Contents

Chapter 5

Multiple Random Variables

5.1 Two random variables

5.1.1 Continuous random variables:

We are given a probability space $(S, \mathcal{F}, P)$, and we define two random variables $X(\omega)$ and $Y(\omega)$ as follows:

Figure 5.1: Two r.v.'s $X(\omega)$ and $Y(\omega)$ mapping into R^2-plane.

(cf.)
Note that for the case of two random variables, each element $\omega \in S$ is being mapped into a point in the R^2-plane, whereas a single r.v. maps each $\omega \in S$ into a point on R^1-line!!!

Definition 5.1 The joint probability distribution function of two random variables $X(\omega)$ and $Y(\omega)$ is denoted and defined as follows:

$$F_{XY}(x, y) \triangleq P\{\omega \mid (X(\omega) \leq x) \cap (Y(\omega) \leq y)\}$$

Figure 5.2: The event defining the joint PDF $F_{XY}(x, y)$.

Note:

(i) The event defining the joint PDF is in *area*.

(ii) $F_{XY}(x, y)$ is a 2-dimensional surface as a function of x, and y on the xy-plane.

(iii) Recall the definition of PDF for a single r.v. $\ni$:

$$F_X(x) \triangleq P\{\omega \mid X(\omega) \leq x\}$$

where the event defining the PDF is in *interval*.

Definition 5.2 The joint probability density function of two random variables $X(\omega)$ and $Y(\omega)$ is denoted and defined as follows:

$$f_{XY}(x, y) \triangleq \frac{\partial^2}{\partial x \partial y} F_{XY}(x, y)$$

Remark: Notice that the relationship between the joint PDF and the joint p.d.f. is *differentiation/integration*, and thus the joint PDF can be expressed in terms of the joint p.d.f. as follows:

$$F_{XY}(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{XY}(\alpha, \beta) d\alpha d\beta$$

Properties of $F_{XY}(x, y)$:

1. $F_{XY}(x, y)$ in both of x and y at $-\infty$ is zero:

$$F_{XY}(-\infty, -\infty) = 0$$

2. $F_{XY}(x, y)$ in one of x or y at $-\infty$ is also zero:

$$F_{XY}(x, -\infty) = F_{XY}(-\infty, y) = 0 \quad \forall\, x, y$$

3. $F_{XY}(x, y)$ in both of x and y at ∞ is unity:

$$F_{XY}(\infty, \infty) = 1$$

4. If we let one of x or y be ∞, we get the PDF of y and x respectively, i.e.:

$$F_{XY}(x, \infty) = F_X(x)$$

$$F_{XY}(\infty, y) = F_Y(y)$$

   and we call these PDF's the **"marginal dsitributions"**.

5. $F_{XY}(x, y)$ is a *non-decreasing* function of x and y.

6. $F_{XY}(x, y)$ is *right-hand continuous* in both of x and y.

7. For $F_{XY}(x, y)$ to be a **valid** joint PDF, it must satisfy the following inequality:

$$F_{XY}(x_2, y_2) - F_{XY}(x_1, y_2) - F_{XY}(x_2, y_1) + F_{XY}(x_1, y_1) \geq 0$$
$$\forall\, x_2 \geq x_1, \quad y_2 \geq y_1$$

Brief proof:

1. Notice that:

$$F_{XY}(-\infty,-\infty) = P\{(X \le -\infty) \cap (Y \le -\infty)\} = P(\phi \cap \phi) = P(\phi) = 0$$

2. We have:

$$F_{XY}(x,-\infty) = P\{(X \le x) \cap (Y \le -\infty)\} = P\{(X \le x) \cap \phi\} = P(\phi) = 0$$

and

$$F_{XY}(-\infty,y) = P\{(X \le -\infty) \cap (Y \le y)\} = P\{\phi \cap (Y \le y)\} = P(\phi) = 0$$

3. This is so since:

$$F_{XY}(\infty,\infty) = P\{(X \le \infty) \cap (Y \le \infty)\} = P(S \cap S) = P(S) = 1$$

4. Notice that:

$$\begin{aligned} F_{XY}(x,\infty) &= P\{\omega \mid (X(\omega) \le x) \cap (Y(\omega) \le \infty)\} \\ &= P\{\omega \mid (X(\omega) \le x) \cap S\} \\ &= P\{\omega \mid X(\omega) \le x\} \\ &\triangleq F_X(x) \end{aligned}$$

Similarly, we can prove that $F_{XY}(\infty, y) = F_Y(y)$ as well.

5. This is implicitly indicated in the process of proving the property 7 below.

6. We omit, but you can prove this property in a similar way as in the case of single random variable, and the property is due to the inequality sign($\le$) in the definition of the joint PDF. If we had defined it using the strict inequality sign($<$), it would have been *left-hand continuous*.

7. From the figure below, we have the following probability of the shaded area:

Figure 5.3: The range space of X and Y.

$$P\{\omega \mid (x_1 < X(\omega) \leq x_2) \cap (y_1 < Y(\omega) \leq y_2)\}$$

$$\begin{aligned} &= P\{\omega|\, (X(\omega) \leq x_2) \cap (Y(\omega) \leq y_2)\} - P\{\omega|\, (X(\omega) \leq x_1) \cap (Y(\omega) \leq y_2)\} \\ &\quad -P\{\omega|\, (X(\omega) \leq x_2) \cap (Y(\omega) \leq y_1)\} + P\{\omega|\, (X(\omega) \leq x_1) \cap (Y(\omega) \leq y_1)\} \\ &\triangleq F_{XY}(x_2, y_2) - F_{XY}(x_1, y_2) - F_{XY}(x_2, y_1) + F_{XY}(x_1, y_1) \\ &\geq 0 \quad \text{(from the axiom \#1 of probability} \ni\text{: } P(\cdot) \geq 0) \end{aligned}$$

(cf.) Notice that we have implicitly used the fact that the *disjoint* areas in R^2-space correspond to the *mutually exclusive* events!!!

