Contents

Chapter 4

PROBABILITY DISTRIBUTIONS FOR MORE THAN ONE RANDOM VARIABLE

4.1 What are Bivariate Random Variables?

Necessaity: Examples (Bivariate)

1. *A systen w/ random inputs or random components:*

   The joint probability that the input and output are in certain ranges at a specific time instant.

2. *Shooting a projectile at a target:*

   Derscription of the impact point needs two variables, in either Cartesian or polar coordinates.

Remark:
Once the concept of two r.v.'s(i.e. bivariate), the extension to more than two r.v.'s(i.e. multivariate) is relatively easy.....

In summary: We deal with the probability that *a pair of random variables* $\ni$: X *and* Y *are in certain region of the plane.*

4.2 Bivariate CDF

Definition 4.1 Bivariate or joint cdf*:*
Given two r.v.'s X and Y, we define their bivariate, or joint cdf as:

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

Note:
This corresponds to the joint probability that $X$and Y are in the lower lefthand portion of the xy-plane. (See figure4-1(a) below)

Figure 4.1: Definitions pertinent to the joint cdf: (a)defining joint cdf (b) probability of X and Y being in a rectangular region.

Properties the joint cdf: [1]

1. If either one of x or y is minus infinity, the joint cdf is zero:

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

2. If both x and $y$are infinity, the joint cdf is unity:

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

[1] Joint cdf can be defined for both *continuous* and *discrete* bivariate r.v.'s, of which the discrete case will be discussed in Section4-4.

3. If we set one of the variable to be infinity, we get the cdf of the other r.v.:

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

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

We call it the *marginal cdf*.....

4. The joint cdf is a nondecreasing function of its arguments:

$$F_{XY}(x_1,y_1) \geq F_{XY}(x_0,y_0) \quad \text{for} \quad x_1 \geq x_0 \text{ or } y_1 \geq y_0$$

5. The joint cdf is continuous from the right on either variable:

$$\lim_{x\to x_0^+} F_{XY}(x,y) = F_{XY}(x_0,y)$$

$$\lim_{y\to y_0^+} F_{XY}(x,y) = F_{XY}(x,y_0)$$

FACT:
The probability of X and Y lying in a certain range is found from the following relationship:

$$\begin{aligned}
& P(x_1 \leq X \leq x_2, y_1 \leq Y \leq y_2) \\
= \;& P(x_1 \leq X \leq x_2, Y \leq y_2) - P(x_1 \leq X \leq x_2, Y \leq y_1) \\
= \;& P(X \leq x_2, Y \leq y_2) - P(X \leq x_1, Y \leq y_2) \\
& \qquad - P(X \leq x_2, Y \leq y_1) + P(X \leq x_1, Y \leq y_1) \\
= \;& F_{XY}(x_2,y_2) - F_{XY}(x_1,y_2) - F_{XY}(x_2,y_1) + F_{XY}(x_1,y_1)
\end{aligned}$$

Verification: assignment...(via probability axioms and figure4-1(b))

Example 4.1

Consider the function

$$\tilde{F}(x, y) = [1 - \exp(-x)][1 - \exp(-y)]u(x)u(y)$$

(a) Is it suitable for a joint cdf? Explain why.

(b) Find the probability that the r.v.'s lie in the rectangle

$$-10 \leq X \leq 2 \quad 3 \leq Y \leq 5$$

Solution:

(a) Yes, since it meets all required properties to be a joint cdf. (see figure4-2.)

(b) According to the FACT mentioned above, we have:

$$\begin{aligned} & P(-10 \leq X \leq 2, 3 \leq Y \leq 5) \\ = \; & F_{XY}(2, 5) - F_{XY}(-10, 5) - F_{XY}(2, 3) + F_{XY}(-10, 3) \end{aligned}$$

where the value of each term can be calculated as:

$$\begin{aligned} F_{XY}(2, 5) &= (1 - e^{-2})(1 - e^{-5}) = 0.859 \\ F_{XY}(-10, 5) &= 0 \\ F_{XY}(2, 3) &= (1 - e^{-2})(1 - e^{-3}) = 0.822 \\ F_{XY}(-10, 3) &= 0 \end{aligned}$$

Thus the desired probability is:

$$P(-10 \leq X \leq 2 \quad 3 \leq Y \leq 5) = 0.037$$

Figure 4.2: Plot of the bivariate cdf of Example 4-1.

4.3 Bivariate PDF

Definition 4.2 Bivariate or joint pdf*:*
The bivariate, or joint probability density function of two r.v.'s X and Y is defined in terms of their joint cdf as:

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

Note: We assume that the joint cdf is *everywhere differentiable* in this definition.

Example 4.2

Find the joint pdf corresponding to the cdf in Example 4-1.

Solution:

Carrying out the differentiation, we obtain

$$f_{XY}(x, y) = e^{-x}e^{-y}u(x)u(y) = e^{-(x+y)}u(x)u(y)$$

which is plotted in figure4-3.

Figure 4.3: The joint pdf of Example 4-2.

Properties the joint pdf: [2]

1. The joint cdf in terms of the joint pdf can be expressed as follows:

$$F_{XY}(x,y) = \int_{-\infty}^{y}\int_{-\infty}^{x} f_{XY}(u,v)dudv$$

2. The volume under the joint pdf is unity, i.e.:

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

3. The joint pdf is non-negative, i.e.:

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

4. The marginal cdf in terms of the joint pdf is in the following forms:

$$F_X(x) = \int_{-\infty}^{\infty}\int_{-\infty}^{x} f_{XY}(u,v)dudv$$

$$F_Y(y) = \int_{-\infty}^{y}\int_{-\infty}^{\infty} f_{XY}(u,v)dudv$$

5. By differentiating above double integrals w.r.t. x and y respectively, we obtain:

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

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

where these $f_X(x)$ and $f_Y(y)$ are called *the marginal pdf*s.

