Stochastic variables and processes

In this chapter, we're going to examine stochastic variables and stochastic processes. First, we look at one stochastic variable. Then, we examine multiple stochastic variables. Finally, we're going to examine stochastic processes. What kind of properties can such processes have?

1 Stochastic variables and their properties

1.1 Probability distribution and probability density functions

Let's examine a **stochastic variable** $\bar{x}$, also called a **random variable**. A stochastic variable can be seen as a normal variable x, but with uncertainty concerning its value. For example, in $1/3$ of the cases its value may be 2, but in the other $2/3$ of the cases, its value may be 3.

Every random variable has an associated **probability distribution function** $F_{\bar{x}}(x)$, also known as the **cumulative distribution function**. This function is defined as the probability that $\bar{x} \leq x$. So, in formal notation,

$$F_{\bar{x}}(x) = \Pr\{\bar{x} \leq x\}. \tag{1.1}$$

Such a function has several evident properties. We have $F_{\bar{x}}(-\infty) = 0$ and $F_{\bar{x}}(+\infty) = 1$. Also, the function is monotonically increasing. So, if $a \leq b$ then also $F_{\bar{x}}(a) \leq F_{\bar{x}}(b)$.

There is also the **probability density function** $f_{\bar{x}}(x)$, abbreviated as PDF. (Note that PDF does not mean probability distribution function!) The PDF is defined as the derivative of the probability distribution function. So,

$$f_{\bar{x}}(x) = \frac{dF_{\bar{x}}(x)}{dx}. \tag{1.2}$$

It immediately follows that $f_{\bar{x}}(x) \geq 0$. We also have

$$\int_{-\infty}^{\infty} f_{\bar{x}}(x) = 1, \qquad \int_{-\infty}^{b} f_{\bar{x}}(x) = F_{\bar{x}}(b) \quad \text{and} \quad \int_{-a}^{b} f_{\bar{x}}(x) = F_{\bar{x}}(b) - F_{\bar{x}}(a). \tag{1.3}$$

1.2 Moments of distributions

Often, it is very hard, if not impossible, to exactly determine $F_{\bar{x}}(x)$ and $f_{\bar{x}}(x)$. But we may try to determine other quantities. For example, we have defined the i**th moment** of the PDF as

$$m_i = \mathrm{E}\left\{\bar{x}^i\right\} = \int_{-\infty}^{\infty} x^i f_{\bar{x}}(x)\,dx. \tag{1.4}$$

So, we have $m_0 = 1$. Also, $m_1 = \mu_{\bar{x}}$ is the **mean** or **average** of $\bar{x}$. A similar and even more important quantity is the i**th central moment** m_i'. It is defined as

$$m_i' = \mathrm{E}\left\{(\bar{x} - \mu_{\bar{x}})^i\right\} = \int_{-\infty}^{\infty} (x - \mu_{\bar{x}})^i f_{\bar{x}}(x)\,dx. \tag{1.5}$$

Now we have $m_1' = 0$. Also, $m_2' = \sigma_{\bar{x}}^2$ is the **variance** of the stochastic process. The square root of the variance, being $\sigma_{\bar{x}}$, is called the **standard deviation**.

1.3 The normal distribution

There is one very important and common type of distribution. This is the **normal distribution**, also known as the **Gaussian distribution**. Its PDF is given by

$$f_{\bar{x}}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}. \tag{1.6}$$

Quite trivially, this distribution has $\mu = \mu_{\bar{x}}$ as mean and $\sigma^2 = \sigma_{\bar{x}}^2$ as variance.

The **central limit theorem** states that the PDF of a process, caused by a large number of other random processes, approximates the PDF of the Gaussian distribution. Such situations often occur in real life. So, in the remainder of this summary, we will mostly assume that unknown stochastic processes are normally distributed. All that is then left for us to do is to determine the mean $\mu_{\bar{x}}$ and the variance $\sigma_{\bar{x}}^2$.

2 Multiple stochastic variables

2.1 Definitions for multiple random variables

Let's examine the case where we have two random variables $\bar{x}$ and $\bar{y}$. The **joint probability distribution function** $F_{\bar{x}\bar{y}}(x, y)$ is now defined as

$$F_{\bar{x}\bar{y}}(x, y) = \Pr\left\{\bar{x} \leq x \wedge \bar{y} \leq y\right\}, \tag{2.1}$$

where the $\wedge$ operator means 'and'. We thus have $F_{\bar{x}\bar{y}}(-\infty, b) = F_{\bar{x}\bar{y}}(a, -\infty) = 0$, $F_{\bar{x}\bar{y}}(+\infty, +\infty) = 1$, $F_{\bar{x}\bar{y}}(a, +\infty) = F_{\bar{x}}(a)$ and $F_{\bar{x}\bar{y}}(+\infty, b) = F_{\bar{y}}(b)$.

