9) Joint Probability Distributions

9.1 Joint distributions of discrete random variables

Example: 9.1 Consider an experiment of flipping two coins.

The sample space is

$$\Omega = \{(H,H), (H,T), (T,H), (T,T)\}$$

All these outcomes are equally likely. Now introduce two random variables:

N: equal to total number of heads

S: 1 if second coin shows heads
0 otherwise

This summarised in following table:

ω	(H,H)	(H,T)	(T,H)	(T,T)
S	1	0	1	0
N	2	1	1	0

Calculating probability mass function of S:

Image: $S(\Omega) = \{0,1\}$

$$\{S=0\} = \{(H,T),(T,T)\}$$

$$\{S=1\} = \{(H,H),(T,H)\} = \{S=0\}^c$$

and thus

$$p_S(0) = P(S=0) = P(\{(H,T),(T,T)\}) = 1/2$$

$$p_S(1) = P(S=1) = P(\{(H,H),(T,H)\}) = 1/2$$

$$p_S(s) = \begin{cases} 1/2 & \text{if } s \in \{0,1\} \\ 0 & \text{otherwise} \end{cases}$$

Calculating mass function for N

Image: $N(\Omega) = \{0, 1, 2\}$

$\{N=0\} = \{(T,T)\}$

$\{N=1\} = \{(H,T), (T,H)\}$

$\{N=2\} = \{(H,H)\}$

Thus

$p_N(0) = p_N(\{T,T\}) = 1/4$

$p_N(1) = p_N(\{(H,T), (T,H)\}) = 1/2$

$p_N(2) = p_N(\{H,H\}) = 1/4$

$$p_N(n) = \begin{cases} 1/4 & n \in \{0,2\} \\ 1/2 & n \in \{1\} \\ 0 & n \notin \{0,1,2\} \end{cases}$$

But now we can also consider events defined in terms of both random variables, simultaneously, like

$$\{S=0 \text{ and } N=1\}$$

For convenience replace "and" with comma ","

$$\{S=0,\ N=1\}$$

We find

$$\begin{aligned}\{S=0, N=1\} &= \{S=0\} \cap \{N=1\} \\ &= \{(H,T),(T,T)\} \cap \{(H,T),(T,H)\} \\ &= \{(H,T)\}\end{aligned}$$

And thu

$$P(S=0, N=1) = P(\{H,T\}) = 1/4$$

Defn 9.2: Given two discrete random variables X and Y, function

$p_{X,Y} : \mathbb{R}^2 \to \mathbb{R}$ is defined by

$$\boxed{\begin{aligned} p_{X,Y}(x,y) &= P(X^{-1}(x) \cap Y^{-1}(y)) \\ &= P(\{X=x\} \cap \{Y=y\}) \\ &= P(X=x, Y=y) \end{aligned}}$$

Example 9.1 (continued): As calculated before

$$p_{S,N}(0,1) = P(S=0, N=1) = P(\{H,H\}) = 1/4$$

Calculating others similarly using defn 9.2

$$p_{S,N}(0,0) = P(S=0, N=1) = P(\{H,T\}) = 1/4$$

$$p_{S,N}(1,1) = P(S=1, N=1) = P(\{T,H\}) = 1/4$$

$$p_{S,N}(1,2) = P(S=1, N=2) = P(\{H,H\}) = 1/4$$

$$p_{S,N}(0,2) = p_{S,N}(1,0) = P(\emptyset) = 0$$

And clearly

$p_{S,N}(s,n) = 0$ if $s \notin S(\Omega)$ and $n \in N(\Omega)$

The values of $p_{S,N}$ can be expressed in a table.

		n		
		0	1	2
S	0	1/4	1/4	0
	1	0	1/4	1/4

It is convenient to include

		n			
		0	1	2	$p_S(s)$
S	0	1/4	1/4	0	1/2
	1	0	1/4	1/4	1/2
$p_N(n)$		1/4	1/2	1/4	1

Because of the convention of displaying the mass function ~~in~~ of individual random variables in margins, they are also often referred to as the marginal probability mass fn.

