Chapter 7:

Statistical Analysis Data Treatment and Evaluation

Courtesy of the Norman Rockwell Family Agency

❑ The picture here is the Norman Rockwell Saturday Evening Post cover ,The Holdout from February 14, 1959. One of the 12 jurors does not agree with the others, who are trying to convince her.

❑ In the jury room, we can make two types of errors.
  ❑ An innocent person can be convicted, or
  ❑ a guilty person can be set free.

It is a more serious error to convict an innocent person than to acquit a guilty person.

Similarly, in statistical tests to determine whether two quantities are the same, two types of errors are possible:

A type I error occurs when we reject the hypothesis that two quantities are the same, when they are statistically identical.

A type II error occurs when we accept that they are the same when they are not statistically identical.

The characteristics of these errors in statistical testing and the ways we can minimize them are among the subjects of this chapter

The most common applications of statistical data treatment :

1. Defining a numerical interval, *the confidence interval*, around the mean of a set of replicate results within which the population mean can be expected to lie with a certain probability. This interval is related to the standard deviation of the mean.

2. Determining *the number of replicate measurements* required to ensure that an experimental mean falls within a certain range with a given level of probability.

3. *Estimating the probability that* (a) an experimental mean and a true value or (b) *two experimental means are different*. This test is particularly important for discovering systematic errors in a method and determining whether two samples come from the same source.

4. *Determining* at a given probability level whether *the precision of two sets of measurements differs.*

5. *Comparing the means of more than two samples* to determine whether differences in the means are real or the result of random error. This process is known as *analysis of variance*.

6. Deciding whether to *reject or retain a result* that appears to be an outlier in a set of replicate measurements.

7A Confidence intervals

❑ In most quantitative chemical analyses, the true value of the mean, μ, cannot be determined because a huge number of measurements (approaching infinity) would be required.

❑ However, the interval surrounding the experimentally determined mean, x, can be determined within which the population mean μ is expected to lie with a certain degree of probability. This interval is known as the confidence interval. The limits of the interval are called confidence limits.

- For example, we might say that it is 99% probable that the true population mean for a set of potassium measurements lies in the interval 7.25 ± 0.15 % K. Thus, the probability that the mean lies in the interval from 7.10 to 7.40 % K is 99%.

❑ The size of the confidence interval, which is computed from the sample standard deviation, depends on how well the sample standard deviation, s, estimates the population standard deviation, σ.

Finding the confidence interval when σ is known or s is a good estimate of σ

❑ In each of a series of five normal error curves, the relative frequency is plotted as a function of the quantity z. The shaded areas in each plot lie between the values of -z and +z that are indicated to the left and right of the curves.

❑ The numbers within the shaded areas are the percentage of the total area under the curve that is included within these values of z.
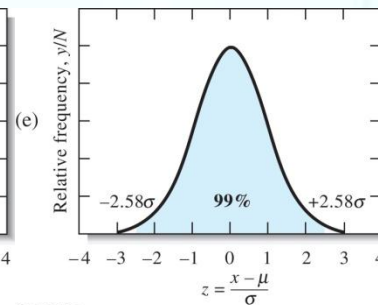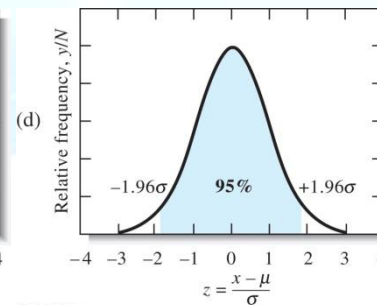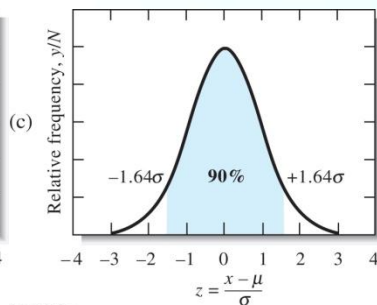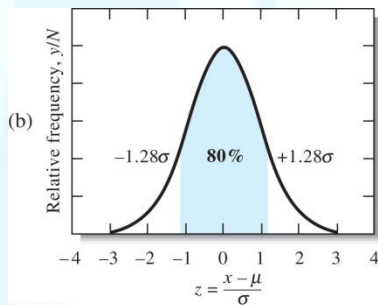
(a) 50% of the area under any Gaussian curve is located between -0.67σ and $+0.67\sigma$;

(b) 80% of the total area lies between -1.28σ and $+1.28\sigma$ and

(c) 90% of the total area lies between $-1.64\ \sigma$ and $+1.64\ \sigma$.

(d)) 95% of the total area lies between $-1.96\ \sigma$ and $+1.96\ \sigma$.

(e)) 99% of the total area lies between $-2.58\ \sigma$ and $+2.58\ \sigma$.
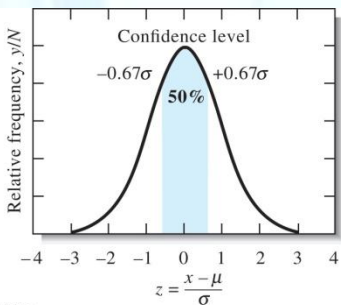
❑ The confidence level (CL) is the probability that the true mean lies within a certain interval and is often expressed as a percentage.

Figure 7-1c the **confidence level** is 90% and the **confidence interval** is from -1.64σ to $+1.64\sigma$

❑ The probability that a result is outside the confidence interval is often called the significance level.

❑ If we make a single measurement x from a distribution of known σ, we can say that the true mean should lie in the interval x ± zσ with a probability dependent on z.

$$CI\ for\mu = x \pm z\sigma$$

❑ However, we rarely estimate the true mean from a single measurement. Instead, we use the experimental mean $\bar{x}$ of N measurements as a better estimate of μ.

$$CI\ for\mu = \bar{x} \pm \frac{z\sigma}{\sqrt{N}}$$

we replace *x* with *x bar* and s with the standard error of the mean, σ/√N,

TABLE 7-1

Confidence Levels for Various Values of z

Confidence Level, %	z
50	0.67
68	1.00
80	1.28
90	1.64
95	1.96
95.4	2.00
99	2.58
99.7	3.00
99.9	3.29

TABLE 7-2

Size of Confidence Interval as a Function of the Number of Measurements Averaged

Number of Measurements Averaged	Relative Size of Confidence Interval
1	1.00
2	0.71
3	0.58
4	0.50
5	0.45
6	0.41
10	0.32

Values of z at various confidence levels are found in Table 7-1.

The relative size of the confidence interval as a function of N is shown in Table 7-2.

EXAMPLE 7-1

Determine the 80% and 95% confidence intervals for (a) the first entry (1108 mg/L glucose) in Example 6-2 (page 107) and (b) the mean value (1100.3 mg/L) for month 1 in the same example. Assume that in each part, $s = 19$ is a good estimate of σ.

Solution

(a) From Table 7-1, we see that $z = 1.28$ and 1.96 for the 80 and 95% confidence levels. Substituting into Equation 7-1 gives

$$80\% \text{ CI} = 1108 \pm 1.28 \times 19 = 1108 \pm 24.3 \text{ mg/L}$$

$$95\% \text{ CI} = 1108 \pm 1.96 \times 19 = 1108 \pm 37.2 \text{ mg/L}$$

From these calculations, we conclude that it is 80% probable that μ, the population mean (and, *in the absence of determinate error,* the true value), lies in the interval 1083.7 to 1132.3 mg/L glucose. Furthermore, the probability is 95% that μ lies in the interval between 1070.8 and 1145.2 mg/L.

(b) For the seven measurements,

$$80\% \text{ CI} = 1100.3 \pm \frac{1.28 \times 19}{\sqrt{7}} = 1100.3 \pm 9.2 \text{ mg/L}$$

$$95\% \text{ CI} = 1100.3 \pm \frac{1.96 \times 19}{\sqrt{7}} = 1100.3 \pm 14.1 \text{ mg/L}$$

Therefore, from the experimental mean ($\bar{x} = 1100.3$ mg/L), we conclude that there is an 80% chance that μ is located in the interval between 1091.1 and 1109.5 mg/L glucose and a 95% chance that it lies between 1086.2 and 1114.4 mg/L glucose. Note that the intervals are considerably smaller when we

EXAMPLE 7-2

How many replicate measurements in month 1 in Example 6-2 are needed to decrease the 95% confidence interval to 1100.3 $\pm$ 10.0 mg/L of glucose?

