भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Probability and Stochastic Processes

Lecture 26: Correlation Coefficient, Cauchy–Schwarz Inequality, Vector Spaces of Random Variables, Conditional Expectations

Karthik P. N.

Assistant Professor, Department of AI

Email: pnkarthik@ai.iith.ac.in

07 November 2025

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Covariance

Definition (Covariance)

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$.
Let X, Y be random variables. Let $\mathbb{E}[X], \mathbb{E}[Y]$ be well defined (i.e., not of the form $\infty - \infty$).
The **covariance** of X and Y is defined as

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

provided the expectation on the right-hand side is well defined (i.e., not $\infty - \infty$).
Furthermore,

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

provided the right-hand side is not of the form $\infty - \infty$.

Remarks:

- $\mathrm{Cov}(X, Y)$ can be negative, positive, or zero
- If $Y = X$, then

$$\mathrm{Cov}(X, X) = \mathbb{E}[X^2] - \left(\mathbb{E}[X]\right)^2 = \mathrm{Var}(X).$$

Uncorrelated Random Variables

Definition (Uncorrelated Random Variables)

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$.
Let X, Y be random variables. Let $\mathbb{E}[X], \mathbb{E}[Y]$ be well defined (i.e., not of the form $\infty - \infty$).
X and Y are said to be **uncorrelated** if

$$\mathrm{Cov}(X, Y) = 0.$$

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Uncorrelatedness and Independence

Theorem (Uncorrelatedness and Independence)

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$.
Let X, Y be random variables. Let $\mathbb{E}[X], \mathbb{E}[Y]$ be well-defined (i.e., not of the form $\infty - \infty$).
If $X \perp\!\!\!\perp Y$, then

$$\mathrm{Cov}(X, Y) = 0.$$

The **converse is not true in general**.

- **Proof:**

$$\mathbb{E}[X\,Y] = \int_{\mathbb{R}^2} x\,y\,\mathrm{d}\mathbb{P}_{X,Y}(x,y) \overset{X \perp\!\!\!\perp Y}{=} \int_{\mathbb{R}^2} x\,y\,\mathrm{d}\mathbb{P}_X(x)\,\mathrm{d}\mathbb{P}_Y(y) = \left(\int_{\mathbb{R}} x\,\mathrm{d}\mathbb{P}_X(x)\right) \cdot \left(\int_{\mathbb{R}} y\,\mathrm{d}\mathbb{P}_Y(y)\right) = \mathbb{E}[X]\,\mathbb{E}[Y]$$

- **Converse not true in general:** Let $X \sim \mathcal{N}(0,1)$, and let $Y = X^2$. Then,

$$\mathbb{E}[XY] = \mathbb{E}[X^3] = 0, \qquad \mathbb{E}[X]\,\mathbb{E}[Y] = 0 \cdot 1 = 0, \qquad \mathrm{Cov}(X, Y) = 0,$$

but $X \not\perp\!\!\!\perp Y$

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Variance of Sum of Two Random Variables

Lemma (Variance of Sum of Two Random Variables)

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$.
Let X, Y be random variables. Let $\mathbb{E}[X], \mathbb{E}[Y]$ be well-defined (i.e., not of the form $\infty - \infty$).
Then,

$$\mathtt{Var}(X+Y) = \mathtt{Var}(X) + \mathtt{Var}(Y) + 2\,\mathtt{Cov}(X, Y).$$

In particular, if X, Y are uncorrelated, then

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

Proof:

- We have

$$\mathtt{Var}(X+Y) = \mathbb{E}\left[\Big(X + Y - \mathbb{E}[X+Y]\Big)^2\right] = \mathbb{E}\left[\Big((X - \mathbb{E}[X]) + (Y - \mathbb{E}[Y])\Big)^2\right].$$

- Expand and apply linearity of expectations

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Correlation Coefficient and Cauchy–Schwarz Inequality

Correlation Coefficient

Definition (Correlation Coefficient)

The correlation coefficient between X and Y is defined as

$$\rho_{X,Y} := \frac{\mathrm{Cov}(X, Y)}{\sqrt{\mathrm{Var}(X) \cdot \mathrm{Var}(Y)}},$$

whenever $\mathrm{Cov}(X, Y)$ is well-defined.

