Mathematics
Academic year 2026-2027
Teacher: P. Gibilisco

Lecture 1 – Monday, August 31, 2026 (14:00-16:30)

The beginner ... should not be discouraged if ... he finds that he does not have the prerequisites for reading the prerequisites.
P.Halmos

Introduction to the course.

An example on the applications of mathematics in economics: risk aversion and concavity.

List of the notions required in advance to follow the course.

Deterministic and random phenomena.

The "law of chance" as an oxymoron.

Axiomatization of geometry and probability: Euclid and Kolmogorov.

The ideas of Kolmogorov: events are represented by sets; probability is a normed measure on them.

Probability spaces and their properties (finitely additive case).

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

No proposition is in itself either probable or improbable, just as no place can be intrinsically distant ...
J.M.Keynes

Conditional probability, pairs of independent events ($A \perp\!\!\!\perp B$ means A and B are independent). Remark: $P(A) = P(A|\Omega)$.

The two pillars of probability theory: on one side analysis and measure theory; on the other side gambling situations, coin-tossing.

Lecture 2 – Tuesday, September 1, 2026 (14:00-16:30)

... the fact is that there is nothing as dreamy and poetic, nothing as radical, subversive, and psychedelic, as mathematics ... Mathematics is the purest of the arts, as well as the most misunderstood.
P.Lockhart

Exercise: prove that $A \perp\!\!\!\perp B \Rightarrow A \perp\!\!\!\perp B^c$

Bayes formula.

Partitions.

Exercise: an urn contains w white balls and r red balls. Choose a ball in the urn and leave it out (without looking at its colour). Choose a second ball. Which is the probability that the second ball is white? (The answer will explains the queue paradox).

Pairwise independence and independence.

A counterexample: pairwise independence does not imply independence. Exercise: prove that

$$p(A|B) + p(A^c|B) = 1,$$

while (in general)

$$p(A|B) + p(A|B^c) \neq 1.$$

(Conditional probability is a probability measure only with respect to the first argument!)

Lecture 3 – Wednesday, September 2, 2026 (14:00-16:30)

The geometric series and its sum.

$$1 + q + q^2 + q^3 + \cdots = \frac{1}{1-q} \quad \text{if} \qquad |q| < 1$$

Taylor polynomial and Taylor series. Example:

$$\frac{1}{1-x} = \sum_{k=0}^{\infty} x^k \quad (|x| < 1)$$

Taylor polynomial and Taylor series. Example:

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} \quad (x \in \mathbb{R})$$

No one shall expel us from the paradise which Cantor has created for us.
D. Hilbert

Countable and uncountable sets. Cardinality of $\mathbb{N}$, $\mathbb{Q}$, $\mathbb{R}$.

Counting subsets of a set (n choose k). Newton binomial formula. The cardinality of the power set: $2^n = \sum_{k=0}^{n} \binom{n}{k}$.

Lecture 4 – Thursday, September 3, 2026 (14:00-16:30)

Counting measure on $\mathbb{N}$ is not continuous.

σ–algebras and σ–additivity for a probability measure.

Probability spaces and their properties (σ–additive case).

Continuity of a probability measure. Equivalence of continuity from above and σ–additivity (no proof!)

Coin flipping and the Bernoulli process. The probability to have k successes flipping a coin n times:

$$\binom{n}{k} p^k (1-p)^{n-k}$$

Image of a function.

Random variables: discrete case

Distribution of a random variable.

Exercise. A fair die is rolled two times. Let us denote by X the first result and by Y the second result. Let $U := X - Y$. Which is the distribution of the random variable U?

Lecture 5 – Friday, September 4, 2026 (11:00-13:30)

Non-negative series with sum equal to one and discrete distributions.

The most important discrete r.v.: Bernoulli, Binomial, Geometric and Poisson distribution.

$$\lim_{n \to +\infty} \left(1 + \frac{c}{n}\right)^n = e^c$$

Convergence of the Bernoulli to the Poisson.

