2.161 Signal Processing: Continuous and Discrete
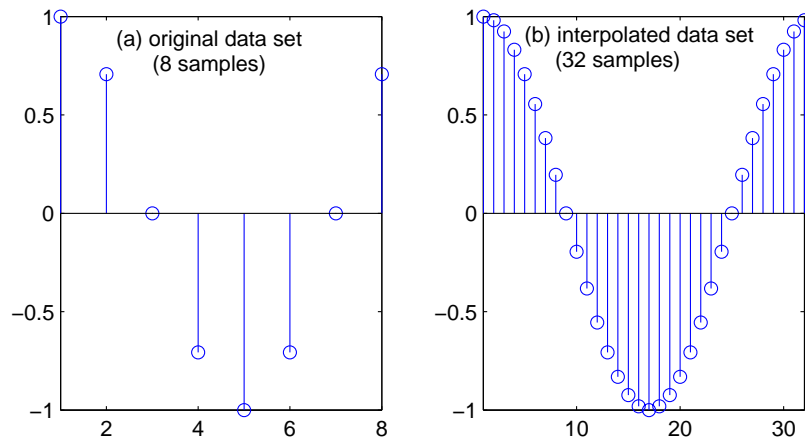Fall 2008

Lecture 21[1]

Reading:

- Class Handout: Interpolation (Up-sampling) and Decimation (Down-sampling)

- Proakis and Manolakis: Secs. 11.1 – 11.5, 12.1

- Oppenheim, Schafer, and Buck: Sec. 4.6, Appendix A

- Stearns and Hush, Ch. 9

1 Interpolation and Decimation

1.1 Up-Sampling (Interpolation) by an Integer Factor

Consider a data set $\{f_n\}$ of length N, where $f_n = f(n\Delta T)$, $n = 0 \ldots N - 1$ and ΔT is the sampling interval. The task is to resample the data at a higher rate so as to create a new data set $\{\hat{f}_n\}$, of length KN, representing samples of the same continuous waveform $f(t)$, sampled at intervals $\Delta T/K$. The following figure shows a cosinusoidal data set with $N = 8$ samples, and resampled with the same data set interpolated by a factor $K = 4$.

1.1.1 A Frequency Domain Method

This method is useful for a finite-sized data record. Consider the DFTs $\{F_m\}$ and $\{\hat{F}_m\}$ of a pair of sample sets $\{f_n\}$ and $\{\hat{f}_n\}$, both recorded from $f(t)$ from $0 \leq t < T$, but with sampling intervals ΔT and $\Delta T/K$ respectively. Let N and KN be the corresponding sample sizes. It is assumed that ΔT has been chosen to satisfy the Nyquist criterion:

- Let $F(j\Omega) = \mathcal{F}\{f(t)\}$ be the Fourier transform of $f(t)$, and let $f(t)$ be sampled at intervals ΔT to produce $f^*(t)$. Then

$$F^*(j\Omega) = \frac{1}{\Delta T} \sum_{n=0}^{\infty} F\left(j\left(\Omega - \frac{2\pi n}{\Delta T}\right)\right) \tag{1}$$

  is periodic with period $2\pi/\Delta T$, and consists of scaled and shifted replicas of $F(j\Omega)$. Let the total sampling interval be T to produce $N = T/\Delta T$ samples.

- If the same waveform $f(t)$ is sampled at intervals $\Delta T/K$ to produce $\hat{f}^*(t)$ the period of its Fourier transform $\hat{F}^*(j\Omega)$ is $2\pi K/\Delta T$ and

$$\hat{F}^*(j\Omega) = \frac{K}{\Delta T} \sum_{n=0}^{\infty} F\left(j\left(\Omega - \frac{2\pi K n}{\Delta T}\right)\right) \tag{2}$$

  which differs only by a scale factor, and an increase in the period. Let the total sampling period be T as above, to generate KN samples.

