VII. Joint Distribution

Instructor: **Yanlin Qi**

Institute of Transportation Studies

Department of Statistics

University of California, Davis

Today

- Start with Joint Distribution Functions
  - Marginal distributions
  - Conditional distributions

Joint Probability Distributions

Joint distributions

Introduction

So far we have considered the distribution of only a single random variable, discrete or continuous.

When two or more random variables are defined on the same sample space, we can describe their joint distribution.

For two random variables X and Y, their joint distribution is denoted as $f_{XY}(x,y)$.

Joint PMFs for Discrete RVs

Joint distributions

- Example 0.1 (Toss two fair dice). Let X denote their sum and Y the absolute value of their difference, which are two discrete random variables.
  - First die → m; Second die → n
  - $X = x = m + n;\ Y = y\ =\ |m - n|$
  - We can find their individual distributions easily:

x	2	3	$\cdots$	12
$P(X = x)$	$\frac{1}{36}$	$\frac{2}{36}$	$\cdots$	$\frac{1}{36}$

$$X = 2\ \rightarrow \{(1,1)\}$$
$$X = 3\ \rightarrow \{(1,2), (2,1)\}$$
$$\cdots$$

y	0	1	$\cdots$	5
$P(Y = y)$	$\frac{6}{36}$	$\frac{10}{36}$	$\cdots$	$\frac{2}{36}$

$$Y = 0$$
$$\rightarrow \{(1,1), (2,2), (3,3), (4,4), (5,5), (6,6)\}$$
$$Y = 1$$
$$\rightarrow \left\{\begin{matrix}(1,2), (2,1), (2,3), (3,2), (3,4), (4,3) \\ (4,5), (5,4), (5,6), (6,5)\end{matrix}\right\}$$
$$\cdots$$

- Now consider X, Y together as a pair (X, Y), or a vectored-valued function.
- Questions:
  - Can (X, Y) attain all the $66 = 11 \times 6$ pairs?

$$\{(x, y) \mid 2 \le x \le 12,\ 0 \le y \le 5\}$$

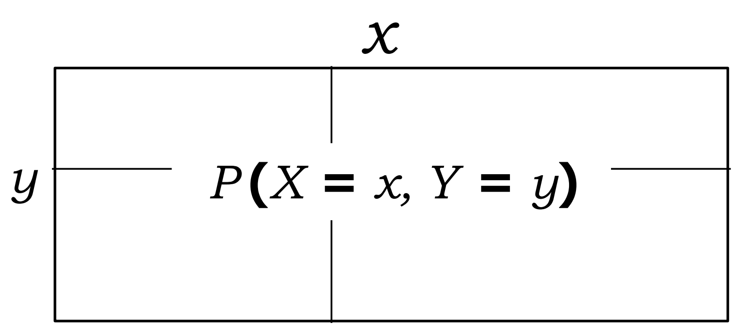
    If not all, identify the subset of feasible pairs.
  - What are the corresponding probabilities for (X, Y) to take those (feasible) pairs as values?

- Answering the above two questions together is equivalent to specifying the joint probability distribution of (X, Y) in terms of range and frequency.

Joint Probability Functions

- Let X and Y be two discrete random variables on the same sample space. We define their joint PMF as a function $f : \mathrm{R}^2 \rightarrow \mathrm{R}$ with

$$f(x,y) = \begin{cases} P(X = x, Y = y), & \text{for all feasible pairs } (x,y) \\ 0, & \text{otherwise} \end{cases}$$



- **Example** Find the joint PMF of *X*, *Y* in the previous example.
  - First die → m; Second die → n
  - $X = x = m + n;\ Y = y\ =\ |m - n|$

$y \backslash x$	2	3	4	5	6	7	8	9	10	11	12
0											
1											
2											
3											
4											
5											

$y \backslash x$	2	3	4	5	6	7	8	9	10	11	12
0	$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$
1											
2											
3											
4											
5											

