POLITECNICO DI MILANO

STEN Doctoral Courses

Advanced Measurement Techniques

Analysis of experimental data

A.Coghe

Department of Energy – Politecnico di Milano, Milano – ITALY

Analysis of experimental data

- **Introduction**
- **Uncertainty analysis**
- **Statistical analysis**
- **The Gaussian error distribution**
- **The Student's t-distribution**
- **Evaluation of uncertainties**
- **Error propagation**
- **Data acquisition**
- **Bibliography**

Analysis of experimental data

Physical processes are described by ***deterministic relationships*** (algebraic, differential or integral equations).

As a consequence, any measurement should provide a ***unique*** and ***repetable*** result.

– –

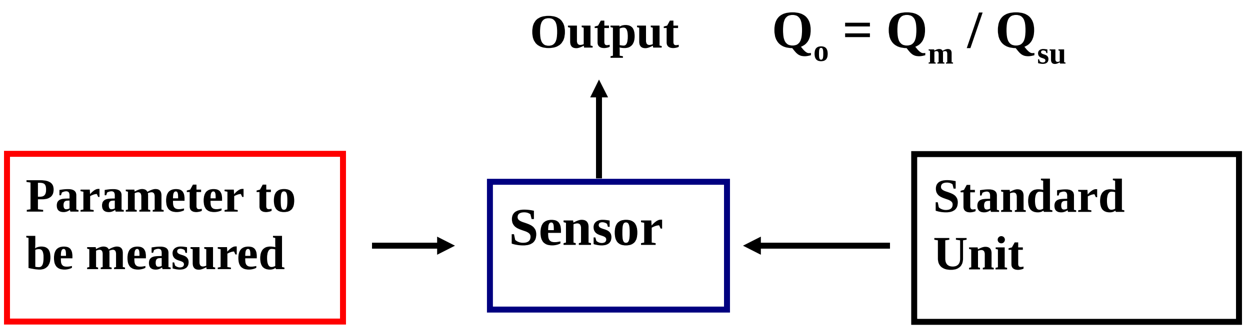
However, it is common experience to find a certain degree of ***uncertainty (random and/or systematic variations)*** in the results of all experiments, regardless of the care which is exerted and the quality of the instruments.

We can bypass these difficulties by introducing probabilistic-statistic concepts in the interpretation of the experimental data.

*To evaluate **static** or **dynamic** errors it is particularly useful to introduce the concepts of **calibration** and **transfer function***

Measurement process

Direct or indirect comparison of a specific property of a system with a reference standard unit



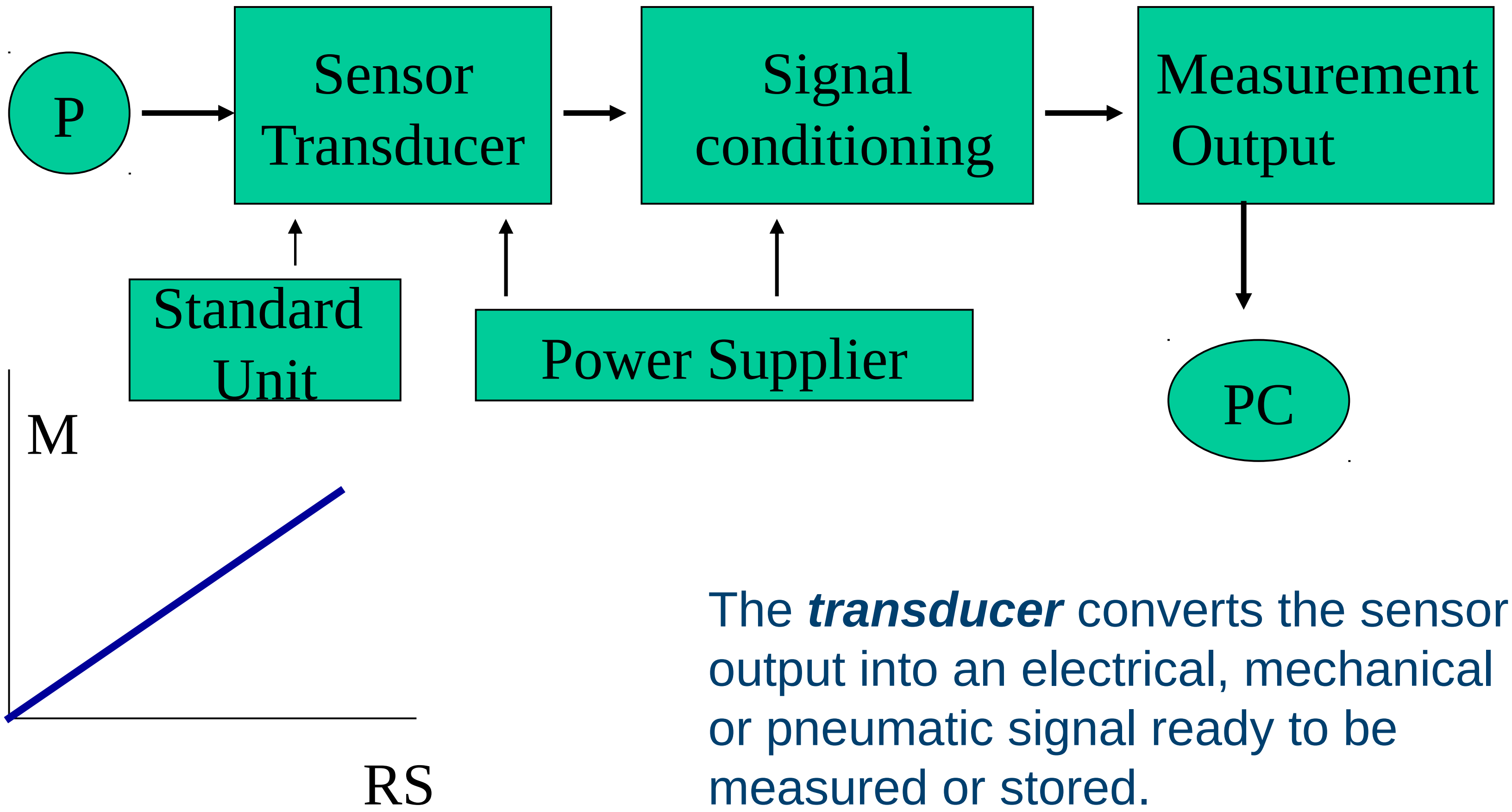
The standard unit should be internationally recognized and univocally defined

Every measurement system needs a ***sensor:*** the element sensitive to variations of the specific parameter to be measured.

The procedure used to evaluate the ratio between the measured quantity and the reference standard is called ***calibration***.

Measurement process

The ***calibration*** procedure must be performed including the entire experimental setup.