Solution

We want the term $\pm \dfrac{z\sigma}{\sqrt{N}}$ to equal ± 10.0 mg/L of glucose.

$$\frac{z\sigma}{\sqrt{N}} = \frac{1.96 \times 19}{\sqrt{N}} = 10.0$$

$$\sqrt{N} = \frac{1.96 \times 19}{10.0} = 3.724$$

$$N = (3.724)^2 = 13.9$$

We thus conclude that 14 measurements are needed to provide a slightly better than 95% chance that the population mean will lie within ± 10 mg/L of glucose of the experimental mean.

Finding the confidence interval when σ is unknown

❑ In case of limitations in time or in the amount of sample available, a single set of replicate measurements must provide not only a mean but also an estimate of precision.

❑ s calculated from a small set of data may be quite uncertain.

❑ Thus, confidence intervals are necessarily broader when we must use a small sample value of s as our estimate of σ.

❑ To account for the variability of s, we use the important statistical parameter t, which is defined in exactly the same way as *z* , except that *s* is substituted for σ.

❑ **For a single measurement with result *x*, we can define *t* as**
$$t = \frac{x - \mu}{s}$$

❑ **For the mean of N measurements**
$$t = \frac{\bar{x} - \mu}{s / \sqrt{N}}$$

The *t* statistic is often called **Student's t**. Student was the name used by W. S. Gossett because Guinness did not allow employees to publish their work, Gossett began to publish his results under the name "Student." He discovered the *t* distribution through mathematical and empirical studies with random numbers.

* Like z, *t* depends on the desired confidence level as well as on the number of degrees of freedom in the calculation of s.

t approaches z as the number of degrees of freedom becomes large.

* The confidence interval for the mean of N replicate measurements can be calculated from t as

$$CI \ for \quad \mu = \bar{x} \pm \frac{ts}{\sqrt{N}}$$

TABLE 7-3

Values of *t* for Various Levels of Probability

Degrees of Freedom	80%	90%	95%	99%	99.9%
1	3.08	6.31	12.7	63.7	637
2	1.89	2.92	4.30	9.92	31.6
3	1.64	2.35	3.18	5.84	12.9
4	1.53	2.13	2.78	4.60	8.61
5	1.48	2.02	2.57	4.03	6.87
6	1.44	1.94	2.45	3.71	5.96
7	1.42	1.90	2.36	3.50	5.41
8	1.40	1.86	2.31	3.36	5.04
9	1.38	1.83	2.26	3.25	4.78
10	1.37	1.81	2.23	3.17	4.59
15	1.34	1.75	2.13	2.95	4.07
20	1.32	1.73	2.09	2.84	3.85
40	1.30	1.68	2.02	2.70	3.55
60	1.30	1.67	2.00	2.62	3.46
∞	1.28	1.64	1.96	2.58	3.29

EXAMPLE 7-3

A clinical chemist obtained the following data for the alcohol content of a sample of blood: % C_2H_5OH: 0.084, 0.089, and 0.079. Calculate the 95% confidence interval for the mean assuming that (a) the three results obtained are the only indication of the precision of the method and that (b), from previous experience on hundreds of samples, we know that the standard deviation of the method $s = 0.005\%$ C_2H_5OH and is a good estimate of σ.

Solution

(a) $\Sigma x_i = 0.084 + 0.089 + 0.079 = 0.252$

   $\Sigma x_i^2 = 0.007056 + 0.007921 + 0.006241 = 0.021218$

   $$s = \sqrt{\frac{0.021218 - (0.252)^2/3}{3-1}} = 0.0050\%\ C_2H_5OH$$

   In this instance, $\bar{x} = 0.252/4 = 0.084$. Table 7-3 indicates that $t = 4.30$ for two degrees of freedom and the 95% confidence level. Thus, using Equation 7-5,

   $$95\%\ CI = \bar{x} \pm \frac{ts}{\sqrt{N}} = 0.084 \pm \frac{4.30 \times 0.0050}{\sqrt{3}}$$

   $$= 0.084 \pm 0.012\%\ C_2H_5OH$$

(b) Because $s = 0.0050\%$ is a good estimate of σ, we can use z and Equation 7-2

   $$95\%\ CI = \bar{x} \pm \frac{z\sigma}{\sqrt{N}} = 0.094 \pm \frac{1.96 \times 0.0050}{\sqrt{3}}$$

   $$= 0.084 \pm 0.006\%\ C_2H_5OH$$

Note that a sure knowledge of σ decreases the confidence interval by a significant amount even though s and σ are identical.

7B Statistical aids to hypothesis testing

❑ The hypothesis tests are used to determine if the results from these experiments support the model.

❑ If they do not support, the hypothesis is rejected.

❑ If agreement is found, the hypothetical model serves as the basis for further experiments.

❑ *Experimental results seldom agree exactly with those predicted from a theoretical model.*

❑ Statistical tests help determine whether a numerical difference is a result of a real difference (a systematic error) or a consequence of the random errors inevitable in all measurements.

❑ Tests of this kind use a *null hypothesis, which assumes that the numerical quantities being compared are the same*.

❑ We then use a probability distribution to calculate the probability that the observed differences are a result of random error.

➤Usually, if the observed difference is *greater than or equal* to the difference that would occur 5 times in 100 by random chance, (a significance level of 0.05), the null hypothesis is considered questionable, and the difference is judged to be significant.

➤Other significance levels, such as 0.01 (1%) or 0.001 (0.1%), may also be adopted, depending on the certainty desired in the judgment.

When expressed as a fraction, the significance level is often given the symbol α. The confidence level, CL, as a percentage is related to α by CL=(1- α) x 100%

Some examples of hypothesis tests that scientists often use include the comparison

(1) the mean of an experimental data set with what is believed to be the true value,

(2) the mean to a predicted or cutoff (threshold) value, and

(3) the means or the standard deviations from two or more sets of data.

The sections that follow consider some of the methods for making these

comparisons.

Comparing an Experimental Mean with a Known Value

❑ In many cases the mean of a data set needs to be compared with a known value.

In such cases, a statistical hypothesis test is used to draw conclusions about the population mean μ and its nearness to the known value, which we call μ_0.

There are two contradictory outcomes in any hypothesis test:

1. The null hypothesis H_0, states that $\mu = \mu_0$.
2. The alternative hypothesis H_a can be stated as:

   • reject the null hypothesis in favor of H_a if $\mu \neq \mu_0$.

   • OR if $\mu < \mu_0$ or $\mu > \mu_0$.

Large Sample z Test

If a large number of results are available so that s is a good estimate of σ, the z test is appropriate. The procedure that is used is summarized below:

1. State the null hypothesis: $H_0: \mu = \mu_0$

2. Form the test statistic:

$$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{N}}$$

3. State the alternative hypothesis H_a and determine the rejection region:

   * For $H_a: \mu \neq \mu_0$, reject H_0 if $z \geq z_{crit}$ or if $z \leq -z_{crit}$ (two-tailed test)

   * For $H_a: \mu > \mu_0$, reject H_0 if $z \geq z_{crit}$ (one-tailed test)

   * For $H_a: \mu < \mu_0$, reject H_0 if $z \leq -z_{crit}$ (one-tailed test)

Note that for H_a: $\mu \neq \mu_0$, we can reject for either a positive value of z or for a negative value of z that exceeds the critical value.

This is called a two-tailed test since rejection can occur for results in either tail of the distribution.

