# 5 Testing statistical hypotheses

In practice, testing statistical hypotheses is one of the most important statistical methods. A *statistical hypothesis* is an assumption (statement) concerning a random sample $V_n$, or the distribution from which the random sample originates. *Testing a statistical hypothesis* is a procedure by which we verify the validity (correctness) of a given hypothesis (assumption). The *null hypothesis* $H_0$ is a statement whose truth we are verifying. In the simplest case it has the form $H_0 : Q = Q_0$, resp. $H_0 : Q - Q_0 = 0$, where $Q$ is a parameter of the population and $Q_0$ is a specific constant. Against the null hypothesis we set up the *alternative hypothesis* $H_1$. In practice, the alternative hypothesis has three forms:

- *Two-sided alternative:* $H_1 : Q \neq Q_0$, which denies the validity of the null hypothesis.
- *Left-sided alternative:* $H_1 : Q < Q_0$, which bounds the range of values of the parameter $Q$ to the left of the value $Q_0$.
- *Right-sided alternative:* $H_1 : Q > Q_0$, which bounds the range of values of the parameter $Q$ to the right of the value $Q_0$.

When testing statistical hypotheses, i. e. when deciding whether to reject or not reject the null hypothesis $H_0$ in favor of the alternative hypothesis $H_1$, we may commit two kinds of errors (Table 5):

- A *Type I error* occurs when we reject the null hypothesis $H_0$ even though it is correct. The probability of this error occurring is denoted $\alpha$ and is called the *significance level* of the test.
- A *Type II error* occurs when we do not reject the null hypothesis $H_0$ even though it is incorrect. The probability of this error occurring is denoted $\beta$ and the value $1 - \beta$ is called the *power of the test*. The power of the test actually represents the probability of rejecting the null hypothesis.

**Table 5:** Errors in testing statistical hypotheses

| | Reality | |
|---|---|---|
| Conclusion | $H_0$ holds | $H_0$ does not hold |
| We reject $H_0$ | Type I error | Correct decision |
| We do not reject $H_0$ | Correct decision | Type II error |

The most ideal situation would be for both $\alpha$ and $\beta$ to be minimal simultaneously. Since this cannot be achieved, in practice, for a fixed $\alpha$ (most often chosen as 0.05; 0.01; 0.1), $\beta$ is minimized.

A *test statistic (test criterion, testing characteristic)* is a value calculated from the sample, which helps us decide whether to reject or not reject the null hypothesis. The range (region) for rejecting the null hypothesis is called the *critical region* $K_\alpha$ and depends on the chosen significance level $\alpha$. If the value of the test statistic falls into the critical region, we reject the null hypothesis $H_0$ and accept the alternative hypothesis $H_1$. Otherwise, we do not reject the null hypothesis $H_0$.

With the development of computing technology, the p-value (from the English p-value) is used when deciding whether to reject or not reject the null hypothesis $H_0$. The *p-value* ($p$) is the smallest possible significance level $\alpha$ at which it is still possible, given the measured values of the random sample, to reject the null hypothesis $H_0$. The following holds:

- if $p < \alpha$, then we reject the null hypothesis $H_0$ and accept the alternative hypothesis $H_1$ at the significance level $\alpha$,
- if $p \geqq \alpha$, then we do not reject the null hypothesis $H_0$ at the significance level $\alpha$.

The procedure for testing hypotheses is as follows:

1. Formulation of the null hypothesis $H_0$ and the alternative hypothesis $H_1$.
2. Choice of the significance level $\alpha$.
3. Calculation of the value of the test statistic based on the given values of the random sample.
4. Determination of the critical region $K_\alpha$.
5. Evaluation of the test, which consists either of rejecting the null hypothesis $H_0$ and accepting the alternative hypothesis $H_1$, or of not rejecting the null hypothesis $H_0$.

The region of non-rejection and rejection of the null hypothesis $H_0$ for a two-sided test is shown in Fig. 19, for a right-sided test it is shown in Fig. 20, and for a left-sided test it is shown in Fig. 21. The values $k_1$ and $k_2$, which separate the region of non-rejection of the null hypothesis $H_0$ from the region of rejection of the null hypothesis $H_0$, are called *critical values.*

There are many statistical tests, which can be divided according to various criteria. The basic division of statistical tests is into parametric and nonparametric tests. *Parametric tests*

