# Téma 4.: Výpočet šikmosti a špičatosti, opakované nezávislé pokusy

**Charakteristika nesymetrie dat:** šikmost $\alpha_3 = \dfrac{\frac{1}{n}\sum_{i=1}^{n}(x_i - m)^3}{s^3}$

Je-li rozložení dat symetrické kolem aritmetického průměru, pak $\alpha_3 = 0$.
Má-li rozložení dat prodloužený pravý konec, jde o kladně zešikmené rozložení, $\alpha_3 > 0$.
Má-li rozložení dar prodloužený levý konec, jde o záporně zešikmené rozložení, $\alpha_3 < 0$.

**Charakteristika koncentrace dat kolem průměru:** špičatost $\alpha_4 = \dfrac{\frac{1}{n}\sum_{i=1}^{n}(x_i - m)^4}{s^4} - 3$

Je-li rozložení dat normální, pak $\alpha_4 = 0$.

Je-li rozložení dat strmější než normální rozložení, pak $\alpha_4 > 0$.

Je-li rozložení dat plošší než normální rozložení, pak $\alpha_4 < 0$.

**Úkol 1.:** Je třeba si uvědomit, že průměr a rozptyl nepopisují rozložení četností jednoznačně. Existují datové soubory, které mají shodný průměr i rozptyl, ale přesto se jejich rozložení četností velmi liší. Tuto skutečnost dobře ilustruje následující příklad: Tři skupiny studentů o počtech 149, 69 a 11 odpovídaly při testu na 10 otázek. Znak X je počet správně zodpovězených otázek. Známe absolutní četnosti znaku X ve všech třech skupinách.

| č. sk. | X | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| 1 | 2 | 5 | 15 | 20 | 25 | 15 | 25 | 20 | 15 | 5 | 2 |
| 2 | 4 | 3 | 2 | 1 | 0 | 49 | 0 | 1 | 2 | 3 | 4 |
| 3 | 1 | 0 | 0 | 0 | 0 | 9 | 0 | 0 | 0 | 0 | 1 |

Vypočtěte průměr, rozptyl, šikmost a špičatost počtu správně zodpovězených otázek ve všech třech skupinách. Nakreslete sloupkové diagramy absolutních četností.

**Návod:** Načteme datový soubor body_ve_3_sk.sta, který má 4 proměnné X, SK1, SK2, SK3 a 11 případů. V 1. sloupci jsou varianty znaku X (tj. 0 až 10), v dalších sloupcích pak absolutní četnosti daného počtu bodů v jednotlivých skupinách.

V tabulce Popisné statistiky zadáme Proměnná X a klepneme na tlačítko vah, abychom program upozornili, že budeme pracovat s daty zadanými pomocí absolutních četností. Zadáme Proměnná vah SK1, zaškrtneme Stav Zapnuto, OK.

Ve volbě Popisné statistiky zaškrtneme Průměr, Rozptyl, Šikmost, Špičatost – Výpočet. Dále pro znak X nakreslíme sloupkový diagram. Tytéž úkoly provedeme s váhovými proměnnými SK2 a SK3.

1. skupina (X váženo pomocí SK1)

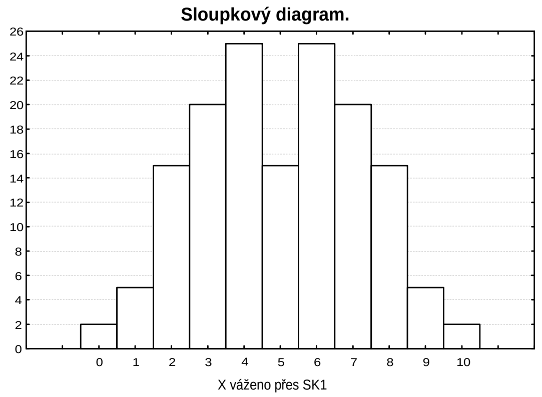
| | Popisné statistiky (tri_skupiny.sta) | | | |
|---|---|---|---|---|
| Proměnná | Průměr | Rozptyl | Šikmost | Špičatost |
| X | 5,000000 | 5,000000 | -0,000000 | -0,759500 |

