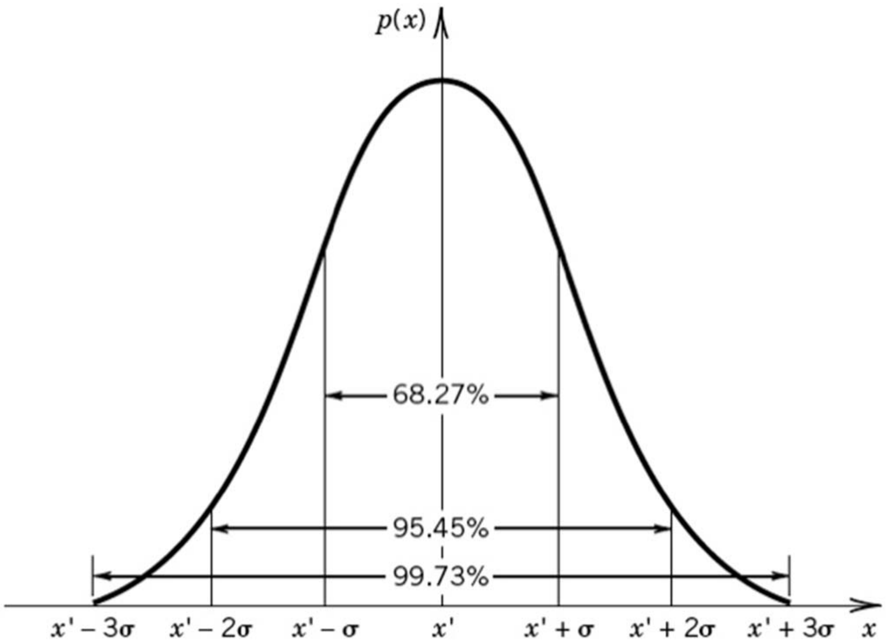
Instrumentation & Measurements MECH 430

Chapter 4 Probability and Statistics



4.1 Introduction

- For a given set of measurement, we want to be able to quantify:
  – A single representative value that best characterize the average of the data set
  – A representative value that provides a measure of the variation in the measured data set
  – How well the single average value represents the true average value of the variable

4.1 Introduction

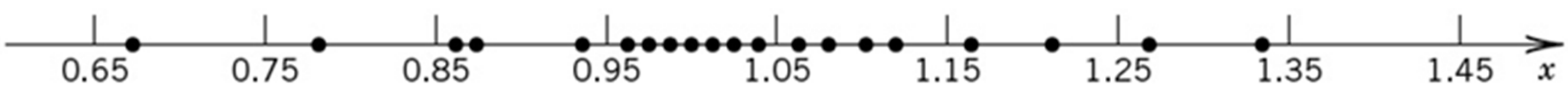
- In this chapter we will see:
  - Statistical characteristics of a data set.
  - Histogram of measured data
  - Probability density functions to describe the behavior of a variable.
  - Confidence interval about the measured mean
  - Perform regression analysis to quantify the confidence interval

4.2 Statistical Measurement

- A sample of data refers to a set of data obtained during repeated measurements of a variable under fixed operating conditions. Systematic errors are considered zero.
- The measurement problem will try to estimate the true value of the variable (x'), base on repeated measurement of (x).
- The true value is the mean of all possible values of x.

Probability Density Functions

- When obtaining a set of data from independent measurements (even under identical conditions), random scatter in the data values will occur.
- As such, the measured variable behaves as a *random variable*.
- When represented by discrete values, the set of data points is called *discrete random variable*.
- The variable will tend to assume one preferred value know as *central tendency*.



Probability Density Functions

- The frequency with which the measured variable assumes a particular value or interval of values is described by its probability density.
- The x axis is divided between the maximum and the minimum measured values of x into K intervals.
- Let n_j be the number of samples that falls in one interval, and $2\delta x$ the width of the interval.
- For small number of samples, K should be chosen such that $n_j > 5$ for at least one interval.

Probability Density Functions

- To estimate the number of intervals k :

$$K = 1.87(N-1)^{0.4} + 1$$

- The resulting plot is called histogram of the variable.
- The histogram is a way of viewing both the tendency and the probability density of a variable.
- The frequency distribution is $f_j = n_j / N$