---

[2] **(Cf.) The Leibnitz Rule:**

Let

$$g(x) = \int_{\alpha(x)}^{\beta(x)} f(x,u)du$$

where $f(x,u)$ is a *continuous* function w.r.t. x and u, then the derivative of $g(x)$ can be expressed in the following form:

$$\frac{dg(x)}{dx} = f\left(x,\beta(x)\right)\frac{d\beta(x)}{dx} - f\left(x,\alpha(x)\right)\frac{d\alpha(x)}{dx} + \int_{\alpha(x)}^{\beta(x)} \frac{\partial}{\partial x} f(x,u)du$$

FACT:
The probability of X and Y lying in a certain range, in terms of the joint pdf, can be found from the following relationship:

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

Verification: assignment

Example 4.3

Consider the function of two variables:

$$f_{XY}(x, y) = \begin{cases} Axy, & 0 < x < y,\ 0 < y < 1 \\ 0, & \text{otherwise} \end{cases}$$

(a) Find A such that this is a proper pdf.

(b) Find the probability that $0 < X < 0.5$ and $0.5 < Y < 1$.

(c) Obtain the marginal pdf's for X and Y

Solution:

(a) Since the volume under the joint pdf is unity, we compute

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{XY}(x, y) dx dy = \int_0^1 \int_0^y Axy dx dy = \frac{A}{2} \int_0^1 y^3 dy = \frac{A}{8} = 1$$

from which we get $A = 8$.

Figure: Integration region to obtain A.

(b) From the FACT mentioned above, the probability can be calculated as:

$$P(0 < X < 0.5,\ 0.5 < Y < 1) = 8\int_{0.5}^{1}\int_{0}^{0.5} xydxdy = 0.375$$

(c) The marginal pdf of X is obtained by integrating the joint pdf over all y as:

$$f_X(x) = \int_x^1 8xydy = 8x\left.\frac{y^2}{2}\right|_x^1 = 4x(1-x^2),\ \ 0 < x < 1$$

and zero elsewhere. (Refer the above figure to deduce the integration limits...)
Likewise, the marginal pdf for Y can be obtained by iontegrating the joint pdf overall x as:

$$f_Y(y) = \int_0^y 8xydx = 8y\left.\frac{x^2}{2}\right|_0^y = 4y^3,\ \ 0 < y < 1$$

and zero elsewhere.

Assignment: Verify that both $f_X(x)$ and $f_Y(y)$ integrate to 1, as they should.

Example 4.4 *Self study*

4.4 Discrete Random Variable Pairs

Suppose pairs of r.v.'s (X, Y) take on the *exhaustive* set of values $\{(x_i, y_j),\ i = 1, 2, \cdots, m,\ j = 1, 2, \cdots, n\}$ with associated probabilities $\{P_{ij},\ i = 1, 2, \cdots, m,\ j = 1, 2, \cdots, n\}$.

$\Longrightarrow$ Since the set of values are *exhaustive*. we have:

$$\sum_{i=1}^{m}\sum_{j=1}^{n} P_{ij} = 1$$

$\Longrightarrow$ The probability mass function consists of point above the xy-plane.

Figure 4.4: The probability mass function for the case of $m = n = 3$.

Example 4.5

Let X and Y be the numbers shown on a pair of fair dice thrown. Find the joint probability mass function of X and Y.

Solution:

Assuming that the dice roll independently to each other, we have the joint probability mass function as:

$$P(X = i,\ Y = j) = P(X = i)P(Y = j) = \frac{1}{6} \times \frac{1}{6} = \frac{1}{36},\ \ i, j = 1, 2, \cdots, 6$$

A plot of this pmf consists of lines $\frac{1}{36}$ high at points $(1, 1), (1, 2), \cdots, (6, 6)$ on the xy-plane.

Corresponding joint cdf can then be found from the double summation as:

$$F_{XY}(x,y) = \begin{cases} \sum_{i\leq x}\sum_{j\leq y}\frac{1}{36}, & i,j\leq 6 \\ 1, & i \text{ and } j > 6 \end{cases}$$

Assignment: Plot this joint cdf as vertical planes parallel to the $x$and y axis...

Example 4.6

Referring to the rolling of a pair of dice, determine the joint probability mass function of two discrete r.v.'s defined as follows:

$$X = \text{sum of numbers shown}, \quad Y = \text{difference of numbers shown}$$

Solution:

Construct a table of sums and differences as in Table 4-1, from which we deduce the set of joint probabilities listed in Table 4-2, which define the joint probability mass function.

Table 4-1: Sum and differences(in parentheses) when a pair of dice are thrown.

Table 4-2: The joint probability mass function of Example 4-6.

The marginal pmf of X and Y may be obtained by summing over the columns and the rows, respectively. The probability mass function of X, for example, is given as [3]:

$$
\begin{aligned}
&P(X=2)=\tfrac{1}{36}, \quad P(X=3)=\tfrac{2}{36}, \quad P(X=4)=\tfrac{3}{36}, \quad P(X=5)=\tfrac{4}{36},\\
&P(X=6)=\tfrac{5}{36}, \quad P(X=7)=\tfrac{6}{36}, \quad P(X=8)=\tfrac{5}{36}, \quad P(X=9)=\tfrac{4}{36}\\
&P(X=10)=\tfrac{3}{36}, \quad P(X=11)=\tfrac{2}{36}, \quad P(X=12)=\tfrac{1}{36},
\end{aligned}
$$

[3]When plotted, this would look similar to Figure 3-5, except that we have lines here in place of impulses.

4.5 Conditional CDF and Conditional PDF

Often, it might be convenient or necessary to *condition* the outcomes of r.v.'s on the occurrence of a hypothesis $\ni$:

1. Rain or no rain
2. Noise alone present or signal+noise present
3. Random variable Z being in a certain region

Recall: The conditional probability (Chapter 2)

Definition 4.3 The Conditional CDF*:*
Given two r.v.'s X and Y, let $A = \{X \le x\}$, and $B = \{Y \in R_y\}$, then the conditional cdf of X given the event $\{Y \in R_y\}$ is denoted and defined as the following conditional probability:

$$F_X(x|Y \in R_y) \triangleq P(X \le x|Y \in R_y) = \frac{P(X \le x,\, Y \in R_y)}{P(Y \in R_y)}$$

Extending the idea, we take the events A and B to be:

$$A = \{x < X \le x + \Delta x\}$$
$$B = \{y < Y \le y + \Delta y\}, \quad |\Delta x|, |\Delta y| \ll 1$$

Then, the conditional probability of A given B becomes:

$$\begin{aligned} P(x < X \le x + \Delta x | y < Y \le y + \Delta y) &= \frac{P(x < X \le x + \Delta x,\, y < Y \le y + \Delta y)}{P(y < Y \le y + \Delta y)} \\ &\simeq \frac{f_{XY}(x,y)\Delta x \Delta y}{f_Y(y)\Delta y} \\ &= f_{X|Y}(x|y)\Delta x \end{aligned}$$

where the approximation comes from the *mean value theorem*.

