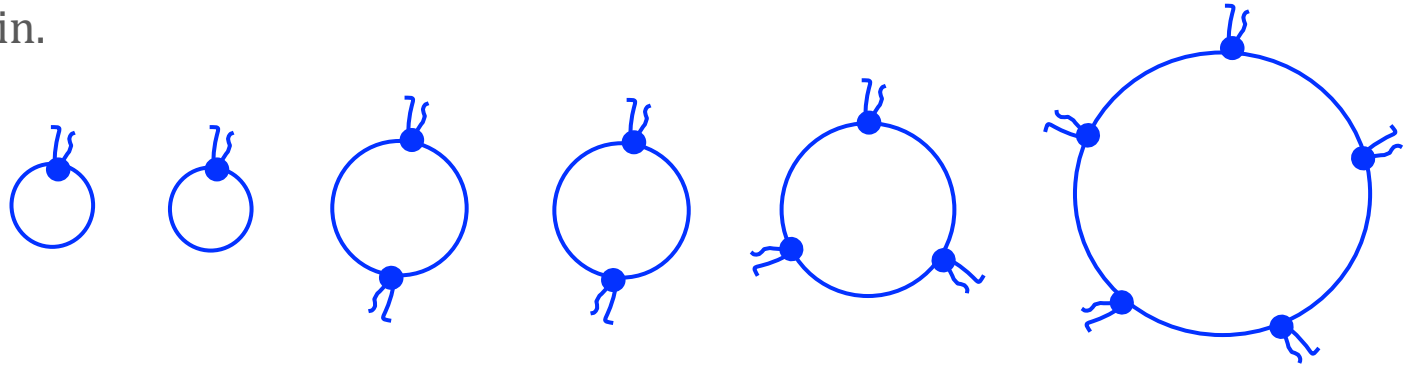
Average Number of Loops



Suppose you tie two free ends uniformly at random and continue until no more free ends remain.

Example outcome for $n = 14$:

6 loops formed



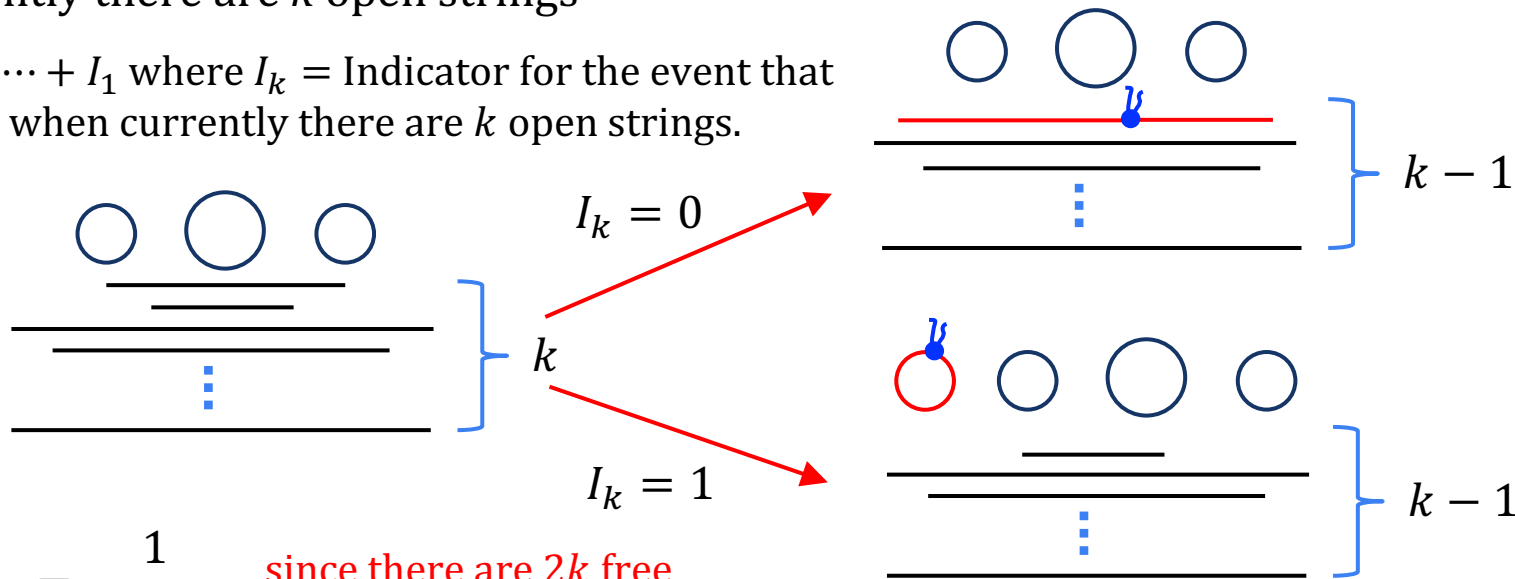
Question: What is the **expected** number of loops formed?

Average Number of Loops

L = Number of Loops formed. $\mathbb{E}[L]$?

