ISEG - Lisbon School of Economics and Management
Statistics I
(Date of this version: 1/04/2018)
Handout 3 – Multivariate Random Variables

Issues covered:

- 3.1 Multivariate random variables
- 3.2 Joint cumulative distribution function
- 3.3 Marginal Cumulative distribution function
- 3.4 Independence of jointly distributed random variables
- 3.5 Jointly distributed discrete random variables
- 3.6 Jointly distributed continuous random variables
- 3.7 Conditional probabilities

3.1 Multivariate random variables

Multivariate random variable or random vector or k- dimensional random variables: A random variable $k-$ dimensional is a function with domain $\mathbf{S}$ and codomain $\mathbb{R}^k$:

$$(X_1, \ldots, X_k) : s \in \mathbf{S} \rightarrow (X_1(s), \ldots, X_k(s)) \in \mathbb{R}^k \ .$$

The function $(X_1(s), \ldots, X_k(s))$ is usually written for simplicity as $(X_1, \ldots, X_k)$.

Remark: If $k = 2$ we have the bivariate random variable or two dimensional random variable

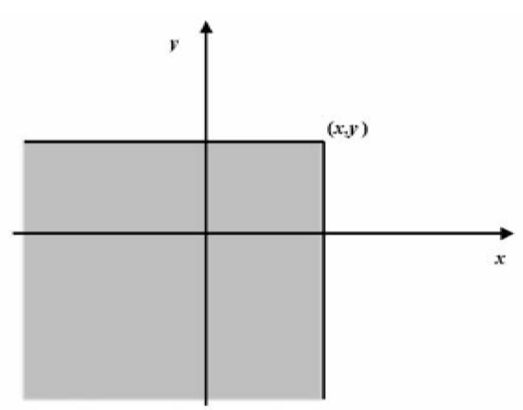
$$(X, Y) : s \in \mathbf{S} \rightarrow (X(s), Y(s)) \in \mathbb{R}^2 \ .$$

3.2 Joint cumulative distribution function

Joint cumulative distribution function: Let (X, Y) be a bivariate random variable. The real function of two real variables with domain $\mathbb{R}^2$ and defined by

$$F_{X,Y}(x, y) = P(X \leq x, Y \leq y)$$

is the joint cumulative distribution function of the two dimensional random variable (X, Y) .



Region defined by the inequalities $X \leq x$ and $Y \leq y$

Properties of the joint cumulative distribution function

1. $0 \leq F_{X,Y}(x,y) \leq 1$
2. $F_{X,Y}(x,y)$ is non decreasing with respect to x and y :
   (a) $\Delta_x > 0 \Rightarrow F_{X,Y}(x+\Delta x, y) \geq F_{X,Y}(x,y)$
   (b) $\Delta_y > 0 \Rightarrow F_{X,Y}(x, y+\Delta y) \geq F_{X,Y}(x,y)$
3. $\lim\limits_{x\to-\infty} F_{X,Y}(x,y) = 0$, $\lim\limits_{y\to-\infty} F_{X,Y}(x,y) = 0$ and $\lim\limits_{x\to+\infty}\lim\limits_{y\to+\infty} F_{X,Y}(x,y) = 1$
4. Let I be a rectangle defined as $I = (x_1, x_2] \times (y_1, y_2]$, where $\times$ denotes the Cartesian product of sets[1] then $P((X,Y) \in I) = P(x_1 < X \leq x_2, y_1 < Y \leq y_2) = F_{X,Y}(x_2,y_2) - F_{X,Y}(x_1,y_2) - F_{X,Y}(x_2,y_1) + F_{X,Y}(x_1,y_1)$.
5. $F_{X,Y}(x,y)$ is right continuous with respect to x and y: $\lim\limits_{x\to a^+} F_{X,Y}(x,y) = F_{X,Y}(a,y)$ and $\lim\limits_{y\to b^+} F_{X,Y}(x,y) = F_{X,Y}(x,b)$.

