27

The t-test

In many applications the quantity of interest can be represented by the expectation of the model distribution. In some of these applications one wants to know whether this expectation deviates from some a priori specified value. This can be investigated by means of a statistical test, known as the t-test. We consider this test both under the assumption that the model distribution is normal and without the assumption of normality. Furthermore, we discuss a similar test for the slope and the intercept in a simple linear regression model.

27.1 Monitoring the production of ball bearings

A production line in a large industrial corporation are set to produce a specific type of steel ball bearing with a diameter of 1 millimeter. In order to check the performance of the production lines, a number of ball bearings are picked at the end of the day and their diameters are measured. Suppose we observe 20 diameters of ball bearings from the production lines, which are listed in Table 27.1. The average diameter is $\bar{x}_{20} = 1.03$ millimeter. This clearly deviates from the target value 1, but the question is whether the difference can be attributed to chance or whether it is large enough to conclude that the production line is producing ball bearings with a wrong diameter. To answer this question, we model the dataset as a realization of a random sample $X_1, X_2, \ldots, X_{20}$ from a probability distribution with expected value μ. The parameter μ represents the diameter of ball bearings produced by the produc-

Table 27.1. Diameters of ball bearings.

1.018	1.009	1.042	1.053	0.969	1.002	0.988	1.019	1.062	1.032
1.072	0.977	1.062	1.044	1.069	1.029	0.979	1.096	1.079	0.999

tion lines. In order to investigate whether this diameter deviates from 1, we test the null hypothesis $H_0 : \mu = 1$ against $H_1 : \mu \neq 1$.

This example illustrates a situation that often occurs: the data $x_1, x_2, \ldots, x_n$ are a realization of a random sample $X_1, X_2, \ldots, X_n$ from a distribution with expectation μ, and we want to test whether μ equals an a priori specified value, say μ_0. According to the law of large numbers, $\bar{X}_n$ is close to μ for large n. This suggests a test statistic based on $\bar{X}_n - \mu_0$; realizations of $\bar{X}_n - \mu_0$ close to zero are in favor of the null hypothesis. Does $\bar{X}_n - \mu_0$ suffice as a test statistic?

In our example, $\bar{x}_n - \mu_0 = 1.03 - 1 = 0.03$. Should we interpret this as small? First, note that under the null hypothesis $\mathrm{E}\left[\bar{X}_n - \mu_0\right] = \mu - \mu_0 = 0$. Now, if $\bar{X}_n - \mu_0$ would have standard deviation 1, then the value 0.03 is within one standard deviation of $\mathrm{E}\left[\bar{X}_n - \mu_0\right]$. The "$\mu \pm$ a few σ" rule on page 185 then suggests that the value 0.03 is not exceptional; it must be seen as a small deviation. On the other hand, if $\bar{X}_n - \mu_0$ has standard deviation 0.001, then the value 0.03 is 30 standard deviations away from $\mathrm{E}\left[\bar{X}_n - \mu_0\right]$. According to the "$\mu \pm$ a few σ" rule this is very exceptional; the value 0.03 must be seen as a large deviation. The next quick exercise provides a concrete example.

QUICK EXERCISE 27.1 Suppose that $\bar{X}_n$ is a normal random variable with expectation 1 and variance 1. Determine $\mathrm{P}\left(\bar{X}_n - 1 \geq 0.03\right)$. Find the same probability, but for the case where the variance is $(0.01)^2$.

This discussion illustrates that we must standardize $\bar{X}_n - \mu_0$ to incorporate its variation. Recall that

$$\mathrm{Var}\left(\bar{X}_n - \mu_0\right) = \mathrm{Var}\left(\bar{X}_n\right) = \frac{\sigma^2}{n},$$

where σ^2 is the variance of each X_i. Hence, standardizing $\bar{X}_n - \mu_0$ means that we should divide by $\sigma/\sqrt{n}$. Since σ is unknown, we substitute the sample standard deviation S_n for σ. This leads to the following test statistic for the null hypothesis $H_0 : \mu = \mu_0$:

$$T = \frac{\bar{X}_n - \mu_0}{S_n/\sqrt{n}}.$$

Values of T close to zero are in favor of $H_0 : \mu = \mu_0$. Large positive values of T suggest that $\mu > \mu_0$ and large negative values suggest that $\mu < \mu_0$; both are evidence against H_0.

