Random Variables: Variance
Random Variables: Variance

1. Variance
2. Distributions
Variance

Flip a coin:

If H you make a dollar. If T you lose a dollar.

Let \( X \) be the RV indicating how much money you make.

\[
E(X) = 0.
\]

Flip a coin:

If H you make a million dollars. If T you lose a million dollars.

Let \( Y \) be the RV indicating how much money you make.

\[
E(Y) = 0.
\]

Any other measures???

What else that's informative can we say?
Flip a coin: If H you make a dollar. If T you lose a dollar.
Flip a coin: If H you make a dollar. If T you lose a dollar. Let $X$ be the RV indicating how much money you make.
Variance

Flip a coin: If H you make a dollar. If T you lose a dollar. Let $X$ be the RV indicating how much money you make. $E(X) =$

Any other measures??

What else that's informative can we say?
Flip a coin: If H you make a dollar. If T you lose a dollar. Let $X$ be the RV indicating how much money you make. $E(X) = 0$. 
Flipped a coin: If H you make a dollar. If T you lose a dollar. Let $X$ be the RV indicating how much money you make. $E(X) = 0$.

Flipped a coin:
Flip a coin: If H you make a dollar. If T you lose a dollar. Let $X$ be the RV indicating how much money you make. $E(X) = 0$.

Flip a coin: If H you make a million dollars. If T you lose a million dollars.
Flip a coin: If H you make a dollar. If T you lose a dollar. Let $X$ be the RV indicating how much money you make. $E(X) = 0$.

Flip a coin: If H you make a million dollars. If T you lose a million dollars. Let $Y$ be the RV indicating how much money you make.
Variance

Flip a coin: If H you make a dollar. If T you lose a dollar. Let $X$ be the RV indicating how much money you make. $E(X) = 0$.

Flip a coin: If H you make a million dollars. If T you lose a million dollars. Let $Y$ be the RV indicating how much money you make. $E(Y) =$
Flip a coin: If H you make a dollar. If T you lose a dollar. Let $X$ be the RV indicating how much money you make. $E(X) = 0$.

Flip a coin: If H you make a million dollars. If T you lose a million dollars. Let $Y$ be the RV indicating how much money you make. $E(Y) = 0$. 
Variance

Flip a coin: If H you make a dollar. If T you lose a dollar. Let $X$ be the RV indicating how much money you make. $E(X) = 0$.

Flip a coin: If H you make a million dollars. If T you lose a million dollars. Let $Y$ be the RV indicating how much money you make. $E(Y) = 0$.

Any other measures???
Flip a coin: If H you make a dollar. If T you lose a dollar. Let \( X \) be the RV indicating how much money you make. \( E(X) = 0 \).

Flip a coin: If H you make a million dollars. If T you lose a million dollars. Let \( Y \) be the RV indicating how much money you make. \( E(Y) = 0 \).

Any other measures???
What else that’s informative can we say?
Variance

The variance measures the deviation from the mean value.

Definition:

The variance of $X$ is $\sigma^2(X) := \text{var}[X] = E[(X - E[X])^2]$.

$\sigma(X)$ is called the standard deviation of $X$. 

Variance

The variance measures the deviation from the mean value.

Definition:

The variance of $X$ is $\sigma^2(X) := \text{var}[X] = E[(X - E[X])^2]$.

$\sigma(X)$ is called the standard deviation of $X$. 

![Diagram showing two bell curves with different variances. One labeled Var = 1, the other Var = 10.}]
Variance

The variance measures the deviation from the mean value.

\[
\sigma^2(X) := \text{var}[X] = E[(X - E[X])^2].
\]

\(\sigma(X)\) is called the standard deviation of \(X\).
Variance

The variance measures the deviation from the mean value.

**Definition:** The *variance* of $X$ is
**Variance**

The variance measures the deviation from the mean value.

**Definition:** The variance of $X$ is

$$\sigma^2(X) := var[X] = E[(X - E[X])^2].$$
The variance measures the deviation from the mean value.

**Definition:** The variance of $X$ is

$$\sigma^2(X) := \text{var}[X] = E[(X - E[X])^2].$$

$\sigma(X)$ is called the standard deviation of $X$. 
Variance and Standard Deviation

Fact:

\[ var[X] = E[X^2] - E[X]^2. \]
Fact:


Indeed:

$$\text{var}(X) = E[(X - E[X])^2]$$
Fact:

\[ var[X] = E[X^2] - E[X]^2. \]

Indeed:

\[
\begin{align*}
var(X) &= E[(X - E[X])^2] \\
&= E[X^2 - 2XE[X] + E[X]^2]
\end{align*}
\]
Variance and Standard Deviation

Fact:

\[
\]

Indeed:

\[
\text{var}(X) = E[(X - E[X])^2] \\
= E[X^2 - 2XE[X] + E[X]^2] \\
= E[X^2] - E[2XE[X]] + E[E[X]^2]
\]
Variance and Standard Deviation

Fact:

\[ \text{var}[X] = E[X^2] - E[X]^2. \]

Indeed:

\[
\text{var}(X) = E[(X - E[X])^2] \\
= E[X^2 - 2XE[X] + E[X]^2] \\
= E[X^2] - E[2XE[X]] + E[E[X]^2] \text{ by linearity}
\]
Variance and Standard Deviation

Fact:

\[ \text{var}[X] = E[X^2] - E[X]^2. \]

Indeed:

\[
\begin{align*}
\text{var}(X) & = E[(X - E[X])^2] \\
& = E[X^2 - 2XE[X] + E[X]^2] \\
& = E[X^2] - E[2XE[X]] + E[E[X]^2] \text{ by linearity} \\
& = E[X^2] - 2E[X]E[X] + E[X]^2, \\
\end{align*}
\]
Variance and Standard Deviation

Fact:

\[ \text{var}[X] = E[X^2] - E[X]^2. \]

Indeed:

\[
\begin{align*}
\text{var}(X) &= E[(X - E[X])^2] \\
&= E[X^2 - 2XE[X] + E[X]^2] \\
&= E[X^2] - E[2XE[X]] + E[E[X]^2] \text{ by linearity} \\
&= E[X^2] - 2E[X]E[X] + E[X]^2, \\
&= E[X^2] - E[X]^2.
\end{align*}
\]
Example

Consider $X$ with

$$X = \begin{cases} -1, & \text{w. p. 0.99} \\ 99, & \text{w. p. 0.01}. \end{cases}$$
Example

Consider $X$ with

$$X = \begin{cases} -1, & \text{w. p. 0.99} \\ 99, & \text{w. p. 0.01}. \end{cases}$$

Then

$$E[X] = -1 \times 0.99 + 99 \times 0.01 = 0.$$
Consider $X$ with

$$X = \begin{cases} 
-1, & \text{w. p. 0.99} \\
99, & \text{w. p. 0.01}.
\end{cases}$$

Then

$$E[X] = -1 \times 0.99 + 99 \times 0.01 = 0.$$  
$$E[X^2] = (-1)^2 \times 0.99 + (99)^2 \times 0.01 \approx 100.$$
Example

Consider $X$ with

$$X = \begin{cases} 
-1, & \text{w. p. 0.99} \\
99, & \text{w. p. 0.01}.
\end{cases}$$

Then

$$E[X] = -1 \times 0.99 + 99 \times 0.01 = 0.$$  
$$E[X^2] = (-1)^2 \times 0.99 + (99)^2 \times 0.01 \approx 100.$$  
$$Var(X) \approx 100 \implies \sigma(X) \approx 10.$$
A simple example

This example illustrates the term ‘standard deviation.’
A simple example

This example illustrates the term ‘standard deviation.’

\[ Pr = 0.5 \quad \sigma \quad \mu \quad \sigma \quad Pr = 0.5 \]

\[ \mu - \sigma \quad \mu \quad \mu + \sigma \]
A simple example

This example illustrates the term ‘standard deviation.’

Consider the random variable $X$ such that

$$X = \begin{cases} 
\mu - \sigma, & \text{w.p. } 1/2 \\
\mu + \sigma, & \text{w.p. } 1/2.
\end{cases}$$
A simple example

This example illustrates the term ‘standard deviation.’

Consider the random variable $X$ such that

$$X = \begin{cases} 
\mu - \sigma, & \text{w.p. } 1/2 \\
\mu + \sigma, & \text{w.p. } 1/2. 
\end{cases}$$

Then, $E[X] = \mu$ and $E[(X - E[X])^2] = \sigma^2$. 
A simple example

This example illustrates the term ‘standard deviation.’

Consider the random variable \(X\) such that

\[
X = \begin{cases} 
  \mu - \sigma, & \text{w.p. } 1/2 \\
  \mu + \sigma, & \text{w.p. } 1/2.
\end{cases}
\]

Then, \(E[X] = \mu\) and \(E[(X - E[X])^2] = \sigma^2\). Hence,

\[\text{var}(X) = \sigma^2\text{ and } \sigma(X) = \sigma.\]
Properties of variance.

1. $\text{Var}(cX) = c^2 \text{Var}(X)$, where $c$ is a constant.
Properties of variance.

1. $\text{Var}(cX) = c^2 \text{Var}(X)$, where $c$ is a constant.
   Scales by $c^2$. 
Properties of variance.

1. $\text{Var}(cX) = c^2 \text{Var}(X)$, where $c$ is a constant. Scales by $c^2$.
2. $\text{Var}(X + c) = \text{Var}(X)$, where $c$ is a constant.
Properties of variance.

1. $\text{Var}(cX) = c^2 \text{Var}(X)$, where $c$ is a constant. Scales by $c^2$.
2. $\text{Var}(X + c) = \text{Var}(X)$, where $c$ is a constant. Shifts center.
Properties of variance.

1. $\text{Var}(cX) = c^2 \text{Var}(X)$, where $c$ is a constant. Scales by $c^2$.

2. $\text{Var}(X + c) = \text{Var}(X)$, where $c$ is a constant. Shifts center.

Proof:

$$\text{Var}(cX) = E((cX)^2) - (E(cX))^2$$
Properties of variance.

1. \( \text{Var}(cX) = c^2 \text{Var}(X) \), where \( c \) is a constant. Scales by \( c^2 \).

2. \( \text{Var}(X + c) = \text{Var}(X) \), where \( c \) is a constant. Shifts center.

Proof:

\[
\text{Var}(cX) = E((cX)^2) - (E(cX))^2 \\
= c^2 E(X^2) - c^2 (E(X))^2
\]
Properties of variance.

1. \( \text{Var}(cX) = c^2 \text{Var}(X) \), where \( c \) is a constant. 
   Scales by \( c^2 \).

2. \( \text{Var}(X + c) = \text{Var}(X) \), where \( c \) is a constant. 
   Shifts center.

Proof:

\[
\text{Var}(cX) = E((cX)^2) - (E(cX))^2 \\
= c^2 E(X^2) - c^2 (E(X))^2 = c^2 (E(X^2) - E(X)^2)
\]
Properties of variance.

1. \( \text{Var}(cX) = c^2 \text{Var}(X) \), where \( c \) is a constant. Scales by \( c^2 \).

2. \( \text{Var}(X + c) = \text{Var}(X) \), where \( c \) is a constant. Shifts center.

Proof:

\[
\begin{align*}
\text{Var}(cX) &= E((cX)^2) - (E(cX))^2 \\
&= c^2 E(X^2) - c^2 (E(X))^2 \\
&= c^2 (E(X^2) - E(X)^2) \\
&= c^2 \text{Var}(X)
\end{align*}
\]
Properties of variance.

1. $\text{Var}(cX) = c^2 \text{Var}(X)$, where $c$ is a constant.
   Scales by $c^2$.

2. $\text{Var}(X + c) = \text{Var}(X)$, where $c$ is a constant.
   Shifts center.

Proof:

$$\text{Var}(cX) = E((cX)^2) - (E(cX))^2$$
$$= c^2 E(X^2) - c^2 (E(X))^2 = c^2 (E(X^2) - E(X)^2)$$
$$= c^2 \text{Var}(X)$$

$$\text{Var}(X + c) = E((X + c - E(X + c))^2)$$
Properties of variance.

1. \( \text{Var}(cX) = c^2 \text{Var}(X) \), where \( c \) is a constant. 
   Scales by \( c^2 \).

2. \( \text{Var}(X + c) = \text{Var}(X) \), where \( c \) is a constant. 
   Shifts center.

Proof:

\[
\begin{align*}
\text{Var}(cX) &= E((cX)^2) - (E(cX))^2 \\
&= c^2 E(X^2) - c^2 (E(X))^2 \\
&= c^2 (E(X^2) - E(X)^2) \\
&= c^2 \text{Var}(X) \\
\text{Var}(X + c) &= E((X + c - E(X + c))^2) \\
&= E((X + c - E(X) - c)^2)
\end{align*}
\]
Properties of variance.

1. \( Var(cX) = c^2 Var(X) \), where \( c \) is a constant.  
   Scales by \( c^2 \).

2. \( Var(X + c) = Var(X) \), where \( c \) is a constant.  
   Shifts center.

Proof:

\[
Var(cX) = E((cX)^2) - (E(cX))^2
\]
\[
= c^2 E(X^2) - c^2 (E(X))^2
= c^2 (E(X^2) - E(X)^2)
= c^2 Var(X)
\]

\[
Var(X + c) = E((X + c - E(X + c))^2)
\]
\[
= E((X + c - E(X) - c)^2)
= E((X - E(X))^2)
\]
Properties of variance.

1. \( \text{Var}(cX) = c^2 \text{Var}(X) \), where \( c \) is a constant. Scales by \( c^2 \).

2. \( \text{Var}(X + c) = \text{Var}(X) \), where \( c \) is a constant. Shifts center.

Proof:

\[
\begin{align*}
\text{Var}(cX) &= E((cX)^2) - (E(cX))^2 \\
&= c^2 E(X^2) - c^2 (E(X))^2 = c^2 (E(X^2) - E(X)^2) \\
&= c^2 \text{Var}(X)
\end{align*}
\]

\[
\begin{align*}
\text{Var}(X + c) &= E((X + c - E(X + c))^2) \\
&= E((X + c - E(X) - c)^2) \\
&= E((X - E(X))^2) = \text{Var}(X)
\end{align*}
\]
Properties of variance.

1. \( \text{Var}(cX) = c^2 \text{Var}(X) \), where \( c \) is a constant. Scales by \( c^2 \).

2. \( \text{Var}(X + c) = \text{Var}(X) \), where \( c \) is a constant. Shifts center.

Proof:

\[
\begin{align*}
\text{Var}(cX) &= E((cX)^2) - (E(cX))^2 \\
&= c^2 E(X^2) - c^2 (E(X))^2 = c^2 (E(X^2) - E(X)^2) \\
&= c^2 \text{Var}(X)
\end{align*}
\]

\[
\begin{align*}
\text{Var}(X + c) &= E((X + c - E(X + c))^2) \\
&= E((X + c - E(X) - c)^2) \\
&= E((X - E(X))^2) = \text{Var}(X)
\end{align*}
\]
Theorem: If $X$ and $Y$ are independent, then
$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y).$$

Proof: Since shifting the random variables does not change their variance, let us subtract their means. That is, we assume that $E(X) = 0$ and $E(Y) = 0$. Then, by independence,
$$E(XY) = E(X)E(Y) = 0.$$

Hence,
$$\text{Var}(X) = E(X^2), \quad \text{Var}(Y) = E(Y^2).$$

Thus,
$$\text{Var}(X + Y) = E((X + Y)^2) = E(X^2 + 2XY + Y^2) = E(X^2) + 2E(XY) + E(Y^2) = E(X^2) + E(Y^2) = \text{Var}(X) + \text{Var}(Y).$$
Variance of sum of two independent random variables

**Theorem:**
If $X$ and $Y$ are independent, then

$$Var(X + Y) = Var(X) + Var(Y).$$
Variance of sum of two independent random variables

**Theorem:**
If $X$ and $Y$ are independent, then

$$Var(X + Y) = Var(X) + Var(Y).$$

**Proof:**
Since shifting the random variables does not change their variance, let us subtract their means.
Variance of sum of two independent random variables

**Theorem:**
If $X$ and $Y$ are independent, then

$$Var(X + Y) = Var(X) + Var(Y).$$

**Proof:**
Since shifting the random variables does not change their variance, let us subtract their means.

That is, we assume that $E(X) = 0$ and $E(Y) = 0$. 
Variance of sum of two independent random variables

**Theorem:**
If $X$ and $Y$ are independent, then

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y).$$

**Proof:**
Since shifting the random variables does not change their variance, let us subtract their means.

That is, we assume that $E(X) = 0$ and $E(Y) = 0$.

