CH11 = CH6 + CH7,8,9

Sirindhorn International Institute of Technology
Thammasat University
School of Information, Computer and Communication Technology

ECS315 2019/1 Part V Dr.Prapun

11 Multiple Random Variables

One is often interested not only in individual random variables, but also in relationships between two or more random variables. Furthermore, one often wishes to make inferences about one random variable on the basis of observations of other random variables.

Example 11.1. If the experiment is the testing of a new medicine, the researcher might be interested in cholesterol level, blood pressure, and the glucose level of a test person.

11.1 A Pair of Discrete Random Variables

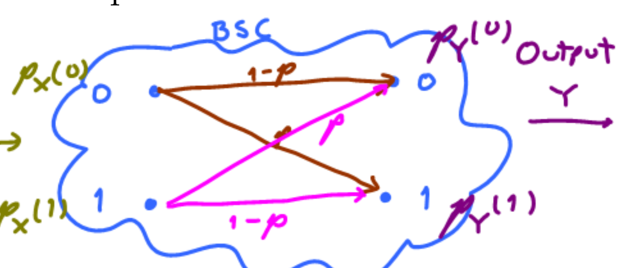
In this section, we consider two discrete random variables, say X and Y, simultaneously.

11.2. The analysis are different from Section 9.2 in two main aspects. First, there may be *no* deterministic relationship (such as $Y = g(X)$) between the two random variables. Second, we want to look at both random variables as a whole, not just X alone or Y alone.

Example 11.3. Communication engineers may be interested in the input X and output Y of a communication channel.

pmf of X $p_X(x)$

Input X



pmf of Y $p_Y(y)$

Problem: $p_X(x)$ and $p_Y(y)$ do not "show" the relationship between X and Y.

11.4. Recall that, in probability, "," means "and". For example,

$$P[X=x, Y=y] = P[X=x \text{ and } Y=y]$$

and

$$\begin{aligned} P[3 \le X < 4, Y < 1] &= P[3 \le X < 4 \text{ and } Y < 1] \\ &= P[X \in [3,4) \text{ and } Y \in (-\infty, 1)]. \end{aligned}$$

In general, the event

$$[\text{"Some condition(s) on } X\text{"}, \text{"Some condition(s) on } Y\text{"}]$$

is the same as the intersection of two events:

$$[\text{"Some condition(s) on } X\text{"}] \cap [\text{"Some condition(s) on } Y\text{"}]$$

which simply means both statements happen.

More technically,

$$[X \in B, Y \in C] = [X \in B \text{ and } Y \in C] = [X \in B] \cap [Y \in C]$$

and

$$\begin{aligned} P[X \in B, Y \in C] &= P[X \in B \text{ and } Y \in C] \\ &= P([X \in B] \cap [Y \in C]). \end{aligned}$$

Remark: Linking back to the original sample space, this shorthand actually says

$$\begin{aligned} [X \in B, Y \in C] &= [X \in B \text{ and } Y \in C] \\ &= \{\omega \in \Omega : X(\omega) \in B \text{ and } Y(\omega) \in C\} \\ &= \{\omega \in \Omega : X(\omega) \in B\} \cap \{\omega \in \Omega : Y(\omega) \in C\} \\ &= [X \in B] \cap [Y \in C]. \end{aligned}$$

11.5. The concept of conditional probability can be straightforwardly applied to discrete random variables. For example,

$$P\left[\text{“Some condition(s) on } X\text{”} \mid \text{“Some condition(s) on } Y\text{”}\right] \quad (28)$$

is the conditional probability $P(A|B)$ where

$$A = [\text{“Some condition(s) on } X\text{”}] \text{ and}$$
$$B = [\text{“Some condition(s) on } Y\text{”}].$$

Recall that $P(A|B) = P(A \cap B)/P(B)$. Therefore,

$$P\left[X = x|\, Y = y\right] = \frac{P\left[X = x \text{ and } Y = y\right]}{P\left[Y = y\right]},$$

and

$$P\left[3 \le X < 4|\, Y < 1\right] = \frac{P\left[3 \le X < 4 \text{ and } Y < 1\right]}{P\left[Y < 1\right]}$$