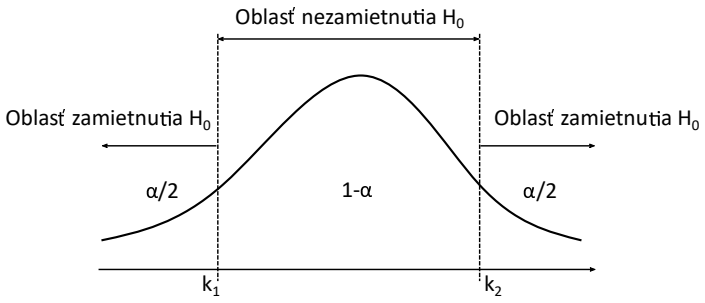



**Figure 19:** Region of non-rejection and rejection of the null hypothesis for a two-sided test

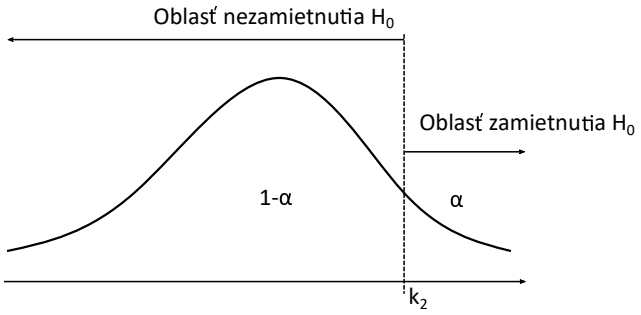



**Figure 20:** Region of non-rejection and rejection of the null hypothesis for a right-sided test

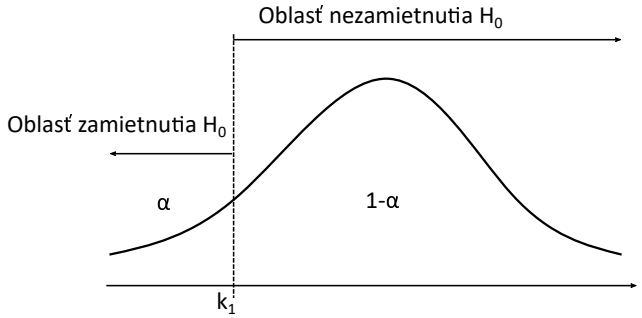



**Figure 21:** Region of non-rejection and rejection of the null hypothesis for a left-sided test

test an unknown parameter of the distribution from which the random sample $V_n$ originates. *Nonparametric tests* test the type of distribution or some property of it. The most commonly used nonparametric tests include *symmetry tests* (e.g., the sign test, the Wilcoxon two-sample test), *goodness-of-fit tests* (e.g., Pearson's test, the Kolmogorov–Smirnov test, the Anderson–Darling test, the Shapiro–Wilk normality test) and *tests for extreme values* (e.g., Dixon's test, Grubbs' test). We will define all of the tests mentioned in the following chapters.

## 5.1 One-sample parametric tests of parameters of the normal distribution

Consider a random sample $V_n$ that comes from a population with a normal distribution with mean value $\mu$ and standard deviation $\sigma$. We write $V_n \sim N(\mu, \sigma^2)$. In this chapter we will address the following tests:

- test of the mean value $\mu$, if
  - we know the variance $\sigma^2$ (one-sample Z-test),
  - we do not know the variance $\sigma^2$ (one-sample t-test),
- test of the variance $\sigma^2$, if
  - we know the mean value $\mu$,
  - we do not know the mean value $\mu$.

The general procedure for choosing a test of the mean value for one sample is shown in Fig. 22.

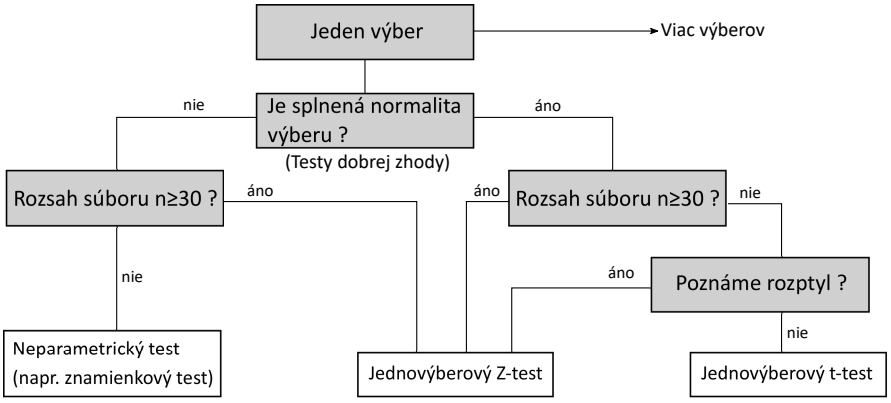



**Figure 22:** Procedure for choosing a test of the mean value for one sample

### 5.1.1 One-sample test of the mean value $\mu$, where we know the variance $\sigma^2$ (one-sample Z-test)

1. Hypotheses:

$$H_0 : \mu = \mu_0 \quad (\mu_0 \text{ is a known constant}),$$
$$H_1 : \mu \neq \mu_0 \quad (\text{resp. } \mu < \mu_0, \text{ resp. } \mu > \mu_0).$$

2. The test statistic is

$$Z = \frac{\overline{x} - \mu_0}{\sigma} \sqrt{n} \sim N(0,1), \tag{96}$$

where $\overline{x}$ is the sample arithmetic mean and $n$ is the size of the random sample.

