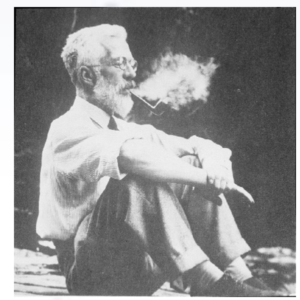
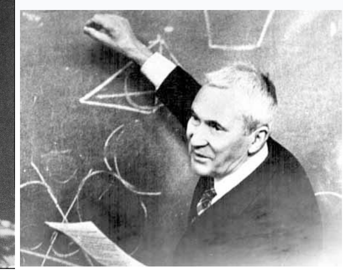
Applied Statistics

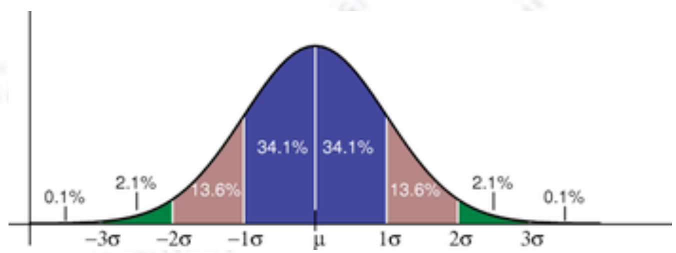
Error propagation





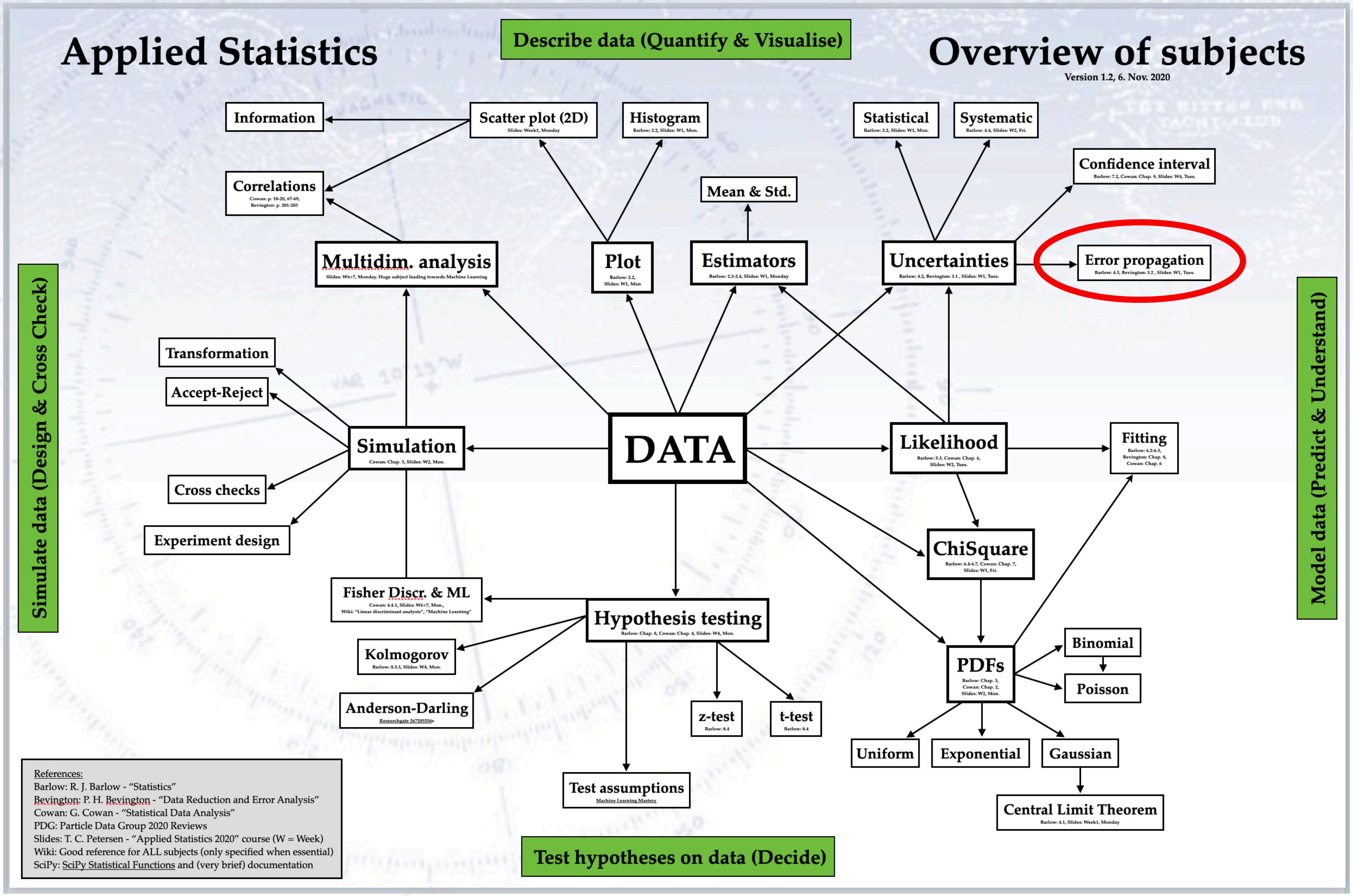



Troels C. Petersen (NBI)



"Statistics is merely a quantisation of common sense"

Error propagation



Error propagation

