Ukázková řešení domácích cvičení úterní skupinka

7. 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 nejvýše 50 letech? (Některé kardiostimulátory používají plutonium-238 jako zdroj energie. `https://en.wikipedia.org/wiki/Plutonium-238#Nuclear_powered_pacemakers`)

Řešení: Využijeme toho, že n.v. $X \sim Exp(\lambda)$ má distribuční funkci $F_X(x) = 1 - e^{-\lambda x}$ a střední hodnotu $1/\lambda$. Máme zadáno, že $F_X(87.7) = 0.5$. Platí tedy $e^{-\lambda \cdot 87.7} = 0.5$ a $\lambda = (\log 2)/87.7 \doteq 0.0079$. (Do dalších rovnic budeme ale dosazovat přesnou hodnotu, abychom nekumulovali zaokrouhlovací chybu.)

Odsud tedy dostáváme střední hodnotu $\mathbb{E}(X) = 1/\lambda = 87.7/\log 2 \doteq 127$ let.

Pro určení devadesátého percentilu musíme vyřešit rovnici

$$0.9 = F_X(x) = 1 - e^{-\lambda x}.$$

Odsud dostáváme $x = -(\log 0.1)/\lambda = 87.7\frac{\log 10}{\log 2} \doteq 291$ let.

Konečně, podíl atomů, které se rozpadnou po nejvýše 50 letech je roven $P(X \leq 50)$

$$P(X \leq 50) = F_X(50) = 1 - e^{-50 \cdot 87.7/\log 2} \doteq 0.326,$$

čili cca třetina.

---

8. 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?

Řešení: Označme skutečnou délku skoku S a chybu měření C. Víme tedy, že $C \sim N(0, 0.01)$ a že $S + C = 9$. Zajímá nás $P(S > 8.95)$, což je totéž jako $P(C < 0.05) = P(C < \sigma_C/2) = \Phi(0.5) \doteq 0.69$.

Záludná chyba vznikne převodem na centimetry. Pak víme, že $S' + C' = 900$ a zajímá nás $P(S' > 895)$. Je ale třeba správně přepočíst, co víme o C: platí $\sigma_{C'} = 100\sigma_C$, ale $\mathrm{var}(C') = 100^2 \,\mathrm{var}(C)$.

Několik lidí řešilo úlohu pomocí integrálu z hustoty, např. $\int_{-\infty}^{0.5} \varphi(t)dt$ (nebo stejný integrál jinak přeškálovaný, tj. t v metrech nebo centimetrech). To je v zásadě správně – ale pak je třeba říct, jak zjistíme hodnotu toho integrálu. Pokud numericky na počítači, tak jsme rovnou mohli numericky zjistit hodnotu distribuční funkce a snížit počet míst, kde lze udělat chybu.

---

9. Buďte $X, Y, Z \sim Exp(\lambda)$ nezávislé náhodně veličiny.

(a) Jaké je rozdělení $X + Y$?

(b) Jaké je rozdělení $X + Y + Z$?

Řešení:Z přednášky víme, že hustota všech našich veličin je dána vzorcem $f(x) = \lambda e^{-\lambda x}$ pro $x \geq 0$ (a 0 jinak). Podle konvolučního vzorce je hustota $X+Y$ pro $t \geq 0$ rovna

$$\begin{aligned} g(t) &= \int_{-\infty}^{\infty} f(x) f(t-x) dx \\ &= \int_0^t f(x) f(t-x) dx \\ &= \int_0^t \lambda e^{-\lambda x} \lambda e^{-\lambda(t-x)} dx \\ &= \int_0^t \lambda^2 e^{-\lambda t} dx \\ &= t \lambda^2 e^{-\lambda t} \end{aligned}$$

Pro $t < 0$ je zjevně $g(t) = 0$. To je vzorec, který definuje rozdělení $Gamma(2, \lambda)$.