Example 4.1

- Construct the histogram for the data in table 4.1

Table 4.1 Sample of Random Variable x

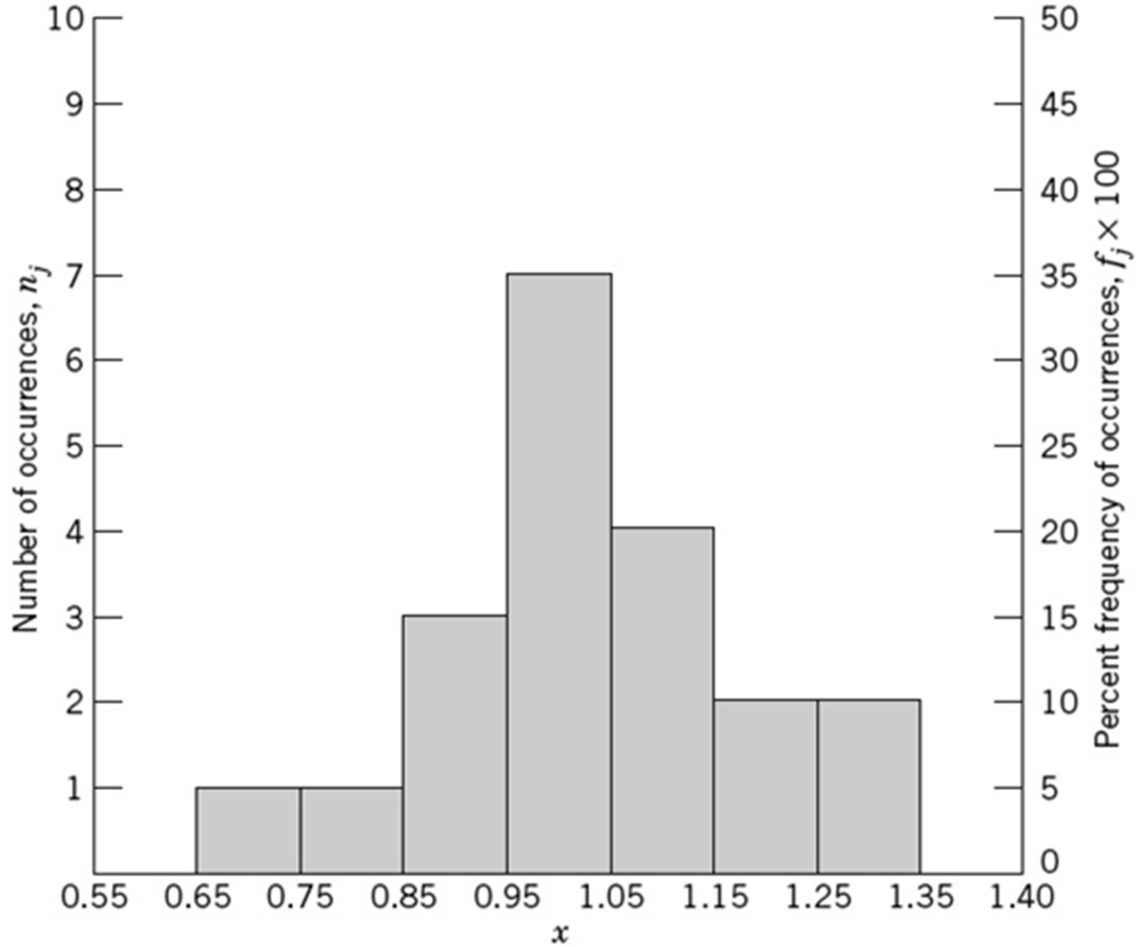
i	x_i	i	x_i
1	0.98	11	1.02
2	1.07	12	1.26
3	0.86	13	1.08
4	1.16	14	1.02
5	0.96	15	0.94
6	0.68	16	1.11
7	1.34	17	0.99
8	1.04	18	0.78
9	1.21	19	1.06
10	0.86	20	0.96

j	Interval	n_j	$f_j = n_j/\text{N}$
1	$0.65 \le x_i < 0.75$	1	0.05
2	$0.75 \le x_i < 0.85$	1	0.05
3	$0.85 \le x_i < 0.95$	3	0.15
4	$0.95 \le x_i < 1.05$	7	0.35
5	$1.05 \le x_i < 1.15$	4	0.20
6	$1.15 \le x_i < 1.25$	2	0.10
7	$1.25 \le x_i < 1.35$	2	0.10

$$N = \sum_{j=1}^{K} n_j$$

$$100 * \sum_{j=1}^{K} f_j = 100\%$$

Example 4.1



Number of occurrences, n_j
Percent frequency of occurrences, $f_j \times 100$
10
9
8
7
6
5
4
3
2
1
50
45
40
35
30
25
20
15
10
5
0
0.55
0.65
0.75
0.85
0.95
1.05
1.15
1.25
1.35
1.40
x

Probability Density Functions

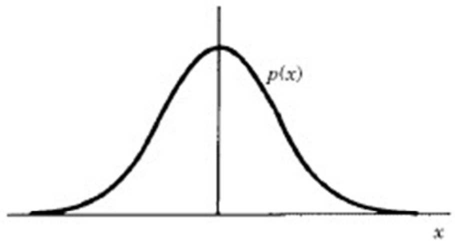
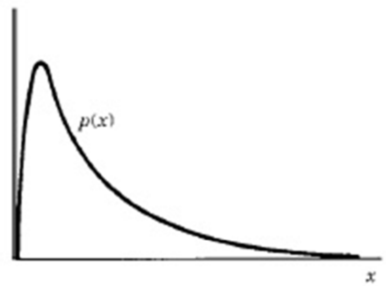
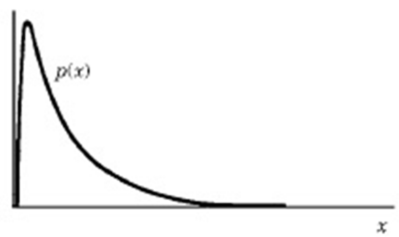
- The probability density function p(x) is:

$$p(x) = \lim_{N \to \infty, \delta x \to 0} \frac{n_j}{N(2\delta x)}$$

- The probability density function defines the probability that a measured variable might assume a particular value upon any individual measurement.
- It provides the central tendency of the variable.

