SOP 23

Statistical techniques used in quality assessment

1. Scope and field of application

This procedure describes various statistical calculations used in quality assessment. Calculations are detailed which allow the computation of the mean and standard deviation of a set of values, the computation of a standard deviation from a set of duplicate measurements, the computation of the confidence interval for a mean, the examination of the values of two means or of two standard deviations to assess if they are significantly different at some chosen level of probability, and the computation of the least-squares estimates of the slope and intercept of a straight line.

2. Principle

These calculations are based on statistical principles, specifically on the normal distribution. More details of the relevant statistical background are given in the bibliography.

3. Procedure

3.1 Estimation of the mean and standard deviation from a series of measurements

Given n measurements,

$$x_1, x_2, x_3, \ldots, x_n,$$

the mean, $\bar{x}$, is given by

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad (1)$$

and an estimate of the standard deviation, s, is given by

$$s = \left(\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1} \right)^{1/2} . \qquad (2)$$

3.2 Estimation of the standard deviation from the difference of sets of duplicate measurements

Given k differences of duplicate measurements,

$$d_1, d_2, d_3, \ldots, d_k ,$$

an estimate of the standard deviation, s, is given by

$$s_R = \left(\frac{\sum_{i=1}^{k} d_i^2}{2k} \right)^{1/2} . \tag{3}$$

This is a measure of the short-term standard deviation, or repeatability of measurements[1].

3.3 Confidence interval for a mean

The formula for use is

$$\bar{x} \pm \frac{ts}{\sqrt{n}} \tag{4}$$

where

$\bar{x}$ = sample mean,
n = number of measurements on which the mean is based,
s = estimate of the standard deviation[2],
t = Student's t value, *i.e.*, the probability factor for the desired confidence limit and the number of degrees of freedom associated with s. (For numerical values, see Table 1 in the Annexe to this procedure.)

3.4 Comparing values of two means

Case 1. No reason to believe that the standard deviations differ.

Step 1: Choose α, the desired probability level (*i.e.*, the significance level) of the test.

Step 2: Calculate a pooled standard deviation from the two estimates to obtain a better estimate of the standard deviation:

---

1 The International Organization for Standardization (ISO) applies two descriptions of precision: (1) the *reproducibility*, the closeness of agreement between individual results obtained with the same method but under different conditions (*e.g.*, in different laboratories) and (2) the *repeatability*, the closeness of agreement between successive results obtained with the same method and under the same conditions.

2 If $\bar{x}$ and s are based on the same data set, the number of degrees of freedom, df = $n - 1$. However, if s is based on additional evidence, such as a system under statistical control (judged by a control chart), then the degrees of freedom on which the estimate of s is based may be used to determine t. In such a case, one can calculate a confidence interval for even a single measurement.

$$s_\mathrm{p} = \left(\frac{\nu_\mathrm{A} s_\mathrm{A}^2 + \nu_\mathrm{B} s_\mathrm{B}^2}{\nu_\mathrm{A} + \nu_\mathrm{B}} \right)^{1/2} \tag{5}$$

where ν_A and ν_B are the number of degrees of freedom associated with s_A and s_B, respectively. s_p will thus be based on $\nu_\mathrm{A} + \nu_\mathrm{B}$ degrees of freedom.

Step 3: Calculate the uncertainty, U, of the differences

$$U = t s_\mathrm{p} \left(\frac{1}{n_\mathrm{A}} + \frac{1}{n_\mathrm{B}} \right)^{1/2} \tag{6}$$

where t is the appropriate Student's t value.

Step 4: Compare $\Delta = \left| \bar{x}_\mathrm{A} - \bar{x}_\mathrm{B} \right|$ with U. If $\Delta \le U$, there is no reason to believe that the means disagree.

Case 2. The standard deviations differ significantly (see section 3.5).

Step 1: Choose α, the significance level of the test.

Step 2: Compute the estimated variance of each mean using the individual estimates of the standard deviations,

$$V_\mathrm{A} = s_\mathrm{A}^2 / n_\mathrm{A}, \qquad V_\mathrm{B} = s_\mathrm{B}^2 / n_\mathrm{B} \; . \tag{7}$$

Step 3: Compute the effective number of degrees of freedom[3]:

$$f^* = \frac{\left(V_\mathrm{A} + V_\mathrm{B}\right)^2}{\dfrac{V_\mathrm{A}^2}{n_\mathrm{A}+1} + \dfrac{V_\mathrm{B}^2}{n_\mathrm{B}+1}} - 2 \; . \tag{8}$$

Step 4: Calculate the uncertainty, U, of the differences

$$U = t^* \sqrt{V_\mathrm{A} + V_\mathrm{B}} \tag{9}$$

where t^* is the effective value of t based on f^* degrees of freedom and the chosen significance level, α (Table 1 in the Annexe).

