

Chapter 4

Most of the problems in this chapter require the calculation of a data set's basic statistical characteristics, such as its mean, median, range, standard deviation, or variance. Although equations for these calculations are highlighted in the solution to the first problem, for the remaining problems, both here and elsewhere in this text, such values simply are provided. Be sure you have access to a scientific calculator, a spreadsheet program, such as Excel, or a statistical software program, such as R, and that you know how to use it to complete these most basic of statistical calculations.

1. The mean is obtained by adding together the mass of each quarter and dividing by the number of quarters; thus

$$\begin{aligned}\bar{X} &= \frac{\sum_{i=1}^n X_i}{n} \\ &= \frac{5.683 + 5.549 + \dots + 5.554 + 5.632}{12} \\ &= 5.583 \text{ g}\end{aligned}$$

To find the median, we first order the data from the smallest mass to the largest mass

$$\begin{array}{cccccc} 5.536 & 5.539 & 5.548 & 5.549 & 5.551 & 5.552 \\ 5.552 & 5.554 & 5.560 & 5.632 & 5.683 & 5.684 \end{array}$$

and then, because there is an even number of samples, take the average of the $n/2$ and the $n/2+1$ values; thus

$$\tilde{X} = \frac{X_6 + X_7}{2} = \frac{5.552 + 5.552}{2} = 5.552 \text{ g}$$

The range is the difference between the largest mass and the smallest mass; thus

$$w = X_{\text{largest}} - X_{\text{smallest}} = 5.684 - 5.536 = 0.148 \text{ g}$$

The standard deviation for the data is

$$\begin{aligned}s &= \sqrt{\frac{\sum_i^n (X_i - \bar{X})^2}{n - 1}} \\ &= \sqrt{\frac{(5.683 - 5.583)^2 + \dots + (5.632 - 5.583)^2}{12 - 1}} \\ &= 0.056 \text{ g}\end{aligned}$$

The variance is the square of the standard deviation; thus

$$s^2 = (0.056)^2 = 3.1 \times 10^{-3}$$

As a reminder, if we have an odd number of data points, then the median is the middle data point in the rank-ordered data set, or, more generally, the value of the $(n+1)/2$ data point in the rank-ordered data set where n is the number of values in the data set.

The variance in this case has units of g^2 , which is correct but not particularly informative in an intuitive sense; for this reason, we rarely attach a unit to the variance. See Rumsy, D. J. *Journal of Statistics Education* **2009**, 17(3) for an interesting argument that the variance should be excluded for summary statistics.

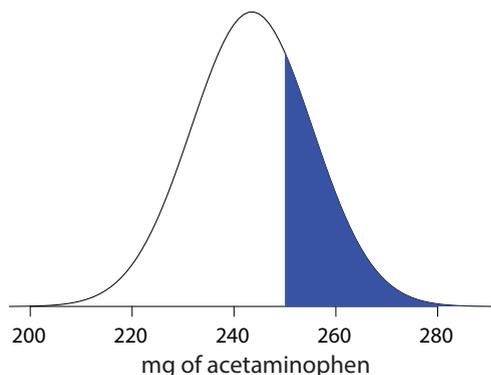


Figure SM4.1 Normal distribution curve for Problem 4.2 given a population with a mean of 243.5 mg and a standard deviation of 11.9 mg; the area in blue is the probability that a random sample has more than 250.0 mg of acetaminophen.

2. (a) The values are as follows:

mean: 243.5 mg

median: 243.4 mg

range: 37.4 mg

standard deviation: 11.9 mg

variance: 141

(b) We are interested in the area under a normal distribution curve that lies to the right of 250 mg, as shown in Figure SM4.1. Because this limit is greater than the mean, we need only calculate the deviation, z , and look up the corresponding probability in Appendix 3; thus,

$$z = \frac{X - \mu}{\sigma} = \frac{250 - 243.5}{11.9} = 0.546$$

From Appendix 3 we see that the probability is 0.2946 when z is 0.54 and 0.2912 when z is 0.55. Interpolating between these values gives the probability for a z of 0.546 as

$$0.2946 - 0.6(0.2946 - 0.2912) = 0.2926$$

Based on our experimental mean and standard deviation, we expect that 29.3% of the tablets will contain more than 250 mg of acetaminophen.

3. (a) The means and the standard deviations for each of the nominal dosages are as follows:

nominal dosage	mean	std. dev.
100-mg	95.56	2.16
60-mg	55.47	2.11
30-mg	26.85	1.64
10-mg	8.99	0.14

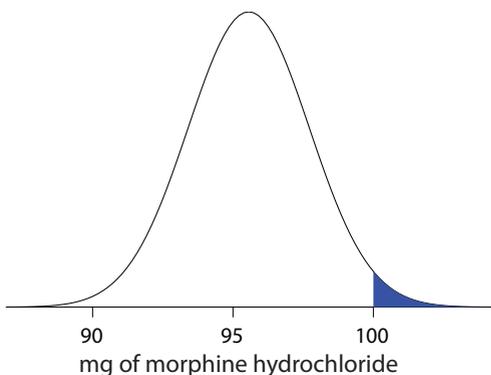


Figure SM4.2 Normal distribution curve for Problem 4.3 given a population with a mean of 95.56 mg and a standard deviation of 2.16 mg; the area in blue is the probability that a random sample has more than 100.0 mg of morphine hydrochloride.