Similarly, the **joint probability density function** (joint PDF) $f_{\bar{x}\bar{y}}(x, y)$ is defined as

$$f_{\bar{x}\bar{y}}(x, y) = \frac{\partial^2 F_{\bar{x}\bar{y}}(x, y)}{\partial x \partial y}. \tag{2.2}$$

The joint PDF has as properties

$$\int_{-\infty}^{a}\int_{-\infty}^{b} f_{\bar{x}\bar{y}}(x, y)\,dx\,dy = F_{\bar{x}\bar{y}}(a, b), \qquad \int_{-\infty}^{\infty} f_{\bar{x}\bar{y}}(x, y)\,dy = f_{\bar{x}}(x) \quad \text{and} \quad \int_{-\infty}^{\infty} f_{\bar{x}\bar{y}}(x, y)\,dx = f_{\bar{y}}(y). \tag{2.3}$$

It may occur that the value of one of the two random variables is known. Let's suppose that it is given that $\bar{y} = y_1$. We can then find the **conditional distribution** of $\bar{x}$ given $\bar{y}$ using

$$f_{\bar{x}}(x|\bar{y} = y_1) = \frac{f_{\bar{x}\bar{y}}(x, y_1)}{f_{\bar{y}}(y_1)}. \tag{2.4}$$

2.2 Moments of joint distributions

The **joint moment** m_{ij} of two random variables $\bar{x}$ and $\bar{y}$ is defined as

$$m_{ij} = \mathrm{E}\left\{\bar{x}^i \bar{y}^j\right\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^i y^j f_{\bar{x}\bar{y}}(x, y)\,dx\,dy. \tag{2.5}$$

The sum $n = i + j$ is called the **order** of the joint moment. It can be noted that $m_{10} = \mu_{\bar{x}}$ and $m_{01} = \mu_{\bar{y}}$. Also, the second order moment m_{11} is called the **average product** $R_{\bar{x}\bar{y}}$.

Of course, there is also a **joint central moment** m'_{ij}. It is defined as

$$m'_{ij} = \mathrm{E}\left\{(\bar{x} - \mu_{\bar{x}})^i (\bar{y} - \mu_{\bar{y}})^j\right\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (\bar{x} - \mu_{\bar{x}})^i (\bar{y} - \mu_{\bar{y}})^j f_{\bar{x}\bar{y}}(x, y)\,dx\,dy. \tag{2.6}$$

The second order joint moments m'_{20} and m'_{02} are equal to the variances of $\bar{x}$ and $\bar{y}$: $m'_{20} = \sigma_{\bar{x}}^2$ and $m'_{02} = \sigma_{\bar{y}}^2$. The other second order joint moment m'_{11} is called the **covariance** $C_{\bar{x}\bar{y}}$. It satisfies

$$m'_{11} = C_{\bar{x}\bar{y}} = R_{\bar{x}\bar{y}} - \mu_{\bar{x}}\mu_{\bar{y}} = m_{11} - m_{10}m_{01}. \tag{2.7}$$

Finally, we can define the **correlation** $K_{\bar{x}\bar{y}}$ as

$$K_{\bar{x}\bar{y}} = \frac{C_{\bar{x}\bar{y}}}{\sigma_{\bar{x}}\sigma_{\bar{y}}} = \frac{m'_{11}}{\sqrt{m_{10}m_{01}}}. \tag{2.8}$$

2.3 Properties of multiple random variables

Let's examine two random variables $\bar{x}$ and $\bar{y}$. We say that $\bar{x}$ and $\bar{y}$ are...

- **orthogonal** if $\mathrm{E}\{\bar{x}\bar{y}\} = 0$.
- **fully correlated** if $\mathrm{E}\{\bar{x}\bar{y}\} = \mu_{\bar{x}}\mu_{\bar{y}} \pm \sigma_{\bar{x}}\sigma_{\bar{y}}$ or, equivalently, if $K_{\bar{x}\bar{y}} = \pm 1$.
- **uncorrelated** if $\mathrm{E}\{\bar{x}\bar{y}\} = \mathrm{E}\{\bar{x}\}\mathrm{E}\{\bar{y}\} = \mu_{\bar{x}}\mu_{\bar{y}}$ or, equivalently, if $K_{\bar{x}\bar{y}} = 0$. We now have $C_{\bar{x}\bar{y}} = 0$ and $\sigma_{\bar{x}+\bar{y}}^2 = \sigma_{\bar{x}}^2 + \sigma_{\bar{y}}^2$. Also, $\bar{x} - \mu_{\bar{x}}$ and $\bar{y} - \mu_{\bar{y}}$ are orthogonal.
- **independent** if $f_{\bar{x}\bar{y}}(x, y) = f_{\bar{x}}(x)f_{\bar{y}}(y)$. This also implies that $\bar{x}$ and $\bar{y}$ are uncorrelated. (Though the converse is not always true.)

