Coupon Collecting: Fun with harmonic numbers!
Coupon Collecting: Fun with harmonic numbers!
Memoryless Property.
Coupon Collecting: Fun with harmonic numbers!
Memoryless Property.
Law of the unconscious statistician. (Hmm.)
Coupon Collecting: Fun with harmonic numbers!
Memoryless Property.
Law of the unconscious statistician. (Hmmm.)
Variance/ Covariance.
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: \( X_1 = 1 \).
Time to collect coupons

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

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

$Pr[\text{“get second coupon”|“got milk ”}]$
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 coupon after getting first.

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

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

$E[X_2]$?
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 coupon after getting first.

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

\(E[X_2]\)? 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 coupon after getting first.

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

$E[X_2]$? 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 coupon after getting first.

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

$E[X_2]$? 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 coupon after getting first.

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

$E[X_2]$? 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 coupon after getting first.

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

$E[X_2]$? 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 coupon after getting first.

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

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

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

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

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

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

$Pr[\text{“getting $i$th coupon”}|\text{“got $i-1$rst 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 coupon after getting first.

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

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

$Pr[\text{“getting } i \text{th coupon|“got } i-1 \text{rst coupons”}] = \frac{n-(i-1)}{n} = \frac{n-i+1}{n}$

$E[X_i]$
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 coupon after getting first.

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

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

$Pr[\text{"getting $i$th coupon"|"got $i-1$rst 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 coupon after getting first.

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

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

$Pr[\text{“getting } i \text{th coupon”} | \text{“got } i-1\text{rst coupons”}] = \frac{n-(i-1)}{n} = \frac{n-i+1}{n}$

$E[X_i] = \frac{1}{p} = \frac{n}{n-i+1}$,
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 coupon after getting first.

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

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

$Pr[\text{“getting $i$th coupon”}|\text{“got $i-1$rst coupons”}] = \frac{n-(i-1)}{n} = \frac{n-i+1}{n}$

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

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

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

$Pr[\text{“getting $i$th coupon”}|\text{“got $i-1$rst 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 coupon after getting first.

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

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

$Pr[\text{“getting $i$th coupon|“got $i-1$rst 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 coupon after getting first.

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

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

$Pr[\text{“getting } i \text{th coupon|“got } (i-1) \text{rst coupons”}] = \frac{n-(i-1)}{n} = \frac{n-i+1}{n}$

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

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

$$= n(1 + \frac{1}{2} + \cdots + \frac{1}{n}) =: nH(n)$$
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 coupon after getting first.

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

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

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

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

\[
E[X] = E[X_1] + \cdots + E[X_n] = \frac{n}{n} + \frac{n}{n-1} + \frac{n}{n-2} + \cdots + \frac{n}{1} \\
= n\left(1 + \frac{1}{2} + \cdots + \frac{1}{n}\right) =: nH(n) \approx n(\ln n + \gamma)
\]
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

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

A good approximation is

\[ H(n) \approx \ln(n) + \gamma \]

where \( \gamma \approx 0.58 \) (Euler-Mascheroni constant).
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

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

par-a-dox
\[\text{ˈpərəˌdāks}/\]
noun

A statement or proposition that, despite sound (or apparently sound) reasoning from acceptable premises, leads to a conclusion that seems senseless, logically unacceptable, or self-contradictory.
"A potentially serious conflict between quantum mechanics and the general theory of relativity known as the information paradox"

- A seemingly absurd or self-contradictory statement or proposition that when investigated or explained may prove to be well founded or true.
  "In a paradox, he has discovered that stepping back from his job has increased the rewards he gleans from it"

  synonyms: contradiction, contradiction in terms, self-contradiction, inconsistency, incongruity; More

- A situation, person, or thing that combines contradictory features or qualities.
  "The mingling of deciduous trees with elements of desert flora forms a fascinating ecological paradox"
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.

Video.

\[
\begin{align*}
nx &= 1 - x \\
\Rightarrow x &= 1/(n + 1)
\end{align*}
\]
The cards have width 2.
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.

Video.
Calculating $E[g(X)]$: LOTUS
Calculating $E[g(X)]$: LOTUS

Let $Y = g(X)$. 
Calculating $E[g(X)]:$ LOTUS

Let $Y = g(X)$. Assume that we know the distribution of $X$. 

\[ \text{Method 1: We calculate the distribution of } Y. \]

\[ \text{Method 2: We use the following result. Called "Law of the unconscious statistician."} \]

\[ \text{Theorem: } E[g(X)] = \sum x g(x) \Pr[X = x]. \]

\[ \text{Proof: } E[g(X)] = \sum \omega g(X(\omega)) \Pr[\omega] = \sum x \sum \omega \in g^{-1}(x) g(x) \Pr[\omega] = \sum x \sum \omega \in g^{-1}(x) g(x) \Pr[\omega] = \sum x g(x) \Pr[X = x]. \]
Calculating $E[g(X)]:$ LOTUS

Let $Y = g(X).$ Assume that we know the distribution of $X.$

We want to calculate $E[Y].$
Calculating $E[g(X)]:$ LOTUS

Let $Y = g(X)$. Assume that we know the distribution of $X$.

We want to calculate $E[Y]$.

**Method 1:**
Calculating $E[g(X)]:$ LOTUS

Let $Y = g(X)$. Assume that we know the distribution of $X$.

We want to calculate $E[Y]$.

**Method 1:** We calculate the distribution of $Y$:
Calculating $E[g(X)]:$ LOTUS

Let $Y = g(X).$ Assume that we know the distribution of $X.$

We want to calculate $E[Y].$

**Method 1:** We calculate the distribution of $Y$: 

$$Pr[Y = y] = Pr[X \in g^{-1}(y)] \text{ where } g^{-1}(x) = \{x \in \mathbb{R} : g(x) = y\}.$$
Calculating $E[g(X)]:$ LOTUS

Let $Y = g(X).$ Assume that we know the distribution of $X.$

We want to calculate $E[Y].$

**Method 1:** We calculate the distribution of $Y:$

$$Pr[Y = y] = Pr[X \in g^{-1}(y)]$$

where $g^{-1}(x) = \{x \in \mathbb{R} : g(x) = y\}.$

This is typically rather tedious!
Calculating $E[g(X)]$: LOTUS

Let $Y = g(X)$. Assume that we know the distribution of $X$. We want to calculate $E[Y]$.

**Method 1:** We calculate the distribution of $Y$:

$$Pr[Y = y] = Pr[X \in g^{-1}(y)] \text{ where } g^{-1}(x) = \{x \in \mathbb{R} : g(x) = y\}.$$  

This is typically rather tedious!

**Method 2:** We use the following result.
Calculating $E[g(X)]:$ LOTUS

Let $Y = g(X)$. Assume that we know the distribution of $X$.

We want to calculate $E[Y]$.

**Method 1:** We calculate the distribution of $Y$:

$$Pr[Y = y] = Pr[X \in g^{-1}(y)]$$

where $g^{-1}(x) = \{ x \in \mathbb{R} : g(x) = y \}$.

This is typically rather tedious!

**Method 2:** We use the following result.

Called “Law of the unconscious statistician.”

**Theorem:**

$$E[g(X)] = \sum_x g(x) Pr[X = x].$$
Calculating $E[g(X)]:$ LOTUS

Let $Y = g(X)$. Assume that we know the distribution of $X$.

We want to calculate $E[Y]$.

**Method 1:** We calculate the distribution of $Y$:

$$Pr[Y = y] = Pr[X \in g^{-1}(y)] \text{ where } g^{-1}(x) = \{x \in \mathbb{R} : g(x) = y\}.$$ 

This is typically rather tedious!

**Method 2:** We use the following result.

Called “Law of the unconscious statistician.”

**Theorem:**

$$E[g(X)] = \sum_x g(x)Pr[X = x].$$

**Proof:**
Calculating $E[g(X)]:$ LOTUS

Let $Y = g(X)$. Assume that we know the distribution of $X$.

We want to calculate $E[Y]$.

Method 1: We calculate the distribution of $Y$:

$$Pr[Y = y] = Pr[X \in g^{-1}(y)]$$

where $g^{-1}(x) = \{x \in \mathbb{R} : g(x) = y\}$.

This is typically rather tedious!

Method 2: We use the following result.

Called “Law of the unconscious statistician.”

Theorem:

$$E[g(X)] = \sum_x g(x)Pr[X = x].$$

Proof:

$$E[g(X)] = \sum_\omega g(X(\omega))Pr[\omega]$$
Calculating $E[g(X)]$: LOTUS

Let $Y = g(X)$. Assume that we know the distribution of $X$. We want to calculate $E[Y]$.

**Method 1:** We calculate the distribution of $Y$:

$$Pr[Y = y] = Pr[X ∈ g^{-1}(y)]$$

where $g^{-1}(x) = \{x ∈ \mathbb{R} : g(x) = y\}$. This is typically rather tedious!

**Method 2:** We use the following result.

Called “Law of the unconscious statistician.”

**Theorem:**

$$E[g(X)] = \sum_x g(x)Pr[X = x].$$

**Proof:**

$$E[g(X)] = \sum_\omega g(X(\omega))Pr[\omega] = \sum_x \sum_{\omega ∈ X^{-1}(x)} g(X(\omega))Pr[\omega]$$
Calculating $E[g(X)]:$ LOTUS

Let $Y = g(X)$. Assume that we know the distribution of $X$.

We want to calculate $E[Y]$.

**Method 1:** We calculate the distribution of $Y$:

$$Pr[Y = y] = Pr[X \in g^{-1}(y)]$$

where $g^{-1}(x) = \{x \in \mathbb{R} : g(x) = y\}$.

This is typically rather tedious!

**Method 2:** We use the following result.

**Called “Law of the unconscious statistician.”**

**Theorem:**

$$E[g(X)] = \sum_x g(x) Pr[X = x].$$

**Proof:**

$$E[g(X)] = \sum_\omega g(X(\omega)) Pr[\omega] = \sum_x \sum_{\omega \in X^{-1}(x)} g(X(\omega)) Pr[\omega]$$

$$= \sum_x \sum_{\omega \in X^{-1}(x)} g(x) Pr[\omega]$$
Calculating $E[g(X)]:$ LOTUS

Let $Y = g(X)$. Assume that we know the distribution of $X$.

We want to calculate $E[Y]$.

**Method 1:** We calculate the distribution of $Y$:

$$ Pr[Y = y] = Pr[X \in g^{-1}(y)] \text{ where } g^{-1}(x) = \{x \in \mathcal{R} : g(x) = y\}.$$  

This is typically rather tedious!

**Method 2:** We use the following result.

Called “Law of the unconscious statistician.”

**Theorem:**

$$ E[g(X)] = \sum_x g(x) Pr[X = x]. $$

**Proof:**

$$ E[g(X)] = \sum_{\omega} g(X(\omega)) Pr[\omega] = \sum_x \sum_{\omega \in X^{-1}(x)} g(X(\omega)) Pr[\omega] $$

$$ = \sum_x \sum_{\omega \in X^{-1}(x)} g(x) Pr[\omega] = \sum_x g(x) \sum_{\omega \in X^{-1}(x)} Pr[\omega] $$
Calculating $E[g(X)]$: LOTUS

Let $Y = g(X)$. Assume that we know the distribution of $X$.