Example 5.1

Is $F_{XY}(x, y)$ given below a valid joint PDF?

$$F_{XY}(x, y) = \begin{cases} 0, & x < 0, \text{ or } x + y < 1 \text{ or } y < 0 \\ 1, & \text{elsewhere} \end{cases}$$

Figure 5.4: $F_{XY}(x, y)$ in xy-plane.

Solution:

We can check that all the properties from 1 to 6 are satisfied, but the probability of the shaded area A in above figure is:

$$
\begin{aligned}
& P\{\omega \mid (x_1 < X(\omega) \le x_2) \cap (y_1 < Y(\omega) \le y_2)\} \\
= \; & F_{XY}(x_2, y_2) - F_{XY}(x_1, y_2) - F_{XY}(x_2, y_1) + F_{XY}(x_1, y_1) \\
& \quad (\text{let } x_1 = \tfrac{1}{2}, \;\; x_2 = 1, \;\; y_1 = \tfrac{1}{4}, \;\; y_2 = \tfrac{3}{4}) \\
= \; & F_{XY}(1, \frac{3}{4}) - F_{XY}(\frac{1}{2}, \frac{3}{4}) - F_{XY}(1, \frac{1}{4}) + F_{XY}(\frac{1}{2}, \frac{1}{4}) \\
= \; & 1 - 1 - 1 + 0 \\
= \; & -1 \\
< \; & 0 \quad (\text{wrong!!!})
\end{aligned}
$$

which means that the property 7 is violated.

Therefore, above $F_{XY}(x, y)$ CANNOT be a valid joint distribution function...

Properties of $f_{XY}(x, y)$:

1. The joint density is non-negative for all x and y:

$$f_{XY}(x, y) \geq 0$$

2. The volume under the joint p.d.f. is always unity:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{XY}(x, y) dx dy = 1$$

3. The **marginal** PDF and p.d.f. of $Y(\omega)$ can respectively be obtained by the integrations below:

$$F_Y(y) = \int_{-\infty}^{\infty} \int_{-\infty}^{y} f_{XY}(\alpha, \beta) d\beta d\alpha$$
$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x, y) dx$$

4. Similarly, the **marginal** PDF and p.d.f. of $X(\omega)$ can be obtained respectively by the integrations below:

$$F_X(x) = \int_{-\infty}^{x} \int_{-\infty}^{\infty} f_{XY}(\alpha, \beta) d\beta d\alpha$$
$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x, y) dy$$

5. The probability of a rectangle in R^2-space can be calculated using the joint p.d.f. as:

$$P\{(x_1 < X \leq x_2) \cap (y_1 < Y \leq y_2)\} = \int_{x_1}^{x_2} \int_{y_1}^{y_2} f_{XY}(x, y) dy dx$$

In general, the probability of an event such that r.v.'s X and Y mapping into any area A in R^2-space is as follows:

$$P\{(X, Y) \in A\} = \int_A \int f_{XY}(x, y) dy dx$$

Figure 5.5: Any area A in XY-plane.

Brief proof:

1. This is because the joint PDF is non-decreasing function of x and y, and $f_{XY}(x,y)$ is the derivative of $F_{XY}(x,y)$ w.r.t. x and y.

2. Notice that from the differentiation/integration relation of the joint PDF and p.d.f., we have:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x,y)dxdy \equiv F_{XY}(\infty,\infty) = 1$$

3. From the property of the joing PDF, we know that the marginal PDF of Y is $F_Y(y) = F_{XY}(\infty, y)$, thus:

$$\begin{aligned} F_Y(y) &= F_{XY}(\infty, y) \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{y} f_{XY}(\alpha,\beta)d\beta d\alpha \end{aligned}$$

Therefore, by taking the derivative of $F_Y(y)$ w.r.t. y, we get the density of Y as:

$$\begin{aligned} f_Y(y) \triangleq \frac{d}{dy}F_Y(y) &= \int_{-\infty}^{\infty}\frac{d}{dy}\left\{\int_{-\infty}^{y} f_{XY}(\alpha,\beta)d\beta\right\}d\alpha \quad \text{(Leibnitz rule)} \\ &= \int_{-\infty}^{\infty}\frac{dy}{dy}\cdot f_{XY}(\alpha,y)d\alpha \quad \text{(Leibnitz rule)} \\ &= \int_{-\infty}^{\infty} f_{XY}(x,y)dx \end{aligned}$$

: called "marginal density"

4. This can be proved in the same manner as in 3.

5. *Assignment* : Express the probability in terms of the joint PDF, and use the relation between the joint PDF and p.d.f..

