DECEMBER 7, 2014

LECTURE 12
JOINT (BIVARIATE) DISTRIBUTIONS, MARGINAL DISTRIBUTIONS, INDEPENDENCE

So far we have considered one random variable at a time. However, in economics we are typically interested in relationships between several variables. Therefore, we need to extend the concept of distributions in a way that will allow us to describe the *joint* behavior of several random variables. Most of the ideas and methods can be illustrated using only two variables. In this case distributions are called *bivariate*.

1 Joint and marginal distributions

Suppose that we have two discrete random variables X and Y defined on the same sample space Ω, which comes with a probability function P:

$$\begin{aligned} X &= X(\omega), \\ Y &= Y(\omega). \end{aligned}$$

In that case, we would be interested in describing their joint behavior:

$$P(X \in A, Y \in B) = P\left(\{\omega \in \Omega : X(\omega) \in A\} \cap \{\omega \in \Omega : Y(\omega) \in B\}\right),$$

where A and B are some subsets of $\mathbb{R}$.

Definition 1. (Joint PMF) Let X and Y be two discrete random variables defined on the same probability space. Let $S_X = \{x_1, x_2, \ldots\}$ denote the support of X and $S_Y = \{y_1, y_2, \ldots\}$ denote the support of Y. The *joint PMF* of X and Y is defined as

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

for $x \in S_X$ and $y \in S_Y$.

Remark. 1. The comma in $P(X = x, Y = y)$ stands for "and", i.e. the intersection of the events $\{\omega \in \Omega : X(\omega) = x\}$ and $\{\omega \in \Omega : Y(\omega) = y\}$.

2. The supports of X and Y can have different number of elements or be countably infinite.

3. The joint PMF must satisfy the following properties:

$$0 \leq p_{X,Y}(x, y) \leq 1$$

for any $x \in S_X, y \in S_Y$. Also,

$$\sum_{x \in S_X} \sum_{y \in S_Y} p_{X,Y}(x, y) = 1.$$

4. The joint PMF can be used to compute the probabilities of events defined through conditions relating X and Y. For example, consider the event of X being equal to Y: $\{\omega \in \Omega : X(\omega) = Y(\omega)\}$. We have

$$P(X = Y) = \sum \sum_{x \in S_X, y \in S_Y : x = y} p_{X,Y}(x, y).$$

Example. In this example, we have two random variables: earnings per share E (during the last 12 months) and the price S of a share. The joint PMF is described in the following table:

Table 1: Example of a joint PMF

		Price of a share (S)		
		\$100	\$250	\$400
Earnings per share (E)	\$10	$\frac{2}{6}$	$\frac{1}{6}$	0
	\$20	0	$\frac{2}{6}$	$\frac{1}{6}$

For example,

$$\begin{aligned} P(E = 10, S = 250) &= p_{E,S}(10, 250) = \frac{1}{6}, \\ P(E = 20, S = 250) &= p_{E,S}(20, 250) = \frac{2}{6}. \end{aligned}$$

In this example, we can use the joint PMF to compute the probability that the *price-earnings* ratio is under 20:

$$\begin{aligned} P(S/E < 20) &= P(E = 10, S = 100) + P(E = 20, S = 100) + P(E = 20, S = 250) \\ &= p_{E,S}(10, 100) + p_{E,S}(20, 100) + p_{E,S}(20, 250) \\ &= \frac{2}{6} + 0 + \frac{2}{6} \\ &= \frac{2}{3}. \end{aligned}$$