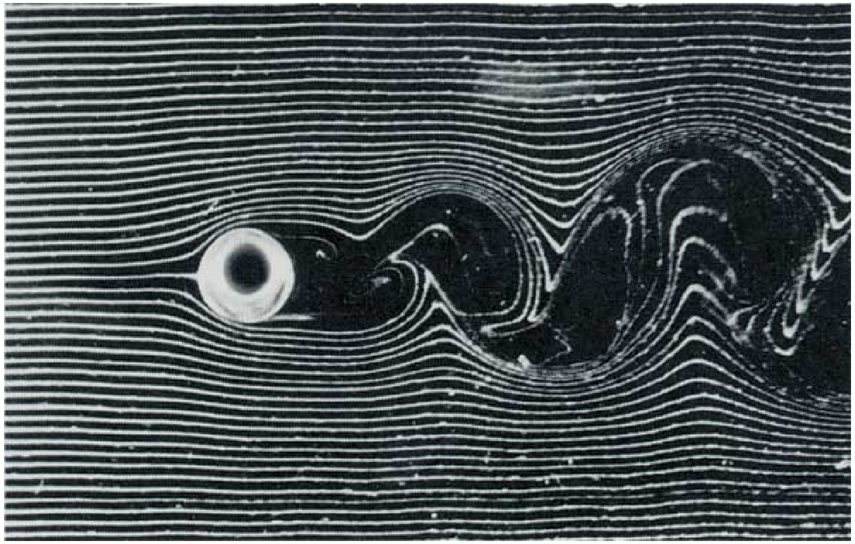


The ***transducer*** converts the sensor output into an electrical, mechanical or pneumatic signal ready to be measured or stored.

Types of Experimental Errors

The observation of any physical phenomenon and the measurement process require a ***close interaction*** of the measuring sensor and the process under examination.

This interaction can be assimilated to an ***energy exchange*** which inevitably perturbs the system under observation and should be minimized, although it is difficult to quantify the possible error.



An example is the introduction of a cylindrical probe in a flow field to measure pressure, velocity or temperature.

Alternatively, the perturbed flow field past any obstacle will produce random outputs (statistical errors)

Measurement uncertainty

The ***perfect (ideal) measurement*** is impossible!

There may be several types of errors that cause uncertainty in the experimental measurements:

1. ***Fixed errors***, called ***systematic*** or ***bias errors***: which will cause repeated readings to be in error by roughly the same amount, for some unknown reason. → Calibration is needed!
2. ***Random (statistical) errors***, which may be caused by parameter fluctuations, random electronic fluctuations in the instruments, various external influences (electronic noise, radiofrequencies, vibrations, temperature fluctuations,….).

Random errors usually follow a certain statistical distribution, *but not always.*

In many cases it is very difficult to distinguish between fixed and random errors.

Statistical Analysis

In presence of *random errors*, individual results of a measurement process will vary from each other. We define:

Frequency of occurrence $$f_k = \frac{n_k}{N}$$

Arithmetic mean $$\bar{x} = \frac{1}{N}\sum_{1}^{N} x_i$$

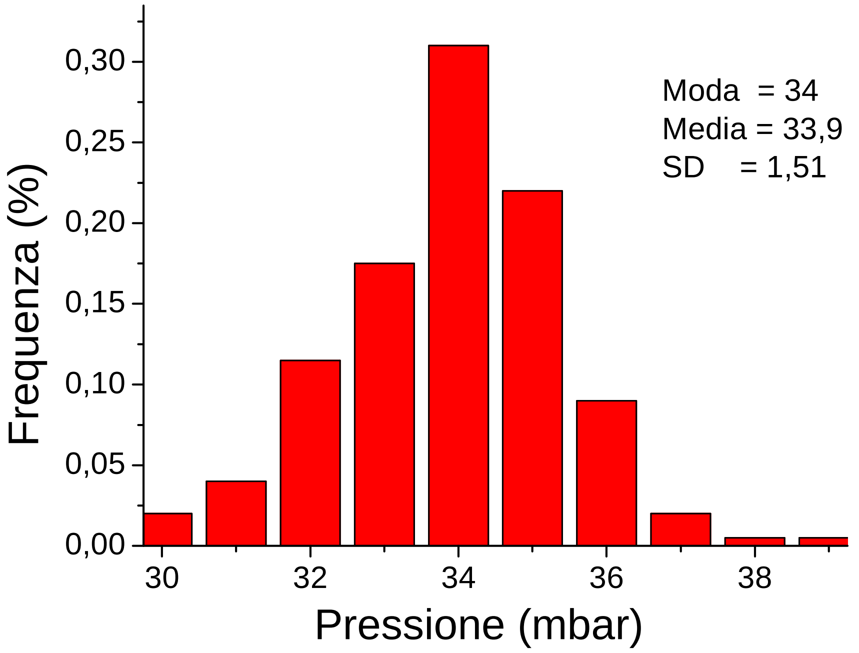
Median: the value that divides the data points in half

Moda: the value at maximum frequency

Deviation $$x_i - \bar{x}$$

Standard deviation (rms)

$$\sigma = \left[\frac{1}{N}\sum_{1}^{N}\left(x_i - \bar{x}\right)^2\right]^{1/2}$$



Variance σ^2

Probability concepts

For a large number of samples taken to describe the population, we can define

Probability function

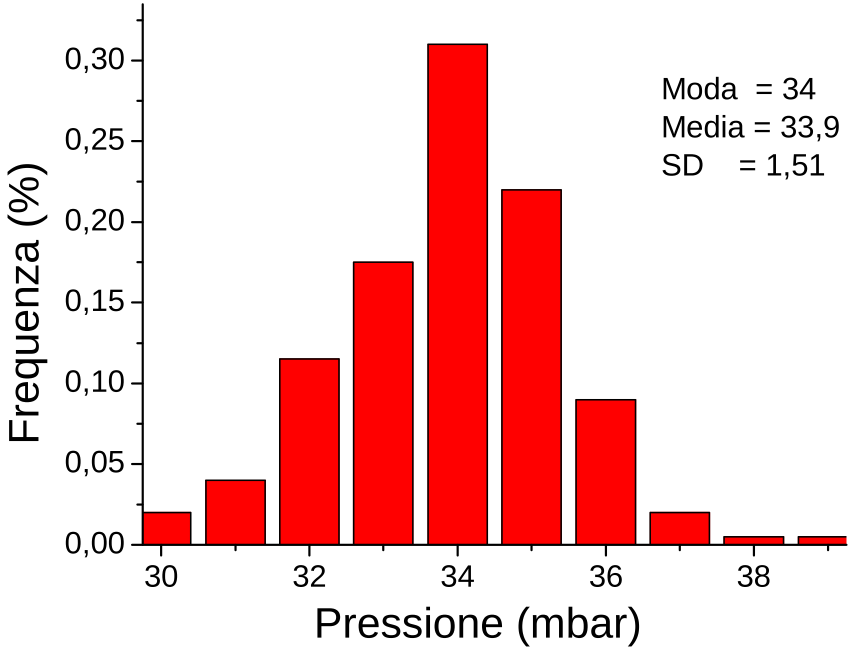
$$P(x) = \lim_{N\to\infty,\Delta x\to 0} \frac{n_k}{N\Delta x}$$

