

---

## Randomness

We describe as “random” effects that are too complex to precisely analyze in practice, or simply unknown:

- physical noise: thermal, mechanical, acoustic, radio/radar
- somebody’s decisions made from data unknown (to us): aircraft pilotage, human voice

But we know *something* about the constraints: system bandwidth, physical dependencies, vocal tract properties.

We describe:

**value constraints** as probability

**dependencies** as conditional probability  $\longrightarrow$  correlation

---

## Description of a random variable

$\xi$  is a random variable taking some value  $x \in X$ ; for each value  $x$  there is some probability that  $\xi = x$ .

We may imagine  $\xi$  as the ensemble of all possible values together with their probabilities, or as the “set of all possible experiments”.

**CDF** stochastic variable  $\xi$  is described by  $F_\xi(x)$  – the probability that  $\xi \leq x$ ;  $F_\xi(x)$  is a Cumulative Density Function (or Cumulative Distribution Function)

**PDF** Probability Density Function is more intuitive, defined as  $f_\xi(x) = \frac{dF_\xi(x)}{dx}$

**expectation** or probabilistic mean, or mean value:  $\mu_\xi = E[\xi] = \int_{-\infty}^{\infty} x f_\xi(x) dx$

**mean square (MS) value**  $P_\xi = E[\xi^2(t)] = \int_{-\infty}^{\infty} x^2 f_\xi(x) dx$

**variance** (mean power of variable component)  $\sigma_\xi^2 = E[(\xi - \mu_\xi)^2] = \int_{-\infty}^{\infty} [x - \mu_\xi(t)]^2 f_\xi(x) dx$

**covariance** is a measure of relation between *two* variables  $\xi$  and  $\eta$ :  $cov(\xi, \eta) = E[\xi \cdot \eta]$

---

## Stochastic signal basics (discrete time)

$\xi[n]$  – a sequence of stochastic variables  $\xi(n)$  (as a DT signal  $x[n]$  is a sequence of numbers  $x(n)$ )

**signal or process:**  $\xi[n]$  – a set of all possible realizations  $x[n]$

**realization:**  $x[n]$  one sequence, being particular member of the set  $\xi[n]$

**process value** at the moment  $n_1$ ,  $\xi(n_1)$  is a stochastic variable, described with its PDF  $f_\xi(x(n_1); n_1)$

→ for the full description of  $\xi(n)$  we need all the possible multidimensional (joint) PDF's  $f_\xi(x(n_1), x(n_2), \dots; n_1, n_2, \dots)$

**practical view:** we narrow our interest to the two-dimensional PDF  $f_\xi(x(n_1), x(n_2); n_1, n_2)$  to be able to tell the relation between the process values at two points in time.

For the stochastic (random) signal, we use the same description with expectation (called *mean*, or precisely *probabilistic mean*), MS value, variance, covariance - but they are in general dependent on time:  $\mu_\xi(n)$ , etc.

---

## Complex signal

Complex signal

$$\xi[n] = \xi_R[n] + j \xi_I[n]$$

**Example:** complex filtering (convolution of complex signals)

$$h[n] = h_R[n] + j h_I[n]$$
$$\eta[n] = h[n] * \xi[n] \longrightarrow \eta(n) = \sum_{m=0}^n h(m) \xi(n-m)$$

$$\eta_R[n] = h_R[n] * \xi_R[n] - h_I[n] * \xi_I[n]$$
$$\eta_I[n] = h_I[n] * \xi_R[n] + h_R[n] * \xi_I[n]$$

---

## We describe complex random signals with....

- *Mean value*

$$\mu_{\xi}(n) = E [\xi(n)] = E [\xi_R(n)] + jE [\xi_I(n)] = \mu_{\xi_R}(n) + j\mu_{\xi_I}(n)$$

