Biosignal Analysis
Biosignal Processing Methods

Medical Informatics 1
WS 2005/2006

JH van Bemmel, MA Musen:
Handbook of medical informatics, Springer 1997

Biosignal Analysis

Introduction

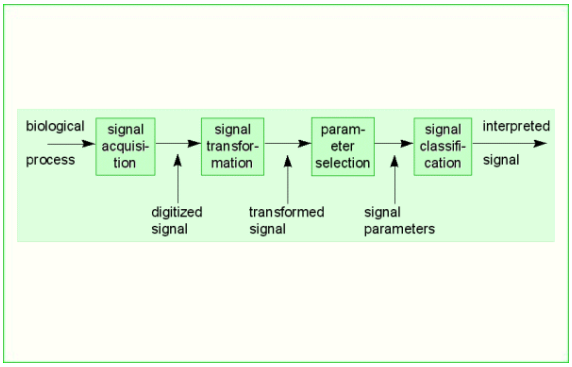


Fig. 8.1: The four stages of biosignal processing

Types of Signals

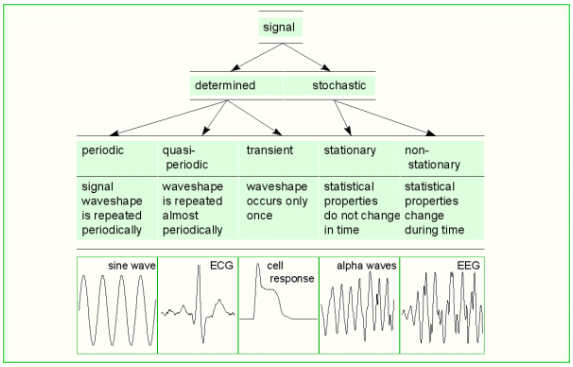


Fig. 8.2: Types of biological signals

Types of Signals

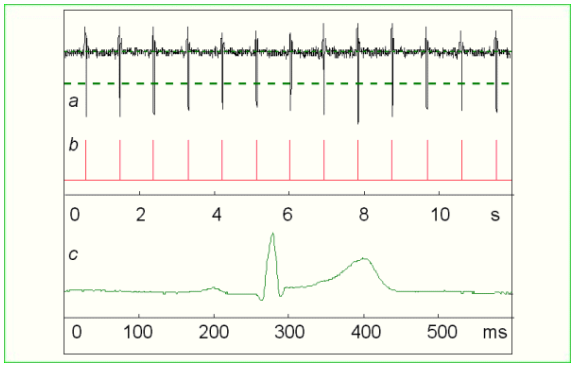


Figure 8.3: Impulse series

Analog-to Digital Conversion

08.01 Bandwidths, Amplitude Ranges, and Quantization of Some Frequently Used Biosignals

Signal	*Bandwidth (Hz)*	*Amplitude range*	*Quantization (bits)*
Electroencephalogram	0.2-50	600 V	4-6
Electrooculogram	0.2-15	10 mV	4-6
Electrocardiogram	0.15-150	10 mV	10-12
Electromyogram	20-8000	10 mV	4-8
Blood pressure	0-60	400 mm Hg	8-10
Spirogram	0-40	10 L	8-10
Phonocardiogram	5-2000	80 dB	8-10

Table 8.1. Bandwidths, Amplitude Ranges, and Quantization of Some Frequently Used Biosignals.

Analog-to Digital Conversion

08.01 Sampling of Signals: How Often?

Without exception, all biosignals are analog signals. Processing of biosignals by computers therefore requires discretization (i.e., sampling and quantification). This panel explains the *sampling process* without referring to formulas.

The sampling theorem mathematically phrased by Shannon and Nyquist states that a signal must be sampled at a rate at least twice the rate of the highest-frequency component present in the signal. If we use a sampling rate that is too low, the signal is distorted. If we obey the sampling theorem, the complete *syntactic information content* of the signal is retained. This is illustrated by the following example.

An EEG usually contains statistical, more or less sine wave-shaped fluctuations that may occur at a rate of up to 30 times/second. This can also be expressed by saying that the EEG contains frequencies up to 30 Hz. Higher frequencies may also be present (e.g., from other signal sources) but these are generally not of semantic interest.

