CHEM 454L – Excel Exercise 1

Replicate Measurements & Standard Deviation

1. Open an Excel spreadsheet.

2. Point with the mouse pointer to the tab labeled Sheet, and rename it Average.

3. Deposit the label 'n' in cell Al, and the label 'data' in B1.

4. In cell A3 deposit the number 1, and in cell A4 the number 2.

5. Place the pointer in cell A3 (it should be a heavy cross), depress the mouse button, move the pointer to cell A4, then release the button. Both cells (A3 + A4) should now be activated.

6. Grab the common handle of cells A3 + A4 (when the mouse pointer has the shape of a plus sign) and drag the cells by this handle down to cell 302. This will establish *N*- values from 1 to 300 in column A. Or: in cell A4 use the instruction = A3 +1, and copy this down to cell A302.

7. Click on Tools, then on Data Analysis, and in the Data Analysis dialog box select Random Number Generation.

8. In the Random Number Generation dialog box that now appears, select 'Normal' as the Distribution, 10 as the Mean, and 1 as the Standard Deviation. Furthermore, specify the Output Range as B3:B302. Click OK.

9. Select block A3 :B302, e.g., by first using the mouse pointer to activate a small block such as A3:B6, and by then, while keeping the Shift depressed, keying in End followed by $\downarrow$.

10. Select Insert $\rightarrow$ Chart, and use the ChartWizard to select an XY plot, showing individual data points without a connecting line. **Complete the graph**.

11. We notice that the data all cluster around aY-value of about 10, but that individual points can lie quite a bit farther from that average value: occasionally a point will lay more than 2 or 3 standard deviations from the average. That shows the true nature of such a distribution; if we consider a sufficiently large number of such data, we will find that about 68% of them lie within one standard deviation from the mean, but that the remainder, about one-third of all points, lie further away.

12. Click on the numbers with the X- or Y-axis, right-click, choose Format Axis, and select the Font and Scale to your liking. You might consider using 16 point regular Times New Roman, and restricting the Y-scale to the range from 7 to 13, but please make your own decisions. You may also want to delete the series marker (which usually appears in a separate box to the right of the graph) by clicking on it to highlight it, and by then using the Delete key to remove it.

13. Activate cell C4, and deposit in it the instruction = AVERAGE(B3:B5) which is equivalent to the instruction = (B3 + B4 + B5) /3. Verify in an empty cell that the instruction indeed calculates the average, then erase your verification lest it will show as an odd point in one of the graphs you will make.

14. Activate cell D4, and make it carry the instruction = STDEV(B3:B5), which calculates the standard deviation. Again verify that STDEV indeed calculates correctly, then erase that test.

15. Highlight the area C3:D5, grab its common handle, and pull that handle all the way down to cell D302. Column C should now have 100 data, each the average of three successive data points in column B, while column D will now contain the corresponding standard deviations.

16. Take a look at the averages and the standard deviations: they will show considerable fluctuations, even though all numbers were calculated from Gaussian noise with a constant standard deviation.

17. The easiest way to 'take a look' at these averages and their standard deviations is, of course, to **plot them**. Activate cell A3, then drag the pointer to A302, depress the Ctrl key, move the mouse pointer over to cell C302, and drag the pointer to cell C3. This will highlight the data in columns A and C you want to graph.

18. Click on Insert → Chart, and answer the ChartWizard to make the graph showing the three-point averages. **Plot it**.

19. Click on the chart, then on one of the data points in it, then right-click, and select Format Data Series. In the resulting dialog box go for the Y Error Bars tab, under Display select Both, and push the radio button for Custom. Then deposit in the two boxes labeled + and - the identical instruction: =AVERAGE!D3:D302, and use the OK button to enter these instructions. You should now obtain a graph in which all three-point averages are specifically labeled with their corresponding standard deviations.

20. Activate cell *E8* and let it calculate the average of the 10 consecutive data in B3:B12.

21. Similarly, in cell F8, compute the corresponding standard deviation.

22. Highlight block E3:F12, grab it by its handle, and copy it down to F302. You should now have 30 averages of 10 points each, with their standard deviations.

23. Use columns A and E to make a **graph**, on a new sheet.

24. Add error bars to the averages plotted, using as before the Format Data Series dialog box, where you select the Y Error Bars tab, then specify the Error Amount under Custom as = AVERAGE!F3:F302 in both directions.

25. In column G calculate the averages of 30 consecutive data points and in column H the corresponding standard deviations. **Plot** the resulting thirty-point averages, with their individual error bars.

26. Finally, calculate the average and standard deviation of all 300 data points.

27. Determine the smallest value of the standard deviations in the three-point samples with the instruction = MIN(D3:D302), and for the corresponding largest value with = MAX(D3:D302). Likewise find the extreme values for the standard deviations of the ten-point samples in column F, and for the thirty-point samples in column H.

28. Record these values, then save the spreadsheet, and close it.

29. **Turn in all spreadsheets and graphs on Monday April 4 at the start of class.**

Note: Exercise adapted from: How to Use Excel in Analytical Chemistry and in General Scientific Data Analysis by Robert de Levie, Cambridge University Press.