Mean

$$\bar{x} = \int_{-\infty}^{\infty} x\, P(x)\, dx$$

$$\bar{x} = \lim_{N\to\infty} \frac{1}{N} \sum_{1}^{N} x_i$$

$$\bar{x} = \lim_{T\to\infty} \frac{1}{T} \int_{0}^{T} x(t)dt$$



The last two definitions are equivalent only for a ***stationary*** process

Gaussian or Normal distribution

The ***Central Limit Theorem*** asserts that the contribution of a large number of independent random variables acting together will produce a ***Normal distribution***

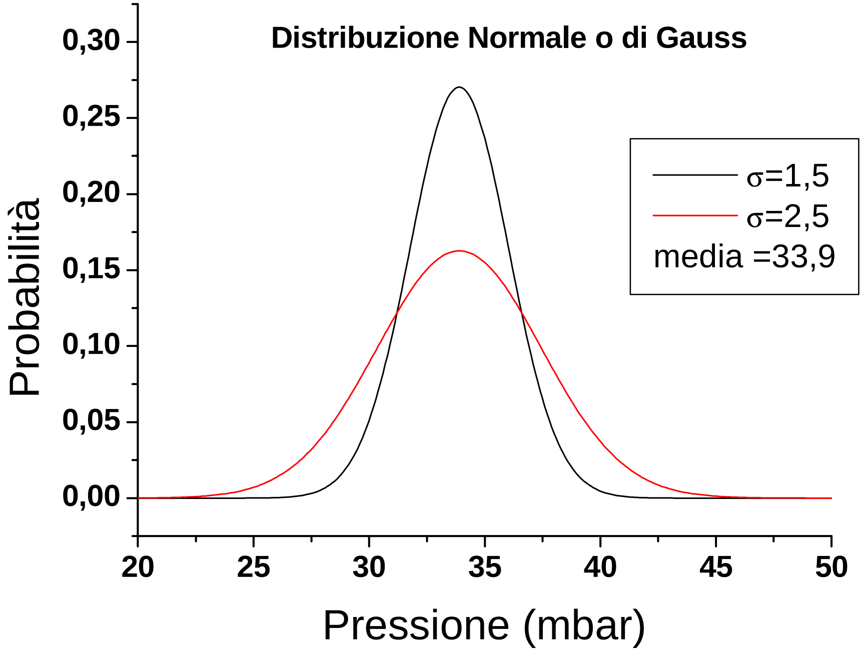
$$P(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\bar{x})^2}{2\sigma^2}\right]$$

In this expression $\bar{x}$ and σ are the only parameters needed to define a family of normal distributions.

$\bar{x}$ is the most probable result

$\sigma$is the standard deviation

For a Normal distribution, Mean, Median and Moda are coincident!



Gaussian or Normal error distribution

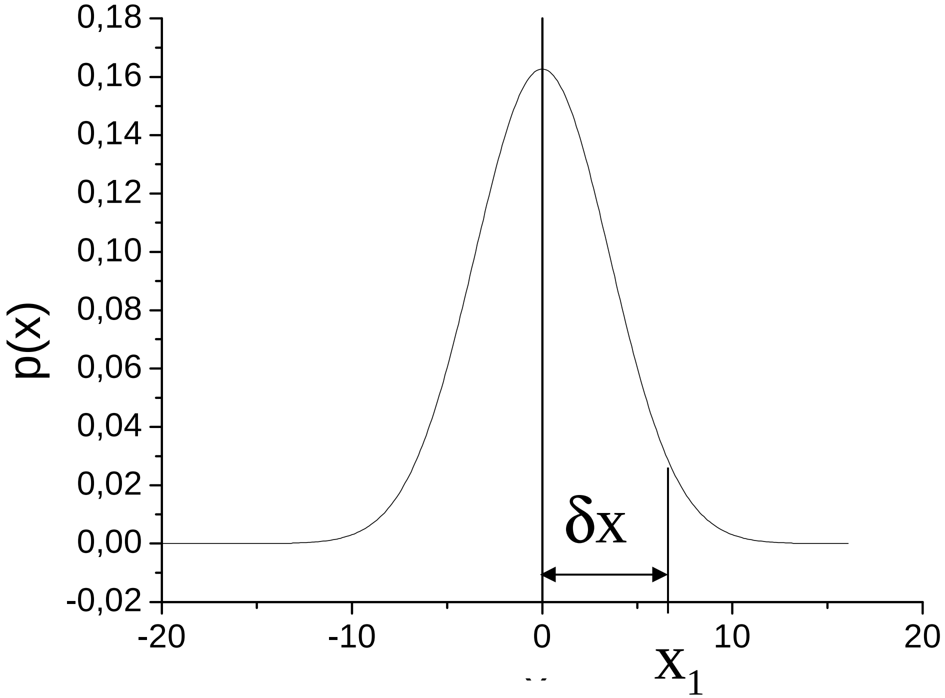
The probability density P(x) is normalized so that the total area under the curve is unity.

$$\int_{-\infty}^{\infty} P(x)\,dx = 1$$

The maximum probability occurs at $x = \bar{x}$

$$P(\bar{x}) = \frac{1}{\sigma\sqrt{2\pi}}$$

$P(\bar{x})$ is a measure of **precision** of the data since it has a larger value for smaller values of σ.



Gaussian or Normal error distribution

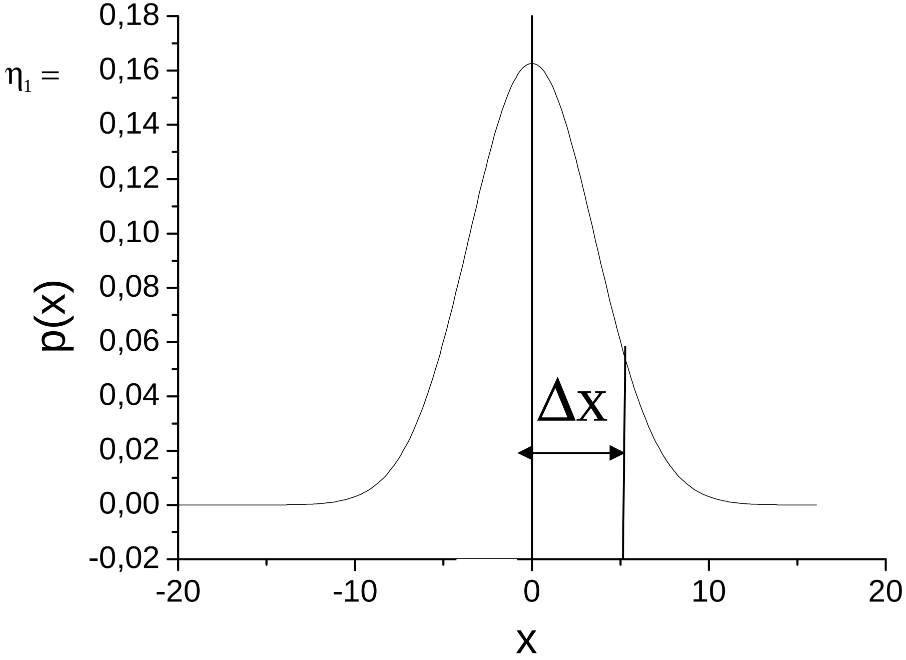
The probability that a measurement will fall within a certain range Δx of the mean is

$$P(-\Delta x \le \bar{x} \le \Delta x) = 2\int_0^{\Delta x} P(x)\,dx$$

Making the substitution

$$\eta = \frac{(x-\bar{x})}{\sigma} \;\Rightarrow\; dx = \sigma\, d\eta$$

$$P(\eta) = \frac{1}{\sqrt{2\pi}}\int_{-\eta_1}^{\eta_1} e^{-\frac{\eta^2}{2}}\, d\eta$$