The sampling theorem then prescribes that we should sample the EEG at least at 2 x 30 = 60 Hz to keep all signal properties. Table 8.1 gives the frequency bandwidths of interest and the most commonly used sampling rates for some biosignals. For instance, for ECGs (bandwidth, 0.15-150 Hz) a sampling rate above the Shannon frequency (500 Hz) is most often used. If we obey the rule of the sampling theorem it is, in principle, possible to restore the original analog signal by digital-to-analog conversion.

Analog-to Digital Conversion

08.02 Sampling of Signals: How Accurate?

When sampling a signal, we use an analog-to-digital converter (A-D converter or ADC). Samples are taken at a rate at least twice the rate of the highest-frequency component contained in the signal (i.e., the mixture of signal plus noise, unless the noise has been filtered out beforehand), and the samples are quantitated and expressed as numbers. The latter is always done with a limited accuracy and may, in principle, add so-called quantization noise to the sampled signal. This quantization noise should generally not exceed the noise that is already present in the signal, or, as expressed in more general terms, discretization by the ADC should not increase the information entropy (see Chapter 2); syntactic and semantic signal properties should be left intact.

The degree of quantization can be expressed as the number of quantization steps for the range of possible amplitude values. If the signal amplitude spans a range of A volts (e.g., from -A/2 to +A/2) and the quantization step is Δq, then the number of quantization steps is $m = A/\Delta q$.

In practice, let m be a power of 2: $m = 2^n$, so that the quantization of the ADC can be expressed in n bits. For instance, an ADC with an accuracy of 10 bits can discern $2^{10} = 1024$ different amplitude levels, resulting in a resolution of about 0.1%, expressed as a percentage of the signal range A. An ADC that delivers samples with 8-bit accuracy ($2^8 = 256$ steps) is called an 8-bit ADC. A 1-bit ADC only determines the sign of the signal (or whether it is larger or lower than some threshold).

For most biosignals a 6- to 12-bit ADC is sufficient; a 12-bit ADC implies a resolution of 1/4096 (less than 0.025%), related to a signal-to-noise ratio which is far superior to that attainable with most signal transducers.

Analog-to Digital Conversion

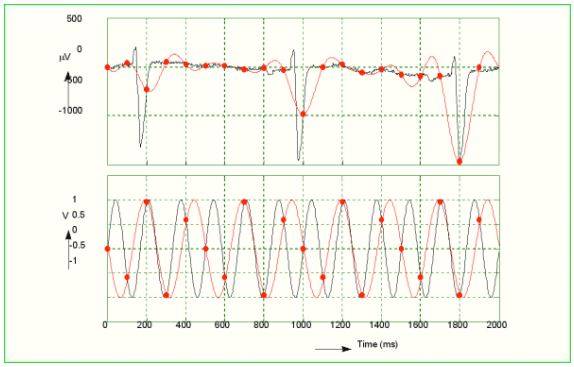


Figure 8.4: Effect of sampling frequencies

Application Areas of Biosignal Analysis

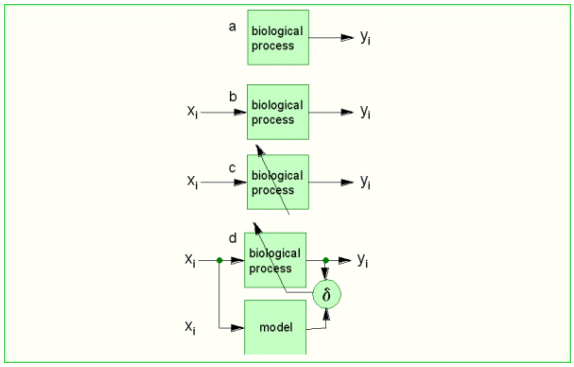


Figure 8.5: Four different situations in biosignal processing: output signal, evoked signal, provocative test, process modelling.

Biosignal Processing Methods

Signal-Amplitude Properties

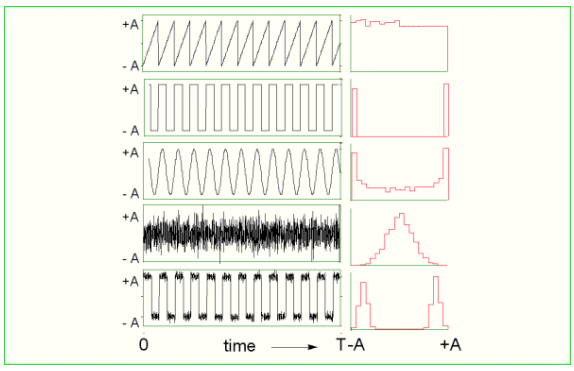


Figure 25.1: Amplitude distribution functions (density distribution function (ddf))