(b) We are interested in the area under a normal distribution curve that lies to the right of each tablet's nominal dosage, as shown in Figure SM4.2 for tablets with a nominal dosage of 100-mg. Because the nominal dosage is greater than the mean, we need only calculate the deviation, z , for each tablet and look up the corresponding probability in Appendix 3. Using the 100-mg tablet as an example, the deviation is

$$z = \frac{X - \mu}{\sigma} = \frac{100 - 95.56}{2.16} = 2.06$$

for which the probability is 0.0197; thus, we expect that 1.97% of tablets drawn at random from this source will exceed the nominal dosage. The table below summarizes results for all four sources of tablets.

nominal dosage	z	% exceeding nominal dosage
100-mg	2.06	1.97
60-mg	2.15	1.58
30-mg	1.92	2.74
10-mg	7.21	—

For tablets with a 10-mg nominal dosage, the value of z is sufficiently large that effectively no tablet is expected to exceed the nominal dosage.

4. The mean and the standard deviation for the eight spike recoveries are 99.5% and 6.3%, respectively. As shown in Figure SM4.3, to find the expected percentage of spike recoveries in the range 85%–115%, we find the percentage of recoveries that exceed the upper limit by calculating z and using Appendix 3 to find the corresponding probability

$$z = \frac{X - \mu}{\sigma} = \frac{115 - 99.5}{6.3} = 2.46 \text{ or } 0.695\%$$

and the percentage of recoveries that fall below the lower limit

$$z = \frac{X - \mu}{\sigma} = \frac{85 - 99.5}{6.3} = -2.30 \text{ or } 1.07\%$$

Subtracting these two values from 100% gives the expected probability of spike recoveries between 85%–115% as

$$100\% - 0.695\% - 1.07\% = 98.2\%$$

5. (a) Substituting known values for the mass, the gas constant, the temperature, the pressure, and the volume gives the compound's formula weight as

$$FW = \frac{(0.118 \text{ g})(0.082056 \frac{\text{L} \cdot \text{atm}}{\text{mol} \cdot \text{K}})(298.2 \text{ K})}{(0.724 \text{ atm})(0.250 \text{ L})} = 16.0 \text{ g/mol}$$

To estimate the uncertainty in the formula weight, we use a propagation of uncertainty. The relative uncertainty in the formula weight is

$$\frac{u_{FW}}{FW} = \sqrt{\left(\frac{0.002}{0.118}\right)^2 + \left(\frac{0.000001}{0.082056}\right)^2 + \left(\frac{0.1}{298.2}\right)^2 + \left(\frac{0.005}{0.724}\right)^2 + \left(\frac{0.005}{0.250}\right)^2} = 0.0271$$

which makes the absolute uncertainty in the formula weight

$$u_{FW} = 0.0271 \times 16.0 \text{ g/mol} = 0.43 \text{ g/mol}$$

The formula weight, therefore, is $16.0 \pm 0.4 \text{ g/mol}$.

(b) To improve the uncertainty in the formula weight we need to identify the variables that have the greatest individual uncertainty. The relative uncertainties for the five measurements are

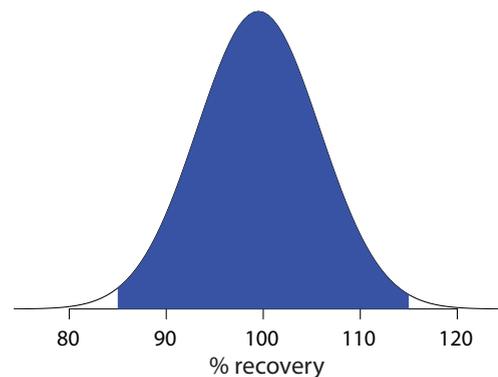


Figure SM4.3 Normal distribution curve for Problem 4.4 given a population with a mean of 99.5% and a standard deviation of 6.3%; the area in blue is the probability that a spike recovery is between 85% and 115%.

$$\text{mass: } 0.002/0.118 = 0.017$$

$$\text{gas constant: } 0.000001/0.082056 = 1.22 \times 10^{-5}$$

$$\text{temperature: } 0.1/298.2 = 3.4 \times 10^{-4}$$

$$\text{pressure: } 0.005/0.724 = 0.007$$

$$\text{volume: } 0.005/0.250 = 0.020$$

Of these variables, the two with the largest relative uncertainty are the mass in grams and the volume in liters; these are the measurements where an improvement in uncertainty has the greatest impact on the formula weight's uncertainty.

6. (a) The concentration of Mn^{2+} in the final solution is

$$\frac{0.250 \text{ g}}{0.1000 \text{ L}} \times \frac{1000 \text{ mg}}{\text{g}} \times \frac{10.00 \text{ mL}}{500.0 \text{ mL}} = 50.0 \text{ mg/L}$$

To estimate the uncertainty in concentration, we complete a propagation of uncertainty. The uncertainties in the volumes are taken from Table 4.2; to find the uncertainty in the mass, however, we must account for the need to tare the balance. Taking the uncertainty in any single determination of mass as ± 1 mg, the absolute uncertainty in mass is

$$u_{\text{mass}} = \sqrt{((0.001)^2 + (0.001)^2)} = 0.0014 \text{ g}$$

The relative uncertainty in the concentration of Mn^{2+} , therefore, is

$$\frac{u_C}{C} = \sqrt{\left(\frac{0.0014}{0.250}\right)^2 + \left(\frac{0.00008}{0.1000}\right)^2 + \left(\frac{0.02}{10.00}\right)^2 + \left(\frac{0.20}{500.0}\right)^2} = 0.00601$$

which makes the relative uncertainty in the concentration

$$u_C = 0.00601 \times (50.0 \text{ ppm}) = 0.3 \text{ ppm}$$

The concentration, therefore, is 50.0 ± 0.3 ppm.

(b) No, we cannot improve the concentration's uncertainty by measuring the HNO_3 with a pipet instead of a graduated cylinder. As we can see from part (a), the volume of HNO_3 does not affect our calculation of either the concentration of Mn^{2+} or its uncertainty.

