CS70

Coupon Collecting: Fun with harmonic numbers!

CS70

Coupon Collecting: Fun with harmonic numbers!

Memoryless Property.

CS70

Coupon Collecting: Fun with harmonic numbers!

Memoryless Property.

Law of the unconscious statistician. (Hmmm.)

CS70

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[“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[“get second coupon”|“got ~~milk~~ 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[“get second coupon”|“got ~~milk~~ 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[“get second coupon”|“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[“get second coupon”|“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.

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

$E[X_2]$? 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 ~~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[“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.

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

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

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

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 ~~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[“getting ith coupon|“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[“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[“getting ith coupon|“got $i-1$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[“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[“getting ith 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["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["getting ith coupon|"got $i-1$rst coupons"] $= \frac{n-(i-1)}{n} = \frac{n-i+1}{n}$

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

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

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

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

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

$$E[X] = E[X_1] + \cdots + E[X_n] =$$

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

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

$$\begin{aligned} E[X] &= E[X_1] + \cdots + E[X_n] = \frac{n}{n} + \frac{n}{n-1} + \frac{n}{n-2} + \cdots + \frac{n}{1} \\ &= n(1 + \frac{1}{2} + \cdots + \frac{1}{n}) =: nH(n) \end{aligned}$$

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 ~~milk~~ first coupon"$] = \frac{n-1}{n}$

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

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

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

$$
\begin{aligned}
E[X] &= E[X_1] + \cdots + E[X_n] = \frac{n}{n} + \frac{n}{n-1} + \frac{n}{n-2} + \cdots + \frac{n}{1} \\
&= n(1 + \frac{1}{2} + \cdots + \frac{1}{n}) =: nH(n) \approx n(\ln n + \gamma)
\end{aligned}
$$

Review: Harmonic sum

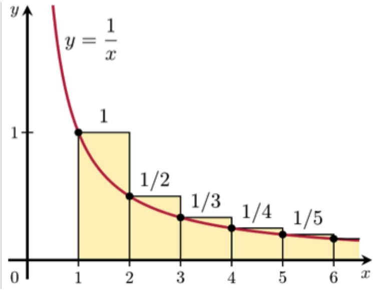
$$H(n) = 1 + \frac{1}{2} + \cdots + \frac{1}{n} \approx \int_1^n \frac{1}{x} dx = \ln(n).$$

.

Review: Harmonic sum

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

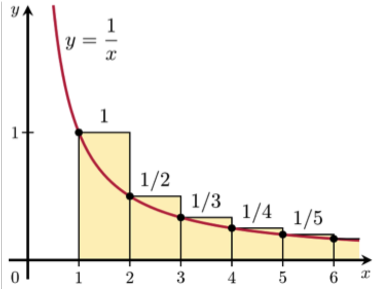
.



Review: Harmonic sum

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

.



A good approximation is

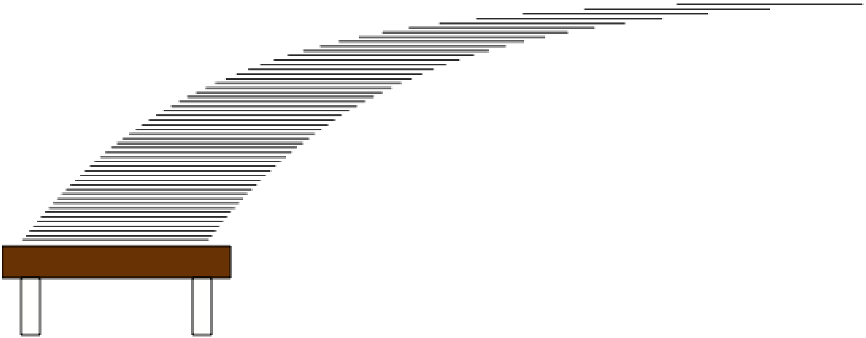
$$H(n) \approx \ln(n) + \gamma \text{ where } \gamma \approx 0.58 \text{ (Euler-Mascheroni constant).}$$

Harmonic sum: Paradox

Consider this stack of cards (no glue!):

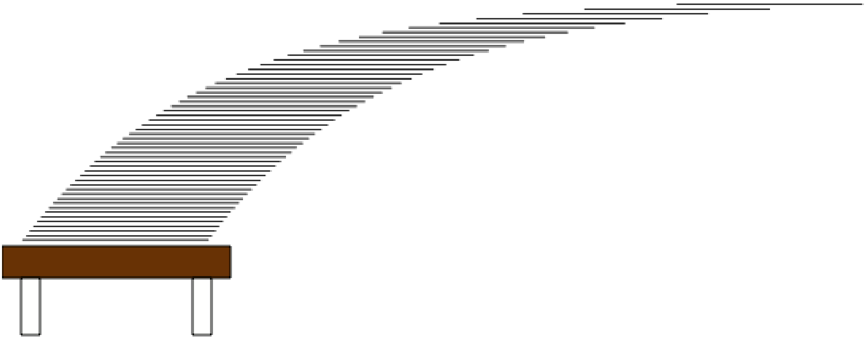
Harmonic sum: Paradox

Consider this stack of cards (no glue!):



Harmonic sum: Paradox

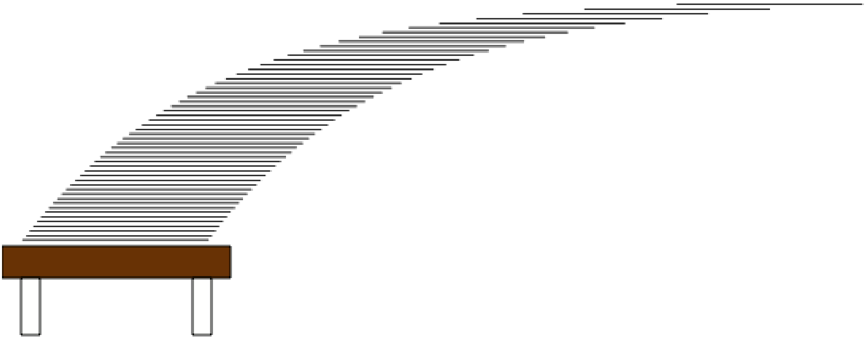
Consider this stack of cards (no glue!):



