Instrumentation and data acquisition Spring 2010

Lecture 5: Noise, precision and accuracy

Zheng-Hua Tan

Department of Electronic Systems
Aalborg University, Denmark
zt@es.aau.dk

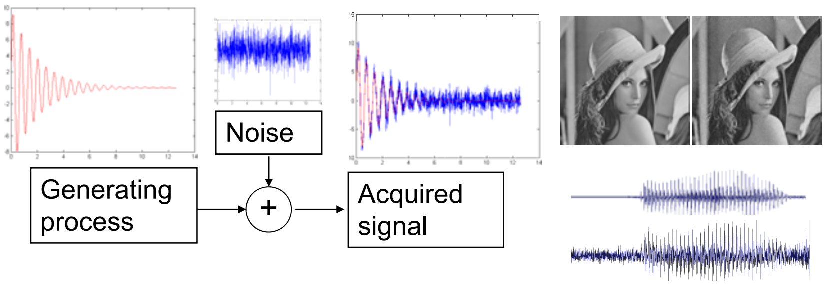
Thanks to Christian Fischer Pedersen for providing some of the slides.

Part I: The normal distribution

- The normal distribution
- Digital noise generation
- Precision and accuracy of data acquisition
- Data acquisition system

Signal acquisition

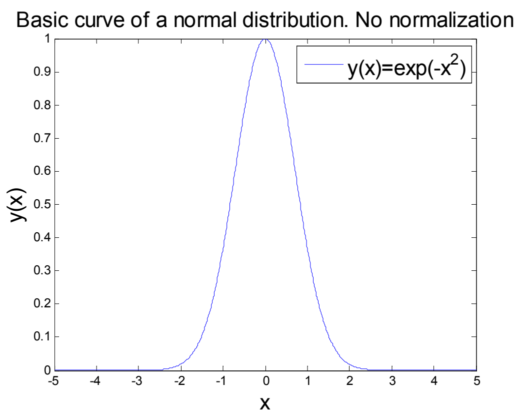
- When acquiring a physical signal, the particular measurement could be affected by many factors: electronic (random) noise, interfering signals and inaccuracies.



- Statistics and probability for characterizing signal & noise, reducing the noise, etc.
  - The normal distribution
  - Precision and accuracy

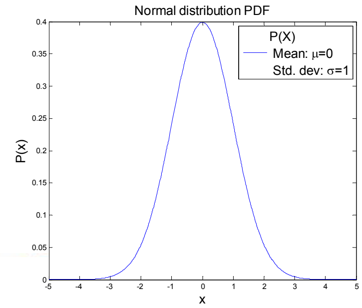
The normal distribution

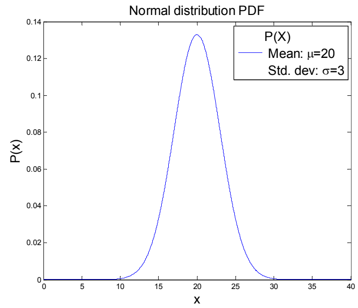
- Normal distribution occur frequently in nature
- Signals from random processes usually have a bell-shaped PDF, called a normal (Gauss) distribution, or a Gaussian.



The normal distribution

- Raw un-normalized curve: $y(x) = e^{-x^2}$
- To get the complete normal distribution PDF
  - Add adjustable mean, μ
  - Add adjustable standard deviation, σ
  - Normalize the curve – area under curve equals to 1
- Normal distribution, PDF: $P(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-(x-\mu)^2/2\sigma^2}$





Cumulative distribution function

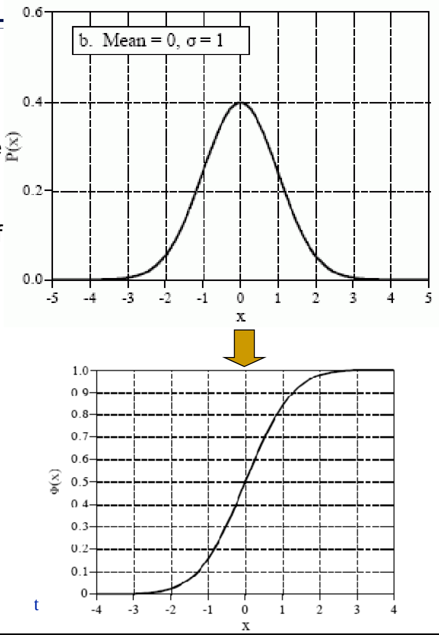
→ Integral of PDF gives probability of samples being within a certain range

