Note 18

Section 4.3. Conditional Distributions

Q · How does the distribution of a RV, say X, change given information of another RV, say Y?

· How to systematically work with such "conditional distributions"?

1 Conditional PMFs

DEF Suppose X & Y have a joint PMF. Then the conditional PMF of X given $Y=y$ is defined by

$$p_{X|Y}(x|y) := P(X=x \mid Y=y) = \frac{p_{X,Y}(x,y)}{p_Y(y)},$$

Why?

provided $p_Y(y) > 0$.

Ex 1 Let X & Y have the joint PMF:

$$p_{X,Y}(x,y) = \frac{x+y}{21}, \qquad x = 1,2,3, \quad y = 1,2.$$

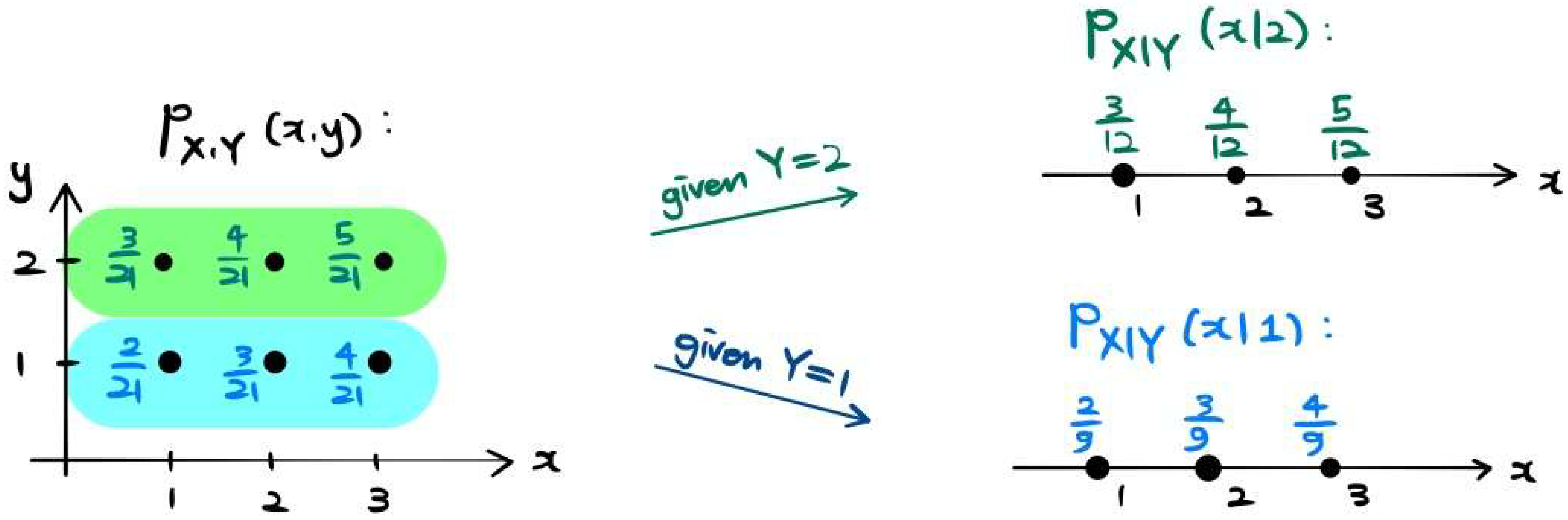
Then the marginal PMF $p_Y(y)$ is

$$p_Y(y) = \sum_{x=1}^{3} p_{X,Y}(x,y) = \sum_{x=1}^{3} \frac{x+y}{21} = \frac{6+3x}{21}, \qquad y = 1,2.$$

So the conditional PMF of X given $Y=y$, for $y = 1, 2$, is

$$p_{X|Y}(x|y) = \frac{p_{X,Y}(x,y)}{p_Y(y)} = \frac{(x+y)/21}{(3y+6)/21} = \frac{x+y}{3y+6}, \qquad x = 1,2,3.$$

