SP

Stochastic Processes

SP.1 RANDOM VARIABLES

Consider an event ω, possible outcome of a random experiment. The *probability* of this event, denoted as $\Pr(\omega)$, may be intuitively considered to coincide with the limit, when the number of trials tends to infinity, of the ratio between the number of times that ω occurred and the number of experiments. If all possible outcomes are denoted by $\omega_i,\ i = 1 \dots, N$, then

$$0 \leq \Pr(\omega_i) \leq 1, \qquad \sum_{i=1}^{N} \Pr(\omega_i) = 1. \tag{SP.1.1}$$

Definition SP.1.1 (Independent events) – The events A, B, C, are defined as mutually independent if

$$\Pr(ABC) = \Pr(A)\Pr(B)\Pr(C) \tag{SP.1.2}$$

where ABC denotes the joint occurrence of A, B and C.

The definition of *conditional probability* can be used to describe the probability of occurrence of non independent events. The probability of occurrence of event A after event B is denoted by $\Pr(A|B)$ and is given by

$$\Pr(A|B) = \frac{\Pr(AB)}{\Pr(B)}. \tag{SP.1.3}$$

Of course, if A and B are independent $\Pr(A|B) = \Pr(A)$.

A *scalar random variable* or *stochastic variable* $x = x(\omega)$ is a real–valued function whose value depends on the outcome ω of a random experiment. The value taken by a random variable is called its realization.

More generally a *vector–valued random variable* or *vector–valued stochastic variable* $x = x(\omega) = [x_1, \dots, x_n]^T$ is a vector of random variables. Random vectors can be characterized by their *probability density function* (PDF) or *probability distribution function* defined as follows.

Definition SP.1.2 (Probability Density Function) – The *probability density function* $\mathrm{p}_x(\cdot)$ of a continuos vector–valued random variable $x = [x_1, \ldots, x_n]^T$ at $x = \xi = [\xi_1, \ldots, \xi_n]^T$ is

$$\mathrm{p}_x(\xi)\,\mathrm{d}\xi = \Pr\left(\cap_{i=1}^{n}(\xi_i \le x_i \le \xi_i + \mathrm{d}\xi_i)\right) \tag{SP.1.4}$$

where the intersection symbol denotes joint events.

A common notation for $\mathrm{p}_x(x)$ is $\mathrm{p}(x)$.

Definition SP.1.3 (Probability Distribution Function) – The *probability distribution function* $\mathrm{P}_x(\cdot)$ of a continuos vector–valued random variable $x = [x_1, \ldots, x_n]^T$ at $x = \xi = [\xi_1, \ldots, \xi_n]^T$ is

$$\mathrm{P}_x(\xi) = \Pr(x \le \xi) = \int_{-\infty}^{\xi_1}\int_{-\infty}^{\xi_2}\cdots\int_{-\infty}^{\xi_n} \mathrm{p}(x)\,\mathrm{d}x. \tag{SP.1.5}$$

The relationship between density and distribution functions is, on the basis of (SP.1.3),

$$\mathrm{p}(x) = \frac{\partial^n \mathrm{P}(x)}{\partial x_1 \partial x_2 \ldots \partial x_n} \tag{SP.1.6}$$

where $\mathrm{P}(x)$ denotes $\mathrm{P}_x(x)$.

Definition SP.1.4 (Expectation of a random vector) – The *expectation* or mean of the random vector x is defined as

$$\mathrm{E}[x] = \bar{x} = \int_{R^n} x\ \mathrm{p}(x)\,\mathrm{d}x \tag{SP.1.7}$$

where $\int_{R^n} x\ \mathrm{p}(x)\,\mathrm{d}x = \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} x\ \mathrm{p}(x)\,\mathrm{d}x$ denotes the integration between $-\infty$ and $+\infty$ of n real variables.

The expectation of a random vector can be considered as the limit, in a probabilistic sense, of the average of the observations when their number tends to infinity. The expectation of a function $g(x)$ of the random variable x can be obtained from the PDF of x

$$\mathrm{E}[g(x)] = \int_{R^n} g(x)\,\mathrm{p}(x)\,\mathrm{d}x. \tag{SP.1.8}$$

Definition SP.1.5 (Covariance matrix of a random vector) – The *covariance matrix* of a random vector x is defined as

$$\Sigma_x = \mathrm{E}\left[(x-\bar{x})(x-\bar{x})^T\right] = \int_{R^n}\big(x-\bar{x}\big)\big(x-\bar{x}\big)^T \mathrm{p}(x)\,\mathrm{d}x. \tag{SP.1.9}$$

The covariance matrix is always non negative definite.

Remark SP.1.1 – In the case of scalar variables, Definition (SP.1.5) leads to the non negative number $\sigma_x^2 = \int_{-\infty}^{+\infty} (x - \bar{x})^2 \, \mathrm{p}(x) \, \mathrm{d}x$ called *variance* of x.