$y \backslash x$	2	3	4	5	6	7	8	9	10	11	12
0	$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$
1		$\frac{2}{36}$		$\frac{2}{36}$		$\frac{2}{36}$		$\frac{2}{36}$		$\frac{2}{36}$	
2											
3											
4											
5											

$y \backslash x$	2	3	4	5	6	7	8	9	10	11	12
0	$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$
1		$\frac{2}{36}$		$\frac{2}{36}$		$\frac{2}{36}$		$\frac{2}{36}$		$\frac{2}{36}$	
2			$\frac{2}{36}$		$\frac{2}{36}$		$\frac{2}{36}$		$\frac{2}{36}$		
3				$\frac{2}{36}$		$\frac{2}{36}$		$\frac{2}{36}$			
4					$\frac{2}{36}$		$\frac{2}{36}$				
5						$\frac{2}{36}$					

Properties of Joint Probability Functions

- Any joint PMF $f(x, y) : \mathrm{R}^2 \to \mathrm{R}$, must satisfy (and vice versa)

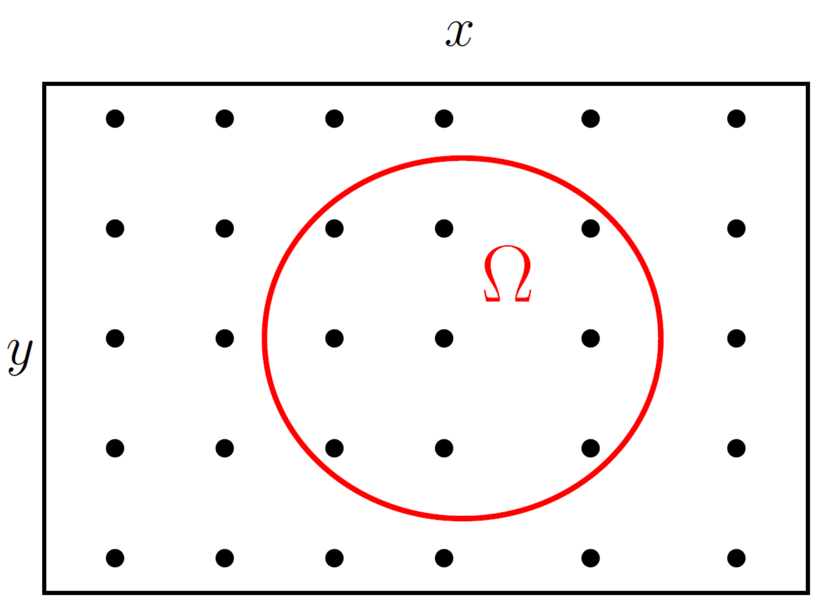
1. $f(x, y) \geq \mathbf{0}$ for all $x, y \in \mathbf{R}$

2. $$\sum_x \sum_y f_{XY}(x, y) = 1$$

3. $$f_{XY}(x, y) = P(X = x, Y = y)$$

❖ *Theorem* 0.1. Let X, Y be two discrete random variables with joint PMF $f(x, y)$. Then for any region $\Omega \subset \mathbb{R}^2$,

$$P((X, Y) \in \Omega) = \sum_{(x,y)\in\Omega} f(x, y)$$



- **Example** (Toss 2 fair dice, cont'd). Find the following probabilities:
- $P\,(X \;\le\; 4, Y \;\le\; 2)$

$$= \frac{6}{36}$$

(sum of top-left 3 × 3 block of joint PMF)

y \ x	2	3	4	5	6	7	8	9	10	11	12
0	$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$
1		$\frac{2}{36}$		$\frac{2}{36}$		$\frac{2}{36}$		$\frac{2}{36}$		$\frac{2}{36}$	
2			$\frac{2}{36}$		$\frac{2}{36}$		$\frac{2}{36}$		$\frac{2}{36}$		
3				$\frac{2}{36}$		$\frac{2}{36}$		$\frac{2}{36}$			
4					$\frac{2}{36}$		$\frac{2}{36}$				
5						$\frac{2}{36}$					

