is an increasing sequence which does not converge to a finite number. In general, any sequence $((x_n, y_n) : n \geq 1)$ can be considered such that $\min(x_n, y_n) \uparrow \infty$. Then the same argument gives that

$$\lim_{n \to \infty} F(x_n, y_n) = 1.$$

This is equivalent to say

$$\lim_{\min(x,y) \uparrow \infty} F(x, y) = 1.$$

We now show that $\lim_{y \downarrow -\infty} F(x, y) = 0$. The same argument works to show $\lim_{x \downarrow -\infty} F(x, y) = 0$.

We can use again the continuity property of probability to establish that

$$\lim_{n \to \infty} F(x, -n) = \mathbb{P}\left(\bigcap_{n=1}^{\infty} \left((-\infty, x] \times (-\infty, n] \right) \right)$$

$$= \mathbb{P}(\{\phi\}) = 0$$

~~again~~ where we have used the fact that $\mathbb{P}(\{\phi\}) = 0$.

Finally, we have to show that

$$F(x_1 + \delta_1, x_2 + \delta_2) - F(x_1 + \delta_1, x_2) - F(x_1, x_2 + \delta_2) + F(x_1, x_2) \geq 0.$$

where $\delta_i > 0$ for $i=1,2$. Note that

$$\mathbb{P}\left(\{x_1 < X \le x_1+\delta_1\} \cap \{x_2 < Y \le x_2+\delta_2\}\right)$$

$$= \mathbb{P}\left(x_1 < X \le x_1+\delta_1 ;\ x_2 < Y \le x_2+\delta_2\right)$$

$$= \mathbb{P}\left((X,Y) \in (x_1, x_1+\delta_1] \times (x_2, x_2+\delta_2]\right)$$

$$= \mathbb{P}\left(\left(\{X \le x_1+\delta_1\} \setminus \{X \le x_1\}\right) \cap \{x_2 < Y \le x_2+\delta_2\}\right)$$

$$= \mathbb{P}\left(\{X \le x_1+\delta_1\} \cap \{X \le x_1\}^c \cap \{x_2 < Y \le x_2+\delta_2\}\right)$$

$$= \mathbb{P}\left[\left(\{X \le x_1+\delta_1\} \cap \{x_2 < Y \le x_2+\delta_2\}\right) \setminus \left(\{X \le x_1\} \cap \{x_2 < Y \le x_2+\delta_2\}\right)\right]$$

$$= \mathbb{P}\left(\{X \le x_1+\delta_1\} \cap \{x_2 < Y \le x_2+\delta_2\}\right) - \mathbb{P}\left(\{X \le x_1\} \cap \{x_2 < Y \le x_2+\delta_2\}\right)$$

$$= \mathbb{P}\left(\{X \le x_1+\delta_1\} \cap \{Y \le x_2+\delta_2\}\right) - \mathbb{P}\left(\{X \le x_1+\delta_1\} \cap \{Y \le x_2\}\right) - \mathbb{P}\left(\{X \le x_1\} \cap \{Y \le x_2+\delta_2\}\right) + \mathbb{P}\left(\{X \le x_1\} \cap \{X \le x_2\}\right)$$

$$\cancel{= F(x_1+\delta_1}$$

$$= \mathbb{F}(x_1+\delta_1, x_2+\delta_2) - \mathbb{F}(x_1+\delta_1, x_2) - \mathbb{F}(x_1, x_2+\delta_2)$$
$$- \mathbb{F}(x_1, x_2).$$

The claim follows as

$$(x_1, x_1+\delta_1] \times (x_2, x_2+\delta_2] \in \mathcal{B}(\mathbb{R}^2)$$

(check it yourself)

and so it follows that

$$\mathbb{P}\big((X, Y) \in (x_1, x_1+\delta_1] \times (x_2, x_2+\delta_2] \big) \geq 0.$$

Therefore, the claim follows.

In the proof, we have used the fact that

$$\big(\{X \leq x_1+\delta_1\} \cap \{x_2 < Y \leq x_2+\delta_2\} \big)$$
$$\setminus \big(\{X \leq x_1\} \cap \{x_2 < Y \leq x_2+\delta_2\} \big)$$
$$= \big(\{X \leq x_1+\delta_1\} \cap \{x_2 < Y \leq x_2+\delta_2\} \big)$$
$$\cap \big(\{X > x_1\} \cup \{Y \notin (x_2, x_2+\delta_2]\} \big)$$
$$= \big(\{X \leq x_1+\delta_1\} \cap \{x_2 < Y \leq x_2+\delta_2\} \cap \{X > x_1\} \big)$$
$$\cup \underbrace{\big(\{X \leq x_1+\delta_1\} \cap \{x_2 < Y \leq x_2+\delta_2\} \cap \{Y \notin (x_2, x_2+\delta_2]\} \big)}_{\phi}$$
$$= \{X \leq x_1+\delta_1\} \cap \{x_2 < Y \leq x_2+\delta_2\} \cap \{X > x_1\}$$

There are other ~~o~~ facts similar to this which have been used in the derivation. So I am omitting the other details and conclude the proof. ▨

2-dimensional random vector is called <u>bivariate random vector</u> and the corresponding distribution function is called <u>bivariate distribution function</u>.

