Table 3.3

Rejection Quotient, Q, at Different Confidence Limits[a]

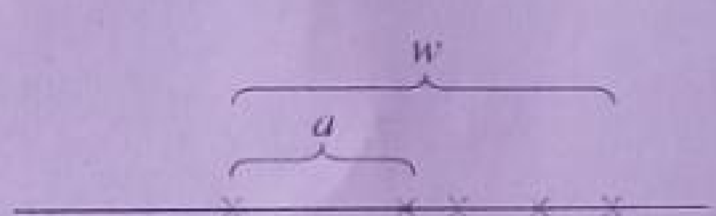
No. of Observations	Confidence Level		
	Q_{90}	Q_{95}	Q_{99}
3	0.941	0.970	0.994
4	0.765	0.829	0.926
5	0.642	0.710	0.821
6	0.560	0.625	0.740
7	0.507	0.568	0.680
8	0.468	0.526	0.634
9	0.437	0.493	0.598
10	0.412	0.466	0.568
15	0.338	0.384	0.475
20	0.300	0.342	0.425
25	0.277	0.317	0.393
30	0.260	0.298	0.372

[a] Adapted from D. B. Rorabacher, *Anal. Chem.*, 63 (1991) 139.

If $Q_{calc} > Q_{table}$, then the data point may be an outlier and could be discarded. In practice, it is a good idea to make more measurements to be sure, especially if the decision is a close one.

is determining what the range should be. If it is too small, then perfectly good data will be rejected; and if it is too large, then erroneous measurements will be retained too high a proportion of the time. The Q **test** is, among the several suggested tests, one of the most statistically correct for a fairly small number of observations and is recommended when a test is necessary. The ratio Q is calculated by arranging the data in increasing or decreasing order of numbers. If you have a large data set, you will find the Data/Sort function in Excel very helpful to arrange the data either in ascending or descending order. The difference between the suspect number and its nearest neighbor (a) is divided by the range (w), the range being the difference between the highest number and the lowest number. Referring to the figure in the margin, $Q = a/w$. This ratio is compared with tabulated values of Q. If it is equal to or greater than the tabulated value, the suspected observation can be rejected. The tabulated values of Q at the 90, 95, and 99% confidence levels are given in Table 3.3. If Q exceeds the tabulated value for a given number of observations and a given confidence level, the questionable measurement may be rejected with, for example, 95% confidence that some definite error is in this measurement.

$$Q = \frac{a}{w}$$



"And now the sequence of events in no particular order."—Dan Rather, television news anchor

Example 3.20

The following set of chloride determinations on separate aliquots of a pooled serum were reported: 103, 106, 107, and 114 meq/L. One value appears suspect. Determine if it can be ascribed to accidental error, at the 95% confidence level.

Solution

The suspect result is 114 meq/L. It differs from its nearest neighbor, 107 meq/L, by 7 meq/L. The range is 114 to 103, or 11 meq/L. Q is therefore $7/11 = 0.64$. The tabulated value for four observations is 0.829. Since the calculated Q is less than the tabulated Q, the suspected number may be ascribed to random error and should not be rejected. ■

Table 3.2

Values of F at the 95% Confidence Level

	$\nu_1 = 2$	3	4	5	6	7	8	9	10	15	20	30
$\nu_2 = 2$	19.0	19.2	19.2	19.3	19.3	19.4	19.4	19.4	19.4	19.4	19.4	19.5
3	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.70	8.66	8.62
4	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.86	5.80	5.75
5	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.62	4.56	4.50
6	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	3.94	3.87	3.81
7	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.51	3.44	3.38
8	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.22	3.15	3.08
9	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.01	2.94	2.86
10	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.85	2.77	2.70
15	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.40	2.33	2.25
20	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.20	2.12	2.04
30	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.01	1.93	1.84

where $s_1^2 > s_2^2$. There are two different degrees of freedom, ν_1 and ν_2, where degrees of freedom is defined as $N - 1$ for each case.

If the calculated F value from Equation 3.10 exceeds a tabulated F value at the selected confidence level, then there is a significant difference between the variances of the two methods. A list of F values at the 95% confidence level is given in Table 3.2.

