

Mnohorozměrný test homogenity pro prostorově korelovaná data

ROBUST 2016

Tereza Šimková

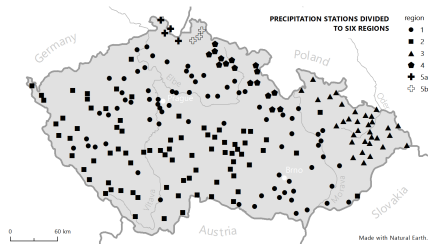
Technická univerzita v Liberci
Fakulta přírodovědně-humanitní a pedagogická
Katedra aplikované matematiky

13. 9. 2016

- regionální frekvenční analýza (RFA): dosažení přesnějších odhadů vysokých kvantilů analyzováním dat v celém homogenním regionu
- homogenní region: množina stanic s přibližně stejným rozdělením až na měřítko
- 2 hlavní kroky RFA: 1) identifikace homogenních regionů
2) odhad vysokých kvantilů
- parametrický test (Chebana a Ouarda, 2007) a neparametrický test (Masselot, Chebana a Ouarda, 2016) založený na L-momentech
- testy nezohledňují závislosti mezi daty stanic

Tabulka: Výsledky testování homogenity

Region	Míra heterogenity	Závěr
1	1.4884	příjemně homogenní
2	1.1316	příjemně homogenní
3	-1.3857	homogenní
4	0.7111	homogenní
5a	-1.5404; -1.5305	homogenní
5b	-0.7635	homogenní



Obrázek: Rozdělení meteorologických stanic do 6 regionů

- ilustrovat vliv závislosti mezi daty stanic na výše uvedené testy
- navrhnout zobecnění zohledňující závislosti mezi daty stanic

MNOHOROZMĚRNÉ L-MOMENTY

- náhodný vektor $X = (X_1, \dots, X_n)$
- $\Lambda_1 = EX$
- $\Lambda_k = (\lambda_{k[ij]})_{i,j=1}^n, k \geq 2$
- k -tý L-komoment n. v. X_i vzhledem k n. v. X_j :

$$\lambda_{k[ij]} = \text{cov}(X_i, P_{k-1}^*(F_j(X_j)))$$

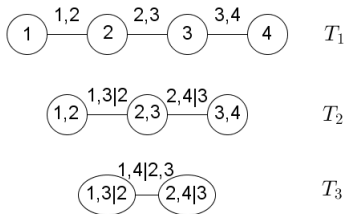
MNOHOROZMĚRNÉ TESTY HOMOGENITY ZALOŽENÉ NA L-MOMENTECH

$$V_{||\cdot||} = \left[\left(\sum_{i=1}^N n_i \| \Lambda_2^{*(i)} - \bar{\Lambda}_2^* \|^2 \right) / \left(\sum_{i=1}^N n_i \right) \right]^{1/2}$$

- parametrický test (Chebana a Ouarda, 2007)
 - data homogenních regionů získána užitím kopulí a kappa rozdělení
 - míra heterogenity $H_{||\cdot||} = (V_{||\cdot||} - \mu) / \sigma$
 - region je homogenní, jestliže $H_{||\cdot||} < 1$, přijatelně homogenní, jestliže $1 \leq H_{||\cdot||} < 2$, a rozhodně heterogenní, jestliže $H_{||\cdot||} \geq 2$
- neparametrický permutační test (Masselot, Chebana a Ouarda, 2016)
 - data homogenních regionů získána permutací původních dat
 - p – hodnota = $\frac{1}{N_{sim}} \#\{V_{||\cdot||}^{(j)} > V_{||\cdot||}\}$
- modifikace parametrického testu
 - data homogenních regionů získána užitím D-vine kopulí a kappa rozdělení

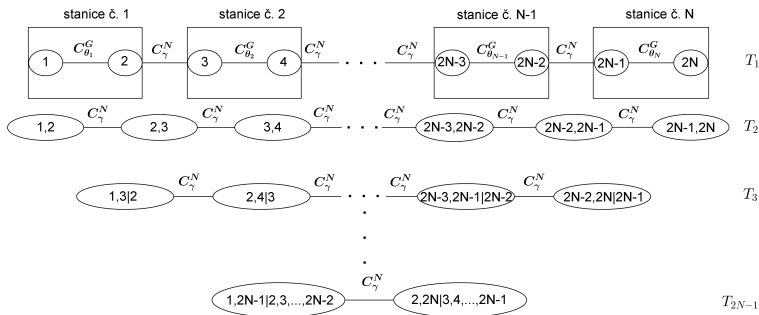
D-VINE KOPULE

- kopule je definována jako vícerozměrná distribuční funkce, jejíž jednorozměrné distribuční funkce jsou rovnoměrné na $[0, 1]$
- vine kopule umožňují modelovat závislosti jen pomocí dvourozměrných kopulí, princip je založen na rozkladu hustoty náhodného vektoru na součin marginálních hustot a hustot dvourozměrných kopulí
- d -rozměrná D-vine kopule: $d - 1$ stromů T_i , i -tý strom z $d + 1 - i$ stromů a $d - i$ hran



Obrázek: Grafická struktura čtyřrozměrné D-vine kopule

- zjištění vlivu korelace mezi daty na testy a odhad síly navrženého testu
- 100 reprezentativních regionů určitého typu o N stanicích vygenerováno užitím $2N$ -rozměrné D-vine kopule

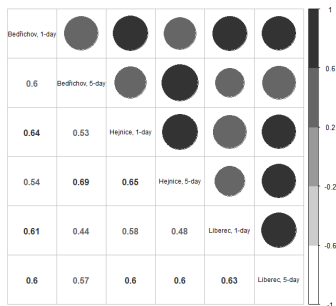


Obrázek: Grafická struktura D-vine kopule použité v simulaci

- korelace snižuje míru heterogenity a může vést k mylným závěrům
- empirická chyba 1. druhu a síla v případě nekorelovaného regionu je stejná pro parametrický test i jeho modifikaci
- navržená modifikace dává hodnotu míry heterogenity odpovídající typu regionu (na rozdíl od původního)
- v případě nekorelovaného regionu dosahuje nejvyšší síly neparametrický permutační test
- v případě korelovaného regionu dosahuje nejvyšší síly navržená modifikace
- odhad chyby 1. druhu navržené modifikace nepřekračuje 7 %

APLIKACE

- použitelnost navržené modifikace je ilustrována na reálných meteorologických datech regionu 5b (Hejnice, Bedřichov, Liberec)
- parametrický test: homogenní ($H_{||\cdot||} = -0.5632$)
- neparametrický permutační test: nelze zamítnout hypotézu o homogenitě ($p - hodnota = 0.702$)
- navržená modifikace permutačního testu: heterogenní ($H_{||\cdot||} = 3.5256$)



Děkuji za pozornost!

- F. Chebana and T. B. M. J. Ouarda, "Multivariate L-moment homogeneity test", *Water Resources Research*, vol. 43, W08406, doi:10.1029/2006WR005639, 2007.
- P. Masselot, F. Chebana, and T. B. M. J. Ouarda, "Fast and direct nonparametric procedures in the L-moment homogeneity test", *Stochastic Environmental Research and Risk Assessment*, doi:10.1007/s00477-016-1248-0, 2016.