5.1.2 Discrete random variables:

We begin with a specific example of two discrete random variables:

Example 5.2

For the two random variables X and Y defined below, find the joint probability distribution function $F_{XY}(x, y)$.

$$S = \{\omega_1, \omega_2, \omega_3\}$$

$$\begin{cases} X(\omega_1) = 1, \quad Y(\omega_1) = 1 \\ X(\omega_2) = 2, \quad Y(\omega_2) = 1 \\ X(\omega_3) = 3, \quad Y(\omega_3) = 3 \end{cases}$$

where $P(\omega_1) = 0.2$, $P(\omega_2) = 0.3$, and $P(\omega_3) = 0.5$.

Figure 5.6: The sample space S and r.v.'s X, Y mapping into XY-plane.

Solution:

From the definition of the joint PDF, we have:

$$F_{XY}(x, y) = P\{\omega \mid (X(\omega) \leq x) \cap (Y(\omega) \leq y)\}$$

(i) $x = 0,\ y = 0$: $\quad F_{XY}(0, 0) = P(\phi) = 0$

(ii) $x = 1,\ y = 0$: $\quad F_{XY}(1, 0) = P(\phi) = 0$

(iii) $x = 1,\ y = 1$: $\quad F_{XY}(1, 1) = P(\{\omega_1\}) = 0.2$

$$\vdots$$

$$\vdots$$

Let

$$p(x,y) \triangleq P\{\omega \mid (X(\omega)=x) \cap (Y(\omega)=y)\}$$

then, the joint PDF can be expressed as a sum of the weighted, and shifted 2-dimensional unit step function (or surface), similary to the case of single discrete r.v., i.e.:

$$F_{XY}(x,y) = p(1,1)u(x-1)u(y-1) + p(2,1)u(x-2)u(y-1) + p(3,3)u(x-3)u(y-3)$$

In general, if there $\exists$ NM points in R^2-space such as $(x_i, y_i), \quad i = 1,2,\ldots,N, \quad j = 1,2,\ldots,M$:

Figure 5.7: NM points in XY-plane being mapped by two discrete r.v.'s X, Y.

Then, the joint PDF can be represented in the following **fixed** formula:

$$F_{XY}(x,y) = \sum_{i=1}^{N}\sum_{j=1}^{M} p(x_i,y_j)u(x-x_i)u(y-y_j)$$

where $p(x_i,y_j) \triangleq P\{\omega \mid (X(\omega)=x_i) \cap (Y(\omega)=y_j)\}$

Corresponding joint p.d.f. for discrete two r.v.'s is also in the fixed form of:

$$\begin{aligned}
f_{XY}(x,y) &\triangleq \frac{\partial^2}{\partial x \partial y} F_{XY}(x,y) \\
&= \frac{\partial^2}{\partial x \partial y} \left\{ \sum_{i=1}^{N} \sum_{j=1}^{M} p(x_i, y_j) u(x - x_i) u(y - y_j) \right\} \\
&= \sum_{i=1}^{N} \sum_{j=1}^{M} p(x_i, y_j) \frac{\partial^2}{\partial x \partial y} \left\{ u(x - x_i) u(y - y_j) \right\} \\
&= \sum_{i=1}^{N} \sum_{j=1}^{M} p(x_i, y_j) \delta(x - x_i) \delta(y - y_j)
\end{aligned}$$

where $\delta(\cdot)$ is the Dirac delta function.

Figure 5.8: An example of the joint p.d.f. for two discrete r.v.'s X, Y.

Summary:
Given a probability space $(S, \mathcal{F}, P)$, the joint PDF and the joint p.d.f. of two random variables X and Y are as follows, regardless of whether they are *continuous* or *discrete*:

$$\begin{cases} F_{XY}(x,y) \triangleq P\left\{\omega \mid \left(X(\omega) \le x\right) \cap \left(Y(\omega) \le y\right)\right\} \\ f_{XY}(x,y) \triangleq \frac{\partial^2}{\partial x \partial y} F_{XY}(x,y) \end{cases}$$

5.2 Conditional distribution and conditional density between two random variables

We now consider the concept of the conditional distribution and density functions of a r.v. $X(\omega)$ given a value of another r.v. $Y(\omega)$, i.e. $Y = y$, where the joint PDF and joint p.d.f. $F_{XY}(x, y)$ and $f_{XY}(x, y)$ are known. [1]

Figure 5.9: Two r.v.'s X and Y mapping into R^2-space.

Here, an element $\omega \in S$ maps into a point (x, y) in R^2-plane via two r.v.'s $X(\omega)$ and $Y(\omega)$ as:

$$X(\omega) \longrightarrow x$$

$$Y(\omega) \longrightarrow y$$

Recall:

Let two events A and B be as follows:

$$A = \{\omega \mid X(\omega) \leq x\}$$

$$B = \{\omega \mid X(\omega) \in \mathcal{R}\}$$

where $\mathcal{R}$ is the set of real numbers, and thus B is some kind of event related to the r.v. $Y(\omega)$.

[1]The one dimensional slice(or cut) image of $F_{XY}(x, y)$ along the line $Y = y$.

Then, the conditional distribution function and the conditional density function of $X(\omega)$ based on the event B are defined respectively as follows:

$$\begin{aligned} F_{X|Y}(x|B) &\triangleq P\left[\{X(\omega) \leq x\} \mid B\right] \\ &= \frac{P\left[\{\omega \mid (X(\omega) \leq x) \cap B\}\right]}{P(B)} \end{aligned} \tag{5.1}$$

$$f_{X|Y}(x|y) \triangleq \frac{d}{dx} F_{X|Y}(x|B) \tag{5.2}$$

(cf.) You may have to check that (5.1) and (5.2) are valid definitions.

