8. Confidence interval estimation

In confidence interval parameter estimation based on the sample $\boldsymbol{V} = (X_1, X_2, \dots, X_n)$, a **confidence interval** $[l, u]$ is determined for which it is trusted with a **confidence coefficient** $(1-\alpha)$ or a **risk coefficient** α that it contains the true value of the estimated parameter q: $P(l \le q \le u) = 1-\alpha$. Confidence intervals can be **two-sided** $(l \le q \le u)$ or **left** $(l \le q)$ and **right** $(q \le u)$ **one-sided**. **Error** of interval estimation is $|l-q|$ or $|u-q|$.

Mean m: distribution of X normal, σ known $\rightarrow Z = \frac{\langle X\rangle_n - m}{\sigma/\sqrt{n}}$, $z_{\alpha/2}$: $\Phi(z_{\alpha/2}) = (1-\alpha)/2$

$$\langle X\rangle_n - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < m < \langle X\rangle_n + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

Mean m: distribution of X any, σ unknown, $n > 30 \rightarrow Z = \frac{\langle X\rangle_n - m}{S/\sqrt{n}}$

$$\langle X\rangle_n - z_{\alpha/2}\frac{S}{\sqrt{n}} < m < \langle X\rangle_n + z_{\alpha/2}\frac{S}{\sqrt{n}}$$

Mean m: distribution of X normal, σ unknown, $n < 30 \rightarrow T = \frac{\langle X\rangle_n - m}{S/\sqrt{n}}$, $t_{n-1;\,\alpha/2}$: from the table

$$\langle X\rangle_n - t_{n-1;\,\alpha/2}\frac{S}{\sqrt{n}} < m < \langle X\rangle_n + t_{n-1;\,\alpha/2}\frac{S}{\sqrt{n}}$$

Variance σ^2: distribution of X normal $\rightarrow \chi^2 = \frac{(n-1)S^2}{\sigma^2}$, $\chi^2_{n-1;\,\alpha/2}$, $\chi^2_{n-1;\,1-\alpha/2}$: from the table

$$\frac{(n-1)S^2}{\chi^2_{n-1;\,\alpha/2}} < \sigma^2 < \frac{(n-1)S^2}{\chi^2_{n-1;\,1-\alpha/2}}$$

Proportion of the population p: distribution of X binomial, it can be approximated by a normal

$$\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} < p < \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

Sum/difference of means $m_1 \pm m_2$: distributions of X_1, X_2 normal, σ_1, σ_2 known

$$\langle X_1\rangle_n \pm \langle X_2\rangle_n - z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} < m_1 \pm m_2 < \langle X_1\rangle_n \pm \langle X_2\rangle_n + z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

Sum/diff. of means $m_1 \pm m_2$: X_1, X_2 normal, σ_1, σ_2 unknown, $n_1, n_2 < 30$; $S_p = \sqrt{\frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}}$

$$\langle X_1\rangle_n \pm \langle X_2\rangle_n - t_{n_1+n_2-2;\,\alpha/2} S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} < m_1 \pm m_2 < \langle X_1\rangle_n \pm \langle X_2\rangle_n + t_{n_1+n_2-2;\,\alpha/2} S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

Sum/difference of population proportions $p_1 \pm p_2$: distributions of X_1, X_2 binomial, can be approximated by normal

$$\hat{p}_1 \pm \hat{p}_2 - z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} < p_1 \pm p_2 < \hat{p}_1 \pm \hat{p}_2 + z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

8. Confidence interval estimation – problems

1. Flow time of certain product has been measured in a workshop for ten selected pieces. The resulting values were (in minutes): 17, 21, 14, 23, 20, 24, 19, 19, 25 and 18. It is assumed that the flow time of the studied product is normally distributed. Determine point estimators of mean and standard deviation of the flow time. R: $\hat{m}$ = 20 min, $\hat{\sigma}$ = 3.37 min

2. It is assumed that the standard deviation of the flow time is 3.4 min.

   a. Based on the measurements in the above problem estimate the confidence interval on mean of the flow time. R: 17.9 min ≤ m ≤ 22.1 min at α = 0.05

   b. What should be the sample size to have the error of confidence interval estimation below 1 min? R: n ≥ 45 at α = 0,05

3. Let us assume that the standard deviation of the flow time is not known. Based on the measurements in Problem 1 estimate the confidence intervals on mean and on standard deviation of the flow time. R: 17.6 min ≤ m ≤ 22.4 min and 2.32 min ≤ σ ≤ 6.15 min at α = 0,05

4. Using the spot welding machine 100 spots are welded among which 42 weld spots are defective.

   a. Determine point estimator and the confidence interval on the proportion of the defective weld spots. R: $\hat{p}$ = 0.42 and 0.32 ≤ p ≤ 0.52

   b. What should be the sample size to have the error of confidence interval estimation below 0.05? R: n ≥ 375 at α = 0.05

5. In turning of the product the technology requires a successive application of two different turning inserts. From a sample of 40 inserts of type A and 50 inserts of type B it is found that the insert A can remove on average 1000 mm^3 of material with standard deviation of 150 mm^3 before it must be replaced, while the insert B can remove on average 1400 mm^3 of material with standard deviation of 200 mm^3. Estimate the confidence interval on the average of the total volume of material removed using the randomly selected inserts of type A nad B.
   R: 2328 $mm^3 \leq \mu_A + \mu_B \leq$ 2472 mm^3

6. Steel rollers are manufactured by extrusion through dies A and B. On a sample of fifteen rollers extruded through the die A an average diameter of 28.7 mm with standard deviation of 0.45 mm is determined while on a sample of twelve rollers extruded through the die B an average diameter of 27.9 mm with standard deviation of 0.50 mm is determined. Estimate the confidence interval on the difference of the average roller diameters. Assume that both diameters of the extruded rollers are normally distributed.
   R: 0.423 mm $\leq \mu_A - \mu_B \leq$ 1.177 mm at $\alpha = 0.05$

8. Confidence interval estimation – additional problems

1. Roughness (R_a) of a turned surface has been measured at nine locations on the shaft. The following values were obtained (in μm): 7, 12, 11, 5, 12, 9, 7, 8 and 10. We assume that the surface roughness is normally distributed. Determine point estimators of roughness mean and standard deviation. R: $\hat{m}$ = 9 μm, $\hat{\sigma}$ = 2.45 μm

2. It is assumed that the standard deviation of roughness is 2.5 μm.

    a. Based on the measurements in the above problem estimate the confidence interval on mean of the roughness. R: 7.4 μm ≤ m ≤ 10.6 μm at α = 0.05

    b. What should be the sample size to have the error of confidence interval estimation below 1 μm? R: n ≥ 24 at α = 0.05

3. Let us assume that the standard deviation of roughness in unknown. Based on the measurements in Problem 1 estimate the confidence interval on mean and standard deviation of roughness. R: 7.1 μm ≤ m ≤ 10.9 μm in 1.7 μm ≤ σ ≤ 4.7 μm pri α = 0.05

4. At the end of production line the products are inspected. Among the 200 inspected 120 are flawless and the rest had to be repaired.

    a. Determine point estimator and confidence interval of flawless products proportion. R: $\hat{p}$ = 0.60 and 0.53 ≤ p ≤ 0.67

    b. What should be the sample size to have the error of confidence interval estimation below 0.03? R: n ≥ 1025 at α = 0.05