Suppose currently there are k open strings

$L = I_n + I_{n-1} + \dots + I_1$ where I_k = Indicator for the event that a loop is formed when currently there are k open strings.



$$\mathbb{P}(I_k = 1) = \frac{k}{\binom{2k}{2}} = \frac{1}{2k-1} \quad \text{since there are } 2k \text{ free ends and } k \text{ open strings}$$

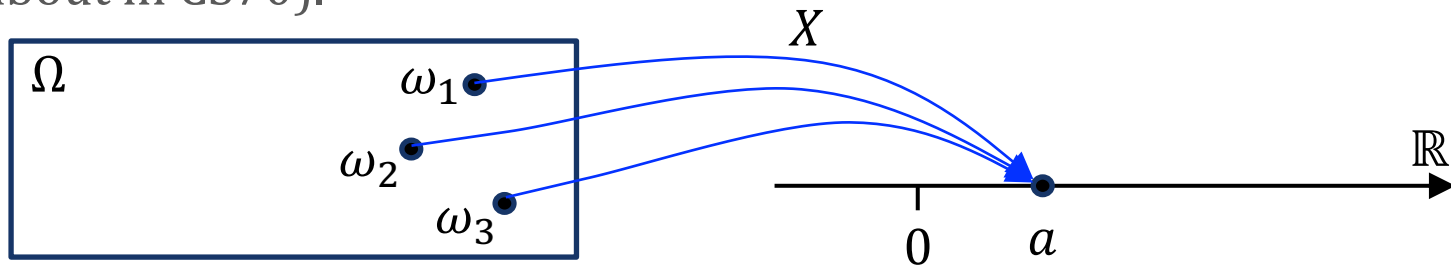
$$\mathbb{E}[L] = \mathbb{E}[I_n + I_{n-1} + \dots + I_1] = \mathbb{E}[I_n] + \mathbb{E}[I_{n-1}] + \dots + \mathbb{E}[I_1] = \sum_{k=1}^n \mathbb{P}[I_k = 1] = 1 + \frac{1}{3} + \dots + \frac{1}{2n-1}$$

which is $\sim \frac{1}{2} \ln(n)$ for large n

Review

- A random variable is a **function** $X: \Omega \rightarrow \mathbb{R}^n$ (we will mostly deal with the $n = 1$ case) satisfying some technical conditions (which you will not need to worry about in CS70).

Probability measure \mathbb{P} is defined on the set of subsets of Ω



- Let $\mathcal{A} = \{a \in \mathbb{R} \mid X(\omega) = a \text{ for some } \omega \in \Omega\}$ denote the **range of X** .
- For a given $a \in \mathcal{A}$, the **event “ $X = a$ ”** is defined as the **pre-image**

$$X^{-1}(a) = \{\omega \in \Omega \mid X(\omega) = a\} \subseteq \Omega.$$

Example: Toss a fair coin twice. $X(\omega) = \#$ number of **Heads** in $\omega \in \Omega$.

$$\text{“}X = 2\text{”} = X^{-1}(2) = \{(H, H)\}$$

$$\text{“}X = 1\text{”} = X^{-1}(1) = \{(H, T), (T, H)\}$$

$$\text{“}X = 0\text{”} = X^{-1}(0) = \{(T, T)\}$$

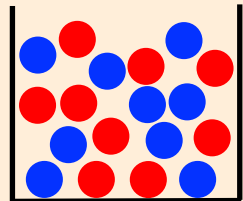
$$\mathbb{P}(X = 2) = \mathbb{P}(X = 0) = \frac{1}{4}, \text{ and } \mathbb{P}(X = 1) = \frac{1}{2}.$$

Review

Sample once. Success probability $p = B/N$.

Example (Bernoulli): $I \sim \text{Bernoulli}(p)$, for $0 < p < 1$.

$$\mathbb{P}(I = a) = \begin{cases} p, & \text{if } a = 1, \\ 1 - p, & \text{if } a = 0. \end{cases}$$



Sample n times **with** replacement.

Example (Binomial): $X \sim \text{Binomial}(n, p)$, where $n \in \mathbb{Z}_+$ and $0 < p < 1$.

$$\text{For } k = 0, \dots, n, \quad \mathbb{P}(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}.$$

Sample n times **without** replacement.

Example (Hypergeometric): $Y \sim \text{Hypergeometric}(N, B, n)$

$$\text{For } \max(0, n + B - N) \leq k \leq \min(n, B), \quad \mathbb{P}(Y = k) = \frac{\binom{B}{k} \binom{N-B}{n-k}}{\binom{N}{n}}.$$

Review



Sample 2 balls at random

X = label of the first ball

Y = label of the second ball

Case 1: Sample with replacement

$X = 1$ $X = 2$ $X = 3$ Row sums

$Y = 1$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{3}$
$Y = 2$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{3}$
$Y = 3$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{3}$

Column sums $\frac{1}{3}$ $\frac{1}{3}$ $\frac{1}{3}$

X, Y are independent and identically distributed (i.i.d.)

Case 2: Sample without replacement

$X = 1$ $X = 2$ $X = 3$ Row sums

$Y = 1$	0	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{3}$
$Y = 2$	$\frac{1}{6}$	0	$\frac{1}{6}$	$\frac{1}{3}$
$Y = 3$	$\frac{1}{6}$	$\frac{1}{6}$	0	$\frac{1}{3}$

Column sums $\frac{1}{3}$ $\frac{1}{3}$ $\frac{1}{3}$

X, Y are **not** independent, but identically distributed ⁹

Review

- Indicator Random Variables
- For $i = 1, \dots, n$, let

$$I_i(\omega) = \begin{cases} 1, & \text{if the } i\text{th trial in } \omega \text{ is success,} \\ 0, & \text{otherwise.} \end{cases}$$

- For $X \sim \text{Binomial}(n, p)$, $X = I_1 + \dots + I_n$, where I_1, \dots, I_n are **mutually independent Bernoulli(p)** random variables.
- For $Y \sim \text{Hypergeometric}(N, B, n)$, $Y = I_1 + \dots + I_n$, where I_1, \dots, I_n are **exchangeable Bernoulli(p)** random variables with $p = B/N$.


Definition (Exchangeability): A collection of random variables X_1, \dots, X_n with common range \mathcal{A} is said to be exchangeable if $(X_{\pi(1)}, \dots, X_{\pi(n)})$ has the same joint distribution as (X_1, \dots, X_n) for every permutation π of $\{1, \dots, n\}$.

