Conditional Distribution Functions and Expectation

(i) Understand probability models conditioned by event or random variable

(ii) Able to compute the conditional distribution functions for one or two random variables

(iii) Able to compute expected values for the conditional distributions

(v) Able to apply conditional probability models to solve problems

Distribution Conditioned by Event

The concept of conditional probability can be extended to the cumulative distribution function (CDF), probability mass function (PMF) and probability density function (PDF) as they are probability measures.

For discrete random variable (RV) X and event A, following (1.7), the conditional PMF of X given A is defined as:

$$P_{X|A}(x) = P(X = x|A) = \begin{cases} \dfrac{P(X = x)}{P(A)}, & x \in A \\ 0, & \text{otherwise} \end{cases} \tag{4.1}$$

The conditional CDF is:

$$F_{X|A}(x) = P(X \leq x|A) = \frac{P(X \leq x \text{ and } A)}{P(A)} \tag{4.2}$$

For continuous RV X, the conditional CDF is same as (4.2) and conditional PDF is:

$$p_{X|A}(x) = \frac{d}{dx} F_{X|A}(x) = \begin{cases} \frac{P(X=x)}{P(A)}, & x \in A \\ 0, & \text{otherwise} \end{cases} \quad (4.3)$$

Example 4.1

A Web distributes instructional videos on bicycle repair. The length of a video is a RV X in min. with PMF:

$$p_X(x) = \begin{cases} 0.15, & x = 1,2,3,4 \\ 0.1, & x = 5,6,7,8 \\ 0, & \text{otherwise} \end{cases}$$

Suppose the Web has two servers, one for videos shorter than 5 min. and the other for videos with at least 5 min. What is the PMF of video length in the second server?

Let A be the event when the second server is chosen, which corresponds to $x \in \{5,6,7,8\}$. Hence we have:

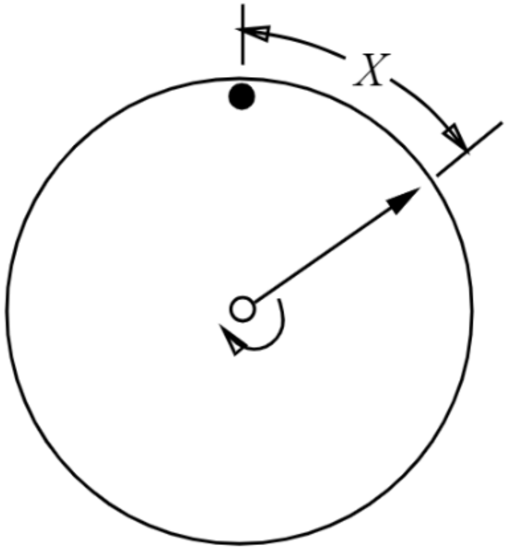
$$P(A) = 0.1 + 0.1 + 0.1 + 0.1 = 0.4$$

According to (4.1),

$$P_{X|A}(x) = \frac{P(X = x \text{ and } A)}{P(A)} = \begin{cases} 0.1/0.4 = 0.25, & x = 5,6,7,8 \\ 0, & \text{otherwise} \end{cases}$$

<u>Example 4.2</u>

Consider a pointer-spinning experiment such that the pointer will stop at a point with uniform distribution on the circumference whose length is 1. Find the conditional PDF of the pointer position for spins in which the pointer stops on the left side of the circle.



Let RV X be the pointer position. It is clear that $X \sim \mathcal{U}(0,1)$ with PDF $p(x) = u(x) - u(x-1)$.

Denote A as the event when the pointer stops on the left side of the circle. Without loss of generality, we may assume that $x \in (0.5, 1)$ corresponds to A. Hence we have:

$$P(A) = \int_{0.5}^{1} p(x)dx = \int_{0.5}^{1} 1dx = 0.5$$

Analogous to (4.1), the conditional PDF is

$$P_{X|A}(x) = \frac{P(X = x \text{ and } A)}{P(A)} = \begin{cases} 1/0.5 = 2, & x \in (0.5, 1) \\ 0, & \text{otherwise} \end{cases}$$

<u>Example 4.3</u>

Suppose the time in integer minutes you wait for a bus is a discrete RV X with PMF:

$$P_X(x) = \begin{cases} 0.05, & x = 1, 2, \cdots, 20 \\ 0, & \text{otherwise} \end{cases}$$

Suppose the bus has not arrived by the 8th minute. What is the conditional PMF of your waiting time X?