Video: Solver

The F-test is available in Excel. The preferred method is to install the "Data Analysis" Add-In. The videos Solver, and Data Analysis Regression on the website of the book discusses installation of Add-Ins. It is also discussed at the end of Section 3.24. After installation, the Data Analysis icon will appear in the top right-hand corner of the tool bar. Click on this icon. A drop-down menu will appear. Select "F-test Two samples for Variance". Another box will appear. Select your first set of data as "Variable 1 Range" and the second set as "Variable 2 Range". You can select the confidence level you want by choosing the "alpha" value. An alpha value of 0.05 means a 95% confidence level, 0.01 connotes 99% confidence level; the default is an alpha of 0.05. The F-test results will appear in a box that will occupy 3 columns and 10 rows, giving Mean, Variance, df, F, P one tail, and F Critical one tail for Variables 1 and 2 (see also 4. t-test to Compare Different Samples on page 94). Click on the output range button and specify a cell in the box that will constitute the top left corner of the output. Your F-value should always be >1; if it is <1, interchange the data: select your second set of data as "Variable 1 Range" and the first set as "Variable 2 Range". The output will specify the critical value of F.

Video: Data Analysis Regression

Example 3.16

You are developing a new colorimetric procedure for determining the glucose content of blood serum. You have chosen the standard Folin-Wu procedure with which to compare your results. From the following two sets of replicate analyses on the same sample, determine whether the variance of your method differs significantly from that of the standard method.

The Folin-Wu method, now rarely used, uses the reaction of glucose with cupric ion (Cu^{2+}) in an alkaline medium to reduce phosphomolybdate to molybdenum blue, sometimes called heteropoly blue. Try this problem using F-test in Excel.

3.11 The Confidence Limit—How Sure Are You?

The true value falls within the confidence limit, estimated using t at the desired confidence level.

Calculation of the standard deviation for a set of data provides an indication of the precision inherent in a particular procedure or analysis. But unless there is a large amount of data, it does not by itself give any information about how close the experimentally determined mean $\bar{x}$ might be to the true mean value μ. Statistical theory, though, allows us to estimate the range within which the true value might fall, within a given probability, defined by the experimental mean and the standard deviation. This range is called the **confidence interval**, and the limits of this range are called the **confidence limit**. The likelihood that the true value falls within the range is called the **probability**, or **confidence level**, usually expressed as a percent. The confidence limit, in terms of the standard deviation, s (Equation 3.2), is given by

$$\boxed{\text{Confidence limit} = \bar{x} \pm \frac{ts}{\sqrt{N}}} \tag{3.9}$$

where t is a statistical factor that depends on the number of degrees of freedom and the confidence level desired. The number of degrees of freedom is one less than the number of measurements. Values of t at different confidence levels and degrees of freedom ν are given in Table 3.1. Note that the confidence limit is simply the product of t and the standard deviation of the mean $(s/\sqrt{N})$, also called the standard error of the mean. **(The confidence limit for a single observation ($N = 1$), x, is given by $x \pm ts$. This is larger than that of the mean by a factor $\sqrt{N}$.)**

Example 3.15

A soda ash sample is analyzed in the analytical chemistry laboratory by titration with standard hydrochloric acid. The analysis is performed in triplicate with the

Table 3.1

Values of t for ν Degrees of Freedom for Various Confidence Levels[a]

	Confidence Level			
ν	90%	95%	99%	99.5%
1	6.314	12.706	63.657	127.32
2	2.920	4.303	9.925	14.089
3	2.353	3.182	5.841	7.453
4	2.132	2.776	4.604	5.598
5	2.015	2.571	4.032	4.773
6	1.943	2.447	3.707	4.317
7	1.895	2.365	3.500	4.029
8	1.860	2.306	3.355	3.832
9	1.833	2.262	3.250	3.690
10	1.812	2.228	3.169	3.581
15	1.753	2.131	2.947	3.252
20	1.725	2.086	2.845	3.153
25	1.708	2.060	2.787	3.078
∞	1.645	1.960	2.576	2.807

[a] $\nu = N - 1$ = degrees of freedom.