Probability Density Functions

Table 4.2 Standard Statistical Distributions and Relations to Measurements

Distribution	Applications	Mathematical Representation	Shape
Normal	Most physical properties that are continuous or regular in time or space. Variations due to random error.	$p(x) = \frac{1}{\sigma(2\pi)^{1/2}} \exp\left[-\frac{1}{2}\frac{(x-x')^2}{\sigma^2}\right]$	
Log normal	Failure or durability projections; events whose outcomes tend to be skewed toward the extremity of the distribution.	$p(x) = \frac{1}{\pi\sigma(2\pi)^{1/2}} \exp\left[-\frac{1}{2}\ln\frac{(x-x')^2}{\sigma^2}\right]$	
Poisson	Events randomly occurring in time; $p(x)$ refers to probability of observing x events in time t. Here λ refers to x'.	$p(x) = \frac{e^{-\lambda}\lambda^x}{x!}$	

Probability Density Functions

- Regardless of the type of distribution assumed by a variable, if the variable shows a central tendency, it can be described and quantified through its mean and variance.

$$x' = \lim_{T\to\infty} \frac{1}{T}\int_0^T x(t)dt \qquad \sigma^2 = \lim_{T\to\infty} \frac{1}{T}\int_0^T [x(t) - x']^2 dt$$

$$x' = \lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} x_i \qquad \sigma^2 = \lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} (x_i - x')^2$$

- The standard deviation is defined as the square root of the variance.

4.3 Infinite Statistics

- Many measurands common to engineering are described by a *normal* or *Gaussian* distribution where data will scatter symmetrically about some central tendency.
- The PDF for a random variable x having a normal distribution is:

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} exp\left[-\frac{1}{2}\frac{(x-x')^2}{\sigma^2}\right]$$

where x' is the true mean value and σ^2 is the true variance of x

4.3 Infinite Statistics

- The probability *P(x)* that the random variable x will assume a value within the interval *x'±δx* is given by the area under *p(x)*.

$$P(x' - \delta x \le x \le x' + \delta x) = \int_{x'-\delta x}^{x'+\delta x} p(x)\,dx$$

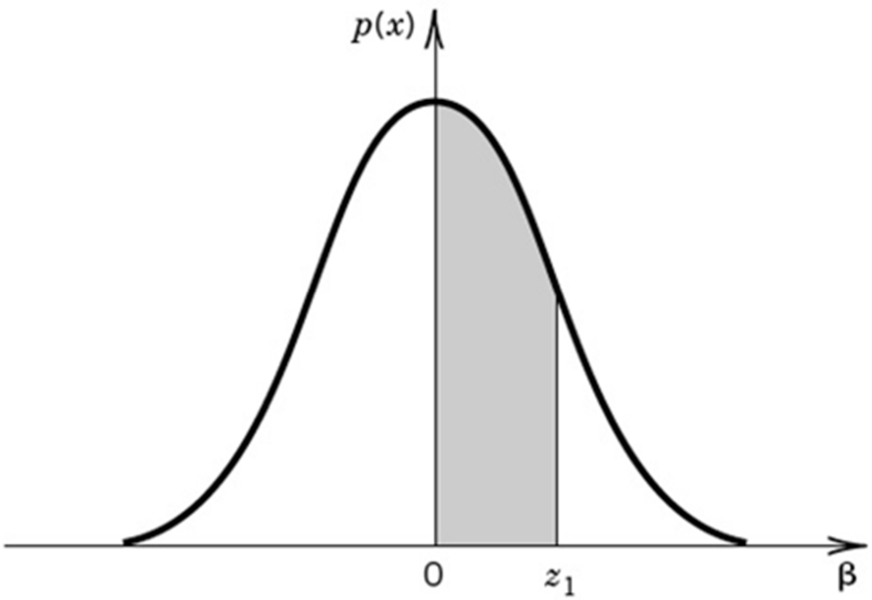
- *Using the following transformation:*

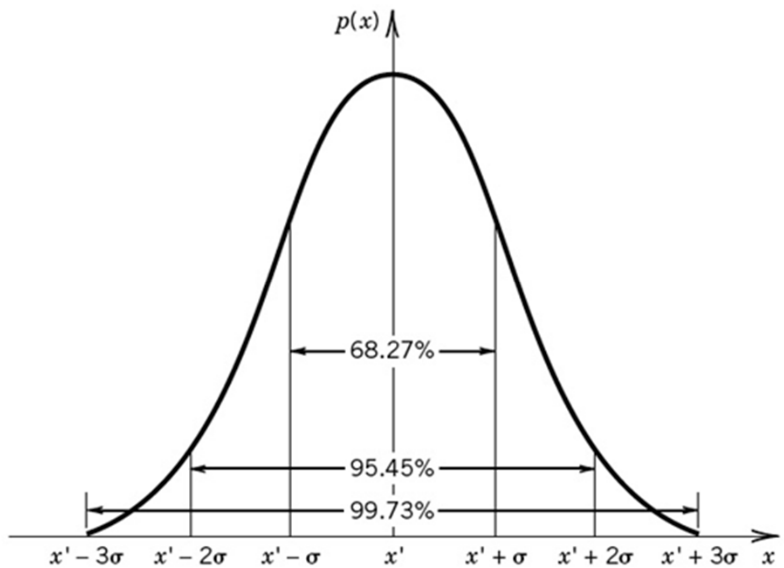
$$\beta = \frac{(x - x')}{\sigma} \qquad dx = \sigma d\beta \qquad \boldsymbol{z_1 = \frac{(x_1 - x')}{\sigma}}$$