The simultaneous consideration of more random vectors is often necessary. In the case of two random vectors $x \in \mathcal{R}^n$ and $y \in \mathcal{R}^m$, the probability of the joint occurrence of values in given ranges is described by the *joint probability distribution function*

$$\mathrm{P}_{xy}(\xi, \eta) = \Pr\left(\mathrm{x} \leq \xi, \ \mathrm{y} \leq \eta\right) \tag{SP.1.10}$$

with $\xi \in \mathcal{R}^n$ and $\eta \in \mathcal{R}^m$. The corresponding *joint probability density function* is

$$\mathrm{p}(xy) = \frac{\partial \mathrm{P}(x, y)}{\partial x_1 \partial x_2 \ldots \partial x_n \partial y_1 \partial y_2 \ldots \partial y_m} \tag{SP.1.11}$$

where $\mathrm{P}(x, y)$ denotes $\mathrm{P}_{xy}(x, y)$.

Definition SP.1.6 (Statistically independent random vectors) – Two random vectors x and y are defined *statistically independent* if

$$\mathrm{P}(x, y) = \mathrm{P}(x)\, \mathrm{P}(y), \tag{SP.1.12a}$$

$$\mathrm{p}(x, y) = \mathrm{p}(x)\, \mathrm{p}(y). \tag{SP.1.12b}$$

As a consequence of Definition (SP.1.6), if x and y are independent, then

$$\mathrm{E}\left[\, x\, y^T \right] = \mathrm{E}\left[\, x \,\right] \mathrm{E}\left[\, y^T \,\right]. \tag{SP.1.13}$$

Two random vectors are said to be *uncorrelated* if

$$\mathrm{E}\left[\, x\, y^T \right] = \mathrm{E}\left[\, x \,\right] \mathrm{E}\left[\, y^T \,\right] \tag{SP.1.14}$$

so that two independent random vector are uncorrelated but the inverse is not necessarily true. A measure of the correlation between two random vectors is given by their *cross–covariance* matrix defined as follows.

Definition SP.1.7 (Cross–covariance matrix of two random vectors) – The *cross–covariance matrix* Σ_{xy} of the random vectors $x \in \mathcal{R}^n$ and $y \in \mathcal{R}^m$ is defined as

$$\Sigma_{xy} = \mathrm{E}\left[(x - \bar{x})(y - \bar{y})^T\right] = \tag{SP.1.15}$$

$$\int_{\mathcal{R}^n} \int_{\mathcal{R}^m} (x - \bar{x})\left(y - \bar{y}\right)^T \mathrm{p}(x, y) \, \mathrm{d}x \, \mathrm{d}y = \mathrm{E}\left[\, x\, y^T \right] - \mathrm{E}\left[\, \bar{x} \,\right] \mathrm{E}\left[\, \bar{y}^T \,\right].$$

Remark SP.1.2 – In the case of scalar variables x and y, Definition (SP.1.7) leads to a scalar σ_{xy} called *covariance* of x and y.

Remark SP.1.3 – Σ_x and Σ_{xy} have the following structures:

$$\Sigma_x = \begin{bmatrix} \sigma_{x_1}^2 & \sigma_{x_1 x_2} & \cdots & \sigma_{x_1 x_n} \\ \sigma_{x_2 x_1} & \sigma_{x_2}^2 & \cdots & \sigma_{x_2 x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{x_n x_1} & \sigma_{x_n x_2} & \cdots & \sigma_{x_n}^2 \end{bmatrix}, \quad \Sigma_{xy} = \begin{bmatrix} \sigma_{x_1 y_1} & \sigma_{x_1 y_2} & \cdots & \sigma_{x_1 y_m} \\ \sigma_{x_2 y_1} & \sigma_{x_2 y_2} & \cdots & \sigma_{x_2 y_m} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{x_n y_1} & \sigma_{x_n y_2} & \cdots & \sigma_{x_n y_m} \end{bmatrix}.$$

Remark SP.1.4 – If x and y are independed or uncorrelated random vectors then $\Sigma_{xy} = 0$ ($\sigma_{xy} = 0$ in the scalar case).

Definition SP.1.8 (Correlation coefficient of two random variables) – The correlation coefficient ρ of the scalar random variables x and y is the real number

$$\rho_{xy} = \frac{\sigma_{xy}}{\sqrt{\sigma_x^2 \sigma_y^2}}. \tag{SP.1.16}$$

If x and y are uncorrelated, $\rho_{xy} = 0$.

Definition SP.1.9 (Correlation matrix of a random vector) – The correlation matrix $\rho(x)$ of a random vector $x = [x_1, \ldots, x_n]^T$ is defined as

$$\rho(x) = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{bmatrix} \quad \text{where} \quad c_{ij} = \frac{\sigma_{x_i x_j}}{\sqrt{\sigma_{x_i}^2 \sigma_{x_j}^2}}. \tag{SP.1.17}$$

Remark SP.1.5 – It can be noted that

- $\rho(x)$ is a symmetric matrix;
- $c_{ii} = 1, \ \forall i$;
- $c_{ij} \leq 1$ for $i \neq j$ and $c_{ij} = 0$ if x_i and x_j are uncorrelated;
- $\rho(x)$ is non negative definite.

When a set of N observed vectors x_i is available, it is possible to define the *sample mean*

$$\mu_x = \frac{1}{N} \sum_{i=1}^{N} x_i \tag{SP.1.18}$$

and the *sample (co)variance*

$$S_x = \frac{1}{N} \sum_{i=1}^{N} \left(x_i - \mu_x\right)\left(x_i - \mu_x\right)^T . \tag{SP.1.19}$$

SECTIONS	MODULES	QUESTIONS	HOME PAGE
PREV. MODULE	FAQ	TUTOR	NEXT MODULE