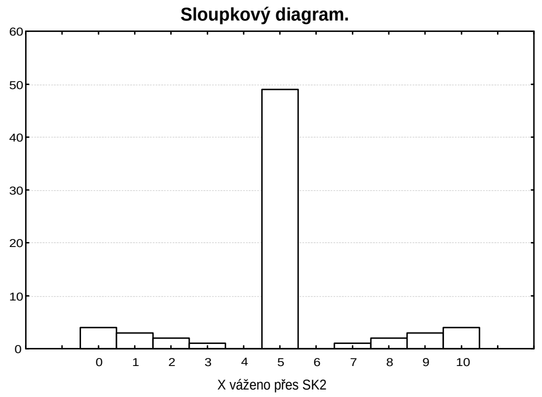
2. skupina (X váženo pomocí  SK2)

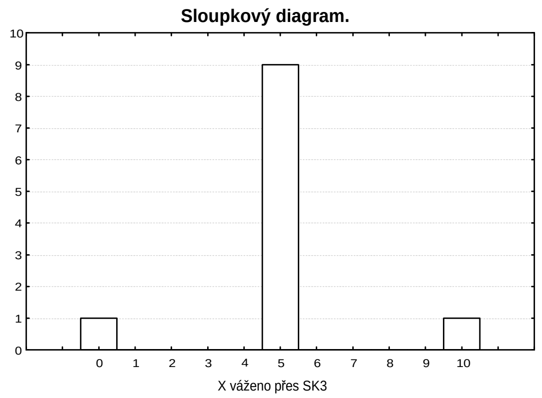
| Proměnná | Popisné statistiky (tri_skupiny.sta) | | | |
|---|---|---|---|---|
| | Průměr | Rozptyl | Šikmost | Špičatost |
| X | 5,000000 | 5,000000 | -0,000000 | 1,291133 |

3. skupina (X váženo pomocí SK3)

| Proměnná | Popisné statistiky (tri_skupiny.sta) | | | |
|---|---|---|---|---|
| | Průměr | Rozptyl | Šikmost | Špičatost |
| X | 5,000000 | 5,000000 | -0,000000 | 5,000000 |




Všechny tři skupiny mají týž průměr, rozptyl a šikmost, liší se pouze ve špičatosti. Sloupkové diagramy počtu správně zodpovězených otázek v každé ze tří uvažovaných skupin mají naprosto odlišný vzhled.

# Opakované nezávislé pokusy

Opakované nezávisle provádíme týž náhodný pokus a sledujeme nastoupení jevu, kterému říkáme úspěch. V každém z těchto pokusů nastává úspěch s pravděpodobností $\vartheta$, $0 < \vartheta < 1$.

**a) Binomické rozložení pravděpodobností**

Pravděpodobnost, že v prvních n pokusech úspěch nastane právě x-krát ( $0 \le x \le n$ ):

$$P_n(x) = \binom{n}{x} \vartheta^x (1-\vartheta)^{n-x}.$$

K výpočtu v systému STATISTICA slouží funkce Binom(x; $\vartheta$; n)

Pravděpodobnost, že v prvních n pokusech úspěch nastane nejvýše $x_1$-krát ( $0 \le x_1 \le n$ ):

$$\sum_{x=0}^{x_1} P_n(x).$$

K výpočtu v systému STATISTICA slouží funkce IBinom($x_1$; $\vartheta$; n)

Pravděpodobnost, že v prvních n pokusech úspěch nastane aspoň $x_0$-krát ( $0 \le x_0 \le n$ ):

$$\sum_{x=x_0}^{n} P_n(x).$$

Výpočet lze provést takto: 1 - IBinom($x_0$ - 1; $\vartheta$; n)

