Class Exercise 3 – Random Vectors, Conditional Expectation

Introduction: Functions of Random Vectors

For two random variables:

Definition of the problem:

Given two random variables X and Y with PDF $f_{XY}(x,y)$.

Let us denote: $$V = h(X,Y) \\ W = g(X,Y)$$

What is the PDF of V and W?

The solution (assuming a countable amount of solutions to the above system of equations):

For every pair of values (V ,W), let us solve the system of equations:

$$v = h(x,y)$$
$$w = g(x,y)$$

and denote the solutions:

$$\{(x_i, y_i)\}_{i=1}^{n}$$

Let us define the Jacobian of the transformation:

$$\frac{\partial(v,w)}{\partial(x,y)} = \det\begin{bmatrix} \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \\ \frac{\partial w}{\partial x} & \frac{\partial w}{\partial y} \end{bmatrix} \qquad ; \qquad \frac{\partial(x,y)}{\partial(v,w)} = \det\begin{bmatrix} \frac{\partial x}{\partial v} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial v} & \frac{\partial y}{\partial w} \end{bmatrix}$$

The joint PDF of V and W is defined as:

$$f_{VW}(v,w) \underset{(1)}{=} \sum_{i=1}^{n} f_{XY}(x_i, y_i) \left|\frac{\partial(x,y)}{\partial(v,w)}\right|_{x=x_i, y=y_i} \underset{(2)}{=} \sum_{i=1}^{n} \frac{f_{XY}(x_i, y_i)}{\left|\frac{\partial(v,w)}{\partial(x,y)}\right|_{x=x_i, y=y_i}}$$

[!] Notice that in method (1), the Jacobian is a function of (v,w), whereas in method (2), the Jacobian is a function of (x,y), and we place in it the solutions of the equation (which are functions of (v,w)).

For a random vector of any length:

Given the random vector $\underline{X}$ of length n, let us create the random vector $\underline{Y}$ as follows:

$$\begin{cases} Y_1 = g_1(X_1, X_2, ..., X_n) \\ Y_2 = g_2(X_1, X_2, ..., X_n) \\ \vdots \\ Y_n = g_n(X_1, X_2, ..., X_n) \end{cases} \Leftrightarrow \quad \underline{Y} = [Y_1 ... Y_n]^T = \underline{g}(\underline{X})^T$$

Thus, the joint PDF of vector $\underline{Y}$ satisfies:

$$f_{\underline{Y}}(\underline{y}) = \sum_i f_{\underline{X}}(\underline{x}_i) \left| \frac{\partial(x_1, x_2, ..., x_n)}{\partial(y_1, y_2, ..., y_n)} \right|_{\underline{x}=\underline{x}_i} = \sum_i \frac{f_{\underline{X}}(\underline{x}_i)}{\left| \frac{\partial(y_1, y_2, ..., y_n)}{\partial(x_1, x_2, ..., x_n)} \right|_{\underline{x}=\underline{x}_i}}$$

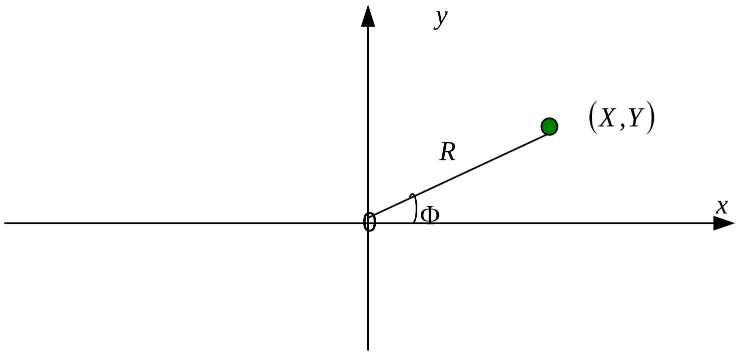
where $\underline{x}_i$ are solutions of the system of equations $\underline{y} = \underline{g}(\underline{x})$ (a countable number of solutions is required)

$$\frac{\partial(y_1, y_2, ..., y_n)}{\partial(x_1, x_2, ..., x_n)} = \begin{vmatrix} \frac{\partial g_1(\underline{x})}{\partial x_1} \cdots \frac{\partial g_1(\underline{x})}{\partial x_n} \\ \vdots \qquad\quad \vdots \\ \frac{\partial g_n(\underline{x})}{\partial x_1} \cdots \frac{\partial g_n(\underline{x})}{\partial x_n} \end{vmatrix}$$

Problem 1:

Consider the real and independent random variables X, Y, with Gaussian distribution $N(0, \sigma^2)$.

