Section 7.1

Conditioning a Random Variable by an Event

Example 7.1

Let N equal the number of bytes in an email. A conditioning event might be the event I that the email contains an image. A second kind of conditioning would be the event $\{N > 100{,}000\}$, which tells us that the email required more than 100,000 bytes. Both events I and $\{N > 100{,}000\}$ give us information that the email is likely to have many bytes.

Definition 7.1 Conditional CDF

Given the event B *with* $\mathsf{P}[B] > 0$, *the* conditional cumulative distribution function *of* X *is*

$$F_{X|B}(x) = \mathsf{P}\left[X \leq x|B\right].$$

Definition 7.2 Conditional PMF Given an Event

Given the event B with $\mathsf{P}[B] > 0$, *the* conditional probability mass function *of X is*

$$P_{X|B}(x) = \mathsf{P}\left[X = x|B\right].$$

Definition 7.3 Conditional PDF Given an Event

For a random variable X *and an event* B *with* $\mathrm{P}[B] > 0$, *the* conditional PDF of X given B *is*

$$f_{X|B}(x) = \frac{dF_{X|B}(x)}{dx}.$$

Theorem 7.1

For a random variable X and an event $B \subset S_X$ with $\mathrm{P}[B] > 0$, the conditional PDF of X given B is

$$\text{Discrete:} \qquad P_{X|B}(x) = \begin{cases} \dfrac{P_X(x)}{\mathrm{P}[B]} & x \in B, \\ 0 & \text{otherwise} \end{cases}$$

$$\text{Continuous:} \quad f_{X|B}(x) = \begin{cases} \dfrac{f_X(x)}{\mathrm{P}[B]} & x \in B, \\ 0 & \text{otherwise.} \end{cases}$$

Example 7.2 Problem

A website distributes instructional videos on bicycle repair. The length of a video in minutes X has PMF

$$P_X(x) = \begin{cases} 0.15 & x = 1,2,3,4, \\ 0.1 & x = 5,6,7,8, \\ 0 & \text{otherwise.} \end{cases} \tag{7.3}$$

Suppose the website has two servers, one for videos shorter than five minutes and the other for videos of five or more minutes. What is the PMF of video length in the second server?

Example 7.2 Solution

We seek a conditional PMF for the condition $x \in L = \{5, 6, 7, 8\}$. From Theorem 7.1,

$$P_{X|L}(x) = \begin{cases} \dfrac{P_X(x)}{\mathrm{P}[L]} & x = 5, 6, 7, 8, \\ 0 & \text{otherwise.} \end{cases} \tag{7.4}$$

From the definition of L, we have

$$\mathrm{P}[L] = \sum_{x=5}^{8} P_X(x) = 0.4. \tag{7.5}$$

With $P_X(x) = 0.1$ for $x \in L$,

$$P_{X|L}(x) = \begin{cases} 0.1/0.4 = 0.25 & x = 5, 6, 7, 8, \\ 0 & \text{otherwise.} \end{cases} \tag{7.6}$$

Thus the lengths of long videos are equally likely. Among the long videos, each length has probability 0.25.

Example 7.3 Problem

For the pointer-spinning experiment of Example 4.1, find the conditional PDF of the pointer position for spins in which the pointer stops on the left side of the circle.

Example 7.3 Solution

Let L denote the left side of the circle. In terms of the stopping position, $L = [1/2, 1)$. Recalling from Example 4.4 that the pointer position X has a uniform PDF over $[0, 1)$,

$$\mathrm{P}[L] = \int_{1/2}^{1} f_X(x)\, dx = \int_{1/2}^{1} dx = 1/2. \tag{7.7}$$

Therefore,

$$f_{X|L}(x) = \begin{cases} 2 & 1/2 \le x < 1, \\ 0 & \text{otherwise.} \end{cases} \tag{7.8}$$

Example 7.4 Problem

Suppose X, the time in integer minutes you wait for a bus, has the discrete uniform PMF

$$P_X(x) = \begin{cases} 1/20 & x = 1, 2, \ldots, 20, \\ 0 & \text{otherwise.} \end{cases} \tag{7.9}$$

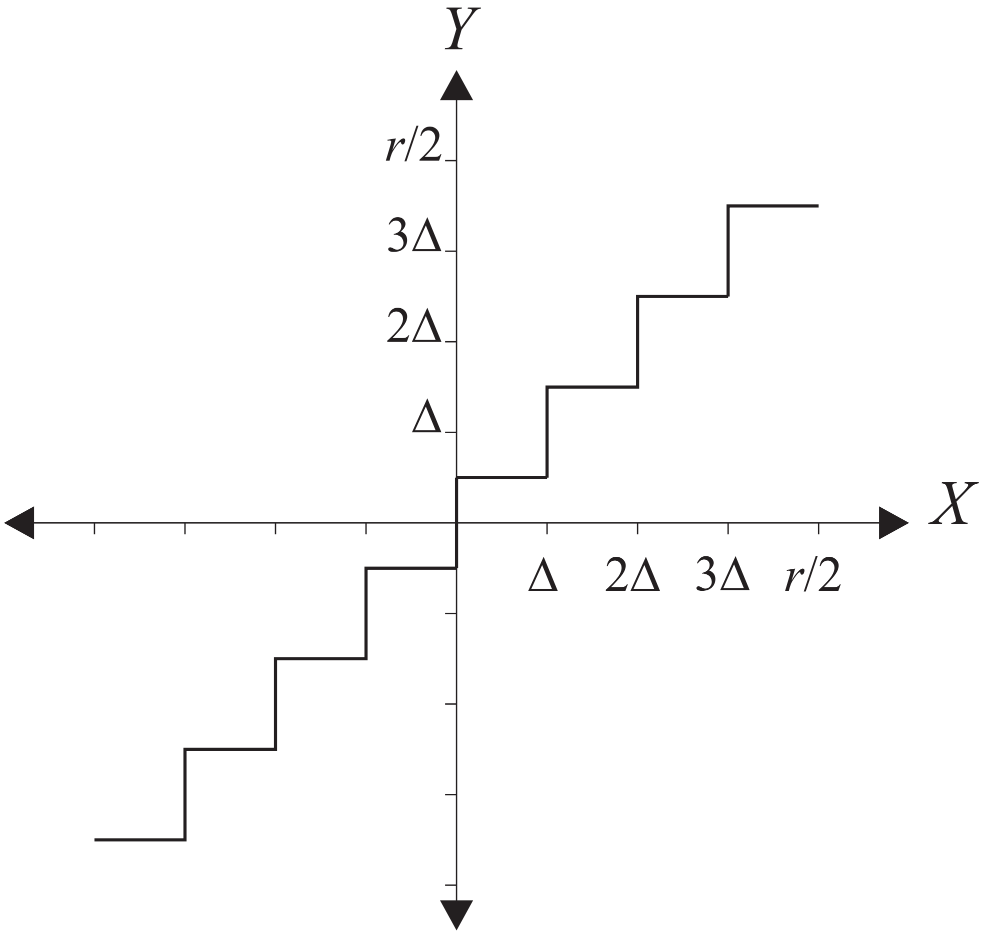
Suppose the bus has not arrived by the eighth minute; what is the conditional PMF of your waiting time X?

Example 7.4 Solution

Let A denote the event $X > 8$. Observing that $\text{P}[A] = 12/20$, we can write the conditional PMF of X as

$$P_{X|X>8}(x) = \begin{cases} \dfrac{1/20}{12/20} = \dfrac{1}{12} & x = 9, 10, \ldots, 20, \\ 0 & \text{otherwise.} \end{cases} \tag{7.10}$$

Example 7.5 Problem



The continuous uniform $(-r/2, r/2)$ random variable X is processed by a b-bit uniform quantizer to produce the quantized output Y. Random variable X is rounded to the nearest quantizer level. With a b-bit quantizer, there are $n = 2^b$ quantization levels. The quantization step size is $\Delta = r/n$, and Y takes on values in the set

$$Q_Y = \left\{y_{-n/2}, y_{-n/2+1}, \dots, y_{n/2-1}\right\} \tag{7.11}$$

where $y_i = \Delta/2 + i\Delta$. This relationship is shown for $b = 3$ in the figure on the left. Given the event B_i that $Y = y_i$, find the conditional PDF of X given B_i.

Example 7.5 Solution

In terms of X, we observe that $B_i = \{i\Delta \leq X < (i+1)\Delta\}$. Thus,

$$\mathrm{P}[B_i] = \int_{i\Delta}^{(i+1)\Delta} f_X(x)\, dx = \frac{\Delta}{r} = \frac{1}{n}. \tag{7.12}$$

By Definition 7.3,

$$f_{X|B_i}(x) = \begin{cases} \dfrac{f_X(x)}{\mathrm{P}[B_i]} & x \in B_i, \\ 0 & \text{otherwise,} \end{cases} = \begin{cases} 1/\Delta & i\Delta \leq x < (i+1)\Delta, \\ 0 & \text{otherwise.} \end{cases} \tag{7.13}$$

Given B_i, the conditional PDF of X is uniform over the ith quantization interval.

