LECTURE 6

§11. JOINT DISTRIBUTION FUNCTIONS

Two random variables, $\eta_1(\omega)$ and $\eta_2(\omega)$, are said to be jointly distributed if they are defined as functions on the same probability space. It is then possible to make joint probability statements about $\eta_1(\omega)$ and $\eta_2(\omega)$ (that is, probability statements about the simultaneous behavior of the two random variables). In order to deal with such probabilities, we define, for any two random variables $\eta_1(\omega)$ and $\eta_2(\omega)$, the *Joint distribution function* of $\eta_1(\omega)$ and $\eta_2(\omega)$ by the formula

$$F(x_1, x_2) = P\left(\omega : \ \eta_1(\omega) \le x_1 \bigcap \eta_2(\omega) \le x_2\right), \quad x_1, x_2 \in \mathbb{R}^1. \tag{22}$$

PROPERTIES OF JOINT DISTRIBUTION FUNCTION

Property 1. $F(x_1, x_2)$ is a nondecreasing function by each argument.

Property 2. $F(x_1, x_2) \to 1$ as $x_1 \to +\infty$ and $x_2 \to +\infty$.

Property 3. $F(x_1, x_2) \to 0$ as either x_1 or x_2 tends to $-\infty$.

Property 4. $F(x_1, x_2)$ is right continuous by each argument.
Verification of these properties is left as an exercise because they can be proved as the corresponding properties for distribution function of a random variable.
Therefore, Properties 1 — 4 are necessary conditions for $G(x_1, x_2)$ to be a joint distribution function.

Theorem 3. *Let $F(x_1, x_2)$ be a joint distribution function of a vector $(\eta_1(\omega), \eta_2(\omega))$. Then*

$$P\left\{\omega : \ a_1 < \eta_1(\omega) \le b_1 \bigcap a_2 < \eta_2(\omega) \le b_2\right\} = F(b_1, b_2) - F(a_1, b_2) - F(b_1, a_2) + F(a_1, a_2), \tag{23}$$

whenever, $a_1 < b_1$ and $a_2 < b_2$, i. e. probability that a random point $(\eta_1(\omega), \eta_2(\omega))$ belongs to rectangle $(a_1, b_1] \times (a_2, b_2]$ is equal to the algebraic sum of the values of joint distribution function at the vertices of the rectangle ($F(b_1, b_2)$ and $F(a_1, a_2)$ with plus, but $F(b_1, a_2)$ and $F(a_1, b_2)$ with minus).

The theorem is easily proved geometrically.

Let us give an example that Properties 1 — 4 are not sufficient for a function $G(x_1, x_2)$ to be a joint distribution function of a random variables $\eta_1(\omega)$ and $\eta_2(\omega)$.

Let

$$G(x_1, x_2) = \begin{cases} 0 & \text{if either } x_1 < 0 \text{ or } x_2 < 0 \\ & \text{or } x_1 \geq 0, x_2 \geq 0 \text{ and } x_1 + x_2 < 1 \\ & \\ 1 & \text{if } x_1 \geq 0, x_2 \geq 0 \text{ and } x_1 + x_2 \geq 1. \end{cases}$$

It is not difficult to verify that the function satisfies Properties 1 — 4. Let us assume that the function $G(x_1, x_2)$ is a joint distribution function for some random vector (η_1, η_2). Then

$$P\left\{\omega\colon\ \frac{1}{2} < \eta_1(\omega) \leq 1 \bigcap 0 < \eta_2(\omega) \leq \frac{1}{2}\right\} =$$

$$= G\left(1, \frac{1}{2}\right) - G(1, 0) - G\left(\frac{1}{2}, \frac{1}{2}\right) + G\left(\frac{1}{2}, 0\right) = 1 - 1 - 1 + 0 = -1.$$

By Axiom of probability this probability must be nonnegative.

