An Invitation to Random Functions:

A hypothetic lecture on a hypothetic Milagro Seminar.*

Roman Fleysher Lazar Fleysher

May 7, 1999

I can not make my peace with the randomness doctrine:
I can not abide the notion of purposelessness and blind
chance in nature. And yet, I do not know what to put in
its place for the quieting of my mind.
Lewis Thomas
On the uncertainty of science.

1 Introduction.

The notion of random variables can be generalized to the notion of random functions. We will start with an introduction and comparison of the characteristics of random functions with the ones of random variables.

A random variable X is a variable which, as the result of a trial, can take one, certain but unknown value x. *Example.* In a series of trials, the random variable X took the following values: $1, 5, 7, 3, 2, 5, 4$. Full description of a random variable is provided by the probability distribution function $\rho(x)$ which tells us "how often" one might observe a particular realization x of the random variable X. Probability distribution function, in turn, can be characterized by its mean (m_x), dispersion (D_x) and other higher moments.

A generalization of the notion of the random variable is the notion of the random function. As the result of a trial, random function $X(t)$ can take

*This paper does not constitute the full lecture notes, on contrary, it reflects only a fraction of the presented material and is truly an invitation.

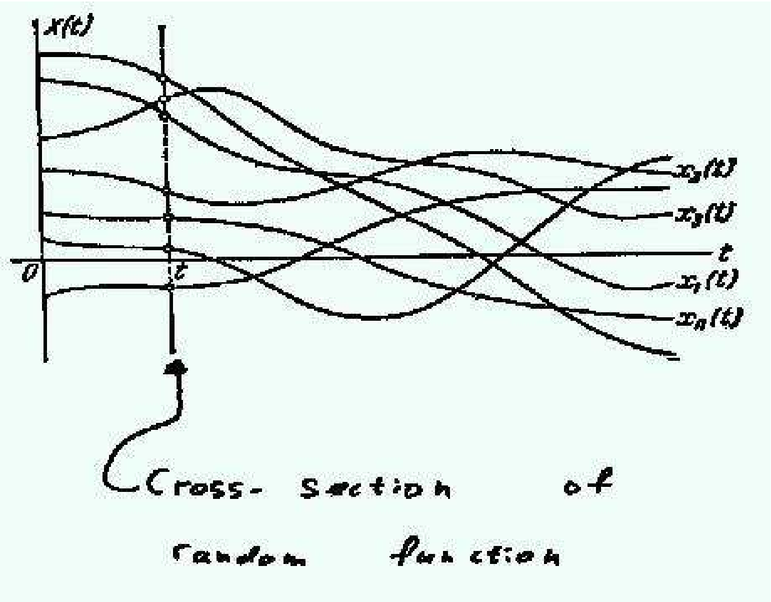

Figure 1:

one certain but unknown realization: regular function $x(t)$. If the argument of a function is time, then the term *"stochastic process"* is usually used. If there are several arguments, then the random function is usually referred as the *"random field"*. Figure 1 provides an example of several realizations of a random function. The cross-section of this family at some $t = const$ is readily interpreted as a random variable $X(t = conts)$. Thus, the random function can be interpreted as a collection of random variables. Since a random variable is described by the probability distribution function, its generalization $\rho_1(x_1, t_1)$ will describe the random function (or more precisely, its cross-section at $t_1 = const$). Joint distribution functions $\rho_2(x_1, t_1; x_2, t_2)$, $\rho_3(x_1, t_1; x_2, t_2; x_3, t_3)$,... will provide additional information about the process. [1]

The generalization of the mean will be the mean function $m_x(t)$ constructed in such a way that at each $t = const$ it is equal to the mean of the random variable $X(t = const)$. Dispersion will also become a function:

[1]If it is enough to specify $\rho_2(x_1, t_1; x_2, t_2)$ to describe the random process, then it is called a process *without consequences*. Physically, it means that the future is determined by the present and does not depend on the past.

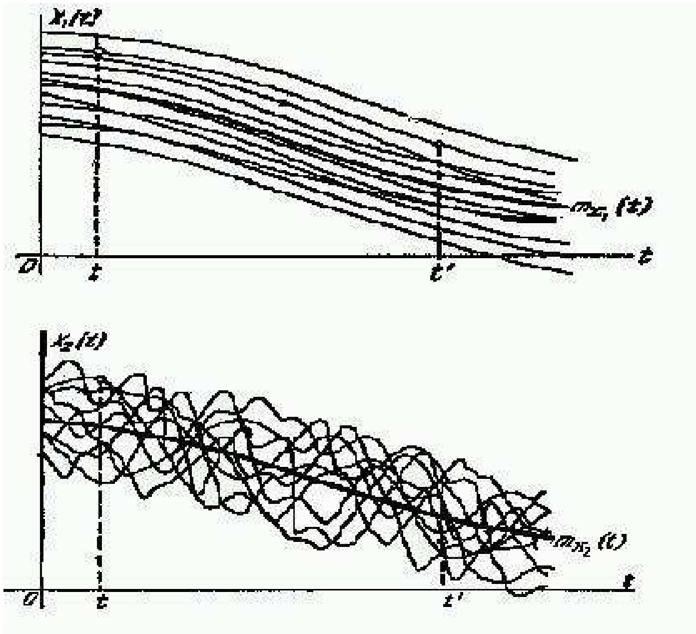

Figure 2:

$$m_x(t) = \langle X(t) \rangle$$

$$D_x(t) = \left\langle (X(t) - m_x(t))^2 \right\rangle$$

Expectation value $m_x(t)$ and dispersion $D_x(t)$ characterize the "one-point" distribution function $\rho_1(x_1, t_1)$. At the lowest level, "two-point" distribution function $\rho_2(x_1, t_1; x_2, t_2)$ can be characterized by the autocorrelation function $K_x(t, t')$, the analogue of the correlation matrix of two random variables:

$$K_x(t, t') = \langle X(t)X(t') \rangle - m_x(t)m_x(t')$$

$$K_x(t, t) = D_x(t)$$

$$K_x(t, t') = K_x(t', t)$$

Thus, $m_x(t)$ and $K_x(t, t')$ provide the same level of description (second order) of a random function as the mean and dispersion for a random variable.

Figure 2 showers two different random functions with similar means and dispersions. The difference is clearly due to correlation properties.

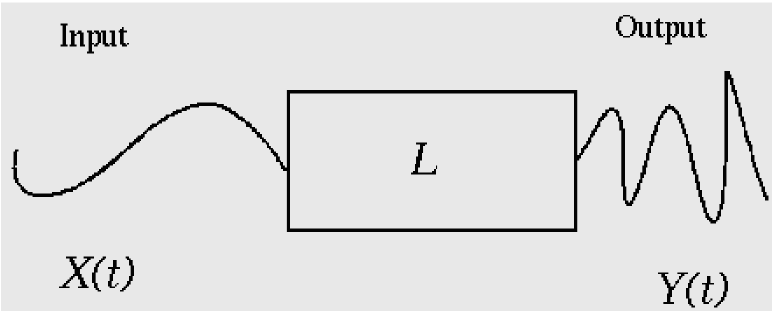

Figure 3:

2 Linear Transformations of Random Functions.

Let L be a linear operator, then it is interchangeable with the averaging operation $\langle\rangle$. If the input of a device, described by the operator L, is a random function $X(t)$, then the output will generally be some other random function $Y(t)$ (figure 3):

$$Y(t) = LX(t)$$

Let us determine the characteristics of $Y(t)$ assuming that characteristics of $X(t)$ are known. The expectation value, by linearity of L:

$$m_y(t) = \langle LX(t)\rangle = L\,\langle X(t)\rangle = Lm_x(t)$$

Example. If $L = \frac{d}{dt}$ — differential operator, then $m_y(t) = \frac{d}{dt}m_x(t)$.
Autocorrelation function:

$$K_y(t,t') = \langle Y(t)Y(t')\rangle - m_y(t)m_y(t')$$

$$= \langle L_tX(t)L_{t'}X(t')\rangle - L_tm_x(t)L_{t'}m_x(t')$$

$$= L_tL_{t'}\,\langle X(t)X(t')\rangle - L_tL_{t'}m_x(t)m_x(t')$$

$$K_y(t,t') = L_tL_{t'}K_x(t,t')$$

Here, subscript t or t' denote the variable with respect to which the given operator L acts.

Example. $L = \frac{d}{dt}$, therefore

$$K_y(t,t') = \frac{\partial}{\partial t}\frac{\partial}{\partial t'}K_x(t,t')$$

3 Example of a Linear Filter.

Let us assume that the input of a differentiating machine is a random function $X(t)$ with expectation $\sin(t)$ and correlation function $D_x e^{-\alpha(t-t')^2}$. What are the characteristics of the output random function $Y(t)$?