In general, n-dimensional random vector is called <u>multivariate random vector</u> ~~or~~ ~~multivariate~~, and their distribution function is called <u>multivariate distribution function</u>.

<u>Bivariate discrete random vector.</u>

A bivariate random vector (X, Y) is said to be discrete type if it takes on pairs of values belonging to a countable set of pairs E with probability 1. The elements in E can be enumerated as

$$E = \{ (x_i, y_j) : i \in \mathbb{N},\ j \in \mathbb{N} \}.$$

Each member $((x_i, y_j))$ is called a point of jump of the random vector (X, Y) if.

$$p_{i,j} = \mathbb{P}\left(\{X = x_i\} \cap \{Y = y_j\}\right) > 0; \quad (i,j) \in \mathbb{N} \times \mathbb{N}$$

and $p_{i,j}$ is the mass of the bivariate distribution at the jump point (x_i, y_j). If (X, Y) is a bivariate random vector, ~~the~~ of discrete-type, then

$$\sum_{i\geq 1}^{\infty} \sum_{j\geq 1}^{\infty} p_{i,j} = \sum_{(x_i, y_j) \in E} \mathbb{P}(X = x_i ; Y = y_j) = 1.$$

The probabilities $(p_{i,j} : i \in \mathbb{N}, j \in \mathbb{N}) = (\mathbb{P}(X = x_i, Y = y_j) : (i,j) \in \mathbb{N} \times \mathbb{N})$ are called probability mass function. The set $E = \{(x_i, y_j) : \mathbb{P}(X = x_i, Y = y_j) > 0\}$ is called the support of the distribution function ~~of~~ $F_{X,Y} : \mathbb{R}^2 \to [0,1]$ of the vector (X, Y), where

$$F_{X,Y}(u, v) = \sum_{i\geq 1;\ x_i \leq u}^{\infty} \sum_{j\geq 1;\ y_j \leq v}^{\infty} \mathbb{P}(X = x_i, Y = y_j)$$

$$= \sum_{i\geq 1:\ x_i \leq u}^{\infty} \sum_{j\geq 1:\ y_j \leq v}^{\infty} p_{i,j}.$$

Theorem

A collection of non-negative numbers $(p_{i,j} : (i,j) \in \mathbb{N} \times \mathbb{N})$ satisfying $\sum_{i\geq 1}^{\infty} \sum_{j\geq 1}^{\infty} p_{i,j} = 1$ is the probability mass function of some random vector.

Proof: The proof is left as an exercise.

Hint: We can define the distribution function associated to the probability mass function $(p_{i,j} : (i,j) \in \mathbb{N} \times \mathbb{N})$ and check that the distribution function satisfies the necessary properties.

Definition (Bivariate random vector of continuous type).

A two-dimensional random vector is said to be of continuous type if there exists a non-negative function $f(\cdot,\cdot)$ such that for every pair $(x,y) \in \mathbb{R}^2$ we have

$$\mathbb{F}(x,y) = \int_{-\infty}^{x} du \int_{-\infty}^{y} dv \, f(u,v).$$

and $\mathbb{F} : \mathbb{R}^2 \to [0,1]$ is a distribution function.

The function $f : \mathbb{R}^2 \to [0,\infty)$ is called joint density function of (X,Y). If the ~~density~~ joint density function $f : \mathbb{R}^2 \to [0,\infty)$ is continuous, then

$$\frac{\partial^2 \mathbb{F}(x,y)}{\partial x \, \partial y} = f(x,y).$$

Theorem:

If $f : \mathbb{R}^2 \to [0,\infty)$ ~~kero~~ is a function satisfying $\int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy\, f(x,y) = 1$, then f is a joint probability density function.

Marginal mass/density function.

Let (X, Y) be a discrete bivariate random vector with probability mass function $(p_{i,j} : (i,j) \in \mathbb{N} \times \mathbb{N})$.

Then we have

$$\mathbb{P}(X = x_i) = \sum_{j=1}^{\infty} \mathbb{P}(X = x_i, Y = y_j) = \sum_{j=1}^{\infty} p_{i,j}$$

using countable additivity of probability.