Theorem 9.3: Let X and Y be discrete random variables. The probability mass functions of X and Y can be obtained as

$$\boxed{p_X(x) = \sum_{y \in Y(\Omega)} p_{XY}(x,y)}$$

$$\boxed{p_Y(y) = \sum_{x \in X(\Omega)} p_{X,Y}(x,y)}$$

proof: This is just a consequence of the fact that the collection of events

$$\{\{Y=y\} \mid y \in Y(\Omega)\}$$

is a partition of the sample space. i.e.

$$\bigcup_{y \in Y(\Omega)} \{Y = y_x\} = \Omega \qquad \text{and}$$

$$\{Y = y_1\} \cap \{Y = y_2\} = \phi \quad \text{if} \quad y_1 \neq y_2$$

Thus we can write the event $\{X = x\}$ as a disjoint union,

$$\{X = x\} = \{X = x\} \cap \Omega$$

distribution law

$$= \left(\{X = x\} \cap \bigcup_{y \in Y(\Omega)} \{Y = y\} \right)$$

$$= \bigcup_{y \in Y(\Omega)} \left(\{X \in x\} \cap \{Y \in y\} \right)$$

disjoint union, look at chapter 2 for explanation

applying axiom (P3) to this

Therefore by axiom (P3)

$$p_X(x) = P\left(\bigcup_{y \in Y(\Omega)} \left(\{X = x\} \cap \{Y = y\} \right) \right)$$

$$= \sum_{y \in Y(\Omega)} P\left(\left(\{X = x\} \cap \{Y = y\} \right) \right)$$

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

$$= \sum_{y \in Y(\Omega)} p_{XY}(x, y)$$

The second identity follows similarly with X and Y interchanged.

$\blacksquare$

Joint mass functions have 2 defining properties.

Properties of Joint mass functions :

(jm1): $p_{XY}(x,y) \geq 0 \quad \forall x,y \in \mathbb{R}$

In particular,

$p_{XY}(x,y) = 0$ unless $x \in X(\Omega)$ and $y \in Y(\Omega)$

(jm2): $$\sum_{x \in X(\Omega)} \sum_{y \in Y(\Omega)} p_{XY}(x,y) = 1$$

Next introduce the joint distribution function as an alternative way of specifying probability distributions.

This has an advantage over probability mass function as it will also work for continuous random variables.

Defn 9.4: Let X and Y be random variables. The **joint distribution function** of X and Y is the function $F_{X,Y}:\mathbb{R}^2 \to \mathbb{R}$ defined by

$$\boxed{\begin{aligned} F_{X,Y}(x,y) &= P\left(X^{-1}(-\infty,x] \cap (-\infty,y]\right) \\ &= P(X \le x, Y \le y) \end{aligned}}$$

Example: 9.1 (continued) The joint distribution function of S and N is

$$F_{S,N}(s,n) = P(s \le S, n \le N) = \begin{cases} 0 & \text{if } s<0 \text{ or } n<0 \\ 1/4 & \text{if } 0 \le s \text{ or } 0 \le n < 1 \\ 1/2 & \text{if } 0 \le s < 1 \text{ or } 1 \le n \\ 3/4 & \text{if } 1 \le s, 1 \le n < 2 \\ 1 & \text{if } 1 \le s, 2 \le n \end{cases}$$

The joint distribution function of 2 discrete random variables is a two-dimensional step function.

fixing n, one sided inequality for n.

Using tables to determine distribution function:

$P(1 \le S, 1 \le n < 2)$

$P(n \ge 0, 1 \le s < 2)$

$= P(N=0, S=0) + P(N=0, S=1)$

$= 1/4 + 0 = 1/4$

$P(1 \le n < 2, S \ge 1)$ or $P(1 \le n, 1 \le s < 2)$

$= P(N=1, S=0) + P(N=1 + S=1)$

$+ P(N=0, S=0)$

$+ P(N=0, S=1)$

$= \frac{1}{4} + \frac{1}{4} + \frac{1}{4} + 0$

$= 3/4$

$P(S \ge 1, 2 \le n < 3)$ or $P(n \ge 2, 1 \le S < 2)$

$= P(S=0, N=0) + P(S=0, N=1)$

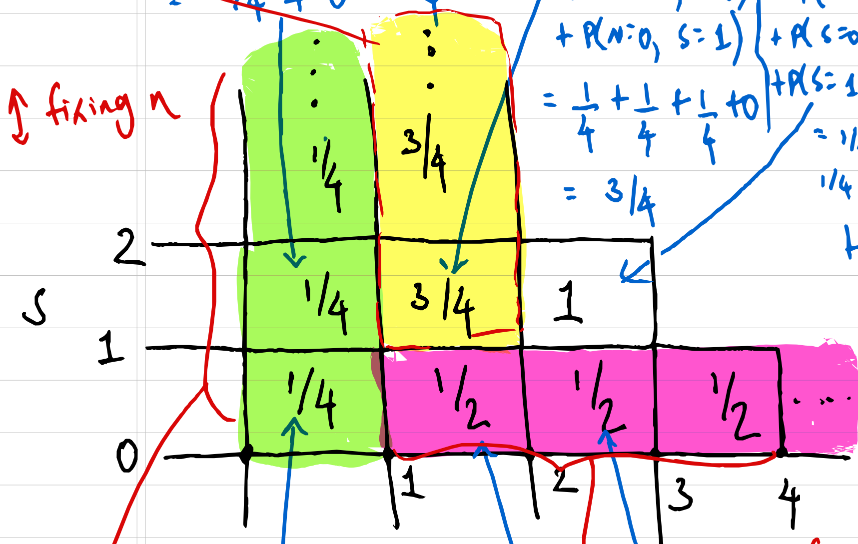
$+ P(S=0, N=2) + P(S=1, N=0)$

$+ P(S=1, N=1) + P(S=1, N=2)$

$= 1/4 + 1/4 + 0 + 0 +$

$1/4 + 1/4 = 1$

Here left side line inclusive right side line exclusive



$P(0 \le S, 0 \le 1 < n)$ n

or

$P(0 \le n, 0 \le S < 1) = \frac{1}{4}$

$P(0 \le S, 2 \le n < 3)$

$= P(S=0, N=1) + P(S=0, N=1) + P(S=0, N=2)$

$= 1/4 + 1/4 + 0 = 1/2$

$P(S \le 0, 1 \le n < 2) = P(S=0 \cap (N=0 \cup N=1))$

$= \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$

$= P(S=0, N=0)$

$+ P(S=0, N=1)$

$P(S \ge 0, 0 \le n < 1)$

$P(n \ge 1, 0 \le S < 1)$

To get marginal distribution function from joint distribution function, use the theorem.

Theorem 9.5: Let X and Y be random variables and let $F_{X,Y}$ be be their joint distribution function. Then their (marginal) distribution function can be obtained as

$$\boxed{\begin{aligned} F_X(x) &= \lim_{y\to\infty} F_{X,Y}(x,y) \\ F_Y(y) &= \lim_{x\to\infty} F_{X,Y}(x,y) \end{aligned}}$$

This is true because

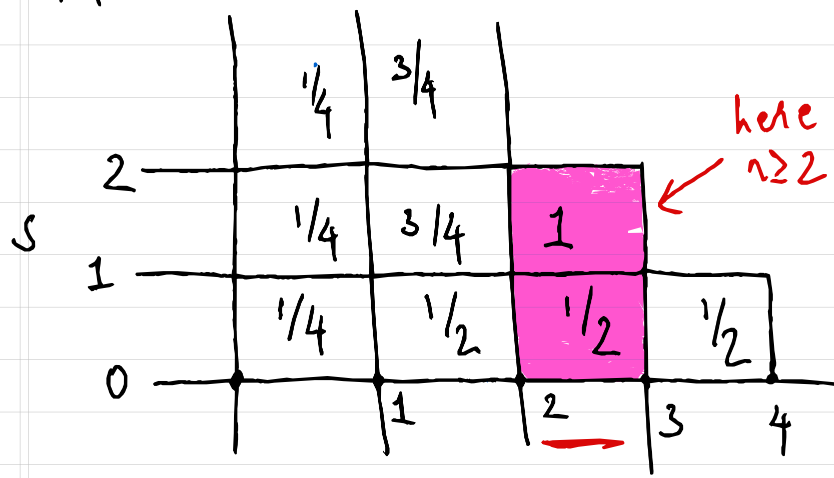
$$\begin{aligned} F_X(x) = P(X \le x) &= P(X \le x, Y \le \infty) \\ &= \lim_{y\to\infty} P(X \le x, Y \le y) \\ &= \lim_{y\to\infty} F_{X,Y}(x,y) \end{aligned}$$