The distribution of $\eta_1(\omega)$ can be obtained from the joint distribution function as follows:

$$F_{\eta_1}(x_1) = P(\omega\colon\ \eta_1(\omega) \leq x_1) = P\left(\omega\colon\ \bigcup_{n=1}^{\infty} [(\eta_1(\omega) \leq x_1) \bigcap (\eta_2(\omega) \leq x_{2n})]\right) =$$
$$= \lim_{x_2 \to +\infty} P\left(\omega\colon\ (\eta_1(\omega) \leq x_1) \bigcap (\eta_2(\omega) \leq x_2)\right) = \lim_{x_2 \to +\infty} F(x_1, x_2) \equiv F(x_1, +\infty),$$

where x_{2n} is a monotone increasing sequence which tends to $+\infty$.

Proof: The proof follows from the relation

$$\{\omega\colon\ \eta_1(\omega) \leq x_1\} = \bigcup_{n=1}^{+\infty} \{\omega\colon\ (\eta_1(\omega) \leq x_1) \bigcap (\eta_2(\omega) \leq x_{2n})\}$$

and Property of probability.

Similarly, the distribution function of $\eta_2(\omega)$ is given by

$$F_{\eta_2}(x_2) = P(\omega\colon\ \eta_2(\omega) \leq x_2) = \lim_{x_1 \to +\infty} F(x_1, x_2) \equiv F(+\infty, x_2).$$

The distribution functions F_{η_1} and F_{η_2} are sometimes referred to as the *Marginal* distributions of $\eta_1(\omega)$ and $\eta_2(\omega)$.

In the case when $\eta_1(\omega)$ and $\eta_2(\omega)$ are both discrete random variables, it is convenient to define the *joint probability mass function* of $\eta_1(\omega)$ and $\eta_2(\omega)$ by

$$p(x,y) = P\left\{\omega:\ \eta_1(\omega) = x \bigcap \eta_2(\omega) = y\right\}.$$

The probability mass function of $\eta_1(\omega)$ can be obtained from $p(x,y)$ by

$$p_{\eta_1(\omega)}(x) = P(\omega:\ \eta_1(\omega) = x) = \sum_{y:\ p(x,y)>0} p(x,y).$$

Similarly

$$p_{\eta_2(\omega)}(y) = P(\omega:\ \eta_2(\omega) = y) = \sum_{x:\ p(x,y)>0} p(x,y).$$

We say that $\eta_1(\omega)$ and $\eta_2(\omega)$ are *jointly continuous* if there exists a function $f(x_1,x_2)$ defined for all real x_1 and x_2, having the property

$$F(x_1,x_2) = \int_{-\infty}^{x_1}\int_{-\infty}^{x_2} f(x,y)\,dx\,dy. \tag{24}$$

The function $f(x_1,x_2)$ is called the *joint density function* of $\eta_1(\omega)$ and $\eta_2(\omega)$.
It follows upon differentiation, that

$$f(x_1,x_2) = \frac{\partial^2}{\partial x_1\,\partial x_2}F(x_1,x_2) \tag{25}$$

whenever the partial derivatives are defined.
If $\eta_1(\omega)$ and $\eta_2(\omega)$ are jointly continuous, they are individually continuous, and their probability density functions can be obtained as follows:

$$P\{\omega:\ \eta_1(\omega) \le x_1\} = P\left\{\omega:\ \eta_1(\omega) \le x_1 \bigcap -\infty < \eta_2(\omega) < +\infty\right\} =$$
$$= \int_{-\infty}^{x_1}\int_{-\infty}^{+\infty} f(x,y)\,dx\,dy = \int_{-\infty}^{x_1} f_{\eta_1}(x)\,dx$$

where

$$f_{\eta_1}(x) = \int_{-\infty}^{+\infty} f(x,y)\,dy \tag{26}$$

is thus the density function of $\eta_1(\omega)$. Similarly, the density function of $\eta_2(\omega)$ is given by

$$f_{\eta_2}(x) = \int_{-\infty}^{+\infty} f(x,y)\,dx. \tag{27}$$

It follows from (23) that

$$P\left\{\omega:\ \ x<\eta_1(\omega)\le x+dx\bigcap y<\eta_2(\omega)\le y+dy\right\}=\int_x^{x+dx}\int_y^{y+dy}f(x_1,x_2)\,dx_1\,dx_2=$$
$$=f(x,y)\,dx\,dy+o(dx\,dy)$$

when dx and dy are small and $f(x_1,x_2)$ is continuous at (x,y). Hence $f(x,y)$ is a measure of how likely it is that the random vector $(\eta_1(\omega),\eta_2(\omega))$ will be near (x,y).

