Essential statistics for quality assurance (II)

Vicki Barwick

Overview

- *t*-tests
- *F*-test
- Analysis of variance (ANOVA)
- Excel data analysis tools

Typical questions

- Comparison of the mean of a data set with a known value
  - e.g. are the results from the analysis of a CRM significantly different from the certified value? One-sample *t*-test
- Comparison of the means of two independent data sets
  - e.g. is there any significant difference between the results produced by two analysts? Two-sample *t*-test
- Comparison of pairs of data obtained from two treatments applied once each to a range of different test samples
  - e.g. is there any significant difference between the results produced by two different test methods? Paired-sample *t*-test
- Comparison of the standard deviations of two independent data sets
  - e.g. is there any significant difference between the precision produced by two methods? *F*-test

One sample *t*-test

Alternative Hypothesis	*t*	Tests for
Not equal to x_0 (two-tailed)	$t = \frac{\lvert \bar{x} - x_0 \rvert}{s/\sqrt{n}}$	Any difference?
Greater than x_0 (one-tailed)	$t = \frac{(\bar{x} - x_0)}{s/\sqrt{n}}$	Exceeding reference value/ upper limit
Less than x_0 (one-tailed)	$t = \frac{(x_0 - \bar{x})}{s/\sqrt{n}}$	Below reference value/ lower limit

Significance: $t > t_{crit}$

One-sample *t*-test
Example

- Validation of a method for the determination of arsenic in effluent – analysis of a certified reference material (CRM)
  - mean = 33.9 µg L^{-1} (n = 11), s = 0.63 µg L^{-1}
  - certified value = 32.4 µg L^{-1}
- State the question
  - does the mean of the results from the analysis of the CRM differ significantly from the certified value?
- Select the test
  - we are comparing a mean value with a reference value – one-sample *t*-test
- Choose level of significance
  - 5% significance (α = 0.05, 95% confidence)
- Decide number of tails
  - two-tailed test (interested in a difference either direction)

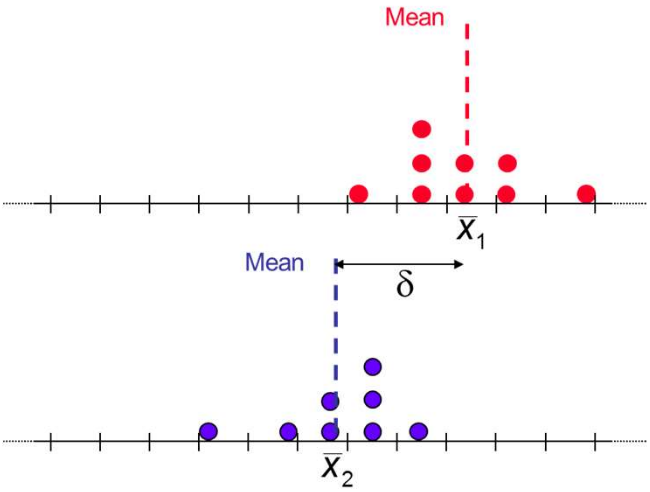
One-sample *t*-test
Example (continued)

- Calculate degrees of freedom
  - degrees of freedom $\nu = n-1 = 10$
- Obtain critical value
  - 5% significance, two-tailed test, 10 degrees of freedom
    - $t_{0.05,10} = 2.228$
- Calculate test statistic from experimental data

$$t = \frac{|\bar{x} - x_0|}{s/\sqrt{n}} = \frac{|33.9 - 32.4|}{0.63/\sqrt{11}} = 7.897$$

- Compare the test statistic with the critical value
  - $t > t_{0.05,10}$ the mean is significantly different from the certified value

Significance testing between sets of data
Two-sample *t*-test



$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\left[\frac{s_1^2(n_1 - 1) + s_2^2(n_2 - 1)}{n_1 + n_2 - 2}\right]}}$$

If $n_1 = n_2$:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2 + s_2^2}{n}}}$$

$$\nu = n_1 + n_2 - 2$$

(Assumes equal variance*)

