# Math 3501 - Probability and Statistics I

4.1 - Bivariate distributions of the discrete type

# Joint probability mass function

**Definition (Joint probability mass function)**

Let $X$ and $Y$ be two random variables defined on a discrete sample space, and let $S$ denote the corresponding two-dimensional space of $X$ and $Y$, the two random variables of the discrete type.

The *joint probability mass function* (abbreviated joint pmf) of $X$ and $Y$, denoted $f(x, y)$, is defined as

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

**Properties**:

(a) $0 \leq f(x, y) \leq 1$ for all $(x, y) \in \mathbb{R}^2$

(b) $\displaystyle\sum_{(x,y)\in S} f(x, y) = 1$

(c) $P[(X, Y) \in A] = \displaystyle\sum_{(x,y)\in A} f(x, y)$, where $A \subset \mathbb{R}^2$

analogous to 1-dim case (Sec 2.1)

## Example

Roll a pair of fair 6-faced dice. Let $X$ denote the smaller and $Y$ the larger outcome on the dice.

Determine the joint pmf of $X$ and $Y$.

$X$ = r.v. giving the smaller score

$Y$ = r.v. " " larger score

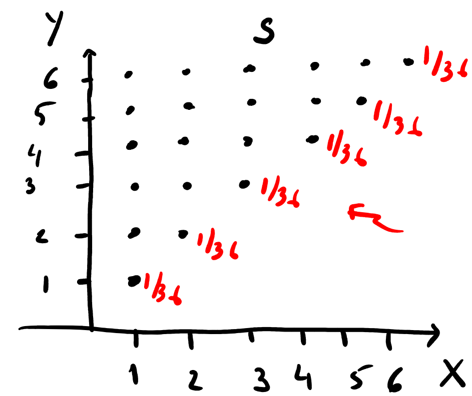
e.g.

if scores are $(2,3)$, $X=2$ and $Y=3$

if scores are $(4,4)$, $X=4$ and $Y=4$

The space of $X$ and $Y$ is

$$S = \{ (i,j) : i,j \in \{1,2,\dots,6\} \text{ and } i \leq j \} \longrightarrow$$



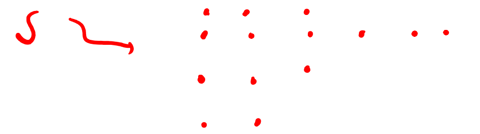

$$f(2,3) = P(X=2, Y=3) = P(\{(2,3),(3,2)\}) = \frac{2}{36} = \frac{1}{18}$$

$$f(4,4) = P(X=4, Y=4) = P(\{(4,4)\}) = \frac{1}{36}$$

if $i < j \implies f(i,j) = \frac{2}{36}$

if $i = j \implies f(i,j) = \frac{1}{36}$

$$\boxed{f(x,y) = \begin{cases} \frac{2}{36} & \text{if } i,j \in \{1,\dots,6\} \text{ and } i<j \\ \frac{1}{36} & \text{if } i,j \in \{1,\dots,6\} \text{ and } i=j \end{cases}}$$




## Marginal probability mass function

### Definition (Marginal probability mass function)

Let $X$ and $Y$ have the joint probability mass function $f(x, y)$ with space $S$.

The probability mass function of $X$, called the *marginal probability mass function of $X$*, is defined by

$$f_X(x) = \sum_y f(x, y) = P(X = x) \;, \quad x \in S_X \;,$$

where the summation is taken over all possible $y$ values for each given $x \in S_X$.

Similarly, the *marginal probability mass function of $Y$* is defined by

$$f_Y(y) = \sum_x f(x, y) = P(Y = y) \;, \quad y \in S_Y \;,$$

where the summation is taken over all possible $x$ values for each given $y \in S_Y$.

