The Lecture Contains:

Uncertainty Analysis

Errors in experimentation fall into two categories: bias and random. Bias errors are related to a consistent difference between the signal generated by a probe and that sensed by the measurement system. Most modern measurement systems permit self-calibration using a known in-built signal generator and bias errors can be easily minimized. Random errors arise from a host of uncontrollable factors that simultaneously affect the experiment. These include room temperature changes, supply voltage fluctuations, air currents, building vibration and roughness of nominally smooth surfaces. Random errors are a part of any experiment and these render the experimental data as fundamentally irreproducible. However, in any good experiment, measured data will have a stable mean value about which readings obtained in definite runs of the experiment are distributed. Such a distribution is called *scatter*; considerable effort must be expended to control and reduce scatter. This requires simultaneous improvement in the quality of the measurement systems and the test cell in which the experiment is performed. Experiments dedicated to determining scatter must be carried out to estimate its magnitude. Such experiments involve performing runs on: different days at identical room temperatures; on different apparatus that look alike; with measurement systems having identical specifications and with mild mis-orientation of the probe with respect to the flow. Data obtained from such experiments would result in a distribution of values about the mean.

Previous Next

Methods of analyzing scatter are described later in this lecture. Using these methods the value of a variable, say velocity u, is reported as

$$u = \bar{u} \pm \delta u, \quad 95\% \quad confidence \ \ interval$$

Here $\bar{u}$ is the mean value of u, δu the uncertainty in the value of $\bar{u}$ as determined from the scatter experiments and is usually expressed as a percentage. Confidence interval (CI) denotes the probability that a fresh experiment which determines u produces a value between $\bar{u} + \delta u$ and $\bar{u} - \delta u$.

A note of interest: Instruments of the earlier generation were limited by their *least count* . With substantial progress in science and technology, modern instruments can measure very small changes, say in resistance and voltage. Consequently, scatter in the experimental data arises from the influence of environmental factors on the flow field, rather than the measurement system itself.

Previous Next

Error Propagation

Associated with each measured variable x_i is an uncertainty δ_i originating from scatter. Uncertainty in universal constants such as acceleration due to gravity and fluid properties can be treated as negligible in comparison to δ_i. It is quite common to construct a new quantity y from the measured quantities x_i. For example, a combination of pitot static tube and manometer measures a pressure difference $(= x)$ and velocity $(= y)$ is recovered through a square root formula. It is of interest to determine the uncertainty δy in terms of those in x, namely δ_i. Since

$$y = f(x_i) \quad i = 1,2,\dots n$$

the following mathematical identity holds:

$$\delta y \leq \sum_i \left|\frac{\partial f}{\partial x_i}\right| |\delta_i|$$

In most applications, δy thus determined is substantially larger than the true uncertainty. A closer unbiased estimate used in engineering is

$$\delta y = \left(\sum_i \left(\frac{\partial f}{\partial x_i}\delta_i\right)^2\right)^{1/2} \qquad (1)$$

In many instances, the partial derivative of function f with x is of order unity; in any case it serves effectively as a scale factor. Hence, $\delta y \geq |\delta_i|$ for each i, and Equation 1 is called the error propagation formula.

Previous Next

Examples of use of Formula 1 are as follows:

Example: Consider the measurement of pressure in a gas through measurement of its temperature T and density ρ. Let the uncertainties in T and ρ be δT and $\delta \rho$ where p is determined from the ideal gas law $p = \rho RT$. Hence the uncertainty in p is

$$\delta p = ((RT\delta\rho)^2 + (\rho R\delta T)^2)^{1/2}$$

Here we have used $\partial p/\partial \rho = RT$ and $\partial p/\partial T = \rho R$. Further

$$\frac{\delta p}{\rho} = \left\{\left(\frac{\delta\rho}{\rho}\right)^2 + \left(\frac{\delta T}{T}\right)^2\right\}^{1/2}$$

Suppose the uncertainty in $\delta\rho$ is specified as 1% and δT as 2^o at $T = 27^o$. Note that T in the expression above is absolute temprature. Using $\delta\rho/\rho = 0.01$ and $\delta T/T = 2/300,\ \delta p/p$ is calculated as 0.0119, i.e. 1.19%.