7. The weight of the sample taken is the difference between the container's original weight and its final weight; thus, the mass is

$$\text{mass} = 23.5811 \text{ g} - 22.1559 \text{ g} = 1.4252 \text{ g}$$

and its absolute uncertainty is

$$u_{\text{mass}} = \sqrt{(0.0001)^2 + (0.0001)^2} = 0.00014 \text{ g}$$

The molarity of the solution is

There is no particular need to tare the balance when we weigh by difference if the two measurements are made at approximately the same time; this is the usual situation when we acquire a sample by this method. If the two measurements are separated by a significant period of time, then we should tare the balance before each measurement and then include the uncertainty of both tares when we calculate the absolute uncertainty in mass.

$$\frac{1.4252 \text{ g}}{0.1000 \text{ L}} \times \frac{1 \text{ mol}}{121.34 \text{ g}} = 0.1175 \text{ M}$$

The relative uncertainty in this concentration is

$$\frac{u_C}{C} = \sqrt{\left(\frac{0.00014}{1.4252}\right)^2 + \left(\frac{0.01}{121.34}\right)^2 + \left(\frac{0.00008}{0.1000}\right)^2} = 0.00081$$

and the absolute uncertainty in the concentration is

$$u_C = 0.00081 \times (0.1175 \text{ M}) = 0.000095 \text{ M}$$

The concentration, therefore, is $0.1175 \pm 0.0001 \text{ M}$.

8. The mean value for n measurements is

$$\begin{aligned} \bar{X} &= \frac{\sum_i^n X_i}{n} \\ &= \frac{X_1 + X_2 + \cdots + X_{n-1} + X_n}{n} \\ &= \frac{1}{n} \{X_1 + X_2 + \cdots + X_{n-1} + X_n\} \end{aligned}$$

If we let the absolute uncertainty in the measurement of X_i equal σ , then a propagation of uncertainty for the sum of n measurements is

$$\begin{aligned} \sigma_{\bar{X}} &= \frac{1}{n} \sqrt{(\sigma)_1^2 + (\sigma)_2^2 + \cdots + (\sigma)_{n-1}^2 + (\sigma)_n^2} \\ &= \frac{1}{n} \sqrt{n(\sigma)^2} = \frac{\sqrt{n}}{n} \sigma = \frac{\sigma}{\sqrt{n}} \end{aligned}$$

9. Because we are subtracting \bar{X}_B from \bar{X}_A , a propagation of uncertainty of their respective uncertainties shows us that

$$\begin{aligned} u_{|\bar{X}_B - \bar{X}_A|} &= \sqrt{\left(\frac{t_{\text{exp}} s_A}{\sqrt{n_A}}\right)^2 + \left(\frac{t_{\text{exp}} s_B}{\sqrt{n_B}}\right)^2} \\ &= \sqrt{\frac{t_{\text{exp}}^2 s_A^2}{n_A} + \frac{t_{\text{exp}}^2 s_B^2}{n_B}} \\ &= \sqrt{t_{\text{exp}}^2 \left(\frac{s_A^2}{n_A} + \frac{s_B^2}{n_B}\right)} \\ &= t_{\text{exp}} \sqrt{\left(\frac{s_A^2}{n_A} + \frac{s_B^2}{n_B}\right)} \end{aligned}$$

10. To have a relative uncertainty of less than 0.1% requires that we satisfy the following inequality

$$\frac{0.1 \text{ mg}}{x} \leq 0.001$$

where x is the minimum mass we need to take. Solving for x shows that we need to weigh out a sample of at least 100 mg.

11. It is tempting to assume that using the 50-mL pipet is the best option because it requires only two transfers to dispense 100.0 mL, providing

fewer opportunities for a determinate error; although this is true with respect to determinate errors, our concern here is with indeterminate errors. We can estimate the indeterminate error for each of the three methods using a propagation of uncertainty. When we use a pipet several times, the total volume dispensed is

$$V_{total} = \sum_i^n V_i$$

for the which the uncertainty is

$$u_{V_{total}} = \sqrt{(u_{V_1})^2 + (u_{V_2})^2 + \dots + (u_{V_{n-1}})^2 + (u_{V_n})^2} = \sqrt{n(u_{V_i})^2}$$

The uncertainties for dispensing 100.0 mL using each pipet are:

$$50\text{-mL pipet: } u_{V_{total}} = \sqrt{2(0.05)^2} = 0.071 \text{ mL}$$

$$25\text{-mL pipet: } u_{V_{total}} = \sqrt{4(0.03)^2} = 0.060 \text{ mL}$$

$$10\text{-mL pipet: } u_{V_{total}} = \sqrt{10(0.02)^2} = 0.063 \text{ mL}$$

where the uncertainty for each pipet are from Table 4.2. Based on these calculations, if we wish to minimize uncertainty in the form of indeterminate errors, then the best option is to use a 25-mL pipet four times.

12. There are many ways to use the available volumetric glassware to accomplish this dilution. Shown here are the optimum choices for a one-step, a two-step, and a three-step dilution using the uncertainties from Table 4.2. For a one-step dilution we use a 5-mL volumetric pipet and a 1000-mL volumetric flask; thus

$$\frac{u_C}{C} = \sqrt{\left(\frac{0.01}{5.00}\right)^2 + \left(\frac{0.30}{1000.0}\right)^2} = 0.0020$$

For a two-step dilution we use a 50-mL volumetric pipet and a 1000-mL volumetric flask followed by a 50-mL volumetric pipet and a 500-mL volumetric flask; thus

$$\frac{u_C}{C} = \sqrt{\left(\frac{0.05}{50.00}\right)^2 + \left(\frac{0.30}{1000.0}\right)^2 + \left(\frac{0.05}{50.00}\right)^2 + \left(\frac{0.20}{500.0}\right)^2} = 0.0015$$

Finally, for a three-step dilution we use 50-mL volumetric pipet and a 100-mL volumetric flask, a 50-mL volumetric flask and a 500-mL volumetric flask, and a 50-mL volumetric pipet and a 500-mL volumetric flask; thus

$$\frac{u_C}{C} = \sqrt{\left(\frac{0.05}{50.00}\right)^2 + \left(\frac{0.08}{100.0}\right)^2 + \left(\frac{0.05}{50.00}\right)^2 + \left(\frac{0.20}{500.0}\right)^2 + \left(\frac{0.05}{50.00}\right)^2 + \left(\frac{0.20}{500.0}\right)^2} = 0.0020$$

The smallest uncertainty is obtained with the two-step dilution.