For the ball bearing data one finds that $s_n = 0.0372$, so that

$$t = \frac{\bar{x}_n - \mu_0}{s_n/\sqrt{n}} = \frac{1.03 - 1}{0.0372/\sqrt{20}} = 3.607.$$

This is clearly different from zero, but the question is whether this difference is large enough to reject $H_0 : \mu = 1$. To answer this question, we need to know the probability distribution of T under the null hypothesis. Note that under the null hypothesis $H_0 : \mu = \mu_0$, the test statistic

$$T = \frac{\bar{X}_n - \mu_0}{S_n/\sqrt{n}}$$

is the studentized mean (see also Chapter 23)

$$\frac{\bar{X}_n - \mu}{S_n/\sqrt{n}}.$$

Hence, *under the null hypothesis*, the probability distribution of T is the *same* as that of the studentized mean.

27.2 The one-sample t-test

The classical assumption is that the dataset is a realization of a random sample from an $N(\mu, \sigma^2)$ distribution. In that case our test statistic T turns out to have a t-distribution under the null hypothesis, as we will see later. For this reason, the test for the null hypothesis $H_0 : \mu = \mu_0$ is called the *(one-sample) t-test*. Without the assumption of normality, we will use the bootstrap to approximate the distribution of T. For large sample sizes, this distribution can be approximated by means of the central limit theorem. We start with the first case.

Normal data

Suppose that the dataset $x_1, x_2, \ldots, x_n$ is a realization of a random sample $X_1, X_2, \ldots, X_n$ from an $N(\mu, \sigma^2)$ distribution. Then, according to the rule on page 349, the studentized mean has a $t(n-1)$ distribution. An immediate consequence is that, under the null hypothesis $H_0 : \mu = \mu_0$, also our test statistic T has a $t(n-1)$ distribution. Therefore, if we test $H_0 : \mu = \mu_0$ against $H_1 : \mu \neq \mu_0$ at level α, then we must reject the null hypothesis in favor of $H_1 : \mu \neq \mu_0$, if

$$T \leq -t_{n-1,\alpha/2} \quad \text{or} \quad T \geq t_{n-1,\alpha/2}.$$

Similar decision rules apply to alternatives $H_1 : \mu > \mu_0$ and $H_1 : \mu < \mu_0$. Suppose that in the ball bearing example we test $H_0 : \mu = 1$ against $H_1 : \mu \neq 1$ at level $\alpha = 0.05$. From Table B.2 we find $t_{19,0.025} = 2.093$. Hence, we must reject if $T \leq -2.093$ or $T \geq 2.093$. For the ball bearing data we found $t = 3.607$, which means we reject the null hypothesis at level $\alpha = 0.05$.

Alternatively, one might report the one-tailed p-value corresponding to the observed value t and compare this with $\alpha/2$. The one-tailed p-value is either a right or a left tail probability, which must be computed by means of the $t(n-1)$ distribution. In our ball bearing example the one-tailed p-value is the right tail probability $\mathrm{P}(T \geq 3.607)$. From Table B.2 we see that this probability is between 0.0005 and 0.0010, which is smaller than $\alpha/2 = 0.025$ (to be precise, by means of a statistical software package we found $\mathrm{P}(T \geq 3.607) = 0.00094$). The data provide strong enough evidence against the null hypothesis, so that it seems sensible to adjust the settings of the production line.

Quick exercise 27.2 Suppose that the data in Table 27.1 are from two separate production lines. The first ten measurements have average 1.0194 and standard deviation 0.0290, whereas the last ten measurements have average 1.0406 and standard deviation 0.0428. Perform the t-test $H_0 : \mu = 1$ against $H_1 : \mu \neq 1$ at level $\alpha = 0.01$ for both datasets separately, assuming normality.

Nonnormal data

Draw a rectangle with height h and width w (let us agree that $w > h$), and within this rectangle draw a square with sides of length h (see Figure 27.1). This creates another (smaller) rectangle with horizontal and vertical sides of

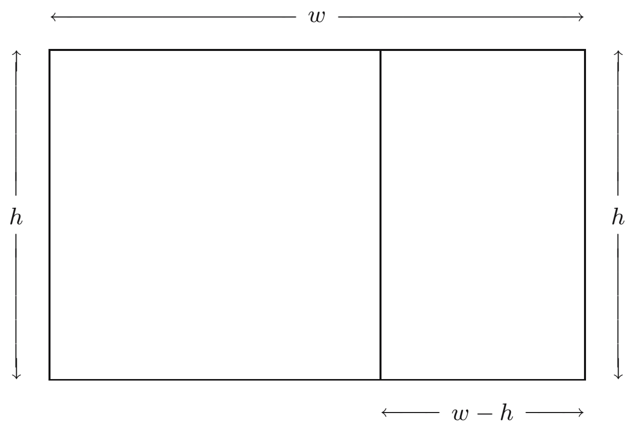

Fig. 27.1. Rectangle with square within.

lengths $w - h$ and h. A large rectangle with a vertical-to-horizontal ratio that is equal to the horizontal-to-vertical ratio for the small rectangle, i.e.,

$$\frac{h}{w} = \frac{w-h}{h},$$

was called a "golden rectangle" by the ancient Greeks, who often used these in their architecture. After solving for h/w, we obtain that the height-to-width

Table 27.2. Ratios for Shoshoni rectangles.

