Topics:

- Introduction to random variables.
- Binomial random variables.
- Expectation of a discrete random variable.
- Wright-Fisher Model.
Bernoulli trials and Bernoulli random variables.

For a given $0 \leq p \leq 1$, a Bernoulli trial is an experiment with exactly two possible outcomes, success and failure, in which the probability of success is $p$ and probability of failure is $1 - p$.

Here, the sample space $S$ consists of the two outcomes, success and failure, and

$$P(\text{success}) = p \quad \text{and} \quad P(\text{failure}) = 1 - p$$

- **Example.** Consider tossing a coin such that it will fall heads up with some probability $p$, and tails up with probability $1 - p$.

Bernoulli random variable $X$ counts the number of successes after one Bernoulli trial, and thus, can equal either 0 or 1, with probabilities $1 - p$ and $p$.

Here, $P(X = 0) = 1 - p$ and $P(X = 1) = p$. 
Bernoulli trials

- **Example.** Consider performing independent Bernoulli trials, each with probability $p$ of success and probability $1 - p$ of failure. Let $X$ be a random variable representing the number of successes in two Bernoulli trials. Find $P(X = k)$ for $k = 0, 1, 2$.

- **Solution.** Let
  
  $S_1 = \{\text{trial 1} \Rightarrow \text{success}\}$ and $F_1 = \overline{S}_1 = \{\text{trial 1} \Rightarrow \text{failure}\}$
  
  $S_2 = \{\text{trial 2} \Rightarrow \text{success}\}$ and $F_2 = \overline{S}_2 = \{\text{trial 2} \Rightarrow \text{failure}\}$.

  Then $P(S_1) = P(S_2) = p$ and $P(F_1) = P(F_2) = 1 - p$, and

  $P(X = 0) = P(\text{no successes}) = P(F_1 \cap F_2) = P(F_1)P(F_2) = (1 - p)^2$

  $P(X = 1) = P(\text{one success and one failure}) = P(S_1 \cap F_2) + P(F_1 \cap S_2) = 2p(1 - p)$

  $P(X = 2) = P(\text{two successes}) = P(S_1 \cap S_2) = P(S_1)P(S_2) = p^2$

  by independence.
**Bernoulli trials.**

- **Example.** Consider performing independent Bernoulli trials, each with probability \( p \) of success and probability \( 1 - p \) of failure. Let \( X \) be a random variable representing the number of successes in three Bernoulli trials. Find \( P(X = k) \) for \( k = 0, 1, 2, 3 \).

- **Solution.** Let

\[
S_1 = \{\text{trial 1 } \Rightarrow \text{ success}\} \quad \text{and} \quad F_1 = \overline{S}_1 = \{\text{trial 1 } \Rightarrow \text{ failure}\}
\]

\[
S_2 = \{\text{trial 2 } \Rightarrow \text{ success}\} \quad \text{and} \quad F_2 = \overline{S}_2 = \{\text{trial 2 } \Rightarrow \text{ failure}\}
\]

\[
S_3 = \{\text{trial 3 } \Rightarrow \text{ success}\} \quad \text{and} \quad F_3 = \overline{S}_3 = \{\text{trial 3 } \Rightarrow \text{ failure}\}.
\]

Then, by independence,

\[
P(X = 0) = P(F_1 \cap F_2 \cap F_3) = \binom{3}{0}(1 - p)^3
\]

\[
P(X = 1) = P(S_1 \cap F_2 \cap F_3) + P(F_1 \cap S_2 \cap F_3) + P(F_1 \cap F_2 \cap S_3) = \binom{3}{1}p(1 - p)^2
\]

\[
P(X = 2) = P(S_1 \cap S_2 \cap F_3) + P(S_1 \cap F_2 \cap S_3) + P(F_1 \cap S_2 \cap S_3) = \binom{3}{2}p^2(1 - p)
\]

\[
P(X = 3) = P(S_1 \cap S_2 \cap S_3) = \binom{3}{3}p^3
\]
Bernoulli trials

- **Example.** Consider performing independent Bernoulli trials, each with probability $p$ of success and probability $1 - p$ of failure. Let $X$ be a random variable representing the number of successes in $n$ Bernoulli trials. Find $P(X = k)$ for $k = 0, 1, \ldots, n$.