Theorem 7.2

For random variable X resulting from an experiment with partition

$$B_1, \ldots, B_m,$$

$$\text{Discrete:} \qquad P_X(x) = \sum_{i=1}^{m} P_{X|B_i}(x)\,\mathsf{P}[B_i]$$

$$\text{Continuous:} \quad f_X(x) = \sum_{i=1}^{m} f_{X|B_i}(x)\,\mathsf{P}[B_i]$$

Proof: Theorem 7.2

The theorem follows directly from Theorem 1.9 with $A = \{X = x\}$ for discrete X or $A = \{x < X \leq x + dx\}$ when X is continuous.

Example 7.6 Problem

Let X denote the number of additional years that a randomly chosen 70-year-old person will live. If the person has high blood pressure, denoted as event H, then X is a geometric ($p = 0.1$) random variable. Otherwise, if the person's blood pressure is normal, event N, X has a geometric ($p = 0.05$) PMF. Find the conditional PMFs $P_{X|H}(x)$ and $P_{X|N}(x)$. If 40 percent of all 70-year-olds have high blood pressure, what is the PMF of X?

Example 7.6 Solution

The problem statement specifies the conditional PMFs in words. Mathematically, the two conditional PMFs are

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

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

Since H, N is a partition, we can use Theorem 7.2 to write

$$\begin{aligned} P_X(x) &= P_{X|H}(x)\,\mathrm{P}[H] + P_{X|N}(x)\,\mathrm{P}[N] \\ &= \begin{cases} (0.4)(0.1)(0.9)^{x-1} + (0.6)(0.05)(0.95)^{x-1} & x = 1, 2, \ldots, \\ 0 & \text{otherwise}. \end{cases} \end{aligned} \tag{7.14}$$

Example 7.7 Problem

Random variable X is a voltage at the receiver of a modem. When symbol "0" is transmitted (event B_0), X is the Gaussian $(-5, 2)$ random variable. When symbol "1" is transmitted (event B_1), X is the Gaussian $(5, 2)$ random variable. Given that symbols "0" and "1" are equally likely to be sent, what is the PDF of X?

Example 7.7 Solution

The problem statement implies that $\mathsf{P}[B_0] = \mathsf{P}[B_1] = 1/2$ and

$$f_{X|B_0}(x) = \frac{1}{2\sqrt{2\pi}} e^{-(x+5)^2/8}, \qquad f_{X|B_1}(x) = \frac{1}{2\sqrt{2\pi}} e^{-(x-5)^2/8}. \tag{7.15}$$

By Theorem 7.2,

$$\begin{aligned} f_X(x) &= f_{X|B_0}(x)\,\mathsf{P}[B_0] + f_{X|B_1}(x)\,\mathsf{P}[B_1] \\ &= \frac{1}{4\sqrt{2\pi}} \left(e^{-(x+5)^2/8} + e^{-(x-5)^2/8} \right). \end{aligned} \tag{7.16}$$

Problem 7.7.1 asks the reader to graph $f_X(x)$ to show its similarity to Figure 4.3.

Quiz 7.1(A)

On the Internet, data is transmitted in packets. In a simple model for World Wide Web traffic, the number of packets N needed to transmit a Web page depends on whether the page has graphic images. If the page has images (event I), then N is uniformly distributed between 1 and 50 packets. If the page is just text (event T), then N is uniform between 1 and 5 packets. Assuming a page has images with probability 1/4, find the

(a) conditional PMF $P_{N|I}(n)$

(b) conditional PMF $P_{N|T}(n)$

(c) PMF $P_N(n)$

(d) conditional PMF $P_{N|N\leq 10}(n)$

Quiz 7.1(A) Solution

(a) From the problem statement, we learn that the conditional PMF of N given the event I is

$$P_{N|I}(n) = \begin{cases} 0.02 & n = 1, 2, \ldots, 50, \\ 0 & \text{otherwise.} \end{cases}$$

(b) Also from the problem statement, the conditional PMF of N given the event T is

$$P_{N|T}(n) = \begin{cases} 0.2 & n = 1, \ldots, 5, \\ 0 & \text{otherwise.} \end{cases}$$

(c) The problem statement tells us that $\text{P}[T] = 1 - \text{P}[I] = 3/4$. From Theorem 7.2, we find the PMF of N is

$$P_N(n) = P_{N|T}(n)\,\text{P}[T] + P_{N|I}(n)\,\text{P}[I] = \begin{cases} 0.155 & n = 1, \ldots, 5, \\ 0.005 & n = 6, \ldots, 50, \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

[Continued]

(d) First we find

$$\mathrm{P}[N \le 10] = \sum_{n=1}^{10} P_N(n) = (0.155)(5) + (0.005)(5) = 0.80. \tag{2}$$

By Theorem 7.1, the conditional PMF of N given $N \le 10$ is

$$P_{N|N\le 10}(n) = \begin{cases} \frac{P_N(n)}{\mathrm{P}[N\le 10]} & n \le 10, \\ 0 & \text{otherwise,} \end{cases}$$

$$= \begin{cases} \frac{0.155}{0.8} = 0.19375 & n = 1, \ldots, 5, \\ \frac{0.005}{0.8} = 0.00625 & n = 6, \ldots, 10, \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

Quiz 7.1(B)

Y is a continuous uniform $(0, 10)$ random variable. Find the following:

(a) $\mathrm{P}[Y \leq 6]$

(b) the conditional PDF $f_{Y|Y\leq 6}(y)$

(c) $\mathrm{P}[Y > 8]$

(d) the conditional PDF $f_{Y|Y>8}(y)$

Quiz 7.1(B) Solution

From the problem statement,

$$f_Y(y) = \begin{cases} 1/10 & 0 < y < 10, \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

Using this PDF and Definition 7.3, the parts are straightforward.

(a) $\mathrm{P}[Y \leq 6] = \int_{-\infty}^{6} f_Y(y)\,dy = \int_0^6 (1/10)\,dy = 0.6\,.$

(b) From Definition 7.3, the conditional PDF of Y given $Y \leq 6$ is

$$\begin{aligned} f_{Y|Y\leq 6}(y) &= \begin{cases} \frac{f_Y(y)}{\mathrm{P}[Y\leq 6]} & y \leq 6, \\ 0 & \text{otherwise,} \end{cases} \\ &= \begin{cases} 1/6 & 0 \leq y \leq 6, \\ 0 & \text{otherwise.} \end{cases} \end{aligned} \tag{2}$$

(c) The probability $Y > 8$ is

$$\mathrm{P}\left[Y > 8\right] = \int_8^{10} \frac{1}{10}\,dy = 0.2. \tag{3}$$

(d) From Definition 7.3, the conditional PDF of Y given $Y > 8$ is

$$\begin{aligned} f_{Y|Y>8}(y) &= \begin{cases} \frac{f_Y(y)}{\mathrm{P}[Y>8]} & y > 8, \\ 0 & \text{otherwise,} \end{cases} \\ &= \begin{cases} \frac{1}{2} & 8 < y \leq 10, \\ 0 & \text{otherwise.} \end{cases} \end{aligned} \tag{4}$$

Conditional Expected Value Given an Event

Theorem 7.3

Discrete X:

(a) For any $x \in B$, $P_{X|B}(x) \geq 0$.

(b) $\sum_{x \in B} P_{X|B}(x) = 1$.

(c) The conditional probability that X is in the set C is

$$\mathrm{P}[C|B] = \sum_{x \in C} P_{X|B}(x).$$

Continuous X:

(a) For any $x \in B$, $f_{X|B}(x) \geq 0$.

(b) $\int_B f_{X|B}(x)\, dx = 1$.

(c) The conditional probability that X is in the set C is

$$\mathrm{P}[C|B] = \int_C f_{X|B}(x)\, dx.$$

Definition 7.4 Conditional Expected Value

The conditional expected value *of random variable X given condition B is*

$$\textit{Discrete:} \qquad \mathsf{E}[X|B] = \sum_{x \in B} x P_{X|B}(x)$$

$$\textit{Continuous:} \quad \mathsf{E}[X|B] = \int_{-\infty}^{\infty} x f_{X|B}(x)\, dx.$$

Theorem 7.4

For a random variable X resulting from an experiment with partition $B_1, \ldots, B_m$,

$$\mathsf{E}[X] = \sum_{i=1}^{m} \mathsf{E}[X|B_i]\,\mathsf{P}[B_i].$$

Proof: Theorem 7.4

When X is discrete, $\mathsf{E}[X] = \sum_x x P_X(x)$, and we can use Theorem 7.2 to write

$$\begin{aligned}\mathsf{E}[X] &= \sum_x x \sum_{i=1}^{m} P_{X|B_i}(x)\,\mathsf{P}[B_i] \\ &= \sum_{i=1}^{m} \mathsf{P}[B_i] \sum_x x P_{X|B_i}(x) = \sum_{i=1}^{m} \mathsf{P}[B_i]\,\mathsf{E}[X|B_i]\,. \end{aligned} \tag{7.17}$$

When X is continuous, the proof uses the continuous version of Theorem 7.2 and follows the same logic, with the summation over x replaced by integration.

