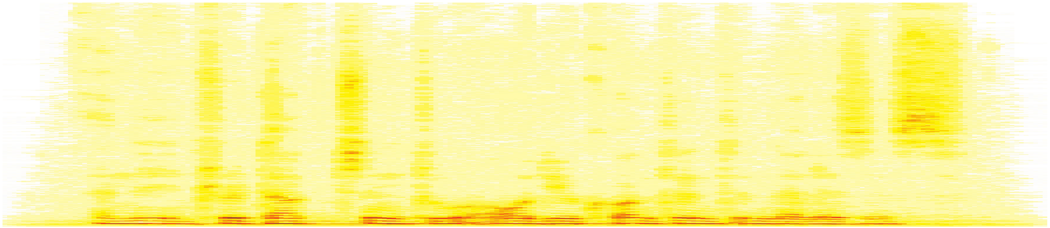
Introduction to Audio Content Analysis

Module 3.1: Feature Extraction — Statistical Features

alexander lerch



Georgia Tech Center for Music Technology
College of Design

corresponding textbook section

Chapter 3 — Instantaneous Features: pp. 35–41

- **lecture content**
  - introduction to statistical features
    - mean
    - variance and standard deviation
    - quantiles
- **learning objectives**
  - give examples of where statistical features can be used
  - describe the meaning of the introduced statistical features

corresponding textbook section

Chapter 3 — Instantaneous Features: pp. 35–41

- **lecture content**
  - introduction to statistical features
    - mean
    - variance and standard deviation
    - quantiles
- **learning objectives**
  - give examples of where statistical features can be used
  - describe the meaning of the introduced statistical features

- statistical features are...
  - numerical descriptors of statistical properties of the signal or its PDF
  - general features that are not audio-related
- **usage**
  - often not directly applied to audio signal but to feature representations of it

- from signal x:

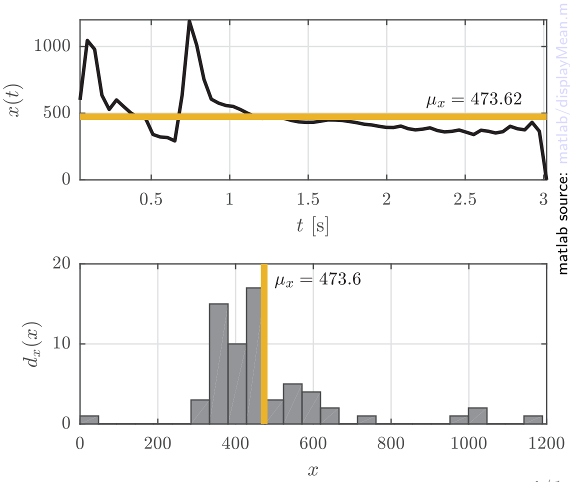
$$\mu_x(n) = \frac{1}{\mathcal{K}} \sum_{i=i_s(n)}^{i_e(n)} x(i)$$

- from distribution p_x:

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

- from unnormalized distribution d_x:

$$\mu_x(n) = \frac{\sum\limits_{x=-\infty}^{\infty} x \cdot d_x(x)}{\sum\limits_{x=-\infty}^{\infty} d_x(x)}$$



- from signal x:

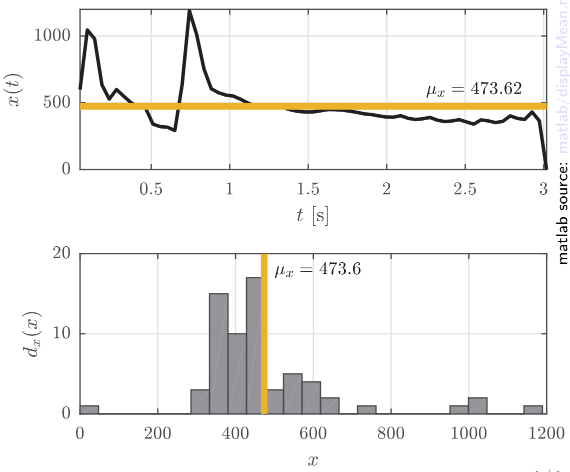
$$\mu_x(n) = \frac{1}{\mathcal{K}} \sum_{i=i_s(n)}^{i_e(n)} x(i)$$

- from distribution p_x:

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

- from unnormalized distribution d_x:

$$\mu_x(n) = \frac{\sum\limits_{x=-\infty}^{\infty} x \cdot d_x(x)}{\sum\limits_{x=-\infty}^{\infty} d_x(x)}$$



- from signal x:

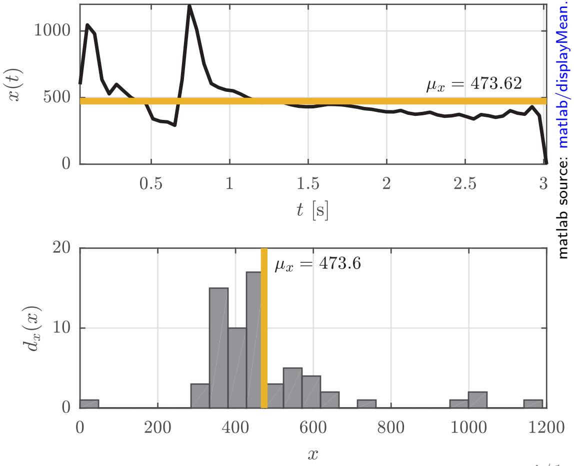
$$\mu_x(n) = \frac{1}{\mathcal{K}} \sum_{i=i_\mathrm{s}(n)}^{i_\mathrm{e}(n)} x(i)$$

- from distribution p_x:

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

- from unnormalized distribution d_x:

$$\mu_x(n) = \frac{\sum\limits_{x=-\infty}^{\infty} x \cdot d_x(x)}{\sum\limits_{x=-\infty}^{\infty} d_x(x)}$$



matlab source: matlab/displayMean.m

- **geometric mean**

$$\begin{aligned} M_x(0,n) &= \sqrt[\mathcal{K}]{\prod_{i=i_s(n)}^{i_e(n)} x(i)} \\ &= \exp\left(\frac{1}{\mathcal{K}}\sum_{i=i_s(n)}^{i_e(n)} \log\left[x(i)\right]\right) \end{aligned}$$

- **harmonic mean**

$$M_x(-1,n) = \frac{\mathcal{K}}{\sum\limits_{i=i_s(n)}^{i_e(n)} {}^{1}/_{x(i)}}$$

- **geometric mean**

$$M_x(0,n) = \sqrt[\mathcal{K}]{\prod_{i=i_s(n)}^{i_e(n)} x(i)} = \exp\left(\frac{1}{\mathcal{K}}\sum_{i=i_s(n)}^{i_e(n)} \log\left[x(i)\right]\right)$$

- **harmonic mean**

$$M_x(-1,n) = \frac{\mathcal{K}}{\sum\limits_{i=i_s(n)}^{i_e(n)} {}^{1}\!/\!{}_{x(i)}}$$

Georgia Tech Center for Music Technology
College of Design

- **generalized mean**

$$M_x(\beta, n) = \sqrt[\beta]{\frac{1}{\mathcal{K}} \sum_{i=i_s(n)}^{i_e(n)} x^{\beta}(i)}$$

  - $\beta = 1$: arithmetic mean
  - $\beta = 2$: quadratic mean
  - $\beta = -1$: harmonic mean
  - $\beta \to 0$: geometric mean
  - $\beta \to -\infty$: minimum
  - $\beta \to \infty$: maximum

- **generalized mean**

$$M_x(\beta, n) = \sqrt[\beta]{\frac{1}{\mathcal{K}} \sum_{i=i_\mathrm{s}(n)}^{i_\mathrm{e}(n)} x^\beta(i)}$$

- $\beta = 1$: arithmetic mean
- $\beta = 2$: quadratic mean
- $\beta = -1$: harmonic mean
- $\beta \to 0$: geometric mean
- $\beta \to -\infty$: minimum
- $\beta \to \infty$: maximum

centroid: *center of gravity*

$$v_{\mathrm{C}}(n) = \frac{\sum\limits_{i=i_{\mathrm{s}}(n)}^{i_{\mathrm{e}}(n)} \big(i - i_{\mathrm{s}}(n)\big) \cdot x(i)}{\sum\limits_{i=i_{\mathrm{s}}(n)}^{i_{\mathrm{e}}(n)} x(i)}$$

centroid: *center of gravity*

$$v_\mathrm{C}(n) = \frac{\sum\limits_{i=i_\mathrm{s}(n)}^{i_\mathrm{e}(n)} \big(i - i_\mathrm{s}(n)\big) \cdot x(i)}{\sum\limits_{i=i_\mathrm{s}(n)}^{i_\mathrm{e}(n)} x(i)}$$

Why does this look familiar?

centroid: *center of gravity*

$$v_{\mathrm{C}}(n) = \frac{\sum\limits_{i=i_{\mathrm{s}}(n)}^{i_{\mathrm{e}}(n)} \left(i - i_{\mathrm{s}}(n)\right) \cdot x(i)}{\sum\limits_{i=i_{\mathrm{s}}(n)}^{i_{\mathrm{e}}(n)} x(i)}$$

Why does this look familiar?