# Independence

## Definition (Independence)

The random variables $X$ and $Y$ are *independent* if

$$P(X = x, Y = y) = P(X = x)P(Y = y) \quad \text{for all } x \in \overbrace{S_x}^{\text{space of } X} \text{ and } y \in \overbrace{S_y}^{\text{space of } Y}$$

or, equivalently,

$$\boxed{f(x, y) = f_X(x) f_Y(y) \quad \text{for all } x \in S_x \text{ and } y \in S_y\ .}$$

Otherwise, $X$ and $Y$ are said to be *dependent*.

In words: X and Y are independent if their joint pmf may be written as the product of their marginals

## Example (Previous example continued)

Roll a pair of fair dice. Let $X$ denote the smaller and $Y$ the larger outcome on the dice.

6-faced

Determine the marginal pmfs of $X$ and $Y$ and determine whether $X$ and $Y$ are independent.

joint pmf

| $f_Y(y)$ | $y$ | | | | | | |
|---|---|---|---|---|---|---|---|
| 11/36 | 6 | 2/36 | 2/36 | 2/36 | 2/36 | 2/36 | 1/36 |
| 9/36 | 5 | 2/36 | 2/36 | 2/36 | 2/36 | 1/36 | |
| 7/36 | 4 | 2/36 | 2/36 | 2/36 | 1/36 | | |
| 5/36 | 3 | 2/36 | 2/36 | 1/36 | | | |
| 3/36 | 2 | 2/36 | 1/36 | | | | |
| 1/36 | 1 | 1/36 | | | | | |
| | $x$ | 1 | 2 | 3 | 4 | 5 | 6 |
| | $f_X(x)$: | 11/36 | 9/36 | 7/36 | 5/36 | 3/36 | 1/36 |

← from last day

?

sum of values above $x=1$

$S_X = \{1, 2, \dots, 6\}$

$S_Y = \{1, 2, \dots, 6\}$

Marginals:

$$f_X(x) = \sum_y f(x,y) = \begin{cases} 11/36 & \text{if } x=1 \\ 9/36 & \text{if } x=2 \\ 7/36 & \text{if } x=3 \\ 5/36 & \text{if } x=4 \\ 3/36 & \text{if } x=5 \\ 1/36 & \text{if } x=6 \end{cases}$$

$$f_Y(y) = \sum_x f(x,y) = \begin{cases} 1/36 & \text{if } y=1 \\ 3/36 & \text{if } y=2 \\ 5/36 & \text{if } y=3 \\ \vdots & \vdots \\ 11/36 & \text{if } y=6 \end{cases}$$

Recall :

$X$ and $Y$ are independent if $f(x,y) = f_X(x) \cdot f_Y(y)$ for all $x \in S_X$ and $y \in S_Y$

To see that this condition fails, it is enough to find a pair of values for which $f(x,y) \neq f_X(x) \cdot f_Y(y)$

For instance if we take $x=2$ and $y=3$, we have:

$$f(2,3) = \frac{2}{36} \quad \text{while} \quad f_X(2) \cdot f_Y(3) = \frac{9}{36} \cdot \frac{5}{36} = \frac{45}{(36)^2} \neq \frac{2}{36}$$

$\Rightarrow$ $X$ and $Y$ are not independent.

Alternative approach would be: to take, for instance, $x=6$ and $y=1$, for which

$$\to \quad f(6,1) = 0 \neq f_X(6) \cdot f_Y(1) = \frac{1}{36} \cdot \frac{1}{36} = \left(\frac{1}{36}\right)^2 \Rightarrow \text{Not independent!}$$

INSIGHT FROM THIS EXAMPLE: $X$ and $Y$ cannot be independent if $\boxed{S \neq S_X \times S_Y}$

## Example

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

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

Decide if $X$ and $Y$ are independent.

