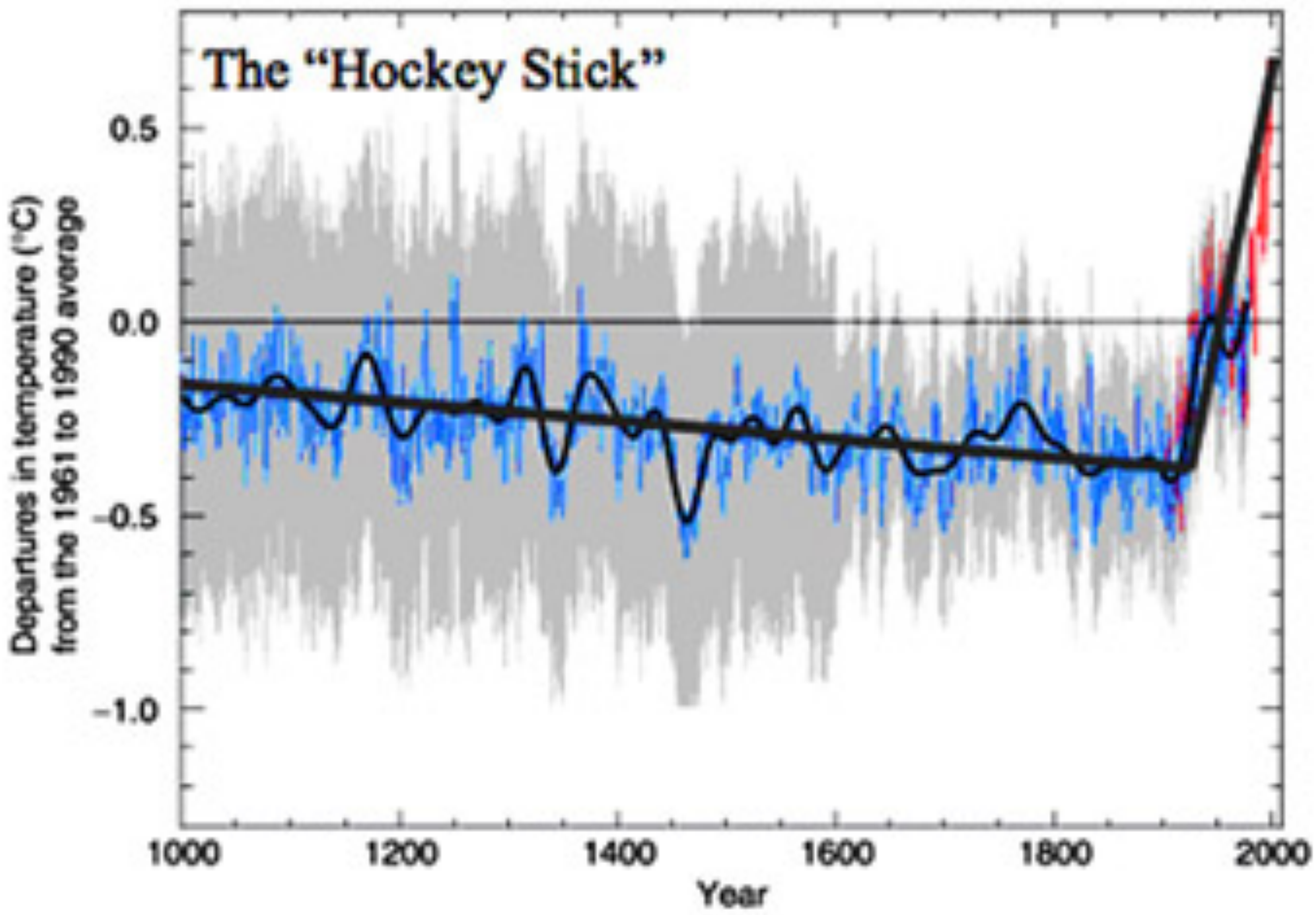
Data Uncertainty, Data Fitting and Error Propagation

