भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Probability and Stochastic Processes

Uncorrelatedness and Independence, Cauchy–Schwartz Inequality, Vector Spaces of Random Variables, The $\mathcal{L}^2(\Omega, \mathscr{F}, \mathbb{P})$ Space, $\mathcal{L}^p$ Spaces, Conditional Expectations

Karthik P. N.
Assistant Professor, Department of AI
Email: pnkarthik@ai.iith.ac.in

21/24 October 2024

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Uncorrelatedness and Independence

Uncorrelatedness and Independence

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

Theorem (Uncorrelatedness and Independence)

If $X \perp\!\!\!\perp Y$, then
$$\mathrm{Cov}(X, Y) = 0.$$

The converse is not true in general.
For example, consider
$$X \sim \mathcal{N}(0, 1).$$

Let $Y = X^2$. Then, it is easy to verify that $\mathbb{E}[XY] = \mathbb{E}[X^3] = 0$, and $\mathbb{E}[X]\mathbb{E}[Y] = 0$.
Therefore, $\mathrm{Cov}(X, Y) = 0$, but $X \not\perp\!\!\!\perp Y$.

Proof – 1

Suppose X, Y are simple random variables, $X \perp\!\!\!\perp Y$.

- Let X and Y have the canonical representations

$$X = \sum_{i=1}^{n} a_i \, \mathbf{1}_{A_i}, \qquad Y = \sum_{j=1}^{m} b_j \, \mathbf{1}_{B_j}.$$

- Then, XY will have the canonical representation

$$XY = \sum_{i=1}^{n} \sum_{j=1}^{m} a_i \, b_j \, \mathbf{1}_{A_i \cap B_j}$$

- Then,

$$\mathbb{E}[XY] = \sum_{i=1}^{n} \sum_{j=1}^{m} a_i \, b_j \, \mathbb{P}(A_i \cap B_j) \overset{X \perp\!\!\!\perp Y}{=} \sum_{i=1}^{n} \sum_{j=1}^{m} a_i \, b_j \, \mathbb{P}(A_i) \cdot \mathbb{P}(B_j) = \mathbb{E}[X] \cdot \mathbb{E}[Y].$$

Proof – 2

Suppose X and Y are non-negative random variables, $X \perp\!\!\!\perp Y$.

- There exists $\{X_n\}_{n=1}^{\infty}, \{Y_n\}^{\infty}$ such that X_n, Y_n simple for each n,

$$X_n \uparrow X, \qquad Y_n \uparrow Y \quad \text{pointwise.}$$

- Recall: for each n,

$$X_n = \frac{\lfloor 2^n X \rfloor}{2^n} \mathbf{1}_{\{X<n\}} + n\,\mathbf{1}_{\{X \geq n\}}, \qquad Y_n = \frac{\lfloor 2^n Y \rfloor}{2^n} \mathbf{1}_{\{Y<n\}} + n\,\mathbf{1}_{\{Y \geq n\}}.$$

- What can we say about X_n and Y_n? Ans: $X_n \perp\!\!\!\perp Y_n$ for all n.
- We have

$$\mathbb{E}[XY] \overset{\text{MCT}}{=} \lim_{n\to\infty} \mathbb{E}[X_n Y_n] = \lim_{n\to\infty} \mathbb{E}[X_n]\cdot\mathbb{E}[Y_n] = \lim_{n\to\infty} \mathbb{E}[X_n] \cdot \lim_{n\to\infty} \mathbb{E}[Y_n] = \mathbb{E}[X]\cdot\mathbb{E}[Y].$$

Variance of Sum of Two Random Variables

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$.

Let X, Y be random variables with respect to $\mathscr{F}$.

Lemma (Variance of Sum of Two Random Variables)

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

Variance of Sum of Two Random Variables

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$.
Let X, Y be random variables with respect to $\mathscr{F}$.

Lemma (Variance of Sum of Two Random Variables)

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

Remarks

- More generally,

$$\mathrm{Var}(X_1 + \cdots + X_n) = \sum_{i=1}^{n} \mathrm{Var}(X_i) + 2\sum_{i<j} \mathrm{Cov}(X_i, X_j).$$

- If X, Y are uncorrelated, then

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

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Correlation Coefficient and Cauchy–Schwartz Inequality

Correlation Coefficient

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$.
Let X, Y be random variables with respect to $\mathscr{F}$.