1. Find the joint PDF of the random variables (R, Φ) defined below:



In other words: R is the length of vector (X,Y) and Φ is the angle that the vector creates with the positive side of the x-axis.

2. What is the marginal distribution of R and of Φ?

3. Are (R,Φ) independent?

<u>Solution</u>:

1.

$$\begin{cases} R^2 = X^2 + Y^2 \\ \Phi = \tan^{-1}\left(\dfrac{Y}{X}\right) \end{cases} \rightarrow \begin{cases} X = R\cos(\Phi) \\ Y = R\sin(\Phi) \end{cases}$$

We got a unique solution for X and Y. Let us use the first option of the solution for the transformation of the variables.

According to the proposition, the relationship between the PDF of (X,Y) and the PDF of (R,Φ) is:

$$f_{R\Phi}(r,\phi) = \frac{f_{XY}(x,y)}{\left|\dfrac{\partial(r,\phi)}{\partial(x,y)}\right|} = f_{XY}(x,y)\left|\frac{\partial(x,y)}{\partial(r,\phi)}\right|$$

where, instead of x and y, one shall place $x = r\cos\phi, y = r\sin\phi$.

In our case, it is simpler to calculate the Jacobian of the inverse transformation:

$$\begin{aligned} &x(r,\phi) = r\cos\phi \\ &y(r,\phi) = r\sin\phi \\ &\left|\frac{\partial(x,y)}{\partial(r,\phi)}\right| = \left|\det\begin{pmatrix} \cos(\phi) & -r\sin(\phi) \\ \sin(\phi) & r\cos(\phi) \end{pmatrix}\right| = r \end{aligned}$$

Consequently:

$$f_{R\Phi}(r,\phi) = f_{XY}\left(r\cos(\phi), r\sin(\phi)\right)\cdot r \qquad r \geq 0, \quad 0 < \phi \leq 2\pi$$

In our case, the random variables X,Y are independent and have identical distribution $N(0,\sigma^2)$, and, therefore:

$$f_{XY}(x,y) = f_X(x)\cdot f_Y(y) = \frac{1}{2\pi\sigma^2}e^{-\frac{x^2+y^2}{2\sigma^2}}$$

$$\Rightarrow f_{XY}(r\cos\phi, r\sin\phi) = \frac{1}{2\pi\sigma^2}e^{-\frac{r^2}{2\sigma^2}}$$

And finally: $f_{R\Phi}(r,\phi) = r \cdot f_{XY}(r\cos\phi, r\sin\phi) = \dfrac{1}{2\pi}\dfrac{r}{\sigma^2}e^{-\frac{r^2}{2\sigma^2}} \qquad r>0,\ \ 0<\phi\le 2\pi$

2.

$$f_R(r) = \int_{-\infty}^{\infty} f_{R\Phi}(r,\phi)d\phi = \int_0^{2\pi}\frac{1}{2\pi}\frac{r}{\sigma^2}e^{-\frac{r^2}{2\sigma^2}}d\phi = \frac{r}{\sigma^2}e^{-\frac{r^2}{2\sigma^2}}, \qquad r>0$$

$$f_\Phi(\phi) = \int_{-\infty}^{\infty} f_{R\Phi}(r,\phi)dr = \int_0^{\infty}\frac{1}{2\pi}\frac{r}{\sigma^2}e^{-\frac{r^2}{2\sigma^2}}dr = \frac{1}{2\pi}, \qquad 0<\phi<2\pi$$

Namely:

$$\Phi \sim U(-\pi,\pi] \qquad\qquad R \sim Rayleigh(\sigma)$$

3. It holds that: $f_{R\Phi}(r,\phi) = f_R(r)\cdot f_\Phi(\phi)$, i.e. (R,Φ) are independent.

Generally speaking: (R,Φ) are independent iff one can write their joint PDF as the product: $f_{R\Phi}(r,\phi) = g(r)\cdot h(\phi)$ (g and h are not necessarily density functions!)

Problem 2:

Consider the random variable S defined as $S = \sum_{k=1}^{N} X_k$, where $\{X_k\}_{k=1}^{\infty}$ are random variables independent of N, which is a random variable that takes the values $\{1,2,....\}$

Additionally, it is given:

$$E\{X_k\} = \eta$$

$$Cov\{X_i,X_j\} = \begin{cases}\sigma^2 & i=j\\ 0 & i\ne j\end{cases}$$

Express the expectation and variance of S with the expectation and variance of N.

