PDF

Probability
Density
Func

Chapter 7

Noise

Noise

The noise components of a signal can have different origins. Sometimes noise is **man-made,** e.g., artifacts from **switching instruments** or 50 - 60-Hz hum **originating from power lines**.

Other noise sources are random in nature, such as **thermal noise** originating from **resistors** in the measurement chain.

Random noise is intrinsically unpredictable but it can be described by statistics. From a measurement point of view we can have noise that is introduced as a result of the measurement procedure itself, either producing systematic bias (e.g., measuring the appetite after dinner) or random measurement noise (e.g., thermal noise added by recording equipment).

If we consider a measurement M as a function of the measured process x and some additive noise N, the i-th measurement can be defined as

Eq. (1) $M_i = x_i + N_i \qquad x_i = 0.8x_{i-1} + 3.5$

Noise …

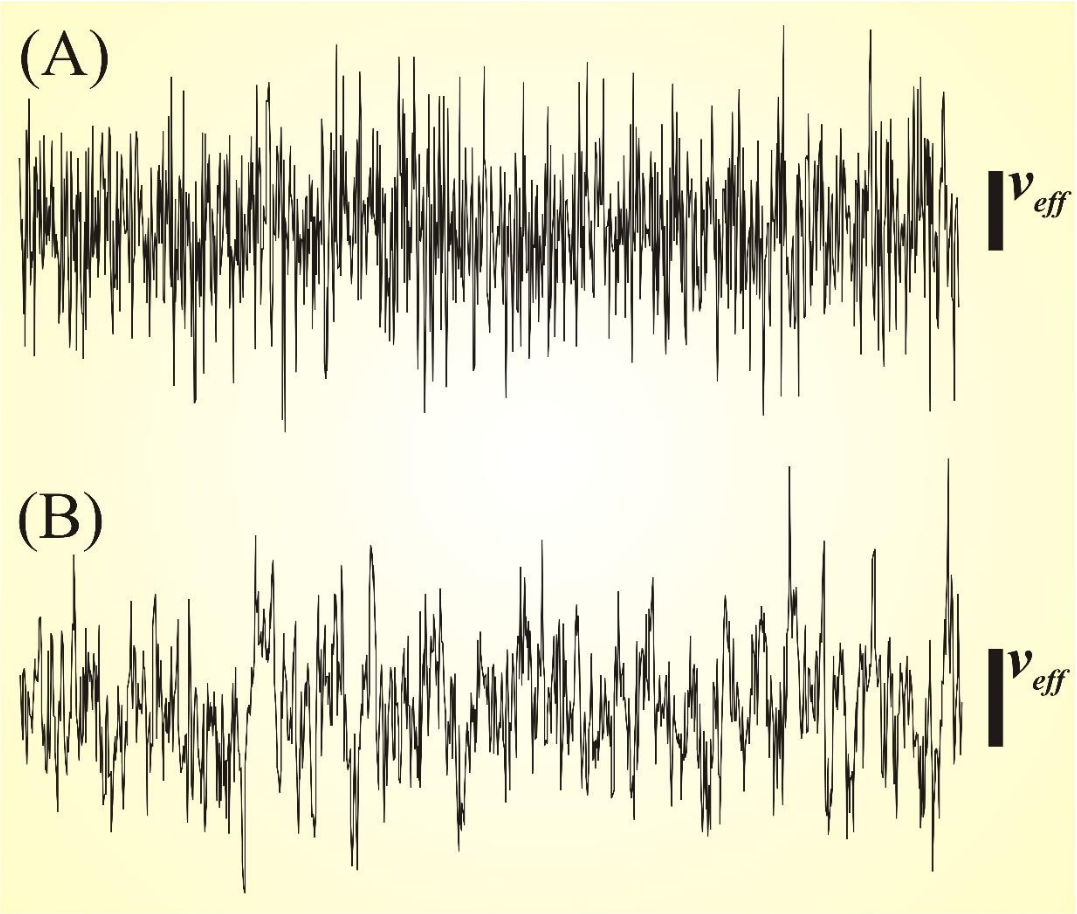
Noise may be intrinsic to the process under investigation. This **dynamical noise** is not an independent additive term associated with the measurement, but instead **interacts with the process** itself. For example, **temperature fluctuations** during the measurement of **cellular membrane potential** not only add unwanted variations to the voltage reading, **they physically influence the actual processes** that determine the potential. If we consider appropriately small time steps, we can imagine the noise at one time step contributing to a change in the state at the next time step. Thus one way to represent dynamical noise D affecting process x is:

$$M_i = x_i + N_i$$

Eq. (2) $$x_i = [0.8\, x_{i-1} + 3.5] + D_{i-1}$$

Stocastic

Noise ...