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

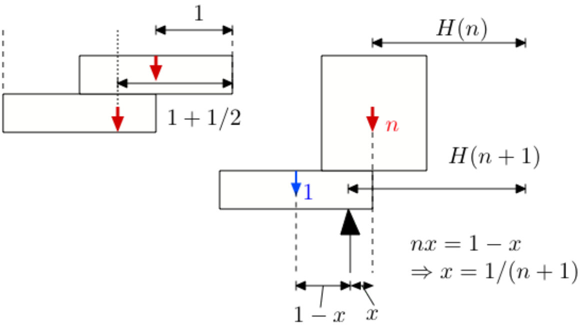
/ˈperəˌdäks/

noun

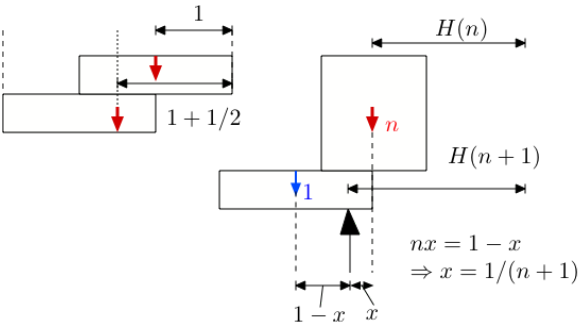
a statement or proposition that, despite sound (or apparently sound) reasoning from acceptable premises, leads to a conclusion that seems senseless, logically unacceptable, or self-contradictory.
"a potentially serious conflict between quantum mechanics and the general theory of relativity known as the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

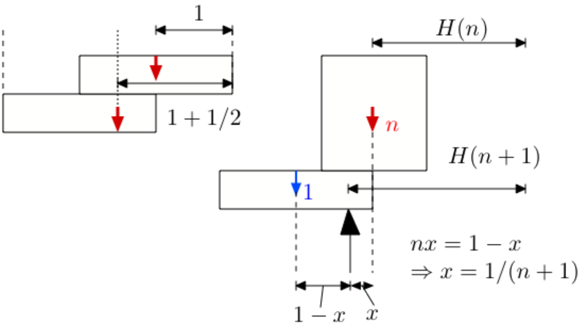


Stacking



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.

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:

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

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:

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

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:

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

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:

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

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:

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

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:

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

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:

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

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:

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

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:

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

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:

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

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:

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

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:

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

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:

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

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:

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

$\square$

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.

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

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.

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

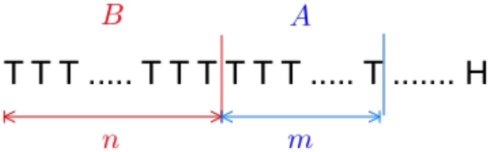
(A) Distribution of $X \sim G(p)$: $Pr[X = i] = (1-p)^{i-1}p$.

Geometric Distribution: Memoryless - Interpretation

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

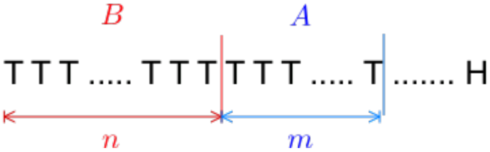
Geometric Distribution: Memoryless - Interpretation

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



Geometric Distribution: Memoryless - Interpretation

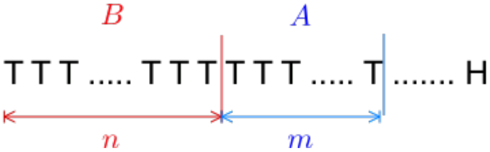
$$Pr[X > n+m | X > n] = Pr[X > m], m, n \geq 0.$$



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

Geometric Distribution: Memoryless - Interpretation

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

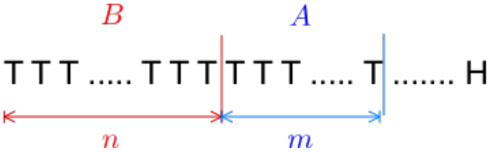


$$Pr[X > n+m | X > n] = Pr[A|B] = Pr[A'] = Pr[X > m].$$

A': is m coin tosses before heads.
$A|B$: m 'more' coin tosses before heads.

Geometric Distribution: Memoryless - Interpretation

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



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

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:

$$Pr[X > n+m | X > n] \quad =$$

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:

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

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:

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

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:

$$\begin{aligned} Pr[X > n+m | X > n] &= \frac{Pr[X > n+m \text{ and } X > n]}{Pr[X > n]} \\ &= \frac{Pr[X > n+m]}{Pr[X > n]} \\ &= \frac{(1-p)^{n+m}}{(1-p)^n} = \end{aligned}$$

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:

$$\begin{aligned} Pr[X > n+m | X > n] &= \frac{Pr[X > n+m \text{ and } X > n]}{Pr[X > n]} \\ &= \frac{Pr[X > n+m]}{Pr[X > n]} \\ &= \frac{(1-p)^{n+m}}{(1-p)^n} = (1-p)^m \end{aligned}$$

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

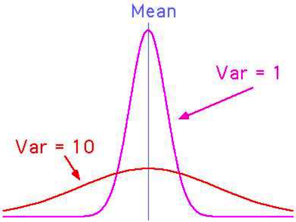
$$Pr[X > n+m | X > n] = Pr[X > m], m, n \geq 0.$$

Proof:

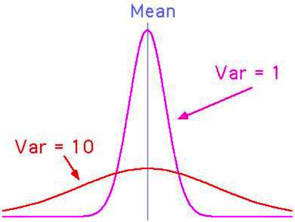
$$\begin{aligned}
Pr[X > n+m | X > n] &= \frac{Pr[X > n+m \text{ and } X > n]}{Pr[X > n]} \\
&= \frac{Pr[X > n+m]}{Pr[X > n]} \\
&= \frac{(1-p)^{n+m}}{(1-p)^n} = (1-p)^m \\
&= Pr[X > m].
\end{aligned}$$

Variance

Variance

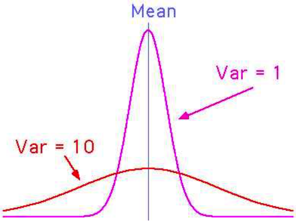


Variance



The variance measures the deviation from the mean value.

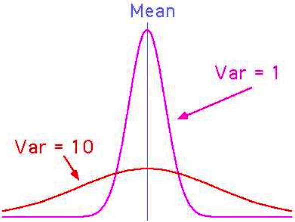
Variance



The variance measures the deviation from the mean value.

Definition: The variance of X is

Variance

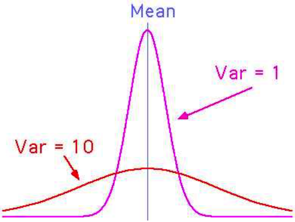


The variance measures the deviation from the mean value.

Definition: The variance of X is

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

Variance



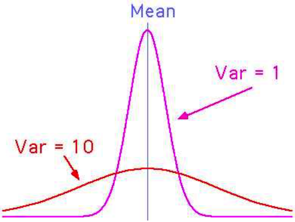
The variance measures the deviation from the mean value.

Definition: The variance of X is

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

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

Variance



The variance measures the deviation from the mean value.