Imagine that y is a function of x_i

$$y(x_i)$$

and that we wish to find the error on y from the errors on x_i .

$$\sigma(x_i) = 0.8$$
$$\sigma(y(x_i)) = ?$$

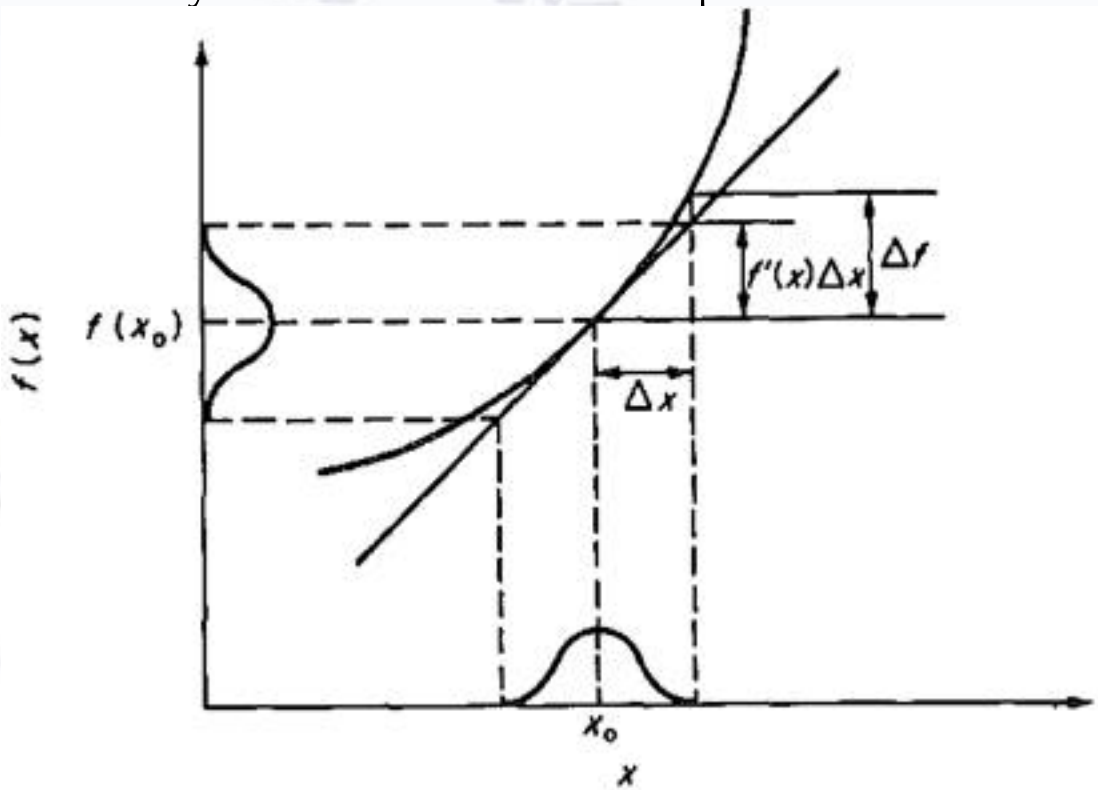
Error propagation

Imagine that y is a function of x_i

$$y(x_i)$$

and that we wish to find the error on y from the errors on x_i .