Remark: $\rho_{X,Y}$ can be positive, negative, or zero

The Cauchy–Schwarz Inequality

Theorem (Cauchy–Schwarz Inequality)

For any two random variables X and Y,

$$-1 \leq \rho_{X,Y} \leq 1.$$

Furthermore, the following hold.

1. If $\rho_{X,Y} = 1$, then there exists $a > 0$ such that

$$\mathbb{P}\left(\left\{\frac{Y - \mathbb{E}[Y]}{X - \mathbb{E}[X]} = a\right\}\right) = 1.$$

2. If $\rho_{X,Y} = -1$, then there exists $a < 0$ such that

$$\mathbb{P}\left(\left\{\frac{Y - \mathbb{E}[Y]}{X - \mathbb{E}[X]} = a\right\}\right) = 1.$$

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Proof of CS Inequality

- Define $\widetilde{X}, \widetilde{Y}$ as
$$\widetilde{X} := X - \mathbb{E}[X], \qquad \widetilde{Y} := Y - \mathbb{E}[Y].$$
- The following holds:
$$\mathbb{E}\left[\left(\widetilde{X} - \frac{\mathbb{E}[\widetilde{X}\,\widetilde{Y}]}{\mathbb{E}[(\widetilde{Y})^2]}\,\widetilde{Y}\right)^2\right] \geq 0.$$
- Expanding the inner squared term and using linearity of expectations, we arrive at the CS inequality
- **Equality in CS inequality:**
$$\mathbb{E}\left[\left(\widetilde{X} - \frac{\mathbb{E}[\widetilde{X}\,\widetilde{Y}]}{\mathbb{E}[(\widetilde{Y})^2]}\,\widetilde{Y}\right)^2\right] = 0 \quad \Longrightarrow \quad \mathbb{P}\left(\widetilde{X} = \frac{\mathbb{E}[\widetilde{X}\,\widetilde{Y}]}{\mathbb{E}[(\widetilde{Y})^2]}\,\widetilde{Y}\right) = 1.$$
- Let $a \in \mathbb{R}$ be defined as
$$a := \left(\frac{\mathbb{E}[\widetilde{X}\,\widetilde{Y}]}{\mathbb{E}[(\widetilde{Y})^2]}\right)^{-1} = \frac{\mathbb{E}[(\widetilde{Y})^2]}{\mathbb{E}[\widetilde{X}\,\widetilde{Y}]} = \frac{\mathrm{Var}(Y)}{\mathrm{Cov}(X,Y)} = \frac{\sqrt{\mathrm{Var}(Y)}}{\rho_{X,Y}\,\sqrt{\mathrm{Var}(X)}} \quad \begin{cases} > 0, & \rho_{X,Y} = 1, \\ < 0, & \rho_{X,Y} = -1. \end{cases}$$

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Inequalities

Markov's Inequality

Theorem (Markov's Inequality)

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$.
Let X be a **non-negative** random variable with $\mathbb{E}[X] < +\infty$. Then,

$$\forall\, \alpha > 0, \qquad \mathbb{P}(\{X > \alpha\}) \leq \frac{\mathbb{E}[X]}{\alpha}.$$

Remarks:

- Markov's inequality only applies to non-negative random variables
- The inequality is useful only for $\alpha > \mathbb{E}[X]$

Markov's Inequality

Theorem (Markov's Inequality)

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$.
Let X be a **non-negative** random variable with $\mathbb{E}[X] < +\infty$. Then,

$$\forall\, \alpha > 0, \qquad \mathbb{P}(\{X > \alpha\}) \leq \frac{\mathbb{E}[X]}{\alpha}.$$