We want to calculate $E[Y]$.

**Method 1:** We calculate the distribution of $Y$:

$$Pr[Y = y] = Pr[X \in g^{-1}(y)] \text{ where } g^{-1}(x) = \{x \in \mathbb{R} : g(x) = y\}.$$

This is typically rather tedious!

**Method 2:** We use the following result.

Called “Law of the unconscious statistician.”

**Theorem:**

$$E[g(X)] = \sum_x g(x)Pr[X = x].$$

**Proof:**

$$E[g(X)] = \sum_\omega g(X(\omega))Pr[\omega] = \sum_x \sum_{\omega \in X^{-1}(x)} g(X(\omega))Pr[\omega]$$

$$= \sum_x \sum_{\omega \in X^{-1}(x)} g(x)Pr[\omega] = \sum_x g(x) \sum_{\omega \in X^{-1}(x)} Pr[\omega]$$

$$= \sum_x g(x)Pr[X = x].$$
Poll.

Which is LOTUS?
Poll.

Which is LOTUS?

(A) \( E[X] = \sum_{x \in \text{Range}(X)} g(x) \Pr[g(X) = g(x)] \)
(B) \( E[X] = \sum_{x \in \text{Range}(X)} g(x) \Pr[X = x] \)
(C) \( E[X] = \sum_{x \in \text{Range}(g)} x \Pr[g(X) = x] \)
Geometric Distribution.

Geometric Distribution.

Experiment: flip a coin with heads prob. $p$. until Heads.
Random Variable $X$: number of flips.
And distribution is:
Geometric Distribution.

Experiment: flip a coin with heads prob. $p$. until Heads.
Random Variable $X$: number of flips.

And distribution is:

(A) $X \sim G(p) : Pr[X = i] = (1 - p)^{i-1}p$.
(B) $X \sim B(p, n) : Pr[X = i] = \binom{n}{i} p^i (1 - p)^{n-i}$.
Geometric Distribution.


And distribution is:

(A) $X \sim G(p)$: $Pr[X = i] = (1 - p)^{i-1}p$.
(B) $X \sim B(p, n)$: $Pr[X = i] = \binom{n}{i} p^i (1 - p)^{n-i}$.

(A) Distribution of $X \sim G(p)$: $Pr[X = i] = (1 - p)^{i-1}p$. 
Geometric Distribution: Memoryless - Interpretation

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

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

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

\[ Pr[X > n + m | X > n] = Pr[A|B] = Pr[A'] = Pr[X > m]. \]
**Geometric Distribution: Memoryless - Interpretation**

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

\[
Pr[X > n + m \mid X > n] = Pr[A \mid B] = Pr[A'] = Pr[X > m].
\]

\(A'\): is \(m\) coin tosses before heads.
\(A \mid B\): \(m\) ’more’ coin tosses before heads.
Geometric Distribution: Memoryless - Interpretation

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

Pr[X > n + m | X > n] = Pr[A|B] = Pr[A′] = Pr[X > m].

A’: is m coin tosses before heads.
A|B: m ‘more’ coin tosses before heads.

The coin is memoryless, therefore, so is X.

Independent coin: Pr[H ′|anyprevioussetofcointosses′] = p
Geometric Distribution: Memoryless by derivation.

Let $X$ be $G(p)$. Then, for $n \geq 0$,
Geometric Distribution: Memoryless by derivation.

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

$$Pr[X > n] = Pr[ \text{first } n \text{ flips are } T] =$$
Geometric Distribution: Memoryless by derivation.

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

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

Let $X$ be $G(p)$. 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 by derivation.

Let $X$ be $G(p)$. 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:

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

Let \( X \) be \( G(p) \). 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 by derivation.

Let \( X \) be \( G(p) \). 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 by derivation.

Let $X$ be $G(p)$. 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 by derivation.

Let $X$ be $G(p)$. 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 by derivation.

Let $X$ be $G(p)$. 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:**

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

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

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

$$= Pr[X > m].$$
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 variance values. The curve on the left has Var = 1, and the curve on the right has Var = 10.]}
The variance measures the deviation from the mean value.
Variance

The variance measures the deviation from the mean value.

**Definition:** The variance of $X$ is
The variance measures the deviation from the mean value.

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

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

Fact:

\[ \text{var}[X] = E[X^2] - E[X]^2. \]
Variance and Standard Deviation

Fact:


Indeed:

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

Fact:

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

Indeed:

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

Fact:

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

Indeed:

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

**Fact:**

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

**Indeed:**

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

Fact:


Indeed:

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

$$= E[X^2 - 2XE[X] + E[X]^2]$$

$$= E[X^2] - 2E[X]E[X] + E[X]^2, \text{ by linearity}$$

A simple example

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

This example illustrates the term ‘standard deviation.’

\[Pr = 0.5 \quad \mu - \sigma \quad \mu \quad \mu + \sigma \quad Pr = 0.5\]
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 $(X - E[X])^2 = \sigma^2$. 
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 $(X - E[X])^2 = \sigma^2$. Hence,

$$
\text{var}(X) = \sigma^2 \text{ and } \sigma(X) = \sigma.
$$
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

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

Consider $X$ with

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

Then

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

Consider $X$ with

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

Then

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

Consider $X$ with

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

Then

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

Also,

$$E(|X|) = 1 \times 0.99 + 99 \times 0.01 = 1.98.$$
Example

Consider $X$ with

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

Then

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

Also,

$$E(|X|) = 1 \times 0.99 + 99 \times 0.01 = 1.98.$$ 

Thus, $\sigma(X) = \sqrt{E[(X - E(X))^2]} \neq E[|X - E[X]|]$!
Example

Consider $X$ with

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

Then

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

Also,

$$E(|X|) = 1 \times 0.99 + 99 \times 0.01 = 1.98.$$  

Thus, $\sigma(X) = \sqrt{E[(X - E(X))^2]} \neq E[|X - E[X]|]$!

Exercise: How big can you make $\frac{\sigma(X)}{E[|X - E[X]|]}$?
Assume that $Pr[X = i] = \frac{1}{n}$ for $i \in \{1, \ldots, n\}$. Then
Uniform

Assume that $Pr[X = i] = \frac{1}{n}$ for $i \in \{1, \ldots, n\}$. Then

$$E[X] = \sum_{i=1}^{n} i \times Pr[X = i] = \frac{1}{n} \sum_{i=1}^{n} i$$
Uniform

Assume that $Pr[X = i] = \frac{1}{n}$ for $i \in \{1, \ldots, n\}$. Then

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

$$= \frac{1}{n} \frac{n(n+1)}{2} = \frac{n+1}{2}.$$
Uniform

Assume that \( Pr[X = i] = \frac{1}{n} \) for \( i \in \{1, \ldots, n\} \). Then

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

\[
= \frac{1}{n} \frac{n(n+1)}{2} = \frac{n+1}{2}.
\]

Also,

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

Assume that $Pr[X = i] = \frac{1}{n}$ for $i \in \{1, \ldots, n\}$. Then

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

$$= \frac{1}{n} \frac{n(n+1)}{2} = \frac{n+1}{2}.$$

Also,

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

$$= \frac{1}{n} \frac{(n)(n+1)(n+2)}{6} = \frac{1+3n+2n^2}{6},$$

as you can verify.
Uniform

Assume that $Pr[X = i] = \frac{1}{n}$ for $i \in \{1, \ldots, n\}$. Then

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

$$= \frac{1}{n} \frac{n(n+1)}{2} = \frac{n+1}{2}.$$

Also,

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

$$= \frac{1}{n} \frac{(n)(n+1)(n+2)}{6} = \frac{1+3n+2n^2}{6},$$

as you can verify.
Uniform

Assume that $Pr[X = i] = \frac{1}{n}$ for $i \in \{1, \ldots, n\}$. Then

\[
E[X] = \sum_{i=1}^{n} i \times Pr[X = i] = \frac{1}{n} \sum_{i=1}^{n} i = \frac{1}{n} \frac{n(n+1)}{2} = \frac{n+1}{2}.
\]

Also,

\[
E[X^2] = \sum_{i=1}^{n} i^2 Pr[X = i] = \frac{1}{n} \sum_{i=1}^{n} i^2 = \frac{1}{n} \frac{(n)(n+1)(n+2)}{6} = \frac{1 + 3n + 2n^2}{6}, \text{ as you can verify.}
\]

This gives

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

(Sort of $\int_{0}^{1/2} x^2 dx = \frac{x^3}{3}$.)
Variance of geometric distribution.

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

$\sigma(X) = \sqrt{1 - p} p \approx E[X]$ 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 $n \geq 1$. 

$E[X] = \frac{1}{p}$.

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

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

$\sigma(X) = \sqrt{1 - p \frac{p}{2^2}} \approx E[X]$ 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 $n \geq 1$. Recall $E[X] = 1/p$. 

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

$E[X^2] = \frac{p}{2} + \frac{2p(1 - p)}{3} + \frac{3p(1 - p)^2}{5} + \ldots$

$\sigma(X) = \sqrt{1 - p/p} \approx E[X] = 1/p$ 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 $E[X] = 1/p$.

\[
E[X^2] = p + 4p(1 - p) + 9p(1 - p)^2 + ... \\
-(1 - p)E[X^2] = -[p(1 - p) + 4p(1 - p)^2 + ...] \\
pE[X^2] = p + 3p(1 - p) + 5p(1 - p)^2 + ...
\]
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 $E[X] = 1/p$.

\[
E[X^2] = p + 4p(1 - p) + 9p(1 - p)^2 + ... \\
-(1 - p)E[X^2] = -[p(1 - p) + 4p(1 - p)^2 + ...] \\
pE[X^2] = p + 3p(1 - p) + 5p(1 - p)^2 + ... \\
= 2(p + 2p(1 - p) + 3p(1 - p)^2 + ..) \quad E[X]^2! \\
-(p + p(1 - p) + p(1 - p)^2 + ...) \quad Distribution.
\]
**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 $E[X] = 1/p$.

\[ E[X^2] = p + 4p(1 - p) + 9p(1 - p)^2 + \ldots \]
\[-(1 - p)E[X^2] = -[p(1 - p) + 4p(1 - p)^2 + \ldots] \]
\[ pE[X^2] = p + 3p(1 - p) + 5p(1 - p)^2 + \ldots \]
\[ = 2(p + 2p(1 - p) + 3p(1 - p)^2 + \ldots) \quad E[X]^2 \]
\[-(p + p(1 - p) + p(1 - p)^2 + \ldots) \quad \text{Distribution.} \]
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 $E[X] = 1/p$.

\[
\begin{align*}
E[X^2] &= p + 4p(1 - p) + 9p(1 - p)^2 + 
\quad \cdots \\
-(1 - p)E[X^2] &= -[p(1 - p) + 4p(1 - p)^2 + 
\quad \cdots ] \\
pE[X^2] &= p + 3p(1 - p) + 5p(1 - p)^2 + 
\quad \cdots \\
&= 2(p + 2p(1 - p) + 3p(1 - p)^2 + \cdots) \quad E[X]^2! \\
&= (p + p(1 - p) + p(1 - p)^2 + \cdots) \quad \text{Distribution.}
\end{align*}
\]
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 \(E[X] = 1/p\).