Signal-Amplitude Properties

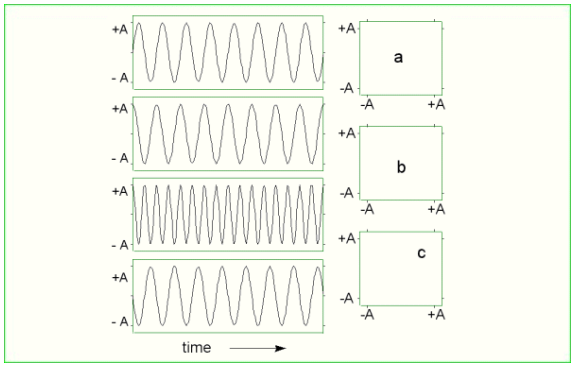


Figure 25.2: Examples of 2D amplitude distributions

Signal-Amplitude Properties

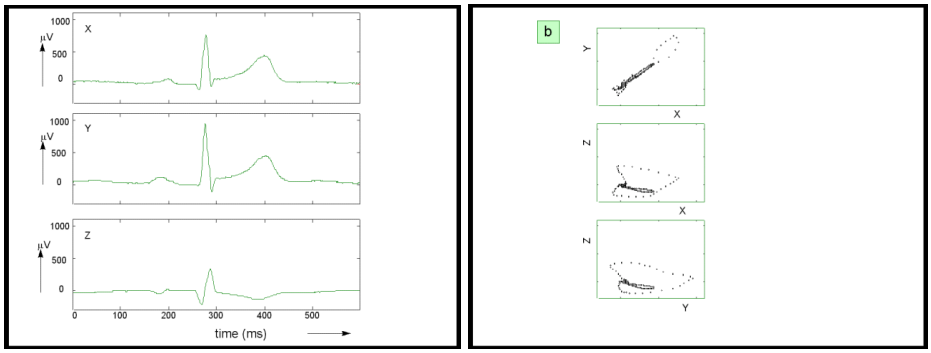


Figure 25.3: Vectorcardiogram

Signal-Amplitude Properties

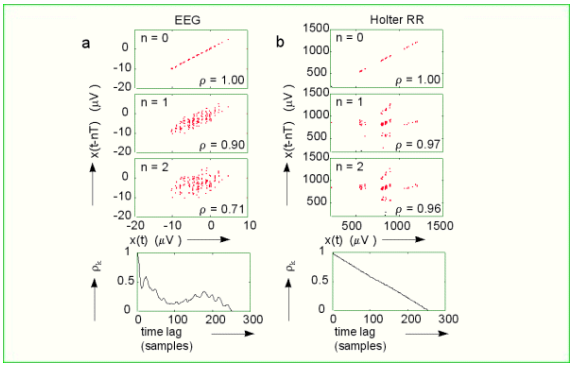


Fig. 25.4: Two-dimensional amplitude ddfs for EEG amplitudes and RR intervals

25.01 One-Dimensional Distributions, Means and Variances

In many instances we want to follow how signal properties change as a function of time in order to detect trends in the underlying biological processes. In that case we compute parameters for short observation periods T, in which the periods T may show some overlap. During such an observation period we consider the signal to be stationary, and determine for each period T parameter values that describe the signal. In this Panel we derive parameter values from amplitude ddfs, which may or may not have a stable means, which may be narrow or wide, and which may or may not be symmetric. If the signal can be described as:

$x(t)$, for $t = t_0$ to $t = t_0 + T$, where T is the observation period,

then the sampled signal can be expressed as:

x_i, with $i = 1, 2, \ldots N$, the sample number.

