MP-208
Optimal Filtering with Aerospace Applications
Section 2.6: Random Vectors

Prof. Dr. Davi Antônio dos Santos
Instituto Tecnológico de Aeronáutica
www.professordavisantos.com

São José dos Campos - SP
2023

Contents

1. Definition
2. Probability Distribution Function
3. Probability Density Function
4. Expected Value
5. Independence
6. Conditional Distribution
7. Important Rules
8. Gaussian Vector

Definition. . .

Definition

Random Vector:

Consider a probability space $(\Omega, \mathcal{F}, P)$. A continuous random vector is a map $\mathbf{X} : \Omega \to \mathbb{R}^n$. The function $\mathbf{X}$ is arbitrary, except for the following restriction:

$$A(\mathbf{x}) \triangleq \{\omega \in \Omega : \mathbf{X}(\omega) \leq \mathbf{x}\} \in \mathcal{F}, \ \forall \mathbf{x} \in \mathbb{R}^n$$

$$\text{and} \quad P(A(\infty, \infty, \ldots, \infty)) = 1$$

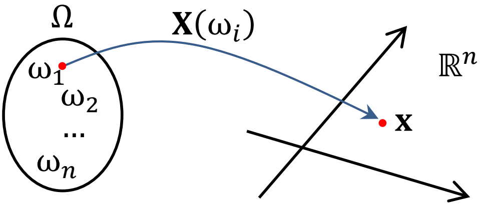
Remark:

We often abbreviate:

- $\{\omega \in \Omega : \mathbf{X}(\omega) \leq \mathbf{x}\}$ by $\{\mathbf{X} \leq \mathbf{x}\}$
- $\{\omega \in \Omega : \mathbf{x}_1 \leq \mathbf{X}(\omega) \leq \mathbf{x}_2\}$ by $\{\mathbf{x}_1 \leq \mathbf{X} \leq \mathbf{x}_2\}$
- etc.

Definition

Illustration of a Random Vector:



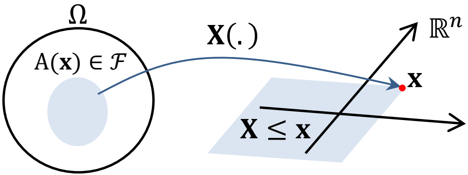
Probability Distribution Function . . .

Probability Distribution Function

Definition:

The probability (cumulative) distribution function (cdf) of the random vector $\mathbf{X}$ is defined by:

$$F_{\mathbf{X}}(\mathbf{x}) \triangleq P\left(A(\mathbf{x})\right), \quad \forall \mathbf{x} \in \mathbb{R}^n$$

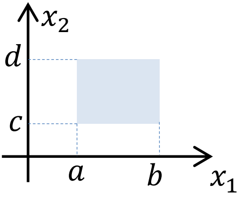


Probability Distribution Function

Properties:

For simplicity, consider $n = 2$ and denote the vector components as in $\mathbf{x} = (x_1, x_2)$ and $\mathbf{X} = (X_1, X_2)$.

1. $F_{\mathbf{X}}(\infty, \infty) = 1$, $F_{\mathbf{X}}(-\infty, x_2) = 0$, $F_{\mathbf{X}}(x_1, -\infty) = 0$.
2. $P(\{a \leq X_1 \leq b\} \cap \{c \leq X_2 \leq d\}) =$
   $F_{\mathbf{X}}(b, d) - F_{\mathbf{X}}(a, d) - F_{\mathbf{X}}(b, c) + F_{\mathbf{X}}(a, c)$.



Remark:

We often abbreviate:

- $F_{\mathbf{X}}(\mathbf{x})$ by $F(\mathbf{x})$

Probability Density Function . . .

Probability Density Function

Definition:

The probability density function (pdf) of a random vector $\mathbf{X}$ is defined as:

$$f_{\mathbf{X}}(\mathbf{x}) \triangleq \frac{\partial^n F_{\mathbf{X}}(\mathbf{x})}{\partial x_1 \partial x_2 \ldots \partial x_n}$$

We often abbreviate $f_{\mathbf{X}}(\mathbf{x})$ by $f(\mathbf{x})$.