Theorem 7.5

The conditional expected value of $Y = g(X)$ given condition B is

Discrete: $$\mathsf{E}[Y|B] = \mathsf{E}[g(X)|B] = \sum_{x \in B} g(x) P_{X|B}(x)$$

Continuous: $$\mathsf{E}[Y|B] = \mathsf{E}[g(X)|B] = \int_{-\infty}^{\infty} g(x) f_{X|B}(x)\, dx.$$

Definition 7.5 Conditional Variance and Standard Deviation

The conditional variance of X given event B is

$$\text{Var}[X|B] = \text{E}\left[\left(X - \mu_{X|B}\right)^2 |B\right] = \text{E}\left[X^2|B\right] - \mu_{X|B}^2.$$

The conditional standard deviation is $\sigma_{X|B} = \sqrt{\text{Var}[X|B]}$.

Example 7.8 Problem

Find the conditional expected value, the conditional variance, and the conditional standard deviation for the long videos defined in Example 7.2.

Example 7.8 Solution

$$\mathrm{E}[X|L] = \mu_{X|L} = \sum_{x=5}^{8} x P_{X|L}(x) = 0.25 \sum_{x=5}^{8} x = 6.5 \text{ minutes}. \tag{7.18}$$

$$\mathrm{E}\left[X^2|L\right] = 0.25 \sum_{x=5}^{8} x^2 = 43.5 \text{ minutes}^2. \tag{7.19}$$

$$\mathrm{Var}[X|L] = \mathrm{E}\left[X^2|L\right] - \mu_{X|L}^2 = 1.25 \text{ minutes}^2. \tag{7.20}$$

$$\sigma_{X|L} = \sqrt{\mathrm{Var}[X|L]} = 1.12 \text{ minutes}. \tag{7.21}$$

Quiz 7.2(A)

Continuing Quiz 7.1(A), find

(a) $\mathrm{E}[N|N \leq 10]$,

(b) $\mathrm{Var}[N|N \leq 10]$.

Quiz 7.2(A) Solution

We refer to the solution of Quiz 7.1(A) for $P_{N|N\le 10}(n)$.

(a) Given $P_{N|N\le 10}(n)$, calculating a conditional expected value is the same as for any other expected value except we use the conditional PMF.

$$\begin{aligned} \mathrm{E}[N|N \le 10] &= \sum_n n P_{N|N\le 10}(n) \\ &= \sum_{n=1}^{5} 0.19375n + \sum_{n=6}^{10} 0.00625n = 3.15625. \end{aligned} \tag{1}$$

(b) For the conditional variance, we first find the conditional second moment

$$\begin{aligned} \mathrm{E}\left[N^2|N \le 10\right] &= \sum_n n^2 P_{N|N\le 10}(n) \\ &= \sum_{n=1}^{5} 0.19375n^2 + \sum_{n=6}^{10} 0.00625n^2 \\ &= 0.19375(55) + 0.00625(330) = 12.719. \end{aligned} \tag{2}$$

The conditional variance is

$$\begin{aligned} \mathrm{Var}[N|N \le 10] &= \mathrm{E}\left[N^2|N \le 10\right] - (\mathrm{E}[N|N \le 10])^2 \\ &= 12.719 - (3.156)^2 = 2.757. \end{aligned} \tag{3}$$

Quiz 7.2(B)

Continuing Quiz 7.1(B), find

(a) $\mathrm{E}[Y|Y \leq 6]$,

(b) $\mathrm{Var}[Y|Y \leq 6]$.

Quiz 7.2(B) Solution

We refer to the solution of Quiz 7.1(B) for the conditional PDFs $f_{Y|Y\leq 6}(y)$ and $f_{Y|Y>8}(y)$.

(a) From $f_{Y|Y\leq 6}(y)$, the conditional expectation is

$$\mathrm{E}[Y|Y \leq 6] = \int_{-\infty}^{\infty} y f_{Y|Y\leq 6}(y)\, dy = \int_0^6 \frac{y}{6}\, dy = 3. \tag{1}$$

(b) From the conditional PDF $f_{Y|Y>8}(y)$, we see that given $Y > 8$, Y is conditionally a continuous uniform $(a = 8, b = 10)$ random variable. Thus,

$$\mathrm{Var}[Y|Y > 8] = (b-a)^2/12 = 1/3. \tag{2}$$

Conditioning Two Random Variables by an Event

Definition 7.6 Conditional Joint PMF

For discrete random variables X and Y and an event B with $\mathsf{P}[B] > 0$, *the* conditional joint PMF *of X and Y given B is*

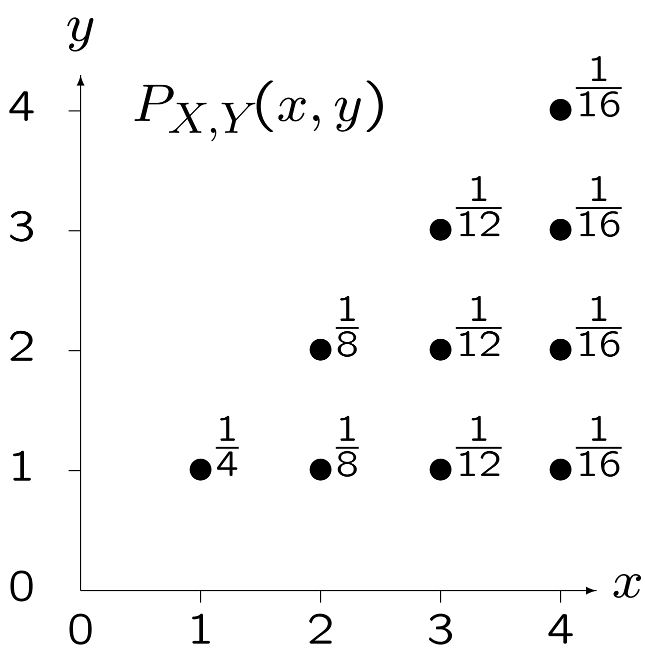
$$P_{X,Y|B}(x, y) = \mathsf{P}\left[X = x, Y = y|B\right].$$

Theorem 7.6

For any event B, a region of the X, Y plane with $\mathrm{P}[B] > 0$,

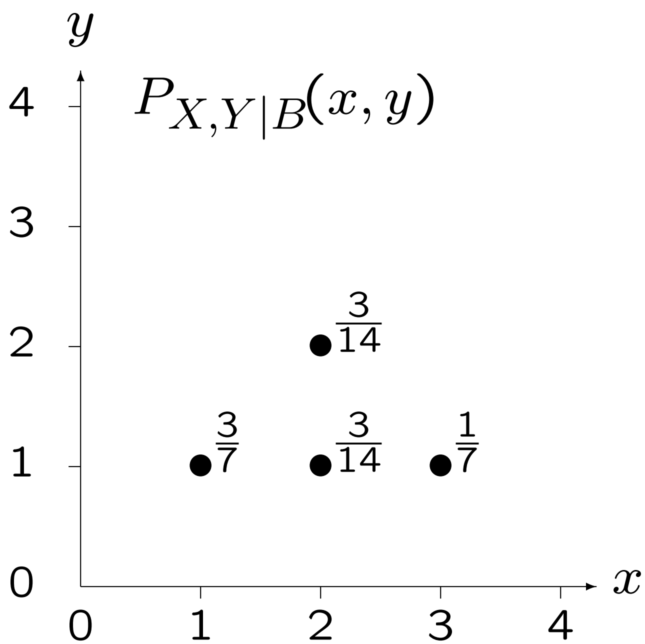
$$P_{X,Y|B}(x,y) = \begin{cases} \dfrac{P_{X,Y}(x,y)}{\mathrm{P}[B]} & (x,y) \in B, \\ 0 & \text{otherwise.} \end{cases}$$

Example 7.9 Problem



Random variables X and Y have the joint PMF $P_{X,Y}(x,y)$ as shown. Let $B = \{X + Y \leq 4\}$ and find the conditional PMF $P_{X,Y|B}(x,y)$.

Example 7.9 Solution



Event $B = \{(1,1),(2,1),(2,2),(3,1)\}$ consists of all points (x,y) such that $x+y \leq 4$. By adding up the probabilities of all outcomes in B, we find

$$\mathrm{P}[B] = P_{X,Y}(1,1) + P_{X,Y}(2,1) + P_{X,Y}(2,2) + P_{X,Y}(3,1) = \frac{7}{12}.$$

The conditional PMF $P_{X,Y|B}(x,y)$ is shown on the left.