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

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

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

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

\[
-(p + p(1 - p) + p(1 - p)^2 + \ldots) \quad \text{Distribution.}
\]

\[
pE[X^2] = 2E[X] - 1
\]
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 $E[X] = 1/p$.

\[
\begin{align*}
E[X^2] &= p + 4p(1 - p) + 9p(1 - p)^2 + ... \\
-(1 - p)E[X^2] &= -[p(1 - p) + 4p(1 - p)^2 + ...] \\
pE[X^2] &= p + 3p(1 - p) + 5p(1 - p)^2 + ... \\
&= 2(p + 2p(1 - p) + 3p(1 - p)^2 + ...) \quad E[X]! \\
&= -(p + p(1 - p) + p(1 - p)^2 + ...) \quad \text{Distribution.} \\
pE[X^2] &= 2E[X] - 1 \\
&= 2\left(\frac{1}{p}\right) - 1
\end{align*}
\]
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 $E[X] = 1/p$.

\[
E[X^2] = p + 4p(1 - p) + 9p(1 - p)^2 + \ldots
\]
\[
-(1 - p)E[X^2] = -[p(1 - p) + 4p(1 - p)^2 + \ldots]
\]
\[
pE[X^2] = p + 3p(1 - p) + 5p(1 - p)^2 + \ldots
\]
\[
= 2(p + 2p(1 - p) + 3p(1 - p)^2 + \ldots) \quad E[X]!
\]
\[
-(p + p(1 - p) + p(1 - p)^2 + \ldots) \quad \text{Distribution.}
\]
\[
pE[X^2] = 2E[X] - 1
\]
\[
= 2\left(\frac{1}{p}\right) - 1 = \frac{2 - p}{p}
\]
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 $E[X] = 1/p$.

\[
\begin{align*}
E[X^2] &= p + 4p(1 - p) + 9p(1 - p)^2 + \ldots \\
-(1 - p)E[X^2] &= -[p(1 - p) + 4p(1 - p)^2 + \ldots] \\
pE[X^2] &= p + 3p(1 - p) + 5p(1 - p)^2 + \ldots \\
&= 2(p + 2p(1 - p) + 3p(1 - p)^2 + \ldots) \quad E[X]! \\
&= (p + p(1 - p) + p(1 - p)^2 + \ldots) \quad Distribution. \\
pE[X^2] &= 2E[X] - 1 \\
&= 2\left(\frac{1}{p}\right) - 1 = \frac{2 - p}{p} \\
\implies E[X^2] &= (2 - p)/p^2
\end{align*}
\]
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 $E[X] = 1/p$.

\[
\begin{align*}
E[X^2] & = p + 4p(1 - p) + 9p(1 - p)^2 + ... \\
-(1 - p)E[X^2] & = -[p(1 - p) + 4p(1 - p)^2 + ...] \\
pE[X^2] & = p + 3p(1 - p) + 5p(1 - p)^2 + ... \\
& = 2(p + 2p(1 - p) + 3p(1 - p)^2 + ..) \quad E[X]! \\
& -(p + p(1 - p) + p(1 - p)^2 + ...) \quad \text{Distribution.} \\
pE[X^2] & = 2E[X] - 1 \\
& = 2\left(\frac{1}{p}\right) - 1 = \frac{2 - p}{p}
\end{align*}
\]

$\implies E[X^2] = (2 - p)/p^2$ and $\text{var}[X] = E[X^2] - E[X]^2$
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 $E[X] = 1/p$.

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

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

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

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

\[
- (p + p(1 - p) + p(1 - p)^2 + ...) Distribution.
\]

\[
pE[X^2] = 2E[X] - 1
\]

\[
= 2\left(\frac{1}{p}\right) - 1 = \frac{2 - p}{p}
\]

\[
\implies E[X^2] = \frac{(2 - p)}{p^2} \text{ and var}[X] = E[X^2] - E[X]^2 = \frac{2-p}{p^2} - \frac{1}{p^2}
\]
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 $E[X] = 1/p$.

\[
\begin{align*}
E[X^2] &= p + 4p(1 - p) + 9p(1 - p)^2 + \ldots \\
-(1 - p)E[X^2] &= -[p(1 - p) + 4p(1 - p)^2 + \ldots] \\
pE[X^2] &= p + 3p(1 - p) + 5p(1 - p)^2 + \ldots \\
&= 2(p + 2p(1 - p) + 3p(1 - p)^2 + \ldots) \\
&= -p + p(1 - p) + p(1 - p)^2 + \ldots \\
pE[X^2] &= 2E[X] - 1 \\
&= 2\left(\frac{1}{p}\right) - 1 = \frac{2 - p}{p} \\
\end{align*}
\]

\[\implies E[X^2] = \frac{2 - p}{p^2} \text{ and }\]
\[\text{var}[X] = E[X^2] - E[X]^2 = \frac{2 - p}{p^2} - \frac{1}{p^2} = \frac{1 - p}{p^2}.\]

$\sigma(X) = \frac{\sqrt{1 - p}}{p}$
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 \(E[X] = 1/p\).

\[
\begin{align*}
E[X^2] &= p + 4p(1 - p) + 9p(1 - p)^2 + ... \\
-(1 - p)E[X^2] &= -[p(1 - p) + 4p(1 - p)^2 + ...] \\
pE[X^2] &= p + 3p(1 - p) + 5p(1 - p)^2 + ... \\
&= 2(p + 2p(1 - p) + 3p(1 - p)^2 + ...) \quad E[X]! \\
&= -(p + p(1 - p) + p(1 - p)^2 + ...) \quad \text{Distribution.}
\end{align*}
\]

\[
\begin{align*}
pE[X^2] &= 2E[X] - 1 \\
&= 2\left(\frac{1}{p}\right) - 1 = \frac{2 - p}{p}
\end{align*}
\]

\[\implies E[X^2] = \frac{(2 - p)}{p^2} \text{ and}\]

\[\text{var}[X] = E[X^2] - E[X]^2 = \frac{2-p}{p^2} - \frac{1}{p^2} = \frac{1-p}{p^2}.
\]

\[\sigma(X) = \frac{\sqrt{1-p}}{p} \approx E[X] \text{ when } p \text{ is small(ish)}.\]
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 $E[X] = 1/p$.

\[
\begin{align*}
E[X^2] &= p + 4p(1 - p) + 9p(1 - p)^2 + \ldots \\
-(1 - p)E[X^2] &= -[p(1 - p) + 4p(1 - p)^2 + \ldots] \\
pE[X^2] &= p + 3p(1 - p) + 5p(1 - p)^2 + \ldots \\
&= 2(p + 2p(1 - p) + 3p(1 - p)^2 + \ldots)E[X]! \\
&= (p + p(1 - p) + p(1 - p)^2 + \ldots)E[X]! \\
pE[X^2] &= 2E[X] - 1 \\
&= 2\left(\frac{1}{p}\right) - 1 = \frac{2 - p}{p}
\end{align*}
\]

\[
\Rightarrow E[X^2] = \frac{2 - p}{p^2} \quad \text{and} \quad \text{var}[X] = E[X^2] - E[X]^2 = \frac{2 - p}{p^2} - \frac{1}{p^2} = \frac{1 - p}{p^2}.
\]

$\sigma(X) = \frac{\sqrt{1 - p}}{p} \approx E[X]$ when $p$ is small(ish).
Fixed points.

Number of fixed points in a random permutation of $n$ items.
Fixed points.

Number of fixed points in a random permutation of $n$ items. “Number of students that get homework back.”
Fixed points.

Number of fixed points in a random permutation of $n$ items. “Number of student that get homework back.”

$$X = X_1 + X_2 \cdots + X_n$$
Fixed points.

Number of fixed points in a random permutation of $n$ items. “Number of student that get homework back.”

$X = X_1 + X_2 \cdots + X_n$

where $X_i$ is indicator variable for $i$th student getting hw back.
Fixed points.

Number of fixed points in a random permutation of $n$ items. “Number of student that get homework back.”

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

where $X_i$ is indicator variable for $i$th student getting hw back.

$$E(X^2) = \sum_i E(X_i^2) + \sum_{i\neq j} E(X_iX_j).$$

$$= +$$
Fixed points.

Number of fixed points in a random permutation of \( n \) items. “Number of student that get homework back.”

\[ X = X_1 + X_2 \cdots + X_n \]

where \( X_i \) is indicator variable for \( i \)th student getting hw back.

\[
E(X^2) = \sum_i E(X_i^2) + \sum_{i \neq j} E(X_i X_j).
\]

\[
= +
\]

\[
E(X_i^2) = 1 \times Pr[X_i = 1] + 0 \times Pr[X_i = 0]
\]
Fixed points.

Number of fixed points in a random permutation of \( n \) items. “Number of student that get homework back.”

\[ X = X_1 + X_2 \cdots + X_n \]

where \( X_i \) is indicator variable for \( i \)th student getting hw back.

\[
E(X^2) = \sum_i E(X_i^2) + \sum_{i \neq j} E(X_i X_j).
\]

\[
E(X_i^2) = 1 \times Pr[X_i = 1] + 0 \times Pr[X_i = 0] = \frac{1}{n}
\]
Fixed points.

Number of fixed points in a random permutation of $n$ items.
“Number of student that get homework back.”

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

where $X_i$ is indicator variable for $i$th student getting hw back.

$$E(X^2) = \sum_i E(X_i^2) + \sum_{i \neq j} E(X_i X_j).$$

$$= \quad +$$

$$E(X_i^2) = 1 \times Pr[X_i = 1] + 0 \times Pr[X_i = 0]$$

$$= \frac{1}{n}$$
Fixed points.

Number of fixed points in a random permutation of $n$ items.
“Number of student that get homework back.”

\[ X = X_1 + X_2 \cdots + X_n \]

where $X_i$ is indicator variable for $i$th student getting hw back.

\[
E(X^2) = \sum_i E(X_i^2) + \sum_{i \neq j} E(X_i X_j).
\]

\[
= n \times \frac{1}{n} + \frac{1}{n}.
\]

\[
E(X_i^2) = 1 \times Pr[X_i = 1] + 0 \times Pr[X_i = 0]
= \frac{1}{n}.
\]
Fixed points.

Number of fixed points in a random permutation of \( n \) items. “Number of student that get homework back.”

\[
X = X_1 + X_2 \cdots + X_n
\]

where \( X_i \) is indicator variable for \( i \)th student getting hw back.

\[
E(X^2) = \sum_i E(X_i^2) + \sum_{i \neq j} E(X_i X_j).
\]

\[
= n \times \frac{1}{n} +
\]

\[
E(X_i^2) = 1 \times Pr[X_i = 1] + 0 \times Pr[X_i = 0]
\]

\[
= \frac{1}{n}
\]

\[
E(X_i X_j) = 1 \times Pr[X_i = 1 \cap X_j = 1] + 0 \times Pr[“anything else’”]
\]
Fixed points.

Number of fixed points in a random permutation of $n$ items. “Number of student that get homework back.”