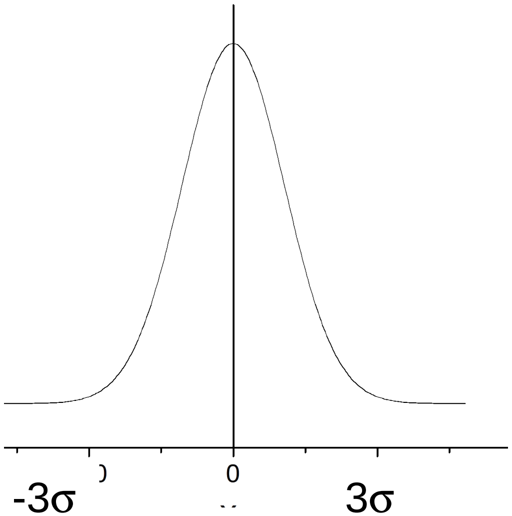
Error function

$$\frac{1}{\sqrt{2\pi}} e^{-\eta^2/2}$$

Gaussian or Normal error distribution

Values of the Gaussian normal error distribution

$$\frac{1}{\sqrt{2\pi}}e^{-\eta^2/2}$$



η	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	39894	39892	39886	39876	39862	39844	39822	39797	39767	39733
0.1	39695	39654	39608	39559	39505	39448	39387	39322	39253	39181
0.2	39104	39024	38940	38853	38762	38667	38568	38466	38361	38251
0.3	38139	38023	37903	37780	37654	37524	37391	37255	37115	36973
0.4	36827	36678	36526	36371	36213	36053	35889	35723	35553	35381
0.5	35207	35029	34849	34667	34482	34294	34105	33912	33718	33521
0.6	33322	33121	32918	32713	32506	32297	32086	31875	31659	31443
0.7	31225	31006	30785	30563	30339	30114	29887	29658	29430	29200
0.8	28969	28737	28504	28269	28034	27798	27562	27324	27086	26848
0.9	26609	26369	36129	25888	25647	25406	25164	24923	24681	24439
1.0	24197	23955	23713	23471	23230	22988	22747	22506	22265	22025
1.1	21785	21546	21307	21069	20831	20594	20357	20121	19886	19652
1.2	19419	19186	18954	18724	18494	18265	18037	17810	17585	17360
1.3	17137	16915	16694	16474	16256	16038	15822	15608	15395	15183
1.4	14973	14764	14556	14350	14146	13943	13742	13542	13344	13147
1.5	12952	12758	12566	12376	12188	12001	11816	11632	11450	11270
1.6	11092	10915	10741	10567	10396	10226	10059	09893	09728	09566
1.7	09405	09246	09089	08933	08780	08628	08478	08329	08183	08038
1.8	07895	07754	07614	07477	07341	07206	07074	06943	06814	06687
1.9	06562	06438	06316	06195	06077	05959	05844	05730	05618	05508
2.0	05399	05292	05186	05082	04980	04879	04780	04682	04586	04491
2.1	04398	04307	04217	04128	04041	03955	03871	03788	03706	03626
2.2	03547	03470	03394	03319	03246	03174	03103	03034	02965	02898
2.3	02833	02768	02705	02643	02582	02522	02463	02406	02349	02294
2.4	02239	02186	02134	02083	02033	01984	01936	01888	01842	01797
2.5	01753	01709	01667	01625	01585	01545	01506	01468	01431	01394
2.6	01358	01323	01289	01256	01223	01191	01160	01130	01100	01071
2.7	01042	01014	00987	00961	00935	00909	00885	00861	00837	00814
2.8	00792	00770	00748	00727	00707	00687	00668	00649	00631	00613
2.9	00595	00578	00562	00545	00530	00514	00499	00485	00470	00457
3.0	00443									
3.5	008727									
4.0	0001338									
4.5	0000160									
5.0	000001487									

Gaussian or Normal error distribution

Integrals of the Gaussian normal eror function

$$\frac{1}{\sqrt{2\pi}}\int_{-\eta_1}^{\eta_1} e^{-\frac{\eta^2}{2}}\,d\eta =$$

$$\frac{2}{\sqrt{2\pi}}\int_{0}^{\eta_1} e^{-\frac{\eta^2}{2}}\,d\eta$$

$$\mathrm{erf}\left(\eta_1\right) = \left(\frac{1}{\sqrt{\pi}}\right)\int_{-\eta_1}^{+\eta_1} e^{-\eta^2}\,d\eta$$

