Míry závislosti

Monika Matoušková

25. března 2020

Co je to míra závislosti

- Míra závislosti měří závislost dvou náhodných veličin
- Asi nejjednodušší je korelace

Definice

X, Y náhodné veličiny takové, že $\mathrm{var} X \in (0, \infty)$, $\mathrm{var}\, Y \in (0, \infty)$. Definujeme korelaci X, Y jako $\rho(X, Y) = \frac{\mathbb{E}XY - \mathbb{E}X\mathbb{E}Y}{\sqrt{\mathrm{var} X}\sqrt{\mathrm{var}\, Y}}$.

- $\rho(X, Y) \in [-1, 1]$
- Pro korelaci platí: X, Y nezávislé $\implies \rho(X, Y) = 0$
- Neplatí však obrácená implikace

Vlastnosti míry závislosti

- Chtěli bychom, aby dobrá míra závislosti $\delta(X, Y)$ měla tyto vlastnosti:
  1. $\delta(X, Y)$ je definována pro libovolné dvě nedegenerované náhodné veličiny X, Y
  2. $\delta(X, Y) = \delta(Y, X)$
  3. $\delta(X, Y) \in [0, 1]$
  4. $\delta(X, Y) = 0 \iff X, Y$ jsou nezávislé
  5. $\delta(X, Y) = 1$ pokud platí $X = g(Y)$ nebo $Y = f(X)$
  6. Pokud f, g jsou bijekce, pak $\delta(f(X), g(Y)) = \delta(X, Y)$
  7. Pokud X, Y mají sdružené normální rozdělení, pak $\delta(X, Y) = |\rho(X, Y)|$
- V bodu 4 a 5 uvažujeme f, g borelovsky měřitelné funkce

Maximální korelace

- $|\rho(X, Y)|$ splňuje pouze vlastnosti 2 a 3 (a samozřejmě 7)
- Všechny tyto vlastnosti splňuje maximální korelace

Definice

X, Y náhodné veličiny, definujeme maximální korelaci jako $R(X, Y) = \sup \rho(f(X), g(Y))$, kde supremum bereme přes všechny borelovsky měřitelné funkce $f, g : \mathbb{R} \to \mathbb{R}$ takové, že $0 < \operatorname{var} f(X) < \infty$, $0 < \operatorname{var} g(Y) < \infty$.

- Platí vztah $R(X, Y) \geq |\rho(X, Y)|$

Kdy umíme maximální korelaci spočítat

- Díky tomu, že $R(X, Y)$ splňuje body 5 a 7 umíme maximální korelaci spočítat v případě, že jedna náhodná veličina je funkcí druhé nebo mají sdružené normální rozdělení
- Platí, že pokud X, Y nabývají pouze dvou hodnot skoro jistě, pak $R(X, Y) = |\rho(X, Y)|$
- Pokud ale X nabývá dvou hodnot a Y už nabývá tří hodnot, lze najít případ, kdy platí $0 < |\rho(X, Y)| < R(X, Y) < 1$
- Existují ještě další velmi specifické případy, kdy umíme $R(X, Y)$ spočítat

Cíl práce

- Popsat vlastnosti některých měr závislosti a některé vlastnosti dokázat
- Uvést příklady, jako například kdy platí $0 < |\rho(X, Y)| < R(X, Y) < 1$

Zdroje

- Rényi, Alfréd. ON MEASURES OF DEPENDENCE. *Acta Mathematica Academiae Scientiarum Hungarica*, 10(3-4):441-451, 1959.

Děkuji za pozornost.