→Integral of the PDF is called cumulative distribution function, CDF

→High resolution numerical integration of the Gaussian PDF, e.g. -10σ≤x≤10σ

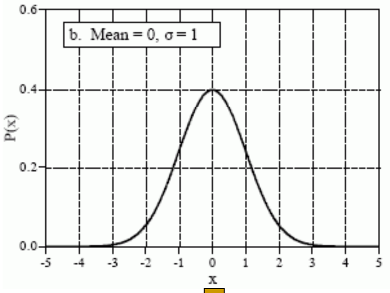
→Store results in table for looking up probabilities (see next slide)

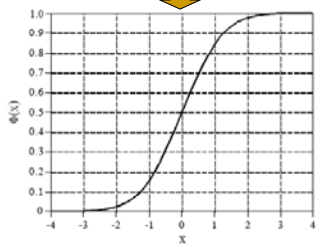
→Φ(x) (*phi*) is the probability that the value of a normally distributed signal at a randomly chosen time will be less than x (expressed in standard deviations referenced to the mean).



Cumulative distribution function

For example, Φ(-1) has a value of 0.1587, indicating that there is a 15.87% probability that the value of the signal will be between -∞ and one standard deviation below the mean, at any randomly chosen time.





x	Φ(x)	x	Φ(x)
-3.4	.0003	0.0	.5000
-3.3	.0005	0.1	.5398
-3.2	.0007	0.2	.5793
-3.1	.0010	0.3	.6179
-3.0	.0013	0.4	.6554
-2.9	.0019	0.5	.6915
-2.8	.0026	0.6	.7257
-2.7	.0035	0.7	.7580
-2.6	.0047	0.8	.7881
-2.5	.0062	0.9	.8159
-2.4	.0082	1.0	.8413
-2.3	.0107	1.1	.8643
-2.2	.0139	1.2	.8849
-2.1	.0179	1.3	.9032
-2.0	.0228	1.4	.9192
-1.9	.0287	1.5	.9332
-1.8	.0359	1.6	.9452
-1.7	.0446	1.7	.9554
-1.6	.0548	1.8	.9641
-1.5	.0668	1.9	.9713
-1.4	.0808	2.0	.9772
-1.3	.0968	2.1	.9821
-1.2	.1151	2.2	.9861
-1.1	.1357	2.3	.9893
-1.0	.1587	2.4	.9918
-0.9	.1841	2.5	.9938
-0.8	.2119	2.6	.9953
-0.7	.2420	2.7	.9965
-0.6	.2743	2.8	.9974
-0.5	.3085	2.9	.9981
-0.4	.3446	3.0	.9987
-0.3	.3821	3.1	.9990
-0.2	.4207	3.2	.9993
-0.1	.4602	3.3	.9995
0.0	.5000	3.4	.9997

Gaussian Mixture Model

- **GMM: convex combination of Gaussian densities**
- **Gaussian density for d-dimesional space:**

$$\phi(\boldsymbol{x};\theta) = \frac{1}{(2\pi)^{d/2}|\boldsymbol{\Sigma}|^{1/2}}\exp\{\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^t\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\}$$
$$\boldsymbol{\mu} = E[\boldsymbol{x}]$$
$$\boldsymbol{\Sigma} = E[(\boldsymbol{x}-\boldsymbol{\mu})^t(\boldsymbol{x}-\boldsymbol{\mu})]$$