From this result, we have the following definition of the conditional pdf of X given $Y = y$.

Definition 4.4 The conditional PDF:
The conditional pdf of X given $Y = y$ is defined as follows:

$$f_{X|Y}(x|y) \triangleq \frac{f_{XY}(x,y)}{f_Y(y)}$$

Interpretation: $f_{X|Y}(x|y)\Delta x$ represents the probability that r.v. X is in a Δx interval around x, given that Y is in a small Δy interval around y!!!

Likewise, we define the conditional pdf of Y given $X = x$ to be:

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

Properties of conditional cdf and pdf:
(same as those of regular cdf and pdf...)

(e.g.)

$$\int_{-\infty}^{\infty} f_{X|Y}(x|y)dx = \int_{-\infty}^{\infty} \frac{f_{XY}(x,y)}{f_Y(y)}dx = \frac{f_Y(y)}{f_Y(y)} = 1$$

Example 4.7

Obtain the conditional pdf of X given Y for the joint pdf of Example 4-3.

Solution:

From example 4-3, we already have the joint pdf and the marginal pdf, and thus we obtain:

$$f_{X|Y}(x|y) = \frac{f_{XY}(x,y)}{f_Y(y)} = \begin{cases} \frac{2x}{y^2}, & 0 < x < y,\ 0 < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

Example 4.8

A certain signal may or may not be present in noise, where the voltage at the output of the detector has the following pdf's respectively depending on the presence of the signal.

$$f_{X|S}(x|S) = \frac{e^{-(x-2)^2/8}}{\sqrt{8\pi}}$$

$$f_{X|\bar{S}}(x|\bar{S}) = \frac{e^{-x^2/8}}{\sqrt{8\pi}}$$

A threshold of 1 volt is set at the output of the detector, and the decision that the signal was present is made if the detector output is greater than the threshold.

(a) What is the probability that, if present, the signal is not detected (called the *probability of miss*)?

(b) What is the probability that if the signal was not present, the decision is that it was present (called the *probability of false alarm*)?

Solution:

(a) This corresponds to the probability that the r.v. X does not cross the threshold given the signal was present, which, in terms of the first conditional pdf, is as follows [4]:

$$\begin{aligned} P_{\text{miss}} &= \int_{-\infty}^{1} \frac{e^{-(x-2)^2/8}}{\sqrt{8\pi}} dx = \int_{-\infty}^{-0.5} \frac{e^{-u^2/2}}{\sqrt{2\pi}} du = \int_{0.5}^{\infty} \frac{e^{-u^2/2}}{\sqrt{2\pi}} du \\ &= Q(0.5) = 0.309 \end{aligned}$$

Figure: The pdf's $f_{X|S}(x|S)$ and $f_{X|\bar{S}}(x|\bar{S})$.

[4] A change of variable $u = (x-2)/2$ and the symmetricity of the integrand along with the table of Q-function has been used.

(b) The probability of the false alarm is the probability that noise alone crosses the threshold, which, in terms of the second pdf above, is as follows:

$$P_{\mathrm{FA}} = \int_1^{\infty} \frac{e^{-x^2/8}}{\sqrt{8\pi}} dx = \int_{0.5}^{\infty} \frac{e^{-u^2/2}}{\sqrt{2\pi}} du = Q(0.5) = 0.309$$

Note:

(1) High error probabilities. (Typical values for communication systems are $10^{-2} \sim 10^{-8}$)

(2) $P_{\mathrm{miss}} = P_{\mathrm{FA}}$ since the threshold is set halfway b/w two mean values(i.e. 0 and 2), and the Gaussian pdf is even.

(3) How can we make these error probabilities smaller?

(i) Increase the mean of the signal present pdf(i.e. 2).

(ii) Decrease the variance of both pdf's(i.e. 4), which corresponds to decreasing the noise.

Example 4.9 *Self study*

4.6 Statistically Indenpendent Random Variables

Recall: Two events A and B are called to be statistically independent if:

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

Applying directly to two r.v.'s via definition of the joint cdf, we have:

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

From the definition of the joint pdf, and by differentiating the above w.r.t. x and y, we have:

$$f_{XY}(x, y) = f_X(x)f_Y(y)$$

Or, from the definition of the conditional pdf, it follows:

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

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

Summary: Any one of these statements defines statistically independent random variables; all are equivalent.

Example 4.10

The r.v.'s discussed in example 4-1 and 4-2 are statistically independent, since their joint cdf and pdf's factor, i.e.

$$\begin{aligned} F_{XY}(x,y) &= [1-\exp(-x)][1-\exp(-y)]u(x)u(y) \\ &= [1-\exp(-x)]u(x)[1-\exp(-y)]u(y) \\ &= F_X(x)F_Y(y) \end{aligned}$$

and

$$f_{XY}(x,y) = e^{-(x+y)}u(x)u(y) = e^{-x}u(x)e^{-y}u(y) = f_X(x)f_Y(y)$$

Example 4.11

The r.v.'s X and Y in example 4-3 are NOT statistically independent, since the joint pdf is:

$$f_{XY}(x,y) = \begin{cases} 8xy, & 0<x<y,\ 0<y<1 \\ 0, & \text{otherwise} \end{cases}$$

whereas the marginal pdf's for X and Y are respectively as:

$$f_X(x) = 4x(1-x^2),\ \ 0<x<1$$

$$f_Y(y) = 4y^3,\ \ 0<y<1$$

and therefore:

$$f_{XY}(x,y) \neq f_X(x)f_Y(y)$$

4.7 Averages of Functions of Two Random Variables

Recall: Given a function of two r.v.'s. e.g. $g(X, Y)$, it is also a r.v..

Extending the definition of the expectation for functions of a single r.v., we have:

1. Continuous r.v.:

$$E[g(X,Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y) f_{XY}(x,y)dxdy$$

where $f_{XY}(x, y)$ is the joint pdf of X and Y.

2. Discrete r.v.:

$$E[g(X,Y)] = \sum_{i=1}^{m}\sum_{j=1}^{n} g(x_i, y_j) P(X = x_i, Y = y_j)$$

where $P(X = x_i, Y = y_j)$ is the joint pmf of X and Y.