Visually,



For instance,

$$P(X=2 \mid Y=1) = P_{X|Y}(2|1) = \frac{3}{9} = \frac{1}{3}.$$

Similarly, $P_{Y|X}(y|x)$ can be computed as

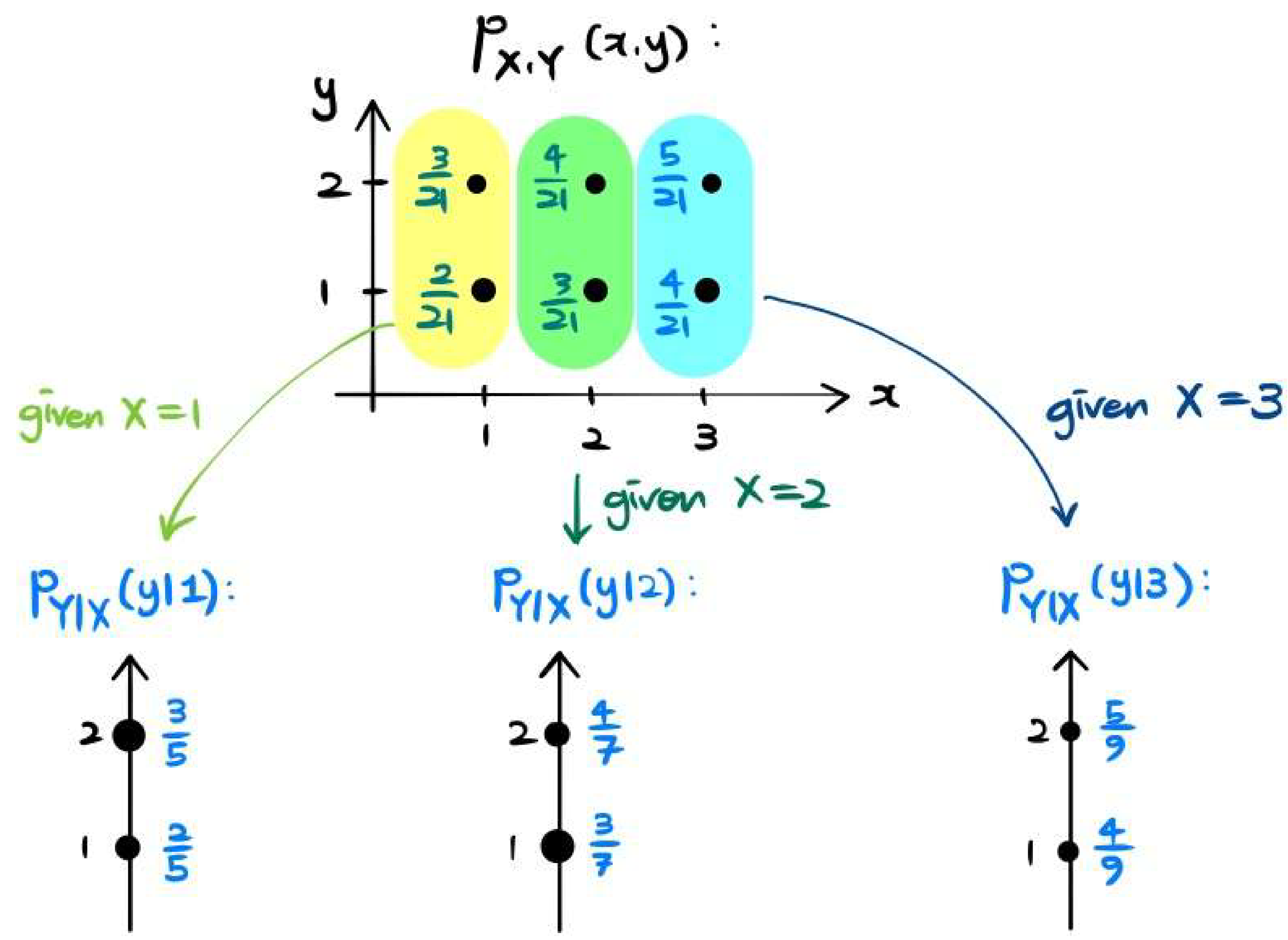
$$P_{Y|X}(y|x) = P(Y=y \mid X=x) = \frac{P_{X,Y}(x,y)}{P_X(x)}$$

$$= \frac{(x+y)/21}{(2x+3)/21} = \frac{x+y}{2x+3}, \quad y=1,2, \quad \text{when } x=1,2,3.$$

For example,

$$P(Y=1 \mid X=3) = P_{Y|X}(1|3) = \frac{4}{9}.$$

Visually,



$\square$

NOTE Each conditional PMF is a PMF. Indeed, given $P_X(x) > 0$,

(a) $0 \le P_{Y|X}(y|x) \le 1$ for any y,

(b) $$\sum_y P_{Y|X}(y|x) = \sum_y \frac{P_{X,Y}(x,y)}{P_X(x)} = \frac{1}{P_X(x)} \underbrace{\sum_y P_{X,Y}(x,y)}_{=P_X(x)} = 1$$

- The previous observation tells that we can compute

$$P(Y \in A \mid X = x) = \sum_{y \in A} P_{Y|X}(y|x).$$

$\boxed{2}$ Conditional Expectation given $X = x$.