Now, let the event B be specifically as:

$$B = \{\omega \mid y - \Delta y < Y(\omega) \leq y + \Delta y\}$$

Then, the conditional distribution in (5.1) becomes:

$$F_{X|Y}(x|B) = \frac{P\left[\{\omega \mid (X(\omega) \leq x) \cap (y - \Delta y < Y(\omega) \leq y + \Delta y)\}\right]}{P\left(\omega \mid y - \Delta y < Y(\omega) \leq y + \Delta y\right)} \tag{5.3}$$

Here in (5.3), the numerator and the denominator can each be claculated as:

$$\begin{aligned} \text{numerator} &= P\left[\{\omega \mid (X(\omega) \leq x) \cap (y - \Delta y < Y(\omega) \leq y + \Delta y)\}\right] \\ &= \int_{-\infty}^{x} \int_{y-\Delta y}^{y+\Delta y} f_{XY}(u, v) dv du \end{aligned} \tag{5.4}$$

and

$$\begin{aligned} \text{denominator} &= P\left(\omega \mid y - \Delta y < Y(\omega) \leq y + \Delta y\right) \\ &= \int_{y-\Delta y}^{y+\Delta y} f_Y(v) dv \end{aligned} \tag{5.5}$$

Inserting (5.4) and (5.5) into (5.3), we get:

$$F_{X|Y}\left(x \mid y - \Delta y < Y(\omega) \le y + \Delta y\right) = \frac{\int_{-\infty}^{x} \int_{y-\Delta y}^{y+\Delta y} f_{XY}(u,v)dvdu}{\int_{y-\Delta y}^{y+\Delta y} f_Y(v)dv} \tag{5.6}$$

Case #1: X and Y are both continuous r.v.'s

In this case, as $\Delta y \rightarrow 0$ the integrals in (5.6) can be approximated to the following expressions by the **mean value theorem**:

As $\Delta y \rightarrow 0$, we have:

$$\int_{y-\Delta y}^{y+\Delta y} f_{XY}(u,v)dv = f_{XY}(u,y) \cdot 2\Delta y$$

$$\int_{y-\Delta y}^{y+\Delta y} f_Y(v)dv = f_Y(y) \cdot 2\Delta y$$

Figure 5.10: Approximation of integral by the mean value theorem.

Therefore, from (5.6), we get:

$$\begin{aligned} F_{X|Y}(x|Y=y) \overset{\text{let}}{=} F_{X|Y}(x|y) &= \lim_{\Delta y \to 0} F_{X|Y}\left(x \mid y - \Delta y < Y(\omega) \le y + \Delta y\right) \\ &= \frac{\int_{-\infty}^{x} f_{XY}(u,y) \cdot 2\Delta y \, du}{f_Y(y) \cdot 2\Delta y} \\ &= \frac{\int_{-\infty}^{x} f_{XY}(u,y)du}{f_Y(y)} \end{aligned}$$

Also from (5.2), the conditional p.d.f. of X given $Y = y$ is in the following form:

$$f_{X|Y}(x|Y=y) \overset{\text{let}}{=} f_{X|Y}(x|y) = \frac{d}{dx} F_{X|Y}(x|y)$$

$$= \frac{\frac{d}{dx}\int_{-\infty}^{x} f_{XY}(u,y)du}{f_Y(y)}$$

$$= \frac{f_{XY}(x,y)}{f_Y(y)} \qquad \text{(by the Leibnitz rule)}$$

Case #2: X and Y are both discrete r.v.'s

In this case, recall that the joint p.d.f. of $X(\omega)$ and $Y(\omega)$, and the marginal p.d.f. of $Y(\omega)$ are in the following fixed forms:

$$f_{XY}(x,y) = \sum_{i=1}^{N}\sum_{j=1}^{M} p(x_i, y_j)\delta(x-x_i)\delta(y-y_j)$$

and

$$f_Y(y) = \sum_{j=1}^{M} p(y_j)\delta(y-y_j)$$

Applying above two expressions to (5.6) and 5.2), we will eventually obtain the conditional PDF and p.d.f. of X given $Y = y_k$ as follows:

$$F_{X|Y}(x|Y=y_k) = \frac{\sum_{i=1}^{N} p(x_i y_k)u(x-x_i)}{p(y_k)}$$

$$f_{X|Y}(x|Y=y_k) = \frac{\sum_{i=1}^{N} p(x_i y_k)\delta(x-x_i)}{p(y_k)}$$

proof: assignment [2]

[2]Or, we can directly apply (5.3) to obtain the conditional distribution function.

Example 5.3

Given the joint p.d.f. of two r.v.'s X and Y below, find the conditional p.d.f. of Y given $X = x$, i.e. $f_{Y|X}(y|X = x)$.

$$f_{XY}(x, y) = \begin{cases} 2, & 0 \le x \le y \le 1 \\ 0, & \text{elsewhere} \end{cases}$$

Figure 5.11: The joint p.d.f. $f_{XY}(x, y)$.

