mmzm_np

September 29, 2016

1 Náhodná proměnná

- obor hodnot = všechny možné případy
- konečný/spočetný (diskrétní)

(rozlišovat NP versus hodnota NP)
v případě spojitých náhod. proměnných (obor reálných čísel) zavádíme *hustotu pravděpodobnosti*

$$f(x) = lim_{dx \to 0} \frac{P(x \leq \xi \leq x + dx)}{dx}$$

- je nenulová, ale může být >1
- lze rozšířit i na diskrétní proměnné (KO:jak?)

příklad kvantové mechaniky : spektrum energií může být zčásti diskrétní, zčásti spojité
výhodné popsání **distribuční funkcí** $F(x) = P(\xi < x)$ - neklesajici, od 0 do 1

1.0.1 funkce (= transformace) NP

- přenos z intervalů původní proměnné do nové proměnné
- je-li transformace $y = h(x)$ vzájemně jednoznačná

pro hustotu g(y) v nové proměnné

$$g(y) = f(x)\frac{dx}{dy} = \frac{f(x)}{|h'(x)|} = \frac{f(h^{-1}(y))}{|h'(h_i(y))|}$$

h^{-1} je inverzní k h

1.0.2 v případě více proměnných (náhodný vektor)

$$F(x_1, x_2, ...x_n) = P(\xi_1 < x_1 \wedge \xi_2 < x_2 \wedge ...\xi_n < x_n)$$

souvisí s hustotou NP analogicky s 1-D případem
některé proměnné lze "odintegrovat" (tzv. *marginalizace*)
zbude-li jediná proměnná, jde o *marginální rozdělení* (projekce do daného směru)

$$F_{\xi_1}(x_1) = F_\xi(x_1, \infty, ..., \infty)$$

fixování jedné komponenty (či více) vytváří *řez rozdělovací funkce* (podmíněné rozdělení) (nutno ji normovat pomocí marginálního rozdělení)

$$f_p(x_1, x_2, ...|\xi_n = x_0) = \frac{f(x_1, x_2, ..., x_{n0})}{f_{\xi_n}(x_{n0})}$$

- nezávislost komponent

$$F_\xi(x_1, x_2) = F_{\xi_1}(x_1)F_{\xi_2}(x_2)$$

2 Vlastnosti náhodných proměnných

2.0.3 charakteristiky

očekávaná hodnota funkce g náhodné proměnné

$$E(g) = \int_\Omega g(X)f(X)dX$$

střední hodnota (matem. očekávání = expektance) $E(\xi)$
event. pracujeme s očekávanou hodnotou z funkce NP - např.

$$D(\xi) = E\{[x - E(\xi)]^2\} = \int_{-\infty}^{\infty}\left[x - \int_{-\infty}^{\infty} xf(x)dx\right]^2 f(x)dx$$

disperze - σ^2 - jeden z centrálních momentů

- algebraické $\nu_k = E(\xi^k)$
- centrální $\mu_k = E((\xi - \nu_1)^k)$

asymetrie ("skewness", 3. řád) - $\gamma_1 = \frac{\mu_3}{\sqrt{\mu_2^3}}$
exces (a.k.a. "špičatost", 4. řád) - $\gamma_2 = \frac{\mu_4}{\mu_2^2} - 3$
korekce zavedena, aby pro normální rozdělení $\gamma_1 = \gamma_2 = 0$

2.0.4 momenty funkcí více proměnných

analogická definice střední hodnoty (či střední hodnoty funkce)
dostáváme nyní i smíšené momenty:

- smíšený druhý centrální moment (kovariace, korelační moment)

$$D(\xi_1, \xi_2) = E([\xi_1 - E(\xi_1)][\xi_2 - E(\xi_2)]) = E(\xi_1\xi_2) - E(\xi_1)E(\xi_2)$$

odtud korelační koeficient

$$\rho(\xi_1, \xi_2) = D(\xi_1, \xi_2)/\sqrt{D(\xi_1)D(\xi_2)}$$

pro nezávislé vektory nulový, max. 1 pro plně korelované ("úměrné")

