# STATISTICS WITH APPLIED PROBABILITY

## Custom eBook for STA258

Nishan Mudalige
Nurlana Alili
Bryan Su

CANADA

# Statistics with Applied Probability
# Custom eBook for STA258

Nishan Mudalige
*Department of Mathematical and Computational Sciences*
*University of Toronto Mississauga*

Nurlana Alili
*University of Toronto Mississauga*

Bryan Xu
*University of Toronto Mississauga*

First edition: August 2025

Mudalige, M.; Alili, N.; Xu, B.
University of Guelph Bookstore Press

University of Toronto Mississauga,
Mississauga, Ontario, Canada

# Contents

# Chapter 0

# Overview

Uncertainty is an inherent part of everyday life. We all face questions regarding uncertainty such as whether classes will go ahead as planned on any given day; will a flight leave on time; will a student pass a certain course? Uncertainties might also change depending on other factors, such as whether classes will still go ahead as planned when there is a snow warning in effect; if a flight is delayed can a person still manage to make their connection; will a student pass their course considering that the instructor is known to be a tough grader?

The ability to quantify uncertainty using rigorous mathematics is a powerful and useful tool. Calculating uncertainty on an intuitive level is something that is hard-wired in our DNA, such as the decision to fight or flight depending on a given set of circumstances. However we cannot always make such intuitive decisions based purely on hunches and gut feelings. Fortunes have been lost based on someone having a good feeling about something. If we have information available, we should make the best prediction possible using this information. For instance if we wanted to invest a lot of money in a company, we should use all available data such as past sales, market and industry trends, leadership ability of the CEO, forward looking statements etc. and with all this information we can then predict whether our investment will be profitable.

In order for companies to survive and remain competitive in todays environment it is essential to monitor industry trends and read markets properly. Companies that don't adapt and stick to an outdated business model tend to pay the price. At the other end of the spectrum, companies that understand the needs of the consumer, build their product around the consumer and keep evolving their product offerings based on consumer trends tend to perform well and remain competitive.

Statistics is the science of uncertainty and it is clearly a very useful subject for business. In this book you will be given an introduction to statistics and you will learn the framework as well as the language required at the introductory level. The material may be daunting at times, but the more you get familiar with the subject the more comfortable you will become with it. As business students, doing well in a statistics course will give you a competitive edge since the ability to interpret and perform quantitative analytics are skills that are highly desired by many employers.

# Chapter 1

# Introduction

## 1.1 Basics

Intuitively, statistics can be considered the science of uncertainty. Formally,

**Definition 1.1** (Statistics).
*Statistics is the science of collecting, classifying, summarizing, analyzing and interpreting data.*

more information goes here
anything

# Chapter 4

# Normal Approximation to the Binomial Distribution

## 4.1 Definitions and Setup

**Definition 4.1.1 (Statistic)**

A *statistic* is a function of the observable random variables in a sample and known constants.
Since statistics are functions of the random variables observed in a sample, they themselves are random variables. As such, all statistics have a corresponding probability distribution, which we refer to as their *sampling distribution*.

**Review from STA256**

**Bernoulli Distribution:**
A Bernoulli trial is a single experiment with two outcomes:

- Success: $X = 1$ with probability $p$
- Failure: $X = 0$ with probability $1 - p$

| $X = x$ | $0$ | $1$ |
|---|---|---|
| $P(X = x)$ | $1 - p$ | $p$ |

The probability mass function (PMF) is:

$$f(x) = p^x(1-p)^{1-x}, \quad x \in \{0, 1\}$$

**Binomial Distribution:**
A binomial distribution arises from $n$ independent Bernoulli trials. Let:

$$X = \text{number of successes in } n \text{ trials}$$

Then:

$$X \sim \text{Binomial}(n, p)$$

where:

- Each trial results in either success (with probability $p$) or failure (with probability $1-p$)
- $X \in \{0, 1, \ldots, n\}$

The PMF is:

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$$

**Moment Generating Function (MGF):**
The moment generating function (MGF) of a random variable $X$ is defined as:

$$M_X(t) = \mathbb{E}[e^{tX}]$$

The MGF uniquely characterizes the distribution of $X$ (if it exists in an open interval around 0), and it can be used to compute moments such as the mean and variance.