Definition: The variance of X is

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

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

Variance and Standard Deviation

Fact:

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

Variance and Standard Deviation

Fact:

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

Indeed:

$$var(X) \quad = \quad E[(X - E[X])^2]$$

Variance and Standard Deviation

Fact:

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

Indeed:

$$\begin{aligned} var(X) &= E[(X - E[X])^2] \\ &= E[X^2 - 2XE[X] + E[X]^2) \end{aligned}$$

Variance and Standard Deviation

Fact:

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

Indeed:

$$\begin{aligned} var(X) &= E[(X - E[X])^2] \\ &= E[X^2 - 2XE[X] + E[X]^2) \\ &= E[X^2] - 2E[X]E[X] + E[X]^2, \end{aligned}$$

Variance and Standard Deviation

Fact:

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

Indeed:

$$\begin{aligned} var(X) &= E[(X - E[X])^2] \\ &= E[X^2 - 2XE[X] + E[X]^2) \\ &= E[X^2] - 2E[X]E[X] + E[X]^2, \text{ by linearity} \end{aligned}$$

Variance and Standard Deviation

Fact:

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

Indeed:

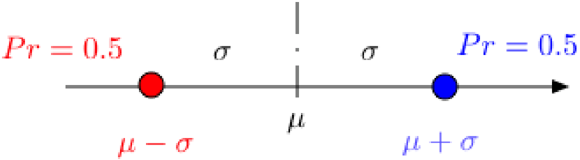
$$\begin{aligned} var(X) &= E[(X - E[X])^2] \\ &= E[X^2 - 2XE[X] + E[X]^2) \\ &= E[X^2] - 2E[X]E[X] + E[X]^2, \text{ by linearity} \\ &= E[X^2] - E[X]^2. \end{aligned}$$

A simple example

This example illustrates the term ‘standard deviation.’

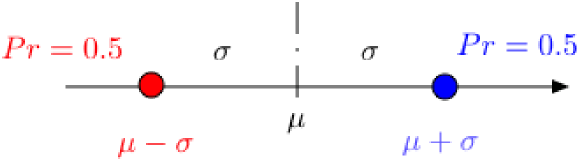
A simple example

This example illustrates the term ‘standard deviation.’



A simple example

This example illustrates the term ‘standard deviation.’

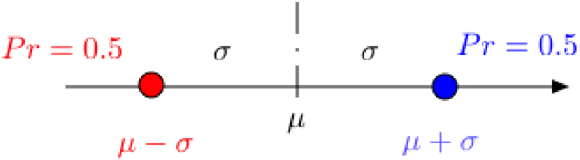


Consider the random variable X such that

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

A simple example

This example illustrates the term ‘standard deviation.’



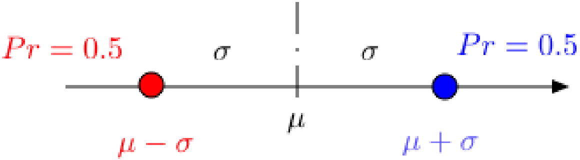
Consider the random variable X such that

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

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,

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

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

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

Example

Consider X with

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

Then

$$\begin{aligned} E[X] &= -1 \times 0.99 + 99 \times 0.01 = 0. \\ E[X^2] &= 1 \times 0.99 + (99)^2 \times 0.01 \approx 100. \end{aligned}$$

Example

Consider X with

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

Then

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

Example

Consider X with

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

Then

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

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

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

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

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

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 $\frac{\sigma(X)}{E[|X-E[X]|]}$?

Uniform

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

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

Uniform

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

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

Uniform

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

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

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

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

Also,

$$\begin{aligned} E[X^2] &= \sum_{i=1}^{n} i^2 Pr[X = i] = \frac{1}{n} \sum_{i=1}^{n} i^2 \\ &= \frac{1}{n} \frac{(n)(n+1)(n+2)}{6} = \frac{1+3n+2n^2}{6}, \end{aligned}$$

Uniform

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

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

Also,

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

Uniform

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

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

Also,

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

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.

Variance of geometric distribution.

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

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 $E[X] = 1/p$.

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 $E[X] = 1/p$.

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

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 $E[X] = 1/p$.

$$\begin{aligned} E[X^2] &= p + 4p(1-p) + 9p(1-p)^2 + ... \\ -(1-p)E[X^2] &= -[p(1-p) + 4p(1-p)^2 + ...] \end{aligned}$$

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 $E[X] = 1/p$.

$$\begin{aligned}
E[X^2] &= p + 4p(1-p) + 9p(1-p)^2 + ... \\
-(1-p)E[X^2] &= \quad -[p(1-p) + 4p(1-p)^2 + ...] \\
pE[X^2] &= p + 3p(1-p) + 5p(1-p)^2 + ...
\end{aligned}$$

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 $E[X] = 1/p$.

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

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 $E[X] = 1/p$.

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

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 $E[X] = 1/p$.

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

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 $E[X]=1/p$.

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

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 $E[X] = 1/p$.

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

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 $E[X] = 1/p$.

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

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 $E[X]=1/p$.

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

$\implies E[X^2]=(2-p)/p^2$

Variance of geometric distribution.

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

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

$\Longrightarrow E[X^2]=(2-p)/p^2$ and
$var[X]=E[X^2]-E[X]^2$

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 $E[X] = 1/p$.

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

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

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 $E[X] = 1/p$.

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

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

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 $E[X] = 1/p$.

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

$\Longrightarrow E[X^2] = (2-p)/p^2$ and
$var[X] = E[X^2] - E[X]^2 = \frac{2-p}{p^2} - \frac{1}{p^2} = \frac{1-p}{p^2}$.
$\sigma(X) = \frac{\sqrt{1-p}}{p} \approx E[X]$ when p is small(ish).

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 $E[X]=1/p$.

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

$\Longrightarrow E[X^2]=(2-p)/p^2$ and
$var[X]=E[X^2]-E[X]^2=\frac{2-p}{p^2}-\frac{1}{p^2}=\frac{1-p}{p^2}$.
$\sigma(X)=\frac{\sqrt{1-p}}{p} \approx E[X]$ when p is small(ish).

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 student 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 ith 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 ith student getting hw back.

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

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 ith student getting hw back.

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

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

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 ith student getting hw back.

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

$$\begin{aligned} E(X_i^2) &= 1 \times Pr[X_i = 1] + 0 \times Pr[X_i = 0] \\ &= \frac{1}{n} \end{aligned}$$

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 ith 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 ith student getting hw back.

$$\begin{aligned} E(X^2) &= \sum_i E(X_i^2) + \sum_{i \neq j} E(X_i X_j). \\ &= n \times \frac{1}{n} + \end{aligned}$$

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

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 ith student getting hw back.

$$
\begin{aligned}
E(X^2) &= \sum_i E(X_i^2) + \sum_{i \neq j} E(X_i X_j). \\
&= n \times \frac{1}{n} +
\end{aligned}
$$

$E(X_i^2) = 1 \times Pr[X_i = 1] + 0 \times Pr[X_i = 0]$
$= \frac{1}{n}$
$E(X_i X_j) = 1 \times Pr[X_i = 1 \cap X_j = 1] + 0 \times Pr[\text{"anything else"}]$

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 ith student getting hw back.

$$\begin{aligned} E(X^2) &= \sum_i E(X_i^2) + \sum_{i \neq j} E(X_i X_j). \\ &= n \times \frac{1}{n} + \end{aligned}$$

