Téma 3: 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 p, $0 < p < 1$.

a) Binomické rozložení pravděpodobností

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

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

K výpočtu v systému STATISTICA slouží funkce Binom(x; p; 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; p; 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 - IBinom(x_0 - 1; p; 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; p; n) - IBinom(x_0 - 1; p; 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

a) nejvýše 6,
b) aspoň 6,
c) právě 6,
d) od dvou do pěti?

Řešení:

Počet pokusů: n = 30, pravděpodobnost úspěchu: $p = 0{,}12$

ad a)

$$\sum_{x=0}^{6} P_n(x) = \sum_{x=0}^{6} P_{30}(x) = \sum_{x=0}^{6} \binom{30}{x} 0{,}12^x 0{,}88^{30-x} = \mathrm{IBinom}(6;0{,}12;30) = 0{,}939$$

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

ad b)

$$\sum_{x=6}^{n} P_n(x) = \sum_{x=6}^{30} P_{30}(x) = 1 - \sum_{x=0}^{5} P_{30}(x) = 1 - \sum_{x=0}^{5} \binom{30}{x} 0{,}12^x 0{,}88^{30-x} = 1 - \mathrm{IBinom}(5;0{,}12;30) = 0{,}143$$

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

ad c)

$$P_n(x=6) = P_{30}(6) = \binom{30}{6} 0{,}12^6 0{,}88^{24} = \mathrm{Binom}(6;0{,}12;30) = 0{,}082$$

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

ad d)

$$\sum_{x=2}^{5} P_n(x) = \sum_{x=2}^{5} P_{30}(x) = \sum_{x=0}^{5} P_{30}(x) - \sum_{x=0}^{1} P_{30}(x) = \sum_{x=0}^{5} \binom{30}{x} 0{,}12^x 0{,}88^{30-x} - \sum_{x=0}^{1} \binom{30}{x} 0{,}12^x 0{,}88^{30-x} = \mathrm{IBinom}(5;0{,}12;30) - \mathrm{IBinom}(1;0{,}12;30) = 0{,}7469$$

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 případu.
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:

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

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

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; ϑ)

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; ϑ)

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

Řešení:

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

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

Pravděpodobnost, že figurku nasadíme nejpozději při třetím hodu, 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%.

2. Opakované závislé pokusy

Hypergeometrické rozložení pravděpodobností
Máme N objektů, mezi nimi je M objektů označeno $0 < M < N$. Náhodně bez vracení vybereme n objektů ($0 < n < N$).
Pravděpodobnost, že ve výběru je právě x označených objektů ($\max(0, M-N+n) \le x \le \min(n, M)$):

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

Výpočet lze provést takto: Combin(M;x)* Combin(N-M;n-x)/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.
a) Jaká je pravděpodobnost, že žádná nebude cibulka žlutých tulipánů?
b) Jaká je pravděpodobnost, že jsme zasadili všech 5 cibulek žlutých tulipánů?
c) 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)

$$P_{15,5,8}(0) = \frac{\binom{5}{0}\binom{10}{8}}{\binom{15}{8}} = \mathrm{Combin}(10;8)/\mathrm{Combin}(15;8) = 0{,}00\ldots$$

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

ad b)

$$P_{15,5,8}(5) = \frac{\binom{5}{5}\binom{10}{3}}{\binom{15}{8}} = \mathrm{Combin}(10;3)/\mathrm{Combin}(15;8) = 0{,}018\ldots$$

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

ad c)

$$1-P_{5,8}(0)-P_{5,8}(1)= \ldots$$

[illegible]

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