*Note – there is also an 'unequal variance' version of the test

Two-sample *t*-test
Example

- 2 methods for determining selenium in cabbage are being compared
- 16 test portions are selected from the same cabbage sample
- 8 portions are analysed using each method
- Is there any significant difference between the means of the results obtained using the 2 methods (95% confidence)?

	n	Mean $\bar{x}$ (mg/100 g)	Standard deviation s (mg/100 g)
Method 1	8	0.199	0.0123
Method 2	8	0.155	0.00810

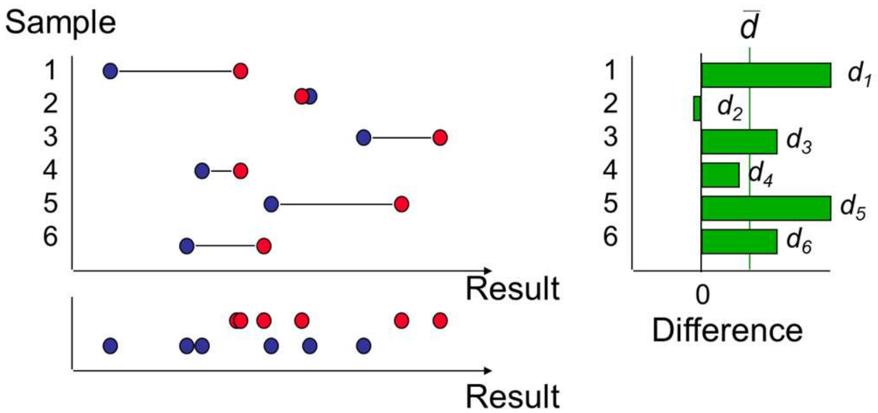
Two-sample t-test
Example

- Comparing 2 independent estimates of the mean, variances of datasets are not significantly different – two-sample t-test assuming equal variance
- 95% confidence
- Two-tailed test – is there a difference between the mean values?
- Degrees of freedom: $\nu = n_1+n_2-2 = 14$
- Critical value: $t_{0.05,14} = 2.145$ (two-tailed)
- Calculate test statistic from experimental data

$$t = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\frac{s_1^2 + s_2^2}{n}}} = \frac{|0.199 - 0.155|}{\sqrt{\frac{0.0123^2 + 0.00810^2}{8}}} = 8.450$$

$t>t_{0.05,14}$ there is a significant difference between the means

Significance testing between paired samples
Paired sample t-test



$$t = \frac{\bar{d}}{s(d)/\sqrt{n}}$$

Paired *t*-test
Example

- 2 methods for determining GMO in maize are being compared
- 6 different samples of maize analysed
- Each sample divided into 2 parts – one half analysed using Method A, the other half analysed using Method B
- Is there any significant difference between the results obtained using the 2 methods (95% confidence)?
- The data are paired

	n	Mean difference $\bar{d}$ (%GMO by mass)	Standard deviation of differences of differences $s(d)$ (%GMO by mass)
Method A-B	6	-0.0688	0.0226

Paired-sample *t*-test
Example

- Comparing pairs of data – paired *t*-test
- 95% confidence
- Two-tailed test – is there a difference between the results obtained using the 2 methods?
- Degrees of freedom: $\nu = n_{pairs} - 1 = 5$
- Critical value: $t_{0.05,5} = 2.571$ (two-tailed)
- Calculate test statistic from experimental data

$$t = \frac{|\bar{d}|}{s(d)/\sqrt{n}} = \frac{|-0.0688|}{0.0226/\sqrt{6}} = 7.457$$

$t > t_{0.05,5}$ there is a significant difference between the results obtained from method A and method B