Previous Next

Module 1: Introduction to Experimental Techniques

Lecture 6: Uncertainty analysis

Example: A battery supplies a current I to a resistor R with a potential difference E. The percentage uncertainties in E I and R are known to be identical. Should the power drawn from the battery ($= W$) be calculated as EI or I^2R. We have

$$W = EI, \qquad \frac{\delta W}{W} = \left\{\left(\frac{\delta E}{E}\right)^2 + \left(\frac{\delta I}{I}\right)^2\right\}^{1/2}$$

and

$$W = I^2R \quad \frac{\delta W}{W} = \left\{4\left(\frac{\delta I}{I}\right)^2 + \left(\frac{\delta R}{R}\right)^2\right\}^{1/2}$$

Clearly, the first formula is preferred since it indicates a lower level of uncertainty in W.

Previous Next

Example: Determine the uncertainty in velocity as a function of uncertainty in the head h when a pitot static tube is used for air velocity measurement along with an inclined tube manometer. Here, h is the displacement of the manometric liquid along the tube. The governing equation is

$$\frac{1}{2}\rho u^2 = \rho_l g\ h\ \sin\theta\ (\ assuming\ \rho_l \gg\ \rho(air))$$

where θ is the angle at which the arm of the manometer is inclined with respect to the horizontal and ρ_l the density of the manometric liquid. We assume ρ_l and g to be precisely known and constant. Hence

$$u = c\sqrt{h\ \sin\theta}\ = c\sqrt{h\theta}\ \ for\ small\ \theta$$

$$\frac{\delta u}{u} = \frac{1}{4}\left\{\left(\frac{\delta h}{h}\right)^2 + \left(\frac{\delta\theta}{\theta}\right)^2\right\}^{1/2}$$

Since θ is small, any uncertainty in its value will lead to large values of δu. Hence in practice θ is not used as a measured quantity for determining velocity. Instead the manometer is calibrated inclusive of θ and u is obtained as

$$u = c_\theta\sqrt{h}$$

Previous Next

Analysis of Scatter

We now return to the question of determining the uncertainty in a variable arising from scatter. Let x_i, $i = 1, \dots n$ be n different readings of the variable x obtained from n distinct but nominally identical (similar) experiments. The mean $\bar{x}$ and variance v for the set x_i are formally defined as

$$\bar{x} = \frac{1}{n}\sum_i x_i, \quad v = \frac{1}{n-1} \sum_i (x_i - \bar{x})^2$$

and the standard deviation σ is equal to $v^{1/2}$. The quantity $n - 1$ in the definition of v arises from the loss of one degree of freedom in forming the sum due to the presence of $\bar{x}$. If n is large, the probability distribution function B_x of the variable x can be assumed to be Gaussian in view of the central limit theorem (refer Probability Density Function Approach of Lecture 3). Hence, the scatter of values x about the mean $\bar{x}$ follows the formula

$$B_x = \frac{1}{\sigma\sqrt{2\pi}} \exp(-(x-\bar{x})^2/2\sigma^2)$$

Previous Next

Here $B_x dx$ is interpreted as the fraction of the total number of readings that falls between x and $x + dx$. Hence the probability that any given value is less than x_1 is

$$g(x_1) = \int_{-\infty}^{x_1} B_x \, dx$$

and the probability that it is between x_1 and x_2 is

$$g(x_2) - g(x_1) = \int_{x_1}^{x_2} B_x \, dx$$

The function $p(z) = g(z) - g(0)$ where the dimensionless variable z is $z = (x - \bar{x})/\sigma$ is available in tabular form (Table A1) and Figure 1.15

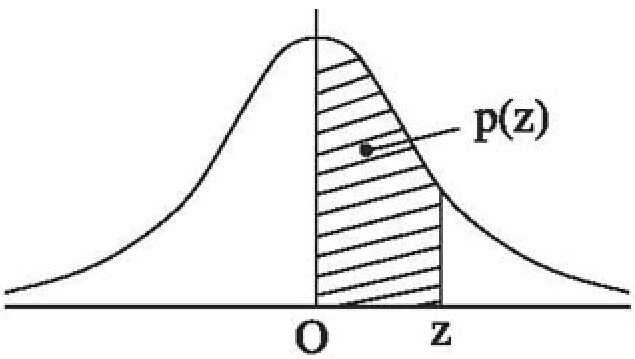

Figure 1.15: Definition of $p(z)$ for Gaussian PDF (Table A1).

Example: Given $\bar{x} = 5$ and $\sigma = 1$ find the probability that the value of x in a new experiment is (1) between 4.5 and 5.5, (2) between 4.5 and 5.75, (3) less than 6.5 and (4) between 6 and 7.

