6. cvičení

Náhodné veličiny

Úloha 1 (Narozeninová party)
Jste ve skupině s dalšími 500 lidmi. Nechť X je počet lidí, kteří mají stejné narozeniny jako vy. (Narozeniny jednotlivých lidí jsou nezávislé rovnoměrné náhodné. Lidé narození 29. února mají vstup zakázán :))

a) Spočítejte $P[X = 1]$, tedy $p_X(1)$.

b) Jaké je rozdělení veličiny X?

c) Kolik je $\mathbb{E}(X)$?

d) Aproximujte $p_X(1)$ pomocí Poissonova rozdělení.

Řešení
Rozdělení je $X \sim \mathrm{Bin}(500, 1/365)$

a) $P[X = 1] = \binom{500}{1} \cdot \frac{1}{365} \cdot \left(\frac{364}{365}\right)^{499} \doteq 0.34844$.

b) Viz výše

c) $\mathbb{E}[X] = np = \frac{500}{365}$

d) Dobrá aproximace je $\lambda = np = \frac{500}{365}$, tedy $P[X = 1] = \frac{\lambda}{1} e^{-\lambda} = \lambda \cdot e^{-\lambda} = \frac{500}{365} \cdot e^{-\frac{500}{365}} \doteq 0.34814$

Úloha 2 (Pět klíčů)
Na kroužku máme pět klíčů, jeden z nich je správný, ale my nevíme jaký. Zkoušíme otevřít dveře, X říká, kolikátým pokusem se nám to povede. Jaké rozdělení má X a kolik je $\mathbb{E}(X)$?

a) Po každém pokusu se nám kroužek vysmekne, a vybíráme vždy znovu náhodně.

b) Vybíráme v náhodném pořadí, ale každý klíč jenom jednou (můžeme si je poznačit).

c) Jak se výsledek změní, pokud jsou správně 2 klíče z 10?

Řešení a) $X \sim \mathrm{Geom}(1/5), \mathbb{E}[X] = 5$

b) Y má rovnoměrné rozdělení, $\mathbb{E}[Y] = 3$

c) Pro X nic, pro Y: $p_Y(1) = 2/10, p_Y(2) = \frac{8}{10} \cdot \frac{2}{9}, p_Y(k) = \frac{2}{10-k} \cdot \prod_{j=1}^{k-1} \frac{9-j}{11-j}$, střední hodnota je $\frac{11}{3}$.

Úloha 3 (Buďme uniformní)
Nechť X je uniformně náhodná mocnina dvojky z $\{2^a, 2^{a+1}, \dots, 2^b\}$. Vyjádřete X pomocí uniformního rozdělení na množině $\{a, a+1, \dots, b\}$. Určete $\mathbb{E}(X)$ a $\mathrm{var}(X)$.

Řešení
Buď $Y \sim \mathrm{Unif}(a, b)$. Pak $X = 2^Y$. $\mathbb{E}X = \sum_{i=a}^{b} \frac{2^i}{b-a+1} = \frac{2^{b+1}-2^a}{b-a+1}$.
$\mathbb{E}[X^2] = \sum_{i=a}^{b} \frac{2^{2i}}{b-a+1} = \frac{4^{b+1}-4^a}{3(b-a+1)}$ $\mathrm{var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \frac{4^{b+1}-4^a}{3(b-a+1)} - \left(\frac{2^{b+1}-2^a}{b-a+1}\right)^2 = \frac{4^a(a-b-4)+3\cdot 2^{a+b+2}+4^{b+1}(-a+b-2)}{3(a-b-1)^2}$

Náhodné vektory

Úloha 4 (Měl s sebou jen balíček karet)
Ze standardního balíčku s 52 kartami vytáhneme dvě karty. Označíme X počet vytažených es, Y počet králů (v balíčku jsou čtyři esa a čtyři králové). Určete sdruženou pravděpodobnostní funkci $p_{X,Y}$ a také marginální pravděpodobnostní funkce p_X, p_Y.

