NMAI059 Pravděpodobnost a statistika 1
2. přednáška

Robert Šámal

Přehled

Diskrétní náhodné veličiny

Příklady diskrétních n.v.

Střední hodnota

Bonusy

Náhodná proměnná

Často nás zajímá číslo dané výsledkem náhodného pokusu.

- Hodíme na terč a změříme vzdálenost od středu.
- Házíme kostkou, dokud nepadne šestka, ale pak si všimneme jenom toho, kolik hodů to trvalo.
- U quicksortu (algoritmus na třídění) měříme počet kroků (v závislosti na náhodné volbě pivotu).

Definice

Mějme pravděpodobnostní prostor $(\Omega, \mathcal{F}, P)$. *Funkci* $X : \Omega \to \mathbb{R}$ *nazveme diskrétní náhodná veličina (discrete random variable), pokud* $Im(X)$ *(obor hodnot* X*) je spočetná množina a pokud pro všechna reálná* x *platí*

$$\{\omega \in \Omega : X(\omega) = x\} \in \mathcal{F}.$$

Pravděpodobnostní funkce

Definice

Pravděpodobnostní funkce (probability mass function, pmf) diskrétní náhodné veličiny X je funkce $p_X : \mathbb{R} \to [0,1]$ *taková, že*

$$p_X(x) = P(X = x).$$

- Označíme-li S obor hodnot X a $Q(A) := \sum_{x \in A} p_X(x)$, tak $(S, \mathcal{P}(S), Q)$ je diskrétní pravděpodobnostní prostor.
- Pro $S = \{s_i : i \in I\}$ spočetnou množinu reálných čísel a $c_i \in [0,1]$ splňující $\sum_{i \in I} c_i = 1$ existuje pravděpodobnostní prostor a diskrétní n.v. X na něm taková, že $p_X(s_i) = c_i$ pro $i \in I$.

Přehled

Bernoulliho/alternativní rozdělení

- $X =$ počet orlů při jednom hodu nespravedlivou mincí.
- Značíme $X \sim Bern(p)$.

- Dáno $p \in [0, 1]$.
- $p_X(1) = p$
- $p_X(0) = 1 - p$
- $p_X(k) = 0$ pro $k \neq 0, 1$

- Pro libovolný jev $A \in \mathcal{F}$ definujeme *indikátorovou n.v.* I_A:
- $I_A(\omega) = 1$ pokud $\omega \in A$, $I_A(\omega) = 0$ jinak.
- $I_A \sim Bern(P(A))$

Binomiální rozdělení

- $X =$ počet orlů při n hodech nespravedlivou mincí.
- Značíme $X \sim Bin(n, p)$.

- Dáno $p \in [0, 1]$.
- $p_X(k) = \binom{n}{k} p^k (1-p)^{n-k}$ pro $k \in \{0, 1, \ldots, n\}$

TODO: OBR.

Poissonovo rozdělení

- Značíme $X \sim Pois(\lambda)$.

- Dáno reálné $\lambda > 0$.
- $p_X(k) = \frac{\lambda^k}{k!} e^{-\lambda}$

- $Pois(\lambda)$ je limitou $Bin(n, \lambda/n)$

- X popisuje např. počet emailů, které dostaneme za jednu hodinu.

TODO: OBR.

Poissonovo paradigma

- $A_1, \dots, A_n$ jsou (skoro-)nezávislé jevy s $P(A_i) = p_i$, $\lambda = \sum_i p_i$. Nechť n je velké, každé z p_i malé. Pak přibližně platí
$$\sum_{i=1}^{n} I_{A_i} \sim Pois(\lambda).$$

Geometrické rozdělení

- $X =$ kolikátým hodem mincí padl první orel.
- Značíme $X \sim Geom(p)$.

- Dáno $p \in [0, 1]$.
- $p_X(k) = (1-p)^{k-1}p$

- Někdy se tomuto rozdělení říká posunuté geometrické, a za normální geometrické se považuje rozdělení $X - 1$, tj. počet neúspěšných hodů.

Přehled

Střední hodnota

Definice

Pokud X je diskrétní n.v., tak její střední hodnota (expectation) je označována $\mathbb{E}(X)$ a definována

$$\mathbb{E}(X) = \sum_{x \in Im(X)} x \cdot P(X = x),$$

pokud součet má smysl.

LOTUS

- Pro reálnou funkci g a diskrétní n.v. X je $Y = g(X)$ také diskrétní n.v.

Věta (LOTUS)

Pokud X je diskrétní n.v. a g reálná funkce, tak

$$\mathbb{E}(g(X)) = \sum_{x \in Im(X)} g(x)P(X = x)$$

pokud součet má smysl.

Vlastnosti $\mathbb{E}$

Věta

Nechť X, Y jsou diskrétní n.v. a $a, b \in \mathbb{R}$.

1. *Pokud $P(X \geq 0) = 1$ a $\mathbb{E}(X) = 0$, tak $P(X = 0) = 1$.*
2. *Pokud $\mathbb{E}(X) \geq 0$ tak $P(X \geq 0) > 0$.*
3. $\mathbb{E}(a \cdot X + b) = a \cdot \mathbb{E}(X) + b$.
4. $\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$.

Podmíněná střední hodnota

Definice

Pokud X je diskrétní n.v. a $P(B) > 0$, tak podmíněná střední hodnota X za předpokladu B (conditional expectation of X given B) je

$$\mathbb{E}(X \mid B) = \sum_{x \in Im(X)} x \cdot P(X = x \mid B),$$

pokud součet má smysl.

Rozbor všech možností

Věta

Pokud $B_1, B_2, \dots$ *je rozklad* Ω *a* $A \in \mathcal{F}$*, tak*

$$\mathbb{E}(X) = \sum_i \mathbb{E}(X \mid B_i) P(B_i),$$

kdykoliv má součet smysl. (Sčítance s $P(B_i) = 0$ *považujeme za 0.)*

Rozbor všech možností

Přehled

Bertrand's paradox

Simpsons's paradox