Definition (Correlation Coefficient)

The correlation coefficient of X and Y is defined as

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

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Correlation Coefficient

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$.
Let X, Y be random variables with respect to $\mathscr{F}$.

Definition (Correlation Coefficient)

The correlation coefficient of X and Y is defined as

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

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

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

The Cauchy–Schwartz Inequality

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$.

Theorem (Cauchy–Schwartz 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
$$\mathbb{P}(\{Y - \mathbb{E}[Y] = a\,(X - \mathbb{E}[X])\}) = 1.$$
2. If $\rho_{X,Y} = -1$, then there exists $a < 0$ such that
$$\mathbb{P}(\{Y - \mathbb{E}[Y] = a\,(X - \mathbb{E}[X])\}) = 1.$$

Proof – 1

$\widetilde{X} := X - \mathbb{E}[X], \qquad \widetilde{Y} := Y - \mathbb{E}[Y].$

Clearly,

$$\mathbb{E}\left[\left(\widetilde{X} - \frac{\mathbb{E}[\widetilde{X}\,\widetilde{Y}]}{\mathbb{E}[(\widetilde{Y})^2]}\,\widetilde{Y}\right)^2\right] \geq 0.$$

Expanding the inner squared term and using linearity of expectations, we arrive at the Cauchy–Schwartz inequality.

Proof – 2

Equality in Cauchy–Schwartz inequality:

$$\mathbb{E}\left[\left(\widetilde{X} - \frac{\mathbb{E}[\widetilde{X}\,\widetilde{Y}]}{\mathbb{E}[(\widetilde{Y})^2]}\,\widetilde{Y}\right)^2\right] = 0 \quad \implies \quad \mathbb{P}\left(\widetilde{X} = \frac{\mathbb{E}[\widetilde{X}\,\widetilde{Y}]}{\mathbb{E}[(\widetilde{Y})^2]}\,\widetilde{Y}\right) = 1.$$

- Let $a := \left(\frac{\mathbb{E}[\widetilde{X}\,\widetilde{Y}]}{\mathbb{E}[(\widetilde{Y})^2]}\right)^{-1}$.
- If $\rho_{X,Y} = 1$, then $a > 0$.
- If $\rho_{X,Y} = -1$, then $a < 0$.

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Vector Spaces of Random Variables

Vector Spaces of Random Variables

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$

- Fix $p = 2$

Vector Spaces of Random Variables

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$

- Fix $p = 2$

- Fix $1 \leq p < +\infty$

Vector Spaces of Random Variables

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$

- Fix $p = 2$
- We say $X \in \mathcal{L}^2(\Omega, \mathscr{F}, \mathbb{P})$ if

$$\mathbb{E}[X^2] < +\infty.$$

- Fix $1 \leq p < +\infty$

Vector Spaces of Random Variables

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$

- Fix $p = 2$
- We say $X \in \mathcal{L}^2(\Omega, \mathscr{F}, \mathbb{P})$ if

$$\mathbb{E}[X^2] < +\infty.$$

- Fix $1 \leq p < +\infty$
- We say $X \in \mathcal{L}^p(\Omega, \mathscr{F}, \mathbb{P})$ if

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

Vector Spaces of Random Variables

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$

- Fix $p = 2$
- We say $X \in \mathcal{L}^2(\Omega, \mathscr{F}, \mathbb{P})$ if

$$\mathbb{E}[X^2] < +\infty.$$

- Define 2-norm of X as

$$\|X\|_2 := (\mathbb{E}[X^2])^{1/2}.$$

- Fix $1 \leq p < +\infty$
- We say $X \in \mathcal{L}^p(\Omega, \mathscr{F}, \mathbb{P})$ if

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

Vector Spaces of Random Variables

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$

- Fix $p = 2$
- We say $X \in \mathcal{L}^2(\Omega, \mathscr{F}, \mathbb{P})$ if

$$\mathbb{E}[X^2] < +\infty.$$

- Define 2-norm of X as

$$\|X\|_2 := (\mathbb{E}[X^2])^{1/2}.$$

- Fix $1 \leq p < +\infty$
- We say $X \in \mathcal{L}^p(\Omega, \mathscr{F}, \mathbb{P})$ if