- We have learned that :
  - ▷ $P_{Y|X}(y|x)$ describes the dist. of Y, given $X = x$.
  - ▷ Conditional prob. of events in terms of Y, given $X = x$, can be computed by $P_{Y|X}(y|x)$.

- This suggests : We may consider the expected value of $u(Y)$ given $X = x$. Indeed, we define the conditional expectation as :

$$E[u(Y) \mid X = x] = \sum_{y} u(y) P_{Y|X}(y|x).$$

- Two special cases :
  - ▷ the conditional mean of Y, given $X = x$, is

$$E(Y \mid X = x) = \sum_{y} y P_{Y|X}(y|x).$$

  - ▷ the conditional variance of Y, given X, is

$$\begin{aligned} Var(Y \mid X = x) &:= E[(Y - E(Y \mid X = x))^2 \mid X = x] \\ &= E(Y^2 \mid X = x) - E(Y \mid X = x)^2. \end{aligned}$$

Ex 1 (continued) Recall $P_{Y|X}(y|x) = \frac{x+y}{2y+3}$, $x = 1, 2, 3$, $y = 1, 2$. Then

$$E(Y \mid X = 2) = \sum_{y=1}^{2} y P_{Y|X}(y|2) = (1) \cdot \left(\frac{3}{7}\right) + (2) \cdot \left(\frac{4}{7}\right) = \frac{11}{7}$$

and

$$E(Y^2 \mid X = 2) = \sum_{y=1}^{2} y^2 \cdot P_{Y|X}(y|2) = (1)^2 \cdot \left(\frac{3}{7}\right) + (2)^2 \cdot \left(\frac{4}{7}\right) = \frac{19}{7}$$

So we get

$$\begin{aligned} Var(Y \mid X=2) &= E(Y^2 \mid X=2) - E(Y \mid X=2)^2 \\ &= \frac{19}{7} - \left(\frac{11}{7}\right)^2 = \frac{12}{49}. \end{aligned}$$

$\boxed{3}$ Conditional Expectation as Random Variable

- We may think of $E(Y \mid X=x)$ as "expectation of Y, knowing that $X=x$ is observed".

$$X = \boxed{?} \xrightarrow[X \text{ is observed}]{\cdots\cdots} X = \boxed{x} \xrightarrow{\cdots\cdots} [\text{condi. exp. of } Y] = E(Y \mid X=x).$$

- Then it makes sense to consider "expectation of Y given X, before X is observed". Then

$$\begin{aligned} &X = x_1, \text{ w/ prob } p_X(x_1) \quad \cdots\cdots\to \quad E(Y \mid X=x_1) \\ &X = x_2, \text{ w/ prob } p_X(x_2) \quad \cdots\cdots\to \quad E(Y \mid X=x_2) \\ &X = x_3, \text{ w/ prob } p_X(x_3) \quad \cdots\cdots\to \quad E(Y \mid X=x_3) \\ &\qquad\vdots \qquad\qquad\qquad\qquad\qquad\qquad\qquad \vdots \end{aligned}$$

So, what we get is a random variable that takes the value $E(Y \mid X=x)$ with when $X=x$ occurs.

DEF The conditional expectation of Y given X, denoted by

$$E(Y \mid X),$$

is the random variable such that:

$$E(Y \mid X) = E(Y \mid X=x) \quad \text{whenever } X=x \text{ occurs.}$$

Note $E(Y|X)$ is always a function of X.

Ex 1 (continued) Recall that $p_{Y|X}(y|x) = \frac{x+y}{2y+3}$, $y = 1, 2$ when $x = 1, 2, 3$.