The following theorem we cite without proof.

Theorem 4 (About Joint Distribution Function). *Let a function $G(x_1,x_2)$, $x_1,x_2\in\mathbb{R}^1$ satisfy the properties 1 — 4 and, in addition, the condition*

$$G(b_1,b_2)-G(a_1,b_2)-G(b_1,a_2)+G(a_1,a_2)\ge 0,\qquad \text{for any } a_1<b_1 \text{ and } a_2<b_2. \tag{28}$$

Then there exist a probability space (Ω,P) and a random vector $(\eta_1(\omega),\eta_2(\omega))$ for which joint distribution function coincides with given function $G(x_1,x_2)$, that is

$$P(\omega:\quad \eta_1(\omega)\le x_1\bigcap\eta_2(\omega)\le x_2)=G(x_1,x_2).$$

Therefore, for giving an example of a random vector we have to cite a function which satisfies the properties 1 — 4 and the condition (28).
It should be noted that it is not easy to verify the condition (28) for a function $G(x_1,x_2)$.

Remark 6. We can also define joint distributions for n random variables in exactly the same manner as we did for $n=2$. For instance, the joint distribution function $F(x_1,x_2,...,x_n)$ of the n random variables $\eta_1(\omega)$, $\eta_2(\omega)$, ..., $\eta_n(\omega)$ is defined by

$$F(x_1,x_2,...,x_n)=P\left(\omega:\ \ \eta_1(\omega)\le x_1\bigcap\eta_2(\omega)\le x_2\bigcap...\bigcap\eta_n(\omega)\le x_n\right).$$

§12. SOME REMARKS ABOUT JOINT DENSITY FUNCTIONS

Remark 7. We note that a joint density function has two properties

$$1)\quad f(x_1,x_2)\ge 0; \tag{29}$$

$$2)\quad \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}f(x_1,x_2)\,dx_1\,dx_2=1. \tag{30}$$

These conditions can be proved in the same manner as we did for $n=1$ (compare with Properties 1, 2 in §20).

Remark 8. It should be noted that if a function $g(x_1, x_2)$ $x_1, x_2 \in \mathbb{R}^1$ satisfies the conditions

$$g(x_1, x_2) \geq 0 \quad \text{and} \quad \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x_1, x_2)\, dx_1\, dx_2 = 1$$

then there exist a probability space (Ω, P) and two random variables $\eta_1(\omega)$ and $\eta_2(\omega)$ such that their joint density function $f(x_1, x_2)$ coincides with $g(x_1, x_2)$, that is

$$P\left\{\omega : \ \eta_1(\omega) \leq x_1 \bigcap \eta_2(\omega) \leq x_2\right\} = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} g(x, y)\, dx\, dy. \tag{31}$$

Remark 9. If $f(x_1, x_2)$ is a joint density function for a random vector $(\eta_1(\omega), \eta_2(\omega))$ then

$$\begin{aligned} &P\left\{\omega : \ a_1 < \eta_1(\omega) \leq b_1 \bigcap a_2 < \eta_2(\omega) \leq b_2\right\} = P\left\{\omega : \ a_1 \leq \eta_1(\omega) \leq b_1 \bigcap a_2 \leq \eta_2(\omega) \leq b_2\right\} = \\ &= \int_{a_1}^{b_1} \int_{a_2}^{b_2} f(x, y)\, dx\, dy. \end{aligned} \tag{32}$$

(32) follows from Theorem 4 and (24).

§13. INDEPENDENT RANDOM VARIABLES

In this section we define the notion of independent random variables. We consider the case of two jointly distributed random variables.

Let $\eta_1(\omega)$ and $\eta_2(\omega)$ be jointly distributed random variables, with individual distribution functions $F_{\eta_1}(x)$, $F_{\eta_2}(x)$, respectively, and joint distribution function $F(x_1, x_2)$.

