Introduction to **Audio Content Analysis**

module 3.1: input representation — signals

alexander lerch

Georgia Tech Center for Music Technology
College of Design

corresponding textbook section

section 3.1

- **lecture content**
  - deterministic & periodic signals
  - Fourier Series
  - random signals
  - statistical signal description
  - digital signals

- **learning objectives**
  - name basic signal categories
  - discuss the nature of periodic signals with respect to harmonics
  - give a short description of meaning and use of the Fourier Series
  - list common descriptors for properties of a random signal

corresponding textbook section

section 3.1

- **lecture content**
  - deterministic & periodic signals
  - Fourier Series
  - random signals
  - statistical signal description
  - digital signals
- **learning objectives**
  - name basic signal categories
  - discuss the nature of periodic signals with respect to harmonics
  - give a short description of meaning and use of the Fourier Series
  - list common descriptors for properties of a random signal

audio signals
signal categories

- **deterministic signals**:
  predictable: future shape of the signal can be known (example: sinusoidal)
- **random signals**:
  unpredictable: no knowledge can help to predict what is coming next (example: white noise)

"real-world" audio signals can be modeled as time-variant combination of

- (quasi-)periodic parts
- (quasi-)random parts

audio signals
signal categories

- **deterministic signals**:
  predictable: future shape of the signal can be known (example: sinusoidal)
- **random signals**:
  unpredictable: no knowledge can help to predict what is coming next (example: white noise)

"real-world" audio signals can be modeled as time-variant combination of

- (quasi-)periodic parts
- (quasi-)random parts

audio signals

signal categories

- **deterministic signals**:
  predictable: future shape of the signal can be known (example: sinusoidal)
- **random signals**:
  unpredictable: no knowledge can help to predict what is coming next (example: white noise)

"real-world" audio signals can be modeled as time-variant combination of

- (quasi-)periodic parts
- (quasi-)random parts

audio signals

periodic signals 1/5

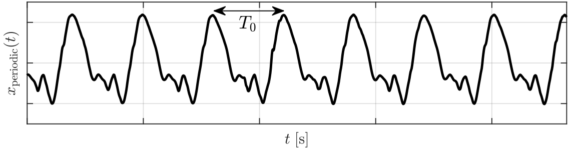
periodic signals: most prominent examples of deterministic signals

$$\begin{aligned} x(t) &= x(t + T_0) \\ f_0 &= \frac{1}{T_0} = \frac{\omega_0}{2\pi} \end{aligned}$$

audio signals
periodic signals 1/5

periodic signals: most prominent examples of deterministic signals

$$x(t) = x(t + T_0)$$
$$f_0 = \frac{1}{T_0} = \frac{\omega_0}{2\pi}$$



matlab source: plotPeriodic.m

audio signals

periodic signals 2/5

periodic signal $\Rightarrow$ representation in **Fourier series**[1]

$$x(t) = \sum_{k=-\infty}^{\infty} a_k e^{\mathrm{j}\omega_0 kt}$$

- $\omega_0 = 2\pi \cdot f_0$
- $k\omega_0$: integer multiples of the lowest frequency
- $e^{\mathrm{j}\omega_0 kt} = \cos(\omega_0 kt) + \mathrm{j}\sin(\omega_0 kt)$
- a_k: Fourier coefficients — amplitude of each component

$$a_k = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} x(t) e^{-\mathrm{j}\omega_0 kt}\, dt$$

[1] Jean-Baptiste Joseph Fourier, 1768–1830

audio signals
periodic signals 2/5

periodic signal ⇒ representation in **Fourier series**[1]

$$x(t) = \sum_{k=-\infty}^{\infty} a_k e^{j\omega_0 kt}$$

- $\omega_0 = 2\pi \cdot f_0$
- $k\omega_0$: integer multiples of the lowest frequency
- $e^{j\omega_0 kt} = \cos(\omega_0 kt) + j\sin(\omega_0 kt)$
- a_k: Fourier coefficients — amplitude of each component

$$a_k = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} x(t) e^{-j\omega_0 kt}\, dt$$