Definition 7.7 Conditional Joint PDF

Given an event B with $\mathsf{P}[B] > 0$, the conditional joint probability density function of X and Y is

$$f_{X,Y|B}(x,y) = \begin{cases} \dfrac{f_{X,Y}(x,y)}{\mathsf{P}[B]} & (x,y) \in B, \\ 0 & \textit{otherwise.} \end{cases}$$

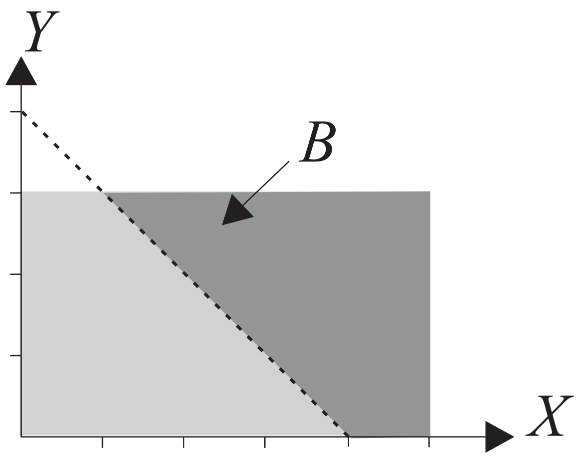
Example 7.10 Problem

X and Y are random variables with joint PDF

$$f_{X,Y}(x,y) = \begin{cases} 1/15 & 0 \leq x \leq 5, 0 \leq y \leq 3, \\ 0 & \text{otherwise.} \end{cases} \tag{7.22}$$

Find the conditional PDF of X and Y given the event $B = \{X + Y \geq 4\}$.

Example 7.10 Solution



We calculate $P[B]$ by integrating $f_{X,Y}(x,y)$ over the region B.

$$\begin{aligned} P[B] &= \int_0^3 \int_{4-y}^5 \frac{1}{15}\,dx\,dy \\ &= \frac{1}{15}\int_0^3 (1+y)\,dy = 1/2. \end{aligned} \tag{7.23}$$

Definition 7.7 leads to the conditional joint PDF

$$f_{X,Y|B}(x,y) = \begin{cases} 2/15 & 0 \le x \le 5, 0 \le y \le 3, x+y \ge 4, \\ 0 & \text{otherwise.} \end{cases} \tag{7.24}$$

Theorem 7.7 Conditional Expected Value

For random variables X and Y and an event B of nonzero probability, the conditional expected value of $W = g(X,Y)$ given B is

Discrete: $$\mathrm{E}[W|B] = \sum_{x \in S_X} \sum_{y \in S_Y} g(x,y) P_{X,Y|B}(x,y)$$

Continuous: $$\mathrm{E}[W|B] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y) f_{X,Y|B}(x,y)\, dx\, dy.$$

Example 7.11 Problem

Continuing Example 7.9, find the conditional expected value and the conditional variance of $W = X + Y$ given the event $B = \{X + Y \leq 4\}$.

Example 7.11 Solution

We recall from Example 7.9 that $P_{X,Y|B}(x,y)$ has four points with nonzero probability: $(1,1)$, $(1,2)$, $(1,3)$, and $(2,2)$. Their probabilities are 3/7, 3/14, 1/7, and 3/14, respectively. Therefore,

$$\begin{aligned}
\mathsf{E}[W|B] &= \sum_{x,y} (x+y) P_{X,Y|B}(x,y) \\
&= 2\left(\frac{3}{7}\right) + 3\left(\frac{3}{14}\right) + 4\left(\frac{1}{7}\right) + 4\left(\frac{3}{14}\right) = \frac{41}{14}. \qquad (7.25)
\end{aligned}$$

Similarly,

$$\begin{aligned}
\mathsf{E}\left[W^2|B\right] &= \sum_{x,y} (x+y)^2 P_{X,Y|B}(x,y) \\
&= 2^2\left(\frac{3}{7}\right) + 3^2\left(\frac{3}{14}\right) + 4^2\left(\frac{1}{7}\right) + 4^2\left(\frac{3}{14}\right) = \frac{131}{14}. \qquad (7.26)
\end{aligned}$$

The conditional variance is

$$\mathrm{Var}[W|B] = \mathsf{E}\left[W^2|B\right] - (\mathsf{E}[W|B])^2 = \frac{131}{14} - \left(\frac{41}{14}\right)^2 = \frac{153}{196}. \qquad (7.27)$$

Quiz 7.3(A)

Random variables L and X have joint PMF

$P_{L,X}(l,x)$	$x = 40$	$x = 60$
$l = 1$	0.15	0.1
$l = 2$	0.3	0.2
$l = 3$	0.15	0.1

(7.28)

For random variable $V = LX$, we define the event $A = \{V > 80\}$. Find the conditional PMF $P_{L,X|A}(l,x)$. What are $E[V|A]$ and $\text{Var}[V|A]$?

Quiz 7.3(A) Solution

Since the event $V > 80$ occurs only for the pairs $(L, X) = (2, 60)$, $(L, X) = (3, 40)$ and $(L, X) = (3, 60)$,

$$\begin{aligned} \mathrm{P}[A] = \mathrm{P}[V > 80] &= P_{L,X}(2, 60) + P_{L,X}(3, 40) + P_{L,X}(3, 60) \\ &= 0.45. \end{aligned} \tag{1}$$

By Definition 7.6,

$$P_{L,X|A}(l, X) = \begin{cases} \frac{P_{L,X}(l,x)}{\mathrm{P}[A]} & lx > 80, \\ 0 & \text{otherwise.} \end{cases}$$

We can represent this conditional PMF in the following table:

$P_{L,X\|A}(l,x)$	$x = 40$	$x = 60$
$l = 1$	0	0
$l = 2$	0	4/9
$l = 3$	1/3	2/9

[Continued]

The conditional expectation of V can be found from the conditional PMF.

$$\begin{aligned} \mathrm{E}[V|A] &= \sum_l \sum_x lx P_{L,X|A}(l,x) \\ &= (120)\frac{4}{9} + (120)\frac{1}{3} + (180)\frac{2}{9} = 133\ \frac{1}{3}. \end{aligned} \tag{2}$$

For the conditional variance $\mathrm{Var}[V|A]$, we first find the conditional second moment

$$\begin{aligned} \mathrm{E}\left[V^2|A\right] &= \sum_l \sum_x (lx)^2 P_{L,X|A}(l,x) \\ &= (120)^2\frac{4}{9} + (120)^2\frac{1}{3} + (180)^2\frac{2}{9} = 18,400. \end{aligned} \tag{3}$$

It follows that

$$\mathrm{Var}[V|A] = \mathrm{E}\left[V^2|A\right] - (\mathrm{E}[V|A])^2 = 622\,\frac{2}{9} \tag{4}$$

Quiz 7.3(B)

Random variables X and Y have the joint PDF

$$f_{X,Y}(x,y) = \begin{cases} xy/4000 & 1 \le x \le 3, 40 \le y \le 60, \\ 0 & \text{otherwise.} \end{cases} \tag{7.29}$$

For random variable $W = XY$, we define the event $B = \{W > 80\}$. Find the conditional joint PDF $f_{X,Y|B}(x,y)$. What are $E[W|B]$ and $Var[W|B]$?

Quiz 7.3(B) Solution

For continuous random variables X and Y, we first calculate the probability of the conditioning event.

$$\mathrm{P}[B] = \iint_B f_{X,Y}(x,y)\,dx\,dy = \int_{40}^{60}\int_{80/y}^{3} \frac{xy}{4000}\,dx\,dy. \tag{1}$$

A little calculus yields

$$\begin{aligned}\mathrm{P}[B] &= \int_{40}^{60} \frac{y}{4000}\left(\left.\frac{x^2}{2}\right|_{80/y}^{3}\right) dy \\ &= \int_{40}^{60} \frac{y}{4000}\left(\frac{9}{2} - \frac{3200}{y^2}\right) dy = \frac{9}{8} - \frac{4}{5}\ln\frac{3}{2}.\end{aligned} \tag{2}$$

In fact, $\mathrm{P}[B] \approx 0.801$. The conditional PDF of X and Y is

$$f_{X,Y|B}(x,y) = \begin{cases} \frac{f_{X,Y}(x,y)}{\mathrm{P}[B]} & (x,y) \in B, \\ 0 & \text{otherwise}, \end{cases} = \begin{cases} Kxy & \substack{40 \le y \le 60,\\ 80/y \le x \le 3,} \\ 0 & \text{otherwise}. \end{cases} \tag{3}$$

where $K = (4000\,\mathrm{P}[B])^{-1}$. The conditional expectation of W given event B is[Continued]

Quiz 7.3(B) Solution (Continued 2)

$$\mathsf{E}[W|B] = \iint xy f_{X,Y|B}(x,y)\,dx\,dy = \int_{40}^{60}\int_{80/y}^{3} Kx^2y^2\,dx\,dy. \tag{4}$$

These next steps are just calculus:

$$\begin{aligned}\mathsf{E}[W|B] &= \frac{K}{3}\int_{40}^{60} y^2x^3\Big|_{x=80/y}^{x=3}\,dy \\ &= \frac{K}{3}\int_{40}^{60}\left(27y^2 - 80^3/y\right)\,dy = \frac{K}{3}\left(9y^3 - 80^3\ln y\right)\Big|_{40}^{60} \approx 120.78.\end{aligned} \tag{5}$$

The conditional second moment of K given B is

$$\mathsf{E}\left[W^2|B\right] = \iint (xy)^2 f_{X,Y|B}(x,y)\,dx\,dy = \int_{40}^{60}\int_{80/y}^{3} Kx^3y^3\,dx\,dy. \tag{6}$$

With a final bit of calculus,

$$\begin{aligned}\mathsf{E}\left[W^2|B\right] &= \frac{K}{4}\int_{40}^{60} y^3x^4\Big|_{x=80/y}^{x=3}\,dy \\ &= \frac{K}{4}\int_{40}^{60}\left(81y^3 - 80^4/y\right)\,dy = \frac{K}{4}\left(\frac{81}{4}y^4 - 80^4\ln y\right)\Bigg|_{40}^{60} \\ &\approx 16{,}116.10.\end{aligned} \tag{7}$$

It follows that $\text{Var}[W|B] = \mathsf{E}[W^2|B] - (\mathsf{E}[W|B])^2 \approx 1528.30$.