Tabulated values

$$\left(1/2\right)\mathrm{erf}\left(\eta_1/\sqrt{2}\right)$$

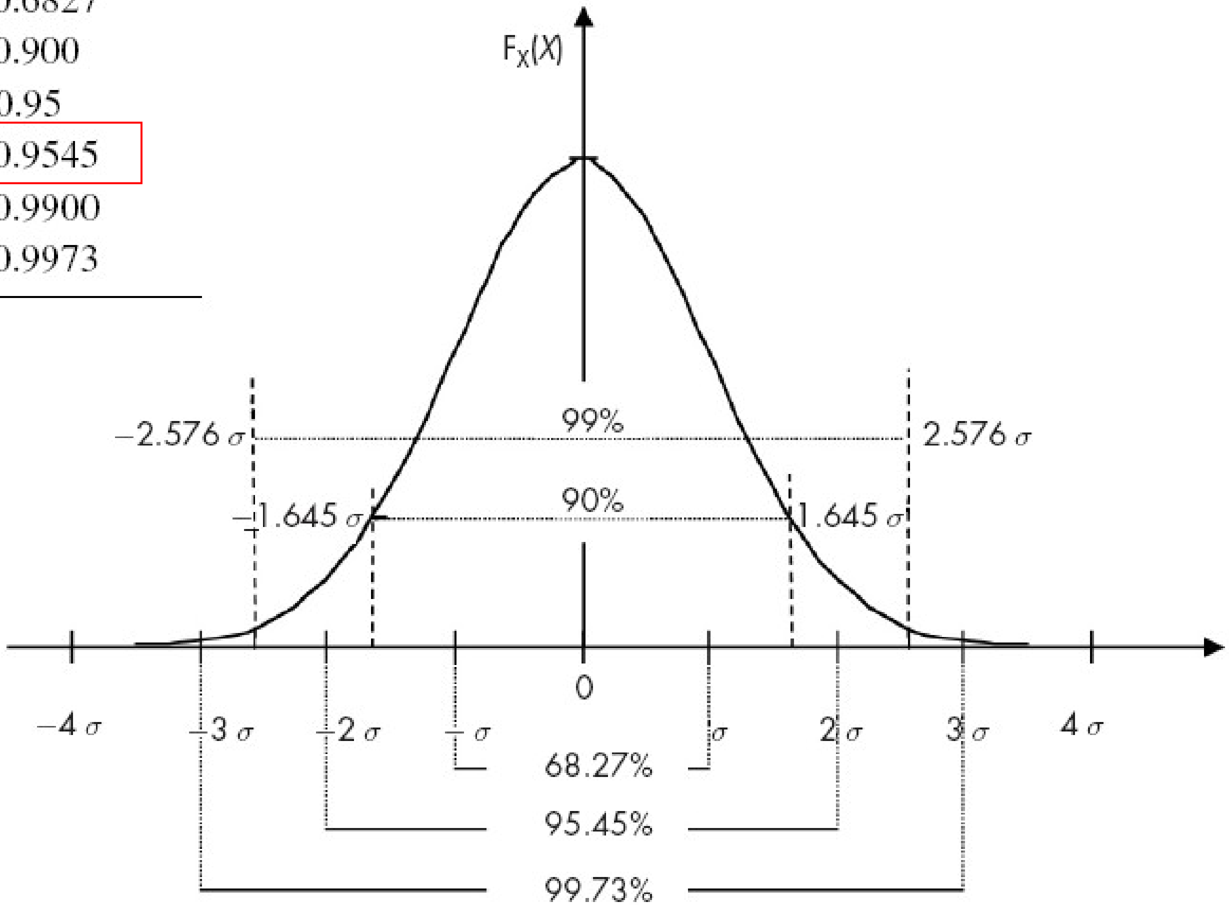
η_1	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0000	00399	00798	01197	01595	01994	02392	02790	03188	03586
0.1	03983	04380	04776	05172	05567	05962	06356	06749	07142	07355
0.2	07926	08317	08706	09095	09483	09871	10257	10642	11026	11409
0.3	11791	12172	12552	12930	13307	13683	14058	14431	14803	15173
0.4	15554	15910	16276	16640	17003	17364	17724	18082	18439	18793
0.5	19146	19497	19847	20194	20450	20884	21226	21566	21904	22240
0.6	22575	22907	23237	23565	23891	24215	24537	24857	25175	25490
0.7	25084	26115	26424	26730	27035	27337	27637	27935	28230	28524
0.8	28814	29103	29389	29673	29955	30234	30511	30785	31057	31327
0.9	31594	31859	32121	32381	32639	32894	33147	33398	33646	33891
1.0	34134	34375	34614	34850	35083	35313	35543	35769	35993	36214
1.1	36433	36650	36864	37076	37286	37493	37698	37900	38100	38298
1.2	38493	38686	38877	39065	39251	39435	39617	39796	39973	40147
1.3	40320	40490	40658	40824	40988	41198	41308	41466	41621	41774
1.4	41924	42073	42220	42364	42507	42647	42786	42922	43056	43189
1.5	43319	43448	43574	43699	43822	43943	44062	44179	44295	44408
1.6	44520	44630	44738	44845	44950	45053	45154	45254	45352	45449
1.7	45543	45637	45728	45818	45907	45994	46080	46164	46246	46327
1.8	46407	46485	46562	46638	46712	46784	46856	46926	46995	47062
1.9	47128	47193	47257	47320	47381	47441	47500	47558	47615	47670
2.0	47725	47778	47831	47882	47932	47962	48030	48077	48124	48169
2.1	48214	48257	48300	48341	48382	48422	48461	48500	48537	48574
2.2	48610	48645	48679	48713	48745	48778	48809	48840	48870	48899
2.3	48928	48956	48983	49010	49036	49061	49086	49111	49134	49158
2.4	49180	49202	49224	49245	49266	49286	49305	49324	49343	49361
2.5	49379	49296	49413	49430	49446	49461	49477	49492	49506	49520
2.6	49534	49547	49560	49573	49585	49598	49609	49621	49632	49643
2.7	49653	49664	49674	49683	49693	49702	49711	49720	49728	49736
2.8	49744	49752	49760	49767	49774	49781	49788	49795	49801	49807
2.9	49813	49819	49825	49831	49836	49841	49846	49851	49856	49861
3.0	49865									
3.5	4997674									
4.0	4999683									
4.5	4999966									
5.0	4999997133									

Gaussian or Normal distribution

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

Entro	Area
$\mu \pm \sigma$	0.6827
$\mu \pm 1.6450\sigma$	0.900
$\mu \pm 1.960\sigma$	0.95
$\mu \pm 2\sigma$	0.9545
$\mu \pm 2.576\sigma$	0.9900
$\mu \pm 3\sigma$	0.9973

Confidence interval



E.O. Doebelin – Strumenti e Metodi di Misura (2° ed.)

Finite Statistics

Previous analysis refers to an infinitely large number of samples **($N \to \infty$).**

In many circumstances it is not possible to collect a very large data sets.

For small sets of data, the estimated statistical parameters are biased and we have to use different definitions to obtain an unbiased value.

A data sets of N statistically independent samples is considered to have N degrees of freedom **(N finite).**

If the data are randomly distributed around the mean value, a relationship exists

$$\sum_{1}^{N}(x_i - \overline{x}) \equiv 0$$

which reduces the degrees of freedom to N-1. Therefore, the unbiased definition of the ***standard deviation*** becomes

