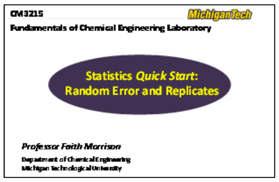
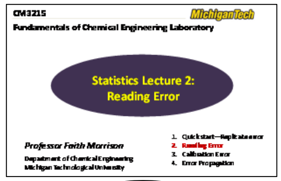
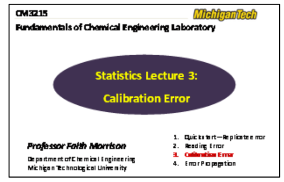
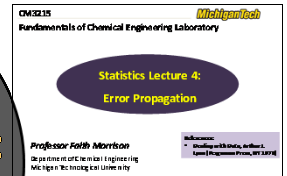
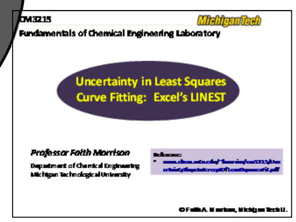






Engineering Error Analysis: 5 Practical Lessons





Professor Faith Morrison

Department of Chemical Engineering
Michigan Technological University

CM3215

Fundamentals of Chemical Engineering Laboratory

Statistics *Quick Start*: Random Error and Replicates

Professor Faith Morrison

Department of Chemical Engineering
Michigan Technological University

1. **Quick start—Replicate error**
2. Reading Error
3. Calibration Error
4. Error Propagation
5. Least Squares Curve Fitting

Measurements are affected by errors

(uncertainty)

There are two general categories of errors (uncertainties) in experimental measurements:

- Systematic errors
- Random errors

3

Measurements are affected by errors

(uncertainty)

Systematic errors

1. Has same sign and magnitude for identical conditions
2. Must be checked for, identified, eliminated, randomized

Sources:

- Calibration of instruments
- Reading error (resolution, coarse scale)
- Consistent operator error
- Failure to produce experimental conditions assumed in an analysis (e.g. steady state, isothermal, well mixed, pure component, etc.)

Solutions:

- Recalibrate
- Improve instrument resolution
- Apply correction for identified error
- Improve procedures, experimental design
- Shift to other methods
- Take data in random order; rotate operators

4

Measurements are affected by errors

(uncertainty)

Random errors

1. Varies in sign and magnitude for identical conditions
2. May be due to the instrument or the process being measured
3. Must be understood and communicated with results

Sources:

- Random process, instrument fluctuations
- Randomized systematic trends (e.g. operator identity, thermal drift)
- Rare events

Solutions:

- Replicate and average
- Improve measurement methods, practices
- Isolate from rare events

5

Measurements are affected by errors

(uncertainty)

- We never stop looking for and fixing random and systematic errors in real experimental data.
- We use **statistical methods** to *measure*, *reduce*, and *communicate* the random errors that we cannot eliminate.

6

Obtaining a Good Estimate of a Quantity

- Measure the quantity several times – replicates
- The average value is a good estimate of the quantity we are measuring

$x_1, x_2, x_3, x_4, x_5 \ldots x_n$ Replicates

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$ Sample mean

$\bar{x}$ is a Good* Estimate of x

**small print*: This is true if we take enough replicates and if only random errors are present; see sources on statistics

7

Measurements are affected by errors (uncertainty)

$\bar{x}$ is a Good Estimate of x

Now, How precisely do we know x?

Sample mean

Answer: (we discuss where this comes from later)

$E(x) = \bar{x} \pm 2e_s$ with 95% confidence

Expected value of x

Standard Error

Use "2" for $n \geq 7$)

The expected value of x sits centered within an interval of twice the standard error of x

Three sources of standard error:
- **Replicate errors**
- **Reading errors**
- **Calibration errors**

8

Measurements are affected by errors (uncertainty)

$\bar{x}$ is a Good Estimate of x

Sample mean

Now, How precisely do we know x?

Answer: (we discuss more later)

$$E(x) = \bar{x} \pm 2e_s \text{ with 95\% confidence}$$

Expected value of x

Standard Error

Use "2" for $n \geq 7$)

The expected value of x sits centered within an interval of twice the standard error of x

Three sources of standard error:

- **Replicate errors**
- ~~**Reading errors**~~
- ~~**Calibration errors**~~

For this week's *Quick Start*, we consider only random errors

9

One source of standard error…

Obtaining a Good Estimate of **Replicate Errors**

- Measure the quantity several times – i.e. take *replicates*
- The average value is a good estimate of the quantity

