Plotting/Statistics Assignment 2
CM3215 Chem Eng Fundamentals Lab

Prof. Faith A. Morrison

Michigan Technological University
Department of Chemical Engineering

Learning Objectives of the Exercise *modified 21Jan2016*

Experimental data always have some degree of uncertainty associated with them. We use *statistics* and *significant figures* to appropriately express our confidence in the precision and/or accuracy of our experimental data. In this exercise, we use CM3215 data from a past semester to practice making these error analysis/statistics calculations.

We have the following objectives for this assignment:

- Learn to make replicate error calculations with Excel
- Learn how to determine the *95% confidence interval (CI) of the mean* based on the Student's t distribution for data replicates.
- Learn how to determine the *95% prediction interval (PI) of the next value of x*.
- Learn what the difference is between the 95% CI and the 95% PI.

The Exercise

Instructions: This is an individual assignment. Submit your answers to the following questions by ***9:05am*** ***2:05pm*** ***Friday 22 January 2016*** *in Homework Box A. Please attach a memo of transmittal to your submission. Note that w*e only report one significant figure on final answers of error (or two sig figs if the digit is a "1" or a "2").

1. For the data in Table 1, what is the best estimate of the density of sugar solutions as a function of concentration? Display your results in both tabular form and in a graph. Excel sorting commands are helpful in organizing the data.
2. What is the 95% confidence interval of the mean at each concentration, assuming only random error is present? Display your results in both tabular form and as error bars in a graph.
3. Based on the current data, if new values of density are measured at each concentration, what is the interval within which you expect the next values of density to fall? This is called the 95% prediction interval of the next value of x. ~~Display your results on the graph.~~[1]
4. Using LINEST in Excel, what are the slope and intercept of an ordinary least squares fit of a line to the $\rho(c)$ data given? Report both slope and intercept with their 95% confidence intervals. You will want to use the handout on the web at www.chem.mtu.edu/~fmorriso/cm3215/UncertaintySlopeInterceptOfLeastSquaresFit.pdf .
5. Please answer this question in your submission: What does the 95% confidence intervals mean for the current data? What do the 95% PI mean for the current data?

Note that there are instructions on the web on how to put error bars on graphs in Excel: www.chem.mtu.edu/~fmorriso/cm3215/2014WordFigureErrorBars95CI.pdf

[1] This is called the 95% prediction interval. We did not cover this yet, but it's quite straightforward. The 95% prediction interval is $\bar{x} \pm t_{0.025,n-1}s$.

Table 1: Student data on the density of aqueous sugar solutions of various concentrations; taken Fall 2013 at room temperature.

data point	wt%	ρ g/cm^3	data point	wt%	ρ g/cm^3	data point	wt%	ρ g/cm^3
1	10	1.02	13	45	1.19	25	60	1.27
2	60	1.28	14	50	1.24	26	20	1.08
3	40	1.17	15	20	1.14	27	30	1.11
4	30	1.12	16	40	1.22	28	65	1.3
5	65	1.31	17	30	1.1144	29	45	1.202
6	45	1.21	18	65	1.2894	30	50	1.246
7	50	1.23	19	40	1.23	31	20	1.07
8	10	1.03	20	30	1.11	32	40	1.2
9	60	1.28	21	65	1.3	33	30	1.12
10	20	1.1	22	45	1.19	34	50	1.23
11	30	1.1	23	50	1.26	35	10	1.03
12	65	1.28	24	10	1.02	36	30	1.12
						37	50	1.23
						38	10	1.03

Table 2: 95% confidence intervals for the Student's t distribution, calculated from Excel; n is the number of replicates taken and ν is the number of degrees of freedom ($\nu = n - 1$).

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