3. The critical region $K_\alpha$ for various alternatives is given in the following table:

| $H_0$ | $H_1$ | $K_\alpha$ |
|---|---|---|
| $\mu = \mu_0$ | $\mu \neq \mu_0$ | $\left(-\infty; -u_{1-\frac{\alpha}{2}}\right) \cup \left(u_{1-\frac{\alpha}{2}}; \infty\right)$ |
| $\mu = \mu_0$ | $\mu > \mu_0$ | $(u_{1-\alpha}; \infty)$ |
| $\mu = \mu_0$ | $\mu < \mu_0$ | $(-\infty; -u_{1-\alpha})$ |

**Example 5.1** *The length of a manufactured part should have a value of 10 cm. 9 parts were randomly selected, whose lengths [in cm] were 10.24; 10.12; 9.91; 10.19; 9.78; 10.14; 9.86; 10.17; 10.05. We assume that the random sample was drawn from a population with a normal distribution, whose variance is 0.0016.*

*a) At the significance level $\alpha$=0.05, let us test whether the length of the parts differs significantly from the required length.*

*b) At the significance level $\alpha$=0.01, let us test whether the length of the parts is smaller than the required length.*

*c) At the significance level $\alpha$=0.01, let us test whether the length of the parts is larger than the required length.*

**Solution:** For testing we use the one-sample Z-test, because we know the variance of the random sample. The sample size is $n = 9$ and the standard deviation is $\sigma = \sqrt{0,0016} = 0,04$.

a) We will test

$$H_0 : \mu = 10,$$
$$H_1 : \mu \neq 10.$$

The null hypothesis states that the lengths of the parts do not differ significantly from the required length. Against the alternative hypothesis, which states that the lengths of the parts differ significantly from the required length. We determine the value of the test statistic $Z$ using relation (96)

$$Z = \frac{\overline{x} - 10}{0,04} \cdot \sqrt{9},$$

where the sample arithmetic mean $\overline{x}$ is calculated using relation (71)

$$\overline{x} = \frac{1}{9}\left(10,24 + 10,12 + 9,91 + 10,19 + 9,78 + 10,14 + 9,86 + 10,17 + 10,05\right) = 10,0511.$$

The value of the test statistic is then $Z = 3,8325$. The quantile of the standard normal distribution is $u_{1-\frac{\alpha}{2}} = u_{0,975} = 1,96$ (Table 28). The critical region is

$$K_{0,05} = (-\infty; -1,96) \cup (1,96; \infty) \Rightarrow Z \in K_{0,05}.$$

The value of the test statistic $Z = 3,8325$ falls into the critical region $K_{0,05}$, therefore we reject the null hypothesis $H_0$ and accept the alternative hypothesis $H_1$. We can assume that, at the significance level $\alpha$=0.05, the lengths of the parts differ significantly from the required value.

b) We will test

$$H_0 : \mu = 10,$$
$$H_1 : \mu < 10.$$

The null hypothesis states that the lengths of the parts do not differ significantly from the required length. Against the alternative hypothesis, which states that the lengths of the parts are smaller than the required length. We calculate the value of the test statistic $Z$ in the same way as in case a), and thus $Z = 3,8325$. The quantile of the standard normal distribution is $u_{1-\alpha} = u_{0,99} = 2,3264$ (Table 28). The critical region is

$$K_{0,01} = (-\infty; -2,3264) \Rightarrow Z \notin K_{0,01}.$$

The value of the test statistic $Z = 3,8325$ does not fall into the critical region $K_{0,01}$, therefore we do not reject the null hypothesis $H_0$. We can assume that, at the significance level $\alpha$=0.01, the lengths of the parts are not smaller than the required value.

c) We will test

$$H_0 : \mu = 10,$$
$$H_1 : \mu > 10.$$

The null hypothesis states that the lengths of the parts do not differ significantly from the required length. Against the alternative hypothesis, which states that the lengths of the parts are larger than the required length. We calculate the value of the test statistic $Z$ in the same way as in case a), and thus $Z = 3,8325$. The quantile of the standard normal distribution is $u_{1-\alpha} = u_{0,99} = 2,3264$ (Table 28). The critical region is

$$K_{0,01} = (2,3264; \infty) \Rightarrow Z \in K_{0,01}.$$

The value of the test statistic $Z = 3,8325$ falls into the critical region $K_{0,01}$, therefore we reject the null hypothesis $H_0$ and accept the alternative hypothesis $H_1$. We can assume that, at the significance level $\alpha$=0.01, the lengths of the parts are larger than the required value.

### 5.1.2 One-sample test of the mean value $\mu$, where we do not know the variance $\sigma^2$

**(one-sample t-test)**

1. Hypotheses:

$$H_0 : \mu = \mu_0 \quad (\mu_0 \text{ is a known constant}),$$
$$H_1 : \mu \neq \mu_0 \quad (\text{resp. } \mu < \mu_0, \text{ resp. } \mu > \mu_0).$$

2. The test statistic is

$$t = \frac{\overline{x} - \mu_0}{s} \sqrt{n} \sim t(n-1), \tag{97}$$

where $\overline{x}$ is the sample arithmetic mean, $s$ is the sample standard deviation and $n$ is the size of the random sample.

3. The critical region $K_\alpha$ for various alternatives is given in the following table:

| $H_0$ | $H_1$ | $K_\alpha$ |
|---|---|---|
| $\mu = \mu_0$ | $\mu \neq \mu_0$ | $\left(-\infty; -t_{1-\frac{\alpha}{2}}(n-1)\right) \cup \left(t_{1-\frac{\alpha}{2}}(n-1); \infty\right)$ |
| $\mu = \mu_0$ | $\mu > \mu_0$ | $(t_{1-\alpha}(n-1); \infty)$ |
| $\mu = \mu_0$ | $\mu < \mu_0$ | $(-\infty; -t_{1-\alpha}(n-1))$ |

***Remark:*** If the size of the random sample is sufficiently large ($n \geqq 30$), we proceed in the same way, but we approximate the Student t-distribution by the standard normal distribution, as was done in the previous test (one-sample Z-test).