$$\sigma = \sqrt{\frac{1}{(N-1)}\sum_{1}^{N}\left(x_i - \overline{x}\right)^2}$$

Student's t-Distribution

For small data sets, Student developed a distribution function

$$f(t) = \frac{K_0}{\left(1+\frac{t^2}{N-i}\right)^{N/2}}$$

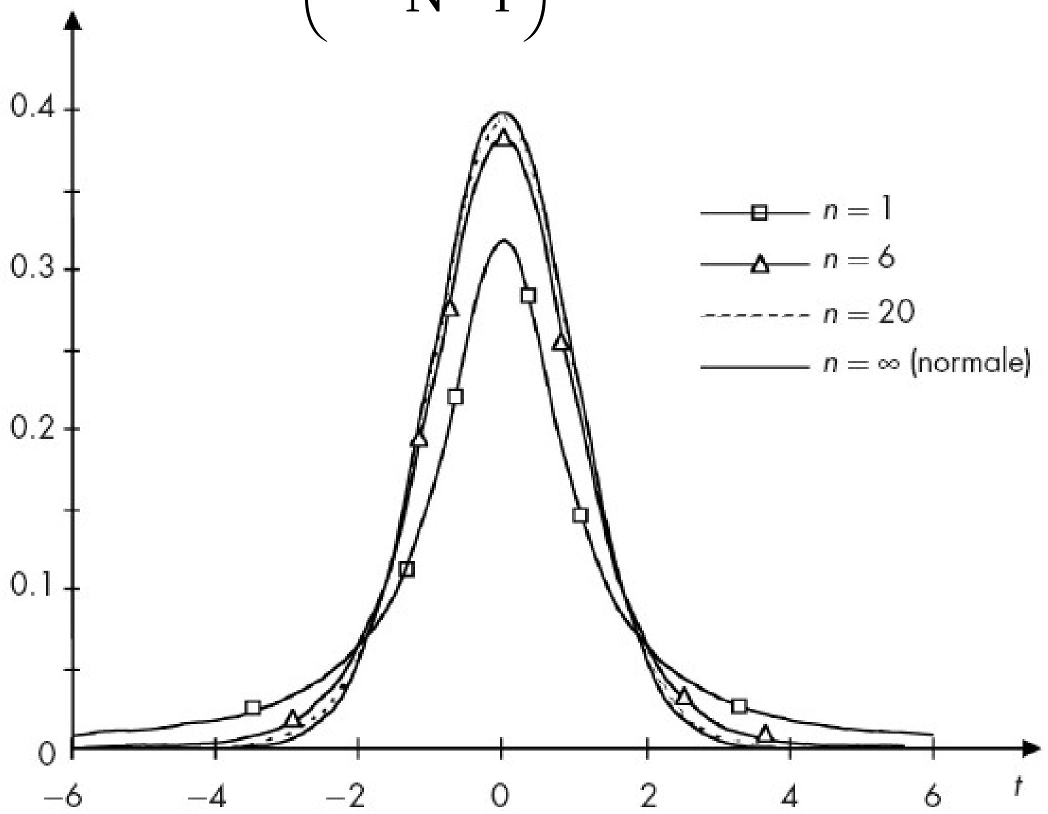
N = number of samples

$$t = \frac{\overline{x}-X}{\sigma}\sqrt{N}$$

$\overline{x}$ mean of N samples

X mean of normal population which the samples are taken from

K_0 constant which depends on N and the degrees of freedom.



Student's t-Distribution

Values of Student's t-distribution as function of the degrees of freedom and the probability

The Student's distribution is different from the normal one only for N < 60

At 95% probability t tends to the value z = 2 of the normal distribution.

Degrees of freedom ν	t_{50}	t_{80}	t_{90}	t_{95}	t_{98}	t_{99}	$t_{99.9}$
1	1.000	3.078	6.314	12.706	31.821	63.657	636.619
2	0.816	1.886	2.920	4.303	6.965	9.925	31.598
3	0.765	1.638	2.353	3.182	4.541	5.841	12.941
4	0.741	1.533	2.132	2.776	3.747	4.604	8.610
5	0.727	1.476	2.015	2.571	3.365	4.032	6.859
6	0.718	1.440	1.943	2.447	3.143	3.707	5.959
7	0.711	1.415	1.895	2.365	2.998	3.499	5.405
8	0.706	1.397	1.860	2.306	2.896	3.355	5.041
9	0.703	1.383	1.833	2.262	2.821	3.250	4.781
10	0.700	1.372	1.812	2.228	2.764	3.169	4.587
11	0.697	1.363	1.796	2.201	2.718	3.106	4.437
12	0.695	1.356	1.782	2.179	2.681	3.055	4.318
13	0.694	1.350	1.771	2.160	2.650	3.012	4.221
14	0.692	1.345	1.761	2.145	2.624	2.977	4.140
15	0.691	1.341	1.753	2.131	2.602	2.947	4.073
16	0.690	1.337	1.746	2.120	2.583	2.921	4.015
17	0.689	1.333	1.740	2.110	2.567	2.898	3.965
18	0.688	1.330	1.734	2.101	2.552	2.878	3.922
19	0.688	1.328	1.729	2.093	2.539	2.861	3.883
20	0.687	1.325	1.725	2.086	2.528	2.845	3.850
21	0.686	1.323	1.721	2.080	2.518	2.831	3.819
22	0.686	1.321	1.717	2.074	2.508	2.819	3.792
23	0.685	1.319	1.714	2.069	2.500	2.807	3.767
24	0.685	1.318	1.711	2.064	2.492	2.797	3.745
25	0.684	1.316	1.708	2.060	2.485	2.787	3.725
26	0.684	1.315	1.706	2.056	2.479	2.779	3.707
27	0.684	1.314	1.703	2.052	2.473	2.771	3.690
28	0.683	1.313	1.701	2.048	2.467	2.763	3.674
29	0.683	1.311	1.699	2.045	2.462	2.756	3.659
30	0.683	1.310	1.697	2.042	2.457	2.750	3.646
40	0.681	1.303	1.684	2.021	2.423	2.704	3.551
60	0.679	1.296	1.671	2.000	2.390	2.660	3.460
120	0.677	1.289	1.658	1.980	2.358	2.617	3.373
∞	0.674	1.282	1.645	1.960	2.326	2.576	3.291

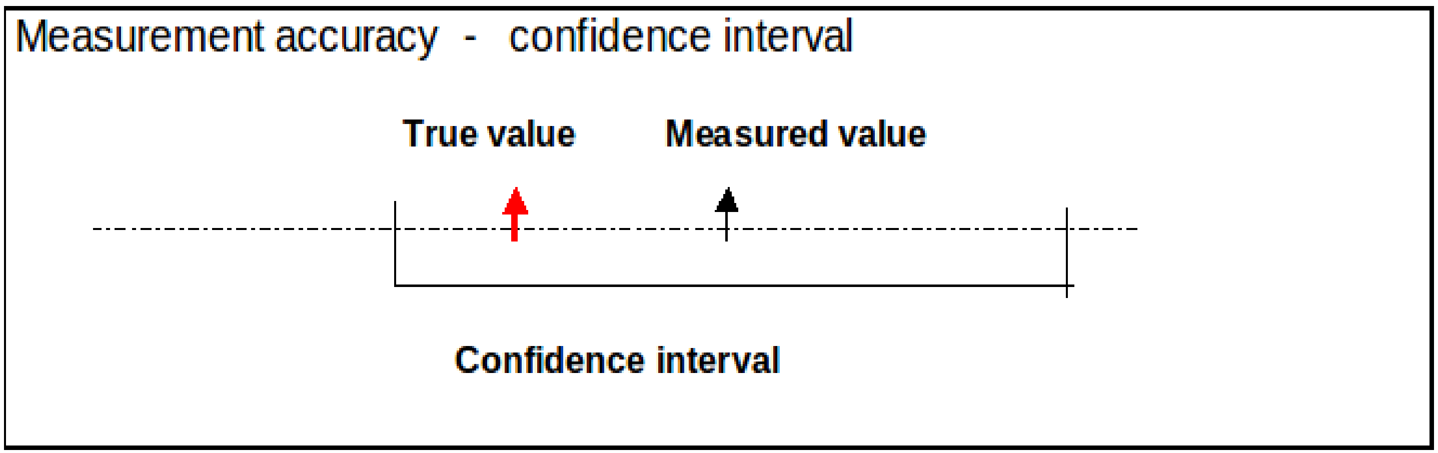
Confidence interval

The confidence interval represents the probability that the mean value will lie within a certain range of σ values and is given by

$$x_{\upsilon} = \bar{x} \pm t_{\nu,P}\sigma_{\bar{x}} \quad (P\%)$$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{N}}$$

Standard deviation of the mean



The % confidence level must be clearly defined.