Similarly for F_Y

Example 9.1 (continued)

$$F_S(s) = \lim_{n\to\infty} F_{S,N}(s,n) = F_{S,N}(s,2)$$

because $F_{S,N}(s,n) = F_{S,N}(s,2) \ \forall \ n \geq 2$, we find



here values $n \geq 2$ of $F_S(s) = F_{S,N}(s,2)$

$$F_S = \begin{cases} 0 & s < 0 \\ 1/2 & 0 \leq s < 1 \\ 2 & s \geq 1 \end{cases}$$

Reasoning

$$F_S(s) = \lim_{n\to\infty} F_{S,N}(s,n) = \lim_{n\to\infty} P(S \leq s,\ N \leq n)$$

(since $P(N \leq \infty) = 1$
$= P(N \leq 2)$)

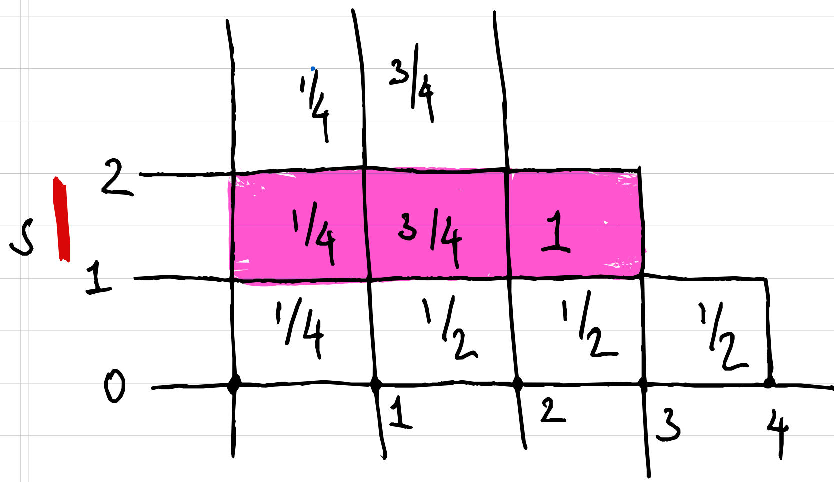
$$= P(S \leq s,\ N \leq \infty)$$

$$= P(S \leq s,\ N \leq 2)$$

Similarly:

$$F_N(n) = \lim_{s\to\infty} F_{S,N}(s,n) = F_{S,N}(1,n)$$

because $F_{S,N}(s,n) = F_{S,N}(1,n)$ $\underline{\forall s \geq 1}$, we find,



So

$$F_N(n) = \begin{cases} 0 & \text{if } n < 0 \\ 1/4 & \text{if } 0 \leq n < 1 \\ 3/4 & \text{if } 1 \leq n < 2 \\ 1 & \text{if } n \geq 2 \end{cases}$$

