# Continuous Random Variables: Joint PDFs, Conditioning, Expectation and Independence

Berlin Chen
Department of Computer Science & Information Engineering
National Taiwan Normal University

Reference:

**-** D. P. Bertsekas, J. N. Tsitsiklis, *Introduction to Probability* , Sections 3.4-3.6

# Multiple Continuous Random Variables (1/2)

- Two continuous random variables $X$ and $Y$ associated with a common experiment are **jointly continuous** and can be described in terms of a **joint PDF** $f_{X,Y}$ satisfying

$$\mathbf{P}((X,Y)\in B)=\iint_{(x,y)\in B} f_{X,Y}(x,y)dxdy$$

  - $f_{X,Y}$ is a nonnegative function
  - Normalization Probability $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)dxdy=1$

- Similarly, $f_{X,Y}(a,c)$ can be viewed as the "probability per unit area" in the vicinity of $(a,c)$

$$\mathbf{P}(a\le X\le a+\delta, c\le Y\le c+\delta)$$

$$=\int_a^{a+\delta}\int_c^{c+\delta} f_{X,Y}(x,y)dxdy = f_{X,Y}(a,c)\cdot\delta^2$$

  - Where $\delta$ is a small positive number

# Multiple Continuous Random Variables (2/2)

- Marginal Probability

$$\mathbf{P}(X \in A) = \mathbf{P}(X \in A \text{ and } Y \in (-\infty, \infty))$$
$$= \int_{X \in A} \int_{-\infty}^{\infty} f_{X,Y}(x, y)\,dy\,dx$$

  – We have already defined that

$$\mathbf{P}(X \in A) = \int_{X \in A} f_X(x)\,dx$$

  - We thus have the marginal PDF

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)\,dy$$

Similarly

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)\,dx$$

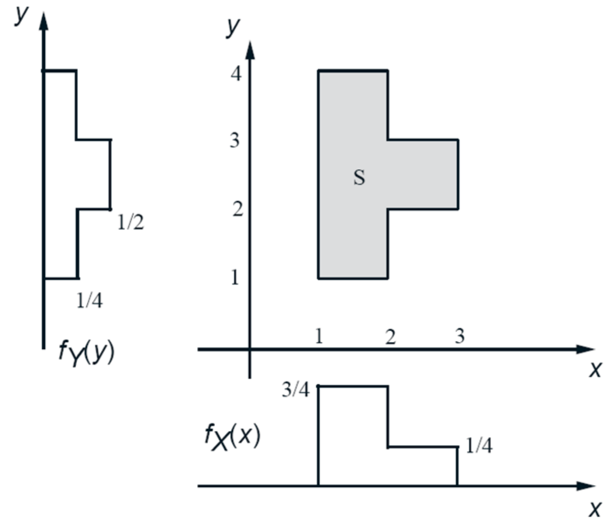
# An Illustrative Example

- **Example 3.10. Two-Dimensional Uniform PDF**. We are told that the joint PDF of the random variables $X$ and $Y$ is a constant $c$ on an area $S$ and is zero outside. Find the value of $c$ and the marginal PDFs of $X$ and $Y$.

The corresponding uniform joint PDF on an area $S$ is defined to be (cf. Example 3.9)

$$f_{X,Y}(x,y) = \begin{cases} \dfrac{1}{\text{Size of area S}}, & \text{if } (x,y) \in S \\ 0, & \text{otherwise} \end{cases}$$

$$\Rightarrow f_{X,Y}(x,y) = \frac{1}{4} \quad \text{for } (x,y) \in S$$




for $1 \le x \le 2$

$$\Rightarrow f_X(x) = \int_1^4 f_{X,Y}(x,y)dy = \int_1^4 \frac{1}{4} dy = \frac{3}{4}$$

for $2 \le x \le 3$

$$\Rightarrow f_X(x) = \int_2^3 f_{X,Y}(x,y)dy = \int_2^3 \frac{1}{4} dy = \frac{1}{4}$$

for $1 \le y \le 2$

$$\Rightarrow f_Y(y) = \int_1^2 f_{X,Y}(x,y)dx = \int_1^2 \frac{1}{4} dx = \frac{1}{4}$$

for $2 \le y \le 3$

$$\Rightarrow f_Y(y) = \int_1^3 f_{X,Y}(x,y)dx = \int_1^3 \frac{1}{4} dx = \frac{1}{2}$$

for $3 \le y \le 4$

$$\Rightarrow f_Y(y) = \int_1^2 f_{X,Y}(x,y)dx = \int_1^2 \frac{1}{4} dx = \frac{1}{4}$$