0.693	0.749	0.654	0.670	0.662	0.672	0.615	0.606	0.690	0.628
0.668	0.611	0.606	0.609	0.601	0.553	0.570	0.844	0.576	0.933

Source: C. Dubois (ed.). *Lowie's selected papers in anthropology*, 1960.

ratio h/w is equal to the "golden number" $(\sqrt{5}-1)/2 \approx 0.618$. The data in Table 27.2 represent corresponding h/w ratios for rectangles used by Shoshoni Indians to decorate their leather goods. Is it reasonable to assume that they were *also* using golden rectangles? We examine this by means of a t-test.

The observed ratios are modeled as a realization of a random sample from a distribution with expectation μ, where the parameter μ represents the true esthetic preference for height-to-width ratios of the Shoshoni Indians. We want to test

$$H_0 : \mu = 0.618 \quad \text{against} \quad H_1 : \mu \neq 0.618.$$

For the Shoshoni ratios, $\bar{x}_n = 0.6605$ and $s_n = 0.0925$, so that the value of the test statistic is

$$t = \frac{\bar{x}_n - 0.618}{s_n/\sqrt{n}} = \frac{0.6605 - 0.618}{0.0925/\sqrt{20}} = 2.055.$$

Closer examination of the data indicates that the normal distribution is not the right model. For instance, by definition the height-to-width ratios h/w are always between 0 and 1. Because some of the data points are also close to right boundary 1, the normal distribution is inappropriate. If we *cannot* assume a normal model distribution, we can *no longer* conclude that our test statistic has a $t(n-1)$ distribution under the null hypothesis.

Since there is no reason to assume any other particular type of distribution to model the data, we approximate the distribution of T under the null hypothesis. Recall that this distribution is the same as that of the studentized mean (see the end of Section 27.1). To approximate its distribution, we use the empirical bootstrap simulation for the studentized mean, as described on page 351. We generate 10 000 bootstrap datasets and for each bootstrap dataset $x_1^*, x_2^*, \ldots, x_n^*$, we compute

$$t^* = \frac{\bar{x}_n^* - 0.6605}{s_n^*/\sqrt{n}}.$$

In Figure 27.2 the kernel density estimate and empirical distribution function are displayed for 10 000 bootstrap values t^*. Suppose we test $H_0 : \mu = 0.618$ against $H_1 : \mu \neq 0.618$ at level $\alpha = 0.05$. In the same way as in Section 23.3, we find the following bootstrap approximations for the critical values:

$$c_l^* = -3.334 \quad \text{and} \quad c_u^* = 1.644.$$

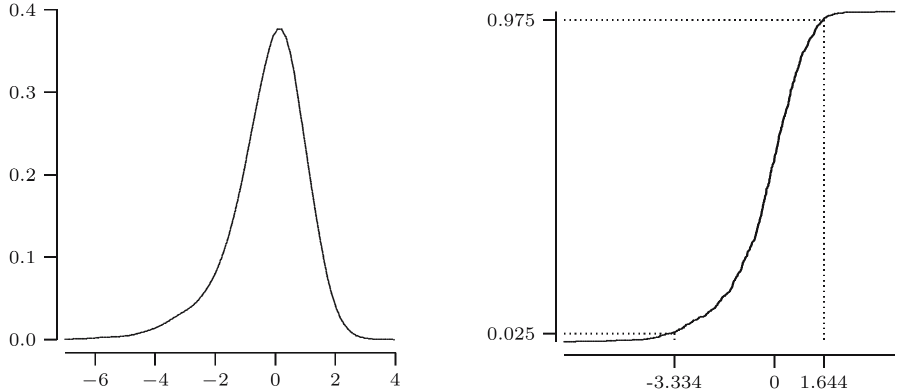

Fig. 27.2. Kernel density estimate and empirical distribution function of 10 000 bootstrap values t^*.

