



Introduction to **Audio Content Analysis**

module 3.1: input representation — signals

alexander lerch

introduction

overview

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



introduction

overview

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

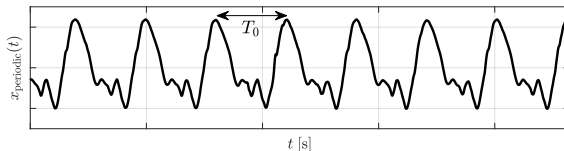
$$x(t) = x(t + T_0)$$
$$f_0 = \frac{1}{T_0} = \frac{\omega_0}{2\pi}$$

audio signals

periodic signals 1/5

periodic signals: most prominent examples of deterministic signals

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



audio signals

periodic signals 2/5

periodic signal \Rightarrow 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

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

- $\omega_0 = 2\pi \cdot f_0$
- $k\omega_0$: 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

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

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

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

¹ Jean-Baptiste Joseph Fourier, 1768–1830

audio signals

periodic signals 2/5

periodic signal \Rightarrow representation in **Fourier series**¹

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

- $\omega_0 = 2\pi \cdot f_0$
- $k\omega_0$: 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

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

- $\omega_0 = 2\pi \cdot f_0$
- $k\omega_0$: 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 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 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

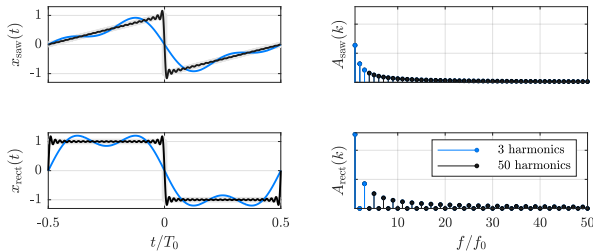


audio signals

periodic signals 4/5

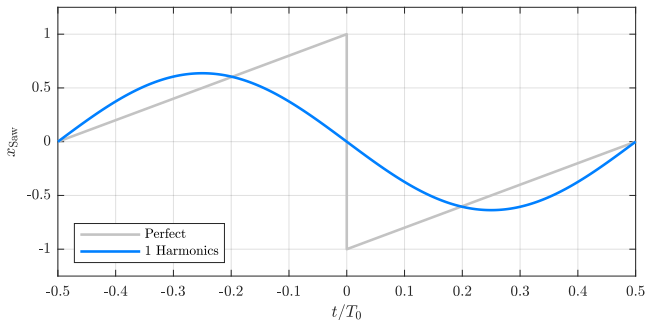
reconstruction of periodic signals with limited number of sinusoidals:

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



audio signals

periodic signals 5/5

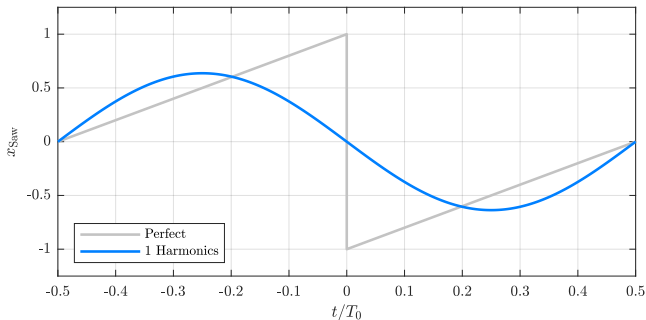


1 3 10 25 50 harmonics

🔊 🔊 🔊 🔊 🔊

audio signals

periodic signals 5/5

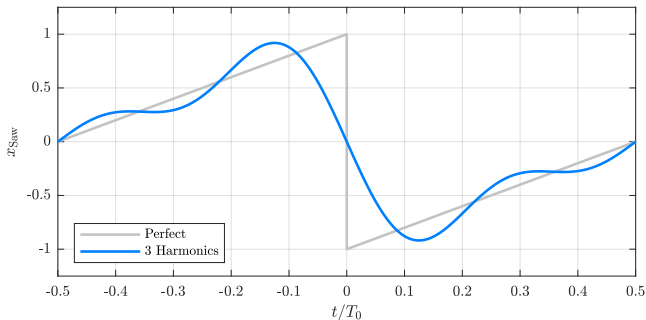


1 3 10 25 50 harmonics

🔊 🔊 🔊 🔊 🔊

audio signals

periodic signals 5/5

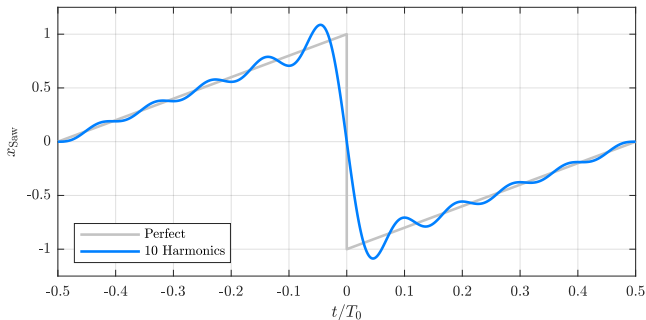


1 3 10 25 50 harmonics

🔊 🔊 🔊 🔊 🔊

audio signals

periodic signals 5/5

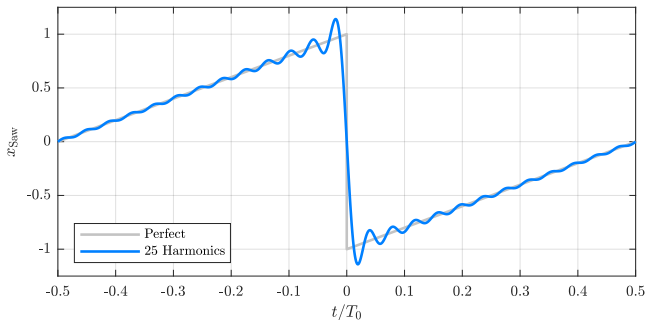


1 3 10 25 50 harmonics

Speaker icons are shown below each number, indicating the audio output for each harmonic count.

audio signals

periodic signals 5/5

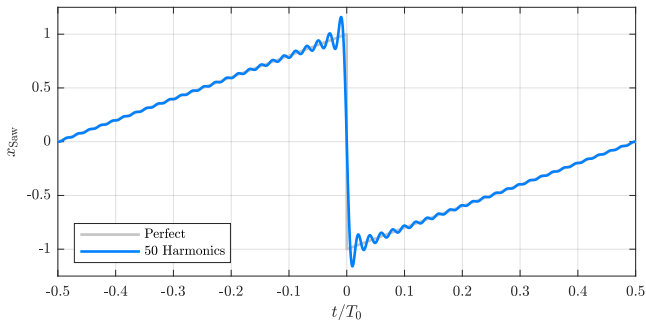


1 3 10 25 50 harmonics

Speaker icons are shown below the numbers 1, 3, 10, 25, and 50, indicating that clicking on these numbers would play the corresponding audio signal.

