Note 16

Section 4.1.

1 Review

① Joint PMF: If both X and Y are discrete RVs, then

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

Emphasize which RVs are involved

is called the joint PMF of X and Y.

② Properties of joint PMF:

(1) $0 \le p(x,y) \le 1$ for any x, y.

(2) $\sum_{x,y} p(x,y) = 1$

(3) $P((X,Y) \in A) = \sum_{(x,y)\in A} p(x,y)$ for any $A \subseteq S$ (= space of (X,Y))

③ The PMF of a single RV alone is called marginal PMF. The marginal PMF of X can be computed by

$$p_X(x) = P(X=x) = \sum_y p_{X,Y}(x,y) \qquad \text{(Why?)}$$

and likewise for the marginal PMF of Y:

$$p_Y(y) = P(Y=y) = \sum_x p_{X,Y}(x,y).$$

④ X and Y are called independent if

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

Otherwise, they are called dependent.

<u>Note</u>

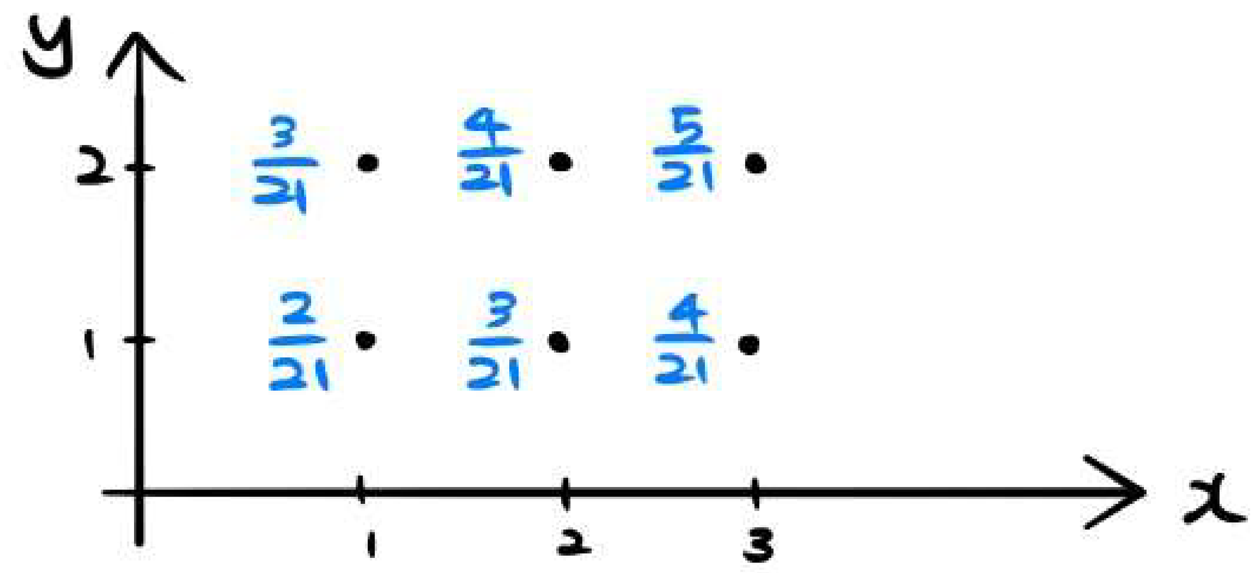
In other words, X & Y are independent iff $\{X=x\}$ and $\{Y=y\}$ are independent events for any x and y.