3.3 Marginal Cumulative distribution function

The (marginal) cumulative distribution functions of X and Y can be obtained form the Joint cumulative distribution functions of (X,Y) :

- The Marginal cumulative distribution function of X :$F_X(X \leq x) = P(X \leq x, Y \leq +\infty) = \lim\limits_{y\to+\infty} F_{X,Y}(x,y)$.
- The Marginal cumulative distribution function of Y :$F_Y(Y \leq y) = P(X \leq +\infty, Y \leq y) = \lim\limits_{x\to+\infty} F_{X,Y}(x,y)$.

Remark: The joint cumulative distribution function uniquely determines the marginal distributions, but the opposite is not true.

3.4 Independence of jointly distributed random variables

Definition: The jointly distributed random variables X and Y are said to be independent if and only if for any two sets $B_1 \in \mathbb{R}$, $B_2 \in \mathbb{R}$ we have

$$P(X \in B_1, Y \in B_2) = P(X \in B_1)P(Y \in B_2)$$

Remark: Independence implies that $F_{X,Y}(x,y) = F_X(x)F_Y(y)$, for any $(x,y) \in \mathbb{R}^2$.

Theorem: If X and Y are independent random variables and if $h(X)$ and $g(Y)$ are two functions of X and Y respectively, then the random variables $U = h(X)$ and $V = g(Y)$ are also independent random variables.

[1] The Cartesian product of two sets A and B is given by $A \times B = \{(x,y) : x \in A, y \in B\}$

3.5 Jointly distributed discrete random variables

Let $D_{(X,Y)}$ be the set of of discontinuities of the joint cumulative distribution function $F_{(X,Y)}(x, y)$, that is

$$D_{(X,Y)} = \left\{(x, y) \in \mathbb{R}^2 : P(X = x, Y = y) > 0\right\}$$

Definition: (X, Y) is a two dimensional discrete random variable if and only if

$$\sum_{(x,y)\in D_{(X,Y)}} P\left(X = x, Y = y\right) = 1.$$

Remark: As in the univariate case, a multivariate discrete random variable can take a finite number of possible values (x_i, y_j), where $i = 1, 2, ..., k_1$ and $j = 1, 2, ..., k_2$, where k_1 and k_2 are finite integers, or a countably infinite (x_i, y_j), where $i = 1, 2, ...$ and $j = 1, 2, ...$ For the sake of generality we consider the latter case. That is $D_{(X,Y)} = \{(x_i, y_j), i = 1, 2, ..., j = 1, 2, ...\}$

Definition: (Joint probability distribution/ function) If X and Y are discrete random variables, then the function given by

$$f_{X,Y}(x, y) = P\left(X = x, Y = y\right)$$

for $(x, y) \in D_{(X,Y)}$ is called the joint probability function of (X, Y) or joint probability distribution of the random variables X and Y.

Hence we have the following theorem.

Theorem: A bivariate function $f_{X,Y}(x, y)$ can serve as joint probability distribution of the pair of discrete random variables X and Y if and only if its values satisfy the conditions:

1. $f_{X,Y}(x, y) \geq 0$ for any $(x, y) \in \mathbb{R}^2$
2. $\sum_{(x,y)\in D_{(X,Y)}} f_{X,Y}(x, y) = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} f_{X,Y}(x_i, y_j) = 1$

Example: Let X and Y be the random variables representing the population of monthly wages of husbands and wives in a particular community. Say, there are only three possible monthly wages in euros: 0, 1000, 2000. The joint probability distribution is

	X	0	1000	2000
Y				
0		0.05	0.15	0.10
1000		0.10	0.10	0.30
2000		0.05	0.05	0.10

For example

$$\begin{aligned} f_{X,Y}(2000, 1000) &= P(X = 2000, Y = 1000) \\ &= 0.30 \end{aligned}$$

gives the probability that a husband earns 2000 euros and the wife earns 1000 euros.

