

## S 2 Charakteristika základních výběrových technik a teoretická rozložení četností

### TEORIE

Základním typem úvahy ve statistice bývá úsudek z části na celek čili z určitého, tzv. výběrového souboru na soubor základní. *Jinými slovy – na to abychom poznali, zda je víno v sudu kvalitní, nemusíme vypít celý sud.*

**Základní soubor** ..... souhrn všech jedinců u kterých bychom měli šetření provádět (např.  $X_i$  dětí pátých tříd v ČR)

**Výběrový soubor** ..... na základě randomizace (náhodného výběru) získáme omezený počet jedinců  $x_i$ , kteří reprezentují vlastnosti a charakteristiky **celého** základního výběru. Náhodný výběr získáme losováním, pomocí tabulky náhodných čísel, nebo použitím generátoru náhodných čísel.

**Rozsah souboru** ..... počet prvků základního (N) a výběrového (n) souboru

Stanovení rozsahu náhodného výběru:

Hlavním požadavkem na výběrové šetření mimo jeho reprezentativnost je odpovídající rozsah výběru (počet vybraných prvků). Vypočítá se podle vzorce:

$$n = t_p^2 \frac{\sigma^2}{\Delta^2}$$

Kde  $t_p^2 = 1,96$  při 0,05 nebo 2,58 při 0,01 hladině pravděpodobnosti

$$\sigma^2 = s^2 \frac{n}{n-1}$$

$\Delta$  je požadovaná přesnost měření (odhad). Je dána polovičním intervalem spolehlivosti:

$$\mu = \bar{x} \pm t_p \sqrt{\frac{s}{n-1}}$$

$\bar{x}$ ,  $s$ ,  $n$  jsou hodnoty získané v předvýzkumech

### PŘÍKLAD

Počet 12ti-letých chlapců v ČR je 45 000. Hodnoty předvýzkumu testování výkonnosti ve skoku dalekém jsou:  $n = 121$ ,  $\bar{x} = 169$ ,  $s = 20$ . Stanov počet prvků náhodného výběru, aby byla zajištěna reprezentativnost a výsledky byly statisticky významné pro základní soubor.

$$\mu = 169 \pm 1,96 \sqrt{\frac{20}{120}} = 169 \pm 0,8 \quad \text{tj: interval spolehlivosti je } 168,2 - 169,8 \doteq 2 \quad \Delta = 1$$

$$\sigma^2 = 400 \frac{121}{120} = 403,33 \quad n = 1,96^2 \frac{403,33}{1^2} = 1549,33$$

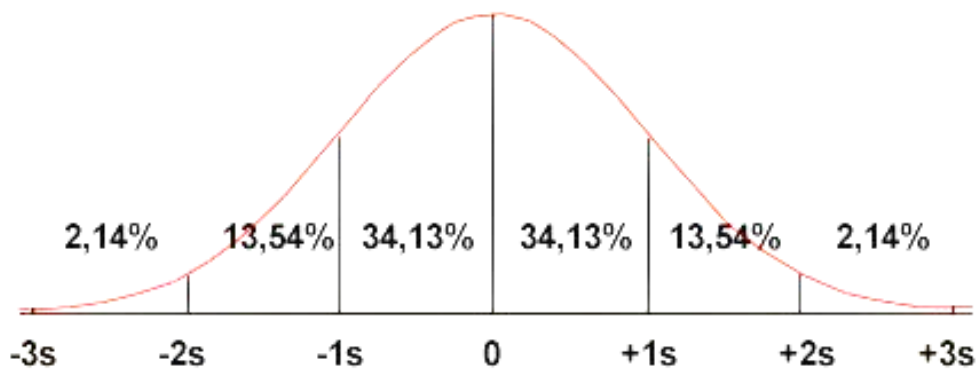
Výběrový soubor bude mít rozsah  $n = 1\,549$  probandů.

## TEORIE

### Teoretická rozložení četností

#### 1. Normální rozložení

##### Model normálního rozdělení četností



Znaky Gaussovy křivky:

- symetrická podle osy
- stejnoměrný zvonovitý tvar
- vrchol křivky je totožný s  $\bar{x}$ ,  $Mo$ ,  $Me$
- $R \doteq 6s$
- v intervalu  $x \pm 1s$  leží přibližně 68% všech případů
- v intervalu  $x \pm 2s$  leží přibližně 95% všech případů -
- v intervalu  $x \pm 3s$  leží přibližně 99% všech případů

**Normální rozložení četností je jedním z předpokladů použití parametrických statistických metod a postupů, které budou prezentovány v dalších částech.**

Existují další typy rozložení četností např:

- chí kvadrát rozložení
- F rozložení
- logaritmické rozložení

Pokud námi naměřená data vykazují tento typ rozložení, je nutné použít alternativních (neparametrických) metod.

## ÚKOL

Počet dětí osmých tříd základních škol v Ústeckém kraji je 8 740 (z toho 4 285 dívek). Hodnoty předvýzkumu testování výkonnosti v leh sedu chlapci:  $n = 138$   $\bar{x} = 39,9$   $s = 10,4$  dívky:  $n = 131$   $\bar{x} = 33,5$   $s = 7,5$ . Stanov počet prvků náhodného výběru, aby byla zajištěna reprezentativnost a výsledky byly statisticky významné pro základní soubor. Muži počítají hodnoty pro chlapce, ženy pro dívky.