## 4.2 Bernoulli Distribution

Consider the random experiment of rolling a die once. Define the random variable:

$$X_i = \begin{cases} 1 & \text{if the } i\text{-th roll is a six,} \\ 0 & \text{otherwise} \end{cases}$$

Then $X_i \sim \text{Bernoulli}(p)$, where $p = P(\text{rolling a six})$.
The expected value and variance of a Bernoulli random variable are:

$$\mu = \mathbb{E}(X) = p, \quad \sigma^2 = \text{Var}(X) = p(1-p)$$

## 4.3 Sampling Distribution of the Sum and MGF Derivation

Consider determining the sampling distribution of the sample total:

$$T_n = X_1 + X_2 + \cdots + X_n$$

Suppose $X_i \overset{iid}{\sim} \text{Bernoulli}(p)$. Then the moment-generating function of $T_n$ is:

$$\begin{aligned} M_{T_n}(t) &= \mathbb{E}[e^{tT_n}] \\ &= \mathbb{E}\left[e^{t(X_1+X_2+\cdots+X_n)}\right] \\ &= \mathbb{E}\left[e^{tX_1}e^{tX_2}\ldots e^{tX_n}\right] \quad \text{(independence)} \\ &= \mathbb{E}[e^{tX_1}] \cdot \mathbb{E}[e^{tX_2}] \cdots \mathbb{E}[e^{tX_n}] \\ &= M_{X_1}(t) \cdot M_{X_2}(t) \cdots M_{X_n}(t) \\ &= \left[pe^t + (1-p)\right]^n \end{aligned}$$

Since this is the MGF of a binomial random variable with parameters $n$ and $p$, we conclude:

$$T_n \sim \text{Binomial}(n, p)$$

**Example: Binomial Distribution from Die Rolls**

We can think of rolling a die $n$ times as an example of the binomial setting. Each roll gives either a six (a "success") or a number different from six (a "failure").
Knowing the outcome of one roll doesn't tell us anything about the others, so the $n$ rolls are independent.
If we call a six a success, then:

- The probability of success on each trial is $p = P(\text{rolling a six}) = \frac{1}{6}$
- The probability of failure is $1 - p = \frac{5}{6}$

Let $Y$ be the number of sixes rolled in $n$ trials. Then $Y \sim \text{Binomial}(n, p)$, and the distribution of $Y$ is called a **binomial distribution**.

## 4.4 Binomial Distribution Summary

A random variable $Y$ is said to have a **binomial distribution** based on $n$ trials with success probability $p$ if and only if:

$$p(y) = \frac{n!}{y!(n-y)!} p^y (1-p)^{n-y}, \quad y = 0, 1, 2, \ldots, n \quad \text{and} \quad 0 \le p \le 1$$

The expected value and variance of a binomial random variable are:

$$\mu = \mathbb{E}(Y) = np, \quad \sigma^2 = \text{Var}(Y) = np(1-p)$$

### 4.4.1 Visualizing the PMF of Binomial Distributions

**R Output:**

*Image will be inserted here for Slide 11*

**Probability Mass Functions (PMFs) for increasing $n$:**

The following plots display the probability mass functions (PMFs) for a binomial distribution with $p = \frac{1}{6}$ and increasing values of $n$. As $n$ increases, the binomial distribution begins to resemble a normal distribution.

- Slide 12: Image will be inserted here
- Slide 13: Image will be inserted here
- Slide 14: Image will be inserted here
- Slide 15: Image will be inserted here

- Slide 16: Image will be inserted here
- Slide 17: Image will be inserted here
- Slide 18: Image will be inserted here
- Slide 19: Image will be inserted here
- Slide 20: Image will be inserted here
- Slide 21: Image will be inserted here
- Slide 22: Image will be inserted here
- Slide 23: Image will be inserted here

## 4.5 Sampling Distribution of a Sample Proportion and the Normal Approximation

Draw a *Simple Random Sample (SRS)* of size $n$ from a large population that contains proportion $p$ of "successes". Let $\hat{p}$ be the ***sample proportion*** of successes:

$$\hat{p} = \frac{\text{number of successes in the sample}}{n}$$

Then:

- The **mean** of the sampling distribution of $\hat{p}$ is $p$.
- The **standard deviation** of the sampling distribution is $\sqrt{\frac{p(1-p)}{n}}$.