$S = S_X \times S_Y$


$S_X = \{1,2,3\}$

$S_Y = \{1,2\}$

$S = S_X \times S_Y =$

$= \{(1,1),(1,2),(2,1),(2,2),(3,1),(3,2)\}$

Let us determine the marginal pmfs:

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

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

Observe that $f(1,1) = \frac{2}{21}$ while $f_X(1) \cdot f_Y(1) = \frac{5}{21} \cdot \frac{9}{21} = \frac{45}{(21)^2} \neq \frac{2}{21}$

$\Rightarrow$ $X$ and $Y$ are not independent because $f(x,y) \neq f_X(x) \cdot f_Y(y)$

## Example

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

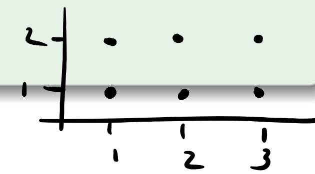
$$f(x,y) = \frac{xy^2}{30}\ , \quad x = 1,2,3, \quad y = 1,2\ .$$

Decide if $X$ and $Y$ are independent.

$S_X = \{1,2,3\}$

$S_Y = \{1,2\}$

$S = S_X \times S_Y$




Let us determine the marginal pmf s of $X$ and $Y$:

$$f_X(x) = \sum_{y=1}^{2} f(x,y) = \sum_{y=1}^{2} \frac{xy^2}{30} = \frac{x}{30} + \frac{4x}{30} = \frac{5x}{30} = \frac{x}{6}\ , \quad x = 1,2,3$$

$$f_Y(y) = \sum_{x=1}^{3} f(x,y) = \sum_{x=1}^{3} \frac{xy^2}{30} = \frac{y^2}{30} + \frac{2y^2}{30} + \frac{3y^2}{30} = \frac{6y^2}{30} = \frac{y^2}{5}\ , \quad y = 1,2$$

We conclude that $X$ and $Y$ are independent since

$$f_X(x) \cdot f_Y(y) = \frac{x}{6} \cdot \frac{y^2}{5} = \frac{xy^2}{30} = f(x,y) \text{ for all } \underbrace{x = 1,2,3}_{x \in S_X} \text{ and } \underbrace{y = 1,2}_{y \in S_Y}$$

**Notes**:

- If the support $S$ of $X$ and $Y$ is not the product set

$$\{(x, y) : x \in S_X, y \in S_Y\} \ , \quad \leftarrow \text{ "rectangular lattice"}$$

  then $X$ and $Y$ must be dependent (no need to check any other condition).

- If the support $S$ of $X$ and $Y$ is equal to the product set

$$\{(x, y) : x \in S_X, y \in S_Y\} \ , \quad \leftarrow \text{ "rectangular lattice"}$$

  the condition

$$f(x, y) = f_X(x) f_Y(y) \quad \text{for all } x \in S_x \text{ and } y \in S_y \ .$$

  must still be checked to decide whether $X$ and $Y$ are independent.

# Mathematical expectation

**Definition (Mathematical expectation)**

Let $X_1$ and $X_2$ be random variables of the discrete type with the joint pmf $f(x_1, x_2)$ on the space $S$, and let $u(X_1, X_2)$ be a function of these two random variables.

The *mathematical expectation* (or *expected value*) of $u(X_1, X_2)$ is given by

$$E[u(X_1, X_2)] = \sum_{(x_1,x_2)\in S} u(x_1, x_2)\, f(x_1, x_2) \ ,$$

double summation over $(x_1, x_2) \in S$

provided the sum on the right is absolutely convergent.

**Note**: If $Y = u(X_1, X_2)$ is a random variable with pmf $g(y)$ on the space $S_Y$ it holds that

$$E[u(x_1,x_2)] = \sum_{(x_1,x_2)\in S} u(x_1, x_2)\, f(x_1, x_2) = \sum_{y\in S_Y} yg(y) = E[Y]$$

**Special cases**:

(a) If $u(X_1, X_2) = X_i$, $i = 1, 2$, then

$$E[u(X_1, X_2)] = E[X_i] = \mu_i$$

is called the mean of $X_i$.

(b) If $u(X_1, X_2) = (X_i - \mu_i)^2$, $i = 1, 2$, then

$$E[u(X_1, X_2)] = E\left[(X_i - \mu_i)^2\right] = \mathrm{Var}(X_i) = \sigma_i^2$$

is called the variance of $X_i$.