Conditioning by a Random Variable

Definition 7.8 Conditional PMF

For any event $Y = y$ *such that* $P_Y(y) > 0$*, the* conditional PMF *of* X *given* $Y = y$ *is*

$$P_{X|Y}(x|y) = \mathrm{P}\left[X = x|Y = y\right].$$

Theorem 7.8

For discrete random variables X and Y with joint PMF $P_{X,Y}(x, y)$, and x and y such that $P_X(x) > 0$ and $P_Y(y) > 0$,

$$P_{X|Y}(x|y) = \frac{P_{X,Y}(x, y)}{P_Y(y)}, \qquad P_{Y|X}(y|x) = \frac{P_{X,Y}(x, y)}{P_X(x)}.$$

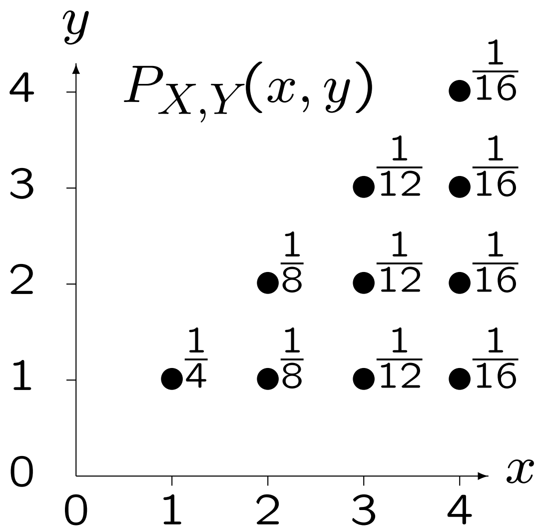
Proof: Theorem 7.8

Referring to Definition 7.8, Definition 1.5, and Theorem 5.4, we observe that

$$P_{X|Y}(x|y) = \mathsf{P}[X = x|Y = y] = \frac{\mathsf{P}[X = x, Y = y]}{\mathsf{P}[Y = y]} = \frac{P_{X,Y}(x, y)}{P_Y(y)}. \quad (7.30)$$

The proof of the second part is the same with X and Y reversed.

Example 7.12 Problem



Random variables X and Y have the joint PMF $P_{X,Y}(x,y)$, as given in Example 7.9 and repeated in the accompanying graph. Find the conditional PMF of Y given $X = x$ for each $x \in S_X$.

Example 7.12 Solution

To apply Theorem 7.8, we first find the marginal PMF $P_X(x)$. By Theorem 5.4, $P_X(x) = \sum_{y \in S_Y} P_{X,Y}(x,y)$. For a given $X = x$, we sum the nonzero probablities along the vertical line $X = x$. That is,

$$P_X(x) = \begin{cases} 1/4 & x = 1, \\ 1/8 + 1/8 & x = 2, \\ 1/12 + 1/12 + 1/12 & x = 3, \\ 1/16 + 1/16 + 1/16 + 1/16 & x = 4, \\ 0 & \text{otherwise}, \end{cases} = \begin{cases} 1/4 & x = 1, \\ 1/4 & x = 2, \\ 1/4 & x = 3, \\ 1/4 & x = 4, \\ 0 & \text{otherwise}. \end{cases} \tag{7.31}$$

Theorem 7.8 implies that for $x \in \{1, 2, 3, 4\}$,

$$P_{Y|X}(y|x) = \frac{P_{X,Y}(x,y)}{P_X(x)} = 4P_{X,Y}(x,y). \tag{7.32}$$

[Continued]

Example 7.12 Solution (Continued 2)

For each $x \in \{1,2,3,4\}$, $P_{Y|X}(y|x)$ is a different PMF.

$$P_{Y|X}(y|1) = \begin{cases} 1 & y = 1, \\ 0 & \text{otherwise.} \end{cases} \qquad P_{Y|X}(y|2) = \begin{cases} 1/2 & y \in \{1,2\}, \\ 0 & \text{otherwise.} \end{cases}$$

$$P_{Y|X}(y|3) = \begin{cases} 1/3 & y \in \{1,2,3\}, \\ 0 & \text{otherwise.} \end{cases} \qquad P_{Y|X}(y|4) = \begin{cases} 1/4 & y \in \{1,2,3,4\}, \\ 0 & \text{otherwise.} \end{cases}$$

Given $X = x$, Y is conditionally a discrete uniform $(1, x)$ random variable.

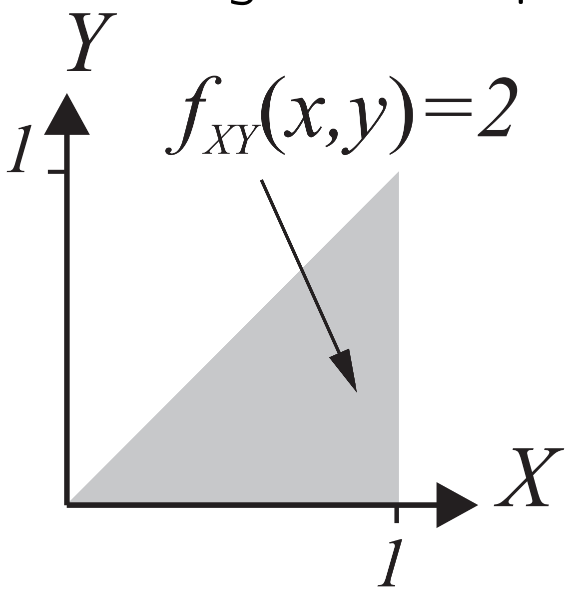
Definition 7.9 Conditional PDF

For y such that $f_Y(y) > 0$, the conditional PDF of X given $\{Y = y\}$ is

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

Example 7.13 Problem

Returning to Example 5.8, random variables X and Y have joint PDF



$$f_{X,Y}(x,y) = \begin{cases} 2 & 0 \le y \le x \le 1, \\ 0 & \text{otherwise.} \end{cases} \tag{7.34}$$

For $0 \le x \le 1$, find the conditional PDF $f_{Y|X}(y|x)$. For $0 \le y \le 1$, find the conditional PDF $f_{X|Y}(x|y)$.

Example 7.13 Solution

For $0 \leq x \leq 1$, Theorem 5.8 implies

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\ dy = \int_0^x 2\,dy = 2x. \tag{7.35}$$

The conditional PDF of Y given X is

$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \begin{cases} 1/x & 0 \leq y \leq x \leq 1, \\ 0 & \text{otherwise.} \end{cases} \tag{7.36}$$

Given $X = x$, we see that Y is the uniform $(0, x)$ random variable. For $0 \leq y \leq 1$, Theorem 5.8 implies

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\ dx = \int_y^1 2\,dx = 2(1-y). \tag{7.37}$$

Furthermore, Equation (7.33) implies

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \begin{cases} 1/(1-y) & y \leq x \leq 1, \\ 0 & \text{otherwise.} \end{cases} \tag{7.38}$$

Conditioned on $Y = y$, we see that X is the uniform $(y, 1)$ random variable.

Theorem 7.9

For discrete random variables X and Y with joint PMF $P_{X,Y}(x, y)$, and x and y such that $P_X(x) > 0$ and $P_Y(y) > 0$,

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

Theorem 7.10

For continuous random variables X and Y with joint PDF $f_{X,Y}(x,y)$, and x and y such that $f_X(x) > 0$ and $f_Y(y) > 0$,

$$f_{X,Y}(x,y) = f_{Y|X}(y|x)\, f_X(x) = f_{X|Y}(x|y)\, f_Y(y)\,.$$

Example 7.14 Problem

Let R be the uniform $(0, 1)$ random variable. Given $R = r$, X is the uniform $(0, r)$ random variable. Find the conditional PDF of R given X.