Since for the Shoshoni data the value 2.055 of the test statistic is greater than 1.644, we reject the null hypothesis at level 0.05. Alternatively, we can also compute a bootstrap approximation of the one-tailed p-value corresponding to 2.055, which is the right tail probability $\mathrm{P}(T \geq 2.055)$. The bootstrap approximation for this probability is:

$$\frac{\text{number of } t^* \text{values greater than or equal to } 2.055}{10\,000} = 0.0067.$$

Hence $\mathrm{P}(T \geq 2.055) \approx 0.0067$, which is smaller than $\alpha/2 = 0.025$. The value 2.055 should be considered as exceptionally large, and we reject the null hypothesis. The esthetic preference for height-to-width ratios of the Shoshoni Indians differs from that of the ancient Greeks.

Large samples

For large sample sizes the distribution of the studentized mean can be approximated by a standard normal distribution (see Section 23.4). This means that for large sample sizes the distribution of the t-test statistic under the null hypothesis can also be approximated by a standard normal distribution. To illustrate this, recall the Old Faithful data. Park rangers in Yellowstone National Park inform the public about the behavior of the geyser, such as the expected time between successive eruptions and the length of the duration of an eruption. Suppose they claim that the expected length of an eruption is 4 minutes (240 seconds). Does this seem likely on the basis of the data from Section 15.1? We investigate this by testing $H_0 : \mu = 240$ against $H_1 : \mu \neq 240$ at level $\alpha = 0.001$, where μ is the expectation of the model distribution. The value of the test statistic is

$$t = \frac{\bar{x}_n - 240}{s_n/\sqrt{n}} = \frac{209.3 - 240}{68.48/\sqrt{272}} = -7.39.$$

The one-tailed p-value $\mathrm{P}(T \leq -7.39)$ can be approximated by $\mathrm{P}(Z \leq -7.39)$, where Z has an $N(0,1)$ distribution. From Table B.1 we see that this probability is smaller than $\mathrm{P}(Z \leq -3.49) = 0.0002$. This is smaller than $\alpha/2 = 0.0005$, so we reject the null hypothesis at level 0.001. In fact the p-value is much smaller: a statistical software package gives $\mathrm{P}(Z \leq -7.39) = 7.5 \cdot 10^{-14}$. The data provide overwhelming evidence against $H_0 : \mu = 240$, so that we conclude that the expected length of an eruption is different from 4 minutes.

QUICK EXERCISE 27.3 Compute the critical region K for the test, using the normal approximation, and check that $t = -7.39$ falls in K.

In fact, if we would test $H_0 : \mu = 240$ against $H_1 : \mu < 240$, the p-value corresponding to $t = -7.39$ is the left tail probability $\mathrm{P}(T \leq -7.39)$. This probability is very small, so that we also reject the null hypothesis in favor of this alternative and conclude that the expected length of an eruption is smaller than 4 minutes.

27.3 The t-test in a regression setting

Is calcium in your drinking water good for your health? In England and Wales, an investigation of environmental causes of disease was conducted. The annual mortality rate (percentage of deaths) and the calcium concentration in the drinking water supply were recorded for 61 large towns. The data in Table 27.3 represent the annual mortality rate averaged over the years 1958–1964, and the calcium concentration in parts per million. In Figure 27.3 the 61 paired measurements are displayed in a scatterplot. The scatterplot shows a slight downward trend, which suggests that higher concentrations of calcium lead to lower mortality rates. The question is whether this is really the case or if the slight downward trend should be attributed to chance.

To investigate this question we model the mortality data by means of a simple linear regression model with normally distributed errors, with the mortality rate as the dependent variable y and the calcium concentration as the independent variable x:

$$Y_i = \alpha + \beta x_i + U_i \quad \text{for } i = 1, 2, \ldots, 61,$$

where $U_1, U_2, \ldots, U_{61}$ is a random sample from an $N(0, \sigma^2)$ distribution. The parameter β represents the change of the mortality rate if we increase the calcium concentration by one unit. We test the null hypothesis $H_0 : \beta = 0$ (calcium has no effect on the mortality rate) against $H_1 : \beta < 0$ (higher concentration of calcium reduces the mortality rate).

This example illustrates the general situation, where the dataset

$$(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$$

Table 27.3. Mortality data.