Solution:

We will first find the expected value and variance of S given the event N=n.

$$E(S\mid N=n) = E(X_1+\cdots+X_N\mid N=n) = E(X_1+\cdots+X_n\mid N=n)$$

$$= \sum_{i=1}^{n}E(X_i\mid N=n) = \sum_{i=1}^{n}E(X_i) = nE(X_i) = n\eta$$

$$
\begin{aligned}
&Var(S \mid N=n) = Var(X_1+\cdots+X_N \mid N=n) \\
&= Var(X_1+\cdots+X_n \mid N=n) = Var(X_1+\cdots+X_n) \\
&\overset{(1)}{=} Cov(X_1+\cdots+X_n, X_1+\cdots+X_n) = \sum_{i=1}^{n}\sum_{j=1}^{n} Cov(X_i, X_j) = \sum_{i=1}^{n}\sum_{j=1}^{n} \sigma^2 \delta_{ij} = n\sigma^2
\end{aligned}
$$

where we make use of the fact that variable N is independent of all the collection $\{X_k\}_{k=1}^{\infty}$ **(it is not enough to assume that it is independent of each X_k separately! Why?)**.

<u>Another way of calculating (1)</u> – in our case, the variables $\{X_k\}_{k=1}^{\infty}$ are **uncorrelated**, and, therefore, the variance of their sum is the sum of the variances.

Simply put, all in all we got:

$$
\begin{aligned}
E(S \mid N) &= N\eta \\
Var(S \mid N) &= N\sigma^2
\end{aligned}
$$

The second moment of S given N is, thus:

$$
E(S^2 \mid N) = [E(S \mid N)]^2 + Var(S \mid N) = N^2\eta^2 + N\sigma^2
$$

Let us now use the Law of Total Expectation (also known as the smoothing theorem) and get the first two moments of S.

<u>**Reminder:** **("Smoothing Theorem")**</u>

$$
\begin{aligned}
E(E(X \mid Y)) &= E(X) \\
E(E(X \mid Y, Z) \mid Z) &= E(X \mid Z)
\end{aligned}
$$

$$
E\{S\} = E\{E\{S \mid N\}\} = E\{N\eta\} = \eta \cdot E\{N\}
$$

$$
E\{S^2\} = E\{E\{S^2 \mid N\}\} = \eta^2 \cdot E\{N^2\} + \sigma^2 \cdot E\{N\}
$$

From here, the variance of S is:

$$
Var\{S\} = E\{S^2\} - E^2\{S\} = \eta^2\left(E\{N^2\} - E^2\{N\}\right) + \sigma^2 E\{N\} = \eta^2 Var\{N\} + \sigma^2 E\{N\}
$$

Extra Questions

Problem 3

Given are the following random variables (which are independent):

$$X \sim Unif\left([-\tfrac{1}{2}, \tfrac{1}{2}]\right) \quad Y \sim Unif\left([-\tfrac{1}{2}, \tfrac{1}{2}]\right)$$

1. Calculate:

$$E\{X\}, E\{X^2\}, E\{XY\}$$

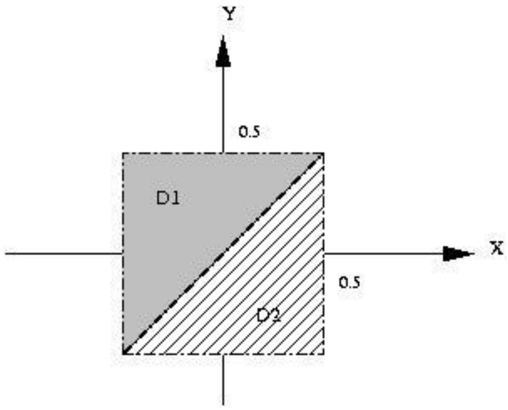
2. Find the joint PDF $f_{XY}(x,y)$ and draw the section in which it is not zero.

3. Calculate:

$$E\left\{\left\|[X \quad Y]\right\|^2\right\} = E\left\{X^2 + Y^2\right\}$$

4. Let us define the following events:

- D1 – the pair (X,Y) belongs to the triangle marked as D1 in the figure to the left.
- D2 – the pair (X,Y) belongs to the triangle marked as D2 in the figure.



Find the values of:

$$E\left\{X^2 + Y^2 \mid D_1\right\}, E\left\{X^2 + Y^2 \mid D_2\right\}$$

(use symmetry). A numerical answer is required.

Now, the rotation transformation of variables X,Y is defined:

$$\begin{bmatrix} W \\ V \end{bmatrix} = R(\theta) \begin{bmatrix} X \\ Y \end{bmatrix}$$

where:

$$R(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

In the following sections, assume $\theta = 45^o$.

