Náhodná veličina a její rozdělení

24.10.2017

1. Do volební urny bylo vhozeno a volebních lístků strany A a $N-a$ lístků jiných stran. Před volební místností zcela náhodně zastavíme n osob a zeptáme se jich, koho volili (a všichni nám to popravdě sdělí). Označme jako X počet osob, které jsme oslovili a které volily stranu A.
   (a) Určete rozdělení náhodné veličiny X.
   (b) Na základě výsledku našeho průzkumu (tj. náhodné veličiny X) chceme odhadnout volební výsledek strany A. Předpokládejme, že lístky jsou v urně zastoupeny v poměru 3:7 a že jsme oslovili $n = 20$ osob. S jakou pravděpodobností bude náš odhad X/n ležet v intervalu $[0.25, 0.35]$? Napište výsledek nejprve obecně a poté vyčíslete pro $N = 100$ a $N = 1000$.
   (c) Zakreslete graf pravděpodobností X a jeho distribuční funkci. Vyznačte na obou grafech pravděpodobnost z (b).

2. Předpokládejme, že můžeme oslovit všechny občany (tj. ne jen ty z dané volební místnosti) a že opět náhodně vybereme n lidí. Nyní lze tedy předpokládat, že každý z oslovených je s pravděpodobností $p = 0.3$ voličem strany A. Označme jako Y náhodnou veličinu udávající počet osob, které jsme oslovili a které volí stranu A.
   (a) Odvoďte rozdělení náhodné veličiny Y.
   (b) Spočtěte pro $n = 20$, s jakou pravděpodobností bude Váš odhad Y/n v intervalu $[0.25, 0.35]$.
   (c) Porovnejte výsledky (b) v 1 a 2. Jaký je jejich vzájemný vztah?
   (d) Formulujte oba příklady 1 a 2 pomocí urny s kuličkami různých barev.

3. Předpokládejte, že máme k dispozici perfektní generátor náhodných čísel z intervalu $[0, 1]$. Označme jako X náhodnou veličinu, která nám udává vygenerované číslo.
   (a) Jak se liší rozdělení této veličiny od veličin z příkladů 1 a 2? Jakým způsobem popíšeme rozdělení X? Zapište a nakreslete graf.
   (b) Nakreslete také graf distribuční funkce. V obou obrázcích zakreslete pravděpodobnost, s jakou dostaneme číslo menší než 0.5. S jakou pravděpodobností dostaneme přesně 0.5?
   (c) Zkuste navrhnout, jak pomocí realizací X simulovat situace z příkladů 1 a 2.
   (d) Získané náhodné číslo X umocníme na druhou a dostaneme tak jiné náhodné číslo Y. Jaké je rozdělení Y?

4. Doba strávená čekáním na příjezd vlaku (v minutách) je náhodná veličina s hustotou
$$f(x) = \begin{cases} ce^{-x/5}, & x \geq 0, \\ 0, & x < 0. \end{cases}$$
   (a) Určete konstantu $c > 0$, tak aby f byla hustota.
   (b) Určete distribuční funkci F a načrtněte ji.
   (c) Jaká je pravděpodobnost, že budete na vlak čekat déle než 5 minut? Vyznačte v grafu hustoty a v grafu distribuční funkce.
   (d) S jakou pravděpodobností bude doba strávená čekáním mezi 2 a 5 min?
   (e) Aktuálně čekáte 5 min. Jaká je pravděpodobnost, že budete celkově čekat déle než 10 min?
   (f) Na vlak čekáte $n = 5$ dní po sobě. S jakou pravděpodobností jste nikdy nečekali méně než 5 min?

Opakování

Náhodná veličina: **Náhodná veličina** X je měřitelné zobrazení (funkce) z prostoru $(\Omega, \mathcal{A})$ do $(\mathbb{R}, \mathcal{B})$. Jednotlivým prvkům $\omega \in \Omega$ tedy přiřazuje reálná čísla $X(\omega)$.

- **Rozdělení** náhodné veličiny X
  - popisuje pravděpodobnosti $\mathsf{P}(X \in B) = \mathsf{P}(\{\omega : X(\omega) \in B\})$ pro všechny množiny $B \in \mathcal{B}$,
  - je jednoznačně určeno **distribuční funkcí**, která je funkcí reálné proměnné $x \in \mathbb{R}$ a je definovaná jako
  $$F(x) = \mathsf{P}(X \leq x).$$
  Distribuční funkce je vždy neklesající, zprava spojitá s limitou 0 v $-\infty$ a limitou 1 v ∞.
  - Platí $\mathsf{P}(X \in (a, b]) = F(b) - F(a)$ pro libovolné $a < b$.
- Nabývá-li náhodná veličina X s kladnou pravděpodobností **nejvýše spočetně** mnoha (tj. konečně nebo spočetně) hodnot $x_1, x_2, \ldots$, říkáme, že má **diskrétní rozdělení**.
  - Rozdělení X je charakterizováno pravděpodobnostmi $p_k = \mathsf{P}(X = x_k)$, $k = 1, 2, \ldots$ a platí $\sum_k p_k = 1$.
  - **Distribuční funkce** je po částech konstantní, skokovitá se skoky o velikosti p_k v bodech x_k.
- Jestliže existuje funkce $f \geq 0$ taková, že $F(x) = \int_{-\infty}^{x} f(t)dt$, pak říkáme, že X má **spojité rozdělení**. Funkce f se nazývá **hustota**.
  - Spojitá náhodná veličina nabývá **nespočetně mnoha** hodnot z nějakého podintervalu $\mathbb{R}$.
  - Rozdělení veličiny X je charakterizováno hustotou $f \geq 0$. Pro každou $B \in \mathcal{B}$ je pak
  $$\mathsf{P}(X \in B) = \int_B f(x)dx.$$
  Speciálně:
  - $1 = \mathsf{P}(X \in \mathbb{R}) = \int_{-\infty}^{\infty} f(x)dx$,
  - **distribuční funkce** F je spojitá na $\mathbb{R}$ a lze ji spočítat jako $F(x) = \int_{-\infty}^{x} f(t)dt$,
  - pro libovolné $a \in \mathbb{R}$ je $\mathsf{P}(X = a) = \int_{\{a\}} f(t)dt = 0$,
  - je-li $a < b$, pak
  $$\mathsf{P}(a < X < b) = F(b) - F(a) = \int_a^b f(t)dt.$$