The joint PMF describes the joint behavior (distribution) of two or more random variables. In particular, it contains the information on the distribution of each random variable *individually* or *marginal distributions*. Suppose that you are given the joint PMF of X and Y: $p_{X,Y}(x, y)$, $x \in S_X = \{x_1, \ldots, x_n\}$ and $y \in S_Y = \{y_1, \ldots, y_m\}$. Consider the following question: for some $x \in S_X$, what is the probability that $X = x_1$ (regardless of the value of Y)? To answer that question, note that Y introduces a partition on the sample space Ω:

$$\Omega = \{Y = y_1\} \cup \{Y = y_2\} \cup \ldots \cup \{Y = y_m\}.$$

Therefore, the event $\{X = x_1\}$ can be represented as

$$\{X = x_1\} = \{X = x_1, Y = y_1\} \cup \{X = x_1, Y = y_2\} \cup \ldots \cup \{X = x_1, Y = y_m\}.$$

Since the events in a partition are mutually independent, we have:

$$\begin{aligned} p_X(x_1) &= P(X = x_1) \\ &= P(X = x_1, Y = y_1) + P(X = x_1, Y = y_2\} + \ldots + P(X = x_1, Y = y_m) \\ &= p_{X,Y}(x_1, y_1) + p_{X,Y}(x_1, y_2) + \ldots + p_{X,Y}(x_1, y_m) \\ &= \sum_{y \in S_Y} p_{X,Y}(x_1, y). \end{aligned}$$

This operation can be performed for every value of $x \in S_X$. The resulting function is called the *marginal* PMF of X.

Definition 2. (Marginal PMF) Let $p_{X,Y}(x, y)$ be the joint PMF of two discrete random variables X and Y, $x \in S_X$ and $y \in S_Y$. The marginal PMF of X is given by

$$p_X(x) = \sum_{y \in S_Y} p_{X,Y}(x, y).$$

The marginal PMF of Y is given by

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

Remark. 1. To find the marginal PMF of X from the joint PMF of X and Y, one computes the sums of the joint PMF values over all possible values of Y for each given $x \in S_X$.

2. It is easy to see that the marginal PMF is a PMF: it is strictly positive and

$$\begin{aligned} \sum_{x \in S_X} p_X(x) &= \sum_{x \in S_X} \left(\sum_{y \in S_Y} p_{X,Y}(x, y) \right) \\ &= 1, \end{aligned}$$

where the equality in the second line holds by the properties of the joint PMF.

Example. Consider the example in Table 1. To obtain the marginal PMF of E, one should take sums across the columns of the table, i.e. compute row totals:

$$\begin{aligned}
p_E(10) &= p_{E,S}(10,100) + p_{E,S}(10,250) + p_{E,S}(10,400) \\
&= \frac{2}{6} + \frac{1}{6} + 0 \\
&= \frac{1}{2}, \\
p_E(20) &= \frac{1}{2}.
\end{aligned}$$

To obtain the marginal PMF of S, one should take sums across the rows of the table, i.e. compute column totals:

$$\begin{aligned}
p_S(100) &= p_{E,S}(10,100) + p_{E,S}(20,100) \\
&= \frac{2}{6} + 0 \\
&= \frac{1}{3} \\
p_S(250) &= p_{E,S}(10,250) + p_{E,S}(20,250) \\
&= \frac{1}{6} + \frac{2}{6} \\
&= \frac{1}{2}, \\
p_S(400) &= \frac{1}{6}.
\end{aligned}$$

Note that marginal probabilities describe the distributions of the respective variables, and therefore they cannot be used to find the probabilities of events defined in terms of both E and S. For example, we cannot use $p_E(10) = 1/2$ and $p_S(100) = 1/3$ to find $P(E = 10, S = 100)$. Also, we cannot answer any questions concerning the price-earnings ratio.

The concept of the CDF can be extended to the case of two or more random variables. In the bivariate case, we have the following definition.