audio signals

periodic signals 5/5

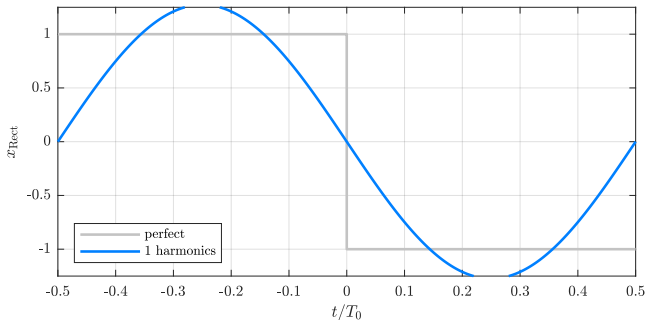


1 3 10 25 50 harmonics

🔊 🔊 🔊 🔊 🔊

audio signals

periodic signals 5/5

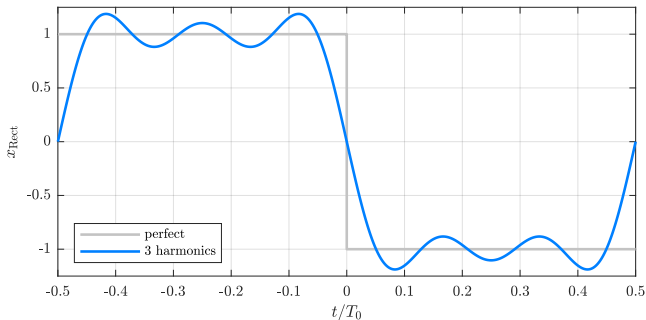


1 3 10 25 50 harmonics

🔊 🔊 🔊 🔊 🔊

audio signals

periodic signals 5/5

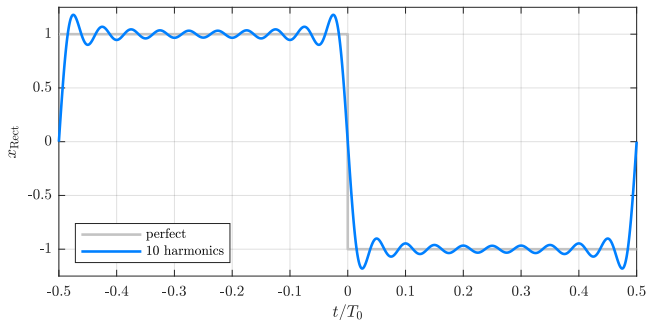


1 3 10 25 50 harmonics

🔊 🔊 🔊 🔊 🔊

audio signals

periodic signals 5/5

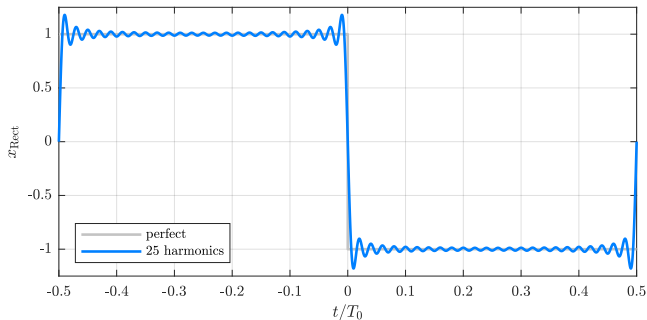


1 3 10 25 50 harmonics

🔊 🔊 🔊 🔊 🔊

audio signals

periodic signals 5/5

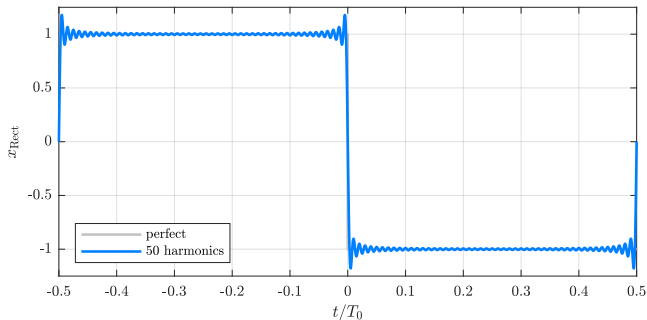


1 3 10 25 50 harmonics

🔊 🔊 🔊 🔊 🔊

audio signals

periodic signals 5/5



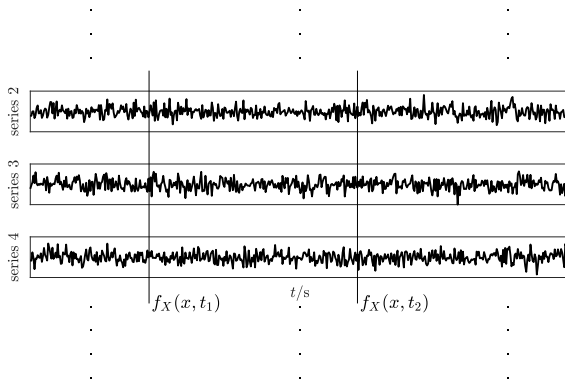
1 3 10 25 50 harmonics

🔊 🔊 🔊 🔊 🔊

audio signals

random process 1/2

random process: ensemble of random series



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

$$p_x(x) \geq 0, \text{ and}$$
$$\int_{-\infty}^{\infty} p_x(x) dx = 1$$

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

$$p_x(x) \geq 0, \text{ and}$$
$$\int_{-\infty}^{\infty} p_x(x) dx = 1$$

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

$$p_x(x) \geq 0, \text{ and}$$
$$\int_{-\infty}^{\infty} p_x(x) dx = 1$$

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:

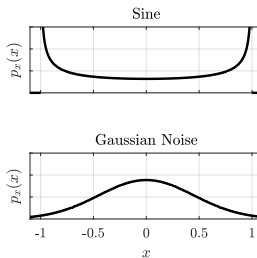
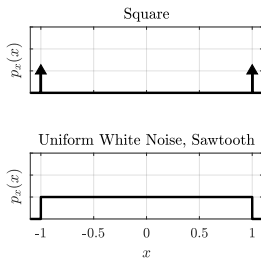
- square wave
- sawtooth wave
- sine wave
- white noise (uniform, gaussian)
- DC

statistical signal description

PDF examples

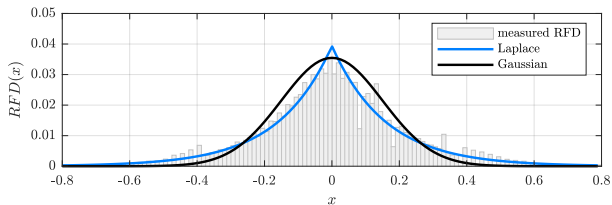


What is the PDF of the following prototype signals:



statistical signal description

RFD: real world signals



statistical signal description

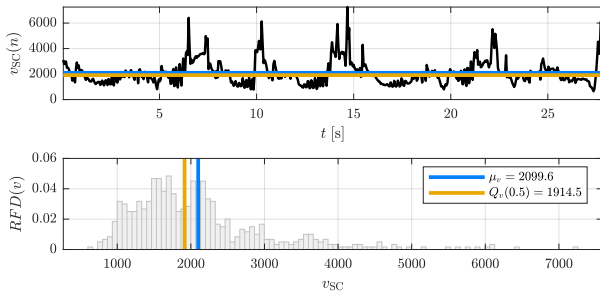
arithmetic mean

- from time series x :

$$\mu_x(n) = \frac{1}{K} \sum_{i=i_s(n)}^{i_e(n)} x(i)$$

- from distribution p_x :

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



statistical signal description

geometric & harmonic mean

■ geometric mean

$$\begin{aligned} \text{Mg}_v &= \sqrt[\mathcal{N}]{\prod_0^{\mathcal{N}-1} v(n)} \\ &= \exp \left(\frac{1}{\mathcal{N}} \sum_0^{\mathcal{N}-1} \log(v(n)) \right). \end{aligned}$$

■ harmonic mean

$$\text{Mh}_v = \frac{\mathcal{N}}{\sum_0^{\mathcal{N}-1} 1/v(n)}.$$

statistical signal description

geometric & harmonic mean

■ geometric mean

$$\begin{aligned} \text{Mg}_v &= \sqrt[\mathcal{N}]{\prod_0^{\mathcal{N}-1} v(n)} \\ &= \exp \left(\frac{1}{\mathcal{N}} \sum_0^{\mathcal{N}-1} \log(v(n)) \right). \end{aligned}$$