Independence for discrete random variables.

Expected value for discrete r.v.

Linearity and positivity for the expected value.

How the Bernoulli, Binomial, Geometric and Poisson distributions arise flipping a coin (Bernoulli process).

Mean value for the Bernoulli and binomial distribution.

A counterexample. Consider a box where there are three balls with the numbers $-1, 0, 1$. Choose randomly a ball and let X be the random variable which represent the result so that:

$$P(X = -1) = P(X = 0) = P(X = 1) = \frac{1}{3}.$$

Prove that $X \not\perp\!\!\!\perp X^2$.

Lecture 6 – Monday, September 7, 2026 (14:00-16:30)

$$\mathbb{E}(g(X)) = \sum_k g(x_k) P(X = x_k)$$

Moments of a r.v.. Variance.

Covariance and its properties.

Standard deviation.

Correlation coefficient and scale invariance.

$$X \perp Y \qquad \Rightarrow \qquad \mathrm{E}(XY) = \mathrm{E}(X) \cdot \mathrm{E}(Y)$$

$$X \perp Y \qquad \Rightarrow \qquad \mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$$

$$X \perp Y \qquad \Rightarrow \qquad \mathrm{Cov}(X, Y) = 0$$

A counterexample. Consider a box where there are three balls with the numbers $-1, 0, 1$. Choose randomly a ball and let X be the random variable which represent the result so that:

$$P(X = -1) = P(X = 0) = P(X = 1) = \frac{1}{3}.$$

Prove that $X \not\perp X^2$ while $\text{Cov}(X, X^2) = 0$ (this is true in general for X and X^2 when X has a distribution which is symmetric w.r.t. to 0).

This prove that:

$$X \perp Y \qquad \not\Leftarrow \qquad \text{Cov}(X, Y) = 0$$

Variance for the Bernoulli and Binomial distribution.
The derivative of the geometric series.
Waiting times for the Bernoulli process and the geometric distribution.
Expectation of the geometric distribution.

Lecture 7 – Tuesday, September 8, 2026 (14:00-16:30)

Cumulative distribution function.
Continuous random variables ($P(X = x) = 0 \, \forall x \in \mathbb{R}$, continuity of F_X).
Densities and absolutely continuous random variable.
The uniform density.
Exponential density.
Moments and variance for absolutely continuous random variables.
Mean value and variance for the uniform distribution.
Mean value and variance for the exponential distribution.
Cumulative distribution function for the uniform distribution.
If X is an a.c. r.v. with density f_X then

$$F_X' = f_X$$

Lecture 8 – Wednesday, September 9, 2026 (14:00-16:30)

A mathematician is someone to whom the facts that

$$2 + 2 = 4$$

and

$$\int_{\mathbb{R}} e^{-x^2} \, dx = \sqrt{\pi}$$

are equally obvious.
Lord Kelvin

Cumulative distribution function for the exponential distribution.

The Gaussian density

$$\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$$

Functions whose antiderivative cannot be expressed as an elementary function.

$$\int_{\mathbb{R}} e^{-x^2}\,\mathrm{d}x = \sqrt{\pi} \qquad\qquad \text{(No proof)}$$

Density of $aX + b$ from the density of X.

$$\int_{\mathbb{R}} x^2 e^{-\frac{x^2}{2}}\,\mathrm{d}x = \sqrt{2\pi} \qquad\qquad \text{Integration by parts}$$

Mean and variance for the gaussian density.

The Gaussian density: the general case ($X \sim \mathcal{N}(\mu, \sigma^2)$ means $X = \sigma Z + \mu$ where $Z \sim \mathcal{N}(0, 1)$.)