Properties:

1. Suppose $g(X, Y) = g_1(X)g_2(Y)$, and X and Y are statistically indenpendent, then:

$$\begin{aligned} E[g(X,Y)] = E[g_1(X)g_2(Y)] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g_1(x)g_2(y)dxdy \\ &= \int_{-\infty}^{\infty} g_1(x)f_X(x)dx \int_{-\infty}^{\infty} g_2(y)f_Y(y)dy \\ &= E[g_1(X)]E[g_2(Y)] \end{aligned}$$

2. If $g(X,Y) = a_1 g_1(X,Y) + a_2 g_2(X,Y)$ where a_1 and a_2 are constants, then:

$$\begin{aligned}
& E[a_1 g_1(X,Y) + a_2 g_2(X,Y)] \\
= & \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} [a_1 g_1(x,y) + a_2 g_2(x,y)] f_{XY}(x,y) dx dy \\
= & \int_{-\infty}^{\infty} a_1 g_1(x,y) f_{XY}(x,y) dx dy + \int_{-\infty}^{\infty} a_2 g_2(x,y) f_{XY}(x,y) dx dy \\
= & a_1 E[g_1(X,Y)] + a_2 E[g_2(X,Y)]
\end{aligned}$$

This is called the *linearity property*, and it holds whether or not X and Y are statistically independent.

Example 4.12

Given the joint pdf:

$$f_{XY}(x,y) = \frac{\alpha\beta}{4} \exp(-\alpha|x| - \beta|y|)$$

(a) Find the expectation of $g(X,Y) = XY$.
(b) Obtain the expectation of $h(X,Y) = X^2 + Y^2$.

Solution:

(a) Since the joint pdf can be expressed as $f_{XY}(x,y) = f_X(*x) f_Y(y)$, X and Y are statistically indenpendent, and thus:

$$E(XY) = E(X)E(Y) = 0$$

Note that $E[X] = E[Y] = 0$ since each pdf is an even function.

(b) From the linearity property above, we have:

$$E(X^2 + Y^2) = E(X^2) + E(Y^2)$$

where $E[X^2]$ can be calculated as:

$$\begin{aligned}
E(X^2) &= \int_{-\infty}^{\infty} x^2 \frac{\alpha}{2} \exp(-\alpha|x|) dx \\
&= \alpha \int_0^{\infty} x^2 \exp(-\alpha x) dx = \frac{1}{\alpha^2} \int_0^{\infty} u^2 \exp(-u) du = \frac{2}{\alpha^2}
\end{aligned}$$

Similarly, we get $E(Y^2) = \frac{2}{\beta^2}$, and therefore we have:

$$E(X^2 + Y^2) = \frac{2}{\alpha^2} + \frac{2}{\beta^2}$$

4.8 Some Special Averages of Functions of Two Random Variables

Joint Moments

The mn-th joint moments of two r.v.'s X and Y are defined as:

$$m_{mn} = E[X^m Y^n], \quad m, n = 1, 2, \cdots$$

(cf)

1. Special cases are means of $X$and Y, obtained by setting $m = 1$, $n = 0$ and $m = 0$, $n = 1$ respectively.
2. If X and Y are independent, then $m_{mn} = E[X^m]E[Y^n]$, $\forall m, n$

Example 4.13

Find the joint moments of the r.v.'s whose joint pmf ius shown in fig. 4-4, if

$$x_1 = 1,\ x_2 = 2,\ x_3 = 3,\ y_1 = 3,\ y_2 = 3,\ y_3 = 4$$

Solution:

Inserting the given values into the definition of the joint moments, we have:

$$E(X^m Y^n) = 1^m 3^n \times 0.2 + 2^m 3^n \times 0.6 + 3^m 4^n \times 0.2$$

and several special cases are provided in Table 4-3.

Table 4.3 Joint moments for X and Y whose joint pmf is shown in figure 4-4.

Joint Central Moments

The mn-th joint central moments of two r.v.'s X and Y are defined as:

$$\mu_{mn} = E[(X - \mu_X)^m (Y - \mu_Y)^n], \quad m, n = 1, 2, \cdots$$

where μ_X and μ_Y are the means of X and Y respectively. The special cases of $m = 2$, $n = 0$ and $m = 0$, $n = 2$ give the variances σ_X^2 and σ_Y^2 of X and Y, respectively.

Covariance

This is a special case of the joint central moments with $m = n = 1$, i.e.:

$$C_{XY} = E[(X - \mu_X)(Y - \mu_Y)]$$

By expanding the expectation, thsi can be put into another form as:

$$C_{XY} = E[XY] - \mu_X \mu_Y = R_{XY} - \mu_X \mu_Y$$

where R_{XY} is called the *correlation* of X and Y, which corresponds to a special case of the joint moments for the case of $m = n = 1$.

(cf) Two r.v.'s X and Y are called *uncorrelated* if:

$$C_{XY} = 0$$

or equivalently if:

$$R_{XY} = E[X] \cdot E[Y]$$

Correlation Coefficient

The correlation coefficient is defined as:

$$\rho_{XY} = \frac{C_{XY}}{\sigma_X \sigma_Y}$$

where C_{XY} denores the covariance of X and Y.

properties of ρ_{XY}

1. Its absolute value is bounded by unity:

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

proof: Consider the following non-negative quantity:

$$E\left[\left(\frac{X-\mu_X}{\sigma_X} \pm \frac{Y-\mu_Y}{\sigma_Y}\right)^2\right] \geq 0$$

By expanding the square, and taking the expectation term by term, we get:

$$1 \pm 2\rho_{XY} + 1 \geq 0$$

which is equivalent to the above statement.

Assignment: detailed derivation

2. If X and Y are statistically independent, then:

$$\rho_{XY} = 0$$

proof: We show $C_{XY} = 0$, from which it follows that $\rho_{XY} = 0$. For statistically independent r.v.'s X and Y, the covariance can be shown to be zero as:

$$\begin{aligned} C_{XY} = E[(X-\mu_X)(Y-\mu_Y)] &= [E(X)-\mu_X][E(Y)-\mu_Y] \\ &= (\mu_X-\mu_X)(\mu_Y-\mu_Y) \\ &= 0 \end{aligned}$$

(cf)

(a) Two r.v.'s X and Y are *uncorrelated* if $\rho_{XY} = 0$.

