Diskrétní náhodné veličiny

Velmi často nám nejde ani tak o zjištění, zda daný jev nastal či nenastal, ale zajímá nás výsledek pokusu ve formě čísla. Budeme uvažovat celočíselné výsledky náhodného pokusu, např. počet líců v posloupnosti hodů mincí nebo počet šestek při hodu několika kostkami.

Definice: *Diskrétní (celočíselná) náhodná veličina* je funkce X definovaná na Ω s hodnotami v $\mathbb{Z}$, která splňuje $\{\omega \in \Omega : X(\omega) = k\} \in \mathcal{A}$ pro každé $k \in \mathbb{Z}$. Pravděpodobnosti $p_k = P(\{\omega \in \Omega : X(\omega) = k\}) = P(X = k)$ určují *rozdělení pravděpodobnosti* celočíselné náhodné veličiny X.

Poznámka: Pokud $X(\omega) \neq k$ pro všechna $\omega \in \Omega$, je $\{\omega \in \Omega : X(\omega) = k\} = \emptyset$, což je vždy prvek $\mathcal{A}$, v tomto případě je $p_k = 0$. Pravděpodobnosti $p_k = P(X = k)$ jsou nezáporné a $\sum_{k\in\mathbb{Z}} p_k = 1$.

Příklad: Uvažujme dva hody mincí. Potom $\Omega = \{LL, LR, RL, RR\}$ a náhodná veličina, která označuje počet líců, je dána jako $X(LL) = 2$, $X(LR) = 1$, $X(RL) = 1$, $X(RR) = 0$. Podle klasické pravděpodobnosti zřejmě $P(X = 2) = 1/4$, $P(X = 1) = 1/2$ a $P(X = 0) = 1/4$. Pravděpodobnost, že padne aspoň jeden líc, je $P(X \geq 1) = 3/4$.

Zadání: V případě, kdy dva hráči hrají ruskou ruletu s maximálním možným počtem 6 výstřelů, určete rozdělení počtu výstřelů prvního hráče.

Řešení: Z materiálu o nezávislosti víme, že $P(Z) = 1/6$, $P(NZ) = 5/6^2$, $P(NNZ) = 5^2/6^3$, $P(NNNZ) = 5^3/6^4$, $P(NNNNZ) = 5^4/6^5$, $P(NNNNNZ) = 5^5/6^6$. Proto $P(X = 1) = 1/6 + 5/6^2 = 11/36 \doteq 0{,}306$, $P(X = 2) = 5^2/6^3 + 5^3/6^4 = 275/6^4 \doteq 0{,}212$ a $P(X = 3) = 5^4/6^5 + 5^5/6^6 + 5^6/6^6 = 22\,500/6^6 \doteq 0{,}482$.

Definice: Nechť X je celočíselná náhodná veličina. Jestliže řada $\sum_{k=-\infty}^{\infty} kp_k$ konverguje absolutně, označíme její součet symbolem $\mathbb{E}X$ a nazveme jej *střední hodnotou* náhodné veličiny X. Pokud je X nezáporná (tj. $P(X \geq 0) = \sum_{k=0}^{\infty} p_k = 1$, neboli $p_k = 0$ pro záporná k), pak definujeme $\mathbb{E}X = \sum_{k=0}^{\infty} kp_k$. To v sobě zahrnuje případ, kdy řada diverguje (pak $\mathbb{E}X = \infty$).

Definice: Nechť X je celočíselná náhodná veličina a $g : \mathbb{Z} \to \mathbb{R}$ je reálná funkce celočíselné proměnné. Pokud řada $\sum_{k\in\mathbb{Z}} g(k)p_k$ konverguje absolutně, označíme její součet $\mathbb{E}g(X)$. Je-li g nezáporná funkce, pak můžeme položit $\mathbb{E}g(X) = \sum_{k=1}^{\infty} g(-k)p_{-k} + \sum_{k=0}^{\infty} g(k)p_k$, výsledkem může být nezáporné reálné číslo nebo nekonečno.

Poznámka: Nabývá-li g jen celočíselných hodnot, můžeme se na $\mathbb{E}g(X)$ dívat jako na střední hodnotu náhodné veličiny $Y = g(X)$.

Zadání: Uvažujme sázky na součet ok při hodu dvěma kostkami. Je možné sázet na tři možnosti: součet bude pod 7, právě 7, nebo přes 7. Při správném odhadu prvního a posledního případu se vyplácí dvojnásobek vsazené částky, v prostředním případě pak pětinásobek. Jaký je střední zisk při této hře, když vsadíme 1 korunu na „pod 7“ a 1 korunu na „právě 7“?

Řešení: Označme X zisk z této hry při uvedených sázkách. Protože součet pod 7 nebo přes 7 padne s pravděpodobností 15/36, zatímco součet 7 s pravděpodobností 6/36, je $P(X = 0) = 15/36$, $P(X = 3) = 6/36$ a $P(X = -2) = 15/36$. Střední zisk činí

$$\mathbb{E}X = 0 \cdot \frac{15}{36} + 3 \cdot \frac{6}{36} - 2 \cdot \frac{15}{36} = -\frac{12}{36} = -\frac{1}{3}.$$

Lehce se přesvědčíme, že jakékoli rozdělení vsazených 2 korun mezi 3 možné sázky vede ve střední hodnotě ke ztrátě $-1/3$.

