

Proiect

Serii de Timp cu Aplicații*

Se consideră o serie de timp $x = \{x_1, \dots, x_n\}$ formată din cel puțin 150 de măsurători (observații; $n \geq 150$) ale unui proces stohastic (atât trendul cât și efectul sezonier nu trebuie să fie nule). Scopul acestui proiect este de a analiza această serie (folosind R) prin aplicarea metodei netezirii exponențiale triple. Cerințe:

- se va porni cu seria de timp x restricționată la primele 90% dintre măsurători $\{x_1, \dots, x_m\}$ (adică $m = \lceil 9n/10 \rceil$) și se va face o previziune pentru restul de 10% (adică pentru valorile $\{x_{m+1}, \dots, x_n\}$); se va alege varianta de descompunere optimă (componentă sezonieră aditivă, respectiv multiplicativă) în sensul minimizării sumei erorilor de previziune;
- se vor calcula, doar pentru partea previzionată (ultimele 10% dintre valori), erorile de estimare: $MSE = \frac{\sum_{i=m+1}^n (x_i - \hat{x}_i)^2}{n-m}$, $MAE = \frac{\sum_{i=m+1}^n |x_i - \hat{x}_i|}{n-m}$ și $MAPE = \frac{100 \sum_{i=m+1}^n |x_i - \hat{x}_i| / |x_i|}{n-m}$;
- se vor reprezenta, pe același grafic (folosind culori diferite), atât seria de timp inițială, cât și cea netezită (n valori);
- se va repeta procedeul pornind de la seria de timp x restricționată la primele 75% dintre măsurători;
- comparați rezultatele obținute și încercați le interpretați.

Observații. • Nu se impun restricții în modalitatea de alegere a seriilor de timp. Datele statistice ar putea fi preluate, spre exemplu, din librăria de date *Time Series Data Library* creată de către Robin John Hyndman (Profesor de Statistică la Universitatea Monash, Melbourne, Australia) la adresa <http://data.is/TSDLdemo/> și care conține circa 800 de serii de timp. În caz că folosiți baza de date respectivă vă rog să o citați conform solicitărilor autorului (informații în acest sens se găsesc la adresa <http://robjhyndman.com/TSDL/>).

- Proiectul este individual; la alegerea datelor statistice vă rog să vă consultați între voi pentru a nu studia aceleași serii de timp în proiecte diferite.
- Ca formă de prezentare, proiectul va conține:
 - o pagină separată cu numele dumneavoastră, anul/grupa, adresa de e-mail și titlul proiectului.
 - cerințele proiectului.

*Master: Matematici Financiare, Anul universitar: 2016-2017

- câteva informații de ordin general despre seriile de timp ce urmează a fi analizate, indicând precis sursa de unde au fost preluate.
 - textele programelor **R**, cu câteva explicații referitoare la pașii mai importanți ce i-ați parcurs; explicațiile pot fi incluse ca și comentarii (precedate de simbolul **#**) în fiecare linie a acestor programe; fișierele corespunzătoare (cu extensia **.R**) se vor copia pe un CD (alături de seriile de timp folosite; de obicei fișiere de tip text - extensia **.txt**) care va fi atașat proiectului.
 - graficele obținute; se va preciza ce reprezintă fiecare.
 - se va întocmi un tabel cu două linii (corespunzătoare celor două abordări) și cu trei coloane conținând valorile MSE, MAE și MAPE.
 - se va întocmi un tabel cu două linii (corespunzătoare celor două abordări) și cu trei coloane conținând valorile concrete ale constantelor α , β și γ determinate în urma minimizării sumei erorilor de previziune.
 - interpretarea rezultatelor obținute.
- Data limită de predare: ultima oră de curs din semestru.
 - Pentru orice fel de problemă legată de acest proiect sau de tematica de curs/laborator vă rog să mă contactați la ore sau prin e-mail.

SUCCES !

CONF. DR. DAN POPOVICI