5. Find the joint PDF $f_{WV}(w,v)$ and draw the section in which it is not zero. Are they independent?

6. Calculate:

$$E\left\{\left\|[W \quad V]\right\|^2\right\} = E\left\{W^2 + V^2\right\}$$

Solution:

1.

$$E\{X\} = \int_{-0.5}^{0.5} x f_x dx = 0$$

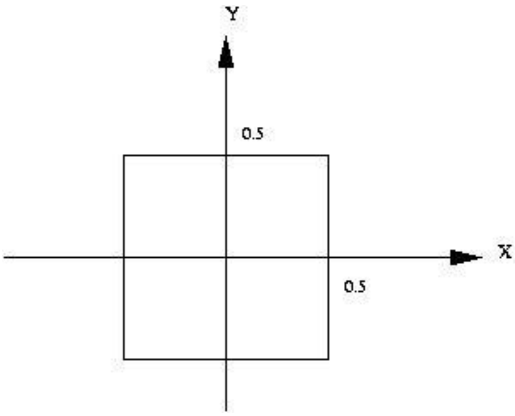
$$E\{X^2\} = \int_{-0.5}^{0.5} x^2 f_x dx = \int_{-0.5}^{0.5} x^2 dx = \frac{x^3}{3}\Bigg|_{-0.5}^{0.5} = \frac{1}{3}\left[\frac{1}{8} - \left(-\frac{1}{8}\right)\right] = \frac{1}{12}$$

$$E\{xy\} \underset{a}{=} E\{x\}E\{y\} = 0$$

The equality a is due to the fact that X and Y are independent.

2. Since X and Y are independent:

$$f_{x,y}(x,y) = f_x(x)\cdot f_y(y) = \begin{cases} 1 & if\ -0.5 \le X \le 0.5 \cap -0.5 \le Y \le 0.5 \\ 0 & otherwise \end{cases}$$



3.

$$E\{X^2 + Y^2\} = E\{X^2\} + E\{Y^2\} = 2\frac{1}{12} = \frac{1}{6}$$

4. Notice that for any point (x_1, y_1) in triangle D1 that satisfies $x_1{}^2 + y_1{}^2 = d$, there exists a point (x_2, y_2) in triangle D2 that satisfies $x_2{}^2 + y_2{}^2 = d$. Together with the fact that the conditional distribution inside each of the triangles in uniform, we conclude:

$$E\left\{X^2 + Y^2 \mid D1\right\} = E\left\{X^2 + Y^2 \mid D2\right\}$$

By applying the smoothing theorem on $E\{X^2 + Y^2\}$:

$$E\left\{X^2+Y^2\right\} = E\left\{E\left\{X^2+Y^2 \mid D\right\}\right\} \underset{a}{=} \frac{1}{2}E\left\{X^2+Y^2 \mid D1\right\} + \frac{1}{2}E\left\{X^2+Y^2 \mid D2\right\} \underset{b}{=} E\left\{X^2+Y^2 \mid D1\right\} \underset{c}{=} \frac{1}{6}$$

The equality a is because the triangles D1 and D2 have the same area and, therefore:

$$\Pr(D = D1) = \Pr(D = D2) = \frac{1}{2}$$

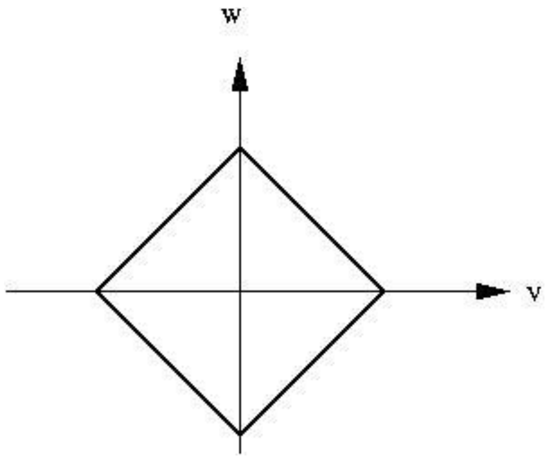
The equality b is due to:

$$E\{X^2 + Y^2 \mid D1\} = E\{X^2 + Y^2 \mid D2\}$$

And the equality c is true from the results of section 3.

$$E\{X^2 + Y^2\} = \frac{1}{6}$$

5. Since the matrix $R(\theta)$ rotates the plane by $\theta = 45^o$, we get the following PDF:



$$f_{WV}(w,v) = \begin{cases} 1 & if -\frac{1}{\sqrt{2}} \le v \le \frac{1}{\sqrt{2}} \cap \left(|v| - \frac{1}{\sqrt{2}}\right) \le w \le \left(\frac{1}{\sqrt{2}} - |v|\right) \\ 0 & otherwise \end{cases}$$