Review

Arithmetic average of 1, 1, 1, 3, 3, 4

$$= \frac{1 + 1 + 1 + 3 + 3 + 4}{6} = \frac{1 \times 3 + 3 \times 2 + 4 \times 1}{6} = 1 \times \frac{3}{6} + 3 \times \frac{2}{6} + 4 \times \frac{1}{6}$$

Value Frequency



Definition (Expectation). The expectation of a random variable X with range \mathcal{A} is defined as

$$\mathbb{E}[X] = \sum_{\omega \in \Omega} X(\omega) \mathbb{P}(\omega) = \sum_{a \in \mathcal{A}} a \mathbb{P}(X = a),$$

provided that $\sum_{a \in \mathcal{A}} |a| \mathbb{P}(X = a) < \infty$ (absolutely convergent).

Theorem (Linearity of Expectation): For any two random variables X and Y on the same probability space and fixed **constants** α and β ,

$$\mathbb{E}[\alpha X + \beta Y] = \alpha \mathbb{E}[X] + \beta \mathbb{E}[Y].$$

(Note: this holds irrespective of whether X and Y are independent.)

Examples

Example (Bernoulli): $I \sim \text{Bernoulli}(p)$, for $0 < p < 1$.

$$\mathbb{P}(I = a) = \begin{cases} p, & \text{if } a = 1, \\ 1 - p, & \text{if } a = 0. \end{cases}$$

$$\mathbb{E}[I] = 1 \cdot \mathbb{P}(I = 1) + 0 \cdot \mathbb{P}(I = 0) = p$$

Example (Binomial): $X \sim \text{Binomial}(n, p)$, where $n \in \mathbb{Z}_+$ and $0 < p < 1$.

$$\mathbb{P}(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}, \text{ for } k = 0, \dots, n.$$

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

Cumbersome calculation

$$\mathbb{E}[X] = \mathbb{E}[I_1 + \dots + I_n] = \mathbb{E}[I_1] + \dots + \mathbb{E}[I_n] = np$$

Much easier

Linearity of expectation

Examples

Example (Fixed points, Lecture 15):

Ω = set of all permutations of $\{1, \dots, n\}$.

Suppose $\mathbb{P}[\omega] = \frac{1}{|\Omega|} = \frac{1}{n!}$ for all $\omega \in \Omega$

Let $X(\omega)$ = denote the number of fixed points of $\omega \in \Omega$.

What is $\mathbb{E}[X]$?

$i:$	1	2	3	4	5	...	n
Permutation $\pi_i:$	7	1	3	8	5	...	2

Fixed points

Let $I_i(\omega) = \begin{cases} 1, & \text{if } i \text{ is a fixed point of } \omega \in \Omega, \\ 0, & \text{otherwise.} \end{cases}$

By linearity of expectation

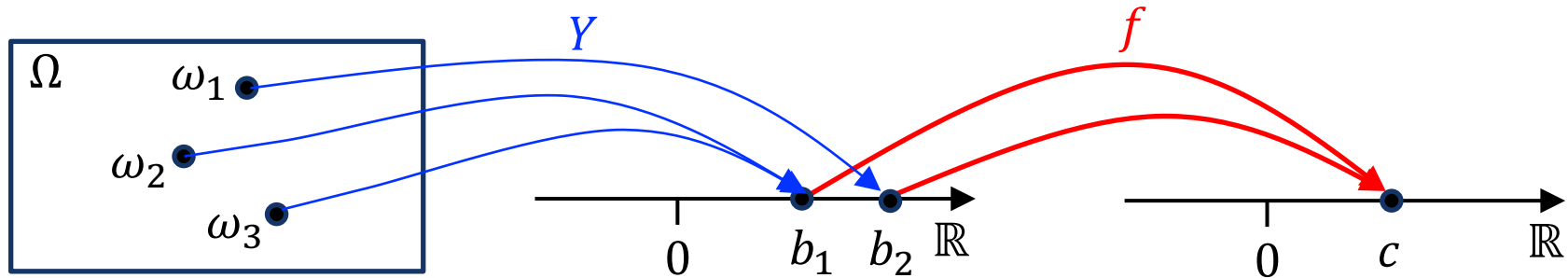
Then, $X = I_1 + \dots + I_n$ and $\mathbb{E}[X] = \mathbb{E}[I_1 + \dots + I_n] = \mathbb{E}[I_1] + \dots + \mathbb{E}[I_n]$

These indicators are NOT independent

$\mathbb{P}[I_i = 1] = \frac{(n-1)!}{n!} = \frac{1}{n}$ and $\mathbb{E}[I_i] = \frac{1}{n}$ for all $i = 1, \dots, n$.

Therefore, $\mathbb{E}[X] = 1$. Does not depend on n .

Functions of Random Variables



- Suppose $Y: \Omega \rightarrow \mathbb{R}$ is a random variable.
- Suppose the range (or image) of Y is $Y(\Omega) = \mathcal{B}$.
- $f: \mathcal{B} \rightarrow \mathbb{R}$, a function satisfying a certain technical condition which we will not need to worry about in CS70.
- Then, $f(Y): \Omega \rightarrow \mathbb{R}$ is also a random variable with range $\mathcal{C} = f(\mathcal{B})$.
- $\mathbb{P}(f(Y) = c) = \sum_{b \in \mathcal{B}: f(b)=c} \mathbb{P}(Y = b)$
- $\mathbb{E}[f(Y)] = \sum_{c \in \mathcal{C}} c \mathbb{P}(f(Y) = c) = \sum_{c \in \mathcal{C}} c \sum_{b \in \mathcal{B}: f(b)=c} \mathbb{P}(Y = b)$
 $= \sum_{c \in \mathcal{C}} \sum_{b \in \mathcal{B}: f(b)=c} f(b) \mathbb{P}(Y = b) = \sum_{b \in \mathcal{B}} f(b) \mathbb{P}(Y = b).$

Conditional Expectation

Let X and Y be random variables defined on the same probability space, taking values in \mathcal{A} and \mathcal{B} , respectively.