Properties:

1. $f_{\mathbf{X}}(\mathbf{x}) \geq 0$.
2. $\displaystyle\int_{\mathbb{R}^n} f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} = 1$.
3. $P\left(\{\mathbf{X} \in R\}\right) = \displaystyle\int_R f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x}$, where $R \subset \mathbb{R}^n$ is a given set.

Marginal Statistics

Definition:

In multiple random variable theory, a marginal statistic (*e.g.*, pdf or cdf) is a statistic that characterizes only part of the random variables.

Consider $n = 2$ and denote the joint cdf by $F_{\mathbf{X}}(\mathbf{x}) \equiv F_{X_1X_2}(x_1, x_2)$.

Marginal Distribution:

- $F_{X_1}(x_1) = F_{X_1X_2}(x_1, \infty)$ is the marginal distribution of X_1.
- $F_{X_2}(x_2) = F_{X_1X_2}(\infty, x_2)$ is the marginal distribution of X_2.

Marginal Density:

- $f_{X_1}(x_1) = \int_{-\infty}^{\infty} f_{X_1X_2}(x_1, x_2) dx_2$ is the marginal density of X_1.
- $f_{X_2}(x_2) = \int_{-\infty}^{\infty} f_{X_1X_2}(x_1, x_2) dx_1$ is the marginal density of X_2.

Expected Value. . .

Expected Value

Definition:

The expected value, expectation or mean of a random vector $\mathbf{X} : \Omega \to \mathbb{R}^n$ is defined by:

$$\mathbf{m}_{\mathbf{X}} = E(\mathbf{X}) \triangleq \int_{\mathbb{R}^n} \mathbf{x} f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x}$$

Properties:

Consider a constant vector $\mathbf{a} \in \mathbb{R}^n$, a constant matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$, and the random vectors $\mathbf{X}_i : \Omega \to \mathbb{R}^n$, $i = 1, \ldots, m$.

1. If $\mathbf{X}_1(\omega) = \mathbf{a}, \forall \omega \in \Omega$, then $E(\mathbf{X}_1) = \mathbf{a}$.
2. $E(\mathbf{A}\mathbf{X}_1) = \mathbf{A}E(\mathbf{X}_1)$.
3. $E(\mathbf{X}_1 + \mathbf{X}_2 + \cdots + \mathbf{X}_n) = E(\mathbf{X}_1) + E(\mathbf{X}_2) + \cdots + E(\mathbf{X}_n)$.
4. $E(g(\mathbf{X})) \triangleq \int_{\mathbb{R}^n} g(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x}$.

Expected Value

Covariance of Random Variables:

The cross covariance of two random variables $X_i : \Omega \to \mathbb{R}$ and $X_j : \Omega \to \mathbb{R}$ is given by:

$$C_{ij} \triangleq E\left((X_i - m_i)(X_j - m_j)\right) = E(X_i X_j) - m_i m_j$$

where m_i and m_j are the expected values of X_i and X_j, respectively. If $i = j$, we have the autocovariance C_{ii} of X_i.

Remark:

Two random variables X_i and X_j are said uncorrelated if $C_{ij} = 0$. Note that this is equivalent to

$$E(X_i X_j) = E(X_i)E(X_j)$$

Expected Value

Covariance of Random Vectors:

The cross covariance of two random vectors $\mathbf{X}_i : \Omega \to \mathbb{R}^n$ and $\mathbf{X}_j : \Omega \to \mathbb{R}^m$ is given by:

$$\mathbf{C}_{ij} \triangleq E\left((\mathbf{X}_i - \mathbf{m}_i)(\mathbf{X}_j - \mathbf{m}_j)^{\mathrm{T}}\right) = E(\mathbf{X}_i \mathbf{X}_j^{\mathrm{T}}) - \mathbf{m}_i \mathbf{m}_j^{\mathrm{T}}$$

where $\mathbf{m}_i$ and $\mathbf{m}_j$ are the expected values of $\mathbf{X}_i$ and $\mathbf{X}_j$, respectively.

