EE5110: Probability Foundations for Electrical Engineers **July-November 2015**

Lecture 13: Conditional Distributions and Joint Continuity

Lecturer: Dr. Krishna Jagannathan *Scribe: Subrahmanya Swamy P*

13.1 Conditional Probability for Discrete Random Variables

If X and Y are discrete random variables, then the range of the map $(X(\cdot), Y(\cdot))$ is a countable subset of $\mathbb{R}^2$. This is because the Cartesian product of two countable sets is countable (Why?). Hence, $(X(.), Y(.))$ is a discrete random variable on $\mathbb{R}^2$. We will see later that we do not have similar result when X and Y are continuous random variables i.e., if X and Y are marginally continuous random variables, they need not be jointly continuous.

The joint pmf of discrete random variables X and Y is defined as:

$$p_{X,Y}(x,y) = \mathbb{P}(X = x, Y = y), \ x, y \in \mathbb{R}.$$

The joint pmf uniquely specifies the joint law. In particular, for any $B \in \mathcal{B}(\mathbb{R}^2)$,

$$\mathbb{P}_{X,Y}(B) = \sum_{x,y \in B} p_{X,Y}(x,y). \tag{13.1}$$

An example of two discrete random variables is shown in Figure (13.1).

13.1.1 Conditional pmf

Now, we define the conditional pmf for discrete random variables.

Definition 13.1 *Let X and Y be discrete random variables defined on $(\Omega, \mathbb{F}, \mathbb{P})$. Conditional probability of X given Y is defined as:*

$$p_{X|Y}(x|y) = \mathbb{P}(X = x | Y = y) = \frac{\mathbb{P}(X = x, Y = y)}{\mathbb{P}(Y = y)} = \frac{p_{X,Y}(x,y)}{p_Y(y)} \ \text{where } p_Y(y) > 0.$$

The following theorem characterizes independence of discrete random variables in terms of the conditional pmf.

Theorem 13.2 *The following statements are equivalent for discrete random variables X and Y:*

(a) X, Y are independent.

(b) For all $x, y \in \mathbb{R}$, $\{X = x\}$ and $\{Y = y\}$ are independent.

(c) For all $x, y \in \mathbb{R}$, $\mathbb{P}_{X,Y}(x,y) = \mathbb{P}_X(x)\mathbb{P}_Y(y)$.

(d) For all $x, y \in \mathbb{R}$ such that $p_X(y) > 0, p_{X|Y}(x|y) = p_X(x)$.

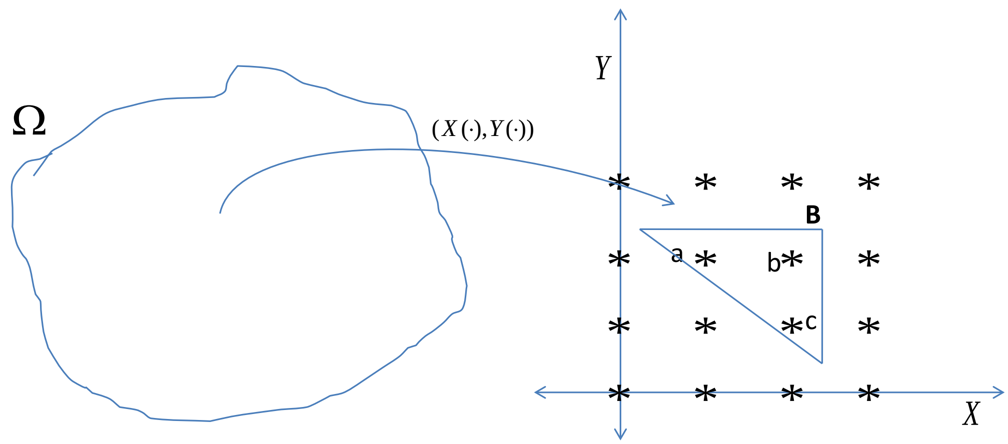

Figure 13.1: An example of two discrete random variables. The probability measure assigned to any Borel set B on $\mathbb{R}2$ can be obtained by summing the joint pmf over the set B; see (13.1).

