Introduction to Probability
Lecture 7: Independence, Covariance and Correlation
Mateja Jamnik, Thomas Sauerwald

University of Cambridge, Department of Computer Science and Technology
e-mail: {mateja.jamnik,thomas.sauerwald}@cl.cam.ac.uk
Independence of Random Variables

Definition of Independence

Two random variables $X$ and $Y$ are independent if for all values $a, b$:

$$P [ X \leq a, Y \leq b ] = P [ X \leq a ] \cdot P [ Y \leq b ].$$
Independence of Random Variables

Definition of Independence

Two random variables $X$ and $Y$ are independent if for all values $a, b$:

$$P[X \leq a, Y \leq b] = P[X \leq a] \cdot P[Y \leq b].$$

For two discrete random variables, an equivalent definition is:

$$P[X = a, Y = b] = P[X = a] \cdot P[Y = b].$$
Independence of Random Variables

Definition of Independence

Two random variables $X$ and $Y$ are independent if for all values $a, b$:

$$P[X \leq a, Y \leq b] = P[X \leq a] \cdot P[Y \leq b].$$

For two discrete random variables, an equivalent definition is:

$$P[X = a, Y = b] = P[X = a] \cdot P[Y = b].$$

This is useless for continuous random variables.

Remark
All these definitions extend in the natural way to more than two variables!
Independence of Random Variables

Two random variables $X$ and $Y$ are independent if for all values $a, b$:

$$ P[X \leq a, Y \leq b] = P[X \leq a] \cdot P[Y \leq b]. $$

Definition of Independence

For two discrete random variables, an equivalent definition is:

$$ P[X = a, Y = b] = P[X = a] \cdot P[Y = b]. $$

This is useless for continuous random variables.

This definition covers the discrete and continuous case!
Independence of Random Variables

Two random variables $X$ and $Y$ are independent if for all values $a, b$:

$$\mathbb{P}[X \leq a, Y \leq b] = \mathbb{P}[X \leq a] \cdot \mathbb{P}[Y \leq b].$$

Definition of Independence

For two discrete random variables, an equivalent definition is:

$$\mathbb{P}[X = a, Y = b] = \mathbb{P}[X = a] \cdot \mathbb{P}[Y = b].$$

This is useless for continuous random variables.

Remark

Using the joint probability distribution, the above is equivalent to for all $a, b$,

$$F(a, b) = F_X(a) \cdot F_Y(b).$$
Independence of Random Variables

Definition of Independence

Two random variables $X$ and $Y$ are independent if for all values $a, b$:

$$P[X \leq a, Y \leq b] = P[X \leq a] \cdot P[Y \leq b].$$

For two discrete random variables, an equivalent definition is:

$$P[X = a, Y = b] = P[X = a] \cdot P[Y = b].$$

This is useless for continuous random variables.

Remark

Using the joint probability distribution, the above is equivalent to for all $a, b$,

$$F(a, b) = F_X(a) \cdot F_Y(b).$$

All these definitions extend in the natural way to more than two variables!
Factorisation

The definition of independence of \( X \) and \( Y \) implies the following factorisation formula: for any “suitable” sets \( A \) and \( B \),

\[
P[ X \in A, Y \in B ] = P[ X \in A ] \cdot P[ Y \in B ]
\]
The definition of independence of \( X \) and \( Y \) implies the following factorisation formula: for any “suitable” sets \( A \) and \( B \),

\[
P [ X \in A, Y \in B ] = P [ X \in A ] \cdot P [ Y \in B ]
\]

For continuous distributions one obtains by differentiating both sides in the formula for the joint distribution:

\[
f_{X, Y}(x, y) = f_X(x) \cdot f_Y(y)
\]
Factorisation

The definition of independence of $X$ and $Y$ implies the following factorisation formula: for any “suitable” sets $A$ and $B$,

$$P[X \in A, Y \in B] = P[X \in A] \cdot P[Y \in B]$$

For continuous distributions one obtains by differentiating both sides in the formula for the joint distribution:

$$f_{X,Y}(x,y) = f_X(x) \cdot f_Y(y)$$

Example

Let $X$ and $Y$ be two independent variables. Let $I = (a, b]$ be any interval and define $U := 1_{X \in I}$ and $V := 1_{Y \in I}$. Prove $U$ and $V$ are independent.
A table is ruled with equidistant, parallel lines a distance $D$ apart.
Buffon’s Needle Problem (1/2)

- A table is ruled with equidistant, parallel lines a distance $D$ apart.
- A needle of length $L$ is thrown randomly on the table.
A table is ruled with equidistant, parallel lines a distance $D$ apart.

A needle of length $L$ is thrown randomly on the table.

What is the probability that the needle will intersect one of the two lines?
Buffon’s Needle Problem (1/2)

- A table is ruled with equidistant, parallel lines a distance $D$ apart.
- A needle of length $L$ is thrown randomly on the table.
- What is the probability that the needle will intersect one of the two lines?

Source: Ross, Probability 8th ed.
Buffon’s Needle Problem (1/2)

- A table is ruled with equidistant, parallel lines a distance $D$ apart.
- A needle of length $L$ is thrown randomly on the table.
- What is the probability that the needle will intersect one of the two lines?