Remark: We can calculate any probability using this function. For instance $P\left((x, y) \in B\right) = \sum_{(x,y)\in B} f_{X,Y}(x, y)$

Definition: (Joint cumulative distribution function) If X and Y are discrete random variables, the function given by

$$F_{X,Y}(x\ y) = \sum_{s \le x} \sum_{t \le y} f_{X,Y}(s,t)$$

for $(x, y) \in \mathbb{R}^2$ is called the joint distribution function or joint cumulative distribution of X and Y.

Example. (Cont) The joint cumulative probability distribution is

Y \ X	0	1000	2000
0	0.05	0.2	0.3
1000	0.15	0.4	0.8
2000	0.2	0.5	1

For instance

$$\begin{aligned} F_{X,Y}(1000, 1000) = P(X = 0, Y = 0) + P(X = 0, Y = 1000) \\ + P(X = 1000, Y = 0) + P(X = 1000, Y = 1000) \end{aligned}$$

Marginal probability distribution/function: The marginal probability function is another name for the probability function and it can be computed from the joint probability function. The marginal probability can be computed in the following way. If Y and X can take values $y_1, y_2, \ldots,$ and $,x_1, x_2 \ldots,$ respectively, then

$$\begin{aligned} P(X = x) &= \begin{cases} \sum_{y \in D_y} f(x,y) = \sum_{i=1}^{\infty} P(X = x, Y = y_i) & \text{for } x \in D_x \\ 0 & \text{for } x \notin D_x \end{cases} . \\ P(Y = y) &= \begin{cases} \sum_{x \in D_x} f(x,y) = \sum_{i=1}^{\infty} P(X = x_i, Y = y) \ . & \text{for } y \in D_y \\ 0 & \text{for } y \notin D_y \end{cases} \end{aligned}$$

where D_x and D_y are the range of X and Y respectively.

Example (cont):

$$\begin{aligned} P(X = x) &= P(X = x, Y = 0) + P(X = x, Y = 1000) \\ &\quad + P(X = x, Y = 2000) \\ P(Y = y) &= P(X = 0, Y = y) + P(X = 1000, Y = y) \\ &\quad + P(X = 2000, Y = y) \end{aligned}$$

Applying these formulas we have:

Y \ X	0	1000	2000	$P(Y = y)$
0	0.05	0.15	0.10	0.30
1000	0.10	0.10	0.30	0.50
2000	0.05	0.05	0.10	0.20
$P(X = x)$	0.20	0.30	0.50	1

Independence of random variables: Two random variables X and Y are independent if and only if

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

Example (cont): In the previous example

$$P(X = 0, Y = 0) = 0.05$$

and

$$P(X = 0)P(Y = 0) = 0.20 \times 0.30 = 0.06$$

thus X and Y are not independent.

Independence of discrete multivariate random variables: k random variables $X_1, \ldots, X_k$ are independent if and only if

$$P(X_1 = a_1, X_2 = a_2, \ldots, X_k = a_k) = P(X_1 = a_1)\, P(X_2 = a_2) \ldots, P(X_k = a_k)\,.$$

for all $(a_1, a_2, \ldots, a_k) \in \mathbb{R}^k$ such that $P(X_1 = a_1, X_2 = a_2, \ldots, X_k = a_k) > 0$.

3.6 Jointly distributed continuous random variables

Definition: (Two-dimensional continuous random variable) (X, Y) is a two-dimensional continuous random variable with a joint cumulative distribution function $F_{X,Y}(x, y)$, if and only if X and Y are continuous random variables and there is a non-negative real function $f_{X,Y}(x, y)$, such that

$$F_{X,Y}(x, y) = P(X \le x, Y \le y) = \int_{-\infty}^{y} \int_{-\infty}^{x} f_{X,Y}(t, s)dtds$$

The function $f_{X,Y}(x, y)$ is the joint (probability) density of X and Y.

Example: Given the joint probability density function

$$f_{X,Y}(x, y) = \begin{cases} (x + y) & \text{for } (x, y) \in (0, 1) \times (0, 1) \\ 0 & \textit{otherwise} \end{cases}$$

Compute $F_{X,Y}(x, y)$.