Remark:

Two random vectors $\mathbf{X}_i$ and $\mathbf{X}_j$ are said uncorrelated if $\mathbf{C}_{ij} = \mathbf{0}_{n \times m}$. Note that this is equivalent to

$$E\left(\mathbf{X}_i \mathbf{X}_j^{\mathrm{T}}\right) = E(\mathbf{X}_i) E(\mathbf{X}_j)^{\mathrm{T}}$$

Independence. . .

Independence

Definition:

The random variables $X_i : \Omega \to \mathbb{R}, i = 1, ..., n,$ are (statistically) independent if the events $\left\{\omega \in \Omega : X_i(\omega) \leq x_i\right\}, i = 1, ..., n,$ are independent. From this, one can obtain:

$$F_{\mathbf{X}}(\mathbf{x}) = F_{X_1}(x_1)F_{X_2}(x_2)...F_{X_n}(x_n)$$

$$f_{\mathbf{X}}(\mathbf{x}) = f_{X_1}(x_1)f_{X_2}(x_2)...f_{X_n}(x_n)$$

Similarly, if two (or more) random vectors $\mathbf{X}_1$ and $\mathbf{X}_2$ are independent, then:

$$F_{\mathbf{X}_1\mathbf{X}_2}(\mathbf{x}_1, \mathbf{x}_2) = F_{\mathbf{X}_1}(\mathbf{x}_1)F_{\mathbf{X}_2}(\mathbf{x}_2)$$

$$f_{\mathbf{X}_1\mathbf{X}_2}(\mathbf{x}_1, \mathbf{x}_2) = f_{\mathbf{X}_1}(\mathbf{x}_1)f_{\mathbf{X}_2}(\mathbf{x}_2)$$

Conditional Distribution…

Conditional Distribution

Conditional Probability Density Function:

Consider two random vectors $\mathbf{X}_1 : \Omega \to \mathbb{R}^n$ and $\mathbf{X}_2 : \Omega \to \mathbb{R}^m$ with joint pdf $f_{\mathbf{X}_1\mathbf{X}_2}(\mathbf{x}_1, \mathbf{x}_2)$. The conditional pdf of $\mathbf{X}_1$ given the event $\{\mathbf{X}_2 = \mathbf{x}_2\}$ is

$$f_{\mathbf{X}_1|\mathbf{X}_2}(\mathbf{x}_1|\mathbf{x}_2) = \frac{f_{\mathbf{X}_1\mathbf{X}_2}(\mathbf{x}_1, \mathbf{x}_2)}{f_{\mathbf{X}_2}(\mathbf{x}_2)}$$

for $f_{\mathbf{X}_2}(\mathbf{x}_2) > 0$.

Remarks:

- A conditional probability can be computed from the conditional pdf:

$$P\left(\{\mathbf{X}_1 \in R_1\} \,|\{\mathbf{X}_2 = \mathbf{x}_2\}\right) = \int_{R_1} f_{\mathbf{X}_1|\mathbf{X}_2}(\mathbf{x}_1|\mathbf{x}_2) d\mathbf{x}_1$$

- If $\mathbf{X}_1$ and $\mathbf{X}_2$ are independent, then $f_{\mathbf{X}_1|\mathbf{X}_2}(\mathbf{x}_1|\mathbf{x}_2) = f_{\mathbf{X}_1}(\mathbf{x}_1)$.

Conditional Distribution

Conditional Expectation:

Consider two random vectors $\mathbf{X}_1 : \Omega \to \mathbb{R}^n$ and $\mathbf{X}_2 : \Omega \to \mathbb{R}^m$ with joint pdf $f_{\mathbf{X}_1\mathbf{X}_2}(\mathbf{x}_1, \mathbf{x}_2)$. Analogous to the unconditional expected value, we define the conditional expectation of $g(\mathbf{X}_1)$ given the event $\{\mathbf{X}_2 = \mathbf{x}_2\}$ as

$$E(g(\mathbf{X}_1)|\{\mathbf{X}_2 = \mathbf{x}_2\}) \triangleq \int_{\mathbb{R}^n} g(\mathbf{x}_1) f_{\mathbf{X}_1|\mathbf{X}_2}(\mathbf{x}_1|\mathbf{x}_2) d\mathbf{x}_1$$