Definition (Conditional Expectation): For any $b \in \mathcal{B}$ with $\mathbb{P}[Y = b] > 0$, the conditional expectation of X given $Y = b$ is

$$\mathbb{E}[X | Y = b] = \sum_{a \in \mathcal{A}} a \mathbb{P}(X = a | Y = b).$$

Define $f(Y) = \mathbb{E}[X | Y]$, which takes value $\mathbb{E}[X | Y = b]$ when $Y = b$. Then,

$$\mathbb{E}[f(Y)] = \sum_{b \in \mathcal{B}} f(b) \mathbb{P}[Y = b] = \underbrace{\sum_{b \in \mathcal{B}} \mathbb{E}[X | Y = b] \mathbb{P}[Y = b]}_{\mathbb{E}[X]}$$

Law of Total Expectation

Theorem (Law of Total Expectation): If $\mathbb{E}[|X|] < \infty$,

$$\mathbb{E}[\mathbb{E}[X | Y]] = \mathbb{E}[X].$$

$$\mathbb{E}[\mathbb{E}[X | Y]] = \sum_{b \in \mathcal{B}} \mathbb{E}[X | Y = b] \mathbb{P}[Y = b]$$

$$= \sum_{b \in \mathcal{B}} \sum_{a \in \mathcal{A}} a \mathbb{P}[X = a | Y = b] \mathbb{P}[Y = b]$$

Interchange the order of summation

$$= \sum_{a \in \mathcal{A}} a \sum_{b \in \mathcal{B}} \mathbb{P}[X = a | Y = b] \mathbb{P}[Y = b]$$

By the Total Probability Rule

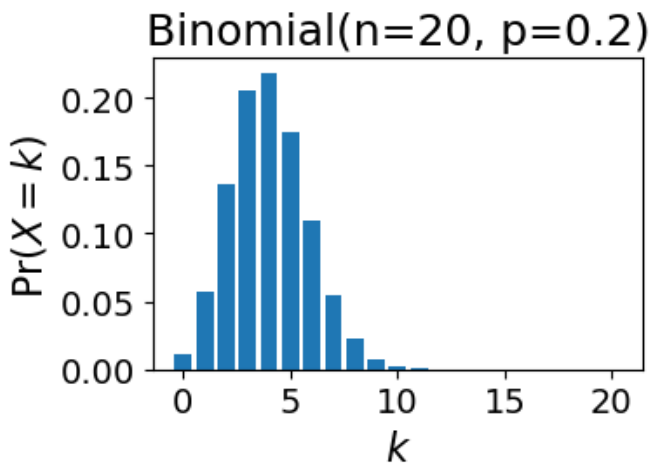
$$= \sum_{a \in \mathcal{A}} a \mathbb{P}[X = a]$$

Q: Why is this a useful result?

$$= \mathbb{E}[X].$$



Variance



200 independent draws from Binomial(20, 0.2):

```
[7 4 4 4 3 8 5 6 3 3 6 3 4 2 4 7 2 4 3 2 3 2 3 3 4
2 5 5 1 5 1 4 4 4 6 6 4 4 2 8 5 7 7 3 6 3 6 4 3 3 3
4 2 5 5 5 1 4 3 3 2 4 2 3 4 3 3 4 1 3 2 4 3 2 3 7 7
7 9 6 6 4 8 5 6 2 2 2 5 4 3 3 7 4 4 3 5 3 2 4 4 2 6
1 6 3 5 3 6 3 3 1 3 1 5 3 3 3 6 6 7 5 1 5 2 3 3 5 4
4 4 6 7 5 4 5 3 1 4 2 2 6 6 3 2 1 2 8 1 4 2 1 3 5 3
4 2 1 2 7 7 2 5 5 2 1 4 1 5 6 4 3 2 5 2 5 5 3 2 6 5
7 6 2 3 5 4 8 5 7 3 4 4 4 5 6 1 6 7 4]
```

Sample average = 3.96, which is quite close to $\mathbb{E}[X] = np = 4$.

Next week, we will prove a precise relationship between sample averages and expectations.

Signed deviations from $\mathbb{E}[X] = 4$:

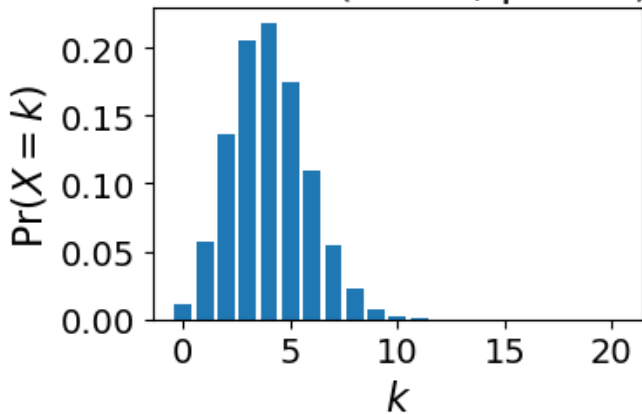
```
[ 3, 0, 0, 0, -1, 4, 1, 2, -1, -1, 2, -1, 0, -2, 0, 3, -2, 0, -1, -2, -1, -2, -1, -1, 0, -2, 1, 1,
-3, 1, -3, 0, 0, 0, 2, 2, 0, 0, -2, 4, 1, 3, 3, -1, 2, -1, 2, 0, -1, -1, -1, 0, -2, 1, 1, 1, -3, 0,
-1, -1, -2, 0, -2, -1, 0, -1, -1, 0, -3, -1, -2, 0, -1, -2, -1, 3, 3, 3, 5, 2, 2, 0, 4, 1, 2, -2,
-2, -2, 1, 0, -1, -1, 3, 0, 0, -1, 1, -1, -2, 0, 0, -2, 2, -3, 2, -1, 1, -1, 2, -1, -1, -3, -1, -3,
1, -1, -1, -1, 2, 2, 3, 1, -3, 1, -2, -1, -1, 1, 0, 0, 0, 2, 3, 1, 0, 1, -1, -3, 0, -2, -2, 2, 2,
-1, -2, -3, -2, 4, -3, 0, -2, -3, -1, 1, -1, 0, -2, -3, -2, 3, 3, -2, 1, 1, -2, -3, 0, -3, 1, 2, 0,
-1, -2, 1, -2, 1, 1, -1, -2, 2, 1, 3, 2, -2, -1, 1, 0, 4, 1, 3, -1, 0, 0, 0, 1, 2, -3, 2, 3, 0]
```

Average signed deviations = -0.04

← Will tend to zero as the sample size approaches ∞

Variance

Binomial($n=20, p=0.2$)



200 independent draws from Binomial(20, 0.2):

```
[7 4 4 4 3 8 5 6 3 3 6 3 4 2 4 7 2 4 3 2 3 2 3 3 4
2 5 5 1 5 1 4 4 4 6 6 4 4 2 8 5 7 7 3 6 3 6 4 3 3 3
4 2 5 5 5 1 4 3 3 2 4 2 3 4 3 3 4 1 3 2 4 3 2 3 7 7
7 9 6 6 4 8 5 6 2 2 2 5 4 3 3 7 4 4 3 5 3 2 4 4 2 6
1 6 3 5 3 6 3 3 1 3 1 5 3 3 3 6 6 7 5 1 5 2 3 3 5 4
4 4 6 7 5 4 5 3 1 4 2 2 6 6 3 2 1 2 8 1 4 2 1 3 5 3
4 2 1 2 7 7 2 5 5 2 1 4 1 5 6 4 3 2 5 2 5 5 3 2 6 5
7 6 2 3 5 4 8 5 7 3 4 4 4 5 6 1 6 7 4]
```

Sample average = 3.96, which is quite close to $\mathbb{E}[X] = np = 4$.

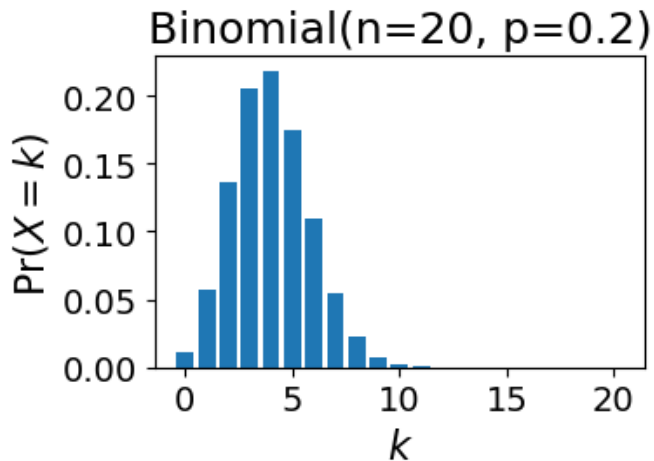
Next week, we will prove a precise relationship between sample averages and expectations

Squared deviations from $\mathbb{E}[X] = 4$:

```
[ 9, 0, 0, 0, 1, 16, 1, 4, 1, 1, 4, 1, 0, 4, 0, 9, 4, 0, 1, 4, 1, 4, 1, 1, 0, 4, 1, 1, 9, 1, 9,
0, 0, 0, 4, 4, 0, 0, 4, 16, 1, 9, 9, 1, 4, 1, 4, 0, 1, 1, 1, 0, 4, 1, 1, 1, 9, 0, 1, 1, 4, 0, 4,
1, 0, 1, 1, 0, 9, 1, 4, 0, 1, 4, 1, 9, 9, 9, 25, 4, 4, 0, 16, 1, 4, 4, 4, 4, 1, 0, 1, 1, 9, 0, 0,
1, 1, 1, 4, 0, 0, 4, 4, 9, 4, 1, 1, 1, 4, 1, 1, 9, 1, 9, 1, 1, 1, 1, 4, 4, 9, 1, 9, 1, 4, 1, 1,
1, 0, 0, 0, 4, 9, 1, 0, 1, 1, 9, 0, 4, 4, 4, 4, 1, 4, 9, 4, 16, 9, 0, 4, 9, 1, 1, 1, 0, 4, 9, 4,
9, 9, 4, 1, 1, 4, 9, 0, 9, 1, 4, 0, 1, 4, 1, 4, 1, 1, 1, 4, 4, 1, 9, 4, 4, 1, 1, 0, 16, 1, 9, 1,
0, 0, 0, 1, 4, 9, 4, 9, 0]
```

Average squared deviations = 3.32

Variance



200 independent draws from Binomial(20, 0.2):

```
[7 4 4 4 3 8 5 6 3 3 6 3 4 2 4 7 2 4 3 2 3 2 3 3 4
2 5 5 1 5 1 4 4 4 6 6 4 4 2 8 5 7 7 3 6 3 6 4 3 3 3
4 2 5 5 5 1 4 3 3 2 4 2 3 4 3 3 4 1 3 2 4 3 2 3 7 7
7 9 6 6 4 8 5 6 2 2 2 5 4 3 3 7 4 4 3 5 3 2 4 4 2 6
1 6 3 5 3 6 3 3 1 3 1 5 3 3 3 6 6 7 5 1 5 2 3 3 5 4
4 4 6 7 5 4 5 3 1 4 2 2 6 6 3 2 1 2 8 1 4 2 1 3 5 3
4 2 1 2 7 7 2 5 5 2 1 4 1 5 6 4 3 2 5 2 5 5 3 2 6 5
7 6 2 3 5 4 8 5 7 3 4 4 4 5 6 1 6 7 4]
```

Sample average = 3.96, which is quite close to $\mathbb{E}[X] = np = 4$.