Example 0.3 (Toss 2 fair dice, cont'd). Find the following probabilities:

- $P(X \geq 11,\, Y \leq 2)$

$$= \frac{3}{36} \text{ (sum of top-right } 3 \times 2 \text{ block of joint PMF table)}$$

y \ x	2	3	4	5	6	7	8	9	10	11	12
0	$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$
1		$\frac{2}{36}$		$\frac{2}{36}$		$\frac{2}{36}$		$\frac{2}{36}$		$\frac{2}{36}$	
2			$\frac{2}{36}$		$\frac{2}{36}$		$\frac{2}{36}$		$\frac{2}{36}$		
3				$\frac{2}{36}$		$\frac{2}{36}$		$\frac{2}{36}$			
4					$\frac{2}{36}$		$\frac{2}{36}$				
5						$\frac{2}{36}$					

Example 0.3 (Toss 2 fair dice, cont'd). Find the following probabilities:

- $P(X \leq 5)$

$$= \frac{10}{36} \text{ (sum of first four columns of joint PMF table)}$$

$y \backslash x$	2	3	4	5	6	7	8	9	10	11	12
0	$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$
1		$\frac{2}{36}$		$\frac{2}{36}$		$\frac{2}{36}$		$\frac{2}{36}$		$\frac{2}{36}$	
2			$\frac{2}{36}$		$\frac{2}{36}$		$\frac{2}{36}$		$\frac{2}{36}$		
3				$\frac{2}{36}$		$\frac{2}{36}$		$\frac{2}{36}$			
4					$\frac{2}{36}$		$\frac{2}{36}$				
5						$\frac{2}{36}$					

Example 0.3 (Toss 2 fair dice, cont'd). Find the following probabilities:

- $P(Y \leq 1)$

$$= \frac{16}{36} \text{ (sum of first two rows of joint PMF table)}$$

$y \backslash x$	2	3	4	5	6	7	8	9	10	11	12
0	$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$
1		$\frac{2}{36}$		$\frac{2}{36}$		$\frac{2}{36}$		$\frac{2}{36}$		$\frac{2}{36}$	
2			$\frac{2}{36}$		$\frac{2}{36}$		$\frac{2}{36}$		$\frac{2}{36}$		
3				$\frac{2}{36}$		$\frac{2}{36}$		$\frac{2}{36}$			
4					$\frac{2}{36}$		$\frac{2}{36}$				
5						$\frac{2}{36}$					