$x_1, x_2, x_3, x_4, x_5 \dots x_n$ Replicates

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$ Sample mean

$$s^2 = \frac{1}{n-1}\sum_{i=1}(x_i - \bar{x})^2$$ Sample variance

s Sample standard deviation

$$e_s = \frac{s}{\sqrt{n}}$$ *Standard error of the mean of replicates

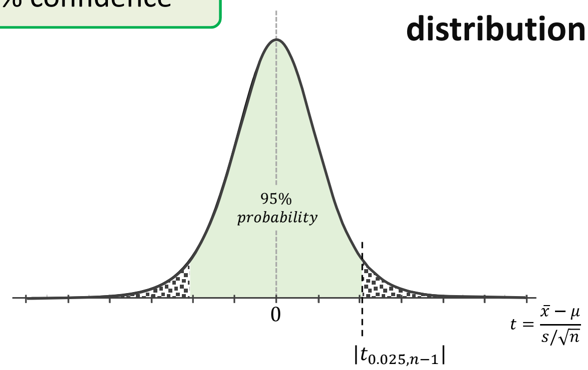
10

*see sources on statistics

$E(x) = \bar{x} \pm 2e_s$ with 95% confidence

Student's t distribution

Where does the "2" come from?



95% Confidence Interval (α=0.05) for n replicates		Student's T Distribution =T.INV.2T(0.05,n-1)											Standard Normal Dist
	n=	2	3	4	5	6	7	8	9	10	20	50	
	ν=n-1	1	2	3	4	5	6	7	8	9	19	49	(infinity)
α/2=0.025	$\lvert t_{0.025,n-1}\rvert$=	12.71	4.30	3.18	2.78	2.57	2.45	2.36	2.31	2.26	2.09	2.01	1.96
1 sig fig:	$\lvert t_{0.025,n-1}\rvert$=	13	4	3	3	3	**2**	**2**	**2**	**2**	**2**	**2**	**2**

www.chem.mtu.edu/~fmorriso/cm3215/ConfidenceIntervals.pdf

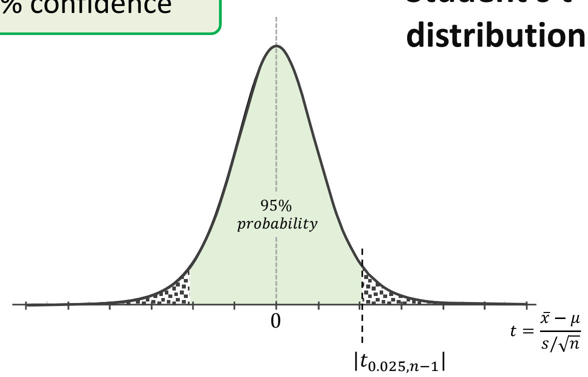
11

$E(x) = \bar{x} \pm 2e_s$ with 95% confidence

Student's t distribution

Where does the "2" come from?

95% of the probability resides within about $\pm 2e_s$ of the expected value



95% Confidence Interval (α=0.05) for n replicates		Student's T Distribution =T.INV.2T(0.05,n-1)											Standard Normal Dist
	n=	2	3	4	5	6	7	8	9	10	20	50	
	ν=n-1	1	2	3	4	5	6	7	8	9	19	49	(infinity)
α/2=0.025	$\lvert t_{0.025,n-1}\rvert$=	12.71	4.30	3.18	2.78	2.57	2.45	2.36	2.31	2.26	2.09	2.01	1.96
1 sig fig:	$\lvert t_{0.025,n-1}\rvert$=	13	4	3	3	3	**2**	**2**	**2**	**2**	**2**	**2**	**2**

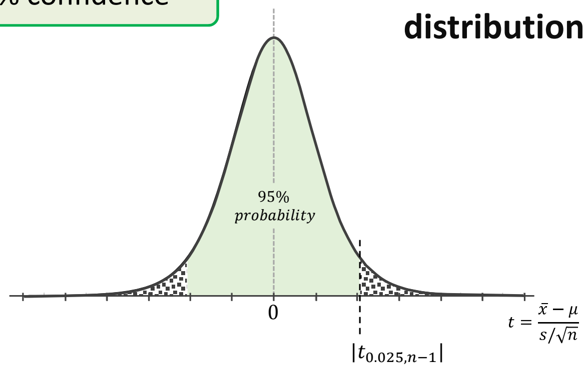
www.chem.mtu.edu/~fmorriso/cm3215/ConfidenceIntervals.pdf

12

$E(x) = \bar{x} \pm 2e_s$ with 95% confidence

Student's t distribution

Where does the "2" come from?