- We consider the DFTs to be sampled representations of a single period of $F^*(j\Omega)$ and $\hat{F}^*(j\Omega)$. The equivalent line spacing in the DFT depends only on the total duration of the sample set T, and is $\Delta\Omega = 2\pi/T$ in each case:
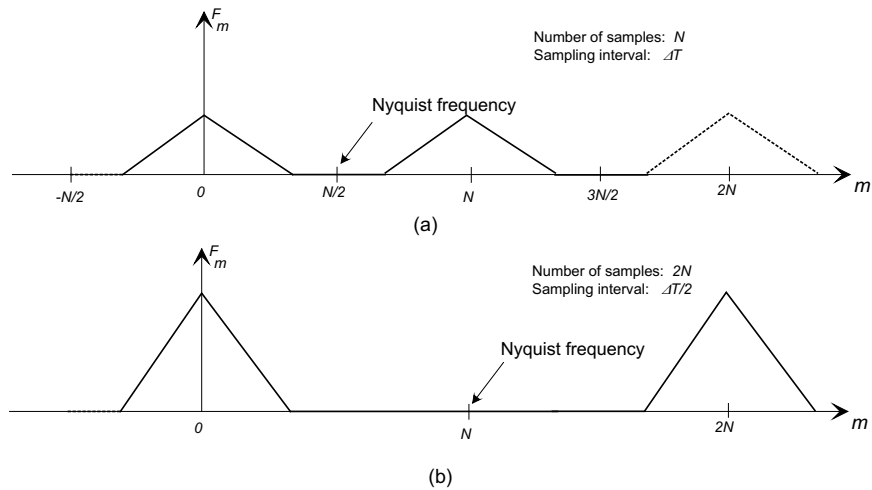
$$F_m = F^*\left(j\left(\frac{2\pi m}{T}\right)\right), \quad m = 0, 1, \ldots N - 1$$

$$\hat{F}_m = \hat{F}^*\left(j\left(\frac{2\pi m}{T}\right)\right), \quad m = 0, 1, \ldots KN - 1.$$

  From Eqs. (1) and (2) the two DFTs $\{F_m\}$ and $\left\{\hat{F}_m\right\}$ are related:

$$\hat{F}_m = \begin{cases} K F_m & m = 0, 1, \ldots, N/2 - 1 \\ 0 & m = N/2, \ldots, NK - N/2 - 1 \\ K F_{m-(K-1)N} & m = NK - N/2, \ldots, KN - 1 \end{cases}$$

- The effect of increasing N (or decreasing ΔT) in the sample set, while maintaining $T = N\Delta T$ constant, is to increase the length of the DFT by raising the effective Nyquist frequency Ω_N.

$$\Omega_N = \frac{\pi}{\Delta T} = \frac{N\pi}{T}$$

(a)

(b)

The above figure demonstrates these effects by schematically, by comparing the DFT of (a) a data set of length N derived by sampling at intervals ΔT, and (b) a data set of length $2N$ resulting from sampling at intervals $\Delta T/2$. The low frequency region of both spectra are similar, except for a scale factor, and the primary difference lies in the "high frequency" region, centered around the Nyquist frequency, in which all data points are zero.

The above leads to an algorithm for the interpolation of additional points into a data set, by a constant factor K:

1. Take the DFT of the original data set to create $\{F_m\}$ of length N.

2. Insert $(K-1)N$ zeros into the *center* of the DFT to create a length KN array.

3. Take the IDFT of the expanded array, and scale the sequence by a factor K.

1.1.2 A Time-Domain Method

We now examine an interpolation scheme that may implemented on a sample-by-sample basis in real-time using time domain processing alone. As before, assume that the process $f(t)$ is sampled at intervals ΔT, generating a sequence $\{f_n\} = \{f(n\Delta T)\}$. Now assume that $K-1$ zeros are inserted between the samples to form a sequence $\{\hat{f}_k\}$ at intervals $\Delta T/K$.

