Lecture 05.03 Rigorous uncertainty analysis

confidence

sample standard deviation

sample mean of standard deviations

sample mean

sample standard deviation of sample means

We have learned about *confidence* from a statistical point-of-view. Based on sample variability, we might have $P\%$ confidence in a given measurement. As we have learned, an estimate of the variability of the random variable X—in this case, the measurand—is given by its *sample standard deviation* S_X. In the case of multiple samples, its best estimate was the *sample mean of standard deviations* $\overline{S_X}$. The best estimate of the value of X (that is, it's mean μ_X) was the *sample mean* $\overline{X}$. In the case of multiple samples, it was the sample mean of sample means $\overline{\overline{X}}$. Finally, the best estimate of the variability of the mean was the *sample standard deviation of sample means* $S_{\overline{X}}$. Further recall the nice estimate of $S_{\overline{X}}$ from a *single sample* with size N:

random error

random uncertainty

confidence intervals

As we can see, as sample size N increases, $S_{\overline{X}}$ decreases. This type of error in a measurement is called *random error* and gives rise to *random uncertainty* u_r, related to what we have called *confidence intervals* about the best estimate of the mean, such as, for a single sample of size N,

$$x \in \overline{X} \pm t_{\nu,P} S_{\overline{X}} \qquad (P\%) \tag{05.2}$$

where $\nu = N-1$ is the degree of freedom and $P\%$ is our confidence based on the probability $P\%$ of a Student random variable X taking a value x within $t_{\nu,P}$ standard deviations of the mean. The random uncertainty is a half-interval

$$u_r = t_{\nu,P}\, s \qquad (P\%) \tag{05.3}$$

standard random uncertainty

where $s = S_{\overline{X}}$ (68%) is the *standard random uncertainty*, which is simply one standard deviation of the means.

bias

systematic error

systematic uncertainty

systematic standard uncertainty

However, error can arise from more than randomness. Other sources arise that *bias* the measured values—say, up or down—from the mean. This is called *systematic error* and generates *systematic uncertainty* u_b (because *bias*) and *systematic standard uncertainty* b that has 68% confidence. Let a measurement instrument's manual list an elemental error B, which (unless otherwise stated in the manual) is assigned a 95% confidence; the systematic standard uncertainty is $b = B/2$. Assuming a large sample was used to estimate B, we might report an uncertainty associated with that error to be $u_b = 2b$ with 95% confidence (we are assuming a Gaussian distribution, but the distribution shape has little effect).

Let's consider systematic error a bit more, through an example. A scale might be systematically reading high (I know my scale does, especially around the holidays). This can be identified and mitigated by *calibration* to a standard. The National Institute of Standards and Technology (NIST) calibrates weights. Let's say you have a 10 kg NIST calibrated weight (such an object is called a *standard*) with one-normal-standard deviation confidence $\pm 200 \cdot 10^{-9}$ kg. Let's say I weigh it $N = 10$ times on my scale and the sample mean $\bar{x} = 10.5$ kg and sample standard deviation $S_x = 0.3$ kg. The calibration allows me to adjust the bias on my scale by $10 - 10.5 = -0.5$ kg.

calibration

standard

However, there remain two systematic uncertainties associated with my scale's bias: (1) NIST's standard uncertainty $b_{std} = 200 \cdot 10^{-9}$ kg due to NIST's measurement of the standard 10 kg weight and (2) our calibration standard uncertainty

The systematic standard uncertainties are combined in the usual RSS way (although the calibration uncertainty clearly dominates):

A measurement sample of size $M = 23$ of an object with unknown mass m is then performed with the calibrated scale. The sample mean is $\overline{m} = 9.04$ kg and sample standard deviation $S_m = 1$ kg. How confident can we be in the result? Certainly both the systematic and random certainties must contribute. Before we can consider their combined effect, let's compute the standard random uncertainty:

We combine the systematic and random uncertainties in the usual RSS way:

Since these are *standard* uncertainties, we must find its effective degree of freedom ν before assigning it a confidence with a Student t-score. How can we estimate this from these different sources of uncertainty, each with their own degree of freedom? A method of estimating the effective degree of freedom is given by (Figliola and Beasley, 2015) and is presented in an equivalent form here. Let a measurement have J random standard uncertainties s_j with corresponding degrees of freedom ν_{s_j}; further, let it have K systematic standard uncertainties b_k with corresponding degrees of freedom ν_{b_k}; then the effective degree of freedom is

$$\nu = \frac{\left(\sum_{j=1}^{J} s_j^2 + \sum_{k=1}^{K} b_k^2\right)^2}{\sum_{j=1}^{J} s_j^4/\nu_{s_j} + \sum_{k=1}^{K} b_k^4/\nu_{b_k}}. \tag{05.4}$$

From above, we have $J = 1$ and $K = 2$ and standard uncertainties given in Table 05.1. This gives $\nu = 29.0$. That's close enough to 30 to call it "large" and assign a 95% confidence uncertainty

So, using our 95% confidence uncertainty for our interval, our best estimate for the mass is

Table 05.1: summary of standard uncertainties for a fictional mass measurements.

uncertainty	value s or b	deg. of freedom
s	0.209 kg	22
b_{std}	$200 \cdot 10^{-9}$ kg	∞
b_{cal}	0.0949 kg	9

05.03.1 An extensive example

Consider a temperature measurement made with a linear calibrated temperature-voltage transducer. The calibration data is given as `t_cal` (units C) and `v_cal` (units V). The measurement voltage sample is given as a time series `v_a` (units V) versus `time_a` (units s), where we can assume relatively constant measurement conditions and a stationary process.

The voltmeter (used for calibration and for data) and the thermometer (used for calibration) have the systematic uncertainties defined below.

```
bv_1 = .1; % V ... voltmeter absolute uncertainty
bv_2 = .05; % V ... voltmeter linearity uncertainty
bt_1 = .05; % C ... thermometer absolute uncertainty
```

05.03.1.1 Calibration curve and its uncertainty

Let's first consider the calibration data.

```
disp('sample data (time,voltage)')
disp([time_a;v_a]')
```

```
sample data (time,voltage)
         0    5.7959
    1.8182    5.5286
    3.6364    5.2110
    5.4545    5.4191
```

```
    7.2727    5.6164
    9.0909    5.6746
   10.9091    5.4349
   12.7273    5.8535
   14.5455    5.6782
   16.3636    5.7058
   18.1818    5.5820
   20.0000    6.3077
```

Let's perform a linear regression analysis on the calibration data to find a *calibration curve*. The standard uncertainty of a polynomial regression of order $m-1$ and data with values $\tilde{y}_i$ approximating calibration curve values y_i with sample size N is (Figliola and Beasley, 2015, Equation 4.35)

$$s_{\text{fit}} = \sqrt{\frac{\sum_{i=1}^{N} (\tilde{y}_i - y_i)^2}{\nu}} \tag{05.5}$$

where the degree of freedom $\nu = N - (m+1)$. For a linear fit, $m = 2$.

```
pf_cal = polyfit(v_cal,t_cal,1)
k_trans = pf_cal(1); % this is the transducer gain
p_cal = polyval(pf_cal,v_cal);
d_cal = p_cal - t_cal;
nu_cal = length(d_cal)-(2+1)
s_cal = sqrt((sum(d_cal.^2))/nu_cal)
```

```
pf_cal =
    4.9994   -0.2028
nu_cal =
    22
s_cal =
    0.4712
```

```
h = figure;
p = plot(v_cal,t_cal,'x'); hold on
p2 = plot(v_cal,p_cal,'r-');
grid on;
xlabel('voltage (V)')
ylabel('temperature (C)')
hgsave(h,'figures/temp');
```

05.03.1.2 Random uncertainty

Another source of random error is the finite sample size. It can be computed, in the usual way, as the sample standard deviation of the sample means. And first, of course, the sample `v_a` must be passed through the calibration curve `pf_cal`.

```
t_a = polyval(pf_cal,v_a);
mu_t = mean(t_a)
s_mu_t = std(t_a)/length(t_a)
nu_a = length(t_a)-1
```

```
mu_t =
   28.0473
s_mu_t =
    0.1132
nu_a =
    11
```

The total random uncertainty is the root-sum-square (RSS) combination of the calibration and finite sample size uncertainties.