*Image will be inserted here (Slide 24)*

According to the Central Limit Theorem (CLT), the sampling distribution of a sample proportion becomes approximately normal as the sample size increases.
That is:

$$\hat{p} \sim \mathcal{N}\left(p, \sqrt{\frac{p(1-p)}{n}}\right)$$

This approximation is most accurate when both $np \geq 10$ and $n(1-p) \geq 10$.
These are called the **success-failure conditions**.

*Key Point:* When the success-failure conditions are met, the normal approximation to the sampling distribution of $\hat{p}$ can be used for probability calculations.

### Conditions for Using the Normal Approximation

Suppose $X \sim \text{Binomial}(n, p)$. Then:

$$\mu = np, \quad \sigma^2 = np(1-p)$$

**Binomial probabilities can be approximated by the normal distribution:**

$$X \approx \mathcal{N}(np,\, np(1-p))$$

This approximation is *useful for large* $n$ and valid under the following conditions:

**Standard Conditions**

The binomial setting holds (i.e., independent trials, fixed $n$, same probability $p$) and

$$np \geq 10 \quad \text{and} \quad np(1-p) \geq 10$$

Alternatively, a more conservative criterion for using the normal approximation is:

$$n > 9 \cdot \left( \frac{\max(p,\, 1-p)}{\min(p,\, 1-p)} \right)$$

These ensure that the binomial distribution is sufficiently symmetric and smooth to approximate with the normal distribution.

## 4.6 Sampling Distribution of $\hat{p}$

**Bernoulli Distribution (Binomial with $n = 1$)**

$$X_i = \begin{cases} 1 & \text{if the } i\text{-th roll is a six} \\ 0 & \text{otherwise} \end{cases}$$

$$\mu = \mathbb{E}(X_i) = p, \quad \sigma^2 = \text{Var}(X_i) = p(1-p)$$

Let $\hat{p}$ be our estimate of $p$. Note that $\hat{p} = \frac{1}{n} \sum_{i=1}^{n} X_i = \bar{X}$. Let $\hat{p} = \frac{\#\text{ successes } (X)}{\text{sample size } (n)}$
Recall that for $X \sim \text{Binomial}(n, p)$:

$$X \dot{\sim} \mathcal{N}(np, np(1-p))$$

Let $\hat{p} = \frac{X}{n}$

**Mean of $\hat{p}$:**

$$\mathbb{E}(\hat{p}) = \mathbb{E}\left(\frac{X}{n}\right) = \frac{1}{n} \cdot \mathbb{E}(X) = \frac{1}{n} \cdot np = p$$

**Variance of $\hat{p}$:**

$$\text{Var}(\hat{p}) = \text{Var}\left(\frac{X}{n}\right) = \frac{1}{n^2} \cdot \text{Var}(X) = \frac{1}{n^2} \cdot np(1-p) = \frac{p(1-p)}{n}$$

By the Central Limit Theorem (CLT), for sufficiently large $n$:

$$\hat{p} \sim \mathcal{N}\left(p, \frac{p(1-p)}{n}\right)$$

**Standardization of $\hat{p}$:**

$$Z = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}}$$

If $n$ is large, then by the Central Limit Theorem:

$$\bar{X} \approx \mathcal{N}\left(\mu, \frac{\sigma}{\sqrt{n}}\right) \quad \Rightarrow \quad \hat{p} \sim \mathcal{N}\left(p, \sqrt{\frac{p(1-p)}{n}}\right)$$

**Example: Normal Approximation for Proportions**

**Problem:**

In the last election, a state representative received 52% of the votes cast. One year after the election, the representative organized a survey that asked a random sample of 300 people whether they would vote for him in the next election. If we assume that his popularity has not changed, what is the probability that more than half the sample would vote for him?

**Solution 1 (using Normal Approximation)**

We want to determine the probability that the sample proportion is greater than 50%. In other words, we want to find $P(\hat{p} > 0.50)$.

We know that the sample proportion $\hat{p}$ is roughly Normally distributed with mean $p = 0.52$ and standard deviation

$$\sqrt{p(1-p)/n} = \sqrt{(0.52)(0.48)/300} = 0.0288.$$

Thus, we calculate

$$P(\hat{p} > 0.50) = P\left(\frac{\hat{p} - p}{\sqrt{p(1-p)/n}} > \frac{0.50 - 0.52}{0.0288}\right)$$

$$= P(Z > -0.69) = 1 - P(Z < -0.69) \quad \text{(Z is symmetric)}$$

$$= P(Z > -0.69) = 1 - P(Z > 0.69)$$

$$= 1 - 0.2451 = 0.7549.$$

If we assume that the level of support remains at 52%, the probability that more than half the sample of 300 people would vote for the representative is 0.7549.

