# STA2001 Probability and Statistics (I)

## Lecture 14

Tianshi Chen

The Chinese University of Hong Kong, Shenzhen

# Chapter 4. Bivariate Distribution

## Section 4.1 Bivariate Distribution of Discrete Type

# Motivation

Very often, we are interested to study two random experiments jointly, each of whose outcome is a scalar, or a random experiment whose outcome is a pair of two scalars.

1. observe college students to obtain information such as height $x$ and weight $y$.
2. observe high school students to obtain information such as rank $x$ and score of college entrance examination $y$.

# Motivation

Very often, we are interested to study two random experiments jointly, each of whose outcome is a scalar, or a random experiment whose outcome is a pair of two scalars.

1. observe college students to obtain information such as height $x$ and weight $y$.
2. observe high school students to obtain information such as rank $x$ and score of college entrance examination $y$.

- a random experiment whose outcome is a scalar, $\rightarrow$ univariate RV
- two random experiments jointly each of whose outcome is a scalar, or a random experiment whose outcome is a pair of two scalars, $\rightarrow$ bivariate RV

# Bivariate RV

## Definition

Let $(X, Y)$ be a pair of RVs with their range denoted by $\overline{S} \subseteq R^2$. Then $(X, Y)$ or $X$ and $Y$ is said to be a bivariate RV. If $\overline{S}$ is finite or countably infinite, then $(X, Y)$ is said to be a discrete bivariate RV.

Moreover, let $\overline{S_X} \subseteq R$ and $\overline{S_Y} \subseteq R$ denote the range of $X$ and $Y$, respectively.

$$
\begin{aligned}
\overline{S} &= \{\text{all possible values of } (X, Y)\} \\
\overline{S_X} &= \{\text{all possible values of } X\} = \left\{x | (x, y) \in \overline{S}\right\} \\
\overline{S_Y} &= \{\text{all possible values of } Y\} = \left\{y | (x, y) \in \overline{S}\right\}
\end{aligned}
$$

Then, it holds that

$$
\overline{S} \subseteq \overline{S_X} \times \overline{S_Y} = \left\{(x, y) | x \in \overline{S_X}, y \in \overline{S_Y}\right\}
$$

# Example 1, Page 134

Roll a pair of 4-sided fair dice and let $X$ denote the smaller and $Y$ the larger outcome of the pair of dice. For instance , if the outcome is (3,2) or (2,3), then $X = 2$, $Y = 3$.

# Example 1, Page 134

Roll a pair of 4-sided fair dice and let $X$ denote the smaller and $Y$ the larger outcome of the pair of dice. For instance , if the outcome is (3,2) or (2,3), then $X = 2$, $Y = 3$.

Sample space $\overline{S} \subseteq \overline{S_X} \times \overline{S_Y}$:

# Example 1, Page 134

Roll a pair of 4-sided fair dice and let $X$ denote the smaller and $Y$ the larger outcome of the pair of dice. For instance , if the outcome is (3,2) or (2,3), then $X = 2$, $Y = 3$.

Sample space $\overline{S} \subseteq \overline{S_X} \times \overline{S_Y}$:

$$\overline{S}_X = \{1,2,3,4\}, \overline{S}_Y = \{1,2,3,4\}$$

$$\overline{S} = \left\{ \begin{array}{cccc} (1,1), & (1,2), & (1,3), & (1,4), \\ & (2,2), & (2,3), & (2,4), \\ & & (3,3), & (3,4), \\ & & & (4,4) \end{array} \right\}$$

# Joint pmf

**Definition**

The function $f(x, y) : \overline{S} \to (0, 1]$ is called the joint probability mass function (joint pmf) of $X$ and $Y$ or $(X, Y)$, if

1. $f(x, y) > 0$ for $(x, y) \in \overline{S}$,
2. $\sum_{(x,y)\in\overline{S}} f(x, y) = 1$,
3. For $A \subseteq \overline{S}$,

$$P[(X, Y) \in A] \stackrel{\Delta}{=} P(\{(X, Y) \in A\}) = \sum_{(x,y)\in A} f(x, y)$$

which defines the probability function for a set $A$. In particular, taking $A = \{(x, y)\}$ yields the probability of $X = x$ and $Y = y$, i.e.,

$$P(X = x, Y = y) = f(x, y)$$

# Example 1, Page 134

Question:

$$P(X = 2, Y = 3) = ?,\ \ P(X = 2, Y = 2) = ?$$

# Example 1, Page 134

Question:

$$P(X = 2, Y = 3) = ?,\ \ P(X = 2, Y = 2) = ?$$

$$P(X = 2, Y = 3) = \frac{1}{16} + \frac{1}{16} = \frac{2}{16}$$

$$P(X = 2, Y = 3) = \frac{1}{16}$$

Question: What is the joint pmf $f(x, y)$?