(a) Initial data set, N=8

(b) Data set with two samples interpolated between samples, N=24
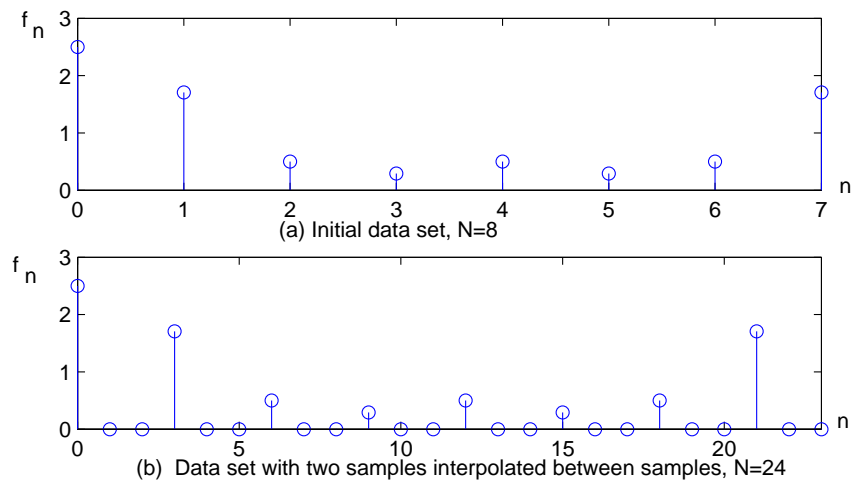
This is illustrated above, where a data record with $N = 8$ samples has been expanded by a factor $K = 3$ to form a new data record of length $N = 24$ formed by inserting two zero samples between each of the original data points. We now examine the effect of inserting $K - 1$ samples with amplitude 0 after each sample. The DFT of the original data set is

$$F_m = \sum_{n=0}^{N-1} f_n e^{-j\frac{2\pi mn}{N}}, \qquad m = 0 \ldots N - 1$$
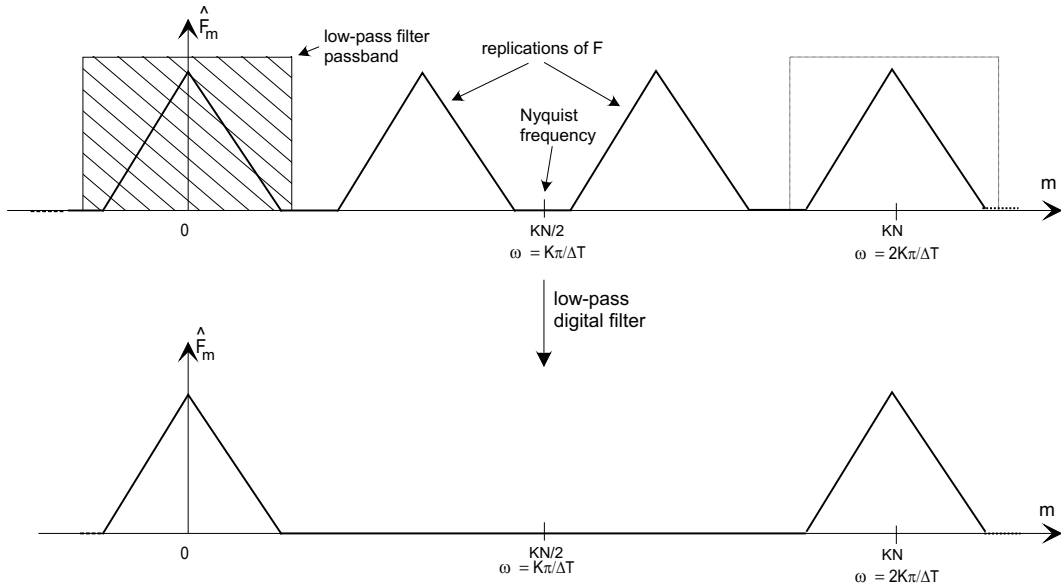
and for the extended data set $\{\hat{f}_n\}$, $n = 0 \ldots KN - 1$