- Use "~2" for $n \geq 7$)
- Use this table or Excel for all other n

95% Confidence Interval (α=0.05) for n replicates		Student's T Distribution =T.INV.2T(0.05,n-1)											Standard Normal Dist
	n=	2	3	4	5	6	7	8	9	10	20	50	
	ν=n-1	1	2	3	4	5	6	7	8	9	19	49	(infinity)
α/2=0.025	$\lvert t_{0.025,n-1}\rvert$=	12.71	4.30	3.18	2.78	2.57	2.45	2.36	2.31	2.26	2.09	2.01	1.96
1 sig fig:	$\lvert t_{0.025,n-1}\rvert$=	13	4	3	3	3	**2**	**2**	**2**	**2**	**2**	**2**	**2**

www.chem.mtu.edu/~fmorriso/cm3215/ConfidenceIntervals.pdf

EXAMPLE 1.1 Based on the following seven data points we have on Blue Fluid 175 density, what is the density ρ and the 95% confidence interval based on replicate error?

i	X_i
1	1.7348
2	1.7465
3	1.7359
4	1.83
5	1.74688
6	1.74412
7	1.73173

You try.

EXAMPLE 1.1 Based on the following seven data points we have on Blue Fluid 175 density, what is the density and the 95% confidence interval based on replicate error?

i	X_i
1	1.7348
2	1.7465
3	1.7359
4	1.83
5	1.74688
6	1.74412
7	1.73173

	ρ	
1	1.7348	g/cm^3
2	1.7465	g/cm^3
3	1.7359	g/cm^3
4	1.83	g/cm^3
5	1.74688	g/cm^3
6	1.74412	g/cm^3
7	1.73173	g/cm^3
mean	1.752847	g/cm^3
variance	0.001194	$(g/cm^3)^2$
std dev	0.034553	g/cm^3
std err	0.013060	g/cm^3

Excal:
AVERAGE()
VAR.S()
STDEV.S()

What is the answer for $\rho = ? \pm ?$ with 95% confidence (replicate)?

EXAMPLE 1.1 Based on the following seven data points we have on Blue Fluid 175 density, what is the density and the 95% confidence interval based on replicate error?

Answer:

	ρ	
1	1.7348	g/cm^3
2	1.7465	g/cm^3
3	1.7359	g/cm^3
4	1.83	g/cm^3
5	1.74688	g/cm^3
6	1.74412	g/cm^3
7	1.73173	g/cm^3
mean	1.752847	g/cm^3
variance	0.001194	$(g/cm^3)^2$
std dev	0.034553	g/cm^3
std err	0.013060	g/cm^3

Excel:
AVERAGE()
VAR.S()
STDEV.S()

EXAMPLE 1.1 Based on the following seven data points we have on Blue Fluid 175 density, what is the density and the 95% confidence interval based on replicate error?

Answer:
$\rho = 1.75 \pm 0.03 g/cm^3$ (95%CI)

	ρ	
1	1.7348	g/cm^3
2	1.7465	g/cm^3
3	1.7359	g/cm^3
4	1.83	g/cm^3
5	1.74688	g/cm^3
6	1.74412	g/cm^3
7	1.73173	g/cm^3
mean	1.752847	g/cm^3
variance	0.001194	$(g/cm^3)^2$
std dev	0.034553	g/cm^3
std err	0.013060	g/cm^3

Excal:
AVERAGE()
VAR.S()
STDEV.S()

17

Significant Figures on Error

Common rules:

- Usually use one significant figure on error

$$\rho_A = 1.722 \pm 0.005 g/cm^3$$

- If the digit is 1 or 2, you may include two digits (to avoid round-off error)

$$\rho_B = 1.9431 \pm 0.0015 g/cm^3$$

Note: do not truncate numbers used in intermediate calculations.

18

Significant Figures on Error

Common rules:

- Usually use one significant figure on error

$$\rho_A = 1.722 \pm 0.005 g/cm^3$$

Once you know the uncertainty, truncate the expected value appropriately:

- If the digit is 1 or 2, you may include two digits (to avoid round-off error)

$$\rho_B = 1.9431 \pm 0.0015 g/cm^3$$

Note: do not truncate numbers used in intermediate calculations.

19

EXAMPLE 1.1 Based on the following seven data points we have on Blue Fluid 175 density, what is the density and the 95% confidence interval based on replicate error?