Let $X$ be the distance of the middle point of the needle to the closest parallel line. Needle intersects a line if hypotenuse of the triangle is less than $L/2$, i.e.,

$$\frac{X}{\cos(\theta)} < \frac{L}{2} \iff X < \frac{L}{2} \cos(\theta).$$
Buffon’s Needle Problem (1/2)

- A table is ruled with equidistant, parallel lines a distance \( D \) apart.
- A needle of length \( L \) is thrown randomly on the table.
- What is the probability that the needle will intersect one of the two lines?

Let \( X \) be the distance of the middle point of the needle to the closest parallel line. Needle intersects a line if hypotenuse of the triangle is less than \( L/2 \), i.e.,

\[
\frac{X}{\cos(\theta)} < \frac{L}{2} \quad \iff \quad X < \frac{L}{2} \cos(\theta).
\]

We assume that \( X \in [0, D/2] \) and \( \theta \in [0, \pi/2] \) are independent and uniform.
Buffon’s Needle Problem (1/2)

- A table is ruled with equidistant, parallel lines a distance $D$ apart.
- A needle of length $L$ is thrown randomly on the table.
- What is the probability that the needle will intersect one of the two lines?

Let $X$ be the distance of the middle point of the needle to the closest parallel line. Needle intersects a line if hypotenuse of the triangle is less than $L/2$, i.e.,

$$\frac{X}{\cos(\theta)} < \frac{L}{2} \iff X < \frac{L}{2} \cos(\theta).$$

We assume that $X \in [0, D/2]$ and $\theta \in [0, \pi/2]$ are independent and uniform.

Can be thought of as: 1. Sample the middle point of needle, 2. Sample the angle.
Buffon’s Needle Problem (2/2)

Let us compute the probability that the line intersects:

\[ P \left[ X < \frac{L}{2} \cdot \cos(\theta) \right] \]
Let us compute the probability that the line intersects:

$$P \left[ X < \frac{L}{2} \cdot \cos(\theta) \right] = \int \int_{x < (L/2) \cos y} f_{X,\theta}(x, y) \, dx \, dy$$
Let us compute the probability that the line intersects:

\[
P \left[ X < \frac{L}{2} \cdot \cos(\theta) \right] = \int \int f_{X,\theta}(x, y) \, dx \, dy \\
= \int \int f_X(x) f_\theta(y) \, dx \, dy
\]
Let us compute the probability that the line intersects:

\[
P \left[ X < \frac{L}{2} \cdot \cos(\theta) \right] = \iint_{x < (L/2) \cos y} f_{X,\theta}(x, y) \, dx \, dy
\]

\[
= \iint_{x < (L/2) \cos y} f_X(x) f_{\theta}(y) \, dx \, dy
\]

\[
= \frac{4}{\pi D} \int_0^{\pi/2} \int_0^{L/2 \cos(y)} dx \, dy
\]
Let us compute the probability that the line intersects:

\[
P \left[ X < \frac{L}{2} \cdot \cos(\theta) \right] = \iint_{x < (L/2) \cos y} f_{X,\theta}(x, y) \, dx \, dy
\]

\[
= \iint_{x < (L/2) \cos y} f_X(x) f_{\theta}(y) \, dx \, dy
\]

\[
= \frac{4}{\pi D} \int_0^{\pi/2} \int_0^{L/2 \cos(y)} dxdy
\]

\[
= \frac{4}{\pi D} \int_0^{\pi/2} \frac{L}{2} \cos(y) dy
\]
Let us compute the probability that the line intersects:

\[
P \left[ X < \frac{L}{2} \cdot \cos(\theta) \right] = \int \int_{x < (L/2) \cos y} f_{X,\theta}(x, y) \, dx \, dy
\]

\[
= \int \int_{x < (L/2) \cos y} f_X(x) f_\theta(y) \, dx \, dy
\]

\[
= \frac{4}{\pi D} \int_0^{\pi/2} \int_0^{L/2 \cos(y)} dx \, dy
\]

\[
= \frac{4}{\pi D} \int_0^{\pi/2} \frac{L}{2} \cos(y) \, dy
\]

\[
= \frac{2L}{\pi D}.
\]
Buffon’s Needle Problem (2/2)

Let us compute the probability that the line intersects:

\[
P \left[ X < \frac{L}{2} \cdot \cos(\theta) \right] = \int \int_{x<(L/2) \cos y} f_{X,\theta}(x, y) \, dx \, dy
\]

\[
= \int \int_{x<(L/2) \cos y} f_X(x) f_{\theta}(y) \, dx \, dy
\]

\[
= \frac{4}{\pi D} \int_0^{\pi/2} \int_0^{L/2 \cos(y)} dx \, dy
\]

\[
= \frac{4}{\pi D} \int_0^{\pi/2} L/2 \cos(y) \, dy
\]

\[
= \frac{2L}{\pi D}.
\]

This gives us a method to estimate \(\pi\)!
Covariance

**Definition of Covariance**

Let $X$ and $Y$ be two random variables. The covariance is defined as:

$$\text{Cov}[X, Y] = \mathbb{E}[(X - \mathbb{E}[X]) \cdot (Y - \mathbb{E}[Y])].$$
Covariance

Let $X$ and $Y$ be two random variables. The covariance is defined as:

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

Interpretation:

If $\text{Cov}[X, Y] > 0$ and $X$ has a realisation larger (smaller) than $E[X]$, then $Y$ will likely have a realisation larger (smaller) than $E[Y]$.

If $\text{Cov}[X, Y] < 0$, then it is the other way around.

Using the linearity of expectation rule, one has the equivalent definition:

$$\text{Cov}[X, Y] = E[X \cdot Y] - E[X] \cdot E[Y].$$

Alternative Formula

Note that $\text{Cov}[X, X] = V[X]$.