**R code (Normal approximation)** Just type in the following:

```
1 - pnorm(0.50, mean = 0.52, sd = 0.0288)
## [1] 0.7562982
```

Recall that, `pnorm` will give you the area to the left of 0.50, for a Normal distribution with mean 0.52 and standard deviation 0.0288.

**Solution 2 (using Binomial)**

We want to determine the probability that the sample proportion is greater than 50%. In other words, we want to find $P(\hat{p} > 0.50)$. We know that $n = 300$ and $p = 0.52$.
Thus, we calculate

$$\begin{aligned}
P(\hat{p} > 0.50) &= P\left(\frac{\sum_{i=1}^{n} x_i}{n} > 0.50\right) \\
&= P\left(\sum_{i=1}^{300} x_i > 150\right) \\
&= 1 - P\left(\sum_{i=1}^{300} x_i \leq 150\right) \\
&\text{(it can be shown that } Y = \sum_{i=1}^{300} x_i \text{ has a Binomial distribution with} \\
&n = 300 \text{ and } p = 0.52) \\
&= 1 - F_Y(150)
\end{aligned}$$

**R code (using Binomial distribution )** Just type in the following:

```
1- pbinom(150, size = 300, prob = 0.52);
## [1] 0.7375949
```

Recall that, `pbinom` will give you the CDF at 150, for a Binomial distribution with $n = 300$ and $p = 0.52$.

**Solution 3 (using continuity correction)**

We have that $n = 300$ and $p = 0.52$. Thus, we calculate

$$\begin{aligned}
P(\hat{p} > 0.50) &= P\left(\frac{\sum_{i=1}^{n} x_i}{n} > 0.50\right) \\
&= P\left(\sum_{i=1}^{300} x_i > 150\right) \\
&= 1 - P\left(\sum_{i=1}^{300} x_i \leq 150\right) \\
&\text{(it can be shown that } Y = \sum_{i=1}^{300} x_i \text{ has a Binomial distribution with} \\
&n = 300 \text{ and } p = 0.52). \\
&\approx 1 - P\left(\sum_{i=1}^{300} x_i \leq 150.5\right) \quad \text{(continuity correction)}
\end{aligned}$$

$$
\begin{aligned}
&= 1 - P\left(\frac{\sum_{i=1}^{300} x_i}{n} \le \frac{150.5}{300}\right) \\
&= 1 - P(\hat{p} \le 0.5017) \\
&= 1 - P\left(Z \le -0.6354\right) \quad \text{(Why?)}
\end{aligned}
$$

**R code (Normal approximation with continuity correction)** Just type in the following:

```
1 - pnorm(0.5017, mean = 0.52, sd = 0.0288)
## [1] 0.7374216
```

Recall that, `pnorm` will give you the area to the left of 0.5017, for a Normal distribution with mean 0.52 and standard deviation 0.0288.

## 4.7 Continuity Correction

Suppose that $Y$ has a Binomial distribution with $n = 20$ and $p = 0.4$. We will find the exact probabilities that $Y \le y$ and compare these to the corresponding values found by using two Normal approximations. One of them, when $X$ is Normally distributed with $\mu_X = np$ and $\sigma_X = \sqrt{np(1-p)}$. The other one, $W$, a shifted version of $X$.

For example,

$$P(Y \le 8) = 0.5955987$$

As previously stated, we can think of $Y$ as having approximately the same distribution as $X$.

$$P(Y \le 8) \approx P(X \le 8) = P\left[\frac{X - np}{\sqrt{np(1-p)}} \le \frac{8-8}{\sqrt{20(0.4)(0.6)}}\right] = P(Z \le 0) = 0.5$$

$$P(Y \le 8) \approx P(W \le 8.5) = P\left[\frac{W - np}{\sqrt{np(1-p)}} \le \frac{8.5-8}{\sqrt{20(0.4)(0.6)}}\right] = P(Z \le 0.2282) = 0.5902615$$

**Example: Normal Approximation with Continuity Correction (Exact Probability)**

Fifty-one percent of adults in the U. S. whose New Year's resolution was to exercise more achieved their resolution. You randomly select 65 adults in the U. S. whose resolution was to exercise more and ask each if he or she achieved that resolution. What is the probability that exactly forty of them respond yes?