**Example 5.2** *The length of a manufactured part should have a value of 10 cm. 9 parts were randomly selected, whose lengths [in cm] were 10.24; 10.12; 9.91; 10.19; 9.78; 10.14; 9.86; 10.17; 10.05. We assume that the random sample was drawn from a population with a normal distribution.*

*a) At the significance level $\alpha$=0.05, let us test whether the lengths of the parts differ significantly from the required length.*

*b) At the significance level $\alpha$=0.01, let us test whether the lengths of the parts are smaller than the required length.*

*c) At the significance level $\alpha$=0.01, let us test whether the lengths of the parts are larger than the required length.*

**Solution:** For testing we use the one-sample t-test, because we do not know the variance of the random sample. The sample size is $n = 9$.

a) We will test

$$H_0 : \mu = 10,$$
$$H_1 : \mu \neq 10.$$

The null hypothesis states that the lengths of the parts do not differ significantly from the required length. Against the alternative hypothesis, which states that the lengths of the parts differ significantly from the required length. We determine the value of the test statistic $t$ using relation (97)

$$t = \frac{\overline{x} - 10}{s} \cdot \sqrt{9},$$

where the sample arithmetic mean $\overline{x}$ is calculated using relation (71)

$$\overline{x} = \frac{1}{9}\,(10,24 + 10,12 + 9,91 + 10,19 + 9,78 + 10,14 + 9,86 + 10,17 + 10,05) = 10,0511$$

and the sample standard deviation $s$ is calculated using relation (77), where the sample variance $s^2$ is given by (73)

$$s^2 = \frac{1}{8}\big[(10,24 - 10,0511)^2 + (10,12 - 10,0511)^2 + (9,91 - 10,0511)^2 + (10,19 - 10,0511)^2 +$$
$$+(9,78 - 10,0511)^2 + (10,14 - 10,0511)^2 + (9,86 - 10,0511)^2 + (10,17 - 10,0511)^2 +$$
$$+(10,05 - 10,0511)^2\big] = 0,0265 \Rightarrow s = \sqrt{s^2} = 0,1627.$$

The value of the test statistic is then $t = 0,9426$. The quantile of the Student t-distribution is $t_{1-\frac{\alpha}{2}}(n-1) = t_{0,975}(8) = 2,306$ (Table 29). The critical region is

$$K_{0,05} = (-\infty; -2,306) \cup (2,306; \infty) \Rightarrow t \notin K_{0,05}.$$

The value of the test statistic $t = 0,9426$ does not fall into the critical region $K_{0,05}$, therefore we do not reject the null hypothesis $H_0$. We can assume that, at the significance level $\alpha$=0.05, the lengths of the parts do not differ significantly from the required value.

b) We will test

$$H_0 : \mu = 10,$$
$$H_1 : \mu < 10.$$

The null hypothesis states that the lengths of the parts do not differ significantly from the required length. Against the alternative hypothesis, which states that the lengths of the parts are smaller than the required length. We calculate the value of the test statistic $t$ in the same way as in case a), and thus $t = 0,9426$. The quantile of the Student t-distribution is $t_{1-\alpha}(n-1) = t_{0,99}(8) = 2,896$ (Table 29). The critical region is

$$K_{0,01} = (-\infty; -2,896) \Rightarrow t \notin K_{0,01}.$$

The value of the test statistic $t = 0,9426$ does not fall into the critical region $K_{0,01}$, therefore we do not reject the null hypothesis $H_0$. We can assume that, at the significance level $\alpha$=0.01, the lengths of the parts are not smaller than the required value.

c) We will test

$$H_0 : \mu = 10,$$
$$H_1 : \mu > 10.$$

The null hypothesis states that the lengths of the parts do not differ significantly from the required length. Against the alternative hypothesis, which states that the lengths of the parts are larger than the required length. We calculate the value of the test statistic $t$ in the same way as in case a), and thus $t = 0,9426$. The quantile of the Student t-distribution is $t_{1-\alpha}(n-1) = t_{0,99}(8) = 2,896$ (Table 29). The critical region is

$$K_{0,01} = (2,896; \infty) \Rightarrow t \notin K_{0,01}.$$

The value of the test statistic $t = 0,9426$ does not fall into the critical region $K_{0,01}$, therefore we do not reject the null hypothesis $H_0$. We can assume that, at the significance level $\alpha = 0,01$, the lengths of the parts are not larger than the required value.

### 5.1.3 One-sample test of the variance $\sigma^2$, where we know the mean value $\mu$

1. Hypotheses:

$$H_0 : \sigma^2 = \sigma_0^2 \quad (\sigma_0^2 \text{ is a known constant}),$$
$$H_1 : \sigma^2 \neq \sigma_0^2 \quad (\text{resp. } \sigma^2 < \sigma_0^2, \text{ resp. } \sigma^2 > \sigma_0^2).$$

2. The test statistic is

$$\chi^2 = \frac{n\, s_0^2}{\sigma_0^2} \sim \chi^2(n), \tag{98}$$

where $n$ is the size of the random sample and $s_0^2$ is the sample variance.