Then, by independence,

$$E(XY) = E(X)E(Y) = 0.$$
Variance of sum of two independent random variables

**Theorem:**
If $X$ and $Y$ are independent, then

$$Var(X + Y) = Var(X) + Var(Y).$$

**Proof:**
Since shifting the random variables does not change their variance, let us subtract their means.

That is, we assume that $E(X) = 0$ and $E(Y) = 0$.

Then, by independence,

$$E(XY) = E(X)E(Y) = 0.$$ 

Thus,

$$Var(X) = E(X^2), Var(Y) = E(Y^2).$$
Variance of sum of two independent random variables

**Theorem:**
If $X$ and $Y$ are independent, then

$$Var(X + Y) = Var(X) + Var(Y).$$

**Proof:**
Since shifting the random variables does not change their variance, let us subtract their means.

That is, we assume that $E(X) = 0$ and $E(Y) = 0$.

Then, by independence,

$$E(XY) = E(X)E(Y) = 0.$$

Hence,

$$Var(X) = E(X^2), \ Var(Y) = E(Y^2).$$

Hence,

$$var(X + Y) = E((X + Y)^2).$$
Variance of sum of two independent random variables

**Theorem:**
If $X$ and $Y$ are independent, then

$$Var(X + Y) = Var(X) + Var(Y).$$

**Proof:**
Since shifting the random variables does not change their variance, let us subtract their means.

That is, we assume that $E(X) = 0$ and $E(Y) = 0$.

Then, by independence,

$$E(XY) = E(X)E(Y) = 0.$$

$$Var(X) = E(X^2), \ Var(Y) = E(Y^2).$$

Hence,

$$\text{var}(X + Y) = E((X + Y)^2) = E(X^2 + 2XY + Y^2)$$
Variance of sum of two independent random variables

**Theorem:**
If $X$ and $Y$ are independent, then

$$Var(X + Y) = Var(X) + Var(Y).$$

**Proof:**
Since shifting the random variables does not change their variance, let us subtract their means.

That is, we assume that $E(X) = 0$ and $E(Y) = 0$.

Then, by independence,

$$E(XY) = E(X)E(Y) = 0.$$

Hence,

$$Var(X) = E(X^2), Var(Y) = E(Y^2).$$

Hence,

$$\text{var}(X + Y) = E((X + Y)^2) = E(X^2 + 2XY + Y^2) = E(X^2) + 2E(XY) + E(Y^2)$$
Variance of sum of two independent random variables

**Theorem:**
If $X$ and $Y$ are independent, then

$$Var(X + Y) = Var(X) + Var(Y).$$

**Proof:**
Since shifting the random variables does not change their variance, let us subtract their means.

That is, we assume that $E(X) = 0$ and $E(Y) = 0$.

Then, by independence,

$$E(XY) = E(X)E(Y) = 0.$$

Hence,

$$Var(X) = E(X^2), \ Var(Y) = E(Y^2).$$

Hence,

$$\text{var}(X + Y) = E((X + Y)^2) = E(X^2 + 2XY + Y^2)$$
$$= E(X^2) + 2E(XY) + E(Y^2) = E(X^2) + E(Y^2)$$
Variance of sum of two independent random variables

**Theorem:**
If $X$ and $Y$ are independent, then

$$Var(X + Y) = Var(X) + Var(Y).$$

**Proof:**
Since shifting the random variables does not change their variance, let us subtract their means.

That is, we assume that $E(X) = 0$ and $E(Y) = 0$.

Then, by independence,

$$E(XY) = E(X)E(Y) = 0.$$

$$Var(X) = E(X^2), Var(Y) = E(Y^2).$$

Hence,

$$\begin{align*}
var(X + Y) &= E((X + Y)^2) = E(X^2 + 2XY + Y^2) \\
&= E(X^2) + 2E(XY) + E(Y^2) = E(X^2) + E(Y^2) \\
&= var(X) + var(Y).
\end{align*}$$
Theorem:
If $X, Y, Z, \ldots$ are pairwise independent, then
$$\text{var}(X + Y + Z + \cdots) = \text{var}(X) + \text{var}(Y) + \text{var}(Z) + \cdots.$$ 

Proof:
Since shifting the random variables does not change their variance, let us subtract their means. That is, we assume that $E[X] = E[Y] = \cdots = 0$. Then, by independence,
Also, $E[XZ] = E[YZ] = \cdots = 0$. Hence,
$$\text{var}(X + Y + Z + \cdots) = E((X + Y + Z + \cdots)^2) = E(X^2 + Y^2 + Z^2 + \cdots + 2XY + 2XZ + 2YZ + \cdots) = E(X^2) + E(Y^2) + E(Z^2) + \cdots + 0 + \cdots + 0 = \text{var}(X) + \text{var}(Y) + \text{var}(Z) + \cdots.$$
Variance of sum of independent random variables

**Theorem:**
If $X, Y, Z, \ldots$ are pairwise independent, then

$$\text{var}(X + Y + Z + \cdots) = \text{var}(X) + \text{var}(Y) + \text{var}(Z) + \cdots.$$
Variance of sum of independent random variables

Theorem:
If $X, Y, Z, \ldots$ are pairwise independent, then

$$var(X + Y + Z + \cdots) = var(X) + var(Y) + var(Z) + \cdots.$$ 

Proof:
Since shifting the random variables does not change their variance, let us subtract their means.
Variance of sum of independent random variables

**Theorem:**
If $X, Y, Z, \ldots$ are pairwise independent, then

$$\text{var}(X + Y + Z + \cdots) = \text{var}(X) + \text{var}(Y) + \text{var}(Z) + \cdots.$$  

**Proof:**
Since shifting the random variables does not change their variance, let us subtract their means.
That is, we assume that $E[X] = E[Y] = \cdots = 0.$
Variance of sum of independent random variables

**Theorem:**
If $X, Y, Z, \ldots$ are pairwise independent, then

\[
\text{var}(X + Y + Z + \cdots) = \text{var}(X) + \text{var}(Y) + \text{var}(Z) + \cdots.
\]

**Proof:**
Since shifting the random variables does not change their variance, let us subtract their means.

That is, we assume that $E[X] = E[Y] = \cdots = 0$.

Then, by independence,

\[
\]
Variance of sum of independent random variables

**Theorem:**
If $X, Y, Z, \ldots$ are pairwise independent, then

$$\text{var}(X + Y + Z + \cdots) = \text{var}(X) + \text{var}(Y) + \text{var}(Z) + \cdots.$$ 

**Proof:**
Since shifting the random variables does not change their variance, let us subtract their means.

That is, we assume that $E[X] = E[Y] = \cdots = 0$.

Then, by independence,

Also, $E[XZ] = E[YZ] = \cdots = 0.$
Variance of sum of independent random variables

**Theorem:**
If $X, Y, Z, \ldots$ are pairwise independent, then

$$\text{var}(X + Y + Z + \cdots) = \text{var}(X) + \text{var}(Y) + \text{var}(Z) + \cdots.$$ 

**Proof:**
Since shifting the random variables does not change their variance, let us subtract their means.

That is, we assume that $E[X] = E[Y] = \cdots = 0$.

Then, by independence,

$$E[XY] = E[X]E[Y] = 0. \text{ Also, } E[XZ] = E[YZ] = \cdots = 0.$$ 

Hence,

$$\text{var}(X + Y + Z + \cdots) = E((X + Y + Z + \cdots)^2)$$
Variance of sum of independent random variables

**Theorem:**
If \( X, Y, Z, \ldots \) are pairwise independent, then

\[
\text{var}(X + Y + Z + \cdots) = \text{var}(X) + \text{var}(Y) + \text{var}(Z) + \cdots.
\]

**Proof:**
Since shifting the random variables does not change their variance, let us subtract their means.

That is, we assume that \( E[X] = E[Y] = \cdots = 0 \).

Then, by independence,

\[
E[XY] = E[X]E[Y] = 0. \text{ Also, } E[XZ] = E[YZ] = \cdots = 0.
\]

Hence,

\[
\text{var}(X + Y + Z + \cdots) = E((X + Y + Z + \cdots)^2)
= E(X^2 + Y^2 + Z^2 + \cdots + 2XY + 2XZ + 2YZ + \cdots)
\]
Variance of sum of independent random variables

**Theorem:**
If $X, Y, Z, \ldots$ are pairwise independent, then

$$\text{var}(X + Y + Z + \cdots) = \text{var}(X) + \text{var}(Y) + \text{var}(Z) + \cdots.$$ 

**Proof:**
Since shifting the random variables does not change their variance, let us subtract their means.

That is, we assume that $E[X] = E[Y] = \cdots = 0$.

Then, by independence,

$$E[XY] = E[X]E[Y] = 0. \text{ Also, } E[XZ] = E[YZ] = \cdots = 0.$$ 

Hence,

$$\text{var}(X + Y + Z + \cdots) = E((X + Y + Z + \cdots)^2)$$ 
$$= E(X^2 + Y^2 + Z^2 + \cdots + 2XY + 2XZ + 2YZ + \cdots)$$ 
$$= E(X^2) + E(Y^2) + E(Z^2) + \cdots + 0 + \cdots + 0$$
Variance of sum of independent random variables

**Theorem:**
If $X, Y, Z, \ldots$ are pairwise independent, then

$$\text{var}(X + Y + Z + \cdots) = \text{var}(X) + \text{var}(Y) + \text{var}(Z) + \cdots.$$

**Proof:**
Since shifting the random variables does not change their variance, let us subtract their means.

That is, we assume that $E[X] = E[Y] = \cdots = 0$.

Then, by independence,

$$E[XY] = E[X]E[Y] = 0. \text{ Also, } E[XZ] = E[YZ] = \cdots = 0.$$

Hence,

$$\text{var}(X + Y + Z + \cdots) = E((X + Y + Z + \cdots)^2) = E(X^2 + Y^2 + Z^2 + \cdots + 2XY + 2XZ + 2YZ + \cdots) = E(X^2) + E(Y^2) + E(Z^2) + \cdots + 0 + \cdots + 0 = \text{var}(X) + \text{var}(Y) + \text{var}(Z) + \cdots.$$
Distributions

- Bernoulli
- Binomial
- Uniform
- Geometric
Bernoulli

Flip a coin, with heads probability $p$. 

Distribution:

$$X = \begin{cases} 
1 \text{ w.p. } p \\
0 \text{ w.p. } 1-p 
\end{cases}$$

$$E[X] = p$$

$$E[X^2] = 1 \times p + 0 \times (1-p) = p$$

$$Var[X] = E[X^2] - (E[X])^2 = p - p^2$$

Notice that:

$p = 0 \implies Var(X) = 0$

$p = 1 \implies Var(X) = 0$
Bernoulli

Flip a coin, with heads probability $p$.
Random variable $X$: 1 is heads, 0 if not heads.
Bernoulli

Flip a coin, with heads probability \( p \).

Random variable \( X \): 1 is heads, 0 if not heads.

\( X \) has the Bernoulli distribution.
Bernoulli

Flip a coin, with heads probability $p$.
Random variable $X$: 1 is heads, 0 if not heads.
$X$ has the Bernoulli distribution.

**Distribution:**

$$X = \begin{cases} 
1 & \text{w.p. } p \\
0 & \text{w.p. } 1 - p 
\end{cases}$$
Bernoulli

Flip a coin, with heads probability $p$.
Random variable $X$: 1 is heads, 0 if not heads.
$X$ has the Bernoulli distribution.

Distribution:

$$X = \begin{cases} 
1 & \text{w.p. } p \\
0 & \text{w.p. } 1 - p 
\end{cases}$$

$$E[X] = p$$
Bernoulli

Flip a coin, with heads probability $p$.
Random variable $X$: 1 is heads, 0 if not heads.
$X$ has the Bernoulli distribution.

Distribution:

$$X = \begin{cases} 1 \text{ w.p. } p \\ 0 \text{ w.p. } 1 - p \end{cases}$$

$$E[X] = p$$

$$E[X^2] =$$
Bernoulli

Flip a coin, with heads probability $p$.
Random variable $X$: 1 is heads, 0 if not heads.
$X$ has the Bernoulli distribution.

Distribution:

$$X = \begin{cases} 
1 & \text{w.p. } p \\
0 & \text{w.p. } 1 - p
\end{cases}$$

$$E[X] = p$$

$$E[X^2] = 1^2 \times p + 0^2 \times (1 - p)$$
Bernoulli

Flip a coin, with heads probability $p$.

Random variable $X$: 1 is heads, 0 if not heads.

$X$ has the Bernoulli distribution.

Distribution:

$$X = \begin{cases} 
1 & \text{w.p. } p \\
0 & \text{w.p. } 1 - p
\end{cases}$$

$$E[X] = p$$

$$E[X^2] = 1^2 \times p + 0^2 \times (1 - p) = p$$
Bernoulli

Flip a coin, with heads probability $p$.
Random variable $X$: 1 is heads, 0 if not heads.
$X$ has the Bernoulli distribution.

Distribution:

\[ X = \begin{cases} 
1 & \text{w.p. } p \\
0 & \text{w.p. } 1 - p 
\end{cases} \]

\[
E[X] = p
\]

\[
E[X^2] = 1^2 \times p + 0^2 \times (1 - p) = p
\]

\[
Var[X] = \]

Bernoulli

Flip a coin, with heads probability $p$.

Random variable $X$: 1 is heads, 0 if not heads.

$X$ has the Bernoulli distribution.

**Distribution:**

$$X = \begin{cases} 
1 & \text{w.p. } p \\
0 & \text{w.p. } 1 - p
\end{cases}$$

$$E[X] = p$$

$$E[X^2] = 1^2 \times p + 0^2 \times (1 - p) = p$$

$$Var[X] = E[X^2] - (E[X])^2 = p - p^2 = p(1 - p)$$
Bernoulli

Flip a coin, with heads probability $p$.
Random variable $X$: 1 is heads, 0 if not heads.
$X$ has the Bernoulli distribution.

Distribution:

$$X = \begin{cases} 
1 & \text{w.p. } p \\
0 & \text{w.p. } 1 - p 
\end{cases}$$

$$E[X] = p$$

$$E[X^2] = 1^2 \times p + 0^2 \times (1 - p) = p$$

$$Var[X] = E[X^2] - (E[X])^2 = p - p^2 = p(1 - p)$$

Notice that:
Bernoulli

Flip a coin, with heads probability $p$.

Random variable $X$: 1 is heads, 0 if not heads.

$X$ has the Bernoulli distribution.

**Distribution:**

$$X = \begin{cases} 
1 & \text{w.p. } p \\
0 & \text{w.p. } 1 - p 
\end{cases}$$

$$E[X] = p$$

$$E[X^2] = 1^2 \times p + 0^2 \times (1 - p) = p$$

$$Var[X] = E[X^2] - (E[X])^2 = p - p^2 = p(1 - p)$$

Notice that:

$p = 0$
Bernoulli

Flip a coin, with heads probability \( p \).
Random variable \( X \): 1 is heads, 0 if not heads.
\( X \) has the Bernoulli distribution.

Distribution:

\[
X = \begin{cases} 
1 & \text{w.p. } p \\
0 & \text{w.p. } 1 - p 
\end{cases}
\]

\[ E[X] = p \]

\[ E[X^2] = 1^2 \times p + 0^2 \times (1 - p) = p \]

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

Notice that:

\( p = 0 \implies \text{Var}(X) = 0 \)
Bernoulli

Flip a coin, with heads probability $p$.

Random variable $X$: 1 is heads, 0 if not heads.