$X = X_1 + X_2 \cdots + X_n$

where $X_i$ is indicator variable for $i$th student getting hw back.

$$E(X^2) = \sum_i E(X_i^2) + \sum_{i \neq j} E(X_i X_j).$$

$$= n \times \frac{1}{n} +$$

$$E(X_i^2) = 1 \times Pr[X_i = 1] + 0 \times Pr[X_i = 0]$$

$$= \frac{1}{n}$$

$$E(X_i X_j) = 1 \times Pr[X_i = 1 \cap X_j = 1] + 0 \times Pr[\text{“anything else”}]$$

$$= 1 \times \frac{(n-2)!}{n!}$$
Fixed points.

Number of fixed points in a random permutation of \( n \) items. “Number of student that get homework back.”

\[ X = X_1 + X_2 \cdots + X_n \]

where \( X_i \) is indicator variable for \( i \)th student getting hw back.

\[
E(X^2) = \sum_i E(X_i^2) + \sum_{i \neq j} E(X_iX_j).
\]

\[
= n \times \frac{1}{n} + \]

\[
E(X_i^2) = 1 \times Pr[X_i = 1] + 0 \times Pr[X_i = 0]
\]

\[
= \frac{1}{n}
\]

\[
E(X_iX_j) = 1 \times Pr[X_i = 1 \cap X_j = 1] + 0 \times Pr[\text{“anything else”}]
\]

\[
= 1 \times \frac{(n-2)!}{n!} = \frac{1}{n(n-1)}
\]
Fixed points.

Number of fixed points in a random permutation of $n$ items. “Number of student that get homework back.”

$X = X_1 + X_2 \cdots + X_n$

where $X_i$ is indicator variable for $i$th student getting hw back.

$$E(X^2) = \sum_i E(X_i^2) + \sum_{i \neq j} E(X_iX_j).$$

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

$$E(X_i^2) = 1 \times Pr[X_i = 1] + 0 \times Pr[X_i = 0]$$

$$= \frac{1}{n}$$

$$E(X_iX_j) = 1 \times Pr[X_i = 1 \cap X_j = 1] + 0 \times Pr[\text{“anything else”}]$$

$$= 1 \times \frac{(n-2)!}{n!} = \frac{1}{n(n-1)}$$
Fixed points.

Number of fixed points in a random permutation of $n$ items. “Number of student that get homework back.”

$X = X_1 + X_2 \cdots + X_n$

where $X_i$ is indicator variable for $i$th student getting hw back.

$$E(X^2) = \sum_i E(X_i^2) + \sum_{i \neq j} E(X_iX_j).$$

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

$$= \quad 1 + 1 = 2.$$

$E(X_i^2) = 1 \times Pr[X_i = 1] + 0 \times Pr[X_i = 0]$

$$= \quad \frac{1}{n}$$

$E(X_iX_j) = 1 \times Pr[X_i = 1 \cap X_j = 1] + 0 \times Pr[“anything else’’]$  

$$= \quad 1 \times \frac{(n-2)!}{n!} = \frac{1}{n(n-1)}$
Fixed points.

Number of fixed points in a random permutation of \( n \) items. “Number of student that get homework back.”

\[ X = X_1 + X_2 \cdots + X_n \]

where \( X_i \) is indicator variable for \( i \)th student getting hw back.

\[
E(X^2) = \sum_i E(X_i^2) + \sum_{i \neq j} E(X_i X_j).
\]

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

\[
= 1 + 1 = 2.
\]

\[
E(X_i^2) = 1 \times Pr[X_i = 1] + 0 \times Pr[X_i = 0]
\]

\[
= \frac{1}{n}
\]

\[
E(X_i X_j) = 1 \times Pr[X_i = 1 \cap X_j = 1] + 0 \times Pr[“anything else”]
\]

\[
= 1 \times \frac{(n-2)!}{n!} = \frac{1}{n(n-1)}
\]

\[
Var(X) = E(X^2) - (E(X))^2 = 2 - 1 = 1.
\]
Poll: fixed points.

What's true?

(A) \(X_i\) and \(X_j\) are independent.
(B) \(E[X_iX_j] = Pr[X_iX_j = 1]\)
(C) \(Pr[X_iX_j] = \frac{(n-2)!}{n!}\)
(D) \(X_i^2 = X_i\).
Poll: fixed points.

What’s true?

(A) $X_i$ and $X_j$ are independent.
(B) $E[X_i X_j] = Pr[X_i X_j = 1]$
(C) $Pr[X_i X_j] = \frac{(n-2)!}{n!}$
(D) $X_i^2 = X_i$. 
Variance: binomial.

\[ E[X^2] = \sum_{i=0}^{n} i^2 \binom{n}{i} p^i (1-p)^{n-i}. \]
Variance: binomial.

\[
E[X^2] = \sum_{i=0}^{n} i^2 \binom{n}{i} p^i (1-p)^{n-i}.
\]
Variance: binomial.

\[ E[X^2] = \sum_{i=0}^{n} i^2 \binom{n}{i} p^i (1-p)^{n-i}. \]

= Really??!!#... 

Too hard!
Variance: binomial.

\[ E[X^2] = \sum_{i=0}^{n} i^2 \binom{n}{i} p^i (1 - p)^{n-i}. \]

= Really???!!##...

Too hard!

Ok..
Variance: binomial.

\[ E[X^2] = \sum_{i=0}^{n} i^2 \binom{n}{i} p^i (1-p)^{n-i}. \]

= Really???!!#...  

Too hard!  
Ok.. fine.
Variance: binomial.

\[ E[X^2] = \sum_{i=0}^{n} i^2 \binom{n}{i} p^i (1 - p)^{n-i}. \]

\[ = \text{Really???!!##...} \]

Too hard!

Ok.. fine.

Let’s do something else.
Variance: binomial.

\[ E[X^2] = \sum_{i=0}^{n} i^2 \binom{n}{i} p^i (1 - p)^{n-i}. \]

= Really??!!#... 

Too hard!
Ok.. fine.
Let’s do something else.
Maybe not much easier...
Variance: binomial.

\[
E[X^2] = \sum_{i=0}^{n} i^2 \binom{n}{i} p^i (1 - p)^{n-i}.
\]

= Really??!!#... 

Too hard!
Ok.. fine.
Let’s do something else.
Maybe not much easier...but there is a payoff.
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. \( Var(cX) = c^2 \ Var(X) \), where \( c \) is a constant. Scales by \( c^2 \).

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

Proof:

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

\[
\text{Var}(cX) = E((cX)^2) - (E(cX))^2 \\
= c^2 E(X^2) - c^2 (E(X))^2 \\
= c^2 (E(X^2) - (E(X))^2) \\
= c^2 \text{Var}(X)
\]
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:

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

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

\[
\text{Var}(X + c) = E((X + c - E(X + c))^2) \\
= E((X + c - E(X) - c)^2)
\]
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:

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

\[
\text{Var}(X + c) = E((X + c - E(X + c))^2) \\
= E((X + c - E(X) - c)^2) \\
= E((X - E(X))^2)
\]
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:

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

\[
\text{Var}(X + c) = E((X + c - E(X + c))^2) \\
= E((X + c - E(X) - c)^2) \\
= E((X - E(X))^2) = \text{Var}(X)
\]
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:**

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

\[
Var(X + c) = E((X + c - E(X + c))^2) = E((X + c - E(X) - c)^2) = E(\color{red}{(X - E(X))^2}) = Var(X)
\]
Independent random variables.

Independent: $P[X = a, Y = b] = Pr[X = a]Pr[Y = b]$
Independent random variables.

Independent: $P[X = a, Y = b] = Pr[X = a]Pr[Y = b]$

Independent random variables.

Independent: \( P[X = a, Y = b] = Pr[X = a]Pr[Y = b] \)


\[
E[XY] = \sum_a \sum_b a \times b \times Pr[X = a, Y = b] \\
= \sum_a \sum_b a \times b \times Pr[X = a]Pr[Y = b] \\
= (\sum_a aPr[X = a])(\sum_b bPr[Y = b]) \\
= E[X]E[Y]
\]
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,
\[ E(XY) = E(X)E(Y) = 0. \]
Hence,
\[ \text{var}(X + Y) = E((X + Y)^2) = E(X^2 + 2XY + Y^2) = E(X^2) + 2E(XY) + E(Y^2) = E(X^2) + E(Y^2) = \text{var}(X) + \text{var}(Y). \]
Variance of sum of two independent random variables

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

$$Var(X + Y) = Var(X) + Var(Y).$$
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.
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 \( E(Y) = 0 \).
**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,

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

$$\text{var}(X + Y) = E((X + Y)^2).$$
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,

$$var(X + Y) = E((X + Y)^2) = E(X^2 + 2XY + Y^2)$$
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,

$$\text{var}(X + Y) = E((X + Y)^2) = E(X^2 + 2XY + Y^2)$$

$$= E(X^2) + 2E(XY) + E(Y^2)$$
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,

$$var(X + Y) = E((X + Y)^2) = E(X^2 + 2XY + Y^2)$$

$$= E(X^2) + 2E(XY) + E(Y^2) = E(X^2) + E(Y^2).$$
**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,

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

Hence,

$$var(X + Y) = E((X + Y)^2) = E(X^2 + 2XY + Y^2)$$

$$= E(X^2) + 2E(XY) + E(Y^2) = E(X^2) + E(Y^2)$$

$$= var(X) + var(Y).$$
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,
\[ E[XY] = E[X]E[Y] = 0. \]
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

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

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

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,

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

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,


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,

\[
\text{var}(X + Y + Z + \cdots) = E((X + Y + Z + \cdots)^2)
\]
\[
= E(X^2 + Y^2 + Z^2 + \cdots + 2XY + 2XZ + 2YZ + \cdots)
\]
\[
= E(X^2) + E(Y^2) + E(Z^2) + \cdots + 0 + \cdots + 0
\]
\[
= \text{var}(X) + \text{var}(Y) + \text{var}(Z) + \cdots.
\]
Variance of Binomial Distribution.

Flip coin with heads probability $p$. 
Variance of Binomial Distribution.

Flip coin with heads probability $p$. 
$X$ - how many heads?
Variance of Binomial Distribution.

Flip coin with heads probability \( p \).
\( X \)- how many heads?

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

Flip coin with heads probability $p$. $X$ - how many heads?

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

$$E(X_i^2)$$
Variance of Binomial Distribution.

Flip coin with heads probability $p$. $X$: how many heads?

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

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

Flip coin with heads probability $p$. $X$ - how many heads?

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

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

Flip coin with heads probability $p$. $X$ - how many heads?

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

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

$$Var(X_i) = p - (E(X))^2$$
Variance of Binomial Distribution.

Flip coin with heads probability $p$. $X$- how many heads?

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

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

Flip coin with heads probability $p$. $X$- how many heads?

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

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

Flip coin with heads probability $p$. $X$ - how many heads?

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

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

$$Var(X_i) = p - (E(X))^2 = p - p^2 = p(1 - p).$$

$p = 0$
Variance of Binomial Distribution.

Flip coin with heads probability $p$. $X$- how many heads?

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

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

Flip coin with heads probability \( p \).
\( X \)- how many heads?

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

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

\[
Var(X_i) = p - (E(X))^2 = p - p^2 = p(1 - p).
\]

\[ p = 0 \implies Var(X_i) = 0 \]

\[ p = 1\]
Variance of Binomial Distribution.

Flip coin with heads probability $p$.

$X$ - how many heads?

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

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

$$Var(X_i) = p - (E(X))^2 = p - p^2 = p(1 - p).$$

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

$p = 1 \implies Var(X_i) = 0$
Variance of Binomial Distribution.