# Joint CDFs

- If $X$ and $Y$ are two (either continuous or discrete) random variables associated with the same experiment , their joint cumulative distribution function (Joint CDF) is defined by

$$F_{X,Y}(x, y) = \mathbf{P}(X \leq x, Y \leq y)$$

  - If $X$ and $Y$ further have a joint PDF $f_{X,Y}$ ( $X$ and $Y$ are continuous random variables) , then

$$F_{X,Y}(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(s, t) ds dt$$

And

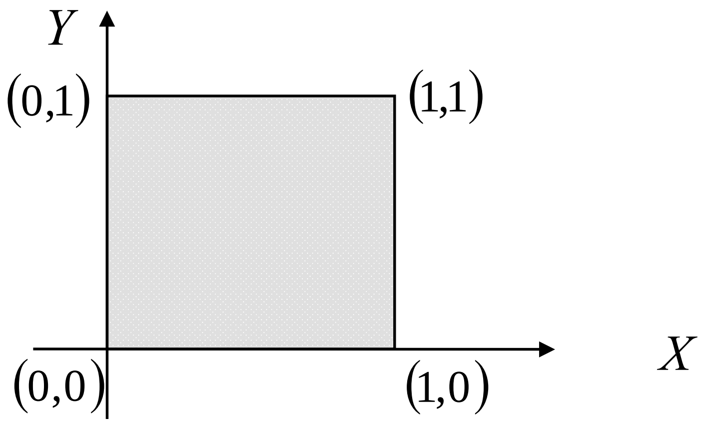
$$f_{X,Y}(x, y) = \frac{\partial^2 F_{X,Y}(x, y)}{\partial x \partial y}$$

If $F_{X,Y}$ can be differentiated at the point $(x, y)$

# An Illustrative Example

- **Example 3.12.** Verify that if *X* and *Y* are described by a uniform PDF on the unit square, then the joint CDF is given by

$$F_{X,Y}(x,y) = \mathbf{P}(X \le x, Y \le y) = xy, \quad \text{for } 0 \le x, y \le 1$$




$$\frac{\partial^2 F_{X,Y}(x,y)}{\partial x \partial y} = 1 = f_{X,Y}(x,y), \quad \text{for all } (x,y) \text{ in the unit square}$$

# Expectation of a Function of Random Variables

- If $X$ and $Y$ are jointly continuous random variables, and $g$ is some function, then $Z = g(X,Y)$ is also a random variable (can be continuous or discrete)
  - The expectation of $Z$ can be calculated by

$$\mathbf{E}[Z] = \mathbf{E}[g(X,Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y) f_{X,Y}(x,y)\,dx\,dy$$

  - If $Z$ is a linear function of $X$ and $Y$, e.g., $Z = aX + bY$, then

$$\mathbf{E}[Z] = \mathbf{E}[aX + bY] = a\mathbf{E}[X] + b\mathbf{E}[Y]$$

    - Where $a$ and $b$ are scalars

We will see in Section 4.1 methods for computing the PDF of $Z$ (if it has one).

# More than Two Random Variables

- The joint PDF of three random variables $X$ , $Y$ and $Z$ is defined in analogy with the case of two random variables

$$\mathbf{P}\left(\left(X,Y,Z\right)\in B\right)=\iiint_{\left(X,Y,Z\right)\in B} f_{X,Y,Z}\left(x,y,z\right)dxdydz$$

  – The corresponding marginal probabilities

$$f_{X,Y}\left(x,y\right)=\int_{-\infty}^{\infty} f_{X,Y,Z}\left(x,y,z\right)dz$$

$$f_{X}\left(x\right)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y,Z}\left(x,y,z\right)dydz$$

- The expected value rule takes the form

$$\mathbf{E}\left[g\left(X,Y,Z\right)\right]=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g\left(x,y,z\right)f_{X,Y,Z}\left(x,y,z\right)dxdy\,dz$$

  – If $g$ is linear (of the form $aX+bY+cZ$ ), then

$$\mathbf{E}\left[aX+bY+cZ\right]=a\mathbf{E}\left[X\right]+b\mathbf{E}\left[Y\right]+c\mathbf{E}\left[Z\right]$$