Proof:

- We may express X as
$$X = X \cdot \mathbf{1}_{\{X>\alpha\}} + X \cdot \mathbf{1}_{\{X\leq\alpha\}}.$$
- Taking expectations on both sides, we get
$$\begin{aligned} \mathbb{E}[X] &= \mathbb{E}[X \cdot \mathbf{1}_{\{X>\alpha\}}] + \mathbb{E}[X \cdot \mathbf{1}_{\{X\leq\alpha\}}] \\ &\geq \mathbb{E}[X \cdot \mathbf{1}_{\{X>\alpha\}}] && \left(\mathbb{E}[X \cdot \mathbf{1}_{\{X\leq\alpha\}}] \geq 0 \text{ as } X \text{ is non-negative}\right) \\ &\geq \alpha \cdot \mathbb{P}(\{X > \alpha\}). \end{aligned}$$

Application of Markov's Inequality

Lemma (Non-Negative Random Variables with Finite Expectation)

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$.
Suppose that $X : \Omega \to [0, +\infty]$ is a **non-negative and extended real-valued** random variable. Then,

$$\mathbb{E}[X] < +\infty \qquad \implies \qquad \mathbb{P}(\{X < +\infty\}) = 1.$$

Proof:

- Notice that

$$\{X = +\infty\} = \bigcap_{n \in \mathbb{N}} \{X > n\} = \lim_{n \to \infty} \{X > n\}.$$

- We then have

$$\begin{aligned} \mathbb{P}(\{X = +\infty\}) &= \lim_{n \to \infty} \mathbb{P}(\{X > n\}) && \text{(continuity of probability)} \\ &\leq \lim_{n \to \infty} \frac{\mathbb{E}[X]}{n} && \text{(Markov's inequality)} \\ &= 0 && \text{(because } \mathbb{E}[X] < +\infty\text{).} \end{aligned}$$

Chebyshev's Inequality

Theorem (Chebyshev's Inequality)

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$.
Let X be a random variable with **finite mean** and **finite variance**. Then,

$$\forall\, \alpha > 0, \qquad \mathbb{P}\left(\left\{\left|X - \mathbb{E}[X]\right| > \alpha\right\}\right) \leq \frac{\mathrm{Var}(X)}{\alpha^2}.$$

Proof:

- We have

$$\begin{aligned}
\mathbb{P}\left(\left\{\left|X - \mathbb{E}[X]\right| > \alpha\right\}\right) &= \mathbb{P}\left(\left\{\left(X - \mathbb{E}[X]\right)^2 > \alpha^2\right\}\right) \\
&\leq \frac{\mathbb{E}\left[\left(X - \mathbb{E}[X]\right)^2\right]}{\alpha^2} && \text{(Markov's inequality applied to } (X - \mathbb{E}[X])^2\text{)} \\
&= \frac{\mathrm{Var}(X)}{\alpha^2}.
\end{aligned}$$

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Jensen's Inequality

Theorem (Jensen's Inequality)

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$. Let $X : \Omega \to \mathbb{R}$ be a random variable.
Let $g : \mathbb{R} \to \mathbb{R}$ be a **convex and differentiable** function, i.e.,

$$g(y) \geq g(x) + g'(x)(y - x) \qquad \forall x, y \in \mathbb{R}.$$

If $\left|\mathbb{E}[X]\right| < +\infty$ and $\left|\mathbb{E}[g(X)]\right| < +\infty$, then

$$\mathbb{E}[g(X)] \geq g(\mathbb{E}[X]).$$

Corollary:

- $\mathbb{E}[X^2] \geq (\mathbb{E}[X])^2, \qquad \mathbb{E}[|X|] \geq \left|\mathbb{E}[X]\right|$
- $\mathbb{E}[\log X] \leq \log \mathbb{E}[X]$ for a non-negative RV X