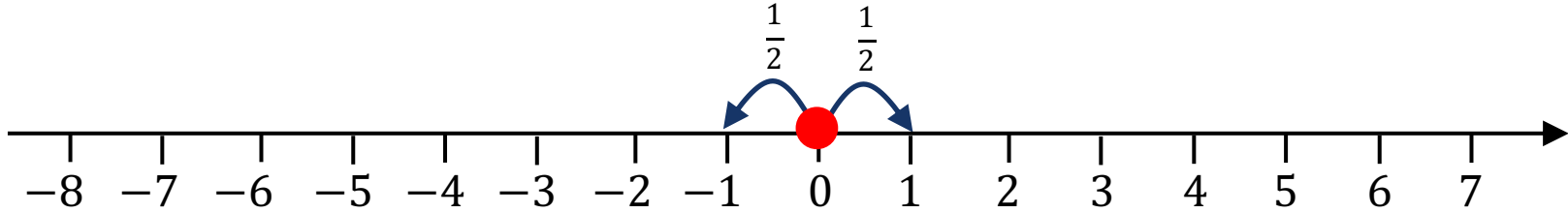
Next week, we will prove a precise relationship between sample average and $\mathbb{E}[X]$.

As the sample size increases, the average squared deviations will approach:

Definition (Variance): Suppose X is a random variable with $\mathbb{E}[X] = \mu$. Then, $\text{Var}(X) = \mathbb{E}[(X - \mu)^2] \geq 0$. (Recall $\mathbb{E}[f(X)]$ and let $f(X) = (X - \mu)^2$.)

The **standard deviation** of X is defined as $\sigma(X) = \sqrt{\text{Var}(X)}$.

Symmetric Random Walk

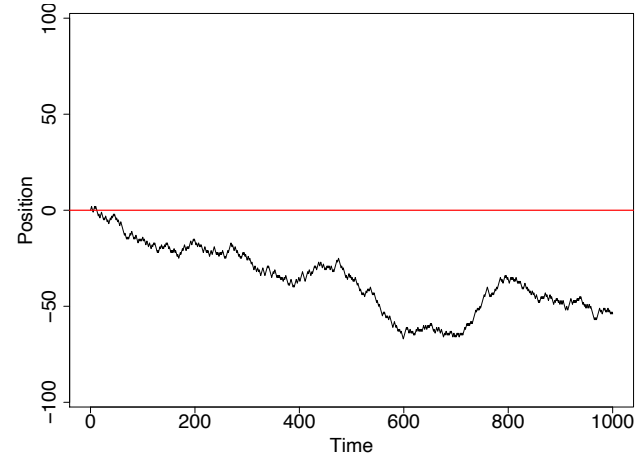
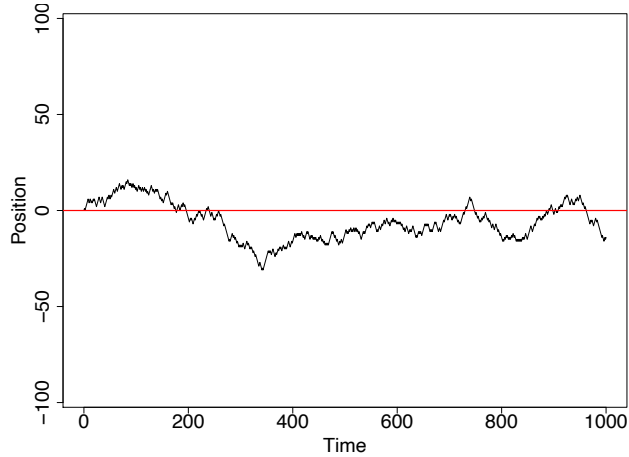
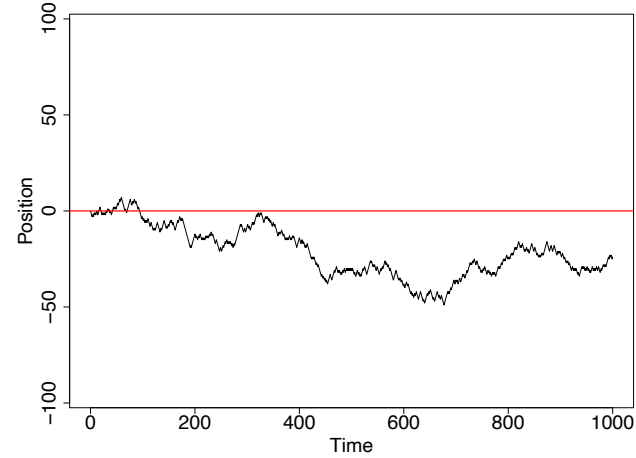
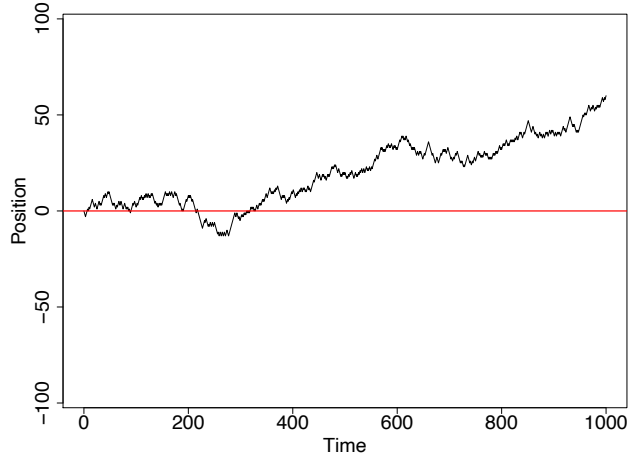


