# Cvičení 3 – Základní pojmy matematické statistiky I

**Přehled vzorců:**

Výběr. průměr: $M = \frac{1}{n}\sum_{i=1}^{n} X_i$ ,výběr. rozptyl: $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - M)^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} X_i^2 - nM^2\right)$,

výběr. distr. funkce: $F_n(x) = \frac{1}{n}\operatorname{card}\{i; X_i \le x\}$, vážený průměr. výběrových rozptylů:

$S_*^2 = \frac{\sum_{j=1}^{r}(n_j - 1)S_j^2}{n - r}$, výběrová kovariance:

$S_{12} = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - M_1)(Y_i - M_2) = \frac{1}{n-1}\left(\sum_{i=1}^{n} X_i Y_i - nM_1M_2\right)$, výběrový koeficient korelace:

$R_{12} = \frac{S_{12}}{S_1 S_2}$

Vlastnosti M: E(M) = $\mu$, $D(M) = \frac{\sigma^2}{n}$. Pro výběr z N($\mu$, $\sigma^2$) platí: M ~ N($\mu$, $\frac{\sigma^2}{n}$)

Meze 100(1-$\alpha$)% intervalu spolehlivosti pro střední hodnotu $\mu$ při známém rozptylu $\sigma^2$ jsou:

$D = M - \frac{\sigma}{\sqrt{n}} u_{1-\alpha/2}$, $H = M + \frac{\sigma}{\sqrt{n}} u_{1-\alpha/2}$.

Stanovení minimálního rozsahu výběru z normálního rozložení tak, aby šířka 100(1-$\alpha$)% intervalu spolehlivosti pro $\mu$ při známém $\sigma^2$ nepřesáhla $\Delta$: $n \ge \frac{4\sigma^2 u_{1-\alpha/2}^2}{\Delta^2}$

**Příklad 1.:** Ve 12 náhodně vybraných prodejnách ve městě byly zjištěny následující ceny určitého výrobku (v Kč): 102, 99, 106, 103, 96, 98, 100, 105, 103, 98, 104, 107. Těchto 12 hodnot považujeme za realizace náhodného výběru $X_1, \ldots, X_{12}$ z rozložení, které má střední hodnotu $\mu$ a rozptyl $\sigma^2$.

a) Vypočtěte realizaci výběrového průměru a výběrového rozptylu.

b) Najděte výběrovou distribuční funkci $F_{12}(x)$ a nakreslete její graf.

**Příklad 2.:** Máme k dispozici výsledky testů ze dvou předmětů zjištěné u osmi náhodně vybraných studentů určitého oboru.

| Číslo studenta | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| Počet bodů v 1. testu | 80 | 50 | 36 | 58 | 42 | 60 | 56 | 68 |
| Počet bodů ve 2. testu | 65 | 60 | 35 | 39 | 48 | 44 | 48 | 61 |

Vypočtěte a interpretujte výběrový koeficient korelace. Pro usnadnění výpočtů máte tyto součty: $\sum_{i=1}^{8} x_i = 450, \sum_{i=1}^{8} y_i = 400, \sum_{i=1}^{8} x_i^2 = 26684, \sum_{i=1}^{8} y_i^2 = 20836, \sum_{i=1}^{8} x_i y_i = 23214$

**Příklad 3.:** Je známo, že týdenní výdaje domácností na určité potravinářské zboží se řídí normálním rozložením se střední hodnotou 90 Kč a směrodatnou odchylkou 14 Kč. Jaká je pravděpodobnost překročení hranice 100 Kč pro průměrné výdaje pěti náhodně vybraných domácností?

**Příklad 4.:** Nechť $X_1, \ldots, X_n$ je náhodný výběr z rozložení N($\mu$, 0,04). Jaký musí být minimální rozsah výběru, aby šířka 95% intervalu spolehlivosti pro $\mu$ nepřesáhla číslo 0,16?