The interval $\Delta T = T/N$ is called the sampling interval, and $f_s = 1/\Delta T$ is the sampling frequency. The amplitude ddf of x_i can be expressed as $f(x)$, in which no time information or sample number information is kept. The mean of a statistical variable x, for which the ddf $f(x)$ is given, may be written as the *expectation* of x, or:

$$E\{x\} = \int_{-\infty}^{\infty} x f(x)\,dx$$

which is also called the *first-order moment* of the ddf $f(x)$. Another way of writing the mean value of the variable x is:

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

which is a good approximation of $E\{x\}$ for large N. Thus, for the observation period T, $E\{x\}$ is equal to $\bar{x}$. For stationary signals, this equation holds for all observation periods T, for nonstationary signals may vary. In a similar way it is possible to write for the variance of the signal (the second-order moment of the ddf):

$$E\{[x - E\{x\}]^2\} = \int_{-\infty}^{\infty} [x - E\{x\}]^2 f(x)\,dx \qquad (1)$$

For a signal with a mean value of zero ($E\{x\} = 0$) this variance becomes:

$$E\{x^2\} = \int_{-\infty}^{\infty} x^2 f(x)\,dx \qquad (2)$$

The dispersion σ, is the square root of the variance. The signal-to-noise ratio (SNR) of a signal that is the sum of a signal plus noise, $x(t) = s(t) + n(t)$ (assuming that $E\{s\}$ and $E\{n\}$ are equal to zero) is defined as the ratio of the variances of signal and noise, S and N, respectively:

$$SNR = S/N = \sigma_s^2 / \sigma_n^2 = E\{s^2\} / E\{n^2\}$$

25.02 Two-Dimensional Distributions and Correlation

Equations 1 and 2 of Panel 25.1 are, in fact, examples of a special second-order moment of the function $f(x)$. As we have seen from Fig. 25.4, it is possible to plot two-dimensional amplitude density distribution functions (2-D ddfs) of signal samples either from two different signals or from the same signal, taken τ seconds apart. As we can see from Section 1.3 and Fig. 25.4, we can compute correlation coefficients from such 2-D ddfs and plot them as a correlation function $\rho(\tau)$.

If we define the 2-D ddf of two samples taken from two signals x and y as $f(x,y)$, then we may write in general terms the second order-moment as:

$$E\{xy\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\, f(x,y)\,dx\,dy. \qquad (3)$$

This value is called the covariance. If we take samples of the signal y at a time lag of τ seconds after the samples of the signal x, we can compute the 2-D ddf for x and y, in the way mentioned in Section 1.2. For each value of $\tau$the correlation coefficient $\rho(\tau)$ can be computed, which is after normalization:

$$\rho_{xy}(\tau) = E\{xy_\tau\} / [E\{x^2\}E\{y_\tau^2\}]^{1/2}.$$

This value is called the cross-correlation. Similarly,

$$\rho_{xx}(\tau) = E\{xx_\tau\} / [E\{x^2\}E\{x_\tau^2\}]^{1/2} = E\{xx_\tau\} / E\{x^2\}$$

is called the autocorrelation, which is equal to one for $\tau = 0$.

Frequency Spectra and Filtering

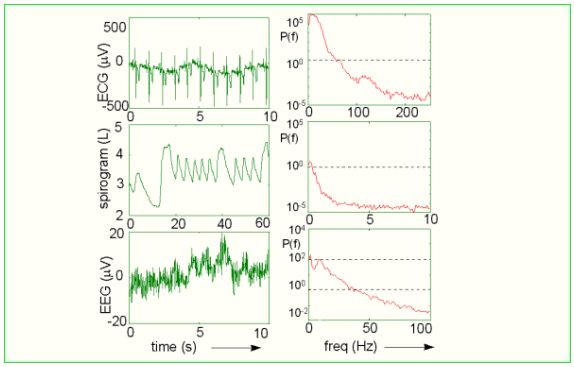


Figure 25.5: Examples of three biological signals with their frequency spectra

Frequency Spectra and Filtering

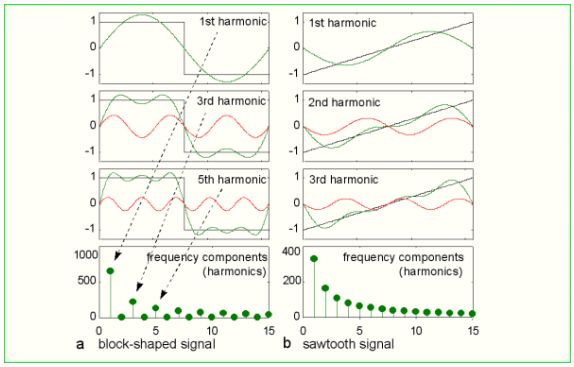
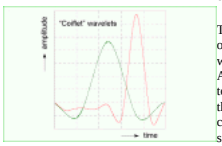


Figure 25.6: Frequency components.