For the 95% confidence level, the probability that z exceeds z_{crit} is 0.025 in each tail or 0.05 total.
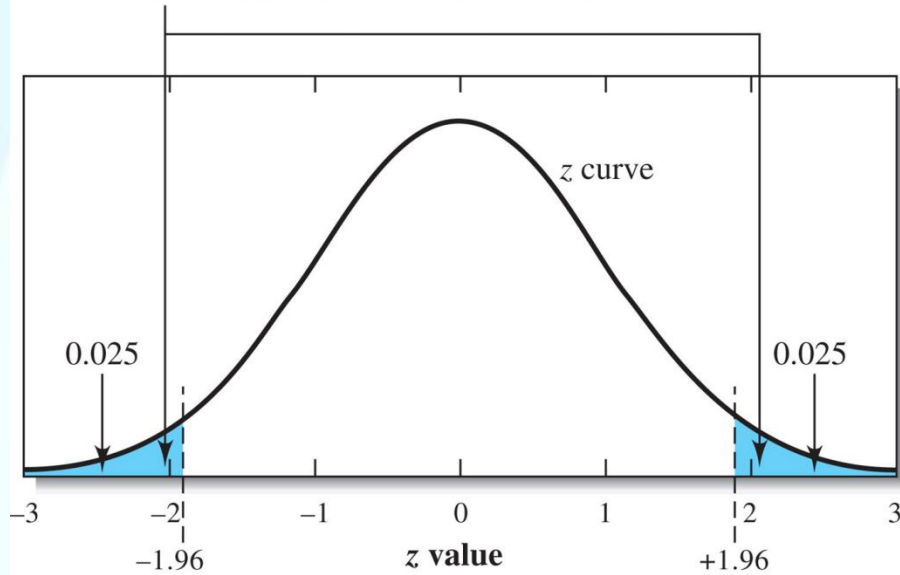
Hence, there is only a 5% probability that random error will lead to a value of $z \geq z_{crit}$ or $z \leq -z_{crit}$. The significance level overall is $\alpha = 0.05$.

If instead our alternative hypothesis is H_a: $\mu > \mu_0$, the test is said to be a one-tailed test. In this case, we can reject only when $z \geq z_{crit}$.

Figure 7-2   Rejection regions for the 95% confidence level.

(a)   Two-tailed test for H_a: $\mu \neq \mu_0$.
Note the critical value of z is 1.96.

(b) One-tailed test for H_a: $\mu > \mu_0$.
The critical value of z is 1.64 so that 95% of the area is to the left of z_{crit} and 5% of the area is to the right.

P value = 0.050 = sum of area in both tails

z curve

0.025

0.025

−3   −2   −1   0   1   2   3
    −1.96   *z* value   +1.96

(a)

© Cengage Learning. All Rights Reserved.

P value = 0.05 = area in upper tail

z curve

0.05

−3   −2   −1   0   1   2   3
            z value   1.64
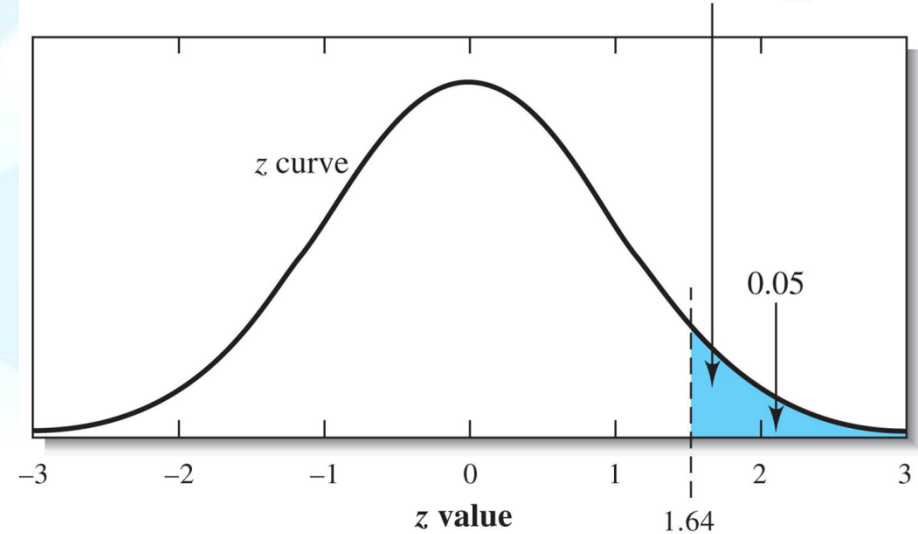
(b)

© Cengage Learning. All Rights Reserved.

z curve

0.05

−3   −2   −1   0   1   2   3

(c) One-tailed test for H_a: $\mu < \mu_0$.
The critical value is again 1.64 so that 5% of the area lies to the left of -z_{crit}.

EXAMPLE 7-4

A class of 30 students determined the activation energy of a chemical reaction to be 116 kJ mol^{-1} (mean value) with a standard deviation of 22 kJ mol^{-1}. Are the data in agreement with the literature value of 129 kJ mol^{-1} at (a) the 95% confidence level and (b) the 99% confidence level? Estimate the probability of obtaining a mean equal to the student value.

Solution

We have enough values that s should be a good estimate of σ. Accordingly, μ_0 is the literature value of 129 kJ mol^{-1} so that the null hypothesis is $\mu = 129$ kJ mol^{-1}. The alternative hypothesis is that $\mu \neq 129$ kJ mol^{-1}. This is thus a two-tailed test. From Table 7-1, $z_{crit} = 1.96$ for the 95% confidence level, and $z_{crit} = 2.58$ for the 99% confidence level. The test statistic is calculated as

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{N}} = \frac{116 - 129}{22/\sqrt{30}} = -3.27$$

Since $z \leq -1.96$, we reject the null hypothesis at the 95% confidence level. Note that, since $z \leq -2.58$, we also reject H_0 at the 99% confidence level. In order to estimate the probability of obtaining a mean value $\mu = 116$ kJ mol^{-1}, we must find the probability of obtaining a z value of 3.27. From Table 7-1, the probability of obtaining a z value this large because of random error is only about 0.2%. All of these results lead us to conclude that the student mean is actually different from the literature value and not just the result of random error.

Small Sample t Test

For a small number of results, we use a similar procedure to the *z* test except that the test statistic is the *t* statistic.

1. State the null hypothesis: H_0: $\mu = \mu_0$

2. Form the test statistic:

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{N}}$$

2. State the alternative hypothesis H_a and determine the rejection region:

- For H_a: $\mu \neq \mu_0$, reject H_0 if $t \geq t_{crit}$ or if $t \leq -t_{crit}$ (two-tailed test)

- For H_a: $\mu > \mu_0$, reject H_0 if $t \geq t_{crit}$ (one-tailed test)

- For H_a: $\mu < \mu_0$, reject H_0 if $t \leq -t_{crit}$ (one-tailed test)

As an illustration, consider the testing for systematic error in an analytical method. In this case, a sample of accurately known composition, such as a standard reference material, is analyzed. Determination of the analyte in the material gives an experimental mean that is an estimate of the population mean.

If the analytical method had no systematic error, or bias, random errors would give the frequency distribution shown by curve A. Method B has some systematic error so that x_B, which estimates μ_B, differs from the accepted value μ_0.

The bias is given by

$$Bias = \mu_B - \mu_0$$
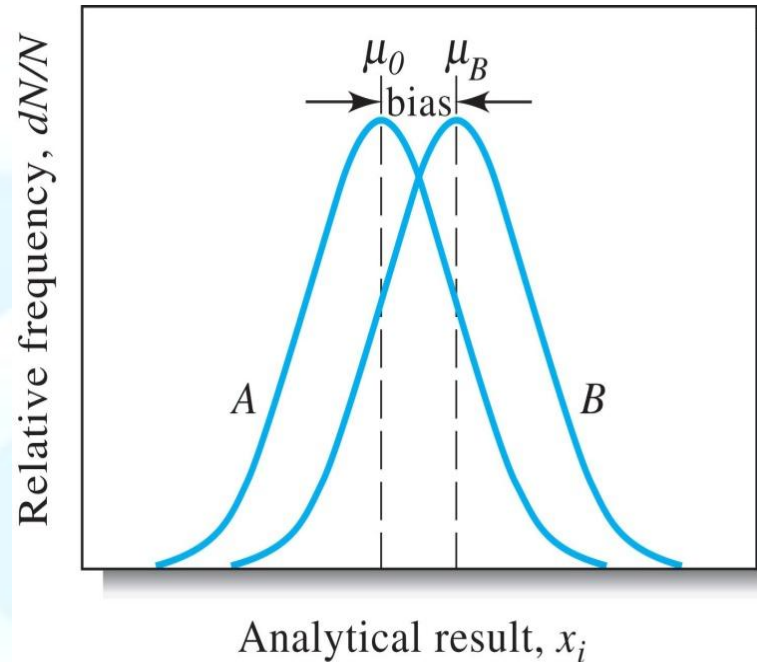

© Cengage Learning. All Rights Reserved.

