**Compararea mediilor a mai mult de două grupuri (analiza varianţei = ANOVA)**

În paginile anterioare am vorbit despre compararea mediilor a două grupuri (exemplul cu TA ale studenţilor la medicina şi ASE), arătând că testele statistice (Student = t) compară atât mediile, cât şi varianţele, şi că ele pot fi aplicate numai dacă distribuţia variabilei numerice (în cazul acela TA) este normală (Gaussiană).

Sunt cazuri în care avem, însa, de comparat între ele mai mult de două grupuri - să presupunem ca vrem să vedem dacă diferă TA ale studenţilor tuturor institutelor din Bucureşti, şi atunci avem tot atâtea grupuri câte institute de învăţământ superior există.
Anul trecut am predat metodologia cercetării ştiinţifice la patru serii de studenţi, iar notele de la testul scris au fost cele din Figura 1 (reprezentare sub formă de *boxplot*).

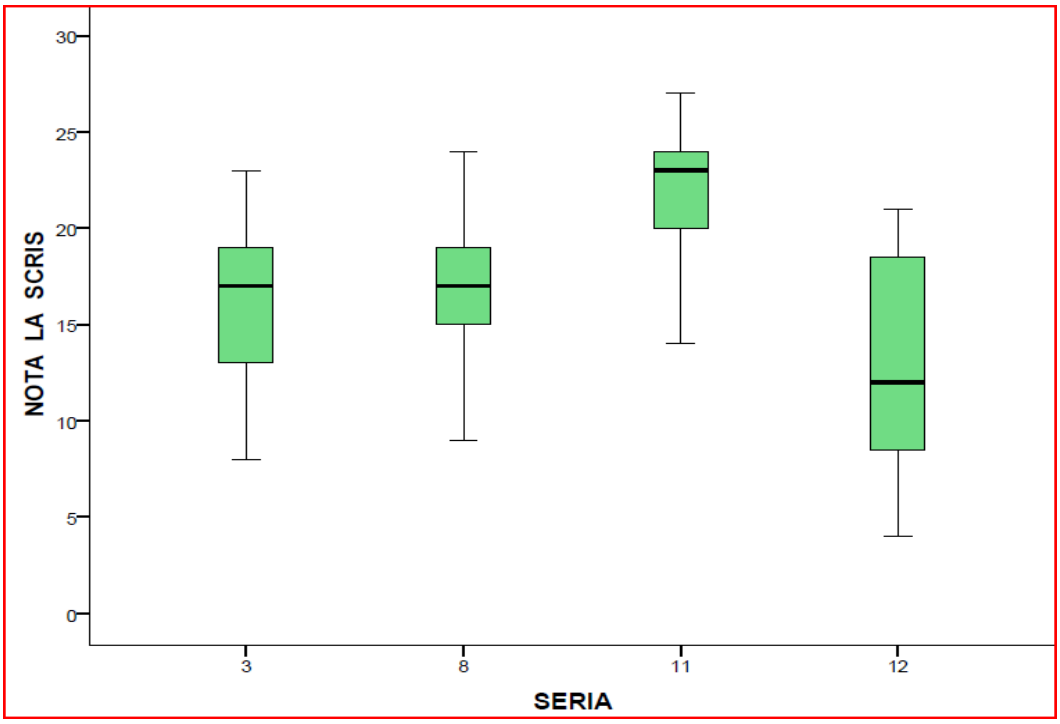


**Figura 1.** Notele a 4 serii de studenţi la testul scris.

Vă aduc aminte dintr-un articol anterior că, la reprezentarea sub formă de *boxplot*, linia centrală reprezintă mediana, marginile orizontale ale dreptunghiului reprezintă cvartilele 25 (inferioară), respectiv 75% (superioară), iar „T”-urile reprezintă limitele. Se observă că *boxplot*-urile sunt relativ simetrice, aşadar distribuţia este probabil gaussiană, deci putem aplica teste statistice parametrice, respectiv ANOVA.