3. The critical region $K_\alpha$ for various alternatives is given in the following table:

| $H_0$ | $H_1$ | $K_\alpha$ |
|---|---|---|
| $\sigma^2 = \sigma_0^2$ | $\sigma^2 \neq \sigma_0^2$ | $\left(0; \chi^2_{\frac{\alpha}{2}}(n)\right) \cup \left(\chi^2_{1-\frac{\alpha}{2}}(n); \infty\right)$ |
| $\sigma^2 = \sigma_0^2$ | $\sigma^2 > \sigma_0^2$ | $\left(\chi^2_{1-\alpha}(n); \infty\right)$ |
| $\sigma^2 = \sigma_0^2$ | $\sigma^2 < \sigma_0^2$ | $\left(0; \chi^2_{\alpha}(n)\right)$ |

**Example 5.3** *The length of a manufactured part should have a value of 10 cm. 9 parts were randomly selected, whose lengths [in cm] were 10.24; 10.12; 9.91; 10.19; 9.78; 10.14; 9.86; 10.17; 10.05. We assume that the random sample was drawn from a population with a normal distribution, whose variability of the part length is expressed by a standard deviation of 0.1 cm.*

*a) At the significance level $\alpha$=0.05, let us test whether the variability of the part length differs significantly from the stated variability.*

*b) At the significance level $\alpha$=0.05, let us test whether the variability of the part length is smaller than the stated variability.*

*c) At the significance level $\alpha$=0.05, let us test whether the variability of the part length is larger than the stated variability.*

**Solution:** For testing we use the one-sample test of the variance $\sigma^2$, where we know the mean value $\mu = 10$. The sample size is $n = 9$. We are given the standard deviation $\sigma_0 = 0,1$ cm, hence $\sigma_0^2 = 0,1^2 = 0,01$.

a) We will test

$$H_0 : \sigma^2 = 0,01,$$
$$H_1 : \sigma^2 \neq 0,01.$$

The null hypothesis states that the variability of the part length does not differ significantly from the required variability. Against the alternative hypothesis, which states that the variability of the part length differs significantly from the stated variability value. We determine the value of the test statistic $\chi^2$ using relation (98)

$$\chi^2 = \frac{9 \cdot s_0^2}{0,01},$$

where the sample variance $s_0^2$ is calculated using relation (75)

$$s_0^2 = \frac{1}{9}\big[(10,24-10)^2 + (10,12-10)^2 + (9,91-10)^2 + (10,19-10)^2 + (9,78-10)^2 +$$

$$+(10,14-10)^2+(9,86-10)^2+(10,17-10)^2+(10,05-10)^2\big] = 0,0235.$$

The value of the test statistic is then $\chi^2 = 21,159$. The quantiles of the $\chi^2$-distribution are $\chi^2_{1-\frac{\alpha}{2}}(n) = \chi^2_{0,975}(9) = 19,023$ and $\chi^2_{\frac{\alpha}{2}}(n) = \chi^2_{0,025}(9) = 2,7$ (Table 30). The critical region is

$$K_{0,05} = (0;2,7) \cup (19,023;\infty) \Rightarrow \chi^2 \in K_{0,05}.$$

The value of the test statistic $\chi^2 = 21,159$ falls into the critical region $K_{0,05}$, therefore we reject the null hypothesis $H_0$ and accept the alternative hypothesis $H_1$. We can assume that, at the significance level $\alpha = 0,05$, the variability of the part length differs significantly from the stated value.

b) We will test

$$H_0 : \sigma^2 = 0,01,$$
$$H_1 : \sigma^2 < 0,01.$$

The null hypothesis states that the variability of the part length does not differ significantly from the required variability. Against the alternative hypothesis, which states that the variability of the part length is smaller than the stated variability value. We calculate the value of the test statistic $\chi^2$ in the same way as in case a), and thus $\chi^2 = 21,159$. The quantile of the $\chi^2$-distribution is $\chi^2_{\alpha}(n) = \chi^2_{0,05}(9) = 3,33$ (Table 30). The critical region is

$$K_{0,05} = (0;3,33) \Rightarrow \chi^2 \notin K_{0,05}.$$

The value of the test statistic $\chi^2 = 21,159$ does not fall into the critical region $K_{0,05}$, therefore we do not reject the null hypothesis $H_0$. We can assume that, at the significance level $\alpha = 0,05$, the variability of the part length is not smaller than the stated variability value.

c) We will test

$$H_0 : \sigma^2 = 0,01,$$
$$H_1 : \sigma^2 > 0,01.$$

The null hypothesis states that the lengths of the parts do not differ significantly from the required length. Against the alternative hypothesis, which states that the lengths of the parts are larger than the required length. We calculate the value of the test statistic $\chi^2$ in the same way as in case a), and thus $\chi^2 = 21,159$. The quantile of the $\chi^2$-distribution is $\chi^2_{1-\alpha}(n) = \chi^2_{0,95}(9) = 16,92$ (Table 30). The critical region is

$$K_{0,05} = (16,92;\infty) \Rightarrow \chi^2 \in K_{0,05}.$$

The value of the test statistic $\chi^2 = 21,159$ falls into the critical region $K_{0,05}$, we reject the null hypothesis $H_0$ and accept the alternative hypothesis $H_1$. We can assume that, at

the significance level $\alpha = 0,05$, the variability of the part length is larger than the stated variability value.