# Example 1, Page 134

Question:

$$P(X=2, Y=3)=?, \ P(X=2, Y=2)=?$$

$$P(X=2, Y=3)=\frac{1}{16}+\frac{1}{16}=\frac{2}{16}$$

$$P(X=2, Y=3)=\frac{1}{16}$$

Question: What is the joint pmf $f(x, y)$?

$$\overline{S}=\left\{\begin{array}{cccc} (1,1) & (1,2) & (1,3) & (1,4) \\ & (2,2) & (2,3) & (2,4) \\ & & (3,3) & (3,4) \\ & & & (4,4) \end{array}\right\}$$

$$f(x, y)=\begin{cases}\frac{2}{16}, & 1 \leq x<y \leq 4 \\ \frac{1}{16}, & 1 \leq x=y \leq 4.\end{cases}$$

# Marginal pmf

**Definition**

Let $(X, Y)$ be a bivariate RV or $X$ and $Y$ be two RVs and have the joint pmf $f(x, y) : \overline{S} \to (0, 1]$. Sometimes, we are interested in the pmf of $X$ or $Y$ alone, which is called the marginal pmf of $X$ or $Y$ and described by

For $x \in \overline{S_X}$,

$$\begin{aligned} f_X(x) = P(X = x) &\overset{\Delta}{=} P\left(\{X = x, Y \in \overline{S_Y}(x)\}\right) \\ &= \sum_{y \in \overline{S_Y}(x)} f(x, y) \end{aligned}$$

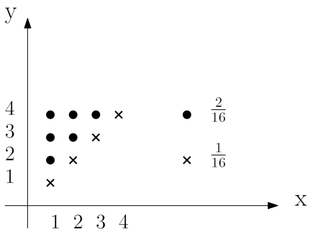
where

$$\overline{S_Y}(x) = \{y | (x, y) \in \overline{S}\} \text{ for the given } x \in \overline{S_X}.$$

# Example 1 [Continued]

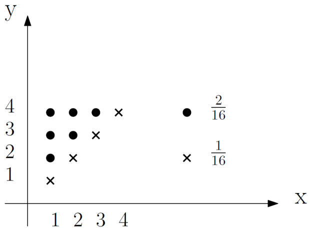
$$f(x,y) = \begin{cases} \frac{2}{16}, & 1 \le x < y \le 4 \\ \frac{1}{16}, & 1 \le x = y \le 4. \end{cases}$$

What is $f_X(x), f_Y(y)$?



# Example 1 [Continued]

$$f(x, y) = \begin{cases} \frac{2}{16}, & 1 \le x < y \le 4 \\ \frac{1}{16}, & 1 \le x = y \le 4. \end{cases}$$




What is $f_X(x), f_Y(y)$?

First, $\overline{S_X} = \overline{S_Y} = \{1, 2, 3, 4\}$.

$$f_X(x) = \sum_{y \in \overline{S_Y}(x)} f(x, y), x \in \overline{S_X} = \{1, 2, 3, 4\}$$

$$\Longrightarrow f_X(1) = \frac{7}{16}, \quad f_X(2) = \frac{5}{16}, \quad f_X(3) = \frac{3}{16}, \quad f_X(4) = \frac{1}{16}$$

# Marginal pmf

**Definition**

Let $(X, Y)$ be a bivariate RV or $X$ and $Y$ be two RVs and have the joint pmf $f(x, y) : \overline{S} \to (0, 1]$. Sometimes, we are interested in the pmf of $X$ or $Y$ alone, which is called the marginal pmf of $X$ or $Y$ and described by

For $y \in \overline{S_Y}$,

$$\begin{aligned} f_Y(y) = P(Y = y) &\overset{\Delta}{=} P\left(\{X \in \overline{S_X}(y), Y = y\}\right) \\ &= \sum_{x \in \overline{S_X}(y)} f(x, y) \end{aligned}$$

