Conditional Expectation

- Say that we observe that $Y=y$ and wish to use this information to predict X. The conditional PMF $P_{X|Y}(x|y)$ or PDF $f_{X|Y}(x|y)$ tells us the distribution, but what about just the average value of X given that $Y=y$?
  - → Examples: * Observe the temperature today and try to predict the average temperature tomorrow.
    * Take a noisy measurement from a reactor and determine the best control input (on average).

- The conditional expected value $\mathbb{E}[X \mid Y=y]$ of X given the event $\{Y=y\}$ is

$$\boxed{\begin{aligned} &\text{Discrete:} && \mathbb{E}[X \mid Y=y] = \sum_{x \in R_X} x \, P_{X|Y}(x|y) \\ &\text{Continuous:} && \mathbb{E}[X \mid Y=y] = \int_{-\infty}^{\infty} x \, f_{X|Y}(x|y) \, dx \end{aligned}}$$

- Recall that the expected value $\mathbb{E}[X]$ is a number. all of the randomness in X is averaged out.

- In contrast, the conditional expected value $\mathbb{E}[X \mid Y = y]$ is a function of y. Roughly speaking, all of the randomness in X that does not depend on Y is averaged out.

  → Sometimes, it helps to write $h(y) = \mathbb{E}[X \mid Y = y]$ to remind ourselves this is a deterministic function of y.

- Recall also that if we plug in the random variable Y into a function $h(y)$, we get $h(Y)$, which is a random variable itself.

- In this sense, we use the notation $\mathbb{E}[X \mid Y]$ to represent $h(Y)$ where $h(y) = \mathbb{E}[X \mid Y = y]$. Thus, $\mathbb{E}[X \mid Y]$ is a random variable.

- The conditional expected value $\mathbb{E}[g(X) \mid Y=y]$ of a function $g(X)$ given the event $\{Y=y\}$ is

$$\boxed{\begin{aligned} \text{Discrete: } \quad & \mathbb{E}[g(X) \mid Y=y] = \sum_{x \in R_X} g(x)\, p_{X|Y}(x|y) \\ \text{Continuous: } \quad & \mathbb{E}[g(X) \mid Y=y] = \int_{-\infty}^{\infty} g(x)\, f_{X|Y}(x|y)\, dx \end{aligned}}$$

- Writing $h(y) = \mathbb{E}[g(X) \mid Y=y]$ helps remind us that $\mathbb{E}[g(X) \mid Y=y]$ is a deterministic function of y.

- The notation $\mathbb{E}[g(X) \mid Y]$ refers to substituting the random variable Y into the function $h(y) = \mathbb{E}[g(X) \mid Y=y]$ and thus $\mathbb{E}[g(X) \mid Y]$ is a random variable.

- **Conditional Expectation Properties:**

→ If X and Y are independent, $\mathbb{E}[X \mid Y=y] = \mathbb{E}[X]$ and
$\mathbb{E}[g(X) \mid Y=y] = \mathbb{E}[g(X)]$.

Why?

$$\mathbb{E}[X \mid Y=y] = \sum_{x \in R_X} x\, P_{X|Y}(x|y) \overset{\text{Independence}}{=} \sum_{x \in R_X} x\, P_X(x) = \mathbb{E}[X].$$

→ Law of Total Expectation: $\boxed{\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X|Y]] \qquad \mathbb{E}[g(X)] = \mathbb{E}[\mathbb{E}[g(X)|Y]]}$

Why?

$$\mathbb{E}[\mathbb{E}[X|Y]] = \mathbb{E}[h(Y)] = \sum_{y \in R_Y} h(y)\, P_Y(y)$$

$$h(y) = \mathbb{E}[X|Y=y] = \sum_{x \in R_X} x\, P_{X|Y}(x|y)$$

$$= \sum_{y \in R_Y} \sum_{x \in R_X} x\, \underbrace{P_{X|Y}(x|y)\, P_Y(y)}_{\downarrow \text{ Multiplication Rule}}$$

$$= \sum_{y \in R_Y} \sum_{x \in R_X} x\, P_{XY}(x,y) = \mathbb{E}[X]$$

• **Example:**

$P_{X\|Y}(x\|y)$		$x=1$	$x=2$	$x=3$
y	1	1	0	0
	2	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{2}$

→ Calculate $\mathbb{E}[X \mid Y=y]$.

$$\mathbb{E}[X \mid Y=y] = \sum_{x \in R_X} x\, P_{X|Y}(x|y) = \begin{cases} \sum_{x \in R_X} x\, P_{X|Y}(x|1) & y=1 \\ \sum_{x \in R_X} x\, P_{X|Y}(x|2) & y=2 \end{cases}$$

$$= \begin{cases} 1\cdot 1 + 2\cdot 0 + 3\cdot 0 & y=1 \\ 1\cdot \frac{1}{4} + 2\cdot\frac{1}{4} + 3\cdot\frac{1}{2} & y=2 \end{cases} = \begin{cases} 1 & y=1 \\ \frac{9}{4} & y=2 \end{cases}$$

→ Calculate $\mathbb{E}[\mathbb{E}[X|Y]]$. We don't have enough information! Need $P_Y(y)$.