### 5.1.4 One-sample test of the variance $\sigma^2$, where we do not know the mean value $\mu$

1. Hypotheses:

$$H_0 : \sigma^2 = \sigma_0^2 \quad (\sigma_0^2 \text{ is a known constant}),$$
$$H_1 : \sigma^2 \neq \sigma_0^2 \quad (\text{resp. } \sigma^2 < \sigma_0^2, \text{ resp. } \sigma^2 > \sigma_0^2).$$

2. The test statistic is

$$\chi^2 = \frac{(n-1)\,s^2}{\sigma_0^2} \sim \chi^2(n-1), \tag{99}$$

where $n$ is the size of the random sample and $s^2$ is the sample variance.

3. The critical region $K_\alpha$ for various alternatives is given in the following table:

| $H_0$ | $H_1$ | $K_\alpha$ |
|---|---|---|
| $\sigma^2 = \sigma_0^2$ | $\sigma^2 \neq \sigma_0^2$ | $\left(0; \chi^2_{\frac{\alpha}{2}}(n-1)\right) \cup \left(\chi^2_{1-\frac{\alpha}{2}}(n-1); \infty\right)$ |
| $\sigma^2 = \sigma_0^2$ | $\sigma^2 > \sigma_0^2$ | $\left(\chi^2_{1-\alpha}(n-1); \infty\right)$ |
| $\sigma^2 = \sigma_0^2$ | $\sigma^2 < \sigma_0^2$ | $(0; \chi^2_{\alpha}(n-1))$ |

**Example 5.4** *From a box, 9 parts were randomly selected, whose lengths [in cm] were 10.24; 10.12; 9.91; 10.19; 9.78; 10.14; 9.86; 10.17; 10.05. We assume that the random sample was drawn from a population with a normal distribution, whose variability of the part length is expressed by a standard deviation of 0.1 cm.*

*a) At the significance level $\alpha$=0.05, let us test whether the variability of the part length differs significantly from the stated variability.*

*b) At the significance level $\alpha$=0.05, let us test whether the variability of the part length is smaller than the stated variability.*

*c) At the significance level $\alpha$=0.05, let us test whether the variability of the part length is larger than the stated variability.*

**Solution:** For testing we use the one-sample test of the variance $\sigma^2$, where we do not know the mean value $\mu$. The sample size is $n = 9$. We are given the standard deviation $\sigma_0 = 0,1$

cm, hence $\sigma_0^2 = 0,1^2 = 0,01$.

a) We will test

$$H_0 : \sigma^2 = 0,01,$$
$$H_1 : \sigma^2 \neq 0,01.$$

The null hypothesis states that the variability of the part length does not differ significantly from the required variability. Against the alternative hypothesis, which states that the variability of the part length differs significantly from the stated variability value. We determine the value of the test statistic $\chi^2$ using relation (99)

$$\chi^2 = \frac{8 \cdot s^2}{0,01},$$

where the sample arithmetic mean $\overline{x}$ is calculated using relation (71)

$$\overline{x} = \frac{1}{9}(10,24 + 10,12 + 9,91 + 10,19 + 9,78 + 10,14 + 9,86 + 10,17 + 10,05) = 10,0511$$

and the sample variance $s^2$ is calculated using relation (73)

$$s^2 = \frac{1}{8}\big[(10,24 - 10,0511)^2 + (10,12 - 10,0511)^2 + (9,91 - 10,0511)^2 + (10,19 - 10,0511)^2 +$$
$$+(9,78 - 10,0511)^2 + (10,14 - 10,0511)^2 + (9,86 - 10,0511)^2 + (10,17 - 10,0511)^2 +$$
$$+(10,05 - 10,0511)^2\big] = 0,0265.$$

The value of the test statistic is then $\chi^2 = 21,1689$. The quantiles of the $\chi^2$-distribution are $\chi^2_{1-\frac{\alpha}{2}}(n-1) = \chi^2_{0,975}(8) = 17,53$ and $\chi^2_{\frac{\alpha}{2}}(n-1) = \chi^2_{0,025}(8) = 2,18$ (Table 30). The critical region is

$$K_{0,05} = (0; 2,18) \cup (17,53; \infty) \Rightarrow \chi^2 \in K_{0,05}.$$

The value of the test statistic $\chi^2 = 21,1689$ falls into the critical region $K_{0,05}$, therefore we reject the null hypothesis $H_0$ and accept the alternative hypothesis $H_1$. We can assume that, at the significance level $\alpha = 0,05$, the variability of the part length differs significantly from the stated value.