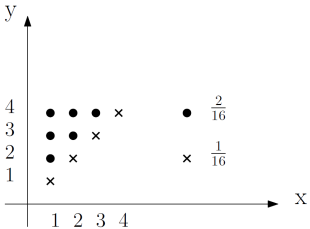
where

$$\overline{S_X}(y) = \left\{x | (x, y) \in \overline{S}\right\} \text{ for the given } y \in \overline{S_Y}.$$

## Example 1 [Continued]

$$f(x,y) = \begin{cases} \frac{2}{16}, & 1 \le x < y \le 4 \\ \frac{1}{16}, & 1 \le x = y \le 4. \end{cases}$$

What is $f_X(x), f_Y(y)$?




First, $\overline{S_X} = \overline{S_Y} = \{1,2,3,4\}$.

$$f_X(x) = \sum_{y \in \overline{S_Y}(x)} f(x,y), x \in \overline{S_X} = \{1,2,3,4\}$$

$$\Longrightarrow f_X(1) = \frac{7}{16}, \quad f_X(2) = \frac{5}{16}, \quad f_X(3) = \frac{3}{16}, \quad f_X(4) = \frac{1}{16}$$

$$f_Y(y) = \sum_{x \in \overline{S_X}(y)} f(x,y), y \in \overline{S_Y} = \{1,2,3,4\}$$

$$\Longrightarrow f_Y(1) = \frac{1}{16}, \quad f_Y(2) = \frac{3}{16}, \quad f_Y(3) = \frac{5}{16}, \quad f_Y(4) = \frac{7}{16}$$

# Remarks on Marginal pmf

It is crucial to understand the following definitions

$\overline{S}, \overline{S_X}, \overline{S_Y}, \overline{S_X}(y), \overline{S_Y}(x)$

$$\begin{aligned}
\overline{S} &= \{\text{all possible values of } (X, Y)\} \\
\overline{S_X} &= \{\text{all possible values of } X\} = \left\{x | (x, y) \in \overline{S}\right\} \\
\overline{S_Y} &= \{\text{all possible values of } Y\} = \left\{y | (x, y) \in \overline{S}\right\}
\end{aligned}$$

$$\begin{aligned}
\overline{S_X}(y) &= \left\{x | (x, y) \in \overline{S}\right\} \text{ for a given y} \in \overline{S_Y} \\
\overline{S_Y}(x) &= \left\{y | (x, y) \in \overline{S}\right\} \text{ for a given x} \in \overline{S_X}
\end{aligned}$$

# Trinomial Distribution

Description: The random experiment has three mutually exclusive and exhaustive outcomes:

- "perfect",
- "second"
- "defective"

We repeat the experiment $n$ independent times, and moreover, the probabilities

- $p_X$: the probability of "perfect",
- $p_Y$: the probability of "second"
- $p_Z$: the probability of "defective"

remain the same for each repetition. Such $n$ repetitions can be called $n$ trinomial trials.

For the $n$ trinomial trials, we are interested in the number of perfects, the number of seconds and the number of defectives.

# Trinomial Distribution

For the $n$ trinomial trials, we let

- $X$ be number of perfects,
- $Y$ be number of seconds,
- $Z = n - X - Y$ be the number of defectives

We are interested in the joint pmf of $(X, Y)$, $f(x, y) : \overline{S} \to \mathbb{R}^2$

- $\overline{S} = \{(x, y) | x + y \leq n, x = 0, 1, \cdots, n, y = 0, 1, \cdots, n\}$
- $f(x, y) = P(X = x, Y = y)$ which is the probability of having $x$ perfects, $y$ seconds, and $n - x - y$ defectives

# Trinomial Distribution

Joint pmf: to calculate $f(x,y) = P(X = x, Y = y)$,

- the probability for each way of having $x$ perfects, $y$ seconds, and $n - x - y$ defectives is

$$p_X^x p_Y^y (1 - p_X - p_Y)^{n-x-y}$$

- the total number of ways of having $x$ perfects, $y$ seconds, and $n - x - y$ defectives is

$$\binom{n}{x, y, n-x-y} = \frac{n!}{x!y!(n-x-y)!}$$

Therefore, the joint pmf for trinomial distribution is

$$f(x,y) = \frac{n!}{x!y!(n-x-y)!} p_X^x p_Y^y (1 - p_X - p_Y)^{n-x-y}, (x,y) \in \overline{S}$$

It's called trinomial distribution because of the trinomial expansion.