(cf.) Notice the above $f_{XY}(x, y)$ satisfies the following property which is required to be a valid p.d.f.:

$$\begin{aligned} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x, y)dxdy &= \int\int_{\text{shaded area}} f_{XY}(x, y)dxdy \\ &= 2 \times 1 \times 1 \times \frac{1}{2} \\ &= 1 \end{aligned}$$

Solution:

The conditional p.d.f. which we want to obtain is as follows:

$$f_{Y|X}(y|X = x) = \frac{f_{XY}(x, y)}{f_X(x)}$$

Since we are given $f_{XY}(x, y)$, we must compute the marginal p.d.f. $f_X(x)$ of $X(\omega)$ which is:

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x, y)dy = \int_x^1 2dy = \begin{cases} 2(1 - x), & 0 \le x \le 1 \\ 0, & \text{otherwise} \end{cases}$$

Figure 5.12: The marginal p.d.f. $f_X(x)$.

Therefore, the conditional p.d.f. becomes: [3] [4]

$$
\begin{aligned}
f_{Y|X}(y|X=x) &= \frac{2}{2(1-x)} \\
&= \begin{cases} 1/(1-x), & (0 \leq)x \leq y \leq 1 \\ undefined, & \text{otherwise} \end{cases}
\end{aligned}
$$

Figure 5.13: The conditional p.d.f. $f_{Y|X}(y|x)$.

Check:

$$(1)\ \int_{-\infty}^{\infty} f_X(x)dx = 1 \times 2 \times \frac{1}{2} = 1$$

$$(2)\ \int_{-\infty}^{\infty} f_{Y|X}(y|x)dy = \int_x^1 \frac{1}{1-x}dy = \frac{1-x}{1-x} = 1$$

[3] In this expression, x is a fixed parameter, NOT a variable!!!

[4] The p.d.f. is not defined for the cases other than $x \leq y \leq 1$, since $f_X(x) = 0$.

5.3 Relationships between two random variables

5.3.1 Statistical independence

Recall: If we are given two *independent* events A and B, then

$$P(A \cap B) = P(A) \cdot P(B)$$

Independent random variables:

Let $X(\omega)$ and $Y(\omega)$ be two *independent* [5] rnadom variables, then we have the following relationships for the joint PDF and the conditional PDF of $X(\omega)$ and $Y(\omega)$:

(1) The joint distribution:

$$\begin{aligned} F_{XY}(x,y) &\triangleq P\{\omega \mid (X(\omega) \leq x) \cap (Y(\omega) \leq y)\} \\ &\overset{\text{let}}{=} P\{\underbrace{(X \leq x)}_{\text{A}} \cap \underbrace{(Y \leq y)}_{\text{B}}\} \\ &= P(X \leq x) \cdot P(Y \leq y) \\ &= F_X(x) \cdot F_Y(y) \qquad (5.7) \end{aligned}$$

(2) The conditional distribution:

$$\begin{aligned} F_{X|Y}(x|Y=y) &\triangleq \frac{P\{(X \leq x) \cap (Y = y)\}}{P(Y=y)} \\ &= \frac{P\,(X \leq x) \cdot P\,(Y = y)}{P(Y=y)} \\ &= F_X(x) \qquad (5.8) \end{aligned}$$

[5] This means that events A and B defined by random variables $X(\omega)$ and $Y(\omega)$ respectively, are independent!!!

(cf.)
Above argument ia not correct in rigorous sense, since $P(Y = y) = 0$ for continuous r.v. $Y(\omega)$. Instead, we could have derived the relation in the following way:

$$\begin{aligned} F_{X|Y}(x|y) &= \frac{\int_{-\infty}^{x} f_{XY}(u,y)du}{f_Y(y)} \\ &= \frac{\int_{-\infty}^{x} f_X(u)du \cdot f_Y(y)}{f_Y(y)} \\ &= \int_{-\infty}^{x} f_X(u)du \\ &= F_X(x) \end{aligned}$$

Differentiating (5.7) and (5.8), we can show the following relationships of the joint p.d.f. and the conditional p.d.f. for *independent* r.v.'s $X(\omega)$ and $Y(\omega)$:

$$\begin{cases} f_{XY}(x,y) = f_X(x) \cdot f_Y(y) \\ f_{X|Y}(x|y) = f_X(x) \end{cases}$$

5.3.2 The correlation of random variables

Correlation:

Definition 5.3 The *correlation* of two random variables $X(\omega)$ and $Y(\omega)$ is denoted and defined as the following mathematical expectation:

$$R_{XY} \triangleq E[XY] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy \cdot f_{XY}(x,y)dxdy$$

Definition 5.4 According to the correlation R_{XY}, we define the following relationships between two r.v.'s $X(\omega)$ and $Y(\omega)$:

(1) If $R_{XY} = 0$, then $X(\omega)$ and $Y(\omega)$ are said to be **orthogonal**.

(2) If $R_{XY} = E[X]E[Y]$, then $X(\omega)$ and $Y(\omega)$ are said to be **uncorrelated**.

Remarks:

(i) If X and Y are *independent*, then X and Y are *uncorrelated*, but NOT vice versa, i.e.

$$\begin{array}{ccc} X \text{ and } Y \text{ are independent} & \xrightarrow{O} & X \text{ and } Y \text{ are uncorrelated} \\ (\quad \cdots \quad) & \xleftarrow{X} & (\quad \cdots \quad) \end{array}$$

(ii) Be careful with the definition of the "uncorrelatedness", i.e. notice that: $R_{XY} = 0$ does NOT indicate that X and Y are uncorrelated!!!

(cf.) Do not be confused between *independence* and *uncorrelatedness*.