[1] Jean-Baptiste Joseph Fourier, 1768–1830

audio signals
periodic signals 2/5

periodic signal $\Rightarrow$ representation in **Fourier series**[1]

$$x(t) = \sum_{k=-\infty}^{\infty} a_k e^{\mathrm{j}\omega_0 k t}$$

- $\omega_0 = 2\pi \cdot f_0$
- $k\omega_0$: integer multiples of the lowest frequency
- $e^{\mathrm{j}\omega_0 kt} = \cos(\omega_0 kt) + \mathrm{j}\sin(\omega_0 kt)$
- a_k: Fourier coefficients — amplitude of each component

$$a_k = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} x(t) e^{-\mathrm{j}\omega_0 kt}\, dt$$

[1] Jean-Baptiste Joseph Fourier, 1768–1830

audio signals
periodic signals 2/5

periodic signal ⇒ representation in **Fourier series**[1]

$$x(t) = \sum_{k=-\infty}^{\infty} a_k e^{\mathrm{j}\omega_0 kt}$$

- $\omega_0 = 2\pi \cdot f_0$
- $k\omega_0$: integer multiples of the lowest frequency
- $e^{\mathrm{j}\omega_0 kt} = \cos(\omega_0 kt) + \mathrm{j}\sin(\omega_0 kt)$
- a_k: Fourier coefficients — amplitude of each component

$$a_k = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} x(t) e^{-\mathrm{j}\omega_0 kt}\, dt$$

[1] Jean-Baptiste Joseph Fourier, 1768–1830

audio signals
periodic signals 2/5

periodic signal ⇒ representation in **Fourier series**[1]

$$x(t) = \sum_{k=-\infty}^{\infty} a_k e^{\mathrm{j}\omega_0 kt}$$

- $\omega_0 = 2\pi \cdot f_0$
- $k\omega_0$: integer multiples of the lowest frequency
- $e^{\mathrm{j}\omega_0 kt} = \cos(\omega_0 kt) + \mathrm{j}\sin(\omega_0 kt)$
- a_k: Fourier coefficients — amplitude of each component

$$a_k = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} x(t) e^{-\mathrm{j}\omega_0 kt}\, dt$$

[1] Jean-Baptiste Joseph Fourier, 1768–1830

Fourier series

- **every** periodic signal can be represented in a Fourier series
- a periodic signal **contains only** frequencies at integer multiples of the fundamental frequency f_0

- Fourier series can only be applied to periodic signals
- Fourier series is analytically elegant but only of limited practical use as the fundamental period has to be known

audio signals

periodic signals 3/5

Fourier series

- **every** periodic signal can be represented in a Fourier series
- a periodic signal **contains only** frequencies at integer multiples of the fundamental frequency f_0

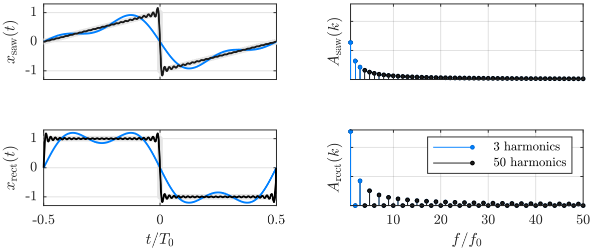
- Fourier series can only be applied to periodic signals
- Fourier series is analytically elegant but only of limited practical use as the fundamental period has to be known

audio signals

periodic signals 4/5

reconstruction of periodic signals with limited number of sinusoidals:

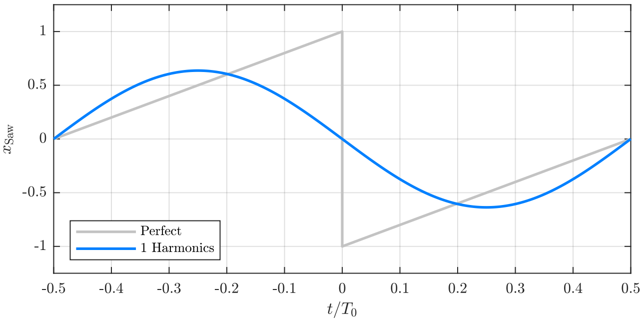
$$\hat{x}(t) = \sum_{k=-\mathcal{K}}^{\mathcal{K}} a_k e^{j\omega_0 kt}$$