Two variables $X, Y$ with $\text{Cov}[X, Y] > 0$ are positively correlated.

Two variables $X, Y$ with $\text{Cov}[X, Y] < 0$ are negatively correlated.

Two variables $X, Y$ with $\text{Cov}[X, Y] = 0$ are uncorrelated.
Covariance

Definition of Covariance

Let $X$ and $Y$ be two random variables. The covariance is defined as:

$$\text{Cov}[X, Y] = \mathbb{E}[(X - \mathbb{E}[X]) \cdot (Y - \mathbb{E}[Y])].$$

Interpretation:

- If $\text{Cov}[X, Y] > 0$ and $X$ has a realisation larger (smaller) than $\mathbb{E}[X]$, then $Y$ will likely have a realisation larger (smaller) than $\mathbb{E}[Y]$. 

Two variables $X$, $Y$ with $\text{Cov}[X, Y] > 0$ are positively correlated.

Two variables $X$, $Y$ with $\text{Cov}[X, Y] < 0$ are negatively correlated.

Two variables $X$, $Y$ with $\text{Cov}[X, Y] = 0$ are uncorrelated.
Covariance

Definition of Covariance

Let $X$ and $Y$ be two random variables. The covariance is defined as:

$$\text{Cov}[X, Y] = \mathbb{E}[(X - \mathbb{E}[X]) \cdot (Y - \mathbb{E}[Y])].$$

Interpretation:
- If $\text{Cov}[X, Y] > 0$ and $X$ has a realisation larger (smaller) than $\mathbb{E}[X]$, then $Y$ will likely have a realisation larger (smaller) than $\mathbb{E}[Y]$.
- If $\text{Cov}[X, Y] < 0$, then it is the other way around.

Two variables $X, Y$ with $\text{Cov}[X, Y] > 0$ are positively correlated.

Two variables $X, Y$ with $\text{Cov}[X, Y] < 0$ are negatively correlated.

Two variables $X, Y$ with $\text{Cov}[X, Y] = 0$ are uncorrelated.
Covariance

Definition of Covariance

Let $X$ and $Y$ be two random variables. The covariance is defined as:

$$\text{Cov}[X, Y] = \mathbb{E}[(X - \mathbb{E}[X]) \cdot (Y - \mathbb{E}[Y])].$$

Interpretation:

- If $\text{Cov}[X, Y] > 0$ and $X$ has a realisation larger (smaller) than $\mathbb{E}[X]$, then $Y$ will likely have a realisation larger (smaller) than $\mathbb{E}[Y]$.
- If $\text{Cov}[X, Y] < 0$, then it is the other way around.

Alternative Formula

Using the linearity of expectation rule, one has the equivalent definition:

$$\text{Cov}[X, Y] = \mathbb{E}[X \cdot Y] - \mathbb{E}[X] \cdot \mathbb{E}[Y].$$
Covariance

Definition of Covariance

Let $X$ and $Y$ be two random variables. The covariance is defined as:

$$\text{Cov}[X, Y] = \mathbb{E}[(X - \mathbb{E}[X]) \cdot (Y - \mathbb{E}[Y])].$$

**Interpretation:**
- If $\text{Cov}[X, Y] > 0$ and $X$ has a realisation larger (smaller) than $\mathbb{E}[X]$, then $Y$ will likely have a realisation larger (smaller) than $\mathbb{E}[Y]$.
- If $\text{Cov}[X, Y] < 0$, then it is the other way around.

Alternative Formula

Using the linearity of expectation rule, one has the equivalent definition:

$$\text{Cov}[X, Y] = \mathbb{E}[X \cdot Y] - \mathbb{E}[X] \cdot \mathbb{E}[Y].$$

- Note that $\text{Cov}[X, X] = \text{Var}[X]$. 

Intro to Probability 6
Covariance

Definition of Covariance

Let $X$ and $Y$ be two random variables. The covariance is defined as:

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

Interpretation:

- If $\text{Cov}[X, Y] > 0$ and $X$ has a realisation larger (smaller) than $E[X]$, then $Y$ will likely have a realisation larger (smaller) than $E[Y]$.
- If $\text{Cov}[X, Y] < 0$, then it is the other way around.

Alternative Formula

Using the linearity of expectation rule, one has the equivalent definition:

$$\text{Cov}[X, Y] = E[X \cdot Y] - E[X] \cdot E[Y].$$

- Note that $\text{Cov}[X, X] = \text{V}[X]$.
- Two variables $X, Y$ with $\text{Cov}[X, Y] > 0$ are positively correlated.
**Covariance**

---

### Definition of Covariance

Let $X$ and $Y$ be two random variables. The **covariance** is defined as:

$$ \text{Cov}[X, Y] = \mathbb{E}[(X - \mathbb{E}[X]) \cdot (Y - \mathbb{E}[Y])] $$

---

**Interpretation:**

- If $\text{Cov}[X, Y] > 0$ and $X$ has a realisation larger (smaller) than $\mathbb{E}[X]$, then $Y$ will likely have a realisation larger (smaller) than $\mathbb{E}[Y]$.
- If $\text{Cov}[X, Y] < 0$, then it is the other way around.