E80 SPRING 2016 – THE ULTIMATE ADVENTURE

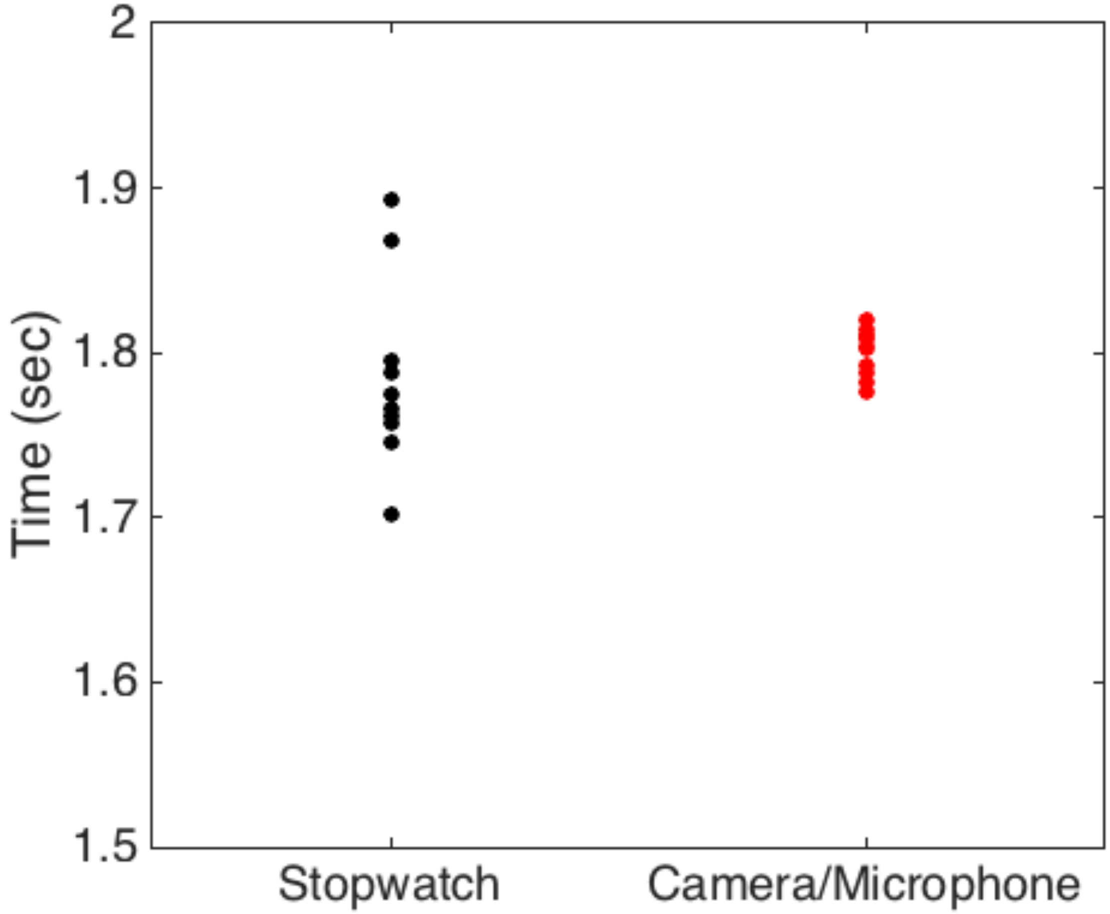
Topics

- Data uncertainty and confidence in measurements
- Data fitting – linear regression
- Error propagation
- Quantization error

Why do we care?