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Jensen's Inequality

Theorem (Jensen's Inequality)

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$. Let $X : \Omega \to \mathbb{R}$ be a random variable.
Let $g : \mathbb{R} \to \mathbb{R}$ be a **convex and differentiable** function, i.e.,

$$g(y) \geq g(x) + g'(x)(y-x) \qquad \forall x, y \in \mathbb{R}.$$

If $\Big|\mathbb{E}[X]\Big| < +\infty$ and $\Big|\mathbb{E}[g(X)]\Big| < +\infty$, then

$$\mathbb{E}[g(X)] \geq g(\mathbb{E}[X]).$$

Proof:

- We have
$$\forall\, \omega \in \Omega, \qquad g(X(\omega)) \geq g(\mathbb{E}[X]) + g'(\mathbb{E}[X])(X(\omega) - \mathbb{E}[X]).$$
- Equivalently, we have
$$g(X) \geq g(\mathbb{E}[X]) + g'(\mathbb{E}[X])(X - \mathbb{E}[X]).$$
- The result is obtained by taking expectations on both sides

Vector Spaces of Random Variables

Absolute Moments and Norms

Definition (Absolute Moments)

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$. Let X be a random variable.
For any $p > 0$, the quantity

$$\mathbb{E}\left[|X|^p\right] = \int_\Omega |X(\omega)|^p \, \mathrm{d}\mathbb{P}(\omega)$$

is called the **pth-order absolute moment of X**.

- If $p \in \mathbb{N}$, the corresponding moments are referred to as the first, second, etc. absolute moments of X
- p can be any non-negative real power

Monotonicity of Absolute Moments

Lemma (Monotonicity of Absolute Moments)

For any $0 < p \leq q$,

$$\mathbb{E}\Big[|X|^q\Big] < +\infty \qquad \implies \qquad \mathbb{E}\Big[|X|^p\Big] < +\infty.$$

Proof:

- We have

$$x^p \leq \begin{cases} 1, & 0 \leq x < 1, \\ x^q, & x \geq 1. \end{cases}$$

- As a result, we can write

$$x^p \leq 1 + x^q \qquad \forall\, x \geq 0.$$

- Applying the above inequality to random variables, we get

$$\forall\, \omega \in \Omega, \quad |X(\omega)|^p \leq 1 + |X(\omega)|^q, \qquad |X|^p \leq 1 + |X|^q.$$

- Taking expectations on either sides, we get

$$\mathbb{E}\Big[|X|^p\Big] \leq 1 + \mathbb{E}\Big[|X|^q\Big] < +\infty.$$

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Hölder's Inequality

Lemma (Hölder's Inequality)

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$. Let X, Y be random variables.
Let $1 < p, q < +\infty$ be numbers such that $\frac{1}{p} + \frac{1}{q} = 1.$ Then, we have

$$\mathbb{E}[|XY|] \leq \left(\mathbb{E}\left[|X|^p\right]\right)^{\frac{1}{p}} \cdot \left(\mathbb{E}\left[|Y|^q\right]\right)^{\frac{1}{q}},$$

with equality if and only if $\mathbb{P}\left(\left\{\frac{|X|^p}{|Y|^q} = a\right\}\right) = 1$ for some $a > 0$.

Remarks:

- Suppose $\mathbb{E}[X]$ and $\mathbb{E}[Y]$ are well-defined
- Taking $X \leftarrow X - \mathbb{E}[X]$, $Y \leftarrow Y - \mathbb{E}[Y]$, $p = q = 2$, we get

$$\mathbb{E}\left[\left|(X - \mathbb{E}[X])\,(Y - \mathbb{E}[Y])\right|\right] \leq \sqrt{\mathrm{Var}(X)} \cdot \sqrt{\mathrm{Var}(Y)}.$$

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Hölder's Inequality

Lemma (Hölder's Inequality)

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$. Let X, Y be random variables.
Let $1 < p, q < +\infty$ be numbers such that $\frac{1}{p} + \frac{1}{q} = 1.$ Then, we have

$$\mathbb{E}[|XY|] \leq \left(\mathbb{E}\left[|X|^p\right]\right)^{\frac{1}{p}} \cdot \left(\mathbb{E}\left[|Y|^q\right]\right)^{\frac{1}{q}},$$

with equality if and only if $\mathbb{P}\left(\left\{\frac{|X|^p}{|Y|^q} = a\right\}\right) = 1$ for some $a > 0$.