(b) If two r.v.'s are *statistically independent*, then they are *uncorrelated*, but NOT vice versa. [5]

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

3. If $Y = aX + b$, where a and b are constants, then:

$$\rho_{XY} = 0$$

proof: We can show that

$$\mu_Y = a\mu_X + b \qquad \sigma_Y = \pm a\sigma_X$$

and thus the covariance can be expressed as:

$$C_{XY} = E\{(X - \mu_X)[aX + b - (a\mu_X + b)]\} = aE[(X - \mu_X)^2] = a\sigma_X^2$$

Substituting these into the definition of the correlation coefficient, we have:

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

Example 4.14

Find the correlation, covariance, and correlation coefficient for the r.v.'s w/ joint pdf given in example 4-3, i.e.

$$f_{XY}(x, y) = \begin{cases} 8xy, & 0 < x < y,\ 0 < y < 1 \\ 0, & \text{otherwise} \end{cases}$$

[5] The only exceptional case is when X and Y are jointly Gaussian. In this case, *uncorrelatedness* implies the *statistical independence*.

Solution:

By definition, the correlation is given as:

$$R_{XY} = E[XY] = \int_0^1 \int_0^y xy(8xy)dxdy = \frac{4}{9}$$

and the means of X and Y are respectively

$$\mu_X = \int_0^1 x[4x(1-x^2)]dx = \frac{8}{15} \quad \text{and} \quad \mu_Y = \int_0^1 y(4y^3)dy = \frac{4}{5}$$

where the marginal pdf's obtained in example 4-3 were used.

The covariance is then:

$$C_{XY} = \frac{4}{9} - \frac{8}{15}\frac{4}{5} = \frac{4}{225}$$

To get correlation coefficient, we first calculate the mean square values as:

$$E[X^2] = \int_0^1 x^2[4x(1-x^2)]dx = \frac{1}{3} \quad \text{and} \quad E[Y^2] = \int_0^1 y^2(4y^3)dy = \frac{2}{3}$$

and the variances are

$$\sigma_X^2 = \frac{1}{3} - \left(\frac{8}{15}\right)^2 = \frac{11}{225} \quad \text{and} \quad \sigma_Y^2 = \frac{2}{3} - \left(\frac{4}{5}\right)^2 = \frac{2}{75}$$

from which the correlation coefficient is given by:

$$\rho_{XY} = \frac{4/225}{(\sqrt{11}/15)(\sqrt{2}/5\sqrt{3})} = 2\sqrt{2/33} = 0.4924$$

Joint Characteristic Function

Generalizing the characteristic functionn of a single r.v. to the bivariate case, the joint characteristic function of two r.v.'s X and Y is defined as:

$$M_{XY}(ju, jv) = E[e^{j(uX+vY)}] = E(e^{juX}e^{jvY})$$

Special cases:

1. If X and Y are statistically independent, the joint characteristic function is the product of their individual characteristic functions, i.e.:

$$M_{XY}(ju, jv)\ E(e^{juX})E(e^{jvY}) = M_X(ju)M_Y(jv)$$

2. Let X and Y be two independent r.v.'s, and defined a new r.v. Z as the sum, i.e.

$$Z = X + Y$$

Then, the characteristic function of the newly defined r.v. Z is also the product of each individual characteristic functions as:

$$\begin{aligned} M_Z(jv) &= E[e^{jv(X+Y)}] = E(e^{jvX}e^{jvY}) \\ &= E(e^{jvX})E(e^{jvY}) = M_X(jv)M_Y(jv) \end{aligned}$$

(cf) Notice the similarity and difference in the form of the characteristic function b/w above two cases.....

Example 4.15

Find the joint characteristic function for the pdf given in example 4-2:

$$f_{XY}(x, y) = e^{-x}e^{-y}u(x)u(y) = e^{-(x+y)}u(x)u(y)$$

Solution:

Since we showed that X and Y are independent in example 4-10, the joint charateristic function can be expressed as the product of marginal characteristic function, where the c.f. of X is:

$$M_X(ju) = \int_{-\infty}^{\infty} e^{jux}e^{-x}u(x)dx = \int_0^{\infty} e^{-(1-ju)x}dx = -\left.\frac{e^{-(1-ju)x}}{1-ju}\right|_0^{\infty} = \frac{1}{1-ju}$$

Similary, we get the c.f. of Y as:

$$M_Y(jv) = \frac{1}{1-jv}$$

Therefore, the joint characteristic function is:

$$M_{XY}(ju, jv) = \frac{1}{(1-ju)(1-jv)}$$

4.9 Joint Gaussian PDF

Two r.v.'s X and Y are called *jointly Gaussian* if their joint pdf is in the following form:

$$f_{XY}(x,y) = \frac{\exp\left[-\frac{(x-\mu_X)^2/\sigma_X^2 - 2r(x-\mu_X)(y-\mu_Y)/\sigma_X\sigma_Y + (y-\mu_Y)^2/\sigma_Y^2}{2(1-r^2)}\right]}{2\pi\sigma_X\sigma_Y\sqrt{1-r^2}}$$

where various parameters in the above joint pdf can be shown to be:

$$\mu_X = E(X) = \text{mean of } X$$

$$\mu_Y = E(Y) = \text{mean of } Y$$

$$\sigma_X^2 = E[(X-\mu_X)^2] = \text{variance of } X$$

$$\sigma_Y^2 = E[(Y-\mu_Y)^2] = \text{variance of } Y$$

$$r = \frac{E[(X-\mu_X)(Y-\mu_Y)]}{\sigma_X\sigma_Y} = \text{correlation coefficient of } X \text{ and } Y$$

proof: assignment

$\Longrightarrow$ $f_{XY}(x,y)$ is a bell-shaped volume above xy-plane centered at $x = \mu_X$ and $y = \mu_Y$ (Figure 4-7)

$\Longrightarrow$ Cuts through $f_{XY}(x,y)$ parallel to xy-plane are ellipses, whose fatness depends on σ_X and σ_Y (Figure 4-6)

Consider a special case via change of variables as:

$$x_n = \frac{x - \mu_X}{\sigma_X} \quad \text{and} \quad y_n = \frac{y - \mu_Y}{\sigma_Y}$$

Then the joint pdf becomes [6]:

$$f_{XY}(x, y) = \frac{\exp[-(x^2 - 2rxy + y^2)/2(1 - r^2)]}{2\pi\sqrt{1 - r^2}}$$

which corresponds to a joint Gaussian pdf for which the r.v.'s have *zero means* and *unit variances*, i.e. normalized bivariate Gaussian pdf.

