Review Problems in Error Analysis and Statistics

1. Estimate the absolute standard deviation and the coefficient of variation for the results of each of the following calculations. Round each result so that it contains only significant figures. (The numbers in brackets are the absolute standard deviations).

a) $6.75(\pm0.03) + 0.843(\pm0.001) - 7.021(\pm0.001) = 0.572 \pm \left[(0.03)^2 + (0.001)^2 + (0.001)^2\right]^{1/2}$

$= 0.572 \pm 0.030$

$= 0.57_2 \pm 0.03_0$

$CV = RSD = \dfrac{0.030}{0.572} \times 100\% = 5.2\%$

b) mass of analyte (mg) $= \dfrac{W\,(mg) \times V\,(mL)}{V_t(mL)}$ $= 20.00 \pm \left[(0.008)^2 + (0.12)^2 + (0.20)^2\right]^{1/2}$

where $W = 500.0 \pm 0.04$ mg $\longrightarrow 0.008\%$

$V_t = 250.0 \pm 0.3$ mL $\longrightarrow 0.12\%$

$V = 10.00 \pm 0.02$ mL $\longrightarrow 0.20\%$

$= 20.00(\pm 0.23\%)$

$= 20.00 \pm 0.047$

$= 20.0_0 \pm 0.04_7$

$CV = RSD = \dfrac{0.047}{20.00} \times 100\% = 0.23\%$

c) $\dfrac{1.43(\pm0.02) \times 10^{-2} - 4.76(\pm0.06) \times 10^{-3}}{24.3(\pm0.7) + 8.06(\pm0.08)}$ $= \dfrac{9.54 \times 10^{-3} \pm \left[(0.0002)^2 + (0.00006)^2\right]^{1/2}}{32.36 \pm \left[(0.7)^2 + (0.08)^2\right]^{1/2}}$

$= \dfrac{9.54 \times 10^{-3} \pm (0.00021)}{32.36 \pm (0.7)}$

$= \dfrac{9.54 \times 10^{-3} \pm (2.2\%)}{32.36 \pm (2.2\%)}$

$= 2.948 \times 10^{-4} \pm \left[(2.2)^2 + (2.2)^2\right]^{1/2}$

$= 2.948 \times 10^{-4} (\pm 3.1\%)$

$= 2.94_8 (\pm 0.09_2) \times 10^{-4}$

$CV = RSD = 3.1\%$

2. Calculate a pooled estimate of the standard deviation from the following analysis for nitrilotriacetic acid, NTA.

sample #	N	[NTA] ppb	mean, x_N	s	$\Sigma \lvert x_i - x_N \rvert^2$
1	4	13, 16, 14, 9	13	2.94	26.0
2	3	38, 37, 38	37.7	0.577	0.67
3	5	25, 29, 23, 29, 26	26.4	2.61	27.2

$$S_{pooled} = \sqrt{\frac{\Sigma\Sigma_i |x_i - x_N|^2}{N_1 + N_2 + N_3 - \#subsets}} = \left[\frac{26.0 + 0.67 + 27.2}{4+3+5-3}\right]^{1/2} = 2.45$$

OR

$$S_{pooled} = \sqrt{\frac{\Sigma_i\, s_i^2 (N_i - 1)}{N_1 + N_2 + N_3 - \#subsets}} = \left[\frac{(2.94)^2(4) + (0.577)^2(2) + (2.61)^2(4)}{4+3+5-3}\right]^{1/2}$$

$$= 2.45$$

3. Calculate the 95% confidence limit for each of the following data sets. What do these limits mean?

A	B	C
2.4	69.94	0.0902
2.1	69.92	0.0884
2.1	69.80	0.0886
2.3		0.1000
1.5		

	A	B	C
mean	2.08	69.887	0.0918
s	0.35	0.076	0.0055

A	B	C
t @ df = 4, 2.78	t @ df = 2, 4.30	t @ df = 3, 3.18
$\frac{ts}{\sqrt{N}} = 0.44$	$\frac{ts}{\sqrt{N}} = 0.19$	$\frac{ts}{\sqrt{N}} = 0.0087$
CL = 2.1 ± 0.4	CL = 69.9 ± 0.2	CL = 0.092 ± 0.009

4. Apply the Q-test (95% CL) to determine whether or not there is a statistical basis for rejection of any outliers in the above data sets.