Flip coin with heads probability $p$. $X$ - how many heads?

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

$$E(X_i^2) = 1^2 \times p + 0^2 \times (1 - p) = p.$$  
$$Var(X_i) = p - (E(X))^2 = p - p^2 = p(1 - p).$$

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

$p = 1 \implies Var(X_i) = 0$
Variance of Binomial Distribution.

Flip a coin with heads probability $p$. How many heads? 

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

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

$$Var(X_i) = p - (E(X))^2 = p - p^2 = p(1 - p).$$

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

$p = 1 \implies Var(X_i) = 0$ 

$X = X_1 + X_2 + \ldots + X_n.$
Variance of Binomial Distribution.

Flip coin with heads probability $p$. 
$X$- how many heads?

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

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

$p = 0 \implies \text{Var}(X) = 0$
$p = 1 \implies \text{Var}(X) = 0$

$X = X_1 + X_2 + \ldots X_n$.

$X_i$ and $X_j$ are independent:
Variance of Binomial Distribution.

Flip coin with heads probability $p$. 
$X$- how many heads?

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

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

$p = 0 \implies \text{Var}(X_i) = 0$
$p = 1 \implies \text{Var}(X_i) = 0$

$X = X_1 + X_2 + \ldots X_n.$

$X_i$ and $X_j$ are independent: $Pr[X_i = 1|X_j = 1] = Pr[X_i = 1].$
Variance of Binomial Distribution.

Flip coin with heads probability $p$. $X$- how many heads?

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

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

$$Var(X_i) = p - (E(X))^2 = p - p^2 = p(1 - p).$$

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

$p = 1 \implies Var(X_i) = 0$

$X = X_1 + X_2 + \ldots X_n.$

$X_i$ and $X_j$ are independent: $Pr[X_i = 1|X_j = 1] = Pr[X_i = 1].$

$$Var(X) = Var(X_1 + \cdots X_n)$$
Variance of Binomial Distribution.

Flip coin with heads probability $p$. $X$- how many heads?

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

$$E(X_i^2) = 1^2 \times p + 0^2 \times (1 - p) = p.$$  
$$Var(X_i) = p - (E(X))^2 = p - p^2 = p(1 - p).$$

$p = 0 \implies Var(X_i) = 0$  
$p = 1 \implies Var(X_i) = 0$

$X = X_1 + X_2 + \ldots X_n$.

$X_i$ and $X_j$ are independent: $Pr[X_i = 1|X_j = 1] = Pr[X_i = 1]$.

$$Var(X) = Var(X_1 + \ldots + X_n) = np(1 - p).$$
Variance of Binomial Distribution.

Flip coin with heads probability \( p \).
\( X \)- how many heads?

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

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

\[
Var(X_i) = p - (E(X))^2 = p - p^2 = p(1 - p).
\]

\[ p = 0 \implies Var(X_i) = 0 \]

\[ p = 1 \implies Var(X_i) = 0 \]

\( X = X_1 + X_2 + \ldots X_n. \)

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

\[
Var(X) = Var(X_1 + \cdots X_n) = np(1 - p).
\]
Definition Poisson Distribution with parameter $\lambda > 0$

$$X = P(\lambda) \iff Pr[X = m] = \frac{\lambda^m}{m!} e^{-\lambda}, m \geq 0.$$
Poisson Distribution: Variance.

**Definition** Poisson Distribution with parameter \( \lambda > 0 \)

\[ X = P(\lambda) \iff Pr[X = m] = \frac{\lambda^m}{m!} e^{-\lambda}, m \geq 0. \]

Mean, Variance?
Poisson Distribution: Variance.

**Definition** Poisson Distribution with parameter $\lambda > 0$

$$X = P(\lambda) \iff Pr[X = m] = \frac{\lambda^m}{m!} e^{-\lambda}, m \geq 0.$$ 

Mean, Variance?

Ugh.
**Definition** Poisson Distribution with parameter $\lambda > 0$

$$X = P(\lambda) \iff \Pr[X = m] = \frac{\lambda^m}{m!} e^{-\lambda}, \ m \geq 0.$$ 

Mean, Variance?

Ugh.

Recall that Poisson is the limit of the Binomial with $p = \lambda/n$ as $n \to \infty$. 

\[ E(X^2) = \lambda + \frac{\lambda^2}{n} \]
Poisson Distribution: Variance.

**Definition** Poisson Distribution with parameter $\lambda > 0$

\[ X = P(\lambda) \iff \Pr[X = m] = \frac{\lambda^m}{m!} e^{-\lambda}, \ m \geq 0. \]

Mean, Variance?

Ugh.

Recall that Poisson is the limit of the Binomial with $p = \lambda / n$ as $n \to \infty$.

Mean: $pn = \lambda$
Definition Poisson Distribution with parameter $\lambda > 0$

$$X = P(\lambda) \iff Pr[X = m] = \frac{\lambda^m}{m!} e^{-\lambda}, m \geq 0.$$ 

Mean, Variance?

Ugh.

Recall that Poission is the limit of the Binomial with $p = \lambda/n$ as $n \to \infty$.

Mean: $pn = \lambda$

Variance: $p(1 - p)n = \lambda - \lambda^2/n \to \lambda$
Poisson Distribution: Variance.

**Definition** Poisson Distribution with parameter $\lambda > 0$

$$X = P(\lambda) \iff Pr[X = m] = \frac{\lambda^m}{m!} e^{-\lambda}, m \geq 0.$$ 

Mean, Variance?

Ugh.

Recall that Poisson is the limit of the Binomial with $p = \lambda / n$ as $n \to \infty$.

Mean: $pn = \lambda$

Variance: $p(1 - p)n = \lambda - \lambda^2 / n \to \lambda$. 
**Definition** Poisson Distribution with parameter $\lambda > 0$

\[ X = \mathcal{P}(\lambda) \iff \Pr[X = m] = \frac{\lambda^m}{m!} e^{-\lambda}, m \geq 0. \]

Mean, Variance?

Ugh.

Recall that Poission is the limit of the Binomial with $p = \lambda / n$ as $n \to \infty$.

Mean: $pn = \lambda$

Variance: $p(1 - p)n = \lambda - \lambda^2 / n \to \lambda$.

$E(X^2)$?
Poisson Distribution: Variance.

**Definition** Poisson Distribution with parameter $\lambda > 0$

$$X = P(\lambda) \iff Pr[X = m] = \frac{\lambda^m}{m!} e^{-\lambda}, m \geq 0.$$  

Mean, Variance?

Ugh.

Recall that Poission is the limit of the Binomial with $p = \lambda/n$ as $n \to \infty$.

Mean: $pn = \lambda$

Variance: $p(1-p)n = \lambda - \lambda^2/n \to \lambda$.

$E(X^2)$? $Var(X) = E(X^2) - (E(X))^2$ or $E(X^2) = Var(X) + E(X)^2$. 
Poisson Distribution: Variance.

**Definition** Poisson Distribution with parameter $\lambda > 0$

$$X = P(\lambda) \iff Pr[X = m] = \frac{\lambda^m}{m!} e^{-\lambda}, m \geq 0.$$  

Mean, Variance?

Ugh.

Recall that Poisson is the limit of the Binomial with $p = \lambda / n$ as $n \to \infty$.

Mean: $pn = \lambda$

Variance: $p(1 - p)n = \lambda - \lambda^2 / n \to \lambda.$

$E(X^2)$? $Var(X) = E(X^2) - (E(X))^2$ or $E(X^2) = Var(X) + E(X)^2$.

$E(X^2) = \lambda + \lambda^2$. 
Covariance

**Definition** The covariance of $X$ and $Y$ is

$$\text{cov}(X, Y) := E[(X - E[X])(Y - E[Y])].$$
Covariance

**Definition** The covariance of $X$ and $Y$ is

$$cov(X, Y) := E[(X - E[X])(Y - E[Y])].$$

**Fact**

$$cov(X, Y) = E[XY] - E[X]E[Y].$$
Covariance

**Definition** The covariance of $X$ and $Y$ is

$$\text{cov}(X, Y) := E[(X - E[X])(Y - E[Y])].$$

**Fact**

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

**Proof:**

Covariance

**Definition** The covariance of $X$ and $Y$ is

$$\text{cov}(X, Y) := E[(X - E[X])(Y - E[Y])].$$

**Fact**

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

**Proof:**
Think about $E[X] = E[Y] = 0$. Just $E[XY]$. \qed
Covariance

**Definition** The covariance of $X$ and $Y$ is

$$\text{cov}(X, Y) := E[(X - E[X])(Y - E[Y])].$$

**Fact**

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

**Proof:**
Covariance

**Definition** The covariance of $X$ and $Y$ is

$$cov(X, Y) := E[(X - E[X])(Y - E[Y])].$$

**Fact**

$$cov(X, Y) = E[XY] - E[X]E[Y].$$

**Proof:**
Think about $E[X] = E[Y] = 0$. Just $E[XY]$. For the sake of completeness.
Covariance

**Definition** The covariance of $X$ and $Y$ is

$$\text{cov}(X, Y) := E[(X - E[X])(Y - E[Y])].$$

**Fact**

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

**Proof:**

Think about $E[X] = E[Y] = 0$. Just $E[XY]$. For the sake of completeness.

\[
= E[XY] - E[X]E[Y].
\]
Correlation

**Definition** The correlation of $X, Y$, $Cor(X, Y)$ is

$$corr(X, Y) : \frac{cov(X, Y)}{\sigma(X)\sigma(Y)}.$$ 

**Theorem:** $-1 \leq corr(X, Y) \leq 1.$

**Proof:** Idea: $(a - b)^2 > 0$
**Definition** The correlation of $X, Y$, $Cor(X, Y)$ is

$$corr(X, Y) : \frac{cov(X, Y)}{\sigma(X)\sigma(Y)}.$$

**Theorem:** $-1 \leq corr(X, Y) \leq 1$.

**Proof:** Idea: $(a - b)^2 > 0 \rightarrow a^2 + b^2 \geq 2ab.$
Definition The correlation of $X, Y$, $Cor(X, Y)$ is

$$corr(X, Y) : \frac{cov(X, Y)}{\sigma(X)\sigma(Y)}.$$ 

Theorem: $-1 \leq corr(X, Y) \leq 1$.
Proof: Idea: $(a - b)^2 > 0 \rightarrow a^2 + b^2 \geq 2ab$.

Correlation

**Definition** The correlation of $X$, $Y$, $\text{Cor}(X, Y)$ is

$$
corr(X, Y) : \frac{\text{cov}(X, Y)}{\sigma(X)\sigma(Y)}.
$$

**Theorem:** $-1 \leq corr(X, Y) \leq 1.$

**Proof:** Idea: $(a - b)^2 > 0 \rightarrow a^2 + b^2 \geq 2ab.$


$\text{Cor}(X, Y) = E[XY]$. 
**Definition** The correlation of $X$, $Y$, $\text{Cor}(X, Y)$ is

$$\text{corr}(X, Y) = \frac{\text{cov}(X, Y)}{\sigma(X)\sigma(Y)}.$$ 

**Theorem:** $-1 \leq \text{corr}(X, Y) \leq 1.$

**Proof:** Idea: $(a - b)^2 > 0 \rightarrow a^2 + b^2 \geq 2ab.$


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

Definition The correlation of $X$, $Y$, $\text{Cor}(X, Y)$ is

$$\text{corr}(X, Y) : \frac{\text{cov}(X, Y)}{\sigma(X)\sigma(Y)}.$$ 

Theorem: $-1 \leq \text{corr}(X, Y) \leq 1$.