Let A be the event $X > 8$, and we get $P(A) = 0.05 \cdot 12 = 0.6$

Hence:

$$P_{X|A}(x) = \begin{cases} 0.05/0.6 = 1/12, & x = 9, 10, \cdots, 20 \\ 0, & \text{otherwise} \end{cases}$$

Analogous to (1.8), law of total probability for PMF/PDF is:

$$P_X(x) = \sum_{n=1}^{N} P_{X|B_n}(x) \cdot P(B_n) \quad (4.4)$$

where events B_n, $n = 1, \cdots, N$, are pairwise disjoint.

Example 4.4

Let X denote the number of additional years that a randomly chosen 70-year-old man will live. If he has abnormal blood pressure, denoted as event A, then X has geometric PMF with $p = 0.1$. If he has normal blood pressure, denoted as event N, then X has geometric PMF with $p = 0.05$. Find the conditional PMFs $P_{X|A}(x)$ and $P_{X|N}(x)$. If 40% of all 70-year-old men have abnormal blood pressure, what is the PMF of X?

According to (2.6), we have 2 conditional PMFs:

$$P_{X|A}(x) = \begin{cases} (0.1)(0.9)^{x-1}, & x = 1, 2, \cdots \\ 0, & \text{otherwise} \end{cases}$$

$$P_{X|N}(x) = \begin{cases} (0.05)(0.95)^{x-1}, & x = 1, 2, \cdots \\ 0, & \text{otherwise} \end{cases}$$

Since $A \cup N = S$ and $A \cap N = \emptyset$, using (4.4) yields:

$$\begin{aligned} P_X(x) &= P_{X|A}(x) \cdot P(A) + P_{X|N}(x) \cdot P(N) \\ &= \begin{cases} (0.4)(0.1)(0.9)^{x-1} + (0.6)(0.05)(0.95)^{x-1}, & x = 1, 2, \cdots \\ 0, & \text{otherwise} \end{cases} \end{aligned}$$

Example 4.5

Let RV X be a received signal. When bits "0" and "1" are transmitted, corresponding to events B_0 and B_1, the received values are $X \sim \mathcal{N}(-5,4)$ and $X \sim \mathcal{N}(5,4)$, respectively. Given that bits "0" and "1" are equally likely to be sent, what is the PDF of X?

It is clear that B_0 and B_1 are pairwise disjoint, and $P(B_0) = P(B_1) = 0.5$. Furthermore, we have:

$$P_{X|B_0}(x) = \frac{1}{2\sqrt{2\pi}} e^{-\frac{1}{8}(x+5)^2}, \quad P_{X|B_1}(x) = \frac{1}{2\sqrt{2\pi}} e^{-\frac{1}{8}(x-5)^2}$$

Applying (4.4) yields:

$$\begin{aligned} P_X(x) &= P_{X|B_0}(x) \cdot P(B_0) + P_{X|B_1}(x) \cdot P(B_1) \\ &= \frac{1}{4\sqrt{2\pi}} \left(e^{-\frac{1}{8}(x+5)^2} + e^{-\frac{1}{8}(x-5)^2} \right) \end{aligned}$$

Conditional Expected Value given Event

Expectation operations can also be extended to the conditional probability functions.

Given that the event A has occurred, the conditional expectation of X can be extended from (2.19) and (2.20) as:

$$\mathbb{E}\{X|A\} = \mu_{X|A} = \sum_{x \in R_x} x P_{X|A}(x) \quad (4.5)$$

$$\mathbb{E}\{X|A\} = \mu_{X|A} = \int_{-\infty}^{\infty} x P_{X|A}(x) dx \quad (4.6)$$