Similarly it follows that

$$\mathbb{P}(Y = y_j) = \sum_{i=1}^{\infty} \mathbb{P}(X = x_i, Y = y_j) = \sum_{i=1}^{\infty} p_{i,j}.$$

Of course, $(\mathbb{P}(X = x_i) : i \in \mathbb{N})$ is a probability mass function of the random variable X and $(\mathbb{P}(Y = y_j) : j \in \mathbb{N})$ is the probability mass function of the random variable Y.

We call each of ~~them~~ them as marginal probability mass function.

From joint probability mass function, we can get the marginal probability mass function of each of the component of the random vector.

We can denote these by

$p_X(i) = \mathbb{P}(X = x_i)$ for every $i \geq 1$.

and $p_Y(j) = \mathbb{P}(Y = y_j)$ for every $j \geq 1$.

If (X, Y) is a bivariate random vector of continuous type ~~then~~ with joint probability density function ~~$(f(x,y)$ $x \in \mathbb{R}$~~ $f_{X,Y} : \mathbb{R}^2 \to [0, \infty)$, then

$$f_X(x) = \int_{-\infty}^{\infty} dy\, f_{X,Y}(x, y) \text{ for every } x \in \mathbb{R}$$

is a probability density function. Similarly

$$f_Y(y) = \int_{-\infty}^{\infty} dx\, f_{X,Y}(x, y) : \mathbb{R} \to [0, \infty)$$

is also a probability density function.

We call $f_X : \mathbb{R} \to [0, \infty)$ and $f_Y : \mathbb{R} \to [0, \infty)$ to be the marginal density function of the random variables X and Y respectively.

Definition (Marginal distribution function)

Let (X, Y) be a random vector with distribution function $\mathbb{F}$. Then the marginal distribution function $\mathbb{F}_X : \mathbb{R} \to [0,1]$ of X and $\mathbb{F}_Y : \mathbb{R} \to [0,1]$ of Y are defined through

$$\mathbb{F}_X(x) = \lim_{y \uparrow \infty} \mathbb{F}(x, y) \quad \text{for every } x \in \mathbb{R}$$

and $$\mathbb{F}_Y(y) = \lim_{x \uparrow \infty} \mathbb{F}(x, y) \quad \text{for every } y \in \mathbb{R}.$$

If (X, Y) ~~be a bivaria~~ is a bivariate random vector of discrete type with probability mass function $(p_{i,j} : (i, j) \in \mathbb{R})$, then

$$\mathbb{F}_X(x) = \sum_{i \geq 1 : x_i \leq x}^{\infty} \sum_{j \geq 1}^{\infty} p_{i,j} = \sum_{i \geq 1 : x_i \leq x}^{\infty} \sum_{j \geq 1} \mathbb{P}(X = x_i, Y = y_j)$$

$$\mathbb{F}_Y(y) = \sum_{i \geq 1}^{\infty} \sum_{j \geq 1 : y_j \leq y}^{\infty} p_{i,j} = \sum_{i \geq 1}^{\infty} \sum_{j \geq 1 : y_j \leq y}^{\infty} \mathbb{P}(X = x_i, Y = y_j).$$

for every $x \in \mathbb{R}$ and $y \in \mathbb{R}$.

Similarly, if (X, Y) is a bivariate random vector of continuous type, then

$$\mathbb{F}_X(x) = \int_{-\infty}^{x} du \int_{-\infty}^{\infty} dv \, f(u, v) \quad \text{for every } x \in \mathbb{R}$$

and $F_Y(y) = \int_{-\infty}^{y} dv \int_{-\infty}^{\infty} du\, f(u,v)$ for every $y \in \mathbb{R}$.

Multivariate random ~~vector variables~~ elements.

- For an n-dimensional ~~multivariate~~ random vector $(X_1, \ldots, X_n)$, we can define k-dimensional distribution function of $(X_{i_1}, X_{i_2}, \ldots X_{i_k})$ for some $k \le n$ and $(i_1, i_2, \ldots, i_k) \in [1:n]^k$ where $[1:n] = \{1, 2, 3, \ldots, n\}$.

$$F_{X_{i_1}, X_{i_2}, \ldots X_{i_k}}(x_{i_1}, x_{i_2}, x_{i_3}, \ldots, x_{i_k}) = \lim_{\substack{\min x_j \uparrow \infty \\ j \in [1:n] \setminus \{i_1, i_2, \ldots, i_k\}}} F_{X_1, \ldots X_n}(x_1, x_2, \ldots, x_n)$$

- We can easily generalize the definitions of bivariate random vector of discrete type and continuous type to n-dimensional random vector of discrete type and continuous type respectively. We can also generalize the concept of joint probability mass and density function and their connections to joint distribution function.

* Here we may have to use concepts from "calculus of several variables".