Where θ is model with mean μ and covariace matrix Σ

- **k-component GMM: linear combination of k-Gaussian pdfs:**

$$f_k(\boldsymbol{x}) = \sum_{j=1}^{k} \pi_j \phi(\boldsymbol{x};\theta_j)$$

Where,

$$\sum_{j=1}^{k} \pi_j = 1 \quad and$$
$$\pi_j \geq 0 : j \in 1, ...k$$

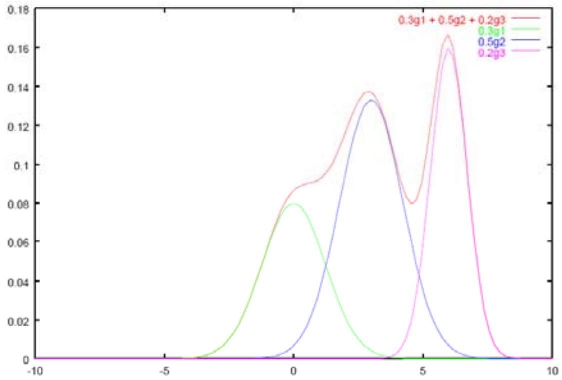
π_j: Mixing Weight

$\phi(\boldsymbol{x};\theta_j)$: Mixture Component

(C. K. Raut, 2004)

Gaussian Mixture Model

- Mixture weight π_j has a form of 'a priori' probability: shows relative importance of each component
- Component Gaussians can have full, diagonal or spherical covariance matrix



(C. K. Raut, 2004)

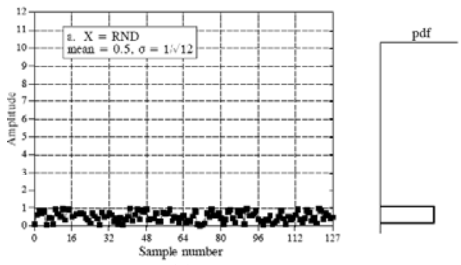
Part II: Digital noise generation

- The normal distribution
- Digital noise generation
- Precision and accuracy of data acquisition
- Data acquisition system

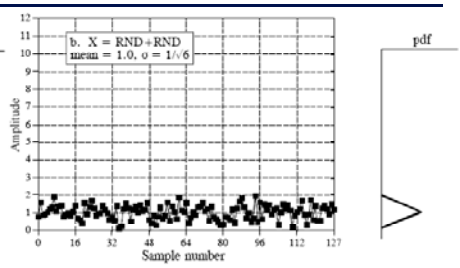
Random noise

- An important topic in both electronics and DSP.
- It limits how small of signals an instrument will measure.
- Digital noise generation is required to test the performance of algorithms that must work in the presence of noise.
  - It relies on random number generator

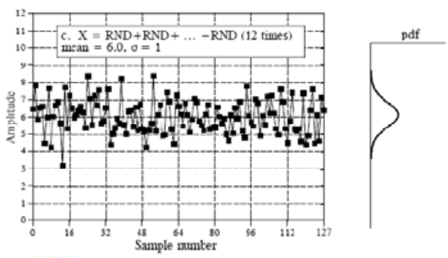
From a uniform distribution to a Gaussian



(1)



(2)



(3)

Adding more and more iid random variables yields a normal distribution.

To generate a normally distributed noise signal with arbitrary mean and σ, for each sample in the signal:

1) Add 12 random numbers
2) Subtract 6 to make the mean zero
3) Multiply by desired σ
4) Add your desired mean

Central limit theorem

- Central limit theorem is the mathematical basis for the aforementioned algorithm (procedure) that creates a normally distributed noise.
- Central limit theorem: A sum of random numbers becomes normally distributed as more and more of the random numbers are added together.
  - This also explains why normally distributed signals are seen so widely in nature.
  - So, when many different random forces are interacting, the resulting PDF is Gaussian/Normal; and this happens often in real physical systems.

