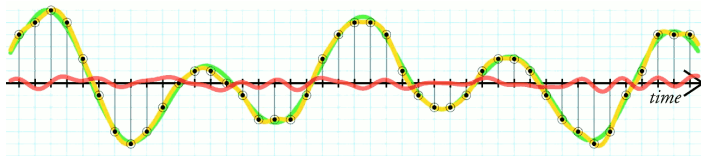# COMS20011 – Data-Driven Computer Science

February 2021
Majid Mirmehdi, Rui Ponte Costa & Dima Damen

Lecture Video #3
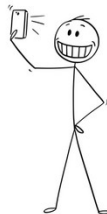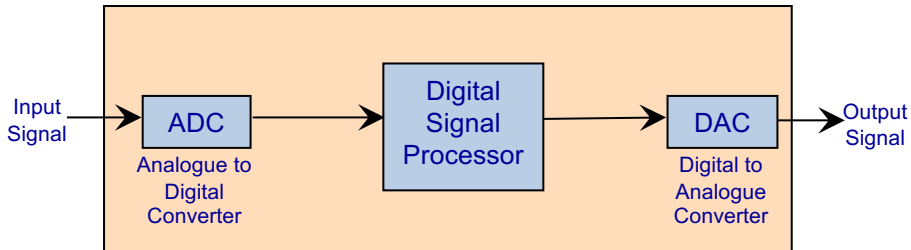
This lecture

Analog Signal

Digital Signal

> ➤ **Data acquisition**
> ➤ Data characteristics: distance measures
> ➤ Data characteristics: summary statistics [*reminder*]
> ➤ Data normalisation and outliers

Data Acquisition – Example data journey

Input Signal → **ADC** (Analogue to Digital Converter) → **Digital Signal Processor** → **DAC** (Digital to Analogue Converter) → Output Signal

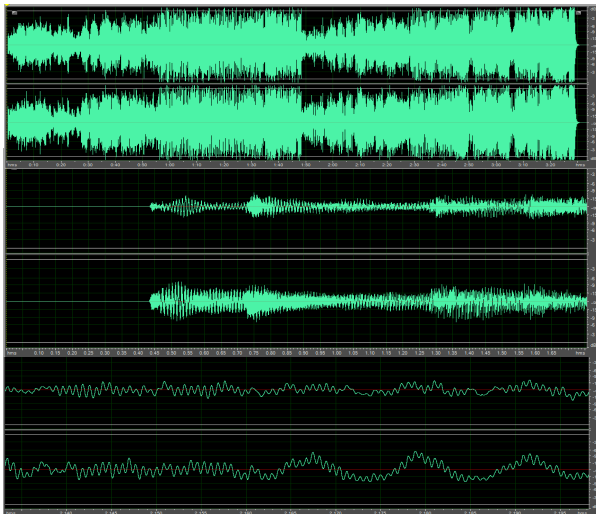# Data Acquisition - Analogue to Digital Conversion

Analogue to Digital conversion involves *Sampling & Quantisation*
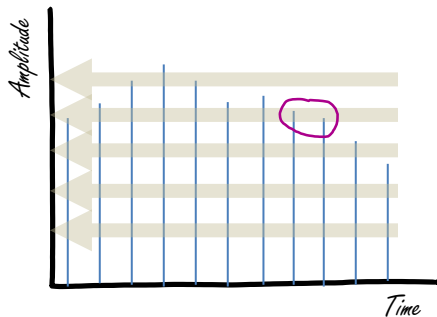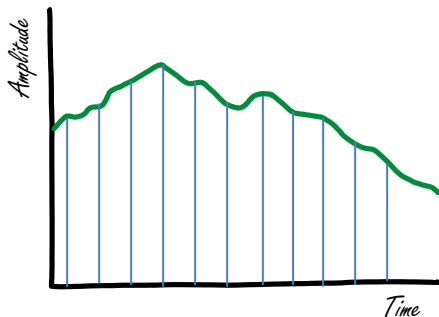
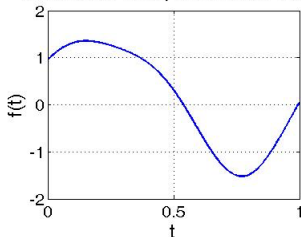e.g. a 1D Audio Signal

low zoom

medium zoom

large zoom

Data Acquisition - Analogue to Digital Conversion

Analogue to Digital conversion involves *Sampling & Quantisation*

Sample and Quantise
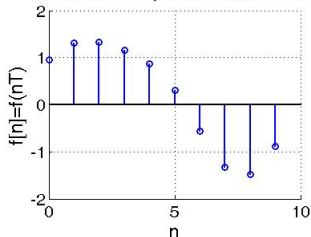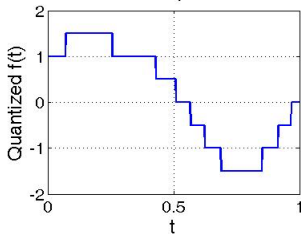
Nyquist-Shannon Sampling Theory
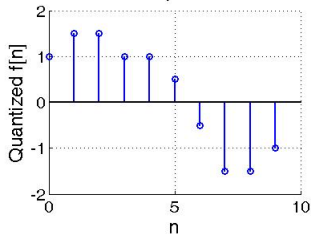
*"An analogue signal containing components up to some maximum frequency **u** (Hz) may be completely reconstructed by regularly spread samples, provided the sampling rate is at least **2u** samples per second"*