Rate	Calcium	Rate	Calcium	Rate	Calcium	Rate	Calcium
1247	105	1466	5	1299	78	1359	84
1392	73	1307	78	1254	96	1318	122
1260	21	1096	138	1402	37	1309	59
1259	133	1175	107	1486	5	1456	90
1236	101	1369	68	1257	50	1527	60
1627	53	1486	122	1485	81	1519	21
1581	14	1625	13	1668	17	1800	14
1609	18	1558	10	1807	15	1637	10
1755	12	1491	20	1555	39	1428	39
1723	44	1379	94	1742	8	1574	9
1569	91	1591	16	1772	15	1828	8
1704	26	1702	44	1427	27	1724	6
1696	6	1711	13	1444	14	1591	49
1987	8	1495	14	1587	75	1713	71
1557	13	1640	57	1709	71	1625	20
1378	71						

Source: M. Hills and the M345 Course Team. *M345 Statistical Methods, Units 3: Examining Straight-line Data*, 1986, Milton Keynes: © Open University, 28. Data provided by Professor M.J.Gardner, Medical Research Council Environmental Epidemiology Research Unit, Southampton.

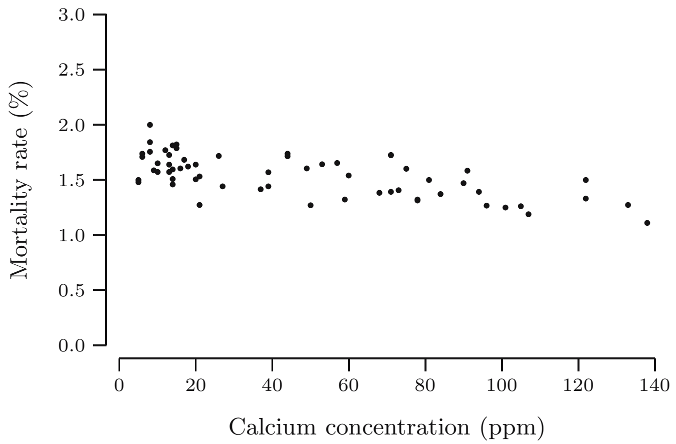

Fig. 27.3. Scatterplot mortality data.

is modeled by a simple linear regression model, and one wants to test a null hypothesis of the form $H_0 : \alpha = \alpha_0$ or $H_0 : \beta = \beta_0$. Similar to the one-sample t-test we will construct a test statistic for each of these null hypotheses. With normally distributed errors, these test statistics have a t-distribution under the null hypothesis. For this reason, for both null hypotheses the test is called a t-test.

The *t*-test for the slope

For the null hypothesis $H_0 : \beta = \beta_0$, we use as test statistic

$$T_b = \frac{\hat{\beta} - \beta_0}{S_b},$$

where $\hat{\beta}$ is the least squares estimator for β (see Chapter 22) and

$$S_b^2 = \frac{n}{n \sum x_i^2 - (\sum x_i)^2}\, \hat{\sigma}^2.$$

In this expression,

$$\hat{\sigma}^2 = \frac{1}{n-2} \sum_{i=1}^{n} (Y_i - \hat{\alpha} - \hat{\beta} x_i)^2$$

is the estimator for σ^2 as introduced on page 332. It can be shown that

$$\mathrm{Var}\left(\hat{\beta} - \beta_0\right) = \frac{n}{n \sum x_i^2 - (\sum x_i)^2}\, \sigma^2,$$

so that the random variable S_b^2 is an estimator for the variance of $\hat{\beta} - \beta_0$. Hence, similar to the test statistic for the one-sample t-test, the test statistic T_b compares the estimator $\hat{\beta}$ with the value β_0 and standardizes by dividing by an estimator for the standard deviation of $\hat{\beta} - \beta_0$. Values of T_b close to zero are in favor of the null hypothesis $H_0 : \beta = \beta_0$. Large positive values of T_b suggest that $\beta > \beta_0$, whereas large negative values of T_b suggest that $\beta < \beta_0$.

Recall that in the case of normal random samples the one-sample t-test statistic has a $t(n-1)$ distribution under the null hypothesis. For the same reason, it is also a fact that in the case of normally distributed errors the test statistic T_b has a $t(n-2)$ distribution under the null hypothesis $H_0 : \beta = \beta_0$.

In our mortality example we want to test $H_0 : \beta = 0$ against $H_0 : \beta < 0$. For the data we find $\hat{\beta} = -3.2261$ and $s_b = 0.4847$, so that the value of T_b is

$$t_b = \frac{-3.2261}{0.4847} = -6.656.$$

If we test at level $\alpha = 0.05$, then we must compare this value with the left critical value $-t_{59,0.05}$. This value is not in Table B.2, but we have that

$$-1.676 = -t_{50,0.05} < -t_{59,0.05}.$$

This means that t_b is much smaller than $-t_{59,0.05}$, so that we reject the null hypothesis at level 0.05. How much evidence the value $t_b = -6.656$ bears against the null hypothesis is expressed by the one-tailed p-value $\mathrm{P}(T_b \le -6.656)$. From Table B.2 we can only see that this probability is smaller than 0.0005. By means of a statistical package we find $\mathrm{P}(T_b \le -6.656) = 5.2 \cdot 10^{-9}$. The data provide overwhelming evidence against the null hypothesis. We conclude that higher concentrations of calcium correspond to lower mortality rates.