- *MS value* (mean power)

$$P_{\xi}(n) = E [\xi(n) \xi^*(n)] = E [|\xi(n)|^2]$$

- *Variance*

$$\sigma_{\xi}^2(n) = E \left\{ [\xi(n) - \mu_{\xi}(n)] [\xi^*(n) - \mu_{\xi}^*(n)] \right\} = E [|\xi(n)|^2 - |\mu_{\xi}(n)|^2]$$

- *Autocorrelation* is a measure of dependency between signal values in different time instants

$$R_{\xi\xi}(n_1, n_2) = E [\xi^*(n_1) \xi(n_2)]$$

- *Autocovariance*

$$C_{\xi\xi}(n_1, n_2) = E \left\{ [\xi^*(n_1) - \mu_{\xi}^*(n_1)] [\xi(n_2) - \mu_{\xi}(n_2)] \right\}$$


---

## Stationarity and ergodicity

Signal is **stationary** when it fulfills the following

$$\begin{aligned}\mu_{\xi}(n) &= \mu_{\xi} = \text{const} \\ R_{\xi\xi}(n_1, n_2) &= R_{\xi\xi}(m), \quad m = n_2 - n_1\end{aligned}$$

Time-domain mean is defined as:  $\langle x[n] \rangle = \lim_{N \rightarrow \infty} \frac{1}{2N+1} \sum_{n=-N}^N x(n)$

Time-domain correlation is defined as:  $\Psi_{xx}(m) = \lim_{N \rightarrow \infty} \frac{1}{2N+1} \sum_{n=-N}^N x^*(n)x(n+m)$

Signal is **ergodic** when it fulfills the following

$$\begin{aligned}\langle x[n] \rangle &= \mu_{\xi} = \text{const} \\ \Psi_{xx}(m) &= R_{\xi\xi}(m), \quad m = n_2 - n_1\end{aligned}$$

In other words, *ergodicity* means that we can draw conclusions on *probabilistic* mean, variance, and autocorrelation from *time-domain* mean, mean power, autocorrelation. We usually have to assume the signal is ergodic, no simple test exists ....

---

## Power spectrum density (PSD)

For a stationary  $\xi[n]$ ,  $C_{\xi\xi}(n_1, n_2) = C_{\xi\xi}(m)$ ,  $m = n_2 - n_1$ .

$$C_{\xi\xi}(m) = E \left\{ [\xi^*(n) - \mu_\xi^*] [\xi(n+m) - \mu_\xi] \right\}$$

If our signal  $\xi[n]$  is zero-mean:  $\mu_\xi = 0$  then  $C_{\xi\xi}(m) = R_{\xi\xi}(m)$

(if not, use  $\xi_1[n] = \xi[n] - \mu_\xi$ ).

Power spectrum density of a stationary discrete signal  $\xi[n]$  (MS convergent if  $\sigma_\xi^2$  is bounded)

$$S_{\xi\xi}(\theta) = \sum_{m=-\infty}^{\infty} C_{\xi\xi}(m) e^{-jm\theta}$$

- Periodic (over  $2\pi$ )
  - (if  $\xi[n]$  is real)  $\rightarrow S_{\xi\xi}(\theta) \geq 0$ , symmetric
  - $C_{\xi\xi}(m) = \frac{1}{2\pi} \int_{-\pi}^{\pi} S_{\xi\xi}(\theta) e^{jm\theta} d\theta$
-

## Autocovariance and PSD estimation

$\xi[n]$  (stationary, ergodic)  $\longrightarrow$  estimation of properties from  $N$  (finite number) of samples of  $x[n]$ .