Zadání: (Petrohradský paradox) Kasino nabízí hru spojenou se sérií hodů mincí. Hráči je vyplacena částka 2^n, pokud rub padne poprvé v n-tém hodu. Můžeme si představit, že na začátku je v banku částka 1 a po každém hodu se zdvojnásobí, hráč získává částku z banku po prvním rubu. Jak velký vstupní poplatek do hry má kasino žádat, aby na hře neprodělávalo?

Řešení: Střední hodnota výplaty je $2 \cdot \frac{1}{2} + 2^2 \cdot \frac{1}{2^2} + \cdots = 1 + 1 + \cdots = \infty$. Proto za žádnou cenu není pro kasino výhodné takovou hru provozovat. Naopak hráč by měl využít příležitost a hrát za každou cenu. Ve skutečnosti se najde jen velmi málo lidí, kteří by byli ochotni vložit do této hry víc než tisíc korun. Paradox spočívá v jasném rozporu mezi očekávanou výhrou a částkou, kterou by lidi byli ochotni zaplatit za vstup do hry. Pochází od Nicolause Bernoulliho (1687–1759), který ho formuloval v roce 1713. Pojmenování ale vychází z přednášky jeho bratrance Daniela Bernoulliho (1700–1782) před Petrohradskou akademií věd v roce 1738. Paradox poukazuje na nedostatečnost střední hodnoty při rozhodování. Problém je, že kasino nemá neomezené prostředky k výplatě, model tak neodráží realitu, a proto výsledky z něho odvozené nejsou relevantní pro praxi.

Předpokládejme, že maximální možná výplata je 2^M (pokud v prvních M hodech nepadne rub, hráč dostává maximální výhru), pak střední hodnota vyplacené částky bude $2 \cdot \frac{1}{2} + \cdots + 2^M \cdot \frac{1}{2^M} + 2^M \cdot \frac{1}{2^M} = M + 1$. Hra se kasinu vyplatí, když poplatek bude větší než $M + 1$.

Zadání: (ruleta) V ruletě se dá sázet na jedno číslo nebo na různé kombinace čísel. Při správném tipu kombinace k čísel vyhrává hráč $(36/k)$-násobek vkladu. Znamená to, že každá vsazená koruna vede ve střední hodnotě k $1 - \frac{36}{k} \cdot \frac{k}{37} = 1/37 \doteq 0{,}027$ korunám pro kasino. Existují různé systémy zaručených výher v ruletě, ale žádný z nich nefunguje. Nejznámější je martingalová strategie, která je založena na sázce na barvu a zdvojnásobení vkladu při prohře. Dejme tomu, že poprvé hráč trefí barvu v n-tém kole, pak vsadil postupně $1, 2, \ldots, 2^{n-1}$ jednotek, což dává dohromady $2^n - 1$ jednotek, ale vyhrál 2^n jednotek. Systém by fungoval, pokud by hráč měl neomezený kapitál a bylo by možno vsadit libovolnou sumu. V praxi ale ani jedna z těchto podmínek neplatí. Nikdo nedisponuje neomezeným kapitálem a kasina mají vždy nastaven určitý limit na maximální povolenou sázku. Předpokládejme, že limit na sázku je 1 000 jednotek a řídíme se martingalovým systémem s počáteční sázkou 1 jednotka. Jaká je střední hodnota výhry?

Řešení: Limitu dosáhneme po 10 prohrách v řadě, v 11. sázce tak použijeme horní limit 1 000. Nechť X značí počet uskutečněných sázek. Potom $P(X = k) = \left(\frac{19}{37}\right)^{k-1} \cdot \frac{18}{37}$ pro $k = 1, \ldots, 10$ a $P(X = 11) = (19/37)^{10}$. K více než 11 sázkám nedojde (buď jsme do té doby vyhráli, nebo už nelze zdvojnásobit vklad). Definujeme-li $g(k) = 2^k - 1$ pro $k = 1, \ldots, 10$ a $g(11) = 2^{10} - 1 + 1\,000 = 2\,023$, pak $g(X)$ značí celkovou vsazenou částku. Střední hodnota je

$$\mathbb{E}g(X) = \sum_{k=1}^{10} (2^k - 1) P(X = k) + 2\,023 P(X = 11) \doteq 12{,}583.$$

Hráč získá 1 jednotku, pokud nedošlo k 11. sázce; když k ní došlo a je vítězná, znamená to ztrátu 23 jednotek; pokud ale není vítězná, vede to ke ztrátě 2 023 jednotek. Proto střední hodnota výhry je

$$1 \cdot \left(1 - \left(\frac{19}{37}\right)^{10}\right) - 23 \cdot \left(\frac{19}{37}\right)^{10} \cdot \frac{18}{37} - 2\,023 \cdot \left(\frac{19}{37}\right)^{10} \cdot \frac{19}{37} \doteq -0{,}340.$$

V dlouhodobém průměru bude podíl prohrané částky ku vsazené $0{,}340/12{,}583 \doteq 0{,}027$, což odpovídá očekávanému zisku kasina z jedné vsazené jednotky.