Proof: $(b) \Leftrightarrow (c)$ and $(c) \Leftrightarrow (d)$ are directly follow from the definitions. Now, we prove equivalence of (a) and (c).
$(a) \Rightarrow (c)$:
X and Y are independent $\Rightarrow \mathbb{P}(X \in B_1, Y \in B_2) = \mathbb{P}(X \in B_1)\mathbb{P}(Y \in B_2)$. Take $B_1 = \{x\}$ and $B_2 = \{y\}$ then the result follows.
$(c) \Rightarrow (a)$:

$$\mathbb{P}(X \in B_1, Y \in B_2) = \sum_{x \in B_1, y \in B_2} p_{X,Y}(x,y) = \sum_{x \in B_1}\sum_{y \in B_2} p_X(x)p_Y(y) = \sum_{x \in B_1} p_X(x) \sum_{y \in B_2} p_Y(y)$$
$$= \mathbb{P}(X \in B_1)\mathbb{P}(Y \in B_2).$$

■

13.2 Jointly Continuous Distributions

Definition 13.3 *Two Random variables X and Y are said to be* jointly continuous, *if the joint probability law $\mathbb{P}_{X,Y}$ is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^2$. That is, for every Borel set $N \subset \mathbb{R}^2$ of Lebesgue measure zero, we have $\mathbb{P}(\{(X,Y) \in N\}) = 0$.*

The Radon-Nikodym theorem for this situation would assert the following

Theorem 13.4 *X, Y are jointly continuous random variables if and only if there exists a measurable function*

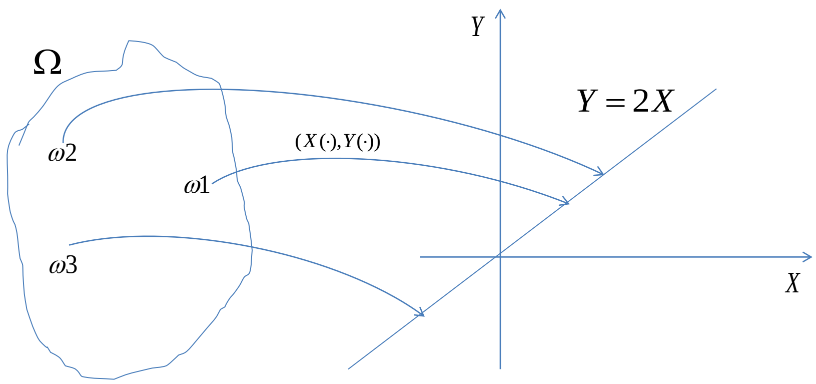

Figure 13.2: $Y = 2X$

$f_{X,Y} : \mathbb{R}^2 \to [0, \infty)$ *such that for any Borel set B on* $\mathbb{R}^2$,

$$\mathbb{P}_{X,Y}(B) = \int_B f_{X,Y} d\lambda,$$

where λ *is Lebesgue measure on* $\mathbb{R}^2$.

In particular, taking $B = (-\infty, x] \times (-\infty, y]$, we have

$$F_{X,Y}(x,y) = \mathbb{P}(\{X \le x, Y \le y\}) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(u,v) dv du, \tag{13.2}$$

where $F_{X,Y}(X,Y)$ and $f_{X,Y}(x,y)$ are the joint cdf and joint pdf respectively. The joint pdf is thus a complete characterization of the joint law, for jointly continuous random variables.

Caution: If X is continuous and Y is continuous, (X,Y) need not be jointly continuous. This can be seen from the following example.

Example: Let $X \sim \mathcal{N}(0,1)$ and $Y = 2X$. i.e. $Y \sim \mathcal{N}(0,4)$. In this case, though X is continuous and Y is continuous, (X,Y) are not jointly continuous.
This can be understood from the Figure 13.2. Each $\omega \in \Omega$ is mapped on to the straight line $Y = 2X$ on $\mathbb{R}^2$. The Lebesgue measure of the line (set), $L = \{(x,y) \in \mathbb{R}^2 : y = 2x\}$ is zero, but the corresponding probability, $\mathbb{P}_{X,Y}(L) = 1$, since every $\omega \in \Omega$ is mapped to this straight line. Thus, from the definition of jointly continuous random variables, X and Y are not jointly continuous.