25.03 Wavelet Analysis

In Fourier analysis, a signal is thought to be composed of sines and cosines. By using the Fourier transform, a signal can be decomposed in these basic functions. The samples of the transformed signal (Fourier coefficients) represent the contribution of sine and cosine functions at different frequencies. A disadvantage of Fourier analysis is that it is difficult to compose a signal that is *limited in time*, by using functions that, by definition, stretch out into infinite time. It is therefore difficult for a Fourier function to approximate sharp changes in a signal. For example, a very simple and time-limited signal, a spike, is decomposed by Fourier transformation into an infinite number of sines and cosines (see also Fig 25.6a for the decomposition of a block signal).

A way to tackle this problem is through wavelet analysis. It uses the same principle as Fourier analysis, namely, that signals are composed of basic functions, called *wavelets*. The most important difference between these wavelets and the sines and cosines used in Fourier analysis is that wavelets are limited in time. The procedure for wavelet analysis is to choose a suitable wavelet prototype function (also called mother wavelet or analyzing wavelet) that meets certain constraints. All composing functions are derived by stretching or scaling the mother wavelet both in time and in amplitude. Using a wavelet transform, the signal is decomposed into these scaled versions of the mother wavelet. In fact, the composing cosines used in Fourier analysis can also be seen as stretched, scaled, and shifted versions of a *mother-cosine*. In Fourier analysis, the composing functions are infinite in the time domain because they represent exactly one frequency. In wavelet analysis, the composing wavelets have a limited extent both in the time domain and in the frequency domain, where contributions from frequencies outside a certain area are negligible.



The most important result of the wavelet transform is the location of the composing wavelets in time. Sharp, time-limited signal parts will be represented by wavelets that are scaled down in duration. As in Fourier analysis, the contribution of the composing wavelets to the signal provides information about the temporal properties of the signal on different time scales. Additionally, the locations of the composing wavelets provide information about the position of a specific signal property. Figure 25.7 shows two examples of a mother wavelet, both from the well-known Coiflet wavelet family.

Frequency Spectra and Filtering

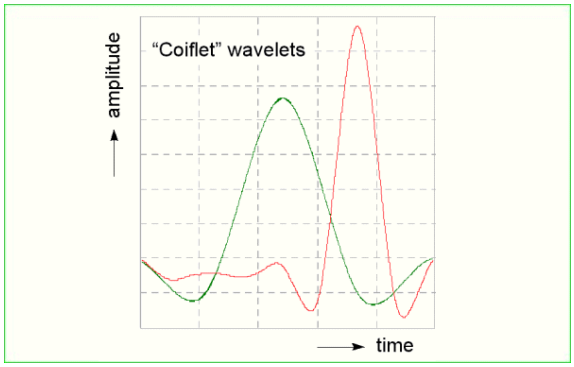


Figure 25.7: Two examples of a mother wavelet.

Frequency Spectra and Filtering

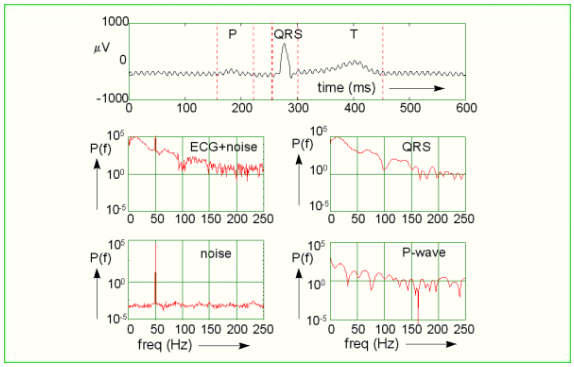


Figure 25.8: Frequency spectra of an ECG.

Frequency Spectra and Filtering

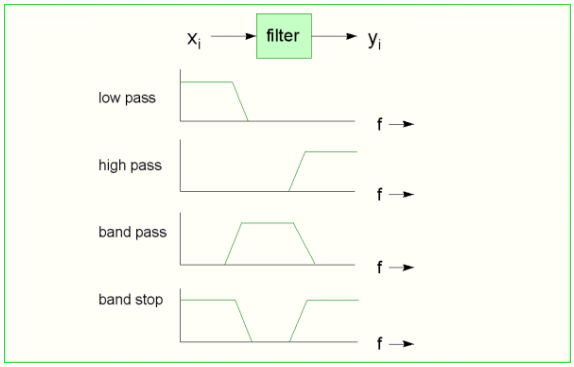


Figure 25.9: Schematic representation of filters.