$$\sigma(x_i) = 0.8$$
$$\sigma(y(x_i)) = ?$$



Error propagation

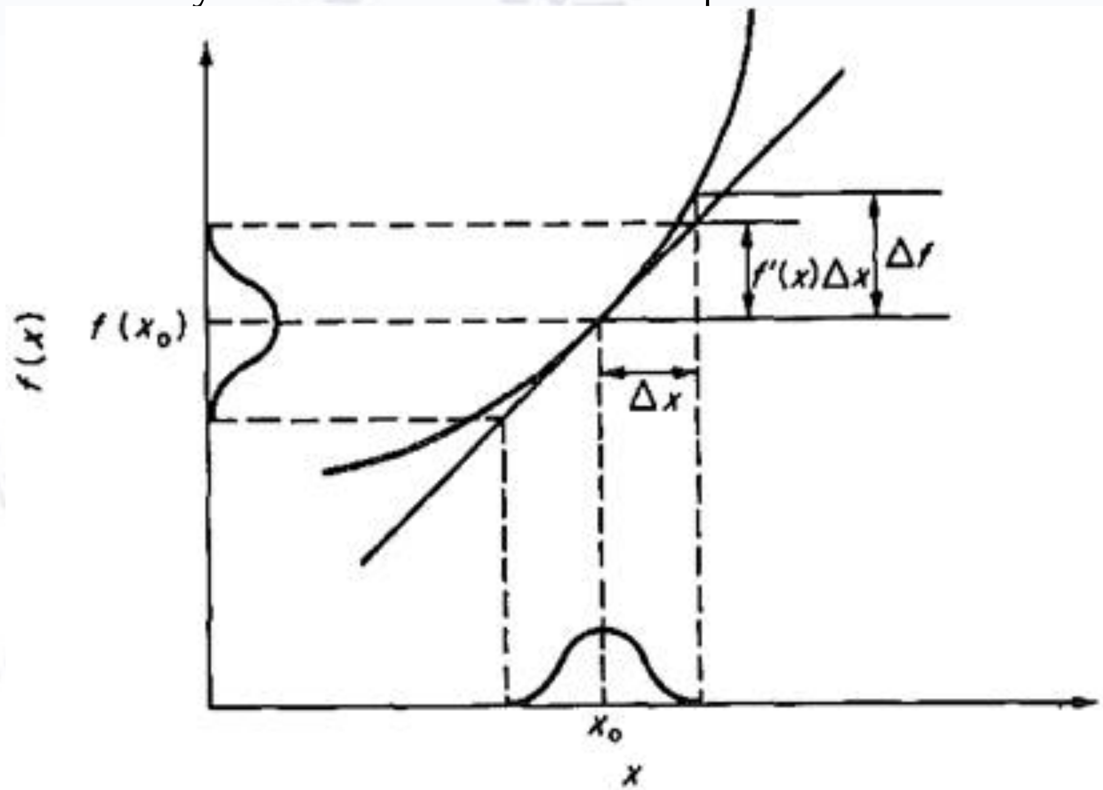
Imagine that y is a function of x_i

$$y(x_i)$$

and that we wish to find the error on y from the errors on x_i .