- **Solution.** Let

  $S_1 = \{\text{trial 1 }\Rightarrow \text{ success}\}$ and $F_1 = \bar{S}_1 = \{\text{trial 1 }\Rightarrow \text{ failure}\}$

  $S_2 = \{\text{trial 2 }\Rightarrow \text{ success}\}$ and $F_2 = \bar{S}_2 = \{\text{trial 2 }\Rightarrow \text{ failure}\}$

  $\vdots$

  $S_n = \{\text{trial }n\Rightarrow \text{ success}\}$ and $F_n = \bar{S}_n = \{\text{trial }n\Rightarrow \text{ failure}\}$.

  Then, for each outcome with $k$ successes and $n-k$ failures, its probability

  $P(\underbrace{SFS\ldots FFS}_{k \ S's \ and \ n-k \ F's}) = P(S_1)P(F_2)P(S_3)P(S_4)\ldots P(F_{n-2})P(F_{n-1})P(S_n) = p^k(1-p)^{n-k}$

  and

  $P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$ for each $k = 0, 1, \ldots, n$

  because there are $\binom{n}{k}$ such outcomes.
Introduction to random variables.

Consider a sample space $S$ and a probability function $P$.

- **Definition.** A function from $S$ to $\mathbb{R}$ is a **random variable**.

- **Example.** Roll two fair dice. Let $X(i, j) = i + j$ for each outcome $(i, j)$ in $S$. Then $X$ is a random variable representing the sum of the digits on the dice.
Introduction to random variables.

- **Example.** Roll two fair dice. Let $X(i, j) = i + j$ for each outcome $(i, j)$ in $S$. Then $X$ is a random variable representing the sum of the digits on the dice.

<table>
<thead>
<tr>
<th></th>
<th>1,1</th>
<th>1,2</th>
<th>1,3</th>
<th>1,4</th>
<th>1,5</th>
<th>1,6</th>
</tr>
</thead>
<tbody>
<tr>
<td>2,1</td>
<td>2,2</td>
<td>2,3</td>
<td>2,4</td>
<td>2,5</td>
<td>2,6</td>
<td></td>
</tr>
<tr>
<td>3,1</td>
<td>3,2</td>
<td>3,3</td>
<td>3,4</td>
<td>3,5</td>
<td>3,6</td>
<td></td>
</tr>
<tr>
<td>4,1</td>
<td>4,2</td>
<td>4,3</td>
<td>4,4</td>
<td>4,5</td>
<td>4,6</td>
<td></td>
</tr>
<tr>
<td>5,1</td>
<td>5,2</td>
<td>5,3</td>
<td>5,4</td>
<td>5,5</td>
<td>5,6</td>
<td></td>
</tr>
<tr>
<td>6,1</td>
<td>6,2</td>
<td>6,3</td>
<td>6,4</td>
<td>6,5</td>
<td>6,6</td>
<td></td>
</tr>
</tbody>
</table>

Here, for example, $X(3, 1) = 4$ and $X(5, 6) = 11$.

We are interested in finding the following probabilities:

$$p(a) = P(X = a) \quad \text{for } a = 2, 3, \ldots, 12$$
Introduction to random variables.

- **Example.** Roll two fair dice. Let \( X(i, j) = i + j \) for each outcome \((i, j)\) in \(S\). Then \(X\) is a random variable representing the sum of the digits on the dice.

We are interested in finding the following probabilities:

\[
P(a) = P(X = a) \quad \text{for} \quad a = 2, 3, \ldots, 12
\]

\[
p(2) = \frac{1}{36}, \quad p(3) = \frac{2}{36}, \quad p(4) = \frac{3}{36}, \quad p(5) = \frac{4}{36}, \quad p(6) = \frac{5}{36}, \quad p(7) = \frac{6}{36}
\]

\[
p(8) = \frac{5}{36}, \quad p(9) = \frac{4}{36}, \quad p(10) = \frac{3}{36}, \quad p(11) = \frac{2}{36}, \quad p(12) = \frac{1}{36}
\]
**Introduction to random variables.**

Let $X$ a **discrete random variable**. That is $X$ assumes a discrete (countable) number of values.