Frequency Spectra and Filtering

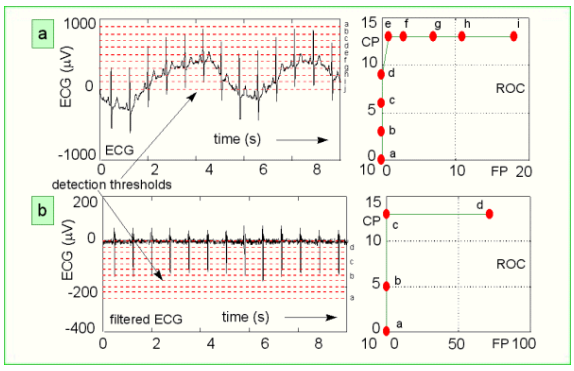


Figure 25.10: ECG and its band-pass filtered version.

Frequency Spectra and Filtering

15.01 Relationships between True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN)

		Decision Model +	Decision Model -	
Truth	+	TP	FN	100%
	-	FP	TN	100%

Table 15.1. Relationships between True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN).

15.02 Illustration of Sensitivity, Specificity, and Predictive Value

		Decision Model +	Decision Model -	
Truth	+	a	b	a + b
	-	c	d	c + d
		a + c	b + d	a + b + c + d

Table 15.2. Illustration of Sensitivity, Specificity, and Predictive Value (see text).

sensitivity=a/(a+b)
specificity=d/(c+d)

ppv=a/(a+c)
npv=d/(b+d)

Frequency Spectra and Filtering

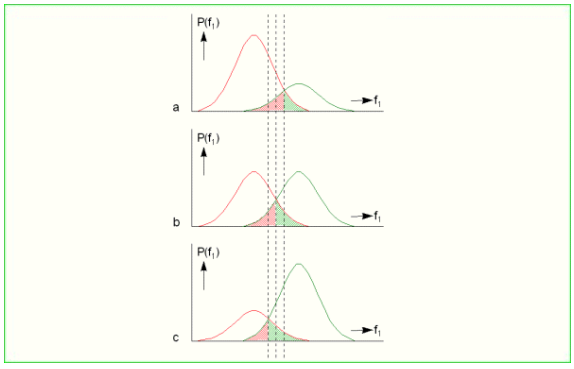


Figure 15.6: Distributions of systolic blood pressure of hypertensive and nonhypertensive people: a) population survey, b) primary care, c) cardiac clinic

Frequency Spectra and Filtering

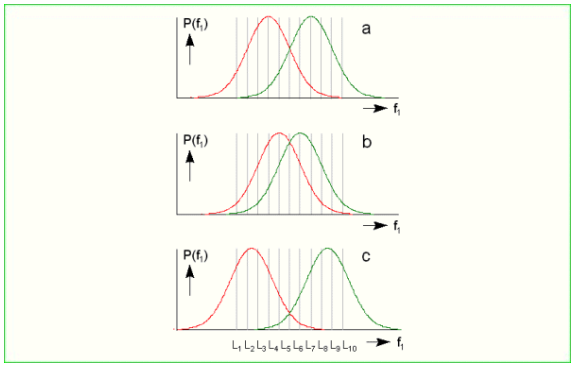


Figure 15.7: Distributions of the primary population of Fig. 15.6.

Frequency Spectra and Filtering

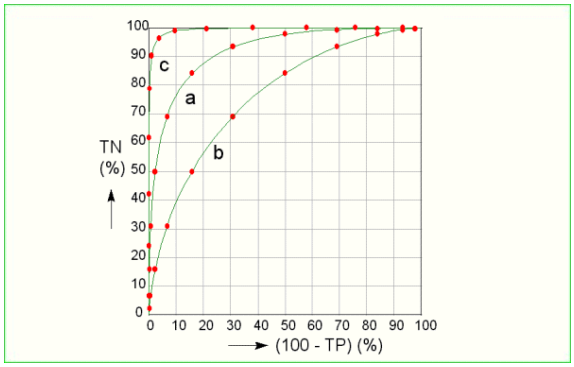


Figure 15.8: ROC curves of the population of Fig. 15.7
TN (100-FP) vs. FN (100-TP).