# Conditioning PDF Given an Event (1/3)

- The conditional PDF of a continuous random variable $X$, given an event $A$
  - If $A$ cannot be described in terms of $X$, the conditional PDF is defined as a nonnegative function $f_{X|A}(x)$ satisfying

$$\mathbf{P}(X \in B|A) = \int_B f_{X|A}(x)dx$$

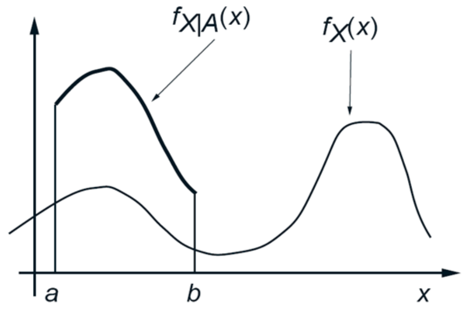
- Normalization property

$$\int_{-\infty}^{\infty} f_{X|A}(x)dx = 1$$

# Conditioning PDF Given an Event (2/3)

– If $A$ can be described in terms of $X$ ( $A$ is a subset of the real line with $\mathbf{P}(X \in A) > 0$), the conditional PDF is defined as a nonnegative function $f_{X|A}(x)$ satisfying

$$f_{X|A}(x) = \begin{cases} \dfrac{f_X(x)}{\mathbf{P}(X \in A)}, & \text{if } x \in A \\ 0, & \text{otherwise} \end{cases}$$




$f_{X|A}$ remains the same shape as $f_X$ except that it is scaled along the vertical axis

- The conditional PDF is zero outside the conditioning event

and for any subset $B$

$$\mathbf{P}(X \in B | X \in A) = \frac{\mathbf{P}(X \in B, X \in A)}{\mathbf{P}(X \in A)} = \frac{\int_{A \cap B} f_X(x)dx}{\mathbf{P}(X \in A)} = \int_{A \cap B} f_{X|A}(x)dx$$

– Normalization Property $\int_{-\infty}^{\infty} f_{X|A}(x)dx = \int_A f_{X|A}(x)dx = 1$

# Conditioning PDF Given an Event (3/3)

- If $A_1, A_2, \ldots, A_n$ are disjoint events with $\mathbf{P}(A_i) > 0$ for each $i$, that form a partition of the sample space, then

$$f_X(x) = \sum_{i=1}^{n} \mathbf{P}(A_i) f_{X|A_i}(x)$$

  – Verification of the above total probability theorem

think of $\{X \leq x\}$ as an event $B$, and use the total probability theorem from Chapter 1

$$\mathbf{P}(X \leq x) = \sum_{i=1}^{n} \mathbf{P}(A_i)\mathbf{P}(X \leq x | A_i)$$

$$\Rightarrow \int_{-\infty}^{x} f_X(t)dt = \sum_{i=1}^{n} \mathbf{P}(A_i) \int_{-\infty}^{x} f_{X|A_i}(t)dt$$

Taking the derivative of both sides with respect to $x$

$$\Rightarrow f_X(x) = \sum_{i=1}^{n} \mathbf{P}(A_i) f_{X|A_i}(x)$$

# Illustrative Examples (1/2)

- **Example 3.13. The exponential random variable is memoryless.**
  - The time *T* until a new light bulb burns out is exponential distribution. John turns the light on, leave the room, and when he returns, *t* time units later, find that the light bulb is still on, which corresponds to the event $A=\{T>t\}$
  - Let *X* be the additional time until the light bulb burns out. What is the conditional PDF of *X* given A ?

$T$ is exponential

$$f_T(t)=\begin{cases}\lambda e^{-\lambda t}, & t>0\\ 0, & \text{otherwise}\end{cases}$$

$$P(T>t)=e^{-\lambda t}$$

$X=T-t,\ A=\{T>t\}$

The conditional CDF of $X$ given $A$ is defined by

$$P(X>x|A)=P(T-t>x|T>t)\ (\text{where } x\geq 0)$$

$$=P(T>t+x|T>t)=\frac{P(T>t+x \text{ and } T>t)}{P(T>t)}$$

$$=\frac{P(T>t+x)}{P(T>t)}$$

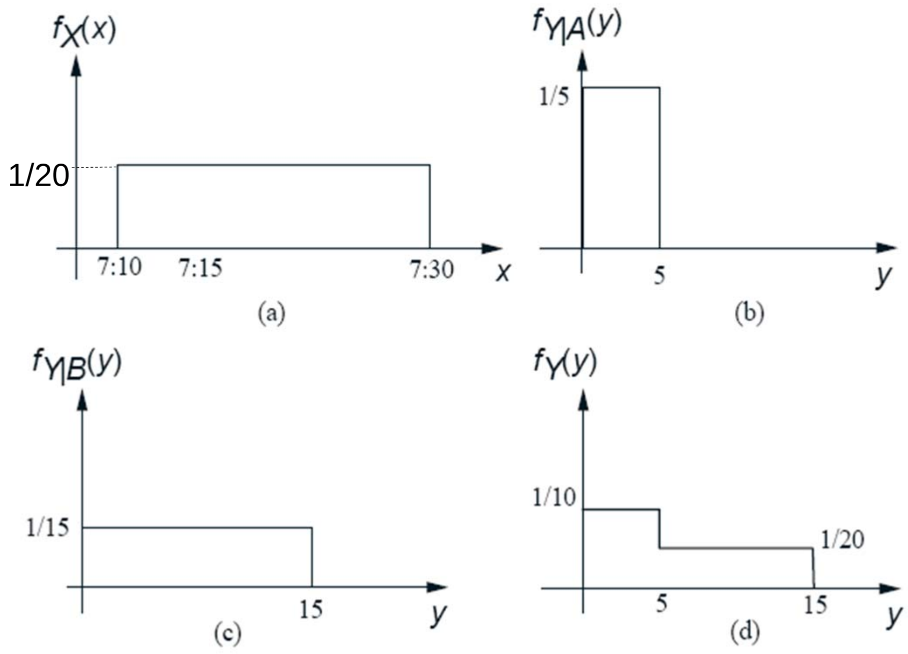
$$=\frac{e^{-\lambda(t+x)}}{e^{-\lambda t}}$$