$$\sigma(x_i) = 0.8$$

$$\sigma(y(x_i)) = \frac{\partial y}{\partial x_i} \times 0.8$$



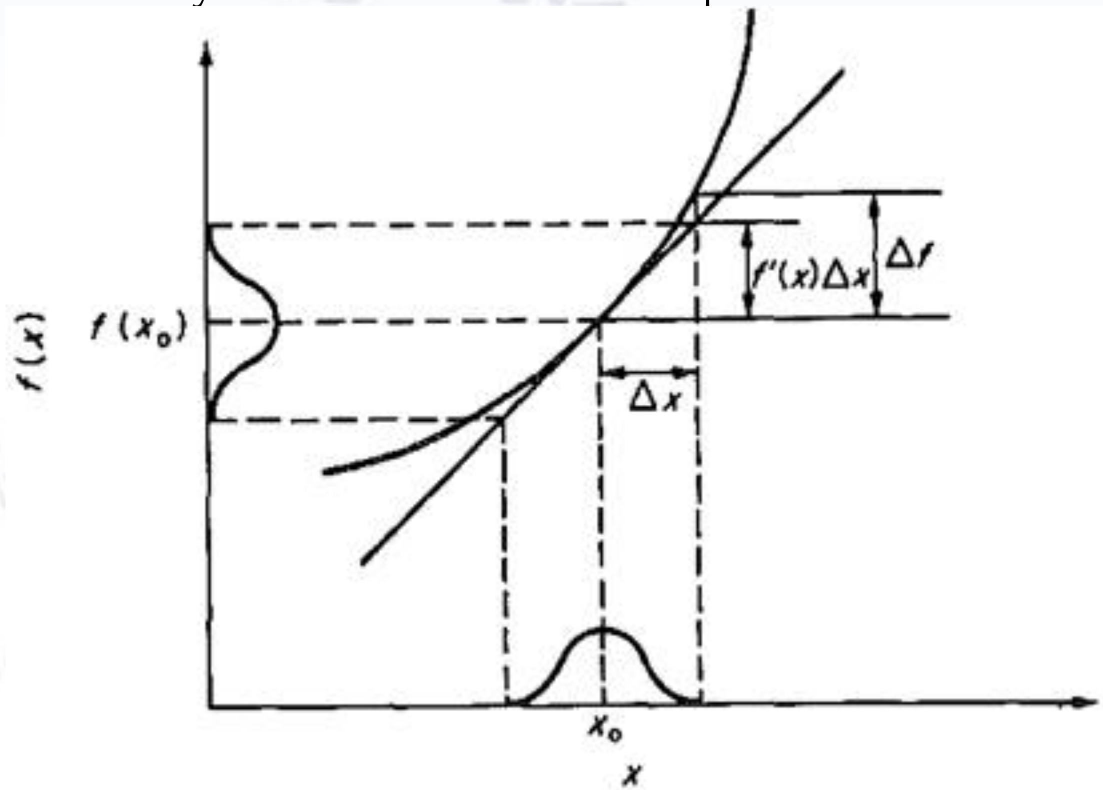
Error propagation

Note, the approximation here:
The derivative of y - dy/dx_i - should be relatively constant.
If not, the error propagation formula breaks down.

and that we wish to find the error on y from the errors on x_i .

$$\sigma(x_i) = 0.8$$

$$\sigma(y(x_i)) = \frac{\partial y}{\partial x_i} \times 0.8$$



Error propagation

Imagine that y is a function of x_i, and that we wish to find the error on y from the errors on x_i. Making a Taylor expansion of the function y gives:

$$y(\bar{x}) \simeq y(\bar{\mu}) + \sum_i^n \frac{\partial y}{\partial x_i}(x_i - \mu_i)$$

In order to get the uncertainty of y as a function of the variables x_i we calculate:

$$\sigma_x^2 = \overline{x^2} - \bar{x}^2 = E[x^2] - E^2[x]$$

$$E[y(\bar{x})] \simeq y(\bar{\mu})$$

$$E[y^2(\bar{x})] \simeq y^2(\bar{\mu}) + \sum_{i,j}^n \left[\frac{\partial y}{\partial x_i}\frac{\partial y}{\partial x_j}\right] V_{ij}$$

Error propagation formula

Subtracting the two formulae, we obtain:

$$\sigma_y^2 = \sum_{i,j}^{n} \left[\frac{\partial y}{\partial x_i} \frac{\partial y}{\partial x_j} \right]_{\bar{x}=\bar{y}} V_{ij}$$

If there are no correlations, only the diagonal (individual errors) enter:

$$\sigma_y^2 = \sum_{i}^{n} \left[\frac{\partial y}{\partial x_i} \right]_{\bar{x}=\bar{y}}^2 \sigma_i^2$$

Error propagation formula

Subtracting the two formulae, we obtain:

Note, that **each term** represents the **individual contributions** of x_i to the uncertainty on y.