More generally, (28) is

$$\begin{aligned}
&= \frac{P\left([\text{“Some condition(s) on } X\text{”}] \cap [\text{“Some condition(s) on } Y\text{”}]\right)}{P\left([\text{“Some condition(s) on } Y\text{”}]\right)} \\
&= \frac{P\left([\text{“Some condition(s) on } X\text{”}, \text{“Some condition(s) on } Y\text{”}]\right)}{P\left([\text{“Some condition(s) on } Y\text{”}]\right)} \\
&= \frac{P[\text{“Some condition(s) on } X\text{”}, \text{“Some condition(s) on } Y\text{”}]}{P[\text{“Some condition(s) on } Y\text{”}]}
\end{aligned}$$

More technically,

$$\begin{aligned}
P\left[X \in B|\, Y \in C\right] = P\left([X \in B]\,|[Y \in C]\right) &= \frac{P\left([X \in B] \cap [Y \in C]\right)}{P\left([Y \in C]\right)} \\
&= \frac{P\left[X \in B, Y \in C\right]}{P\left[Y \in C\right]}.
\end{aligned}$$

Joint probability

$$P(\underbrace{A\cap B}_{\text{Joint event}})$$

$$P[X \ge 2, Y < 3]$$

means "and"

$A = [X \ge 2]$

$B = [Y < 3]$

Conditional probability

$$P(A|B) = \frac{P(A\cap B)}{P(B)}$$

$$P[X \ge 2 | Y < 3] = \frac{P[X \ge 2, Y < 3]}{P[Y < 3]}$$

(a)

Joint probability

$$P(\underbrace{A\cap B}_{\text{Joint event}})$$

$$P[X \in C, Y \in D]$$

$A = [X \in C]$

$B = [Y \in D]$

This represents some event defined from RV X. For ex.,

$[x \geqslant 2] = [X \in [2, \infty)]$

Conditional probability

$$P(A|B) = \frac{P(A\cap B)}{P(B)}$$

$$P[X \in C | Y \in D] = \frac{P[X \in C, Y \in D]}{P[Y \in D]}$$

(b)

Joint probability

$$P(\underbrace{A\cap B}_{\text{Joint event}})$$

Joint pmf

$$P[X = x, Y = y] \equiv p_{X,Y}(x, y)$$

$A = [X = x]$

$B = [Y = y]$

Conditional probability

$$P(A|B) = \frac{P(A\cap B)}{P(B)}$$

Conditional pmf

$$P[X = x | Y = y] = \frac{P[X = x, Y = y]}{P[Y = y]} \equiv p_{X|Y}(x\,|\,y)$$

(c)

Figure 39: Joints events and conditional probabilities for discrete random variables: (a) an example, (b) the general case, (c) an important special case. Case (c) is used to defined the joint pmf and conditional pmf.

Definition 11.6. ***Joint pmf***: If X and Y are two discrete random variables (defined on a same sample space with probability measure P), the function $p_{X,Y}(x,y)$ defined by

Note that $\sum_{(x,y)} p_{X,Y}(x,y) = 1$ all possible pairs (x,y)

$$p_{X,Y}(x,y) = P[X = x, Y = y]$$

"and"

is called the ***joint probability mass function*** of X and Y.

(a) We can visualize the joint pmf via stem plot. See Figure 40.

(b) To evaluate the probability for a statement that involves both X and Y random variables:

Ex. $P[XY > 1]$, $P[X+Y=3]$, $P[X>Y]$

$P[$ statement(s) about X and $Y]$

We first find all pairs (x,y) that satisfy the condition(s) in the statement, and then add up all the corresponding values from the joint pmf .

More technically, we can then evaluate $P[(X,Y) \in R]$ by

$$P[(X,Y) \in R] = \sum_{(x,y):(x,y)\in R} p_{X,Y}(x,y).$$

Example 11.7 (F2011)**.** Consider random variables X and Y whose joint pmf is given by

$$p_{X,Y}(x,y) = \begin{cases} c\,(x+y), & x \in \{1,3\} \text{ and } y \in \{2,4\}, \\ 0, & \text{otherwise.} \end{cases}$$

(a) Check that $c = 1/20$.

"$\Sigma = 1$" $\Rightarrow$ $3c + 5c + 5c + 7c = 1$

$c = 1/20$

Four possible cases

(x,y)	$p_{X,Y}(x,y)$	x^2+y^2
$(1,2)$	$\boxed{3c}$	$1+4=\boxed{5}$
$(1,4)$	$\boxed{5c}$	$1+16=\boxed{17}$
$(3,2)$	$\boxed{5c}$	$9+4=\boxed{13}$
$(3,4)$	$7c$	$9+16=25$

(b) Find $P[X^2 + Y^2 = 13]$.