$Q = \frac{gap}{range}$ compare to Q_{crit}. If $Q_{calc} > Q_{crit}$ then reject

only value to be rejected from above data set is 0.1000 from group C (above) $Q_{calc} = \frac{|0.1000 - 0.0902|}{|0.1000 - 0.0884|} = 0.845$

Q_{crit} for N = 4 is 0.829 $\therefore Q_{calc} > Q_{crit}$

5. The results of 30 analysis to determine the iron content in jet fuel has lead to the following standard deviation, $\sigma = 2.4$ µg/mL.

a) Calculate the 95% confidence limits for the result 18.4 µg/mL if it is based on i) a single measurement and ii) the mean of four measurements.

i) $CL = 18.4 \pm \dfrac{1.96\,(2.4)}{\sqrt{1}} = 18.4 \pm 4.7$ or 18 ± 5 µg/mL

ii) $CL = 18.4 \pm \dfrac{1.96\,(2.4)}{\sqrt{4}} = 18.4 \pm 2.4$ or 18 ± 2 µg/mL

b) How many replicates are needed to decrease the 95% confidence interval to ±1.5 µg/mL?

We want $\dfrac{z\sigma}{\sqrt{N}} \leq 1.5 \quad \therefore \quad \sqrt{N} \geq \dfrac{z\sigma}{1.5} = \dfrac{1.96\,(2.4)}{1.5}$

$= 3.14$

$\therefore N = 9.83$

So $N = 10$ replicates

6. The following data are from the volumetric analysis of calcium ion in a water sample.

$[Ca^{2+}]$ mmol/L: 3.15, 3.25, 3.26

a) What is the 99% confidence limit for these results?

b) What is the 99% confidence limit, if $s \rightarrow \sigma = 0.056$?

mean $= 3.22$, $s = 0.061$

a) t @ $df = 2$ is 9.92

$\therefore [Ca^{2+}] = 3.22\left(\pm \dfrac{ts}{\sqrt{N}} = \dfrac{(9.92)(0.061)}{\sqrt{3}}\right)$

$= 3.2_2 \pm 0.3_6$ mM

b) $z = 2.58$

$\therefore [Ca^{2+}] = 3.22\left(\pm \dfrac{z\sigma}{\sqrt{N}} = \dfrac{(2.58)(0.056)}{\sqrt{3}}\right)$

$= 3.22 \pm 0.08$ mM

7. Two students reported the following results for the analysis of toothpaste for fluoride. Can you conclude whether or not the students analyzed the same brand of toothpaste?

Student A: [F-]
0.391
0.385
0.395

$N_1 = 3$

mean = 0.390; s = 0.00505

Student B:
0.385
0.368
0.372
0.370

$N_2 = 4$

$$S_{pooled} = \left[\frac{(0.00505)^2(2) + (0.00768)^2(3)}{3+4-2}\right]^{1/2} = 0.007$$

mean = 0.374; s = 0.00768

$$|\bar{x}_1 - \bar{x}_2| = \frac{t\, s_p}{\sqrt{\frac{N_1 N_2}{N_1+N_2}}}$$

$$t_{calc} = \frac{|\bar{x}_1 - \bar{x}_2|}{s_p}\sqrt{\frac{N_1 N_2}{N_1+N_2}} = 2.99$$

t_{table} at 95% and df = 5 is 2.57

Since $t_{calc} > t_{table}$ then means are different at 95% confidence level.

8. The Cu content (wt%) of five different ore samples (collected at five locations on Mt Washington) was measured by each of two methods. Do the two analytical techniques give results that are significantly different at the 95% confidence level?

Sample	Method 1	Method 2	$d = x_1 - x_2$
A	0.0134	0.0135	−0.0001
B	0.0144	0.0156	−0.0012
C	0.0126	0.0137	−0.0011
D	0.0125	0.0137	−0.0012
E	0.0137	0.0136	0.0001

$|\bar{d}| = 0.00070$

$S_d = 0.00064$

$$|\bar{d}| = \frac{t\, s_d}{\sqrt{N}}$$

$$t_{calc} = \frac{|\bar{d}|}{S_d}\sqrt{N} = \frac{(0.00070)}{(0.00064)}\sqrt{5} = 2.45$$

t_{table} at 95% and df = 4 is 2.78

Since $t_{calc} < t_{table}$ the two methods are NOT different