$X$ has the Bernoulli distribution.

**Distribution:**

$$X = \begin{cases} 
1 & \text{w.p. } p \\
0 & \text{w.p. } 1 - p 
\end{cases}$$

$$E[X] = p$$

$$E[X^2] = 1^2 \times p + 0^2 \times (1 - p) = p$$

$$Var[X] = E[X^2] - (E[X])^2 = p - p^2 = p(1 - p)$$

Notice that:

$$p = 0 \implies Var(X) = 0$$

$$p = 1$$
Bernoulli

Flip a coin, with heads probability $p$.

Random variable $X$: 1 is heads, 0 if not heads.

$X$ has the Bernoulli distribution.

**Distribution:**

$$X = \begin{cases} 
1 \text{ w.p. } p \\
0 \text{ w.p. } 1 - p 
\end{cases}$$

$$E[X] = p$$

$$E[X^2] = 1^2 \times p + 0^2 \times (1 - p) = p$$

$$Var[X] = E[X^2] - (E[X])^2 = p - p^2 = p(1 - p)$$

Notice that:

$p = 0 \implies Var(X) = 0$

$p = 1 \implies Var(X) = 0$
Jacob Bernoulli
Binomial

Flip $n$ coins with heads probability $p$. 
Binomial

Flip $n$ coins with heads probability $p$.
Random variable: number of heads.
Binomial

Flip $n$ coins with heads probability $p$.
Random variable: number of heads.

Binomial Distribution: $Pr[X = i]$, for each $i$. 

How many sample points in event "$X = i$"?

$i$ heads out of $n$ coin flips $\Rightarrow \binom{n}{i}$

Sample space: $\Omega = \{\text{HHH}...\text{HH}, \text{HHH}...\text{HT},...\}$

What is the probability of $\omega$ if $\omega$ has $i$ heads?

Probability of heads in any position is $p$.
Probability of tails in any position is $(1 - p)$.

So, we get $Pr[\omega] = p^i (1 - p)^{n-i}$.

Probability of "$X = i$" is sum of $Pr[\omega]$, $\omega \in \{X = i\}$.

$Pr[X = i] = \binom{n}{i} p^i (1 - p)^{n-i}, i = 0, 1, ..., n$: Binomial distribution
Binomial

Flip $n$ coins with heads probability $p$.
Random variable: number of heads.

Binomial Distribution: $Pr[X = i]$, for each $i$.

How many sample points in event “$X = i$”?
Binomial

Flip $n$ coins with heads probability $p$.
Random variable: number of heads.

**Binomial Distribution:** $Pr[X = i]$, for each $i$.

How many sample points in event “$X = i$”?
$i$ heads out of $n$ coin flips
Binomial

Flip $n$ coins with heads probability $p$.
Random variable: number of heads.

Binomial Distribution: $Pr[X = i]$, for each $i$.

How many sample points in event “$X = i$”? 
$i$ heads out of $n$ coin flips $\implies \binom{n}{i}$
Binomial

Flip $n$ coins with heads probability $p$.
Random variable: number of heads.

**Binomial Distribution:** $Pr[X = i]$, for each $i$.

How many sample points in event “$X = i$”?  
$i$ heads out of $n$ coin flips $\implies \binom{n}{i}$

Sample space: $\Omega = \{HHH...HH, HHH...HT, \ldots\}$
Binomial

Flip $n$ coins with heads probability $p$.
Random variable: number of heads.

**Binomial Distribution:** $Pr[X = i]$, for each $i$.

How many sample points in event “$X = i$”?  
- $i$ heads out of $n$ coin flips $\implies \binom{n}{i}$

Sample space: $\Omega = \{HHH...HH, HHH...HT, \ldots\}$
What is the probability of $\omega$ if $\omega$ has $i$ heads?
Binomial

Flip \( n \) coins with heads probability \( p \).
Random variable: number of heads.

**Binomial Distribution:** \( Pr[X = i] \), for each \( i \).

How many sample points in event “\( X = i \)”?
\( i \) heads out of \( n \) coin flips \( \implies \binom{n}{i} \)

Sample space: \( \Omega = \{HHH...HH, HHH...HT, \ldots\} \)
What is the probability of \( \omega \) if \( \omega \) has \( i \) heads?
Probability of heads in any position is \( p \).
Binomial

Flip \( n \) coins with heads probability \( p \).
Random variable: number of heads.

Binomial Distribution: \( Pr[X = i] \), for each \( i \).

How many sample points in event “\( X = i \)”?
\( i \) heads out of \( n \) coin flips \( \implies \binom{n}{i} \)

Sample space: \( \Omega = \{HHH...HH, HHH...HT, \ldots\} \)

What is the probability of \( \omega \) if \( \omega \) has \( i \) heads?
Probability of heads in any position is \( p \).
Probability of tails in any position is \( (1 - p) \).
Binomial

Flip $n$ coins with heads probability $p$.
Random variable: number of heads.

Binomial Distribution: $Pr[X = i]$, for each $i$.

How many sample points in event “$X = i$”? 
$i$ heads out of $n$ coin flips $\implies \binom{n}{i}$

Sample space: $\Omega = \{HHH...HH, HHH...HT, \ldots\}$

What is the probability of $\omega$ if $\omega$ has $i$ heads?
Probability of heads in any position is $p$.
Probability of tails in any position is $(1 - p)$.
So, we get $Pr[\omega] = p^i$
Binomial

Flip $n$ coins with heads probability $p$.
Random variable: number of heads.

Binomial Distribution: $Pr[X = i]$, for each $i$.

How many sample points in event “$X = i$”?  
$i$ heads out of $n$ coin flips $\implies \binom{n}{i}$

Sample space: $\Omega = \{HHH\ldots HH, HHH\ldots HT, \ldots\}$

What is the probability of $\omega$ if $\omega$ has $i$ heads?  
Probability of heads in any position is $p$.  
Probability of tails in any position is $(1 - p)$.  
So, we get $Pr[\omega] = p^i(1 - p)^{n-i}$. 
Binomial

Flip $n$ coins with heads probability $p$.
Random variable: number of heads.

**Binomial Distribution**: $Pr[X = i]$, for each $i$.

How many sample points in event “$X = i$”? 
$i$ heads out of $n$ coin flips $\implies \binom{n}{i}$

Sample space: $\Omega = \{HHH...HH, HHH...HT, ...\}$

What is the probability of $\omega$ if $\omega$ has $i$ heads?
Probability of heads in any position is $p$.
Probability of tails in any position is $(1 - p)$.
So, we get $Pr[\omega] = p^i(1 - p)^{n-i}$.

Probability of “$X = i$” is sum of $Pr[\omega]$, $\omega \in \{X = i\}$. 
Binomial

Flip \( n \) coins with heads probability \( p \).
Random variable: number of heads.

**Binomial Distribution:** \( Pr[X = i] \), for each \( i \).

How many sample points in event “\( X = i \)”?
\( i \) heads out of \( n \) coin flips \( \Rightarrow \binom{n}{i} \)

Sample space: \( \Omega = \{HHH...HH, HHH...HT, \ldots\} \)

What is the probability of \( \omega \) if \( \omega \) has \( i \) heads?
Probability of heads in any position is \( p \).
Probability of tails in any position is \( (1 - p) \).
So, we get \( Pr[\omega] = p^i(1 - p)^{n-i} \).

Probability of “\( X = i \)” is sum of \( Pr[\omega] \), \( \omega \in “X = i” \).

\[
Pr[X = i] = \binom{n}{i} p^i(1 - p)^{n-i}, \quad i = 0, 1, \ldots, n: B(n, p) \text{ distribution}
\]
Expectation of Binomial Distribution

Indicator for the $i$-th coin:

$$x_i = \begin{cases} 
1 & \text{if } i\text{th flip is heads} \\
0 & \text{otherwise}
\end{cases}$$
Expectation of Binomial Distribution

Indicator for the $i$-th coin:

$$X_i = \begin{cases} 
1 & \text{if } i\text{th flip is heads} \\
0 & \text{otherwise}
\end{cases}$$

$$E[X_i] = 1 \times Pr["heads"] + 0 \times Pr["tails"]$$
Expectation of Binomial Distribution

Indicator for the $i$-th coin:

$$X_i = \begin{cases} 
1 & \text{if } i \text{th flip is heads} \\
0 & \text{otherwise}
\end{cases}$$

$$E[X_i] = 1 \times Pr[\text{“heads”}] + 0 \times Pr[\text{“tails”}] = p.$$
Expectation of Binomial Distribution

Indicator for the \(i\)-th coin:

\[
X_i = \begin{cases} 
1 & \text{if } i\text{th flip is heads} \\
0 & \text{otherwise}
\end{cases}
\]

\[
E[X_i] = 1 \times Pr[\text{“heads”}] + 0 \times Pr[\text{“tails”}] = p.
\]

Moreover \(X = X_1 + \cdots X_n\) and
Expectation of Binomial Distribution

Indicator for the \( i \)-th coin:

\[
X_i = \begin{cases} 
1 & \text{if \( i \)th flip is heads} \\
0 & \text{otherwise}
\end{cases}
\]

\[
E[X_i] = 1 \times Pr[\text{"heads"}] + 0 \times Pr[\text{"tails"}] = p.
\]

Moreover \( X = X_1 + \cdots + X_n \) and

\[
E[X] = E[X_1] + E[X_2] + \cdots + E[X_n]
\]
Expectation of Binomial Distribution

Indicator for the $i$-th coin:

$$X_i = \begin{cases} 
1 & \text{if $i$th flip is heads} \\
0 & \text{otherwise}
\end{cases}$$

$$E[X_i] = 1 \times Pr[\text{"heads"}] + 0 \times Pr[\text{"tails"}] = p.$$ 

Moreover $X = X_1 + \cdots X_n$ and

$$E[X] = E[X_1] + E[X_2] + \cdots E[X_n] = n \times E[X_i]$$
Expectation of Binomial Distribution

Indicator for the $i$-th coin:

$$X_i = \begin{cases} 
1 & \text{if $i$th flip is heads} \\
0 & \text{otherwise}
\end{cases}$$

$$E[X_i] = 1 \times Pr[\text{"heads"}] + 0 \times Pr[\text{"tails"}] = p.$$ 

Moreover $X = X_1 + \cdots + X_n$ and 

$$E[X] = E[X_1] + E[X_2] + \cdots E[X_n] = n \times E[X_i] = np.$$
Variance of Binomial Distribution.

\[ X_i = \begin{cases} 
1 & \text{if } \text{ith flip is heads} \\
0 & \text{otherwise} 
\end{cases} \]
Variance of Binomial Distribution.

\[ X_i = \begin{cases} 
1 & \text{if } i\text{th flip is heads} \\
0 & \text{otherwise} 
\end{cases} \]

\[ E(X_i^2) \]

\[ \text{Var}(X_i) = p - (E(X_i))^2 = p(1-p) \]
Variance of Binomial Distribution.

\[ X_i = \begin{cases} 
1 & \text{if } \text{ith flip is heads} \\
0 & \text{otherwise}
\end{cases} \]

\[ E(X_i^2) = 1^2 \times p + 0^2 \times (1 - p) \]
Variance of Binomial Distribution.

\[ X_i = \begin{cases} 
1 & \text{if } i\text{th flip is heads} \\
0 & \text{otherwise} 
\end{cases} \]

\[ E(X_i^2) = 1^2 \times p + 0^2 \times (1 - p) = p. \]
Variance of Binomial Distribution.

\[ X_i = \begin{cases} 
1 & \text{if } \text{i}th \text{ flip is heads} \\
0 & \text{otherwise}
\end{cases} \]

\[ E(X_i^2) = 1^2 \times p + 0^2 \times (1 - p) = p. \]
\[ Var(X_i) = p - (E(X_i))^2 \]
Variance of Binomial Distribution.

\[ X_i = \begin{cases} 
1 & \text{if } i\text{th flip is heads} \\
0 & \text{otherwise}
\end{cases} \]

\[
E(X_i^2) = 1^2 \times p + 0^2 \times (1 - p) = p.
\]
\[
Var(X_i) = p - (E(X_i))^2 = p - p^2
\]
Variance of Binomial Distribution.

\[ X_i = \begin{cases} 
1 & \text{if } \text{ith flip is heads} \\
0 & \text{otherwise} 
\end{cases} \]

\[ E(X_i^2) = 1^2 \times p + 0^2 \times (1 - p) = p. \]

\[ \text{Var}(X_i) = p - (E(X_i))^2 = p - p^2 = p(1 - p). \]
Variance of Binomial Distribution.

\[ X_i = \begin{cases} 
1 & \text{if } i\text{th flip is heads} \\
0 & \text{otherwise} 
\end{cases} \]

\[ E(X_i^2) = 1^2 \times p + 0^2 \times (1 - p) = p. \]
\[ \text{Var}(X_i) = p - (E(X_i))^2 = p - p^2 = p(1 - p). \]
\[ X = X_1 + X_2 + \ldots X_n. \]
Variance of Binomial Distribution.

\[ X_i = \begin{cases} 1 & \text{if } i\text{th flip is heads} \\ 0 & \text{otherwise} \end{cases} \]

\[ E(X_i^2) = 1^2 \times p + 0^2 \times (1 - p) = p. \]
\[ \text{Var}(X_i) = p - (E(X_i))^2 = p - p^2 = p(1 - p). \]

\[ X = X_1 + X_2 + \ldots + X_n. \]

\[ X_i \text{ and } X_j \text{ are independent:} \]
Variance of Binomial Distribution.

\[ X_i = \begin{cases} 
1 & \text{if } i\text{th flip is heads} \\
0 & \text{otherwise}
\end{cases} \]

\[ E(X_i^2) = 1^2 \times p + 0^2 \times (1 - p) = p. \]

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

\[ X = X_1 + X_2 + \ldots X_n. \]

\[ X_i \text{ and } X_j \text{ are independent: } Pr[X_i = 1|X_j = 1] = Pr[X_i = 1]. \]
Variance of Binomial Distribution.

\[ X_i = \begin{cases} 
1 & \text{if } i\text{th flip is heads} \\
0 & \text{otherwise} 
\end{cases} \]

\[ E(X_i^2) = 1^2 \times p + 0^2 \times (1 - p) = p. \]

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

\[ X = X_1 + X_2 + \ldots + X_n. \]

\( X_i \) and \( X_j \) are independent: \( \Pr[X_i = 1|X_j = 1] = \Pr[X_i = 1] \).

\[ \text{Var}(X) = \text{Var}(X_1 + \cdots + X_n) \]
Variance of Binomial Distribution.

\[ X_i = \begin{cases} 
1 & \text{if } i\text{th flip is heads} \\
0 & \text{otherwise} 
\end{cases} \]

\[ E(X_i^2) = 1^2 \times p + 0^2 \times (1 - p) = p. \]
\[ \text{Var}(X_i) = p - (E(X_i))^2 = p - p^2 = p(1 - p). \]
\[ X = X_1 + X_2 + \ldots X_n. \]

\[ X_i \text{ and } X_j \text{ are independent: } Pr[X_i = 1|X_j = 1] = Pr[X_i = 1]. \]

\[ \text{Var}(X) = \text{Var}(X_1 + \cdots X_n) = np(1 - p). \]
Uniform Distribution

Roll a six-sided balanced die. Let $X$ be the number of pips (dots). Then $X$ is equally likely to take any of the values $\{1, 2, \ldots, 6\}$. We say that $X$ is uniformly distributed in $\{1, 2, \ldots, 6\}$.

More generally, we say that $X$ is uniformly distributed in $\{1, 2, \ldots, n\}$ if $\Pr[X = m] = \frac{1}{n}$ for $m = 1, 2, \ldots, n$. In that case, $E[X] = \frac{n}{2}(n + 1)$. 