**Note**: The mean $\mu_i$ and the variance $\sigma_i^2$ can be computed from either the joint pmf $f(x_1, x_2)$ or the marginal pmf $f_i(x_i)$, $i = 1, 2$. For instance:

$$\mu_X = \underbrace{E(X) = \sum_x \sum_y x f(x, y)}_{\text{bivariate def of expectation}} = \sum_x x \left[\sum_y f(x, y)\right] = \underbrace{\sum_x x \overset{\text{marginal pmf of } X}{f_X(x)} = E[X]}_{\text{univariate def}}$$

## Example

There are eight similar chips in a bowl: three marked $(0,0)$, two marked $(1,0)$, two marked $(0,1)$, and one marked $(1,1)$.

A player selects a chip at random and is given the sum of the two coordinates in dollars. Determine the expected payoff.

| y \ x | 0 | 1 |
|---|---|---|
| 0 | 3/8 | 2/8 |
| 1 | 2/8 | 1/8 |

joint pmf

$$f(x,y) = \begin{cases} 3/8 & \text{if } (x,y) = (0,0) \\ 2/8 & \text{if } (x,y) \in \{(1,0),(0,1)\} \\ 1/8 & \text{if } (x,y) = (1,1) \end{cases}$$

Player is given $X+Y$

The expected payoff is

$$E[X+Y] = \sum_{(x,y)\in S} (x+y)\cdot f(x,y) = 0\cdot f(0,0) + 1\cdot f(0,1) + 1\cdot f(1,0) + 2\cdot f(1,1)$$

$$= 0 + \frac{2}{8} + \frac{2}{8} + 2\cdot\frac{1}{8} = \frac{6}{8}$$

where $S = \{(0,0),(0,1),(1,0),(1,1)\}$

Harder approach to this example:

1. compute marginals $f_X(x)$ and $f_Y(y)$
2. evaluate $E[X]$ from $f_X(x)$ and $E[Y]$ from $f_Y(y)$
3. sum to get $E[X+Y] = E[X] + E[Y]$

# Math 3501 - Probability and Statistics I

## 4.2 - Covariance and the correlation coefficient

## Covariance and the correlation coefficient

### Definition (Covariance and the correlation coefficient)

Let $X$ and $Y$ be two random variables.

(a) The *covariance of $X$ and $Y$*, denoted either as $\text{Cov}(X, Y)$ or $\sigma_{XY}$, is given by

$$\text{Cov}(X, Y) = E\left[(X - \mu_X)(Y - \mu_Y)\right] ,$$

if $X=Y$ then $\text{Cov}(X,X) = E\left[(X-\mu_X)^2\right] = \text{Var}(X)$

provided the mathematical expectation on the right hand side exists.

(b) Suppose the standard deviations $\sigma_X$ and $\sigma_Y$ are positive and the covariance $\sigma_{XY}$ exists. The *correlation coefficient of $X$ and $Y$*, denoted $\rho$, is defined as

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

**Note**: Recall that

$$\mu_X = E(X) \qquad \mu_Y = E(Y)$$

and

definition

$$\sigma_X^2 = E\left[(X - \mu_X)^2\right] \qquad \sigma_Y^2 = E\left[(Y - \mu_Y)^2\right] .$$

$$\text{Var}(X) \overset{\text{def}}{=} E\left[(X-\mu_X)^2\right] = \underbrace{E[X^2] - (E[X])^2}$$

IMPORTANT PROPERTY

**Properties**:

(a) $\text{Cov}(X, Y) = E[XY] - E[X]E[Y] = E[XY] - \mu_X \mu_Y$.