Proof: Idea: $(a - b)^2 > 0 \rightarrow a^2 + b^2 \geq 2ab$.


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

**Definition** The correlation of $X$, $Y$, $Cor(X, Y)$ is

\[
\text{corr}(X, Y): \frac{\text{cov}(X, Y)}{\sigma(X)\sigma(Y)}.
\]

**Theorem:** $-1 \leq \text{corr}(X, Y) \leq 1$.

**Proof:** Idea: $(a - b)^2 > 0 \rightarrow a^2 + b^2 \geq 2ab$.


$Cor(X, Y) = E[XY]$.

\[
\rightarrow E[XY] \leq 1.
\]

\[
E[(X + Y)^2] = E[X^2] + E[Y^2] + 2E[XY] = 2(1 + E[XY]) \geq 0
\]
**Definition**  The correlation of $X, Y$, $Cor(X, Y)$ is

$$corr(X, Y) = \frac{\text{cov}(X, Y)}{\sigma(X)\sigma(Y)}.$$ 

**Theorem:** $-1 \leq corr(X, Y) \leq 1$. 
**Proof:** Idea: $(a - b)^2 > 0 \rightarrow a^2 + b^2 \geq 2ab$. 


$Cor(X, Y) = E[XY]$. 

\[
\]

\[
\]

Shifting and scaling doesn’t change correlation.
Examples of Covariance

Note that $E[X] = 0$ and $E[Y] = 0$ in these examples. Then $\text{cov}(X, Y) = E[XY]$.

When $\text{cov}(X, Y) > 0$, the RVs $X$ and $Y$ tend to be large or small together. $X$ and $Y$ are said to be positively correlated.

When $\text{cov}(X, Y) < 0$, when $X$ is larger, $Y$ tends to be smaller. $X$ and $Y$ are said to be negatively correlated.

When $\text{cov}(X, Y) = 0$, we say that $X$ and $Y$ are uncorrelated.

Four equally likely pairs of values

$\text{cov}(X, Y) = 1/2$  
$\text{cov}(X, Y) = -1/2$  
$\text{cov}(X, Y) = 0$
Examples of Covariance

Note that $E[X] = 0$ and $E[Y] = 0$ in these examples. Then $\text{cov}(X, Y) = E[XY]$. 

Four equally likely pairs of values

$\text{cov}(X, Y) = 1/2$  $\text{cov}(X, Y) = -1/2$  $\text{cov}(X, Y) = 0$
Examples of Covariance

Note that \( E[X] = 0 \) and \( E[Y] = 0 \) in these examples. Then \( \text{cov}(X, Y) = E[XY] \).

When \( \text{cov}(X, Y) > 0 \), the RVs \( X \) and \( Y \) tend to be large or small together.

\[ \text{cov}(X, Y) = 1/2 \quad \text{cov}(X, Y) = -1/2 \quad \text{cov}(X, Y) = 0 \]
Examples of Covariance

Note that $E[X] = 0$ and $E[Y] = 0$ in these examples. Then $\text{cov}(X, Y) = E[XY]$.

When $\text{cov}(X, Y) > 0$, the RVs $X$ and $Y$ tend to be large or small together. $X$ and $Y$ are said to be positively correlated.
Examples of Covariance

Note that $E[X] = 0$ and $E[Y] = 0$ in these examples. Then $\text{cov}(X, Y) = E[XY]$.

When $\text{cov}(X, Y) > 0$, the RVs $X$ and $Y$ tend to be large or small together. $X$ and $Y$ are said to be positively correlated.

When $\text{cov}(X, Y) < 0$, when $X$ is larger, $Y$ tends to be smaller.
Examples of Covariance

Note that $E[X] = 0$ and $E[Y] = 0$ in these examples. Then $\text{cov}(X, Y) = E[XY]$.

When $\text{cov}(X, Y) > 0$, the RVs $X$ and $Y$ tend to be large or small together. $X$ and $Y$ are said to be **positively correlated**.

When $\text{cov}(X, Y) < 0$, when $X$ is larger, $Y$ tends to be smaller. $X$ and $Y$ are said to be **negatively correlated**.

![Diagram showing four equally likely pairs of values with different covariances](image_url)
Examples of Covariance

Note that $E[X] = 0$ and $E[Y] = 0$ in these examples. Then $\text{cov}(X, Y) = E[XY]$.

When $\text{cov}(X, Y) > 0$, the RVs $X$ and $Y$ tend to be large or small together. $X$ and $Y$ are said to be positively correlated.

When $\text{cov}(X, Y) < 0$, when $X$ is larger, $Y$ tends to be smaller. $X$ and $Y$ are said to be negatively correlated.

When $\text{cov}(X, Y) = 0$, we say that $X$ and $Y$ are uncorrelated.
Examples of Covariance

\[ E[X] = 1 \times 0.15 + 2 \times 0.4 + 3 \times 0.45 = 2.3 \]

\[ E[X^2] = 1^2 \times 0.15 + 2^2 \times 0.4 + 3^2 \times 0.45 = 5.8 \]

\[ E[Y] = 1 \times 0.2 + 2 \times 0.6 + 3 \times 0.2 = 2 \]

\[ E[Y^2] = 1^2 \times 0.2 + 2^2 \times 0.6 + 3^2 \times 0.2 = 4.4 \]

\[ E[XY] = 1 \times 0.05 + 2 \times 0.25 + \cdots + 3 \times 3 = 4.85 \]

\[ \text{cov}(X, Y) = E[XY] - E[X]E[Y] = 0.25 \]

\[ \text{var}[X] = E[X^2] - (E[X])^2 = 0.51 \]

\[ \text{var}[Y] = E[Y^2] - (E[Y])^2 = 0.4 \]

\[ \text{corr}(X, Y) \approx 0.55 \]
Examples of Covariance

\[ E(X) = 1 \times 0.15 + 2 \times 0.4 + 3 \times 0.45 = 2.3 \]
Examples of Covariance

\[ E[X] = 1 \times 0.15 + 2 \times 0.4 + 3 \times 0.45 = 2.3 \]
\[ E[X^2] = 1^2 \times 0.15 + 2^2 \times 0.4 + 3^2 \times 0.45 = 5.8 \]
Examples of Covariance

\[ E[X] = 1 \times 0.15 + 2 \times 0.4 + 3 \times 0.45 = 2.3 \]

\[ E[X^2] = 1^2 \times 0.15 + 2^2 \times 0.4 + 3^2 \times 0.45 = 5.8 \]

\[ E[Y] = 1 \times 0.2 + 2 \times 0.6 + 3 \times 0.2 = 2 \]
Examples of Covariance

\[
E[X] = 1 \times 0.15 + 2 \times 0.4 + 3 \times 0.45 = 2.3
\]

\[
E[X^2] = 1^2 \times 0.15 + 2^2 \times 0.4 + 3^2 \times 0.45 = 5.8
\]

\[
E[Y] = 1 \times 0.2 + 2 \times 0.6 + 3 \times 0.2 = 2
\]

\[
E[Y^2] = 1 \times 0.2 + 4 \times 0.6 + 9 \times 0.2 = 4.4
\]
Examples of Covariance

\[ E[X] = 1 \times 0.15 + 2 \times 0.4 + 3 \times 0.45 = 2.3 \]
\[ E[X^2] = 1^2 \times 0.15 + 2^2 \times 0.4 + 3^2 \times 0.45 = 5.8 \]
\[ E[Y] = 1 \times 0.2 + 2 \times 0.6 + 3 \times 0.2 = 2 \]
\[ E[Y^2] = 1 \times 0.2 + 4 \times 0.6 + 9 \times 0.2 = 4.4 \]
\[ E[XY] = 1 \times 0.05 + 1 \times 2 \times 0.1 + \cdots + 3 \times 3 \times 0.2 = 4.85 \]
Examples of Covariance

$E[X] = 1 \times 0.15 + 2 \times 0.4 + 3 \times 0.45 = 2.3$

$E[X^2] = 1^2 \times 0.15 + 2^2 \times 0.4 + 3^2 \times 0.45 = 5.8$

$E[Y] = 1 \times 0.2 + 2 \times 0.6 + 3 \times 0.2 = 2$

$E[Y^2] = 1 \times 0.2 + 4 \times 0.6 + 9 \times 0.2 = 4.4$

$E[XY] = 1 \times 0.05 + 1 \times 2 \times 0.1 + \cdots + 3 \times 3 \times 0.2 = 4.85$

Examples of Covariance

\[ E[X] = 1 \times 0.15 + 2 \times 0.4 + 3 \times 0.45 = 2.3 \]
\[ E[X^2] = 1^2 \times 0.15 + 2^2 \times 0.4 + 3^2 \times 0.45 = 5.8 \]
\[ E[Y] = 1 \times 0.2 + 2 \times 0.6 + 3 \times 0.2 = 2 \]
\[ E[Y^2] = 1 \times 0.2 + 4 \times 0.6 + 9 \times 0.2 = 4.4 \]
\[ E[XY] = 1 \times 0.05 + 1 \times 2 \times 0.1 + \cdots + 3 \times 3 \times 0.2 = 4.85 \]
\[ \text{cov}(X, Y) = E[XY] - E[X]E[Y] = .25 \]
\[ \text{var}[X] = E[X^2] - E[X]^2 = .51 \]
Examples of Covariance

\[ E[X] = 1 \times 0.15 + 2 \times 0.4 + 3 \times 0.45 = 2.3 \]
\[ E[X^2] = 1^2 \times 0.15 + 2^2 \times 0.4 + 3^2 \times 0.45 = 5.8 \]
\[ E[Y] = 1 \times 0.2 + 2 \times 0.6 + 3 \times 0.2 = 2 \]
\[ E[Y^2] = 1 \times 0.2 + 4 \times 0.6 + 9 \times 0.2 = 4.4 \]
\[ E[XY] = 1 \times 0.05 + 1 \times 2 \times 0.1 + \cdots + 3 \times 3 \times 0.2 = 4.85 \]
\[ \text{cov}(X, Y) = E[XY] - E[X]E[Y] = .25 \]
\[ \text{var}[X] = E[X^2] - E[X]^2 = .51 \]
\[ \text{var}[Y] = E[Y^2] - E[Y]^2 = .4 \]
Examples of Covariance

\[ E[X] = 1 \times 0.15 + 2 \times 0.4 + 3 \times 0.45 = 2.3 \]
\[ E[X^2] = 1^2 \times 0.15 + 2^2 \times 0.4 + 3^2 \times 0.45 = 5.8 \]
\[ E[Y] = 1 \times 0.2 + 2 \times 0.6 + 3 \times 0.2 = 2 \]
\[ E[Y^2] = 1 \times 0.2 + 4 \times 0.6 + 9 \times 0.2 = 4.4 \]
\[ E[XY] = 1 \times 0.05 + 1 \times 2 \times 0.1 + \cdots + 3 \times 3 \times 0.2 = 4.85 \]
\[ \text{cov}(X, Y) = E[XY] - E[X]E[Y] = .25 \]
\[ \text{var}[X] = E[X^2] - E[X]^2 = .51 \]
\[ \text{var}[Y] = E[Y^2] - E[Y]^2 = .4 \]
\[ \text{corr}(X, Y) \approx 0.55 \]
Examples of Covariance

\[
E[X] = 1 \times 0.15 + 2 \times 0.4 + 3 \times 0.45 = 2.3
\]

\[
E[X^2] = 1^2 \times 0.15 + 2^2 \times 0.4 + 3^2 \times 0.45 = 5.8
\]

\[
E[Y] = 1 \times 0.2 + 2 \times 0.6 + 3 \times 0.2 = 2
\]

\[
E[Y^2] = 1 \times 0.2 + 4 \times 0.6 + 9 \times 0.2 = 4.4
\]

\[
E[XY] = 1 \times 0.05 + 1 \times 2 \times 0.1 + \cdots + 3 \times 3 \times 0.2 = 4.85
\]

\[
\]

\[
var[X] = E[X^2] - E[X]^2 = .51
\]

\[
var[Y] = E[Y^2] - E[Y]^2 = .4
\]

\[
corr(X, Y) = 0.55
\]
Properties of Covariance

Properties of Covariance


**Fact**
(a) \( \text{var}[X] = \text{cov}(X, X) \)
Properties of Covariance


**Fact**
(a) \( \text{var}[X] = \text{cov}(X, X) \)
(b) \( X, Y \) independent \( \Rightarrow \) \( \text{cov}(X, Y) = \)
Properties of Covariance


