

Střední chyba průměru

Střední chyba průměru (standardní chyba průměru, *standard error of the mean*, **SEM**, **SE**) je směrodatná odchylka náhodné veličiny **výběrový průměr** dat získaných jako náhodný výběr ze základního souboru.

Zjednodušeně řečeno je to číslo, které označuje, jak moc se asi námi získaný průměr náhodného výběru liší od střední hodnoty základního souboru.

Spočítá se jako

$$\text{SEM} = \frac{\sigma}{\sqrt{n}},$$

kde σ je směrodatná odchylka výběrového souboru a n je množství dat, z nichž průměr počítáme.

V praxi většinou není možné znát směrodatnou odchylku výběrového průměru, protože bychom k tomu museli znát celý základní soubor. Proto se jako střední chyba průměru často označuje její **odhad**, který získáme nahrazením směrodatné odchylky výběrového průměru σ směrodatnou odchylkou realizace náhodného výběru s (veličinu popisující základní soubor odhadneme pomocí čísla, které máme z našeho výběru):

$$\text{SEM} = \frac{s}{\sqrt{n}}.$$

Výběrový průměr jako náhodná veličina

Vezmeme-li z populace náhodné pacienty a budeme u nich měřit nějakou veličinu, řekněme krevní tlak, můžeme z ní vypočítat průměr, což je poměrně rutinní záležitost. **Jedná se však o náhodnou veličinu**, protože stejně tak můžeme náhodně vybrat z populace další pacienty, změřit jim krevní tlak a s největší pravděpodobností získáme jiná čísla, která budou mít jiný průměr.

Výběrový průměr si proto představujeme jako realizaci náhodné veličiny, která by vznikla postupným náhodným vybíráním skupin vzorků ze základního souboru a počítáním průměru v každé jednotlivé skupině.

V praxi se ovšem vzorky nedělí do jednotlivých skupin (až na metodu zvanou *bootstrapping*), protože by jich pak bylo v každé skupině poměrně málo, ale zahrnou se do jedné velké skupiny, z níž se průměr počítá. Pokud tato skupina má n prvků, měli bychom průměr počítat ze souboru X_1, X_2, \dots, X_n . V praxi ale máme skupinu x_1, x_2, \dots, x_n , v níž každé x_i reprezentuje jen jedno číslo z X_i .

Proto číslo, které získáme jako

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

považujeme za realizaci náhodné veličiny výběrového průměru.

Pokud má základní soubor rozptyl σ^2 , pro rozptyl našeho výběrového průměru potom dostáváme:

$$\text{var}(\bar{X}) = \frac{\sigma^2}{n}$$

Směrodatná odchylka našeho výběrového průměru a tím pádem i střední chyba průměru pak bude odmocnina rozptylu:

$$\text{SEM} = \sqrt{\frac{\text{var}(\bar{X})}{n}} = \sqrt{\frac{\sigma^2}{n}} = \frac{\sigma}{\sqrt{n}}.$$

Čím větší počet vzorků, tím menší SEM:

$$\lim_{n \rightarrow \infty} \text{SEM} = 0.$$

Výpočetní technika

- Výpočet SEM v programu Statistica: *Statistika* → *Základní statistika* (SEM není popisná statistika, ale dá se mezi základními statistikami najít, protože se počítá často)
- Výpočet SEM v programu Excel: *Nástroje* → *Analýza dat* → *Popisná statistika* (je nutné přidat modul popisných statistik v možnostech nástrojů)
- Výpočet SEM v programu R:

```
data <- c(1, 2, 3, 4, 5, 6, 7, 8, 9)
sd(data)/sqrt(length(data))
```

Odkazy

Související články

- Míry polohy
- Míry variability

Externí odkazy

- w:en:Bootstrapping (statistics)

Použitá literatura

- KLASCHKA, Jan. *Střední chyba průměru* [přednáška k předmětu Zdravotnická statistika 1,2, obor Všeobecné lékařství, 1. lékařská fakulta Univerzita Karlova]. Praha. 4.4.2011.
- ZVÁRA, Karel. *Biostatistika*. 2. vydání. Praha : Karolinum, 2004. 214 s. s. 74–75. ISBN 80-246-0739-5.