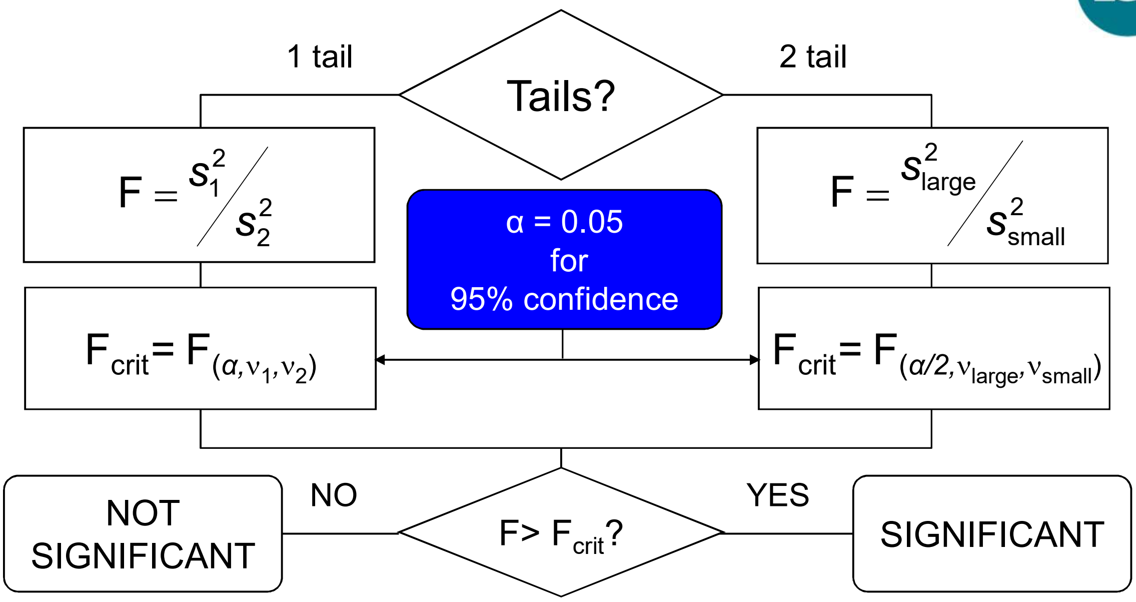
The *F*-test

- To compare the spread, use the ratio of variances:

$$F = {s_1^2}/{s_2^2}$$

- This ratio, the ‘*F*-statistic’, can be compared with values in tables (the ‘*F*-test’)

Rules for the *F*-test



F-test

Example

- 2 methods for determining selenium in cabbage are being compared
- 16 test portions are selected from the same cabbage sample
- 8 portions are analysed using each method
- Is there any significant difference between the precision of the results obtained using the two methods (95% confidence)?

	n	Mean (mg/100 g)	*s* (mg/100 g)
Method 1	8	0.199	0.0123
Method 2	8	0.155	0.00810

F-test

Example

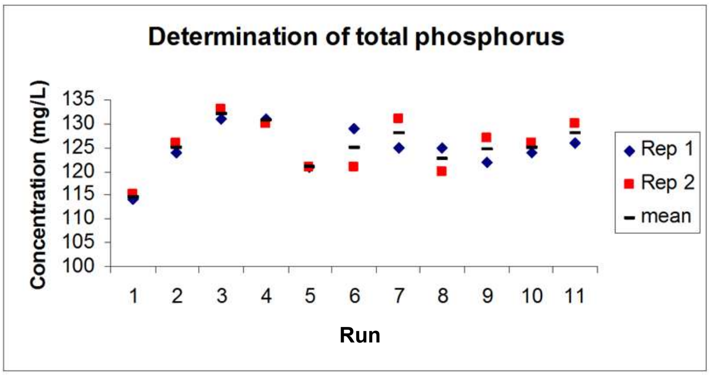
- Comparing variability (standard deviations) – *F*-test
- 95% confidence
- Two-tailed test – is there any difference between the variance of the results obtained using the 2 methods?
- Degrees of freedom: $\nu = n-1 = 7$ for both data sets
- Critical value: $F_{0.025,7,7} = 4.995$ (one-tailed value for $\alpha/2$ to give required two-tailed value)
- Calculate test statistic from experimental data (larger variance as numerator for two-tailed test)

$$F = \frac{0.0123^2}{0.00810^2} = 2.306$$

$F < F_{0.025,7,7}$ there is no significant difference between the variance of the results obtained from 2 methods

