Systems and Signals 344

2025

Lecture 20

Conditional Probability Models: Conditioning by a Random Variable

Sections 7.4-7.6

Recap: Conditioning a RV by an Event

This simplifies when considering whether $x \in B$. For a random variable X and an event $B \subset S_X$ with $P(B) > 0$, the conditional PMF/PDF of X given B is:

Discrete:

$$P_{X|B}(x) = \begin{cases} \frac{P_X(x)}{P(B)} & x \in B \\ 0 & \text{otherwise} \end{cases}$$

Continuous:

$$f_{X|B}(x) = \begin{cases} \frac{f_X(x)}{P(B)} & x \in B \\ 0 & \text{otherwise} \end{cases}$$

Recap: Conditioning Two RVs by an Event

Given the event B with $P(B) > 0$,

Discrete:

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

Continuous:

$$f_{X,Y|B}(x,y) = \begin{cases} \frac{f_{X,Y}(x,y)}{P(B)} & (x,y) \in B \\ 0 & \text{otherwise} \end{cases}$$

Conditioning by a Random Variable

When an experiment produces a pair of random variables X and Y, observing a sample value of one of them provides partial information about the other.

To incorporate this information in the probability model, we derive new probability models: the conditional PMFs $P_{X|Y}(x|y)$ and $P_{Y|X}(y|x)$ for discrete random variables, as well as the conditional PDFs $f_{X|Y}(x|y)$ and $f_{Y|X}(y|x)$ for continuous random variables.

Conditioning by a Random Variable

In the previous lecture, we used the partial knowledge that the outcome of an experiment $(x, y) \in B$ in order to derive a new probability model for the experiment.

Now we turn our attention to the special case in which the partial knowledge consists of the value of one of the random variables: either $B = \{X = x\}$ or $B = \{Y = y\}$.

Learning $\{Y = y\}$ changes our knowledge of random variables X, Y. We now have complete knowledge of Y and modified knowledge of X. From this information, we derive a modified probability model for X.

Conditioning by a Random Variable

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

$$P_{X|Y}(x|y) = P_{X|Y}(x|Y = y) = P(X = x|Y = y)$$

This is sometimes called **point conditioning**.

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 $P_Y(y) > 0$,

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

Conditioning by a Random Variable

For continuous random variables, we don't define the PDF directly, since $P(Y = y) = 0$. Therefore, we we define it as the ratio of the joint PDF and the marginal PDF

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 $f_Y(y) > 0$, respectively,

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

Conditioning by a Random Variable

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 $P_Y(y) > 0$,

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

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 $f_Y(y) > 0$,

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

Conditioning by a Random Variable

Recall that if random variables X and Y are *independent* then

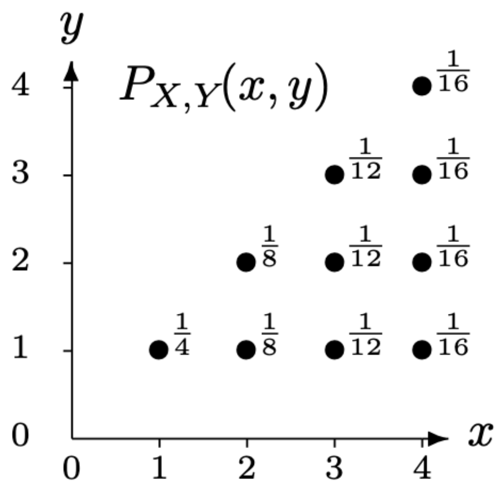
- $P_{X,Y}(x, y) = P_X(x)P_Y(y)$ (discrete)
- $f_{X,Y}(x, y) = f_X(x)f_Y(y)$ (continuous)

Therefore, if X and Y are independent then:

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

Therefore, gaining knowledge about Y does not affect the probability model of X.