---

### Alternative Formula

Using the linearity of expectation rule, one has the equivalent definition:

$$ \text{Cov}[X, Y] = \mathbb{E}[X \cdot Y] - \mathbb{E}[X] \cdot \mathbb{E}[Y]. $$

---

- Note that $\text{Cov}[X, X] = \text{Var}[X]$.
- Two variables $X, Y$ with $\text{Cov}[X, Y] > 0$ are **positively correlated**.
- Two variables $X, Y$ with $\text{Cov}[X, Y] < 0$ are **negatively correlated**.
Covariance

Definition of Covariance

Let \( X \) and \( Y \) be two random variables. The covariance is defined as:

\[
\text{Cov}[X, Y] = \mathbb{E}[(X - \mathbb{E}[X]) \cdot (Y - \mathbb{E}[Y])].
\]

Interpretation:

- If \( \text{Cov}[X, Y] > 0 \) and \( X \) has a realisation larger (smaller) than \( \mathbb{E}[X] \), then \( Y \) will likely have a realisation larger (smaller) than \( \mathbb{E}[Y] \).
- If \( \text{Cov}[X, Y] < 0 \), then it is the other way around.

Alternative Formula

Using the linearity of expectation rule, one has the equivalent definition:

\[
\text{Cov}[X, Y] = \mathbb{E}[XY] - \mathbb{E}[X] \cdot \mathbb{E}[Y].
\]

- Note that \( \text{Cov}[X, X] = \mathbb{V}[X] \).
- Two variables \( X, Y \) with \( \text{Cov}[X, Y] > 0 \) are positively correlated.
- Two variables \( X, Y \) with \( \text{Cov}[X, Y] < 0 \) are negatively correlated.
- Two variables \( X, Y \) with \( \text{Cov}[X, Y] = 0 \) are uncorrelated.
Illustration of 3 Cases for Cov \([ X, Y ]\)

500 outcomes of randomly generated pairs of RVs \((X, Y)\) with different joint distributions

1. What is the covariance (positive, negative, neutral)?

2. Where is the covariance the largest (in magnitude)?

*Fig. 10.1. Some scatterplots.*
Source: Textbook by Dekking
Independence implies Uncorrelated

Let $X$ and $Y$ be two independent random variables. Then $X$ and $Y$ are uncorrelated, i.e., $\text{Cov} [X, Y] = 0$. We give a proof for the discrete case:
Uncorrelated may not imply Independence

Example

Find a (simple) example of two random variables $X$ and $Y$ which are uncorrelated but dependent.

Answer

Let $X$ be uniformly sampled from $\{-1, 0, +1\}$ and $Y := 1_{X=0}$.

$\Rightarrow X \cdot Y = 0$ (for all outcomes), and thus $E[X \cdot Y] = 0$.

Further, $E[X] = 0$ (and $E[Y] = 1/3$), and hence:

$\text{Cov}[X, Y] = E[X \cdot Y] - E[X] \cdot E[Y] = 0$.

On the other hand, $P[X=0] = 1/3$ and $P[Y=0] = 2/3$, and thus $1 = P[X \cdot Y = 0] > P[X=0] \cdot P[Y=0] = 2/9$. 

Example Intro to Probability 9
Uncorrelated may not imply Independence

Example

Find a (simple) example of two random variables $X$ and $Y$ which are uncorrelated but dependent.

Let $X$ be uniformly sampled from $\{-1, 0, +1\}$ and $Y := 1_{X=0}$.

Answer

- Let $X$ be uniformly sampled from $\{-1, 0, +1\}$ and $Y := 1_{X=0}$. 

Find a (simple) example of two random variables $X$ and $Y$ which are un-correlated but dependent.

- Let $X$ be uniformly sampled from $\{-1, 0, +1\}$ and $Y := 1_{X=0}$.
  \[ X \cdot Y = 0 \text{ (for all outcomes), and thus} \]

\[ \Rightarrow \text{ } \]
Uncorrelated may not imply Independence

Example

Find a (simple) example of two random variables $X$ and $Y$ which are uncorrelated but dependent.

- Let $X$ be uniformly sampled from $\{-1, 0, +1\}$ and $Y := 1_{X=0}$. 
  \[ X \cdot Y = 0 \text{ (for all outcomes)}, \text{ and thus} \]
  \[ \mathbb{E}[X \cdot Y] = 0. \]
Uncorrelated may not imply Independence

Example

Find a (simple) example of two random variables $X$ and $Y$ which are uncorrelated but dependent.

- Let $X$ be uniformly sampled from $\{-1, 0, +1\}$ and $Y := 1_{X=0}$.
  $\Rightarrow X \cdot Y = 0$ (for all outcomes), and thus
  \[
  \mathbf{E}[X \cdot Y] = 0.
  \]

- Further, $\mathbf{E}[X] = 0$ (and $\mathbf{E}[Y] = 1/3$), and hence:
Uncorrelated may not imply Independence

Example

Find a (simple) example of two random variables $X$ and $Y$ which are uncorrelated but dependent.

Answer

- Let $X$ be uniformly sampled from $\{-1, 0, +1\}$ and $Y := 1_{X=0}$.
- $\Rightarrow X \cdot Y = 0$ (for all outcomes), and thus $E[X \cdot Y] = 0$.

- Further, $E[X] = 0$ (and $E[Y] = 1/3$), and hence:

$$\text{Cov}[X, Y] = E[X \cdot Y] - E[X] \cdot E[Y] = 0.$$
Find a (simple) example of two random variables $X$ and $Y$ which are un-correlated but dependent.