We are given that $p = 0.51$, $n = 65$, and we want to find $P(X = 40)$ where $X \sim Binomial(n = 65, p = 0.51)$.

**Step 1: Use Normal Approximation** We use normal approximation to the binomial. First, compute the mean and standard deviation:

$\mu = np = 65\times 0.51 = 33.15, \qquad \sigma^2 = np(1-p) = 65\times 0.51\times 0.49 = 16.485, \quad \sigma = \sqrt{16.485} \approx 4.06$

We apply continuity correction:

$$P(X = 40) = P(39.5 \leq X \leq 40.5)$$

$$= P\left(\frac{39.5 - 33.15}{4.06} \leq Z \leq \frac{40.5 - 33.15}{4.06}\right) = P(1.56 \leq Z \leq 1.81)$$

From the standard normal table:

$$= P(Z \leq 1.81) - P(Z \leq 1.56) = 0.0594 - 0.0352 = 0.0242$$

So the approximate probability is:

$$P(X = 40) \approx 0.0242$$

Normal Approximation to Binomial

Let $X = \sum_{i=1}^{n} Y_i$ where $Y_1, Y_2, \ldots, Y_n$ are iid Bernoulli random variables. Note that $X = n\hat{p}$.

1. $n\hat{p}$ is approximately Normally distributed provided that $np \geq 10$ and $n(1-p) \geq 10$.
2. Another criterion is that the Normal approximation is adequate if
$$n > 9\left(\frac{\text{larger of } p \text{ and } q}{\text{smaller of } p \text{ and } q}\right)$$
3. The expected value: $E(\hat{p}) = np$.
4. The variance: $V(\hat{p}) = np(1-p) = npq$.

# Chapter 5

# Law of Large Numbers

## 5.1 Convergence in Probability

**Definition 5.1.1 (Convergence in Probability)**

The sequence of random variables $X_1, X_2, X_3, \ldots, X_n, \ldots$ is said to **converge in probability** to the constant $c$, if for every $\epsilon > 0$,

$$\lim_{n\to\infty} P\left(|X_n - c| \leq \epsilon\right) = 1$$

or equivalently,

$$\lim_{n\to\infty} P\left(|X_n - c| > \epsilon\right) = 0$$

**Notation:** $X_n \xrightarrow{P} c$

This concept plays a key role in the Law of Large Numbers, where the sample mean of independent and identically distributed random variables converges in probability to the population mean as the sample size grows.

**Definition 5.1.2 (Chebyshev's Inequality)**

Let $X$ be a random variable with finite mean $\mu$ and variance $\sigma^2$. Then, for any $k > 0$,

$$P\left(|X - \mu| \geq k\right) \leq \frac{\sigma^2}{k^2}$$

*Using complements:*

$$P\left(|X - \mu| < k\right) \geq 1 - \frac{\sigma^2}{k^2}$$

**Definition 5.1.3 (Weak Law of Large Numbers (WLLN))**

Let $X_1, X_2, \ldots$ be a sequence of independent and identically distributed random variables, each having finite mean $E(X_i) = \mu$ and variance $\text{Var}(X_i) = \sigma^2$. Then, for any $\epsilon > 0$,

$$P\left(\left|\frac{X_1 + X_2 + \cdots + X_n}{n} - \mu\right| \geq \epsilon\right) \to 0 \quad \text{as } n \to \infty$$

**Notation:** $\bar{X}_n \xrightarrow{P} \mu$

**Proof of the Weak Law of Large Numbers (WLLN)**

We aim to show that for every $\epsilon > 0$,

$$\lim_{n\to\infty} P\left(\left|\bar{X}_n - \mu\right| > \epsilon\right) = 0$$

where $\bar{X}_n$ is the sample mean of $n$ independent and identically distributed (i.i.d.) random variables with

$$E(X_i) = \mu, \quad \text{and} \quad \text{Var}(X_i) = \sigma^2.$$

Let

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

By the Central Limit Theorem (CLT), we know that

$$\bar{X}_n \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right).$$

Now, applying **Chebyshev's Inequality**, which states that for any random variable $X$ with mean $\mu$ and variance $\sigma^2$,

$$P\left(|X - \mu| > k\right) \leq \frac{\sigma^2}{k^2} \quad \text{for } k > 0,$$

to $\bar{X}_n$, we set $k = \epsilon$, and obtain:

$$P\left(\left|\bar{X}_n - \mu\right| > \epsilon\right) \leq \frac{\text{Var}(\bar{X}_n)}{\epsilon^2} = \frac{\sigma^2/n}{\epsilon^2} = \frac{\sigma^2}{n\epsilon^2}.$$