Answer:
$\rho = 1.75 \pm 0.03 g/cm^3$ (95%CI)

	ρ	
1	1.7348	g/cm^3
2	1.7465	g/cm^3
3	1.7359	g/cm^3
4	1.83	g/cm^3
5	1.74688	g/cm^3
6	1.74412	g/cm^3
7	1.73173	g/cm^3
mean	1.752847	g/cm^3
variance	0.001194	$(g/cm^3)^2$
std dev	0.034553	g/cm^3
std err	0.013060	g/cm^3

With 95% confidence, using $\pm 2.45 e_s$ $(N = 7)$ or $\pm 2e_s$

Excel:
AVERAGE()
VAR.S()
STDEV.S()

20

Replicate Error Worksheet
CM3215 Fundamentals of Chemical Engineering Lab
Prof. Faith Morrison

MichiganTech
Michigan Technological University
Department of Chemical Engineering

This worksheet guides the user through the calculation of the standard error and 95% confidence interval on a quantity that has been measured n times (replicated). The replicate-error-related standard error e_s may subsequently be used in propagation-of-error calculations of derived quantities.

Replicated Variable, Y:						Units:		
Measured values $Y_1, Y_2, \ldots, Y_n$		Sample Mean, $\bar{Y}$	Sample Variance, s^2	Sample Standard Deviation, s	Standard Error, $e_s = \frac{s}{\sqrt{n}}$	95% Confidence Interval based on n replicates (Student's t distribution)		(include units)
Y_1						$n=1$	n/a	$\pm$
Y_2						$n=2$	$\pm 12.7e_s$	
Y_3						$n=3$	$\pm 4.30e_s$	
Y_4						$n=4$	$\pm 3.18e_s$	
Y_5						$n=5$	$\pm 2.78e_s$	
Y_6						$n=6$	$\pm 2.57e_s$	
Y_7						$n\geq 7$	$\pm 2e_s$	
						∞	$\pm 1.96e_s$	

$$\bar{Y} \equiv \frac{1}{n}\sum_{i=1}^{n} Y_i \qquad s^2 \equiv \frac{1}{(n-1)^2}\sum_{i=1}^{n}(Y_i - \bar{Y})^2$$

11-Sep-14

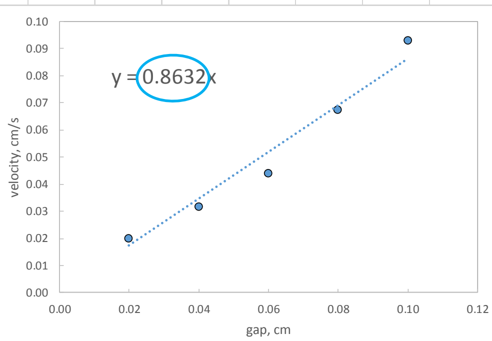
21

www.chem.mtu.edu/~fmorriso/cm3215/ReplicateErrorWorksheet.pdf

If we obtain a parameter from an equation fit, how can we obtain the 95% Confidence Interval? Excel's LINEST

$$\text{Shear rate } \dot{\gamma}(s^{-1}) = 0.8632\ s^{-1} \pm \ ?$$

gap	gap	velocity
mm	cm	cm/s
0.2	0.02	0.0200
0.4	0.04	0.0317
0.6	0.06	0.0440
0.8	0.08	0.0674
1.0	0.10	0.0929



The e_s in this case comes from an error propagation calculation. We'll talk about this soon.

Instructions to do this:
www.chem.mtu.edu/~fmorriso/cm3215/UncertaintySlopeInterceptOfLeastSquaresFit.pdf

22

Measurements are affected by errors

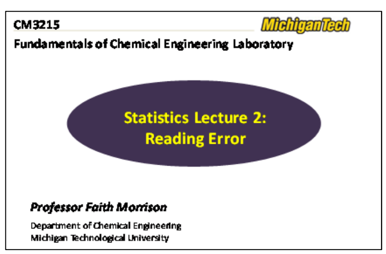
(uncertainty)

Summary:

- Taking replicate measurements is a good way to estimate the value of a quantity affected by random errors
- $\bar{x}$ is a Good Estimate of x
- $E(x) = \bar{x} \pm t_{0.025,n-1} e_s$ with 95% confidence (see table)
- The standard error e_s is obtained from considering:
  1. Replicate errors ($e_s = s/\sqrt{n}$)
  2. Reading errors ($e_s = ?$)
  3. Calibration errors ($e_s = ?$)
- In this *Quick Start* section, we have only considered replicate errors; we consider the reading and calibration errors in subsequent lectures and activities
- Use one significant figure on error limits (unless the digit is 1 or 2)
- When parameters are obtained from a fit, use error propagation to calculate e_s (Excel's LINEST)

CM3215 Fundamentals of Chemical Engineering Laboratory

Next, on to Statistics 2: Reading Error



24