If $x < 0$ or $y < 0$ then $\int_{-\infty}^{y} \int_{-\infty}^{x} f_{X,Y}(t, s)dtds = 0$

Note that for $(x, y) \in (0, 1) \times (0, 1)$ we have

$$\begin{aligned} \int_{-\infty}^{y} \int_{-\infty}^{x} f_{X,Y}(t, s)dtds &= \int_{0}^{y} \int_{0}^{x} (t + s)\, dtds \\ &= \frac{1}{2}xy^2 + \frac{1}{2}x^2y. \end{aligned}$$

If $(x, y) \in (0, 1) \times [1, \infty)$.

$$\begin{aligned} \int_{-\infty}^{y} \int_{-\infty}^{x} f_{X,Y}(t, s)dtds &= \int_{0}^{1} \int_{0}^{x} (t + s)\, dtds \\ &= \frac{1}{2}x + \frac{1}{2}x^2. \end{aligned}$$

If $(x, y) \in [1, \infty) \times (0, 1)$.

$$\begin{aligned} \int_{-\infty}^{y} \int_{-\infty}^{x} f_{X,Y}(t, s) dt ds &= \int_{0}^{y} \int_{0}^{1} (t + s) \, dt ds \\ &= \frac{1}{2} y + \frac{1}{2} y^2 \end{aligned}$$

and if $(x, y) \in [1, \infty) \times [1, \infty)$ we have

$$\begin{aligned} \int_{-\infty}^{y} \int_{-\infty}^{x} f_{X,Y}(t, s) dt ds &= \int_{0}^{1} \int_{0}^{1} (t + s) \, dt ds \\ &= 1 \end{aligned}$$

Hence

$$F_{X,Y}(x, y) = \begin{cases} 0 & \text{for } x \leq 0 \text{ or } y \leq 0 \\ \frac{1}{2} x y^2 + \frac{1}{2} x^2 y & \text{for } (x, y) \in (0, 1) \times (0, 1) \\ \frac{1}{2} x + \frac{1}{2} x^2 & \text{for } (x, y) \in (0, 1) \times [1, \infty) \\ \frac{1}{2} y + \frac{1}{2} y^2 & \text{for } (x, y) \in [1, \infty) \times (0, 1) \\ 1 & \text{for } (x, y) \in [1, \infty) \times [1, \infty) \end{cases}$$

Theorem: A bivariate function can serve as a joint probability density function of a pair of continuous random variables X and Y if its values, $f_{X,Y}(x, y)$, satisfy the conditions:

1. $f_{X,Y}(x, y) \geq 0, (x, y) \in \mathbb{R}^2$
2. $\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f_{X,Y}(x, y) dx dy = 1$.

Remark: Note that $P((X, Y) \in A) = \int \int_A f_{X,Y}(x, y) dx dy$ for any region A in $\mathbb{R}^2$.

Example: Given the joint probability density function

$$f_{X,Y}(x, y) = \begin{cases} \frac{3}{5} x (y + x) & \text{for } (x, y) \in (0, 1) \times (0, 2) \\ 0 & \textit{otherwise} \end{cases}$$

Compute $P((X, Y) \in (0, 1/2) \times (1, 2))$.

Notice that

$$P((X, Y) \in (0, 1/2) \times (1, 2)) = \int_{1}^{2} \int_{0}^{1/2} f_{X,Y}(x, y) dx dy$$

Now

$$\begin{aligned} \int_{0}^{1/2} f_{X,Y}(x, y) dx &= \int_{0}^{1/2} \left(\frac{3}{5} x (y + x) \right) dx \\ &= \frac{3}{40} y + \frac{1}{40} \end{aligned}$$

And consequently

$$\begin{aligned} \int_{1}^{2} \int_{0}^{1/2} f_{X,Y}(x, y) dx dy &= \int_{1}^{2} \left(\frac{3}{40} y + \frac{1}{40} \right) dy \\ &= \frac{11}{80}. \end{aligned}$$