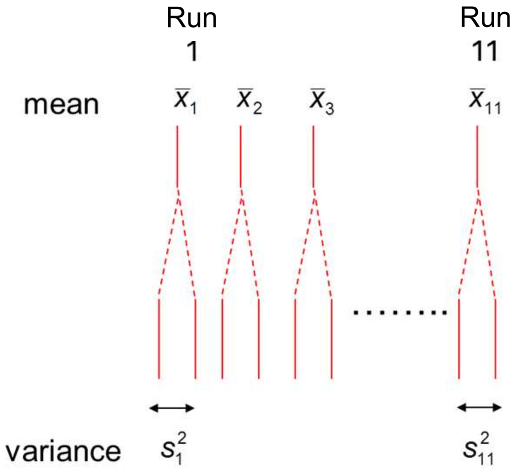
Comparing multiple groups of data

- Variation between duplicates (within-run)
- Variation between runs – measurements made on different days



- Does the variation increase significantly when measurements are made on different days?

Within- and between-group effects
Nested -design



Total variance has contributions from

- Random variation between duplicates (within-run)
- Variation between results obtained in different batches (between-run)

Analysis of variance (ANOVA)

- ANOVA separates different sources of variation
  - e.g. the within- and between-run variation in results
- Different sources of variation can be compared to determine whether they are significantly different
  - e.g. is the between-run variability in results significantly greater than the within-run variability?
- H_0 is that all samples are drawn from same population
- Method validation precision study
  - can be useful to know where variation in results is coming from
    - within-run vs. between-run

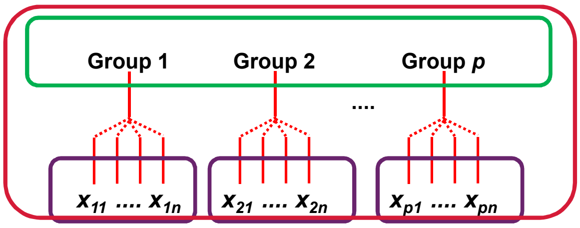
Anatomy of an ANOVA table

Source of variation	Sum of Squares (SS)	ν	Mean Square (MS)	F
Between groups	SS_b	p-1	$MS_b = SS_b/(p-1)$	MS_b/MS_w
Within group (Residuals)	SS_w	N-p	$MS_w = SS_w/(N-p)$	
Total	$SS_{tot} = SS_b+SS_w$	N-1		

No. groups = p

No. replicates = n

Total no. of results = $np = N$



Comparing sources of variation

- Variance contributions compared as Mean Squares

$$Mean\ square\ (MS) = \frac{SS}{\nu}$$

- Mean squares compared using an F-test
  - is the between group MS significantly greater than the within group MS?

$$F = \frac{Between\ group\ MS}{Within\ group\ MS}$$

$F > F_{crit}$ ⇒ differences between groups of data are significant compared to within group variation

ANOVA: Single factor – total phosphorus

ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	459.8182	10	45.98182	5.620	0.004312	2.854
Within Groups	90.00	11	8.181818			
Total	549.8182	21				

$F > F_{crit}$, $P < 0.05$ ⇒ Significant difference between results obtained in different runs

Estimating precision from ANOVA

- Repeatability, s_r (within-run precision)

$$s_r = \sqrt{within\ group\ MS}$$

- The combined precision has contributions from the within and between group variability

$$s_{between} = \sqrt{\frac{between\ group\ MS - within\ group\ MS}{n}}$$

n = no. results per group

$$s_c = \sqrt{s_r^2 + s_{between}^2}$$

- If groups produced by different analysts, different instruments, etc, s_c is the intermediate precision, s_I
- If groups produced by different labs, s_c is the reproducibility standard deviation, s_R