Example 7.14 Solution

The problem definition states that

$$f_R(r) = \begin{cases} 1 & 0 \leq r < 1, \\ 0 & \text{otherwise,} \end{cases} \qquad f_{X|R}(x|r) = \begin{cases} 1/r & 0 \leq x < r, \\ 0 & \text{otherwise.} \end{cases} \tag{7.39}$$

It follows from Theorem 7.10 that the joint PDF of R and X is

$$f_{R,X}(r,x) = f_{X|R}(x|r)\, f_R(r) = \begin{cases} 1/r & 0 \leq x < r < 1, \\ 0 & \text{otherwise.} \end{cases} \tag{7.40}$$

Now we can find the marginal PDF of X from Theorem 5.8. For $0 < x < 1$,

$$f_X(x) = \int_{-\infty}^{\infty} f_{R,X}(r,x)\; dr = \int_x^1 \frac{dr}{r} = -\ln x. \tag{7.41}$$

By the definition of the conditional PDF,

$$f_{R|X}(r|x) = \frac{f_{R,X}(r,x)}{f_X(x)} = \begin{cases} \frac{1}{-r \ln x} & x \leq r \leq 1, \\ 0 & \text{otherwise.} \end{cases} \tag{7.42}$$

Definition 5.4

- We recall from Definition 5.4 that discrete random variables X and Y are independent if $P_{X,Y}(x,y) = P_X(x)P_Y(y)$.

- Similarly, continuous X and Y are independent if $f_{X,Y}(x,y) = f_X(x)f_Y(y)$.

- We apply these observations to Theorems 7.9 and 7.10.

Theorem 7.11

If X and Y are independent,

Discrete:	$P_{X\|Y}(x\|y) = P_X(x)$,	$P_{Y\|X}(y\|x) = P_Y(y)$;
Continuous:	$f_{X\|Y}(x\|y) = f_X(x)$,	$f_{Y\|X}(y\|x) = f_Y(y)$.

Quiz 7.4(A)

The PMF of random variable X satisfies $P_X(0) = 1 - P_X(2) = 0.4$. The conditional probability model for random variable Y given X is

$$P_{Y|X}(y|0) = \begin{cases} 0.8 & y = 0, \\ 0.2 & y = 1, \\ 0 & \text{otherwise}, \end{cases} \qquad P_{Y|X}(y|2) = \begin{cases} 0.5 & y = 0, \\ 0.5 & y = 1, \\ 0 & \text{otherwise}. \end{cases} \tag{7.43}$$

(a) What is the probability model for X and Y? Write the joint PMF $P_{X,Y}(x,y)$ as a table.

(b) If $Y = 0$, what is the conditional PMF $P_{X|Y}(x|0)$?

Quiz 7.4(A) Solution

(a) The joint PMF of X and Y can be found from the marginal and conditional PMFs via $P_{X,Y}(x,y) = P_{Y|X}(y|x)P_X(x)$. Incorporating the information from the given conditional PMFs can be confusing, however. Consequently, we note that X has range $S_X = \{0,2\}$ and Y has range $S_Y = \{0,1\}$. A table of the joint PMF will include all four possible combinations of X and Y. The general form of the table is

$P_{X,Y}(x,y)$	$y=0$	$y=1$
$x=0$	$P_{Y\|X}(0\|0)P_X(0)$	$P_{Y\|X}(1\|0)P_X(0)$
$x=2$	$P_{Y\|X}(0\|2)P_X(2)$	$P_{Y\|X}(1\|2)P_X(2)$

Substituting values from $P_{Y|X}(y|x)$ and $P_X(x)$, we have

$P_{X,Y}(x,y)$	$y=0$	$y=1$
$x=0$	(0.8)(0.4)	(0.2)(0.4)
$x=2$	(0.5)(0.6)	(0.5)(0.6)

,

which simplifies to

$P_{X,Y}(x,y)$	$y=0$	$y=1$
$x=0$	0.32	0.08
$x=2$	0.3	0.3

.

(b) From the joint PMF $P_{X,Y}(x,y)$, we can calculate $P_Y(0) = 0.32 + 0.3 = 0.62$ and the conditional PMF

$$P_{X|Y}(x|0) = \frac{P_{X,Y}(x,0)}{P_Y(0)} = \begin{cases} \frac{0.32}{0.62} = \frac{16}{31} & x = 0, \\ \frac{0.3}{0.62} = \frac{15}{31} & x = 2, \\ 0 & \text{otherwise.} \end{cases}$$

Quiz 7.4(B)

The PDF of Y given X are

$$f_X(x) = \begin{cases} 3x^2 & 0 < x \leq 1, \\ 0 & \text{otherwise}, \end{cases} \qquad f_{Y|X}(y|x) = \begin{cases} 2y/x^2 & 0 \leq y \leq x, \\ 0 & \text{otherwise}. \end{cases} \tag{7.44}$$

(a) What is the probability model for X and Y? Find $f_{X,Y}(x,y)$.

(b) If $Y = 1/2$, what is the conditional PDF $f_{X|Y}(x|1/2)$?

Quiz 7.4(B) Solution

(a) The joint PDF of X and Y is

$$f_{X,Y}(x,y) = f_{Y|X}(y|x)\, f_X(x) = \begin{cases} 6y & 0 \le y \le x, 0 \le x \le 1, \\ 0 & \text{otherwise.} \end{cases}$$

(b) To find $f_{X|Y}(x|1/2)$, we first find

$$f_Y(1/2) = \int_{-\infty}^{\infty} f_{X,Y}(x, 1/2)\, dx.$$

For this integral, we keep in mind that $f_{X,Y}(x,y)$ is nonzero for $y \le x \le 1$, Specifically, for $y = 1/2$, we integrate over $1/2 \le x \le 1$:

$$f_Y(1/2) = \int_{1/2}^{1} 6(1/2)\, dx = 3/2. \tag{1}$$

For $1/2 \le x \le 1$, the conditional PDF of X given $Y = 1/2$ is

$$f_{X|Y}(x|1/2) = \frac{f_{X,Y}(x,1/2)}{f_Y(1/2)} = \frac{6(1/2)}{3/2} = 2. \tag{2}$$

For $x < 1/2$ or $x > 1$, $f_{X|Y}(x|1/2) = 0$. Thus given $Y = 1/2$, the X has the continuous uniform $(1/2, 1)$ PDF

$$f_{X|Y}(x|1/2) = \begin{cases} 2 & \frac{1}{2} \le x \le 1, \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

Conditional Expected Value Given a Random Variable

Conditional Expected Value

Definition 7.10 of a Function

For any $y \in S_Y$, *the conditional expected value of* $g(X,Y)$ *given* $Y = y$ *is*

Discrete:
$$\mathrm{E}[g(X,Y)|Y=y] = \sum_{x \in S_X} g(x,y) P_{X|Y}(x|y)$$

Continuous:
$$\mathrm{E}[g(X,Y)|Y=y] = \int_{-\infty}^{\infty} g(x,y) f_{X|Y}(x|y)\,dx.$$

Example 7.15 Problem

In Example 7.12, we derived conditional PMFs $P_{Y|X}(y|1)$, $P_{Y|X}(y|2)$, $P_{Y|X}(y|3)$, and $P_{Y|X}(y|4)$. Find $\mathrm{E}[Y|X = x]$ for $x = 1, 2, 3, 4$.

Example 7.15 Solution

In Example 7.12 we found that given $X = x$, Y was a discrete uniform $(1, x)$ random variable. Since a discrete uniform $(1, x)$ random variable has expected value $(1 + x)/2$,

$$\mathrm{E}[Y|X = 1] = \frac{1+1}{2} = 1, \qquad \mathrm{E}[Y|X = 2] = \frac{1+2}{2} = 1.5, \tag{7.45}$$

$$\mathrm{E}[Y|X = 3] = \frac{1+3}{2} = 2, \qquad \mathrm{E}[Y|X = 4] = \frac{1+4}{2} = 2.5. \tag{7.46}$$

Theorem 7.12

For independent random variables X and Y,

(a) $\mathsf{E}[X|Y = y] = \mathsf{E}[X]$ for all $y \in S_Y$,

(b) $\mathsf{E}[Y|X = x] = \mathsf{E}[Y]$ for all $x \in S_X$.

Proof: Theorem 7.12

We present the proof for discrete random variables. By replacing PMFs and sums with PDFs and integrals, we arrive at essentially the same proof for continuous random variables. Since $P_{X|Y}(x|y) = P_X(x)$,

$$\mathrm{E}\left[X|Y=y\right] = \sum_{x \in S_X} x P_{X|Y}(x|y) = \sum_{x \in S_X} x P_X(x) = \mathrm{E}\left[X\right]. \tag{7.47}$$

Since $P_{Y|X}(y|x) = P_Y(y)$,

$$\mathrm{E}\left[Y|X=x\right] = \sum_{y \in S_Y} y P_{Y|X}(y|x) = \sum_{y \in S_Y} y P_Y(y) = \mathrm{E}\left[Y\right]. \tag{7.48}$$

Conditional Expected Values

- When we introduced the concept of expected value in Chapters 3 and 4, we observed that $\mathsf{E}[X]$ is a property of the probability model of X. This is also true for $\mathsf{E}[X|B]$ when $\mathsf{P}[B] > 0$.

- The situation is more complex when we consider $\mathsf{E}[X|Y = y]$, the conditional expected value given a random variable.

- In this case, the conditional expected value is a different number for each possible observation $y \in S_Y$.

- This implies that $\mathsf{E}[X|Y = y]$ is a function of the random variable Y.

- We use the notation $\mathsf{E}[X|Y]$ to denote this function of the random variable Y.