Example 7.12

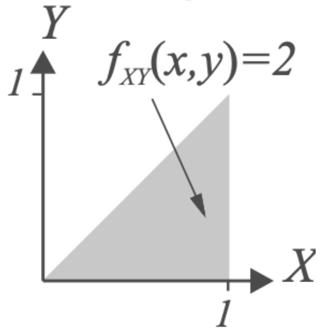


Random variables X and Y have the joint PMF $P_{X,Y}(x,y)$, as given in Example 7.9 and repeated in the accompanying graph. Find the conditional PMF of Y given $X = x$ for each $x \in S_X$.

Find only $P_{Y|X}(y|3)$

Example 7.13

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



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

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

Example 7.14

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.

Conditional Expected Value Given a RV

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

Discrete

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

Continuous

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

Conditional Expected Value Given a RV

A special case of the previous slide when $g(x, y) = x$, then $W = X$

Discrete

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

Continuous

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

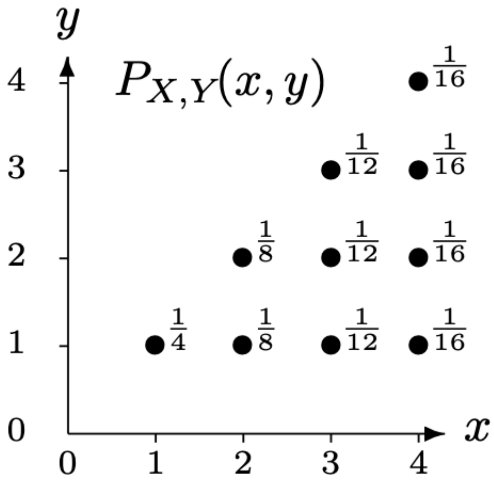
Conditional Expected Value Given a RV

Note that typically $\mathbf{E}[X|Y=y]$ is a **function** of y
and $\mathbf{E}[Y|X=x]$ is a **function** of x.

However, if X and Y are *independent*, then no additional information is provided, and therefore:

- $\mathbf{E}[X|Y=y] = \mathbf{E}[X]$ for all $y = S_Y$
- $\mathbf{E}[Y|X=x] = \mathbf{E}[Y]$ for all $x = S_X$

Example 7.12



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 $x \in S_X$.

For Example 7.12 find $\mathbf{E}[Y|X = 3]$

Example 7.16

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

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

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

Conditional Expected Value Given a RV

Since, for example $\mathbf{E}[X|Y=y]$ is a function of y, we can also write $\mathbf{E}[X|Y]$ to denote it is a function of random variable Y. But since a function of a random variable is another random variable, that means that $\mathbf{E}[X|Y]$ is also a random variable.

We can therefore also calculate $\mathbf{E}[\mathbf{E}[X|Y]]$, by first calculating $g(y)=\mathbf{E}[X|Y=y]$ and then $\mathbf{E}[g(Y)]$.

This is called **iterated expectation** and:

$$\mathbf{E}[\mathbf{E}[X|Y]] = \mathbf{E}[X]$$

In general:

$$\mathbf{E}[\mathbf{E}[g(X)|Y]] = \mathbf{E}[g(X)]$$

Bivariate Gaussian Random Variables: Conditional PDFs

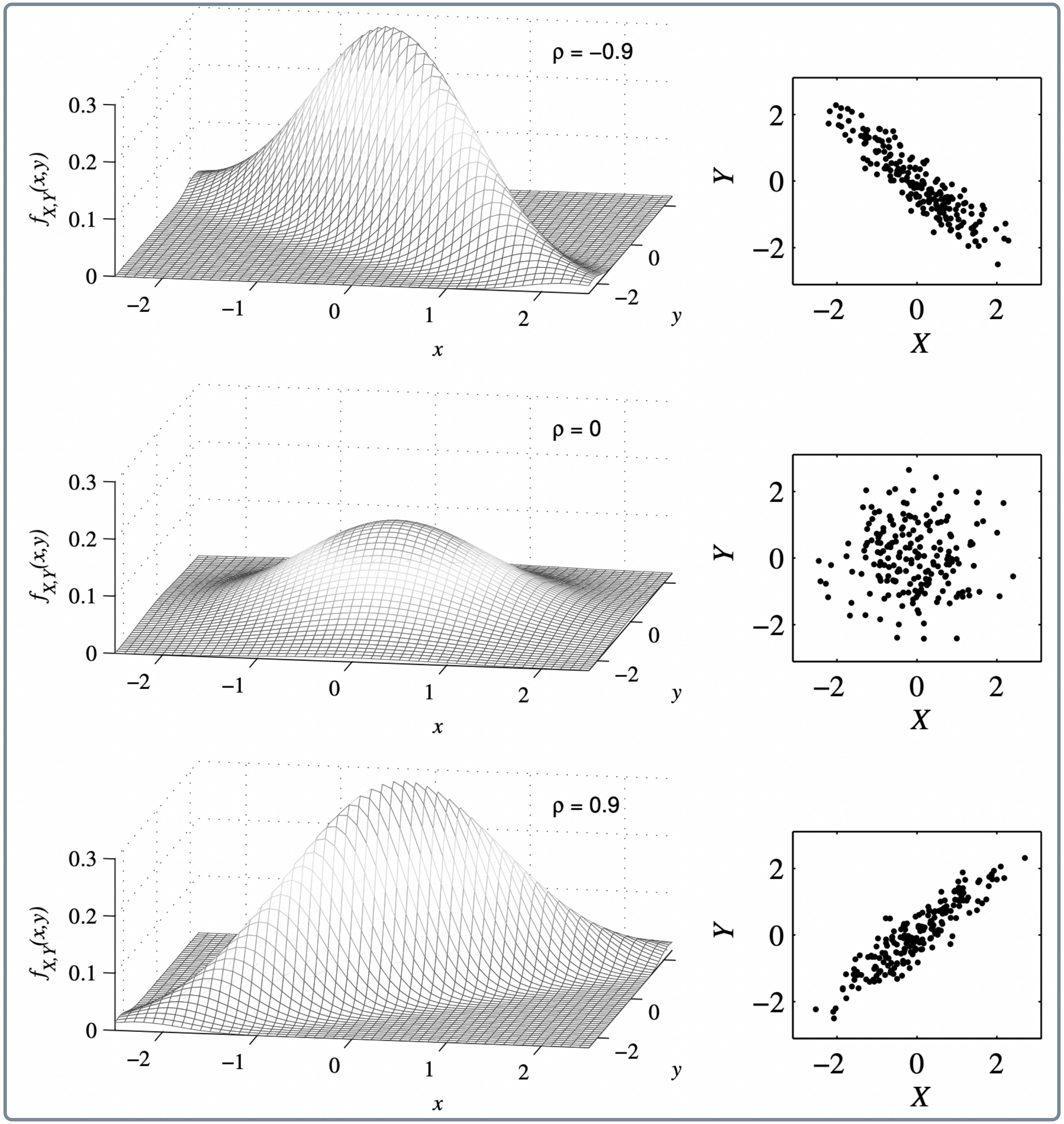
For bivariate Gaussian random variables X *and* Y*, the conditional PDFs* $f_{X|Y}(x|y)$ *and* $f_{Y|X}(y|x)$ *are Gaussian.*

$\mathrm{Var}[X|Y] \leq \mathrm{Var}[X]$ *and*
$\mathrm{Var}[Y|X] \leq \mathrm{Var}[Y]$*.*

Bivariate Gaussian Random Variables

Random variables X and Y have a **bivariate Gaussian PDF** with parameters $\mu_X, \mu_Y, \sigma_X > 0, \sigma_Y > 0,$ and $\rho_{X,Y}$, satisfying $-1 < \rho_{X,Y} < 1$ if

$$f_{X,Y}(x,y) = \frac{\exp\left[-\frac{1}{2\left(1-\rho_{X,Y}^2\right)}\left(\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - \frac{2\rho_{X,Y}(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y} - \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right)\right]}{2\pi\,\sigma_X\sigma_Y\sqrt{1-\rho_{X,Y}^2}}$$