b) We will test

$$H_0 : \sigma^2 = 0,01,$$
$$H_1 : \sigma^2 < 0,01.$$

The null hypothesis states that the variability of the part length does not differ significantly from the required variability. Against the alternative hypothesis, which states that the variability of the part length is smaller than the stated variability value. We calculate the value

of the test statistic $\chi^2$ in the same way as in case a), and thus $\chi^2 = 21,1689$. The quantile of the $\chi^2$-distribution is $\chi^2_\alpha(n-1) = \chi^2_{0,05}(8) = 2,73$ (Table 30). The critical region is

$$K_{0,05} = (0; 2,73) \Rightarrow \chi^2 \notin K_{0,05}.$$

The value of the test statistic $\chi^2 = 21,1689$ does not fall into the critical region $K_{0,05}$, therefore we do not reject the null hypothesis $H_0$. We can assume that, at the significance level $\alpha = 0,05$, the variability of the part length is not smaller than the stated variability value.

c) We will test

$$H_0 : \sigma^2 = 0,01,$$
$$H_1 : \sigma^2 > 0,01.$$

The null hypothesis states that the lengths of the parts do not differ significantly from the required length. Against the alternative hypothesis, which states that the lengths of the parts are larger than the required length. We calculate the value of the test statistic $\chi^2$ in the same way as in case a), and thus $\chi^2 = 21,1689$. The quantile of the $\chi^2$-distribution is $\chi^2_{1-\alpha}(n-1) = \chi^2_{0,95}(8) = 15,51$ (Table 30). The critical region is

$$K_{0,05} = (15,51; \infty) \Rightarrow \chi^2 \in K_{0,05}.$$

The value of the test statistic $\chi^2 = 21,1689$ falls into the critical region $K_{0,05}$, we reject the null hypothesis $H_0$ and accept the alternative hypothesis $H_1$. We can assume that, at the significance level $\alpha = 0,05$, the variability of the part length is larger than the stated variability value.

## 5.2 One-sample parametric test of a parameter of the exponential distribution

Consider a random sample $V_n$ that comes from a population with an exponential distribution with parameter $\lambda$, $\lambda > 0$.

1. Hypotheses:

$$H_0 : \lambda = \lambda_0 \quad (\lambda_0 \text{ is a known constant}),$$
$$H_1 : \lambda \neq \lambda_0 \quad (\text{resp. } \lambda < \lambda_0, \text{ resp. } \lambda > \lambda_0).$$

2. The test statistic is

$$\chi^2 = \frac{2n\overline{x}}{\lambda_0} \sim \chi^2(2n), \tag{100}$$

where $\overline{x}$ is the sample arithmetic mean and $n$ is the size of the random sample.

3. The critical region $K_\alpha$ for various alternatives is given in the following table:

| $H_0$ | $H_1$ | $K_\alpha$ |
|---|---|---|
| $\lambda = \lambda_0$ | $\lambda \neq \lambda_0$ | $\left(0; \chi^2_{\frac{\alpha}{2}}(2n)\right) \cup \left(\chi^2_{1-\frac{\alpha}{2}}(2n); \infty\right)$ |
| $\lambda = \lambda_0$ | $\lambda > \lambda_0$ | $\left(\chi^2_{1-\alpha}(2n); \infty\right)$ |
| $\lambda = \lambda_0$ | $\lambda < \lambda_0$ | $(0; \chi^2_\alpha(2n))$ |

**Example 5.5** *For 100 randomly selected passengers, the waiting time for the bus [in min.] was observed. The results are given in the table:*

| *average waiting time $x_j$ [in min.]* | 0,5 | 2,5 | 4,5 | 6,5 | 8,5 |
|---|---|---|---|---|---|
| *number of people $n_j$* | 60 | 20 | 11 | 5 | 4 |

*We assume that the waiting time has an exponential distribution with parameter $\lambda$. The carrier claims that the average waiting time for the bus is two minutes. Let us decide whether they are telling the truth.*

*a) At the significance level $\alpha$=0.05, let us test whether the average waiting time of the selected passengers differs significantly from the waiting time given by the carrier.*

*b) At the significance level $\alpha$=0.01, let us test whether the average waiting time of the selected passengers is greater than the waiting time given by the carrier.*

*c) At the significance level $\alpha$=0.01, let us test whether the average waiting time of the selected passengers is smaller than the waiting time given by the carrier.*

**Solution:** For testing we use the one-sample parametric test of a parameter of the exponential distribution. The sample size is $n = 100$.

a) We will test

$$H_0 : \lambda = 2,$$
$$H_1 : \lambda \neq 2.$$

The null hypothesis states that the average waiting time $\lambda$ does not differ significantly from the waiting time given by the carrier. Against the alternative hypothesis, which states that the average waiting time differs significantly from the waiting time given by the carrier. We determine the value of the test statistic $\chi^2$ using relation (100)

$$\chi^2 = \frac{200 \cdot \overline{x}}{2},$$

where the sample arithmetic mean is calculated using relation (72)

$$\overline{x} = \frac{1}{100}\big(0,5 \cdot 60 + 2,5 \cdot 20 + 4,5 \cdot 11 + 6,5 \cdot 5 + 8,5 \cdot 4\big) = 1,96.$$

The value of the test statistic is then $\chi^2 = 196$. The quantiles of the $\chi^2$-distribution are $\chi^2_{1-\frac{\alpha}{2}}(2n) = \chi^2_{0,975}(200) = 241,06$ and $\chi^2_{\frac{\alpha}{2}}(2n) = \chi^2_{0,025}(200) = 162,73$ (Table 32). The critical region is

$$K_{0,05} = (0; 162,73) \cup (241,06; \infty) \Rightarrow \chi^2 \notin K_{0,05}.$$