The lines of constant *probability density*(i.e. family of ellipses) can then be expressed as follows [7]:

$$x^2 - 2rxy + y^2 = K$$

Figure 4.5: Equal probability density contours for a joint Gaussian pdf w/ $\mu_X = \mu_Y = 0$, and $\sigma_X = \sigma_Y = 1$.

$\Longrightarrow$ To get an idea of the effect of the standard deviations of X and Y, consider the contours shown below where $\sigma_X = 1.25$ and $\sigma_Y = 1$.

Figure 4.6: Equal probability density contours for a joint Gaussian pdf w/ $\mu_X = \mu_Y = 0$, and $\sigma_X = 1.25$ & $\sigma_Y = 1$.

[6] We drop the subscripts for notational convenience.

[7] These family of ellipses have their major and minor axes NOT parallel to x and y coordinates. To make them parallel, the x and y axes must be rotated by $\theta = \pm\frac{\pi}{4}$. Read textbook.....

Figure 4.7: Three dimensional representation of normalized bivariate Gaussian pdf's.

Now, consider the case when $r = 0$, i.e. X and Y are *uncorrelated.*

Then, the joint pdf become:

$$\begin{aligned} f_{XY}(x,y) &= \frac{\exp[-(x-\mu_X)^2/2\sigma_X^2 - (y-\mu_Y)^2/2\sigma_Y^2]}{2\pi\sigma_X\sigma_Y} \\ &= \frac{\exp[-(x-\mu_X)^2/2\sigma_X^2]\exp[-(y-\mu_Y)^2/2\sigma_Y^2]}{\sqrt{2\pi\sigma_X^2}\sqrt{2\pi\sigma_Y^2}} \\ &= f_X(x)\cdot f_Y(y) \end{aligned}$$

which means that $X$and Y are *statistically independent.*

conclusion: *Uncorrelated* Gaussian random variables are also *statistically independent*, which is NOT generally true.

Example 4.16

Measurements on a random voltage at two time instants(denoted as X and Y respectively) showed that their means are zero and the variances are 4 watts. Assuming X and Y are uncorrelated Gaussian r.v.'s, write down the joint pdf.

Solution:

We are given that:

$$\mu_X = \mu_Y = 0$$
$$\sigma_X^2 = \sigma_Y^2 = 4$$

and since X and Y are uncorrelated Gaussian r.v.'s, they are also statistically independent, and therefore the joint pdf becomes:

$$f_{XY}(x,y) = \frac{e^{-(x^2+y^2)/2(4)}}{2\pi(4)} = \frac{e^{-(x^2+y^2)/8}}{8\pi}$$

4.10 Functions of Two Random Variables

Consider a function of two r.v.'s, say

$$Z = g(X, Y)$$

$\Longrightarrow$ Given the joint pdf $f_{XY}(x, y)$, we wish to find the pdf of Z

$\Longrightarrow$ We first find the dcf of Z as:

$$F_Z(z) = P(Z \leq z) = P[g(X, Y) \leq z]$$

$\Longrightarrow$ Differentiate the dcf to find the pdf of Z

Remark: The condition $g(X, Y) \leq z$ defines some region in xy-plane, and we have to integrate the joint pdf $f_{XY}(x, y)$ of X and Y over this region.

Example 4.17

Let the transformation be:

$$Z = \sqrt{X^2 + Y^2}$$

Find the pdf of Z assuming that X and Y are independent Gaussian r.v.'s w/ zero means and variances σ^2.

Solution:

We evaluate

$$P(\sqrt{X^2+Y^2} \le z) = F_Z(z) = \int_{R(z)} \int \frac{e^{-(x^2+y^2)/2\sigma^2}}{2\pi\sigma^2} dxdy$$

where the region $R(z)$ is the area inside the circle(figure 4-8):

$$x^2 + y^2 = z^2$$

We change the cartesian coordinate to polar coordinate via the change of variables as follows:

$$x = r\cos\theta \quad \text{and} \quad y = r\sin\theta$$

or,

$$r = \sqrt{x^2+y^2} \quad \text{and} \quad \theta = \tan^{-1}\frac{y}{x}$$

where the differential area $dxdy$ transforms to $rdrd\theta$, and the region $R(z)$ becomes as the region corresponding to $0 \le r \le z$ and $0 \le \theta \le 2\pi$.

Thus the cdf $F_Z(z)$ becomes:

$$\begin{aligned} F_Z(z) &= \int_{\theta=0}^{2\pi} \int_0^z \frac{e^{-r^2/2\sigma^2}}{2\pi\sigma^2} rdrd\theta = \int_0^{2\pi} \left[\frac{e^{-r^2/2\sigma^2}}{2\pi}\right]_0^z d\theta \\ &= \int_0^{2\pi} (1 - e^{-z^2/2\sigma^2}) \frac{d\theta}{2\pi} = 1 - e^{-z^2/2\sigma^2}, \quad z \ge 0 \end{aligned}$$

This is called a *Rayleigh* pdf [8].

Figure 4.8: Area of integration.

Figure 4.9: Rayleigh pdf w/ $\sigma = 1$.

[8] The mean is $E(Z) = (\pi/2)^{1/2}\sigma$, while the variance is $\sigma_Z^2 = (2-\pi/2)\sigma^2$: derivation (assignment).

Example 4.18

Given the transformation:

$$Z = X + Y$$

Find the pdf of Z in terms of f_X and f_Y, where X and Y are independent r.v.'s.

Solution:

We first evaluate the cdf $F_Z(z)$ as follows:

$$F_Z(z) = P(X + Y \leq z) = \int_{-\infty}^{\infty} \int_{-\infty}^{z-y} f_{XY}(u, v) du dv$$

where the area of integration is shown in figure 4-10.

Differentiating and using statistical independence, we obtain [9] (*assignment*)

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

Figure 4.10: Integration region for the pdf of sum of two independent r.v.'s.

Remark: The pdf of the sum of two *statistically independent* r.v's is the convolution [10] of their separate pdf's!!!

Suppose both X and Y are uniformly distributed as $U[-0.5, 0.5]$, then the pdf of Z is as follows: (*assignment*)

$$f_Z(z) = \begin{cases} 0, & z < -1 \\ 1 + z, & -1 \leq z \leq 0 \\ 1 - z, & 0 < z \leq 1 \\ 0, & z > 1 \end{cases}$$

Example 4.19 *Self study*

[9] You may need to use the Leibnitz rule here...