$\boxed{2}$ Bivariate Distribution, continued

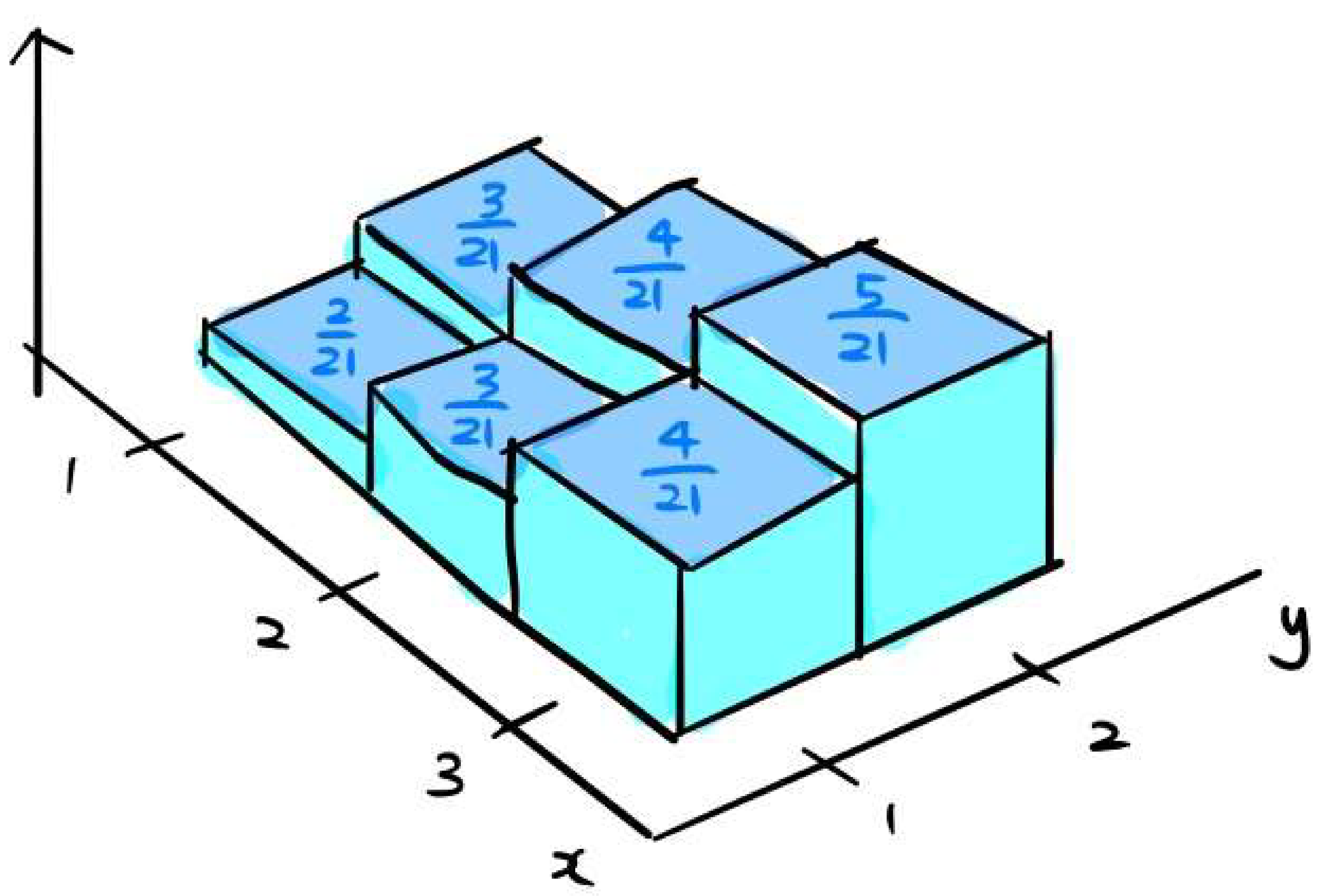
① Representing joint PMF

Ex Let $p_{X,Y}(x,y) = \dfrac{x+y}{21}$, $\quad x = 1, 2, 3, \quad y = 1, 2.$

(1) We may plot the space of (X, Y) labeled with values of $p_{X,Y}$:



(2) We may plot a 3D probability histogram:



(3) We may draw a table:

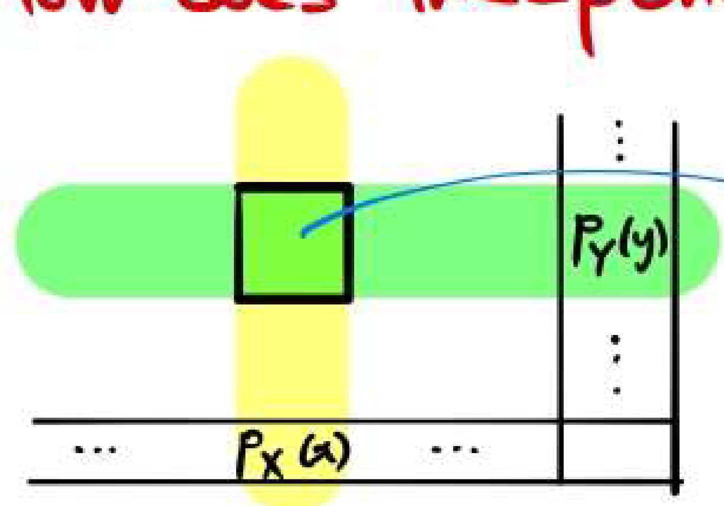
Y \ X	1	2	3	total
1	2/21	3/21	4/21	9/21
2	3/21	4/21	5/21	12/21
total	5/21	7/21	9/21	1

marginal PMF of Y.

marginal PMF of X

Note: Q. How does independence manifest in the table?

A.



X & Y independent

$\Updownarrow$

$p_{X,Y}(x,y) = p_X(x)p_Y(y)$ holds.

② Expected value of a function of 2 RVs

PROP If

- X and Y are discrete RVs,
- $u(x,y)$: real-valued function,

Then $u(X,Y)$ is another RV. Its expected value can be computed by

$$E[u(X,Y)] = \sum_{x,y} u(x,y)\, P_X(x,y).$$

Ex The expected values of X and Y can be computed by:

$$E(X) = \sum_{x,y} x\, P_X(x,y) \quad \text{and} \quad E(Y) = \sum_{x,y} y\, P_X(x,y).$$

(Why?)

Ex Let the joint PMF of X and Y be given by

Y \ X	1	2
0	0.1	0.3
1	0.4	0.2

Then

- $E(X) = 1\cdot(0.1) + 1\cdot(0.4) + 2\cdot(0.3) + 2\cdot(0.2) = 1.5.$
- $$\begin{aligned} E[X^2 - XY] &= (1^2 - 1\cdot 0)\cdot(0.1) + (1^2 - 1\cdot 1)\cdot(0.4) \\ &\quad + (2^2 - 2\cdot 0)\cdot(0.3) + (2^2 - 2\cdot 1)\cdot(0.2) \\ &= 0.1 + 0 + 1.2 + 0.4 \\ &= 1.7. \end{aligned}$$

$\square$

Ex Variance of $X-2Y$ can be computed by:

$$Var(X-2Y) = E[(X-2Y)^2] - (E[X-2Y])^2$$

3 Multinomial Distribution

- Setting:
  - ▷ Each trial yields an outcome of:

$$\begin{cases} \text{category } 1 \text{ w/ prob. } p_1 \\ \quad \text{—''—} \quad 2 \text{ w/ prob. } p_2 \\ \qquad \vdots \\ \quad \text{—''—} \quad k \text{ w/ prob. } p_n, \end{cases}$$

    Of course, $0 \le p_i \le 1$ and $p_1 + \cdots + p_k = 1$.
  - ▷ Perform n indep. trials.
  - ▷ $X_1 = [\# \text{ of outcomes of category } 1]$

    $\vdots$

    $X_k = [\text{———— '' ————} \; k]$.

DEF $(X_1, \cdots, X_k)$ is said to have a multinomial distribution.

