5. cvičení z PSt — 2. a 5.11.2020

Spojité náhodné veličiny

- $Exp(\lambda)$ má hustotu $\lambda e^{-\lambda x}$, střední hodnotu $1/\lambda$ a rozptyl $1/\lambda^2$.
- $N(0,1)$ má hustotu $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$, střední hodnotu 0 a rozptyl 1.

x	-4	-3	-2	-1	0	1	2	3	4
$\Phi(x)$	0.00003	0.00135	0.02275	0.15866	0.500000	0.84135	0.97725	0.99865	0.99997

Další hodnoty viz `https://en.wikipedia.org/wiki/Standard_normal_table` – sekce Cumulative.

1. Pan Chen Cheng navštívil Prahu a v uniformně náhodný čas se objeví na Staroměstském náměsstí. Každou celou hodinu od 9:00 do 23:00 se na orloji objevuje 12 figur apoštolů.

(a) Jaká je pravděpodobnost, že pan Cheng uvidí apoštoly, aniž by čekal déle než 15 minut.

(b) Co když pan Cheng přijde na Staroměstské náměstí v uniformně náhodném čase po poledni, tj. 12:00–24:00?

2. Metrový klacek rozlomíme na dva kusy – lomem v uniformně náhodném bodě. Buď X délka delší části.

(a) Jaké je rozdělení X?

(b) Určete $\mathbb{E}(X)$, $var(X)$.

3. Předpokládejme, že u poštovní přepážky trvá vyřízení jednoho zákazníka čas, který má exponenciální rozdělení a střední hodnotu 4 minuty.

(a) Jaký je parametr λ, jaká je distribuční funkce?

(b) Jaká je pravděpodobnost, že budeme čekat více než 4 minuty?

(c) Jaká je pravděpodobnost, že budeme čekat něco mezi 3 a 5 minutami?

4. Říkáme, že náhodná veličina X (resp. její rozdělení) *nemá paměť*, pokud

$$P(X > s + t \mid X \geq s) = P(X > t)$$

pro $s, t \geq 0$. Jinými slovy, doba, kterou jsme již čekali, nemá vliv na dobu, kterou budeme ještě čekat. Na druhém cvičení jsme viděli, že geometrické rozdělení nemá paměť. Ukažte, že ani exponenciální rozdělení nemá paměť. Platí dokonce, že je to jediné spojité rozdělení na kladných čísel bez paměti (a geometrické je jediné diskrétní bez paměti), ale to dokazovat nemusíte.

5. Najděte analogii „pravidla 3σ“, neboli spočtěte $P(|X - \mathbb{E}(X)| < c \cdot \sigma_X)$ ($c = 1, 2, 3$), pokud

(a) X má uniformní rozdělení,

(b) $X \sim Exp(1)$,

(c) $X \sim Exp(2)$.

6. Pro jistý problém máme k dispozici dva algoritmy A a B. Algoritmus C spočívá v tom, že si náhodně vybereme, který z algoritmů A, B spustíme – A bude mít pravděpodobnost p, B pravděpodobnost $1 - p$. Dobu běhu A, B, C chápeme jako náhodné veličiny, označíme je X, Y, Z.

(a) Určete F_Z pomocí F_X, F_Y.

(b) Pokud jsou X, Y spojité, určete f_Z pomocí f_X, f_Y.

7. Nechť $X_i \sim Exp(\lambda_i)$ pro $i = 1, \dots, n$ jsou nezávislé náhodné veličiny. Označme $M = \min(X_1, \dots, X_n)$. Ukažte, že $M \sim Exp(\lambda_1 + \dots + \lambda_n)$.

8. (a) Nechť X je diskrétní nebo spojitá náhodná veličina a $X \geq 0$ s.j. Pokud $\mathbb{E}(X)$ existuje, tak $\mathbb{E}(X) \geq 0$. Dokažte.

(b) Nechť Y, Z jsou diskrétní nebo spojité náhodné veličiny a $Y \leq Z$ s.j. Pokud $\mathbb{E}(Y)$, $\mathbb{E}(Z)$ existují, tak $\mathbb{E}(Y) \leq \mathbb{E}(Z)$. Dokažte.

9. Nechť $Z \sim N(0,1)$. Pomocí tabulky funkce Φ ověřte pravidlo 3σ, neboli spočtěte
(a) $P(|Z| \leq 1)$
(b) $P(|Z| \leq 2)$
(c) $P(|Z| \leq 3)$
(d) Přepište, co to znamená pro n.v. $X \sim N(\mu, \sigma^2)$.

10. Budeme modelovat množství sněhu, který bude na Silvestra v lyžarském areálu Ještěd, pomocí normálního rozdělení se střední hodnotou 40 (centrimetrů) a směrodatnou odchylkou 10.
(a) Jaká je pravděpodobnost, že nám model určí zápornou hodnotu sněhové pokrývky?
(b) Jaká je pravděpodobnost, že sněhu napadne 50–70 cm?