[10] Recall that the characteristic function of Z is the product of individual charateristic functions, i.e. $M_Z(jv) = M_X(jv) M_Y(jv)$. Can you explain it from the viewpoints of the above result...?

4.11 Transformation of a Pair of Random Variables

Suppose we are given a pair of r.v.'s X and Y, whose joint pdf is known. We define another pair of r.v.'s U and V via the following transformations:

$$U = g(X, Y)$$
$$V = h(X, Y)$$

$\Longrightarrow$ The probability of (X, Y) being in some region R in xy-plane is:

$$P[(X, Y) \in R] = \int_R \int f_{XY}(x, y) dxdy$$

$\Longrightarrow$ Assuming (1) transformation is *one-to-one*, and the region R in xy-plane transforms into some region R^* in uv-plane; (2) inverse transformations $x = g^{-1}(u, v)$ and $y = h^{-1}(u, v)$ have continuous *1st partial derivatives*, we have [11]:

$$\begin{aligned} \int_R \int f_{XY}(x, y) dxdy &= \int_{R^*} \int f_{XY}[g^{-1}(u, v), h^{-1}(u, v)] \left| \frac{\partial(x, y)}{\partial(u, v)} \right| dudv \\ &\triangleq \int_{R^*} \int f_{UV}(u, v) dudv \end{aligned}$$

where

$$J = \frac{\partial(x, y)}{\partial(u, v)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} \frac{\partial g^{-1}(u,v)}{\partial u} & \frac{\partial g^{-1}(u,v)}{\partial v} \\ \frac{\partial h^{-1}(u,v)}{\partial u} & \frac{\partial h^{-1}(u,v)}{\partial v} \end{vmatrix}$$

is called the *Jacobian*.

Remark: Therefore, the pdf of U and V in terms of the pdf of X and Y is as follows:

$$f_{UV}(u, v) = f_{XY}[g^{-1}(u, v), h^{-1}(u, v)] \left| \frac{\partial(x, y)}{\partial(u, v)} \right|$$

[11] From the calculus of transformation of variables in double integrals (Kreyzig, 1988)

Example 4.20

Suppose two Gaussian r.v.'s X and Y have zero mean, variance of 2, and correlation coefficient 0.7. Determine the joint pdf of U and V defined as follows:

$$U = X - Y \quad \text{and} \quad Y = X + 2Y$$

Solution:

We are given that the joint pdf of X and Y is:

$$\begin{aligned} f_{XY}(x,y) &= \frac{\exp\left\{-\frac{(x^2/2)-[2(0.7)xy/2]+(y^2/2)}{2(1-(0.7)^2)}\right\}}{2\pi(2)\sqrt{1-(0.7)^2}} \\ &= 0.111e^{-(0.495x^2-0.686xy+0.495y^2)} \end{aligned}$$

Solving the transformation w.r.t. x and y, we obtain:

$$x = \frac{2}{3}u + \frac{1}{3}v \quad \text{and} \quad y = -\frac{1}{3}u + \frac{1}{3}v$$

The Jacobian is

$$J\begin{pmatrix} x & y \\ u & v \end{pmatrix} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} \frac{2}{3} & \frac{1}{3} \\ -\frac{1}{3} & \frac{1}{3} \end{vmatrix} = \frac{1}{3}$$

Therefore, the joint pdf of U and V is as follows:

$$\begin{aligned} f_{UV}(u,v) &= \frac{0.111}{3}\exp\{-[0.495(0.667u+0.333v)^2 - 0.686(0.667u+0.333v) \\ &\quad (-0.333u+0.333v) + 0.495(-0.333u+0.333v)^2]\} \\ &= 0.037\exp\{-(0.423u^2+0.333uv+0.0325v^2)\} \end{aligned}$$

Example 4.21

Given two indpendent, zero mean Gaussian r.v.'s X and Y w/ variances σ^2, find the joint pdf of the following r.v.'s represented as polar coordinates:

$$R = \sqrt{X^2 + Y^2} \quad \text{and} \quad \Theta = \tan^{-1} \frac{Y}{X}$$

Solution:

The transformation from polar to cartesian coordinates is:

$$x = r\cos\theta \quad \text{and} \quad y = r\sin\theta, \;\; r \geq 0, \;\; 0 \leq \theta < 2\pi$$

The Jacobian is then:

$$J = \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} = r$$

Using the known values of $\mu_X = \mu_Y = 0$ and $\sigma_X^2 = \sigma_Y^2 = \sigma^2$, along with the Jacobian J, the joint pdf of R and Θ can be expressed as follows:

$$f_{R\Theta}(r, \theta) = \frac{r}{2\pi\sigma^2} e^{-r^2/2\sigma^2}, \quad r \geq 0, \; 0 \leq \theta < 2\pi$$

To obtain the marginal pdf's, we integrate the above joint pdf over θ and r respectively:

$$f_R(r) = \frac{r}{\sigma^2} e^{-r^2/2\sigma^2}, \; r \geq 0 \quad \text{and} \quad f_\Theta(\theta) = \frac{1}{2\pi}, \; 0 \leq \theta < 2\pi$$

Note:

(a) R is the Rayleigh r.v., whereas Θ is the uniform r.v..

(b) R and Θ are statistically independent, since $f_{R\Theta}(r, \theta) = f_R(r) \cdot f_\Theta(\theta)$.

4.12 Sum of Independent Random Variables: The Central Limit Theorem

Recall: The pdf of the sum of two independent r.v.'s is the *convolution* of their respective pdf's...

$\Longrightarrow$ Carrying this on to four r.v.'s by defining

$$Z_1 = X_1 + X_2 \quad \text{and} \quad Z_2 = X_3 + X_4$$

$\Longrightarrow$ Then, $f_{Z_1} = f_{X_1} * f_{X_2}$ and $f_{Z_2} = f_{X_3} * f_{X_4}$

$\Longrightarrow$ The pdf of four independent r.v.'s can be found by convolving the f_{Z_1} and f_{Z_2}:

$$Z = Z_1 + Z_2 = X_1 + X_2 + X_3 + X_4$$

$$f_Z = f_{Z_1} * f_{Z_2} = f_{X_1} * f_{X_2} * f_{X_3} * f_{X_4}$$

$\Longrightarrow$ Can be continued to more than four r.v.'s...

