Basic Statistics and error propagation

$$\bar{x} = \frac{\sum_i x_i}{n}$$

$$s = \sqrt{\frac{\sum_i (x_i - \bar{x})^2}{n - 1}}$$

$$\sigma_n = \frac{\sigma}{\sqrt{n}}$$

TABLE 3-1 Summary of rules for propagation of uncertainty

Function	Uncertainty	Function[a]	Uncertainty[b]
$y = x_1 + x_2$	$e_y = \sqrt{e_{x_1}^2 + e_{x_2}^2}$	$y = x^a$	$\%e_y = a\%e_x$
$y = x_1 - x_2$	$e_y = \sqrt{e_{x_1}^2 + e_{x_2}^2}$	$y = \log x$	$e_y = \frac{1}{\ln 10}\frac{e_x}{x} \approx 0.434\,29\,\frac{e_x}{x}$
$y = x_1 \cdot x_2$	$\%e_y = \sqrt{\%e_{x_1}^2 + \%e_{x_2}^2}$	$y = \ln x$	$e_y = \frac{e_x}{x}$
$y = \frac{x_1}{x_2}$	$\%e_y = \sqrt{\%e_{x_1}^2 + \%e_{x_2}^2}$	$y = 10^x$	$\frac{e_y}{y} = (\ln 10)e_x \approx 2.302\,6\,e_x$
		$y = e^x$	$\frac{e_y}{y} = e_x$

a. x represents a variable and a represents a constant that has no uncertainty.

b. e_x/x is the relative error in x and $\%e_x$ is $100 \times e_x/x$.

Harris, *Quantitative Chemical Analysis,* 8e

Student's T-test

$$t_{\text{calculated}} = \frac{|\bar{x}_1 - \bar{x}_2|}{s_{\text{pooled}}}\sqrt{\frac{n_1 n_2}{n_1 + n_2}}$$

$$s_{\text{pooled}} = \sqrt{\frac{\sum_{\text{set 1}}(x_i - \bar{x}_1)^2 + \sum_{\text{set 2}}(x_j - \bar{x}_2)^2}{n_1 + n_2 - 2}} = \sqrt{\frac{s_1^2(n_1 - 1) + s_2^2(n_2 - 1)}{n_1 + n_2 - 2}}$$

$$t_{\text{calculated}} = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{s_1^2/n_1 + s_2^2/n_2}}$$

$$\text{Degrees of freedom} = \frac{(s_1^2/n_1 + s_2^2/n_2)^2}{\frac{(s_1^2/n_1)^2}{n_1 - 1} + \frac{(s_2^2/n_2)^2}{n_2 - 1}}$$

$$t_{\text{calculated}} = \frac{|\bar{d}|}{s_d}\sqrt{n}$$

$$s_d = \sqrt{\frac{\sum(d_i - \bar{d})^2}{n - 1}}$$

TABLE 4-2 **Values of Student's *t***

Degrees of freedom	Confidence level (%)						
	50	90	95	98	99	99.5	99.9
1	1.000	6.314	12.706	31.821	63.656	127.321	636.578
2	0.816	2.920	4.303	6.965	9.925	14.089	31.598
3	0.765	2.353	3.182	4.541	5.841	7.453	12.924
4	0.741	2.132	2.776	3.747	4.604	5.598	8.610
5	0.727	2.015	2.571	3.365	4.032	4.773	6.869
6	0.718	1.943	2.447	3.143	3.707	4.317	5.959
7	0.711	1.895	2.365	2.998	3.500	4.029	5.408
8	0.706	1.860	2.306	2.896	3.355	3.832	5.041
9	0.703	1.833	2.262	2.821	3.250	3.690	4.781
10	0.700	1.812	2.228	2.764	3.169	3.581	4.587
15	0.691	1.753	2.131	2.602	2.947	3.252	4.073
20	0.687	1.725	2.086	2.528	2.845	3.153	3.850
25	0.684	1.708	2.060	2.485	2.787	3.078	3.725
30	0.683	1.697	2.042	2.457	2.750	3.030	3.646
40	0.681	1.684	2.021	2.423	2.704	2.971	3.551
60	0.679	1.671	2.000	2.390	2.660	2.915	3.460
120	0.677	1.658	1.980	2.358	2.617	2.860	3.373
∞	0.674	1.645	1.960	2.326	2.576	2.807	3.291

In calculating confidence intervals, σ may be substituted for s in Equation 4-6 if you have a great deal of experience with a particular method and have therefore determined its "true" population standard deviation. If σ is used instead of s, the value of t to use in Equation 4-6 comes from the bottom row of Table 4-2.

Values of t in this table apply to two-tailed tests illustrated in Figure 4-9a. The 95% confidence level specifies the regions containing 2.5% of the area in each wing of the curve. For a one-tailed test, we use values of t listed for 90% confidence. Each wing outside of t for 90% confidence contains 5% of the area of the curve.