Timing Ball Drops

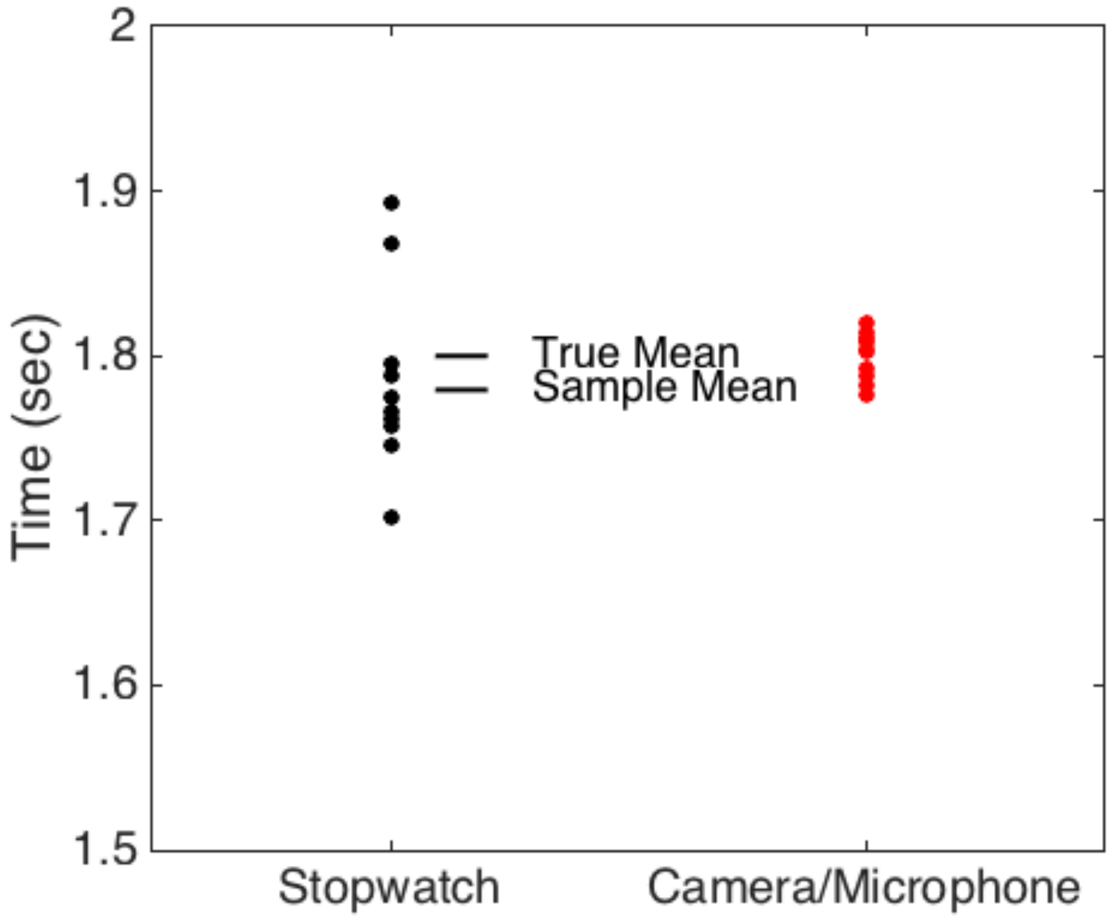


Timing Ball Drops

What is the mean value?

How certain are you?

Timing Ball Drops



Measured Value

Our data do not let us calculate the **true mean** μ

We calculate the **sample mean**

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

Measured Value

Our data do not let us calculate the **error** in each measurement

$$\varepsilon_i = x_i - \mu$$

We calculate the **residuals**

$$e_i = x_i - \bar{x}$$

How certain are we? $\bar{x} \pm \lambda$

Use **estimated standard error** and Student's *t*-test to find **confidence interval**

$$\lambda = tESE = \frac{tS}{\sqrt{N}}$$

Sample standard deviation $S = \sqrt{S^2}$

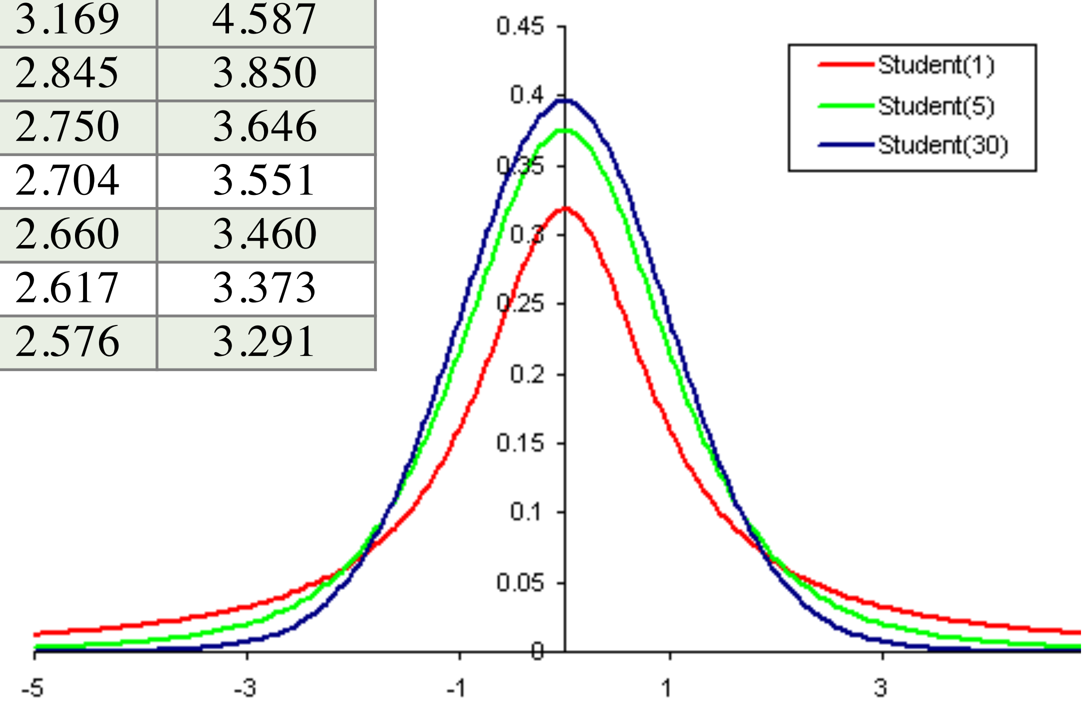
Sample variance

$$S^2 \equiv \frac{1}{N-1}\sum_{i=1}^{N} e_i^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(x_i - \bar{x}\right)^2$$

