

Praktiskie darbi lekciju kursam “Matemātikas statistikas izvēlētas nodaļas”

Jānis Valeinis
Maijs 2021

Neparametriskie testi, blīvuma funkcijas neparametriskā gludināšana un neparametriskā regresija

Uzdevumi:

1. Salīdzināt t-testu un *wilcox* testu dažādiem datiem (ģenerētiem un iebūvētiem). Pārbaudīt, kā izlecēji ietekmē ticamības intervālus! Var salīdzināt abus testus ar pārklājuma precizitātes palīdzību. Aplūkot Koši sadalījumu, vai t-tests strādā šajā gadījumā?
2. Ielasīt programmā R internetā ieliktos datus, kuri tika analizēti lekcijā. Uzzīmēt histogrammu. Veikt neparametrisku blīvuma funkcijas novērtēšanu. Mainīt dažādus kodolus un joslas platumu, izdarīt attiecīgus secinājumus. Izvēlēties joslas platumu pēc dažādām metodēm (krosvalidācijas un citām), kas iebūvētas programmā R.
3. Programmā R uzrakstīt programmu, kas realizē neparametrisko Nadaraya-Watson kodolu regresiju. Dažādiem iebūvētiem datiem (piemēram, cars un LIDAR datiem) veikt kodolu regresiju. Pievienot arī lokālo lineāro regresijas novērtējumu kas uzlabo kodolu regresiju robežās. Pārbaudīt uz dažādiem datiem, vai tiešām iegūts uzlabojums. Joslas platumu izvēlēties pēc iebūvētajām komandām programmā R.
4. **Mājasdarbs.** Izveidot programmas kodu, kas realizē krosvalidācijas funkciju programmā R blīvuma funkcijas novērtēšanai. Atrast joslas platuma krosvalidācijas novērtējumu joslas platumam datiem kas ielikti serverī un konstruēt attiecīgo blīvuma funkcijas novērtējumu, pievienot to histogrammai. Salīdzināt ar citām metodēm.
5. **Mājasdarbs.** Ielādēt programmā R CMB datus no Larry Wasserman mājaslapas. Veikt polinomiālo un neparametrisko kodolu Nadaraya-Watson regresiju. Pārbaudīt nosacījumus, veikt diagnostiku. Izdarīt attiecīgus secinājumus. Neparametriskai regresijai novērtēt joslas platumu ar krosvalidācijas metodi un salīdzināt ar citām metodēm.