Harris, *Quantitative Chemical Analysis*, 8e

$$\text{Confidence interval} = \bar{x} \pm \frac{ts}{\sqrt{n}}$$

TABLE 4-4 **Critical values of $F = s_1^2/s_2^2$ at 95% confidence level**

Degrees of freedom for s_2	Degrees of freedom for s_1													
	2	3	4	5	6	7	8	9	10	12	15	20	30	∞
2	19.0	19.2	19.2	19.3	19.3	19.4	19.4	19.4	19.4	19.4	19.4	19.4	19.5	19.5
3	9.55	9.28	9.12	9.01	8.94	8.89	8.84	8.81	8.79	8.74	8.70	8.66	8.62	8.53
4	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	5.75	5.63
5	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.50	4.36
6	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.81	3.67
7	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.58	3.51	3.44	3.38	3.23
8	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.08	2.93
9	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.86	2.71
10	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.84	2.77	2.70	2.54
11	3.98	3.59	3.36	3.20	3.10	3.01	2.95	2.90	2.85	2.79	2.72	2.65	2.57	2.40
12	3.88	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.47	2.30
13	3.81	3.41	3.18	3.02	2.92	2.83	2.77	2.71	2.67	2.60	2.53	2.46	2.38	2.21
14	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46	2.39	2.31	2.13
15	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.25	2.07
16	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35	2.28	2.19	2.01
17	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.31	2.23	2.15	1.96
18	3.56	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19	2.11	1.92
19	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23	2.16	2.07	1.88
20	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.04	1.84
30	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.84	1.62
∞	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.67	1.57	1.46	1.00

Critical values of F for a one-tailed test of the hypothesis that $s_1 > s_2$*. There is a 5% probability of observing F above the tabulated value.*

You can compute F for a chosen level of confidence with the Excel function FINV(probability,deg_freedom1,deg_freedom2). The statement "=FINV(0.05,7,6)" reproduces the value F = 4.21 in this table. The statement "=FINV(0.1,7,6)" gives F = 3.01 for 90% confidence.

Harris, *Quantitative Chemical Analysis,* 8e

TABLE 4-5 Critical values of *G* for rejection of outlier

Number of observations	G (95% confidence)
4	1.463
5	1.672
6	1.822
7	1.938
8	2.032
9	2.110
10	2.176
11	2.234
12	2.285
15	2.409
20	2.557

$G_{calculated}$ = |questionable value − mean|/s. If $G_{calculated} > G_{table}$, the value in question can be rejected with 95% confidence. Values in this table are for a one-tailed test, as recommended by ASTM.

SOURCE: *ASTM E 178-02 Standard Practice for Dealing with Outlying Observations, http://webstore.ansi.org; F. E. Grubbs and G. Beck, Technometrics* ***1972,*** *14, 847.*

Harris, *Quantitative Chemical Analysis,* 8e

$$G_{\text{calculated}} = \frac{|\text{questionable value} - \bar{x}|}{s}$$

$$F_{\text{calculated}} = \frac{s_1^2}{s_2^2}$$

Least Squares fitting and Standard addition

$$y = mx + b$$

$$\text{Uncertainty in } x\ (= s_x) = \frac{s_y}{|m|}\sqrt{\frac{1}{k} + \frac{1}{n} + \frac{(y - \bar{y})^2}{m^2 \sum (x_i - \bar{x})^2}}$$

$$s_y = \sqrt{\frac{\sum (d_i^2)}{n - 2}}$$

$$[\mathrm{X}]_\mathrm{f} = [\mathrm{X}]_\mathrm{i}\left(\frac{V_\mathrm{o}}{V}\right) \qquad [\mathrm{S}]_\mathrm{f} = [\mathrm{S}]_\mathrm{i}\left(\frac{V_\mathrm{S}}{V}\right)$$

$$\frac{[\mathrm{X}]_\mathrm{i}}{[\mathrm{S}]_\mathrm{f} + [\mathrm{X}]_\mathrm{f}} = \frac{I_\mathrm{X}}{I_{\mathrm{S+X}}}$$

$$\text{Standard deviation of } x\text{-intercept} = \frac{s_y}{|m|}\sqrt{\frac{1}{n} + \frac{\bar{y}^2}{m^2 \sum (x_i - \bar{x})^2}}$$

QA and internal standard

$$\% \text{ recovery} = \frac{C_{\text{spiked sample}} - C_{\text{unspiked sample}}}{C_{\text{added}}} \times 100$$

$$\text{Lower limit of quantitation} \equiv \frac{10s}{m}$$

$$\text{Minimum detectable concentration} = \frac{3s}{m}$$

$$y_{\text{dl}} = y_{\text{blank}} + 3s$$

$$y_{\text{sample}} - y_{\text{blank}} = m \times \text{sample concentration}$$

$$R^2 = \frac{[\sum(x_i - \bar{x})(y_i - \bar{y})]^2}{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}$$

$$\frac{A_X}{[X]} = F\left(\frac{A_S}{[S]}\right)$$