Then $E(Y|X)$ is the random variable taking values:

$$\begin{cases} E(Y|X=1) = \frac{8}{5} & \text{w/ prob. } p_X(1) = \frac{5}{21} \\ E(Y|X=2) = \frac{11}{7} & \text{w/ prob. } p_X(2) = \frac{7}{21} \\ E(Y|X=3) = \frac{14}{9} & \text{w/ prob. } p_X(3) = \frac{9}{21}. \end{cases}$$

- There are two useful theorems involving conditional expectations:

THM 1 (Law of Total Expectation)

$$E[E(Y|X)] = E(Y).$$

Pf)

$$\begin{aligned} E[E(Y|X)] &= \sum_x E(Y|X=x)\, p_X(x) && \text{(from the def. of } E(Y|X)\text{)} \\ &= \sum_x \left(\sum_y y\, p_{Y|X}(y|x) \right) p_X(x) && \text{(from the def. of } E(Y|X=x)\text{)} \\ &= \sum_{x,y} y\, \underbrace{p_{Y|X}(y|x)\, p_X(x)}_{= p_{X,Y}(x,y)} && \text{(from the def. of } p_{Y|X}(y|x)\text{)} \\ &= \sum_{x,y} y\, p_{X,Y}(x,y) \\ &= E(Y). \end{aligned}$$

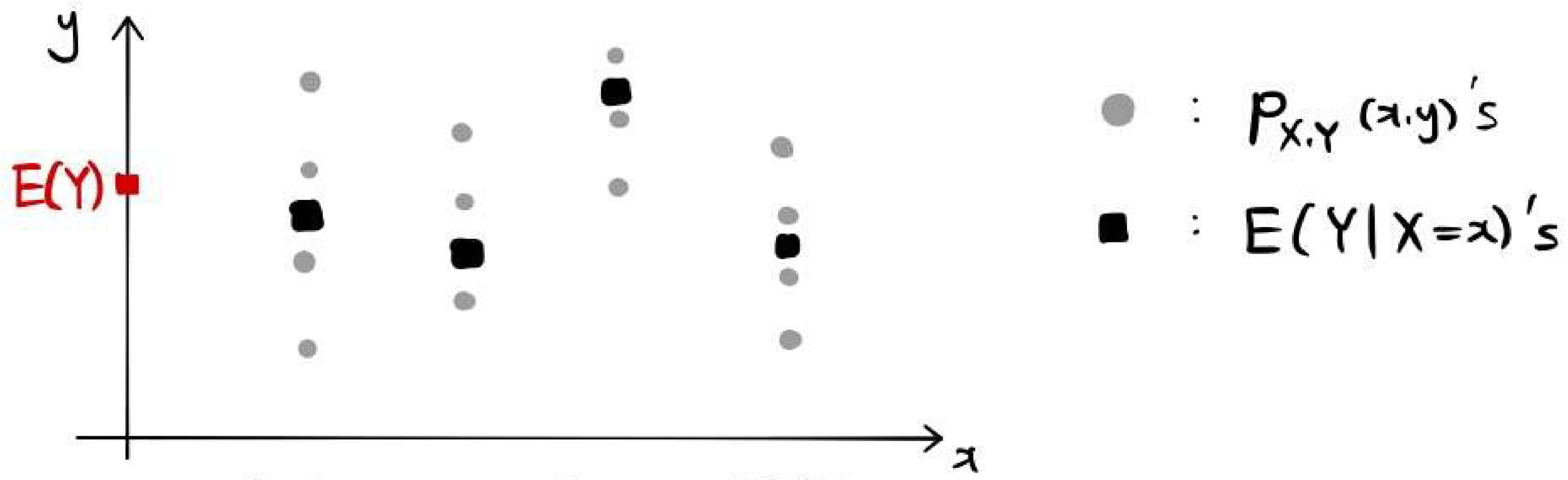
$\square$

Note The proof suggests another way of stating LTE:

$$E(Y) = \sum_x E(Y|X=x)\, P(X=x).$$

Compare this with the Law of Total Probability!

Figure



A diagram explaining LTE

DEF The conditional variance of Y given X, denoted by

$$Var(Y|X)$$

is the random variable such that:

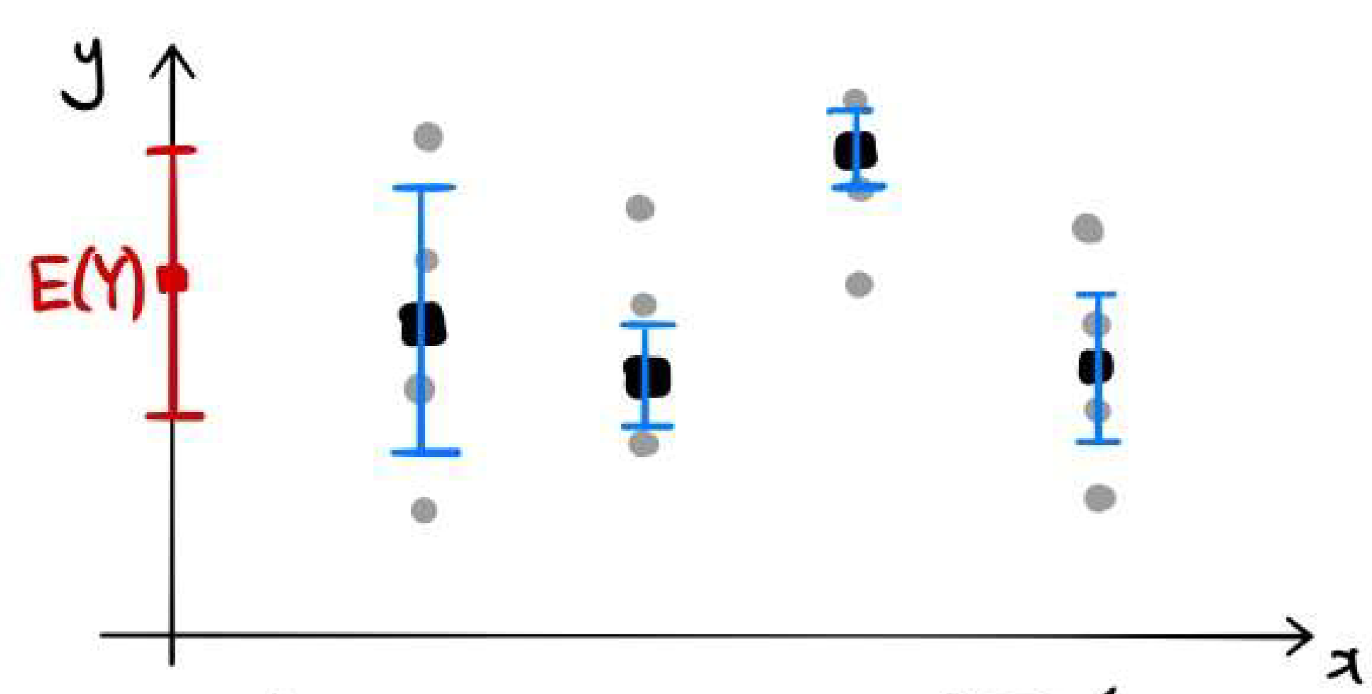
$$Var(Y|X) = Var(Y|X=x) \qquad \text{whenever} \quad X = x.$$

THM 2 (Law of Total Variance)

$$Var(Y) = E[Var(Y|X)] + Var[E(Y|X)]$$

- The proof can be done in a similar spirit.

Figure



A diagram explaining LTV

Var(Y) = [mean of cond. vars I] + [var. of cond. exp.s ■]

Ex 2 Roll a fair die and let X be the outcome. Then flip a biased coin (with H's appearing w/ prob. p) X times and let

$$Y = [\# \text{ of H's}].$$

Then the conditional dist. of Y given X is $b(X, p)$, and so,

$$\begin{aligned} E(Y) &= E(E(Y|X)) \\ &= E(Xp) \\ &= \tfrac{7}{2}p. \qquad (\because\ E(X) = \tfrac{7}{2}) \end{aligned}$$

Moreover,

$$\begin{aligned} Var(Y) &= E(Var(Y|X)) + Var(E(Y|X)) \\ &= E(Xp(1-p)) + Var(Xp) \\ &= \tfrac{7}{2}p(1-p) + \tfrac{35}{12}p^2 \qquad (\because\ Var(X) = \tfrac{35}{12}) \end{aligned}$$

$\square$

THM3 (Pulling out what is known)

$$E[u(X)Y|X] = u(X)E(Y|X).$$

Pf) Given $X = x$, we know $u(X) = u(x)$. So, whenever $X = x$ is observed,

truly a constant !

$$E[u(X)Y \mid X=x] = E[u(x)Y \mid X=x] = u(x)E(Y|X=x). \qquad \square$$

Ex 2 (continued)

$$\begin{aligned} E(XY) &= E[E(XY|X)] = E[XE(Y|X)] \\ &= E[X \cdot Xp] = \tfrac{91}{6}p \qquad (\because\ E(X^2) = \tfrac{91}{6}) \end{aligned}$$