$E(X_i^2) = 1 \times Pr[X_i = 1] + 0 \times Pr[X_i = 0]$
$\qquad = \frac{1}{n}$
$E(X_i X_j) = 1 \times Pr[X_i = 1 \cap X_j = 1] + 0 \times Pr[\text{"anything else"}]$
$\qquad = 1 \times \frac{(n-2)!}{n!}$

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 ith student getting hw back.

$$\begin{aligned} E(X^2) &= \sum_i E(X_i^2) + \sum_{i \neq j} E(X_i X_j). \\ &= n \times \frac{1}{n} + \end{aligned}$$

$E(X_i^2) = 1 \times Pr[X_i = 1] + 0 \times Pr[X_i = 0]$
$= \frac{1}{n}$
$E(X_i X_j) = 1 \times Pr[X_i = 1 \cap X_j = 1] + 0 \times Pr[\text{"anything else"}]$
$= 1 \times \frac{(n-2)!}{n!} = \frac{1}{n(n-1)}$

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 ith student getting hw back.

$$\begin{aligned} E(X^2) &= \sum_i E(X_i^2) + \sum_{i \neq j} E(X_i X_j). \\ &= n \times \frac{1}{n} + (n)(n-1) \times \frac{1}{n(n-1)} \end{aligned}$$

$E(X_i^2) = 1 \times Pr[X_i = 1] + 0 \times Pr[X_i = 0]$
$= \frac{1}{n}$
$E(X_i X_j) = 1 \times Pr[X_i = 1 \cap X_j = 1] + 0 \times Pr[\text{"anything else"}]$
$= 1 \times \frac{(n-2)!}{n!} = \frac{1}{n(n-1)}$

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 ith student getting hw back.

$$\begin{aligned} E(X^2) &= \sum_i E(X_i^2) + \sum_{i \neq j} E(X_i X_j). \\ &= n \times \frac{1}{n} + (n)(n-1) \times \frac{1}{n(n-1)} \\ &= 1 + 1 = 2. \end{aligned}$$

$E(X_i^2) = 1 \times Pr[X_i = 1] + 0 \times Pr[X_i = 0]$
$= \frac{1}{n}$
$E(X_i X_j) = 1 \times Pr[X_i = 1 \cap X_j = 1] + 0 \times Pr[\text{"anything else"}]$
$= 1 \times \frac{(n-2)!}{n!} = \frac{1}{n(n-1)}$

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 ith student getting hw back.

$$\begin{aligned} E(X^2) &= \sum_i E(X_i^2) + \sum_{i \neq j} E(X_i X_j). \\ &= n \times \frac{1}{n} + (n)(n-1) \times \frac{1}{n(n-1)} \\ &= 1 + 1 = 2. \end{aligned}$$

$E(X_i^2) = 1 \times Pr[X_i = 1] + 0 \times Pr[X_i = 0]$
$\quad = \frac{1}{n}$
$E(X_i X_j) = 1 \times Pr[X_i = 1 \cap X_j = 1] + 0 \times Pr[\text{``anything else''}]$
$\quad = 1 \times \frac{(n-2)!}{n!} = \frac{1}{n(n-1)}$

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

Poll: fixed points.

What's true?

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) $X_i^2 = X_i$.

Variance: binomial.

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

Variance: binomial.

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

Variance: binomial.

$$
\begin{aligned}
E[X^2] &= \sum_{i=0}^{n} i^2 \binom{n}{i} p^i (1-p)^{n-i}. \\
&= \text{Really???!!##...}
\end{aligned}
$$

Too hard!

Variance: binomial.

$$\begin{aligned} E[X^2] &= \sum_{i=0}^{n} i^2 \binom{n}{i} p^i (1-p)^{n-i}. \\ &= \text{Really???!!\#\#...} \end{aligned}$$

Too hard!

Ok..

Variance: binomial.

$$\begin{aligned} E[X^2] &= \sum_{i=0}^{n} i^2 \binom{n}{i} p^i (1-p)^{n-i}. \\ &= \text{Really???!!\#\#...} \end{aligned}$$

Too hard!

Ok.. fine.

Variance: binomial.

$$
\begin{aligned}
E[X^2] &= \sum_{i=0}^{n} i^2 \binom{n}{i} p^i (1-p)^{n-i}. \\
&= \text{Really???!!\#\#...}
\end{aligned}
$$

Too hard!

Ok.. fine.
Let's do something else.

Variance: binomial.

$$\begin{aligned} E[X^2] &= \sum_{i=0}^{n} i^2 \binom{n}{i} p^i (1-p)^{n-i}. \\ &= \text{Really???!!\#\#...} \end{aligned}$$

Too hard!

Ok.. fine.
Let's do something else.
Maybe not much easier...

Variance: binomial.

$$
\begin{aligned}
E[X^2] &= \sum_{i=0}^{n} i^2 \binom{n}{i} p^i (1-p)^{n-i}. \\
&= \text{Really???!!\#\#...}
\end{aligned}
$$

Too hard!

Ok.. fine.
Let's do something else.
Maybe not much easier...but there is a payoff.

Properties of variance.

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

Properties of variance.

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

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.

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:

$$Var(cX) = E((cX)^2) - (E(cX))^2$$

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:

$$\begin{aligned} Var(cX) &= E((cX)^2) - (E(cX))^2 \\ &= c^2 E(X^2) - c^2 (E(X))^2 \end{aligned}$$

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:

$$\begin{aligned} Var(cX) &= E((cX)^2) - (E(cX))^2 \\ &= c^2 E(X^2) - c^2 (E(X))^2 = c^2 (E(X^2) - E(X)^2) \end{aligned}$$

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:

$$\begin{aligned} Var(cX) &= E((cX)^2) - (E(cX))^2 \\ &= c^2 E(X^2) - c^2 (E(X))^2 = c^2 (E(X^2) - E(X)^2) \\ &= c^2 Var(X) \end{aligned}$$

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:

$$\begin{aligned}
Var(cX) &= E((cX)^2) - (E(cX))^2 \\
&= c^2 E(X^2) - c^2 (E(X))^2 = c^2 (E(X^2) - E(X)^2) \\
&= c^2 Var(X) \\
Var(X + c) &= E((X + c - E(X + c))^2)
\end{aligned}$$

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:

$$\begin{aligned}
Var(cX) &= E((cX)^2) - (E(cX))^2 \\
&= c^2 E(X^2) - c^2(E(X))^2 = c^2(E(X^2) - E(X)^2) \\
&= c^2 Var(X) \\
Var(X + c) &= E((X + c - E(X + c))^2) \\
&= E((X + c - E(X) - c)^2)
\end{aligned}$$

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:

$$
\begin{aligned}
Var(cX) &= E((cX)^2) - (E(cX))^2 \\
&= c^2 E(X^2) - c^2 (E(X))^2 = c^2 (E(X^2) - E(X)^2) \\
&= c^2 Var(X) \\
Var(X + c) &= E((X + c - E(X + c))^2) \\
&= E((X + c - E(X) - c)^2) \\
&= E((X - E(X))^2)
\end{aligned}
$$

