Two Dimensional Random Variables (Continued)

Definition 3.3

1. (X, Y) is a two-dimensional **discrete** random variable if the possible values of $(X(s), Y(s))$ are finite or countable infinite.

   i.e. the possible values of $(X(s), Y(s))$ may be represented as
   $$(x_i, y_j), i = 1, 2, 3, \cdots; j = 1, 2, 3, \cdots$$

2. (X, Y) is a two-dimensional **continuous** random variable if the possible values of $(X(s), Y(s))$ can assume all values in some region of the Euclidean plane $\mathbb{R}^2$.

ST2334 Probability and Statistics CYM 2-dim RVs & cond prob dist 3-6

To judge whether a two dimensional random vector (X, Y) is discrete or continuous, we can view X and Y separately.

✓ If both X and Y are discrete random variables, we say (X, Y) is a discrete random vector.

✓ Likewise, if both X and Y are continuous random variables, we say (X, Y) is a continuous random vector.

✓ Certainly, there are other cases. For example, X is discrete, but Y is continuous, or Y is neither a discrete nor a continuous random variable. But these are not the main focus of this module.

3.2.1 Joint Probability Function for Discrete RVs

Definition 3.4

- Let (X, Y) be a 2-dimensional **discrete** random variable defined on the sample space of an experiment. With each possible value (x_i, y_j), we associate a number $f_{X,Y}(x_i, y_j)$ representing $\Pr(X = x_i, Y = y_j)$ and satisfying the following conditions:

1. $f_{X,Y}(x_i, y_j) \geq 0$ for all $(x_i, y_j) \in R_{X,Y}$.
2. $\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} f_{X,Y}(x_i, y_j) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \Pr(X = x_i, Y = y_j) = 1$ (3.1)

Equation (3.1) on this page of the lecture slide essentially requires that the summation over all $f(x_i, y_j) > 0$ equals 1. It can be equivalently written as

$$\sum_{(x_i,y_j): f_{X,Y}(x_i,y_j)>0} f_{X,Y}(x_i, y_j) = 1.$$

Note that in this case, $f_{X,Y}(x_i, y_j)$ may not be defined for some x_i and y_j; see the distribution given on page 3-20. So, in this case, if you would like to add $i = 0, 1, 2, 3$ and $j = 0, 1, 2, 3$ freely, you need use 0 to replace those $f_{X,Y}(x, y)$ who does not have a point mass on (x, y).

Solution to Example 3 (Continued)

The above p.f. are given explicitly in the following table.

x	***y***				Row Total
	0	**1**	**2**	**3**	
0	0	3/84	6/84	1/84	10/84
1	4/84	24/84	12/84	0	40/84
2	12/84	18/84	0	0	30/84
3	4/84	0	0	0	4/84
Column Total	20/84	45/84	18/84	1/84	1

Joint pdf for Continuous RVs (Continued)

1. $f_{X,Y}(x,y) \geq 0$ for all $(x,y) \in R_{X,Y}$.
2. $$\iint_{(x,y)\in R_{X,Y}} f_{X,Y}(x,y)dx\,dy = 1$$

or

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)dx\,dy = 1.$$

ST2334 Probability and Statistics CYM 2-dim RVs & cond prob dist 3-22

✓ In most cases of this module, when we do the bivariate integration, the integration region is a rectangular; therefore, the variables x and y can be integrated separately; and the order of which is integrated first does not matter. See examples 3-24, 3-28, and 3-29.

✓ However, we need to bear in mind that there are cases under which the integration region is NOT a rectangular, so that x and y can not move freely for a unified expression of $f_{X,Y}(x,y)$. See the example given on pages 3-25, 3-26, and 3-27 of the lecture slides: the region is defined by straight lines such as a triangle or a trapezium.

✓ There are also even more difficult cases that the integration region is defined by more complicated shapes. Deep understanding of the multi-variate integration and extra caution might be needed.

Let's use an example to illustrate: find the value $c > 0$ (a constant not depending on x and

y), such that

$$f_{X,Y}(x,y) = \begin{cases} c & x^2+y^2 \le 1 \\ 0 & \text{otherwise} \end{cases}$$

is a joint pdf. Note that $R_{X,Y} = \left\{(x,y)\Big| x^2+y^2 \le 1\right\}$ in this example is a plate with radius 1 as shown in the figure below. Based on the slide, to make $f_{X,Y}(x,y)$ satisfy the criteria a joint pdf, we need

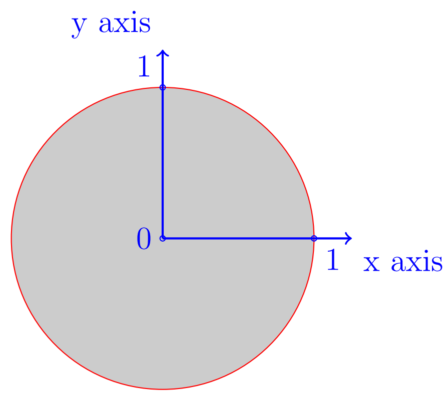
$$\begin{aligned} 1 &= \int\int_{x^2+y^2\le 1} c dxdy = c\int_{-1}^{1}\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} dydx = 2c\int_{-1}^{1}\sqrt{1-x^2}dx \quad \text{set } x = sin\theta \\ &= 2c\int_{-\pi/2}^{\pi/2}\cos^2(\theta)d\theta = c\pi, \end{aligned}$$

where the last "=" is because

$$\begin{aligned} \int_{-\pi/2}^{\pi/2}\cos^2(\theta)d\theta &= \int_{-\pi/2}^{\pi/2}\left\{1-\sin^2(\theta)\right\}d\theta = \pi - \int_{-\pi/2}^{\pi/2}\sin^2(\theta)d\theta \\ \int_{-\pi/2}^{\pi/2}\cos^2(\theta)d\theta &= \int_{-\pi/2}^{\pi/2}\cos(\theta)\cos(\theta)d\theta = \sin(\theta)\cos(\theta)\Big|_{-\pi/2}^{\pi/2} + \int_{-\pi/2}^{\pi/2}\sin^2(\theta)d\theta \\ &= \int_{-\pi/2}^{\pi/2}\sin^2(\theta)d\theta, \end{aligned}$$

where the second equation above used integration by part.