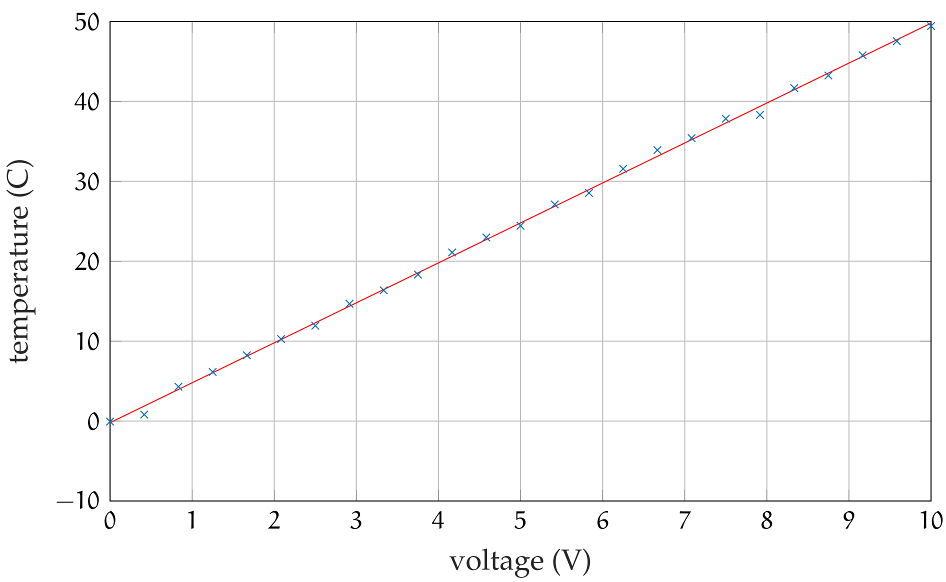

Figure 05.1: voltage-temperature transducer calibration data with its linear fit.

```
s = sqrt(s_cal^2+s_mu_t^2)
```

```
s =
    0.4846
```

05.03.1.3 Systematic uncertainty

The total systematic uncertainty is an RSS combination of those systematic uncertainties described in the problem statement. The transducer gain, found from the calibration curve, can be used to convert voltage uncertainties to temperature uncertainties.

```
b = sqrt(2*(k_trans*bv_1)^2+2*(k_trans*bv_2)^2+bt_1^2)
```

```
b =
    0.7921
```

Note the factors of two. These are due to the voltmeter's use in the calibration and in the sample.

05.03.1.4 Total uncertainty

The total *standard* uncertainty is the RSS combination of the standard random and systematic uncertainties.

```
u_t = sqrt(s^2+b^2)
```

```
u_t =
    0.9286
```

In order to assign a confidence interval via a t-score, we can use Equation 05.4 to compute the effective degree of freedom of the standard uncertainty. Given no information to the contrary, we assume the degree of freedom for each systematic uncertainty is high.

```
nu_t = (u_t^2)^2/(s_cal^4/nu_cal+s_mu_t^4/nu_a)
```

```
nu_t =
  329.5059
```

This is much greater than 30, so we can assume the distribution is Gaussian and use a z-score. Let's assign a 95% confidence uncertainty.

```
u_t_95 = 2*u_t
```

```
u_t_95 =
    1.8571
```

So a confidence interval for the estimate of the temperature is as follows.

```
mu_t_int = mu_t + [-1,1]*u_t_95
```

This is the result of our full uncertainty analysis.