Ex (Binomial dist.) If $k=2$, then
- ▷ Support of $(X_1, X_2) = \{(0,n), (1,n-1), \cdots, (n,0)\}$
- ▷ Joint PMF: If $x_1 + x_2 = n$,

$$p(x_1, x_2) = P(X_1 = x_1) = \binom{n}{x_1} p_1^{x_1} (1-p_1)^{n-x_1} = \frac{n!}{x_1! \, x_2!} p_1^{x_1} p_2^{x_2}$$

- ▷ Note X_1 is $b(n, p_1)$ and X_2 is $b(n, p_2)$.

Ex (Trinomial dist.)
- ▷ Support of $(X_1, X_2, X_3) = \left\{ (x_1, x_2, x_3) : \begin{array}{l} x_1, x_2, x_3 \text{ non-negative integers} \\ \text{satisfying } x_1 + x_2 + x_3 = n \end{array} \right\}$
- ▷ Joint PMF: If $x_1 + x_2 + x_3 = n$,

$$p(x_1, x_2, x_3) = \frac{n!}{x_1! \, x_2! \, x_3!} p_1^{x_1} p_2^{x_2} p_3^{x_3}.$$

- ▷ Note Each X_i is $b(n, p_i)$.

Section 4.2 Correlation Coefficient

1 Correlation Coefficient

DEF • If $\mu_X = E(X)$ and $\mu_Y = E(Y)$, then

$$\mathrm{Cov}(X, Y) = \sigma_{XY} := E[(X-\mu_X)(Y-\mu_Y)]$$

is called the covariance of X and Y.

• If $\sigma_X = \sqrt{\mathrm{Var}(X)}$ and $\sigma_Y = \sqrt{\mathrm{Var}(Y)}$ are positive, then

$$\rho = \frac{\mathrm{Cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$$

is called the correlation coefficient of X and Y.

PROP (a) $\mathrm{Cov}(X,Y) = E(XY) - E(X)E(Y) = E(XY) - \mu_X\mu_Y$.

(b) $\mathrm{Cov}(X,X) = \mathrm{Var}(X)$.

Pf) (a)

$$\begin{aligned}\mathrm{Cov}(X,Y) &= E[(X-\mu_X)(Y-\mu_Y)] \\ &= E[XY - \mu_X Y - \mu_Y X + \mu_X\mu_Y] \\ &= E(XY) - \mu_X \underbrace{E(Y)}_{=\mu_Y} - \mu_Y \underbrace{E(X)}_{=\mu_X} + \mu_X\mu_Y \\ &= E(XY) - \mu_X\mu_Y.\end{aligned}$$

(b) $\mathrm{Cov}(X,X) = E[(X-\mu_X)(X-\mu_X)] = E[(X-\mu_X)^2] = \mathrm{Var}(X)$. $\square$

Ex Let X and Y have the joint PMF

$$p(0,0) = 0.2, \quad p(0,1) = 0.3, \quad p(1,1) = 0.3, \quad p(1,2) = 0.2$$

Then

Y \ X	0	1	
0	0.2	0	0.2
1	0.3	0.3	0.6
2	0	0.2	0.2
	0.5	0.5	

Y marginal

X marginal

- $E(X) = (0)\cdot 0.5 + (1)\cdot 0.5 = 0.5$
- $E(Y) = (0)\cdot 0.2 + (1)\cdot 0.6 + (2)\cdot 0.2 = 1$
- $Var(X) = E(X^2) - \mu_X^2 = 0.25 \Rightarrow \sigma_X = 0.5$
- $Var(Y) = E(Y^2) - \mu_Y^2 = 0.4 \Rightarrow \sigma_Y = \sqrt{0.4} \approx 0.632$
- $Cov(X,Y) = E(XY) - \mu_X \mu_Y$
  $= \big((0)\cdot(0)\cdot 0.2 + (0)\cdot(1)\cdot 0.3 + (1)\cdot(1)\cdot 0.3 + (1)\cdot(2)\cdot 0.2\big) - 0.5$
  $= 0.2.$
- $\rho = Cov(X,Y) / \sigma_X \sigma_Y = \sqrt{0.4} \approx 0.632.$ $\square$

<u>Next Time</u>: Linear Square Regression Line, conditional distribution.