QUICK EXERCISE 27.4 The data in Table 27.3 can be separated into measurements for towns at least as far north as Derby and towns south of Derby. For the data corresponding to 35 towns at least as far north as Derby, one finds $\hat{\beta} = -1.9313$ and $s_b = 0.8479$. Test $H_0 : \beta = 0$ against $H_0 : \beta < 0$ at level 0.01, i.e., compute the value of the test statistic and report your conclusion about the null hypothesis.

The *t*-test for the intercept

We test the null hypothesis $H_0 : \alpha = \alpha_0$ with test statistic

$$T_a = \frac{\hat{\alpha} - \alpha_0}{S_a}, \tag{27.1}$$

where $\hat{\alpha}$ is the least squares estimator for α and

$$S_a^2 = \frac{\sum x_i^2}{n \sum x_i^2 - (\sum x_i)^2} \hat{\sigma}^2,$$

with $\hat{\sigma}^2$ defined as before. The random variable S_a^2 is an estimator for the variance

$$\mathrm{Var}(\hat{\alpha} - \alpha_0) = \frac{\sum x_i^2}{n \sum x_i^2 - (\sum x_i)^2} \sigma^2.$$

Again, we compare the estimator $\hat{\alpha}$ with the value α_0 and standardize by dividing by an estimator for the standard deviation of $\hat{\alpha} - \alpha_0$. Values of T_a close to zero are in favor of the null hypothesis $H_0 : \alpha = \alpha_0$. Large positive values of T_a suggest that $\alpha > \alpha_0$, whereas large negative values of T_a suggest that $\alpha < \alpha_0$. Like T_b, in the case of normal errors, the test statistic T_a has a $t(n-2)$ distribution under the null hypothesis $H_0 : \alpha = \alpha_0$.

As an illustration, recall Exercise 17.9 where we modeled the volume y of black cherry trees by means of a linear model without intercept, with independent variable $x = d^2 h$, where d and h are the diameter and height of the trees. The scatterplot of the pairs $(x_1, y_1), (x_2, y_2), \ldots, (x_{31}, y_{31})$ is displayed in Figure 27.4. As mentioned in Exercise 17.9, there are physical reasons to leave out the intercept. We want to investigate whether this is confirmed by the data. To this end, we model the data by a simple linear regression model *with* intercept

$$Y_i = \alpha + \beta x_i + U_i \quad \text{for } i = 1, 2, \ldots, 31,$$

where $U_1, U_2, \ldots, U_{31}$ are a random sample from an $N(0, \sigma^2)$ distribution, and we test $H_0 : \alpha = 0$ against $H_1 : \alpha \neq 0$ at level 0.10. The value of the test statistic is

$$t_a = \frac{-0.2977}{0.9636} = -0.3089,$$

and the left critical value is $-t_{29,0.05} = -1.699$. This means we cannot reject the null hypothesis. The data do not provide sufficient evidence against $H_0 : \alpha = 0$, which is confirmed by the one-tailed p-value $\mathrm{P}(T_a \leq -0.3089) = 0.3798$ (computed by means of a statistical package). We conclude that the intercept does not contribute significantly to the model.

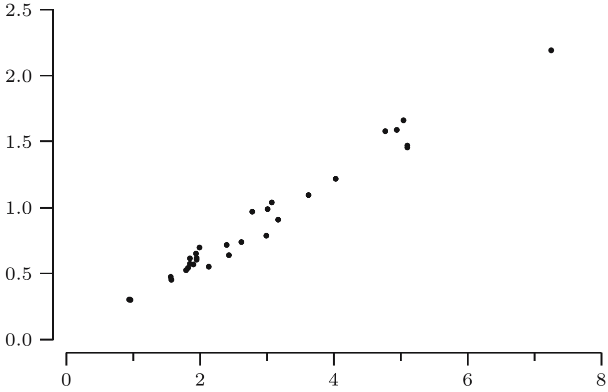

Fig. 27.4. Scatterplot of the black cherry tree data.