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

- Define p-norm of X as

$$\|X\|_p := (\mathbb{E}[|X|^p])^{1/p}.$$

Vector Spaces of Random Variables

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$

- Fix $p = 2$
- We say $X \in \mathcal{L}^2(\Omega, \mathscr{F}, \mathbb{P})$ if

$$\mathbb{E}[X^2] < +\infty.$$

- Define 2-norm of X as

$$\|X\|_2 := (\mathbb{E}[X^2])^{1/2}.$$

- $\mathcal{L}^2(\Omega, \mathscr{F}, \mathbb{P})$ is a normed vector space over $\mathbb{R}$
  - $X, Y \in \mathcal{L}^2 \implies X + Y \in \mathcal{L}^2$
  - $X \in \mathcal{L}^2, \alpha \in \mathbb{R} \implies \alpha X \in \mathcal{L}^2$

- Fix $1 \leq p < +\infty$
- We say $X \in \mathcal{L}^p(\Omega, \mathscr{F}, \mathbb{P})$ if

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

- Define p-norm of X as

$$\|X\|_p := (\mathbb{E}[|X|^p])^{1/p}.$$

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Vector Spaces of Random Variables

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$

- Fix $p = 2$
- We say $X \in \mathcal{L}^2(\Omega, \mathscr{F}, \mathbb{P})$ if
$$\mathbb{E}[X^2] < +\infty.$$
- Define 2-norm of X as
$$\|X\|_2 := (\mathbb{E}[X^2])^{1/2}.$$
- $\mathcal{L}^2(\Omega, \mathscr{F}, \mathbb{P})$ is a normed vector space over $\mathbb{R}$
  - $X, Y \in \mathcal{L}^2 \implies X + Y \in \mathcal{L}^2$
  - $X \in \mathcal{L}^2, \alpha \in \mathbb{R} \implies \alpha X \in \mathcal{L}^2$

- Fix $1 \leq p < +\infty$
- We say $X \in \mathcal{L}^p(\Omega, \mathscr{F}, \mathbb{P})$ if
$$\mathbb{E}[|X|^p] < +\infty.$$
- Define p-norm of X as
$$\|X\|_p := (\mathbb{E}[|X|^p])^{1/p}.$$
- $\mathcal{L}^p(\Omega, \mathscr{F}, \mathbb{P})$ is a normed vector space over $\mathbb{R}$
  - $X, Y \in \mathcal{L}^p \implies X + Y \in \mathcal{L}^p$
  - $X \in \mathcal{L}^p, \alpha \in \mathbb{R} \implies \alpha X \in \mathcal{L}^p$

$\mathcal{L}^2$ Space in More Depth

Consider the space $\mathcal{L}^2 = \mathcal{L}^2(\Omega, \mathscr{F}, \mathbb{P})$.

- If $X, Y \in \mathcal{L}^2$, then $|\mathbb{E}[X]| < +\infty$, $|\mathbb{E}[Y]| < +\infty$
- If $X, Y \in \mathcal{L}^2$, then $|\mathrm{Cov}(X, Y)| < +\infty$ (Cauchy–Schwartz)

Let $\mathcal{S} \subseteq \mathcal{L}^2$ be defined as

$$\mathcal{S} := \left\{ X \in \mathcal{L}^2 : \mathbb{E}[X] = 0 \right\}.$$

Define $\langle\ ,\ \rangle : \mathcal{S} \times \mathcal{S} \to \mathbb{R}$ as

$$\langle X, Y \rangle := \mathbb{E}[XY], \qquad X, Y \in \mathcal{S}.$$

Vector Spaces of Random Variables

Note the following properties of $\langle\ ,\ \rangle$:

- $\langle X, Y\rangle = \langle Y, X\rangle$

Note

$\langle\ ,\ \rangle : \mathcal{S} \times \mathcal{S} \to \mathbb{R}$ is NOT an inner product operator (because of bullet number 3 above).

Vector Spaces of Random Variables

Note the following properties of $\langle \; , \; \rangle$:

- $\langle X, Y \rangle = \langle Y, X \rangle$
- $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$

Note

$\langle \; , \; \rangle : \mathcal{S} \times \mathcal{S} \to \mathbb{R}$ is NOT an inner product operator (because of bullet number 3 above).

Vector Spaces of Random Variables

Note the following properties of $\langle \, , \, \rangle$:

- $\langle X, Y \rangle = \langle Y, X \rangle$
- For all $a, b \in \mathbb{R}$ and $X, Y, Z \in \mathcal{S}$,

$$\langle aX + bY, Z \rangle = a \langle X, Z \rangle + b \langle Y, Z \rangle.$$

- $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$

Note

$\langle \, , \, \rangle : \mathcal{S} \times \mathcal{S} \to \mathbb{R}$ is NOT an inner product operator (because of bullet number 3 above).