Definition 3. (Joint CDF) Let X and Y be two random variables. Their joint CDF is defined as

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

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

Below we show some of the properties of the joint CDF:

$$\begin{aligned}
\lim_{y\to-\infty} F_{X,Y}(x,y) &= \lim_{y\to-\infty} P(X\le x, Y\le y) \\
&= P(X\le x, Y< -\infty) \\
&\le P(Y<-\infty) \\
&= 0, \\
\lim_{y\to\infty} F_{X,Y}(x,y) &= \lim_{y\to\infty} P(X\le x, Y\le y) \\
&= P(X\le x, Y<\infty) \\
&= P(X\le x) \\
&= F_X(x). \qquad (1)
\end{aligned}$$

The second result shows that we can always obtain the marginal CDF from the joint CDF.

All the concepts described above can be naturally extended to the case of continuous random variables.

Definition 4. The joint distribution of X and Y is continuous if the joint CDF $F_{X,Y}(x,y)$ is continuous and differentiable in both x and y.

Definition 5. (Joint PDF) Let X and Y be two continuous random variables with a joint CDF $F_{X,Y}$. The joint PDF of X and Y is defined as

$$f_{X,Y}(x,y) = \frac{\partial^2 F_{X,Y}(x,y)}{\partial x \partial y}.$$

Remark. The joint PDF satisfies the following properties:

1. $f_{X,Y}(x,y)\ge 0$ for all $x\in\mathbb{R}$ and $y\in\mathbb{R}$.
2. $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)dydx = 1$.
3. $F_{X,Y}(x,y) = P(X\le x, Y\le y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f_{X,Y}(u,v)dvdu$.
4. $P((X,Y)\in A) = \int\int_A f_{X,Y}(x,y)dydx$, where A is a set on a plane (in $\mathbb{R}^2$).

Similarly to the discrete case, the marginal PDF can be computed from the joint PDF.

Theorem 6. *Let X and Y be continuously distributed with the joint PDF $f_{X,Y}(x,y)$. The marginal PDF of X is given by*

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

Remark. To compute the marginal PDF of X, one has to "integrate out" y for every value of x.

Proof. Let $F_{X,Y}$ denote the joint CDF of X and Y. From (1),

$$\begin{aligned} F_X(x) &= \lim_{y\to\infty} F_{X,Y}(x,y) \\ &= \lim_{y\to\infty} \int_{-\infty}^{x}\int_{-\infty}^{y} f_{X,Y}(u,v)dvdu \\ &= \int_{-\infty}^{x}\int_{-\infty}^{\infty} f_{X,Y}(u,v)dvdu. \end{aligned}$$

Next,

$$\begin{aligned} f_X(x) &= \frac{dF_X(x)}{dx} \\ &= \frac{d}{dx}\left(\int_{-\infty}^{x}\int_{-\infty}^{\infty} f_{X,Y}(u,v)dvdu\right) \\ &= \int_{-\infty}^{\infty} f_{X,Y}(x,v)dv, \end{aligned}$$

where the result in the last line follows from the fact that

$$\frac{d}{dx}\left(\int^{x} h(u)du\right) = h(x).$$

□

2 Independence

We say that two (or more) random variables are independent if changes in one random variable are not going to affect the distribution of other random variables. Formally, we have the following definition.

Definition 7. (Independence) (a) Let X and Y be two discrete random variables with supports S_X and S_Y respectively. We say that X and Y are independent if

$$P(X = x, Y = y) = P(X = x)P(Y = y) \text{ for all } x \in S_X, y \in S_Y,$$

i.e.

$$p_{X,Y}(x,y) = p_X(x)p_Y(y) \text{ for all } x \in S_X, y \in S_Y,$$

where $p_{X,Y}$ is the joint PMF of X and Y, p_X is the marginal PMF of X, and p_Y is the marginal PMF of Y.

(b) Let X and Y be two continuous random variables distributed with a joint PDF $f_{X,Y}$. Let f_X and f_Y be the PDFs of X and Y respectively. We say that X and Y are independent if

$$f_{X,Y}(x,y) = f_X(x)f_Y(y) \text{ for all } x \in \mathbb{R}, y \in \mathbb{R}.$$

Example. Consider the joint distribution described in Table 1. We have

$$\begin{aligned} p_{E,S}(10,100) &= \frac{1}{3}, \\ p_E(10) \times p_S(100) &= \frac{1}{2} \times \frac{1}{3}. \end{aligned}$$

Hence, E and S are not independent.