We could have also calculated it directly from the formula of a linear transformation of a random vector:

$$f_{WV}(w,v) = \frac{1}{|\det(R(\theta))|} f_{x,y}\left(R^{-1}(\theta)\begin{bmatrix} w \\ v \end{bmatrix}\right)$$

$$R(\theta) = \begin{vmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{vmatrix} = \cos^2\theta + \sin^2\theta = 1$$

$$R^{-1}(\theta) = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$$

$$f_{WV}(w,v) = \frac{1}{1} f_{x,y}\left(\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} w \\ v \end{bmatrix}\right) = f_{x,y}(w\cos\theta + v\sin\theta, v\cos\theta - w\sin\theta) =$$

$$= f_x(w\cos\theta + v\sin\theta) f_y(v\cos\theta - w\sin\theta) =$$

$$= \begin{cases} 1 & -\frac{1}{2} \le w\cos\theta + v\sin\theta \le \frac{1}{2} \cap -\frac{1}{2} \le v\cos\theta - w\sin\theta \le \frac{1}{2} \\ 0 & \end{cases}$$

By substituting $\theta = 45^o$, we get the same result we got before.

W and V are dependent. For example, if $W = \frac{1}{\sqrt{2}}$, then the value of V is set deterministically $V = 0$, namely $f_{V|W}(v \mid w) \neq f_v(v)$.

6. The transformation is a rotation transformation (orthonormal transformation), and, therefore, the squared distance from any point to the origin does not change. To see this, consider a point (w_0, v_0) created by the transformation on (x_0, y_0):

$$\begin{bmatrix} w_0 \\ v_0 \end{bmatrix} = R(\theta) \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$$

$$w_0^{\ 2} + v_0^{\ 2} = \begin{bmatrix} w_0 & v_0 \end{bmatrix} \begin{bmatrix} w_0 \\ v_0 \end{bmatrix} = \begin{bmatrix} x_0 & y_0 \end{bmatrix} R(\theta)^T R(\theta) \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} =$$

$$= \begin{bmatrix} x_0 & y_0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} = x_0^{\ 2} + y_0^{\ 2}$$

We indeed got that the squared distance from any point to the origin does not change. Since the squared distance from any point to the origin does not change after the transformation, the expectation of the squared distance does not change as well, thus:

$$E\{W^2 + V^2\} = E\{X^2 + Y^2\} = \frac{1}{6}$$

Problem 4:

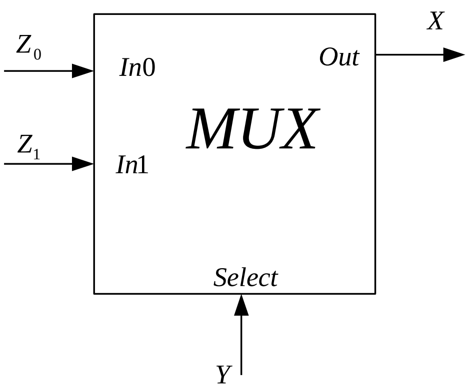
Consider a multiplexer with two information inputs (In0, In1), one control input (Select) and an output (Out). In our example, the information inputs will be the random variables Z_0, Z_1, the Select input will be random variable Y and the multiplexer's output will be denoted as the random variable X:

$$Z_1 \sim N\left(\mu, \sigma^2\right) \qquad Z_0 \sim Poisson\left(\lambda\right) \qquad Y \sim Ber\left(\varepsilon\right) = \begin{cases} 0 & w.p\ 1-\varepsilon \\ 1 & w.p\ \varepsilon \end{cases} \qquad 0 < \varepsilon < 1, \lambda > 0$$

where Z_0, Z_1, Y are independent.

The output X is chosen from Z_0 and Z_1 according to Y:

$$X = \begin{cases} Z_0 & if\ \ Y = 0 \\ Z_1 & if\ \ Y = 1 \end{cases}$$



Find the expectation of X and $f_X(x)$.

Solution:

Notice that $X = Z_Y$ holds.

