# Probability Structure of Non-Stationary Data

Non-stationary probability density $p(x, t_1)$

of a non-stationary random process $\{x(t)\}$

$$p(x, t_1) = \lim_{\Delta x \to 0} \frac{\text{Prob}\left[x < x(t_1) \leq x + \Delta x\right]}{\Delta x}$$

For any $t$ $p(x,t)$ has these basic properties

$$\int_0^{\infty} p(x,t)\, dx = \underline{1}$$

$$\mu_x(t) = E\left[x(t)\right] = \int_{-\infty}^{+\infty} x \cdot p(x,t)\, dx$$

$$\Psi_x^2(t) = E\left[x^2(t)\right] = \int x^2 \cdot p(x,t)\, dx$$

and

$$\sigma_x^2(t) = E\left[\left(x(t) - \mu_x(t)\right)^2\right] = \Psi_x^2(t) - \mu_x^2(t)$$

These are the same equations as for stationary processes for which $p(x,t) = p(x)$ only, that is, all statistical ideas and concept now become time-dependent

If the non-stationary random process $\{x(t)\}$ is Gaussian at $t = t_1$ then

$$p(x, t_1) = \frac{1}{\sigma_x(t_1)\sqrt{2\pi}} \cdot \exp - \frac{[x - \mu_x(t_1)]^2}{2\,\sigma_x^2(t_1)}$$

So the mean $\mu_x(t_1)$ at time $t_1$ and variance $\sigma_x^2(t_1)$ at time $t_1$ may become useful

Higher order probability functions are defined similar, for example, second order non-stationary pdf of $x(t_1)$ and $x(t_2)$ are

$$p(x_1, t_1; x_2, t_2) = \lim_{\substack{\Delta x_1 \to 0 \\ \Delta x_2 \to 0}} \frac{\text{Prob}\,[(x_1 < x_1(t) \leq x_1 + \Delta x_1) \text{ and } (x_2 < x_2(t) \leq x_2 + \Delta x_2)]}{\Delta x_1 \, \Delta x_2}$$

and

$$R_{xx}(t_1, t_2) = E[x(t_1)\, x(t_2)] = \iint x_1 \cdot x_2 \cdot p(x_1, t_1; x_2, t_2)\, dx_1\, dx_2$$

Example (to show the complexity)

Time Averaged pdf :

Consider

$$x(t) = \begin{cases} x_1(t) & \text{for } 0 \le t \le T/2 \\ x_2(t) & \text{for } T/2 \le t \le T \end{cases}$$

with Gaussian pdf, e.g.,

$$p(x,t) = \frac{1}{\sqrt{2\pi}} \cdot \begin{cases} \frac{1}{\sigma_1} \exp\left(-x^2 / 2\sigma_1^2\right) & |t| \le T/2 \\ \frac{1}{\sigma_2} \exp\left(-x^2 / 2\sigma_2^2\right) & |t| \ge T/2 \end{cases}$$