- Let $X$ be uniformly sampled from $\{-1, 0, +1\}$ and $Y := 1_{X=0}$.
- $X \cdot Y = 0$ (for all outcomes), and thus
  \[ E[X \cdot Y] = 0. \]
- Further, $E[X] = 0$ (and $E[Y] = 1/3$), and hence:
  \[ \text{Cov}[X, Y] = E[X \cdot Y] - E[X] \cdot E[Y] = 0. \]
- On the other hand, $P[X = 0] = 1/3$ and $P[Y = 0] = 2/3$, and thus
Uncorrelated may not imply Independence

Example

Find a (simple) example of two random variables \(X\) and \(Y\) which are uncorrelated but dependent.

**Answer**

- Let \(X\) be uniformly sampled from \(\{-1, 0, +1\}\) and \(Y := 1_{X=0}\).

\[X \cdot Y = 0\] (for all outcomes), and thus

\[E[ X \cdot Y ] = 0.\]

- Further, \(E[ X ] = 0\) (and \(E[ Y ] = 1/3\)), and hence:

\[\text{Cov}[X, Y] = E[ X \cdot Y ] - E[ X ] \cdot E[ Y ] = 0.\]

- On the other hand, \(P[ X = 0 ] = 1/3\) and \(P[ Y = 0 ] = 2/3\), and thus

\[1 = P[ X \cdot Y = 0 ] > P[ X = 0 ] \cdot P[ Y = 0 ] = 2/9.\]

Variance of Sums and Covariances

Variance of Sum Formula

For any two random variables \( X, Y \),

\[
V[X + Y] = V[X] + V[Y] + 2 \cdot \text{Cov}[X, Y].
\]

Hence if \( X \) and \( Y \) are uncorrelated variables,

\[
\]

For any random variables \( X_1, X_2, \ldots, X_n \):

\[
V[nX_i] = nV[X_i] + 2 \cdot n \sum_{i=1}^{n} \text{Cov}[X_i, X_{i+1}],
\]
Variance of Sums and Covariances

- **Variance of Sum Formula**
  - For any two random variables $X$, $Y$,
    \[
    V[X + Y] = V[X] + V[Y] + 2 \cdot \text{Cov}[X, Y].
    \]
  - Hence if $X$ and $Y$ are uncorrelated variables,
    \[
    \]
Variance of Sums and Covariances

Variance of Sum Formula

- For any two random variables $X$, $Y$,

$$V[X + Y] = V[X] + V[Y] + 2 \cdot \text{Cov}[X, Y].$$

- Hence if $X$ and $Y$ are uncorrelated variables,


Generalisation of the case where $X$ and $Y$ are even independent!
Variance of Sums and Covariances

Variance of Sum Formula

- For any two random variables $X$, $Y$,

$$V[X + Y] = V[X] + V[Y] + 2 \cdot \text{Cov}[X, Y].$$

- Hence if $X$ and $Y$ are uncorrelated variables,


- For any random variables $X_1, X_2, \ldots, X_n$:

$$V\left[ \sum_{i=1}^{n} X_i \right] = \sum_{i=1}^{n} V[X_i] + 2 \cdot \sum_{i=1}^{n} \sum_{j=i+1}^{n} \text{Cov}[X_i, X_j].$$

Generalisation of the case where $X$ and $Y$ are even independent!
Recall the example where $X \in \{-1, 0, +1\}$ uniformly and $Y := 1_{X=0}$. Compute $V[X + Y]$.

Answer

\[ V[X] = \frac{1}{3} \cdot (1 - 1)^2 + \frac{2}{3} \cdot (0 - 1)^2 = \frac{2}{3} \]

\[ V[Y] = \frac{1}{3} \cdot (1 - 0)^2 + \frac{2}{3} \cdot (0 - 0)^2 = \frac{1}{3} \]

\[ V[X + Y] = V[X] + V[Y] + 2 \cdot Cov[X, Y] = \frac{2}{3} + \frac{1}{3} + 0 = \frac{8}{9} \]
Correlation Coefficient: Normalising the Covariance

The definition of covariance is not scaling invariant:

\[ \text{Cov}[X, Y] \text{ increases by a factor of } \alpha. \]

\[ \Rightarrow \text{Even if } X \text{ and } Y \text{ both increase by } \alpha, \text{ then } \text{Cov}[X, Y] \text{ will change.} \]

(Exercise: It changes by?)

Let \( X \) and \( Y \) be two random variables. The correlation coefficient \( \rho(X, Y) \)

is defined as:

\[ \rho(X, Y) = \frac{\text{Cov}[X, Y]}{\text{Var}[X] \cdot \text{Var}[Y]}. \]

If \( \text{Var}[X] = 0 \) or \( \text{Var}[Y] = 0 \), then it is defined as 0.

Correlation Coefficient Properties:

1. The correlation coefficient is scaling-invariant, i.e.,
   \[ \rho(X, Y) = \rho(\alpha \cdot X, \beta \cdot Y) \]
   for any \( \alpha, \beta > 0 \).

2. For any two random variables \( X, Y \),
   \[ \rho(X, Y) \in [-1, 1]. \]
The definition of covariance is not scaling invariant:
- If $X$ increases by a factor of $\alpha$, then $\text{Cov}[X, Y]$ increase by a factor of $\alpha$. 