$$P(-z_1 \le \beta \le +z_1) = \frac{1}{\sqrt{2\pi}} \int_{-z_1}^{z_1} e^{-\beta^2/2} d\beta = \frac{2}{\sqrt{2\pi}} \int_{0}^{z_1} e^{-\beta^2/2} d\beta$$

4.3 Infinite Statistics

- The half value of the probability is tabulated for the interval defined by z_1.





- The probability that x have a value between $x' \pm z_1\sigma$ is $2P(z_1)x100$.

Table 4.3 Probability Values for Normal Error Function

One-Sided Integral Solutions for $p(z_1) = \frac{1}{(2\pi)^{1/2}} \int_0^{z_1} e^{-\beta^2/2} d\beta$

$z_1 = \frac{x_1 - x'}{\sigma}$	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1809	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4758	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4799	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.49865	0.4987	0.4987	0.4988	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990

Example 4.3

- It is known that the statistics of a voltage signal are given by x'=8.5v and σ^2 =2.25V^2. If a single measurement of the voltage is made, determine the probability that the measured value is between 10.0 and 11.5V.

4.3 Infinite Statistics

- *Normal* or *Gaussian*.
- The PDF for a random variable x having a normal distribution is:

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} exp\left[-\frac{1}{2}\frac{(x-x')^2}{\sigma^2}\right]$$

$$P(x'-\delta x \le x \le x'+\delta x) = \int_{x'-\delta x}^{x'+\delta x} p(x)dx$$

$$\beta = \frac{(x-x')}{\sigma} \qquad dx = \sigma d\beta \qquad \boldsymbol{z_1} = \frac{(\boldsymbol{x_1}-\boldsymbol{x'})}{\boldsymbol{\sigma}}$$

$$P(-z_1 \le \beta \le +z_1) = \frac{1}{\sqrt{2\pi}}\int_{-z_1}^{z_1} e^{-\beta^2/2} d\beta = \frac{2}{\sqrt{2\pi}}\int_{0}^{z_1} e^{-\beta^2/2} d\beta$$

4.3 Infinite Statistics

- The probability that x is between $x' \pm z_1\sigma$ is $2P(z_1)x100$.

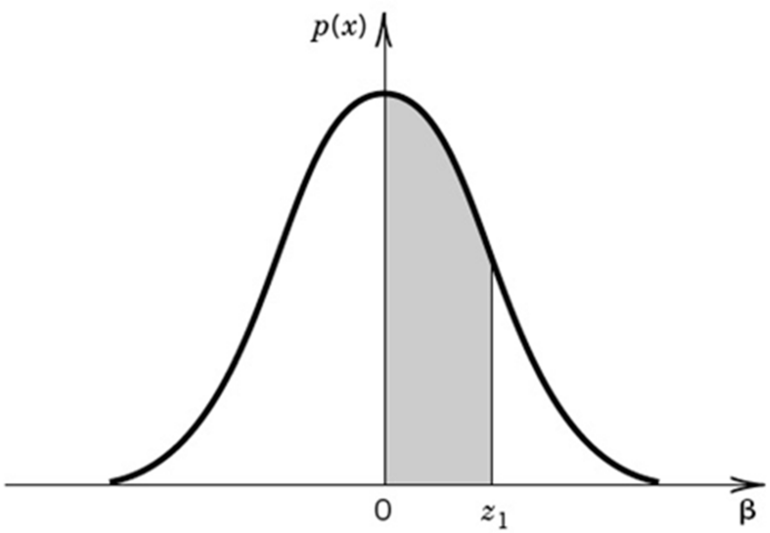


Table 4.3 Probability Values for Normal Error Function

One-Sided Integral Solutions for $p(z_1) = \frac{1}{(2\pi)^{1/2}} \int_0^{z_1} e^{-\beta^2/2} d\beta$

$z_1 = \frac{x_1 - x'}{\sigma}$	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1809	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4758	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4799	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.49865	0.4987	0.4987	0.4988	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990

Example

- A professor's exam scores are approximately distributed normally with mean 80 and standard deviation 5.
- What is the probability that a student scores an 82 or less?

$$\text{Prob}(X \le 82) \quad = \text{Prob}(Z \le (82-80)/5)$$
$$= \text{Prob}(Z \le .40) = 0.5 + 0.1554 = .6554$$

- What is the probability that a student scores a 90 or more?

$$\text{Prob}(X \ge 90) \quad = \text{Prob}(Z \ge (90-80)/5)$$
$$= \text{Prob}(Z \ge 2.00)$$
$$= 0.5 - 0.4772 = .0228$$

Example

- What is the probability that a student scores a 74 or less?