**Fact**
(a) \( \text{var}[X] = \text{cov}(X, X) \)
(b) \( X, Y \) independent \( \Rightarrow \) \( \text{cov}(X, Y) = 0 \)
Properties of Covariance


**Fact**

(a) \(\text{var}[X] = \text{cov}(X, X)\)
(b) \(X, Y\) independent \(\Rightarrow \text{cov}(X, Y) = 0\)
(c) \(\text{cov}(a + X, b + Y) = \text{cov}(X, Y)\)
Properties of Covariance


Fact
(a) \( \text{var}[X] = \text{cov}(X, X) \)
(b) \( X, Y \) independent \( \Rightarrow \) \( \text{cov}(X, Y) = 0 \)
(c) \( \text{cov}(a + X, b + Y) = \text{cov}(X, Y) \)
(d) \( \text{cov}(aX + bY, cU + dV) = ac \cdot \text{cov}(X, U) + ad \cdot \text{cov}(X, V) \)
\[ + bc \cdot \text{cov}(Y, U) + bd \cdot \text{cov}(Y, V). \]
Properties of Covariance


**Fact**

(a) \( \text{var}[X] = \text{cov}(X, X) \)
(b) \( X, Y \) independent \( \Rightarrow \text{cov}(X, Y) = 0 \)
(c) \( \text{cov}(a + X, b + Y) = \text{cov}(X, Y) \)
(d) \( \text{cov}(aX + bY, cU + dV) = ac \cdot \text{cov}(X, U) + ad \cdot \text{cov}(X, V) \)
\[ + bc \cdot \text{cov}(Y, U) + bd \cdot \text{cov}(Y, V). \]

**Proof:**

(a)-(b)-(c) are obvious.
Properties of Covariance


**Fact**
(a) \( \text{var}[X] = \text{cov}(X, X) \)
(b) \( X, Y \) independent \( \Rightarrow \) \( \text{cov}(X, Y) = 0 \)
(c) \( \text{cov}(a + X, b + Y) = \text{cov}(X, Y) \)
(d) \( \text{cov}(aX + bY, cU + dV) = ac \cdot \text{cov}(X, U) + ad \cdot \text{cov}(X, V) \)
\[ +bc \cdot \text{cov}(Y, U) + bd \cdot \text{cov}(Y, V). \]

**Proof:**
(a)-(b)-(c) are obvious.
(d) In view of (c), one can subtract the means and assume that the RVs are zero-mean.
Properties of Covariance


Fact
(a) \( \text{var}[X] = \text{cov}(X, X) \)
(b) \( X, Y \) independent \( \Rightarrow \) \( \text{cov}(X, Y) = 0 \)
(c) \( \text{cov}(a + X, b + Y) = \text{cov}(X, Y) \)
(d) \( \text{cov}(aX + bY, cU + dV) = ac \cdot \text{cov}(X, U) + ad \cdot \text{cov}(X, V) \)
\quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad + bc \cdot \text{cov}(Y, U) + bd \cdot \text{cov}(Y, V). \)

Proof:
(a)-(b)-(c) are obvious.
(d) In view of (c), one can subtract the means and assume that the RVs are zero-mean. Then,
\[ \text{cov}(aX + bY, cU + dV) = E[(aX + bY)(cU + dV)] \]
Properties of Covariance


Fact
(a) \( \text{var}[X] = \text{cov}(X, X) \)
(b) \( X, Y \) independent \( \Rightarrow \) \( \text{cov}(X, Y) = 0 \)
(c) \( \text{cov}(a + X, b + Y) = \text{cov}(X, Y) \)
(d) \( \text{cov}(aX + bY, cU + dV) = ac \cdot \text{cov}(X, U) + ad \cdot \text{cov}(X, V) + bc \cdot \text{cov}(Y, U) + bd \cdot \text{cov}(Y, V) \).

Proof:
(a)-(b)-(c) are obvious.
(d) In view of (c), one can subtract the means and assume that the RVs are zero-mean. Then,

\[
\text{cov}(aX + bY, cU + dV) = E[(aX + bY)(cU + dV)] \\
= ac \cdot E[XU] + ad \cdot E[XV] + bc \cdot E[YU] + bd \cdot E[YV]
\]
Properties of Covariance


Fact
(a) $$\text{var}[X] = \text{cov}(X, X)$$
(b) $$X, Y$$ independent $$\Rightarrow \text{cov}(X, Y) = 0$$
(c) $$\text{cov}(a + X, b + Y) = \text{cov}(X, Y)$$
(d) $$\text{cov}(aX + bY, cU + dV) = ac \cdot \text{cov}(X, U) + ad \cdot \text{cov}(X, V)$$
$$+ bc \cdot \text{cov}(Y, U) + bd \cdot \text{cov}(Y, V).$$

Proof:
(a)-(b)-(c) are obvious.
(d) In view of (c), one can subtract the means and assume that the RVs are zero-mean. Then,

$$\text{cov}(aX + bY, cU + dV) = E[(aX + bY)(cU + dV)]$$
$$= ac \cdot E[XU] + ad \cdot E[XV] + bc \cdot E[YU] + bd \cdot E[YV]$$
$$= ac \cdot \text{cov}(X, U) + ad \cdot \text{cov}(X, V) + bc \cdot \text{cov}(Y, U) + bd \cdot \text{cov}(Y, V).$$
Properties of Covariance


Fact

(a) \( \text{var}[X] = \text{cov}(X, X) \)
(b) \( X, Y \) independent \( \Rightarrow \text{cov}(X, Y) = 0 \)
(c) \( \text{cov}(a + X, b + Y) = \text{cov}(X, Y) \)
(d) \( \text{cov}(aX + bY, cU + dV) = ac \cdot \text{cov}(X, U) + ad \cdot \text{cov}(X, V) \)
\[ + bc \cdot \text{cov}(Y, U) + bd \cdot \text{cov}(Y, V). \]

Proof:

(a)-(b)-(c) are obvious.
(d) In view of (c), one can subtract the means and assume that the RVs are zero-mean. Then,

\[
\text{cov}(aX + bY, cU + dV) = E[(aX + bY)(cU + dV)]
\]
\[
= ac \cdot E[XU] + ad \cdot E[XV] + bc \cdot E[YU] + bd \cdot E[YV]
\]
\[
= ac \cdot \text{cov}(X, U) + ad \cdot \text{cov}(X, V) + bc \cdot \text{cov}(Y, U) + bd \cdot \text{cov}(Y, V).
\]
Summary

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

Fact:
\[ \text{var} [aX + b] = a^2 \text{var} [X] \]

Sum:
\[ X, Y, Z \text{ pairwise ind.} \Rightarrow \text{var} [X+Y+Z] = \cdots \]
Summary

Variance

- **Variance:** $\text{var}[X] := E[(X - E[X])^2] = E[X^2] - E[X]^2$
- **Fact:** $\text{var}[aX + b] = a^2 \text{var}[X]$
Summary

Variance:
- Fact: \( \text{var}[aX + b] = a^2 \text{var}[X] \)
- Sum: \( X, Y, Z \) pairwise ind. \( \Rightarrow \text{var}[X + Y + Z] = \cdots \)
Variance

- **Variance:** \( \text{var}[X] := E[(X - E[X])^2] = E[X^2] - E[X]^2 \)
- **Fact:** \( \text{var}[aX + b] = a^2 \text{var}[X] \)
- **Sum:** \( X, Y, Z \) pairwise ind. \( \Rightarrow \text{var}[X + Y + Z] = \cdots \)
Random Variables so far.

Probability Space: $\Omega$, $Pr : \Omega \rightarrow [0, 1]$, $\sum_{\omega \in \Omega} Pr(\omega) = 1$. 
Random Variables so far.

Probability Space: $\Omega$, $Pr : \Omega \rightarrow [0, 1]$, $\sum_{\omega \in \Omega} Pr(\omega) = 1$.
Random Variables: $X : \Omega \rightarrow R$. 
Random Variables so far.

Probability Space: $\Omega$, $Pr : \Omega \rightarrow [0, 1]$, $\sum_{\omega \in \Omega} Pr(\omega) = 1$.
Random Variables: $X : \Omega \rightarrow R$.
Associated event: $Pr[X = a] = \sum_{\omega : X(\omega) = a} Pr(\omega)$
Random Variables so far.

Probability Space: $\Omega$, $Pr : \Omega \rightarrow [0, 1]$, $\sum_{\omega \in \Omega} Pr(\omega) = 1$.
Random Variables: $X : \Omega \rightarrow \mathbb{R}$.
  Associated event: $Pr[X = a] = \sum_{\omega : X(\omega) = a} Pr(\omega)$
$X$ and $Y$ independent $\iff$ all associated events are independent.
Random Variables so far.

Probability Space: $\Omega, Pr : \Omega \rightarrow [0, 1], \sum_{\omega \in \Omega} Pr(\omega) = 1.$
Random Variables: $X : \Omega \rightarrow \mathbb{R}.$

Associated event: $Pr[X = a] = \sum_{\omega : X(\omega) = a} Pr(\omega)$

$X$ and $Y$ independent $\iff$ all associated events are independent.

Expectation: $E[X] = \sum_{a} aPr[X = a] = \sum_{\omega \in \Omega} X(\omega) Pr(\omega).$
Random Variables so far.

Probability Space: $\Omega$, $Pr : \Omega \rightarrow [0, 1]$, $\sum_{\omega \in \Omega} Pr(\omega) = 1$.  

Random Variables: $X : \Omega \rightarrow R$.  

Associated event: $Pr[X = a] = \sum_{\omega : X(\omega) = a} Pr(\omega)$  

$X$ and $Y$ independent $\iff$ all associated events are independent.  