On the other hand, if X and Y are jointly continuous, their marginals are necessarily continuous. To see this, note that

$$\begin{aligned}\mathbb{P}(\{X \leq x, Y \leq y\}) &= \int_{-\infty}^{x}\int_{-\infty}^{y} f_{X,Y}(u,v)dvdu \\ \Rightarrow \mathbb{P}(X \leq x) &= \int_{-\infty}^{x}\left(\int_{-\infty}^{\infty} f_{X,Y}(u,v)dv\right)du = \int_{-\infty}^{x} f_X(u)du. \end{aligned} \tag{13.3}$$

In (13.3), it is clear that the inner integral in the parentheses produces a non-negative measurable function of u. Thus, (13.3) asserts that the marginal CDF of X can be written as the integral of a non-negative measurable function, which can be identified as the marginal pdf f_X. Thus, X is continuous and f_X given by the inner integral in (13.3) is the pdf of X. A similar argument holds for the marginal pdf of Y.

13.3 Independence of Jointly Continous Random Variables

For any two random variables X and Y, they are said to be independent iff,

$$F_{X,Y}(x,y) = F_X(x)F_Y(y) \quad \forall x, y \in \mathbb{R}$$

Applying this definition, for the particular case of jointly continuous random variables,

$$\begin{aligned}\int_{-\infty}^{x}\int_{-\infty}^{y} f_{X,Y}(u,v)dvdu &= \left(\int_{-\infty}^{x} f_X(u)du\right)\left(\int_{-\infty}^{y} f_Y(v)dv\right) \\ &= \int_{-\infty}^{x}\int_{-\infty}^{y} f_X(u)f_Y(v)dvdu.\end{aligned}$$

Since the above equality holds for all $x, y \in \mathbb{R}$, the integrands must be equal almost every where, i.e.,

$$f_{X,Y}(x,y) = f_X(x)f_Y(y) \quad \forall\ x,y \in \mathbb{R}$$

except possibly on a subset of $\mathbb{R}^2$ of Lebesgue measure zero. Indeed, the above condition can be seen to be both necessary and sufficient for the independence of two jointly continuous random variables.

13.4 Conditional pdf for jointly continuous random variables

We would like to define the conditional cdf, $F_{X|Y}(x|y) \approx \mathbb{P}(X \leq x|Y = y)$. But the event, $\{Y = y\}$ has zero probability $\forall y$, when Y is continuous! To overcome this technical difficulty,we proceed by conditioning on a Y taking values in a small interval $(y, y + \epsilon)$, and then take the limit $\epsilon \downarrow 0$. More concretely, let us consider the following derivation, which motivates the definition of the conditional pdf for the jointly continuous case.

Informal Motivation
We can approximately define the conditional CDF of X, given that Y takes a value "close to y" as

$$\begin{aligned} F_{X|Y}(x|y) &= \mathbb{P}\left(X \leq x \mid y \leq Y \leq y+\epsilon\right) \quad \text{(for small } \epsilon) \\ &= \frac{\mathbb{P}\left(\{X \leq x\} \cap \{y \leq Y \leq y+\epsilon\}\right)}{\mathbb{P}\left(\{y \leq Y \leq y+\epsilon\}\right)} \\ &= \frac{F_{X,Y}(x, y+\epsilon) - F_{X,Y}(x,y)}{F_Y(y+\epsilon) - F_Y(y)} \\ &= \frac{\frac{F_{X,Y}(x,y+\epsilon)-F_{X,Y}(x,y)}{\epsilon}}{\frac{F_Y(y+\epsilon)-F_Y(y)}{\epsilon}}. \end{aligned}$$

As $\epsilon \to 0$, the RHS looks like, $\frac{\text{partial derivative of } F_{X,Y} \text{ w.r.t } y}{\text{derivate of } F_Y \text{ w.r.t } y}$.
This motivates the following definition for the conditional cdf and conditional pdf.

Definition 13.5 *a) The Conditional cdf of X given Y is defined as follows:*

$$F_{X|Y}(x|y) = \int_{-\infty}^{x} \frac{f_{X,Y}(u,y)}{f_Y(y)} du.$$

b) The Conditional pdf of X given Y is defined as follows:

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} \quad \text{for any } y \text{ such that } f_Y(y) > 0.$$

c) The Conditional probability of an event $A \in \mathcal{B}(\mathbb{R})$ given $Y = y$ is defined by

$$\begin{aligned} \mathbb{P}(X \in A \mid Y = y) &= \int_A f_{X|Y}(v|y) dv \\ &= \int_{-\infty}^{\infty} \mathbb{I}_A(v) f_{X|Y}(v|y) dv, \end{aligned}$$

where $\mathbb{I}_A(x)$ is indicator function for the event $\{x \in A\}$.