Uniform Distribution

Roll a six-sided balanced die. Let $X$ be the number of pips (dots).
Uniform Distribution

Roll a six-sided balanced die. Let $X$ be the number of pips (dots). Then $X$ is equally likely to take any of the values \{1, 2, ..., 6\}.
Uniform Distribution

Roll a six-sided balanced die. Let $X$ be the number of pips (dots). Then $X$ is equally likely to take any of the values $\{1, 2, \ldots, 6\}$. We say that $X$ is *uniformly distributed* in $\{1, 2, \ldots, 6\}$. 
Roll a six-sided balanced die. Let $X$ be the number of pips (dots). Then $X$ is equally likely to take any of the values $\{1, 2, \ldots, 6\}$. We say that $X$ is \textit{uniformly distributed} in $\{1, 2, \ldots, 6\}$.

More generally, we say that $X$ is uniformly distributed in $\{1, 2, \ldots, n\}$ if $\Pr[X = m] = 1/n$ for $m = 1, 2, \ldots, n$. 
Uniform Distribution

Roll a six-sided balanced die. Let $X$ be the number of pips (dots). Then $X$ is equally likely to take any of the values $\{1, 2, \ldots, 6\}$. We say that $X$ is uniformly distributed in $\{1, 2, \ldots, 6\}$.

More generally, we say that $X$ is uniformly distributed in $\{1, 2, \ldots, n\}$ if $Pr[X = m] = 1/n$ for $m = 1, 2, \ldots, n$. In that case,

$$E[X] = \sum_{m=1}^{n} mPr[X = m]$$
Roll a six-sided balanced die. Let $X$ be the number of pips (dots). Then $X$ is equally likely to take any of the values \{1, 2, \ldots, 6\}. We say that $X$ is \textit{uniformly distributed} in \{1, 2, \ldots, 6\}.

More generally, we say that $X$ is uniformly distributed in \{1, 2, \ldots, n\} if $Pr[X = m] = 1/n$ for $m = 1, 2, \ldots, n$. In that case,

$$E[X] = \sum_{m=1}^{n} mPr[X = m] = \sum_{m=1}^{n} m \times \frac{1}{n}$$
Roll a six-sided balanced die. Let $X$ be the number of pips (dots). Then $X$ is equally likely to take any of the values \{1, 2, \ldots, 6\}. We say that $X$ is uniformly distributed in \{1, 2, \ldots, 6\}.

More generally, we say that $X$ is uniformly distributed in \{1, 2, \ldots, n\} if $Pr[X = m] = 1/n$ for $m = 1, 2, \ldots, n$.

In that case,

$$E[X] = \sum_{m=1}^{n} mPr[X = m] = \sum_{m=1}^{n} m \times \frac{1}{n} = \frac{1}{n} \frac{n(n+1)}{2} = \frac{n+1}{2}.$$
Variance of Uniform

\[ E[X] = \frac{n + 1}{2}. \]
Variance of Uniform

\[ E[X] = \frac{n+1}{2}. \]

Also,

\[ E[X^2] = \sum_{i=1}^{n} i^2 \Pr[X = i] = \frac{1}{n} \sum_{i=1}^{n} i^2 \]
Variance of Uniform

\[ E[X] = \frac{n+1}{2}. \]

Also,

\[ E[X^2] = \sum_{i=1}^{n} i^2 \Pr[X = i] = \frac{1}{n} \sum_{i=1}^{n} i^2 = \frac{1 + 3n + 2n^2}{6}, \]
Variance of Uniform

\[ E[X] = \frac{n+1}{2}. \]

Also,

\[
E[X^2] = \sum_{i=1}^{n} i^2 \Pr[X = i] = \frac{1}{n} \sum_{i=1}^{n} i^2
\]

\[
= \frac{1 + 3n + 2n^2}{6}, \text{ as you can verify.}
\]
Variance of Uniform

\[ E[X] = \frac{n+1}{2}. \]

Also,

\[ E[X^2] = \sum_{i=1}^{n} i^2 Pr[X = i] = \frac{1}{n} \sum_{i=1}^{n} i^2 = \frac{1 + 3n + 2n^2}{6}, \]

as you can verify.

This gives

\[ \text{var}(X) = \frac{1 + 3n + 2n^2}{6} - \frac{(n+1)^2}{4} = \frac{n^2 - 1}{12}. \]
Geometric Distribution

Let's flip a coin with $\Pr[H] = p$ until we get $H$. For instance:

- $\omega_1 = H$,
- $\omega_2 = TH$,
- $\omega_3 = TTH$,
- $\omega_n = TTT\cdots TH$.

Note that $\Omega = \{\omega_n, n = 1, 2, \ldots\}$.

Let $X$ be the number of flips until the first $H$. Then, $X(\omega_n) = n$.

Also, $\Pr[X = n] = (1 - p)^{n-1}p$, $n \geq 1$. 
Geometric Distribution

Let’s flip a coin with $Pr[H] = p$ until we get $H$. 

Note that $\Omega = \{\omega_n, n = 1, 2, \ldots\}$. 

Let $X$ be the number of flips until the first $H$. Then, $X(\omega_n) = n$. 

Also, $Pr[X = n] = \left(1 - p\right)^{n-1} p$, $n \geq 1$. 

Geometric Distribution

Let’s flip a coin with $Pr[H] = p$ until we get $H$. 

For instance:

$\omega_1 = H$,

$\omega_2 = T H$,

$\omega_3 = T T H$,

$\omega_n = T T T T \cdots T H$.

Note that $\Omega = \{\omega_n, n = 1, 2, \ldots\}$.

Let $X$ be the number of flips until the first $H$. Then, $X(\omega_n) = n$.

Also, $Pr[X = n] = (1 - p)^{n-1} p$, $n \geq 1$. 


Geometric Distribution

Let’s flip a coin with $Pr[H] = p$ until we get $H$.

For instance:

$\omega_1 = H$, or
Geometric Distribution

Let's flip a coin with $Pr[H] = p$ until we get $H$.

For instance:

\[ \omega_1 = H, \text{ or } \omega_2 = T \, H, \text{ or } \omega_3 = T \, T \, H, \ldots \]

Let $X$ be the number of flips until the first $H$. Then, $X(\omega_n) = n$. Also,

\[ Pr[ X = n ] = (1 - p)^{n-1} p, \quad n \geq 1. \]
Geometric Distribution

Let’s flip a coin with $Pr[H] = p$ until we get $H$.

For instance:

\[ \omega_1 = H, \text{ or } \]
\[ \omega_2 = T H, \text{ or } \]
\[ \omega_3 = T T H, \text{ or } \]
Geometric Distribution

Let's flip a coin with $Pr[H] = p$ until we get $H$.

For instance:

$\omega_1 = H$, or
$\omega_2 = T\ H$, or
$\omega_3 = T\ T\ H$, or
$\omega_n = T\ T\ T\ T\ \cdots\ T\ H$. 
Geometric Distribution

Let’s flip a coin with \( Pr[H] = p \) until we get \( H \).

For instance:

\[
\begin{align*}
\omega_1 &= H, \text{ or} \\
\omega_2 &= T \ H, \text{ or} \\
\omega_3 &= T \ T \ H, \text{ or} \\
\omega_n &= T \ T \ T \ T \ \cdots \ T \ H.
\end{align*}
\]

Note that \( \Omega = \{ \omega_n, n = 1, 2, \ldots \} \).
Geometric Distribution
Let’s flip a coin with $Pr[H] = p$ until we get $H$.

For instance:

$\omega_1 = H$, or
$\omega_2 = T \ H$, or
$\omega_3 = T \ T \ H$, or
$\omega_n = T \ T \ T \ T \ \cdots \ T \ H$.

Note that $\Omega = \{\omega_n, n = 1, 2, \ldots\}$.

Let $X$ be the number of flips until the first $H$. Then, $X(\omega_n) =$
Geometric Distribution

Let’s flip a coin with $Pr[H] = p$ until we get $H$.

For instance:

$\omega_1 = H$, or
$\omega_2 = T \ H$, or
$\omega_3 = T \ T \ H$, or
$\omega_n = T \ T \ T \ T \ \cdots \ T \ H$.

Note that $\Omega = \{ \omega_n, n = 1, 2, \ldots \}$.

Let $X$ be the number of flips until the first $H$. Then, $X(\omega_n) = n$. 
Geometric Distribution

Let’s flip a coin with $Pr[H] = p$ until we get $H$.

For instance:

$\omega_1 = H$, or
$\omega_2 = T\ H$, or
$\omega_3 = T\ T\ H$, or
$\omega_n = T\ T\ T\ T\ \cdots\ T\ H$.

Note that $\Omega = \{\omega_n, n = 1, 2, \ldots\}$.

Let $X$ be the number of flips until the first $H$. Then, $X(\omega_n) = n$.

Also,

$Pr[X = n] =$
Geometric Distribution

Let’s flip a coin with \( Pr[H] = p \) until we get \( H \).

For instance:

\[
\begin{align*}
\omega_1 &= H, \text{ or} \\
\omega_2 &= T \ H, \text{ or} \\
\omega_3 &= T \ T \ H, \text{ or} \\
\omega_n &= T \ T \ T \ T \ \cdots \ T \ H.
\end{align*}
\]

Note that \( \Omega = \{ \omega_n, n = 1, 2, \ldots \} \).

Let \( X \) be the number of flips until the first \( H \). Then, \( X(\omega_n) = n \).

Also,

\[
Pr[X = n] = (1 - p)^{n-1} p, \ n \geq 1.
\]
Geometric Distribution

$$Pr[X = n] = (1 - p)^{n-1} p, n \geq 1.$$
Geometric Distribution

\[ Pr[X = n] = (1 - p)^{n-1} p, \quad n \geq 1. \]
Geometric Distribution

\[ Pr[X = n] = (1 - p)^{n-1} p, \quad n \geq 1. \]
Geometric Distribution

\[ Pr[X = n] = (1 - p)^{n-1} p, \quad n \geq 1. \]

Note that

\[ \sum_{n=1}^{\infty} Pr[X = n] = \]

\[ \frac{1}{1 - p}. \]
Geometric Distribution

\[ Pr[X = n] = (1 - p)^{n-1} p, \ n \geq 1. \]

Note that

\[ \sum_{n=1}^{\infty} Pr[X = n] = \sum_{n=1}^{\infty} (1 - p)^{n-1} p = \]
Geometric Distribution

\[ Pr[X = n] = (1 - p)^{n-1} p, \quad n \geq 1. \]

Note that

\[
\sum_{n=1}^{\infty} Pr[X = n] = \sum_{n=1}^{\infty} (1 - p)^{n-1} p = p \sum_{n=1}^{\infty} (1 - p)^{n-1}
\]
Geometric Distribution

$$Pr[X = n] = (1 - p)^{n-1} p, n \geq 1.$$  

Note that

$$\sum_{n=1}^{\infty} Pr[X = n] = \sum_{n=1}^{\infty} (1 - p)^{n-1} p = p \sum_{n=1}^{\infty} (1 - p)^{n-1} = p \sum_{n=0}^{\infty} (1 - p)^n.$$
Geometric Distribution

\[ Pr[X = n] = (1 - p)^{n-1} p, \quad n \geq 1. \]

Note that

\[
\sum_{n=1}^{\infty} Pr[X = n] = \sum_{n=1}^{\infty} (1 - p)^{n-1} p = p \sum_{n=1}^{\infty} (1 - p)^{n-1} = p \sum_{n=0}^{\infty} (1 - p)^n.
\]

We want to analyze \( S := \sum_{n=0}^{\infty} a^n \) for \( |a| < 1 \).
Geometric Distribution

\[ Pr[X = n] = (1 - p)^{n-1} p, \ n \geq 1. \]

Note that

\[
\sum_{n=1}^{\infty} Pr[X = n] = \sum_{n=1}^{\infty} (1 - p)^{n-1} p = p \sum_{n=1}^{\infty} (1 - p)^{n-1} = p \sum_{n=0}^{\infty} (1 - p)^n.
\]

We want to analyze \( S := \sum_{n=0}^{\infty} a^n \) for \(|a| < 1\). \( S = \frac{1}{1-a} \).
Geometric Distribution

\[ Pr[X = n] = (1 - p)^{n-1} p, \quad n \geq 1. \]

Note that

\[
\sum_{n=1}^{\infty} Pr[X = n] = \sum_{n=1}^{\infty} (1 - p)^{n-1} p = p \sum_{n=1}^{\infty} (1 - p)^{n-1} = p \sum_{n=0}^{\infty} (1 - p)^n.
\]

We want to analyze \( S := \sum_{n=0}^{\infty} a^n \) for \( |a| < 1 \). \( S = \frac{1}{1-a} \). Indeed,

\[
S = 1 + a + a^2 + a^3 + \cdots
\]
Geometric Distribution

\[ Pr[X = n] = (1 - p)^{n-1} p, \quad n \geq 1. \]

Note that

\[
\sum_{n=1}^{\infty} Pr[X = n] = \sum_{n=1}^{\infty} (1 - p)^{n-1} p = p \sum_{n=1}^{\infty} (1 - p)^{n-1} = p \sum_{n=0}^{\infty} (1 - p)^n.
\]

We want to analyze \( S := \sum_{n=0}^{\infty} a^n \) for \(|a| < 1\). \( S = \frac{1}{1-a} \). Indeed,

\[
S = 1 + a + a^2 + a^3 + \cdots
\]
\[
aS = a + a^2 + a^3 + a^4 + \cdots
\]
Geometric Distribution

\[ Pr[X = n] = (1 - p)^{n-1} p, \; n \geq 1. \]

Note that

\[
\sum_{n=1}^{\infty} Pr[X = n] = \sum_{n=1}^{\infty} (1 - p)^{n-1} p = p \sum_{n=1}^{\infty} (1 - p)^{n-1} = p \sum_{n=0}^{\infty} (1 - p)^n.
\]

We want to analyze \( S := \sum_{n=0}^{\infty} a^n \) for \( |a| < 1 \). \( S = \frac{1}{1-a} \). Indeed,

\[
S = 1 + a + a^2 + a^3 + \cdots
\]

\[
aS = a + a^2 + a^3 + a^4 + \cdots
\]

\[
(1-a)S = 1 + a - a + a^2 - a^2 + \cdots = 1.
\]
Geometric Distribution

\[ Pr[X = n] = (1 - p)^{n-1} p, \ n \geq 1. \]

Note that

\[
\sum_{n=1}^{\infty} Pr[X = n] = \sum_{n=1}^{\infty} (1 - p)^{n-1} p = p \sum_{n=1}^{\infty} (1 - p)^{n-1} = p \sum_{n=0}^{\infty} (1 - p)^n.
\]

We want to analyze \( S := \sum_{n=0}^{\infty} a^n \) for \(|a| < 1\). \( S = \frac{1}{1-a} \). Indeed,

\[
S = 1 + a + a^2 + a^3 + \cdots \\
aS = a + a^2 + a^3 + a^4 + \cdots \\
(1-a)S = 1 + a - a + a^2 - a^2 + \cdots = 1.
\]

Hence,

\[
\sum_{n=1}^{\infty} Pr[X = n] =
\]
Geometric Distribution

\[ Pr[X = n] = (1 - p)^{n-1}p, \ n \geq 1. \]

Note that

\[
\sum_{n=1}^{\infty} Pr[X = n] = \sum_{n=1}^{\infty} (1 - p)^{n-1}p = p \sum_{n=1}^{\infty} (1 - p)^{n-1} = p \sum_{n=0}^{\infty} (1 - p)^n.
\]

We want to analyze \( S := \sum_{n=0}^{\infty} a^n \) for \( |a| < 1 \). \( S = \frac{1}{1-a} \). Indeed,

\[
S = 1 + a + a^2 + a^3 + \cdots
\]
\[
aS = a + a^2 + a^3 + a^4 + \cdots
\]
\[
(1 - a)S = 1 + a - a + a^2 - a^2 + \cdots = 1.
\]

Hence,

\[
\sum_{n=1}^{\infty} Pr[X = n] = p \frac{1}{1 - (1 - p)} = \]

Geometric Distribution

\[ Pr[X = n] = (1 - p)^{n-1} p, \quad n \geq 1. \]

Note that

\[ \sum_{n=1}^{\infty} Pr[X = n] = \sum_{n=1}^{\infty} (1 - p)^{n-1} p = p \sum_{n=1}^{\infty} (1 - p)^{n-1} = p \sum_{n=0}^{\infty} (1 - p)^n. \]