$= 5c = \frac{5}{20} = \frac{1}{4}$

(c) $P[X^2 + Y^2 < 20]$ $= 3c + 5c + 5c = 13c = 13/20$

In most situation, it is much more convenient to focus on the "important" part of the joint pmf. To do this, we usually present the joint pmf (and the conditional pmf) in their matrix forms:

$$\mathbf{P}_{X,Y} = \begin{array}{c|cc} x \backslash y & 2 & 4 \\ \hline 1 & 3c & 5c \\ 3 & 5c & 7c \end{array}$$

Definition 11.8. When both X and Y take finitely many values (both have finite supports), say $S_X = \{x_1, \ldots, x_m\}$ and $S_Y = \{y_1, \ldots, y_n\}$, respectively, we can arrange the probabilities $p_{X,Y}(x_i, y_j)$ in an $m \times n$ matrix

$$\begin{array}{cc}
x \backslash y & \begin{array}{cccc} y_1 & y_2 & \cdots & y_n \end{array} \\
\begin{array}{c} x_1 \\ x_2 \\ \vdots \\ x_m \end{array} & \begin{bmatrix} p_{X,Y}(x_1, y_1) & p_{X,Y}(x_1, y_2) & \ldots & p_{X,Y}(x_1, y_n) \\ p_{X,Y}(x_2, y_1) & p_{X,Y}(x_2, y_2) & \ldots & p_{X,Y}(x_2, y_n) \\ \vdots & \vdots & \ddots & \vdots \\ p_{X,Y}(x_m, y_1) & p_{X,Y}(x_m, y_2) & \ldots & p_{X,Y}(x_m, y_n) \end{bmatrix}
\end{array}. \tag{29}$$

- We shall call this matrix the ***joint pmf matrix***.
- The sum of all the entries in the matrix is one.

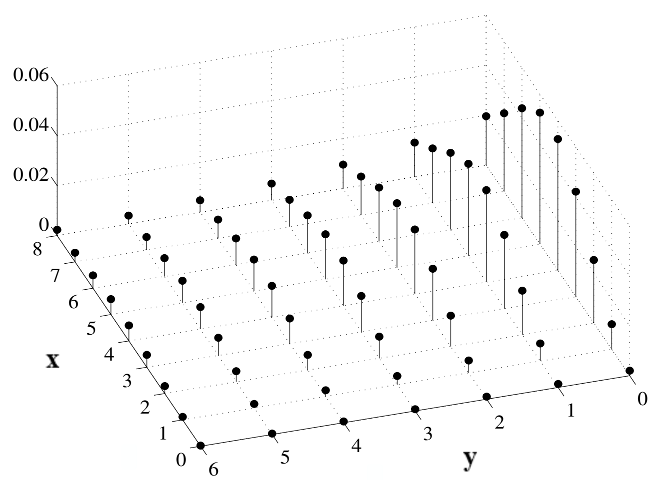


Figure 40: Example of the plot of a joint pmf. [9, Fig. 2.8]

- $p_{X,Y}(x, y) = 0$ if[51] $x \notin S_X$ or $y \notin S_Y$. In other words, we don't have to consider the x and y outside the supports of X and Y, respectively.

[51]To see this, note that $p_{X,Y}(x, y)$ cannot exceed $p_X(x)$ because $P(A \cap B) \le P(A)$. Now, suppose at $x = a$, we have $p_X(a) = 0$. Then $p_{X,Y}(a, y)$ must also $= 0$ for any y because it cannot exceed $p_X(a) = 0$. Similarly, suppose at $y = a$, we have $p_Y(a) = 0$. Then $p_{X,Y}(x, a) = 0$ for any x.

11.9. From the joint pmf, we can find $p_X(x)$ and $p_Y(y)$ by

$$p_X(x) = \sum_y p_{X,Y}(x, y) \tag{30}$$

$$p_Y(y) = \sum_x p_{X,Y}(x, y) \tag{31}$$

In this setting, $p_X(x)$ and $p_Y(y)$ are call the ***marginal pmfs*** (to distinguish them from the joint one).