From joint to marginal

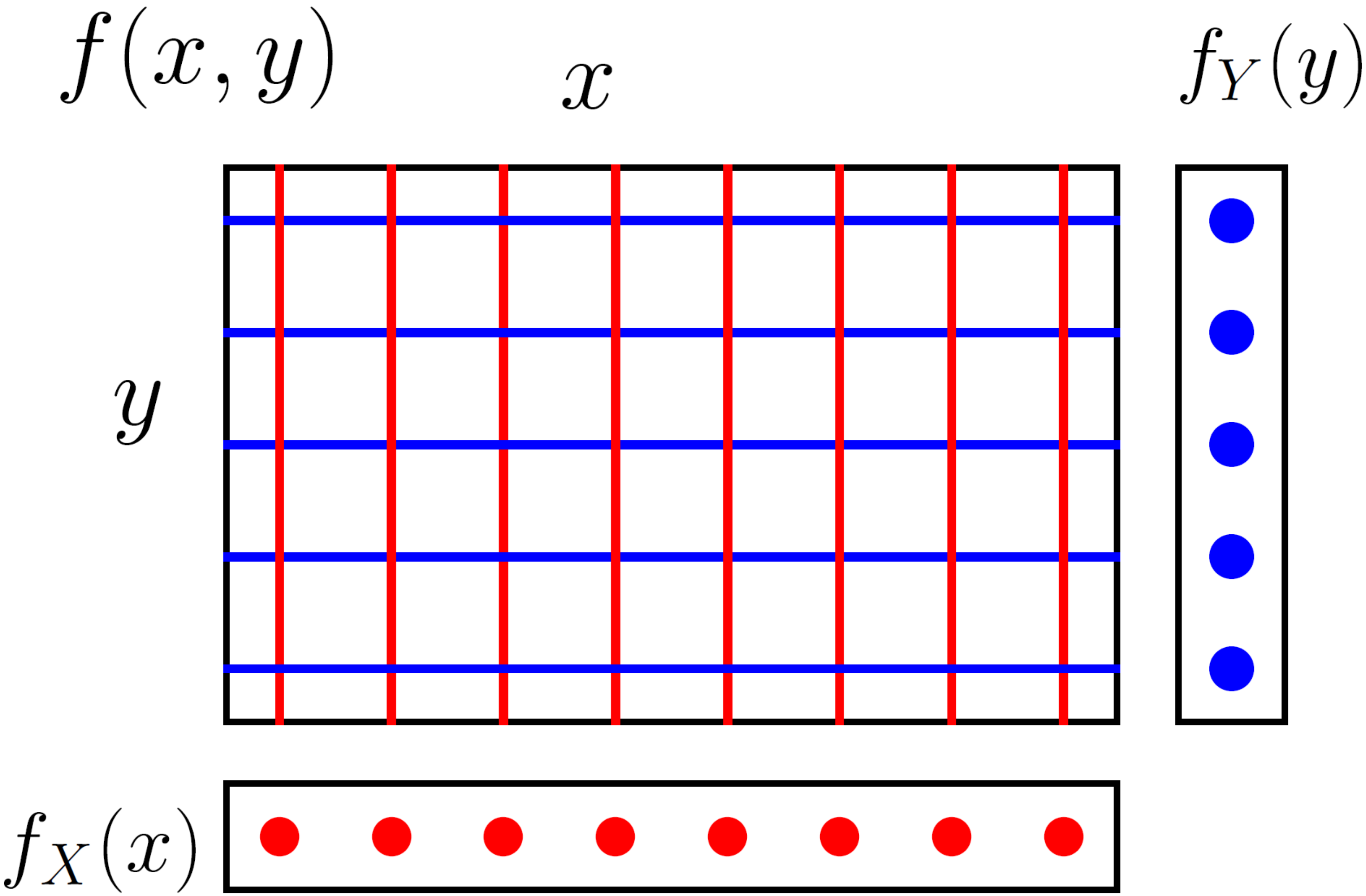
- For any two discrete random variables X, Y that have a joint distribution, we call their individual PMFs $f_X(x), f_Y(y)$ the **marginal PMFs.**

Proposition. Let $f(x, y)$ be the joint PMF for X, Y . Then

$$f_X(x) = \sum_y f(x, y), \quad \text{and} \quad f_Y(y) = \sum_x f(x, y).$$

Proof. This is just the Law of Total Probability:

$$\underbrace{P(X = x)}_{f_X(x)} = \sum_y \underbrace{P(X = x, Y = y)}_{f(x,y)}.$$


$f(x, y)$
x
$f_Y(y)$
y
$f_X(x)$

y \ x	2	3	4	5	6	7	8	9	10	11	12	$f_Y(y$
0	$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$	$\frac{6}{36}$
1		$\frac{2}{36}$		$\frac{2}{36}$		$\frac{2}{36}$		$\frac{2}{36}$		$\frac{2}{36}$		$\frac{10}{36}$
2			$\frac{2}{36}$		$\frac{2}{36}$		$\frac{2}{36}$		$\frac{2}{36}$			$\frac{8}{36}$
3				$\frac{2}{36}$		$\frac{2}{36}$		$\frac{2}{36}$				$\frac{6}{36}$
4					$\frac{2}{36}$		$\frac{2}{36}$					$\frac{4}{36}$
5						$\frac{2}{36}$						$\frac{2}{36}$
$f_X(x)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	