Time series including measurement noise
(A) and a combination of dynamical and measurement noise (B).

Figure 3.1 Signal processing for Neuroscientist, second edition by Wim van Drongelen

you can see that the dynamical noise (due to the **correlation between sequential values**) creates slower trends when compared to the time series with only additive noise. It must be noted here that in many cases, a dynamic noise term is used to represent a random process simply because often we do not know all of the details necessary to accurately represent the entire range of complex interactions in a physiological system. In this sense, the random process compensates for our lack of detailed knowledge by giving us a statistical proxy for what we don't know about the system

Note: The process in Eq. (1) is deterministic, only its measurement is corrupted by noise. However, although the process in Eq.(2) includes an deterministic component, it is a so-called stochastic process because a noise component is part of the process itself.

NOISE STATISTICS

- One common way to characterize a random process is by its probability density function (PDF), describing the probability p(x) that particular values of x(t) occur.
- For instance, if we create a function to describe the probability of each outcome of a fair roll of a single die, we would have the possible observations 1, 2, 3, 4, 5, 6. In this case, each of the six possible observations occurs with a probability p(1), p(2), … … , p(6), each equal to 1/6. This would result in a PDF that is 1/6 for each of the values 1 through 6 and 0 for all other values.
- The PDF for the fair die is shown **A**. This example can be extended to continuous variables, and such an example of a variable that ranges between 0 and 6 is shown in **B**. In this example all values within the range are equally likely to occur. Often this is not the case; the most well-known PDF is the normal distribution shown in **C**, reflecting a process where most values are close to the mean and extreme values (either posit r.

$$F(x) = \int_{-\infty}^{x} p(y)dy$$

Cumulative function

$$\int_{-\infty}^{\infty} p(x)dx = 1$$

$$\mathscr{F}(x) = 1 - F(x) = \int^{\infty} p(y)dy$$

Survival function

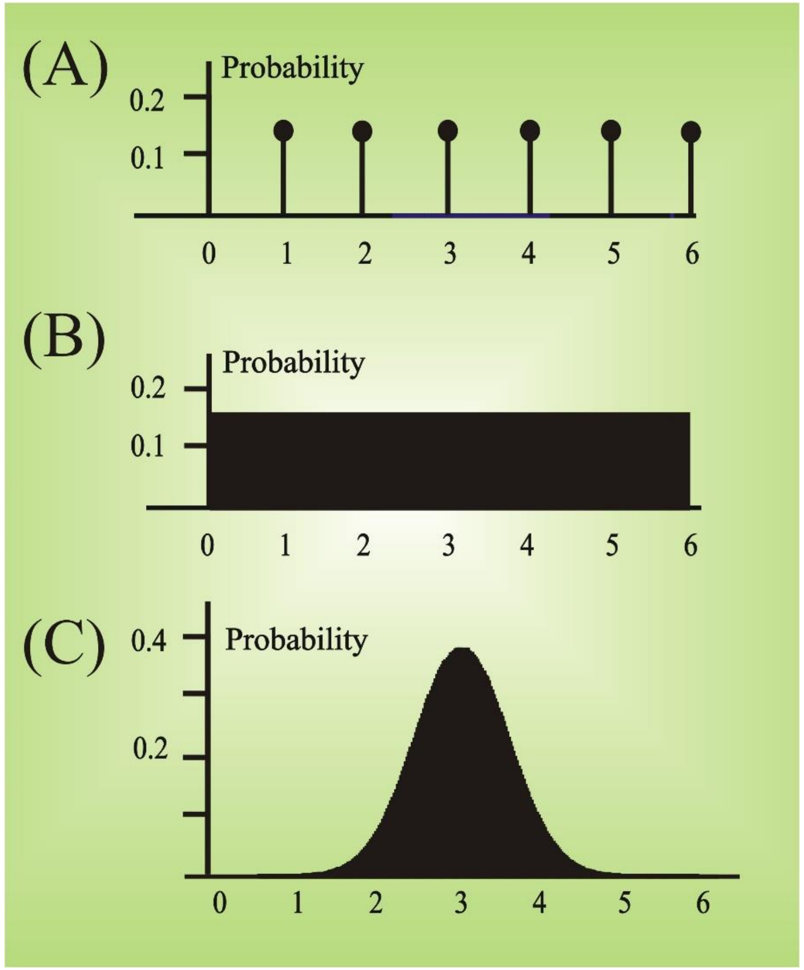


Figure 3.2 Signal processing for Neuroscientist, second edition by Wim van Drongelen

Ensemble

If one observes a random process over time, one can obtain sample functions, series of measured values representing one instance of the random process .A collection of these sample functions forms an ensemble.