Součet v druhé části si napíšeme jako $(X+Y)+Z$. Použijeme tedy konvoluční vzorec pro $g = f_{X+Y}$ a $f = f_Z$. Pro $t \geq 0$ obdobným výpočtem jako výše získáváme

$$\begin{aligned} h(t) &= \int_{-\infty}^{\infty} g(x) f(t-x) dx \\ &= \int_0^t g(x) f(t-x) dx \\ &= \int_0^t x \lambda^2 e^{-\lambda x} \lambda e^{-\lambda(t-x)} dx \\ &= \int_0^t x \lambda^3 e^{-\lambda t} dx \\ &= \frac{t^2}{2} \lambda^3 e^{-\lambda t}, \end{aligned}$$

pro $t < 0$ je zase $h(t) = 0$. (Proč?) Všimněme si, že jsme dostali vzorec rozdělení $Gamma(3, \lambda)$. Není těžké ověřit, že podobným postupem dostaneme pro součet n exponenciálně rozdělených veličin hustotu $\frac{t^{n-1}}{(n-1)!} \lambda^n e^{-\lambda t}$, což je rozdělení $Gamma(n, \lambda)$. (Také se používá název Erlangeovo rozdělení, pro Gamma rozdělení, jehož první parametr je přirozené číslo.)

Ze vzorců pro hustotu by šlo v obou částech přímočaře odvodit i distribuční funkci, to ale nebylo třeba. Šlo by i naopak odvodit jinými postupy napřed distribuční funkci (a z ní derivací případně hustotu), ale vede to ke kompolikovanějším výpočtům. Někteří řešili obecnější úlohu: $X \sim Exp(\lambda_1, Y \sim Exp(\lambda_2)$. I pro takové zadání lze najít rozdělení $X+Y$, jen je výpočet komplikovanější (a navíc vzorec vyjde jiný pro $\lambda_1 \neq \lambda_2$ a jiný pro $\lambda_1 = \lambda_2$).

Několik lidí úlohu vyřešilo tím, že jsme si na přednášce říkali, že součet Exp rozdělení je Gamma. To jsme si sice říkali (a je dobře, že jste to v poznámkách našli), ale cílem tohoto cvičení (jakož celého oddílu v zadání cvičení) bylo, vyzkoušet si použití konvolučního vzorce.

---

10. Máme náhodný výběr $X_1, \ldots, X_n \sim Pois(\lambda)$.

(a) Navrhněte bodový odhad λ momentovou metodou.

(b) Navrhněte bodový odhad λ metodou maximální věrohodnosti.

(c) Spočtěte střední kvadratickou odchylku (MSE).

Řešení:

(a) Z přednášky víme, že $m_1(\lambda) = \mathbb{E}(X_i) = \lambda$. Také víme, že nestranný odhad prvního momentu (střední hodnoty) je výběrový průměr, neboli $\widehat{m_1(\lambda)} = \bar{X}_n$. Odsud tedy náš odhad $\hat{\lambda} = \bar{X}_n$.

(b) Ze vzorce pro Poissonovo rozdělení je $L(x_i; \lambda) = P(X_i = x_i) = e^{-\lambda} \frac{\lambda^{x_i}}{x_i!}$. Bude se nám hodit pracovat s logaritmem tohoto výrazu, tj. s $\ell(x_i; \lambda) = -\lambda + x_i \log(\lambda) - \log(x_i!)$. Derivace podle proměnné λ (x_i je konstanta): $\ell'(x_i; \lambda) = -1 + x_i/\lambda - 0$. Metoda max. věrohodnosti po nás chce, abychom nalezli maximum součtu $\ell(x; \lambda) = \sum_i \ell(x_i; \lambda)$. Nalezneme napřed λ, kde je derivace nulová:

$$\ell'(x; \lambda) = \sum_{i=1}^{n} (-1 + x_i/\lambda) = -n + \frac{1}{\lambda} \sum_{i=1}^{n} x_i$$

Položíme-li poslední vzorec roven nule, dostáváme opět odhad $\hat{\lambda} = \sum_{i=1}^{n} x_i/n$. Pak si ještě můžeme všimnout, že v tomto bodě se opravdu nabývá maxima (a ne minima či inflexního bodu): buď pomocí derivace (kladná nalevo od toho bodu, záporná napravo), nebo pomocí limity pro $\lambda \to 0_-$ a pro $\lambda \to \infty$.

(c) Použijeme vzorec $MSE = bias^2 + \mathrm{var}(\bar{X}_n)$. Z principu momentové metody je nalezený odhad nestranný, tj. $bias = 0$. Potřebujeme tedy spočítat rozptyl. Z vlastností Poissonova rozdělení je $\mathrm{var}(X_i) = \lambda$, tj. $\mathrm{var}(X_1 + \cdots + X_n) = n\lambda$ a $\mathrm{var}(\bar{X}_n) = n\lambda/n^2 = \lambda/n$.