Conditional PMFs

- Consider the following question:
- **Example** (Toss 2 fair dice)**.** Suppose we are told that the sum is $X = 6$. What is the (conditional) distribution of Y?

y \ x	2	3	4	5	6	7	8	9	10	11	12	$f_Y(y$
0	$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$	$\frac{6}{36}$
1		$\frac{2}{36}$		$\frac{2}{36}$		$\frac{2}{36}$		$\frac{2}{36}$		$\frac{2}{36}$		$\frac{10}{36}$
2			$\frac{2}{36}$		$\frac{2}{36}$		$\frac{2}{36}$		$\frac{2}{36}$			$\frac{8}{36}$
3				$\frac{2}{36}$		$\frac{2}{36}$		$\frac{2}{36}$				$\frac{6}{36}$
4					$\frac{2}{36}$		$\frac{2}{36}$					$\frac{4}{36}$
5						$\frac{2}{36}$						$\frac{2}{36}$
$f_X(x)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	

$$P(Y = y|X = 6) = \frac{P(X = 6, Y = y)}{P(X = 6)}$$

$$P(Y = 0|X = 6) = \frac{P(X = 6, Y = 0)}{P(X = 6)} = \frac{\frac{1}{36}}{\frac{5}{36}} = \frac{1}{5}$$

$$P(Y = 2|X = 6) = \frac{P(X = 6, Y = 2)}{P(X = 6)} = \frac{\frac{2}{36}}{\frac{5}{36}} = \frac{2}{5}$$

$$P(Y = 4|X = 6) = \frac{P(X = 6, Y = 4)}{P(X = 6)} = \frac{\frac{2}{36}}{\frac{5}{36}} = \frac{2}{5}$$

Answer

y	0	2	4
$P(Y = y \mid X = 6)$	$\frac{1}{5}$	$\frac{2}{5}$	$\frac{2}{5}$

Definition

- Let X, Y be two discrete random variables with joint PMF $f(x, y)$. The conditional PMF of Y given $X = x$ (with $f_X(x) \neq 0$) is defined as