Covariance:

Definition 5.5 The *covariance* of two random variables $X(\omega)$ and $Y(\omega)$ is denoted and defined as the following joint central moment:

$$C_{XY} \triangleq E\left[(X - m_X)(Y - m_Y)\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x - m_X)(y - m_Y) \cdot f_{XY}(x, y)dxdy$$

Note:
The uncorrelatedness between teo r.v.'s X and Y can be defined interms of the covariance as follows:

If the covariance of X and Y is $C_{XY} = 0$, then X and Y are uncorrelated

proof: assignment(easy!)

Correlation coefficient:

Definition 5.6 The *correlation coefficient* of two random variables $X(\omega)$ and $Y(\omega)$ is denoted and defined as the following normalized joint central moment:

$$\rho_{XY} \triangleq E\left[\left(\frac{X - m_X}{\sigma_X}\right)\left(\frac{Y - m_Y}{\sigma_Y}\right)\right] = \frac{C_{XY}}{\sigma_X \sigma_Y}$$

FACT: The maginitude of the correlation coefficient is less than or equal to unity, i.e.:

$$|\rho_{XY}| \leq 1$$

Proof:
Let an expectation A defined as follows:

$$A \triangleq E\left[\{\alpha(X - m_X) + (Y - m_Y)\}^2\right]$$

where α is an any real number. [6]
Then, since A is an expectation of a square term, it must be non-negative, i.e. $A \geq 0$.
Now, we have:

$$\begin{aligned} A &= E\left[\alpha^2(X - m_X)^2 + 2\alpha(X - m_X)(Y - m_Y) + (Y - m_Y)^2\right] \\ &= \alpha^2 E\left[(X - m_X)^2\right] + 2\alpha E\left[(X - m_X)(Y - m_Y)\right] + E\left[(Y - m_Y)^2\right] \\ &= \alpha^2\sigma_X^2 + 2\alpha C_{XY} + \sigma_Y^2 \\ &\geq 0 \quad \forall\, \alpha \\ & \quad \text{(should be)} \end{aligned}$$

Therefore, the discriminant must be as follows:

$$\frac{D}{4} = C_{XY}^2 - \sigma_X^2\sigma_Y^2 \leq 0$$

from which it follows:

$$\frac{C_{XY}^2}{\sigma_X^2\sigma_Y^2} \leq 1 \quad \longrightarrow \quad -1 \leq \rho_{XY} \leq 1$$

[6] α is called the *Lagrange multiplier*.

Example 5.4

Let a new random variable Y be as:

$$Y = cX$$

where c is a real constant.

Then find the mean m_Y, variance σ_Y^2 of the newly defind r.v. Y, and the correlation coefficient ρ_{XY} between X and Y.

Solution:

(i) Mean m_Y:

$$m_Y = E[Y] = E[cX] = c \cdot E[X] = c \cdot m_X$$

(ii) Variance σ_Y^2:

$$\begin{aligned} \sigma_Y^2 = E[Y^2] - m_Y^2 &= E[c^2 X^2] - c^2 m_X^2 \\ &= c^2 \left\{ E[X^2] - m_X^2 \right\} \\ &= c^2 \sigma_X^2 \end{aligned}$$

(iii) Correlation coefficient ρ_{XY}:
The covariance C_{XY} is:

$$\begin{aligned} C_{XY} = E[(X - m_X)(Y - m_Y)] &= E(x - m_X)(cX - c \cdot m_X)] \\ &= c \cdot E\left[(X - m_X)^2\right] \\ &= c \cdot \sigma_X^2 \end{aligned}$$

Therefore, the correlation coefficient becomes:

$$\rho_{XY} = \frac{C_{XY}}{\sigma_X \sigma_Y} = \frac{c \cdot \sigma_X^2}{\pm c \cdot \sigma_X^2} = \pm 1$$

Note: Notice that depending on the sign of the constant c, ρ_{XY} respectively is:

$$\begin{cases} +1, & \text{if } c > 0 \\ -1, & \text{if } c < 0 \end{cases}$$

Joint characteristic function:

Definition 5.7 The joint characteristic function of two randoma variables X and Y is denoted and defined as the following mathematical expectation:

$$\begin{aligned}\Phi(\omega_1,\omega_2) &\triangleq E\left[e^{j(\omega_1 X+\omega_2 Y)}\right] \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{j(\omega_1 x+\omega_2 y)} f_{XY}(x,y)dxdy\end{aligned}$$

Note:

(i) The definition of joint characteristic function is similar to the two dimensional inverse Fourier transform.

(ii) Based on the similarity mentioned in (i), the joint p.d.f. $f_{XY}(x,y)$ can be obtained from $\Phi(\omega_1,\omega_2)$ as :

$$f_{XY}(x,y) = \frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \Phi(\omega_1,\omega_2)e^{-j(\omega_1 x+\omega_2 y)}d\omega_1 d\omega_2$$

(iii) If two r.v.'s X and Y are independent, then the joint characteristic function becomes:

$$\Phi(\omega_1,\omega_2) = \Phi_X(\omega_1)\cdot\Phi_Y(\omega_2)$$

proof: assignment

5.4 Sum of two random variables

5.4.1 The distribution and density functions

Let X and Y be two (independent) random variables, and suppose the joint and marginal p.d.f.'s $f_{XY}(x, y)$, $f_X(x)$, and $f_Y(y)$ are given. Define a new random variable W as the sum of the given two r.v.'s, i.e.:

$$W \triangleq X + Y$$

Then, determine the probability distribution and density functions $F_W(w)$ and $f_W(w)$ of the newly defined r.v. W.