Step 5: Compare $\Delta = \left| \bar{x}_\mathrm{A} - \bar{x}_\mathrm{A} \right|$ with U. If $\Delta \le U$, there is no reason to believe that the means disagree.

3.5 Comparing estimates of a standard deviation (F test)

This test may be used to decide whether there is sufficient reason to believe that two estimates of a standard deviation are significantly different. It consists of

[3] A number of expressions exist in the literature for this calculation, with some authors even arguing that such a pooling of the variances is inappropriate. The expression used here comes from Taylor (1987).

calculating the ratio of the variances and comparing it with tabulated values. Unless the computed ratio is larger than the tabulated value, there is no reason to believe that the respective standard deviations are significantly different.

The F ratio is calculated as

$$F = \frac{s_{\mathrm{L}}^2}{s_{\mathrm{S}}^2} \tag{10}$$

where s_{L} is the larger value and s_{S} is the smaller of the two estimates under consideration. The critical value of F will depend on the significance level chosen and on the degrees of freedom associated with s_{L} and s_{S} (see Table 2 in the Annexe).

3.6 Computation of least-squares estimates

For the linear model,

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \tag{11}$$

where x is essentially without error (for data with errors in x and y—see York, 1966) and the error ε_i is normally distributed with a constant variance, least-squares estimates of the coefficients, β_0 and β_1, are given by the expressions

$$\beta_1 = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sum_i (x_i - \bar{x})^2} , \tag{12}$$

$$\beta_0 = \bar{y} - \beta_1 \bar{x} . \tag{13}$$

An estimate of the experimental error variance is then given by

$$s^2 = \frac{\sum_i (y_i - \beta_0 - \beta_1 x_i)^2}{n-2} \tag{14}$$

and estimates of the standard errors of the coefficients by

$$\mathrm{S.E.}(\beta_0) = s\left(\frac{1}{n} + \frac{\bar{x}^2}{\sum_i (x_i - \bar{x})^2}\right)^{1/2} , \tag{15}$$

$$\mathrm{S.E.}(\beta_1) = \frac{s}{\left(\sum_i (x_i - \bar{x})^2\right)^{1/2}} . \tag{16}$$

4. Example calculations

4.1 Estimation of the mean and standard deviation from a series of measurements

Given the following 9 measurements:

1977.67, 1977.98, 1977.29, 1978.60, 1979.48,
1979.14, 1979.33, 1979.95, 1979.99,

the mean is 1978.83 and the standard deviation is 0.99.

4.2 Estimation of the standard deviation from the difference of sets of duplicate measurements

Given 10 pairs of measurements:

1976.8, 1979.3,   1978.9, 1979.6,   1979.6, 1979.8
1978.3, 1978.6,   1981.2, 1979.8,   1977.6, 1977.8
1976.2, 1976.8,   1978.6, 1977.0,   1976.6, 1978.9
1978.3, 1978.9,

the standard deviation calculated using

$$s_R = \left(\frac{\sum_{i=1} d_i^2}{2k} \right)^{1/2}$$

is 0.93.

4.3 Confidence interval for a mean

The 95% confidence interval for the mean calculated in section 4.1 is

$$1978.83 \pm \frac{(2.306)(0.99)}{\sqrt{9}} = 1978.83 \pm 0.76$$

4.4 Comparing values for two means

Case 1. No reason to believe that the standard deviations differ.

$$\bar{x}_A = 1978.78, \quad s_A = 0.93, \quad n_A = 9$$
$$\bar{x}_B = 1981.74, \quad s_B = 0.87, \quad n_B = 18$$

Step 1: Require 95% confidence in decision.

Step 2: Pooled standard deviation:

$$s_p = \left(\frac{8(0.93)^2 + 17(0.87)^2}{8+17} \right)^{1/2}$$
$$= 0.89.$$

Step 3: Calculate U:

$$U = 2.060(0.89)\left(\frac{1}{9} + \frac{1}{18}\right)^{1/2}$$
$$= 0.75.$$

Step 4: As Δ (= 1981.74 − 1978.78 = 2.96) is larger than U, the means disagree at the 95% confidence level.

Case 2. The standard deviations differ significantly.

$$\bar{x}_A = 1978.78, \quad s_A = 0.93, \quad n_A = 9$$
$$\bar{x}_B = 1981.74, \quad s_B = 2.75, \quad n_B = 16$$

Step 1: Require 95% confidence in decision.