If there are no correlations, only the diagonal (individual errors) enter:

$$\sigma_y^2 = \sum_i^n \left[\frac{\partial y}{\partial x_i} \right]_{\bar{x}=\bar{y}}^2 \sigma_i^2$$

Specific error propagation formula
Addition

Specific formula:

$$x = u + v$$

$$\sigma_x^2 = \sigma_u^2 + \sigma_v^2 + 2V_{uv}$$

General formula:

$$x = au + bv$$

$$\sigma_x^2 = a^2\sigma_u^2 + b^2\sigma_v^2 + 2abV_{uv}$$

"When adding numbers, their errors add in quadrature"

Specific error propagation formula

Multiplication

$$x = uv$$

$$\sigma_x^2 = (v\sigma_u)^2 + (u\sigma_v)^2 + 2uvV_{uv}$$

Dividing by x^2 to get relative terms, we obtain:

$$\frac{\sigma_x^2}{x^2} = \frac{\sigma_u^2}{u^2} + \frac{\sigma_v^2}{v^2} + 2\frac{V_{uv}}{uv}$$

"When multiplying numbers, their RELATIVE errors add in quadrature"

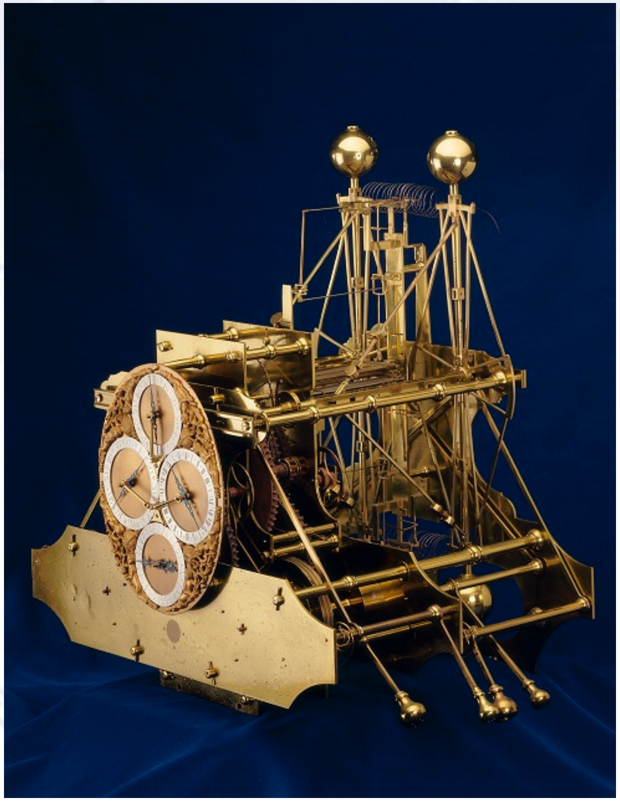
Error propagation at work...

John Harrison (24 March 1693 – 24 March 1776)

British clockmaker extraordinaire

“Won” the Longitude Act prize (3 sec/day).

Harrison's first sea clock (H1)