Consider events B_n, $n = 1, \cdots, N$, with $B_1 \cup B_2 \cup \cdots \cup B_N = S$ and $B_m \cap B_n = \emptyset$, $m \neq n$, we have:

$$\mathbb{E}\{X\} = \sum_{n=1}^{N} \mathbb{E}\{X|B_n\} P(B_n) \quad (4.7)$$

Without loss of generality, for discrete RV, the proof is obtained using (4.4) and (4.5):

$$\begin{aligned}\mathbb{E}\{X\} &= \sum_{x\in R_x} x\left(\sum_{n=1}^{N} P_{X|B_n}(x)P(B_n)\right)\\ &= \sum_{n=1}^{N} P(B_n)\left(\sum_{x\in R_x} xP_{X|B_n}(x)\right) = \sum_{n=1}^{N} P(B_n)\mathbb{E}\{X|B_n\}\end{aligned}$$

Example 4.6

Compute $\mathbb{E}\{X\}$ of the geometric RV with parameter p using conditional expectation.

Assign S as the event of getting success in the first trial and recall that the geometric RV is the number of trials until the first success occurs. Applying (4.7) yields:

$$
\begin{aligned}
\mathbb{E}\{X\} &= \mathbb{E}\{X|S\}P(S) + \mathbb{E}\{X|\overline{S}\}P(\overline{S}) \\
&= 1 \cdot p + (\mathbb{E}\{X\} + 1) \cdot (1 - p) \\
&\Rightarrow \mathbb{E}\{X\} = \frac{1}{p}
\end{aligned}
$$

Note that $\mathbb{E}\{X|S\} = 1$ because when the first trial gives a success, X must be 1. While $\mathbb{E}\{X|\overline{S}\}$ contains 1 due to the first failure and $\mathbb{E}\{X\}$ accounts for the expected value starting from the second trial because of independence.

As in (2.24)-(2.25), (4.5)-(4.6) can be generalized to a function of RV X, i.e., $Y = g(X)$:

$$
\mathbb{E}\{Y|A\} = \mathbb{E}\{g(X)|A\} = \sum_{x \in R_x} g(x) P_{X|A}(x) \tag{4.8}
$$

$$
\mathbb{E}\{Y|A\} = \mathbb{E}\{g(X)|A\} = \int_{-\infty}^{\infty} g(x) P_{X|A}(x) dx \tag{4.9}
$$

Analogous to (2.23), the conditional variance of X given A is:

$$\mathrm{var}(X|A) = \mathbb{E}\{(X - \mu_{X|A})^2|A\} = \mathbb{E}\{X^2|A\} - \mu_{X|A}^2 \qquad (4.10)$$

Example 4.7

Find the conditional expected value and conditional variance for the video length in the second server in Example 4.1.

$$\mathbb{E}\{X|A\} = \mu_{X|A} = \sum_{x=5}^{8} x P_{X|A}(x) = 0.25 \sum_{x=5}^{8} x = 6.5 \text{ (min.)}$$

$$\mathbb{E}\{X^2|A\} = 0.25 \sum_{x=5}^{8} x^2 = 43.5 \text{ (min.}^2\text{)}$$

$$\mathrm{var}(X|A) = \mathbb{E}\{X^2|A\} - \mu_{X|A}^2 = 1.25 \text{ (min.}^2\text{)}$$

Conditional Joint Distributions given Event

For RVs X and Y, and an event A, the conditional joint PMF/PDF of X and Y given A is:

$$P_{XY|A}(x,y) = P(X=x, Y=y|A) = \begin{cases} \dfrac{P_{XY}(x,y)}{P(A)}, & (x,y) \in A \\ 0, & \text{otherwise} \end{cases} \quad (4.11)$$