13. The mean is the average value. If each measurement, X_i , is changed by the same amount, ΔX , then the total change for n measurement is $n\Delta X$ and the average change is $n\Delta X/n$ or ΔX . The mean, therefore, changes by ΔX . When we calculate the standard deviation

$$s = \sqrt{\frac{(X_i - \bar{X})^2}{n - 1}}$$

the important term is the summation in the numerator, which consists of the difference between each measurement and the mean value

$$(X_i - \bar{X})^2$$

Because both X_i and \bar{X} change by ΔX , the value of $X_i - \bar{X}$ becomes

$$X_i + \Delta X - (\bar{X} + \Delta X) = X_i - \bar{X}$$

which leaves unchanged the numerator of the equation for the standard deviation; thus, changing all measurements by ΔX has no effect on the standard deviation.

14. Answers to this question will vary with the object chosen. For a simple, regularly shaped object—a sphere or cube, for example—where you can measure the linear dimensions with a caliper, Method A should yield a smaller standard deviation and confidence interval than Method B. When using a mm ruler to measure the linear dimensions of a regularly shaped object, the two methods should yield similar results. For an object that is irregular in shape, Method B should yield a smaller standard deviation and confidence interval.
15. The isotopic abundance for ^{13}C is 1.11%; thus, for a molecule to average at least one atom of ^{13}C , the total number of carbon atoms must be at least

$$N = \frac{\mu}{p} = \frac{1}{0.0111} = 90.1$$

which we round up to 91 atoms. The probability of finding no atoms of ^{13}C in a molecule with 91 carbon atoms is given by the binomial distribution; thus

$$P(0,91) = \frac{91!}{0!(91 - 0)!} (0.0111)^0 (1 - 0.0111)^{91-0} = 0.362$$

and 36.2% of such molecules will not contain an atom of ^{13}C .

16. (a) The probability that a molecule of cholesterol has one atom of ^{13}C is

$$P(1,27) = \frac{27!}{1!(27 - 0)!} (0.0111)^1 (1 - 0.0111)^{27-1} = 0.224$$

or 22.4%. (b) From Example 4.10, we know that $P(0,27)$ is 0.740. Because the total probability must equal one, we know that

$$P(\geq 2, 27) = 1.000 - P(0, 27) - P(1, 27)$$

$$P(\geq 2, 27) = 1.00 - 0.740 - 0.224$$

$$P(\geq 2, 27) = 0.036$$

and 3.6% of cholesterol molecules will have two or more atoms of ^{13}C .

17. The mean and the standard deviation for the eight samples are, respectively, 16.883% w/w Cr and 0.0794% w/w Cr. The 95% confidence interval is

$$\begin{aligned}\mu &= \bar{X} \pm \frac{ts}{\sqrt{n}} = 16.883 \pm \frac{(2.365)(0.0794)}{\sqrt{8}} \\ &= 16.883 \pm 0.066\% \text{ w/w Cr}\end{aligned}$$

Based on this one set of experiments, and in the absence of any determinate errors, there is a 95% probability that the actual %w/w Cr in the reference material is in the range 16.817–16.949% w/w Cr.

18. (a) The mean and the standard deviation for the nine samples are 36.1 ppt and 4.15 ppt, respectively. The null hypothesis and the alternative hypothesis are

$$H_0: \bar{X} = \mu \quad H_A: \bar{X} \neq \mu$$

The test statistic is t_{exp} , for which

$$t_{\text{exp}} = \frac{|\mu - \bar{X}| \sqrt{n}}{s} = \frac{|40.0 - 36.1| \sqrt{9}}{4.15} = 2.82$$

The critical value for $t(0.05, 8)$ is 2.306. Because t_{exp} is greater than $t(0.05, 8)$, we reject the null hypothesis and accept the alternative hypothesis, finding evidence, at $\alpha = 0.05$, that the difference between \bar{X} and μ is too great to be explained by random errors in the measurements.

(b) Because concentration, C , and signal are proportional, we can use concentration in place of the signal when calculating detection limits. For σ_{mb} we use the standard deviation for the method blank of 0.16 ppt, and for σ_A we use the standard deviation of 4.15 ppt from part (a); thus

$$C_{\text{DL}} = C_{\text{mb}} + z\sigma_{\text{mb}} = 0.16 + (3.00)(1.20) = 3.76 \text{ ppt}$$

$$\begin{aligned}C_{\text{LOI}} &= C_{\text{mb}} + z\sigma_{\text{mb}} + z\sigma_A \\ &= 0.16 + (3.00)(1.20) + (3.00)(4.15) = 16.21 \text{ ppt}\end{aligned}$$

$$C_{\text{LOQ}} = C_{\text{mb}} + 10\sigma_{\text{mb}} = 0.16 + (10.00)(1.20) = 12.16 \text{ ppt}$$

19. The mean and the standard deviation are, respectively, 0.639 and 0.00082. The null hypothesis and the alternative hypothesis are

$$H_0: \bar{X} = \mu \quad H_A: \bar{X} \neq \mu$$

The test statistic is t_{exp} , for which

$$t_{\text{exp}} = \frac{|\mu - \bar{X}| \sqrt{n}}{s} = \frac{|0.640 - 0.639| \sqrt{7}}{0.00082} = 3.23$$

The critical value for $t(0.01,6)$ is 3.707. Because t_{exp} is less than $t(0.01,6)$, we retain the null hypothesis, finding no evidence, at $\alpha = 0.01$, that there is a significant difference between \bar{X} and μ .