ρ = −0.9
ρ = 0
ρ = 0.9
$f_{X,Y}(x,y)$
x
y
Y
X

Bivariate Gaussian Random Variables: Conditional PDFs

Recall from before, that the bivariate Gaussian can be expressed as:

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

where

$$\tilde{\mu}_Y(x) = \mu_Y + \rho_{X,Y}\frac{\sigma_Y}{\sigma_X}(x-\mu_X), \quad \tilde{\sigma}_Y = \sigma_Y\sqrt{1-\rho_{X,Y}^2}$$

Bivariate Gaussian Random Variables: Conditional PDFs

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

Since $f_{X,Y}(x,y) = f_X(x) f_{Y|X}(y|x)$ and the first half is simply the marginal PDF $f_X(x)$, it thus means that the second half is the conditional PDF $f_{Y|X}(y|x)$.

Bivariate Gaussian Random Variables: Conditional PDFs

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

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

Given $X = x$, the conditional expected value and variance of Y are:

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

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

The inverse is also true for $f_{X|Y}(x|y)$...

Bivariate Gaussian Random Variables: Conditional PDFs

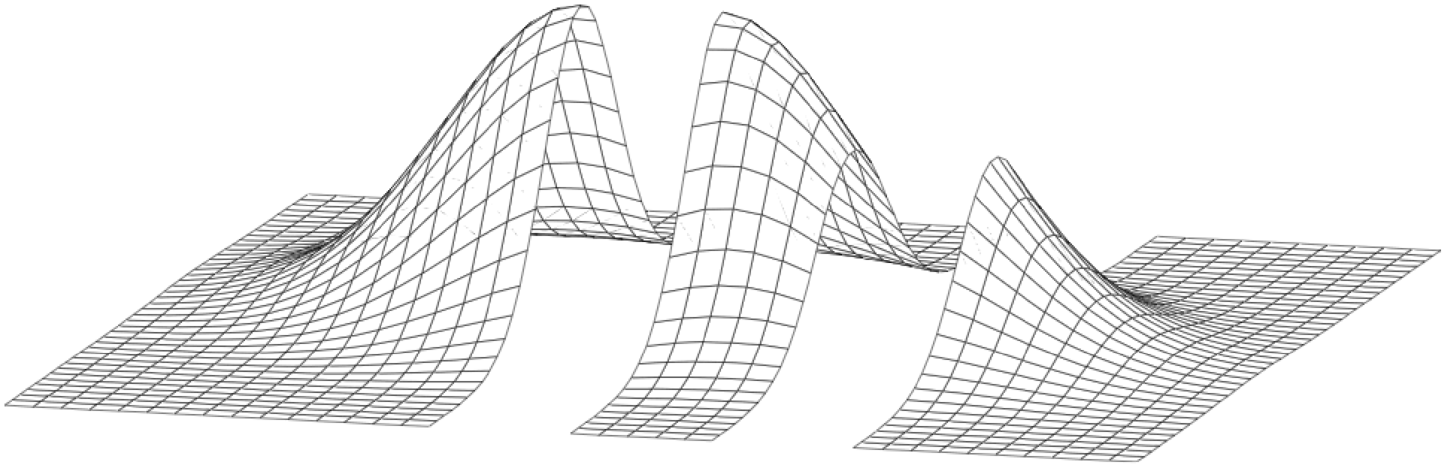


Figure 7.1 Cross-sectional view of the joint Gaussian PDF with $\mu_X = \mu_Y = 0$, $\sigma_X = \sigma_Y = 1$, and $\rho_{X,Y} = 0.9$. Theorem 7.15 confirms that the bell shape of the cross section occurs because the conditional PDF $f_{Y|X}(y|x)$ is Gaussian.

$X = x$ (with $f_X(x)$ also a constant) on each cross section, showing that the conditional probability model of Y is also Gaussian.

Bivariate Gaussian Random Variables: Conditional PDFs

For any pair of random variables (that are not identical), $|\rho_{X,Y}| < 1$. Therefore,

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

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

This means 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.

Thus, learning the value of, for example, Y reduces our uncertainty regarding X.