Correlation Coefficient: Normalising the Covariance

Let $X$ and $Y$ be two random variables. The correlation coefficient $\rho(X, Y)$ is defined as:

$$\rho(X, Y) = \frac{\text{Cov}[X, Y]}{\sigma_X \sigma_Y}.$$ 

If $\sigma_X = 0$ or $\sigma_Y = 0$, then it is defined as 0.
The definition of covariance is not scaling invariant:

- If $X$ increases by a factor of $\alpha$, then $\text{Cov}[X, Y]$ increase by a factor of $\alpha$.

$\Rightarrow$ Even if $X$ and $Y$ both increase by $\alpha$, then $\text{Cov}[X, Y]$ will change.

(Exercise: It changes by?)
The definition of covariance is not scaling invariant:

- If \( X \) increases by a factor of \( \alpha \), then \( \text{Cov} [ X, Y ] \) increase by a factor of \( \alpha \).

\[ \Rightarrow \text{Even if } X \text{ and } Y \text{ both increase by } \alpha, \text{ then } \text{Cov} [ X, Y ] \text{ will change.} \]

(Exercise: It changes by?)

---

Correlation Coefficient

Let \( X \) and \( Y \) be two random variables. The correlation coefficient \( \rho(X, Y) \) is defined as:

\[
\rho(X, Y) = \frac{\text{Cov} [ X, Y ]}{\sqrt{\text{V}[X] \cdot \text{V}[Y]}}.
\]

If \( \text{V}[X] = 0 \) or \( \text{V}[Y] = 0 \), then it is defined as 0.
Correlation Coefficient: Normalising the Covariance

The definition of covariance is not scaling invariant:

- If \( X \) increases by a factor of \( \alpha \), then \( \text{Cov}[X, Y] \) increase by a factor of \( \alpha \).

\[ \Rightarrow \] Even if \( X \) and \( Y \) both increase by \( \alpha \), then \( \text{Cov}[X, Y] \) will change.

(Exercise: It changes by?)

Let \( X \) and \( Y \) be two random variables. The correlation coefficient \( \rho(X, Y) \) is defined as:

\[ \rho(X, Y) = \frac{\text{Cov}[X, Y]}{\sqrt{\text{V}[X] \cdot \text{V}[Y]}}. \]

If \( \text{V}[X] = 0 \) or \( \text{V}[Y] = 0 \), then it is defined as 0.

Properties:
Correlation Coefficient: Normalising the Covariance

The definition of covariance is not scaling invariant:

- If $X$ increases by a factor of $\alpha$, then $\text{Cov}[X, Y]$ increase by a factor of $\alpha$.

$\Rightarrow$ Even if $X$ and $Y$ both increase by $\alpha$, then $\text{Cov}[X, Y]$ will change. (Exercise: It changes by?)

---

**Correlation Coefficient**

Let $X$ and $Y$ be two random variables. The correlation coefficient $\rho(X, Y)$ is defined as:

$$
\rho(X, Y) = \frac{\text{Cov}[X, Y]}{\sqrt{\text{V}[X] \cdot \text{V}[Y]}}.
$$

If $\text{V}[X] = 0$ or $\text{V}[Y] = 0$, then it is defined as 0.

**Properties:**

1. The correlation coefficient is scaling-invariant, i.e.,
   $$
   \rho(X, Y) = \rho(\alpha \cdot X, \beta \cdot Y) \text{ for any } \alpha, \beta > 0.
   $$
The definition of covariance is not scaling invariant:
- If $X$ increases by a factor of $\alpha$, then $\text{Cov} [ X, Y ]$ increase by a factor of $\alpha$.

$\Rightarrow$ Even if $X$ and $Y$ both increase by $\alpha$, then $\text{Cov} [ X, Y ]$ will change.
(Exercise: It changes by?)

Let $X$ and $Y$ be two random variables. The correlation coefficient $\rho(X, Y)$ is defined as:

$$\rho(X, Y) = \frac{\text{Cov} [ X, Y ]}{\sqrt{\text{V}[X] \cdot \text{V}[Y]}}.$$ 

If $\text{V}[X] = 0$ or $\text{V}[Y] = 0$, then it is defined as $0$.

Properties:
1. The correlation coefficient is scaling-invariant, i.e.,
   $$\rho(X, Y) = \rho(\alpha \cdot X, \beta \cdot Y)$$ for any $\alpha, \beta > 0$.
2. For any two random variables $X, Y$, $\rho(X, Y) \in [-1, 1]$. 
Verify that the correlation coefficients’ range satisfies $\rho(X, Y) \in [-1, 1].$
Example

Verify that the correlation coefficients’ range satisfies $\rho(X, Y) \in [-1, 1]$.

- We will only prove $\rho(X, Y) \geq -1$ (the other direction follows in analogous way).
Verify that the correlation coefficients’ range satisfies $\rho(X, Y) \in [-1, 1]$.

---

We will only prove $\rho(X, Y) \geq -1$ (the other direction follows in analogous way).

Let $\sigma^2_x$ and $\sigma^2_y$ denote the variances of $X$ and $Y$, and $\sigma_x$ and $\sigma_y$ their standard deviations.
Range of the Correlation Coefficient

Verify that the correlation coefficients’ range satisfies $\rho(X, Y) \in [-1, 1]$.

- We will only prove $\rho(X, Y) \geq -1$ (the other direction follows in analogous way).
- Let $\sigma_x^2$ and $\sigma_y^2$ denote the variances of $X$ and $Y$, and $\sigma_x$ and $\sigma_y$ their standard deviations.
- Then:

$$0 \leq V \left[ \frac{X}{\sigma_x} + \frac{Y}{\sigma_y} \right]$$
Verify that the correlation coefficients’ range satisfies $\rho(X, Y) \in [-1, 1]$.