It is common practice to assume 95% ($\rightarrow t = 2$).

Minimum number of samples

From

$$x_{\upsilon} - \overline{x} = \pm t_{\nu,P} \frac{\sigma}{\sqrt{N}} \quad (P = 95\%)$$

We can deduce

$$N \approx \left[\frac{t_{\nu,P}\sigma}{\frac{\left(x_{\upsilon} - \overline{x}\right)}{2}} \right]^2 \qquad (P = 95\%)$$

If N > 60, then $t \rightarrow 2$ and:

$$N \approx \left[\frac{4\sigma}{\left(x_{\upsilon} - \overline{x}\right)} \right]^2 \quad (P = 95\%)$$

Experiment planning

Particular questions which should be asked in the initial phases of experiment planning:

- ✓ What I am looking for?
- ✓ What primary variables shall be investigated?
- ✓ What ranges of the primary variables should be covered?
- ✓ How many samples should be taken to ensure good precision?
- ✓ What instrument accuracy is required?
- ✓ What frequency response must the instrument have?
- ✓ What safety precautions are necessary?
- ✓ Is the control of the surrounding temperature, humidity, vibration,… necessary?
- ✓ What about recording and computer processing of the data?

Evaluation of systematic uncertainty

Differential Pressure Transducer

Characteristic	Measure
Pressure ranges (psid)	1000, 2000, 3000, 5000, 7500, 10000
Accuracy	±0.25 % full scale
Linearity	±0.15 % full scale
Hysteresis	±0.10 % full scale
Non-repeatability	±0.05 % full scale
Output (standard)	4 mA to 20 mA
Max. line pressure	Full scale capacity +2000 psi
Resolution	Infinite

± 2.5 psid

± 1.5 psid

± 1.0 psid

± 0.5 psid

$$e_T = \pm\sqrt{e_1^2 + e_2^2 + \dots + e_n^2} = \sqrt{(2.5)^2 + (1.5)^2 + (1.0)^2 + (0.5)^2}$$

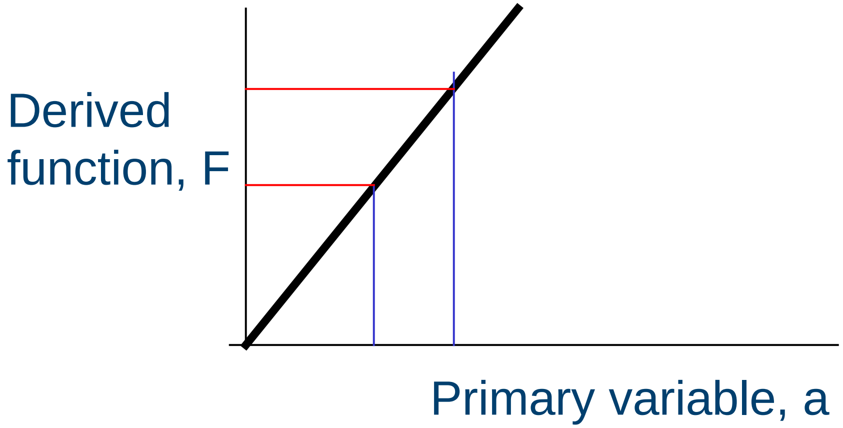
$$\frac{e_T}{P} = \pm 3.12/1000 = \pm 0.312 \ \%$$

Relative to full scale

Uncertainty propagation

Suppose a set of measurements is made which must be used to calculate some given function of the primary measurements.

We wish to estimate the uncertainty of the derived function on the basis of the uncertainties in the primary measurements.



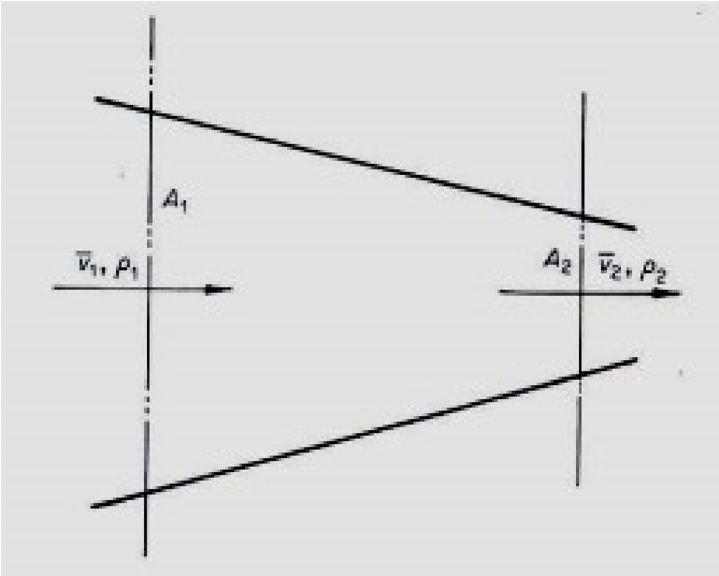
F = m a

If the function is linearly related to the primary variables

$$e_a = \pm\sqrt{e_1^2 + e_2^2 + ... + e_n^2} \quad \rightarrow \quad e_F = m\, e_a$$

Uncertainty propagation

When the measured quantity, such as the **flow rate**, is not linearly reated to the primary variables