from which one can obtain the conditional mean $\mathbf{m}_{\mathbf{X}_1|\mathbf{X}_2}$ and the conditional autocovariance $\mathbf{C}_{\mathbf{X}_1\mathbf{X}_1|\mathbf{X}_2}$, by making, respectively, $g(\mathbf{X}_1) = \mathbf{X}_1$ and $g(\mathbf{X}_1) = (\mathbf{X}_1 - \mathbf{m}_{\mathbf{X}_1|\mathbf{X}_2})(\mathbf{X}_1 - \mathbf{m}_{\mathbf{X}_1|\mathbf{X}_2})^{\mathrm{T}}$.

Remarks:

- We often abbreviate $E(g(\mathbf{X}_1)|\{\mathbf{X}_2 = \mathbf{x}_2\})$ by $E\left(g(\mathbf{X}_1)|\mathbf{X}_2\right)$.
- $E\left(E(g(\mathbf{X}_1)|\mathbf{X}_2)\right) = E\left(g(\mathbf{X}_1)\right)$.

Important Rules...

Important Rules

Bayes Rule:

Consider two random vectors $\mathbf{X}_1 : \Omega \to \mathbb{R}^n$ and $\mathbf{X}_2 : \Omega \to \mathbb{R}^m$. The Bayes Rule (or Theorem) says that:

$$f_{\mathbf{X}_1|\mathbf{X}_2}(\mathbf{x}_1|\mathbf{x}_2) = \frac{f_{\mathbf{X}_2|\mathbf{X}_1}(\mathbf{x}_2|\mathbf{x}_1) f_{\mathbf{X}_1}(\mathbf{x}_1)}{f_{\mathbf{X}_2}(\mathbf{x}_2)}$$

where the marginal pdf $f_{\mathbf{X}_2}(\mathbf{x}_2) > 0$ is given by:

$$f_{\mathbf{X}_2}(\mathbf{x}_2) = \int_{\mathbb{R}^n} f_{\mathbf{X}_2|\mathbf{X}_1}(\mathbf{x}_2|\mathbf{x}_1) f_{\mathbf{X}_1}(\mathbf{x}_1) d\mathbf{x}_1$$

Remark:

This result is useful for formulating Bayesian particle filters.

Important Rules

Chain Rule:

Consider the random vectors $\mathbf{X}_i : \Omega \to \mathbb{R}^{n_i}, i = 1, ..., N,$. The chain rule says that:

$$\begin{aligned} f(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_N) &= f(\mathbf{x}_N | \mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_{N-1}) \times \\ & \quad f(\mathbf{x}_{N-1} | \mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_{N-2}) \times \ldots \\ & \quad \ldots \times f(\mathbf{x}_2 | \mathbf{x}_1) \times f(\mathbf{x}_1) \end{aligned}$$

Remark:

This result is also useful for formulating Bayesian particle filters!

Gaussian Vector...

Gaussian Vector

Definition:

A Gaussian vector is a random vector $\mathbf{X} : \Omega \to \mathbb{R}^n$ with joint pdf given by:

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{(2\pi)^{n/2} \det^{1/2}(\mathbf{C})} \exp\left\{ -\frac{1}{2} (\mathbf{x} - \mathbf{m})^{\mathrm{T}} \mathbf{C}^{-1} (\mathbf{x} - \mathbf{m}) \right\}$$

where

$$\mathbf{m} = E(\mathbf{X})$$

$$\mathbf{C} = E\left((\mathbf{X} - \mathbf{m}) (\mathbf{X} - \mathbf{m})^{\mathrm{T}} \right)$$

Remark:

We denote $\mathbf{X} \sim \mathcal{N}(\mathbf{m}, \mathbf{C})$.

References...

References

- Papoulis, A.; Pillai, S. U. **Probability, Random Variables, and Stochastic Processes**. New York: McGraw-Hill, 2002.
- Ross, S. M. **A First Course in Probability**. New York: Prentice Hall, 2002.