27.4 Solutions to the quick exercises

27.1 If Y has an $N(1, 1)$ distribution, then $Y - 1$ has an $N(0, 1)$ distribution. Therefore, from Table B.1: $\mathrm{P}(Y - 1 \geq 0.03) = 0.4880$. If Y has an $N(1, (0.01)^2)$ distribution, then $(Y - 1)/0.01$ has an $N(0, 1)$ distribution. In that case,

$$\mathrm{P}(Y - 1 \geq 0.03) = \mathrm{P}\left(\frac{Y - 1}{0.01} \geq 3\right) = 0.0013.$$

27.2 For the first and last ten measurements the values of the test statistic are

$$t = \frac{1.0194 - 1}{0.0290/\sqrt{10}} = 2.115 \quad \text{and} \quad t = \frac{1.0406 - 1}{0.0428/\sqrt{10}} = 3.000.$$

The critical value $t_{9,0.025} = 2.262$, which means we reject the null hypothesis for the second production line, but not for the first production line.

27.3 The critical region is of the form $K = (-\infty, c_l] \cup [c_u, \infty)$. The right critical value c_u is approximated by $z_{0.0005} = t_{\infty,0.0005} = 3.291$, which can be found in Table B.2. By symmetry of the normal distribution, the left critical value c_l is approximated by $-z_{0.0005} = -3.291$. Clearly, $t = -7.39 < -3.291$, so that it falls in K.

27.4 The value of the test statistic is

$$t_b = \frac{-1.9313}{0.8479} = -2.2778.$$

The left critical value is equal to $-t_{33,0.01}$, which is not in Table B.2, but we see that $-t_{33,0.01} < -t_{40,0.01} = -2.423$. This means that $-t_{33,0.01} < t_b$, so that we cannot reject $H_0 : \beta = 0$ against $H_0 : \beta < 0$ at level 0.01.

27.5 Exercises

27.1 We perform a t-test for the null hypothesis $H_0 : \mu = 10$ by means of a dataset consisting of $n = 16$ elements with sample mean 11 and sample variance 4. We use significance level 0.05.

a. Should we reject the null hypothesis in favor of $H_1 : \mu \neq 10$?

b. What if we test against $H_1 : \mu > 10$?

27.2 ⊡ The Cleveland Casting Plant is a large highly automated producer of gray and nodular iron automotive castings for Ford Motor Company. One process variable of interest to Cleveland Casting is the pouring temperature of molten iron. The pouring temperatures (in degrees Fahrenheit) of ten crankshafts are given in Table 27.4. The target setting for the pouring temperature is set at 2550 degrees. One wants to conduct a test at level $\alpha = 0.01$ to determine whether the pouring temperature differs from the target setting.

Table 27.4. Pouring temperatures of ten crankshafts.

2543	2541	2544	2620	2560
2559	2562	2553	2552	2553

a. Formulate the appropriate null hypothesis and alternative hypothesis.

b. Compute the value of the test statistic and report your conclusion. You may assume a normal model distribution and use that the sample variance is 517.34.

27.3 Table 27.5 lists the results of tensile adhesion tests on 22 U-700 alloy specimens. The data are loads at failure in MPa. The sample mean is 13.71 and the sample standard deviation is 3.55. You may assume that the data originated from a normal distribution with expectation μ. One is interested in whether the load at failure exceeds 10 MPa. We investigate this by means of a t-test for the null hypothesis $H_0 : \mu = 10$.

a. What do you choose as the alternative hypothesis?

b. Compute the value of the test statistic and report your conclusion, when performing the test at level 0.05.

Table 27.5. Loads at failure of U-700 specimens.

19.8	18.5	17.6	16.7	15.8
15.4	14.1	13.6	11.9	11.4
11.4	8.8	7.5	15.4	15.4
19.5	14.9	12.7	11.9	11.4
10.1	7.9			

Source: C.C. Berndt. Instrumented Tensile adhesion tests on plasma sprayed thermal barrier coatings. *Journal of Materials Engineering* II(4): 275-282, Dec 1989.

27.4 Consider the coal data from Table 23.2, where 22 gross calorific value measurements are listed for Daw Mill coal coded 258GB41. We modeled this dataset as a realization of a random sample from an $N(\mu, \sigma^2)$ distribution with μ and σ unknown. We are planning to buy a shipment if the gross calorific value exceeds 31.00 MJ/kg. The sample mean and sample variance of the data are $\bar{x}_n = 31.012$ and $s_n = 0.1294$. Perform a t-test for the null hypothesis $H_0 : \mu = 31.00$ against $H_1 : \mu > 31.00$ using significance level 0.01, i.e., compute the value of the test statistic, the critical value of the test, and report your conclusion.