The value of the test statistic $\chi^2 = 196$ does not fall into the critical region $K_{0,05}$, therefore we do not reject the null hypothesis $H_0$. We can assume that, at the significance level $\alpha = 0,05$, the average waiting time does not differ significantly from the waiting time given by the carrier.

b) We will test

$$H_0 : \lambda = 2,$$
$$H_1 : \lambda > 2.$$

The null hypothesis states that the average waiting time $\lambda$ does not differ significantly from the required time. Against the alternative hypothesis, which states that the average waiting time is greater than the waiting time given by the carrier. We calculate the value of the test statistic $\chi^2$ in the same way as in case a), and thus $\chi^2 = 196$. The quantile of the $\chi^2$-distribution is $\chi^2_{1-\alpha}(2n) = \chi^2_{0,99}(200) = 249,45$ (Table 32). The critical region is

$$K_{0,01} = (249,45; \infty) \Rightarrow \chi^2 \notin K_{0,01}.$$

The value of the test statistic $\chi^2 = 196$ does not fall into the critical region $K_{0,01}$, therefore we do not reject the null hypothesis $H_0$. We can assume that, at the significance level $\alpha = 0,01$, the average waiting time is not greater than the waiting time given by the carrier.

c) We will test

$$H_0 : \lambda = 2,$$
$$H_1 : \lambda < 2.$$

The null hypothesis states that the average waiting time $\lambda$ does not differ significantly from the required time. Against the alternative hypothesis, which states that the average waiting time is smaller than the waiting time given by the carrier. We calculate the value of the test statistic $\chi^2$ in the same way as in case a), and thus $\chi^2 = 196$. The quantile of the $\chi^2$-distribution is $\chi^2_{\alpha}(2n) = \chi^2_{0,01}(200) = 156,43$ (Table 32). The critical region is

$$K_{0,01} = (0; 156,43) \Rightarrow \chi^2 \notin K_{0,01}.$$

The value of the test statistic $\chi^2 = 196$ does not fall into the critical region $K_{0,01}$, therefore we do not reject the null hypothesis $H_0$. We can assume that, at the significance level $\alpha = 0,01$, the average waiting time is not smaller than the waiting time given by the carrier.

## 5.3 Two-sample parametric tests of parameters of the normal distribution

Consider two random samples $V_{n_1}$ and $V_{n_2}$. The random sample $V_{n_1}$ has size $n_1$ and comes from a population with a normal distribution with mean value $\mu_1$ and standard deviation $\sigma_1$. We write $V_{n_1} \sim N(\mu_1, \sigma_1^2)$. The random sample $V_{n_2}$ has size $n_2$ and comes from a population with a normal distribution with mean value $\mu_2$ and standard deviation $\sigma_2$. We write $V_{n_2} \sim N(\mu_2, \sigma_2^2)$.

When testing we must take into account whether the random samples come from dependent or independent populations. In the case of *independent populations* we assume that a sample from one population does not depend on the sample from the other population. Conversely, for *dependent populations* we assume that a sample from one population depends on the sample from the other population.

If we do not reject the null hypothesis about the equality of the given parameters, there is no statistically significant difference between the tested parameters. The general procedure for choosing a test of mean values for two samples is shown in Fig. 23.

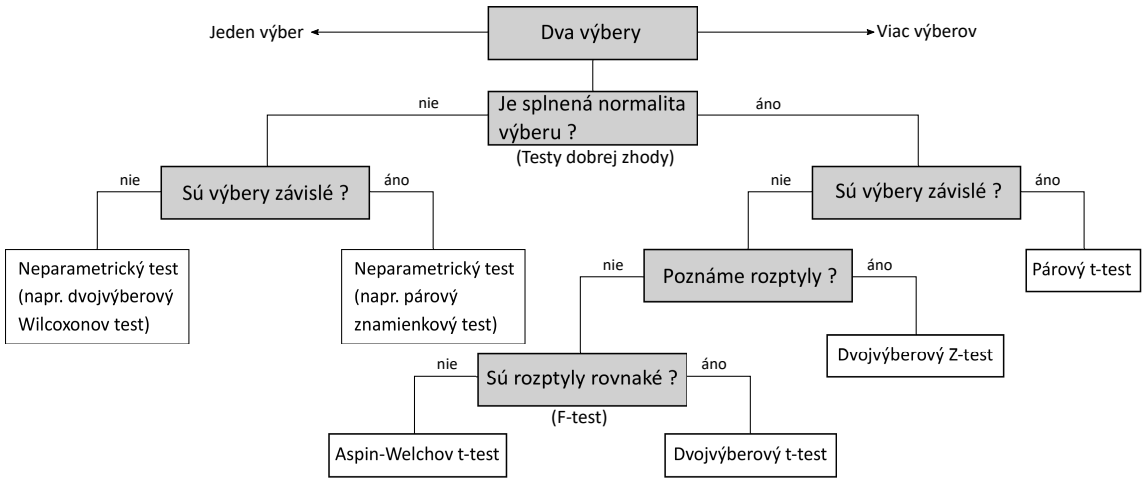



**Figure 23:** Procedure for choosing a test of mean values for two samples