We want to analyze \( S := \sum_{n=0}^{\infty} a^n \) for \( |a| < 1 \). \( S = \frac{1}{1-a} \). Indeed,

\[
\begin{align*}
S &= 1 + a + a^2 + a^3 + \cdots \\
\quad aS &= a + a^2 + a^3 + a^4 + \cdots \\
(1-a)S &= 1 + a - a + a^2 - a^2 + \cdots = 1.
\end{align*}
\]

Hence,

\[ \sum_{n=1}^{\infty} Pr[X = n] = p \frac{1}{1 - (1 - p)} = 1. \]
Geometric Distribution: Expectation

\[ X \sim \text{Geom}(p), \text{ i.e., } Pr[X = n] = (1 - p)^{n-1} p, n \geq 1. \]
Geometric Distribution: Expectation

\[ X \sim Geom(p), \text{ i.e., } Pr[X = n] = (1 - p)^{n-1} p, \ n \geq 1. \]

One has

\[ E[X] = \sum_{n=1}^{\infty} nPr[X = n] = \sum_{n=1}^{\infty} n(1 - p)^{n-1} p. \]
Geometric Distribution: Expectation

\[ X \sim \text{Geom}(p), \text{ i.e., } Pr[X = n] = (1 - p)^{n-1} p, n \geq 1. \]

One has

\[ E[X] = \sum_{n=1}^{\infty} nPr[X = n] = \sum_{n=1}^{\infty} n(1 - p)^{n-1} p. \]

Thus,

\[ E[X] = p \]
Geometric Distribution: Expectation

\[ X \sim Geom(p), \text{ i.e., } Pr[X = n] = (1 - p)^{n-1} p, n \geq 1. \]

One has

\[ E[X] = \sum_{n=1}^{\infty} nPr[X = n] = \sum_{n=1}^{\infty} n(1 - p)^{n-1} p. \]

Thus,

\[ E[X] = p + 2(1 - p)p \]
Geometric Distribution: Expectation

\[ X \sim \text{Geom}(p), \text{ i.e., } Pr[X = n] = (1 - p)^{n-1}p, \ n \geq 1. \]

One has

\[ E[X] = \sum_{n=1}^{\infty} nPr[X = n] = \sum_{n=1}^{\infty} n(1 - p)^{n-1}p. \]

Thus,

\[ E[X] = p + 2(1 - p)p + 3(1 - p)^2p \]
Geometric Distribution: Expectation

\[ X \sim Geom(p), \text{ i.e., } Pr[X = n] = (1 - p)^{n-1}p, \ n \geq 1. \]

One has

\[ E[X] = \sum_{n=1}^{\infty} nPr[X = n] = \sum_{n=1}^{\infty} n(1 - p)^{n-1}p. \]

Thus,

\[ E[X] = p + 2(1 - p)p + 3(1 - p)^2p + 4(1 - p)^3p + \cdots \]
Geometric Distribution: Expectation

\( X \sim Geom(p) \), i.e., \( Pr[X = n] = (1 - p)^{n-1} p, \ n \geq 1. \)

One has

\[
E[X] = \sum_{n=1}^{\infty} nPr[X = n] = \sum_{n=1}^{\infty} n(1 - p)^{n-1} p.
\]

Thus,

\[
E[X] = p + 2(1 - p)p + 3(1 - p)^2 p + 4(1 - p)^3 p + \ldots
\]

\[
(1 - p)E[X] = (1 - p)p + 2(1 - p)^2 p + 3(1 - p)^3 p + \ldots
\]
Geometric Distribution: Expectation

\( X \sim \text{Geom}(p), \) i.e., \( \Pr[X = n] = (1 - p)^{n-1}p, \ n \geq 1. \)

One has

\[
E[X] = \sum_{n=1}^{\infty} n\Pr[X = n] = \sum_{n=1}^{\infty} n(1 - p)^{n-1}p.
\]

Thus,

\[
E[X] = p + 2(1 - p)p + 3(1 - p)^2p + 4(1 - p)^3p + \cdots
\]

\[
(1 - p)E[X] = (1 - p)p + 2(1 - p)^2p + 3(1 - p)^3p + \cdots
\]

\[
pE[X] = p + (1 - p)p + (1 - p)^2p + (1 - p)^3p + \cdots
\]
Geometric Distribution: Expectation

\( X \sim \text{Geom}(p) \), i.e., \( Pr[X = n] = (1 - p)^{n-1}p, n \geq 1 \).

One has

\[
E[X] = \sum_{n=1}^{\infty} nPr[X = n] = \sum_{n=1}^{\infty} n(1 - p)^{n-1}p.
\]

Thus,

\[
E[X] = p + 2(1 - p)p + 3(1 - p)^2p + 4(1 - p)^3p + \cdots
\]

\[
(1 - p)E[X] = (1 - p)p + 2(1 - p)^2p + 3(1 - p)^3p + \cdots
\]

\[
pE[X] = p + (1 - p)p + (1 - p)^2p + (1 - p)^3p + \cdots
\]

by subtracting the previous two identities
Geometric Distribution: Expectation

\[ X \sim Geom(p), \text{ i.e., } Pr[X = n] = (1 - p)^{n-1}p, \, n \geq 1. \]

One has

\[ E[X] = \sum_{n=1}^{\infty} nPr[X = n] = \sum_{n=1}^{\infty} n(1 - p)^{n-1}p. \]

Thus,

\[ E[X] = p + 2(1 - p)p + 3(1 - p)^2p + 4(1 - p)^3p + \cdots \]

\[ (1 - p)E[X] = (1 - p)p + 2(1 - p)^2p + 3(1 - p)^3p + \cdots \]

\[ pE[X] = p + (1 - p)p + (1 - p)^2p + (1 - p)^3p + \cdots \]

by subtracting the previous two identities

\[ = \sum_{n=1}^{\infty} (1 - p)^{n-1}p = \]
Geometric Distribution: Expectation

\(X \sim Geom(p), \text{ i.e., } \Pr[X = n] = (1 - p)^{n-1}p, n \geq 1.\)

One has

\[ E[X] = \sum_{n=1}^{\infty} n \Pr[X = n] = \sum_{n=1}^{\infty} n (1 - p)^{n-1}p. \]

Thus,

\[ E[X] = p + 2(1 - p)p + 3(1 - p)^2p + 4(1 - p)^3p + \cdots \]
\[ (1 - p)E[X] = (1 - p)p + 2(1 - p)^2p + 3(1 - p)^3p + \cdots \]
\[ pE[X] = p + (1 - p)p + (1 - p)^2p + (1 - p)^3p + \cdots \]

by subtracting the previous two identities

\[ = \sum_{n=1}^{\infty} (1 - p)^{n-1}p = \sum_{n=1}^{\infty} \Pr[X = n] = \]
Geometric Distribution: Expectation

\[ X \sim \text{Geom}(p), \text{ i.e., } \Pr[X = n] = (1 - p)^{n-1}p, n \geq 1. \]

One has

\[
E[X] = \sum_{n=1}^{\infty} n \Pr[X = n] = \sum_{n=1}^{\infty} n(1 - p)^{n-1}p.
\]

Thus,

\[
E[X] = p + 2(1 - p)p + 3(1 - p)^2p + 4(1 - p)^3p + \ldots
\]

\[
(1 - p)E[X] = (1 - p)p + 2(1 - p)^2p + 3(1 - p)^3p + \ldots
\]

\[
pE[X] = p + (1 - p)p + (1 - p)^2p + (1 - p)^3p + \ldots
\]

by subtracting the previous two identities

\[
= \sum_{n=1}^{\infty} (1 - p)^{n-1}p = \sum_{n=1}^{\infty} \Pr[X = n] = 1.
\]
Geometric Distribution: Expectation

$X \sim \text{Geom}(p)$, i.e., $\Pr[X = n] = (1 - p)^{n-1}p, n \geq 1$.

One has

$$E[X] = \sum_{n=1}^{\infty} nPr[X = n] = \sum_{n=1}^{\infty} n(1 - p)^{n-1}p.$$ 

Thus,

$$E[X] = p + 2(1 - p)p + 3(1 - p)^2p + 4(1 - p)^3p + \cdots$$

$$(1 - p)E[X] = (1 - p)p + 2(1 - p)^2p + 3(1 - p)^3p + \cdots$$

$pE[X] = p + (1 - p)p + (1 - p)^2p + (1 - p)^3p + \cdots$

by subtracting the previous two identities

$$= \sum_{n=1}^{\infty} (1 - p)^{n-1}p = \sum_{n=1}^{\infty} Pr[X = n] = 1.$$

Hence,

$$E[X] = \frac{1}{p}. $$
Experiment: Get coupons at random from $n$ until collect all $n$ coupons.
Coupon Collectors Problem.

**Experiment**: Get coupons at random from $n$ until collect all $n$ coupons.
**Outcomes**: \{123145..., 56765...\}
Coupon Collectors Problem.

**Experiment:** Get coupons at random from $n$ until collect all $n$ coupons.

**Outcomes:** \{123145..., 56765...\}

**Random Variable:** $X$ - length of outcome.
Experiment: Get coupons at random from $n$ until collect all $n$ coupons.

Outcomes: $\{123145..., 56765...\}$

Random Variable: $X$ - length of outcome.

Before: $Pr[X \geq n \ln 2n] \leq \frac{1}{2}$. 
Coupon Collectors Problem.

**Experiment:** Get coupons at random from \( n \) until collect all \( n \) coupons.

**Outcomes:** \{123145..., 56765...\}

**Random Variable:** \( X \) - length of outcome.

Before: \( Pr[X \geq n\ln 2n] \leq \frac{1}{2} \).

Today: \( E[X] \)?
Time to collect coupons

$X$-time to get $n$ coupons.
Time to collect coupons

$X$ - time to get $n$ coupons.

$X_1$ - time to get first coupon.
Time to collect coupons

$X$-time to get $n$ coupons.

$X_1$ - time to get first coupon. Note: $X_1 = 1$. 

$E(X) = 1$. 

$E(X_2) = \frac{1}{p} = \frac{n}{n-1}$. 

$E(X_i) = \frac{1}{p} = \frac{n}{n-i+1}$, for $i = 1, 2, ..., n$. 

$E[X] = E[X_1] + \cdots + E[X_n] = n + n - 1 + n - 2 + \cdots + 1 = n^2 \approx n \ln n + \gamma n$.
Time to collect coupons

$X$-time to get $n$ coupons.

$X_1$ - time to get first coupon. Note: $X_1 = 1$. $E(X_1) = 1$. 

Note:

$E(X_1) = 1.$
Time to collect coupons

$X$-time to get $n$ coupons.

$X_1$ - time to get first coupon. Note: $X_1 = 1$. $E(X_1) = 1$.

$X_2$ - time to get second (distinct) coupon after getting first.
**Time to collect coupons**

$X$ - time to get $n$ coupons.

$X_1$ - time to get first coupon. Note: $X_1 = 1$. $E(X_1) = 1$.

$X_2$ - time to get second (distinct) coupon after getting first.

$Pr[\text{“get second distinct coupon”}|\text{“got first coupon”}]$
Time to collect coupons

$X$-time to get $n$ coupons.

$X_1$ - time to get first coupon. Note: $X_1 = 1$. $E(X_1) = 1$.

$X_2$ - time to get second (distinct) coupon after getting first.

$Pr[\text{“get second distinct coupon”}|\text{“got first coupon”}] = \frac{n-1}{n}$
Time to collect coupons

$X$-time to get $n$ coupons.

$X_1$ - time to get first coupon. Note: $X_1 = 1$. $E(X_1) = 1$.

$X_2$ - time to get second (distinct) coupon after getting first.

$Pr[\text{“get second distinct coupon”|“got first coupon”}] = \frac{n-1}{n}$

$E[X_2]$?
Time to collect coupons

$X$ - time to get $n$ coupons.

$X_1$ - time to get first coupon. Note: $X_1 = 1$. $E(X_1) = 1$.

$X_2$ - time to get second (distinct) coupon after getting first.

$Pr[\text{“get second distinct coupon”}|\text{“got first coupon”}] = \frac{n-1}{n}$

$E[X_2]$? Geometric
Time to collect coupons

$X$-time to get $n$ coupons.

$X_1$ - time to get first coupon. Note: $X_1 = 1. E(X_1) = 1$.

$X_2$ - time to get second (distinct) coupon after getting first.

$Pr[\text{“get second distinct coupon”}|\text{“got first coupon”}] = \frac{n-1}{n}$

$E[X_2]$? Geometric!
Time to collect coupons

$X$-time to get $n$ coupons.

$X_1$ - time to get first coupon. Note: $X_1 = 1$. $E(X_1) = 1$.

$X_2$ - time to get second (distinct) coupon after getting first.

$Pr[\text{“get second distinct coupon”|“got first coupon”}] = \frac{n-1}{n}$

$E[X_2]$? Geometric ! !
Time to collect coupons

$X$-time to get $n$ coupons.

$X_1$ - time to get first coupon. Note: $X_1 = 1$. $E(X_1) = 1$.

$X_2$ - time to get second (distinct) coupon after getting first.

$Pr[“get second distinct coupon”|“got first coupon”] = \frac{n-1}{n}$

$E[X_2]$? Geometric ! ! !
Time to collect coupons

$X$-time to get $n$ coupons.

$X_1$ - time to get first coupon. Note: $X_1 = 1$. $E(X_1) = 1$.

$X_2$ - time to get second (distinct) coupon after getting first.

$Pr[\text{“get second distinct coupon”} | \text{“got first coupon”}] = \frac{n-1}{n}$

$E[X_2]$? Geometric !!! $\implies E[X_2] = \frac{1}{p} =$
Time to collect coupons

$X$-time to get $n$ coupons.

$X_1$ - time to get first coupon. Note: $X_1 = 1$. $E(X_1) = 1$.

$X_2$ - time to get second (distinct) coupon after getting first.

$Pr[\text{“get second distinct coupon”|“got first coupon”}] = \frac{n-1}{n}$

$E[X_2]$? Geometric !!! $\implies E[X_2] = \frac{1}{p} = \frac{1}{n-1}$
Time to collect coupons

- **$X$-time to get $n$ coupons.**

  - $X_1$ - time to get first coupon. Note: $X_1 = 1$. $E(X_1) = 1$.
  
  - $X_2$ - time to get second (distinct) coupon after getting first.

  $Pr[\text{“get second distinct coupon”} | \text{“got first coupon”}] = \frac{n-1}{n}$

  $E[X_2]?$ Geometric ! ! ! $\implies E[X_2] = \frac{1}{p} = \frac{1}{\frac{n-1}{n}} = \frac{n}{n-1}.$
Time to collect coupons

- time to get \( n \) coupons.

\( X_1 \) - time to get first coupon. Note: \( X_1 = 1. \) \( E(X_1) = 1. \)

\( X_2 \) - time to get second (distinct) coupon after getting first.

\[ \Pr[\text{“get second distinct coupon”} | \text{“got first coupon”}] = \frac{n-1}{n} \]

\( E[X_2]? \) Geometric ! ! ! \( \implies E[X_2] = \frac{1}{p} = \frac{1}{n-1} = \frac{n}{n-1}. \)

\[ \Pr[\text{“getting } i \text{th distinct coupon”} | \text{“got } i - 1 \text{ distinct coupons”}] \]
Time to collect coupons

- Time to get 1 coupon.
- Time to get second (distinct) coupon after getting first.

\[ \Pr[\text{"get second distinct coupon"}|\text{"got first coupon"}] = \frac{n-1}{n} \]

\[ E[X_2] = \frac{1}{p} = \frac{1}{n-1} = \frac{n}{n-1}. \]

\[ \Pr[\text{"getting } i \text{th distinct coupon"}|\text{"got } i-1 \text{ distinct coupons"}] = \frac{n-(i-1)}{n} = \frac{n-i+1}{n} \]
Time to collect coupons

$X$-time to get $n$ coupons.