  $\text{Prob}(X \leq 74) \quad = \text{Prob}(Z \leq (74\text{-}80)/5)$

  $\quad = \text{Prob}(Z \leq -1.20)$

- If your table does not have negatives, use:

  $\text{Prob}(Z \leq -1.20) = \text{Prob}(Z \geq 1.20) = 0.5 - .3849 = .1151$

Example

- What is the probability that a student scores between 78 and 88?

  Prob(78 ≤ X ≤ 88) = Prob((78-80)/5 ≤ Z ≤ (88-80)/5)

  = Prob(-0.40 ≤ Z ≤ 1.60)

  = Prob(Z ≤ 1.60) -Prob(Z ≤ -0.40)

  = 0.4452 +0.1554= .6006

4.4 Finite Statistics

- If N measurements of x have been made and N is a finite value, the statistical value obtained from such data set will be regarded as estimates of the true values.

- Finite sized data sets can provide the statistical estimates know as the sample mean value, the sample variance and the sample standard deviation.

4.4 Finite Statistics

- While infinite statistics (what we have seen so far) describes the true behavior of a variable, *finite statistics* describes the behavior of the finite data set.
- Here we talk about:
  - The sample mean, $$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i$$
  - The sample variance, $$S_x^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(x_i - \bar{x}\right)^2$$
  - The sample standard deviation $$S_x = \left(\frac{1}{N-1}\sum_{i=1}^{N}\left(x_i - \bar{x}\right)^2\right)^{1/2}$$

4.4 Finite Statistics

- The degrees of freedom in the data set "*v*" is given as N-1.
- The predictive utility of infinite statistics can be extended to data sets for finite sample size with some modification.
- For finite sets, the z variable does not provide a reliable weight estimate.
- To compensate for the difference, the sample variance is weighted as follows:

$$x_i = \bar{x} \pm t_{v,P} S_x$$

4.4 Finite Statistics

- When sample sizes are small ($N < 31$) the z variable is not accurate. Instead we use the student t distribution.

t testimator

$$x_i = \bar{x} \pm t_{v,P} S_x \quad (P\%)$$

v is the number of degrees of freedom (N-1)
P is the desired probability

4.4 Finite Statistics

Table 4.4 Student-*t* Distribution

ν	t_{50}	t_{90}	t_{95}	t_{99}
1	1.000	6.314	12.706	63.657
2	0.816	2.920	4.303	9.925
3	0.765	2.353	3.182	5.841
4	0.741	2.132	2.770	4.604
5	0.727	2.015	2.571	4.032
6	0.718	1.943	2.447	3.707
7	0.711	1.895	2.365	3.499
8	0.706	1.860	2.306	3.355
9	0.703	1.833	2.262	3.250
10	0.700	1.812	2.228	3.169
11	0.697	1.796	2.201	3.106
12	0.695	1.782	2.179	3.055
13	0.694	1.771	2.160	3.012
14	0.692	1.761	2.145	2.977
15	0.691	1.753	2.131	2.947
16	0.690	1.746	2.120	2.921
17	0.689	1.740	2.110	2.898
18	0.688	1.734	2.101	2.878
19	0.688	1.729	2.093	2.861
20	0.687	1.725	2.086	2.845
21	0.686	1.721	2.080	2.831
30	0.683	1.697	2.042	2.750
40	0.681	1.684	2.021	2.704
50	0.680	1.679	2.010	2.679
60	0.679	1.671	2.000	2.660
∞	0.674	1.645	1.960	2.576

Example 4.4

- Consider the data in the table below. Compute the sample statistics for this set. Estimate the interval of values over which 95% of the measurements should be expected to lie.

Table 4.1 Sample of Random Variable x

i	x_i	i	x_i
1	0.98	11	1.02
2	1.07	12	1.26
3	0.86	13	1.08
4	1.16	14	1.02
5	0.96	15	0.94
6	0.68	16	1.11
7	1.34	17	0.99
8	1.04	18	0.78
9	1.21	19	1.06
10	0.86	20	0.96

Example 4.4 (cont.)

- Mean $\bar{x} = 1.02$
- Standard deviation $S_x = 0.16$
- Given the probability 95%, find the corresponding range of data, in other words determined x_i
- $t_i = (x_i - \bar{x})/S_x$ gives $x_i = S_x * t_i + \bar{x}$
- For this example,