matlab source: plotFourierSeries.m

audio signals

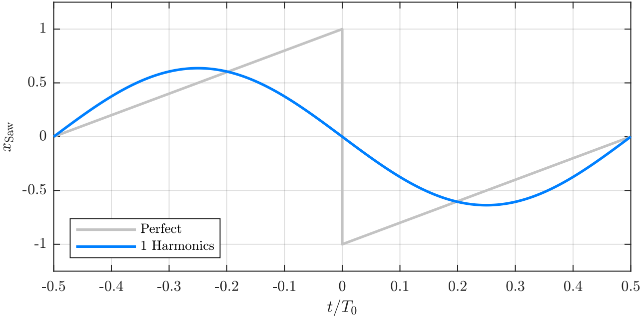
periodic signals 5/5



1 3 1 25 50 harmonics

audio signals

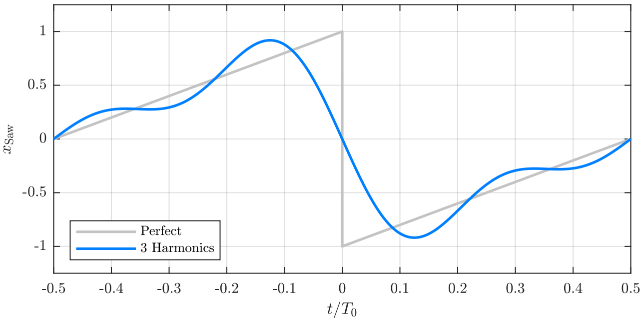
periodic signals 5/5



1 3 1 25 50 harmonics

matlab source: plotAdditiveSynthesis.m

audio signals
periodic signals 5/5

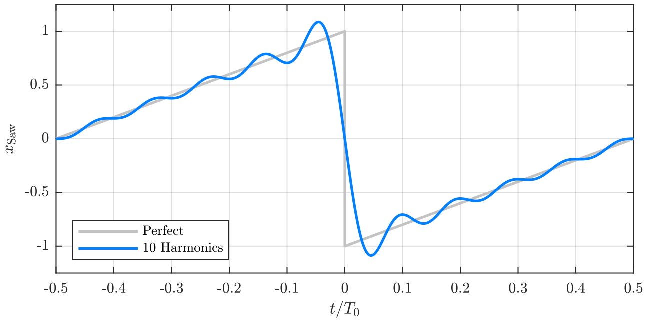


1 3 1 25 50 harmonics

matlab source: plotAdditiveSynthesis.m

audio signals

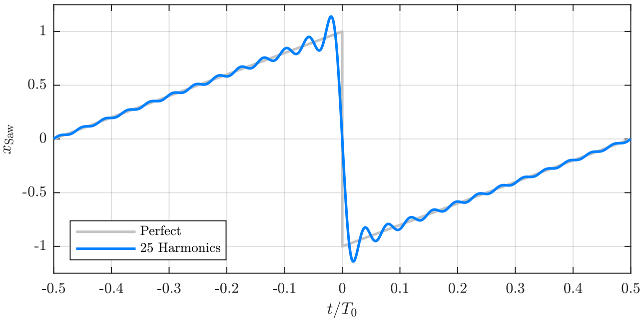
periodic signals 5/5



1 3 1 25 50 harmonics

matlab source: plotAdditiveSynthesis.m

audio signals
periodic signals 5/5

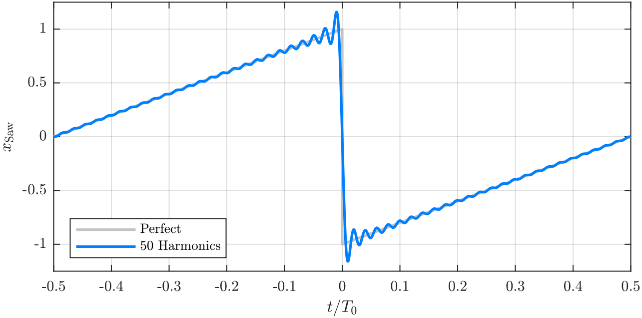


1 3 1 25 50 harmonics

matlab source: plotAdditiveSynthesis.m

audio signals

periodic signals 5/5

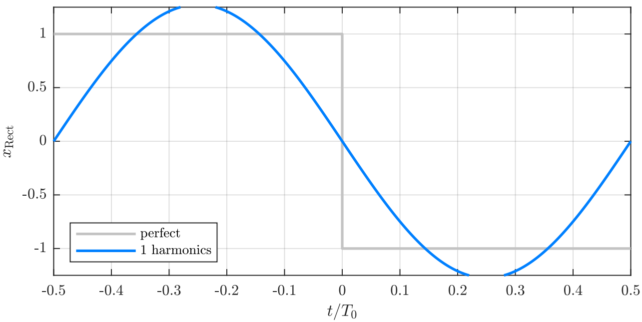


1 3 1 25 50 harmonics

matlab source: plotAdditiveSynthesis.m

audio signals

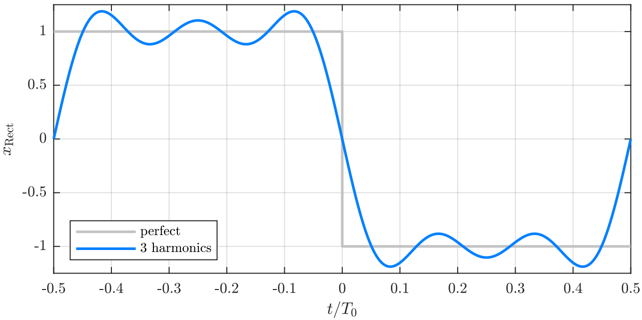
periodic signals 5/5



1 3 10 25 50 harmonics

matlab source: plotAdditiveSynthesis.m

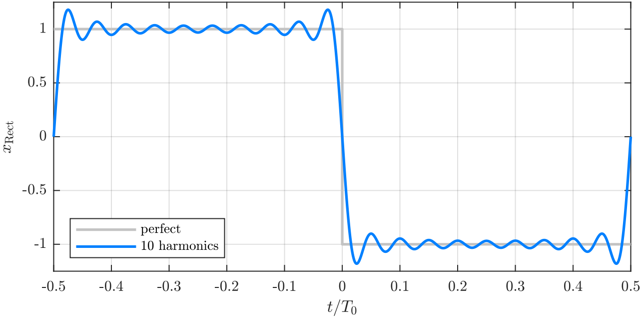
audio signals
periodic signals 5/5



1 3 10 25 50 harmonics

matlab source: plotAdditiveSynthesis.m

audio signals