Nyquist-Shannon Sampling Theory

"An analogue signal containing components up to some maximum frequency u (Hz) may be completely reconstructed by regularly spread samples, provided the sampling rate is at least 2u samples per second"

Also referred to as the Nyquist-Shannon criterion: sampling rate s should be at least twice the highest spatial frequency u.

$$\sin(\frac{2\pi t}{T_m})$$

sampling period $\quad T_m \leq \dfrac{1}{2u}$

equivalent to *sampling rate* $\quad s \geq 2u$
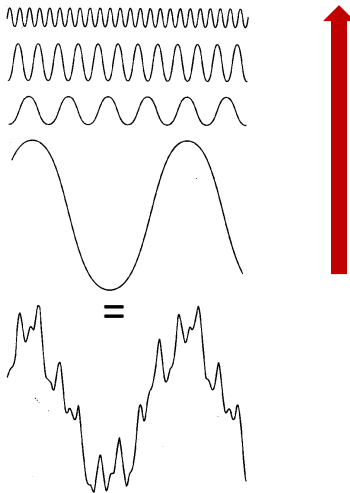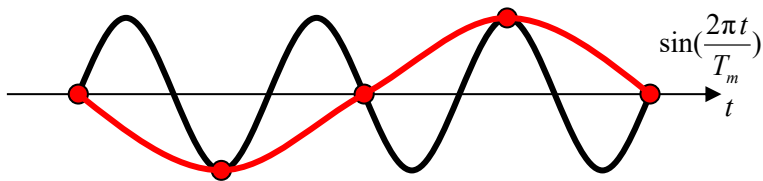
Nyquist-Shannon Sampling Theory

"An analogue signal containing components up to some maximum frequency u (Hz) may be completely reconstructed by regularly spread samples, provided the sampling rate is at least $2u$ samples per second"

Also referred to as the Nyquist-Shannon criterion: sampling rate s should be at least twice the highest spatial frequency u.

$$\sin\left(\frac{2\pi t}{T_m}\right)$$

sampling period $\quad T_m \leq \dfrac{1}{2u}$

equivalent to *sampling rate* $\quad s \geq 2u$

Nyquist-Shannon Sampling Theory

*"An analogue signal containing components up to some maximum frequency **u** (Hz) may be completely reconstructed by regularly spread samples, provided the sampling rate is at least **2u** samples per second"*

Also referred to as the Nyquist-Shannon criterion: sampling rate s should be at least twice the highest spatial frequency u.
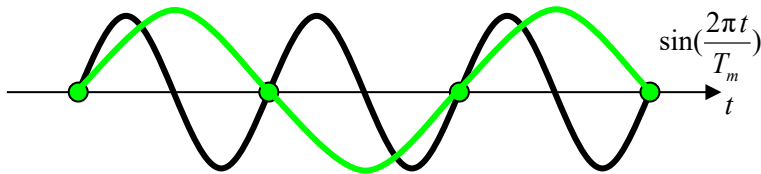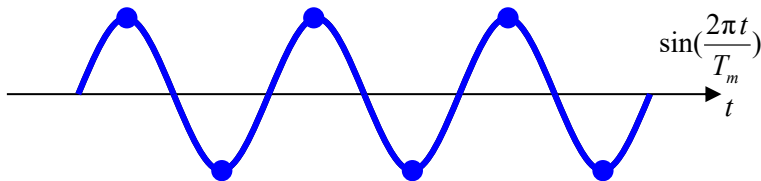

$$\sin(\frac{2\pi t}{T_m})$$

sampling period $\quad T_m \leq \dfrac{1}{2u}$

equivalent to *sampling rate* $\quad s \geq 2u$

Data Acquisition - Analogue to Digital Conversion

Examples of sampling and quantisation of standard audio formats:

- ➢ Speech (e.g. phone call)
    - ➢ Sampling: 8 KHz samples
    - ➢ Quantisation: 8 bits / sample
- ➢ Audio CD
    - ➢ Sampling: 44 KHz samples
    - ➢ Quantisation: 16 bits / sample
    - ➢ Stereo (2 channels)

> Higher sampling and quantisation levels achieves better signal quality, but at the expense of larger memory and storage.

Data Acquisition - Analogue to Digital Conversion

Examples of sampling and quantisation of Images - Multi-Dimensional:

- ➤ Sampling: Resolution in digital photography
- ➤ Quantisation: Representation of each pixel in the image

  - ➤ 8 Mega Pixel Camera: 3264 x 2448 pixels
  - ➤ Colour images: 3 channels: Red, Green, Blue (8 bits per colour)
  - ➤ Greyscale images: 1 channel: intensity $= aR+bG+cB$ *where a+b+c=1.0*
  - ➤ Binary images: Black/White 1 bit per pixel

Sampling – visual example

The effect of sparser sampling…is ALIASING

256x256   64x64   32x32

Anti-aliasing is achieved by filtering to remove frequencies above the Nyquist limit.

Quantisation – visual example

This results from representing a continuously varying function *f(x)* with a discrete one using quantisation levels

16 levels   6 levels   2 levels

Next lecture video

Analog Signal

Digital Signal

- ➤ Data acquisition
- ➤ **Data characteristics: distance measures**
- ➤ Data characteristics: summary statistics [*reminder*]
- ➤ Data normalisation and outliers