Similarly, for $W = g(X,Y)$, (4.8)-(4.9) can be extended to

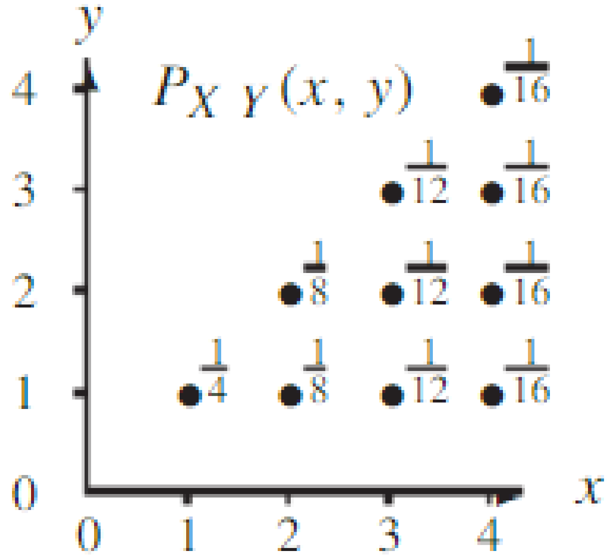
$$\mathbb{E}\{W|A\} = \mathbb{E}\{g(X,Y)|A\} = \sum_{x \in R_x} \sum_{y \in R_y} g(x,y) P_{XY|A}(x,y) \quad (4.12)$$

$$\mathbb{E}\{W|A\} = \mathbb{E}\{g(X,Y)|A\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y) P_{XY|A}(x,y) dx dy \quad (4.13)$$

The conditional variance extends (4.10) straightforwardly.

Example 4.8

Discrete RVs X and Y have the joint PMF $P_{XY}(x,y)$ as shown below. Find the conditional joint PMF $P_{XY|B}(x,y)$ given the event $B \in \{X+Y \leq 4\}$. Then compute the conditional expected value and conditional variance of $W = X+Y$.



It is clear that $B = \{(1,1),(2,1),(2,2),(3,1)\}$. Hence

$$P(B) = p(1,1) + p(2,1) + p(2,2) + (3,1) = \frac{7}{12}$$

$$P_{XY|B}(x,y) = \begin{cases} 3/7, & x=1, y=1 \\ 3/14, & x=2, y=1 \text{ or } y=2 \\ 1/7, & x=3, y=1 \\ 0, & \text{otherwise} \end{cases}$$

Using (4.12), we obtain:

$$\mathbb{E}\{W|B\} = \mu_{W|B} = \sum_{(x,y)\in B} (x+y) P_{XY|B}(x,y)$$

$$= 2\left(\frac{3}{7}\right) + 3\left(\frac{3}{14}\right) + 4\left(\frac{3}{14}\right) + 4\left(\frac{1}{7}\right) = \frac{41}{14}$$

Similarly,

$$\mathbb{E}\{W^2|B\} = \sum_{(x,y)\in B} (x+y)^2 P_{XY|B}(x,y)$$

$$= 2^2\left(\frac{3}{7}\right) + 3^2\left(\frac{3}{14}\right) + 4^2\left(\frac{3}{14}\right) + 4^2\left(\frac{1}{7}\right) = \frac{131}{14}$$

Finally, we have:

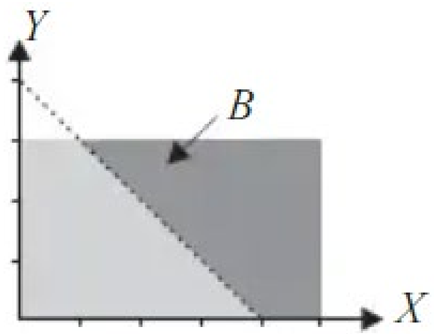
$$\mathrm{var}(W|B) = \mathbb{E}\{W^2|B\} - \mu_{W|B}^2 = \frac{153}{196}$$

<u>Example 4.9</u>

Suppose two RVs X and Y have joint PDF:

$$p(x,y) = \begin{cases} 1/15, & 0 \le x \le 5, 0 \le y \le 3 \\ 0, & \text{otherwise} \end{cases}$$

Find the conditional joint PDF $P_{XY|B}(x,y)$ given the event $B \in \{X + Y \ge 4\}$. Then find the conditional expected value of $W = XY$ given B.



$$P(B) = P(X \geq 4 - Y) = \int_0^3 \int_{4-y}^5 \frac{1}{15} dx dy = \int_0^3 \frac{1+y}{15} dy = \frac{1}{2}$$

Using (4.11), we obtain:

$$P_{XY|B}(x,y) = \begin{cases} \frac{2}{15}, & 0 \leq x \leq 5, 0 \leq y \leq 3, x+y \geq 4 \\ 0, & \text{otherwise} \end{cases}$$

Applying (4.13) yields:

$$\begin{aligned} \mathbb{E}\{XY|B\} &= \int_0^3 \int_{4-y}^5 xy \frac{2}{15} dx dy \\ &= \frac{1}{15} \int_0^3 \left(x^2 \Big|_{4-y}^5 \right) y dy = \frac{1}{15} \int_0^3 (9y + 8y^2 - y^3) dy = \frac{123}{20} \end{aligned}$$

Distribution and Expectation Conditioned by Random Variable

Apart from conditioning by event, probability models can also be conditioned by RV.