Properties:

1. For all points where $f_{X,Y}(x,y)$ is continuous, $f_{X,Y}(x,y) = \frac{\partial^2 F_{X,Y}(x,y)}{\partial x \partial y} = \frac{\partial^2 F_{X,Y}(x,y)}{\partial y \partial x}$.
2. Marginal cumulative distribution functions: A marginal cumulative distribution function is another name for a cumulative distribution function of a single random variable
   (a) Marginal cumulative distribution functions of the random variable X
   $$F_X(x) = \lim_{y \to +\infty} F_{X,Y}(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{+\infty} f_{X,Y}(u,y)dydu,$$
   (b) Marginal cumulative distribution functions of the random variable Y
   $$F_Y(y) = \lim_{x \to +\infty} F_{X,Y}(x,y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{y} f_{X,Y}(x,v)dvdx.$$
3. Marginal density functions: A marginal density function is another name for a density function of a single random variable
   (a) Marginal density functions of the random variable X
   $$f_X(x) = \int_{-\infty}^{+\infty} f_{X,Y}(x,v)dv,$$
   (b) Marginal density functions of the random variable Y
   $$f_Y(y) = \int_{-\infty}^{+\infty} f_{X,Y}(u,y)du.$$

Example: Given the joint probability density function

$$f_{X,Y}(x,y) = \begin{cases} (x+y) & \text{for } (x,y) \in (0,1) \times (0,1) \\ 0 & \textit{otherwise} \end{cases}$$

compute the marginal cumulative distribution function of X and its density function.

Note that

$$\begin{aligned} F_X(x) &= \int_0^x \int_0^1 f_{X,Y}(u,y)dudy \\ &= \int_0^x \int_0^1 (u+y)\, dydu \\ &= \int_0^x \left(u + \frac{1}{2}\right) du \\ &= \frac{1}{2}x + \frac{1}{2}x^2, x \in (0,1). \end{aligned}$$

Also $f_X(x) = \int_0^1 f_{X,Y}(x,v)dv = \int_0^1 (v+x)\, dv = x + \frac{1}{2}, x \in (0,1)$. Alternatively we could obtain the same result by using the formula $f_X(x) = \frac{dF_X(x)}{dx}$.

Independence of random variables: Two random variables X and Y are independent if and only if

$$f_{X,Y}(x,y) = f_X(x) f_Y(y).$$

where $f_X(x)$ and $f_Y(y)$ are the density functions of X and Y respectively.

Example: Given the joint probability density function

$$f_{X,Y}(x,y) = \begin{cases} (4yx) & \text{for } (x,y) \in (0,1)\times(0,1) \\ 0 & \textit{otherwise} \end{cases}$$

Show that X and Y are independent random variables.

Independence of continuous multivariate random variables: k random variables $X_1,\ldots,X_k$ are independent if and only if

$$f_{X_1,\ldots,X_k}(a_1,a_2,\ldots,a_k) = f_{X_1}(a_1)\, f_{X_2}(a_2)\ldots, f_{X_k}(a_k).$$

for all $(a_1,a_2,\ldots,a_k) \in \mathbb{R}^k$, where $f_{X_1,\ldots,X_k}(a_1,a_2,\ldots,a_k)$ is the joint density functions of the random variables $X_1,\ldots,X_k$.

3.7 Conditional probabilities

Discrete random variables

Conditional probability function of Y given X: A *conditional probability function* of a discrete random variable Y given another discrete variable X taking a specific value is defined as

$$\begin{aligned} f_{Y|X=x}(y) &= P(Y=y|X=x) \\ &= \frac{P(Y=y, X=x)}{P(X=x)} \\ &= \frac{f_{X,Y}(x,y)}{f_X(x)}, \quad f_X(x) > 0. \end{aligned}$$

(for fixed x). The *conditional probability* function of X given Y is defined in a similar way

$$f_{X|Y=y}(x) = \frac{f_{X,Y}(x,y)}{f_y(y)}, \quad f_Y(y) > 0.$$

Remark: We can use also an alternative notation $f_{Y|X=x}(y) = f_{Y|X}(y|x)$ and $f_{X|Y=y}(x) = f_{X|Y}(x|y)$.