Student's *t*-test

SIGNIFICANCE LEVEL FOR TWO-TAILED TEST						
df	.20	.10	.05	.02	.01	.001
1	3.078	6.314	12.706	31.821	63.657	636.619
2	1.886	2.920	4.303	6.965	9.925	31.598
3	1.638	2.353	3.182	4.541	5.841	12.941
4	1.533	2.132	2.776	3.747	4.604	8.610
5	1.476	2.015	2.571	3.365	4.032	6.859
10	1.372	1.812	2.228	2.764	3.169	4.587
20	1.325	1.725	2.086	2.528	2.845	3.850
30	1.310	1.697	2.042	2.457	2.750	3.646
40	1.303	1.684	2.021	2.423	2.704	3.551
60	1.296	1.671	2.000	2.390	2.660	3.460
120	1.289	1.658	1.980	2.358	2.617	3.373
∞	1.282	1.645	1.960	2.326	2.576	3.291

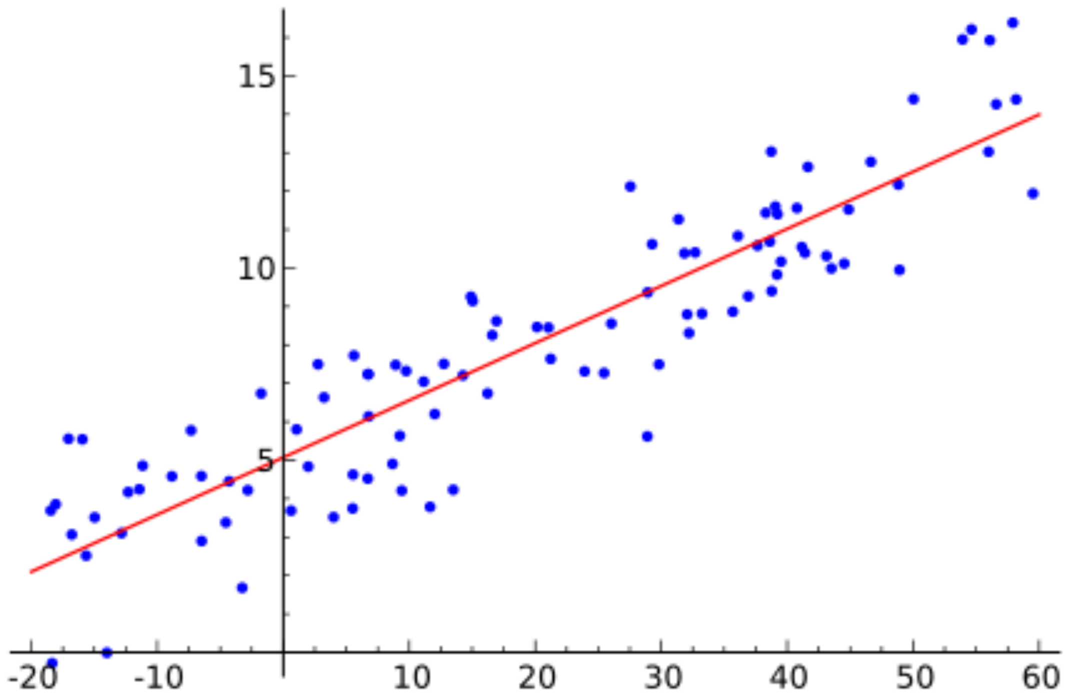
confidence level = 1 - significance level



Linear Regression

How about fitting lines to sets of data? Data points are (x_i,y_i)

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$$



Linear Regression

How about fitting lines to sets of data? Data points are (x_i, y_i)

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$$

Use **linear least squares** fit. Minimizes sum of squared residuals (SSE)

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{N} (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{N} (x_i - \bar{x})^2}$$

How certain are we?

How good are the betas?