Expectation: $E[X] = \sum aPr[X = a] = \sum_{\omega \in \Omega} X(\omega)Pr(\omega)$.  


Poisson: $X \sim P(\lambda)$  

$E(X) = \lambda$, $Var(X) = \lambda$.  

Binomial: $X \sim B(n, p)$  

$E(X) = np$, $Var(X) = np(1 - p)$.  

Uniform: $X \sim U\{1, \ldots, n\}$  

$E[X] = \frac{n + 1}{2}$, $Var(X) = \frac{n^2}{12}$.  

Geometric: $X \sim G(p)$  

$E(X) = \frac{1}{p}$, $Var(X) = \frac{1}{p^2}$.  

Random Variables so far.

Probability Space: $\Omega$, $Pr : \Omega \rightarrow [0, 1]$, $\sum_{\omega \in \Omega} Pr(\omega) = 1$.
Random Variables: $X : \Omega \rightarrow R$.
  Associated event: $Pr[X = a] = \sum_{\omega : X(\omega) = a} Pr(\omega)$
$X$ and $Y$ independent $\iff$ all associated events are independent.
Expectation: $E[X] = \sum_{a} aPr[X = a] = \sum_{\omega \in \Omega} X(\omega)Pr(\omega)$.
Variance: $Var(X) = E[(X - E[X])^2] = E[X^2] - (E(X))^2$
Random Variables so far.

Probability Space: $\Omega$, $Pr : \Omega \rightarrow [0, 1]$, $\sum_{\omega \in \Omega} Pr(\omega) = 1$.
Random Variables: $X : \Omega \rightarrow R$.
  Associated event: $Pr[X = a] = \sum_{\omega : X(\omega) = a} Pr(\omega)$
$X$ and $Y$ independent $\iff$ all associated events are independent.
Expectation: $E[X] = \sum_a a Pr[X = a] = \sum_{\omega \in \Omega} X(\omega) Pr(\omega)$.
Variance: $Var(X) = E[(X - E[X])^2] = E[X^2] - (E(X))^2$
  For independent $X, Y$, $Var(X + Y) = Var(X) + Var(Y)$. 
Random Variables so far.

Probability Space: $\Omega$, $Pr : \Omega \rightarrow [0, 1]$, $\sum_{\omega \in \Omega} Pr(\omega) = 1$.

Random Variables: $X : \Omega \rightarrow R$.

  Associated event: $Pr[X = a] = \sum_{\omega: X(\omega) = a} Pr(\omega)$

$X$ and $Y$ independent $\iff$ all associated events are independent.

Expectation: $E[X] = \sum_{a} aPr[X = a] = \sum_{\omega \in \Omega} X(\omega)Pr(\omega)$.


Variance: $Var(X) = E[(X - E[X])^2] = E[X^2] - (E(X))^2$.

  For independent $X, Y$, $Var(X + Y) = Var(X) + Var(Y)$.

  Also: $Var(cX) = c^2 Var(X)$ and $Var(X + b) = Var(X)$. 
Probability Space: $\Omega, \Pr : \Omega \to [0, 1], \sum_{\omega \in \Omega} \Pr(\omega) = 1.$

Random Variables: $X : \Omega \to \mathbb{R}.$

Associated event: $\Pr[X = a] = \sum_{\omega : X(\omega) = a} \Pr(\omega)$

$X$ and $Y$ independent $\iff$ all associated events are independent.

Expectation: $E[X] = \sum_a a \Pr[X = a] = \sum_{\omega \in \Omega} X(\omega) \Pr(\omega).$


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

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

Also: $\text{Var}(cX) = c^2 \text{Var}(X)$ and $\text{Var}(X + b) = \text{Var}(X)$.

Poisson: $X \sim P(\lambda)$
Random Variables so far.

Probability Space: $\Omega$, $Pr : \Omega \to [0, 1]$, $\sum_{\omega \in \Omega} Pr(\omega) = 1$.

Random Variables: $X : \Omega \to R$.

Associated event: $Pr[X = a] = \sum_{\omega : X(\omega) = a} Pr(\omega)$

$X$ and $Y$ independent $\iff$ all associated events are independent.

Expectation: $E[X] = \sum_a aPr[X = a] = \sum_{\omega \in \Omega} X(\omega)Pr(\omega)$.


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

For independent $X, Y$, $Var(X + Y) = Var(X) + Var(Y)$.

Also: $Var(cX) = c^2 Var(X)$ and $Var(X + b) = Var(X)$.

Poisson: $X \sim P(\lambda)$ $E(X) = \lambda$, $Var(X) = \lambda$. 
Random Variables so far.

Probability Space: $\Omega$, $Pr : \Omega \to [0, 1]$, $\sum_{\omega \in \Omega} Pr(\omega) = 1$.
Random Variables: $X : \Omega \to R$.
Associated event: $Pr[X = a] = \sum_{\omega : X(\omega) = a} Pr(\omega)$
$X$ and $Y$ independent $\iff$ all associated events are independent.
Expectation: $E[X] = \sum_{a} a Pr[X = a] = \sum_{\omega \in \Omega} X(\omega) Pr(\omega)$.
Variance: $\text{Var}(X) = E[(X - E[X])^2] = E[X^2] - (E(X))^2$
For independent $X, Y$, $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$.
Also: $\text{Var}(cX) = c^2 \text{Var}(X)$ and $\text{Var}(X + b) = \text{Var}(X)$.
Poisson: $X \sim P(\lambda)$ $E(X) = \lambda$, $\text{Var}(X) = \lambda$.
Binomial: $X \sim B(n, p)$
Random Variables so far.

Probability Space: $\Omega, \ Pr : \Omega \rightarrow [0, 1], \ \sum_{\omega \in \Omega} Pr(w) = 1$.
Random Variables: $X : \Omega \rightarrow R$.
  Associated event: $Pr[X = a] = \sum_{\omega : X(\omega) = a} Pr(\omega)$
$X$ and $Y$ independent $\iff$ all associated events are independent.
Expectation: $E[X] = \sum_a aPr[X = a] = \sum_{\omega \in \Omega} X(\omega) Pr(\omega)$.
Variance: $Var(X) = E[(X - E[X])^2] = E[X^2] - (E(X))^2$
  For independent $X, Y$, $Var(X + Y) = Var(X) + Var(Y)$.
  Also: $Var(cX) = c^2 Var(X)$ and $Var(X + b) = Var(X)$.
Poisson: $X \sim P(\lambda)$ $E(X) = \lambda$, $Var(X) = \lambda$.
Binomial: $X \sim B(n, p)$ $E(X) = np$, $Var(X) = np(1 - p)$
Probability Space: $\Omega$, $Pr : \Omega \rightarrow [0, 1]$, $\sum_{\omega \in \Omega} Pr(\omega) = 1$.

Random Variables: $X : \Omega \rightarrow R$.

Associated event: $Pr[X = a] = \sum_{\omega : X(\omega) = a} Pr(\omega)$

$X$ and $Y$ independent $\iff$ all associated events are independent.

Expectation: $E[X] = \sum_a a Pr[X = a] = \sum_{\omega \in \Omega} X(\omega) Pr(\omega)$.


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

For independent $X$, $Y$, $Var(X + Y) = Var(X) + Var(Y)$.

Also: $Var(cX) = c^2 Var(X)$ and $Var(X + b) = Var(X)$.

Poisson: $X \sim P(\lambda)$ $E(X) = \lambda$, $Var(X) = \lambda$.

Binomial: $X \sim B(n,p)$ $E(X) = np$, $Var(X) = np(1 - p)$

Uniform: $X \sim U\{1, \ldots, n\}$
Random Variables so far.

Probability Space: $\Omega$, $\Pr : \Omega \rightarrow [0, 1]$, $\sum_{\omega \in \Omega} Pr(\omega) = 1$.

Random Variables: $X : \Omega \rightarrow R$.

Associated event: $Pr[X = a] = \sum_{\omega : X(\omega) = a} Pr(\omega)$

$X$ and $Y$ independent $\iff$ all associated events are independent.

Expectation: $E[X] = \sum_a aPr[X = a] = \sum_{\omega \in \Omega} X(\omega)Pr(\omega)$.


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

For independent $X$, $Y$, $Var(X + Y) = Var(X) + Var(Y)$.

Also: $Var(cX) = c^2 Var(X)$ and $Var(X + b) = Var(X)$.

Poisson: $X \sim P(\lambda)$ $E(X) = \lambda$, $Var(X) = \lambda$.

Binomial: $X \sim B(n, p)$ $E(X) = np$, $Var(X) = np(1 - p)$

Uniform: $X \sim U\{1, \ldots, n\}$ $E[X] = \frac{n+1}{2}$, $Var(X) = \frac{n^2-1}{12}$.
Random Variables so far.

Probability Space: $\Omega, \Pr : \Omega \to [0, 1], \sum_{\omega \in \Omega} \Pr(\omega) = 1$.
Random Variables: $X : \Omega \to \mathbb{R}$.
Associated event: $\Pr[X = a] = \sum_{\omega : X(\omega) = a} \Pr(\omega)$
$X$ and $Y$ independent $\iff$ all associated events are independent.
Expectation: $E[X] = \sum_a a\Pr[X = a] = \sum_{\omega \in \Omega} X(\omega)\Pr(\omega)$.

Variance: $\text{Var}(X) = E[(X - E[X])^2] = E[X^2] - (E(X))^2$
For independent $X, Y$, $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$.
Also: $\text{Var}(cX) = c^2 \text{Var}(X)$ and $\text{Var}(X + b) = \text{Var}(X)$.

Poisson: $X \sim P(\lambda)$ $E(X) = \lambda$, $\text{Var}(X) = \lambda$.
Binomial: $X \sim B(n, p)$ $E(X) = np$, $\text{Var}(X) = np(1 - p)$
Uniform: $X \sim U\{1, \ldots, n\}$ $E[X] = \frac{n+1}{2}$, $\text{Var}(X) = \frac{n^2-1}{12}$.
Geometric: $X \sim G(p)$
Random Variables so far.

Probability Space: $\Omega$, $Pr : \Omega \rightarrow [0, 1]$, $\sum_{\omega \in \Omega} Pr(w) = 1$.
Random Variables: $X : \Omega \rightarrow R$.

Associated event: $Pr[X = a] = \sum_{\omega : X(\omega) = a} Pr(\omega)$

$X$ and $Y$ independent $\iff$ all associated events are independent.

Expectation: $E[X] = \sum_a a Pr[X = a] = \sum_{\omega \in \Omega} X(\omega) Pr(\omega)$.


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

For independent $X, Y$, $Var(X + Y) = Var(X) + Var(Y)$.

Also: $Var(cX) = c^2 Var(X)$ and $Var(X + b) = Var(X)$.

Poisson: $X \sim P(\lambda)$ $E(X) = \lambda$, $Var(X) = \lambda$.
Binomial: $X \sim B(n, p)$ $E(X) = np$, $Var(X) = np(1 - p)$
Uniform: $X \sim U\{1, \ldots, n\}$ $E[X] = \frac{n+1}{2}$, $Var(X) = \frac{n^2-1}{12}$.
Geometric: $X \sim G(p)$ $E(X) = \frac{1}{p}$, $Var(X) = \frac{1-p}{p^2}$.