Vector Spaces of Random Variables

Note the following properties of $\langle \, , \, \rangle$:

- $\langle X, Y \rangle = \langle Y, X \rangle$
- For all $a, b \in \mathbb{R}$ and $X, Y, Z \in \mathcal{S}$,

$$\langle aX + bY, Z \rangle = a \langle X, Z \rangle + b \langle Y, Z \rangle .$$

- $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$
- For all $a, b \in \mathbb{R}$ and $\mathbf{x}, \mathbf{y}, \mathbf{z}$,

$$\langle a\mathbf{x} + b\mathbf{y}, \mathbf{z} \rangle = a \langle \mathbf{x}, \mathbf{z} \rangle + b \langle \mathbf{y}, \mathbf{z} \rangle .$$

Note

$\langle \, , \, \rangle : \mathcal{S} \times \mathcal{S} \to \mathbb{R}$ is NOT an inner product operator (because of bullet number 3 above).

Vector Spaces of Random Variables

Note the following properties of $\langle \ , \ \rangle$:

- $\langle X, Y \rangle = \langle Y, X \rangle$
- For all $a, b \in \mathbb{R}$ and $X, Y, Z \in \mathcal{S}$,

$$\langle aX + bY, Z \rangle = a \langle X, Z \rangle + b \langle Y, Z \rangle.$$

- $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$
- For all $a, b \in \mathbb{R}$ and $\mathbf{x}, \mathbf{y}, \mathbf{z}$,

$$\langle a\mathbf{x} + b\mathbf{y}, \mathbf{z} \rangle = a \langle \mathbf{x}, \mathbf{z} \rangle + b \langle \mathbf{y}, \mathbf{z} \rangle.$$

- $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$ for all $\mathbf{x}$. Furthermore,

$$\langle \mathbf{x}, \mathbf{x} \rangle = 0 \implies \mathbf{x} = \mathbf{0}$$

Note

$\langle \ , \ \rangle : \mathcal{S} \times \mathcal{S} \to \mathbb{R}$ is NOT an inner product operator (because of bullet number 3 above).

Vector Spaces of Random Variables

Note the following properties of $\langle \, , \, \rangle$:

- $\langle X, Y \rangle = \langle Y, X \rangle$
- For all $a, b \in \mathbb{R}$ and $X, Y, Z \in \mathcal{S}$,
$$\langle aX + bY, Z \rangle = a \langle X, Z \rangle + b \langle Y, Z \rangle .$$
- $\langle X, X \rangle \geq 0$ for all $X \in \mathcal{S}$. Furthermore,
$$\langle X, X \rangle = 0 \implies \mathbb{P}(\{X = 0\}) = 1.$$

- $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$
- For all $a, b \in \mathbb{R}$ and $\mathbf{x}, \mathbf{y}, \mathbf{z}$,
$$\langle a\mathbf{x} + b\mathbf{y}, \mathbf{z} \rangle = a \langle \mathbf{x}, \mathbf{z} \rangle + b \langle \mathbf{y}, \mathbf{z} \rangle .$$
- $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$ for all $\mathbf{x}$. Furthermore,
$$\langle \mathbf{x}, \mathbf{x} \rangle = 0 \implies \mathbf{x} = \mathbf{0}$$

Note

$\langle \, , \, \rangle : \mathcal{S} \times \mathcal{S} \to \mathbb{R}$ is NOT an inner product operator (because of bullet number 3 above).

Turning $\langle\ ,\ \rangle$ into an Inner Product

Define the relation $\overset{\mathrm{R}}{\sim}$ on $\mathcal{S} \times \mathcal{S}$ as follows:

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

The above relation satisfies the following properties:

- $X \overset{\mathrm{R}}{\sim} X$ for all $X \in \mathcal{S}$
- For all $X, Y \in \mathcal{S}$,
$$X \overset{\mathrm{R}}{\sim} Y \implies Y \overset{\mathrm{R}}{\sim} X$$
- For all $X, Y, Z \in \mathcal{S}$,
$$X \overset{\mathrm{R}}{\sim} Y, \quad Y \overset{\mathrm{R}}{\sim} Z \quad \implies \quad X \overset{\mathrm{R}}{\sim} Z$$

The relation $\overset{\mathrm{R}}{\sim}$ is an equivalence relation on $\mathcal{S} \times \mathcal{S}$.