Definition 4. Two jointly distributed random variables $\eta_1(\omega)$ and $\eta_2(\omega)$ are independent if their joint distribution function $F(x_1, x_2)$ may be written as the product of their individual distribution functions $F_{\eta_1}(x)$ and $F_{\eta_2}(x)$ in the sense that, for any real numbers x_1 and x_2

$$F(x_1, x_2) = F_{\eta_1}(x_1) \cdot F_{\eta_2}(x_2). \tag{33}$$

Similarly, two jointly continuous random variables $\eta_1(\omega)$ and $\eta_2(\omega)$ are independent if their joint density function $f(x_1, x_2)$ may be written as the product of their individual density functions $f_{\eta_1}(x_1)$ and $f_{\eta_2}(x_2)$ in the sense that, for any real numbers x_1 and x_2

$$f(x_1, x_2) = f_{\eta_1}(x_1) \cdot f_{\eta_2}(x_2). \tag{34}$$

Equation (34) follows from (33) by differentiating both sides of (33) first with respect to x_1 and then with respect to x_2. Equation (33) follows from (34) by integrating both sides of (34).

Similarly, two discrete random variables $\eta_1(\omega)$ and $\eta_2(\omega)$ are independent if their joint probability mass function $p(x, y)$ may be written as the product of their individual probability mass functions $p_{\eta_1}(x)$ and $p_{\eta_2}(y)$ in the sense that, for any x and y

$$p(x, y) = p_{\eta_1}(x) \cdot p_{\eta_2}(y). \tag{35}$$

The equivalence follows because if (33) is satisfied then we obtain (35). Furthermore, if equation (35) is valid, then for any real numbers x_1 and x_2 we obtain

$$\begin{aligned} F(x_1, x_2) &= \sum_{x:\ x \le x_1} \sum_{y:\ y \le x_2} p(x, y) = \sum_{x:\ x \le x_1} \sum_{y:\ y \le x_2} p_{\eta_1}(x) \cdot p_{\eta_2}(y) = \sum_{x:\ x \le x_1} p_{\eta_1}(x) \cdot \sum_{y:\ y \le x_2} p_{\eta_2}(y) = \\ &= F_{\eta_1}(x_1) \cdot F_{\eta_2}(x_2) \end{aligned}$$

and therefore $\eta_1(\omega)$ and $\eta_2(\omega)$ are independent.

Thus, loosely speaking, $\eta_1(\omega)$ and $\eta_2(\omega)$ are independent if knowing the value of one does not change the distribution of the other. Random variables that are not independent are said to be dependent or nonindependent.

Lemma 2. *For any two independent random variables $\eta_1(\omega)$ and $\eta_2(\omega)$ their joint distribution function always satisfies the additional condition (28).*

Proof: As $\eta_1(\omega)$ and $\eta_2(\omega)$ are independent we can rewrite (23) in the following form

$$\begin{aligned} F(b_1, b_2) - F(a_1, b_2) - F(b_1, a_2) + F(a_1, a_2) &= F_{\eta_1}(b_1)\, F_{\eta_2}(b_2) - F_{\eta_1}(a_1)\, F_{\eta_2}(b_2) - \\ - F_{\eta_1}(b_1)\, F_{\eta_2}(a_2) + F_{\eta_1}(a_1)\, F_{\eta_2}(a_2) &= \left[F_{\eta_1}(b_1) - F_{\eta_1}(a_1)\right] \cdot \left[F_{\eta_2}(b_2) - F_{\eta_2}(a_2)\right]. \end{aligned} \tag{36}$$

As $F_{\eta_1}(\cdot)$ and $F_{\eta_2}(\cdot)$ are nondecreasing functions we conclude that the right–hand side of (36) is nonnegative. The proof is complete.

Independent random variables have the following exceedingly important property, the proof of which we leave as an exercise for the reader.

Theorem 5. *Let $\eta_1(\omega)$ and $\eta_2(\omega)$ be independent random variables and $\varphi_1(x)$ and $\varphi_2(x)$ be two continuous functions from $\mathbb{R}^1$ into $\mathbb{R}^1$. Then the random variables $\zeta_1(\omega) = \varphi_1(\eta_1(\omega))$ and $\zeta_2 = \varphi_2(\eta_2(\omega))$ are also independent.*

That is independence of the random variables $\eta_1(\omega)$ and $\eta_2(\omega)$ implies independence of the random variables $\zeta_1(\omega)$ and $\zeta_2(\omega)$.