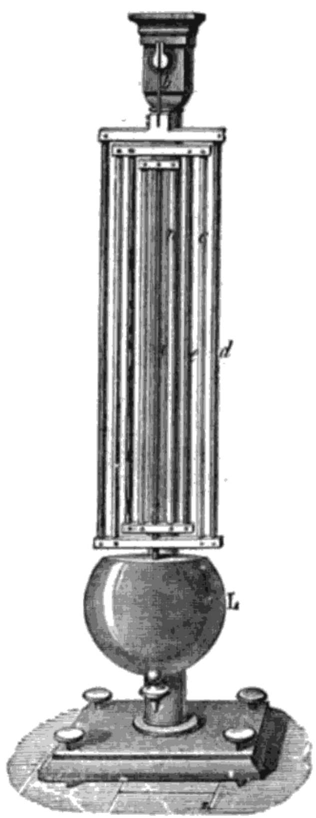


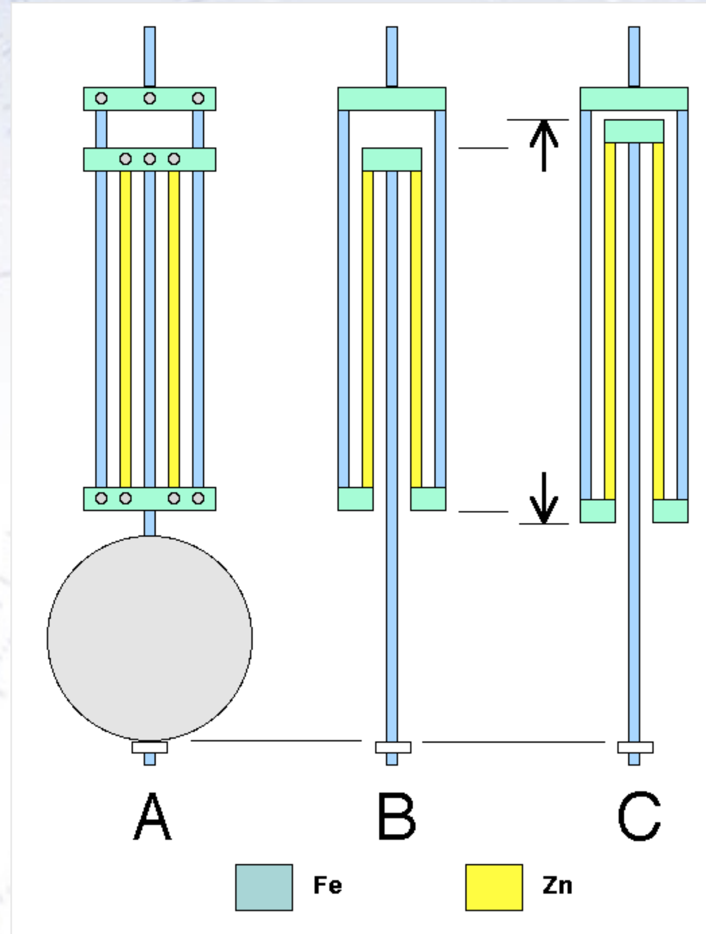
Harrison build H1-H5.

K1 (Copy of H4) was used by James Cook.

Error propagation at work...

Harrison's Gridiron pendulum is designed to cancel the change in length (in fact moment of inertia) with temperature.





Coefficient of thermal expansion:
Iron = 11.8×10^{-6}/C° Zinc = 30.2×10^{-6}/C°

Error propagation at more work...

Analysis of tiny differences in Uranus' orbit from Newtonian prediction led to the prediction and discovery of Neptune!

Continuing with Mercury...

TABLE II. Contributions to the motion of the perihelia of Mercury and the earth.

Cause	m^{-1}	Motion of perihelion: Mercury	Motion of perihelion: Earth
Mercury	6 000 000 ±1 000 000	0″.025±0″.00	−13″.75±2″.3
Venus	408 000 ± 1 000	277.856±0.68	345.49±0.8
Earth	329 390 ± 300	90.038±0.08	
Mars	3 088 000 ± 3 000	2.536±0.00	97.69±0.1
Jupiter	1 047.39± 0.03	153.584±0.00	696.85±0.0
Saturn	3 499 ± 4	7.302±0.01	18.74±0.0
Uranus	22 800 ± 300	0.141±0.00	0.57±0.0
Neptune	19 500 ± 300	0.042±0.00	0.18±0.0
Solar oblateness		0.010±0.02	0.00±0.0
Moon			7.68±0.0
General precession (Julian century, 1850)		5025.645±0.50	5025.65±0.5
Sum		5557.18 ±0.85	6179.1 ±2.5
Observed motion		5599.74 ±0.41	6183.7 ±1.1
Difference		42.56 ±0.94	4.6 ±2.7
Relativity effect		43.03 ±0.03	3.8 ±0.0

Urbain Le Verrier (1811-1877)

Simulating error propagation

Imagine that y is a very complicated function of x_i, perhaps not even parametric (i.e. not a function, but rather a model).

A simple method is to use simulation:

- Choose random values of x_i, corresponding to mean and SD of each x.
- Calculate $y(x_i)$ and record the resulting values.
- The standard deviation (and distribution) of y reflects the impact of x_i.

Note that the distribution of y may NOT be Gaussian, if the error propagation formula breaks down. It is then important to make this clear to the reader.

However, simulation exactly allows one to see to what degree the resulting distribution in y is Gaussian.

Error propagation exercise

Error propagation exercise

Imagine that you measure the Length (L) and Width (W) of a table:

$$L = 3.5 \pm 0.4 \text{ and } W = 0.8 \pm 0.2$$

Now you want to calculate the Perimeter (P), Area (A), and Diagonal (D), along with the uncertainty on these. This problem is an obvious case of applying the error propagation formula.

$L = 3.5 \pm 0.4$

$W = 0.8 \pm 0.2$

But imagine that you were asked to propagate the error through:

```
y1 = log(square(L*tan(W))+sqrt((L-W)/(cos(W)+1.0+L)))
```

Use simulation… not even certain that error propagation formula holds!