Say $P_Y(y) = \begin{cases} \frac{1}{3} & y=1 \\ \frac{2}{3} & y=2 \end{cases}$

$$\mathbb{E}[\mathbb{E}[X|Y]] = \sum_{y \in R_Y} \mathbb{E}[X|Y=y]\, P_Y(y) = \mathbb{E}[X|Y=1]\, P_Y(1) + \mathbb{E}[X|Y=2]\, P_Y(2)$$

$$= 1\cdot\frac{1}{3} + \frac{9}{4}\cdot\frac{2}{3} = \frac{2+9}{6} = \frac{11}{6} = \mathbb{E}[X]$$

- **Example**: X given $Y=y$ is Geometric(y).

$$P_Y(y) = \begin{cases} \frac{1}{4} & y=\frac{1}{2} \\ \frac{3}{4} & y=\frac{2}{3} \end{cases} \qquad P_{X|Y}(x|y) = \begin{cases} y(1-y)^{x-1} & y=\frac{1}{2},\frac{2}{3},\ x=1,2,3,\dots \\ 0 & \text{otherwise} \end{cases}$$

Determine $\mathbb{E}[X|Y=y]$ and $\mathbb{E}[X]$.

$$\mathbb{E}[X|Y=y] = \frac{1}{y} \qquad \mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X|Y]] = \sum_{y\in R_Y} \frac{1}{y} P_Y(y) = \frac{1}{\left(\frac{1}{2}\right)}\cdot\frac{1}{4} + \frac{1}{\left(\frac{2}{3}\right)}\cdot\frac{3}{4}$$

$$= \frac{1}{2} + \frac{9}{8} = \frac{13}{8}$$

Geometric(p) Mean: $\frac{1}{p}$ Law of Total Expectation

- **Example**: X given $Y=y$ is Gaussian$\left(\frac{y^2}{3}, 4\right)$. Y is Exponential(2).

Determine $\mathbb{E}[X|Y=y]$ and $\mathbb{E}[X]$.

$$\mathbb{E}[X|Y=y] = \frac{y^2}{3} \qquad \mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X|Y]] = \mathbb{E}\left[\frac{Y^2}{3}\right] = \frac{1}{3}\cdot\frac{1}{2} = \frac{1}{6}$$

Gaussian(μ,σ^2) Mean: μ

$$\mathbb{E}[Y^2] = \mathrm{Var}[Y] + (\mathbb{E}[Y])^2$$

$$= \frac{1}{2^2} + \frac{1}{2^2} = \frac{1}{2}$$

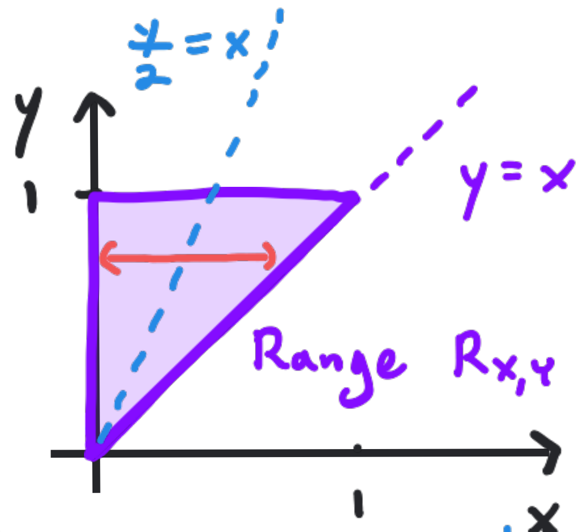
Alternate Variance Formula

$\mathrm{Var}[Y] = \mathbb{E}[Y^2] - (\mathbb{E}[Y])^2$

Exponential(λ) Mean: $\frac{1}{\lambda}$

Variance: $\frac{1}{\lambda^2}$

• **Example:** $f_{X,Y}(x,y) = \begin{cases} 2 & 0 \le x \le y,\ 0 \le y \le 1 \\ 0 & \text{otherwise} \end{cases}$



This makes sense! Average is between 0 and y.

→ Calculate $\mathbb{E}[X \mid Y=y]$ and $\mathbb{E}[X]$.

$\boxed{\text{Step 1}}$ Determine conditional PDF $f_{X|Y}(x|y)$.

→ First, need marginal PDF $f_Y(y)$.

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx = \int_0^y 2\,dx = (2x)\Big|_0^y = \begin{cases} 2y & 0 \le y \le 1 \\ 0 & \text{otherwise} \end{cases}$$

→ Now, plug into conditional PDF formula.

$$f_{X|Y}(x|y) = \begin{cases} \dfrac{f_{X,Y}(x,y)}{f_Y(y)} & (x,y) \in R_{X,Y} \\ 0 & \text{otherwise} \end{cases} = \begin{cases} \dfrac{\cancel{2}}{\cancel{2}y} & 0 \le x \le y,\ 0 \le y \le 1 \\ 0 & \text{otherwise} \end{cases}$$

$\boxed{\text{Step 2}}$ Plug into conditional expectation formula.

$$\mathbb{E}[X \mid Y=y] = \int_{-\infty}^{\infty} x f_{X|Y}(x|y)\,dx = \int_0^y x \cdot \frac{1}{y}\,dx = \left(\frac{1}{2}x^2\right)\Big|_0^y \cdot \frac{1}{y} = \frac{y}{2}$$

$$\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X \mid Y]] = \mathbb{E}\left[\frac{Y}{2}\right] = \int_{-\infty}^{\infty} \frac{y}{2} f_Y(y)\,dy = \int_0^1 y^2\,dy = \left(\frac{1}{3}y^3\right)\Big|_0^1 = \frac{1}{3}$$