Signal-to-Noise Ratio

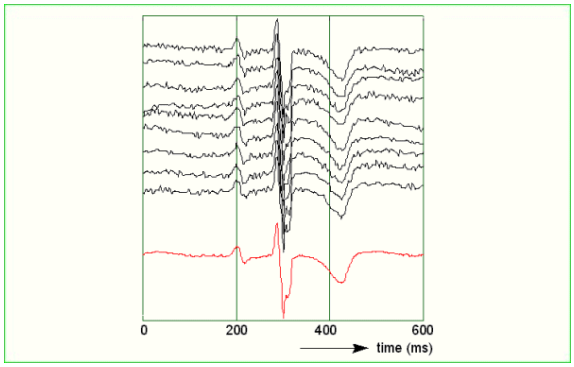


Figure 25.11: Coherent averaging in an ECG recording.

Signal-to-Noise Ratio

25.04 Coherent Averaging

In coherent averaging we compute the sum of, say, K waveforms s_0, which are extracted after detection from a noisy signal $x_0(t) = {}_{s0}(t) + n_0(t)$. The original signal variance is $S_0 = \sigma_{s0}^2$ and the noise variance is $N_0 = \sigma_{n0}^2$, so that the SNR is:

$SNR_0 = S_0/N_0$.

The sum of the K signal waveforms s_0 will result in a waveform s_1 which is K times as large as the original waveform, that is, $s_1 = Ks_0$. The resulting signal dispersion is also K times as large: $\sigma s_1 = K\sigma_{s0}$. The variance of s_1 is then $S_1 = \sigma_{s1}^2 = K^2\sigma_{s0}^2$.

We assume that the noise has a normal distribution. The K noisy waveforms n_0 are also summed to a new noisy signal, n_1. It can be proven that the variance of n_1 is K (and not K^2) times as large as the variance of n_0 so that $N_1 = \sigma n_1^2 = K\,\sigma n_0^2$. The SNR after summation is then:

$SNR_1 = S_1/N_1 = K^2\sigma s_0^2/K\sigma n_0^2 = K\sigma s_0^2/\sigma n_0^2 = KS_0/N_0 = K\ SNR_0$.

This implies that the SNR has improved linearly with the number of summed waveforms.

Signal Detection

25.01 Four Different Detection Situations for the Decision *D* that an event *S* is Present

Situation[a]	*Description*	*S*	*D*
TP	The event is present AND is correctly detected	1	1
FP	The event is not present AND is incorrectly detected	0	1
TN	The event is not present AND is correctly not detected	0	0
FN	The event is present AND is incorrectly not detected	1	0

Table 25.1. Four Different Detection Situations for the Decision *D* that an event *S* is Present.

[a]TP, true positive; TN, true negative; FP, false positive; FN, false negative.

Signal Detection

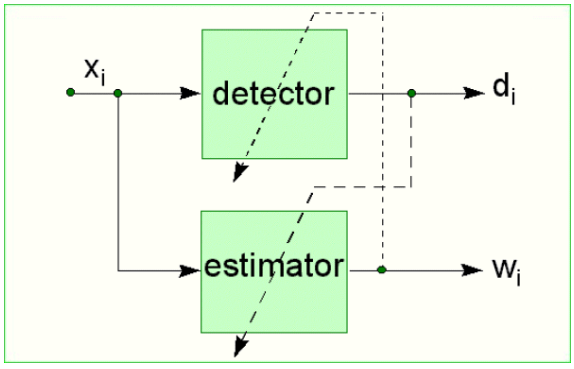


Figure 25.12: Detection and estimation may reinforce each other.

Signal Detection

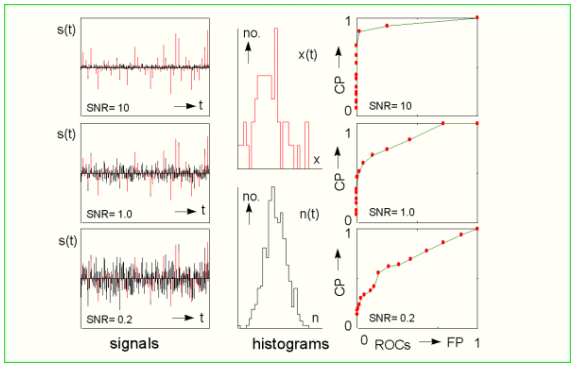


Figure 25.12: An artificial signal of amplitude-modulated impulses.