9.2 Joint distributions of continuous random variables

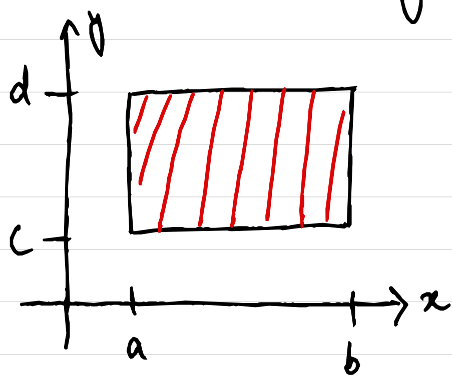
<u>Defn 9.6:</u> We call two random variables X and Y **jointly continuous** if their joint distribution function $F_{X,Y}$ can be written as

$$\boxed{F_{X,Y}(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f_{X,Y}(\hat{x},\hat{y})\, d\hat{y}\, d\hat{x}} \quad \underline{\forall x,y \in \mathbb{R}}$$

for some nonnegative function $\underline{f_{X,Y} : \mathbb{R}^2 \to \mathbb{R}}$.

In this case $f_{X,Y}$ is the joint density function of X and Y.

Example 9.7: Consider the uniform distribution on a rectangle, where <u>probability density is evenly spread over the rectangle.</u>



So the area of rectangle is $(b-a).(d-c)$

The x co-ordinate of a point in rectangle lies in $[a,b]$

The y co-ordinate of a point in rectangle lies in $[c,d]$

Because area is $(b-a)(d-c)$, area is spread evenly, the density function is

$$f_{X,Y}(x,y) = \begin{cases} \dfrac{1}{(b-a)(d-c)} & \text{if } x \in [a,b] \text{ and } y \in [c,d] \\ 0 & \text{otherwise} \end{cases}$$

The joint distribution function for jointly continuous random variables is continuous and even differentiable everywhere. The fundemental theorem of calculus implies (under some mild regularity conditions) that

$\forall (x,y) \in \mathbb{R}^2$,

$$\boxed{\frac{d}{dx}\frac{d}{dy} F_{X,Y}(x,y) = f_{X,Y}(x,y)}$$

Joint density functions have 2 properties.

Properties of Joint density functions:

(jd1) $\quad f_{X,Y}(x,y) \geq 0 \quad \forall\, x,y \in \mathbb{R}$

(jd2) $\quad \displaystyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\, dy\, dx = 1$

Theorem 9.8: If X and Y are jointly continuous random variables with joint density function $f_{X,Y}$ then

$$\boxed{\forall\, a_1, a_2, b_1, b_2 \in \mathbb{R} \quad P(a_1 \leq X \leq b_1,\ a_2 \leq Y \leq b_2) = \int_{a_1}^{b_1}\int_{a_2}^{b_2} f_{X,Y}(x,y)\,dy\,dx}$$

Weak inequality can be replaced with strict inequality on the LHS of the above equation.

Theorem 9.9: Let $f_{X,Y}$ be the joint density of X and Y. Then their (marginal) density functions are

$$\boxed{f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy}$$

$$\boxed{f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx}$$

Example: 9.10 Suppose X and Y have joint density function

$$f_{X,Y}(x,y) = \begin{cases} x e^{-x-y} & \text{for } x \geq 0,\ y \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

The joint distribution function is

$$F_{X,Y}(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(\hat{x},\hat{y})\, d\hat{y}\, d\hat{x}$$

$$= \int_{0}^{x} \int_{0}^{y} \hat{x} e^{-\hat{x}-\hat{y}}\, d\hat{y}\, d\hat{x}$$

$$= \int_{0}^{x} \hat{x} e^{-\hat{x}} \int_{0}^{y} e^{-\hat{y}}\, d\hat{y}\, d\hat{x}$$

$$= (1-e^{-y}) \int_{0}^{x} \hat{x} e^{-\hat{x}}\, dx$$

$$= (1-e^{-y})(1-(1+x)e^{-x})$$

$$= 1-(1+x)e^{-x} - e^{-y} + (1+x)e^{-x-y}$$

We can check our calculation of distribution function by using FTC

$$\frac{d}{dx}\frac{d}{dy} F_{X,Y}(x,y) = \frac{d}{dx}\frac{d}{dy}\left(1-(1+x)e^{-x} - e^{-y} + (1+x)e^{-x-y}\right)$$

$$= \frac{d}{dx}\left(e^{-y} - (1+x)^{-x-y}\right)$$

$$= -e^{-x-y} + (1+x)e^{-x-y}$$

$$= xe^{-x-y} = f_{X,Y}(x,y)$$

Using theorem 9.5, we obtain marginal distribution functions.