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:

$$\begin{aligned}
Var(cX) &= E((cX)^2) - (E(cX))^2 \\
&= c^2 E(X^2) - c^2(E(X))^2 = c^2(E(X^2) - E(X)^2) \\
&= c^2 Var(X) \\
Var(X + c) &= E((X + c - E(X + c))^2) \\
&= E((X + c - E(X) - c)^2) \\
&= E((X - E(X))^2) = Var(X)
\end{aligned}$$

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:

$$\begin{aligned}
Var(cX) &= E((cX)^2) - (E(cX))^2 \\
&= c^2 E(X^2) - c^2 (E(X))^2 = c^2 (E(X^2) - E(X)^2) \\
&= c^2 Var(X) \\
Var(X + c) &= E((X + c - E(X + c))^2) \\
&= E((X + c - E(X) - c)^2) \\
&= E((X - E(X))^2) = Var(X)
\end{aligned}$$

$\square$

Independent random variables.

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

Independent random variables.

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

Fact: $E[XY] = E[X]E[Y]$ for independent random variables.

Independent random variables.

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

Fact: $E[XY] = E[X]E[Y]$ for independent random variables.

$$\begin{aligned} E[XY] &= \sum_a \sum_b a \times b \times Pr[X = a, Y = b] \\ &= \sum_a \sum_b a \times b \times Pr[X = a]Pr[Y = b] \\ &= (\sum_a aPr[X = a])(\sum_b bPr[Y = b]) \\ &= E[X]E[Y] \end{aligned}$$

Variance of sum of two independent random variables

Variance of sum of two independent random variables

Theorem:
If X and Y are independent, then

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

Variance of sum of two independent random variables

Theorem:
If X and Y are independent, then

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

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

Variance of sum of two independent random variables

Theorem:
If X and Y are independent, then

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

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

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

Variance of sum of two independent random variables

Theorem:
If X and Y are independent, then

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

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,

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

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,

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

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,

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

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,

$$\begin{aligned} var(X+Y) &= E((X+Y)^2) = E(X^2 + 2XY + Y^2) \\ &= E(X^2) + 2E(XY) + E(Y^2) \end{aligned}$$

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,

$$\begin{aligned} var(X+Y) &= E((X+Y)^2) = E(X^2 + 2XY + Y^2) \\ &= E(X^2) + 2E(XY) + E(Y^2) = E(X^2) + E(Y^2) \end{aligned}$$

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,

$$\begin{aligned} var(X+Y) &= E((X+Y)^2) = E(X^2 + 2XY + Y^2) \\ &= E(X^2) + 2E(XY) + E(Y^2) = E(X^2) + E(Y^2) \\ &= var(X) + var(Y). \end{aligned}$$

Variance of sum of independent random variables

Variance of sum of independent random variables

Theorem:

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

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

Variance of sum of independent random variables

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

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

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

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

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

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

Then, by independence,

$$E[XY] = E[X]E[Y] = 0.$$

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 $E[X] = E[Y] = \cdots = 0$.

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

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

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

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

Then, by independence,

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

Hence,

$$var(X+Y+Z+\cdots) \quad = \quad E((X+Y+Z+\cdots)^2)$$

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 $E[X] = E[Y] = \cdots = 0$.

Then, by independence,

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

Hence,

$$\begin{aligned} var(X+Y+Z+\cdots) &= E((X+Y+Z+\cdots)^2) \\ &= E(X^2+Y^2+Z^2+\cdots+2XY+2XZ+2YZ+\cdots) \end{aligned}$$

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 $E[X] = E[Y] = \cdots = 0$.

Then, by independence,

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

Hence,

$$\begin{aligned} var(X + Y + Z + \cdots) &= E((X + Y + Z + \cdots)^2) \\ &= E(X^2 + Y^2 + Z^2 + \cdots + 2XY + 2XZ + 2YZ + \cdots) \\ &= E(X^2) + E(Y^2) + E(Z^2) + \cdots + 0 + \cdots + 0 \end{aligned}$$

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 $E[X] = E[Y] = \cdots = 0$.

Then, by independence,

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

Hence,

$$\begin{aligned}
var(X+Y+Z+\cdots) &= E((X+Y+Z+\cdots)^2) \\
&= E(X^2+Y^2+Z^2+\cdots+2XY+2XZ+2YZ+\cdots) \\
&= E(X^2)+E(Y^2)+E(Z^2)+\cdots+0+\cdots+0 \\
&= var(X)+var(Y)+var(Z)+\cdots.
\end{aligned}$$

$\square$

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?

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

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

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

Variance of Binomial Distribution.

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

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

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

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

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

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:

Variance of Binomial Distribution.

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

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

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

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

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

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: $Pr[X_i = 1 | X_j = 1] = Pr[X_i = 1]$.

$$Var(X) = Var(X_1 + \cdots X_n) = np(1-p).$$

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: $Pr[X_i = 1 | X_j = 1] = Pr[X_i = 1]$.

$$Var(X) = Var(X_1 + \cdots X_n) = np(1-p).$$

Poisson Distribution: Variance.

Definition Poisson Distribution with parameter $\lambda > 0$

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

Poisson Distribution: Variance.

Definition Poisson Distribution with parameter $\lambda > 0$

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

Mean, Variance?

Poisson Distribution: Variance.

Definition Poisson Distribution with parameter $\lambda > 0$

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

Mean, Variance?

Ugh.

Poisson Distribution: Variance.

Definition Poisson Distribution with parameter $\lambda > 0$

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

Mean, Variance?

Ugh.

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

Poisson Distribution: Variance.

Definition Poisson Distribution with parameter $\lambda > 0$

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

Mean, Variance?

Ugh.

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

Mean: $pn = \lambda$

Poisson Distribution: Variance.

Definition Poisson Distribution with parameter $\lambda > 0$

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

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 $\lambda > 0$

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

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 $\lambda > 0$

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

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 $\lambda > 0$

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

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 $\lambda > 0$

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

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 $E(X^2) = Var(X) + E(X)^2$.

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

Covariance

Definition The covariance of X and Y is

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

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

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

Fact

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

Proof:
Think about $E[X] = E[Y] = 0$. Just $E[XY]$.

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 $E[XY]$. $\square$

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 $E[XY]$. $\square$ish.

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 $E[XY]$. $\square$ish.

For the sake of completeness.

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 $E[XY]$. $\square$ish.

For the sake of completeness.

$$\begin{aligned}
E[(X - E[X])(Y - E[Y])] &= E[XY - E[X]Y - XE[Y] + E[X]E[Y]] \\
&= E[XY] - E[X]E[Y] - E[X]E[Y] + E[X]E[Y] \\
&= E[XY] - E[X]E[Y].
\end{aligned}$$

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: $(a - b)^2 > 0$

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

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

Simple case: $E[X] = E[Y] = 0$ and $E[X^2] = E[Y^2] = 1$.