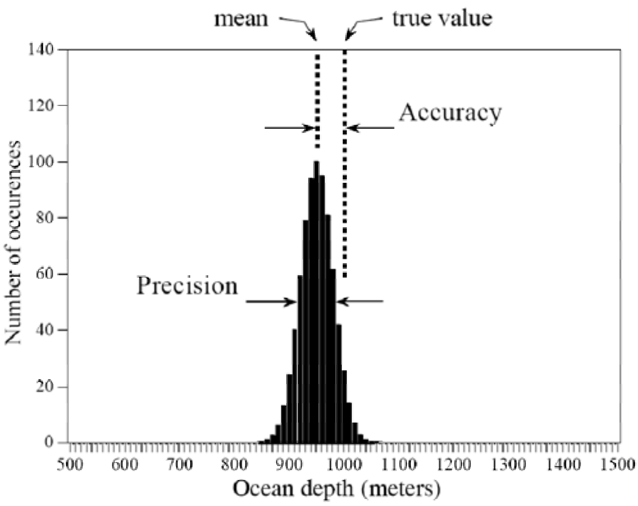
Part III: Precision and accuracy

- The normal distribution
- Digital noise generation
- Precision and accuracy of data acquisition
- Data acquisition system

Precision and accuracy

- Precision and accuracy are terms used to describe systems and methods that measure, estimate or predict
  - Objective: Get TRUE value of parameter
  - Reality: Get MEASURED value of parameter
- Discrepancy between true and measured value can be quantified by accuracy and precision.
  - Accuracy: The mean of the measurements are shifted from the true value. The amount of shift is called accuracy of measurement.
  - Precision: The standard deviation of the measurements indicates the precision, i.e. a slender or broad distribution indicates good or bad precision respectively.

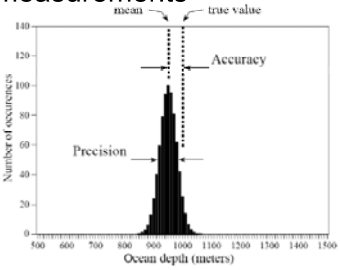
Precision and accuracy – cont.



From: S. W. Smith, "The Scientist and Engineer's Guide to Digital Signal Processing", California Technical Publishing, 1997

Precision and accuracy – cont.

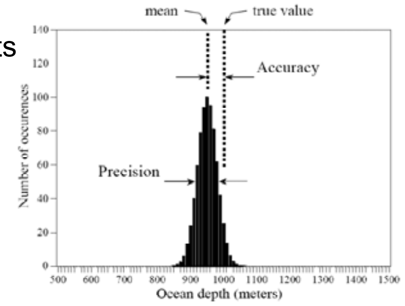
- Situation: Good accuracy & poor precision
  - Shift in mean from true value is small
  - The histogram of measurements is broad
  - Measurements as group: Good
  - Measurements as individuals: Poor
  - Poor repeatability: Measurements do not agree well
  - Improve precision: Average several measurements
- Cause of poor precision
  - Random errors
- Poor precision indicates
  - Presence of random noise



From: S. W. Smith, "The Scientist and Engineer's Guide to Digital Signal Processing", California Technical Publishing, 1997

Precision and accuracy – cont.

- Situation: Poor accuracy & good precision
  - Shift in mean from true value is large
  - The histogram of measurements is slender
  - Successive readings
    - Close in value
    - All have a large error
  - Average several measurements
    - Does not improve accuracy
- Cause of poor accuracy
  - Systematic errors
- Poor accuracy indicates
  - Poor calibration



Precision and accuracy – cont.

- Identification of problem
  - 1) Avg. successive readings yield better measurement?
    - Yes: The error is due to precision
    - No: The error is due to accuracy
  - 2) Calibration corrects the error?
    - Yes: The error is due to accuracy
    - No: The error is due to precision