Pravděpodobnost, že v prvních n pokusech úspěch nastane aspoň $x_0$-krát a nejvýše $x_1$-krát:

$$\sum_{x=x_0}^{x_1} P_n(x).$$

Výpočet lze provést takto: IBinom($x_1$; $\vartheta$; n) - IBinom($x_0$ - 1; $\vartheta$; n)

**Vzorový příklad na binomické rozložení pravděpodobností:** Pravděpodobnost, že se pacient uzdraví z určitého druhu rakoviny, je rovna 0,4. Jestliže byla tato choroba diagnostikována u 15 pacientů, jaká je pravděpodobnost, že

a) se nejvýše 8 uzdraví,
b) se aspoň 10 uzdraví,
c) právě 5 se uzdraví,
d) počet uzdravených bude mezi třemi a osmi?

**Řešení:**

Počet pokusů: n = 15, pravděpodobnost úspěchu: $\vartheta$ = 0,4

ad a)

$$\sum_{x=0}^{x_1} P_n(x) = \sum_{x=0}^{8} P_{15}(x) = \sum_{x=0}^{8} \binom{15}{x} 0{,}4^x 0{,}6^{15-x} = 0{,}905$$

S pravděpodobností 90,5 % se z 15 pacientů uzdraví nejvýše 8.

ad b) $\displaystyle \sum_{x=x_0}^{n} P_n(x) = \sum_{x=10}^{15} P_{15}(x) = 1 - \sum_{x=0}^{9} P_{15}(x) = 1 - \sum_{x=0}^{9} \binom{15}{x} 0{,}4^x 0{,}6^{15-x} = 0{,}0338$

S pravděpodobností 3,38 % se z 15 pacientů uzdraví aspoň 10.

ad c)

$$P_n(x) = P_{15}(5) = \binom{15}{5} 0{,}4^5 0{,}6^{10} = 0{,}1859$$

S pravděpodobností 18,59 % se z 15 pacientů uzdraví právě 5.

ad d)

$$\sum_{x=x_0}^{x_1} P_n(x) = \sum_{x=3}^{8} P_{15}(x) = \sum_{x=0}^{8} P_{15}(x) - \sum_{x=0}^{2} P_{15}(x) = \sum_{x=0}^{8} \binom{15}{x} 0{,}4^x 0{,}6^{15-x} - \sum_{x=0}^{2} \binom{15}{x} 0{,}4^x 0{,}6^{15-x} =$$

$$= 0{,}8778$$

S pravděpodobností 87,78 % bude mezi 15 pacienty uzdravených od 3 do 8.

**Návod: Použití funkcí Binom a IBinom**

Otevřeme nový datový soubor se čtyřmi proměnnými a o jednom případu.
Do Dlouhého jména 1. proměnné napíšeme =IBinom(8;0,4;15).
Do Dlouhého jména 2. proměnné napíšeme =1-IBinom(9;0,4;15).
Do Dlouhého jména 3. proměnné napíšeme =Binom(5;0,4;15).
Do Dlouhého jména 4. proměnné napíšeme =IBinom(8;0,4;15)-IBinom(2;0,4;15).

| | Prom1 =IBinom(8;0,4;15) | Prom2 =1-IBinom(9;0,4;15) | Prom3 =Binom(5;0,4;15) | Prom4 =IBinom(8;0,4;15)-IBinom(2;0,4;15) |
|---|---|---|---|---|
| 1 | 0,904952592 | 0,0338333029 | 0,185937845 | 0,877838591 |

**Příklady k samostatnému řešení**

**Příklad 1.:** Pojišťovna zjistila, že 12 % pojistných událostí je způsobeno vloupáním. Jaká je pravděpodobnost, že mezi 30 náhodně vybranými pojistnými událostmi bude způsobeno vloupáním
a) nejvýše 6,
b) aspoň 6,
c) právě 6,
d) od dvou do pěti?
(Počet pokusů: n = 30, pravděpodobnost úspěchu: $\vartheta = 0{,}12$)
**Výsledek:** ad a) 0,9393, ad b) 0,1431, ad c) 0,0825, ad d) 0,7469