Solution.

$$m_x(t) = \sin(t)$$

$$K_x(t,t') = D_x e^{-\alpha(t-t')^2}$$

$$at\ t = t',\ K_x(t,t) = D_x(t) = D_x = const$$

$$L = \frac{d}{dt}$$

Following the formulae above we obtain:

$$m_y(t) = \frac{d}{dt}m_x(t) = \cos(t)$$

$$K_y(t,t') = \frac{\partial}{\partial t}\frac{\partial}{\partial t'}K_x(t,t') = 2\alpha D_x e^{-\alpha(t-t')^2}(1-2\alpha(t-t')^2)$$

$$at\ t = t',\ D_y(t) = 2\alpha D_x = const$$

An interesting observation can be made. Dispersion D_x describes the "width" of the corridor inside of which the particular realizations of the input function lie. On the output end, the "width" D_y is related not only to D_x, but also to the autocorrelation parameter α ! Thus, if α is small (long range correlation), the corridor of $Y(t)$ is narrow, if $\alpha \to \infty$ (short range correlation) then $D_y \to \infty$! But this means that the output $Y(t)$ has nothing to do with the $\cos(t)$, the expected dependence! This can be easily understood, because making α large brings in wild oscillations (compare top and bottom plots of figure 2). We can not differentiate, can we?

We will not go into details of how this problem might be resolved, but the answer is yes, we still can. This example simply illustrates the fact that the output error D_y depends on the input function correlation properties and that the filter has to be designed accordingly.

It can be shown, that there exists an optimal filter with parameter T which is capable of differentiating the given input under large-α assumption with $D_y < \frac{2D_x}{T}$ accuracy. The limits on the possible values of parameter T are hardware dependent, and in general T can be made as large as needed.

For Milagro readers this might, for example, imply existence of a high accuracy differentiating filter which can be used as an angle fitter. (Angle fitting is differentiating of registered times with respect to x, y, z — the pmt coordinates.)

4 Canonical Representation of Random Functions.

The idea behind the canonical representation is simple: it is to try to separate the randomness from the t-dependence. To do that one considers elementary random functions

$$V_\nu \phi_\nu(t)$$

where $\phi_\nu(t)$ are regular functions, and V_ν are random variables, ν can be discreet or continuous. Example of several realizations of such random basis is given on figure 4. The representation of a random function $X(t)$ is said to be canonical if

$$X(t) = \sum_\nu V_\nu \phi_\nu(t) + m_x(t)$$

where V_ν are such that

$$\langle V_\nu \rangle = 0; \left\langle V_\nu^2 \right\rangle = D_\nu$$

$$\langle V_\nu V_\mu \rangle = 0, \ \mu \neq \nu$$

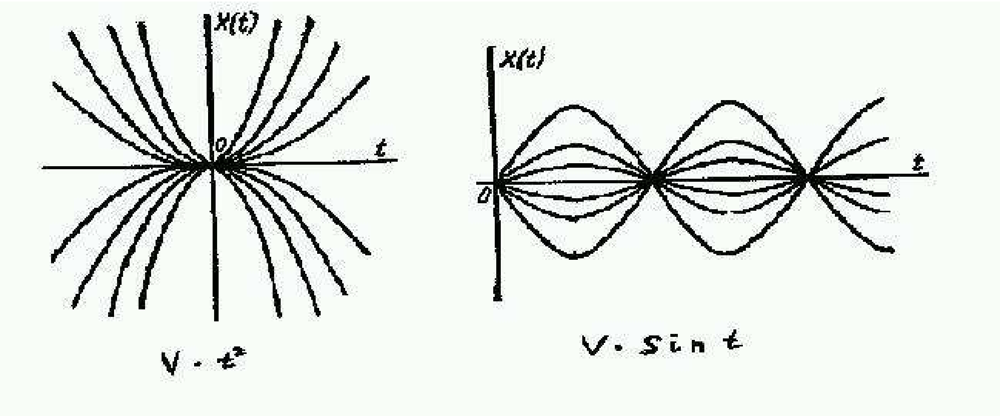

Figure 4:

Then, it is easily seen that the autocorrelation function $K_x(t,t')$ will be represented as:

$$K_x(t,t') = \sum_\nu D_\nu \phi_\nu(t)\phi_\nu(t')$$

Example. If we interpret footprint of showers as realizations of random functions of two types: photon and proton, then one can ask to which family the particular realization belongs. Here we can use difference in correlation properties (clumpiness) as a gamma/hadron identification tool: clumpiness $\rightarrow$ different $K_x(t,t')$ $\rightarrow$ different D_ν. The possibility of separation can be illustrated with the help of figure 2 where two families with different correlation properties can clearly be identified. The algorithms of gamma/hadron identification proposed by Richard[2] and Andy[3] can be viewed in this framework.

Linear transformations of canonical representations can be easily studied and performed, because they are reduced to action of operator L on the basis functions $\phi_\nu(t)$. This opens a practical possibility for the optimal linear filter algorithm search.

[2]Richard Miller, An Image Analysis Approach to Analyzing Milagro Data: Monte Carlo Study of Gamma/Hadron Separation. internal Milagro memo #45

[3]Andy Smith, progress report, MMM Meeting, January 1999.

5 References.

Pugachev, V. S., Theory of random functions and its application to control problems. Oxford ; New York : Pergamon ; Reading, Mass. Addison-Wesley, 1965.