Remarks:

- Finally, we have

$$\left|\mathbb{E}\left[(X - \mathbb{E}[X])\,(Y - \mathbb{E}[Y])\right]\right| \leq \mathbb{E}\left[\left|(X - \mathbb{E}[X])\,(Y - \mathbb{E}[Y])\right|\right] \leq \sqrt{\operatorname{Var}(X)} \cdot \sqrt{\operatorname{Var}(Y)} \qquad \text{(CS inequality)}.$$

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Minkowski's Inequality

Lemma (Minkowski's Inequality)

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$. Let X, Y be random variables.
For any $p \geq 1$,

$$\left(\mathbb{E}\left[|X+Y|^p\right]\right)^{\frac{1}{p}} \leq \left(\mathbb{E}\left[|X|^p\right]\right)^{\frac{1}{p}} + \left(\mathbb{E}\left[|Y|^p\right]\right)^{\frac{1}{p}}.$$

Remark:

- A simple proof of this can be shown using Hölder's inequality

A Semi-Norm

- Let us define

$$\|X\|_p := \left(\mathbb{E}\left[|X|^p\right]\right)^{\frac{1}{p}}, \qquad 1 \leq p < +\infty.$$

- **Does $\|X\|_p$** satisfy all properties of a norm?
- **(Homogeneity)**

  Clearly, for any $\alpha \in \mathbb{R}$,

$$\|\alpha X\|_p = |\alpha| \, \|X\|_p.$$

- **(Triangle Inequality)**

  Thanks to Minkowski's inequality,

$$\|X + Y\|_p \leq \|X\|_p + \|Y\|_p.$$

- **(Non-Negativity)**

  Clearly, $\quad \|X\|_p \geq 0$
- **(Definiteness)**

  However,

$$\|X\|_p = 0 \qquad \Longrightarrow \qquad \mathbb{P}(\{X = 0\}) = 1.$$

Turning $\|\cdot\|_p$ into a Norm – Equivalence Classes

- Let $\mathcal{L}^p(\Omega, \mathscr{F}, \mathbb{P})$ denote the collection of random variables

$$\mathcal{L}_p(\Omega, \mathscr{F}, \mathbb{P}) = \left\{ X : \ \mathbb{E}[|X|^p] = \int_\Omega |X|^p \, \mathrm{d}\mathbb{P} < +\infty \right\}.$$

- Define the relation $\overset{\mathrm{R}}{\sim}$ on $\mathcal{L}^2 \times \mathcal{L}^2$ as follows:

$$X \overset{\mathrm{R}}{\sim} Y \qquad \Longleftrightarrow \qquad \mathbb{P}(\{X = Y\}) = 1.$$

- The above relation satisfies the following properties:

Reflexive property): $X \overset{\mathrm{R}}{\sim} X$ for all $X \in \mathcal{S}$

Symmetry property): $X \overset{\mathrm{R}}{\sim} Y \implies Y \overset{\mathrm{R}}{\sim} X$

Transitive property): $X \overset{\mathrm{R}}{\sim} Y, \quad Y \overset{\mathrm{R}}{\sim} Z \quad \implies \quad X \overset{\mathrm{R}}{\sim} Z$

The relation $\overset{\mathrm{R}}{\sim}$ is an **equivalence relation** on $\mathcal{S} \times \mathcal{S}$. It will partition $\mathcal{S}$ into **equivalence classes**.