# Trinomial Distribution

$$
\begin{aligned}
(a+b+c)^n &= \sum_{x=0}^{n} \binom{n}{x} a^x (b+c)^{n-x} \\
&= \sum_{x=0}^{n} \binom{n}{x} a^x \sum_{y=0}^{n-x} \binom{n-x}{y} b^y c^{n-x-y} \\
&= \sum_{x=0}^{n} \sum_{y=0}^{n-x} \frac{n!}{x!y!(n-x-y)!} a^x b^y c^{n-x-y}
\end{aligned}
$$

Marginal pmf: to calculate $f_X(x)$ or $f_Y(y)$

# Trinomial Distribution

$$(a+b+c)^n = \sum_{x=0}^{n} \left( \begin{array}{c} n \\ x \end{array} \right) a^x (b+c)^{n-x}$$
$$= \sum_{x=0}^{n} \left( \begin{array}{c} n \\ x \end{array} \right) a^x \sum_{y=0}^{n-x} \left( \begin{array}{c} n-x \\ y \end{array} \right) b^y c^{n-x-y}$$
$$= \sum_{x=0}^{n} \sum_{y=0}^{n-x} \frac{n!}{x!y!(n-x-y)!} a^x b^y c^{n-x-y}$$

Marginal pmf: to calculate $f_X(x)$ or $f_Y(y)$

$$f_X(x) = \sum_{y \in \overline{S_Y}(x)} f(x,y) = \sum_{y=0}^{n-x} \left( \begin{array}{c} n \\ x \end{array} \right) \left( \begin{array}{c} n-x \\ y \end{array} \right) p_X^x p_Y^y (1-p_X-p_Y)^{n-x-y}$$
$$= \left( \begin{array}{c} n \\ x \end{array} \right) p_X^x (1-p_X)^{n-x}$$

Without summing, we know $X \sim b(n, p_X)$ and $Y \sim b(n, P_Y)$

# Independent Random Variables

**Definition**

The random variables $X$ and $Y$ are said to be independent if for every $x \in \overline{S_X}$ and $y \in \overline{S_Y}$

$$P(X = x, Y = y) = P(X = x)P(Y = y)$$

or equivalently,

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

$X$ and $Y$ are said to be dependent if otherwise.

When $X$ and $Y$ are independent,

$$\overline{S} = \overline{S_X} \times \overline{S_Y}, \quad \overline{S} \text{ is said to be rectangular}$$

which is a necessary condition for independence of $X$ and $Y$.

# Independent Random Variables

**Definition**

The random variables $X$ and $Y$ are said to be independent if for every $x \in \overline{S_X}$ and $y \in \overline{S_Y}$

$$P(X = x, Y = y) = P(X = x)P(Y = y)$$

or equivalently,

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

The definition of independent RVs has root in the definition of independent events.

$$A = \{X = x, Y \in \overline{S_Y}(x)\}, B = \{X \in \overline{S_X}(y), Y = y\}$$

$X$ and $Y$ are independent if and only if $A$ and $B$ are independent.

# Example 2, Page 135

Let the joint pmf of $X$ and $Y$ be defined by

$$f(x,y) = \frac{x+y}{21}, \quad x = 1,2,3, \quad y = 1,2.$$

$$\overline{S} = \{(x,y)|x = 1,2,3, \quad y = 1,2.\}$$

$$f : \overline{S} \longrightarrow (0,1] \text{ with } \overline{S_X} = \{1,2,3\}, \quad \overline{S_Y} = \{1,2\}.$$

**Question**

Are $X$ and $Y$ independent or dependent?

# Example 2, Page 135

$$f_X(x) = \sum_{y \in \overline{S_Y}(x)} f(x,y) = \sum_{y=1}^{2} \frac{x+y}{21} = \frac{2x+3}{21}, \quad x = 1, 2, 3.$$

$$f_Y(y) = \sum_{x \in \overline{S_X}(y)} f(x,y) = \sum_{x=1}^{3} \frac{x+y}{21} = \frac{3y+6}{21}, \quad y = 1, 2$$

$$f(x,y) = \frac{x+y}{21} \neq \frac{2x+3}{21} \cdot \frac{3y+6}{21} = f_X(x) f_Y(y)$$

$$\Rightarrow X \text{ and } Y \text{ are dependent}$$