3 Stochastic processes

3.1 Basics of stochastic processes

Let's suppose that we are doing an experiment several times. It could occur that the output signal $x(t)$ of the experiment is always the same: it is a **deterministic function**. However, often uncertainty is involved. In this case, the output $x(t)$ is a bit different every time. The output signal $\bar{x}(t)$ is then called a **stochastic function** or a **stochastic process**. At every time τ, the value of $\bar{x}(\tau)$ is a stochastic variable.

Every time we run the experiment, we get a certain output $x(t)$. This output is called a **realization** of the stochastic process $\bar{x}(t)$. The set of all realizations is called the **ensemble** of the process.

There always is a certain chance that a stochastic process $\bar{x}(t)$ results in a certain realization $x(t)$. If this chance is constant in time (that is, the **distribution** of $\bar{x}(t)$ is constant), then we call the process **stationary**. It is very hard, if not impossible, to show that a process is stationary. So it is often simply assumed that stochastic processes are stationary.

3.2 The distribution of stochastic processes

Previously, we talked about a stochastic process $\bar{x}(t)$. Every stochastic process also has a probability distribution and probability density function, which are defined as

$$F_{\bar{x}}(x;t) = \Pr\{\bar{x}(t) \leq x\} \qquad \text{and} \qquad f_{\bar{x}}(x;t) = \frac{\partial F_{\bar{x}}(x;t)}{\partial x}. \tag{3.1}$$

Now let's examine two stochastic processes $\bar{x}(t)$ and $\bar{y}(t)$. The joint distribution of these two processes is defined as

$$F_{\bar{x}\bar{y}}(x, y; t_1, t_2) = \Pr\{\bar{x}(t_1) \leq x \wedge \bar{y}(t_2) \leq y\} \qquad \text{and} \qquad f_{\bar{x}\bar{y}}(x, y; t_1, t_2) = \frac{\partial^2 F_{\bar{x}\bar{y}}(x, y; t_1, t_2)}{\partial x \partial y}. \tag{3.2}$$

Often, it is assumed that the processes $\bar{x}(t)$ and $\bar{y}(t)$ are stationary. This means that not the times t_2 and t_1 theirselves are important, but only the time difference $\tau = t_2 - t_1$. We can thus write

$$F_{\bar{x}\bar{y}}(x, y; \tau) = \Pr\{\bar{x}(t) \leq x \wedge \bar{y}(t+\tau) \leq y\} \qquad \text{and} \qquad f_{\bar{x}\bar{y}}(x, y; \tau) = \frac{\partial^2 F_{\bar{x}\bar{y}}(x, y; \tau)}{\partial x \partial y}. \tag{3.3}$$

3.3 Properties of stochastic processes

Let's suppose that we know the joint distribution function $f_{\bar{x}\bar{y}}(x, y; \tau)$ of two stationary processes $\bar{x}(t)$ and $\bar{y}(t)$. We can now define the **moment function** $m_{ij}(\tau)$ of these processes as

$$m_{ij}(\tau) = \mathrm{E}\left\{\bar{x}(t)^i \bar{y}(t+\tau)^j\right\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^i y^j f_{\bar{x}\bar{y}}(x, y; \tau)\, dx\, dy. \tag{3.4}$$

Similarly, we can define the **central moment function** $m'_{ij}(\tau)$ as

$$m'_{ij}(\tau) = \mathrm{E}\left\{(\bar{x}(t) - \mu_{\bar{x}})^i (\bar{y}(t+\tau) - \mu_{\bar{y}})^j\right\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x - \mu_{\bar{x}})^i (y - \mu_{\bar{y}})^j f_{\bar{x}\bar{y}}(x, y; \tau)\, dx\, dy. \tag{3.5}$$

Several of these moments have special names. The **cross product function** $R_{\bar{x}\bar{y}}(\tau)$ is equal to $m_{11}(\tau)$ and the **cross covariance function** $C_{\bar{x}\bar{y}}(\tau)$ is equal to $m'_{11}(\tau)$. Also, the **cross correlation function** $K_{\bar{x}\bar{y}}(\tau)$ is defined as

$$K_{\bar{x}\bar{y}}(\tau) = \frac{C_{\bar{x}\bar{y}}(\tau)}{\sigma_{\bar{x}}\sigma_{\bar{y}}}. \tag{3.6}$$

Next to these three cross-functions, we also have three auto-functions. They are the **auto product function** $R_{\bar{x}\bar{x}}(\tau)$, the **auto covariance function** $C_{\bar{x}\bar{x}}(\tau)$ and the **auto correlation function** $K_{\bar{x}\bar{x}}(\tau)$. They are defined identically as the cross-functions, with the only difference that we substitute $\bar{y}(t+\tau)$ by $\bar{x}(t+\tau)$.