Let z_1 and z_2 be the limits between which the probability must be determined. The values of z in each case are (1) $z = -0.5$ and 0.5, (2) $z = -0.5$ and 0.75, (3) $z < 1.5$ and (4) z between 1 and 2. The required probabilities are $p\ (z_2) - p\ (z_1)$; hence in (1), this quantity is $0.1915 - (-0.1915) = 0.383$; in (2), $0.2734 - (-0.1915) = 0.4649$; in (3), $0.4332 - (-0.5) = 0.9332$ and in (4), $0.4772 - 0.3413 = 0.1359$.

The most commonly used probability interval called the 95% confidence interval corresponds to $z = \pm 1.96$ since $p(1.96) - p(-1.96) = 2(0.475) = 0.95$. Hence normally distributed scatter can be specified as $\bar{x} \pm 1.96\sigma$ with 95% confidence. For convenience this interval is written as $\bar{x} \pm 2\sigma$. Gaussian distribution for scatter can be assumed if the number of repetitive experiments is greater than 30.

For n points with a mean $\bar{x}$, points that fall outside a probability value of $1 - (1/2n)$ are to be rejected. The mean and the variance $\bar{x}$ and σ must then be recalculated. This is called *Chauvenet's criterion* for statistical rejection of data. The criterion can be applied only once to the data set. It determines the acceptable level of scatter in an experiment arising from a large number of random influences.

Previous Next

Example: Consider the data set (653, 680, 677, 436, 679) with $n = 5$. (A small data set is used here to facilitate a quick arithmetic calculation.) The average $\bar{x} = 625$ and the standard deviation is 106.2. The z values $|(x - \bar{x})/\sigma|$ for the data set are (0.26, 0.52, 0.49, 1.78, 0.51) respectively. The acceptable values of z are those for which $p(z) < 1 - 1/2n = 0.9$, i.e. $z < 1.645$. Hence the value of $x(= 436)$ for which $z = 1.78$ is unacceptable and can be rejected. With $n = 4$ we recompute $\bar{x}$ as 672.2 and σ as 12.89.

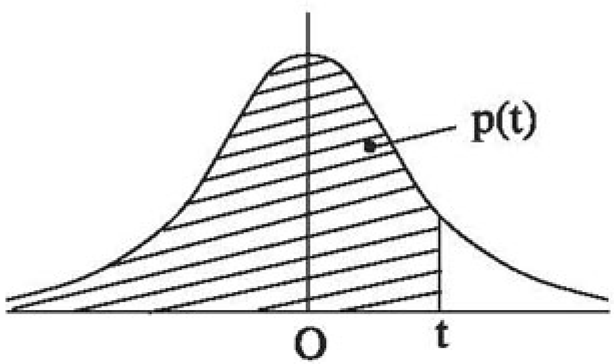

Figure 1.16: Definition of $p(t)$ in Student's- t Distribution (Table A1).

Previous Next

The Gaussian distribution is an acceptable model for scatter if the number of elements in the data set is large. In practice we require $n > 30$. In many experiments it is not possible to repeat a reading such a large number of times either because of cost considerations or because of the nature the experiment itself. If $n > 30$ the 95% confidence interval is written as

$$x = \bar{x} \pm t\sigma$$

where t depends on n. In general t increases as n decreases; for $n > 30$, the limit $t = 1.96$ is reached, the value appropriate for Gaussian distribution. The main effect of having a small number of data points is that the mean is not stationary and changes with n.. If it is assumed that the scatter in the mean is Gaussian it can be shown that the standard deviation in the mean $\bar{x}$ is $\sigma/\sqrt{n}$, where σ is the standard deviation in x. The distribution that takes into account the changes in the mean value as well is called Student's t distribution. Values of t as a function of m the degree of freedom $(= n - 1)$and the percentage confidence interval are also available in tabular form (Table A2, Figure 1.16). Once $t$is determined, the scatter in data is specified as

$$x = \bar{x} \pm t\sigma \quad and \quad \bar{x} = \bar{\bar{x}} \pm t\frac{\sigma}{\sqrt{n}}$$

Module 1: Introduction to Experimental Techniques

Lecture 6: Uncertainty analysis

Example: Specify the 95% confidence interval for a variable x if $n = 9$, $\bar{x}(=\bar{\bar{x}}) = 17$ and $\sigma = 4$. Here $n - 1 = 8$ and $t = 2.31$. Hence

$$x = 17 \pm 9.24\,, \qquad \bar{x} = 17 \pm 3.08$$

Example: Given 4 values obtained by repeating an experiment as 65.3, 68.2, 67.7, 66.4, find the range in which the fifth value will lie. Here $n = 4$, $\bar{x} = 66.9$, $\sigma = 1.308$; for $n - 1 = 3$ and 95% confidence interval $t = 3.18$. We get

$$x = 66.9 \pm 4.15$$

$$\bar{x} = 66.9 \pm 2.075$$

The new value will lie between 71.05 and 62.75 and the new mean will be between 68.97 and 64.825.