- Since a function of a random variable is another random variable, we conclude that $\mathsf{E}[X|Y]$ *is a random variable!*

Definition 7.11 Function

The conditional expected value $\mathrm{E}[X|Y]$ *is a function of random variable* Y *such that if* $Y = y$*, then* $\mathrm{E}[X|Y] = \mathrm{E}[X|Y = y]$.

Example 7.16 Problem

For random variables X and Y in Example 5.8, we found in Example 7.13 that the conditional PDF of X given Y is

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \begin{cases} 1/(1-y) & 0 \leq y \leq x \leq 1, \\ 0 & \text{otherwise.} \end{cases} \qquad (7.49)$$

Find the conditional expected values $\text{E}[X|Y=y]$ and $\text{E}[X|Y]$.

Example 7.16 Solution

Given the conditional PDF $f_{X|Y}(x|y)$, we perform the integration

$$\begin{aligned} \mathsf{E}\left[X|Y=y\right] &= \int_{-\infty}^{\infty} x f_{X|Y}\left(x|y\right)\, dx \\ &= \int_{y}^{1} \frac{1}{1-y} x\, dx = \left.\frac{x^2}{2(1-y)}\right|_{x=y}^{x=1} = \frac{1+y}{2}. \end{aligned} \tag{7.50}$$

Since $\mathsf{E}[X|Y=y] = (1+y)/2$, $\mathsf{E}[X|Y] = (1+Y)/2$.

Iterated Expectation Property

- An interesting property of the random variable E$[X|Y]$ is its expected value E[E$[X|Y]$].

- We find E[E$[X|Y]$] in two steps: First we calculate $g(y) =$ E$[X|Y = y]$, and then we apply Theorem 4.4 to evaluate E$[g(Y)]$.

- This two-step process is known as *iterated expectation*.

Theorem 7.13 Iterated Expectation

$$\mathsf{E}\left[\mathsf{E}\left[X|Y\right]\right] = \mathsf{E}\left[X\right].$$

Proof: Theorem 7.13

We consider continuous random variables X and Y and apply Theorem 4.4:

$$\mathsf{E}\left[\mathsf{E}\left[X|Y\right]\right] = \int_{-\infty}^{\infty} \mathsf{E}\left[X|Y=y\right] f_Y(y)\; dy. \tag{7.51}$$

To obtain this formula from Theorem 4.4, we have used $\mathsf{E}[X|Y=y]$ in place of $g(x)$ and $f_Y(y)$ in place of $f_X(x)$. Next, we substitute the right side of Equation (7.45) for $\mathsf{E}[X|Y=y]$:

$$\mathsf{E}\left[\mathsf{E}\left[X|Y\right]\right] = \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} x f_{X|Y}(x|y)\; dx \right) f_Y(y)\; dy. \tag{7.52}$$

Rearranging terms in the double integral and reversing the order of integration, we obtain

$$\mathsf{E}\left[\mathsf{E}\left[X|Y\right]\right] = \int_{-\infty}^{\infty} x \int_{-\infty}^{\infty} f_{X|Y}(x|y)\, f_Y(y)\; dy\, dx. \tag{7.53}$$

Next, we apply Theorem 7.10 and Theorem 5.8 to infer that the inner integral is $f_X(x)$. Therefore,

$$\mathsf{E}\left[\mathsf{E}\left[X|Y\right]\right] = \int_{-\infty}^{\infty} x f_X(x)\; dx. \tag{7.54}$$

The proof is complete because the right side of this formula is the definition of $\mathsf{E}[X]$. A similar derivation (using sums instead of integrals) proves the theorem for discrete random variables.

Theorem 7.14

$$\mathsf{E}\left[\mathsf{E}\left[g(X)|Y\right]\right] = \mathsf{E}\left[g(X)\right].$$

Quiz 7.5(A)

For random variables A and B in Quiz 7.4(A) find:

(a) $\mathrm{E}[Y|X=2]$,

(b) $\mathrm{Var}[X|Y=0]$.

Quiz 7.5(A) Solution

(a) Given the conditional PMF $P_{Y|X}(y|2)$, it is easy to calculate the conditional expectation

$$\mathrm{E}[Y|X=2] = \sum_{y=0}^{1} y P_{Y|X}(y|2) = (0)(0.5) + (1)(0.5) = 0.5. \tag{1}$$

(b) We can calculate the conditional variance $\mathrm{Var}[X|Y=0]$ using the conditional PMF $P_{X|Y}(x|0)$. First we calculate the conditional expected value

$$\mathrm{E}[X|Y=0] = \sum_{x} x P_{X|Y}(x|0) = 0 \cdot \frac{16}{31} + 2 \cdot \frac{15}{31} = \frac{30}{31}. \tag{2}$$

The conditional second moment is

$$\mathrm{E}\left[X^2|Y=0\right] = \sum_{x} x^2 P_{X|Y}(x|0) = 0^2 \frac{16}{31} + 2^2 \frac{15}{31} = \frac{60}{31}. \tag{3}$$

The conditional variance is then

$$\mathrm{Var}[X|Y=0] = \mathrm{E}\left[X^2|Y=0\right] - (\mathrm{E}[X|Y=0])^2 = 960/961. \tag{4}$$

Quiz 7.5(B)

For random variables X and Y in Quiz 7.4(B), find:

(a) $\mathrm{E}[Y|X = 1/2]$,

(b) $\mathrm{Var}[X|Y = 1/2]$.

Quiz 7.5(B) Solution

(a) From the conditional PDF $f_{Y|X}(y|x)$ given in Quiz 7.4(B),

$$f_{Y|X}(y|1/2) = \begin{cases} 8y & 0 \leq y \leq 1/2, \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

Now we calculate the conditional expected value

$$\mathrm{E}[Y|X = 1/2] = \int_0^{1/2} y(8y)\,dy = 8y^3/3\Big|_0^{1/2} = 1/3. \tag{2}$$

(b) From the solution to Quiz 7.4(B), we see that given $Y = 1/2$, the conditional PDF of X is uniform $(1/2, 1)$. Thus, by the definition of the uniform (a, b) PDF,

$$\mathrm{Var}[X|Y = 1/2] = \frac{(1 - 1/2)^2}{12} = \frac{1}{48}.$$

Bivariate Gaussian Random Variables: Conditional PDFs

Theorem 7.15

If X and Y are the bivariate Gaussian random variables in Definition 5.10, the conditional PDF of Y given X is

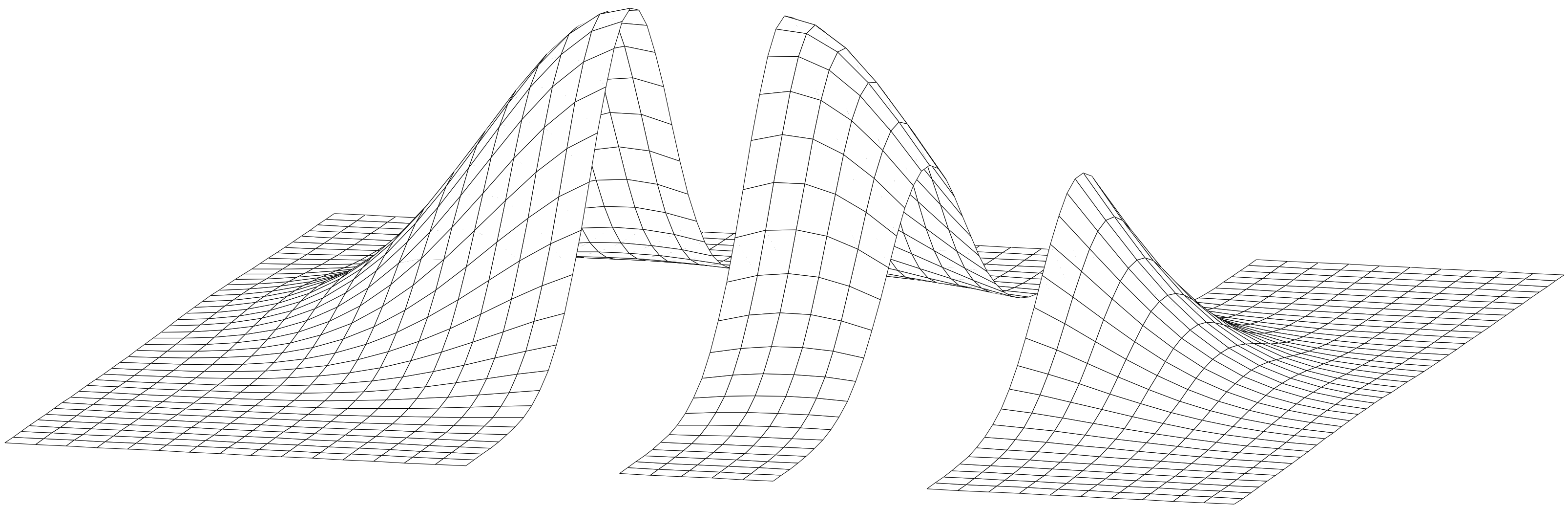
$$f_{Y|X}(y|x) = \frac{1}{\tilde{\sigma}_Y \sqrt{2\pi}} e^{-(y-\tilde{\mu}_Y(x))^2/2\tilde{\sigma}_Y^2},$$

where, given $X = x$, the conditional expected value and variance of Y are

$$\mathrm{E}\left[Y|X = x\right] = \tilde{\mu}_Y(x) = \mu_Y + \rho_{X,Y}\frac{\sigma_Y}{\sigma_X}(x - \mu_X),$$

$$\mathrm{Var}[Y|X = x] = \tilde{\sigma}_Y^2 = \sigma_Y^2(1 - \rho_{X,Y}^2).$$

Figure 7.1



Cross-sectional view of the joint Gaussian PDF with $\mu_X = \mu_Y = 0$, $\sigma_X = \sigma_Y = 1$, and $\rho_{X,Y} = 0.9$. Theorem 7.15 confirms that the bell shape of the cross section occurs because the conditional PDF $f_{Y|X}(y|x)$ is Gaussian.