Řešení

Y\X	0	1	2	$p_Y(i)$
0	$\frac{44\cdot 43}{52\cdot 51} = \frac{473}{663}$	$\frac{4\cdot 44+44\cdot 4}{52\cdot 51} = \frac{88}{663}$	$\frac{4\cdot 3}{52\cdot 51} = \frac{1}{221}$	$\frac{48\cdot 47}{52\cdot 51} = \frac{188}{221}$
1	$\frac{4\cdot 44+44\cdot 4}{52\cdot 51} = \frac{88}{663}$	$\frac{\{}{4} \cdot 4 + 4 \cdot 4\}\{52 \cdot 51\} = \frac{8}{663}$	0	$\frac{4\cdot 48+48\cdot 4}{52\cdot 51} = \frac{32}{221}$
2	$\frac{4\cdot 3}{52\cdot 51} = \frac{1}{221}$	0	0	$\frac{4\cdot 3}{52\cdot 51} = \frac{1}{221}$
$p_X(i)$	$\frac{48\cdot 47}{52\cdot 51} = \frac{188}{221}$	$\frac{4\cdot 48+48\cdot 4}{52\cdot 51} = \frac{32}{221}$	$\frac{4\cdot 3}{52\cdot 51} = \frac{1}{221}$	

Úloha 5 (max(3d4))
Označme X_1, X_2, X_3 výsledky tří nezávislých hodů čtyřstěnnou kostkou (s čísly 1, …, 4).

a) Jaká je pravděpodobnostní funkce $X = X_1$?

b) Jaká je pravděpodobnostní funkce $Y = \max(X_1, X_2)$?

c) Jaká je pravděpodobnostní funkce $Z = \max(X_1, X_2, X_3)$?

d) O kolik se zvýší střední hodnota tím, že můžeme házet třikrát? Neboli o kolik je vyšší $\mathbb{E}(Z)$ než $\mathbb{E}(X)$?

Nápověda: Určete napřed $P(Y \le k)$, $P(Z \le k)$.

Řešení
$p_X(1) = p_X(2) = p_X(3) = p_X(4) = 1/4$
$p_Y(1) = 1/16, p_Y(2) = 3/16, p_Y(3) = 5/16, p_Y(4) = 7/16$
$p_Z(1) = 1/64, p_Y(2) = 7/16, p_Y(3) = 19/64, p_Y(4) = 37/64.$
$\mathbb{E}[X] = 2.5, \mathbb{E}[Z] = \frac{1+14+57+148}{64} = \frac{220}{64} = 3.4375$

Úloha 6 (Počítáme hody kostkou)
Na kostce padne číslo i s pravděpodobností p_i pro $i = 1, \ldots, 6$. Hodíme n-krát a označíme X_i počet hodů, kdy padlo i.

a) Najděte sdruženou pravděpodobnostní funkci pro n.v. X_1, …, X_n.

b) Jaké je marginální rozdělení, tj. rozdělení jednotlivých n.v. X_i?

Řešení

a) Pokud součet X_i není dvacet, je to 0. Pokud součet je 20, pak buďte x_i hodnoty pro X_i.

$$p_{X_1,X_2,X_3,X_4,X_5,X_6}(x_1, x_2, x_3, x_4, x_5, x_6) = \binom{20}{x_1, x_2, x_3, x_4, x_5, x_6} \frac{1}{6^{20}}.$$

b) Binom$(n, 1/6)$

Úloha 7 (Součet Poissonů)
Nechť $X \sim \text{Pois}(\lambda)$ a $Y \sim \text{Pois}(\mu)$ jsou nezávislé náhodné veličiny. Dokažte, že $X + Y \sim \text{Pois}(\lambda + \mu)$.