Lecture 9 – Thursday, September 10, 2026 (14:00-16:30)

Standardized random variables

$$X^* := \frac{X - \mathbb{E}(X)}{\sigma(X)}$$

- The law of large numbers. (Distribution form; no proof yet). $X_1, ..., X_n, ...$ i.i.d. random variables with finite mean μ and variance σ^2. Then

$$\frac{S_n}{n} \longrightarrow \mu \qquad \text{in law}$$

- CLT. (No proof yet). $X_1, ..., X_n, ...$ i.i.d. random variables with finite mean and variance. Then

$$S_n^* \longrightarrow \mathcal{N}(0, 1) \qquad \text{in law}$$

Written in five years, may it last as many thousands.
G. Cardano at the end of Ars Magna

Mac Laurin series for the trigonometric functions.

Introduction to complex numbers.
Exercise: calculate
$$Re\left(\frac{1}{2+3i}\right) = \cdots$$
History of $0, 1, i, e, \pi$.
Complex exponential as a power series.
$$e^{i\theta} = \cos\theta + i\sin\theta$$
The Euler formula.
$$e^{i\pi} + 1 = 0$$
The Fundamental theorem of Algebra.
Complex random variables and their expectations.
A constant random variable is independent from any random variable.
Remark: $X \perp Y$ implies $\mathbb{E}(g(X)h(Y)) = \mathbb{E}(g(X))\mathbb{E}(h(Y))$.
The characteristic function, first properties.

- $\varphi_X(0) = 1$
- $\varphi_{\lambda X}(t) = \varphi_X(\lambda t)$
- $X \perp Y \qquad \Rightarrow \qquad \varphi_{X+Y}(t) = \varphi_X(t)\,\varphi_Y(t)$

Lecture 10 – Tuesday, September 15, 2026 (14:00-16:30)

$e^{i\pi} + 1 = 0$ is for boys,
$i^i = \frac{1}{\sqrt{e^\pi}}$ is for men.
Anonymous

The inversion theorem (no proof!):
$$X, Y \quad \text{have the same distribution} \qquad \text{iff} \qquad \varphi_X(t) = \varphi_Y(t) \qquad \forall t \in \mathbb{R}$$

$$X \sim \text{Poisson}(\lambda) \Longrightarrow \varphi_X(t) = \exp(\lambda(e^{it} - 1))$$
Using the characteristic function prove that
$$X \perp Y, \quad X \sim \text{Poisson}(\lambda), \quad Y \sim \text{Poisson}(\mu) \Longrightarrow X + Y \sim \text{Poisson}(\lambda + \mu)$$
If $X_n \sim B(n,p)$ then $\varphi_X(t) = (pe^{it} + q)^n$. Using this result prove that the sum of independent Bernoulli variables is binomial.

- Convergence in law (in distribution).
- Example: if X_n and X are random variables whose values are natural numbers then convergence in law is equivalent to (no proof)

$$P(X_n = k) \to P(X = k) \qquad \forall k \in \mathbb{N}$$

- Example: if $X_n \sim B(n, \frac{\lambda}{n})$ and $X \sim \text{Poisson}(\lambda)$ then $X_n \Longrightarrow X \quad$ in law.
- Continuity theorem (no proof!):

$$X_n \Longrightarrow X \quad \text{in law} \qquad \text{iff} \qquad \varphi_{X_n}(t) \to \varphi_X(t) \qquad \forall t \in \mathbb{R}$$

- Example: prove that if $X_n \sim B(n, \frac{\lambda}{n})$ and $X \sim \text{Poisson}(\lambda)$ then $X_n \Longrightarrow X \quad$ in law using the characteristic function and the continuity theorem.

Lecture 11 – Thursday, September 17, 2026 (14:00-16:30)

Students don't need a perfect teacher. Students need a happy teacher, who's gonna make them excited to come to school and grow a love for learning.
R. Feynman

The Rest of the Taylor polynomials in x_0 is *negligible near* x_0.
Partial derivatives. Derivation under the integral sign.
Moments and the characteristic function.

$$\varphi_X^{(k)}(0) = i^k \mathbb{E}(X^k)$$

The Taylor formula for $\varphi_X(\cdot)$.