periodic signals 5/5

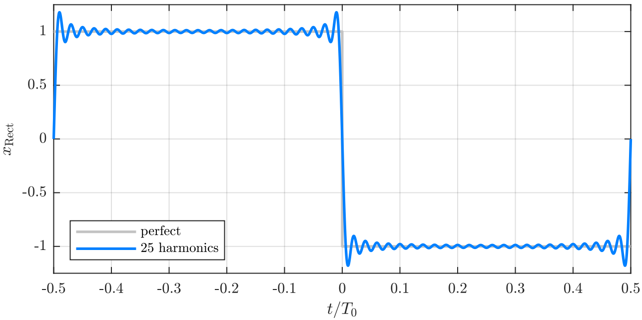


1 3 10 25 50 harmonics

matlab source: plotAdditiveSynthesis.m

audio signals

periodic signals 5/5

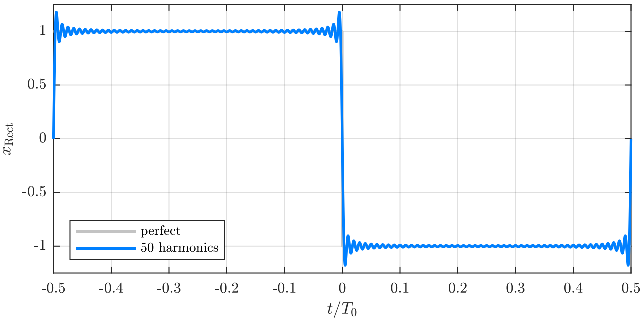


1 3 10 25 50 harmonics

matlab source: plotAdditiveSynthesis.m

audio signals

periodic signals 5/5



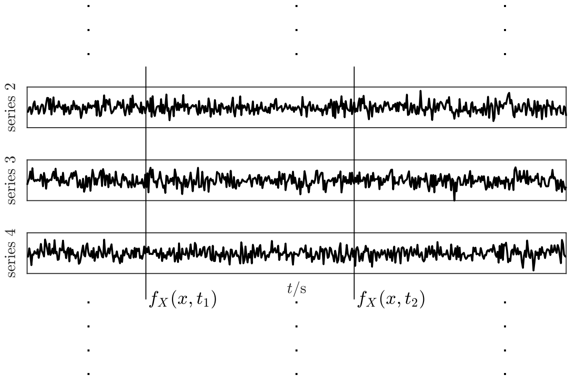
1 3 10 25 50 harmonics

matlab source: plotAdditiveSynthesis.m

audio signals

random process 1/2

random process: ensemble of random series



matlab source: plotRandomProcess.m

audio signals

random process 2/2

random process

- ensemble of random series
- each series represents a *sample* of the process
- the following value is *indetermined*, regardless of any amount of knowledge

- special case: **stationarity**
  statistical properties such as the mean are time invariant
- example: white noise

statistical signal description

probability density function

PDF $p_x(x)$

- x-axis: possible (amplitude) values
- y-axis: probability

$$
\begin{aligned}
p_x(x) &\geq 0, \textit{and} \\
\int\limits_{-\infty}^{\infty} p_x(x)\,dx &= 1
\end{aligned}
$$

RFD—Relative Frequency Distribution (sample of PDF)

