

OPAKOVÁNÍ NÁHODNÝCH VEKTORŮ, EMPIRICKÁ DISTRIBUČNÍ FUNKCE

28.11.2017

-
- Máte celkem 1 000 Kč a jdete nakupovat. V obchodě utratíte zcela náhodně X (v tis. Kč) za jídlo a Y (v tis. Kč) za jiné spotřební zboží. Předpokládejme, že náhodný vektor $(X, Y)'$ má rovnoměrné rozdělení na množině přípustných hodnot, tj. na množině $M = \{(x, y) : 0 \leq x \leq 1, 0 \leq y \leq 1 - x\}$.
 - Zapište hustotu vektoru $(X, Y)'$.
 - Spočítejte marginální rozdělení X a Y .
 - Rozhodněte, zda je částka utracená za jídlo nezávislá na částce utracené za ostatní spotřební zboží. Pokud není nezávislá, spočítejte korelační koeficient.
 - Určete, s jakou pravděpodobností je utratíte za jídlo více než za ostatní spotřební zboží.
 - Spočítejte korelaci veličin $U = 2X$ a $V = Y + 1$.
 - Určete střední hodnotu a rozptyl celkové částky, kterou v obchodě utratíte.
 - Určete střední hodnotu a rozptyl částky vyjadřující, o kolik je X větší než Y .
 - Náhodná veličina X je indikátor jevu, zda vybraná osoba píše primárně levou rukou.
 - Jaké rozdělení má X ? Co vyjadřuje parametr tohoto rozdělení? Načrtněte distribuční funkci tohoto rozdělení.
 - Jaká je střední hodnota a rozptyl X ?
 - Uvažujme náhodný výběr X_1, \dots, X_n z výše uvedeného rozdělení, který je realizován na našem cvičení. Nakreslete empirickou distribuční funkci odpovídající tomuto výběru.
 - Jak bychom na základě empirické distribuční funkce odhadli pravděpodobnost, že je náhodně vybraná osoba levák? Jak bychom odhadli pravděpodobnost, že je to pravák? Napište konkrétní odhady a obecný zápis pomocí X_i .
 - Uvědomte si, že odhady z (d) jsou realizace náhodných veličin. Jaké jsou střední hodnoty a rozptyly těchto veličin?
 - Jak se chová rozptyl odhadů z (d) pro $n \rightarrow \infty$? Kolik bychom museli mít k dispozici pozorování, abychom měli jistotu, že rozptyl zmíněných odhadů bude určitě menší než 10^{-4} ?
 - V jedné základní škole ve třídě 2.B jsou z 28 dětí 3 leváci. Nakreslete empirickou distribuční funkci a vyjádřete odhady z (d) pro tato data. Porovnejte s výsledky z našeho cvičení.
 - Počet studentů, kteří dorazí na cvičení v úterý ráno je náhodná veličina. K dispozici máme 8 jejich realizací a lze předpokládat, že návštěvnosti v jednotlivých týdnech jsou na sobě nezávislé náhodné veličiny.
 - Rozmyslete si, jaké rozdělení bychom mohli teoreticky uvažovat.
 - Skutečné návštěvnosti byly následující: 25, 26, 26, 23, 25, 27, 26, 26. Nakreslete empirickou distribuční funkci spočtenou z výše uvedených dat.
 - Pomocí empirické distribuční funkce odhadněte pravděpodobnost, že počet studentů na cvičení nebude vyšší než je kapacita učebny (24 míst).

OPAKOVÁNÍ

EMPIRICKÁ DISTRIBUČNÍ FUNKCE: Necht' X_1, \dots, X_n jsou nezávislé stejně rozdělení náhodné veličiny (tzv. **náhodný výběr**) z rozdělení s distribuční funkcí F . Pak empirická distribuční funkce \widehat{F}_n je

$$\widehat{F}_n(x) = \frac{1}{n} \sum_{i=1}^n I[X_i \leq x], \quad x \in \mathbb{R}.$$

- Empirická distribuční funkce slouží jako odhad distribuční funkce F .
- Jelikož \widehat{F}_n závisí na X_1, \dots, X_n , je $\widehat{F}_n(x)$ pro každé $x \in \mathbb{R}$ náhodná veličina.
- Pro pevné x má $n\widehat{F}_n(x)$ binomické rozdělení $\text{Bi}(n, F(x))$.