Mohamed Khider University of Biskra
Faculty of Natural and Life Sciences and Earth Sciences — Second year license
Module: Biostatistics — 2025/2026

Solutions TD 03

Exercise Solution 1

1. The point estimates of the mean and the standard deviation of X are given by:

$$\begin{aligned}\overline{X} &= \frac{1}{n}\sum_{i=1}^{n} x_i = 102.7333\\ \hat{\sigma}_c^2 &= \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\overline{X})^2 = 598.9238 \text{ (since the true mean, } \mu\text{, is unknown)}\\ \hat{\sigma}_c &= \sqrt{\hat{\sigma}_c^2} = \sqrt{598.9238} = 24.4729\end{aligned}$$

2. In this question, the formulation of the test to be performed is as follows:

$$H_0 : \ ''\mu = \mu_0'' \text{ versus } H_1 : \ ''\mu \neq \mu_0'.$$

Given the previous sample (sample size < 30, X follows a normal distribution, and the true variance is unknown), the appropriate test is Student's t-test. We have:
• On one hand, the test statistic:

$$t = \frac{\overline{X}-\mu_0}{\hat{\sigma}_c/\sqrt{n}} = \frac{102.7333-110}{24.4729/\sqrt{15}} = -1.1500,$$

• On the other hand, from the Student distribution table:

$$t_\alpha = t_{(n-1,1-\alpha/2)} = t_{(14,1-0.02/2)} = 2.625.$$

Thus, we observe that $t \in [-t_\alpha, t_\alpha]$ (i.e., $-1.1500 \in [-2.625, 2.625]$), which implies that H_0 is accepted. In other words, at a significance level of 2%, the mean height of the children is equal to 110 cm.

Exercise Solution 2 In order to answer the questions of the exercise, we will need the following quantities:
The mean:

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{7}(83+96+99+110+130+95+74) = 98.1429,$$

The variance is known $\sigma = 324$.

1. *The formulation of the test to be carried out in this case is:*

$$H_0 "\mu = \mu_0" \text{ versus } H_1 "\mu < \mu_0", \tag{1}$$

more precisely:

$$H_0 "\mu = 100" \text{ versus } H_1 "\mu < 100". \tag{2}$$

(a) *The test statistic is:* $U = \frac{\overline{X}-\mu_0}{\sqrt{\hat{\sigma}^2/n}}$.

(b) *Its value is:*

$$u = \frac{98.1429-100}{\sqrt{324/7}} = -0.27.$$

(c) We determine u_α from the standard normal table such that:
$P(U < u_\alpha) = 1 - \alpha = 0.95$, hence $u_\alpha = 1.64$, and we decide that:
If $u > -u_\alpha$, then we do not reject H_0;
If $u < -u_\alpha$, then we reject H_0 with a probability $\alpha = 5\%$ of making an error.
In our case, $u = -0.27 > -1.64$, therefore we do not reject H_0 with a probability $\alpha = 5\%$ of error.

Exercise Solution 3

1. Determine a point estimate of the mean and the variance of each sample.
   We have:

$$\begin{aligned} \overline{X} &= \frac{1}{n_1}\sum_{i=1}^{n_1} x_i = 24.5 \quad \text{and} \quad \overline{Y} = \frac{1}{n_2}\sum_{i=1}^{n_2} y_i = 22 \\ \hat{\sigma}_{c,1}^2 &= \frac{1}{n_1 - 1}\sum_{i=1}^{n_1}(x_i - \overline{X})^2 = 0.7880 \text{ and } \hat{\sigma}_{c,2}^2 = \frac{1}{n_2 - 1}\sum_{i=1}^{n_2}(y_i - \overline{Y})^2 = 0.8700 \end{aligned}$$

2. Suppose we want to determine whether the two types of trees have the same mean height.

   a) Give the form of the test to be performed in this case.
   The form of the test is:

$$H_0 : \text{"}\mu_1 = \mu_2\text{"} \text{ versus } H_1 : \text{"}\mu_1 \neq \mu_2\text{"}, \qquad (*)$$

   b) State the necessary conditions to perform this test.
   1) The two samples come from a normal distribution.
   2) The two samples are mutually independent.
   3) The two samples have the same variance, i.e., we must first perform the following test:

$$H_0 : \text{" } \sigma_1^2 = \sigma_2^2 \text{ "} \quad \text{versus} \quad H_1 : \text{" } \sigma_1^2 \neq \sigma_2^2 \text{ "}.$$

   If the two samples have the same variance at a significance level α, we compute the common variance $\hat{\sigma}_c^2$ of the two samples defined by:

$$\hat{\sigma}_c^2 = \frac{(n_1 - 1)\hat{\sigma}_{c,1}^2 + (n_2 - 1)\hat{\sigma}_{c,2}^2}{n_1 + n_2 - 2}.$$

3. If the conditions in 2.*a*) are satisfied, then:
   The two samples come from a normal distribution, are mutually independent, and have the same variance. Therefore, test $(*)$ is performed using Student's t-test for equality of means.

   a) Give the test statistic defined in 2.*a*).
   The test statistic is:

$$T = \frac{\overline{X} - \overline{Y}}{\hat{\sigma}_c\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}.$$

   b) Compute the value of the test statistic, knowing that $\hat{\sigma}_c^2$ is the common variance defined by:

$$\hat{\sigma}_c^2 = \frac{(n_1 - 1)\hat{\sigma}_{c,1}^2 + (n_2 - 1)\hat{\sigma}_{c,2}^2}{n_1 + n_2 - 2} = \frac{(6-1)0.7880 + (5-1)0.8700}{6 + 5 - 2} = 0.8244.$$

   Thus, the value of the statistic T is:

$$t = \frac{24.5 - 22}{0.9080\sqrt{\frac{1}{6} + \frac{1}{5}}} = 4.5469.$$

c) Determine the critical value of the test for a significance level $\alpha = 2\%$.
The critical value is the quantile of order $1 - \alpha/2 = 1 - 0.02/2$ of a Student distribution with degrees of freedom $n_1 + n_2 - 2 = 9$, hence $t_\alpha = 2.821$.

Exercise Solution 4

a) In this exercise, we will focus on a more general case of comparing means when the sample size is strictly greater than two. More specifically, we will examine one-way analysis of variance (ANOVA1), which is the most appropriate in this case. To achieve our objective, we will perform the following test:

$$H_0 : \,''\mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu'' \text{ contre } H_1 : \,''\exists\, i, j \in \{1, 2, 3, 4\} \text{ tel que } \mu_i \neq \mu_j;''.$$

b) The conditions necessary to perform this test are:

- The $p = 4$ samples being compared are independent..
- The quantitative variable studied follows a normal distribution in the $p = 4$ populations compared.
- The $p = 4$ populations compared have the same variance: Homogeneity of variances.