Figure 7-3   Illustration of systematic error in an analytical method.

Curve A is the frequency distribution for the accepted value by a method without bias.

Curve B illustrates the frequency distribution of results by a method that could have a significant bias due to a systematic error.

EXAMPLE 7-5

A new procedure for the rapid determination of sulfur in kerosenes was tested on a sample known from its method of preparation to contain 0.123% S ($\mu_o = 0.123\%$ S). The results for % S were 0.112, 0.118, 0.115, and 0.119. Do the data indicate that there is a bias in the method at the 95% confidence level?

Solution

The null hypothesis is H_0: $\mu = 0.123\%$ S, and the alternative hypothesis is H_a: $\mu \neq 0.123\%$ S.

$$\Sigma x_i = 0.112 + 0.118 + 0.115 + 0.119 = 0.464$$

$$\bar{x} = 0.464/4 = 0.116\% \text{ S}$$

$$\Sigma x_i^2 = 0.012544 + 0.013924 + 0.013225 + 0.014161 = 0.53854$$

$$s = \sqrt{\frac{0.053854 - (0.464)^2/4}{4 - 1}} = \sqrt{\frac{0.000030}{3}} = 0.0032\% \text{ S}$$

The test statistic can now be calculated as

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{N}} = \frac{0.116 - 0.123}{0.032/\sqrt{4}} = -4.375$$

From Table 7-3, we find that the critical value of t for 3 degrees of freedom and the 95% confidence level is 3.18. Since $t \leq -3.18$, we conclude that there is a significant difference at the 95% confidence level and thus bias in the method. Note that, if we were to do this test at the 99% confidence level, $t_{crit} = 5.84$ (Table 7-3). Since $t = -4.375$ is greater than -5.84, we would accept the null hypothesis at the 99% confidence level and conclude there is no difference between the experimental and the accepted values. Note in this case that the outcome depends on the confidence level that is used. As we will see, choice of the confidence level depends on our willingness to accept an error in the outcome. The significance level (0.05 or 0.01) is the probability of making an error by rejecting the null hypothesis (see Section 7B-3).

Comparison of Two Experimental Means

* In many cases, it needs to be determined whether a difference in the means of two sets of data is real or the result of random error.

* In some cases, the results of chemical analyses are used to determine whether two materials are identical.

* In other cases, the results are used to determine whether two analytical methods give the same values or whether two analysts using the same methods obtain the same means.

* Data are often collected in pairs to eliminate one source of variability by focusing on the differences within each pair.

The t Test for Differences in Means

Differences in means can be computed with the z test, modified to take into account a comparison of two sets of data, if we have large numbers of measurements in both data sets.

Typically, both sets contain only a few results, and we must use the t test.

Let us assume that N_1 replicate analyses by analyst 1 yielded a mean value of x_1 and that N_2 analyses by analyst 2 obtained by the same method gave x_2.

According to the null hypothesis, we can write H_0: $\mu_1 = \mu_2$.

The alternative hypothesis is H_a: $\mu_1 \neq \mu_2$, and the test is a two-tailed test.

To get a better estimate of s than given by s_1 or s_2 (estimates of population standard deviation) alone, we use the pooled standard deviation.

The std deviation of the mean of analyst 1 is

$$s_{m1} = \frac{s_1}{\sqrt{N_1}}$$

The variance of the mean of analyst 1 is

$$s_{m1}^2 = s_1^2 \Big/ N_1$$

The variance of the mean of analyst 2 is

$$s_{m2} = \frac{s_2}{\sqrt{N_2}}$$

The variance of the difference s_d^2 between the means is $\quad s_d^2 = s_{m1}^2 + s_{m2}^2$

The std deviation of the difference between the means is

$$\frac{s_d}{\sqrt{N}} = \sqrt{\frac{s_1^2}{N_1} + \frac{s_2^2}{N_2}}$$

* Further assumption that the pooled standard deviation s_{pooled} is a better estimate of σ than s_1 or s_2.

$$\frac{s_d}{\sqrt{N}} = \sqrt{\frac{s^2_{pooled}}{N_1} + \frac{s^2_{pooled}}{N_2}} = s\sqrt{\frac{N_1 + N_2}{N_1 N_2}}$$

The test statistic t is now found from

$$t = \frac{\overline{x_1 - x_2}}{s_{pooled}\sqrt{\frac{N_1 + N_2}{N_1 N_2}}}$$

If there is good reason to believe that the standard deviations of the two data sets differ, the two-sample t test must be used.

However, the significance level for this t test is only approximate, and the number of degrees of freedom is more difficult to calculate.

EXAMPLE 7-6

In a forensic investigation, a glass containing red wine and an open bottle were analyzed for their alcohol content in order to determine whether the wine in the glass came from the bottle. On the basis of six analyses, the average content of the wine from the glass was established to be 12.61% ethanol. Four analyses of the wine from the bottle gave a mean of 12.53% alcohol. The 10 analyses yielded a pooled standard deviation $s_{pooled} = 0.070\%$. Do the data indicate a difference between the wines?

Solution

The null hypothesis is H_0: $\mu_1 = \mu_2$, and the alternative hypothesis is H_a: $\mu_1 \neq \mu_2$. We use Equation 7-7 to calculate the test statistic t.

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s_{pooled}\sqrt{\frac{N_1 + N_2}{N_1 N_2}}} = \frac{12.61 - 12.53}{0.07\sqrt{\frac{6 + 4}{6 \times 4}}} = 1.771$$

The critical value of t at the 95% confidence level for $10 - 2 = 8$ degrees of freedom is 2.31. Since $1.771 < 2.31$, we accept the null hypothesis at the 95% confidence level and conclude that there is no difference in the alcohol content of the wines. The probability of getting a t value of 1.771 can be calculated using the Excel function T.DIST.2T() and is T.DIST.2T(1.771,8) = 0.11. Thus, there is more than a 10% chance that we could get a value this large due to random error.

Paired Data

The paired t test uses the same type of procedure as the normal t test except that we analyze pairs of data and compute the differences, d.

The standard deviation is now the standard deviation of the mean difference. Our null hypothesis is H_0: $\mu_d - \Delta_0$

where Δ_0 is a specific value of the difference to be tested, often zero. The test statistic value is

$$t = \frac{\overline{d} - \Delta_0}{\dfrac{s_d}{\sqrt{N}}}$$

Where $\overline{d} = \dfrac{\sum d_i}{N}$

The alternative hypothesis would be $\mu_d \neq \Delta_0$, $\mu_d > \Delta_0$, $\mu_d < \Delta_0$

EXAMPLE 7-7

A new automated procedure for determining glucose in serum (Method A) is to be compared to the established method (Method B). Both methods are performed on serum from the same six patients in order to eliminate patient-to-patient variability. Do the following results confirm a difference in the two methods at the 95% confidence level?

	Patient 1	Patient 2	Patient 3	Patient 4	Patient 5	Patient 6
Method A glucose, mg/L	1044	720	845	800	957	650
Method B glucose, mg/L	1028	711	820	795	935	639
Difference, mg/L	16	9	25	5	22	11

Solution

Let us now test the appropriate hypotheses. If μ_d is the true average difference between the methods, we want to test the null hypothesis $H_0: \mu_d = 0$ and the alternative hypothesis, $H_a: \mu_d \neq 0$. The test statistic is

$$t = \frac{\bar{d} - 0}{s_d / \sqrt{N}}$$

From the table, $N = 6$, $\sum d_i = 16 + 9 + 25 + 5 + 22 + 11 = 88$, $\sum d_i^2 = 1592$, and $\bar{d} = 88/6 = 14.67$. The standard deviation of the difference s_d is given by Equation 6-5

$$s_d = \sqrt{\frac{1592 - \dfrac{(88)^2}{6}}{6 - 1}} = 7.76$$

and the t statistic is

$$t = \frac{14.67}{7.76/\sqrt{6}} = 4.628$$

From Table 7-3, the critical value of t is 2.57 for the 95% confidence level and 5 degrees of freedom. Since $t > t_{crit}$, we reject the null hypothesis and conclude that the two methods give different results.