- Start at the origin.
- For each time step, walk to the left with probability $\frac{1}{2}$ or to the right with probability $\frac{1}{2}$, independently of all previous steps.
- Position after n steps: $S_n = X_1 + \dots + X_n$, where X_1, \dots, X_n are i.i.d. random variables with

$$X_i = \begin{cases} +1, & \text{with probability } 1/2 \\ -1, & \text{with probability } 1/2 \end{cases}$$

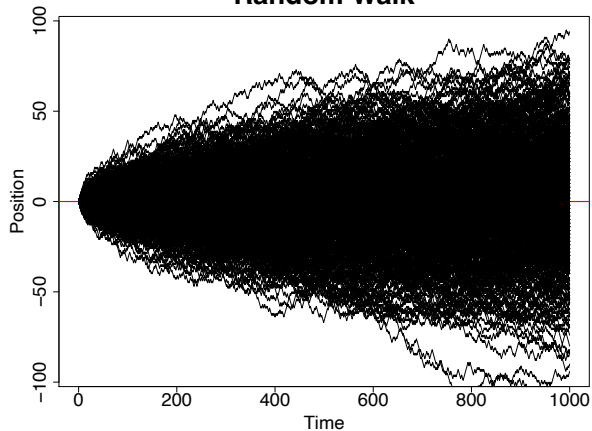
- What is $\mathbb{E}[S_n]$?
- How about $\text{Var}[S_n]$?

Random Walk (Simulations)

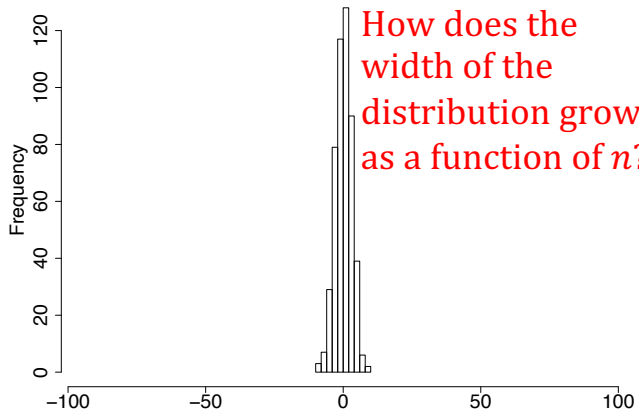


Random Walk (Distributions of S_n)