This figure shows the Observations of the random process characterized by the PDF shown in opposite slide part C . Sample functions are individual “samples” from the larger ensemble. For each trace the **amplitude distribution histogram** is shown on the side in red. (i.e., the vertical axis of this distribution corresponds to the range of amplitude values and the horizontal axis to the number of times this amplitude was present in the associated sample function).

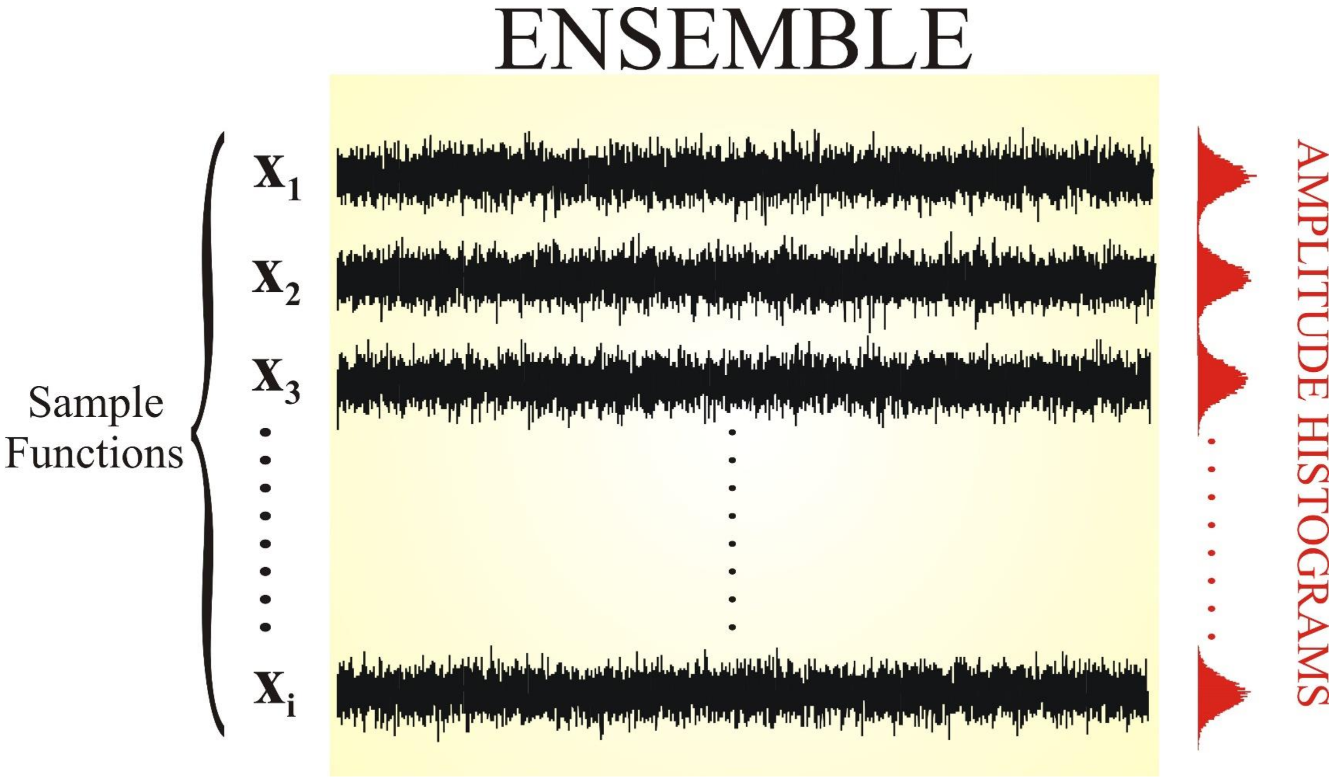


Figure 3.3 Signal processing for Neuroscientist, second edition by Wim van Drongelen

Stationary, ergodicity

The random process is called **stationary** if the **distribution from which x(t) originated doesn't change over time**. In pervious slide, the amplitude distribution is shown for each sample function. The similarity of these distributions makes the assumption of underlying stationarity a reasonable one.

The process is ergodic if any of the **particular sample functions is representative of the whole ensemble**, thus allowing statistics to be obtained from averages over time.

When applying signal processing techniques, the stationarity and ergodicity of signals are frequently (and often implicitly) assumed, and many techniques can be useful, even when these assumptions are not strictly met.