Therefore:

$$E\{X \mid Y=0\}=E\{Z_Y \mid Y=0\}=E\{Z_0 \mid Y=0\}\underset{(1)}{=}E\{Z_0\}=\lambda$$

$$E\{X \mid Y=1\}=E\{Z_Y \mid Y=1\}=E\{Z_1 \mid Y=1\}\underset{(1)}{=}E\{Z_1\}=\mu$$

(1) independence of Z_0, Z_1, Y

$$h(\alpha)=E\{X \mid Y=\alpha\}=\begin{cases}\lambda & \alpha=0\\ \mu & \alpha=1\end{cases} \qquad \Rightarrow E\{X \mid Y\}=h(Y)=\begin{cases}\lambda & w.p\ 1-\varepsilon\\ \mu & w.p\ \ \varepsilon\end{cases}$$

$$E\{X\}=E\{E\{X \mid Y\}\}=E\{h(Y)\}=\varepsilon h(1)+(1-\varepsilon)h(0)=\varepsilon\mu+(1-\varepsilon)\lambda$$

Moreover, we can find the PDF of X. From the given above, we know that:

$$f_{X|Y}(x \mid 0)=\sum_{k=0}^{\infty}\frac{e^{-\lambda}\lambda^k}{k!}\delta(x-k)$$

$$f_{X|Y}(x \mid 1)=\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Consequently, from the Law of Total Probability:

$$f_X(x)=\Pr\{Y=0\}f_{X|Y}(x \mid 0)+\Pr\{Y=1\}f_{X|Y}(x \mid 1)=$$

$$=(1-\varepsilon)\cdot\sum_{k=0}^{\infty}\frac{e^{-\lambda}\lambda^k}{k!}\delta(x-k)+\varepsilon\cdot\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Appendix – Theoretical Background

[1] Random Variables

1.1 Joint CDF

The joint CDF of n random variables is:

$$F_{X_1X_2\ldots X_n}(x_1,x_2,\ldots,x_n)=F_{\underline{X}}(\underline{x})=\Pr\{X_1\le x_1, X_2\le x_2,\ldots,X_n\le x_n\}$$

Characteristics (for two random variables X,Y, but can be easily generalized to n random variables):

- $F_{XY}(\infty,\infty)=1$
- $F_{XY}(-\infty,y)=F_{XY}(x,-\infty)=0$
- $F_{XY}(x,\infty)=F_X(x)$
  $F_{XY}(\infty,y)=F_Y(y)$
- $P\{x_1\le X\le x_2, Y\le y\}=F_{XY}(x_2,y)-F_{XY}(x_1,y)$
  $P\{X\le x, y_1\le Y\le y_2\}=F_{XY}(x,y_2)-F_{XY}(x,y_1)$
- $P\{x_1\le X\le x_2, y_1\le Y\le y_2\}=F_{XY}(x_2,y_2)-F_{XY}(x_1,y_2)-F_{XY}(x_2,y_1)+F_{XY}(x_1,y_1)$

1.2 Joint PDF

The joint PDF of n random variables is:

$$f_{X_1X_2\ldots X_n}(x_1,x_2,\ldots,x_n)=f_{\underline{X}}(\underline{x})=\frac{\partial^n F_{X_1X_2\ldots X_n}(x_1,x_2,\ldots,x_n)}{\partial x_1\ldots\partial x_n}$$

The marginal PDF (of random variable X_1, for example):

$$f_{X_1}(x_1)=\int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty} f_{X_1X_2\ldots X_n}(x_1,x_2,\ldots,x_n)dx_2dx_3\ldots dx_n$$

Characteristics:

The joint PDF has the same known characteristics of the PDF of a single random variable:

- $f_{X_1X_2\ldots X_n}(x_1,x_2,\ldots,x_n)\ge 0 \qquad \forall x_1,\ldots x_n$ and
  $\int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty} f_{X_1X_2\ldots X_n}(x_1,x_2,\ldots,x_n)dx_1\ldots dx_n=1$
- $F_{X_1X_2\ldots X_n}(x_1,x_2,\ldots,x_n)=\int_{-\infty}^{x_1}\ldots\int_{-\infty}^{x_n} f_{X_1X_2\ldots X_n}(\alpha_1,\alpha_2,\ldots,\alpha_n)d\alpha_n\ldots d\alpha_1$

1.3 Conditional PDF

$$f_{X_1,,,X_n|Y_1....Y_k}\left(x_1,...,x_n \mid y_1,....,y_k\right)=\frac{f_{X_1...X_n,Y_1...Y_k}\left(x_1,...,x_n,y_1,....,y_k\right)}{f_{Y_1....Y_k}\left(y_1,....,y_k\right)}$$

$$f_{X_1...X_n,Y_1...Y_k}\left(x_1,...,x_n,y_1,....,y_k\right)=f_{X_1....X_n|Y_1....Y_k}\left(x_1,...,x_n \mid y_1,....,y_k\right)f_{Y_1....Y_k}\left(y_1,....,y_k\right)$$

$$=f_{Y_1....Y_k|X_1...X_n}\left(y_1,....,y_k \mid x_1,...,x_n\right)\cdot f_{X_1...X_n}\left(x_1,...,x_n\right)$$