Root mean squared residual

$$S_e = \sqrt{\frac{SSE}{N-2}} = \sqrt{\frac{\sum_{i=1}^{N} e_i^2}{N-2}}$$

$$df = N - 2$$

$$S_{\beta_0} = S_e \sqrt{\frac{1}{N} + \frac{\overline{x}^2}{\sum_{i=1}^{N}(x_i - \overline{x})^2}}$$

$$\lambda_{\beta_0} = tS_{\beta_0}$$

$$S_{\beta_1} = S_e \sqrt{\frac{1}{\sum_{i=1}^{N}(x_i - \overline{x})^2}}$$

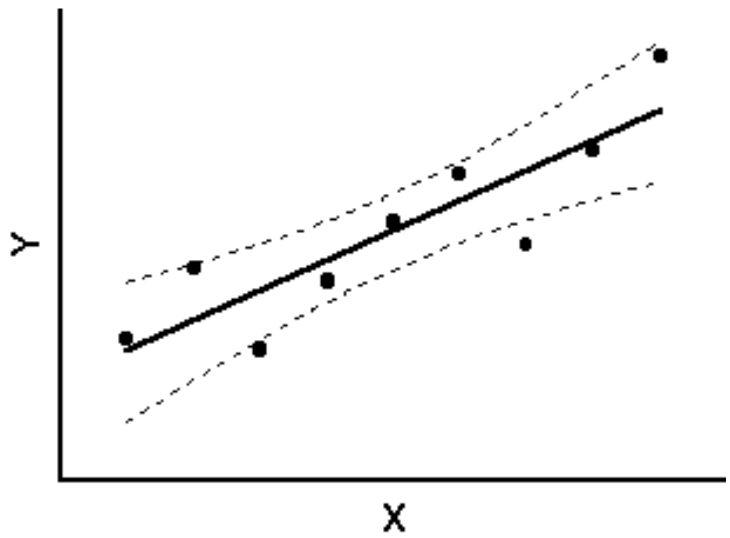
$$\lambda_{\beta_1} = tS_{\beta_1}$$

Usefulness!

- How certain is the calculated value of *y* for a chosen *x*?

$$S_y = S_e \sqrt{\frac{1}{N} + \frac{(x - \bar{x})^2}{\sum_{i=1}^{N} (x_i - \bar{x})^2}}$$

$$\lambda_y = t S_y$$



- If we experimentally set *x*, how will the experimental *y*'s spread? Use S_e like you previously used *S*.

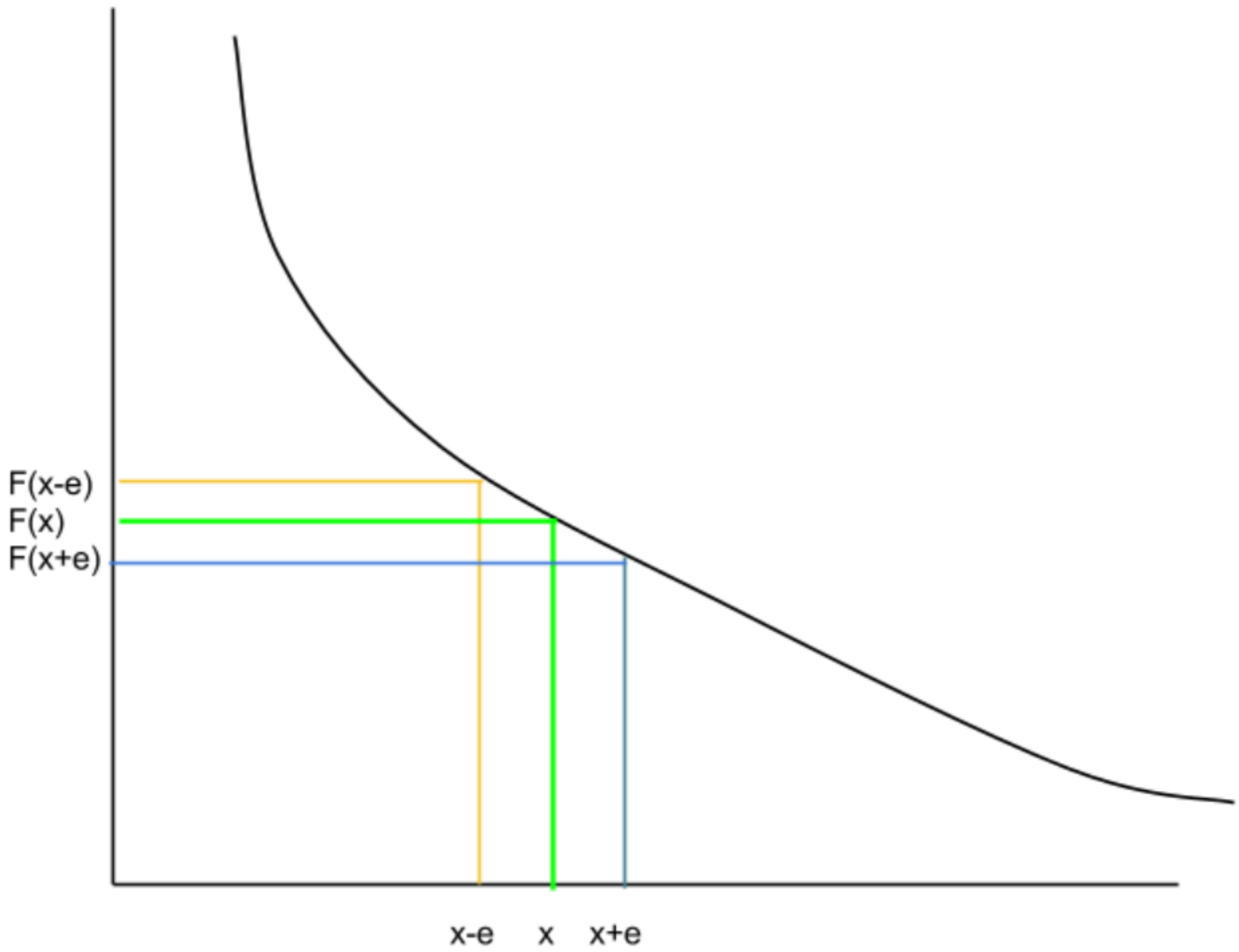
What about functions?