- **Definition.** Function $p(a) = P(X = a)$ is called the **probability mass function** (or distribution function).

- **Definition.** Function $F(a) = P(X \leq a)$ is called the **cumulative distribution function**.

- **Note.** $\sum_{a: p(a) > 0} p(a) = 1$

  In the previous example, $p(2) + p(3) + \cdots + p(12) = 1$.

- **Note.** $0 \leq F(a) \leq 1$

- **Note.** $F(a) = \sum_{x: x \leq a} p(x)$
Binomial random variable. Recall the following example.

- **Example.** Consider performing independent Bernoulli trials, each with probability $p$ of success and probability $1-p$ of failure. Let $X$ be a random variable representing the number of successes in $n$ Bernoulli trials. Find $P(X = k)$ for $k = 0, 1, \ldots, n$.

- **Solution.**

Each outcome with $k$ successes and $n-k$ failures, its probability

$$P(SFSS\ldots FFSS) = p^k (1-p)^{n-k}$$

$k$ $S$'s and $n-k$ $F$'s

and

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

because there are $\binom{n}{k}$ such outcomes.

- **Definition.** The random variable $X$ in the above example is the binomial random variable with parameters $(n, p)$.

Check: $\sum_{k=0}^{n} p(k) = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} = (p + (1-p))^n = 1^n = 1$. 
Expectation of a discrete random variable.

• **Definition.** Let $X$ be a discrete random variable with the probability mass function $p(x)$. Then its expected value is

$$E[X] = \sum_{x: p(x) > 0} x \cdot p(x)$$

• **Example.** Let $X$ be a Bernoulli random variable with parameter $p$. Then

$$p(1) = P(X = 1) = p \quad \text{and} \quad p(0) = P(X = 0) = 1 - p$$

and

$$E[X] = 0 \cdot p(0) + 1 \cdot p(1) = p$$

• **Example.** Roll two fair dice. Let $X$ represent the sum of the digits on the dice. Then

$$E[X] = 2 \cdot p(2) + 3 \cdot p(3) + \cdots + 12 \cdot p(12) = \frac{252}{36} = 7$$
Expectation of a discrete random variable.

- **Definition.** Let $X$ be a discrete random variable with the probability mass function $p(x)$. Then its **expected value** is

$$E[X] = \sum_{x: p(x) > 0} x \cdot p(x)$$

- **Example.** Roll two fair dice. Let $X$ represent the sum of the digits on the dice. Then

$$E[X] = 2 \cdot p(2) + 3 \cdot p(3) + \cdots + 12 \cdot p(12) = \frac{252}{36} = 7$$

This corresponds to a **center of mass** of $p(a)$. 

\[ \text{p(a)} = \text{P}(X=a) \]
**Expectation of a discrete random variable.**

Let $X$ be a *binomial random variable* with parameters $(n, p)$. Then its probability mass function is known to be

$$p(k) = \binom{n}{k} p^k (1 - p)^{n-k} \quad \text{for each } k = 0, 1, \ldots, n$$

- **Definition.** Let $X$ be a discrete random variable with the probability mass function $p(x)$. Then its expected value is

$$E[X] = \sum_{x: p(x) > 0} x \cdot p(x)$$

- **Example.** Let $X$ be a binomial random variable with parameters $(n, p)$. Then

$$E[X] = \sum_{k=0}^{n} k \cdot p(k) = \sum_{k=0}^{n} k \cdot \binom{n}{k} p^k (1 - p)^{n-k} = ?$$
Expectation of a discrete random variable.