[!] Highlight (with regard to continuous and discrete random variables):

Given discrete random variables $B_1,...,B_m$, and continuous random variables $X_1,...,X_k$, the above statements hold. Furthermore, the following also holds:

Law of Total Probability:

$$\Pr\left\{B_1=b_1,B_2=b_2,...,B_m=b_m\right\}=$$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}....\int_{-\infty}^{\infty}\Pr\left\{B_1=b_1,B_2=b_2,...,B_m=b_m \mid X_1=x_1,X_2=x_2,...,X_k=x_k\right\}f_{X_1...X_k}\left(x_1,x_2,..,x_k\right)dx_1dx_2...dx_k$$

$$f_{X_1...X_k}\left(x_{1,}x_2,....,x_k\right)=$$

$$\sum_{b_1}\sum_{b_2}...\sum_{b_m}f_{X_1...X_k|B_1...B_m}\left(x_1,..,x_k \mid b_1,b_2,..b_m\right)\Pr\left\{B_1=b_1,B_2=b_2...,B_m=b_m\right\}$$

Conditional Probability:

$$f_{X_1...X_k|B_1...B_m}\left(x_1,...,x_k \mid b_1,...,b_m\right)=\frac{\Pr\left\{B_1=b_1,B_2=b_2,...,B_m=b_m \mid X_1=x_1,...,X_k=x_k\right\}f_{X_1...X_k}\left(x_1....x_k\right)}{\Pr\left\{B_1=b_1,B_2=b_2,...,B_m=b_m\right\}}$$

$$\Pr\left\{B_1=b_1,B_2=b_2,...,B_m=b_m \mid X_1=x_1,...,X_k=x_k\right\}=$$
$$\frac{f_{X_1...X_k|B_1...B_m}\left(x_1,...,x_k \mid b_1,...,b_m\right)\Pr\left\{B_1=b_1,B_2=b_2,...,B_m=b_m\right\}}{f_{X_1...X_k}\left(x_1....x_k\right)}$$

1.4 Conditional Expectation:

$$E\{X \mid Y=y\}=\int_{-\infty}^{\infty}xf_{X|Y}(x \mid y)dx$$

$$E\{X \mid Y_1=y_1,Y_2=y_2,...,Y_k=y_k\}=\int_{-\infty}^{\infty}xf_{X|Y_1....Y_k}(x \mid y_1,....,y_k)dx$$

***[!]* Highlight (conditional expectation as a function a random variable):**

One can understand the conditional expectation of X on the value of a random variable ($Y=\alpha$)[1] as a deterministic function of this value (we will denote it $h(\bullet)$)

$$h(\alpha)=E\{X\mid Y=\alpha\}$$

Thus, for this function $h(\bullet)$, if we apply it on the random variable Y, we get a new random variable. In other words:

$$h(Y)=E\{X\mid Y\}$$

1.5 Smoothing Theorem:

$$E\ \{g(X,Y)\}=E\{E\{g(X,Y)\mid Y\}\}$$
$$E\{X_1\}=E\{E\{X_1\mid Y_1,\ldots,Y_k\}\}$$

[!] **Highlight (smoothing as a means of calculating conditional expectation):**

Notice that:

$$E\{X_1\mid Y_1,Y_2,..Y_k\}=E\{E\{X_1\mid Z_{1,}Z_2,\ldots,Z_{m,}Y_1,Y_2,..Y_k\}\mid Y_1,Y_2,..Y_k\}$$

1.6 Joint Moments:

- Joint moments:
$$m_{nk}=E\{X^n\cdot Y^k\}$$

- Joint central moments:
$$\mu_{nk}=E\{(X-E\{X\})^n\cdot(Y-E\{Y\})^k\}$$

- Covariance:
$$\sigma_{XY}=Cov(X,Y)=E\{(X-E\{X\})(Y-E\{Y\})\}=E\{XY\}-E\{X\}E\{Y\}=\mu_{11}$$

- Correlation: $E\{XY\}=m_{11}$

- Correlation coefficient: $\rho\equiv\dfrac{\sigma_{XY}}{\sigma_X\sigma_Y}$

1.7 Relationships between Random Variables:

[1] Statistical Independence

Pair of Random Variables:

The random variables X and Y are statistically independent iff:

[1] Namely, the expectation conditioned on event $\{Y=\alpha\}$.