$$F_X(x) = \lim_{y\to\infty} F_{X,Y}(x,y)$$

$$= \lim_{y\to\infty}\left(1-(1+x)e^{-x} - e^{-y} + (1+x)e^{-x-y}\right)$$

$$= 1-(1+x)e^{-x}$$

$$F_Y(y) = \lim_{x\to\infty} F_{X,Y}(x,y)$$

$$= \lim_{x\to\infty} \left(1-(1+x)e^{-x} - e^{-y} + (1+x)e^{-x-y}\right)$$

$$= 1-e^{-y}$$

Use <u>theorem 9.9</u> to obtain marginal density functions.

For example

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

$$= \int_0^{\infty} xe^{-x-y}\, dy$$

$$= xe^{-x}\int_0^{\infty} e^{-y}\, dy = xe^{-x}$$

To check this, we can also obtain density function as a derivative of the distribution function

$$f_X(x) = \frac{d}{dx}\left(1 + (1+x)e^{-x}\right) = -e^{-x} + (1+x)e^{-x}$$

$$= xe^{-x}$$

We get the same result as must be the case.

9.3 More than 2 random variables

Anything we have done till now for 2 random variables generalises to any number of random variables.

9.4 Independant Random variables :

Defn: We call 2 random variables independant if knowing one of them tells us nothing about the value about the other.

Defn 9.11: Two random variables X and Y are independant denoted

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

if

$$\boxed{F_{X,Y}(x,y) = F_X(x) \cdot F_Y(y) \qquad \forall x, y \in \mathbb{R}}$$

You can also check independance by mass functions and density functions.

Theorem 9.2: If X and Y are discrete random variables, they are independant if and only if

$$\boxed{p_{X,Y}(x,y) = p_X(x) \cdot p_Y(y) \qquad \forall x, y \in \mathbb{R}}$$

If X and Y are jointly continuous random variables they are independant if and only if

$$\boxed{f_{X,Y}(x,y) = f_X(x) \cdot f_Y(y) \qquad \forall x, y \in \mathbb{R}}$$

Example: 9.1 (continued) We observe from table 9.1 that for example

$$P_S(1).P_N(0) = \frac{1}{2}\cdot\frac{1}{4} \neq 0 = P_{S,N}(1,0)$$

This one counterexample is enough to show that X and Y are not independant.

Example 9.7: (continued) The density function of x-co-ordinate is

$$f_X(x) = \begin{cases} \frac{1}{b-a} & \text{if } x \in [a,b] \\ 0 & \text{otherwise} \end{cases}$$

Similarly for the y-coordinate

$$f_Y(y) = \begin{cases} \frac{1}{d-c} & \text{if } x \in [c,d] \\ 0 & \text{otherwise} \end{cases}$$

$$F_X(x).F_Y(y) = \begin{cases} \frac{1}{(b-a)(d-c)} & \text{if } x \in [a,b] \text{ and } x \in [c,d] \\ 0 & \text{otherwise} \end{cases}$$

$$= F_{X,Y}(x,y) \quad \forall x,y \in \mathbb{R}.$$

$\Rightarrow X \perp\!\!\!\perp Y$ Hence X and Y are independant.

Example: 9.10 (continued) We observe that

$$F_X(x) . F_Y(y) = \left(1-(1+x)e^{-x}\right) . \left(1-e^{-y}\right)$$
$$= F_{X,Y}(x,y)$$

$\forall\, x, y \in \mathbb{R}$. Hence X and Y are independant.

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

9.5 Propogation of independance

Theorem: 9.13 (Propogation of independance)

Let $X_1, X_2, \ldots, X_n$ be independant random variables and

$$h_1, h_2, h_3, \ldots, h_n : \mathbb{R} \to \mathbb{R}$$

be functions.

Then the random variables

$$h_1(X_1),\ h_2(X_2),\ \ldots,\ h(X_n)$$

are independant.