11. Střední hodnota diskrétní i spojité náhodné veličiny splňuje

$$\boxed{\mathbb{E}(X) = \int_0^\infty P(X > t)dt - \int_{-\infty}^0 P(X < t)dt.}$$

Generování náhodné veličiny, samplování

12. Vzpomeňte si na větu z přednášky. Nechť $U \sim U(0,1)$. Jak vyrobíte náhodnou veličinu
(a) s rozdělením $N(0,1)$?
(b) s Cauchyho rozdělením?
(c) s exponenciálním rozdělením?

13. Nechť $X_1, \ldots, X_n$ mají všechny stejné rozdělení se střední hodnotou μ a rozptylem σ^2. Označme $S_n = (X_1 + \cdots + X_n)/n$.
(a) Určete $\mathbb{E}(S_n)$ a $var(S_n)$.
(b) Ukažte, jak lze počítat S_n z S_{n-1}, X_n a n.
(c) (Bonus) Sestavte program v libovolném jazyce a ověřte pomocí něj hodnotu μ některého rozdělení, o kterém jsme si říkali.

Bonusy

14. Návod, jak ověřit, že hustota normálního rozdělení se opravdu zintegruje na 1. Chceme vypočítat $I = \int_{-\infty}^{\infty} e^{-x^2/2}$, resp. ukázat, že $I = \sqrt{2\pi}$. Ukážeme místo toho, že $I^2 = 2\pi$. Platí totiž

$$I^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-x^2/2} e^{-y^2/2} dxdy$$

a tento dvojný integrál se velmi zjednoduší po převodu do polárních souřadnic.

15. (a) Uvažme seznam států světa a jejich aktuálního počtu obyvatel. Odhadněte, kolik z těchto počtů začíná jedničkou! (Lhostejno, na které pozici na první jednička je.)
(b) Srovnejte s nějakou skutečnou tabulkou, např. na `https://www.worldometers.info/world-population/population-by-country/`.
(c) Promyslete, proč by to tak mohlo být. Případně si přečtěte o *Benfordově zákonu.*

K procvičení

16. Střední doba života harddisku je 4 roky. Přepokládejme, že tato doba je popsána náhodnou veličinou s exponenciálním rozdělením. (To není realistický předpoklad, viz např. `https://www.backblaze.com/blog/how-long-do-disk-drives-last/`.)

(a) Jaká je pravděpodobnost, že disk selže během prvních tří let?
(b) Jaká je pravděpodobnost, že vydrží alespoň 10 let?
(c) Po jaké době se rozbije 10% disků?

17. Plutonium-238 má poločas rozpadu 87.7 let. Jeho rozpad budeme modelovat pomocí exponenciálního rozdělení: pro každý atom budeme čas, za který se rozpadne, považovat za nezávislou náhodnou veličinu s rozdělením $Exp(\lambda)$.
(a) Jaké je λ?
(b) Jaká je střední doba života atomu plutonia-238?
(c) Po jaké době se rozpadne 90% atomů?
(d) Kolik procent atomů se rozpadne po 50 letech? (Některé kardiostimulátory používají plutonium-238 jako zdroj enegerie. `https://en.wikipedia.org/wiki/Plutonium-238#Nuclear_powered_pacemakers`)

18. Doba, za kterou uvidíme meteorit, je exponenciálně rozdělená se střední hodnotou 1 (minuta).
(a) Jaká je pravděpodobnost, že budeme muset čekat více než 5 minut?
(b) Jaká je pravděpodobnost, že se dočkáme za nejvýše jednu minutu?
(c) * Jaké je rozdělení času, kdy uvidíme druhý meteorit? Třetí, ... (Předpokládáme, že jednotlivé meteority jsou navzájem nezávislé.)

19. $X \sim N(0, 1)$, $Y \sim N(1, 4)$.
(a) Najděte lineární funkci $f(t) = a \cdot t + b$, aby $f(Y)$ měla stejnou distribuci jako X.
(b) Spočtěte $P(X \leq 1)$, $P(X > 2)$.
(c) Spočtěte $P(Y < 0)$, $P(Y > 2)$.

20. Buď Y minimum z k uniformně náhodných čísel z intervalu $[0, 1]$. Spočtete $\mathbb{E}(Y)$.

21. Frantovi jsme ve skoku do dálky naměřili 9 metrů, což překonává světový rekord o 5 cm. Při měření jsme se ovšem dopustili chyby s rozdělením $N(0, 0.01)$. Jaká je pravděpodobnost, že byl rekord skutečně překonán?

22. Nechť $X \sim N(0, 1)$ a $Y = |X|$. Určete $\mathbb{E}(Y)$ a $var(Y)$.