(a) Suppose we have the joint pmf matrix in (29). Then, the sum of the entries in the ith row is[52] $p_X(x_i)$, and the sum of the entries in the jth column is $p_Y(y_j)$:

$$p_X(x_i) = \sum_{j=1}^{n} p_{X,Y}(x_i, y_j) \quad \text{and} \quad p_Y(y_j) = \sum_{i=1}^{m} p_{X,Y}(x_i, y_j)$$

(b) In MATLAB, suppose we save the joint pmf matrix as P_XY, then the marginal pmf (row) vectors p_X and p_Y can be found by

```
p_X = (sum(P_XY,2))'
p_Y = (sum(P_XY,1))
```

Example 11.10. Consider the following joint pmf matrix

(given)

$$\begin{array}{c|cccc} x\backslash y & 0 & 1 & 2 & 3 \\ \hline 0 & 0.1 & 0 & 0.2 & 0 \\ 1 & 0 & 0.5 & 0 & 0 \\ 2 & 0 & 0.1 & 0.1 & 0 \end{array} \quad \begin{array}{l} \xrightarrow{\Sigma} 0.3 \\ \xrightarrow{\Sigma} 0.5 \\ \xrightarrow{\Sigma} 0.2 \end{array}$$

$\Sigma\downarrow$: 0.1, 0.6, 0.3, 0

$$p_X(x) = \begin{cases} 0.3, & x=0, \\ 0.5, & x=1 \\ 0.2, & x=2 \\ 0, & \text{otherwise} \end{cases}$$

$$p_Y(y) = \begin{cases} 0.1, & y=0, \\ 0.6, & y=1, \\ 0.3, & y=2, \\ 0, & \text{otherwise} \end{cases}$$

① $P[X=1 \text{ and } Y=1] = 0.5$

$p_{X,Y}(0,2) = 0.2$

② $P[XY > 2] = 0.1 + 0 + 0 = 0.1$

xy:

$$\begin{array}{c|cccc} x\backslash y & 0 & 1 & 2 & 3 \\ \hline 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 2 & 3 \\ 2 & 0 & 2 & 4 & 6 \end{array}$$

③ $p_X(2) \equiv P[X=2] = 0 + 0.1 + 0.1 + 0$
$= 0.2$

---

[52] To see this, we consider $A = [X = x_i]$ and a collection defined by $B_j = [Y = y_j]$ and $B_0 = [Y \notin S_Y]$. Note that the collection $B_0, B_1, \ldots, B_n$ partitions Ω. So, $P(A) = \sum_{j=0}^{n} P(A \cap B_j)$. Of course, because the support of Y is S_Y, we have $P(A \cap B_0) = 0$. Hence, the sum can start at $j = 1$ instead of $j = 0$.

Definition 11.11. The ***conditional pmf*** of X given Y is defined as

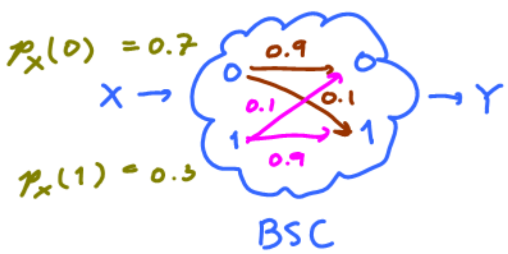
$$p_{X|Y}(x|y) = P[X = x|Y = y]$$

which gives

$$p_{X,Y}(x, y) = p_{X|Y}(x|y)p_Y(y) = p_{Y|X}(y|x)p_X(x). \qquad (32)$$

11.12. Equation (32) is quite important in practice. In most cases, systems are naturally defined/given/studied in terms of their conditional probabilities, say $p_{Y|X}(y|x)$. Therefore, it is important the we know how to construct the joint pmf from the conditional pmf.

Example 11.13. Consider a binary symmetric channel defined in Example 11.3. Suppose the input X to the channel is Bernoulli(0.3). At the output Y of this channel, the crossover (bit-flipped) probability is 0.1. Find the joint pmf $p_{X,Y}(x, y)$ of X and Y.



$p_Y(0) = 0.7 \times 0.9 + 0.3 \times 0.1$

$= 0.63 + 0.03 = 0.66$

$p_Y(1) = 0.34$

$$\begin{array}{c|cc} x \backslash y & 0 & 1 \\ \hline 0 & \boxed{0.63} & \overset{0.07}{\triangle} \\ 1 & 0.03 & 0.27 \end{array}$$

$\square = p_{X,Y}(0,0) \equiv P[\underset{A}{X=0}, \underset{B}{Y=0}] = P(A \cap B)$