$$=e^{-\lambda x}$$

$\therefore$ The conditional PDF of $X$ given the event $A$ is also exponential with parameter $\lambda$.

# Illustrative Examples (2/2)

- **Example 3.14.** The metro train arrives at the station near your home every quarter hour starting at 6:00 AM. You walk into the station every morning between 7:10 and 7:30 AM, with the time in this interval being a uniform random variable. What is the PDF of the time you have to wait for the first train to arrive?




- The arrival time, denoted by $X$, is a uniform random variable over the interval $7:10$ to $7:30$
- Let random varible $Y$ model the waiting time
- Let $A$ be a event
  $A = \{7:10 \le X \le 7:15\}$ (You board the 7:15 train)
- Let $B$ be a event
  $B = \{7:15 < X \le 7:30\}$ (You board the 7:30 train)
- Let $Y$ be uniform conditioned on $A$
- Let $Y$ be uniform conditioned on $B$

Total Probability theorem:

$$P_Y(y) = P(A)P_{Y|A}(y) + P(B)P_{Y|B}(y)$$

$$\text{For } 0 \le y \le 5,\ P_Y(y) = \frac{1}{4}\cdot\frac{1}{5} + \frac{3}{4}\cdot\frac{1}{15} = \frac{1}{10}$$

$$\text{For } 5 < y \le 15,\ P_Y(y) = \frac{1}{4}\cdot 0 + \frac{3}{4}\cdot\frac{1}{15} = \frac{1}{20}$$

# Conditioning one Random Variable on Another

- Two continuous random variables $X$ and $Y$ have a joint PDF. For any $y$ with $f_Y(y) > 0$, the conditional PDF of $X$ given that $Y = y$ is defined by

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

  – Normalization Property $\int_{-\infty}^{\infty} f_{X|Y}(x|y)dx = 1$

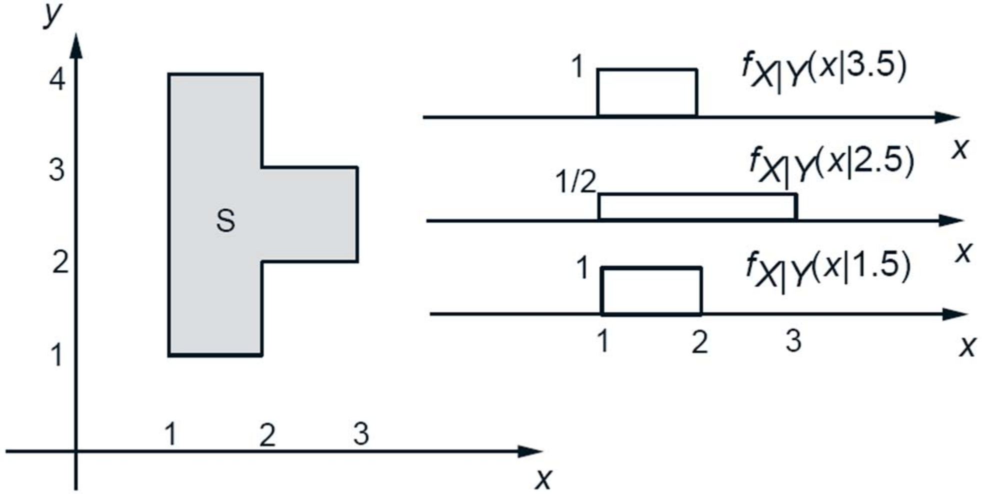
- The marginal, joint and conditional PDFs are related to each other by the following formulas

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

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy. \quad \text{marginalization}$$