**Příklad 2.:** Je pravděpodobnější vyhrát se stejně silným soupeřem tři partie ze čtyř nebo pět partií z osmi, když nerozhodný výsledek je vyloučen a výsledky jsou nezávislé?
(Počet pokusů: n = 4 resp. n = 8, pravděpodobnost úspěchu: $\vartheta = 0{,}5$)
**Výsledek:** ad a) 0,250000, ad b) 0,218750

**Příklad 3.:** Dvacetkrát nezávisle na sobě házíme třemi mincemi. Jaká je pravděpodobnost, že alespoň v jednom hodě padnou tři líce?
(Počet pokusů: n = 20, pravděpodobnost úspěchu: $\vartheta = 0{,}125$)
**Výsledek:** 0,930791

**b) Geometrické rozložení pravděpodobností**

Pravděpodobnost, že prvnímu úspěchu bude předcházet x neúspěchů:

$P(x) = (1-\vartheta)^x \vartheta$.

K výpočtu v systému STATISTICA slouží funkce Geom(x; $\vartheta$)

Pravděpodobnost, že prvnímu úspěchu bude předcházet nejvýše $x_1$ neúspěchů: $\sum_{x=0}^{x_1} P(x)$

K výpočtu v systému STATISTICA slouží funkce IGeom($x_1$; $\vartheta$)

Pravděpodobnost, že prvnímu úspěchu bude předcházet aspoň $x_0$ neúspěchů: $1-\sum_{x=0}^{x_0-1} P(x)$

Výpočet lze provést takto: 1 - IGeom($x_0$-1; $\vartheta$)

**Vzorový příklad na geometrické rozložení pravděpodobností:** Jaká je pravděpodobnost, že při hře „Člověče, nezlob se!“ nasadíme figurku

a) právě při třetím hodu,

b) nejpozději při třetím hodu?

**Řešení:**

Ad a) Počet neúspěchů: x = 2, pravděpodobnost úspěchu: $\vartheta = \frac{1}{6}$

$$P(2) = (1-\vartheta)^2 \vartheta = \left(\frac{5}{6}\right)^2 \frac{1}{6} = \mathrm{Geom}(2;1/6) = 0{,}1157$$

Pravděpodobnost, že figurku nasadíme právě při třetím hodu, je 11,57 %.

Ad b) Počet neúspěchů: x = 0, 1, 2, pravděpodobnost úspěchu: $\vartheta = \frac{1}{6}$

$$\sum_{x=0}^{2} P(x) = \sum_{x=0}^{2} (1-\vartheta)^x \vartheta = \sum_{x=0}^{2} \left(\frac{5}{6}\right)^x \frac{1}{6} = \mathrm{IGeom}(2;1/6) = 0{,}4213$$

Pravděpodobnost, že figurku nasadíme nejpozději při třetím hodu, je 42,13 %.

| | Prom1<br>=Geom(2;1/6) | Prom2<br>=IGeom(2;1/6) |
|---|---|---|
| 1 | 0,115740741 | 0,421296296 |

**Příklady k samostatnému řešení:**

**Příklad 1.:** Studenti biologie zkoumají barvu očí octomilek. Pravděpodobnost, že octomilka má bílou barvu očí, je 0,25, červenou 0,75. Jaká je pravděpodobnost, že až čtvrtá zkoumaná octomilka má bílou barvu očí?

(Počet neúspěchů: x = 3, pravděpodobnost úspěchu: $\vartheta = 0{,}25$)

**Výsledek:** Pravděpodobnost, že až 4. zkoumaná octomilka má bílou barvu očí, je 10,55 %.

**Příklad 2.:** V určitém výrobním procesu je známo, že jeden z deseti výrobků nevyhovuje normám. Jaká je pravděpodobnost, že pátý náhodně vybraný výrobek je prvním, který nevyhovuje?
**Výsledek:** 0,0656