$$\hat{F}_m = \sum_{n=0}^{KN-1} \hat{f}_n e^{-j\frac{2\pi mn}{KN}}, \qquad m = 0 \ldots KN - 1$$
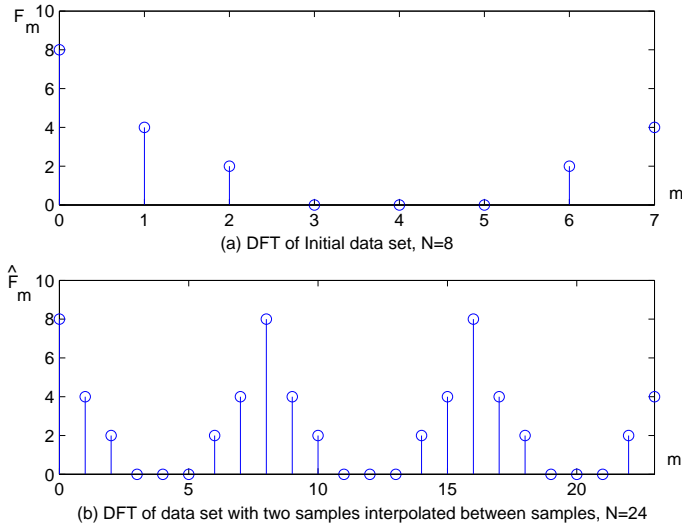
However, only the original N samples contribute to the sum, so that we can write

$$
\begin{aligned}
\hat{F}_m &= \sum_{k=0}^{N-1} \hat{f}_{Kk} e^{-j\frac{2\pi mk}{N}} \\
&= F_m, \qquad m = 0 \ldots KN - 1
\end{aligned}
$$

since $\hat{f}_{Kk} = f_k$. We note that $\{F_m\}$ is periodic with period N, and $\{\hat{F}_m\}$ is periodic with period KN, so that $\{\hat{F}_m\}$ will contain K repetitions of $\{F_m\}$.

The magnitude of the DFTs of the two waveforms is shown above. The effect of inserting the $K - 1$ zeros between the original samples has been to generate a waveform with an equivalent sampling interval of $\Delta T/K$ s, and a Nyquist frequency of $K\pi/\Delta T$ rad/s. The line resolution is unchanged, and the original DFT $\{F_m\}$ is replicated K times within the frequency span of $2K\pi/\Delta T$ rad/s.

(a) DFT of Initial data set, N=8

(b) DFT of data set with two samples interpolated between samples, N=24

The fully interpolated waveform may be reconstructed by elimination of the replications of the original spectral components. While this might be done in the frequency domain, the most common method is to low-pass filter the padded data sequence to retain only the base-band portion of the spectrum as shown below.
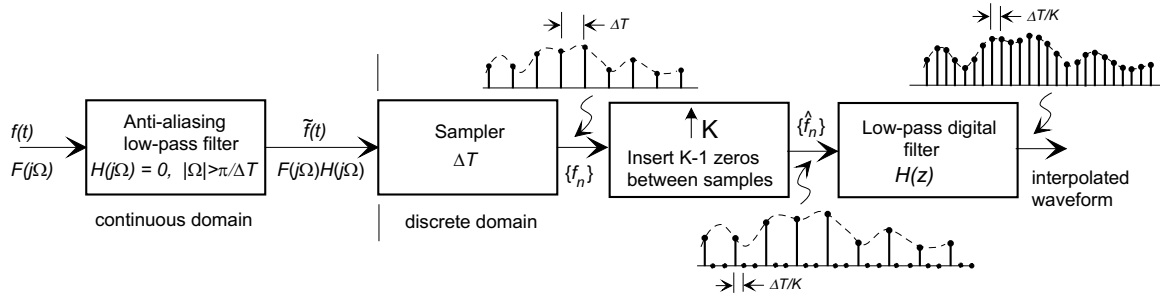