The definitions basically follow those of distribution conditioned by event.

The conditional PMF/PDF of RV X given RV Y is:

$$\begin{aligned} P_{X|Y}(x|y) &= P(X=x|Y=y) \\ &= \frac{P(X=x,Y=y)}{P(Y=y)} = \frac{P_{XY}(x,y)}{P_Y(y)} \end{aligned} \quad (4.14)$$

Similarly, the conditional PMF/PDF of Y given X is

$$P_{Y|X}(y|x) = P(Y=y|X=x) = \frac{P_{XY}(x,y)}{P_X(x)} \quad (4.15)$$

From (4.14) and (4.15), we have:

$$P_{XY}(x,y) = P_Y(y)P_{X|Y}(x|y) = P_X(x)P_{Y|X}(y|x) \tag{4.16}$$

When X and Y are independent, (4.14) becomes:

$$P_{X|Y}(x|y) = \frac{P_{XY}(x,y)}{P_Y(y)} = \frac{P_X(x)P_Y(y)}{P_Y(y)} = P_X(x) \tag{4.17}$$

Similar to (4.4), for discrete RVs, we have:

$$P_X(x) = \sum_{y \in R_y} P_{XY}(x,y) = \sum_{y \in R_y} P_{X|Y}(x|y)P_Y(y) \tag{4.18}$$

Or more generally:

$$P_X(X \in B) = \sum_{y \in R_y} P(X \in B|Y = y)P_Y(y) \tag{4.19}$$

Similar to (4.5)-(4.10), the conditional expected values are:

$$\mathbb{E}\{X|Y=y\} = \mu_{X|Y}(y) = \sum_{x \in R_x} x P_{X|Y}(x|y) \tag{4.20}$$

$$\mathbb{E}\{X|Y=y\} = \mu_{X|Y}(y) = \int_{-\infty}^{\infty} x P_{X|Y}(x|y) dx \tag{4.21}$$

$$\mathbb{E}\{X\} = \sum_{y \in R_y} \mathbb{E}\{X|Y=y\} P_Y(y) \tag{4.22}$$

$$\mathbb{E}\{g(X,Y)|Y=y\} = \sum_{x \in R_x} g(x,y) P_{X|Y}(x|y) \tag{4.23}$$

$$\mathbb{E}\{g(X,Y)|Y=y\} = \int_{-\infty}^{\infty} g(x,y) P_{X|Y}(x|y) dx \tag{4.24}$$

$$\mathrm{var}(X|Y=y) = \mathbb{E}\{X^2|Y=y\} - (\mu_{X|Y}(y))^2 \tag{4.25}$$

Note that conditioning for continuous RVs is treated in a similar manner, e.g.,

$$P(X \in B|Y = y) = \int_B P_{X|Y}(x|y)dx$$

$$P_X(x) = \int_{-\infty}^{\infty} P_{X|Y}(x|y)P_Y(y)dy$$

Example 4.10

Given the joint PDF of two RVs X and Y:

$$P_{XY}(x,y) = ce^{-(x+y)}, \quad 0 \le y \le x,\ 0 \le x < \infty$$

Determine the value of c and $P_{Y|X}(y|x)$.

