

Téma 4: Využití systému STATISTICA při řešení příkladů na opakované pokusy

1. 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 ϑ , $0 < \vartheta < 1$.

a) Binomické rozložení pravděpodobností

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

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

K výpočtu v systému STATISTICA slouží funkce Binom(x ; ϑ ; n)

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

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

K výpočtu v systému STATISTICA slouží funkce IBinom(x_1 ; ϑ ; n)

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

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

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

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: $\text{IBinom}(x_1; \vartheta; n) - \text{IBinom}(x_0 - 1; \vartheta; n)$

Příklad na binomické rozložení pravděpodobností: Pojišťovna zjistila, že 12% pojistných událostí je způsobeno vloupáním. Jaká je pravděpodobnost, že mezi 30 náhodně vybranými pojistnými událostmi bude způsobeno vloupáním

- nejvýše 6,
- aspoň 6,
- právě 6,
- od dvou do pěti?

Řešení:

Počet pokusů: $n = 30$, pravděpodobnost úspěchu: $\vartheta = 0,12$

ad a)

$$\sum_{x=0}^{x_1} P_n(x) = \sum_{x=0}^6 P_{30}(x) = \sum_{x=0}^6 \binom{30}{x} 0,12^x 0,88^{30-x} = \text{IBinom}(6;0,12;30) = 0,9393$$

S pravděpodobností 93,93% bude mezi 30 náhodně vybranými pojistnými událostmi způsobeno vloupáním nejvýše 6 událostí.

ad b)

$$\sum_{x=x_0}^n P_n(x) = \sum_{x=6}^{30} P_{30}(x) = 1 - \sum_{x=0}^5 P_{30}(x) = 1 - \sum_{x=0}^5 \binom{30}{x} 0,12^x 0,88^{30-x} = 1 - \text{IBinom}(5;0,12;30) = 0,1431$$

S pravděpodobností 14,31% bude mezi 30 náhodně vybranými pojistnými událostmi způsobeno vloupáním aspoň 6 událostí.

ad c)

$$P_n(x) = P_{30}(6) = \binom{30}{6} 0,12^6 0,88^{24} = \text{Binom}(6;0,12;30) = 0,0825$$

S pravděpodobností 8,25% bude mezi 30 náhodně vybranými pojistnými událostmi způsobeno vloupáním právě 6 událostí.

ad d)

$$\sum_{x=x_0}^{x_1} P_n(x) = \sum_{x=2}^5 P_{30}(x) = \sum_{x=0}^5 P_{30}(x) - \sum_{x=0}^1 P_{30}(x) = \sum_{x=0}^5 \binom{30}{x} 0,12^x 0,88^{30-x} - \sum_{x=0}^1 \binom{30}{x} 0,12^x 0,88^{30-x} =$$

$$= \text{IBinom}(5;0,12;30) - \text{IBinom}(1;0,12;30) = 0,7469$$

S pravděpodobností 74,69% bude mezi 30 náhodně vybranými pojistnými událostmi způsobeno vloupáním od 2 do 5 událostí.

Návod: Otevřeme nový datový soubor se čtyřmi proměnnými a o jednom případě.

Do Dlouhého jména 1. proměnné napíšeme =IBinom(6;0,12;30).

Do Dlouhého jména 2. proměnné napíšeme =1-IBinom(5;0,12;30).

Do Dlouhého jména 3. proměnné napíšeme =Binom(6;0,12;30).

Do Dlouhého jména 3. proměnné napíšeme =IBinom(5;0,12;30)-IBinom(1;0,12;30).

Upozornění: Podobným způsobem postupujeme při řešení dalších příkladů

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 $\text{Geom}(x; \vartheta)$

Pravděpodobnost, že prvnímu úspěchu bude předcházet nejvýše x_1 neúspěchů:

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

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

Pravděpodobnost, že prvnímu úspěchu bude předcházet aspoň x_0 neúspěchů:

$$1 - \sum_{x=0}^{x_0-1} P(x)$$

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

Příklad na geometrické rozložení pravděpodobností: Jaká je pravděpodobnost, že při hře „Člověče, nezlob se!“ nasadíme figurku nejpozději při třetím hoďu?

Řešení:

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

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

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

Příklad na geometrické rozložení pravděpodobností: 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?

Řešení:

Počet neúspěchů: $x = 3$, pravděpodobnost úspěchu: $\vartheta = 0,25$

$$P(3) = 0,75^3 \cdot 0,25 = \text{Geom}(3; 0,25) = 0,1055$$

Pravděpodobnost, že až čtvrtá zkoumaná octomilka má bílou barvu očí, je 10,55%.

c) Negativní binomické rozložení pravděpodobností

Pravděpodobnost, že k-tému úspěchu bude předcházet x neúspěchů:

$$P_k(x) = \binom{k+x-1}{x} (1-\vartheta)^x \vartheta^k.$$

Výpočet lze provést takto: $\text{Combin}(k+x-1;x) * (1-\vartheta)^x * \vartheta^k$

Pravděpodobnost, že k-tému úspěchu bude předcházet nejvýše x_1 neúspěchů:

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

Pravděpodobnost, že k-tému úspěchu bude předcházet aspoň x_0 neúspěchů:

$$1 - \sum_{x=0}^{x_0-1} P_k(x).$$

Příklad na negativní binomické rozložení pravděpodobností: Jaká je pravděpodobnost, že pro nalezení 3 dárců krevní skupiny A+ budeme muset vyšetřit

- právě 10 osob neznajících svou krevní skupinu,
- aspoň 10 osob neznajících svou krevní skupinu?