What if you want *T*

$$T = \frac{1}{\frac{1}{T_0} + \frac{1}{\beta}\ln\frac{R}{R_0}}$$

and you measure/look up β and *R*?

How error in *x* affects *F*(*x*)



Error Propagation Method 1

For $$F = F(x, y, z, \cdots)$$

$$F - F_{true} = \frac{\partial F}{\partial x}\left(x - x_{true}\right) + \frac{\partial F}{\partial y}\left(y - y_{true}\right) + \frac{\partial F}{\partial z}\left(z - z_{true}\right) + \cdots$$

Let $$\varepsilon_x = x - x_{true}$$

$$\varepsilon_F = \frac{\partial F}{\partial x}\varepsilon_x + \frac{\partial F}{\partial y}\varepsilon_y + \frac{\partial F}{\partial z}\varepsilon_z + \cdots$$

$$\varepsilon_F = \sqrt{\left(\frac{\partial F}{\partial x}\right)^2 \varepsilon_x^2 + \left(\frac{\partial F}{\partial y}\right)^2 \varepsilon_y^2 + \left(\frac{\partial F}{\partial z}\right)^2 \varepsilon_z^2 + \cdots}$$

Thermistor Example

$$T = \frac{1}{\frac{1}{T_0} + \frac{1}{\beta}\ln\frac{R}{R_0}}$$

$$dT = \frac{\left[\left(\frac{1}{\beta^2}\ln\frac{R}{R_0}\right)d\beta - \frac{1}{\beta R}dR\right]}{\left[\frac{1}{T_0} + \frac{1}{\beta}\ln\frac{R}{R_0}\right]^2}$$