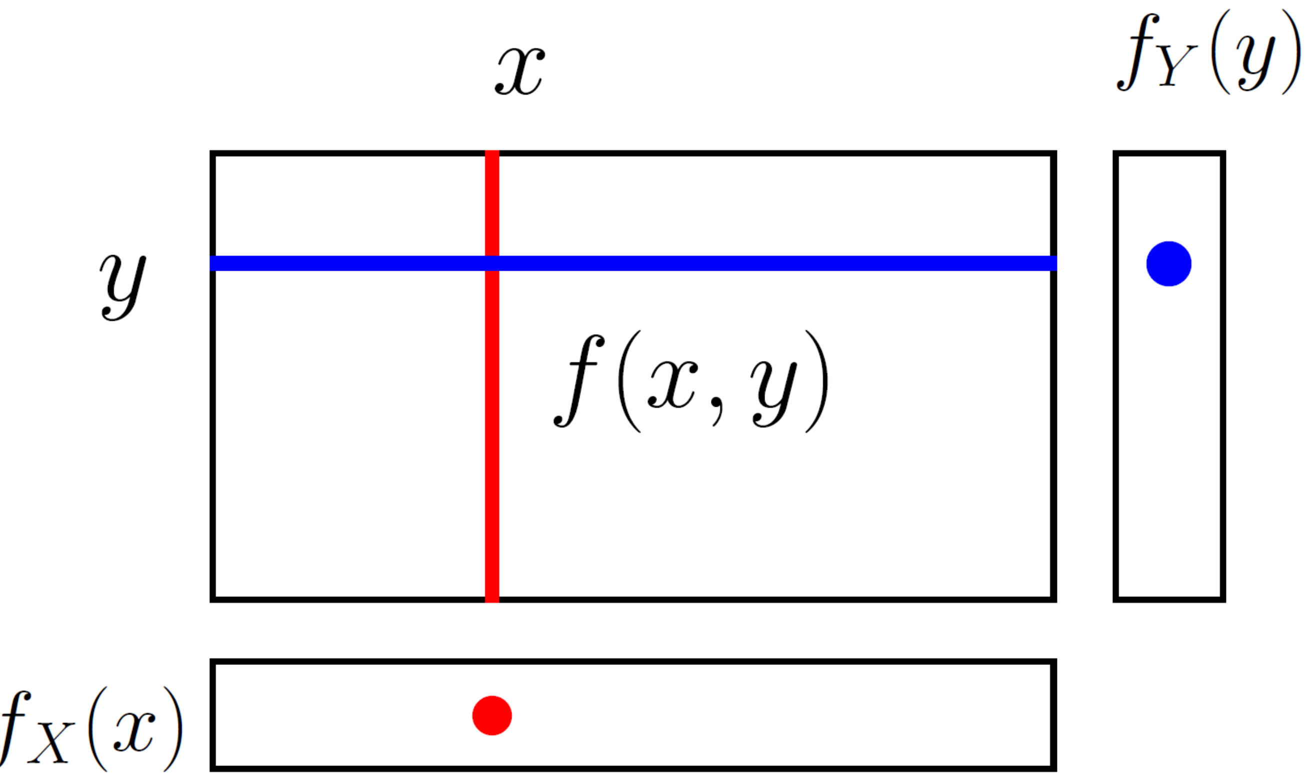
$$f(\underbrace{y}_{\text{variable}} \mid \underbrace{x}_{\text{fixed}}) = \frac{f(x,y)}{f_X(x)}, \quad \text{for all feasible } y$$

Remarks:

- (1) This definition is just based on the conditional probability of events:

$$P(Y = y \mid X = x) = \frac{P(X = x, Y = y)}{P(X = x)}.$$

- (2) For each fixed value x of X, there is a separate conditional distribution for Y at x (x may be regarded as a location parameter).



$$f(x \mid y) = \frac{f(x,y)}{f_Y(y)}$$

$$f(y \mid x) = \frac{f(x,y)}{f_X(x)}$$

Find the following conditional distributions:

Table 1: Conditional pmfs of Y given $X = x$

y \ x	2	3	4	5	6	7	8	9	10	11	12
0	1		$\frac{1}{3}$		$\frac{1}{5}$		$\frac{1}{5}$		$\frac{1}{3}$		1
1		1		$\frac{1}{2}$		$\frac{1}{3}$		$\frac{1}{2}$		1	
2			$\frac{2}{3}$		$\frac{2}{5}$		$\frac{2}{5}$		$\frac{2}{3}$		
3				$\frac{1}{2}$		$\frac{1}{3}$		$\frac{1}{2}$			
4					$\frac{2}{5}$		$\frac{2}{5}$				
5						$\frac{1}{3}$					

- Y given $X = 4$

y	0	2
$f(y \mid x = 4)$	$\frac{1}{3}$	$\frac{2}{3}$

$$f(y \mid x) = \frac{f(x,y)}{f_X(x)}$$

Table 2: Conditional pmfs of X given $Y = y$

$y \backslash x$	2	3	4	5	6	7	8	9	10	11	12
0	$\frac{1}{6}$		$\frac{1}{6}$		$\frac{1}{6}$		$\frac{1}{6}$		$\frac{1}{6}$		$\frac{1}{6}$
1		$\frac{1}{5}$		$\frac{1}{5}$		$\frac{1}{5}$		$\frac{1}{5}$		$\frac{1}{5}$	
2			$\frac{1}{4}$		$\frac{1}{4}$		$\frac{1}{4}$		$\frac{1}{4}$		
3				$\frac{1}{3}$		$\frac{1}{3}$		$\frac{1}{3}$			
4					$\frac{1}{2}$		$\frac{1}{2}$				
5						1					