Řešení:

Počet úspěchů: $k = 3$, pravděpodobnost úspěchu: $\vartheta = \frac{1}{8}$ (protože předpokládáme, že máme 8

krevních skupin (A+, A-, B+, B-, AB+, AB-, 0+, 0-), které se vyskytují se stejnou pravděpodobností).

ad a)

Počet neúspěchů $x = 7$, protože 7 z 10 osob nebude mít krevní skupinu A+

$$P_3(7) = \binom{3+7-1}{7} \left(\frac{7}{8}\right)^7 \left(\frac{1}{8}\right)^3 = \binom{9}{7} \frac{7^7}{8^{10}} = \text{Combin}(9;7) * 7^7 / 8^{10} = 0,0276$$

S pravděpodobností 2,76% musíme vyšetřit právě 10 osob neznajících svou krevní skupinu, abychom našli 3 dárce krevní skupiny A+.

ad b)

Počet neúspěchů $x = 7, 8, 9, \dots$ Přejdeme k opačnému jevu – budeme vyšetřovat nejvýše 9 osob neznajících svou krevní skupinu. Pak počet neúspěchů $x = 0, 1, \dots, 6$.

$$1 - \sum_{x=0}^6 P_3(x) = 1 - \sum_{x=0}^6 \binom{3+x-1}{x} \left(\frac{7}{8}\right)^x \left(\frac{1}{8}\right)^3 = 0,9081$$

S pravděpodobností 90,81% musíme vyšetřit aspoň 10 osob neznajících svou krevní skupinu, abychom našli 3 dárce krevní skupiny A+.

2. Opakované závislé pokusy

Hypergeometrické rozložení pravděpodobností

Máme N objektů, mezi nimi je M objektů označeno $0 \leq M \leq N$. Náhodně bez vracení vybereme n objektů ($0 \leq n \leq N$).

Pravděpodobnost, že ve výběru je právě x označených objektů ($\max\{0, M - N + n\} \leq x \leq \min\{n, M\}$):

$$P_{N,M,n}(x) = \frac{\binom{M}{x} \binom{N-M}{n-x}}{\binom{N}{n}}.$$

Výpočet lze provést takto: $\text{Combin}(M;x) * \text{Combin}(N-M;n-x) / \text{Combin}(N;n)$

Pravděpodobnost, že ve výběru je nejvýše x_1 označených objektů:

$$\sum_{x=\max\{0, M-N+n\}}^{x_1} P_{N,M,n}(x).$$

Pravděpodobnost, že ve výběru je aspoň x_0 označených objektů:

$$\sum_{x=x_0}^{\min\{n, M\}} P_{N,M,n}(x).$$

Příklad na hypergeometrické rozložení pravděpodobností: Koupili jsme 10 cibulek červených tulipánů a 5 cibulek žlutých tulipánů. Zasadili jsme 8 náhodně vybraných cibulek.

- Jaká je pravděpodobnost, že žádná nebude cibulka žlutých tulipánů?
- Jaká je pravděpodobnost, že jsme zasadili všech 5 cibulek žlutých tulipánů?
- Jaká je pravděpodobnost, že aspoň dvě budou cibulky žlutých tulipánů?

Řešení:

Počet objektů: $N = 15$, počet označených objektů: $M = 5$, počet vybraných objektů: $n = 8$
ad a)

$$P_{15,5,8}(0) = \frac{\binom{5}{0} \binom{10}{8}}{\binom{15}{8}} = \frac{\binom{10}{8}}{\binom{15}{8}} = \text{Combin}(10;8) / \text{Combin}(15;8) = 0,007$$

Mezi 8 náhodně vybranými cibulkami se s pravděpodobností 0,7% nevyskytne žádná cibulka žlutých tulipánů.

ad b)

$$P_{15,5,8}(5) = \frac{\binom{5}{5} \binom{10}{3}}{\binom{15}{8}} = \frac{\binom{10}{3}}{\binom{15}{8}} = \text{Combin}(10;3) / \text{Combin}(15;8) = 0,0186$$

S pravděpodobností 1,86% bude mezi 8 náhodně vybranými cibulkami právě 5 cibulek žlutých tulipánů.

ad c)

$$1 - P_{15,5,8}(0) - P_{15,5,8}(1) = 1 - \frac{\binom{5}{0}\binom{10}{8}}{\binom{15}{8}} - \frac{\binom{5}{1}\binom{10}{7}}{\binom{15}{8}} = 1 - \frac{\binom{10}{8}}{\binom{15}{8}} - \frac{5 \cdot \binom{10}{7}}{\binom{15}{8}} =$$

$$= 1 - \text{Combin}(10;8) / \text{Combin}(15;8) - 5 * \text{Combin}(10;7) / \text{Combin}(15;8) = 0,8998$$

S pravděpodobností 89,98% budou mezi 8 náhodně vybranými cibulkami aspoň dvě cibulky žlutých tulipánů.