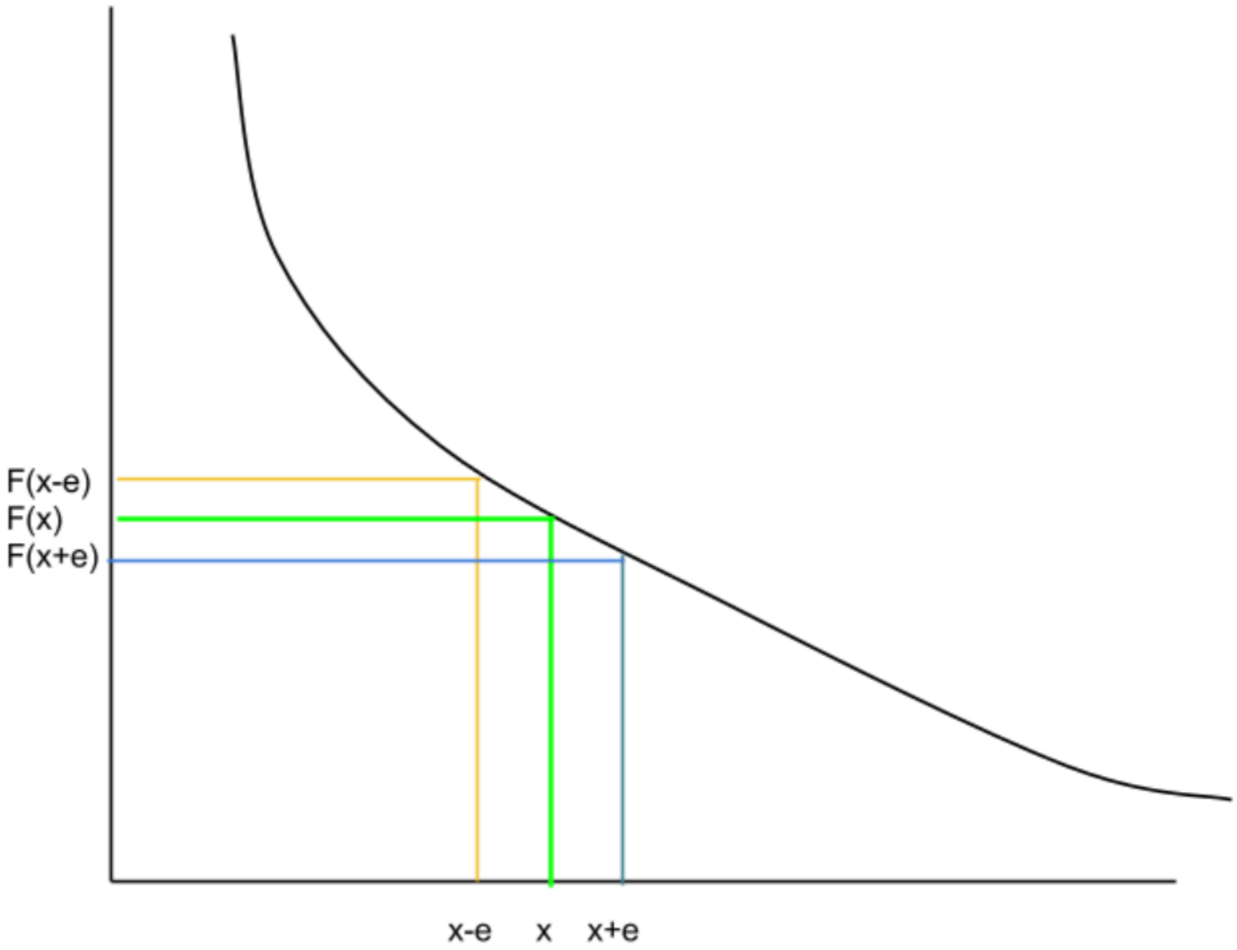
$$dT = T^2\left[\left(\frac{1}{\beta^2}\ln\frac{R}{R_0}\right)d\beta - \frac{1}{\beta R}dR\right]$$

Results

$$e_T = T^2 \left[\left(\frac{1}{\beta^2} \ln \frac{R}{R_0} \right)^2 e_\beta{}^2 + \left(\frac{1}{\beta R} \right)^2 e_R{}^2 \right]^{1/2}$$

	Nom. Value	Error %	Error	Error term
β	4261	1%	42.61	**0.23**
R (Ω)	3,700,000	10%	370000	**1.75**
T (K)	273.14			**1.77**
T_0 (K)	298.15			
R_0 (Ω)	1000000			

How error in *x* affects *F*(*x*)



Error Propagation Method 2

You can do the whole thing on a spreadsheet with no calculus.

$$T = \frac{1}{\frac{1}{T_0} + \frac{1}{\beta}\ln\frac{R}{R_0}}$$

Value	**Nominal**	**β + 1%**	**β – 1%**	**R + 10%**	**R – 10%**
T_0 (K)	298.15	298.15	298.15	298.15	298.15
R_0 (Ω)	1,000,000	1,000,000	1,000,000	1,000,000	1,000,000
β	4261	**4303.61**	**4218.39**	4261	4261
R (Ω)	3,700,000	3,700,000	3,700,000	**4,070,000**	**3,330,000**
T (K)	273.14	**273.37**	**272.91**	**271.49**	**275.00**
ΔT (K)		**0.46**		**-3.52**	
error (K)	**3.55**				
±error (K)	**1.77**				