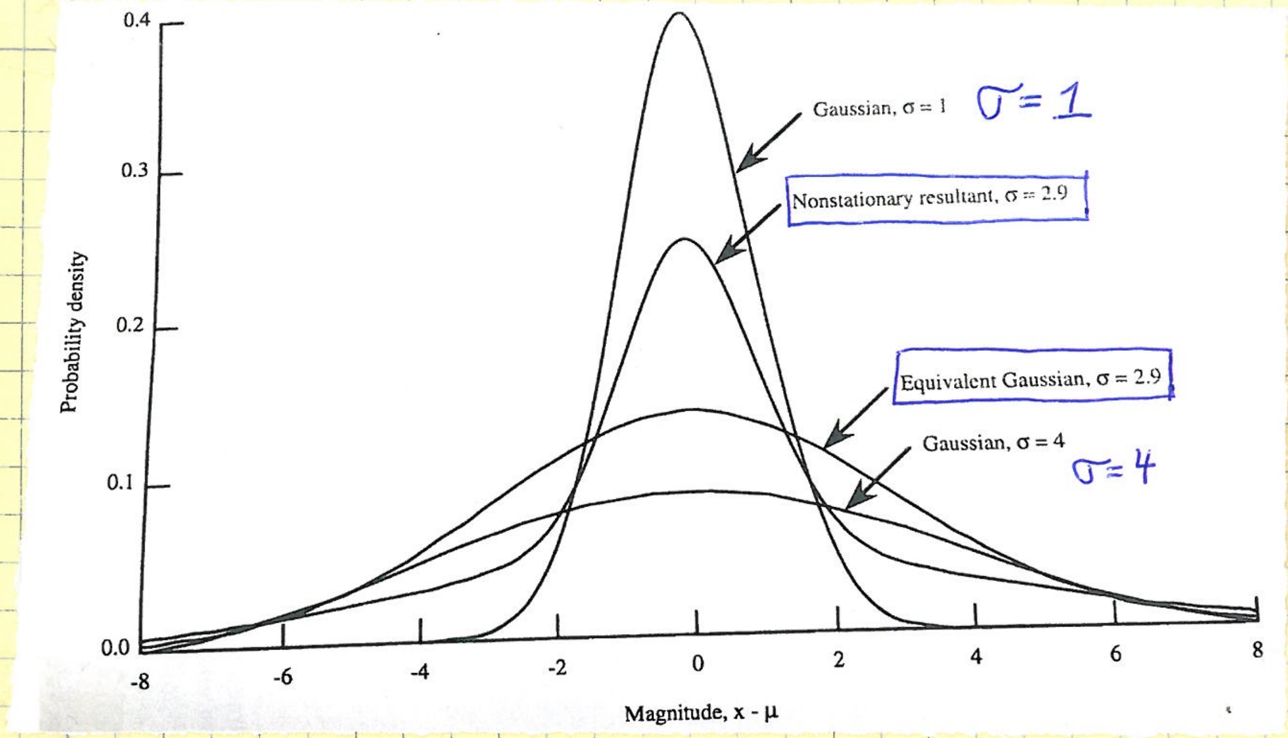
If we ignore this non-stationarity in the computation of a pdf for $x(t)$ $t \in [0, T]$ then we will get an estimated pdf $\hat{p}$ (call this the histogram)

$$\hat{p}(x) = \frac{1}{2\sqrt{2\pi}} \left[ \frac{1}{\sigma_1} \exp[-x^2/\sigma_1^2] + \frac{1}{\sigma_2} \exp[-x^2/\sigma_2^2] \right]$$

Let $\sigma_1^2 = 1$ and $\sigma_2^2 = 16$

$\Downarrow$

$$\hat{\sigma}^2 = (\sigma_1^2 + \sigma_2^2)/2 = 8.5 \qquad \hat{\sigma} \approx 2.9$$




⚠ Two Gaussian distributions do <u>NOT</u> average into a Gaussian distribution.

## Non-stationary Mean Values

$$\hat{\mu}_x(t) = \frac{1}{N}\sum_{i=1}^{N} x_i(t)$$ ensemble average over $N$ ensembles

$$E[\mu_x(t)] = \frac{1}{N}\sum_{i=1}^{N} E[x_i(t)] = \mu_x(t)$$

where $\mu_x(t) = E[x_i(t)]$ is the true mean value of nonstationary process at time $t$

What to do in practice when we only have a single record, that is, $N = 1$?