■ harmonic mean

$$\text{Mh}_v = \frac{\mathcal{N}}{\sum_0^{\mathcal{N}-1} 1/v(n)}.$$

statistical signal description

variance & standard deviation

measure of *spread* of the signal around its mean

■ variance

- from signal block:

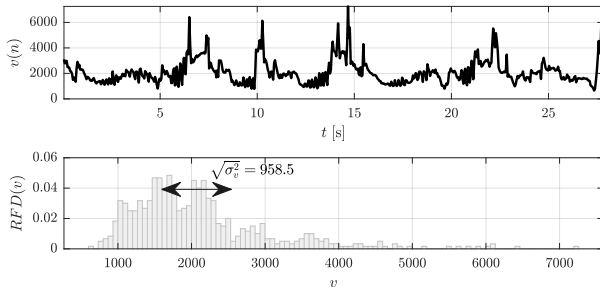
$$\sigma_x^2(n) = \frac{1}{K} \sum_{i=i_s(n)}^{i_e(n)} (x(i) - \mu_x(n))^2$$

- from distribution:

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

■ standard deviation

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



statistical signal description

variance & standard deviation

measure of *spread* of the signal around its mean

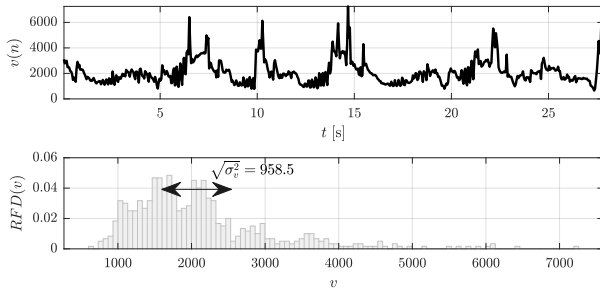
■ variance

- from signal block:

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

- from distribution:

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



■ standard deviation

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

statistical signal description

variance & standard deviation

measure of *spread* of the signal around its mean

■ variance

- from signal block:

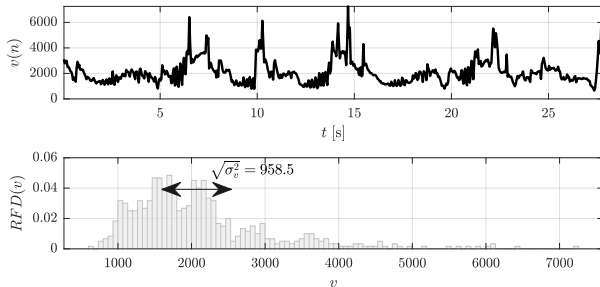
$$\sigma_x^2(n) = \frac{1}{K} \sum_{i=i_s(n)}^{i_e(n)} (x(i) - \mu_x(n))^2$$

- from distribution:

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

■ standard deviation

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



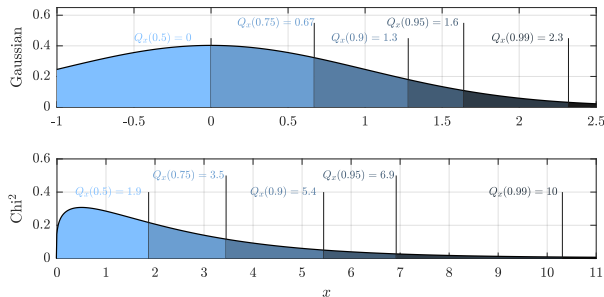
statistical signal description

quantiles & quantile ranges

dividing the PDF into (equal sized) subsets

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

$$\text{with } F_X(x) = \int_{-\infty}^x p_X(y) dy$$



statistical signal description

quantile examples

■ median

$$Q_X(0.5) = \operatorname{argmin} (F_X(x) \leq 0.5)$$

■ quartiles: $Q_X(0.25)$, $Q_X(0.5)$, and $Q_X(0.75)$

■ quantile range, e.g.

$$\Delta Q_X(0.9) = Q_X(0.95) - Q_X(0.05)$$

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