Theorem 7.16

If X and Y are the bivariate Gaussian random variables in Definition 5.10, the conditional PDF of X given Y is

$$f_{X|Y}(x|y) = \frac{1}{\tilde{\sigma}_X\sqrt{2\pi}} e^{-(x-\tilde{\mu}_X(y))^2/2\tilde{\sigma}_X^2},$$

where, given $Y = y$, the conditional expected value and variance of X are

$$\begin{aligned} \mathrm{E}[X|Y = y] &= \tilde{\mu}_X(y) = \mu_X + \rho_{X,Y}\frac{\sigma_X}{\sigma_Y}(y - \mu_Y), \\ \mathrm{Var}[X|Y = y] &= \tilde{\sigma}_X^2 = \sigma_X^2(1 - \rho^2). \end{aligned}$$

Proof: Theorem 5.19

(Theorem 5.19) We define $g(X,Y) = (X-\mu_X)(Y-\mu_Y)/(\sigma_X\sigma_Y)$. From Definition 5.5 and Definition 5.6, we have the following formula for the correlation coefficient of any pair of random variables X and Y:

$$\mathsf{E}\left[g(X,Y)\right] = \frac{\mathsf{E}\left[(X-\mu_X)(Y-\mu_Y)\right]}{\sigma_X\sigma_Y}. \tag{7.57}$$

We will now show that $\mathsf{E}[g(X,Y)] = \rho_{X,Y}$ for bivariate Gaussian random variables X and Y. Using the substitution $f_{X,Y}(x,y) = f_{Y|X}(y|x)f_X(x)$ to evaluate the double integral in the numerator, we obtain

$$\begin{aligned}\mathsf{E}\left[g(X,Y)\right] &= \frac{1}{\sigma_X\sigma_Y}\int_{-\infty}^{\infty}(x-\mu_X)\left(\int_{-\infty}^{\infty}(y-\mu_Y)\,f_{Y|X}(y|x)\;dy\right)f_X(x)\;dx\\ &= \frac{1}{\sigma_X\sigma_Y}\int_{-\infty}^{\infty}(x-\mu_X)\,\mathsf{E}\left[Y-\mu_Y|X=x\right]f_X(x)\;dx.\end{aligned} \tag{7.58}$$

Because $\mathsf{E}[Y|X=x] = \tilde{\mu}_Y(x)$ in Theorem 7.15, it follows that

$$\mathsf{E}\left[Y-\mu_Y|X=x\right] = \tilde{\mu}_Y(x) - \mu_Y = \rho_{X,Y}\frac{\sigma_Y}{\sigma_X}(x-\mu_X). \tag{7.59}$$

Applying Equation (7.59) to Equation (7.58), we obtain

$$\mathsf{E}\left[g(X,Y)\right] = \frac{\rho_{X,Y}}{\sigma_X^2}\int_{-\infty}^{\infty}(x-\mu_X)^2\,f_X(x)\;dx = \rho_{X,Y}, \tag{7.60}$$

because the integral in the final expression is $\mathsf{Var}[X] = \sigma_X^2$.

Gaussian Conditional Variances

- Theorem 5.14 states that for any pair of random variables, $|\rho_{X,Y}| < 1$.

- Introducing this inequality to the formulas for conditional variance in Theorem 7.15 and Theorem 7.16 leads to the following inequalities:

$$\text{Var}[Y|X = x] = \sigma_Y^2(1 - \rho_{X,Y}^2) \leq \sigma_Y^2, \tag{7.61}$$

$$\text{Var}[X|Y = y] = \sigma_X^2(1 - \rho_{X,Y}^2) \leq \sigma_X^2. \tag{7.62}$$

  These formulas state that for $\rho_{X,Y} \neq 0$, learning the value of one of the random variables leads to a model of the other random variable with reduced variance.

Quiz 7.6

Let X and Y be jointly Gaussian $(0, 1)$ random variables with correlation coefficient $1/2$. What is the conditional PDF of X given $Y = 2$? What are the conditional expected value and conditional variance $\mathrm{E}[X|Y = 2]$ and $\mathrm{Var}[X|Y = 2]$?

Quiz 7.6 Solution

Since X and Y are bivariate Gaussian random variables with $\rho = 1/2$, $\mu_X = \mu_Y = 0$, and $\sigma_X = \sigma_Y = 1$, Theorem 7.16 tells us that given $Y = y$, X is conditionally Gaussian with parameters

$$\tilde{\mu}_X(y) = \rho y = \frac{y}{2}, \quad \tilde{\sigma}_X^2 = 1 - \rho^2. \tag{1}$$

For $y = 2$, we have

$$\tilde{\mu}_X = \tilde{\mu}_X(2) = 1 \qquad \tilde{\sigma}_X^2 = 3/4. \tag{2}$$

The conditional PDF of X is

$$f_{X|Y}(x|2) = \frac{1}{\sqrt{2\pi\tilde{\sigma}_X^2}} e^{-(x-\tilde{\mu}_X)^2/2\tilde{\sigma}_X^2} = \frac{1}{\sqrt{3\pi/2}} e^{-2(x-1)^2/3}. \tag{3}$$

Section 7.7

Matlab

Example 7.17 Problem

Repeating Example 7.2, find the conditional PMF for the length X of a video given event L that the video is long with $X \geq 5$ minutes.

Example 7.17 Solution

```
sx=(1:8)';
px=[0.15*ones(4,1);...
        0.1*ones(4,1)];
sxL=unique(find(sx>=5));
pL=sum(finitepmf(sx,px,sxL));
pxL=finitepmf(sx,px,sxL)/pL;
```

With random variable X defined by `sx` and `px` as in Example 3.36, this code solves this problem. The vector `sxL` identifies the event L, `pL` is the probability P$[L]$, and `pxL` is the vector of probabilities $P_{X|L}(x_i)$ for each $x_i \in L$.

Example 7.18 Problem

Write a function `xy = xytrianglerv(m)` that generates m sample pairs (X, Y) in Example 7.13.

Example 7.18 Solution

In Example 7.13, we found that

$$f_X(x) = \begin{cases} 2x & 0 \leq x \leq 1, \\ 0 & \text{otherwise,} \end{cases} \qquad f_{Y|X}(y|x) = \begin{cases} 1/x & 0 \leq y \leq x, \\ 0 & \text{otherwise.} \end{cases} \tag{7.63}$$

```
function xy = xytrianglerv(m);
x=sqrt(rand(m,1));
y=x.*rand(m,1);
xy=[x y];
```

For $0 \leq x \leq 1$, we have that $F_X(x) = x^2$. Using Theorem 6.5 to generate sample values of X, we define $u = F_X(x) = x^2$. Then, for $0 < u < 1$, $x = \sqrt{u}$. By Theorem 6.5, if U is uniform $(0, 1)$, then $\sqrt{U}$ has PDF $f_X(x)$. Next, we observe that given $X = x_i$, Y is the uniform $(0, x_i)$ random variable. Given another uniform $(0, 1)$ random variable U_i, Theorem 6.3(a) states that $Y_i = x_i U_i$ is the uniform $(0, x_i)$ random variable. We implement these ideas in the function `xytrianglerv.m`.

Quiz 7.7

For random variables X and Y with joint PMF $P_{X,Y}(x, y)$ given in Example 7.9, write a Matlab function `xy=dtrianglerv(m)` that generates m sample pairs.

Quiz 7.7 Solution

One straightforward method is to follow the approach of Example 5.26. Instead, we use an alternate approach. First we observe that X has the discrete uniform $(1,4)$ PMF. Also, given $X = x$, Y has a discrete uniform $(1,x)$ PMF. That is,

$$P_X(x) = \begin{cases} 1/4 & x = 1,2,3,4, \\ 0 & \text{otherwise,} \end{cases} \qquad P_{Y|X}(y|x) = \begin{cases} 1/x & y = 1,\ldots,x, \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

Given $X = x$, and an independent uniform $(0,1)$ random variable U, we can generate a sample value of Y with a discrete uniform $(1,x)$ PMF via $Y = \lceil xU \rceil$. This observation prompts the following program:

```
function xy=dtrianglerv(m)
sx=[1;2;3;4];
px=0.25*ones(4,1);
x=finiterv(sx,px,m);
y=ceil(x.*rand(m,1));
xy=[x';y'];
```