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

Simple case: $E[X] = E[Y] = 0$ and $E[X^2] = E[Y^2] = 1$.

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

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

Simple case: $E[X] = E[Y] = 0$ and $E[X^2] = E[Y^2] = 1$.

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

$$E[(X-Y)^2] = E[X^2] + E[Y^2] - 2E[XY] = 2(1 - E[XY]) \geq 0$$

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

Simple case: $E[X] = E[Y] = 0$ and $E[X^2] = E[Y^2] = 1$.

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

$E[(X-Y)^2] = E[X^2] + E[Y^2] - 2E[XY] = 2(1 - E[XY]) \geq 0$
$\rightarrow E[XY] \leq 1$.

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

Simple case: $E[X] = E[Y] = 0$ and $E[X^2] = E[Y^2] = 1$.

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

$E[(X-Y)^2] = E[X^2] + E[Y^2] - 2E[XY] = 2(1 - E[XY]) \geq 0$
$\rightarrow E[XY] \leq 1$.

$E[(X+Y)^2] = E[X^2] + E[Y^2] + 2E[XY] = 2(1 + E[XY]) \geq 0$

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

Simple case: $E[X] = E[Y] = 0$ and $E[X^2] = E[Y^2] = 1$.

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

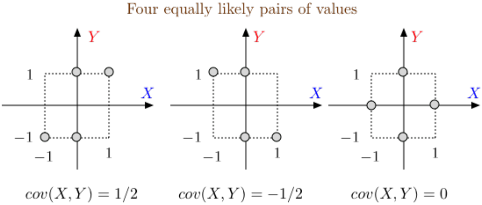
$E[(X-Y)^2] = E[X^2] + E[Y^2] - 2E[XY] = 2(1 - E[XY]) \geq 0$
$\rightarrow E[XY] \leq 1$.

$E[(X+Y)^2] = E[X^2] + E[Y^2] + 2E[XY] = 2(1 + E[XY]) \geq 0$
$\rightarrow E[XY] \geq -1$.

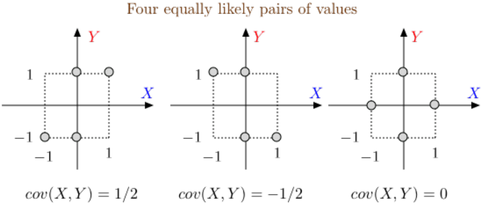
$\square$

Shifting and scaling doesn't change correlation.

Examples of Covariance

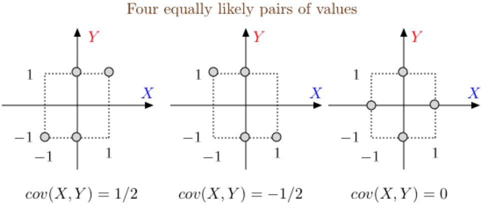


Examples of Covariance



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

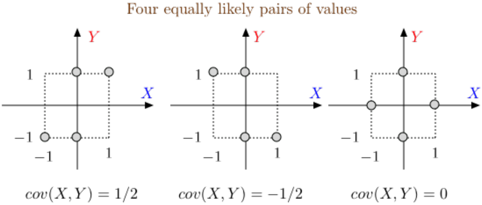
Examples of Covariance



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

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

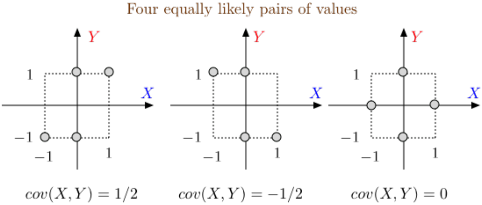
Examples of Covariance



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

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

Examples of Covariance

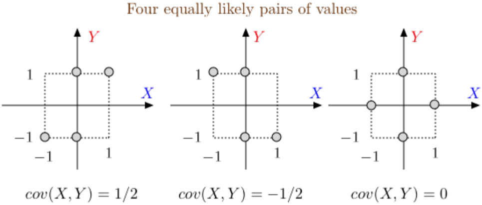


Note that $E[X] = 0$ and $E[Y] = 0$ in these examples. Then $cov(X,Y) = E[XY]$.

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

When $cov(X,Y) < 0$, when X is larger, Y tends to be smaller.

Examples of Covariance

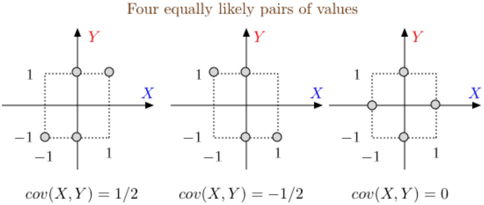


Note that $E[X] = 0$ and $E[Y] = 0$ in these examples. Then $cov(X, Y) = E[XY]$.

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

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

Examples of Covariance



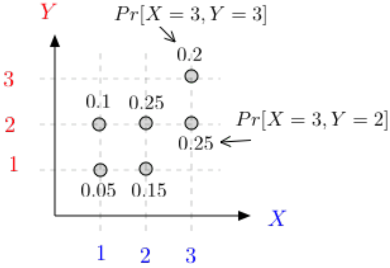
Note that $E[X] = 0$ and $E[Y] = 0$ in these examples. Then $cov(X, Y) = E[XY]$.

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

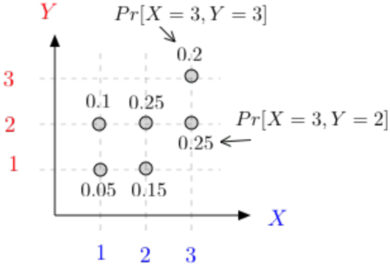
When $cov(X, Y) < 0$, when X is larger, Y tends to be smaller. X and Y are said to be negatively correlated.

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

Examples of Covariance

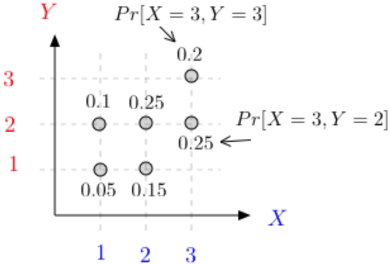


Examples of Covariance



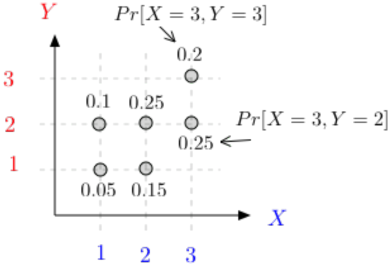
$$E[X] = 1 \times 0.15 + 2 \times 0.4 + 3 \times 0.45 = 2.3$$