histogram of (amplitude) values

statistical signal description

probability density function

PDF $p_x(x)$

- x-axis: possible (amplitude) values
- y-axis: probability

$$\begin{aligned} p_x(x) &\geq 0, \mathit{and} \\ \int\limits_{-\infty}^{\infty} p_x(x)\,dx &= 1 \end{aligned}$$

RFD—Relative Frequency Distribution (sample of PDF)

histogram of (amplitude) values

statistical signal description

probability density function

PDF $p_x(x)$

- x-axis: possible (amplitude) values
- y-axis: probability

$$\begin{aligned} p_x(x) &\geq 0, \textit{and} \\ \int\limits_{-\infty}^{\infty} p_x(x)\,dx &= 1 \end{aligned}$$

RFD—Relative Frequency Distribution (sample of PDF)

histogram of (amplitude) values

statistical signal description

PDF examples

What is the PDF of the following prototype signals:

statistical signal description
PDF examples

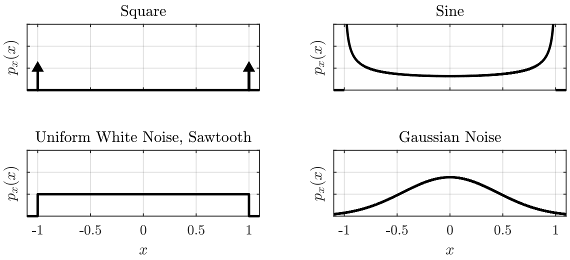
What is the PDF of the following prototype signals:

- square wave
- sawtooth wave
- sine wave
- white noise (uniform, gaussian)
- DC

statistical signal description

PDF examples

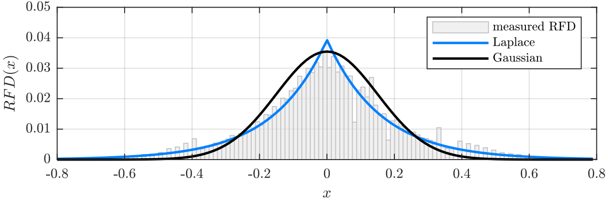
What is the PDF of the following prototype signals:



matlab source: plotPdfExamples.m

statistical signal description

RFD: real world signals



matlab source: plotRfd.m

statistical signal description

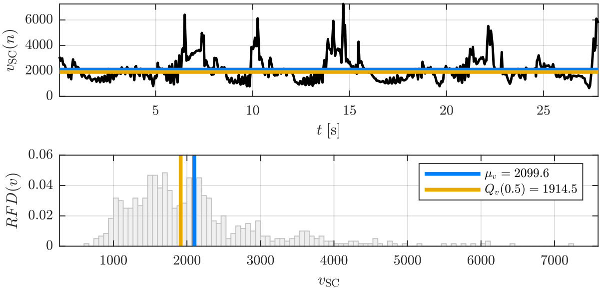
arithmetic mean

- from time series x:

$$\mu_x(n) = \frac{1}{\mathcal{K}} \sum_{i=i_\mathrm{s}(n)}^{i_\mathrm{e}(n)} x(i)$$

- from distribution p_x:

$$\mu_x(n) = \sum_{x=-\infty}^{\infty} x \cdot p_x(x)$$



matlab source: plotArithmeticMean.m

statistical signal description

geometric & harmonic mean

- **geometric mean**

$$\begin{aligned}\mathrm{Mg}_v &= \sqrt[\mathcal{N}]{\prod_{0}^{\mathcal{N}-1} v(n)} \\ &= \exp\left(\frac{1}{\mathcal{N}}\sum_{0}^{\mathcal{N}-1}\log\left(v(n)\right)\right).\end{aligned}$$

- **harmonic mean**

$$\mathrm{Mh}_v = \frac{\mathcal{N}}{\sum\limits_{0}^{\mathcal{N}-1} {}^{1}/{}_{v(n)}}.$$

statistical signal description
geometric & harmonic mean

- **geometric mean**

$$\begin{aligned}\mathrm{Mg}_v &= \sqrt[\mathcal{N}]{\prod_{0}^{\mathcal{N}-1} v(n)} \\ &= \exp\left(\frac{1}{\mathcal{N}}\sum_{0}^{\mathcal{N}-1}\log\left(v(n)\right)\right).\end{aligned}$$

- **harmonic mean**

$$\mathrm{Mh}_v = \frac{\mathcal{N}}{\sum\limits_{0}^{\mathcal{N}-1} {}^{1}/{}_{v(n)}}.$$

statistical signal description
variance & standard deviation

measure of *spread* of the signal around its mean

- **variance**
  - from signal block:

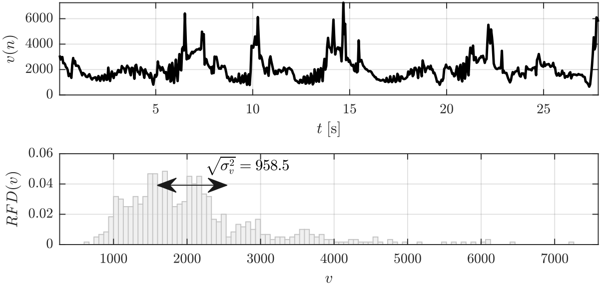
$$\sigma_x^2(n) = \frac{1}{\mathcal{K}} \sum_{i=i_s(n)}^{i_e(n)} \left(x(i) - \mu_x(n)\right)^2$$

  - from distribution:

$$\sigma_x^2(n) = \sum_{x=-\infty}^{\infty} \left(x - \mu_x\right)^2 \cdot p_x(x)$$

- **standard deviation**

$$\sigma_x(n) = \sqrt{\sigma_x^2(n)}$$



matlab source: plotStandardDeviation.m

statistical signal description

variance & standard deviation

measure of *spread* of the signal around its mean

- **variance**
  - from signal block:

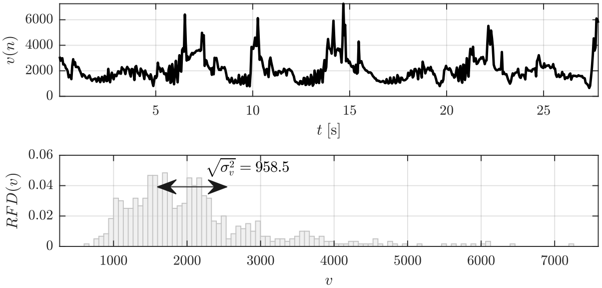
$$\sigma_x^2(n) = \frac{1}{\mathcal{K}} \sum_{i=i_s(n)}^{i_e(n)} \left(x(i) - \mu_x(n)\right)^2$$