1. The PDF $F_W(w)$:

$$\begin{aligned} F_W(w) &= P\left[W \leq w\right] \\ &= P\left[X + Y \leq w\right] \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{w-y} f_{XY}(x, y)dxdy \\ &= \int_{-\infty}^{\infty} f_Y(y)\left\{\int_{-\infty}^{w-y} f_X(x)dx\right\}dy \qquad \text{(if } X \text{ and } Y \text{ are independent)} \end{aligned}$$

Figure 5.14: The integration region in the order of x and y.

(cf.) If we reverse the order of integration we could get another expression or formula as follows;

$$
\begin{aligned}
F_W(w) &= \int_{-\infty}^{\infty}\int_{-\infty}^{w-x} f_{XY}(x,y)dydx \\
&= \int_{-\infty}^{\infty} f_X(x)\left\{\int_{-\infty}^{w-x} f_Y(y)dy\right\}dx \qquad (\text{ if } X \text{ and } Y \text{ are independent})
\end{aligned}
$$

Figure 5.15: The integration region in the order of y and x.

2. The p.d.f. $f_W(w)$:

$$
\begin{aligned}
f_W(w) &= \frac{d}{dw}F_W(w) \\
&= \frac{d}{dw}\int_{-\infty}^{\infty}\int_{-\infty}^{w-y} f_{XY}(x,y)dxdy \\
&= \int_{-\infty}^{\infty}\left\{\frac{d}{dw}\int_{-\infty}^{w-y} f_{XY}(x,y)dx\right\}dy \qquad (\text{by the Leibnitz rule}) \\
&= \int_{-\infty}^{\infty} f_{XY}(w-y,y)dy \qquad (\text{by the Leibnitz rule}) \\
&= \int_{-\infty}^{\infty} f_X(w-y)f_Y(y)dy \qquad (\text{ if } X \text{ and } Y \text{ are independent}) \\
&\triangleq f_Y(w) * f_X(w)
\end{aligned}
$$

: CONVOLUTION INTEGRAL

(cf.) If we reverse the order of integration we could get another expression or formula for the case of independent X and Y as follows;

$$
f_W(w) = \int_{-\infty}^{\infty} f_X(x)f_Y(w-x)dx \triangleq f_X(w) * f_Y(w)
$$

Example 5.5

We are given two *independent* r.v.'s $X(\omega)$ and $Y(\omega)$, whose p.d.f.'s are as follows:

$$f_X(x) = \frac{1}{a}\{u(x) - u(x-a)\}$$

$$f_Y(y) = \frac{1}{b}\{u(x) - u(x-b)\}$$

where $b > a$. That is; X and Y are *uniformly* distributed in the intervals of $[0,a)$ and $[0,b)$ respectively, i.e. $X \sim U[0,a)$ and $Y \sim U[0,b)$.

Then, find the p.d.f. of a new random variable defined as the sum of X amd Y:

$$W \triangleq X + Y$$

Figure 5.16: The p.d.f. $f_X(x)$ and $f_Y(y)$.

Solution:

The p.d.f. $f_W(w)$ is the convolution of $f_Y(w)$ and $f_X(w)$, since X and Y are independent:

$$f_W(w) = f_Y(w) * f_X(w) = \int_{-\infty}^{\infty} f_Y(y) f_X(w-y) dy$$

Figure 5.17: The convolution $f_W(w) = f_Y(w) * f_X(w)$.

(i) $w < 0$:

$$f_W(w) = 0$$

(ii) $0 \le w < a$:

$$f_W(w) = \int_0^w \frac{1}{a} \cdot \frac{1}{b} dy = \frac{1}{ab} w$$

(iii) $a \le w < b$:

$$f_W(w) = \int_{w-a}^{w} \frac{1}{a} \cdot \frac{1}{b} dy = \frac{1}{ab}(w - w + a) = \frac{1}{b}$$

(iv) $b \le w < a + b$:

$$f_W(w) = \int_{w-a}^{b} \frac{1}{a} \cdot \frac{1}{b} dy = \frac{1}{ab}(b - w + a) = -\frac{w}{ab} + \frac{a+b}{ab}$$

(v) $w \ge a + b$:

$$f_W(w) = 0$$

Figure 5.18: The p.d.f. $f_W(w)$.

(cf.)

(1) Note that the integration of $f_W(w)$ over the entire real line is unity:

$$\int_{-\infty}^{\infty} f_W(w) dw = \frac{a}{b} + \frac{b-a}{b} = 1$$

(2) Try $f_W(w) = f_X(w) * f_Y(w)$, and see if you get the same result.
: *assignment*

5.4.2 The characteristic function

Suppose we define a new r.v. Z as the sum of two *independent* r.v.'s X and Y as:

$$Z \triangleq X + Y$$

Then, the characteristic function of the newly defined r.v. Z becomes:

$$\begin{aligned}
\Phi_Z(\omega) &\triangleq E\left[e^{j\omega Z}\right] \\
&= E\left[e^{j\omega(X+Y)}\right] \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{j\omega(x+y)} f_{XY}(x,y)dxdy \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{j\omega x}\cdot e^{j\omega y} f_X(x)\cdot f_Y(y)dxdy \quad \text{(since } X \text{ and } Y \text{ are indep.)} \\
&= \int_{-\infty}^{\infty} e^{j\omega x} f_X(x)dx \cdot \int_{-\infty}^{\infty} e^{j\omega y} f_Y(y)dy \\
&\triangleq \Phi_X(\omega)\cdot\Phi_Y(\omega)
\end{aligned}$$

(cf.) Note that $\Phi_Z(\omega)$ is a one dimensional function of ω, NOT a joint characteristic function: do not be confused!!!