Precision calculation – total phosphorus

- Repeatability, s_r

$$s_r = \sqrt{8.181818} = 2.86\ \mathrm{mg/L}$$

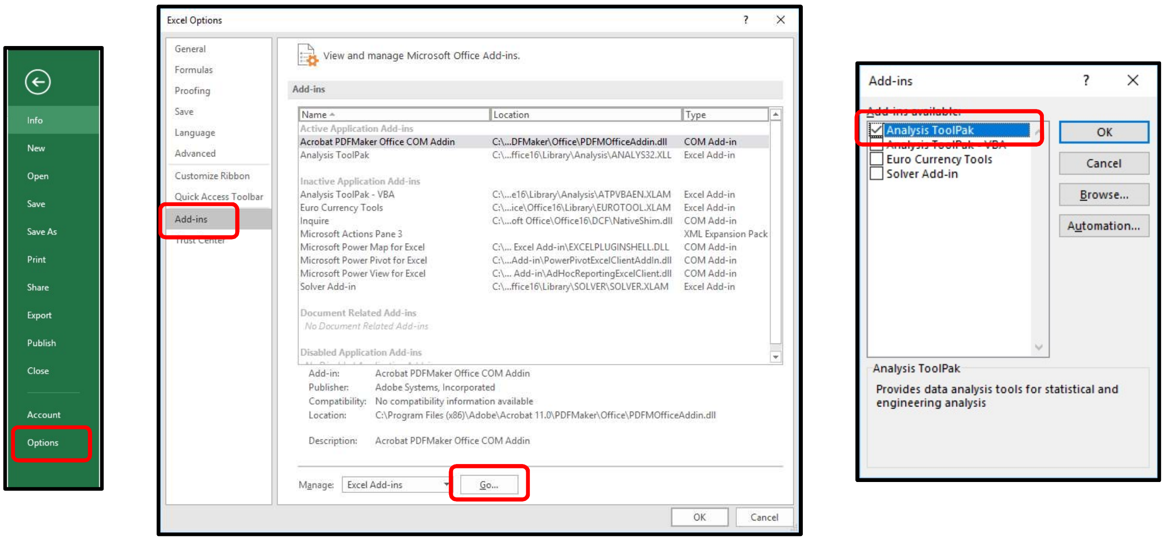
- Between-run standard deviation

$$s_{between} = \sqrt{\frac{45.98182 - 8.181818}{2}} = 4.35\ \mathrm{mg/L}$$

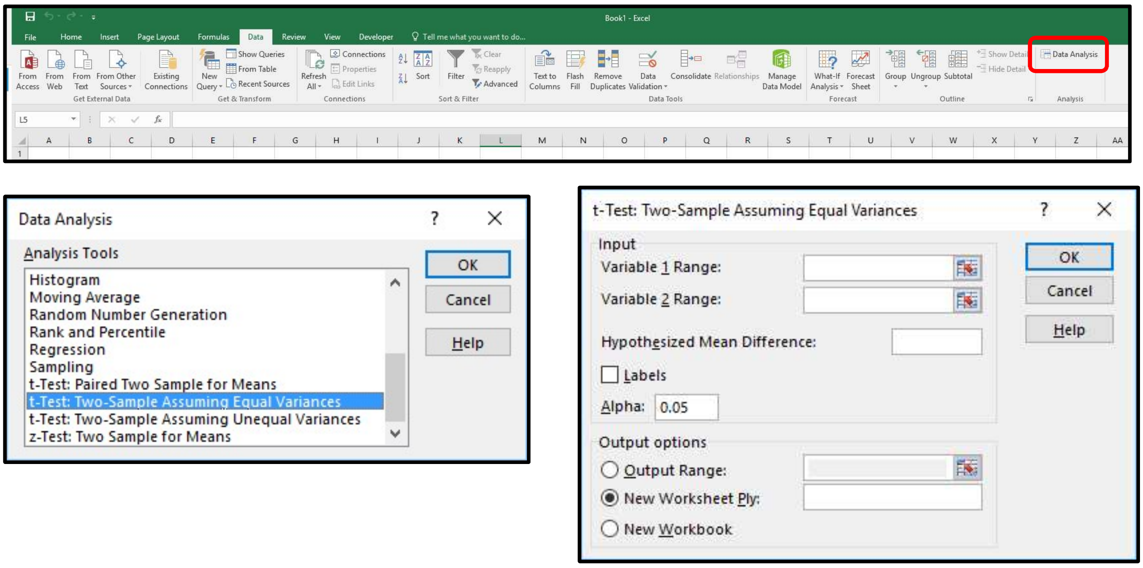
- Intermediate precision, s_I

$$s_c = \sqrt{2.86^2 + 4.35^2} = 5.21\ \mathrm{mg/L}$$

Excel – data analysis tools



Excel – data analysis tools



Any questions?

LGC