Note: $\int_{-\pi/2}^{\pi/2}\cos^2(\theta)d\theta$ **can also be evaluated based on the formula:** $\cos(2\theta) = 2\cos^2(\theta) - 1$. **Please try on your own to work it out.**



Note: when we doing the integration for a two dimensional function in a region which is not a rectangular, we need to be careful that x and y may not move freely! Based on mathematical theory, integrating which variable first won't change the outcome of the integration; however, a right choice of integration order may make the computation of the integration easier; read pages 3-25 to 3-26 carefully for such an example.

Marginal Distributions (Continued)

- For **discrete** case,

$$f_X(x) = \sum_y f_{X,Y}(x,y) \quad \text{and} \quad f_Y(y) = \sum_x f_{X,Y}(x,y)$$

- For **continuous** case,

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

and

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

ST2334 Probability and Statistics CYM 2-dim RVs & cond prob dist 3-31

The practical interpretation of the marginal distribution for X is: focusing on viewing the distribution of X by ignoring the presence of Y. **Note that**

★ **$f_X(x)$ should NOT involve y; and**

★ **it is a pdf/pmf; so it must have all the properties of a pdf/pmf.**

If (X, Y) is discrete, then the marginals are also discrete; likewise, if (X, Y) is continuous, the marginals are also continuous.

The meaning of the formulae for $f_X(x)$ is that "for each given x, integrate (or sum) over all the value of y such that $f_{X,Y}(x, y) > 0$." So, similar to the discussion of pages 4–6 above, we need to take care of the region of y for each x. Referring to the example given in page 5, to derive the marginal distribution for X, we need to compute for every given x,

$$f_X(x) = \int_{y: f_{X,Y}(x,y)>0} f_{X,Y}(x,y)dy.$$

On the other hand, we only need to consider $x \in [-1, 1]$, as if $x \notin [-1, 1]$, $f_{X,Y} = 0$. For each $x \in [-1, 1]$, the region for y should be $y \in [-\sqrt{1-x^2}, \sqrt{1-x^2}]$; therefore

$$f_X(x) = \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \frac{1}{\pi} dy = \frac{2}{\pi}\sqrt{1-x^2} \qquad \text{for } x \in [-1, 1].$$

Page 3-41

Conditional Distribution (Continued)

Definition 3.7 (Continued)

- Then **the conditional distribution of Y given that $X = x$** is given by

$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)}, \qquad \text{if } f_X(x) > 0,$$

for each x within the range of X.

✓ The condition distribution is similar in meaning to the condition probability. It is the distribution of the random variable of Y when the random X is fixed at a certain value x.

✓ It is important to take note that it is a distribution for y, so it must satisfies all the properties of a pdf/pmf in terms of the argument y for every x that it is defined.

✓ It may or may not be a function of x. But it is defined only when x satisfies $f_X(x) > 0$. If it does not depend on x, then we have X and Y independent.

✓ It is not a pdf/pmf for x. So there is NO requirement that $\int_{-\infty}^{\infty} f_{Y|X}(y|x)dx = 1$ when Y is continuous or $\sum_x f_{Y|X}(y|x) = 1$, when Y is discrete.

✓ Can you find $f_{Y|X}(y|x)$ for the example given on page 5?

Pages 3-36 & 3-47

Example 1 (Continued)

- $f_{X,Y}(x,y)$, $f_X(x)$ and $f_Y(y)$ are displayed in the following table

y	x						$f_Y(y)$
	0	1	2	3	4	5	
0	0	0.01	0.02	0.05	0.06	0.08	0.22
1	0.01	0.03	0.04	0.05	0.05	0.07	0.25
2	0.02	0.03	0.05	0.06	0.06	0.07	0.29
3	0.02	0.04	0.03	0.04	0.06	0.05	0.24
$f_X(x)$	0.05	0.11	0.14	0.20	0.23	0.27	1

Example 1 (Continued)

Outcome	HHH	THH	HTH	HHT	TTH	THT	HTT	TTT
(x,y)	(1,3)	(1,2)	(1,2)	(0,2)	(1,1)	(0,1)	(0,1)	(0,0)
$f_{X,Y}(x,y)$	1/8	1/8	1/8	1/8	1/8	1/8	1/8	1/8

- The joint probability distribution of (X, Y) is given in the following table:

x	y				$f_X(x)$
	0	1	2	3	
0	1/8	1/4	1/8	0	1/2
1	0	1/8	1/4	1/8	1/2
$f_Y(y)$	1/8	3/8	3/8	1/8	1

For a discrete random vector (X,Y). The two-dimensional tables as shown in these slides are particularly useful to help us understand the joint, marginal, and the conditional distributions.