1.2 Down-Sampling (Decimation) by an Integer Factor

Decimation by an integer factor K is the reverse of interpolation, that is increasing the sampling interval T by an an integer factor (or decreasing the sampling frequency). At first glance the process seems to be simple: simply retain every Kth sample from the data sequence so that $\tilde{f}_n = f_{nK}$ where $\{\tilde{f}_n\}$ is the down-sampled sequence, as shown in below.
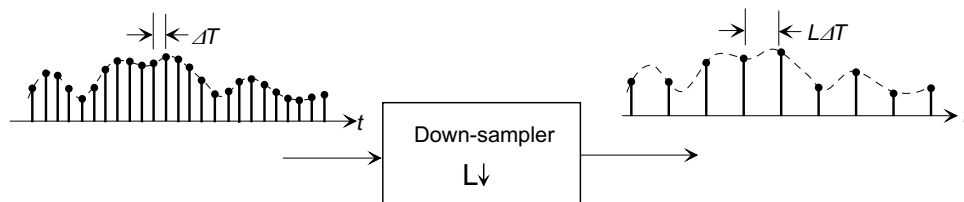
Caution must be taken however to prevent aliasing in the decimated sequence. It is not valid to directly down-sample a sequence directly unless it is known a-priori that the spectrum of the data set is identically zero at frequencies at and above the Nyquist frequency defined by the lower sampling frequency.

In discussing sampling of continuous waveforms, we described the use of a pre-aliasing filter to eliminate (or at least significantly reduce) spectral components that would introduce aliasing into the sampled data set. When down-sampling, the digital equivalent is required: a digital low-pass filter is used to eliminate all spectral components that would cause aliasing in the resampled data set. The complete down-sampling scheme is:
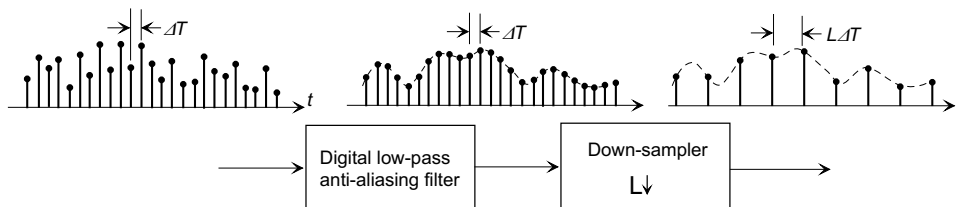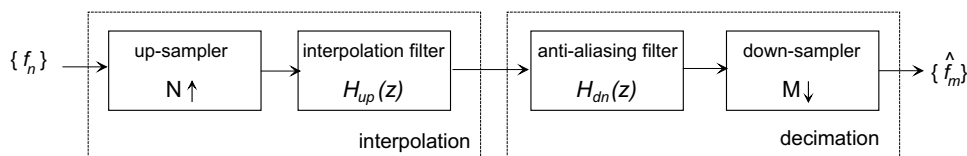

## 1.3   Resampling with a non-integer factor

Assume that the goal is to re-sample a sequence by a non-integer factor p that can be expressed as a rational fraction, that is

$$P = \frac{N}{M}$$

where N and M are positive integers. This can be achieved by (1) interpolation by a factor N, followed by (2) decimation by a factor M, as shown below.


However, since the two low-pass filters are cascaded, they may be replaced with a single filter with a cut-off frequency that is the lower of the two filters, as is shown below
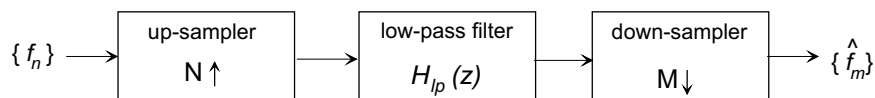
2   Introduction to Random Signals