Part IV: Data acquisition system

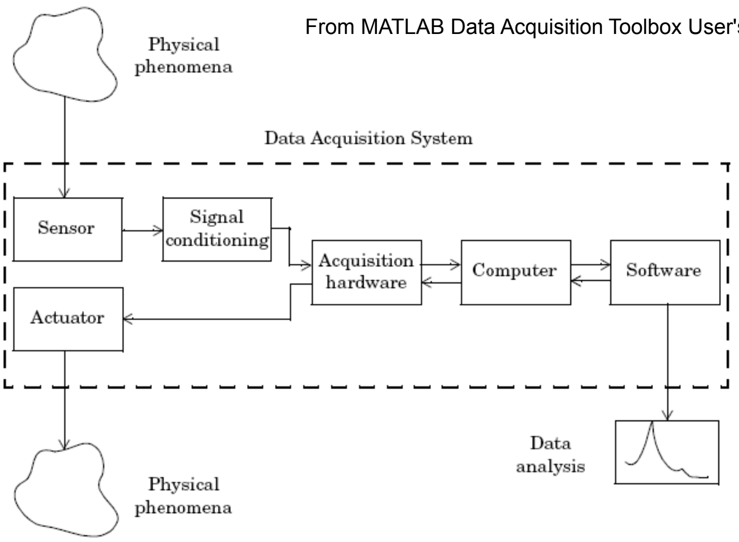
- The normal distribution
- Digital noise generation
- Precision and accuracy of data acquisition
- Data acquisition system

Anatomy of a data acquisition experiment

- System setup
  - Install the hardware and software
  - Attach sensors
- Calibration of hardware
  - Provide a known input to the system and record the output.
  - The rest is to be done by software
- Trials
  - Acquire data
  - To achieve a precise, accurate measurement, you need to perform several data acquisition trials using different hardware and software configurations.

The data acquisition system

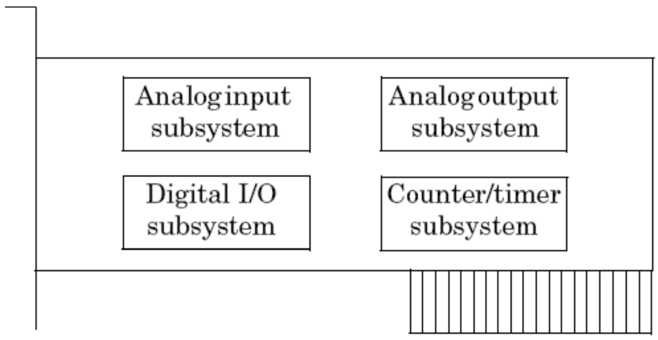
From MATLAB Data Acquisition Toolbox User's Guide



Data acquisition hardware

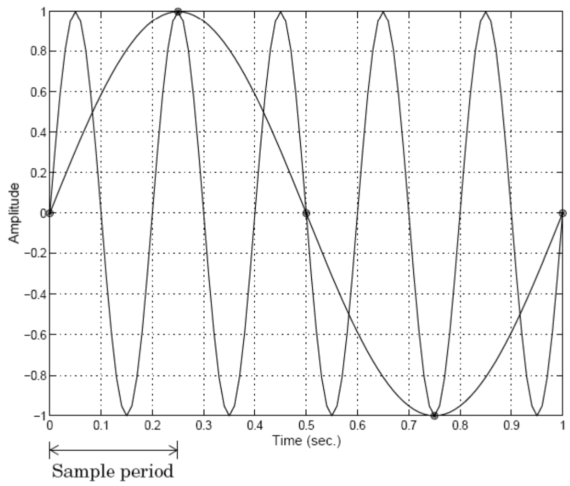
- Multifunction boards

From MATLAB Data Acquisition Toolbox User's Guide



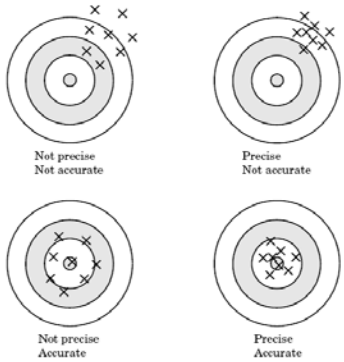
Making quality measurements

- Matching the sensor range and A/D converter range
- Sampling rate



Making quality measurements

- Matching the sensor range and A/D converter range
- Sampling rate
- Accuracy and precision
  - Noise removal



Summary

- The normal distribution
- Digital noise generation
- Precision and accuracy of data acquisition
- Data acquisition system