- **Example.** Let \( X \) be a binomial random variable with parameters \((n, p)\). Then \( E[X] = np \) since

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

\[
= \sum_{k=1}^{n} k \cdot \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} = \sum_{k=1}^{n} \frac{n!}{(k-1)!(n-k)!} p^k (1-p)^{n-k}
\]

\[
= \sum_{j=0}^{n-1} \frac{n!}{j!(n-1-j)!} p^{j+1} (1-p)^{n-1-j} = np \cdot \sum_{j=0}^{n-1} \frac{(n-1)!}{j!(n-1-j)!} p^{j} (1-p)^{n-1-j},
\]

where the new index \( j = k - 1 \). Thus

\[
E[X] = np \cdot \sum_{j=0}^{n-1} \frac{(n-1)!}{j!(n-1-j)!} p^{j} (1-p)^{n-1-j} = np \cdot \sum_{j=0}^{n-1} \binom{n-1}{j} p^{j} (1-p)^{n-1-j}
\]

\[
= np \cdot (p + (1-p))^{n-1} = np \quad \text{by the Binomial theorem}
\]
Binomial random variable.

\[ p(k) = \binom{n}{k} p^k (1-p)^{n-k} \quad \text{for each } k = 0, 1, \ldots, n \quad \text{and} \quad E[X] = np \]
Wright-Fisher Model.

Basic terminology:

- A genetic locus is a location in the genome of an organism. Common example: the sequence of nucleotides that makes up a gene.

- Allele is one of a number of alternative versions of the genetic information encoded at the locus.

Gregor Mendel (1822-1884, founder of modern genetics) discovered that there are alternative forms (alleles) of genes that account for variations in inherited characteristics. For example, the gene for flower color in pea plants used in Mendel’s experiment exists in two forms, one for purple and the other for white.
**Wright-Fisher Model.**

Diploid individuals have two copies of their genetic material in each cell. Consider a genetic locus with two alleles $A$ and $a$ that have the same fitness (ability to survive) in a diploid population of $N$ individuals.

Meiosis is a specialized type of cell division that reduces the chromosome number by half.

Mendel’s Law of Segregation states that every individual organism contains two alleles for each trait, and that these alleles segregate (separate) during meiosis such that each gamete (a cell that fuses with another cell during fertilization) contains only one of the alleles.

Mendel’s Law of Independent Assortment states that alleles for separate traits are passed independently of one another from parents to offspring.
Wright-Fisher Model.

Diploid individuals have two copies of their genetic material in each cell. Consider a genetic locus with two alleles $A$ and $a$ that have the same fitness (ability to survive) in a diploid population of $N$ individuals.

If the two alleles of an inherited pair differ (the heterozygous condition), then one determines the organism’s appearance and is called the dominant allele. Thus, in this simplified model, we think of $N$ individuals as $2N$ copies of the locus, without pairing copies of the locus.

Further simplifications: (i) fixed population size $N$ at all time; (ii) each allele at $2N$ copies of the locus is selected uniformly at random from the previous generation.
Wright-Fisher Model.

**Model:** Let $X_t$ denote the number of alleles $A$ in $t$-th generation. Thus, there are $2N - X_t$ of alleles $a$.

For each of the $2N$ copies of the locus in $(t+1)$-st generation, one of the $2N$ alleles in $t$-th generation is selected uniformly at random. Thus, in these $2N$ Bernoulli trials, allele $A$ comes up with probability

$$p = \frac{X_t}{2N}$$

and allele $a$ comes up with probability $1 - p$.

Hence, given that we know $X_t$, variable $X_{t+1}$ representing the number of alleles $A$ in $(t+1)$-st generation is a **Binomial random variable** with $2N$ trials and $p = \frac{X_t}{2N}$, i.e.

$$P(X_{t+1} = k \mid X_t = m) = \binom{2N}{k} p^k (1 - p)^{2N-k}, \quad \text{where } p = \frac{m}{2N}.$$
Wright-Fisher Model.

Hence, given that we know $X_t$, variable $X_{t+1}$ representing the number of alleles $A$ in $(t+1)$-st generation is a **Binomial random variable** with $2N$ trials and 

\[ p = \frac{X_t}{2N}, \text{ i.e.} \]

\[ P(X_{t+1} = k \mid X_t = m) = \binom{2N}{k} p^k (1 - p)^{2N-k}, \quad \text{where} \quad p = \frac{m}{2N}.\]

Urn model and fixation time: