

10. cvičení z PSt — 24.4.2024

Ještě nezávislé vektory

1. (Buffonova jehla) Na nekonečnou podlahu hodíme náhodně jehlu délky ℓ . Podlaha je z prken, jejichž okraje tvoří rovnoběžné přímký ve vzdálenosti $d \geq \ell$. Určete pravděpodobnost, že jehla bude přesahovat okraj některého prkna.

Celková pravděpodobnost

2. Pro n.n.v. $X \sim U(0, 2)$ a $Y \sim U(0, 1)$ zkoumáme $P(X < Y)$. Řešte několika způsoby:

(a) Přímou z obrázku.

(b) Rozborem možností n.v. Y pomocí vzorce (analogie věty o celkové pravděpodobnosti)

$$P(X < Y) = \int_0^1 f_Y(y)P(X < Y | Y = y)dy.$$

(c) Rozborem možností n.v. X pomocí vzorce

$$P(X < Y) = \int_0^2 f_X(x)P(X < Y | X = x)dx.$$

Konvoluce

3. Buďte $X, Y, Z \sim U(0, 1)$ nezávislé náhodné veličiny.

(a) Jaké je rozdělení $X + Y$? Určete hustotu (dvěma způsoby) – podle konvolučního vzorce i „podle obrázku“.

(b) Jaké je rozdělení $X + Y + Z$? Pro jednoduchost určete hustotní funkci jen na intervalu $[0, 1]$.

(c) Jak výsledek ověřit sámplováním?

Aplikace nerovností

4. Hážíme kostkou, za 1 a 2 dostaneme bod. Označme X počet bodů, které dostaneme po n (nezávislých) hodech. Odhadněte pravděpodobnost, že $X \geq n/2$.

(a) Pomocí Markovovy nerovnosti.

(b) Pomocí Čebyševovy nerovnosti.

(c) Pro konkrétní n , jak lze tuto hodnotu určitě přesně?

5. Statistik chce odhadnout průměrnou výšku h (v metrech) lidí v nějaké populaci, pomocí n nezávislých vzorků $X_1, \dots, X_n$, které vybíráme uniformně náhodně se všech možných lidí. Pro odhad použije výběrový průměr $\bar{X}_n = (X_1 + \dots + X_n)/n$. Odhaduje, že směrodatná odchylka jednoho výběru je nejvýše 1 metr.

(a) Jak velké n má volit, aby směrodatná odchylka $\bar{X}_n$ byla nejvýše 1 cm?

(b) Pro jaké n zajistí Čebyševova nerovnost, že $\bar{X}_n$ se liší od h nejvýše o 5 cm s pravděpodobností alespoň 99 %? (Neboli $P(|\bar{X}_n - h| \leq 5) \geq 0.99$.)

(c) Statistik si všimne, že všichni měření lidé mají výšku v intervalu $(1.4, 2.1)$. Jak má upravit odhad směrodatné odchylky? Jak se změní odpovědi na předchozí otázky?

Soupis vzorečků

- **Konvoluce:** Nechtě X, Y jsou spojitě n.n.v. Pak $S = X + Y$ má hustotu $f_S(s) = \int_{-\infty}^{\infty} f_X(x)f_Y(s-x)dx$.
- **Markovova nerovnost:** $P(X \geq a\mathbb{E}(X)) \leq \frac{1}{a}$ pro $X \geq 0$.
- **Čebyševova nerovnost:** $P(|X - \mathbb{E}(X)| \geq t\sigma_X) \leq \frac{1}{t^2}$.

Nápověda

- 1: Nakreslete obrázek a popište polohu jehly pomocí dvou náhodných proměnných (posun a úhel). Lze řešit jako úlohu na obsah, nebo pomocí sdružené hustoty.
- 3: Použijte vzorec na konci první stránky.
- 4: Jaké je rozdělení X ? Pamatujete si vzoreček na jeho střední hodnotu a rozptyl?
- 5: Jak se zjistí rozptyl součtu nezávislých náhodných veličin? Rozptyl násobku veličiny konstantou?

K procvičení

6. Buďte $X, Y, Z \sim \text{Exp}(\lambda)$ nezávislé náhodné veličiny.
 - (a) Jaké je rozdělení $X + Y$?
 - (b) Jaké je rozdělení $X + Y + Z$?
7. Počítání obsahu kruhu náhodným samplováním. Vygenerujeme náhodný bod ve čtverci (obě souřadnice budou mít rozdělení $U(0, 1)$). Označíme X_i indikátor jevu „ i -tý bod leží ve vepsaném kruhu“.
 - (a) Určete $\mathbb{E}(X_i)$, $\text{var}(X_i)$.
 - (b) Položte $\bar{X}_n = (X_1 + \dots + X_n)/n$. Určete $\mathbb{E}(\bar{X}_n)$ a $\text{var}(\bar{X}_n)$.
 - (c) Pro jaké n čekáte, že dostaneme výsledek správně na jedno desetinné místo? Na dvě, tři, ...?
 - (d) Jiný výpočet obsahu kruhu: $Y_i = \sqrt{1 - U_i^2}$, kde $U_i \sim U(0, 1)$. Uvědomte si, že $\mathbb{E}(Y_i)$ je obsah čtvrtkruhu, tedy $\pi/4$. Jaké je $\text{var}(Y_i)$? Jaké je $\text{var}(\bar{Y}_n)$?
 - (e) Která metoda je přesnější?
8. Víme, že průměrný počet bodů z písemky byl 40 (ze 100). Odhadněte odsud podíl studentů s alespoň 80 body. Vylepšete odhad, pokud víte, že směrodatná odchylka počtu bodů je 10.
9. Metrový klacek rozložíme na tři kusy jedním z níže popsanych způsobů. Pro každý z nich spočítejte, jaká je pravděpodobnost, že ze získaných tří kusů jde sestavit trojúhelník. (Nápověda: napřed si rozmyslete, kdy jsou tři kladná čísla se součtem jedna stranami nějakého trojúhelníku.)
 - (a) Vybereme uniformně náhodně dva body zlomu.
 - (b) Vybereme uniformně náhodně první bod zlomu. Pak totéž uděláme s kusem klacku v pravé ruce.
 - (c) Vybereme uniformně náhodně první bod zlomu. Pak totéž uděláme s větším kusem klacku.
10. Volme uniformně náhodně bod z trojúhelníku s vrcholy v bodech $[0, 0]$, $[0, 1]$ a $[1, 0]$, tj. pravděpodobnost každé podmnožiny je úměrná jejímu obsahu. Označme X, Y souřadnice zvoleného bodu.
 - (a) Najděte sdruženou hustotu $f_{X,Y}$.
 - (b) Najděte marginální hustotu f_Y .
 - (c) Najděte podmíněnou hustotu $f_{X|Y}$.
 - (d) Spočítejte $\mathbb{E}(X | Y = y)$ a podle věty o rozboru možností spočítejte $\mathbb{E}(X)$ (pomocí $\mathbb{E}(Y)$).
 - (e) Spočítejte $\mathbb{E}(X)$ pomocí předchozí části a symetrie.
11. Označme $S = \sum_{k=0}^{30} \binom{100}{k}$. Označme dále $X = \sum_{i=1}^{100} X_i$, kde X_i je ± 1 s pravděpodobnostmi $1/2$ a veličiny $X_1, \dots, X_n$ jsou nezávislé.
 - (a) Vyjádřete S pomocí pravděpodobnosti $P(X \leq x)$ pro vhodné x .
 - (b) Použijte CLV na odhad této pravděpodobnosti.
 - (c) Případně vyčíslíte S vhodným softwarem a srovnejte.

Centrální limitní věta

Nechť $X_1, \dots, X_n$ jsou stejně rozdělené n.n.v. se střední hodnotou μ a rozptylem σ^2 . Označme $Y_n = ((X_1 + \dots + X_n) - n\mu)/(\sqrt{n} \cdot \sigma)$. Pak $Y_n \xrightarrow{d} N(0, 1)$. Neboli

$$\lim_{n \rightarrow \infty} F_{Y_n}(x) = \Phi(x) \quad \text{pro každé } x \in \mathbb{R}.$$

12. Odhadněte $\binom{100}{30}$ pomocí CLV. Nápověda: použijte CLV pro odhad $P(29.5 < X < 30.5)$ pro vhodnou n.v. X . Na druhou stranu pro $P(X = 30)$ máme vzorec $\binom{100}{30}/2^{100}$ z binomického rozdělení. Pokud máte po ruce vhodný stroj, vyčíslíte.