20. The mean and the standard deviation are 76.64 decays/min and 2.09 decays/min, respectively. The null hypothesis and the alternative hypothesis are

$$H_0: \bar{X} = \mu \quad H_A: \bar{X} \neq \mu$$

The test statistic is t_{exp} , for which

$$t_{\text{exp}} = \frac{|\mu - \bar{X}| \sqrt{n}}{s} = \frac{|77.5 - 76.64| \sqrt{12}}{2.09} = 1.43$$

The critical value for $t(0.05,11)$ is 2.2035. Because t_{exp} is less than $t(0.05,11)$, we retain the null hypothesis, finding no evidence, at $\alpha = 0.05$, that there is a significant difference between \bar{X} and μ .

21. The mean and the standard deviation are, respectively, 5730 ppm Fe and 91.3 ppm Fe. In this case we need to calculate μ , which is

$$\begin{aligned} \mu &= \frac{(2.6540 \text{ g sample}) \times \frac{0.5351 \text{ g Fe}}{\text{g sample}} \times \frac{1 \times 10^6 \mu\text{g}}{\text{g}}}{250.0 \text{ mL}} \\ &= 5681 \text{ ppm Fe} \end{aligned}$$

The null hypothesis and the alternative hypothesis are

$$H_0: \bar{X} = \mu \quad H_A: \bar{X} \neq \mu$$

The test statistic is t_{exp} , for which

$$t_{\text{exp}} = \frac{|\mu - \bar{X}| \sqrt{n}}{s} = \frac{|5681 - 5730| \sqrt{4}}{91.3} = 1.07$$

The critical value for $t(0.05,3)$ is 3.182. Because t_{exp} is less than $t(0.05,3)$, we retain the null hypothesis, finding no evidence, at $\alpha = 0.05$, that there is a significant difference between \bar{X} and μ .

22. This problem involves a comparison between two sets of unpaired data. For the digestion with HNO_3 , the mean and the standard deviation are, respectively, 163.8 ppb Hg and 3.11 ppb Hg, and for the digestion with the mixture of HNO_3 and HCl , the mean and the standard deviation are, respectively, 148.3 ppb Hg and 7.53 ppb Hg.

The null hypothesis and the alternative hypothesis are

$$H_0: \bar{X}_{\text{HNO}_3} = \bar{X}_{\text{mix}} \quad H_A: \bar{X}_{\text{HNO}_3} \neq \bar{X}_{\text{mix}}$$

Before we can test these hypotheses, however, we first must determine if we can pool the standard deviations. To do this we use the following null hypothesis and alternative hypothesis

$$H_0: s_{\text{HNO}_3} = s_{\text{mix}} \quad H_A: s_{\text{HNO}_3} \neq s_{\text{mix}}$$

The test statistic is F_{exp} for which

$$F_{\text{exp}} = \frac{s_{\text{mix}}^2}{s_{\text{HNO}_3}^2} = \frac{(7.53)^2}{(3.11)^2} = 5.86$$

The critical value for $F(0.05, 5, 4)$ is 9.364. Because F_{exp} is less than $F(0.05, 5, 4)$, we retain the null hypothesis, finding no evidence, at $\alpha = 0.05$, that there is a significant difference between the standard deviations. Pooling the standard deviations gives

$$s_{\text{pool}} = \sqrt{\frac{(4)(3.11)^2 + (5)(7.53)^2}{5 + 6 - 2}} = 5.98$$

The test statistic for the comparison of the means is t_{exp} , for which

$$\begin{aligned} t_{\text{exp}} &= \frac{|\bar{X}_{\text{HNO}_3} - \bar{X}_{\text{mix}}|}{s_{\text{pool}}} \times \sqrt{\frac{n_{\text{HNO}_3} \times n_{\text{mix}}}{n_{\text{HNO}_3} + n_{\text{mix}}}} \\ &= \frac{|163.8 - 148.3|}{5.98} \times \sqrt{\frac{5 \times 6}{5 + 6}} = 4.28 \end{aligned}$$

with nine degrees of freedom. The critical value for $t(0.05, 9)$ is 2.262. Because t_{exp} is greater than $t(0.05, 9)$, we reject the null hypothesis and accept the alternative hypothesis, finding evidence, at $\alpha = 0.05$, that the difference between the means is significant.

23. This problem involves a comparison between two sets of unpaired data. For the samples of atmospheric origin, the mean and the standard deviation are, respectively, 2.31011 g and 0.000143 g, and for the samples of chemical origin, the mean and the standard deviation are, respectively, 2.29947 g and 0.00138 g.