  $x_{19,95} = S_x * t_{19,95} + \bar{x}$

  From the Student's t distribution table, $t_{19,95} = 2.093$

  thus $x_{19,95} = 0.16*2.093 + 1.02 = 0.33 + 1.02$

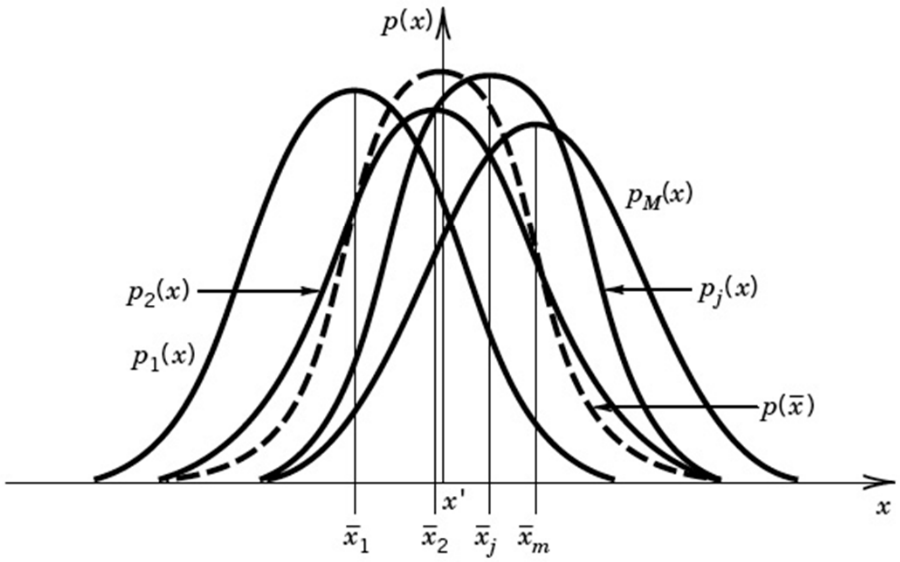
  Interval in which 95% of the data would lie is:1.02 +/- 0.33

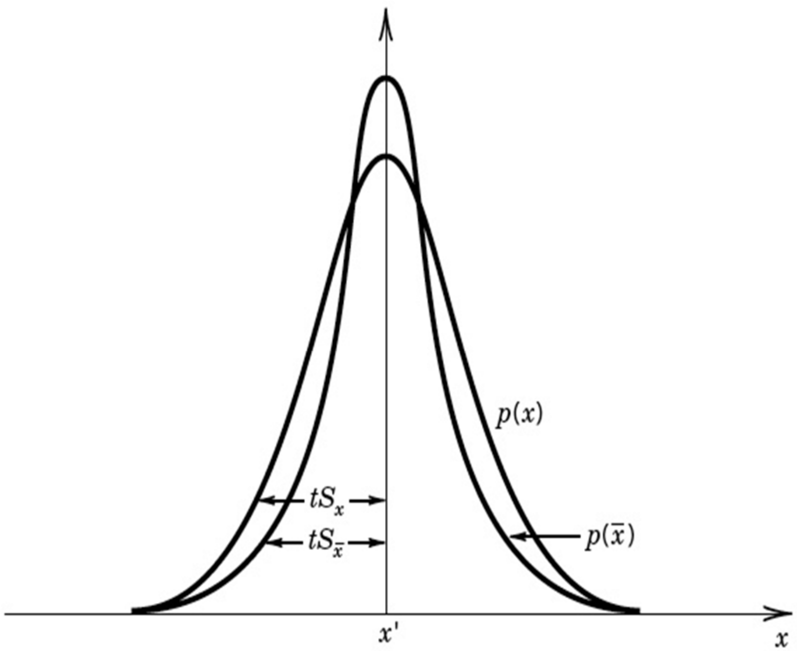
Standard Deviation of the Means

- Each time we take a new sample we expect a different mean and standard deviation.
- The set of mean values will be normally distributed about a central value, even if the PDF of the original function is not itself normal.
- The amount of variation in the sample means depends on the sample variance and the size N.

$$S_{\bar{x}} = \frac{S_x}{\sqrt{N}}$$

Standard Deviation of the Means





True Mean Value

- The standard deviation of the means represents a measure of how well the sample mean represents the true mean.
- In the absence of systematic errors in a measurement, the confidence interval states the true value within a likely range about the sample mean value.

$$x' = \bar{x} \pm t_{v,P} S_{\bar{x}}$$

Example 4.4

- Consider the data in the table below. The mean is 1.02 and the standard variation is 0.16. Estimate the true mean value of the measurand at 95% probability based on this finite data set.

Table 4.1 Sample of Random Variable x

i	x_i	i	x_i
1	0.98	11	1.02
2	1.07	12	1.26
3	0.86	13	1.08
4	1.16	14	1.02
5	0.96	15	0.94
6	0.68	16	1.11
7	1.34	17	0.99
8	1.04	18	0.78
9	1.21	19	1.06
10	0.86	20	0.96

Example 4.4

- The probability that the true Mean lies between 0.94 and 1.1 is 95%.
- Or there are a 95% probability that the true mean will lie between 0.94 and 1.1

(1.02±0.08=1.02±2.93x0.04=1.02±2.93x(0.16/Sqrt(20))

Example 4.4

- Mean $\bar{x} = 1.02$
- Standard of the means deviation $S_{\bar{x}} = 0.04$
- Given the probability 95%, find the corresponding range of mean values, in other words determined $\bar{x}_i$
- $t_i = (\bar{x}_i - \bar{x})/S_{\bar{x}}$ gives $\bar{x}_i = S_{\bar{x}} * t_i + \bar{x}$
- For this example,

$$\bar{x}_{19,95} = S_{\bar{x}} * t_{19,95} + \bar{x}$$

From the Student's t distribution table, $t_{19,95} = 2.093$

thus $\bar{x}_{19,95} = 0.04*2.093 + 1.02 = 0.08 + 1.02$

Interval in which 95% of the means would lie is:1.02 +/- 0.08

Example

- A rotary displacement sensor was tested 16 times and recorded in degrees as follows:

  0.11, 0.12, 0.09, 0.10, 0.10, 0.14, 0.08, 0.08,
  0.13, 0.10, 0.10, 0.12, 0.08, 0.09, 0.11, 0.15.