Note that, if we merely average the results of Method A ($\bar{x}_A = 836.0$ mg/L) and the results of Method B ($\bar{x}_B = 821.3$ mg/L), the large patient-to-patient variation in glucose level gives us large values for s_A (146.5) and s_B (142.7). A comparison of means gives us a test t value of 0.176, and we would accept the null hypothesis. Hence, the large patient-to-patient variability masks the method differences that are of interest. Pairing allows us to focus on the differences.

Errors in Hypothesis Testing

1. A type I error occurs when H_0 is rejected although it is actually true. In some sciences, a type I error is called a false negative. The significance level α gives the frequency of rejecting H_0 when it is true.

2. A type II error occurs when H_0 is accepted and it is actually false. This is sometimes termed a false positive.

    The probability of a type II error is given the symbol β.

    Making α smaller (0.01 instead of 0.05) would appear to minimize the type I error rate. However, decreasing type I error rate increases type II error rate because they are inversely related to each other.

Comparison of Variances

A simple statistical test, called the F test, can be used to test this assumption under the provision that the populations follow the normal (Gaussian) distribution.

The F test is also used in comparing more than two means and in linear regression analysis.

The F test is based on the null hypothesis that the two population variances under consideration are equal, $H_0: \sigma^2_1 = \sigma_2^2$

The test statistic F, which is defined as the ratio of the two sample variances ($F = s^2_1/s^2_2$), is calculated and compared with the critical value of F at the desired significance level.

The null hypothesis is rejected if the test statistic differs too much from unity.

TABLE 7-4

Critical Values of F at the 5% Probability Level (95 % confidence level)

Degrees of Freedom (Denominator)	Degrees of Freedom (Numerator)								
	2	**3**	**4**	**5**	**6**	**10**	**12**	**20**	**∞**
2	19.00	19.16	19.25	19.30	19.33	19.40	19.41	19.45	19.50
3	9.55	9.28	9.12	9.01	8.94	8.79	8.74	8.66	8.53
4	6.94	6.59	6.39	6.26	6.16	5.96	5.91	5.80	5.63
5	5.79	5.41	5.19	5.05	4.95	4.74	4.68	4.56	4.36
6	5.14	4.76	4.53	4.39	4.28	4.06	4.00	3.87	3.67
10	4.10	3.71	3.48	3.33	3.22	2.98	2.91	2.77	2.54
12	3.89	3.49	3.26	3.11	3.00	2.75	2.69	2.54	2.30
20	3.49	3.10	2.87	2.71	2.60	2.35	2.28	2.12	1.84
∞	3.00	2.60	2.37	2.21	2.10	1.83	1.75	1.57	1.00

The F test can be used in either a one-tailed mode or in a two-tailed mode.

For a one-tailed test we test the alternative hypothesis that one variance is greater than the other.

Hence, the variance of the supposedly more precise procedure is placed in the denominator and that of the less precise procedure is placed in the numerator.

The alternative hypothesis is Ha: $\sigma_1^2 = \sigma_2^2$

For a two-tailed test, we test whether the variances are different.

The larger variance always appears in the numerator, which makes the outcome of the test less certain; thus, the uncertainty level of the F values doubles from 5% to 10%.

EXAMPLE 7-8

A standard method for the determination of the carbon monoxide (CO) level in gaseous mixtures is known from many hundreds of measurements to have a standard deviation of 0.21 ppm CO. A modification of the method yields a value for s of 0.15 ppm CO for a pooled data set with 12 degrees of freedom. A second modification, also based on 12 degrees of freedom, has a standard deviation of 0.12 ppm CO. Is either modification significantly more precise than the original?

Solution

We test the null hypothesis $H_0: \sigma_{std}^2 = \sigma_1^2$, where σ_{std}^2 is the variance of the standard method and σ_1^2 the variance of the modified method. The alternative hypothesis is one-tailed, $H_a: \sigma_1^2 < \sigma_{std}^2$. Because an improvement is claimed, the variances of the modifications are placed in the denominator. For the first modification,

$$F_1 = \frac{s_{std}^2}{s_1^2} = \frac{(0.21)^2}{(0.15)^2} = 1.96$$

and, for the second,

$$F_2 = \frac{(0.21)^2}{(0.12)^2} = 3.06$$

For the standard procedure, s_{std} is a good estimate of σ, and the number of degrees of freedom from the numerator can be taken as infinite. From Table 7-4, the critical value of F at the 95% confidence level is $F_{crit} = 2.30$.

Since F_1 is less than 2.30, we cannot reject the null hypothesis for the first modification and conclude that there is no improvement in precision. For the second modification, however, $F_2 > 2.30$. Hence, we reject the null hypothesis and conclude that the second modification does appear to give better precision at the 95% confidence level.

It is interesting to note that if we ask whether the precision of the second modification is significantly better than that of the first, the F test dictates that we must accept the null hypothesis, that is,

$$F = \frac{s_1^2}{s_2^2} = \frac{(0.15)^2}{(0.12)^2} = 1.56$$

In this case, $F_{crit} = 2.69$. Since $F < 2.69$, we must accept H_0 and conclude that the two methods give equivalent precision.

7C Analysis of variance

❑ The methods used for multiple comparisons fall under the general category of analysis of variance, or ANOVA.

❑ These methods use a single test to determine whether there is or is not a difference among the population means rather than pair wise comparisons as is done with the t test.

❑ In case of a potential difference, multiple comparison procedures can be used to identify which specific population means differ from the others.

Experimental design methods take advantage of ANOVA in planning and performing experiments.

ANOVA Concepts

In ANOVA procedures, *difference in several population means* is obtained by *comparing the variances*. For comparing I population means μ_1, μ_2,...μ_I, the null hypothesis H_0 is of the form $H_0 = \mu_1 = \mu_2 = \mu_3 = ... = \mu_I$

The alternative hypothesis H_a: at least two of the μ_i's are different.

The following are typical applications of ANOVA:

1. Is there a difference in the results of five analysts determining calcium by a volumetric method?

2. Will four different solvent compositions have differing influences on the yield of a chemical synthesis?

3. Are the results of manganese determinations by three different analytical methods different?

4. Is there any difference in the fluorescence of a complex ion at six different values of pH?

➢ In each of these situations, the populations have differing values of a common characteristic called a **factor** or sometimes a **treatment**.

➢ In the case of determining calcium by a volumetric method, the factor of interest is the **analyst**.

➢ The different values of the factor of interest are called **levels**.

➢ For the calcium example, there are five levels corresponding to analyst 1, analyst 2, analyst 3, analyst 4, and analyst 5.

➢ The comparisons among the various populations are made by measuring a **response** for each item sampled.