$X_1$ - time to get first coupon. Note: $X_1 = 1$. $E(X_1) = 1$.

$X_2$ - time to get second (distinct) coupon after getting first.

$Pr[\text{“get second distinct coupon”|“got first coupon”}] = \frac{n-1}{n}$

$E[X_2]$? Geometric ! ! ! $\implies E[X_2] = \frac{1}{p} = \frac{1}{\frac{n-1}{n}} = \frac{n}{n-1}$.

$Pr[\text{“getting } i\text{th distinct coupon|“got } i-1\text{ distinct coupons”}]$

$$= \frac{n-(i-1)}{n} = \frac{n-i+1}{n}$$

$E[X_i]$
Time to collect coupons

- time to get \( n \) coupons.

\( X_1 \) - time to get first coupon. Note: \( X_1 = 1 \). \( E(X_1) = 1 \).

\( X_2 \) - time to get second (distinct) coupon after getting first.

\[ P_r[\text{“get second distinct coupon”} | \text{“got first coupon”}] = \frac{n-1}{n} \]

\[ E[X_2]? \text{ Geometric ! ! !} \implies E[X_2] = \frac{1}{p} = \frac{1}{\frac{n-1}{n}} = \frac{n}{n-1}. \]

\[ P_r[\text{“getting } i \text{th distinct coupon”} | \text{“got } i-1 \text{ distinct coupons”}] \]

\[ = \frac{n-(i-1)}{n} = \frac{n-i+1}{n} \]

\[ E[X_i] = \frac{1}{p} \]
Time to collect coupons

$X$-time to get $n$ coupons.

$X_1$ - time to get first coupon. Note: $X_1 = 1$. $E(X_1) = 1$.

$X_2$ - time to get second (distinct) coupon after getting first.

$Pr[\text{"get second distinct coupon"}|\text{"got first coupon"}] = \frac{n-1}{n}$

$E[X_2]$? Geometric ! ! ! $\implies E[X_2] = \frac{1}{p} = \frac{1}{ \frac{n-1}{n} } = \frac{n}{n-1}$.

$Pr[\text{"getting $i$th distinct coupon"}|\text{"got $i-1$ distinct coupons"}]$

$$= \frac{n-(i-1)}{n} = \frac{n-i+1}{n}$$

$E[X_i] = \frac{1}{p} = \frac{n}{n-i+1}$.
Time to collect coupons

$X$-time to get $n$ coupons.

$X_1$ - time to get first coupon. Note: $X_1 = 1$. $E(X_1) = 1$.

$X_2$ - time to get second (distinct) coupon after getting first.

$Pr[\text{“get second distinct coupon”}|\text{“got first coupon”}] = \frac{n-1}{n}$

$E[X_2]$? Geometric ! ! ! $\implies E[X_2] = \frac{1}{p} = \frac{1}{n-1} = \frac{n}{n-1}$.

$Pr[\text{“getting } i\text{th distinct coupon”}|\text{“got } i-1 \text{ distinct coupons”}]$

$= \frac{n-(i-1)}{n} = \frac{n-i+1}{n}$

$E[X_i] = \frac{1}{p} = \frac{n}{n-i+1}$, $i = 1, 2, \ldots, n$. 

$E[X] = E[X_1] + \cdots + E[X_n] = \frac{n}{n} + \frac{n}{n-1} + \cdots + \frac{n}{1} = nH(n) \approx n \left( \ln n + \gamma \right)$
Time to collect coupons

$X$-time to get $n$ coupons.

$X_1$ - time to get first coupon. Note: $X_1 = 1$. $E(X_1) = 1$.

$X_2$ - time to get second (distinct) coupon after getting first.

$Pr[\text{“get second distinct coupon”}|\text{“got first coupon”}] = \frac{n-1}{n}$

$E[X_2]$? Geometric !!! $\implies E[X_2] = \frac{1}{p} = \frac{1}{n-1} = \frac{n}{n-1}$.

$Pr[\text{“getting } i\text{th distinct coupon}|\text{“got } i-1\text{ distinct coupons”}]$

$$= \frac{n-(i-1)}{n} = \frac{n-i+1}{n}$$

$E[X_i] = \frac{1}{p} = \frac{n}{n-i+1}, i = 1, 2, \ldots, n.$

$$E[X] = E[X_1] + \cdots + E[X_n] =$$
Time to collect coupons

- \(X\) - time to get \(n\) coupons.

- \(X_1\) - time to get first coupon. Note: \(X_1 = 1\). \(E(X_1) = 1\).

- \(X_2\) - time to get second (distinct) coupon after getting first.

\[ Pr[“\text{get second distinct coupon”} | “\text{got first coupon’’} ] = \frac{n-1}{n} \]

\[ E[X_2]? \text{ Geometric ! ! !} \implies E[X_2] = \frac{1}{p} = \frac{1}{\frac{n-1}{n}} = \frac{n}{n-1}. \]

\[ Pr[“\text{getting } i\text{th distinct coupon} | “\text{got } i-1\text{ distinct coupons”}] \]

\[ = \frac{n-(i-1)}{n} = \frac{n-i+1}{n} \]

\[ E[X_i] = \frac{1}{p} = \frac{n}{n-i+1}, i = 1, 2, \ldots, n. \]

\[ E[X] = E[X_1] + \cdots + E[X_n] = \frac{n}{n} + \frac{n}{n-1} + \frac{n}{n-2} + \cdots + \frac{n}{1} \]
Time to collect coupons

\(X\)-time to get \(n\) coupons.

\(X_1\) - time to get first coupon. Note: \(X_1 = 1.\ E(X_1) = 1.\)

\(X_2\) - time to get second (distinct) coupon after getting first.

\(Pr[\text{“get second distinct coupon”|“got first coupon”}] = \frac{n-1}{n}\)

\(E[X_2]?\ Geometric \ ! \ ! \ ! \implies E[X_2] = \frac{1}{\rho} = \frac{1}{\frac{n-1}{n}} = \frac{n}{n-1}.\)

\(Pr[\text{“getting } i\text{th distinct coupon|“got } i-1\text{ distinct coupons”}]\)

\[= \frac{n - (i - 1)}{n} = \frac{n - i + 1}{n}\]

\(E[X_i] = \frac{1}{\rho} = \frac{n}{n-i+1}, \ i = 1, 2, \ldots, n.\)

\(E[X] = E[X_1] + \cdots + E[X_n] = \frac{n}{n} + \frac{n}{n-1} + \frac{n}{n-2} + \cdots + \frac{n}{1}\)

\[= n(1 + \frac{1}{2} + \cdots + \frac{1}{n}) =: nH(n)\)
Time to collect coupons

- $X$-time to get $n$ coupons.

- $X_1$ - time to get first coupon. Note: $X_1 = 1$. $E(X_1) = 1$.

- $X_2$ - time to get second (distinct) coupon after getting first.

$PPr[\text{"get second distinct coupon"}\mid \text{"got first coupon"}] = \frac{n-1}{n}$

$E[X_2]$? Geometric!!! $\implies E[X_2] = \frac{1}{p} = \frac{1}{n-1} = \frac{n}{n-1}$.

$PPr[\text{"getting } i\text{th distinct coupon}\mid \text{"got } i-1\text{ distinct coupons"}]$

\[ \frac{n-(i-1)}{n} = \frac{n-i+1}{n} \]

$E[X_i] = \frac{1}{p} = \frac{n}{n-i+1}, i = 1, 2, \ldots, n.$

$E[X] = E[X_1] + \cdots + E[X_n] = \frac{n}{n} + \frac{n}{n-1} + \frac{n}{n-2} + \cdots + \frac{n}{1}$

\[ = n(1 + \frac{1}{2} + \cdots + \frac{1}{n}) =: nH(n) \approx n(\ln n + \gamma) \]
Review: Harmonic sum

\[ H(n) = 1 + \frac{1}{2} + \cdots + \frac{1}{n} \approx \int_1^n \frac{1}{x} \, dx = \ln(n). \]
Review: Harmonic sum

\[ H(n) = 1 + \frac{1}{2} + \cdots + \frac{1}{n} \approx \int_1^n \frac{1}{x} \, dx = \ln(n). \]
Review: Harmonic sum

\[ H(n) = 1 + \frac{1}{2} + \cdots + \frac{1}{n} \approx \int_1^n \frac{1}{x} \, dx = \ln(n). \]

A good approximation is

\[ H(n) \approx \ln(n) + \gamma \quad \text{where} \quad \gamma \approx 0.58 \quad \text{(Euler-Mascheroni constant)}. \]
Harmonic sum: Paradox

Consider this stack of cards (no glue!):
Harmonic sum: Paradox

Consider this stack of cards (no glue!):
Harmonic sum: Paradox

Consider this stack of cards (no glue!):

If each card has length 2, the stack can extend $H(n)$ to the right of the table.
Harmonic sum: Paradox

Consider this stack of cards (no glue!):

If each card has length 2, the stack can extend $H(n)$ to the right of the table. As $n$ increases, you can go as far as you want!
Stacking

The cards have width 2. Induction shows that the center of gravity after \( n \) cards is \( H(n) \) away from the right-most edge.

\[
nx = 1 - x \quad \Rightarrow \quad x = 1/(n + 1)
\]
The cards have width 2.
The cards have width 2. Induction shows that the center of gravity after $n$ cards is $H(n)$ away from the right-most edge.
Geometric Distribution: Memoryless

Let $X$ be $\text{Geom}(p)$. Then, for $n \geq 0$, 

$$
\Pr\left[ X > n \right] = \Pr\left[ \text{first } n \text{ flips are } T \right] = (1 - p)^n.
$$

**Theorem** 

$$
\Pr\left[ X > n + m \mid X > n \right] = \Pr\left[ X > m \right], \quad m, n \geq 0.
$$

**Proof:** 

$$
\Pr\left[ X > n + m \mid X > n \right] = \Pr\left[ X > n \mid X > n \right] \cdot \Pr\left[ X > m \right] = (1 - p)^{n+m} = (1 - p)^n = \Pr\left[ X > m \right].
$$
Geometric Distribution: Memoryless

Let $X$ be $\text{Geom}(p)$. Then, for $n \geq 0$,

$$Pr[X > n] =$$
Geometric Distribution: Memoryless

Let $X$ be $\text{Geom}(p)$. Then, for $n \geq 0$,

$$Pr[X > n] = Pr[ \text{first } n \text{ flips are } T] = \left(1 - p\right)^n.$$
Geometric Distribution: Memoryless

Let $X$ be $Geom(p)$. Then, for $n \geq 0$,

$$Pr[X > n] = Pr[\text{first } n \text{ flips are } T] = (1 - p)^n.$$
Geometric Distribution: Memoryless

Let $X$ be $\text{Geom}(p)$. Then, for $n \geq 0$,

$$ Pr[X > n] = Pr[\text{ first } n \text{ flips are } T] = (1 - p)^n. $$

**Theorem**

$$ Pr[X > n + m | X > n] = Pr[X > m], \ m, n \geq 0. $$
Geometric Distribution: Memoryless

Let $X$ be $\text{Geom}(\rho)$. Then, for $n \geq 0$,

$$Pr[X > n] = Pr[ \text{first } n \text{ flips are } T] = (1 - \rho)^n.$$

**Theorem**

$$Pr[X > n + m | X > n] = Pr[X > m], m, n \geq 0.$$  

**Proof:**

$$Pr[X > n + m | X > n] =$$
Geometric Distribution: Memoryless

Let $X$ be $\text{Geom}(\rho)$. Then, for $n \geq 0$,

$$Pr[X > n] = Pr[ \text{first } n \text{ flips are } T] = (1 - \rho)^n.$$ 

**Theorem**

$$Pr[X > n + m | X > n] = Pr[X > m], \ m, n \geq 0.$$ 

**Proof:**

$$Pr[X > n + m | X > n] = \frac{Pr[X > n + m \text{ and } X > n]}{Pr[X > n]}$$
Geometric Distribution: Memoryless

Let $X$ be $\text{Geom}(p)$. Then, for $n \geq 0$,

$$Pr[X > n] = Pr[\text{first } n \text{ flips are } T] = (1 - p)^n.$$ 

Theorem

$$Pr[X > n + m | X > n] = Pr[X > m], m, n \geq 0.$$ 

Proof:

$$Pr[X > n + m | X > n] = \frac{Pr[X > n + m \text{ and } X > n]}{Pr[X > n]} = \frac{Pr[X > n + m]}{Pr[X > n]}$$
Geometric Distribution: Memoryless

Let $X$ be $Geom(p)$. Then, for $n \geq 0$,

$$Pr[X > n] = Pr[ \text{first } n \text{ flips are } T] = (1 - p)^n.$$ 

Theorem

$$Pr[X > n + m | X > n] = Pr[X > m], \ m, n \geq 0.$$ 

Proof:

$$Pr[X > n + m | X > n] = \frac{Pr[X > n + m \text{ and } X > n]}{Pr[X > n]} = \frac{Pr[X > n + m]}{Pr[X > n]} = \frac{(1 - p)^{n+m}}{(1 - p)^n} =$$
Geometric Distribution: Memoryless

Let $X$ be $\text{Geom}(p)$. Then, for $n \geq 0$,

$$Pr[X > n] = Pr[\text{first } n \text{ flips are } T] = (1 - p)^n.$$

Theorem

$$Pr[X > n + m|X > n] = Pr[X > m], m, n \geq 0.$$ 

Proof:

$$Pr[X > n + m|X > n] = \frac{Pr[X > n + m \text{ and } X > n]}{Pr[X > n]}$$

$$= \frac{Pr[X > n + m]}{Pr[X > n]}$$

$$= \frac{(1 - p)^{n+m}}{(1 - p)^n} = (1 - p)^m$$
Geometric Distribution: Memoryless

Let $X$ be $\text{Geom}(\rho)$. Then, for $n \geq 0$,

$$Pr[X > n] = Pr[\text{first } n \text{ flips are } T] = (1 - \rho)^n.$$ 

Theorem

$$Pr[X > n + m|X > n] = Pr[X > m], \; m, n \geq 0.$$ 

Proof:

$$Pr[X > n + m|X > n] = \frac{Pr[X > n + m \text{ and } X > n]}{Pr[X > n]}$$

$$= \frac{Pr[X > n + m]}{Pr[X > n]}$$

$$= \frac{(1 - \rho)^{n+m}}{(1 - \rho)^n} = (1 - \rho)^m$$

$$= Pr[X > m].$$
The coin is memoryless, therefore, so is \( X \).

\[
Pr[X > n + m | X > n] = Pr[X > m], \ m, n \geq 0.
\]
Geometric Distribution: Memoryless - Interpretation

\[ Pr[X > n + m \mid X > n] = Pr[X > m], \quad m, n \geq 0. \]
Geometric Distribution: Memoryless - Interpretation

\[ Pr[X > n + m|X > n] = Pr[X > m], m, n \geq 0. \]

The coin is memoryless, therefore, so is \( X \).
Geometric Distribution: Yet another look

**Theorem:** For a r.v. $X$ that takes the values $\{0, 1, 2, \ldots\}$, one has

\[ E[X] = \sum_{i=1}^{\infty} Pr[X \geq i]. \]

[See later for a proof.]
Theorem: For a r.v. $X$ that takes the values $\{0, 1, 2, \ldots\}$, one has

$$E[X] = \sum_{i=1}^{\infty} Pr[X \geq i].$$

[See later for a proof.]

If $X = Geom(p)$, then $Pr[X \geq i] = Pr[X > i - 1] = (1 - p)^{i-1}$. 
**Theorem:** For a r.v. $X$ that takes the values $\{0, 1, 2, \ldots\}$, one has

$$E[X] = \sum_{i=1}^{\infty} Pr[X \geq i].$$

[See later for a proof.]

If $X = Geom(p)$, then $Pr[X \geq i] = Pr[X > i - 1] = (1 - p)^{i-1}$. Hence,

$$E[X] = \sum_{i=1}^{\infty} (1 - p)^{i-1} = \sum_{i=0}^{\infty} (1 - p)^i$$
Geometric Distribution: Yet another look

**Theorem:** For a r.v. $X$ that takes the values $\{0, 1, 2, \ldots\}$, one has

$$E[X] = \sum_{i=1}^{\infty} Pr[X \geq i].$$

[See later for a proof.]