$= P(A)P(B|A) = P[X=0]\,P[Y=0|X=0] = 0.7 \times 0.9$

$= 0.63$

$\triangle = p_{X,Y}(0,1) \equiv P[X=0, Y=1] = P[X=0]\,P[Y=1|X=0]$

$= 0.7 \times 0.1 = 0.07$

$$p_{X,Y}(x,y) = \begin{cases} 0.63, & (x,y) = (0,0), \\ 0.07, & (x,y) = (0,1), \\ 0.03, & (x,y) = (1,0), \\ 0.27, & (x,y) = (1,1), \\ 0, & \text{otherwise.} \end{cases}$$

$$p_{X,Y} = \begin{array}{c|cc} x \backslash y & 2 & 4 \\ \hline 1 & 3/20 & 5/20 \\ 3 & 5/20 & 7/20 \end{array} \quad \begin{array}{l} \xrightarrow{\Sigma} 8/20 = 2/5 \\ \xrightarrow{\Sigma} 12/20 = 3/5 \end{array}$$

Exercise 11.14 (F2011)**.** Continue from Example 11.7. Random variables X and Y have the following joint pmf

$$p_{X,Y}(x,y) = \begin{cases} c(x+y), & x \in \{1,3\} \text{ and } y \in \{2,4\}, \\ 0, & \text{otherwise.} \end{cases}$$

(a) Find $p_X(x)$. $= \begin{cases} 2/5, & x=1, \\ 3/5, & x=3 \\ 0, & \text{otherwise.} \end{cases}$

(b) Find $\mathbb{E}X$. $= \sum_x x\, p_X(x) = 1 \times \frac{2}{5} + 3 \times \frac{3}{5} = 11/5 = 2.2$

(c) Find $p_{Y|X}(y|1)$. Note that your answer should be of the form

$= P[\underbrace{Y=y}_{A} \mid \underbrace{X=1}_{B}]$

$= \frac{P(A \cap B)}{P(B)} = P[Y=y, X=1]/P[X=1]$

$$p_{Y|X}(y|1) = \begin{cases} ?, & y=2, \\ ?, & y=4, \\ 0, & \text{otherwise.} \end{cases}$$

$\frac{P[Y=2, X=1]}{P[X=1]} = \frac{3/20}{2/5} = 3/8$

$\frac{P[Y=4, X=1]}{P[X=1]} = \frac{5/20}{2/5} = 5/8$

(d) Find $p_{Y|X}(y|3)$.

Definition 11.15. The ***joint cdf*** of X and Y is defined by

$$F_{X,Y}(x,y) = P[X \le x, Y \le y].$$

Definition 11.16. Two random variables X and Y are said to be ***identically distributed*** if, for every B, $P[X \in B] = P[Y \in B]$.

In words, for any probability statement about X (and only X), if we replace X by Y, we get the same probability.

Example 11.17. Roll a dice twice. Let X be the result from the first roll. Let Y be the result from the second roll.

- X and Y are not the same. (Most of the time, they will be different. By chance, they occasionally take the same value.)
- $P[X > 3] = P[Y > 3]$

Example 11.18. Let $X \sim$ Bernoulli$(1/2)$. Let $Y = X$ and $Z = 1 - X$. Then, all of these random variables are identically distributed.

11.19. The following statements are equivalent:

(a) Random variables X and Y are ***identically distributed***.

(b) For every B, $P[X \in B] = P[Y \in B]$

(c) $p_X(c) = p_Y(c)$ for all c

(d) $F_X(c) = F_Y(c)$ for all c

Definition 11.20. Two random variables X and Y are said to be ***independent*** if the events $[X \in B]$ and $[Y \in C]$ are independent for all sets B and C.

11.21. The following statements are equivalent:

Ex. $X \perp\!\!\!\perp Y$

$\Rightarrow P[\underbrace{X > 3, Y < 7}_{A \cap B}]$

$= P[\underbrace{X > 3}_{A}]\, P[\underbrace{Y < 7}_{B}]$

(a) Random variables X and Y are ***independent***.

(b) $[X \in B] \perp\!\!\!\perp [Y \in C]$ for all B, C.

(c) $P[X \in B, Y \in C] = P[X \in B] \times P[Y \in C]$ for all B, C.

(d) $p_{X,Y}(x, y) = p_X(x) \times p_Y(y)$ for all x, y.

(e) $F_{X,Y}(x, y) = F_X(x) \times F_Y(y)$ for all x, y.

Definition 11.22. Two random variables X and Y are said to be ***independent and identically distributed (i.i.d.)*** if X and Y are both independent and identically distributed.