Factor	Levels	Response
Analyst	Analyst 1, analyst 2, analyst 3, analyst 4, analyst 5	Amount Ca, mmol
Solvent	Composition 1, composition 2, composition 3, composition 4	Synthesis yield, %
Analytical methods	Method 1, method 2, method 3	Concentration Mn, ppm
pH	pH 1, pH 2, pH 3, pH 4, pH 5, pH 6	Fluorescence intensity
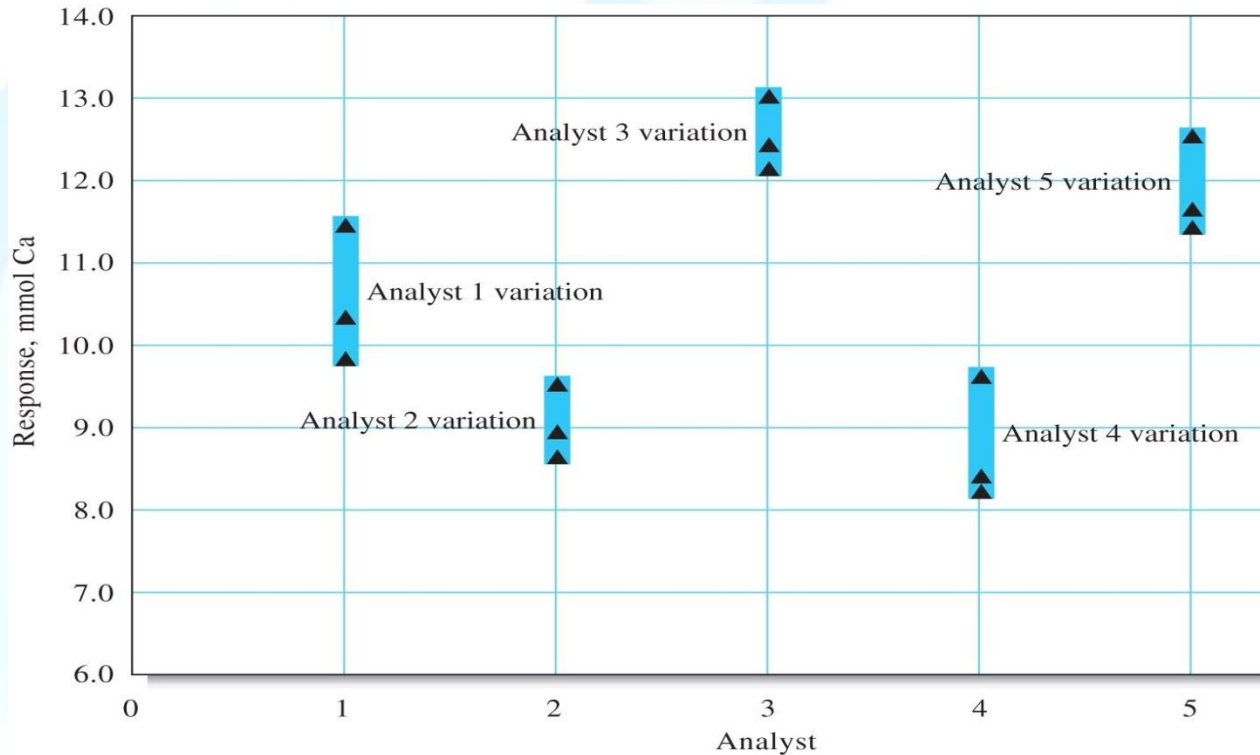
Figure 7-4 Pictorial of the results from the ANOVA study of the determination of calcium by five analysts

❑ The type of ANOVA shown is a single-factor, or one-way, ANOVA.

❑ Each analyst does the determination in triplicate.

❑ *Analyst* is considered a *factor*, while analyst 1, analyst 2, analyst 3, analyst 4, and analyst 5 are *levels of the factor*.
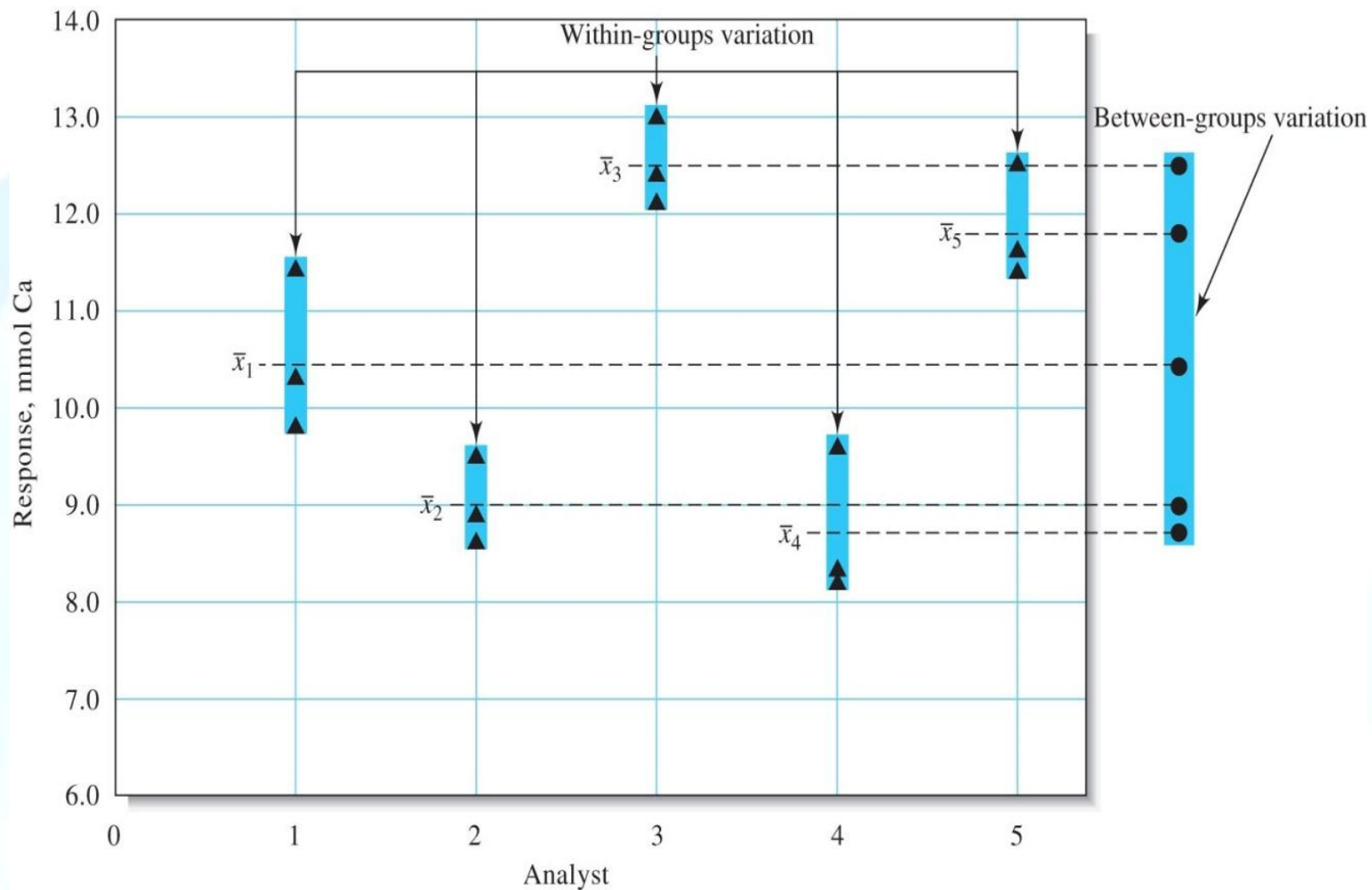
Figure 7-5 This is a pictorial representation of the ANOVA principle.

The results of each analyst are considered a group. The triangles (m) represent individual results, and the circles (d) represent the means.

The variation between the group means is compared to that within groups.

Single-Factor ANOVA

Several quantities are important for testing the null hypothesis.

One of these is the grand average, which is the average of all the data.

$$\overline{\overline{x}} = \left(\frac{N_1}{N}\right)\overline{x}_1 + \left(\frac{N_2}{N}\right)\overline{x}_2 + \left(\frac{N_3}{N}\right)\overline{x}_3 + ... + \left(\frac{N_1}{N}\right)\overline{x}_I$$

where N1 is the number of measurements in group 1, N2 is the number in group 2, and so on.

The grand average can also be found by summing all the data values and dividing by the total number of measurements N.

To calculate the variance ratio needed in the F test, it is necessary to obtain several other quantities called sums of squares:

1. The sum of the squares due to the factor, SSF, is

$$SSF = N_1(\overline{x}_1 - \overline{\overline{x}})^2 + N_2(\overline{x}_2 - \overline{\overline{x}})^2 + N_3(\overline{x}_3 - \overline{\overline{x}})^2 + ... + N_I(\overline{x}_I - \overline{\overline{x}})^2$$

2. The sum of the squares due to error ,SSE, is

$$SSE = \sum_{j=1}^{N_1}(x_{1j} - \overline{x}_1)^2 + \sum_{j=1}^{N_2}(x_{2j} - \overline{x}_2)^2 + \sum_{j=1}^{N_3}(x_{3j} - \overline{x}_3)^2 + ... + \sum_{j=1}^{N_I}(x_{ij} - \overline{x}_I)^2$$

3. These two sums of squares are used to obtain the between-groups variation and the within-groups variation. The error sum of the squares is related to the individual group variances by

$$SSE = (N_1 - 1)s_1^2 + (N_2 - 1)s_2^2 + (N_3 - 1)s_3^2 + ... + (N_I - 1)s_I^2$$

4. The total sum of the squares SST is obtained as the sum of SSF and SSE:

SST = SSF + SSE

* The total sum of the squares SST has $N - 1$ degrees of freedom.