The characteristic functions of a costant random variable $X = c$: $\phi_X(t) = e^{ict}$.

Characteristic function of the gaussian: if $X \sim \mathcal{N}(0,1)$ then $\phi_X(t) = e^{-\frac{t^2}{2}}$. (No proof!)

- LLN. $X_1, ..., X_n, ...$ i.i.d. random variables with finite mean μ and variance σ^2. Then
$$\frac{S_n}{n} \longrightarrow \mu \qquad \text{in law}$$

- CLT. $X_1, ..., X_n, ...$ i.i.d. random variables with finite mean μ and variance σ^2. Then
$$S_n^* \longrightarrow \mathcal{N}(0,1) \qquad \text{in law}$$

- Meaning of the CLT: the sum of (infinitely) many, small, independent effects is (approximately) normal (gaussian).

- CLT reformulated. Let $\Phi(t)$ the c.d.f. of the standard gaussian distribution.

  $X_1, ..., X_n, ...$ i.i.d. random variables with finite mean μ and variance σ^2. Then
$$P(X_1 + \cdots + X_n \leq x) = P\left(S_n^* \leq \frac{x - n\mu}{\sigma\sqrt{n}}\right) \approx \Phi\left(\frac{x - n\mu}{\sigma\sqrt{n}}\right)$$

Characteristic functions of an arbitrary Gaussian. The sum of independent Gaussian is Gaussian.

Lecture 12 – Friday, September 18, 2026 (11:00-13:30)

Simulation 1. – Quiz 7. Exercise 4.
Simulation 2. – Quiz 2.
Simulation 3. – Quiz 1, 6, 7.
Simulation 4. – Quiz 1, 6. Exercise 4.

Lecture 13 – Monday, September 21, 2026 (14:00-16:30)

The Gamma function
$$\Gamma(\alpha) = \int_0^{+\infty} x^{\alpha-1} e^{-x} \mathrm{d}x \qquad\qquad \alpha > 0$$
$$\Gamma(\alpha+1) = \alpha\Gamma(\alpha) \qquad \Gamma(1) = 1 \qquad \Gamma(n) = (n-1)!$$
$X \sim \Gamma(\alpha, \lambda)$ (where $\alpha, \lambda > 0$) if
$$f_X(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} \cdot 1_{(0,\infty)}(x)$$

If

$$c \cdot x^{\alpha-1}e^{-\lambda x} \cdot 1_{(0,\infty)}(x)$$

is a density then

$$c = \frac{\lambda^\alpha}{\Gamma(\alpha)}$$

If X has density f_X then X^2 has density

$$f_{X^2}(x) = \frac{1}{2\sqrt{x}}(f(\sqrt{x}) + f(\sqrt{-x})) \cdot 1_{(0,\infty)}(x)$$

If $X \sim \mathcal{N}(0,1)$ then $X^2 \sim \Gamma(\frac{1}{2}, \frac{1}{2})$.
This implies

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$$

- Moments of the Gamma distribution. If $X \sim \Gamma(\alpha, \lambda)$ then

$$\mathbb{E}(X^n) = \frac{1}{\lambda^n}\frac{\Gamma(\alpha+n)}{\Gamma(\alpha)}$$

$$\mathrm{Var}(X) = \frac{\alpha}{\lambda^2}$$

If $X \sim \Gamma(\alpha, \lambda)$ then

- The characteristic function (no proof):

$$\varphi_X(t) = \left(\frac{\lambda}{\lambda - it}\right)^\alpha$$

- If $X \sim \Gamma(\alpha, \lambda)$, $Y \sim \Gamma(\beta, \lambda)$ and $X \perp Y$ then $X + Y \sim \Gamma(\alpha+\beta, \lambda)$.
- The chi-squared distribution with n degrees of freedom. If $X \sim \Gamma(\frac{n}{2}, \frac{1}{2})$ then we say that $X \sim \chi^2(n)$.
- If $X_1, X_2, ..., X_n$ are independent standard gaussian then $X_1^2 + X_2^2 + \cdots X_n^2 \sim \chi^2(n)$.