- mean value estimate

$$\hat{\mu}_\xi = \frac{1}{N} \sum_{n=0}^{N-1} x(n)$$

- Variance estimate

$$\hat{\sigma}_\xi^2 = \frac{1}{N} \sum_{n=0}^{N-1} [x^*(n) - \hat{\mu}_\xi^*] [x(n) - \hat{\mu}_\xi]$$

- Autocovariance estimate (equal to autocorrelation with  $\mu_\xi = 0$ )

$$\hat{R}_{\xi\xi}(m) = \frac{1}{N} \sum_{n=0}^{N-|m|-1} x^*(n)x(n+m), \quad -(N-1) \leq m \leq N-1 \quad (\text{or, maybe } \frac{1}{N-|m|}?)$$

- PSD estimate

$$\hat{S}_{\xi\xi}(\theta) = \sum_{m=-(N-1)}^{N-1} \hat{R}_{\xi\xi}(m) e^{-jm\theta}$$



## Estimation accuracy

The actual values  $\mu_\xi$ ,  $\xi[n]$ ,  $\hat{\mu}_\xi$ ,  $R_{\xi\xi}(m)$ ,  $S_{\xi\xi}(\theta)$  are constant (=not random)

From different **realizations**  $x[n]$  we obtain different estimates.

Estimates  $\hat{\sigma}_\xi^2$ ,  $\hat{R}_{\xi\xi}(m)$ ,  $\hat{S}_{\xi\xi}(\theta)$  are random  $\longrightarrow$  How to measure the accuracy of estimate?

**bias**  $B = \alpha - E[\hat{\alpha}]$

**variance**  $var[\hat{\alpha}] = \sigma_{\hat{\alpha}}^2 = E \{ [\hat{\alpha}^* - E(\hat{\alpha})^*] [\alpha - E(\hat{\alpha})] \}$

**MS error**  $E[|\hat{\alpha} - \alpha|^2] = B^2 + \sigma_{\hat{\alpha}}^2$

**consistency**  $\lim_{N \rightarrow \infty} var[\hat{\alpha}] \rightarrow 0$  and  $\lim_{N \rightarrow \infty} B[\hat{\alpha}] \rightarrow 0$

If  $\xi[n]$  is stationary and gaussian ...

- mean value estimate  $\hat{\mu}_\xi = \frac{1}{N} \sum_{n=0}^{N-1} x(n) \longrightarrow$  unbiased, with variance  $\sigma_\xi^2/N$
- variance estimate  $\hat{\sigma}_\xi^2 = \frac{1}{N} \sum_{n=0}^{N-1} [x^*(n) - \hat{\mu}_\xi^*] [x(n) - \hat{\mu}_\xi]$   
 $\longrightarrow$  bias  $B[\hat{\sigma}_\xi^2] = \sigma_\xi^2/N$ , variance  $var[\hat{\sigma}_\xi^2] \sim 1/N$  (consistent)

## Autocovariance and PSD estimate properties

$\hat{R}_{\xi\xi}(m) = \frac{1}{N} \sum_{n=0}^{N-|m|-1} x^*(n)x(n+m)$  is biased:

$$E[\hat{R}_{\xi\xi}(m)] = \frac{N-|m|}{N} R_{\xi\xi}(m)$$

$$\text{var}[\hat{R}_{\xi\xi}(m)] \approx \frac{1}{N} \sum_{r=-\infty}^{\infty} \left[ R_{\xi\xi}^2(r) + R_{\xi\xi}(r+m)R_{\xi\xi}(r-m) \right], \quad N \gg m$$

$\hat{S}_{\xi\xi}(\theta) = \sum_{m=-(N-1)}^{N-1} \hat{R}_{\xi\xi}(m)e^{-jm\theta}$ :

$$E[\hat{S}_{\xi\xi}(\theta)] = \sum_{m=-(N-1)}^{N-1} \frac{N-|m|}{N} R_{\xi\xi}(m)e^{-jm\theta}$$

$$\text{var}[\hat{S}_{\xi\xi}(\theta)] = S_{\xi\xi}^2(\theta) \left\{ 1 + \left[ \frac{\sin N\theta}{N \sin \theta} \right]^2 \right\} \quad \text{very large, estimate not consistent!}$$

## Periodogram

Periodogram is a method to estimate PSD that is faster!