* Just as SST is the sum of SSF and SSE, the total number degrees of freedom $N - 1$ can be decomposed into degrees of freedom associated with SSF and SSE.

* Since there are I groups being compared, SSF has $I - 1$ degrees of freedom. This leaves $N - I$ degrees of freedom for SSE. Or,

$$SST = SSF + SSE$$

$$(N - 1) = (I - 1) + (N - I)$$

* Dividing the sums of squares by their corresponding degrees of freedom, we can obtain quantities that are estimates of the between-groups and within-groups variations called mean square values.

* Mean square due to factor levels: $MSF = SSF/I – 1$

* Mean square due to error: $MSE = SSE/N - 1$

* The quantity MSE is an estimate of the variance due to error (σ^2_E), while MSF is an estimate of the error variance plus the between-groups variance ($\sigma^2_E + \sigma^2_F$)

* If the factor has little effect, the between-groups variance should be small compared to the error variance.

* Thus, the two mean squares should be nearly identical under these circumstances. If the factor effect is significant, MSF is greater than MSE.

* The test statistic is the F value, calculated as F = MSF/MSE

Source of Variation	Sum of Squares (SS)	Degrees of Freedom (df)	Mean Square (MS)	Mean Square Estimates	F
Between groups (factor effect)	SSF	$I-1$	$MSF = \dfrac{SSF}{I-1}$	$\sigma^2_E + \sigma^2_F$	$\dfrac{MSF}{MSE}$
Within groups (error)	SSE	$N-I$	$MSE = \dfrac{SSE}{N-I}$	σ^2_E	
Total	SST	$N-1$			

Example 7-9

Five analysts determined calcium by a volumetric method and obtained the amounts (in mmol Ca) shown in the table below. Do the means differ significantly at the 95% confidence level?

Trial No.	Analyst 1	Analyst 2	Analyst 3	Analyst 4	Analyst 5
1	10.3	9.5	12.1	9.6	11.6
2	9.8	8.6	13.0	8.3	12.5
3	11.4	8.9	12.4	8.2	11.4

Solution

First, we obtain the means and standard deviations for each analyst. The mean for analyst 1 is $\bar{x}_1 = (10.3 + 9.8 + 11.4)/3 = 10.5$ mmol Ca. The remaining means are obtained in the same manner: $\bar{x}_2 = 9.0$ mmol Ca, $\bar{x}_3 = 12.5$ mmol Ca, $\bar{x}_4 = 8.7$ mmol Ca, $\bar{x}_5 = 11.833$ mmol Ca. The standard deviations are obtained as described in Section 6B-3. These results are summarized, as follows:

	Analyst 1	Analyst 2	Analyst 3	Analyst 4	Analyst 5
Mean	10.5	9.0	12.5	8.7	11.833
Standard Dev.	0.818535	0.458258	0.458258	0.781025	0.585947

The grand mean is found from Equation 7-8, where $N_1 = N_2 = N_3 = N_4 = N_5 = 3$ and $N = 15$:

$$\bar{\bar{x}} = \frac{3}{15}(\bar{x}_1 + \bar{x}_2 + \bar{x}_3 + \bar{x}_4 + \bar{x}_5) = 10.507 \text{ mmol Ca}$$

$$\bar{\bar{x}} = \left(\frac{N_1}{N}\right)\bar{x}_1 + \left(\frac{N_2}{N}\right)\bar{x}_2 + \left(\frac{N_3}{N}\right)\bar{x}_3 + \ldots + \left(\frac{N_I}{N}\right)\bar{x}_I$$

The between-groups sum of the squares is found from Equation 7-9:

$$\begin{aligned} SSF = {}& 3(10.5 - 10.507)^2 + 3(9.0 - 10.507)^2 + 3(12.5 - 10.507)^2 \\ & + 3(8.7 - 10.507)^2 + 3(11.833 - 10.507)^2 \\ = {}& 33.80267 \end{aligned}$$

$$SSF = N_1(\bar{x}_1 - \bar{\bar{x}})^2 + N_2(\bar{x}_2 - \bar{\bar{x}})^2 + N_3(\bar{x}_3 - \bar{\bar{x}})^2 + \ldots + N_I(\bar{x}_I - \bar{\bar{x}})^2$$

Note that SSF has associated with it $(5 - 1) = 4$ degrees of freedom.
The error sum of squares is easiest to find from the standard deviations and Equation 7-11:

$$\begin{aligned} SSE = {}& 2(0.818535)^2 + 2(0.458258)^2 + 2(0.458258)^2 + \\ & 2(0.781025)^2 + 2(0.585947)^2 \\ = {}& 4.086667 \end{aligned}$$

$$SSE = \sum_{j=1}^{N_1}(x_{1j} - \bar{x}_1)^2 + \sum_{j=1}^{N_2}(x_{2j} - \bar{x}_2)^2 + \sum_{j=1}^{N_3}(x_{3j} - \bar{x}_3)^2 + \ldots + \sum_{j=1}^{N_I}(x_{ij} - \bar{x}_I)^2$$

The error sum of the squares has $(15 - 5) = 10$ degrees of freedom.
We can now calculate the mean square values, MSF and MSE, from Equations 7-13 and 7-14:

$$MSF = \frac{33.80267}{4} = 8.450667$$

MSF=SSF/I-1

$$MSE = \frac{4.086667}{10} = 0.408667$$

MSE = SSE/N - 1

The F value obtained from Equation 7-15 is

F = MSF/MSE

$$F = \frac{8.450667}{0.408667} = 20.68$$

The *F* value obtained from Equation 7-15 is

$$F = \frac{8.450667}{0.408667} = 20.68$$

From Table 7-4, the *F* table, the critical value of *F* at the 95% confidence level for 4 and 10 degrees of freedom is 3.48. Since *F* exceeds 3.48, we reject H_0 at the 95% confidence level and conclude that there is a significant difference among the analysts. The ANOVA table is:

Source of Variation	Sum of Squares (SS)	Degrees of Freedom (df)	Mean Square (MS)	F
Between groups	33.80267	4	8.450667	20.68
Within groups	4.086667	10	0.408667	
Total	37.88933	14		

Determining Which Results Differ

❑ There are several methods to determine which means are significantly different.

❑ The *least significant difference method* is the simplest method in which a difference is calculated that is judged to be the smallest difference that is significant.

❑ The difference between each pair of means is then compared to the least significant difference to determine which means are different.

❑ For an equal number of replicates, N_g, in each group, the least significant difference LSD is calculated as follows:

$$LSD = t\sqrt{\frac{2 \times MSE}{N_g}}$$

where MSE is the mean square for error and the value of t has N – I degrees of freedom.

EXAMPLE 7-10

For the results of Example 7-9, determine which analysts differ from each other at the 95% confidence level.

Solution

First, we arrange the means in increasing order: 8.7, 9.0, 10.5, 11.833, and 12.5. Each analyst did three repetitions, and so we can use Equation 7-16. From Table 7-3, we obtain a t value of 2.23 for the 95% confidence level and 10 degrees of freedom. Application of Equation 7-16 gives us

$$\text{LSD} = 2.23\sqrt{\frac{2 \times 0.408667}{3}} = 1.16 \qquad LSD = t\sqrt{\frac{2 \times MSE}{N_g}}$$

We now calculate the differences in means and compare them to 1.16. For the various pairs:

$$\bar{x}_{largest} - \bar{x}_{smallest} = 12.5 - 8.7 = 3.8 \qquad \text{(a significant difference).}$$
$$\bar{x}_{2ndlargest} - \bar{x}_{smallest} = 11.833 - 8.7 = 3.133 \quad \text{(significant).}$$
$$\bar{x}_{3rdlargest} - \bar{x}_{smallest} = 10.5 - 8.7 = 1.8 \qquad \text{(significant).}$$
$$\bar{x}_{4thlargest} - \bar{x}_{smallest} = 9.0 - 8.7 = 0.3 \qquad \text{(no significant difference).}$$

We then continue to test each pair to determine which are different. From these calculations, we conclude that analysts 3, 5, and 1 differ from analyst 4; analysts 3, 5, and 1 differ from analyst 2; analysts 3 and 5 differ from analyst 1; and analyst 3 differs from analyst 5.