11.23. Being identically distributed does not imply independence. Similarly, being independent, does not imply being identically distributed.

i.d. : ① $P[X \in B] = P[Y \in B]$ for all B

② $p_X(c) = p_Y(c)$ for all c

$$\mathbb{E}[g(X)] = \sum_x g(x) p_X(x) \underset{x \to c}{=} \sum_c g(c) p_X(c) = \sum_c g(c) p_Y(c)$$

$$\underset{c \to y}{=} \sum_y g(y) p_Y(y) = \mathbb{E}[g(Y)]$$

Example 11.24. Roll a dice. Let X be the result. Set $Y = X$. (Note that this is different from Example 11.17. There, X and Y are i.i.d.)

Example 11.25. Suppose the pmf of a random variable X is given by

$$p_X(x) = \begin{cases} 1/4, & x = 3, \\ \alpha, & x = 4, \\ 0, & \text{otherwise.} \end{cases}$$

Let Y be another random variable. Assume that X and Y are i.i.d. → identically distributed

ind.

Find

(a) α, $\quad \sum_x p_X(x) = 1 \Rightarrow \frac{1}{4} + \alpha = 1 \Rightarrow \alpha = \frac{3}{4}$

(b) the pmf of Y, and $\quad p_Y(y) = p_X(y) = \begin{cases} 1/4, & y = 3, \\ 3/4, & y = 4, \\ 0, & \text{otherwise.} \end{cases}$

(c) the joint pmf of X and Y.

$$p_{X,Y}(x,y) \overset{\text{ind}}{=} p_X(x)\, p_Y(y)$$

$p_X(3)$ $\quad$ $p_Y(3)$

$$P_{X,Y} = \begin{array}{c|cc} x \backslash y & 3 & 4 \\ \hline 3 & \frac{1}{4} \times \frac{1}{4} & \frac{1}{4} \times \frac{3}{4} \\ 4 & \frac{3}{4} \times \frac{1}{4} & \frac{3}{4} \times \frac{3}{4} \end{array} = \begin{bmatrix} \frac{1}{16} & \frac{3}{16} \\ \frac{3}{16} & \frac{9}{16} \end{bmatrix}$$

$$p_{X,Y}(x,y) = \begin{cases} 1/16, & (x,y) = (3,3), \\ 3/16, & (x,y) = (3,4) \text{ or } (4,3), \\ 9/16, & (x,y) = (4,4), \\ 0, & \text{otherwise.} \end{cases}$$

Example 11.26. Consider a pair of random variables X and Y whose joint pmf is given by

$$p_{X,Y}(x,y) = \begin{cases} 1/15, & x = 3, y = 1, \\ 2/15, & x = 4, y = 1, \\ 4/15, & x = 3, y = 3, \\ \cancel{\beta}, 8/15 & x = 4, y = 3, \\ 0, & \text{otherwise.} \end{cases}$$

(a) Are X and Y identically distributed?

(b) Are X and Y independent?

(a)

$$\begin{array}{c|cc} x \backslash y & 1 & 3 \\ \hline 3 & 1/15 & 4/15 \\ 4 & 2/15 & 8/15 \end{array} \quad \begin{array}{l} \xrightarrow{\Sigma} 5/15 = 1/3 \\ \xrightarrow{\Sigma} 10/15 = 2/3 \end{array}$$

$$\Sigma \downarrow \quad 1/5 \qquad \Sigma \downarrow \quad 4/5$$

$$p_X(x) = \begin{cases} 1/3, & x = 3, \\ 2/3, & x = 4, \\ 0, & \text{otherwise.} \end{cases}$$

$$p_Y(y) = \begin{cases} 1/5, & y = 1, \\ 4/5, & y = 3, \\ 0, & \text{otherwise.} \end{cases}$$

X and Y are <u>not</u> identically distributed

(For example, $c = 3$

$$\left. \begin{array}{l} p_X(c) = \frac{1}{3} \\ p_Y(c) = \frac{4}{5} \end{array} \right\} \neq \quad)$$

(b) calculate $p_X(x)\, p_Y(y)$

$$p_X(x)\, p_Y(y) \quad \begin{array}{c|cc} x \backslash y & 1 & 3 \\ \hline 3 & 1/3 \times 1/5 & 1/3 \times 4/5 \\ 4 & 2/3 \times 1/5 & 2/3 \times 4/5 \end{array}$$

$$= \begin{bmatrix} 1/15 & 4/15 \\ 2/15 & 8/15 \end{bmatrix}$$

$\swarrow$

same as the joint pmf matrix

$\Downarrow$

$p_{X,Y}(x,y) = p_X(x)\, p_Y(y)$ for any x, y

$\Downarrow$

$X \perp\!\!\!\perp Y$