As the ACF is estimated from the convolution

$$\hat{R}_{\xi\xi}[m] = \begin{cases} \frac{1}{N} x^*[m] * x[-m], & |m| \leq N - 1 \\ 0, & |m| > N - 1 \end{cases}$$

we may rewrite  $\hat{S}_{\xi\xi}(\theta)$  using transforms (in the following  $x_1[m] = x[-m]$ )

$$\hat{S}_{\xi\xi}(\theta) = \frac{1}{N} \cdot X^*(e^{j\theta}) \cdot X_1(e^{j\theta})$$

$$X_1(e^{j\theta}) = \sum_{n=-(N-1)}^0 x_1(n)e^{-jn\theta} = \sum_{n=0}^{N-1} x_1(-n)e^{jn\theta} = \sum_{n=0}^{N-1} [x^*(n)e^{-jn\theta}]^* = [X^*(e^{j\theta})]^*$$

finally

$$\hat{S}_{\xi\xi}(\theta) = \frac{1}{N} \cdot |X(e^{j\theta})|^2$$

Further: we can transform N-sample sections of  $x^*[n]$ ,  $x_1[n]$  and then average periodograms, reducing variance.

---

## Practical implementations of periodogram

- Choose FFT length to avoid cyclic effects
- Average  $K$  segments (of length  $M = N/K$ ) to reduce variance at the cost of bias (*Bartlett procedure*).
- Add overlapping of segments and use non-rectangular window (*Welch procedure*).

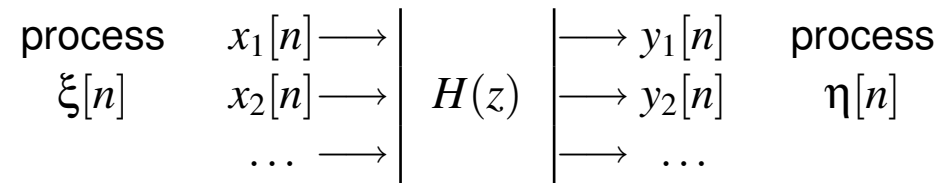
$$\hat{S}_W^i(\theta) = \frac{1}{MF} \left| \sum_{n=0}^{M-1} x^i(n) g(n) e^{-jn\theta} \right|^2, \quad (i = 1, 2, \dots, K \text{ is a segment number})$$

$$F = \frac{1}{M} \sum_{n=0}^{M-1} g^2(n) \quad (\text{energetic normalizing factor})$$

$$\hat{S}_{W\xi\xi}(\theta) = \frac{1}{K} \sum_{i=1}^K \hat{S}_W^i(\theta)$$

---

## Filtering of random signals



For a stationary  $\xi[n]$ :

- mean value:

$$\mu_\eta = \mu_\xi \sum_{n=-\infty}^{\infty} h(n) = \mu_\xi H(e^{j\theta})|_{\theta=0} \quad (1)$$

- autocorrelation

$$R_{\eta\eta}(m) = \sum_{i=-\infty}^{\infty} R_{\xi\xi}(m-i) v(i) \quad \text{where} \quad v(i) = \sum_{k=-\infty}^{\infty} h(k) h(k+i) \quad (2)$$

- power spectrum density

$$S_{\eta\eta}(\theta) = S_{\xi\xi}(\theta) |H(e^{j\theta})|^2 \quad (3)$$

---

## Applications

- Analysis of AD conversion errors
  - Analysis of arithmetic errors in filters
  - Signal modelling → compression (LPC):
    - on the compression side, a filter is tuned so that noise passed through it has the properties of the signal being compressed;
    - only filter coefficients are transmitted;
    - signal is reconstructed from noise passed through a filter
  - System modelling → identification
  - Signal detection → matched filter (presence, time of arrival);
    - a filter matched to signal  $x(n)$  has the impulse response  $h(n) = x(M - n)$ ; such a filter maximizes the S/N ratio if the noise is gaussian and white
    - if the noise is not white, we may *whiten* it with another filter (*whitening filter*): we may *model* the noise as if it was white, but then filtered with a filter with transfer function  $G(z)$ ; then the whitening filter will be  $1/G(z)$ ....
-