The null hypothesis and the alternative hypothesis are

$$H_0: \bar{X}_{\text{atm}} = \bar{X}_{\text{chem}} \quad H_A: \bar{X}_{\text{atm}} \neq \bar{X}_{\text{chem}}$$

Before we can test these hypotheses, however, we first must determine if we can pool the standard deviations. To do this we use the following null hypothesis and alternative hypothesis

$$H_0: s_{\text{atm}} = s_{\text{chem}} \quad H_A: s_{\text{atm}} \neq s_{\text{chem}}$$

The test statistic is F_{exp} for which

$$F_{\text{exp}} = \frac{s_{\text{chem}}^2}{s_{\text{atm}}^2} = \frac{(0.00138)^2}{(0.000143)^2} = 97.2$$

The critical value for $F(0.05, 7, 6)$ is 5.695. Because F_{exp} is less than $F(0.05, 7, 6)$, we reject the null hypothesis and accept the alternative hypothesis that the standard deviations are different at $\alpha = 0.05$. Be-

cause we cannot pool the standard deviations, the test statistic, t_{exp} , for comparing the means is

$$\begin{aligned} t_{\text{exp}} &= \frac{|\bar{X}_{\text{atm}} - \bar{X}_{\text{chem}}|}{\sqrt{\frac{s_{\text{atm}}^2}{n_{\text{atm}}} + \frac{s_{\text{chem}}^2}{n_{\text{chem}}}}} \\ &= \frac{|2.31011 - 2.29947|}{\sqrt{\frac{(0.000143)^2}{7} + \frac{(0.00138)^2}{8}}} = 21.68 \end{aligned}$$

The number of degrees of freedom is

$$\nu = \frac{\left(\frac{(0.000143)^2}{7} + \frac{(0.00138)^2}{8}\right)^2}{\frac{\left(\frac{(0.000143)^2}{7}\right)^2}{7+1} + \frac{\left(\frac{(0.00138)^2}{8}\right)^2}{8+1}} - 2 = 7.21 \approx 7$$

The critical value for $t(0.05,7)$ is 2.365. Because t_{exp} is greater than $t(0.05,7)$, we reject the null hypothesis and accept the alternative hypothesis, finding evidence, at $\alpha = 0.05$, that the difference between the means is significant. Rayleigh observed that the density of N_2 isolated from the atmosphere was significantly larger than that for N_2 derived from chemical sources, which led him to hypothesize the presence of an unaccounted for gas in the atmosphere.

24. This problem involves a comparison between two sets of unpaired data. For the standard method, the mean and the standard deviation are, respectively, $22.86 \mu\text{L}/\text{m}^3$ and $1.28 \mu\text{L}/\text{m}^3$, and for the new method, the mean and the standard deviation are, respectively, $22.51 \mu\text{L}/\text{m}^3$ and $1.92 \mu\text{L}/\text{m}^3$.

The null hypothesis and the alternative hypothesis are

$$H_0: \bar{X}_{\text{std}} = \bar{X}_{\text{new}} \quad H_A: \bar{X}_{\text{std}} \neq \bar{X}_{\text{new}}$$

Before we can test these hypotheses, however, we first must determine if we can pool the standard deviations. To do this we use the following null hypothesis and alternative hypothesis

$$H_0: s_{\text{std}} = s_{\text{new}} \quad H_A: s_{\text{std}} \neq s_{\text{new}}$$

The test statistic is F_{exp} for which

$$F_{\text{exp}} = \frac{s_{\text{new}}^2}{s_{\text{std}}^2} = \frac{(1.92)^2}{(1.28)^2} = 2.25$$

The critical value for $F(0.05,6,6)$ is 5.820. Because F_{exp} is less than $F(0.05,6,6)$, we retain the null hypothesis, finding no evidence, at $\alpha = 0.05$, that there is a significant difference between the standard deviations. Pooling the standard deviations gives

$$s_{\text{pool}} = \sqrt{\frac{(6)(1.28)^2 + (6)(1.92)^2}{7+7-2}} = 1.63$$

The test statistic for the comparison of the means is t_{exp} , for which

$$t_{\text{exp}} = \frac{|\bar{X}_{\text{std}} - \bar{X}_{\text{new}}|}{s_{\text{pool}}} \times \sqrt{\frac{n_{\text{std}} \times n_{\text{new}}}{n_{\text{std}} + n_{\text{new}}}}$$

$$= \frac{|22.86 - 22.51|}{1.63} \times \sqrt{\frac{7 \times 7}{7 + 7}} = 0.40$$

with 12 degrees of freedom. The critical value for $t(0.05, 12)$ is 2.179. Because t_{exp} is less than $t(0.05, 9)$, we retain the null hypothesis, finding no evidence, at $\alpha = 0.05$, that there is a significant difference between new method and the standard method.

25. This problem is a comparison between two sets of paired data. The differences, which we define as (measured – accepted), are

$$0.0001 \quad 0.0013 \quad -0.0003 \quad 0.0015 \quad -0.0006$$

The mean and the standard deviation for the differences are 0.00040 and 0.00095, respectively. The null hypothesis and the alternative hypothesis are

$$H_0: \bar{d} = 0 \quad H_A: \bar{d} \neq 0$$

The test statistic is t_{exp} , for which

$$t_{\text{exp}} = \frac{|\bar{d}| \sqrt{n}}{s} = \frac{|0.00040| \sqrt{5}}{0.00095} = 0.942$$

The critical value for $t(0.05, 4)$ is 2.776. Because t_{exp} is less than $t(0.05, 4)$, we retain the null hypothesis, finding no evidence, at $\alpha = 0.05$, that the spectrometer is inaccurate.

26. This problem is a comparison between two sets of paired data. The differences, which we define as (ascorbic acid – sodium bisulfate), are

$$15 \quad -31 \quad 1 \quad 20 \quad 4 \quad -52 \quad -22 \quad -62 \quad -50$$

The mean and the standard deviation for the differences are –19.7 and 30.9, respectively. The null hypothesis and the alternative hypothesis are

$$H_0: \bar{d} = 0 \quad H_A: \bar{d} \neq 0$$

The test statistic is t_{exp} , for which

$$t_{\text{exp}} = \frac{|\bar{d}| \sqrt{n}}{s} = \frac{|-19.7| \sqrt{9}}{30.9} = 1.91$$

The critical value for $t(0.10, 8)$ is 1.860. Because t_{exp} is greater than $t(0.10, 8)$, we reject the null hypothesis and accept the alternative hypothesis, finding evidence, at $\alpha = 0.10$, that the two preservatives do not have equivalent holding times.