Lecture 14 – Tuesday, September 22, 2026 (14:00-16:30)

Moments of a standard Gaussian distribution:

$$E(X^{2n+1}) = 0 \qquad\qquad E(X^{2n}) = (2n-1)!!$$

Expectation and variance of the distribution $\Gamma(\frac{n}{2}, \frac{1}{2})$.

[Side remark on the meaning of the parameter α in the Gamma distribution (waiting time for a Poisson process).]

An example of a probability spaces where $p(A|B) + p(A|B^c) = 1$ for a fixed event A (not true in general!).

Why we cannot use power series to calculate $\int_{\mathbb{R}} e^{-x^2} \mathrm{d}x$.

Counterexample: S random sign, $X \sim \mathcal{N}(0,1)$, $S \perp X$ then $Y = SX \sim \mathcal{N}(0,1)$. Moreover $P(X+Y=0) = \frac{1}{2}$ therefore $X+Y$ is not Gaussian.

Exercise: $\mathcal{N}(\mu, 1/n) \to \mu$ in law (use the characteristic functions).

Groups and Abelian groups. Real vector spaces.

Lecture 15 – Wednesday, September 23, 2026 (14:00-16:30)

Examples: $\mathbb{R}^n$, real polynomials, continuous functions on the unit interval, random variables on a probability spaces.

Scalar products.

Exercise: prove that $E(X)$ is the constant “nearest" to the random variable X w.r.t. the distance associated to scalar product $\langle X, Y\rangle = E(XY)$.

Cauchy-Schwartz inequality in the plane, for functions, for random variables. Application: $|\rho(X,Y)| \leq 1$.

Lecture 16 – Thursday, September 24, 2026 (14:00-16:30)

- Random vectors, discrete case.
- Marginals.

Exercise. Consider a urn with w white balls and b black balls. You can extract balls with and without replacement.

Replacement case. Define

$$X_1 = \begin{cases} 1 & \text{if the first ball is white} \\ 0 & \text{otherwise} \end{cases} \qquad X_2 = \begin{cases} 1 & \text{if the second ball is white} \\ 0 & \text{otherwise} \end{cases}$$

No replacement case. Define

$$\tilde{X}_1 = \begin{cases} 1 & \text{if the first ball is white} \\ 0 & \text{otherwise} \end{cases} \qquad \tilde{X}_2 = \begin{cases} 1 & \text{if the second ball is white} \\ 0 & \text{otherwise} \end{cases}$$

Prove that the joint distributions of the random vectors (X_1, X_2), $(\tilde{X}_1, \tilde{X}_2)$ are different while the marginal distributions are equal.

One shouldn't never integrate in public.
R. Feynman

Introduction to multiple integrals. The Fubini theorem.
Let $Q = [1, 2] \times [1, 3]$ and $f(x, y) = x^3 \exp(yx^2)$. Show that

$$\int\int_Q f(x,y) = \frac{1}{6}[e^{12} - e^3] - \frac{1}{2}[e^4 - e]$$

Show that to calculate the integral

$$\int\int_Q f(x,y) = \frac{1}{6}[e^{12} - e^3] - \frac{1}{2}[e^4 - e]$$

the order in which we choose the variables matters!!!!
Indeed, integrating by parts

$$\int_1^3 \left(\int_1^2 x^3 e^{yx^2} \mathrm{d}x \right) \mathrm{d}y = \cdots =$$

$$= \int_1^3 \left[\left(\frac{2}{y} - \frac{1}{2y^2} \right) e^{4y} - \frac{1}{2} \left(\frac{1}{y} - \frac{1}{y^2} \right) e^y \right] \mathrm{d}y$$

We cannot go further because functions like

$$\frac{e^{ay}}{y}, \qquad \frac{e^{ay}}{y^2}$$

do not have an antiderivative in terms of "elementary" functions (similarly to the case of e^{-x^2}, $\frac{\sin x}{x}$, $\sin(x^2)$).