The cross correlation function $K_{\bar{x}\bar{y}}(\tau)$ is an indication of the correlation between two stochastic processes $\bar{x}(t)$ and $\bar{y}(t+\tau)$. But you might be wondering, what is the auto correlation function $K_{\bar{x}\bar{x}}(\tau)$ good for? Well, it gives an indication of how much the value of $\bar{x}(t+\tau)$ at time $t+\tau$ depends on the value of $\bar{x}(t)$ at time t. We'll examine how this works.

First, we can note that $K_{\bar{x}\bar{x}}(\tau)$ gives the correlation between the random variables $\bar{x}(t)$ and $\bar{x}(t+\tau)$. Generally, if τ becomes big, then the signals $\bar{x}(t)$ and $\bar{x}(t+\tau)$ will be uncorrelated: $K_{\bar{x}\bar{x}}(\tau)$ will go to zero. But for small (absolute) values of τ, the signals $\bar{x}(t)$ and $\bar{x}(t+\tau)$ are correlated a lot. (Especially if $\tau = 0$, because $K_{\bar{x}\bar{x}}(0) = 1$.) How fast $K_{\bar{x}\bar{x}}(\tau)$ goes to zero now determines how fast the signal $\bar{x}(t)$ loses its influence on $\bar{x}(t+\tau)$.

3.4 Ergodic processes

Let's examine a stochastic process $\bar{x}(t)$. We can examine all possible realizations $x(t)$ of this process. If we then take the (weighted) average of these realization values, we will find the **ensemble average** $\mu_{\bar{x}}(t)$ at time t.

However, in real life, we don't know all realizations of a stochastic process $\bar{x}(t)$. All we have is one realization $x(t)$. The average value μ_x of this realization is called the **time average**. An **ergodic process** is now defined as a process in which these averages are equal. Or, more formally, it is defined as a process for which, for every function $g(x)$, we have

$$\mathrm{E}\{g(\bar{x}(t))\} = \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} g(\bar{x}(t))\, dt. \tag{3.7}$$

In real life, because we only have one realization, we often assume that a process is ergodic. This implies that our single realization is representative for the entire process. In other words, we can use it to derive the process properties. For this, we can use

$$\mu_{\bar{x}} = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} x(t)\,dt, \qquad \sigma_{\bar{x}}^2 = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} (x(t) - \mu_{\bar{x}})^2\,dt = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} x(t)^2\,dt - \mu_{\bar{x}}^2, \quad (3.8)$$

$$R_{\bar{x}\bar{y}}(\tau) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} x(t)y(t+\tau)\,dt, \quad \text{and} \quad C_{\bar{x}\bar{y}}(\tau) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} x(t)y(t+\tau)\,dt - \mu_{\bar{x}}\mu_{\bar{y}}. \quad (3.9)$$

3.5 White noise

One special type of a stochastic process is **white noise** $\bar{w}(t)$. White noise has a zero mean: $\mu_{\bar{w}} = 0$. Next to this, the value of $\bar{w}(t)$ has absolutely no influence on the value of $\bar{w}(t+\tau)$ with $\tau \neq 0$. Thus, $C_{\bar{w}\bar{w}}(\tau) = 0$ for $\tau \neq 0$. To be more precise, the auto covariance function of white noise is defined as

$$C_{\bar{w}\bar{w}}(\tau) = W\delta(\tau), \quad (3.10)$$

where W is called the **intensity** of the white noise and $\delta(\tau)$ is the **Dirac delta function**. However, white noise is only a convenient theoretical trick. In real life, white noise as defined above does not occur. To show this, we can look at the variance $\sigma_{\bar{w}}^2$ of $\bar{w}(t)$. It is given by $\sigma_{\bar{w}}^2 = C_{\bar{w}\bar{w}}(0) = \infty$. This is physically of course impossible. So instead, in real life, we usually call a stochastic process $\bar{x}(t)$ white noise if $C_{\bar{x}\bar{x}}(\tau) \approx 0$ for $|\tau| > \epsilon$ for some sufficiently small ϵ.