Taking the limit as $n \to \infty$, we have:

$$\lim_{n\to\infty} P\left(\left|\bar{X}_n - \mu\right| > \epsilon\right) \leq \lim_{n\to\infty} \frac{\sigma^2}{n\epsilon^2} = 0.$$

Since probabilities are always non-negative, we conclude:

$$\lim_{n\to\infty} P\left(\left|\bar{X}_n - \mu\right| > \epsilon\right) = 0.$$

By the definition of convergence in probability,

$$\bar{X}_n \xrightarrow{P} \mu.$$

$\square$

## Example: Poisson Convergence via WLLN

Let $X_i$, for $i = 1, 2, 3, \ldots$, be independent Poisson random variables with rate parameter $\lambda = 3$. Prove that:

$$\bar{X}_n \xrightarrow{P} 3$$

**Properties of Poisson Distribution:**

$$E(X_i) = \lambda, \quad \mathrm{Var}(X_i) = \lambda$$

In this case, $\lambda = 3$, so:

$$E(X_i) = \mathrm{Var}(X_i) = 3$$

**Proof:**
We know:

$$E\left(\frac{X_1 + X_2 + \cdots + X_n}{n}\right) = 3, \quad \text{and} \quad \mathrm{Var}\left(\frac{X_1 + X_2 + \cdots + X_n}{n}\right) = \frac{3}{n}$$

Applying Chebyshev's Inequality:

$$P\left(\left|\frac{X_1 + X_2 + \cdots + X_n}{n} - 3\right| \geq \epsilon\right) \leq \frac{3}{n\epsilon^2}$$

Taking the limit as $n \to \infty$:

$$P\left(\left|\frac{X_1 + X_2 + \cdots + X_n}{n} - 3\right| \geq \epsilon\right) \to 0$$

**Conclusion:**

$$\bar{X}_n \xrightarrow{P} 3$$

### R Simulation Code (Single Sample Path)

```
n = 10
trial = seq(1, n, by = 1)
sample = rbinom(n, 1, 1/2)

plot(trial, cumsum(sample)/trial, type = "l", ylim = c(0,1), col = "blue")
points(trial, cumsum(sample)/trial, col = "red")
abline(h = 0.5, lty = 2, col = "black")
```

### R Simulation Code (Multiple Sample Paths)

```
n = 100
trial = seq(1, 100, by = 1)

sample1 = rbinom(n, 1, 1/2)
sample2 = rbinom(n, 1, 1/2)
```

```
sample3 = rbinom(n, 1, 1/2)
sample4 = rbinom(n, 1, 1/2)
sample5 = rbinom(n, 1, 1/2)
sample6 = rbinom(n, 1, 1/2)
sample7 = rbinom(n, 1, 1/2)
sample8 = rbinom(n, 1, 1/2)

colors = rainbow(8)

plot(trial, cumsum(sample1)/trial, type = "l", col = colors[1], ylim = c(0,1))
lines(trial, cumsum(sample2)/trial, col = colors[2])
lines(trial, cumsum(sample3)/trial, col = colors[3])
lines(trial, cumsum(sample4)/trial, col = colors[4])
lines(trial, cumsum(sample5)/trial, col = colors[5])
lines(trial, cumsum(sample6)/trial, col = colors[6])
lines(trial, cumsum(sample7)/trial, col = colors[7])
lines(trial, cumsum(sample8)/trial, col = colors[8])
abline(h = 0.5, lty = 2, col = "black")
```

*[Slide Image from Slide 10 will be inserted here — Single Trial Simulation]*
*[Slide Image from Slide 13 will be inserted here — Empirical Probability Explanation]*

**Empirical Probability Insight**

The Law of Large Numbers gives us empirical probabilities. Consider tossing a fair coin. Define the random variable $X$ as:

$$X = \begin{cases} 1 & \text{heads up} \\ 0 & \text{tails up} \end{cases}$$

Then as we sample more and more values of $X$, the sample mean $\bar{X}_n$ converges in probability to $P(\text{heads up})$, that is:

$$\bar{X}_n \xrightarrow{P} P(\text{heads up})$$

# Index