Examples of Covariance



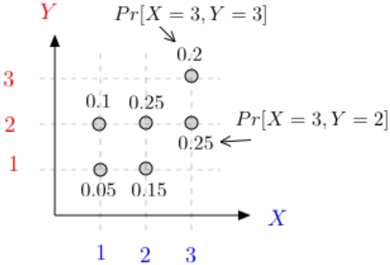
$$E[X] = 1 \times 0.15 + 2 \times 0.4 + 3 \times 0.45 = 2.3$$
$$E[X^2] = 1^2 \times 0.15 + 2^2 \times 0.4 + 3^2 \times 0.45 = 5.8$$

Examples of Covariance



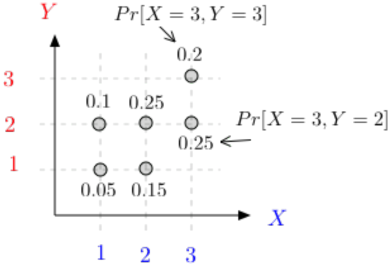
$$E[X] = 1 \times 0.15 + 2 \times 0.4 + 3 \times 0.45 = 2.3$$
$$E[X^2] = 1^2 \times 0.15 + 2^2 \times 0.4 + 3^2 \times 0.45 = 5.8$$
$$E[Y] = 1 \times 0.2 + 2 \times 0.6 + 3 \times 0.2 = 2$$

Examples of Covariance



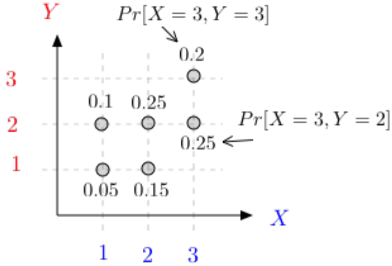
$$E[X] = 1 \times 0.15 + 2 \times 0.4 + 3 \times 0.45 = 2.3$$
$$E[X^2] = 1^2 \times 0.15 + 2^2 \times 0.4 + 3^2 \times 0.45 = 5.8$$
$$E[Y] = 1 \times 0.2 + 2 \times 0.6 + 3 \times 0.2 = 2$$
$$E[Y^2] = 1 \times 0.2 + 4 \times 0.6 + 9 \times 0.2 = 4.4$$

Examples of Covariance



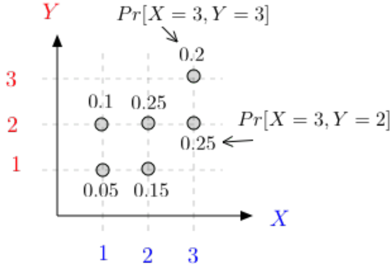
$$E[X] = 1 \times 0.15 + 2 \times 0.4 + 3 \times 0.45 = 2.3$$
$$E[X^2] = 1^2 \times 0.15 + 2^2 \times 0.4 + 3^2 \times 0.45 = 5.8$$
$$E[Y] = 1 \times 0.2 + 2 \times 0.6 + 3 \times 0.2 = 2$$
$$E[Y^2] = 1 \times 0.2 + 4 \times 0.6 + 9 \times 0.2 = 4.4$$
$$E[XY] = 1 \times 0.05 + 1 \times 2 \times 0.1 +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

Examples of Covariance



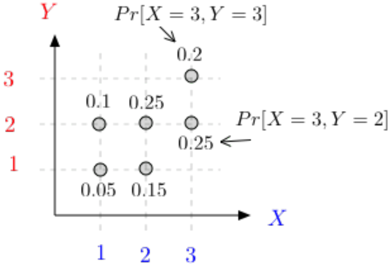
$$E[X] = 1 \times 0.15 + 2 \times 0.4 + 3 \times 0.45 = 2.3$$
$$E[X^2] = 1^2 \times 0.15 + 2^2 \times 0.4 + 3^2 \times 0.45 = 5.8$$
$$E[Y] = 1 \times 0.2 + 2 \times 0.6 + 3 \times 0.2 = 2$$
$$E[Y^2] = 1 \times 0.2 + 4 \times 0.6 + 9 \times 0.2 = 4.4$$
$$E[XY] = 1 \times 0.05 + 1 \times 2 \times 0.1 + \cdots + 3 \times 3 \times 0.2 = 4.85$$
$$cov(X, Y) = E[XY] - E[X]E[Y] = .25$$

Examples of Covariance



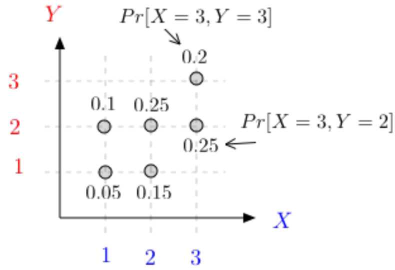
$$E[X] = 1 \times 0.15 + 2 \times 0.4 + 3 \times 0.45 = 2.3$$
$$E[X^2] = 1^2 \times 0.15 + 2^2 \times 0.4 + 3^2 \times 0.45 = 5.8$$
$$E[Y] = 1 \times 0.2 + 2 \times 0.6 + 3 \times 0.2 = 2$$
$$E[Y^2] = 1 \times 0.2 + 4 \times 0.6 + 9 \times 0.2 = 4.4$$
$$E[XY] = 1 \times 0.05 + 1 \times 2 \times 0.1 +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

Examples of Covariance



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

Examples of Covariance



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

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

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

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

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

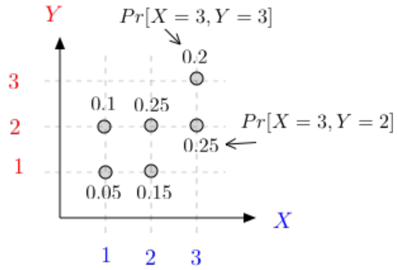
$$cov(X, Y) = E[XY] - E[X]E[Y] = .25$$

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

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

$$corr(X, Y) \approx 0.55$$

Examples of Covariance



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

Properties of Covariance

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

Properties of Covariance

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

Fact
(a) $var[X] = cov(X, X)$

Properties of Covariance

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

Fact
(a) $var[X] = cov(X, X)$
(b) X, Y independent $\Rightarrow cov(X, Y) =$

Properties of Covariance

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

Fact
(a) $var[X] = cov(X,X)$
(b) X, Y independent $\Rightarrow cov(X,Y) = 0$

Properties of Covariance

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

Fact

(a) $var[X] = cov(X,X)$

(b) X,Y independent $\Rightarrow cov(X,Y) = 0$

(c) $cov(a+X, b+Y) = cov(X,Y)$

Properties of Covariance

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

Fact

(a) $var[X] = cov(X,X)$

(b) X, Y independent $\Rightarrow cov(X,Y) = 0$

(c) $cov(a+X, b+Y) = cov(X,Y)$

(d) $$cov(aX+bY, cU+dV) = ac \cdot cov(X,U) + ad \cdot cov(X,V) + bc \cdot cov(Y,U) + bd \cdot cov(Y,V).$$

Properties of Covariance

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

Fact

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

Proof:

(a)-(b)-(c) are obvious.