Example (cont): Consider the joint probability function

	X	0	1000	2000	$P(Y=y)$
Y					
0		0.05	0.15	0.10	0.30
1000		0.10	0.10	0.30	0.50
2000		0.05	0.05	0.10	0.20
$P(X=x)$		0.20	0.30	0.50	1

Compute $P(Y=y|X=0)$, $y = 0, 1000, 2000$.

Note that

$$\begin{aligned} P(Y=0|X=0) &= \frac{P(Y=0, X=0)}{P(X=0)} = \frac{0.05}{0.2} = 0.25 \\ P(Y=1000|X=0) &= \frac{P(Y=1000, X=0)}{P(X=0)} = \frac{0.1}{0.2} = 0.5. \\ P(Y=2000|X=0) &= \frac{P(Y=2000, X=0)}{P(X=0)} = \frac{0.05}{0.2} = 0.25. \end{aligned}$$

Remarks:

- The conditional probability functions satisfy all the properties of probability functions, and therefore $\sum_{i=1}^{\infty} f_{Y|X}(y_i) = 1$.
- If X and Y are independent $f_{Y|X=x}(y) = f_y(y)$ and $f_{X|Y=y}(x) = f_X(x)$

Continuous random variables.

Conditional Probability density function: If $f_{X,Y}(x,y)$ is the joint probability density function of the continuous random variables X and Y and $f_Y(y)$ is the marginal density function of Y, the function given by

$$f_{X|Y=y}(x) = \frac{f_{X,Y}(x,y)}{f_Y(y)}, x \in \mathbb{R} \text{ (for fixed } y\text{)}, f_Y(y) \neq 0$$

is the conditional probability function of X given $\{Y = y\}$. Similarly if $f_X(x)$ is the marginal density function of X

$$f_{Y|X=x}(y) = \frac{f_{X,Y}(x,y)}{f_X(x)}, y \in \mathbb{R} \text{ (for fixed } x\text{)}, f_X(x) \neq 0$$

is the conditional probability function of Y given $\{X = x\}$

Remark: As in the discrete case we can use also the alternative notation $f_{Y|X=x}(y) = f_{Y|X}(y|x)$ and $f_{X|Y=y}(x) = f_{X|Y}(x|y)$.

Remark: Note that

$$P(X \in B|Y = y) = \int_B f_{X|Y=y}(x)\,dx$$

for any $B \subset \mathbb{R}$

Example: Given the joint probability density function

$$f_{X,Y}(x,y) = \begin{cases} (y+x) & \text{for } (x,y) \in (0,1) \times (0,1) \\ 0 & \textit{otherwise} \end{cases}$$

Compute $f_{X|Y=0.5}(x)$ and $P(X \geq 0.7|Y = 0.5)$

$$\begin{aligned} f_{X|Y=0.5}(y) &= \frac{f_{X,Y}(x,0.5)}{f_Y(0.5)} \\ &= x + 0.5, x \in (0,1) \end{aligned}$$

Note that $P(X \geq 0.7|Y = 0.5) = \int_{0.7}^{1} f_{X|Y=0.5}(x)\,dx = \int_{0.7}^{1} (x+0.5)\,dx = 0.405$

Remarks:

1. The conditional density functions of X and Y verify all the properties of a density function of a univariate random variable.
2. Note that we can always decompose a joint density function in the following way

$$f_{X,Y}(x,y) = f_X(x) f_{Y|X=x}(y) = f_Y(y) f_{X|Y=y}(x).$$

3. If X and Y are independent $f_{Y|X=x}(y) = f_Y(y)$ and $f_{X|Y=y}(x) = f_X(x)$.