Equivalence Classes and Inner Product

For any $X \in \mathcal{S}$, let the equivalence class of X under the relation $\overset{\mathrm{R}}{\sim}$ be defined as

$$[X] := \left\{ Z \in \mathcal{S} : \quad \mathbb{P}(\{X = Z\}) = 1 \right\}.$$

Let $\mathcal{C}$ denote the set of all equivalence classes.

- For all $A_1, A_2 \in [X]$ and $B_1, B_2 \in [Y]$,

  $$\langle A_1, B_1 \rangle = \langle A_2, B_2 \rangle.$$

  We denote this common value by $\langle [X], [Y] \rangle$ or simply $\langle X, Y \rangle$
- In the new interpretation of inner products (on equivalence classes), we have

$$\langle X, X \rangle = 0 \quad \Longleftrightarrow \quad \langle [X], [X] \rangle = 0 \quad \Longleftrightarrow \quad [X] = [0].$$

A New Interpretation of $\rho_{X,Y}$

- For $X \in \mathcal{S}$, define

$$\|X\|_2 := \sqrt{\langle X, X\rangle} = \sqrt{\langle [X], [X]\rangle} = \mathbb{E}[X^2].$$

- For $X, Y \in \mathcal{S}$, we have

$$\rho_{X,Y} = \frac{\mathbb{E}[XY]}{\sqrt{\mathbb{E}[X^2] \cdot \mathbb{E}[Y^2]}} = \frac{\langle X, Y\rangle}{\|X\|_2 \cdot \|Y\|_2}.$$

Thus, $\rho_{X,Y}$ represents the cosine of the angle between X and Y

Conditional Expectations

Conditional Expectation

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$.
Let X, Y be random variables with respect to $\mathscr{F}$.

Objective

To define the following quantities:

- $\mathbb{E}[X|\{Y = y\}]$, for any $y \in \mathbb{R}$.
- $\mathbb{E}[X|Y]$.

Programme:
We shall define the above quantities by considering X discrete/continuous, and Y discrete/continuous.

Case 1: X Discrete, Y Discrete

Let X, Y have the joint PMF $p_{X,Y}$.

- Step 1: Conditional PMF of X, conditioned on the event $\{Y = y\}$:

$$p_{X|Y=y}(x) = \frac{p_{X,Y}(x,y)}{p_Y(y)}, \qquad x \in \mathbb{R}.$$

- Step 2: The quantity $\mathbb{E}[X|\{Y = y\}]$ is defined as the expectation with respect to the conditional PMF $p_{X|Y=y}$, i.e.,

$$\mathbb{E}[X|\{Y = y\}] := \sum_{x \in \mathbb{R}} x \cdot p_{X|Y=y}(x).$$

Case 1: X Discrete, Y Discrete

Let X, Y have the joint PMF $p_{X,Y}$.

- Step 3: Define the function $\psi_1 : \mathbb{R} \to \mathbb{R}$ as

$$\psi_1(y) := \begin{cases} \mathbb{E}[X|\{Y=y\}], & p_Y(y) > 0, \\ 0, & p_Y(y) = 0. \end{cases}$$

- Step 4: The quantity $\mathbb{E}[X|Y]$ is simply defined as

$$\mathbb{E}[X|Y] = \psi_1(Y).$$

Case 2: X Continuous, Y Continuous

Let X, Y have the joint PDF $f_{X,Y}$.

- Step 1: Conditional PDF of X, conditioned on the event $\{Y = y\}$:

$$f_{X|Y=y}(x) = \frac{f_{X,Y}(x, y)}{f_Y(y)}, \qquad x \in \mathbb{R}.$$

- Step 2: The quantity $\mathbb{E}[X|\{Y = y\}]$ is defined as the expectation with respect to the conditional PDF $f_{X|Y=y}$, i.e.,

$$\mathbb{E}[X|\{Y = y\}] := \int_{-\infty}^{+\infty} x \cdot f_{X|Y=y}(x).$$

Case 2: X Continuous, Y Continuous

Let X, Y have the joint PDF $p_{X,Y}$.

- Step 3: Define the function $\psi_2 : \mathbb{R} \to \mathbb{R}$ as

$$\psi_2(y) := \begin{cases} \mathbb{E}[X|\{Y = y\}], & f_Y(y) > 0, \\ 0, & f_Y(y) = 0. \end{cases}$$

- Step 4: The quantity $\mathbb{E}[X|Y]$ is simply defined as

$$\mathbb{E}[X|Y] = \psi_2(Y).$$