  - from distribution:

$$\sigma_x^2(n) = \sum_{x=-\infty}^{\infty} (x - \mu_x)^2 \cdot p_x(x)$$

- **standard deviation**

$$\sigma_x(n) = \sqrt{\sigma_x^2(n)}$$



matlab source: plotStandardDeviation.m

statistical signal description
variance & standard deviation

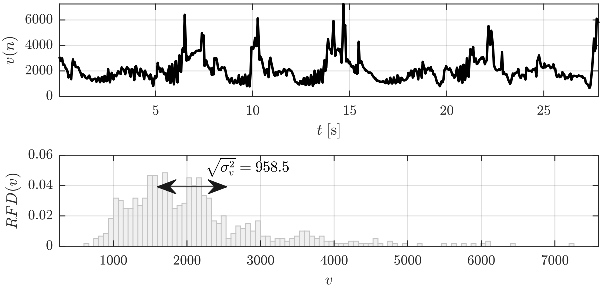
measure of *spread* of the signal around its mean

- **variance**
  - from signal block:

$$\sigma_x^2(n) = \frac{1}{\mathcal{K}} \sum_{i=i_s(n)}^{i_e(n)} \left(x(i) - \mu_x(n)\right)^2$$

  - from distribution:

$$\sigma_x^2(n) = \sum_{x=-\infty}^{\infty} (x - \mu_x)^2 \cdot p_x(x)$$



- **standard deviation**

$$\sigma_x(n) = \sqrt{\sigma_x^2(n)}$$

matlab source: plotStandardDeviation.m

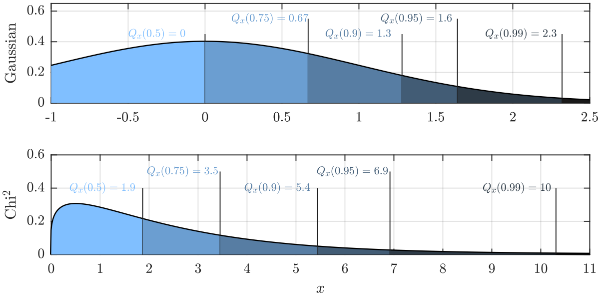
statistical signal description

quantiles & quantile ranges

dividing the PDF into (equal sized) subsets

$$Q_X(c_p) = \operatorname{argmin}\left(F_X(x) \leq c_p\right)$$

$$\textit{with} \quad F_X(x) = \int_{-\infty}^{x} p_x(y)\, dy$$



matlab source: plotPdfQuantiles.m

statistical signal description

quantile examples

- **median**

$$Q_{\mathrm{X}}(0.5) = \mathrm{argmin}\left(F_{\mathrm{X}}(x) \leq 0.5\right)$$

- **quartiles**: $Q_{\mathrm{X}}(0.25)$, $Q_{\mathrm{X}}(0.5)$, and $Q_{\mathrm{X}}x(0.75)$

- **quantile range**, e.g.

$$\Delta Q_{\mathrm{X}}(0.9) = Q_{\mathrm{X}}(0.95) - Q_{\mathrm{X}}(0.05)$$

summary

lecture content

- signals can be categorized into **deterministic and random signals**
  - deterministic signal can be described in a mathematical function
  - random processes can only be described by their general properties
- **periodic signals**
  - periodic signals are probably the most music-related deterministic signal
  - any periodic (pitched) signal is a sum of weighted sinusoidals
  - frequencies *only* at the fundamental frequency and integer multiples
- **random** signals
  - noise, unpredictable
- **real-world** signals
  - can be seen as a time-varying mixture of these two signal categories
- **statistical features**
  - summarize technical signal characteristics in few numerical values
  - may be used on a time domain, frequency domain, or feature domain signal