Baza de date rezultată va arăta ca în Figurile 2 (în programul InStat), sau 3 (în programul SPSS).

| | Group A | Group B | Group C | Group D |
|---|---|---|---|---|
| | seria 3 | seria 8 | seria 11 | seria 12 |
| 1 | | | | |
| 2 | | | | |
| 3 | | | | |
| 4 | | | | |
| 5 | | | | |
| 6 | | | | |
| 7 | | | | |
| 8 | | | | |
| 9 | | | | |
| 10 | | | | |
| 11 | | | | |

**Figura 2.** Baza de date în InStat. Pe orizontală avem seriile de studenţi (variabila nominală), iar în tabel vor fi trecute notele fiecăruia dintre studenţi (variabila numerică, ale cărei medii sunt subiectul comparaţiei). Nu are importanţă ordinea în care le scriem pe verticală.

*Untitled1 [DataSet0] - SPSS Data

File Edit View Data Transforr

1 : nota 26,0

| | nota | seria |
|---|---|---|
| 1 | 26 | 3 |
| 2 | 15 | 8 |
| 3 | 8 | 3 |
| 4 | 23 | 3 |
| 5 | 23 | 8 |
| 6 | 25 | 8 |
| 7 | 19 | 8 |
| 8 | 28 | 11 |
| 9 | 11 | 11 |
| 10 | 9 | 12 |
| 11 | 7 | 12 |
| 12 | 23 | 3 |
| 13 | 14 | 8 |
| 14 | 18 | 12 |
| 15 | 22 | 11 |
| 16 | 13 | 8 |

**Figura 3.** Baza de date în SPSS. În acest program, pe orizontală avem variabilele, în cazul nostru 2: nota (variabila numerică) şi seria (în coloana a doua, care poate lua doar patru valori, fiind vorba de patru serii de studenţi de anul IV).

În ANOVA, testăm ipoteza nulă că între seriile de studenţi nu există o diferenţă în privinţa pregătirii la această materie; ipoteza alternativă este aceea că există, totuşi, diferenţe între serii, şi la o privire sumară a graficului din figura 1, se pare că seria 11 are note ceva mai bune, iar seria 12 are note mai slabe. Rămâne să vedem dacă diferenţele aparente sunt semnificative statistic.

Dacă testul Student (*t*) compara mediile şi varianţele a două grupuri, testul ANOVA compară media şi varianţa totale a studenţilor celor 4 serii puşi la un loc, cu mediile şi varianţele fiecărei serii.

Dacă p rezultat în urma testului (care se mai numeşte şi testul F) este semnificativ statistic, înseamnă că seriile de studenţi nu sunt omogene, şi între ele există diferenţe, fără să ne spună unde se află aceste diferenţe (chiar dacă noi intuim, pe baza graficului din Figura 1). Pentru a vedea unde sunt exact aceste diferenţe şi dacă sunt semnificative statistic, trebuie să facem comparaţii utilizând testul *t* între seriile de studenţi, luate două câte două, deci vom face un număr de „combinări de 4 serii luate câte 2” comparaţii, adică 6. Problema care apare aici este că noi vom face cam multe comparaţii (problema comparaţiilor multiple), şi se ştie că, cu cât faci mai multe comparaţii, cu atât este un risc mai mare de a obţine valori semnificative

statistic numai din întâmplare (ca să obţii un p=0,05, la fiecare 100 de comparaţii, se obţin în medie 5 rezultate semnificative din întâmplare, iar la fiecare 20 de comparaţii, se obţine un rezultat semnificativ din întâmplare, fără să existe o diferenţă reală; aşadar, la 6 comparaţii, câte vom face noi, este foarte probabil să obţinem o diferenţă între două serii fals pozitivă). Pentru a evita ca rezultatele noastre să fie fals pozitive, trebuie să fim mai severi în privinţa p-ului; **corecţia Bonferroni** setează pragul de semnificaţie la p = 0,05/nr. de comparaţii ≈ 0,01. Stabilind acest nou prag de semnificaţie statistică şi comparând seriile două câte două, obţinem p<0,01 (adică diferenţe semnificative statistic) numai între seria 11 şi fiecare dintre celelalte serii, deci putem spune numai că seria 11 a fost semnificativ mai bună decât celelalte, nu şi că seria 12 a fost semnificativ mai slabă, cum am fi putut bănui din figură. Seria a 11-a a dat testul ultima, iar observaţia noastră empirică (înainte de a aplica orice test statistic) a fost aceea că studenţii ei au aflat, într-un mod oarecare, întrebările din test înainte de examen!

NB. Testul *t* (Student) nu este decât un test F, particularizat pentru numai două grupuri.