Řešení
$p_{X+Y}(k) = \sum_{i=0}^{k} \frac{\lambda^i}{i!} e^{-\lambda} \cdot \frac{\mu^{i-k}}{(i-k)!} e^{-\mu} = \sum_{i=0}^{k} \frac{\lambda^i}{i!} e^{-\lambda} \cdot \frac{\mu^{i-k}}{(i-k)!} e^{-\mu} = \frac{e^{-\mu-\lambda}}{k!} \sum_{i=0}^{k} \binom{k}{i} \lambda^i \mu^{i-k} = \frac{(\lambda+\mu)^k}{k!} \cdot e^{-\mu-\lambda}$

Bonusové úlohy

K procvičení

Úloha 8 (Pohled do budoucnosti)
Uvažme skupinu m manželských párů (tj. celkem $2m$ osob). Předpokládejme, že po deseti letech bude každý z těch $2m$ lidí stále naživu s pravděpodobností p, nezávisle na ostatních. Možnosti rozvodů apod. neuvažujeme, tj. páry jsou neměnné.
Označme L množinu lidí, kteří budou po deseti letech naživu a A jejich počet (tj. $A = |L|$). Dále buď B počet párů, kde budou naživu oba; tj. A, B jsou náhodné veličiny splňující $0 \le A \le 2m$ a $0 \le B \le m$. Pro každé $a = 0, \ldots, 2m$ chceme spočítat $\mathbb{E}(B|A = a)$.

a) Uvážíme jednoho konkrétního člověka. Jaká je pravděpodobnost, že bude po deseti letech naživu, pokud víme, že $A = a$? Jinými slovy, pokud ten člověk je x, jaká je $P(x \in L|A = a)$?

b) Uvážíme jeden konkrétní manželský pár. Jaká je pravděpodobnost, že budou oba naživu, pokud víme, že $A = a$?

c) Vyjádřete B jako součet m vhodných indikátorových n.v.

d) Spočítejte $\mathbb{E}(B|A = a)$. (Hint: Linearita střední hodnoty platí i pro podmíněnou střední hodnotu.)

e) Jaké je rozdělení n.v. A? (Buď ho pojmenujte, nebo napište pravděpodobnostní funkci, tj. určete $P(A = a)$.)

f) Pro zvolenou a-prvkovou množinu lidí M: jaká je pravděpodobnost, že je to přesně množina přeživších? Neboli kolik je $P(L = M)$? A kolik $P(L = M|A = a)$?

g) Pro $m = 10$ a $a = 4$ ověřte výsledek samplováním v libovolném programovacím jazyce.

Řešení a) $P(x \in L|A = a) = \frac{\binom{19}{a-1}p^{a-1}\cdot(1-p)^{19-(a-1)}\cdot p}{\binom{20}{a}p^a(1-p)^{20-a}} = \frac{a}{20}$ přes Bayese

b) $P(x, y \in L|A = a) = \frac{\binom{18}{a-2}p^{a-2}\cdot(1-p)^{19-(a-2)}\cdot p^2}{\binom{20}{a}p^a(1-p)^{20-a}} = \frac{a(a-1)}{20\cdot 19}$ přes Bayese

c) $B = \sum_{i=1}^{10} I_{P_i}$

d) $\mathbb{E}(B|A = a) = \frac{a(a-1)}{38}$

e) $A \sim \mathrm{Bin}(20, p)$

f) $P[L = M] = p^{|M|}(1-p)^{20-|M|}$, $P[L = M|A = a] = 0$ pokud $a \neq |M|$ a jinak $P[L = M|A = a] = \frac{1}{\binom{20}{a}}$.

g) Nakóděte si dle libosti

Tahák

- *Sdružená pravděpodobnostní* funkce náhodného vektoru (X, Y) je definována vztahem $p_{X,Y}(x, y) = P(X = x \wedge Y = y)$.
- Připomeňte si, jak z ní zjistit „jednorozměrné funkce“ p_X, p_Y neboli *marginální rozdělení* – představujte si je na *okrajích* tabulky sdruženého rozdělení.
- *Poissonovo rozdělení:* $X \sim \mathrm{Pois}(\lambda)$: $p_X(k) = e^{-\lambda} \cdot \lambda^k/k!$, $\mathbb{E}(X) = \lambda$.
- Konvoluční vzorec pro X, Y nezávislé: $P[X + Y = z] = \sum_{x \in \mathrm{Supp}(X)} P[X = x] \cdot \Pr[Y = z - x]$.