If $X = Geom(p)$, then $Pr[X \geq i] = Pr[X > i - 1] = (1 - p)^{i-1}$. Hence,

$$E[X] = \sum_{i=1}^{\infty} (1 - p)^{i-1} = \sum_{i=0}^{\infty} (1 - p)^i = \frac{1}{1 - (1 - p)} = \frac{1}{p}.$$
Theorem: For a r.v. $X$ that takes the values $\{0, 1, 2, \ldots\}$, one has
\[
E[X] = \sum_{i=1}^{\infty} \Pr[X \geq i].
\]
[See later for a proof.]

If $X = Geom(p)$, then $\Pr[X \geq i] = \Pr[X > i - 1] = (1 - p)^{i-1}$. Hence,
\[
E[X] = \sum_{i=1}^{\infty} (1 - p)^{i-1} = \sum_{i=0}^{\infty} (1 - p)^i = \frac{1}{1 - (1 - p)} = \frac{1}{p}.
\]
A side step: Expected Value of Integer RV

**Theorem:** For a r.v. $X$ that takes values in $\{0, 1, 2, \ldots\}$, one has

$$E[X] = \sum_{i=1}^{\infty} Pr[X \geq i].$$
A side step: Expected Value of Integer RV

**Theorem:** For a r.v. $X$ that takes values in $\{0, 1, 2, \ldots\}$, one has

$$E[X] = \sum_{i=1}^{\infty} Pr[X \geq i].$$

**Proof:** One has

$$E[X] = \sum_{i=1}^{\infty} i \times Pr[X = i]$$
A side step: Expected Value of Integer RV

**Theorem:** For a r.v. $X$ that takes values in $\{0, 1, 2, \ldots\}$, one has

$$E[X] = \sum_{i=1}^{\infty} Pr[X \geq i].$$

**Proof:** One has

\[
E[X] = \sum_{i=1}^{\infty} i \times Pr[X = i]
\]

\[
= \sum_{i=1}^{\infty} i(Pr[X \geq i] - Pr[X \geq i + 1])
\]
Theorem: For a r.v. $X$ that takes values in $\{0, 1, 2, \ldots\}$, one has

$$E[X] = \sum_{i=1}^{\infty} Pr[X \geq i].$$

Proof: One has

$$E[X] = \sum_{i=1}^{\infty} i \times Pr[X = i]$$

$$= \sum_{i=1}^{\infty} i (Pr[X \geq i] - Pr[X \geq i+1])$$

$$= \sum_{i=1}^{\infty} (i \times Pr[X \geq i] - i \times Pr[X \geq i+1])$$
A side step: Expected Value of Integer RV

**Theorem:** For a r.v. $X$ that takes values in $\{0, 1, 2, \ldots\}$, one has

$$E[X] = \sum_{i=1}^{\infty} Pr[X \geq i].$$

**Proof:** One has

\[
E[X] = \sum_{i=1}^{\infty} i \times Pr[X = i]
\]

\[
= \sum_{i=1}^{\infty} i(Pr[X \geq i] - Pr[X \geq i + 1])
\]

\[
= \sum_{i=1}^{\infty} (i \times Pr[X \geq i] - i \times Pr[X \geq i + 1])
\]

\[
= \sum_{i=1}^{\infty} i \times Pr[X \geq i] - \sum_{i=1}^{\infty} i \times Pr[X \geq i + 1]
\]
A side step: Expected Value of Integer RV

**Theorem:** For a r.v. $X$ that takes values in $\{0, 1, 2, \ldots\}$, one has

$$E[X] = \sum_{i=1}^{\infty} Pr[X \geq i].$$

**Proof:** One has

$$E[X] = \sum_{i=1}^{\infty} i \times Pr[X = i]$$

$$= \sum_{i=1}^{\infty} i (Pr[X \geq i] - Pr[X \geq i + 1])$$

$$= \sum_{i=1}^{\infty} (i \times Pr[X \geq i] - i \times Pr[X \geq i + 1])$$

$$= \sum_{i=1}^{\infty} i \times Pr[X \geq i] - \sum_{i=1}^{\infty} i \times Pr[X \geq i + 1]$$

$$= \sum_{i=1}^{\infty} i \times Pr[X \geq i] - \sum_{i=1}^{\infty} (i - 1) \times Pr[X \geq i]$$
A side step: Expected Value of Integer RV

**Theorem:** For a r.v. $X$ that takes values in $\{0, 1, 2, \ldots\}$, one has

$$E[X] = \sum_{i=1}^{\infty} Pr[X \geq i].$$

**Proof:** One has

$$E[X] = \sum_{i=1}^{\infty} i \times Pr[X = i]$$

$$= \sum_{i=1}^{\infty} i (Pr[X \geq i] - Pr[X \geq i + 1])$$

$$= \sum_{i=1}^{\infty} (i \times Pr[X \geq i] - i \times Pr[X \geq i + 1])$$

$$= \sum_{i=1}^{\infty} i \times Pr[X \geq i] - \sum_{i=1}^{\infty} i \times Pr[X \geq i + 1]$$

$$= \sum_{i=1}^{\infty} i \times Pr[X \geq i] - \sum_{i=1}^{\infty} (i - 1) \times Pr[X \geq i] = \sum_{i=1}^{\infty} Pr[X \geq i].$$
**Theorem:** For a r.v. $X$ that takes values in $\{0, 1, 2, \ldots\}$, one has

$$E[X] = \sum_{i=1}^{\infty} Pr[X \geq i].$$

Probability mass at $i$, counted $i$ times.

... Same as $\sum_{i=1}^{\infty} i \times Pr[X = i]$. 

| $Pr[X \geq 1]$ | $Pr[X \geq 2]$ | $Pr[X \geq 3]$ | $\vdots$ |
Variance of geometric distribution.

$X$ is a geometrically distributed RV with parameter $p$. 

Thus, $\Pr[X = n] = (1 - p)^{n-1}p$ for $n \geq 1$.

Recall $E[X] = 1/p$.

$E[X^2] = p + 4p(1 - p) + 9p(1 - p)^2 + \ldots$ 

$pE[X^2] = p + 3p(1 - p) + 5p(1 - p)^2 + \ldots = 2(p + 2p(1 - p) + 3p(1 - p)^2 + \ldots)$ 


$\sigma(X) = \sqrt{1 - p/p^2} \approx E[X]$ when $p$ is small(ish).
Variance of geometric distribution.

$X$ is a geometrically distributed RV with parameter $p$. Thus, $Pr[X = n] = (1 - p)^{n-1} p$ for $n \geq 1$. 

$\sigma(X) = \sqrt{1 - p} p \approx E[X]$ when $p$ is small(ish).
Variance of geometric distribution.

$X$ is a geometrically distributed RV with parameter $p$. Thus, $Pr[X = n] = (1 - p)^{n-1}p$ for $n \geq 1$. Recall $E[X] = 1/p$. 

$E[X^2] = p + 4p(1 - p) + 9p(1 - p)^2 + ... = 1^2$.

$pE[X^2] = p + 4p(1 - p) + 9p(1 - p)^2 + ... = 2$.

$E[X^2] = 2 - (p + 4p(1 - p) + 9p(1 - p)^2 + ...) = 2 - (1/p)$.

$Var(X) = E[X^2] - E[X]^2 = 1 - 1/p^2$.

$\sigma(X) = \sqrt{1 - 1/p^2}$ when $p$ is small(ish).
Variance of geometric distribution.

$X$ is a geometrically distributed RV with parameter $p$. Thus, $Pr[X = n] = (1 - p)^{n-1}p$ for $n \geq 1$. Recall $E[X] = 1/p$.

\[
E[X^2] = p + 4p(1 - p) + 9p(1 - p)^2 + ... \]
Variance of geometric distribution.

$X$ is a geometrically distributed RV with parameter $p$. Thus, $Pr[X = n] = (1 - p)^{n-1}p$ for $n \geq 1$. Recall $E[X] = 1/p$.

\[
E[X^2] = p + 4p(1 - p) + 9p(1 - p)^2 + ... \\
-(1 - p)E[X^2] = -[p(1 - p) + 4p(1 - p)^2 + ...]
\]
Variance of geometric distribution.

$X$ is a geometrically distributed RV with parameter $p$. Thus, $Pr[X = n] = (1 - p)^{n-1}p$ for $n \geq 1$. Recall $E[X] = 1/p$.

\[
E[X^2] = p + 4p(1 - p) + 9p(1 - p)^2 + \ldots
\]
\[
-(1 - p)E[X^2] = -[p(1 - p) + 4p(1 - p)^2 + \ldots]
\]
\[
pE[X^2] = p + 3p(1 - p) + 5p(1 - p)^2 + \ldots
\]

$\sigma(X) = \sqrt{1 - p/p} \approx E[X]$ when $p$ is small(ish).
Variance of geometric distribution.

\( X \) is a geometrically distributed RV with parameter \( p \).
Thus, \( Pr[X = n] = (1 - p)^{n-1} p \) for \( n \geq 1 \). Recall \( E[X] = 1/p \).

\[
E[X^2] = p + 4p(1 - p) + 9p(1 - p)^2 + ... \\
-(1 - p)E[X^2] = -[p(1 - p) + 4p(1 - p)^2 + ...] \\
pE[X^2] = p + 3p(1 - p) + 5p(1 - p)^2 + ... \\
= 2(p + 2p(1 - p) + 3p(1 - p)^2 + ...) \quad E[X]! \\
-(p + p(1 - p) + p(1 - p)^2 + ...) \quad 1.
\]
Variance of geometric distribution.

$X$ is a geometrically distributed RV with parameter $p$. Thus, $Pr[X = n] = (1 - p)^{n-1} p$ for $n \geq 1$. Recall $E[X] = 1/p$.

\[
E[X^2] = p + 4p(1 - p) + 9p(1 - p)^2 + ... \\
-(1 - p)E[X^2] = -[p(1 - p) + 4p(1 - p)^2 + ...] \\
pE[X^2] = p + 3p(1 - p) + 5p(1 - p)^2 + ... \\
= 2(p + 2p(1 - p) + 3p(1 - p)^2 + ..) \quad E[X]^2 \\
-(p + p(1 - p) + p(1 - p)^2 + ...) \quad 1.
\]
Variance of geometric distribution.

$X$ is a geometrically distributed RV with parameter $p$. Thus, $\Pr[X = n] = (1 - p)^{n-1}p$ for $n \geq 1$. Recall $E[X] = 1/p$.

\[
E[X^2] = p + 4p(1 - p) + 9p(1 - p)^2 + ... \\
-(1 - p)E[X^2] = -[p(1 - p) + 4p(1 - p)^2 + ...] \\
pE[X^2] = p + 3p(1 - p) + 5p(1 - p)^2 + ... \\
= 2(p + 2p(1 - p) + 3p(1 - p)^2 + ..) \quad E[X]! \\
-(p + p(1 - p) + p(1 - p)^2 + ...) \quad 1.
\]
Variance of geometric distribution.

$X$ is a geometrically distributed RV with parameter $p$. Thus, $Pr[X = n] = (1 - p)^{n-1}p$ for $n \geq 1$. Recall $E[X] = 1/p$.

\[
E[X^2] = p + 4p(1 - p) + 9p(1 - p)^2 + ... \\
-(1 - p)E[X^2] = -[p(1 - p) + 4p(1 - p)^2 + ...] \\
pE[X^2] = p + 3p(1 - p) + 5p(1 - p)^2 + ... \\
= 2(p + 2p(1 - p) + 3p(1 - p)^2 + ..) \quad E[X]! \\
-(p + p(1 - p) + p(1 - p)^2 + ...) \quad 1. \\
pE[X^2] = 2E[X] - 1 \\
\]
Variance of geometric distribution.

$X$ is a geometrically distributed RV with parameter $p$. Thus, $Pr[X = n] = (1 - p)^{n-1}p$ for $n \geq 1$. Recall $E[X] = 1/p$.

\[
E[X^2] = p + 4p(1 - p) + 9p(1 - p)^2 + ... \\
-(1 - p)E[X^2] = -[p(1 - p) + 4p(1 - p)^2 + ...] \\
pE[X^2] = p + 3p(1 - p) + 5p(1 - p)^2 + ... \\
\quad = 2(p + 2p(1 - p) + 3p(1 - p)^2 + ..) \quad E[X]! \\
\quad = -2(p + p(1 - p) + p(1 - p)^2 + ...) \quad 1. \\
pE[X^2] = 2E[X] - 1 \\
\quad = 2\left(\frac{1}{p}\right) - 1
\]
Variance of geometric distribution.

Let $X$ be a geometrically distributed random variable with parameter $p$. Thus, $\Pr[X = n] = (1 - p)^{n-1} p$ for $n \geq 1$. Recall $E[X] = 1/p$.

\[
E[X^2] = p + 4p(1 - p) + 9p(1 - p)^2 + \ldots
\]

\[-(1 - p)E[X^2] = -[p(1 - p) + 4p(1 - p)^2 + \ldots]
\]

\[pE[X^2] = p + 3p(1 - p) + 5p(1 - p)^2 + \ldots
\]

\[= 2(p + 2p(1 - p) + 3p(1 - p)^2 + \ldots) - (p + p(1 - p) + p(1 - p)^2 + \ldots) = E[X]!
\]

\[ = 2E[X] - 1
\]

\[= 2\left(\frac{1}{p}\right) - 1 = \frac{2 - p}{p}
\]
Variance of geometric distribution.

$X$ is a geometrically distributed RV with parameter $p$. Thus, $\Pr[X = n] = (1 - p)^{n-1}p$ for $n \geq 1$. Recall $E[X] = 1/p$.

\[
\begin{align*}
E[X^2] &= p + 4p(1 - p) + 9p(1 - p)^2 + \ldots \\
-(1 - p)E[X^2] &= -[p(1 - p) + 4p(1 - p)^2 + \ldots] \\
pE[X^2] &= p + 3p(1 - p) + 5p(1 - p)^2 + \ldots \\
&= 2(p + 2p(1 - p) + 3p(1 - p)^2 + \ldots) \quad E[X]! \\
&= 2(p + p(1 - p) + p(1 - p)^2 + \ldots) \quad 1. \\
pE[X^2] &= 2E[X] - 1 \\
&= 2\left(\frac{1}{p}\right) - 1 = \frac{2 - p}{p}
\end{align*}
\]

\[\implies E[X^2] = \frac{(2 - p)}{p^2}\]
Variance of geometric distribution.

$X$ is a geometrically distributed RV with parameter $p$. Thus, $Pr[X = n] = (1 - p)^{n-1}p$ for $n \geq 1$. Recall $E[X] = 1/p$.

\[
E[X^2] = p + 4p(1 - p) + 9p(1 - p)^2 + \ldots
\]
\[-(1 - p)E[X^2] = -[p(1 - p) + 4p(1 - p)^2 + \ldots]
\]
\[pE[X^2] = p + 3p(1 - p) + 5p(1 - p)^2 + \ldots
\]
\[= 2(p + 2p(1 - p) + 3p(1 - p)^2 + \ldots) = 2E[X] - 1
\]
\[\implies E[X^2] = (2 - p)/p^2 \text{ and } var[X] = E[X^2] - E[X]^2
\]
Variances of geometric distribution.

$X$ is a geometrically distributed RV with parameter $p$. Thus, $Pr[X = n] = (1 - p)^{n-1}p$ for $n \geq 1$. Recall $E[X] = 1/p$.

$$E[X^2] = p + 4p(1 - p) + 9p(1 - p)^2 + ...$$

$$-(1 - p)E[X^2] = -[p(1 - p) + 4p(1 - p)^2 + ...]$$

$$pE[X^2] = p + 3p(1 - p) + 5p(1 - p)^2 + ...$$

$$= 2(p + 2p(1 - p) + 3p(1 - p)^2 + ...) \quad E[X]!$$

$$= 2(p + p(1 - p) + p(1 - p)^2 + ...) \quad 1.$$  

$$pE[X^2] = 2E[X] - 1$$

$$= 2\left(\frac{1}{p}\right) - 1 = \frac{2 - p}{p}$$

$\Longrightarrow E[X^2] = (2 - p)/p^2$ and $\text{var}[X] = E[X^2] - E[X]^2 = \frac{2-p}{p^2} - \frac{1}{p^2}$.
Variance of geometric distribution.