Remark:
Notice that $f_X(x)$ and $f_Y(y)$ play similar roles of the input signals and the impulse response of an LTI system, and $\Phi_X(\omega)$ and $\Phi_Y(\omega)$ play roles of their Fourier transforms (i.e. F.T of the input signals and the system's transfer function.)
: All under the assumption that X and Y are independent!!!

Figure 5.19: The sum of two independent r.v.'s vs. an LTI system.

$$f_Z = f_X * f_Y$$

$$\Phi_Z = \Phi_X \cdot \Phi_Y$$

$$\text{where} \quad \Phi_X(\omega) = \int_{-\infty}^{\infty} f_X(x)e^{j\omega x}dx$$

5.5 Generalization to multiple random variables

We can now generalize the concepts discussed in the previous section to the multiple (i.e. more that three r.v.'s) random variable case.

For a probability space $(S, \mathcal{F}, P)$, we are given N random variables, $X_1(\omega)$, $X_2(\omega)$, ..., $X_N(\omega)$ mapping into a point in R^N-space as follows:

Figure 5.20: The sum of two independent r.v.'s vs. an LTI system.

NOTE:
This generalization will be the foundation of formulating the concept of the **random process** in later section!!!

1. The joint probability distribution function:

Definition 5.8 The joint probability distribution function of N random variables $X_1, X_2, \ldots, X_N$ is denoted and defined as the followng probability:

$$\begin{aligned} F_N(x_1, x_2, \ldots, x_N) &\triangleq P\{\omega \mid (X_1(\omega) \le x_1) \cap (X_2(\omega) \le x_2) \cap \cdots (X_N(\omega) \le x_N)\} \\ &= P\left\{\bigcap_{i=1}^{N} (X_i(\omega) \le x_i)\right\} \end{aligned}$$

2. The joint probability density function:

Definition 5.9 The corresponding joint probability density function of N random variables $X_1, X_2, \ldots, X_N$ is denoted and defined as follows:

$$f_N(x_1, x_2, \ldots, x_N) \triangleq \frac{\partial^N}{\partial x_1 \partial x_2 \cdots \partial x_N} F_N(x_1, x_2, \ldots, x_N)$$

(cf.)
Notice that the joint PDF and the joint p.d.f. of N r.v.'s are related by integration/differentiation:

$$F_N(x_1, x_2, \ldots, x_N) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \cdots \int_{-\infty}^{x_N} f_N(\alpha_1, \alpha_2, \ldots, \alpha_N) d\alpha_1 d\alpha_2 \cdots d\alpha_N$$

3. Properties of the joint probability distribution function:

(1) $F_N(x_1, x_2, \ldots, x_N)$ is non-decreasing function of each of its argument.

(2) $F_N(x_1, x_2, \ldots, x_N)$ is right-hand continuous in each of its argument, i.e.:

$$\lim_{\epsilon_k \to 0,\ \epsilon_k > 0} F_N(x_1 + \epsilon_1, x_2 + \epsilon_2, \ldots, x_N + \epsilon_N) = F_N(x_1, x_2, \ldots, x_N)$$

(3) If any one of the arguments is at $-\infty$, the joint PDF is zero, i.e.:

$$F_N(x_1, x_2, \ldots, x_N) = 0 \quad \text{if any } x_k \to -\infty$$

And, of course we have:

$$F_N(-\infty, -\infty, \ldots, -\infty) = 0$$

(4) The joint PDF is unity when all of the arguments are at ∞, i.e.:

$$F_N(\infty, \infty, \ldots, \infty) = 1$$

(5) The *marginal* distribution function can be obtained as follows:

$$F_K(x_1, x_2, \ldots, x_K) = F_N(x_1, x_2, \ldots, x_K, \infty, \infty, \ldots, \infty), \quad \text{where } K < N$$

4. Conditional distribution and density functions:

Definition 5.10 Among the N given random variables, the conditional probability distribution function of K r.v.'s (where $K < N$), given $N - K$ remainig r.v.'s is obtained as follows:

$$F_K(x_1, x_2, \ldots, x_K \mid x_{K+1}, \ldots, x_N)$$

$$= \frac{\int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \cdots \int_{-\infty}^{x_K} f_N(\alpha_1, \alpha_2, \ldots, \alpha_K, x_{K+1}, \ldots, x_N) d\alpha_1 d\alpha_2 \cdots d\alpha_K}{f_{N-K}(x_{K+1}, \ldots, x_N)}$$

Definition 5.11 Corresponding conditional probability density function of K r.v.'s (where $K < N$), given $N - K$ remainig r.v.'s among total of N rnadom variables is then obtained as :

$$f_K(x_1, x_2, \ldots, x_K \mid x_{K+1}, \ldots, x_N) = \frac{f_N(x_1, x_2, \ldots\ldots\ldots, x_N)}{f_{N-K}(x_{K+1}, \ldots, x_N)}$$