- If we know the standard deviation S_x to be 0.01, what are the odds that true value should fall within 5% of the sample mean? (Note: You can assume a normal distribution and use the appended chart).

Example

- To solve this problem we assume that readings are normally distributed. The sample mean is computed as:

$$\overline{X} = \frac{1}{16}\left(0.11 + 0.12 + \ldots + 0.11 + 0.15\right) = 0.10625$$

- Then we use the following to determine z_0 or $t_{\nu P}$:

$$P\left(\overline{X} - \frac{z_o S_x}{\sqrt{N}} \leq x' < \overline{X} + \frac{z_o S_x}{\sqrt{N}}\right) = p$$

- Hence

$$\frac{z_o(0.01)}{\sqrt{16}} = \frac{5}{100}(0.10625) \Rightarrow z_o = 2.125$$

Example

- If we use z_0:
  - interpolate in the z_0 table to find the area under the curve to be 0.4832,
  - the desired area is : 0.4832(2)= 0.9664,
  - therefore the odds that the mean resolution falls within 5% of the sample mean is 96.64%.
- If we use $t_{15,\,P} = 2.125$:
  - interpolate the $t_{\nu,\,P}$ Table to get $P \approx 95\%$

Table 4.3 Probability Values for Normal Error Function

One-Sided Integral Solutions for $p(z_1) = \frac{1}{(2\pi)^{1/2}} \int_0^{z_1} e^{-\beta^2/2} d\beta$

$z_1 = \frac{x_1 - x'}{\sigma}$	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1809	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4758	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4799	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.49865	0.4987	0.4987	0.4988	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990

Table 4.4 Student-*t* Distribution

ν	t_{50}	t_{90}	t_{95}	t_{99}
1	1.000	6.314	12.706	63.657
2	0.816	2.920	4.303	9.925
3	0.765	2.353	3.182	5.841
4	0.741	2.132	2.770	4.604
5	0.727	2.015	2.571	4.032
6	0.718	1.943	2.447	3.707
7	0.711	1.895	2.365	3.499
8	0.706	1.860	2.306	3.355
9	0.703	1.833	2.262	3.250
10	0.700	1.812	2.228	3.169
11	0.697	1.796	2.201	3.106
12	0.695	1.782	2.179	3.055
13	0.694	1.771	2.160	3.012
14	0.692	1.761	2.145	2.977
15	0.691	1.753	2.131	2.947
16	0.690	1.746	2.120	2.921
17	0.689	1.740	2.110	2.898
18	0.688	1.734	2.101	2.878
19	0.688	1.729	2.093	2.861
20	0.687	1.725	2.086	2.845
21	0.686	1.721	2.080	2.831
30	0.683	1.697	2.042	2.750
40	0.681	1.684	2.021	2.704
50	0.680	1.679	2.010	2.679
60	0.679	1.671	2.000	2.660
∞	0.674	1.645	1.960	2.576

4.6 Regression Analysis

- Regression signifies expressing a group of data points by a less perfect state.
- In linear regression, that less perfect state is a line.
- Linear regression can be done via an optimization technique known as *least-squares* method, which minimizes the sum of errors between the true *y* coordinate of a point and its estimated *y* coordinate via the line representing this point (*i.e.*, regression):

$$\Delta^2 = \sum_{i=1}^{n} \left(y_i - y \right)^2 = \sum_{i=1}^{n} \left(y_i - b - mx_i \right)^2$$

$$\begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \\ \vdots & \\ x_n & y_n \end{bmatrix}$$

4.6 Regression Analysis

$$\Delta^2 = \sum_{i=1}^{n} (y_i - y)^2 = \sum_{i=1}^{n} (y_i - b - mx_i)^2$$

- To minimize this quantity, take its derivative (with respect to *m* and *b*) and set it to zero:

$$\frac{\partial \Delta^2}{\partial m} = 0 \quad \text{and} \quad \frac{\partial \Delta^2}{\partial b} = 0$$

Which results:

$$m = \frac{n\sum_{i=1}^{n} x_i y_i - \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} \quad \text{and} \quad b = \frac{\left(\sum_{i=1}^{n} x_i^2\right)\left(\sum_{i=1}^{n} y_i\right) - \left(\sum_{i=1}^{n} x_i y_i\right)\left(\sum_{i=1}^{n} x_i\right)}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}$$