Answer : Interpret $\overset{\text{output from}}{\vee}$ a suitable low-pass filter as the time-variable mean $\mu_x(t)$

Think of the non-stationary process $\{x(t)\}$ as

$$\{x(t)\} = \underbrace{a(t)}_{\text{deterministic function}} + \underbrace{\{u(t)\}}_{\text{stationary random process with zero Mean}}$$

Thus

$$E[x(t)] = E[a(t)+u(t)] = E[a(t)] + \cancel{E[u(t)]}^{\,0}$$
$$= a(t)$$

The estimate is biased, however,

$$\hat{\mu}_x(t) = \int_{t-T/2}^{t+T/2} x(t)\,dt = \int a(t) + u(t)\,dt$$

where $T$ is a short averaging intervall related to low-pass filter cut-off frequency, say

Now

$$E[\hat{\mu}_x] = \int E[a(t)] + E[u(t)]\,dt$$

$$= \int_{t-T/2}^{t+T/2} a(t)\,dt \neq a(t)$$

The bias error can be shown to be $\frac{T^2}{24} \cdot \frac{d^2 a(t)}{dt^2} \propto T^2$

## Non-Stationary Mean Square Values

For $N$ samples or ensembles $x_i(t)$ $\quad t \in [0, T]$, $\quad i = 1, 2, \ldots N$

We ~~form~~ estimate the mean square average

$$\hat{\psi}_x^2(t) = \frac{1}{N} \sum_{i=1}^{N} x_i^2(t)$$

$$E[\hat{\psi}_x^2(t)] = \frac{1}{N} \sum_{i=1}^{N} E[x_i^2(t)] = \psi_x^2(t)$$

where

$$\psi_x^2(t) = E[x_i^2(t)] = \mu_x(t) + \sigma_x^2(t)$$

is the (generally unknown) true mean square value of the non-stationary process $\{x(t)\}$ at time $t$.

What to do in practice where we only have $N = 1$ ?

Assume we can represent our process as

$$\{x(t)\} = \underbrace{a(t)}_{\text{deterministic function}} \cdot \underbrace{\{u(t)\}}_{\text{stationary random process of variance 1 and zero mean}}$$

Then

$$E[x^2(t)] = \psi_x^2(t) = E[a^2(t)\, u^2(t)] = a^2(t)\, \overset{1}{\cancel{E[u^2(t)]}}$$

$$= a^2(t)$$

If variations of $a(t)$ are slow relative to the lowest frequencies of $\{u(t)\}$ then $a^2(t)$ can be extracted via a low-pass filtering operation similar to that used above for extraction of a time-dependent mean $\mu_x(t)$.

Another way to think about this is a "short-term" averaging operation, e.g.

$$\hat{\psi}_x^2(t) = \int_{t-T/2}^{t+T/2} a^2(t)\,dt$$

which indicates, as before, that this is a biased estimator with an bias error $= \dfrac{T^2}{24} \dfrac{1}{a^2(t)} \dfrac{d^2[a^2(t)]}{dt^2} \propto T^2$

## Correlation Structures of Non stationary Data

Consider a pair of non-stationary processes $\{x(t)\}$ and $\{y(t)\}$. The mean values at any fixed time $t$ are

$$\mu_x(t) = E[x(t)] \quad \text{and} \quad \mu_y(t) = E[y(t)]$$

The correlation functions at any pair of fixed times $t_1$ and $t_2$ are <u>defined</u> by their expected values

$$\left.\begin{aligned} R_{xx}(t_1,t_2) &= E[x(t_1)\, x(t_2)] \\ R_{yy}(t_1,t_2) &= \\ R_{xy}(t_1,t_2) &= E[x(t_1)\, y(t_2)] \end{aligned}\right\}$$

for stationary data only $(t_2 - t_1) = \tau$ mattered.

It follows that

$$R_{xx}(t_1,t_2) = R_{xx}(t_2,t_1)$$
$$R_{yy}(t_1,t_2) = R_{yy}(t_2,t_1)$$
$$R_{xy}(t_1,t_2) = R_{xy}(t_2,t_1)$$

and

$$|R_{xy}(t_1,t_2)|^2 \leq R_{xx}(t_1,t_1) \cdot R_{yy}(t_2,t_2)$$