27. This problem is a comparison between two sets of paired data. The differences, which we define as (actual – found), are

$$-1.8 \quad -1.7 \quad 0.2 \quad -0.5 \quad -3.6 \quad -1.7 \quad 1.1 \quad -1.7 \quad 0.3$$

The mean and the standard deviation for the differences are -1.04 and 1.44 , respectively. The null hypothesis and the alternative hypothesis are

$$H_0: \bar{d} = 0 \quad H_A: \bar{d} \neq 0$$

The test statistic is t_{exp} , for which

$$t_{\text{exp}} = \frac{|\bar{d}|\sqrt{n}}{s} = \frac{|-1.04|\sqrt{9}}{1.44} = 2.17$$

The critical value for $t(0.05,8)$ is 2.306 . Because t_{exp} is less than $t(0.10,8)$, we retain the null hypothesis, finding no evidence, at $\alpha = 0.05$, that the analysis for kaolinite is inaccurate.

28. This problem is a comparison between two sets of paired data. The differences, which we define as (electrode – spectrophotometric), are

$$\begin{array}{cccccc} 0.6 & -5.8 & 0.2 & 0.1 & -0.5 & -0.6 \\ 0.1 & -0.5 & -0.7 & -0.3 & 0.3 & 0.1 \end{array}$$

The mean and the standard deviation for the differences are -0.583 and 1.693 , respectively. The null hypothesis and the alternative hypothesis are

$$H_0: \bar{d} = 0 \quad H_A: \bar{d} \neq 0$$

The test statistic is t_{exp} , for which

$$t_{\text{exp}} = \frac{|\bar{d}|\sqrt{n}}{s} = \frac{|-0.583|\sqrt{12}}{1.693} = 1.19$$

The critical value for $t(0.05,11)$ is 2.2035 . Because t_{exp} is less than $t(0.05,11)$, we retain the null hypothesis, finding no evidence, at $\alpha = 0.05$, that the two methods yield different results.

29. This problem is a comparison between two sets of paired data. The differences, which we define as (proposed – standard), are

$$0.19 \quad 0.91 \quad 1.39 \quad 1.02 \quad -2.38 \quad -2.40 \quad 0.03 \quad 0.82$$

The mean and the standard deviation for the differences are -0.05 and 1.51 , respectively. The null hypothesis and the alternative hypothesis are

$$H_0: \bar{d} = 0 \quad H_A: \bar{d} \neq 0$$

The test statistic is t_{exp} , for which

$$t_{\text{exp}} = \frac{|\bar{d}|\sqrt{n}}{s} = \frac{|-0.05|\sqrt{8}}{1.51} = 0.09$$

The critical value for $t(0.05,7)$ is 2.365 . Because t_{exp} is less than $t(0.05,11)$, we retain the null hypothesis, finding no evidence, at $\alpha = 0.05$, that the two methods yield different results. This is not a very satisfying result, however, because many of the individual differ-

ences are quite large. In this case, additional work might help better characterize the improved method relative to the standard method.

30. The simplest way to organize this data is to make a table, such as the one shown here

sample	smallest value	next-to-smallest value	next-to-largest value	largest value
1	21.3	21.5	23.0	23.1
2	12.9	13.5	13.9	14.2
3	15.9	16.0	17.4	17.5

The only likely candidate for an outlier is the smallest value of 12.9 for sample 2. Using Dixon's Q -test, the test statistic, Q_{exp} , is

$$Q_{\text{exp}} = \frac{|X_{\text{out}} - X_{\text{nearest}}|}{X_{\text{largest}} - X_{\text{smallest}}} = \frac{13.5 - 12.9}{14.2 - 12.9} = 0.462$$

which is smaller than the critical value for $Q(0.05, 10)$ of 0.466; thus, there is no evidence using Dixon's Q -test at $\alpha = 0.05$ to suggest that 12.9 is outlier.

To use Grubb's test we need the mean and the standard deviation for sample 2, which are 13.67 and 0.356, respectively. The test statistic, G_{exp} , is

$$G_{\text{exp}} = \frac{|X_{\text{out}} - \bar{X}|}{s} = \frac{|12.9 - 13.67|}{0.356} = 2.16$$

which is smaller than the critical value for $G(0.05, 10)$ of 2.290; thus, there is no evidence using Grubb's test at $\alpha = 0.05$ that 12.9 is an outlier.

To use Chauvenet's criterion we calculate the deviation, z , for the suspected outlier, assuming a normal distribution and using the sample's mean and standard deviation

$$z = \frac{|X_{\text{out}} - \bar{X}|}{s} = \frac{|12.9 - 13.67|}{0.356} = 2.16$$

which, from Appendix 3, corresponds to a probability of 0.0154. The critical value to which we compare this is $(2n)^{-1}$, or $(2 \times 10)^{-1} = 0.05$. Because the experimental probability of 0.0154 is smaller than the theoretical probability of 0.05 for 10 samples, we have evidence using Chauvenet's criterion that 12.9 is an outlier.