- We will only prove $\rho(X, Y) \geq -1$ (the other direction follows in analogous way).
- Let $\sigma^2_x$ and $\sigma^2_y$ denote the variances of $X$ and $Y$, and $\sigma_x$ and $\sigma_y$ their standard deviations.
- Then:

$$0 \leq \mathbb{V} \left[ \frac{X}{\sigma_x} + \frac{Y}{\sigma_y} \right]$$

$$= \mathbb{V} \left[ \frac{X}{\sigma_x} \right] + \mathbb{V} \left[ \frac{Y}{\sigma_y} \right] + 2 \text{Cov} \left[ \frac{X}{\sigma_x}, \frac{Y}{\sigma_y} \right]$$

\[\text{Example Intro to Probability 13}\]
Verify that the correlation coefficients’ range satisfies $\rho(X, Y) \in [-1, 1]$.

We will only prove $\rho(X, Y) \geq -1$ (the other direction follows in analogous way).

Let $\sigma_x^2$ and $\sigma_y^2$ denote the variances of $X$ and $Y$, and $\sigma_x$ and $\sigma_y$ their standard deviations.

Then:

$$0 \leq V\left[\frac{X}{\sigma_x} + \frac{Y}{\sigma_y}\right]$$

$$= V\left[\frac{X}{\sigma_x}\right] + V\left[\frac{Y}{\sigma_y}\right] + 2 \text{Cov}\left[\frac{X}{\sigma_x}, \frac{Y}{\sigma_y}\right]$$

$$= \frac{V[X]}{V[X]} + \frac{V[Y]}{V[Y]} + 2 \cdot \frac{\text{Cov}[X, Y]}{\sigma_x \cdot \sigma_x}$$
Verify that the correlation coefficients’ range satisfies $\rho(X, Y) \in [-1, 1]$.

- We will only prove $\rho(X, Y) \geq -1$ (the other direction follows in analogous way).
- Let $\sigma_x^2$ and $\sigma_y^2$ denote the variances of $X$ and $Y$, and $\sigma_x$ and $\sigma_y$ their standard deviations.
- Then:

$$0 \leq \mathbf{V} \left[ \frac{X}{\sigma_x} + \frac{Y}{\sigma_y} \right]$$

$$= \mathbf{V} \left[ \frac{X}{\sigma_x} \right] + \mathbf{V} \left[ \frac{Y}{\sigma_y} \right] + 2 \mathbf{Cov} \left[ \frac{X}{\sigma_x}, \frac{Y}{\sigma_y} \right]$$

$$= \frac{\mathbf{V}[X]}{\mathbf{V}[X]} + \frac{\mathbf{V}[Y]}{\mathbf{V}[Y]} + 2 \cdot \frac{\mathbf{Cov}[X, Y]}{\sigma_x \cdot \sigma_x}$$

$$= 2 \cdot (1 + \rho(X, Y)).$$
Using Covariance in Gambling/Trading

Example

Suppose a friend is known to predict the daily change of a certain stock price. After thoroughly testing, you are convinced that the covariance of the predicted relative change $X \in [-1, \infty)$ and the actual relative change $Y \in [-1, \infty)$ satisfy $E[X] = 0$ and $\text{Cov}[X, Y] > 0$. Using the information given by $X$ on each day, propose a strategy, by which – at least theoretically – you earn money (in expectation).

Answer
Suppose a friend is known to predict the daily change of a certain stock price. After thoroughly testing, you are convinced that the covariance of the predicted relative change $X \in [-1, \infty)$ and the actual relative change $Y \in [-1, \infty)$ satisfy $E[X] = 0$ and $\text{Cov}[X, Y] > 0$. Using the information given by $X$ on each day, propose a strategy, by which – at least theoretically – you earn money (in expectation).

**Example**

If $X = x > 0$ then we are going to buy $x$ units of stock, sell it on the next day and will have earned (or lost) $x \cdot y$. 
Using Covariance in Gambling/Trading

Example

Suppose a friend is known to predict the daily change of a certain stock price. After thoroughly testing, you are convinced that the covariance of the predicted relative change $X \in [-1, \infty)$ and the actual relative change $Y \in [-1, \infty)$ satisfy $E[X] = 0$ and $\text{Cov}[X, Y] > 0$. Using the information given by $X$ on each day, propose a strategy, by which – at least theoretically – you earn money (in expectation).

Answer

- If $X = x > 0$ then we are going to buy $x$ units of stock, sell it on the next day and will have earned (or lost) $x \cdot y$.
- If $X = x < 0$ then we are going to “short–sell” (This can be regarded as the inverse operation of buying – the details are not important here.) $-x$ units of stock, and on the next day we will have earned (or lost) $x \cdot y$. 
Using Covariance in Gambling/Trading

Example

Suppose a friend is known to predict the daily change of a certain stock price. After thoroughly testing, you are convinced that the covariance of the predicted relative change \( X \in [-1, \infty) \) and the actual relative change \( Y \in [-1, \infty) \) satisfy \( \mathbb{E}[X] = 0 \) and \( \text{Cov}[X, Y] > 0 \). Using the information given by \( X \) on each day, propose a strategy, by which – at least theoretically – you earn money (in expectation).