Applying (3.8), we have

$$\begin{aligned} P_X(x) &= \int_{-\infty}^{\infty} P_{XY}(x,y)dy = \int_0^x ce^{-(x+y)}dy = ce^{-x}\int_0^x e^{-y}dy \\ &= ce^{-x} \cdot -e^{-y}\Big|_0^x = ce^{-x}(1-e^{-x}) = c(e^{-x}-e^{-2x}) \end{aligned}$$

To find c, we utilize:

$$\int_{-\infty}^{\infty} P_X(x)dx = 1 \Rightarrow \int_0^{\infty} c(e^{-x} - e^{-2x})dx = 1 \Rightarrow c\left[-e^{-x} + \frac{e^{-2x}}{2}\right]\Bigg|_0^{\infty} = 1$$

$$\Rightarrow c\left[1 - \frac{1}{2}\right] \Rightarrow c = 2$$

Hence:

$$P_X(x) = 2(e^{-x} - e^{-2x})$$

Finally, applying (4.15) yields:

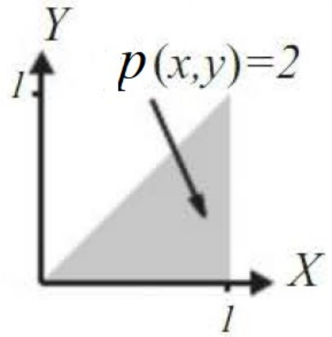
$$P_{Y|X}(y|x) = \frac{P_{XY}(x,y)}{P_X(x)} = \frac{2e^{-(x+y)}}{2(e^{-x} - e^{-2x})} = \frac{e^{-(x+y)}}{e^{-x} - e^{-2x}} = \frac{e^{-y}}{1 - e^{-x}}$$

for $0 \le y \le x,\ 0 \le x < \infty$.

Example 4.11

Suppose the joint PDF of RVs X and Y is:

$$p(x,y)=\begin{cases}2, & 0\le y\le x\le 1\\ 0, & \text{otherwise}\end{cases}$$



Find the conditional PDF $P_{Y|X}(y|x)$ for $0\le x\le 1$ and conditional PDF $P_{X|Y}(x|y)$ for $0\le y\le 1$. Then find the conditional expected value $\mathbb{E}\{X|Y=y\}$.

For $0\le x\le 1$, we compute:

$$P_X(x)=\int_{-\infty}^{\infty}P_{XY}(x,y)dy=\int_0^x 2dy=2x$$

Applying (4.15) yields:

$$P_{Y|X}(y|x) = \frac{P_{XY}(x,y)}{P_X(x)} = \begin{cases} \frac{1}{x}, & 0 \le y \le x \le 1 \\ 0, & \text{otherwise} \end{cases}$$

That is, for a given $X = x$, $Y \sim \mathcal{U}(0, x)$. Similarly, for $0 \le y \le 1$:

$$P_Y(y) = \int_{-\infty}^{\infty} P_{XY}(x,y)dx = \int_y^1 2dx = 2(1-y)$$

$$\Rightarrow P_{X|Y}(x|y) = \frac{P_{XY}(x,y)}{P_Y(y)} = \begin{cases} \frac{1}{1-y}, & 0 \le y \le x \le 1 \\ 0, & \text{otherwise} \end{cases}$$

That is, for a given $Y = y$, $X \sim \mathcal{U}(y, 1)$.

Then we use (4.21):

$$\begin{aligned}\mathbb{E}\{X|Y=y\} &= \int_{-\infty}^{\infty} xP_{X|Y}(x|y)dx \\ &= \int_{y}^{1} \frac{1}{1-y} x dx = \frac{x^2}{2(1-y)}\bigg|_{x=y}^{x=1} = \frac{1+y}{2}\end{aligned}$$

Note that $\mathbb{E}\{X|Y=y\}$ is a function of y. If y is not specified, we can write:

$$\mathbb{E}\{X|Y\} = \frac{1+Y}{2}$$

which is also a RV, namely, a function of Y.

As $\mathbb{E}\{X|Y\}$ is a function of Y, the expected value of $\mathbb{E}\{X|Y\}$ applies on Y only, and the result is:

$$\mathbb{E}\{\mathbb{E}\{X|Y\}\} = \mathbb{E}\{X\} \tag{4.26}$$

This two-step process is known as iterated expectation. Assuming continuous RVs, the proof can be derived as:

$$\begin{aligned}
\mathbb{E}\{\mathbb{E}\{X|Y\}\} &= \int_{-\infty}^{\infty} \mathbb{E}\{X|Y=y\} P_Y(y)dy \\
&= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} x P_{X|Y}(x|y)dx\right] P_Y(y)dy \\
&= \int_{-\infty}^{\infty} x \left[\int_{-\infty}^{\infty} P_{X|Y}(x|y)P_Y(y)dy\right] dx \\
&= \int_{-\infty}^{\infty} x P_X(x)dx = \mathbb{E}\{X\}
\end{aligned}$$