Steinhart-Hart Example

$$T = \frac{1}{A_1 + B_1 \ln\frac{R}{R_{ref}} + C_1\left(\ln\frac{R}{R_{ref}}\right)^2 + D_1\left(\ln\frac{R}{R_{ref}}\right)^3}$$

$$dT = \frac{-\left[dA_1 + (\ln R)dB_1 + (\ln R)^2 dC_1 + (\ln R)^3 dD_1 + \left\{\frac{B_1 + 2C_1\left(\ln\frac{R}{R_{ref}}\right) + 3D_1\left(\ln\frac{R}{R_{ref}}\right)^2}{R}\right\}dR\right]}{\left[A_1 + B_1 \ln\frac{R}{R_{ref}} + C_1\left(\ln\frac{R}{R_{ref}}\right)^2 + D_1\left(\ln\frac{R}{R_{ref}}\right)^3\right]^2}$$

$$dT = -T^2\left[dA_1 + (\ln R)dB_1 + (\ln R)^2 dC_1 + (\ln R)^3 dD_1 + \left\{\frac{B_1 + 2C_1\left(\ln\frac{R}{R_{ref}}\right) + 3D_1\left(\ln\frac{R}{R_{ref}}\right)^2}{R}\right\}dR\right]$$

$$e_T = T^2\left[\left[e_{A_1}\right]^2 + \left[\left(\ln\frac{R}{R_{ref}}\right)e_{B_1}\right]^2 + \left[\left(\ln\frac{R}{R_{ref}}\right)^2 e_{C_1}\right]^2 + \left[\left(\ln\frac{R}{R_{ref}}\right)^3 e_{D_1}\right]^2 + \left[\left\{\frac{B_1 + 2C_1(\ln R) + 3D_1(\ln R)^2}{R}\right\}e_R\right]^2\right]^{1/2}$$

	Nom. Value	error%	error	error term
A_1	3.354016E-03	0.1%	3.35402E-06	**0.318**
B_1	2.460382E-04	1%	2.46038E-06	**-0.104**
C_1	3.405377E-06	1%	3.40538E-08	**0.001**
D_1	1.034240E-07	1%	1.03424E-09	**0.000**
R (Ω)	64080	3%	1922.4	**0.692**
T (K)	308.15			**0.769**
R_{ref}	100,000			

Calculus free style

$$T = \frac{1}{A_1 + B_1 \ln\frac{R}{R_{ref}} + C_1\left(\ln\frac{R}{R_{ref}}\right)^2 + D_1\left(\ln\frac{R}{R_{ref}}\right)^3}$$

Value	Nominal	A_1 +0.1%	A_1 –0.1%	B_1 +1%	B_1 –1%	C_1 +1%	C1 –1%	D_1 +1%	D_1 –1%	R +3%	R –%
A_1	3.354E-03	3.357E-03	3.351E-03	3.354E-03	3.354E-03	3.354E-03	3.354E-03	3.354E-03	3.354E-03	3.354E-03	3.354E-03
B_1	2.460E-04	2.460E-04	2.460E-04	2.485E-04	2.436E-04	2.460E-04	2.460E-04	2.460E-04	2.460E-04	2.460E-04	2.460E-04
C_1	3.405E-06	3.405E-06	3.405E-06	3.405E-06	3.405E-06	3.439E-06	3.371E-06	3.405E-06	3.405E-06	3.405E-06	3.405E-06
D_1	1.034E-07	1.034E-07	1.034E-07	1.034E-07	1.034E-07	1.034E-07	1.034E-07	1.045E-07	1.024E-07	1.034E-07	1.034E-07
R_{ref}	100,000	100,000	100,000	100,000	100,000	100,000	100,000	100,000	100,000	100,000	100,000
R	64,080	64,080	64,080	64,080	64,080	64,080	64,080	64,080	64,080	66,002	62,158
T_{calc}	308.15	307.83	308.47	308.25	308.04	308.15	308.15	308.15	308.15	307.47	308.85
ΔT		-0.64		0.21		0.00		0.00		-1.39	
error	**1.539**										
error ±	**0.769**										

Effect of Finite Precision

Calculate or determine the quantization range, *q*, e.g., for a ±5V, 12-bit DAQ:

$$q = 10\text{V}\frac{1}{2^{12}} = \frac{10\text{V}}{4096} = 0.027\text{V}$$

If $S > \frac{10q}{\sqrt{12}}$ ignore quantization

If $\frac{q}{\sqrt{12}} < S < \frac{10q}{\sqrt{12}}$ include as $S_{used} = \sqrt{S^2 + \frac{q^2}{12}}$

If $\pm q/2$ confidence interval is $S < \frac{q}{\sqrt{12}}$

Resources

Data Analysis Lecture Notes

- http://www.eng.hmc.edu/NewE80/PDFs/DataAnalysisLecNotes2016.pdf

NIST Engineering Statistics Handbook

- http://www.itl.nist.gov/div898/handbook/

ISO Guide to the Expression of Uncertainty in Measurement

- http://www.iso.org/sites/JCGM/GUM/JCGM100/C045315e-html/C045315e.html?csnumber=50461

We DO expect you to do this!