It is also worth noting that the probability of obtaining a particular value is zero; finite probability can be assigned only if a range of values of the measured variable is prescribed.

Table A1: Area $p(z)$ under the Normal Curve

z	0	1	2	3	4	5	6	7	8	9
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0754
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2258	.2291	.2324	.2357	.2380	.2422	.2454	.2486	.2518	.2549
0.7	.2580	.2612	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2996	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4420	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990
3.1	.4990	.4991	.4991	.4991	.4992	.4992	.4992	.4992	.4993	.4993
3.2	.4993	.4993	.4994	.4994	.4994	.4994	.4994	.4995	.4995	.4995

3.3	.4995	.4995	.4995	.4996	.4996	.4996	.4996	.4996	.4996	.4997
3.4	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4998
3.5	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998
3.6	.4998	.4998	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999
3.7	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999
3.8	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999
3.9	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000

Previous Next

Table A 2: Percentile Values for Student's- t Distribution

Degrees of Freedom m

(Shaded area = p)

m	p= .995	.99	.975	.95	.90	.80	.75	.70	.60	.55
01	63.66	31.82	12.71	6.31	3.08	1.376	1.000	.727	.325	.158
02	9.92	6.96	4.30	2.92	1.89	1.061	.816	.617	.289	.142
03	5.84	4.54	3.18	2.35	1.64	.978	.765	.584	.277	.137
04	4.60	3.75	2.78	2.13	1.53	.941	.741	.569	.271	.134
05	4.03	3.36	2.57	2.02	1.48	.920	.727	.559	.267	.132
06	3.71	3.14	2.45	1.94	1.44	.906	.718	.553	.265	.131
07	3.50	3.00	2.36	1.90	1.42	.896	.711	.549	.263	.130
08	3.36	2.90	2.31	1.86	1.40	.889	.706	.546	.262	.130
09	3.25	2.82	2.26	1.83	1.38	.883	.703	.543	.261	.129
10	3.17	2.76	2.23	1.81	1.37	.879	.700	.542	.260	.129
11	3.11	2.72	2.20	1.80	1.36	.876	.697	.540	.260	.129
12	3.06	2.68	2.18	1.78	1.36	.873	.695	.539	.259	.128
13	3.01	2.65	2.16	1.77	1.35	.870	.694	.538	.259	.128
14	2.98	2.62	2.14	1.76	1.34	.868	.692	.537	.258	.128
15	2.95	2.60	2.13	1.75	1.34	.866	.691	.536	.258	.128
16	2.92	2.58	2.12	1.75	1.34	.865	.690	.535	.258	.128
17	2.90	2.57	2.11	1.74	1.33	.863	.689	.534	.257	.128
18	2.88	2.55	2.10	1.73	1.33	.862	.688	.534	.257	.127
19	2.86	2.54	2.09	1.73	1.33	.861	.688	.533	.257	.127
20	2.84	2.53	2.00	1.72	1.32	.860	.687	.533	.257	.127
21	2.83	2.52	2.08	1.72	1.32	.859	.686	.532	.257	.127
22	2.82	2.51	2.07	1.72	1.32	.858	.686	.532	.256	.127
23	2.81	2.50	2.07	1.71	1.32	.858	.685	.532	.256	.127
24	2.80	2.49	2.06	1.71	1.32	.857	.685	.531	.256	.127
25	2.79	2.48	2.06	1.71	1.32	.856	.684	.531	.256	.127
26	2.78	2.48	2.06	1.71	1.32	.856	.684	.531	.256	.127
27	2.77	2.47	2.05	1.70	1.31	.855	.684	.531	.256	.127
28	2.76	2.47	2.05	1.70	1.31	.855	.683	.530	.256	.127

29	2.76	2.46	2.04	1.70	1.31	.854	.683	.530	.256	.127
30	2.75	2.46	2.04	1.70	1.31	.854	.683	.530	.256	.127
40	2.70	2.42	2.02	1.68	1.30	.851	.681	.529	.255	.126
60	2.66	2.39	2.00	1.67	1.30	.848	.679	.527	.254	.126
120	2.62	2.36	1.98	1.66	1.29	.845	.677	.526	.254	.126
∞	2.58	2.33	1.96	1.645	1.28	.842	.674	.524	.253	.126
[p = 0.975 represents 95% confidence interval.]										