Normal domains. Prove that if

$$A = \{(x,y) \in \mathbb{R}^2 | x \in [-1, 1] \quad -1 + x^2 \leq y \leq \sqrt{1 - x^2}\}$$

then

$$\int\int_A xy \, \mathrm{d}x\mathrm{d}y = 0$$

Lecture 17 – Friday, September 25, 2025 (11:00-13:30)

Random vectors: absolutely continuous case.

Proposition $X \perp Y \iff f_{X,Y} = f_X \cdot f_Y$ (No proof.)

Exercise. Suppose that the random vector (X, Y) has the joint density $f_{X,Y}(x, y) = g(x)h(y)$. Then $X \perp Y$.

- Gaussian vectors: the standard case.
- The density of the standard case:

$$\frac{1}{2\pi}\exp\left(-\frac{1}{2}(x^2+y^2)\right)$$

- Gaussian vectors: the general case.
- The density of a bivariate Gaussian (no proof!):

$$f_{X_1,X_2}(x,y) =$$

$$= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\cdot\exp\left(-\frac{1}{2(1-\rho^2)})\left[\left(\frac{x-\mu_1}{\sigma_1}\right)^2 - 2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \left(\frac{y-\mu_2}{\sigma_2}\right)^2\right]\right)$$

- If X, Y are jointly Gaussian random variables then $X \perp Y$ is equivalent to $\text{Cov}(X, Y) = 0$.
- A linear combination of jointly Gaussian random variables is gaussian.
- Remember! S random sign, $X \sim \mathcal{N}(0, 1)$, $S \perp X$ then $Y = SX \sim \mathcal{N}(0, 1)$. Moreover $P(X+Y = 0) = \frac{1}{2}$ therefore $X+Y$ is not Gaussian.

  Therefore X and Y are gaussian but not jointly Gaussian.

Exercise

Let us consider a standard Gaussian vector

$$\begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right)$$

Then

$$P(Z_1 > 0, Z_2 > 0) = \frac{1}{4}, \qquad P(Z_1 > Z_2) = \frac{1}{2}.$$

Lecture 18 – Monday, September 28, 2026 (14:00-16:30)

Simulation 1. – Quiz 3. Exercise 1.
Simulation 1. – Quiz 7. Exercise 1,4
Simulation 3. – Exercise 4.
Simulation 4. – Quiz 7.

- Search for the random variable $\mu(Y)$ nearest to X (w.r.t. the L^2 distance).
- Remember $\mathbb{E}(AB) = \sum_{k,j} a_k b_j P((A = a_k) \cap (B = b_j))$.
- Conditional expectation for discrete random variables.

Lecture 19 – Tuesday, September 29, 2026 (14:00-16:30)

- Let $X_1, X_2 \sim B(1,p)$ and $X_1 \perp\!\!\!\perp X_2$. Prove that

$$E(X_1|X_1 + X_2) = \frac{X_1 + X_2}{2}.$$

- Conditional expectation for discrete random variables (linearity, positivity, constants).

$$\mathbb{E}(g(Y)X|Y) = g(Y)\mathbb{E}(X|Y)$$

- $$\mathbb{E}(g(Y)\mathbb{E}(X|Y)) = \mathbb{E}(g(Y)X)$$

  (Namely: any r.v. $g(Y)$ is orthogonal to the r.v. $(X - \mathbb{E}(X|Y))$.