Example - Calibration

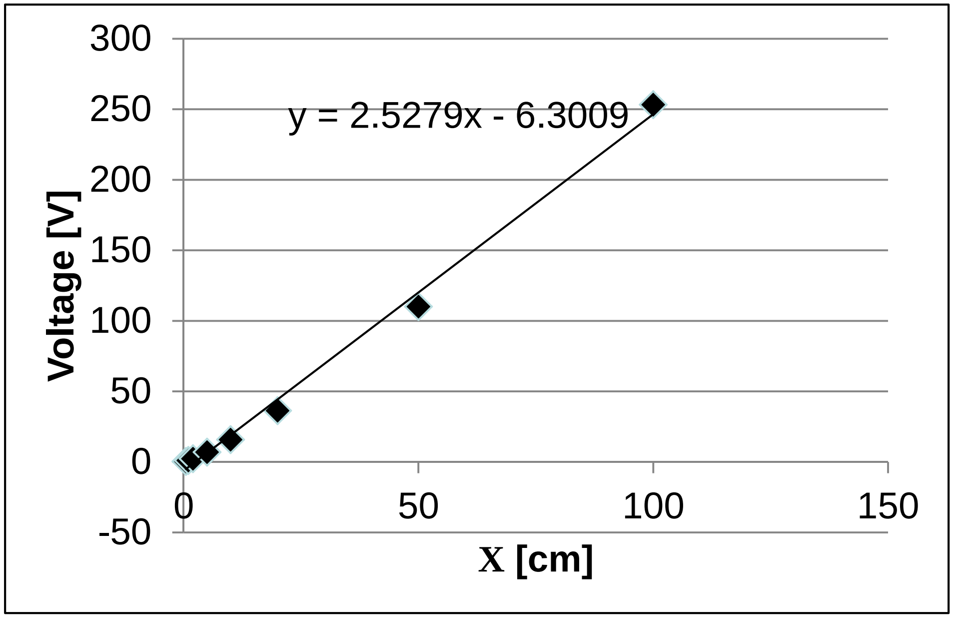
- Given:
  - The voltmeter calibration data:

X [cm]	*Y* [V]
0.5	0.4
1.0	1.0
2.0	2.3
5.0	6.9
10.0	15.8
20.0	36.4
50.0	110.1
100.0	253.2

- Plot this data and calibrate.
- Determine the static sensitivity at:
  - X = 5.0
  - X = 10.0
  - X = 20.0
- For which values is the system more sensitive? Analyze

Solution

- Using Excel:

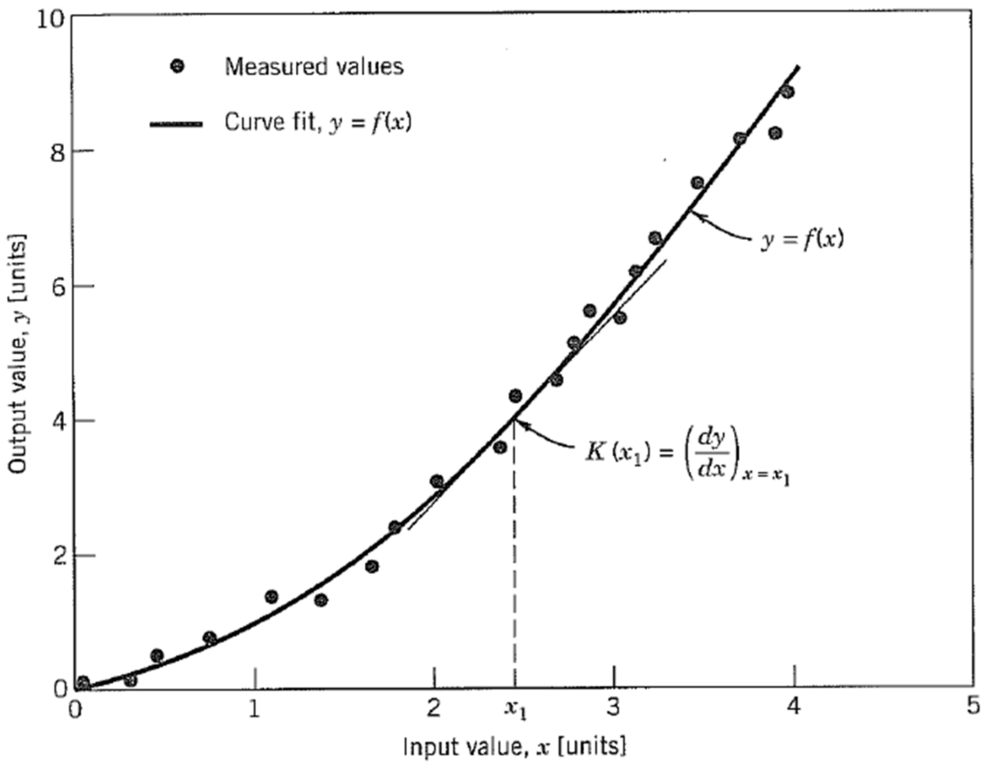


X [cm]	*Y* [V]
0.5	0.4
1.0	1.0
2.0	2.3
5.0	6.9
10.0	15.8
20.0	36.4
50.0	110.1
100.0	253.2

Solution

- Finding the sensitivity at different points:
  - $X = 5.0, \quad S = 1.8$
  - $X = 10.0, S = 2.1$
  - $X = 20.0, S = 2.5$
- As X increases the sensor becomes more sensitive.

Non-linear Calibration Curves



x	***y***

- If the relationship is nonlinear we can set up a ***look-up table*** from the calibration curve.