$$f(x \mid y) = \frac{f(x,y)}{f_Y(y)}$$

- X given $Y = 3$:

x	5	7	9
$f(x \mid y = 3)$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

Independence

- Two discrete random variables X, Y are independent if

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

- *Remark*. This is equivalent to

$$P(X=x, Y=y) = P(X=x)P(Y=Y), \quad \text{for all } x, y.$$

- Two discrete random variables X, Y are independent if all conditional distributions of Y are identical to its marginal distribution:

$$f(y \mid x) = f_Y(y), \quad \text{for all } x, y$$

- **Example** (Toss 2 fair dice)**.** Determine if X (sum) and Y (absolute difference) are independent.

$y \backslash x$	2	3	4	5	6	7	8	9	10	11	12	$f_Y(y)$
0	$\boxed{\frac{1}{36}}$		$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$		$\frac{1}{36}$	$\frac{6}{36}$
1		$\frac{2}{36}$		$\frac{2}{36}$		$\frac{2}{36}$		$\frac{2}{36}$		$\frac{2}{36}$		$\frac{10}{36}$
2			$\frac{2}{36}$		$\frac{2}{36}$		$\frac{2}{36}$		$\frac{2}{36}$			$\frac{8}{36}$
3				$\frac{2}{36}$		$\frac{2}{36}$		$\frac{2}{36}$				$\frac{6}{36}$
4					$\frac{2}{36}$		$\frac{2}{36}$					$\frac{4}{36}$
5						$\frac{2}{36}$						$\frac{2}{36}$
$f_X(x)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	

$$f(x,y) \neq f_X(x) f_Y(y) \qquad f(y \mid x) \neq f_Y(y)$$

Example Are the random variables X, Y independent?

$y \backslash x$	0	1	2	$f_Y(y)$
-1	$\frac{1}{6}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{3}$
1	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{2}{3}$
$f_X(x)$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$	

$$P(X = x, Y = y) = P(X = x)P(Y = Y), \quad \text{for all } x, y.$$

Mean and Variance

- Mean E(X)
- $E[X] = \mu_X = \sum_X x f_X(x)$ (Equation for Single RV X)

$$= \sum_X x\left(\sum_Y f_{XY}(x,y)\right)$$

$$= \sum_X \sum_Y x f_{XY}(x,y)$$

  - $f_X(x) = \sum_Y f_{XY}(x,y)$ Marginal PMF of X

Mean E(Y)

$$E[Y] = \mu_Y = = \sum_X \sum_Y y f_{XY}(x,y)$$

- Step-by-step calculation:
  - First, find the joint PMF $f_{XY}(x,y)$
  - Sum over all y to get the marginal PMF $f_X(x)$
  - Multiply each x by its corresponding $f_X(x)$ and sum over all x to get $E[X]$

- **Expected value of** g(X, Y)
- Consider the following question:
  - Given two discrete random variables X, Y with a joint distribution, what are the expected values of random variables like $X + Y, XY, |X - Y|, (X - Y)^2$?

- **Theorem**
  - Let X, Y be two discrete random variables with a joint PMF $f(x, y)$.
  - Then for any function $g(X, Y)$,

$$\mathrm{E}(g(X, Y)) = \sum_x \sum_y g(x, y) f(x, y).$$

Variance V(X):

$$V(X) = \sigma_X^2 = E[(x-\mu_X)^2] = \sum_X (x-\mu_X)^2 f_X(x)$$

(Equation for Single RV *X*)

$$= \sum_X (x-\mu_X)^2 \left(\sum_Y f_{XY}(x,y)\right)$$

$$= \sum_X \sum_Y (x-\mu_X)^2 f_{XY}(x,y)$$

- Imagine $g(x,y) = (x - \mu_X + 0 \cdot y)^2$

Variance V(Y)