Properties of Covariance

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

Fact

(a) $var[X] = cov(X,X)$

(b) X, Y independent $\Rightarrow cov(X,Y) = 0$

(c) $cov(a + X, b + Y) = cov(X,Y)$

(d) $cov(aX + bY, cU + dV) = ac \cdot cov(X,U) + ad \cdot cov(X,V)$
$\qquad + bc \cdot cov(Y,U) + bd \cdot cov(Y,V).$

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

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

Fact

(a) $var[X] = cov(X,X)$

(b) X, Y independent $\Rightarrow cov(X,Y) = 0$

(c) $cov(a+X, b+Y) = cov(X,Y)$

(d) $cov(aX+bY, cU+dV) = ac \cdot cov(X,U) + ad \cdot cov(X,V) + bc \cdot cov(Y,U) + bd \cdot cov(Y,V).$

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,

$$cov(aX+bY, cU+dV) = E[(aX+bY)(cU+dV)]$$

Properties of Covariance

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

Fact

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

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,

$$\begin{aligned} cov(aX+bY, cU+dV) &= E[(aX+bY)(cU+dV)] \\ &= ac \cdot E[XU] + ad \cdot E[XV] + bc \cdot E[YU] + bd \cdot E[YV] \end{aligned}$$

Properties of Covariance

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

Fact

(a) $var[X] = cov(X,X)$

(b) X, Y independent $\Rightarrow cov(X,Y) = 0$

(c) $cov(a+X, b+Y) = cov(X,Y)$

(d) $cov(aX+bY, cU+dV) = ac \cdot cov(X,U) + ad \cdot cov(X,V)$
$+ bc \cdot cov(Y,U) + bd \cdot cov(Y,V).$

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,

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

Properties of Covariance

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

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

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,

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

$\square$

Summary

$\boxed{\text{Variance}}$

Summary

$\boxed{\text{Variance}}$

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

Summary

Variance

- Variance: $var[X] := E[(X - E[X])^2] = E[X^2] - E[X]^2$
- Fact: $var[aX + b]a^2 var[X]$

Summary

$\boxed{\text{Variance}}$

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

Summary

$\boxed{\text{Variance}}$

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

Random Variables so far.

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

Random Variables so far.

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

Random Variables so far.

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

Random Variables so far.

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

Random Variables so far.

Probability Space: Ω, $Pr : \Omega \to [0,1]$, $\sum_{\omega \in \Omega} Pr(w) = 1$.
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: $E[X] = \sum_a a Pr[X = a] = \sum_{\omega \in \Omega} X(\omega) Pr(\omega)$.

Random Variables so far.

Probability Space: Ω, $Pr : \Omega \to [0,1]$, $\sum_{\omega \in \Omega} Pr(w) = 1$.
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: $E[X] = \sum_a aPr[X = a] = \sum_{\omega \in \Omega} X(\omega)Pr(\omega)$.
Linearity: $E[X + Y] = E[X] + E[Y]$.

Random Variables so far.

Probability Space: Ω, $Pr : \Omega \to [0,1]$, $\sum_{\omega \in \Omega} Pr(w) = 1$.
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: $E[X] = \sum_a aPr[X = a] = \sum_{\omega \in \Omega} X(\omega)Pr(\omega)$.
Linearity: $E[X + Y] = E[X] + E[Y]$.

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

Random Variables so far.

Probability Space: Ω, $Pr : \Omega \to [0,1]$, $\sum_{\omega \in \Omega} Pr(w) = 1$.
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: $E[X] = \sum_a a Pr[X = a] = \sum_{\omega \in \Omega} X(\omega) Pr(\omega)$.
  Linearity: $E[X + Y] = E[X] + E[Y]$.

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: Ω, $Pr : \Omega \to [0,1]$, $\sum_{\omega \in \Omega} Pr(w) = 1$.
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: $E[X] = \sum_a a Pr[X = a] = \sum_{\omega \in \Omega} X(\omega) Pr(\omega)$.
  Linearity: $E[X + Y] = E[X] + E[Y]$.

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

Random Variables so far.

Probability Space: Ω, $Pr : \Omega \to [0,1]$, $\sum_{\omega \in \Omega} Pr(w) = 1$.
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: $E[X] = \sum_a aPr[X = a] = \sum_{\omega \in \Omega} X(\omega)Pr(\omega)$.
  Linearity: $E[X+Y] = E[X] + E[Y]$.

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: $X \sim P(\lambda)$

Random Variables so far.

Probability Space: Ω, $Pr : \Omega \to [0,1]$, $\sum_{\omega \in \Omega} Pr(w) = 1$.
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: $E[X] = \sum_a a Pr[X = a] = \sum_{\omega \in \Omega} X(\omega) Pr(\omega)$.
  Linearity: $E[X+Y] = E[X] + E[Y]$.

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: Ω, $Pr : \Omega \to [0,1]$, $\sum_{\omega \in \Omega} Pr(w) = 1$.
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: $E[X] = \sum_a aPr[X = a] = \sum_{\omega \in \Omega} X(\omega) Pr(\omega)$.
  Linearity: $E[X + Y] = E[X] + E[Y]$.

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

Random Variables so far.

Probability Space: Ω, $Pr : \Omega \to [0,1]$, $\sum_{\omega \in \Omega} Pr(w) = 1$.
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: $E[X] = \sum_a aPr[X = a] = \sum_{\omega \in \Omega} X(\omega) Pr(\omega)$.
Linearity: $E[X + Y] = E[X] + E[Y]$.

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

Random Variables so far.

Probability Space: Ω, $Pr : \Omega \rightarrow [0,1]$, $\sum_{\omega \in \Omega} Pr(w) = 1$.
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: $E[X] = \sum_a a Pr[X = a] = \sum_{\omega \in \Omega} X(\omega) Pr(\omega)$.
Linearity: $E[X+Y] = E[X] + E[Y]$.

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: Ω, $Pr : \Omega \to [0,1]$, $\sum_{\omega \in \Omega} Pr(w) = 1$.
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: $E[X] = \sum_a aPr[X = a] = \sum_{\omega \in \Omega} X(\omega)Pr(\omega)$.
Linearity: $E[X + Y] = E[X] + E[Y]$.

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: Ω, $Pr : \Omega \to [0,1]$, $\sum_{\omega \in \Omega} Pr(w) = 1$.
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: $E[X] = \sum_a aPr[X = a] = \sum_{\omega \in \Omega} X(\omega)Pr(\omega)$.
Linearity: $E[X + Y] = E[X] + E[Y]$.

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: $X \sim G(p)$

Random Variables so far.

Probability Space: Ω, $Pr : \Omega \to [0,1]$, $\sum_{\omega \in \Omega} Pr(w) = 1$.
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: $E[X] = \sum_a aPr[X = a] = \sum_{\omega \in \Omega} X(\omega)Pr(\omega)$.
  Linearity: $E[X + Y] = E[X] + E[Y]$.

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: $X \sim G(p)$ $E(X) = \frac{1}{p}$, $Var(X) = \frac{1-p}{p^2}$