In dealing with physical phenomena and systems we are frequently confronted with *non-deterministic, (stochastic, or random)* signals, where the temporal function can not be described explicitly, nor predicted. Some simple examples are wind loading on a structure, additive noise in a communication system, and speech waveforms. In this brief examination of random phenomena we concentrate on common statistical descriptors of stochastic waveforms and input-output relationships of linear systems excited by random waveforms.
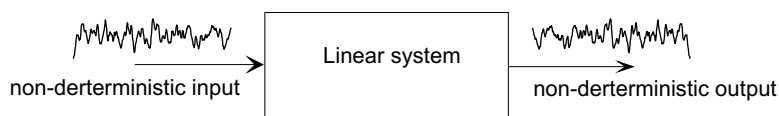
Since we cannot describe $f(t)$, we must use statistical descriptors that capture the essence of the waveform. There are two basic methods of doing this:

(a) Describe the waveform based on temporal measurements, for example define the mean μ of the waveform as
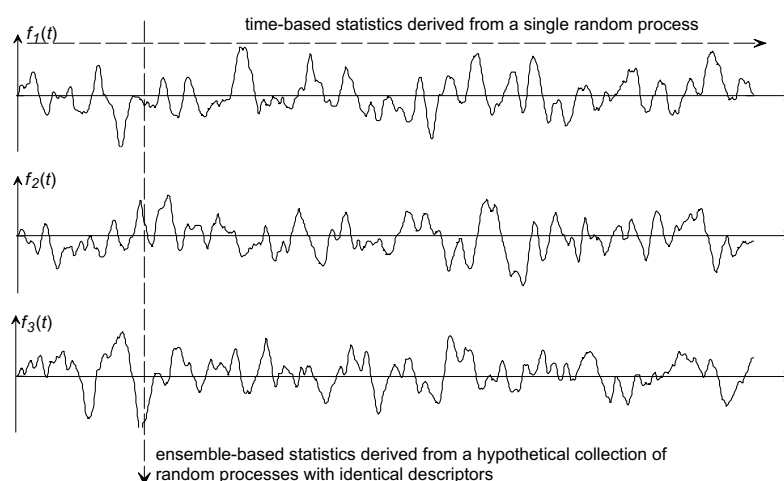
$$\mu = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} f(t) \, dt$$

(b) Conjecture an *ensemble* of random processes $f_i(t)$, $i = 1, \dots N$, with identical statistics and define the descriptors my measurements made across the ensemble at a given time, for example

$$\mu(t) = \lim_{N \to \infty} \sum_{N=1}^{\infty} f_i(t).$$

- A stationary *process* is defined as one whose ensemble statistics are independent of time.

- An *ergodic* process is one in which the temporal statistics are identical to the ensemble statistics.

Clearly, ergodicity impies stationarity.

time-based statistics derived from a single random process

$f_1(t)$

$f_2(t)$

$f_3(t)$

ensemble-based statistics derived from a hypothetical collection of
random processes with identical descriptors

In practice statistical descriptors are usually derived experimentally from measurements. For example, the mean of of a waveform might be estimated from a set of 1000 samples of a waveform and computed as

$$\hat{\mu}_1 = \frac{1}{1000} \sum_{i=1}^{1000} f_i.$$

But $\hat{\mu}$ is not the mean, it is simply an *estimator* of the true mean. If we repeated the experiment, we would come up with a different value $\hat{\mu}_2$. In statistical descriptions we use the terms *expected value*, or *expectation*, designated $E\{x\}$, and say

$$E\{\hat{\mu}_i\} = \mu$$

to indicate that our experimental estimates $\hat{\mu}_i$ will be clustered around the true mean μ.