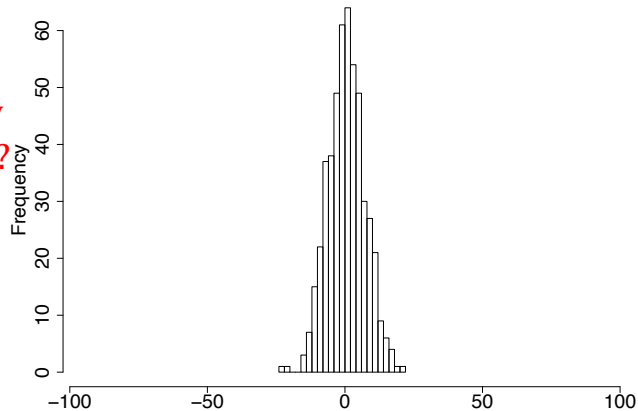
Random Walk



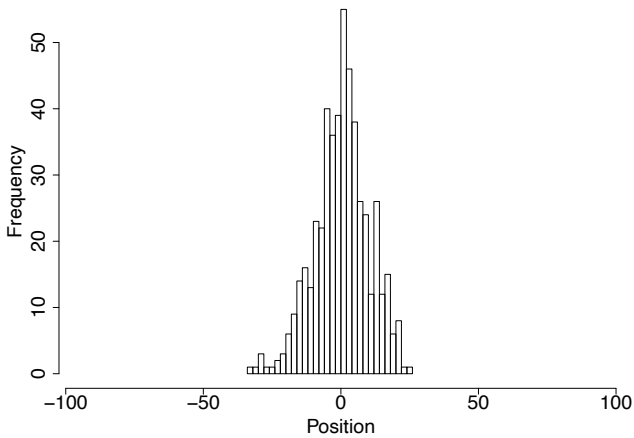
Histogram of S_n for $n = 10$



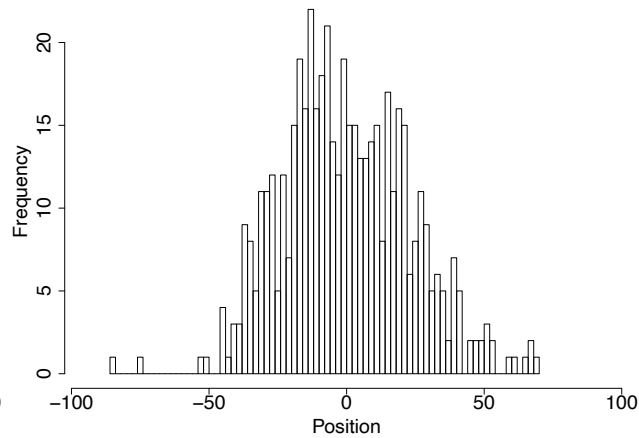
Histogram of S_n for $n = 50$



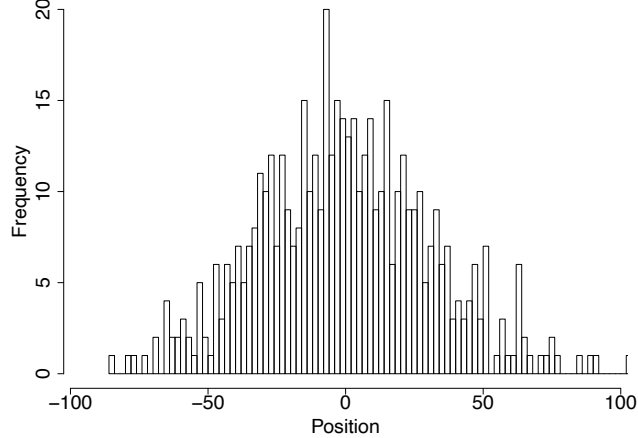
Histogram of S_n for $n = 100$



Histogram of S_n for $n = 500$



Histogram of S_n for $n = 1000$



Expectation of XY

Lemma 1: If X and Y are **independent** random variables on the same probability space, then $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$.

Proof:
$$\mathbb{E}[XY] = \sum_{a \in \mathcal{A}} \sum_{b \in \mathcal{B}} ab \mathbb{P}(X = a, Y = b)$$

$$= \sum_{a \in \mathcal{A}} \sum_{b \in \mathcal{B}} ab \mathbb{P}(X = a)\mathbb{P}(Y = b) = \left[\sum_{a \in \mathcal{A}} a \mathbb{P}(X = a) \right] \left[\sum_{b \in \mathcal{B}} b \mathbb{P}(Y = b) \right] = \mathbb{E}[X]\mathbb{E}[Y].$$
 □

Is the converse true? **No!**

(a, b)	$\mathbb{P}(X = a, Y = b)$
$(-1, 0)$	$1/3$
$(0, 1)$	$1/3$
$(1, 0)$	$1/3$

a	$\mathbb{P}(X = a)$
-1	$1/3$
0	$1/3$
1	$1/3$

b	$\mathbb{P}(Y = b)$
0	$2/3$
1	$1/3$

However,

$$\mathbb{E}[XY] = 0$$

$$\mathbb{E}[X] = 0$$

$$\mathbb{E}[Y] = 1/3$$

$$\left. \begin{array}{l} \mathbb{E}[XY] = 0 \\ \mathbb{E}[X] = 0 \\ \mathbb{E}[Y] = 1/3 \end{array} \right\} \Rightarrow \mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$$

0

1/3

1/3

Are X and Y independent? **No!** $\mathbb{P}(X = 1, Y = 1) \neq \mathbb{P}(X = 1)\mathbb{P}(Y = 1)$

Variance of $X + Y$

Theorem: If X and Y are **independent** random variables on the same probability space, then $\text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y]$.

Proof. Using **linearity of expectation**, we get

$$\begin{aligned}\mathbb{E}[(X + Y)^2] - [\mathbb{E}[X + Y]]^2 &= \mathbb{E}[X^2 + 2XY + Y^2] - [\mathbb{E}[X] + \mathbb{E}[Y]]^2 \\ &= \underbrace{\mathbb{E}[X^2] - [\mathbb{E}[X]]^2}_{\text{Var}(X)} + \underbrace{\mathbb{E}[Y^2] - [\mathbb{E}[Y]]^2}_{\text{Var}(Y)} + \underbrace{2[\mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]]}_{\text{by Lemma 1}}\end{aligned}$$

□

Back to Symmetric Random Walk:

$$\text{Var}(S_n) = \sum_{i=1}^n \text{Var}(X_i) = n \Rightarrow \sigma(S_n) = \sqrt{n}.$$

$S_n = X_1 + \dots + X_n$, where X_1, \dots, X_n are i.i.d. random variables with

$$X_i = \begin{cases} +1, & \text{with probability } 1/2 \\ -1, & \text{with probability } 1/2 \end{cases}$$

$$\text{Var}(X_i) = \mathbb{E}[X_i^2] - (\mathbb{E}[X_i])^2 = (1)^2\mathbb{P}(X_i = 1) + (-1)^2\mathbb{P}(X_i = -1) = 1$$

Variance Examples

Example (Bernoulli): $I \sim \text{Bernoulli}(p)$, for $0 < p < 1$.

$$\mathbb{P}(I = a) = \begin{cases} p, & \text{if } a = 1, \\ 1 - p, & \text{if } a = 0. \end{cases}$$

$$\mathbb{E}[I] = p$$

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

Example (Binomial): $X \sim \text{Binomial}(n, p)$, where $n \in \mathbb{Z}_+$ and $0 < p < 1$.

$$\text{For } k = 0, \dots, n, \quad \mathbb{P}(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}.$$

$X = I_1 + \dots + I_n$, where I_1, \dots, I_n are **mutually independent Bernoulli(p)** random variables.

$$\text{Var}(X) = \sum_{k=1}^n \text{Var}(I_k) = np(1 - p).$$

Covariance and Correlation

Definition (Covariance): The covariance of random variables X and Y is defined as

$$\text{Cov}(X, Y) = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y].$$

Definition (Correlation): The correlation of random variables X and Y with $\sigma(X) > 0$ and $\sigma(Y) > 0$ is defined as

$$\text{Corr}(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma(X)\sigma(Y)}.$$

Theorem: For a pair of random variables X, Y with $\sigma(X) > 0$ and $\sigma(Y) > 0$,
 $-1 \leq \text{Corr}(X, Y) \leq 1$.

See Note 19 for proof.