Example. Consider the joint and marginal PMFs described in the following table:

		support of Y		
		y_1	y_2	marginal PMF of X
Support of X	x_1	$\frac{1}{12}$	$\frac{2}{12}$	$\frac{1}{4}$
	x_2	$\frac{3}{12}$	$\frac{6}{12}$	$\frac{3}{4}$
	marginal PMF of Y	$\frac{1}{3}$	$\frac{2}{3}$	

The center of the table shows the joint PMF of X and Y. For example, $P(X = x_1, Y = y_1) = \frac{1}{12}$, $P(X = x_1, Y = y_2) = \frac{3}{12}$, and etc. On the margins of the table, we have the marginal PMFs of X and Y: $P(X = x_1) = \frac{1}{4}$, $P(X = x_2) = \frac{3}{4}$, and $P(Y = y_1) = \frac{1}{3}$, $P(Y = y_2) = \frac{2}{3}$. In this example, X and Y are independent: for *every* $x \in \{x_1, x_2\}$ and every $y \in \{y_1, y_2\}$, $P(X = x, Y = y) = P(X = x)P(Y = y)$:

$$\begin{aligned} P(X = x_1, Y = y_1) &= \frac{1}{12} = \frac{1}{4} \times \frac{1}{3} = P(X = x_1)P(Y = y_1), \\ P(X = x_1, Y = y_2) &= \frac{2}{12} = \frac{1}{4} \times \frac{2}{3} = P(X = x_1)P(Y = y_2), \\ &\cdots \end{aligned}$$

Note that every entry for the joint PMF in the center of the table is equal to the product of the corresponding marginal PMFs.

Theorem 8. *Suppose that X and Y are independent. Then, for any $A \subset \mathbb{R}$ and $B \subset \mathbb{R}$,*

$$P(X \in A, Y \in B) = P(X \in A)P(Y \in B). \tag{2}$$

Remark. The statement in (2) can actually be used as the definition of independence.

Proof. We will prove the result for continuous X and Y. Let $f_{X,Y}$ be the joint PDF of X and

Y. Let f_X and f_Y denote the marginal PDFs of X and Y respectively.

$$
\begin{aligned}
P(X \in A, Y \in B) &= \int_A \int_B f_{X,Y}(x,y) dy dx \\
&= \int_A \int_B f_X(x) f_Y(y) dy dx \\
&= \int_A f_X(x) \left(\int_B f_Y(y) dy \right) dx \\
&= \left(\int_A f_X(x) dx \right) \left(\int_B f_Y(y) dy \right) \\
&= P(X \in A) P(Y \in B).
\end{aligned}
$$

The proof is analogous in the discrete case: one has to replace integrals with sums and PDFs with PMFs. □

The transformations of independent random variables are also independent.

Theorem 9. *Suppose that X and Y are independent. Define $U = g(X)$ and $V = h(Y)$. Then U and V are also independent.*

Proof. We will provide a proof only for the discrete case. Let S_X and S_Y denote the supports of X and Y respectively. Let u be a point in the support of U, and let v be a point in the support of V. Define $C_u = \{x \in S_X : g(x) = u\}$ and $C_v = \{y \in S_Y : h(y) = v\}$. Here C_u contains all the points in the support of X such that $g(x) = u$ for a chosen value u. The set C_v is defined similarly. We have

$$
\begin{aligned}
P(U = u, V = v) &= P(X \in C_u, Y \in C_v) \\
&= P(X \in C_u) P(Y \in C_v) \\
&= P(U = u) P(V = v),
\end{aligned}
$$

where the equality in the second line holds by Theorem 8 and the independence of X and Y. □