Population parameter – parameter základného súboru

Definition:
A population characteristics, such as population mean, population variance and population standard deviation.

Sample statistics – výberová štatistika

Definition:
A sample characteristics, such as sample mean, sample variance and sample standard deviation. **The value of a sample statistics is used to estimate the value of the population parameter.**

Central limit theorem – centrálna limitná teoréma

Definition:
A theorem that allows us to use the normal probability distribution to approximate the sampling distribution of a sample mean and sample standard deviation whenever the sample size is large.

Point estimate – bodový odhad

Definition:
Point estimate is a single numerical value used as an estimate of a population parameter.

Note:
Point estimator – sample statistics, such as $\bar{x}$, s_1, s_1^2 *that provides a point estimate of the population parameter. Point estimate is more precise than interval estimate but it is not so reliable.*

Properties of point estimators – vlastnosti bodového odhadu

Definition:
Properties of point estimators are as follows:

- **unbiasedness –** a property of point estimator that occurs whenever the expected value of the point estimator is equal to the population parameter it estimates.
- **consistency –** a property of point estimator that occurs whenever larger sample sizes tend to provide point estimates closer to the population parameter.
- **(relative) efficiency –** if we have two unbiased point estimators of the same population parameter, the point estimator with the smaller variance is said to have greater efficiency than the other.

Point estimate of mean – bodový odhad priemeru (strednej hodnoty)

$$\boxed{est\ \mu = \bar{x}}$$

As for a point estimate, population mean (μ) can be estimated by means of a sample mean ($\bar{x}$). A sample mean can be calculated using a function AVERAGE in MS Excel.

Point estimate of variance – bodový odhad rozptylu

$$\boxed{est\ \sigma^2 = s_1^2}$$

As for a point estimate, population variance (σ^2) can be estimated by means of a sample variance (s_1^2). A sample variance can be calculated using a function VAR in MS Excel.

Point estimate of standard deviation – bodový odhad smerodajnej (štandardnej) odchýlky

$$\boxed{est\ \sigma = s_1}$$

As for a point estimate, population standard deviation (σ) can be estimated by means of a sample standard deviation (s_1). A sample standard deviation can be calculated using a function STDEV in MS Excel.

Interval estimate – intervalový odhad

Definition:
Interval estimate is an estimate of population parameter that provides an interval believed to contain the value of the parameter. Interval estimate is not as precise as a point estimate but it is more reliable.

Confidence level – hladina spoľahlivosti

Definition:
The confidence associated with an interval estimate (usually 0.95 or 0.99, it depends on the significance level, i.e. alpha 0.05 or 0.01).

Confidence interval – interval spoľahlivosti

Definition:
An interval estimate for an unknown population parameter.

The width of the confidence interval is related to the significance level, standard error, and n (number of observations) such that the following are true:

- **the higher the percentage of accuracy (significance) desired, the wider the confidence interval**
- **the larger the standard error, the wider the confidence interval**
- **the larger the n, the smaller the standard error, and so the narrower the confidence interval**

All other things being equal, a smaller confidence interval is always more desirable than a larger one because a smaller interval means the population parameter can be estimated more accurately.

Interval estimate of mean - intervalový odhad priemeru (strednej hodnoty)

$$P\left(\bar{x}-\Delta<\mu<\bar{x}+\Delta\right)=1-\alpha$$

where Δ is the sampling error.

- If the sample size is greater than 30 (n>30) then the distribution of the random variable (sample statistic) will be approximated with a normal distribution – N(0,1). The sampling error will be calculated as follows:

$$\Delta = u_{(1-\alpha/2)} \cdot \frac{s_1}{\sqrt{n}}.$$

A critical value ($u_{(1-\alpha/2)}$) will be calculated using a function NORMSINV in MS Excel.

- If the sample size is lower than 30 (n<30) then the distribution of the random variable (sample statistic) will be approximated with a Student t distribution). The sampling error will be calculated as follows:

$$\Delta = t_{(\alpha;n-1)} \cdot \frac{s_1}{\sqrt{n}}.$$

A critical value ($t_{(\alpha;n-1)}$) will be calculated using a function TINV in MS Excel.

Interval estimate of variance - intervalový odhad rozptylu

$$P\left(\frac{(n-1)\cdot s_1^2}{\chi^2_{(\alpha/2;n-1)}}<\sigma^2<\frac{(n-1)\cdot s_1^2}{\chi^2_{(1-\alpha/2;n-1)}}\right)=1-\alpha$$

Critical values for both lower and upper limits ($\chi^2_{(\alpha/2;n-1)}; \chi^2_{(1-\alpha/2;n-1)}$) can be calculated using a function CHIINV in MS Excel.

Interval estimate of standard deviation - intervalový odhad smerodajnej (štandardnej) odchýlky

$$P\left(\sqrt{\frac{(n-1)\cdot s_1^2}{\chi^2_{(\alpha/2;n-1)}}}<\sigma^2<\sqrt{\frac{(n-1)\cdot s_1^2}{\chi^2_{(1-\alpha/2;n-1)}}}\right)=1-\alpha$$

Critical values for both lower and upper limits ($\chi^2_{(\alpha/2;n-1)}; \chi^2_{(1-\alpha/2;n-1)}$) can be calculated using a function CHIINV in MS Excel.

Note: It's much faster to calculate the interval estimate for variance and then calculate the square root of both lower and upper limits 🙂.

Sampling error (standard error, estimation error) – prípustná chyba odhadu

$$\Delta = u_{(1-\alpha/2)} \cdot \frac{s_1}{\sqrt{n}} \quad \text{or} \quad \Delta = t_{(\alpha;n-1)} \cdot \frac{s_1}{\sqrt{n}}$$

The lower the desired sampling error the larger the sample size must be.

Note: Sampling error is influenced by:

- confidence level - we can influence the confidence level
- (sample) standard deviation - we can't influence the variability within the sample (chosen from the collected set of data)
- sample size - we can influence the sample size

Sample size – veľkosť výberového súboru

$$n = u^2_{(1-\alpha/2)} \cdot \frac{s_1^2}{\Delta^2}$$

where,
Δ and s_1 are given in the text of the example.

A critical value ($u_{(1-\alpha/2)}$) will be calculated using a function NORMSINV in MS Excel.