Example:
Let X, Y be jointly continuous with $f_{X,Y}(x,y) = 1$ in the region shown in Figure 13.3. Find all the marginals and conditional distributions.

One can easily verify that $f_{X,Y}(x,y)$ is a valid joint pdf by integrating it over the region.

$$\text{i.e. } \int_0^2 \int_0^1 f_{X,Y}(x,y) dx dy = 1.$$

The marginal pdf of Y can be calculated as follows. In the Figure 13.3, the equation of the line is $y = -2x+2$. So, for a given y, $f_{X,Y}(x,y)$ is non zero only in the range $x \in (0, 1-\frac{y}{2})$. This can be seen from the Figure 13.3.

Thus,

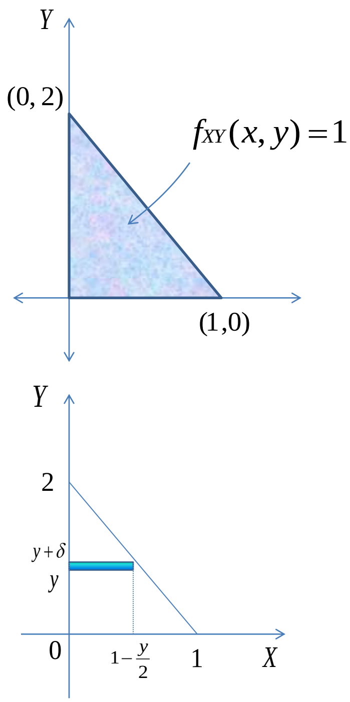

Figure 13.3:

$$f_Y(y) = \int_0^{1-\frac{y}{2}} 1dx = 1 - \tfrac{y}{2} \quad 0 \le y \le 2\ .$$

Similarly, the marginal pdf of X can be calculated as follows. From the line equation, $y = -2x + 2$, we can find that for a given x, $f_{X,Y}(x,y)$ is non zero only in the range $y \in (0, 2-2x)$. Thus,

$$f_X(x) = \int_0^{2-2x} 1dy = 2 - 2x \quad 0 \le x \le 1.$$

The marginals of Y and X have been plotted in Figure 13.4.

Now, lets find the conditional pdf, $f_{X|Y}(x|y)$. Here, we are computing a function of x for a given(fixed) y.

$$\begin{aligned} f_{X|Y}(x|y) &= \frac{f_{X,Y}(x,y)}{f_Y(y)} \\ &= \frac{1}{1-\frac{y}{2}} \\ &= \frac{2}{2-y} \qquad x \in \left(0, 1 - \frac{y}{2}\right). \end{aligned}$$

Though x does not appear in the expression, it appears in the constraint. So this is a conditionally uniform distribution in x i.e., given $\{Y = y\}$, X is uniformly distributed in $x \in \left(0, 1 - \frac{y}{2}\right)$ as $f_{X|Y}(x|y)$ is constant w.r.t x for the specified range.
Similarly, we can find,

$$f_{Y|X}(y|x) = \frac{1}{2-2x} \qquad y \in (0, 2-2x)\,.$$

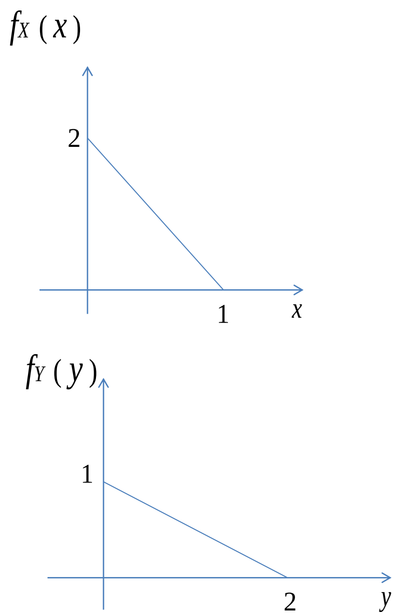

Figure 13.4: Marginal pdf of X and Y

13.5 Exercises

1. Two persons X and Y live in cities A and B but work in cities B and A respectively. Every morning they start for work at a uniformly random time between 9 am and 10 am independent of each other. Both of them travel at the same constant speed and it takes 20 minutes to reach the other city. What is the probability that X and Y meet each other on their way to work ?