# Illustrative Examples (1/2)

- Notice that the conditional PDF $f_{X|Y}(x|y)$ has the same shape as the joint PDF $f_{X,Y}(x,y)$ , because the normalizing factor $f_Y(y)$ does not depend on $x$




$$f_{X|Y}(x|3.5) = \frac{f_{X,Y}(x,3.5)}{f_Y(3.5)} = \frac{1/4}{1/4} = 1$$

$$f_{X|Y}(x|3.5) = \frac{f_{X,Y}(x,2.5)}{f_Y(2.5)} = \frac{1/4}{1/2} = 1/2$$

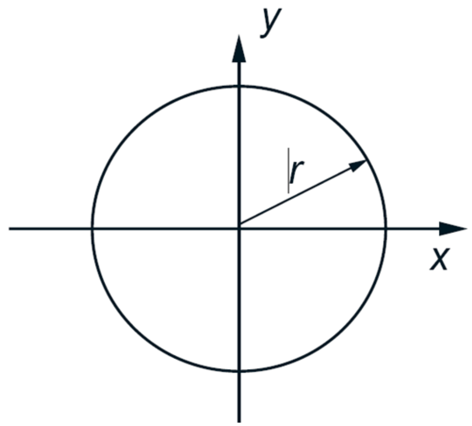
$$f_{X|Y}(x|3.5) = \frac{f_{X,Y}(x,1.5)}{f_Y(1.5)} = \frac{1/4}{1/4} = 1$$

cf. example 3.13

**Figure 3.16:** Visualization of the conditional PDF $f_{X|Y}(x|y)$ .
Let $X$ , $Y$ have a joint PDF which is uniform on the set $S$ . For each fixed $y$ , we consider the joint PDF along the slice $Y = y$ and normalize it so that it integrates to 1

# Illustrative Examples (2/2)

- **Example 3.15. Circular Uniform PDF.** Ben throws a dart at a circular target of radius $r$. We assume that he always hits the target, and that all points of impact $(x, y)$ are equally likely, so that the joint PDF $f_{X,Y}(x,y)$ of the random variables $x$ and $y$ is uniform
  – What is the marginal PDF $f_Y(y)$




$$f_{X,Y}(x,y)=\begin{cases}\dfrac{1}{\text{area of the circle}}, & \text{if } (x,y) \text{ is in the circle} \\ 0, & \text{otherwise}\end{cases}$$

$$=\begin{cases}\dfrac{1}{\pi r^2}, & x^2+y^2\le r^2 \\ 0, & \text{otherwise}\end{cases}$$

$$f_Y(y)=\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx=\int_{x^2+y^2\le r^2}\frac{1}{\pi r^2}\,dx$$

$$=\frac{1}{\pi r^2}\int_{x^2+y^2\le r^2} 1\,dx=\frac{1}{\pi r^2}\int_{-\sqrt{r^2-y^2}}^{\sqrt{r^2-y^2}} 1\,dx$$

$$=\frac{2}{\pi r^2}\sqrt{r^2-y^2},\ if\ |y|\le r$$

(Notice here that PDF $f_Y(y)$ is not uniform)

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

$$=\frac{\frac{1}{\pi r^2}}{\frac{2}{\pi r^2}\sqrt{r^2-y^2}}$$

$$=\frac{1}{2\sqrt{r^2-y^2}},\qquad if\ x^2+y^2\le r^2$$

For each value $y$, $f_{X|Y}(x|y)$ is uniform

# Conditional Expectation Given an Event

- The conditional expectation of a continuous random variable $X$, given an event $A$ ($\mathbf{P}(A) > 0$), is defined by

$$\mathbf{E}[X \mid A] = \int_{-\infty}^{\infty} x f_{X \mid A}(x) dx$$

  – The conditional expectation of a function $g(X)$ also has the form

$$\mathbf{E}[g(X) \mid A] = \int_{-\infty}^{\infty} g(x) f_{X \mid A}(x) dx$$

  – Total Expectation Theorem

$$\mathbf{E}[X] = \sum_{i=1}^{n} \mathbf{P}(A_i) \mathbf{E}[X \mid A_i]$$

and

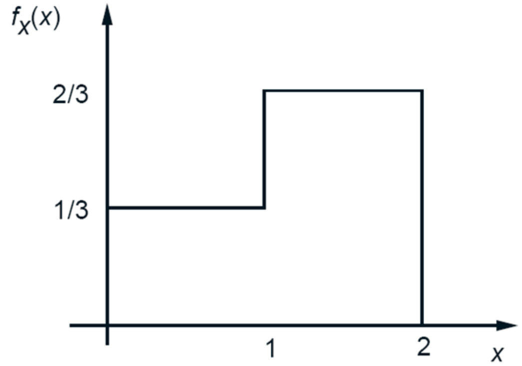
$$\mathbf{E}[g(X)] = \sum_{i=1}^{n} \mathbf{P}(A_i) \mathbf{E}[g(X) \mid A_i]$$