2.1   Ensemble Based Statistics

2.1.1   The Probability Density Function (pdf)

The pdf is strictly an ensemble statistic that describes the distribution of samples x across the amplitude scale. Its definition is

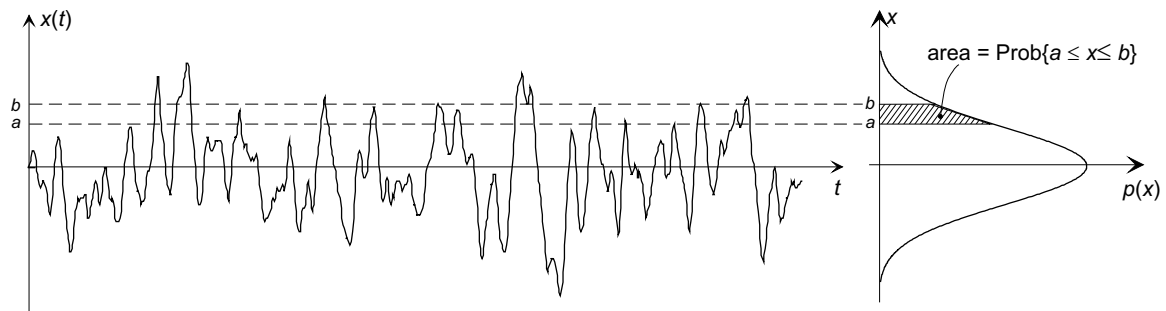$$p(x_a) = \lim_{\Delta x \to 0} \frac{1}{\Delta x} \text{Prob}\left\{x_a \leq x \leq x_a + \Delta x\right\}$$

so that the probability that a single sample lies in the range $a \leq x \leq b$ is

$$\text{Prob}\left\{a \leq x \leq b\right\} = \int_a^b p(x)\,\mathrm{d}x,$$

and we note

$$\text{Prob}\left\{-\infty \leq x \leq \infty\right\} = \int_{-\infty}^{\infty} p(x)\,\mathrm{d}x = 1.$$

For an ergodic process, the pdf can also be described from a single time series

and may be interpreted as the fraction of time that the waveform "dwells" in the range $b \leq f(t) \leq b$.

Two common pdfs are

(a) **The Uniform distribution** a random sample taken from a uniformly distributed random process is equally likely to be found anywhere between a minimum and maximum value.
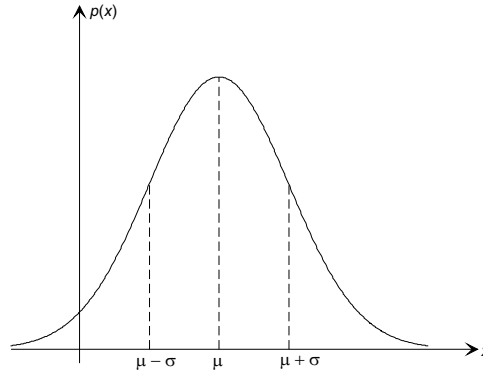
$$p(X) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{elsewhwere.} \end{cases}$$
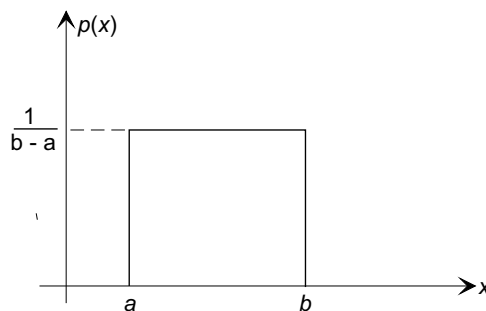
(b) The normal (or gaussian) distribution] The normal distribution defines the well known "bell-shaped-curve" of elementary statistics

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

where μ is the mean of the distribution, and σ^2 is the variance.

Note: The *central limit theorem* of statistics states that any random process that is the sum of a large number of underlying independent random processes, regardless of their distributions will be described by a gaussian distribution.