2. Data is taken on the height and shoe size of a sample of MIT students. *Height* (X) is coded by 3 values: 1 (short), 2 (average), 3 (tall) and *Shoe size* (Y) is coded by 3 values 1 (small), 2 (average), 3 (large). The joint counts are given in the following table:

	X=1	X=2	X =3
Y=1	234	225	84
Y=2	180	453	161
Y=3	39	192	157

   (a) Find the joint and marginal pmf of X and Y.

   (b) Are X and Y independent? Discuss in detail.

3. John is vacationing in Monte Carlo. Each evening, the amount of money he takes to the casino is a random variable X with the pdf

$$f_X(x) = \begin{cases} Cx & 0 < x \leq 100 \\ 0 & \text{elsewhere.} \end{cases}$$

   At the end of each night, the amount Y he returns with is uniformly distributed between zero and twice the amount he came to casino with.

(a) Find the value of C.

(b) For a fixed $\alpha, 0 \leq \alpha \leq 100$, what is the conditional pdf of Y given $X = \alpha$?

(c) If John goes to the casino with α dollars, what is the probability he returns with more than α dollars ?

(d) Determine the joint pdf, $f_{X,Y}(x, y)$, of X and Y as well as the marginal pdf, $f_Y(y)$, of Y.

4. A rod is broken at two points that are chosen uniformly and independently at random. What is the probability that the three resulting pieces form a triangle?

5. Melvin Fooch, a student of probability theory, has found that the hours he spends working (W) and sleeping (S) in preparation for a final exam are random variables described by:

$$f_{W,S}(w, s) = \begin{cases} K, & 10 \leq w + s \leq 20 \text{ and } w \geq 0, s \geq 0 \\ 0, & \textit{elsewhere.} \end{cases}$$

What poor Melvin does not know, and even his best friends will not tell him, is that working only furthers his confusion and that his grade, G, can be described by $G = 2.5(S - W) + 50$.

(a) The instructor has decided to pass Melvin if, on the exam, he achieves $G \geq 75$. What is the probability that this will occur?

(b) Suppose Melvin got a grade greater than or equal to 75 on the exam. Determine the conditional probability that he spent less than one hour working in preparation for this exam.

(c) Are the random variables W and S independent? Justify.

6. Stations A and B are connected by two parallel message channels. A message from A to B is sent over both channels at the same time. Continuous random variables X and Y represent the message delays (in hours) over parallel channels I and II, respectively. These two random variables are independent, and both are uniformly distributed from 0 to 1 hours. A message is considered received as soon as it arrives on any one channel, and it is considered verified as soon as it has arrived over both channels.

(a) Determine the probability that a message is received within 15 minutes after it is sent.

(b) Determine the probability that the message is received but not verified within 15 minutes after it is sent.

(c) If the attendant at B goes home 15 minutes after the message is received, what is the probability that he is present when the message should be verified?

7. Random variables B and C are jointly uniform over a $2l \times 2l$ square centered at the origin, i.e., B and C have the following joint probability density function:

$$f_{B,C}(b, c) = \begin{cases} \frac{1}{4l^2}, & -l \leq b \leq l; -l \leq c \leq l \\ 0, & \textit{elsewhere.} \end{cases}$$

It is given that $l \geq 1$. Find the probability that the quadratic equation $x^2 + 2Bx + C = 0$ has real roots (Answer will be an expression involving l). What is the limit of this probability as $l \to \infty$?

8. (a) Consider four independent rolls of a 6-sided die. Let X be the number of 1's and let Y be the number of 2's obtained. What is the joint PMF of X and Y?

(b) Let X_1, X_2, X_3 be independent random variables, uniformly distributed on $[0, 1]$. Let Y be the median of X_1 , X_2, X_3 (that is the middle of the three values). Find the conditional CDF of X_1, given that $Y = 0.5$. Under this conditional distribution, is X_1 continuous? Discrete?