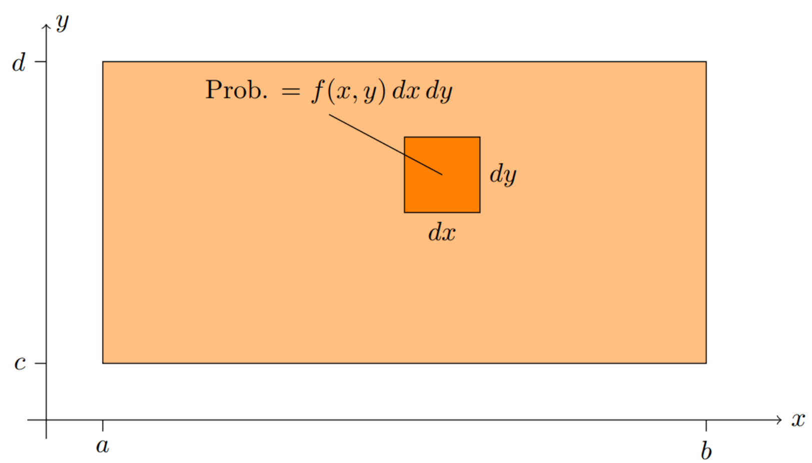
- $V(Y) = \sigma_Y^2 = E[(y-\mu_Y)^2] = \sum_Y \sum_X (y-\mu_Y)^2 f_{XY}(x,y)$

Continuous RVs

- **Definition** The joint PDF $f(x, y)$ describes the likelihood of the random variables X and Y taking on specific values x and y.

- Continuous joint distributions (Joint PDFs)
  - X takes values in [a, b]
  - Y takes values in [c, d]
  - (X, Y) takes values in [a, b] × [c, d]
  - Joint PDF: f(x, y) The value of f(x,y) at any point (x,y) gives the density of the probability at that point

$f(x, y)dxdy$ is the probability of being in the small square.



Joint Probability of Continuous RVS

- The **joint PDF** of the continuous random variables X and Y, is denoted as $f_{XY}(x, y)$ and satisfies

1. $f_{XY}(x, y) \geq 0 \quad \text{for all } x, y$

2. $\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} f_{XY}(x, y)dxdy = 1$

3. For any given area $X \in [a, b]$ and $Y \in [c, d]$:

$$P(a \leq x \leq b, c \leq y \leq d)$$
$$= \int_a^b \int_c^d f_{XY}(u, v)dudv$$

Marginal PDFs

- Marginal PDF of X

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

- Marginal PDF of Y

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

Joint CDF

- If X and Y have a joint PDF $f_{XY}(x, y)$, the joint CDF $F_{XY}(x, y)$ can be obtained by integrating the joint PDF:

$$F_{XY}(x, y) = P(X \leq x, Y \leq y)$$
$$= \int_{-\infty}^{x} \int_{-\infty}^{y} f_{XY}(u, v) du dv$$

Importance of the joint CDF

1. Understanding the dependency structure between two random variables.
2. Calculating the probability that two random variables fall within a specified range.

Conditional Distributions (Consistent with joint PMFs)

- Conditional PDF of Y given $X = x$:

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

- Conditional PDF of X given $Y = y$:

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

- **Independence** (Consistent with joint PMFs)
  - X and Y are independent if the joint PDFs into the product of the marginal PDFs

$$f(x,y) = f_X(x) f_Y(y)$$

Knowledge of one variable does not provide any information about the other variable

- Suppose X and Y are random variables.
- The pdf is $\frac{3}{2}(x^2 + y^2)$, X ∈ [0, 1] and Y ∈ [0, 1].
  1. Show f (x, y) is a valid pdf.
  2. Visualize the event A = ‘X > 0.3 and Y > 0.5’.
  3. Find its probability.
  4. Find F(x, y).
  5. Find the marginal pdf $f_X(x)$.