27.5 ⊞ In the November 1988 issue of *Science* a study was reported on the inbreeding of tropical swarm-founding wasps. Each member of a sample of 197 wasps was captured, frozen, and subjected to a series of genetic tests, from which an inbreeding coefficient was determined. The sample mean and the sample standard deviation of the coefficients are $\bar{x}_{197} = 0.044$ and $s_{197} = 0.884$. If a species does not have the tendency to inbreed, their true inbreeding coefficient is 0. Determine by means of a test whether the inbreeding coefficient for this species of wasp exceeds 0.

a. Formulate the appropriate null hypothesis and alternative hypothesis and compute the value of the test statistic.

b. Compute the p-value corresponding to the value of the test statistic and report your conclusion about the null hypothesis.

27.6 The stopping distance of an automobile is related to its speed. The data in Table 27.6 give the stopping distance in feet and speed in miles per hour of an automobile. The data are modeled by means of simple linear regression model with normally distributed errors, with the square root of the stopping distance as dependent variable y and the speed as independent variable x:

$$Y_i = \alpha + \beta x_i + U_i, \quad \text{for } i = 1, \ldots, 7.$$

For the dataset we find

$$\hat{\alpha} = 5.388, \quad \hat{\beta} = 4.252, \quad s_a = 1.874, \quad s_b = 0.242.$$

Table 27.6. Speed and stopping distance of automobiles.

Speed	20.5	20.5	30.5	30.5	40.5	48.8	57.8
Distance	15.4	13.3	33.9	27.0	73.1	113.0	142.6

Source: K.A. Brownlee. *Statistical theory and methodology in science and engineering.* Wiley, New York, 1960; Table II.9 on page 372.

One would expect that the intercept can be taken equal to 0, since zero speed would yield zero stopping distance. Investigate whether this is confirmed by the data by performing the appropriate test at level 0.10. Formulate the proper null and alternative hypothesis, compute the value of the test statistic, and report your conclusion.

27.7 ⊞ In a study about the effect of wall insulation, the weekly gas consumption (in 1000 cubic feet) and the average outside temperature (in degrees Celsius) was measured of a certain house in southeast England, for 26 weeks before and 30 weeks after cavity-wall insulation had been installed. The house thermostat was set at 20 degrees throughout. The data are listed in Table 27.7. We model the data before insulation by means of a simple linear regression model with normally distributed errors and gas consumption as response variable. A similar model was used for the data after insulation. Given are

$$\begin{array}{ll} \text{Before insulation:} & \hat{\alpha} = 6.8538,\ \hat{\beta} = -0.3932 \text{ and } s_a = 0.1184,\ s_b = 0.0196 \\ \text{After insulation:} & \hat{\alpha} = 4.7238,\ \hat{\beta} = -0.2779 \text{ and } s_a = 0.1297,\ s_b = 0.0252. \end{array}$$

a. Use the data before insulation to investigate whether smaller outside temperatures lead to higher gas consumption. Formulate the proper null and alternative hypothesis, compute the value of the test statistic, and report your conclusion, using significance level 0.05.

b. Do the same for the data after insulation.

Table 27.7. Temperature and gas consumption.

Before insulation		**After insulation**	
Temperature	Gas consumption	Temperature	Gas consumption
−0.8	7.2	−0.7	4.8
−0.7	6.9	0.8	4.6
0.4	6.4	1.0	4.7
2.5	6.0	1.4	4.0
2.9	5.8	1.5	4.2
3.2	5.8	1.6	4.2
3.6	5.6	2.3	4.1
3.9	4.7	2.5	4.0
4.2	5.8	2.5	3.5
4.3	5.2	3.1	3.2
5.4	4.9	3.9	3.9
6.0	4.9	4.0	3.5
6.0	4.3	4.0	3.7
6.0	4.4	4.2	3.5
6.2	4.5	4.3	3.5
6.3	4.6	4.6	3.7
6.9	3.7	4.7	3.5
7.0	3.9	4.9	3.4
7.4	4.2	4.9	3.7
7.5	4.0	4.9	4.0
7.5	3.9	5.0	3.6
7.6	3.5	5.3	3.7
8.0	4.0	6.2	2.8
8.5	3.6	7.1	3.0
9.1	3.1	7.2	2.8
10.2	2.6	7.5	2.6
		8.0	2.7
		8.7	2.8
		8.8	1.3
		9.7	1.5

Source: MDST242 Statistics in Society, Unit 45: Review, 2nd edition, 1984, Milton Keynes: © The Open University, Figures 2.5 and 2.6.