Answer

- If \( X = x > 0 \) then we are going to buy \( x \) units of stock, sell it on the next day and will have earned (or lost) \( x \cdot y \).
- If \( X = x < 0 \) then we are going to “short–sell” (This can be regarded as the inverse operation of buying – the details are not important here.) \(-x\) units of stock, and on the next day we will have earned (or lost) \( x \cdot y \).
- The expected earning/loss after one day equals

\[
\sum_{x \in X} \mathbb{P}[X = x] \cdot \sum_{y \in Y} \mathbb{P}[Y = y \mid X = x] \cdot (x \cdot y)
\]
Using Covariance in Gambling/Trading

Example

Suppose a friend is known to predict the daily change of a certain stock price. After thoroughly testing, you are convinced that the covariance of the predicted relative change \( X \in [-1, \infty) \) and the actual relative change \( Y \in [-1, \infty) \) satisfy \( \mathbb{E}[X] = 0 \) and \( \text{Cov}[X, Y] > 0 \). Using the information given by \( X \) on each day, propose a strategy, by which – at least theoretically – you earn money (in expectation).

Answer

- If \( X = x > 0 \) then we are going to buy \( x \) units of stock, sell it on the next day and will have earned (or lost) \( x \cdot y \).
- If \( X = x < 0 \) then we are going to “short–sell” (This can be regarded as the inverse operation of buying – the details are not important here.) \( -x \) units of stock, and on the next day we will have earned (or lost) \( x \cdot y \).
- The expected earning/loss after one day equals

\[
\sum_{x \in X} P[X = x] \cdot \sum_{y \in Y} P[Y = y \mid X = x] \cdot (x \cdot y)
\]

\[
= \sum_{x \in X, y \in Y} P[X = x, Y = y] \cdot (x \cdot y)
\]
Using Covariance in Gambling/Trading

Example

Suppose a friend is known to predict the daily change of a certain stock price. After thoroughly testing, you are convinced that the covariance of the predicted relative change \( X \in [-1, \infty) \) and the actual relative change \( Y \in [-1, \infty) \) satisfy \( E[X] = 0 \) and \( \text{Cov}[X, Y] > 0 \). Using the information given by \( X \) on each day, propose a strategy, by which – at least theoretically – you earn money (in expectation).

**Answer**

- If \( X = x > 0 \) then we are going to buy \( x \) units of stock, sell it on the next day and will have earned (or lost) \( x \cdot y \).
- If \( X = x < 0 \) then we are going to “short–sell” (This can be regarded as the inverse operation of buying – the details are not important here.) \(-x\) units of stock, and on the next day we will have earned (or lost) \( x \cdot y \).
- The expected earning/loss after one day equals

\[
\sum_{x \in X} P[X = x] \cdot \sum_{y \in Y} P[Y = y \mid X = x] \cdot (x \cdot y)
= \sum_{x \in X, y \in Y} P[X = x, Y = y] \cdot (x \cdot y)
= E[X \cdot Y].
\]
Suppose a friend is known to predict the daily change of a certain stock price. After thoroughly testing, you are convinced that the covariance of the predicted relative change $X \in [-1, \infty)$ and the actual relative change $Y \in [-1, \infty)$ satisfy $E[X] = 0$ and $\text{Cov}[X, Y] > 0$. Using the information given by $X$ on each day, propose a strategy, by which – at least theoretically – you earn money (in expectation).

**Answer**

- If $X = x > 0$ then we are going to buy $x$ units of stock, sell it on the next day and will have earned (or lost) $x \cdot y$.
- If $X = x < 0$ then we are going to “short–sell” (This can be regarded as the inverse operation of buying – the details are not important here.) $-x$ units of stock, and on the next day we will have earned (or lost) $x \cdot y$.
- The expected earning/loss after one day equals

$$\sum_{x \in X} \mathbb{P}[X = x] \cdot \sum_{y \in Y} \mathbb{P}[Y = y \mid X = x] \cdot (x \cdot y) = \sum_{x \in X, y \in Y} \mathbb{P}[X = x, Y = y] \cdot (x \cdot y) = E[X \cdot Y].$$

- Since $E[X] = 0$, it follows that $E[X \cdot Y] = \text{Cov}[X, Y] > 0$. 
Using Covariance in Gambling/Trading

Example

Suppose a friend is known to predict the daily change of a certain stock price. After thoroughly testing, you are convinced that the covariance of the predicted relative change $X \in [-1, \infty)$ and the actual relative change $Y \in [-1, \infty)$ satisfy $E[X] = 0$ and $\text{Cov}[X, Y] > 0$. Using the information given by $X$ on each day, propose a strategy, by which – at least theoretically – you earn money (in expectation).

Answer

- If $X = x > 0$ then we are going to buy $x$ units of stock, sell it on the next day and will have earned (or lost) $x \cdot y$.
- If $X = x < 0$ then we are going to “short–sell” (This can be regarded as the inverse operation of buying – the details are not important here.) $-x$ units of stock, and on the next day we will have earned (or lost) $x \cdot y$.
- The expected earning/loss after one day equals

$$
\sum_{x \in X} P[X = x] \cdot \sum_{y \in Y} P[Y = y \mid X = x] \cdot (x \cdot y)
= \sum_{x \in X, y \in Y} P[X = x, Y = y] \cdot (x \cdot y)
= E[X \cdot Y].
$$

Generalisation: If $E[X] > 0$, we would buy $x - E[X]$ units.

- Since $E[X] = 0$, it follows that $E[X \cdot Y] = \text{Cov}[X, Y] > 0$. 