    - Where $A_1, A_2, \ldots, A_n$ are disjoint events with $\mathbf{P}(A_i) > 0$ for each $i$, that form a partition of the sample space

# An Illustrative Example

- **Example 3.17. Mean and Variance of a Piecewise Constant PDF.** Suppose that the random variable $X$ has the piecewise constant PDF

$$f_X(x) = \begin{cases} 1/3, & \text{if } 0 \le x \le 1, \\ 2/3, & \text{if } 1 \le x \le 2, \\ 0, & \text{otherwise.} \end{cases}$$




Define event $A_1 = \{X \text{ lies in the first interval } [0,1]\}$

event $A_2 = \{X \text{ lies in the second interval } [1,2]\}$

$$\Rightarrow \mathbf{P}(A_1) = \int_0^1 1/3 dx = 1/3, \ \mathbf{P}(A_2) = \int_1^2 2/3 dx = 2/3$$

$$f_{X|A_1}(x) = \begin{cases} \dfrac{f_X(x)}{P(X \in A_1)} = 1, & 0 \le x \le 1 \\ 0, & \text{otherwise} \end{cases} \qquad f_{X|A_2}(x) = \begin{cases} \dfrac{f_X(x)}{P(X \in A_2)} = 1, & 1 \le x \le 2 \\ 0, & \text{otherwise} \end{cases}$$

Recall that the mean and second moment of a uniform random variable over an interval $[a, b]$ is $(a+b)/2$ and $(a^2 + ab + b^2)/3$

$$\Rightarrow \mathrm{E}[X|A_1] = 1/2, \mathrm{E}[X^2|A_1] = 1/3$$

$$\mathrm{E}[X|A_2] = 3/2, \mathrm{E}[X^2|A_2] = 7/3$$

$$\Rightarrow \mathrm{E}[X] = \mathbf{P}(A_1)\mathrm{E}[X|A_1] + \mathbf{P}(A_2)\mathrm{E}[X|A_2]$$

$$= 1/3 \cdot 1/2 + 2/3 \cdot 3/2 = 7/6$$

$$\mathrm{E}[X^2] = \mathbf{P}(A_1)\mathrm{E}[X^2|A_1] + \mathbf{P}(A_2)\mathrm{E}[X^2|A_2]$$

$$= 1/3 \cdot 1/3 + 2/3 \cdot 7/3 = 15/9$$

$$\therefore \operatorname{var}(X) = 15/9 - (7/6)^2 = 11/36$$

# Conditional Expectation Given a Random Variable

- The properties of unconditional expectation carry though, with the obvious modifications, to conditional expectation

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

$$\mathbf{E}\left[g(X) \middle| Y = y\right] = \int_{-\infty}^{\infty} g(x) f_{X|Y}(x|y)\,dx$$

$$\mathbf{E}\left[g(X, Y) \middle| Y = y\right] = \int_{-\infty}^{\infty} g(x, y) f_{X|Y}(x|y)\,dx$$

# Total Probability/Expectation Theorems

- Total Probability Theorem
  - For any event $A$ and a continuous random variable $Y$

$$\mathbf{P}(A) = \int_{-\infty}^{\infty} \mathbf{P}(A \mid Y = y) f_Y(y) dy$$

- Total Expectation Theorem
  - For any continuous random variables $X$ and $Y$

$$\mathbf{E}[X] = \int_{-\infty}^{\infty} \mathbf{E}[X \mid Y = y] f_Y(y) dy$$

$$\mathbf{E}[g(X)] = \int_{-\infty}^{\infty} \mathbf{E}[g(X) \mid Y = y] f_Y(y) dy$$

$$\mathbf{E}[g(X, Y)] = \int_{-\infty}^{\infty} \mathbf{E}[g(X, Y) \mid Y = y] f_Y(y) dy$$

# Independence

- Two continuous random variables $X$ and $Y$ are independent if

$$f_{X,Y}(x, y) = f_X(x) f_Y(y), \quad \text{for all } x,y$$

  – Since that

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

    - We therefore have

$$f_{X|Y}(x|y) = f_X(x), \quad \text{for all } x \text{ and all } y \text{ with } f_Y(y) > 0$$

    - Or

$$f_{Y|X}(y|x) = f_Y(y), \quad \text{for all } y \text{ and all } x \text{ with } f_X(x) > 0$$