- !! existují závislé NP, které mají nulový korel. koeficient: **nekorelovanost** je slabší vlastnost než **nezávislost**

korelační moment lze zavést u n-rozměrného náh. vektoru pro lib. dvojici komponent z marginálního rozdělení (odintegrováním zbylých složek)

$D_{ij} = D(\xi_i, \xi_j)$ - matice (kovarianční, disperzní) je symetrická, na diagonále disperze komponent

Matice je singulární, pokud existuje lineární kombinace složek, která je nulová (jedna komponenta je lineární kombinací jiných). Jinak lze zavést inv. matici D^{-1}, a kromě

korelační matice $\rho_{ij} = \rho(\xi_i, \xi_j) = D_{ij}/\sqrt{D_{ii}D_{jj}}$

spočítat i *globální korelační koeficient* pro danou komponentu, určující její maximální míru korelace s libovolnou lin. kombinací zbylých složek: platí

$$\rho_i = \sqrt{1 - 1/(D_{ii}{D^{-1}}_{ii})}$$

2.0.5 Charakteristická funkce

$$X(t) = E(e^{i\xi t}) = \int e^{ixt} f_\xi(x)dx$$

přičtení konstanty A znamená vynásobení $exp(iAt)$, faktor *a* změní výsledek $X_{a\xi}(t) = X_\xi(at)$

$$X_{\xi+\theta}(t) = X_\xi(t)X_\theta(t)$$

pro normální rozdělení $X(t) = \int e^{ixt}e^{(x-m)^2/2\sigma^2}dx = e^{imt-t^2\sigma^2/2}$

2.0.6 Generující funkce

$$M(t) = E(e^{\xi t}) = \int e^{xt} f_\xi(x)dx$$

rozvojem exponenciály získáme souvislost s momenty (necentrálními)

$$M(t) = E\left[1 + \xi t + \frac{1}{2!}\xi t + \dots\right] = \sum_{n=0} \frac{1}{n}{\mu'}_n t^n$$

odkud lze vyjádřit

$${\mu'}_n = \frac{\partial^n M(t)}{\partial t^n}$$

v bodě $t = 0$

2.1 lineární kombinace NP

$\Theta = \sum_i a_i\xi_i$ +A

- **střední hodnota** $E(\Theta) = \sum_i a_i E(\xi_i) + A$ (důkaz)
- **disperze**

$$D(\Theta) = E(\sum_i \sum_j a_i a_j [\xi_i - E(\xi_i)][\xi_j - E(\xi_j)] = \sum_i a_i^2 D(\xi_i) + \sum_i \sum_{j \neq i} a_i a_j D(\xi_i, \xi_j)$$

pro nekorelované proměnné druhý člen odpadá

- **rozdělení**

hustota pravděp. h(y) nové NP:

1. vynásobení **konstantou** a: $h(y) = f(y/a)/|a|$
2. **součet** $y = x_1 + x_2$

dle distribuční funkce

$$F(y) = \int_{x_1+x_2<y} f(x_1, x_2) dx_1 dx_2 = \int_{-\infty}^{\infty} dx_2 \int_{-\infty}^{y-x_2} f(x_1, x_2) dx_1$$

pro nezávislé proměnné: $f(x_1, x_2) = f_1(x_1) f_2(x_2)$ při zavedení proměnné t se substitucí $x_1 = t - x_2$

$$F(y) = \int_{-\infty}^{y} dt \int_{-\infty}^{\infty} f_2(x_2) f_1(t - x_2) dx_2 = \int_{-\infty}^{y} f(t) dt,$$

kde $f(t) = \int f_1(t-x) f_2(x) dx$ je (Fourierova) konvoluce hustot; analogicky pro součin nezávislých proměnných $y = x_1 * x_2$ dostáváme hustotu pravděpodobnosti jako (Mellinovu) konvoluci

$$f(t) = \int f_1(t/x) f_2(x) dx/|x|$$