→ compare arithmetic mean

measure of *spread* of the signal around its mean

- **variance**
  - from signal block:

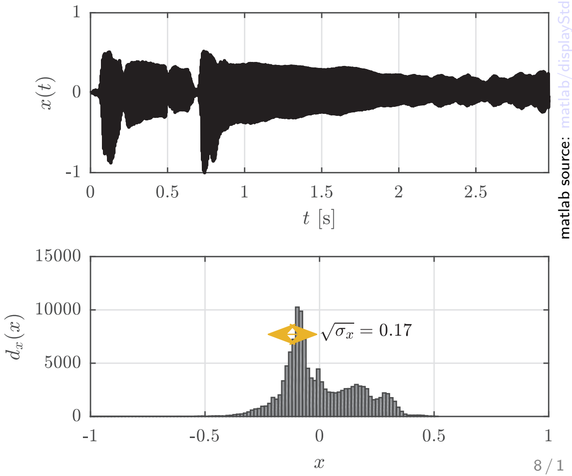
$$\sigma_x^2(n) = \frac{1}{\mathcal{K}} \sum_{i=i_\mathrm{s}(n)}^{i_\mathrm{e}(n)} \left(x(i) - \mu_x(n)\right)^2$$

  - from distribution:

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

- **standard deviation**

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



statistical features

variance & standard deviation

measure of *spread* of the signal around its mean

- **variance**
  - from signal block:

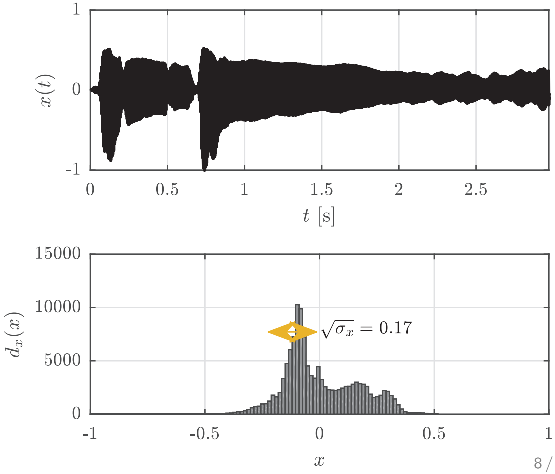
$$\sigma_x^2(n) = \frac{1}{\mathcal{K}} \sum_{i=i_\mathrm{s}(n)}^{i_\mathrm{e}(n)} \left(x(i) - \mu_x(n)\right)^2$$

  - 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: matlab/displayStd.m

measure of *spread* of the signal around its mean

- **variance**
  - from signal block:

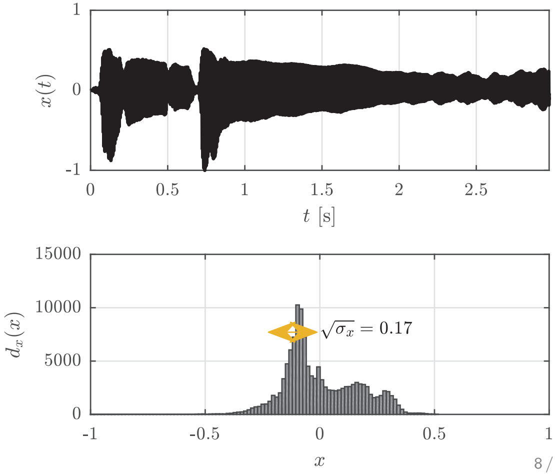
$$\sigma_x^2(n) = \frac{1}{\mathcal{K}} \sum_{i=i_s(n)}^{i_e(n)} \left(x(i) - \mu_x(n)\right)^2$$

  - from distribution:

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

- **standard deviation**

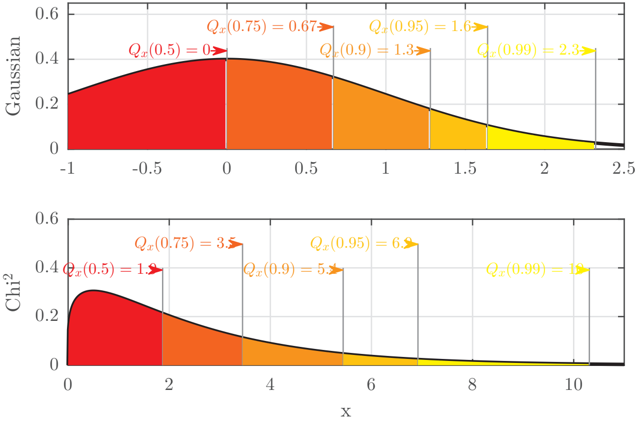
$$\sigma_x(n) = \sqrt{\sigma_x^2(n)}$$



matlab source: matlab/displayStd.m

dividing the PDF into (equal sized) subsets

$$\begin{aligned} Q_{\mathrm{X}}(c_p) &= \mathrm{argmin}\left(F_{\mathrm{X}}(x) \leq c_p\right) \\ \textit{with} \quad F_{\mathrm{X}}(x) &= \int\limits_{-\infty}^{x} p_{\mathrm{x}}(y)\,dy \end{aligned}$$



matlab source: matlab/displayPdfQuantiles.m

Georgia Tech Center for Music Technology
College of Design

- **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.

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

- **statistical features**
  - summarize technical signal characteristics in few numerical values
  - may be used on a time domain, frequency domain, or feature domain signal
- **feature description**
  - *mean*: average value
  - *variance* and *standard deviation*: measure of expected deviation from the mean
  - *quantiles*: numerical pdf shape description