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

(b) Since $\rho = \text{Cov}(X, Y)/\sigma_X \sigma_Y$, we also have $E[XY] = \mu_X \mu_Y + \rho \sigma_X \sigma_Y$.

$$Cov(X,Y) \overset{def}{=} E[(X-\mu_X)(Y-\mu_Y)] = E[XY - \mu_X Y - \mu_Y X + \mu_X \mu_Y]$$

$$= E[XY] - \mu_X \underbrace{E[Y]}_{\mu_Y} - \mu_Y \underbrace{E[X]}_{\mu_X} + \mu_X \mu_Y$$

$$= E[XY] - \mu_X \cdot \mu_Y = E[XY] - E[X] \cdot E[Y]$$

if we set $X = Y$, we get the formula for variance $\underbrace{Var(X)}_{Cov(X,X)} = E[X^2] - (E[X])^2$

$$\rho = \frac{Cov(X,Y)}{\sigma_X \cdot \sigma_Y} \Rightarrow Cov(X,Y) = \rho\, \sigma_X \cdot \sigma_Y \Rightarrow \underbrace{E[XY] - \overbrace{E[X]}^{\mu_X} \cdot \overbrace{E[Y]}^{\mu_Y}}_{Cov(X,Y)} = \rho\, \sigma_X \sigma_Y \Rightarrow$$

$$\Rightarrow E[XY] = \mu_X \cdot \mu_Y + \rho\, \sigma_X \sigma_Y$$

## Example

Let $X$ and $Y$ have the joint pmf

$$f(x,y) = \frac{x+2y}{18}, \quad x = 1,2, \quad y = 1,2\ . \quad \leftarrow \quad S = \{(1,1),(1,2),(2,1),(2,2)\}$$

Determine the covariance of $X$ and $Y$ and the corresponding correlation coefficient.

Recall that $Cov(X,Y) = E[XY] - E[X]\cdot E[Y]$ $\leftarrow$ item a) from previous slide

① $E[XY] = \sum_{(x,y)\in S} x\cdot y\cdot f(x,y) = 1\cdot 1\cdot f(1,1) + 1\cdot 2\cdot f(1,2) + 2\cdot 1\cdot f(2,1) + 2\cdot 2\cdot f(2,2)$

$= \frac{3}{18} + 2\cdot\frac{5}{18} + 2\cdot\frac{4}{18} + 4\cdot\frac{6}{18} = \frac{45}{18}$

② $E[X] = \sum_{(x,y)\in S} x\cdot f(x,y) = 1\cdot f(1,1) + 1\cdot f(1,2) + 2\cdot f(2,1) + 2\cdot f(2,2)$

$= \frac{3}{18} + \frac{5}{18} + 2\cdot\frac{4}{18} + 2\cdot\frac{6}{18} = \frac{28}{18}$

(3) $E[Y] = \sum_{(x,y)\in S} y \cdot f(x,y) = 1 \cdot f(1,1) + 2 \cdot f(1,2) + 1 \cdot f(2,1) + 2 \cdot f(2,2)$

$$= \frac{3}{18} + 2 \cdot \frac{5}{18} + 1 \cdot \frac{4}{18} + 2 \cdot \frac{6}{18}$$

$$= \frac{29}{18}$$

and so $\mathrm{Cov}(X,Y) = E[XY] - E[X] \cdot E[Y] = \frac{45}{18} - \frac{28}{18} \cdot \frac{29}{18} = -\frac{1}{162}$

To compute the correlation coefficient:

$$\rho = \frac{\mathrm{Cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{-1/162}{\sigma_X \sigma_Y}$$

where

$$\sigma_X = \sqrt{\mathrm{Var}(X)} = \sqrt{E[X^2] - (E[X])^2}$$

$$\sigma_Y = \sqrt{\mathrm{Var}(Y)} = \sqrt{E[Y^2] - (E[Y])^2}$$

complete evaluations as HW.

double check

$$= \frac{-1/162}{\sqrt{20/81}\sqrt{77/324}} = \cdots$$

## Proposition

*Let X and Y be independent random variables and suppose that u and v are real-valued functions such that both E[u(X)] and E[v(Y)] exist.*