$$Q \cong K\sqrt{\frac{2\Delta p}{\rho}}$$

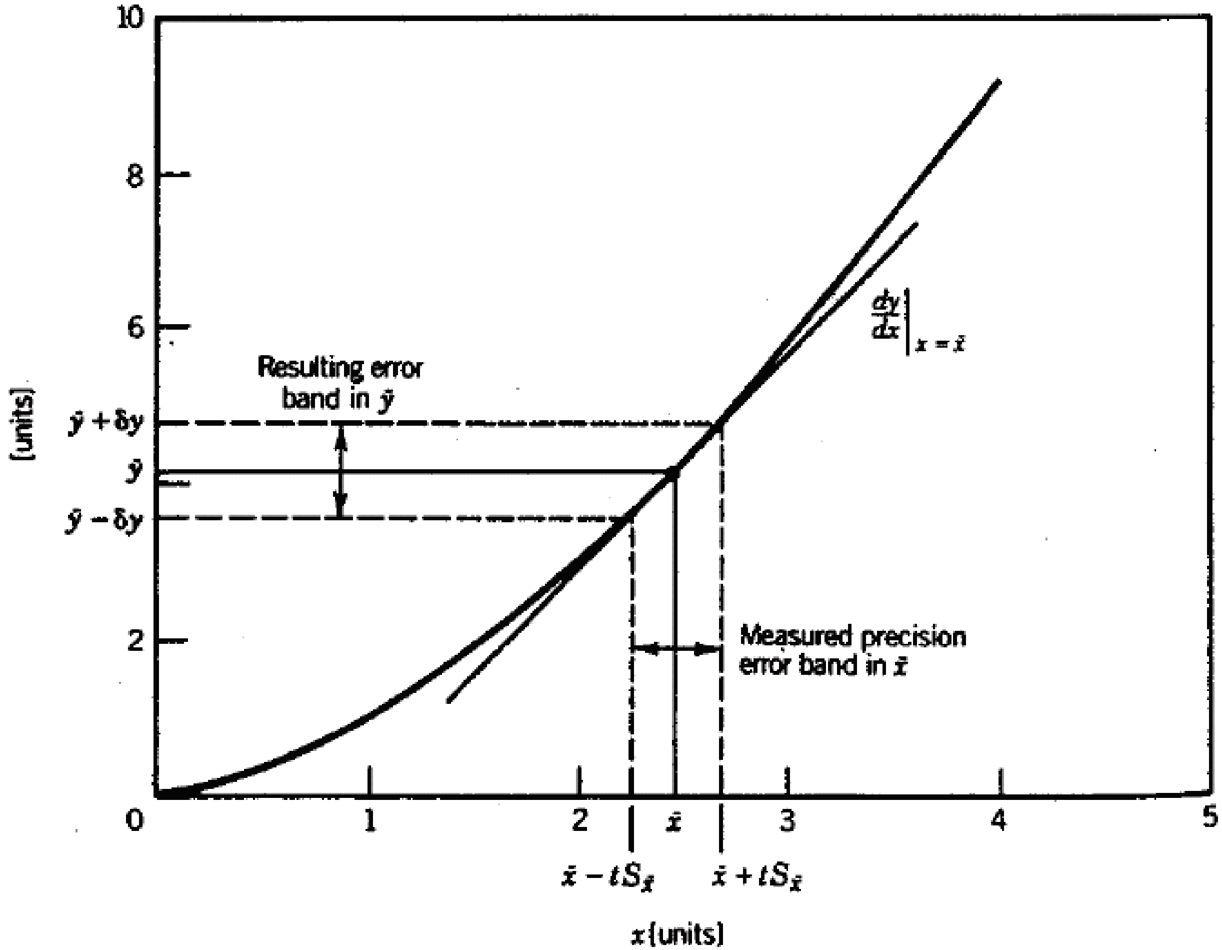
the error in Q should be evaluated by considering an equilibrium fixed condition and assuming small changes (uncertainties) in the primary variables.

Then the changes (uncertainty interval) in the output quantity caused by small changes (errors) in the input variables can be predicted using the numerical values of the partial derivatives of the known function.

Uncertainty propagation

Consider the confidence interval in x $\bar{x} \pm t\sigma_{\bar{x}}$

The corresponding interval in y will be $\bar{y} \pm \delta y = f(\bar{x} \pm t\sigma_{\bar{x}})$



Linear approximation ($t\sigma_x$ is small)

$$\delta y \approx \left(\frac{dy}{dx}\right) t\sigma_{\bar{x}}$$

The derivative may be interpreted as a

sensitivity coefficient

Uncertainty propagation

In general

$$y = f\left(x_1, x_2, \ldots x_n\right)$$

Total error

$$e_T = \pm\sqrt{\sum_1^n\left[\left(\frac{\partial f}{\partial x_i}\right)^2 e_i^2\right]}$$

Relative error

$$\frac{e_y}{\overline{y}} = \pm\sqrt{\sum_1^n\left[\left(\frac{\partial y}{\partial x_i}\right)^2 \frac{e_i^2}{\overline{x}_i^2}\right]}$$

Uncertainty propagation

Consider the pressure drop in a cyrcular duct (diameter D, length L) due to the wall effect: (***Darcy equation***):

$$\Delta p = f\,\frac{L}{D}\,\frac{\rho U^2}{2} \longrightarrow f = 2\Delta p\,\frac{D}{L}\,\frac{1}{\rho U^2}$$

Error in the evaluation of the ***friction coefficient***:

$$e_T = \pm\sqrt{\left(\frac{\partial f}{\partial \Delta p}\right)^2 e_{\Delta p}^2 + \left(\frac{\partial f}{\partial D}\right)^2 e_D^2 + \left(\frac{\partial f}{\partial L}\right)^2 e_L^2 + \left(\frac{\partial f}{\partial \rho}\right)^2 e_\rho^2 + \left(\frac{\partial f}{\partial U}\right)^2 e_U^2}$$

Uncertainty propagation

Absolute value

$$\frac{\partial f}{\partial \Delta p} = 2\frac{D}{L}\frac{1}{\rho U^2}$$

$$\frac{\partial f}{\partial D} = 2\frac{\Delta p}{L}\frac{1}{\rho U^2}$$

$$\frac{\partial f}{\partial L} = -2\frac{\Delta p D}{L^2}\frac{1}{\rho U^2}$$

$$\frac{\partial f}{\partial \rho} = -2\frac{\Delta p D}{L}\frac{1}{\rho^2 U^2}$$

$$\frac{\partial f}{\partial U} = -4\frac{\Delta p D}{L}\frac{1}{\rho U^3}$$

Relative value

$$\frac{1}{f}\frac{\partial f}{\partial \Delta p}e_{\Delta p} = \frac{e_{\Delta p}}{\Delta p}$$

$$\frac{1}{f}\frac{\partial f}{\partial D}e_D = \frac{e_D}{D}$$

$$\frac{1}{f}\frac{\partial f}{\partial L}e_L = -\frac{e_L}{L}$$

$$\frac{1}{f}\frac{\partial f}{\partial \rho}e_\rho = -\frac{e_\rho}{\rho}$$

$$\frac{1}{f}\frac{\partial f}{\partial U}e_U = -2\frac{e_U}{U}$$

Uncertainty propagation

Assuming that all the parameters are affected by a 1% error

$$\frac{e_T}{f} = \pm\sqrt{\left(\frac{e_{\Delta p}}{\Delta p}\right)^2 + \left(\frac{e_D}{D}\right)^2 + \left(\frac{e_L}{L}\right)^2 + \left(\frac{e_\rho}{\rho}\right)^2 + \left(2\frac{e_U}{U}\right)^2}$$

$$\frac{e_T}{f} = \pm\sqrt{4(0.01)^2 + 4(0.01)^2} = \sqrt{8\times(0.01)^2} = 2.83\,\%$$

Procedure

Bibliography

Books

Doebelin E.O., *Strumenti e Metodi di Misura*, McGraw-Hill, 2008 (seconda edizione)

Holman J.P., *Experimental Methods for Engineers*, McGraw-Hill, 1994

Figliola R.S. e Beasley D.E., *Theory and Design for Mechanical Measurements*, J. Wiley, 1991

Cameron Tropea, Alexander L. Yarin, John F. Foss (Eds.), Springer Handbook of Experimental Fluid Mechanics, 2007.

Bendat S.J. and Piersol A.G., Random Data: *Analysis and measurement procedures*, Wiley-Interscience, 1971.

Papers

Moffat R. J., Describing the Uncertainties in Experimental Results, Experimental Thermal and Fluid Science 1988; 1:3-17