7D Detection of gross errors

➢An outlier is a result that is quite different from the others in the data set.

➢It is important to develop a criterion to decide whether to retain or reject the outlying data point.

➢The choice of criterion for the rejection of a suspected result has its perils. If the standard is too strict so that it is quite difficult to reject a questionable result, there is a risk of retaining a spurious value that has an inordinate effect on the mean.

➢If we set a lenient limit and make the rejection of a result easy, we are likely to discard a value that rightfully belongs in the set, thus introducing bias to the data.

➢While there is no universal rule to settle the question of retention or rejection, *the Q test is generally acknowledged to be an appropriate method for making the decision.*

The Q Test

➢ The Q test is a simple, widely used statistical test for deciding whether a suspected result should be retained or rejected.

➢ In this test, the absolute value of the difference between the questionable result x_q and its nearest neighbor x_n is divided by the spread w of the entire set to give the quantity Q:

$$Q = \frac{|x_q - x_n|}{w}$$
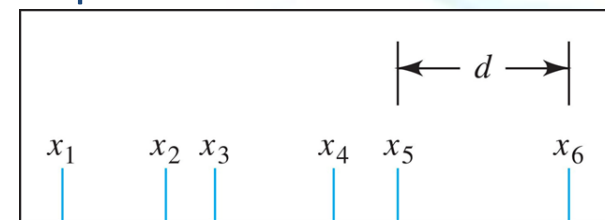
$d = x_6 - x_5$
$w = x_6 - x_1$
$Q = d/w$
If $Q > Q_{crit}$, reject x_6

TABLE 7-5

Critical Values for the Rejection Quotient, Q^*

Number of Observations	Q_{crit} (Reject if $Q > Q_{crit}$)		
	90% Confidence	95% Confidence	99% Confidence
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

The Q test for outliers

EXAMPLE 7-11

The analysis of a city drinking water for arsenic yielded values of 5.60. 5.64, 5.70, 5.69, and 5.81 ppm. The last value appears anomalous; should it be rejected at the 95% confidence level?

Solution

The difference between 5.81 and 5.70 is 0.11 ppm. The spread $(5.81 - 5.60)$ is 0.21 ppm. Thus,

$$Q = \frac{0.11}{0.21} = 0.52$$

For five measurements, Q_{crit} at the 95% confidence level is 0.71. Because $0.52 < 0.71$, we must retain the outlier at the 95% confidence level.

Other Statistical Tests

❑ Statistical rules should be used with extreme caution when applied to samples containing only a few values.

❑ The only valid reason *for rejecting* a result from a small set of data is the *sure knowledge that an error was made* in the measurement process. Else, a cautious approach to rejection of an outlier is wise.

Recommendations for Treating Outliers

1. Reexamine carefully all data relating to the outlying result to see if a gross error could have affected its value. This recommendation demands a properly kept laboratory notebook containing careful notations of all observations.

2. If possible, estimate the precision that can be reasonably expected from the procedure to be sure that the outlying result actually is questionable.

3. Repeat the analysis if sufficient sample and time are available. Agreement between the newly acquired data and those of the original set that appear to be valid will lend weight to the notion that the outlying result should be rejected.

4. If more data cannot be secured, apply the Q test to the existing set to see if the doubtful result should be retained or rejected on statistical grounds.

5. If the Q test indicates retention, consider reporting the median of the set rather than the mean. The median has the great virtue of allowing inclusion of all data in a set without undue influence from an outlying value. In addition, the median of a normally distributed set containing three measurements provides a better estimate of the correct value than the mean of the set after the outlying value has been discarded.

TABLE 7-1
Confidence Levels for Various Values of z

Confidence Level, %	z
50	0.67
68	1.00
80	1.28
90	1.64
95	1.96
95.4	2.00
99	2.58
99.7	3.00
99.9	3.29

TABLE 7-2
Size of Confidence Interval as a Function of the Number of Measurements Averaged

Number of Measurements Averaged	Relative Size of Confidence Interval
1	1.00
2	0.71
3	0.58
4	0.50
5	0.45
6	0.41
10	0.32

TABLE 7-3
Values of t for Various Levels of Probability

Degrees of Freedom	80%	90%	95%	99%	99.9%
1	3.08	6.31	12.7	63.7	637
2	1.89	2.92	4.30	9.92	31.6
3	1.64	2.35	3.18	5.84	12.9
4	1.53	2.13	2.78	4.60	8.61
5	1.48	2.02	2.57	4.03	6.87
6	1.44	1.94	2.45	3.71	5.96
7	1.42	1.90	2.36	3.50	5.41
8	1.40	1.86	2.31	3.36	5.04
9	1.38	1.83	2.26	3.25	4.78
10	1.37	1.81	2.23	3.17	4.59
15	1.34	1.75	2.13	2.95	4.07
20	1.32	1.73	2.09	2.84	3.85
40	1.30	1.68	2.02	2.70	3.55
60	1.30	1.67	2.00	2.62	3.46
∞	1.28	1.64	1.96	2.58	3.29

TABLE 7-4
Critical Values of F at the 5% Probability Level (95 % confidence level)

Degrees of Freedom (Denominator)	Degrees of Freedom (Numerator)								
	2	3	4	5	6	10	12	20	∞
2	19.00	19.16	19.25	19.30	19.33	19.40	19.41	19.45	19.50
3	9.55	9.28	9.12	9.01	8.94	8.79	8.74	8.66	8.53
4	6.94	6.59	6.39	6.26	6.16	5.96	5.91	5.80	5.63
5	5.79	5.41	5.19	5.05	4.95	4.74	4.68	4.56	4.36
6	5.14	4.76	4.53	4.39	4.28	4.06	4.00	3.87	3.67
10	4.10	3.71	3.48	3.33	3.22	2.98	2.91	2.77	2.54
12	3.89	3.49	3.26	3.11	3.00	2.75	2.69	2.54	2.30
20	3.49	3.10	2.87	2.71	2.60	2.35	2.28	2.12	1.84
∞	3.00	2.60	2.37	2.21	2.10	1.83	1.75	1.57	1.00

TABLE 7-5
Critical Values for the Rejection Quotient, Q^*

Number of Observations	Q_{crit} (Reject if $Q > Q_{crit}$)		
	90% Confidence	95% Confidence	99% Confidence
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

TABLE 6-4
Error Propagation in Arithmetic Calculations

Type of Calculation	Example*	Standard Deviation of $y^{\dagger}$	
Addition or subtraction	$y = a + b - c$	$s_y = \sqrt{s_a^2 + s_b^2 + s_c^2}$	(1)
Multiplication or division	$y = a \times b/c$	$\frac{s_y}{y} = \sqrt{\left(\frac{s_a}{a}\right)^2 + \left(\frac{s_b}{b}\right)^2 + \left(\frac{s_c}{c}\right)^2}$	(2)
Exponentiation	$y = a^x$	$\frac{s_y}{y} = x\left(\frac{s_a}{a}\right)$	(3)
Logarithm	$y = \log_{10} a$	$s_y = 0.434\frac{s_a}{a}$	(4)
Antilogarithm	$y = \text{antilog}_{10} a$	$\frac{s_y}{y} = 2.303\, s_a$	(5)

Anova table

Source of Variation	Sum of Squares (SS)	Degrees of Freedom (df)	Mean Square (MS)	Mean Square Estimates	F
Between groups (factor effect)	SSF	$I-1$	$MSF = \dfrac{SSF}{I-1}$	$\sigma_E^2 + \sigma_F^2$	$\dfrac{MSF}{MSE}$
Within groups (error)	SSE	$N-I$	$MSE = \dfrac{SSE}{N-I}$	σ_E^2	
Total	SST	$N-1$			

Suggested Problems to be solved from the end of the chapter 7:

2, 3, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32