Step 2: Compute the estimated variance of each mean:

$$V_A = (0.93)^2 / 9 = 0.0961$$
$$V_B = (2.75)^2 / 16 = 0.4727.$$

Step 3: Compute the effective number of degrees of freedom:

$$f^* = \left[\frac{(0.0961 + 0.4727)^2}{(0.0961)^2/(9+1) + (0.4727)^2/(16+1)} \right] - 2 \approx 21.$$

Step 4: Calculate U:

$$U = 2.08(0.0961 + 0.4727)^{1/2} = 1.57 .$$

Step 5: As Δ (= 1981.74 − 1978.78 = 2.96) is larger than U, the means disagree at the 95% confidence level.

4.5 Comparing estimates of a standard deviation

$$\bar{x}_A = 1978.78, \quad s_A = 0.93, \quad n_A = 9$$
$$\bar{x}_B = 1975.35, \quad s_B = 1.71, \quad n_B = 12$$

Calculate F:

$$F = \frac{(1.71)^2}{(0.93)^2} = 3.38 .$$

The tabulated value of F—with 8 degrees of freedom in the numerator and 11 degrees of freedom in the denominator—is 3.7. As the computed value is smaller than the tabulated value, there is no reason to believe that the two standard deviations are significantly different.

4.6 Computation of least-squares estimates

Given 6 pairs of measurements of *x* and *y*:

0.0	1892
498.8	66537
1001.9	130818
1500.8	195216
2002.5	260068
2497.1	323456

Linear regression gives

$$\beta_0 = 2017.77,$$
$$\beta_1 = 128.765.$$

The error estimates are

$$s = 221.77,$$
$$\text{S.E.}(\beta_0) = 160.55,$$
$$\text{S.E.}(\beta_1) = 0.106.$$

5. Bibliography

Box, G.E.P., Hunter, W.G. and Hunter, J.S. 1978. Statistics for Experimenters. An Introduction to Design, Data Analysis, and Model Building. John Wiley & Sons, 653 pp.

Meier, P.C. and Zünd, R.E. 1993. Statistical Methods in Analytical Chemistry. Wiley Interscience, New York, 321 pp.

Ryan, T.P. 1989. Statistical Methods for Quality Improvement. John Wiley & Sons, Inc., New York, 446 pp.

Taylor, J.K. 1987. Quality Assurance of Chemical Measurements. Lewis Publishers, Inc., Chelsea, 328 pp.

Taylor, J.K. 1990. Statistical Techniques for Data Analysis. Lewis Publishers, Inc., Chelsea, 200 pp.

York, D. 1966. Least-squares fitting of a straight line. *Can. J. Physics* **44**: 1079–1086.

Annexe

Table 1 Student's *t* values for 95% and 99% confidence intervals.

Probability level for two-sided confidence interval		
df[4]	**95%**	**99%**
1	12.706	63.657
2	4.303	9.925
3	3.182	5.841
4	2.776	4.604
5	2.571	4.032
6	2.447	3.707
7	2.365	3.499
8	2.306	3.355
9	2.262	3.250
10	2.228	3.169
11	2.201	3.106
12	2.179	3.055
13	2.160	3.012
14	2.145	2.977
15	2.131	2.947
16	2.120	2.921
17	2.110	2.898
18	2.101	2.878
19	2.093	2.861
20	2.086	2.845
25	2.060	2.787
40	2.021	2.704
60	2.000	2.660
∞	1.960	2.576

[4] degrees of freedom ($n - 1$)

Table 2 Critical values for the *F* test for use in a two-tailed test of equality of standard deviation at 95% level of confidence.

df_D	df_N									
	1	2	4	6	8	10	15	20	30	40
1	648	800	900	937	957	969	983	993	1001	1006
2	38.5	39.0	39.2	39.3	39.4	39.4	39.4	39.4	39.5	39.5
4	12.2	10.6	9.6	9.2	9.0	8.8	8.7	8.6	8.5	8.4
6	8.8	7.3	6.2	5.8	5.6	5.5	5.3	5.2	5.1	5.0
8	7.6	6.1	5.0	4.6	4.4	4.3	4.1	4.0	3.9	3.8
10	6.9	5.5	4.5	4.1	3.8	3.7	3.5	3.4	3.3	3.3
15	6.2	4.8	3.8	3.4	3.2	3.1	2.9	2.8	2.6	2.6
20	5.9	4.5	3.5	3.1	2.9	2.8	2.6	2.5	2.4	2.3
30	5.6	4.2	3.2	2.9	2.6	2.5	2.3	2.2	2.1	2.0
40	5.4	4.0	3.1	2.7	2.5	2.4	2.2	2.1	1.9	1.9
60	5.3	3.9	3.0	2.6	2.4	2.3	2.1	1.9	1.8	1.7
120	5.2	3.8	2.9	2.5	2.3	2.2	1.9	1.8	1.7	1.6
∞	5.0	3.7	2.8	2.4	2.2	2.1	1.8	1.7	1.6	1.5

df_D — degrees of freedom of the variance in the denominator.
df_N — degrees of freedom of the variance in the numerator.