# More Factors about Independence (1/2)

- If two continuous random variables $X$ and $Y$ are independent, then
  - Any two events of the forms $\{X \in A\}$ and $\{Y \in B\}$ are independent

$$\begin{aligned}\mathbf{P}(X \in A, Y \in B) &= \int_{x\in A}\int_{y\in B} f_{X,Y}(x, y)\,dy\,dx \\ &= \int_{x\in A}\int_{y\in B} f_X(x) f_Y(y)\,dy\,dx \\ &= \left[\int_{x\in A} f_X(x)\,dx\right]\left[\int_{y\in B} f_Y(y)\,dy\right] \\ &= \mathbf{P}(X \in A)\mathbf{P}(Y \in B)\end{aligned}$$

  - It also implies that

$$F_{X,Y}(x, y) = \mathbf{P}(X \le x, Y \le y) = \mathbf{P}(X \le x)\mathbf{P}(Y \le y) = F_X(x) F_Y(x)$$

  - The converse statement is also true (See the end-of-chapter problem 28)

# More Factors about Independence (2/2)

- If two continuous random variables $X$ and $Y$ are independent, then
  - $\mathbf{E}[XY] = \mathbf{E}[X]\mathbf{E}[Y]$
  - $\mathrm{var}(X+Y) = \mathrm{var}(X) + \mathrm{var}(Y)$
  - The random variables $g(X)$ and $h(Y)$ are independent for any functions $g$ and $h$
    - Therefore,

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

# Recall: the Discrete Bayes' Rule

- Let $A_1, A_2, \ldots, A_n$ be disjoint events that form a partition of the sample space, and assume that $\mathbf{P}(A_i) \geq 0$, for all $i$ . Then, for any event $B$ such that $\mathbf{P}(B) > 0$ we have

$$\mathbf{P}(A_i|B) = \frac{\mathbf{P}(A_i)\mathbf{P}(B|A_i)}{\mathbf{P}(B)}$$

$$= \frac{\mathbf{P}(A_i)\mathbf{P}(B|A_i)}{\sum_{k=1}^{n}\mathbf{P}(A_k)\mathbf{P}(B|A_k)}$$

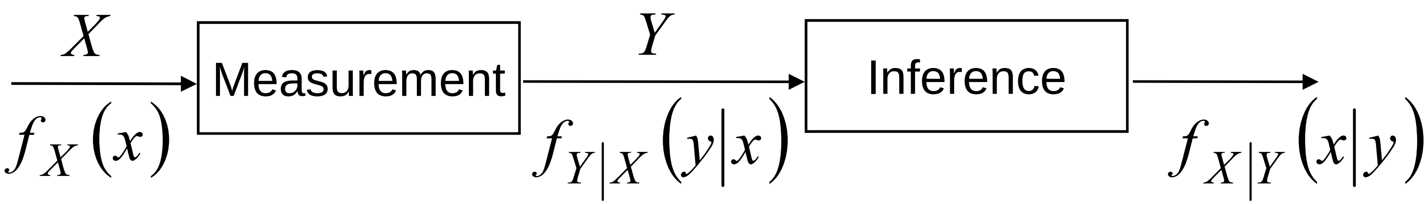
$$= \frac{\mathbf{P}(A_i)\mathbf{P}(B|A_i)}{\mathbf{P}(A_1)\mathbf{P}(B|A_1) + \cdots + \mathbf{P}(A_n)\mathbf{P}(B|A_n)}$$

Multiplication rule

Total probability theorem

# Inference and the Continuous Bayes' Rule

- As we have a model of an underlying but unobserved phenomenon, represented by a random variable $X$ with PDF $f_X$, and we make a noisy measurement $Y$, which is modeled in terms of a conditional PDF $f_{Y|X}$. Once the experimental value of $Y$ is measured, what information does this provide on the unknown value of $X$?




$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{f_X(x) f_{Y|X}(y|x)}{\int_{-\infty}^{\infty} f_X(t) f_{Y|X}(y|t)\,dt}$$

Note that

$f_X f_{Y|X} = f_{X,Y} = f_Y f_{X|Y}$

# Inference and the Continuous Bayes' Rule (2/2)

## Inference about a Discrete Random Variable

- If the unobserved phenomenon is inherently discrete
  - Let $N$ is a discrete random variable of the form $\{N = n\}$ that represents the different discrete probabilities for the unobserved phenomenon of interest, and $p_N$ be the PMF of $N$

$$
\begin{aligned}
\mathbf{P}(N = n \mid Y = y) &\approx \mathbf{P}(N = n \mid y \le Y \le y + \delta) \\
&= \frac{\mathbf{P}(N = n)\mathbf{P}(y \le Y \le y + \delta \mid N = n)}{\mathbf{P}(y \le Y \le y + \delta)} \\
&\approx \frac{p_N(n) f_{Y|N}(y|n)\delta}{f_Y(y)\delta} \\
&= \frac{p_N(n) f_{Y|N}(y|n)}{\sum_i p_N(i) f_{Y|N}(y|i)}
\end{aligned}
$$