$$F_{XY}(x,y) = F_X(x)F_Y(y) \qquad \forall x,y$$
$$F_{X|Y}(x|y) = F_X(x) \qquad \forall x,y$$
$$f_{XY}(x,y) = f_X(x)\cdot f_Y(y) \qquad \forall x,y$$
$$f_{X|Y}(x|y) = f_X(x) \qquad \forall x,y$$

<u>Independence between Random Vectors:</u>

Given the following random vectors: $\underline{X} = [X_1,...,X_n]^T \quad \underline{Y} = [Y_1,...,Y_m]^T$

It is said that they are independent iff:

$$F_{\underline{X},\underline{Y}}(\underline{x},\underline{y}) = F_{\underline{X}}(\underline{x})F_{\underline{Y}}(\underline{y}) \ \forall \underline{x},\underline{y}$$
$$F_{\underline{X}|\underline{Y}}(\underline{x}|\underline{y}) = F_{\underline{X}}(\underline{x}) \ \forall \underline{x},\underline{y}$$
$$f_{\underline{X},\underline{Y}}(\underline{x},\underline{y}) = f_{\underline{X}}(\underline{x})f_{\underline{Y}}(\underline{y}) \ \forall \underline{x},\underline{y}$$
$$f_{\underline{X}|\underline{Y}}(\underline{x}|\underline{y}) = f_{\underline{X}}(\underline{x}) \ \forall \underline{x},\underline{y}$$

<u>[2] Uncorrelated Random Variables:</u>

A pair of random variables for which $\rho_{XY} = 0$ holds is called uncorrelated.

<u>[3] Orthogonality:</u>

A pair of random variables for which $E\{XY\} = 0$ holds is called orthogonal.

<u>Some simple claims:</u>

- For two independent random vectors $\underline{X}$ and $\underline{Y}$, for all $g(\bullet)$ and $f(\bullet)$ deterministic scalar functions, the following holds:
$$E\{g(\underline{X})\cdot f(\underline{Y})\} = E\{g(\underline{X})\}\cdot E\{f(\underline{Y})\}$$
- For two independent random variables X and Y, for all n,k:
$$\mu_{1k} = \mu_{n1} = 0$$
- Statistical independence $\Rightarrow$ no correlation
- No correlation and one of the variables has expectation zero $\Rightarrow$ orthogonality.
- Two uncorrelated random variables that are jointly Gaussian $\Rightarrow$ statistical independence.
- For two random variables X and Y:
$$Var\{X+Y\} = Var\{X\} + 2\cdot Cov(X,Y) + Var\{Y\}$$
and, thus, the variables are uncorrelated:
$$Var\{X+Y\} = Var\{X\} + Var\{Y\}$$

<u>[2] Characteristic Function of Random Vectors:</u>

<u>2.1 Characteristic Function of a Pair of Random Variables</u>

For a pair of random variables X,Y, the first characteristic function is defined as follows:

$$\phi_{XY}(\omega_1,\omega_2) = E\left\{e^{j(\omega_1 X + \omega_2 Y)}\right\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x,y) e^{j(\omega_1 x + \omega_2 y)} dxdy$$

Moreover:

$$f_{XY}(x,y) = \frac{1}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\phi_{XY}(\omega_1,\omega_2) e^{-j(\omega_1 x + \omega_2 y)} d\omega_1 d\omega_2$$

Through the characteristic function, one can create the moment of order p :

$$m_{nk} = \frac{1}{j^p}\frac{\partial^p \phi_{XY}(\omega_1,\omega_2)}{\partial\omega_1^n \partial\omega_2^k}\bigg|_{\omega_1=\omega_2=0}, \qquad n+k=p$$

<u>Characteristics:</u>

- $\phi_{XY}(0,0) = 1$
- $\phi_Y(\omega) = \phi_{XY}(0,\omega)$
  $\phi_Y(\omega) = \phi_{XY}(0,\omega)$

<u>2.2 Characteristic Function of Random Vectors:</u>

For random vector $\underline{X}$ of length n, the first characteristic function is defined as:

$$\phi_{\underline{X}}(\underline{\omega}) = E\left\{\exp\left\{j\underline{\omega}^T \underline{X}\right\}\right\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f_{\underline{X}}(\underline{x})\exp\left\{j\underline{\omega}^T \underline{x}\right\} d\underline{x}$$

(the vector $\underline{\omega}$ is a vector of length n)

$$f_{\underline{X}}(\underline{x}) = \frac{1}{(2\pi)^n}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\phi_{\underline{X}}(\underline{\omega})\exp\left\{-j\underline{\omega}^T \underline{x}\right\} d\underline{\omega}$$

Moments can be created through the first characteristic function similarly to what we saw in 2.1.