Note: The pdf of the sum resembles a Gaussian pdf relatively closely!!!!!

Example 4.22

Suppose all four independent r.v.'s have the same pdf as follows:

$$f_{X_i}(x_i) = \begin{cases} 1, & |x_i| \le \frac{1}{2} \\ 0, & \text{otherwise} \end{cases}$$

Find the pdf of their sum.

Solution:

The convolution by pairs gives the pdf's:

$$f_{Z_i}(z_i) = \begin{cases} 1 - |z_i|, & |z_i| \leq 1 \\ 0, & \text{otherwise} \end{cases} \qquad i = 1, 2$$

Convolution of Z_1 and Z_2 for $z \geq 0$ gives [12]:

$$f_Z(z) = \begin{cases} (1-z) - \frac{1}{3}(1-z)^3 + \frac{1}{6}z^3, & 0 \leq z \leq 1 \\ \frac{1}{6}(2-z)^3, & 1 \leq z \leq 2 \end{cases}$$

(a) The mean and variance of Z are 0 and $\frac{4}{12} = \frac{1}{3}$. (*assignment*)

(b) Camparison of $f_Z(z)$ to the pdf of a Gaussian r.v. w/ mean 0, and variance $\frac{1}{3}$. (figure 4-11)

Figure 4.11: Camparison b/w pdf's of the sum of four independent r.v's and the Gaussian r.v. w/ same mean and variance.

[12] Since the pdf is even, we can use $f_Z(-z) = f_Z(z)$ for $z < 0$.

Central Limit Theorem

Let $X_1, X_2, X_3, \cdots, X_N$ be independent r.v.'s with means $m_1, m_2, m_3, \cdots, m_N$ and standard deviations $\sigma_1, \sigma_2, \sigma_3, \cdots, \sigma_N$ respectively. Then

$$Z = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \frac{X_i - m_i}{\sigma_i}$$

approaches a Gaussian r.v. w/ zero mean and unit variance as N becomes large, provided that:

$$\lim_{N\to\infty} \frac{\sigma_i}{\sigma} = 0$$

Remark:

1. The above condition ensures that no one r.v. in the sum dominates.
2. The pdf of the component r.v. need not be of any specific type.
3. The pdf of the component r.v. need not all be identical.

Example 4.23 *Self study*

4.13 Weak Law of Large Numbers

Weak law of large numbers:

The probability of the average of n independent values of a r.v. X differing from their means $\mu_X = E[X]$ by more than an arbitrary $\epsilon > 0$ goes to zero as $n \to \infty$ [13], i.e.

$$\lim_{n\to\infty} P\left[\left|\frac{1}{n}\sum_{i=1}^{n} X_i - E(X)\right| \geq \epsilon\right] = 0, \quad \epsilon > 0$$

proof:

Consider the Chebyshev's inequality, which is rewritten below [14]:

$$P(|X - \mu_X| \geq \epsilon) \leq \frac{\sigma_X^2}{\epsilon^2}$$

Applying the inequality to the average S_n of n independent r.v.'s:

$$S_n = (1/n)\sum_{i=1}^{n} X_i$$

we obtain:

$$P(|S_n - E(S_n)| \geq \epsilon) \leq \frac{\sigma_{S_n}^2}{\epsilon^2}$$

<u>*Assignment:*</u>

1. The mean of S_n is $E(S_n) = \mu_X$
2. The variance of S_n is σ_X^2

Therefore, the above Chebyshev's inequality becomes:

$$P(|S_n - \mu_X| \geq \epsilon) \leq \frac{\sigma_X^2}{n\epsilon^2}$$

As $n \to \infty$ w/ σ_X and ϵ fixed, the RHS goes to zero.....

q.e.d.

Example 4.24 *Self study*

[13] This means that the sample mean approaches to the actual(or mathematical) average, as the number of samples becomes large...

[14] We obtain this Chebyshev's inequality by letting $k\sigma_X = \epsilon$ in the original inequality.

4.14 Extension to More Than Two Random Variables

Remark: All the definitions given in this chapter regarding joint r.v.'s can be extended directly from two to $N > 2$ r.v.'s.

Let $X_1, X_2, X_3, \cdots, X_N$ be N r.v.'s, then:

(1) The joint cdf:

$$F_N(x_1, x_2, \cdots, x_N) = P(X_1 \leq x_1, X_2 \leq x_2, \cdots, X_N \leq x_N)$$

(2) The joint pdf:

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

(3) The marginal pdf:

$$f_{N-1}(x_1, \cdots, x_{m-1}, x_{m+1}, \cdots, x_N) = \int_{x_m=-\infty}^{\infty} f_N(x_1, x_2, \cdots, x_N) dx_m$$

(4) The conditional pdf:

$$f_N(x_1, x_2, \cdots, x_j | x_{j+1}, \cdots, x_N) = \frac{f_N(x_1, x_2, \cdots, x_N)}{f_{N-j}(x_{j+1}, \cdots, x_N)}$$

(5) The mathematical expectation of a function of r.v.'s:

$$\begin{aligned} & E[g(X_1, X_2, \cdots, X_N)] \\ = \; & \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(X_1, X_2, \cdots, X_N) f_N(x_1, x_2, \cdots, x_N) dx_1 dx_2 \cdots dx_N \end{aligned}$$

FACT:

It is generally difficult to find explicit expressions for joint cdf's and pdf's, except for one case ehich is the N fold jointly Gaussian pdf below:

$$f_{\mathbf{X}}(\mathbf{x}) = |\mathbf{C}|^{-1/2}(2\pi)^{-N/2}\exp\left[\frac{1}{2}(\mathbf{x}-\mathbf{m})^t\mathbf{C}^{-1}(\mathbf{x}-\mathbf{m})\right]$$

where the subscript t denotes the transpose, and $\mathbf{x}$ and $\mathbf{m}$ are column matrices defined as:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} \quad \text{and} \quad \mathbf{m} = \begin{bmatrix} m_1 \\ m_2 \\ \vdots \\ m_N \end{bmatrix}$$

The matrix $\mathbf{C}$ is the covariance matrix w/ elements of:

$$C_{ij} = E[(X_i - m_i)(X_j - m_j)], \quad i, j = 1, 2, \cdots, N$$