$X$ is a geometrically distributed RV with parameter $p$. Thus, $Pr[X = n] = (1 - p)^{n-1}p$ for $n \geq 1$. Recall $E[X] = 1/p$.

\[
E[X^2] = p + 4p(1 - p) + 9p(1 - p)^2 + \ldots
\]
\[
-(1 - p)E[X^2] = -[p(1 - p) + 4p(1 - p)^2 + \ldots]
\]
\[
pE[X^2] = p + 3p(1 - p) + 5p(1 - p)^2 + \ldots
\]
\[
= 2(p + 2p(1 - p) + 3p(1 - p)^2 + \ldots) = 2E[X]!
\]
\[
-(p + p(1 - p) + p(1 - p)^2 + \ldots) = 1.
\]
\[
pE[X^2] = 2E[X] - 1
\]
\[
= 2\left(\frac{1}{p}\right) - 1 = \frac{2 - p}{p}
\]

\[
\Rightarrow E[X^2] = (2 - p)/p^2 \text{ and } var[X] = E[X^2] - E[X]^2 = \frac{2-p}{p^2} - \frac{1}{p^2} = \frac{1-p}{p^2}.
\]
\[
\sigma(X) = \frac{\sqrt{1-p}}{p}
\]
Variance of geometric distribution.

$X$ is a geometrically distributed RV with parameter $p$. Thus, $Pr[X = n] = (1 - p)^{n-1}p$ for $n \geq 1$. Recall $E[X] = 1/p$.

$$E[X^2] = p + 4p(1 - p) + 9p(1 - p)^2 + ...$$

$$-(1 - p)E[X^2] = -[p(1 - p) + 4p(1 - p)^2 + ...]$$

$$pE[X^2] = p + 3p(1 - p) + 5p(1 - p)^2 + ...$$

$$= 2(p + 2p(1 - p) + 3p(1 - p)^2 + ..) \ E[X]!$$

$$-(p + p(1 - p) + p(1 - p)^2 + ...) \ 1.$$  

$$pE[X^2] = 2E[X] - 1$$

$$= 2\left(\frac{1}{p}\right) - 1 = \frac{2 - p}{p}$$

$$\implies E[X^2] = \frac{(2 - p)}{p^2} \text{ and}$$

$$var[X] = E[X^2] - E[X]^2 = \frac{2-p}{p^2} - \frac{1}{p^2} = \frac{1-p}{p^2}.$$  

$$\sigma(X) = \frac{\sqrt{1-p}}{p} \approx E[X] \text{ when } p \text{ is small(ish)}. $$
Variance of geometric distribution.

$X$ is a geometrically distributed RV with parameter $p$. Thus, $Pr[X = n] = (1 - p)^{n-1}p$ for $n \geq 1$. Recall $E[X] = 1/p$.

\[
E[X^2] = p + 4p(1 - p) + 9p(1 - p)^2 + \ldots \\
-(1 - p)E[X^2] = -[p(1 - p) + 4p(1 - p)^2 + \ldots] \\
pE[X^2] = p + 3p(1 - p) + 5p(1 - p)^2 + \ldots \\
= 2(p + 2p(1 - p) + 3p(1 - p)^2 + \ldots) \\
= 2E[X] + p(1 - p) + p(1 - p)^2 + \ldots \\
1.
\]

\[
pE[X^2] = 2E[X] - 1 \\
= 2\left(\frac{1}{p}\right) - 1 = \frac{2 - p}{p}
\]

$\Rightarrow E[X^2] = (2 - p)/p^2$ and $var[X] = E[X^2] - E[X]^2 = \frac{2-p}{p^2} - \frac{1}{p^2} = \frac{1-p}{p^2}$.

$\sigma(X) = \frac{\sqrt{1-p}}{p} \approx E[X]$ when $p$ is small(ish).
Bern \( (p) \):

\[
\begin{align*}
\text{Pr} \left[ X = 1 \right] &= p; \\
E[X] &= p; \\
\text{Var}[X] &= p \left( 1 - p \right); \\
\end{align*}
\]

Bin \( (n, p) \):

\[
\begin{align*}
\text{Pr} \left[ X = m \right] &= \binom{n}{m} p^m \left( 1 - p \right)^{n - m}, \quad m = 0, \ldots, n; \\
E[X] &= np; \\
\text{Var}[X] &= np \left( 1 - p \right); \\
\end{align*}
\]

\( U[1, \ldots, n] \):

\[
\begin{align*}
\text{Pr} \left[ X = m \right] &= \frac{1}{n}, \quad m = 1, \ldots, n; \\
E[X] &= \frac{n + 1}{2}; \\
\text{Var}[X] &= \frac{n^2 - 1}{12}; \\
\end{align*}
\]

Geom \( (p) \):

\[
\begin{align*}
\text{Pr} \left[ X = n \right] &= \left( 1 - p \right)^{n-1} p, \quad n = 1, 2, \ldots; \\
E[X] &= \frac{1}{p}; \\
\text{Var}[X] &= \frac{1}{p^2}; \\
\end{align*}
\]
Review: Distributions

- $Bern(p)$: $Pr[X = 1] = p$;
Review: Distributions

- $Bern(p) : Pr[X = 1] = p$;
  $E[X] = p$;
  $Var[X] = p(1 - p)$;
Review: Distributions

- **Bern**\((p)\): \(\Pr[X = 1] = p\);
  \(E[X] = p\);
  \(Var[X] = p(1 - p)\);

- **Bin**\((n, p)\): \(\Pr[X = m] = \binom{n}{m} p^m (1 - p)^{n-m}, m = 0, \ldots, n\);
Review: Distributions

- **Bern(p)**: $Pr[X = 1] = p$;  
  $E[X] = p$;  
  $Var[X] = p(1 - p)$;

- **Bin(n, p)**: $Pr[X = m] = \binom{n}{m} p^m (1 - p)^{n-m}, m = 0, \ldots, n$;  
  $E[X] = np$;  
  $Var[X] = np(1 - p)$;
Review: Distributions

- **Bern\((p)\):** \(Pr[X = 1] = p;\)
  \(E[X] = p;\)
  \(Var[X] = p(1 − p);\)

- **Bin\((n, p)\):** \(Pr[X = m] = \binom{n}{m} p^m (1 − p)^{n−m}, m = 0, \ldots, n;\)
  \(E[X] = np;\)
  \(Var[X] = np(1 − p);\)

- **U\([1, \ldots, n]\):** \(Pr[X = m] = \frac{1}{n}, m = 1, \ldots, n;\)
Review: Distributions

- **Bern(p)**: \( \Pr[X = 1] = p \);
  \[ E[X] = p \]
  \[ \text{Var}[X] = p(1 - p) \]

- **Bin(n, p)**: \( \Pr[X = m] = \binom{n}{m} p^m (1 - p)^{n-m}, m = 0, \ldots, n \);
  \[ E[X] = np \]
  \[ \text{Var}[X] = np(1 - p) \]

- **U[1, \ldots, n]**: \( \Pr[X = m] = \frac{1}{n}, m = 1, \ldots, n \);
  \[ E[X] = \frac{n+1}{2} \]
  \[ \text{Var}[X] = \frac{n^2 - 1}{12} \]
Review: Distributions

- **Bern(p)**: $Pr[X = 1] = p$;
  
  $E[X] = p$;
  
  $Var[X] = p(1 - p)$;

- **Bin(n, p)**: $Pr[X = m] = \binom{n}{m} p^m (1 - p)^{n-m}, m = 0, \ldots, n$;
  
  $E[X] = np$;
  
  $Var[X] = np(1 - p)$;

- **U[1, ..., n]**: $Pr[X = m] = \frac{1}{n}, m = 1, \ldots, n$;
  
  $E[X] = \frac{n+1}{2}$;
  
  $Var[X] = \frac{n^2-1}{12}$;

- **Geom(p)**: $Pr[X = n] = $
Review: Distributions

- **Bern**($p$) : $\Pr[X = 1] = p$;
  $E[X] = p$;
  $\text{Var}[X] = p(1 - p)$;

- **Bin**($n$, $p$) : $\Pr[X = m] = \binom{n}{m} p^m (1 - p)^{n-m}$, $m = 0, \ldots, n$;
  $E[X] = np$;
  $\text{Var}[X] = np(1 - p)$;

- **$U[1, \ldots, n]$** : $\Pr[X = m] = \frac{1}{n}$, $m = 1, \ldots, n$;
  $E[X] = \frac{n+1}{2}$;
  $\text{Var}[X] = \frac{n^2-1}{12}$;

- **Geom**($p$) : $\Pr[X = n] = (1 - p)^{n-1} p$, $n = 1, 2, \ldots$;
Review: Distributions

- **Bern**$(p)$ : \( Pr[X = 1] = p; \)
  \[ E[X] = p; \]
  \[ Var[X] = p(1 - p); \]

- **Bin**$(n, p)$ : \( Pr[X = m] = \binom{n}{m} p^m (1 - p)^{n-m}, m = 0, \ldots, n; \)
  \[ E[X] = np; \]
  \[ Var[X] = np(1 - p); \]

- **U**$[1, \ldots, n]$ : \( Pr[X = m] = \frac{1}{n}, m = 1, \ldots, n; \)
  \[ E[X] = \frac{n+1}{2}; \]
  \[ Var[X] = \frac{n^2-1}{12}; \]

- **Geom**$(p)$ : \( Pr[X = n] = (1 - p)^{n-1} p, n = 1, 2, \ldots; \)
  \[ E[X] = \frac{1}{p}; \]
  \[ Var[X] = \frac{1-p}{p^2}; \]
Review: Distributions

- **Bern($p$):** $Pr[X = 1] = p$;
  $E[X] = p$;
  $Var[X] = p(1 - p)$;

- **Bin($n, p$):** $Pr[X = m] = \binom{n}{m} p^m (1 - p)^{n-m}, m = 0, \ldots, n$;
  $E[X] = np$;
  $Var[X] = np(1 - p)$;

- **U[$1, \ldots, n$]:** $Pr[X = m] = \frac{1}{n}, m = 1, \ldots, n$;
  $E[X] = \frac{n+1}{2}$;
  $Var[X] = \frac{n^2 - 1}{12}$;

- **Geom($p$):** $Pr[X = n] = (1 - p)^{n-1} p, n = 1, 2, \ldots$;
  $E[X] = \frac{1}{p}$;
  $Var[X] = \frac{1-p}{p^2}$;
Today’s gig: Two envelopes problem.
Today’s gig: Two envelopes problem.

Gigs so far:
1. How to tell random from human.
2. Monty Hall.
5. Simpson’s paradox.

Today:
Gigs so far:

1. How to tell random from human.
2. Monty Hall.
5. Simpson’s paradox.

Today: Two envelopes problem.
Two envelopes

I put $x$ dollars in an envelope, and $2x$ dollars in another envelope, and seal both envelopes.

You pick one at random (you don't know which).

Before you open it you think:

What will happen if I switch?

Well, if I picked the one I picked has $y$ dollars, then the other either $2y$ or $y/2$.

In the first case, I win $y$.

In the second case, I lose $y/2$.

Therefore, in expectation, my net gain is:

$$\frac{1}{2}y - \frac{1}{2}y/2 = \frac{y}{2}.$$

Therefore, I should switch.

Before you open the new envelope you think:

What will happen if I switch?
Two envelopes

I put $x$ dollars in an envelope, and $2x$ dollars in another envelope, and seal both envelopes.

You pick one at random (you don’t know which is which).
I put $x$ dollars in an envelope, and $2x$ dollars in another envelope, and seal both envelopes.

You pick one at random (you don’t know which is which).

Before you open it you think:
Two envelopes

I put $x$ dollars in an envelope, and $2x$ dollars in another envelope, and seal both envelopes.

You pick one at random (you don’t know which is which).

Before you open it you think: What will happen if I switch?
I put $x$ dollars in an envelope, and $2x$ dollars in another envelope, and seal both envelopes.

You pick one at random (you don’t know which is which).

Before you open it you think: What will happen if I switch?

Well, if I picked the one I picked has $y$ dollars, then the other either $2y$ or $\frac{y}{2}$. 
I put $x$ dollars in an envelope, and $2x$ dollars in another envelope, and seal both envelopes.

You pick one at random (you don’t know which is which).

Before you open it you think: What will happen if I switch?

Well, if I picked the one I picked has $y$ dollars, then the other either $2y$ or $\frac{y}{2}$.

In the first case, I win $y$.
I put $x$ dollars in an envelope, and $2x$ dollars in another envelope, and seal both envelopes.

You pick one at random (you don’t know which is which).

Before you open it you think: What will happen if I switch?

Well, if I picked the one I picked has $y$ dollars, then the other either $2y$ or $\frac{y}{2}$.

In the first case, I win $y$. In the second case, I lose $\frac{y}{2}$. 

Therefore, I should switch.
Two envelopes

I put $x$ dollars in an envelope, and $2x$ dollars in another envelope, and seal both envelopes.

You pick one at random (you don’t know which is which).

Before you open it you think: What will happen if I switch?

Well, if I picked the one I picked has $y$ dollars, then the other either $2y$ or $\frac{y}{2}$.

In the first case, I win $y$. In the second case, I lose $\frac{y}{2}$.

Therefore, in expectation, my net gain is:
Two envelopes

I put $x$ dollars in an envelope, and $2x$ dollars in another envelope, and seal both envelopes.

You pick one at random (you don’t know which is which).

Before you open it you think: What will happen if I switch?

Well, if I picked the one I picked has $y$ dollars, then the other either $2y$ or $\frac{y}{2}$.

In the first case, I win $y$. In the second case, I lose $\frac{y}{2}$.

Therefore, in expectation, my net gain is: $\frac{1}{2}y - \frac{1}{2} \frac{y}{2} = \frac{y}{2}$.
Two envelopes

I put $x$ dollars in an envelope, and $2x$ dollars in another envelope, and seal both envelopes.

You pick one at random (you don’t know which is which).

Before you open it you think: What will happen if I switch?

Well, if I picked the one I picked has $y$ dollars, then the other either $2y$ or $\frac{y}{2}$.

In the first case, I win $y$. In the second case, I lose $\frac{y}{2}$.

Therefore, in expectation, my net gain is: $\frac{1}{2}y - \frac{1}{2} \cdot \frac{y}{2} = \frac{y}{2}$.

Therefore, I should switch.
Two envelopes

I put $x$ dollars in an envelope, and $2x$ dollars in another envelope, and seal both envelopes.

You pick one at random (you don’t know which is which).

Before you open it you think: What will happen if I switch?

Well, if I picked the one I picked has $y$ dollars, then the other either $2y$ or $\frac{y}{2}$.

In the first case, I win $y$. In the second case, I lose $\frac{y}{2}$.

Therefore, in expectation, my net gain is: $\frac{1}{2}y - \frac{1}{2} \frac{y}{2} = \frac{y}{2}$.

Therefore, I should switch.

Before you open the new envelope you think:
I put $x$ dollars in an envelope, and $2x$ dollars in another envelope, and seal both envelopes.

You pick one at random (you don’t know which is which).

Before you open it you think: What will happen if I switch?

Well, if I picked the one I picked has $y$ dollars, then the other either $2y$ or $\frac{y}{2}$.

In the first case, I win $y$. In the second case, I lose $\frac{y}{2}$.

Therefore, in expectation, my net gain is: $\frac{1}{2}y - \frac{1}{2} \frac{y}{2} = \frac{y}{2}$.

Therefore, I should switch.

Before you open the new envelope you think: What will happen if I switch?
Summary

Random Variables
Summary

Random Variables

- Variance.
- Distributions.