Many ensemble based statistical descriptors may be described in terms of the pdf, for example

The mean

$$\mathrm{E}\{x\} = \mu = \int_{-\infty}^{\infty} x p(x) dx$$

The variance

$$\mathrm{E}\{(x-\mu)^2\} = \sigma^2 = \int_{-\infty}^{\infty} (x-\mu)^2 p(x) dx$$

and by expanding the integral

$$\sigma^2 = \int_{-\infty}^{\infty} x^2 p(x) dx - 2\mu \int_{-\infty}^{\infty} x p(x) dx + \mu^2 \int_{-\infty}^{\infty} p(x) dx = \mathrm{E}\{x^2\} - \mu^2$$

■ Example 1

Find the mean and variance a random variable that is uniformly distributed between x_0 and $x_0 + \Delta$.
The pdf is

$$p(X) = \begin{cases} \frac{1}{\Delta} & x_0 \leq x \leq x_0 + \Delta \\ 0 & \text{elsewhere.} \end{cases}$$

The mean is

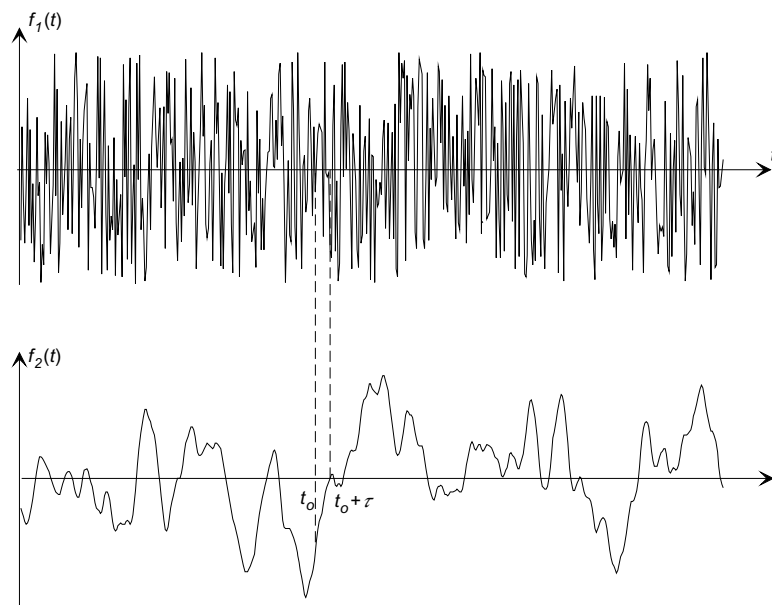$$\mu = \int_{-\infty}^{\infty} xp(x)dx = \int_{x_0}^{x0+\Delta} \frac{x}{\Delta}dx = x_0 + \frac{\Delta}{2}$$

which is the mid-point of the range.

The variance is

$$\sigma^2 = E\left\{x^2\right\} - \mu^2 = \frac{1}{12}\Delta^2$$

2.2 Time-based Statistics

Two stochastic waveforms may have identical pdfs (and hence equal means and variances), but be very different qualitatively, for example

These two waveforms obviously differ in

- spectral content, or

- self-similarity between themselves at a time t_0 and some time τ later.

A random waveform cannot be predicted exactly at any time, but clearly in the above figure the upper waveform (with greater high frequency spectral content) has less self-similarity, or *correlation* with itself, after a delay of τ.

The correlation functions are a measure of the degree to which the value of a function depends upon its past. For infinite duration waveforms the *auto-correlation function* $\phi_{ff}(\tau)$ is defined as

$$\phi_{ff}(\tau) = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} f(t)f(t+\tau)\, dt$$

and is a measure of the self-similarity of the function $f(t)$ at time t and at a time τ later. The *cross-correlation function* $\phi_{fg}(\tau)$ measures the similarity between two different functions $f(t)$ and $g(t)$ at two times τ apart.

$$\phi_{fg}(\tau) = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} f(t)g(t+\tau)\,\mathrm{d}t$$

Note that these definitions must be modified for finite duration waveforms, and if $f(t)$ exists in the interval $T_1 \le t \le T_2$, we define

$$\phi_{ff}(\tau) = \int_{T_1}^{T_2} f(t)f(t+\tau)\,\mathrm{d}t$$