*We have that*

$$E[u(X)v(Y)] = E[u(X)]E[v(Y)] \ .$$

**Note**: The statement above holds in greater generality, not just for discrete random variables.

in particular / consequence : if X and Y are independent, then

$$E[XY] = E[X] \cdot E[Y]$$

follows from Prop by taking $u(x) = X$ and $v(Y) = Y$

Proof :

$$E[u(X) \cdot v(Y)] = \sum_{(x,y) \in S} u(x) \cdot v(y) \cdot f(x,y) \overset{\text{independence}}{=} \sum_{x \in S_X} \sum_{y \in S_Y} u(x) v(y) \cdot f_X(x) \cdot f_Y(y)$$

$$= \sum_{x \in S_X} \Big[ u(x) \cdot f_X(x) \cdot \underbrace{\sum_{y \in S_Y} v(y) \cdot f_Y(y)}_{E[v(Y)]} \Big] = E[v(Y)] \cdot \underbrace{\sum_{x \in S_X} u(x) \cdot f_X(x)}_{E[u(X)]} = E[u(X)] \cdot E[v(Y)]$$

## Proposition

*Let $X$ and $Y$ be random variables with finite means $\mu_X$ and $\mu_Y$.*

*If $X$ and $Y$ are independent, then* $\text{Cov}(X, Y) = 0$.

Proof: $X, Y$ independent $\overset{\text{Previous prop}}{\Longrightarrow}$ $E[XY] = E[X] \cdot E[Y]$ and so

$$\text{Cov}(X,Y) = E[XY] - E[X] \cdot E[Y] = 0$$

**Note**: The converse to the statement above does not hold: two dependent random variables may have zero covariance:

- Take $X$ uniformly distributed on $\{-1, 0, 1\}$ and $Y = X^2$. We have that $\text{Cov}(X, Y) = 0$ even though $X$ and $Y$ are not independent.
- Also check next example.

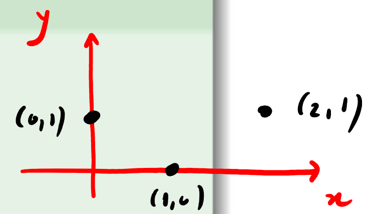
check that $\text{Cov}(X,Y) = 0$ on HW

## Example

Let $X$ and $Y$ have the joint pmf

$$f(x,y) = \frac{1}{3}, \quad (x,y) = (0,1), (1,0), (2,1)\ .$$

Check that $\mathrm{Cov}(X,Y) = 0$ despite $X$ and $Y$ not being independent.




$S \neq S_X \times S_Y$

where $S_X = \{0,1,2\}$

and $S_Y = \{0,1\}$

Since $S \neq S_X \times S_Y$ then $X$ and $Y$ are **NOT** independent.

Despite that, we have:

$$E[XY] = \sum_{(x,y)\in S} x \cdot y \cdot f(x,y) = \overbrace{0\cdot 1 \cdot f(0,1)}^{0} + \overbrace{1 \cdot 0 \cdot f(1,0)}^{0} + 2\cdot 1\, f(2,1) = \frac{2}{3}$$

$$E[X] = \sum_{(x,y)\in S} x \cdot f(x,y) = 0\cdot f(0,1) + 1\cdot f(1,0) + 2\cdot f(2,1) = \frac{1}{3} + \frac{2}{3} = 1$$

$$E[Y] = \sum_{(x,y)\in S} y \cdot f(x,y) = 1\cdot f(0,1) + 0\cdot f(1,0) + 1\cdot f(2,1) = \frac{1}{3} + \frac{1}{3} = \frac{2}{3}$$

The covariance of $X$ and $Y$ is

$$\mathrm{Cov}(X,Y) = E[XY] - E[X]\cdot E[Y]$$

$$= \frac{2}{3} - 1\cdot\frac{2}{3} = 0$$

Despite $X$ and $Y$ being dependent

if we know that $\mathrm{Cov}(X,Y) \neq 0$ then $X$ and $Y$ are not independent

REMEMBER:

$X, Y$ independent $\Longrightarrow$ $\mathrm{Cov}(X,Y) = 0$

BUT

$\mathrm{Cov}(X,Y) = 0$ $\not\Longrightarrow$ $X, Y$ independent