Two common parameters that are estimated from random processes are **mean** and **variance**. If a process is **stationary** and **ergodic**, one can characterize the distribution using any of the sample functions e.g., the estimate of the mean of x over an interval T is

$$\hat{x} = \frac{1}{N}\sum_{i=1}^{N} x_i$$

$$\widehat{Var}(x) = \frac{1}{N}\sum_{i=1}^{N}(x_i - \hat{x})^2$$

SIGNAL TO NOISE RATIO

Generally, any (biomedical) measurement will necessarily be corrupted by some noise. Even if the process itself were noise-free, the measurement chain adds noise components because all analog instruments (amplifiers, analog filters) add, at the very least, a small amount of thermal noise. If the noise component is sufficiently small compared to the signal component, one can still gather reasonable measurements of the signal.
To quantify this **ratio between signal and noise components**, one can (in some cases) determine the amplitude or the power of each component and from those calculate a signal-to-noise ratio.

Mean squared amplitude $ms = \frac{1}{N}\sum_{i=1}^{N} x_i{}^2$

Root mean squared Amplitude $rms = \sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i{}^2}$

$$SNR = \frac{ms(signal)}{ms(noise)}$$

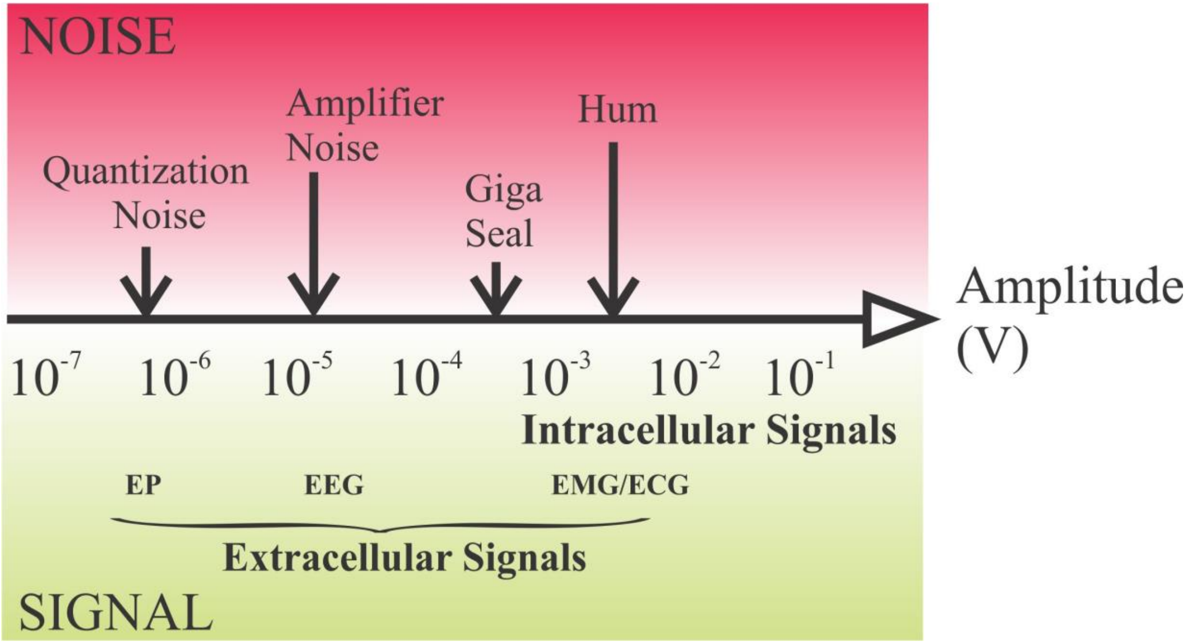
$$SNR = 10log_{10}\frac{ms(signal)}{ms(noise)} \text{ dB}$$

Note that the dB scale does not have a physical dimension; it is simply the logarithm of a ratio. The signal-to-noise ratio (without the log transform) is sometimes used as a **figure of merit (FOM)** by equipment manufacturers.

If this ratio is close to 1, or even <1, signal processing can help to increase SNR in special cases.

Noise sources...

From the results shown in the examples above, it may be clear that with modern equipment, low noise recordings are indeed feasible. However, often the amplitude of the noise is comparable to the amplitude of different types of biopotentials ,indicating that strategies for noise reduction are required. Enemy #1 in any recording of biopotentials (or low-level transducer signals with similar amplitudes) is **hum**. As we will see, hum as a nonrandom noise source may even play a role in spoiling signal averaging results.



Overview of the amplitude of typical biopotentials and different types of noise.

Figure 3.5 Signal processing for Neuroscientist, second edition by Wim van Drongelen