Total probability theorem

# Illustrative Examples (1/2)

- **Example 3.19.** A lightbulb produced by the General Illumination Company is known to have an exponentially distributed lifetime $Y$. However, the company has been experiencing quality control problems. On any given day, the parameter $\Lambda = \lambda$ of the PDF of $Y$ is actually a random variable, uniformly distributed in the interval $[1,\ 3/2]$.
  - If we test a lightbulb and record its lifetime ( $Y = y$ ), what can we say about the underlying parameter $\lambda$ ?

$$f_{Y|\Lambda}(y|\lambda) = \lambda e^{-\lambda y}, \quad y \geq 0, \lambda > 0$$

Conditioned on $\Lambda = \lambda$, $Y$ has a exponential distribution with parameter $\lambda$

$$f_{\Lambda}(\lambda) = \begin{cases} 2, & \text{for } 1 \leq \lambda \leq 3/2 \\ 0, & \text{otherwise} \end{cases}$$

$$f_{\Lambda|Y}(\lambda|y) = \frac{f_{\Lambda}(\lambda) f_{Y|\Lambda}(y|\lambda)}{\int_1^{3/2} f_{\Lambda}(t) f_{Y|\Lambda}(y|t) dt} = \frac{2\lambda e^{-\lambda y}}{\int_1^{3/2} 2te^{-ty}\, dt}, \quad \text{for } 1 \leq \lambda \leq 3/2$$

# Illustrative Examples (2/2)

- **Example 3.20. Signal Detection.** A binary signal $S$ is transmitted, and we are given that $\mathbf{P}(S=1)=p$ and $\mathbf{P}(S=-1)=1-p$ .
  - The received signal is $Y = S + N$, where $N$ is a normal noise with zero mean and unit variance , independent of $S$ .
  - What is the probability that $S=1$ , as a function of the observed value $y$ of $Y$ ?

$$f_{Y|S}(y|s)=\frac{1}{\sqrt{2\pi}\sigma}e^{-(y-s)^2/2}, \text{for } s=1 \text{ and -1, and } -\infty \leq y \leq \infty$$

Conditioned on $S = s$ , $Y$ has a normal distribution with mean $s$ and unit variance

$$\mathbf{P}(S=1|Y=y)=\frac{p_S(1)f_{Y|S}(y|1)}{f_Y(y)}=\frac{p_S(1)f_{Y|S}(y|1)}{p_S(1)f_{Y|S}(y|1)+p_S(-1)f_{Y|S}(y|-1)}$$

$$=\frac{p\frac{1}{\sqrt{2\pi}}e^{-(y-1)^2/2}}{p\frac{1}{\sqrt{2\pi}}e^{-(y-1)^2/2}+(1-p)\frac{1}{\sqrt{2\pi}}e^{-(y+1)^2/2}}$$

$$=\frac{e^{-(y^2+1)/2}\cdot pe^{y}}{e^{-(y^2+1)/2}\cdot pe^{y}+e^{-(y^2+1)/2}\cdot(1-p)e^{-y}}=\frac{pe^{y}}{pe^{y}+(1-p)e^{-y}}$$

# Inference Based on a Discrete Random Variable

- The earlier formula expressing $\mathbf{P}(A|Y=y)$ in terms of $f_{Y|A}(y)$ can be turned around to yield

$$f_{Y|A}(y) = \frac{f_Y(y)\mathbf{P}(A|Y=y)}{\mathbf{P}(A)}$$

$$= \frac{f_Y(y)\mathbf{P}(A|Y=y)}{\int_{-\infty}^{\infty} f_Y(t)\mathbf{P}(A|Y=t)dt}$$

?

$$\mathbf{P}(A)f_{Y|A}(y) = f_Y(y)\mathbf{P}(A|Y=y)$$

$$\Rightarrow \int_{-\infty}^{\infty} \mathbf{P}(A)f_{Y|A}(y)dy = \int_{-\infty}^{\infty} f_Y(y)\mathbf{P}(A|Y=y)dy$$

$$\Rightarrow \mathbf{P}(A) = \int_{-\infty}^{\infty} f_Y(y)\mathbf{P}(A|Y=y)dy \ (\because \text{normalizat ion property} : \int_{-\infty}^{\infty} f_{Y|A}(y)dy = 1)$$

# Recitation

- SECTION 3.4 Joint PDFs of Multiple Random Variables
  - Problems 15, 16
- SECTION 3.5 Conditioning
  - Problems 18, 20, 23, 24
- SECTION 3.6 The Continuous Bayes' Rule
  - Problems 34, 35