Iterated expectation can be generalized to $g(X)$:

$$\mathbb{E}\{g(X)\} = \mathbb{E}\{\mathbb{E}\{g(X)|Y\}\} = \sum_{y \in R_y} \mathbb{E}\{g(X)|Y = y\} P_Y(y)$$

$$\mathbb{E}\{g(X)\} = \mathbb{E}\{\mathbb{E}\{g(X)|Y\}\} = \int_{-\infty}^{\infty} \mathbb{E}\{g(X)|Y = y\} P_Y(y) dy$$

<u>Example 4.12</u>

Apply iterated expectation to find $\mathbb{E}\{X\}$ in Example 4.11.

Recalling the result of $\mathbb{E}\{X|Y\}$, we apply (4.26) to obtain:

$$\mathbb{E}\{X\} = \mathbb{E}\{\mathbb{E}\{X|Y\}\} = \frac{1 + \mathbb{E}\{Y\}}{2} = 1$$

because

$$\mathbb{E}\{Y\} = \int_{-\infty}^{\infty} y P_Y(y) dy = \int_{0}^{1} y \cdot 2(1 - y) dy = 1$$

This aligns with

$$\mathbb{E}\{X\} = \int_{-\infty}^{\infty} x P_X(x) dx = \int_0^1 x \cdot 2x dy = 1$$

Example 4.13

Consider tossing a coin twice and the probability that the outcome is head (H) is p. Let X_1 and X_2 be the number of H, either 0 or 1, in the first and second trials. Assign $W = X_1 - X_2$ and $Y = X_1 + X_2$. Find $P_{WY}(w, y)$, $P_{W|Y}(w|y)$ and $P_{Y|W}(y|w)$.

Denote tail as T. We list out the possible outcomes first.

Outcome	Probability	W	Y
HH	p^2	0	2
HT	$p(1-p)$	1	1
TH	$p(1-p)$	-1	1
TT	$(1-p)^2$	0	0

We can then construct the joint PMF of W and Y:

$P_{WY}(w,y)$	$w=-1$	$w=0$	$w=1$	$P_Y(y)$
$y=0$	0	$(1-p)^2$	0	$(1-p)^2$
$y=1$	$p(1-p)$	0	$p(1-p)$	$2p(1-p)$
$y=2$	0	p^2	0	p^2
$P_W(w)$	$p(1-p)$	$1-2p+2p^2$	$p(1-p)$	

Using (4.14), we have:

$$P_{W|Y}(w|0) = \begin{cases} 1, & w=0 \\ 0, & \text{otherwise} \end{cases}$$

$$P_{W|Y}(w|1) = \begin{cases} 0.5, & w=-1,1 \\ 0, & \text{otherwise} \end{cases}$$

$$P_{W|Y}(w|2) = \begin{cases} 1, & w=0 \\ 0, & \text{otherwise} \end{cases}$$

$$P_{Y|W}(y|-1) = \begin{cases} 1, & y = 1 \\ 0, & \text{otherwise} \end{cases}$$

$$P_{Y|W}(y|0) = \begin{cases} \dfrac{(1-p)^2}{1-2p+2p^2}, & y = 0 \\ \dfrac{p^2}{1-2p+2p^2}, & y = 2 \\ 0, & \text{otherwise} \end{cases}$$

$$P_{Y|W}(y|1) = \begin{cases} 1, & y = 1 \\ 0, & \text{otherwise} \end{cases}$$

References:

1. R. D. Yates and D. J. Goodman, *Probability and Stochastic Processes: A Friendly Introduction for Electrical and Computer Engineer*, Wiley, 2014
2. S. Kay, *Intuitive Probability and Random Processes using MATLAB*, Springer, 2006
3. H. Pishro-Nik, *Introduction to Probability, Statistics, and Random Processes*, Kappa Research LLC, 2014 https://www.probabilitycourse.com