- Law of iterated expectations

$$\mathbb{E}(\mathbb{E}(X|Y)) = \mathbb{E}(X)$$

- $$\mathbb{E}((X-g(Y))^2) = \mathbb{E}((X-\mathbb{E}(X|Y))^2)+\mathbb{E}((g(Y)-\mathbb{E}(X|Y))^2) \geq \mathbb{E}((X-\mathbb{E}(X|Y))^2)$$

  Namely: $\mathbb{E}(X|Y)$ is the random variable $\mu(Y)$ nearest to X (w.r.t. the L^2 distance).

$$\mathbb{E}(g(Y)|Y) = g(Y)$$

$$X \perp Y \Longrightarrow \mathbb{E}(X|Y) = \mathbb{E}(X)$$

Some counterexamples.

$$\mathbb{E}(X|Y) = \mathbb{E}(X) \qquad \Longrightarrow \qquad \mathrm{Cov}(X, Y) = 0$$

The viceversa it is not true (use a symmetric distribution X such that $\mathrm{Cov}(X^2, X) = 0$ but $\mathbb{E}(X^2|X) = X^2 \neq \mathbb{E}(X^2)$).

- S random sign, $X \sim \mathcal{N}(0, 1)$, $S \perp X$ and $Y := SX$ (we proved $Y \sim \mathcal{N}(0, 1)$ and $Y \not\perp X$). Then

$$\mathbb{E}(SX|X) = X\mathbb{E}(S|X) = X\mathbb{E}(S) = 0$$

and

$$\mathbb{E}(SX) = \mathbb{E}(Y) = 0$$

so that $\mathbb{E}(SX|X) = \mathbb{E}(SX)$ but $SX \not\perp X$.

- Conclusion: we have that

$$X \perp Y \qquad \Longrightarrow \qquad \mathbb{E}(X|Y) = \mathbb{E}(X) \qquad \Longrightarrow \qquad \mathrm{Cov}(X, Y) = 0$$

while

$$X \perp Y \qquad \not\Longleftarrow \qquad \mathbb{E}(X|Y) = \mathbb{E}(X) \qquad \not\Longleftarrow \qquad \mathrm{Cov}(X, Y) = 0$$

Lecture 20 – Wednesday, September 30, 2026 (14:00-16:30)

Conditional expectation for jointly Gaussian random variable in some steps.

- If $\mathbb{E}(X) = \mathbb{E}(Y) = 0$ then $(X - \alpha Y)$ and Y are L^2 orthogonal if and only if
$$\alpha = \frac{\mathrm{Cov}(X, Y)}{\mathrm{Var}(Y)}.$$

- Let X, Y be jointly Gaussian, $\mathbb{E}(X) = \mathbb{E}(Y) = 0$ and let

$$Z := X - \frac{\text{Cov}(X,Y)}{\text{Var}(Y)} \cdot Y.$$

Then Z and Y are independent (and $\mathbb{E}(Z) = 0$).

- X, Y jointly Gaussian and $\mathbb{E}(X) = \mathbb{E}(Y) = 0$ implies

$$\mathbb{E}(X|Y) = \frac{\text{Cov}(X,Y)}{\text{Var}Y} \cdot Y = \rho \frac{\sigma_X}{\sigma_Y} \cdot Y$$

- Let $k, h, \alpha \in \mathbb{R}$. Then

$$\text{Cov}(X,Y) = \text{Cov}(X+k, Y+h)$$
$$\mathbb{E}(U|Y) = \mathbb{E}(U|Y+\alpha)$$

- X, Y jointly Gaussian implies

$$\mathbb{E}(X|Y) = \frac{\text{Cov}(X,Y)}{\text{Var}Y}(Y - \mathbb{E}(Y)) + \mathbb{E}(X) = \rho \frac{\sigma_X}{\sigma_Y}(Y - \mu_Y) + \mu_X$$

(Once you calculate a conditional expectation check the law of iterated expectation.)

- The conditional variance

$$\text{Var}(X|Y) = \mathbb{E}((X - \mathbb{E}(X|Y))^2|Y)$$

$$\text{Var}(X|Y) = \mathbb{E}(X^2|Y) - \mathbb{E}(X|Y)^2$$

- The law of total variance

$$\text{Var}(X) = \mathbb{E}(\text{Var}(X|Y)) + \text{Var}(\mathbb{E}(X|Y))$$