At this point, you may be asking yourself what to make of these seemingly contradictory results, in which two tests suggest that 12.9 is not an outlier and one test suggests that it is an outlier. Here it is helpful to keep in mind three things. First, Dixon's Q -test and Grubb's test require us to pick a particular confidence level, α , and make a decision based on that confidence level. When using Chauvenet's

criterion, however, we do not assume a particular confidence level; instead, we simply evaluate the probability that the outlier belongs to a normal distribution described by the sample's mean and standard deviation relative to a predicted probability defined by the size of the sample. Second, although Q_{exp} and G_{exp} are not large enough to identify 12.9 as an outlier at $\alpha = 0.05$, their respective values are not far removed from their respective critical values (0.462 vs. 0.466 for Dixon's Q -test and 2.16 vs. 2.290 for Grubb's test). Both tests, for example, identify 12.9 as an outlier at $\alpha = 0.10$. Third, and finally, for the reasons outlined in the text, you should be cautious when rejecting a possible outlier based on a statistical test only. All three of these tests, however, suggest that we should at least take a closer look at the measurement that yielded 12.9 as a result.

31. (a) The mean is 1.940, the median is 1.942 (the average of the 31st and the 32nd rank ordered values rounded to four significant figures), and the standard deviation is 0.047.

(b) Figure SM4.4 shows a histogram for the 60 results using bins of size 0.02. The resulting distribution is a reasonably good approximation to a normal distribution, although it appears to have a slight skew toward smaller Cu/S ratios.

(c) The range $\bar{X} \pm 1s$ extends from a Cu/S ratio of 1.893 to 1.987. Of the 62 experimental results, 44 or 71% fall within this range. This agreement with the expected value of 68.26% for a normal distribution is reasonably good.

(d) For a deviation of

$$z = \frac{2.000 - 1.940}{0.047} = 1.28$$

the probability from Appendix 3 that a Cu/S ratio is greater than 2 is 10.03%. Of the 62 experimental results, three or 4.8% fall within this range. This is a little lower than expected for a normal distribution, but consistent with the observation from part (b) that the data are skewed slightly toward smaller Cu/S ratios.

(e) The null hypothesis and the alternative hypothesis are

$$H_0: \bar{X} = 2.000 \quad H_A: \bar{X} < 2.000$$

Note that the alternative hypothesis here is one-tailed as we are interested only in whether the mean Cu/S ratio is significantly less than 2. The test statistic, t_{exp} , is

$$t_{\text{exp}} = \frac{|1.940 - 2.000| \sqrt{62}}{0.047} = 10.0$$

As t_{exp} is greater than the one-tailed critical value for $t(0.05, 61)$, which is between 1.65 and 1.75, we reject the null hypothesis and

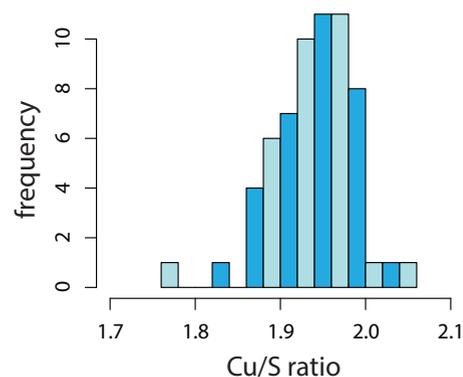


Figure SM4.4 Histogram for the data in problem 31. Each bar in has a width of 0.02. For example, the bar on the far left includes all Cu/S ratios from 1.76 to 1.78, which includes the single result of 1.764.

accept the alternative hypothesis, finding evidence that the Cu/S ratio is significantly less than its expected stoichiometric ratio of 2.

32. Although answers for this problem will vary, here are some details you should address in your report. The descriptive statistics for all three data sets are summarized in the following table.

statistic	sample X	sample Y	sample Z
mean	24.56	27.76	23.75
median	24.55	28.00	23.52
range	1.26	4.39	5.99
std dev	0.339	1.19	1.32
variance	0.115	1.43	1.73

The most interesting observation from this summary is that the spread of values for sample X—as given by the range, the standard deviation, and the variance—is much smaller than that for sample Y and for sample Z.

Outliers are one possible explanation for the difference in spread among these three samples. Because the number of individual results for each sample is greater than the largest value of n for the critical values included in Appendix 6 for Dixon's Q -test and in Appendix 7 for Grubb's test, we will use Chauvenet's criterion; the results are summarized in the following table.

statistic	sample X	sample Y	sample Z
possible outlier	23.92	24.41	28.79
z	1.89	2.63	3.83
probability	0.0294	0.0043	0.0000713

For 18 samples, the critical probability is $(2 \times 18)^{-1}$ or 0.0277; thus, we have evidence that there is an outlier in sample Y and in sample Z, but not in sample X. Removing these outliers and recalculating the descriptive statistics gives the results in the following table.

statistic	sample X	sample Y	sample Z
mean	24.56	27.74	23.45
median	24.55	28.00	23.48
range	1.26	3.64	1.37
std dev	0.339	0.929	0.402
variance	0.115	0.863	0.161

The spread for sample Y still seems large relative to sample X, but the spread for sample Z now seems similar to sample X. An F -test of the variances using the following null hypothesis and alternative hypothesis

$$H_0: s_1 = s_2 \quad H_A: s_1 \neq s_2$$

gives an F_{exp} of 5.340 when comparing sample Y to sample Z, and of 1.406 when comparing sample Z to sample X. Comparing these values to the critical value for $F(0.05,17,17)$, which is between 2.230 and 2.308, suggests that our general conclusions are reasonable.

The mean values for the three samples appear different from each other. A t -test using the following null hypothesis and alternative hypothesis

$$H_0: \bar{X}_1 = \bar{X}_2 \quad H_A: \bar{X}_1 \neq \bar{X}_2$$

gives a t_{exp} of 13.30 when comparing sample Y to sample X, which is much greater than the critical value for $t(0.05,20)$ of 2.086. The value of t_{exp} when comparing sample Z to sample X is 8.810, which is much greater than the critical value for $t(0.05,33)$, which is between 2.042 and 2.086.

This process of completing multiple significance tests is not without problems, for reasons we will discuss in Chapter 14 when we consider analysis of variance.

