

Syllabus

Modele Ierarhice Liniare*

Iunie 2014

Instructor: Constantin Manuel Bosâncianu
Afiere: Universitatea Centrală Europeană, Budapesta
Email: Bosancianu_Constantin@ceu-budapest.edu
Website: <https://manuelbosancianu.github.io>¹

Introducere

Scopul modulului este de a oferi participanților un set de cunoștințe de bază despre o categorie de modele statistice foarte frecvent folosite în ultimii 10-15 ani în științele sociale. Modelele ierarhice liniare sunt aplicabile situațiilor în care eșantionul este format din observații (alegători, studenți, companii, partide etc.) grupate în unități superioare din punct de vedere ierarhic (țări, școli, regiuni etc). În aceste situații, cercetătorul poate fi interesat de a identifica efectul asupra fenomenului de interes a variabilelor măsurate atât la nivelul observațiilor cât și la nivelul unităților. Modelele ierarhice liniare (cunoscute și ca modele multinivel, modele cu efecte mixte, sau modele cu coeficienți aleatorii) permit testarea teoriilor care includ factori mășurați la mai multe niveluri de analiză (indivizi, regiuni, țări), oferind astfel o imagine mult mai complexă privind realitatea socială și legăturile dintre atitudini și comportamente individuale și factori contextuali. Modelele au o aplicabilitate foarte largă în mediul academic și non-academic, fiind folosite pentru analize a factorilor ce determină intenția de vot, a efectelor politicilor publice din domeniul educațional, a factorilor ce influențează tratamentul eficient al pacienților în sistemul medical, sau în cadrul studiilor epidemiologice sau biostatistice.

Cursul presupune faptul că participanții au cunoștințe solide de regresie OLS și logistică, precum și informație de bază privind interacțiuni în cadrul regresiei și tehnici de estimare. Dat fiind timpul limitat de care dispunem, nu vom putea face o recapitulare a acestor noțiuni de bază. Dacă aceste subiecte nu sunt încă stăpânite, vă rog să-mi trimiteți un email și vă pot îndruma către câteva texte scurte care tratează aceste subiecte. Este foarte important ca aceste noțiuni de bază să fie clare pentru a putea trece la un subiect statistic mai avansat, cum este cel al modelelor ierarhice.

Bibliografia oferită mai jos reprezintă minimul de cunoștințe teoretice necesare pentru a începe folosirea acestor modele în cadrul proiectelor individuale ale participanților. Textele obligatorii sunt la un nivel introductiv pentru fiecare dintre subiectele săptămânale, iar cele oferite pe lista opțională oferă o perspectivă mai amănunțită asupra acestor subiecte. Sfatul oferit participanților este de a parcurge textul obligatoriu din fiecare săptămână, precum și măcar textele suplimentare de la o săptămână de al cărei subiect sunt interesați. Timpul alocat nu permite învățarea și a unei aplicații statistice care poate estima aceste modele, deși sunt încrezător că pentru aplicațiile de tipul SPSS sau Stata, interfața

*Modul pregătit pentru Școala de Vară de Metode Aplicate, ediția 2014 (2 Iunie - 26 Iulie), Facultatea de Științe Politice, Administrative și ale Comunicării, Universitatea Babeș-Bolyai, Cluj Napoca, România.

¹Am încercat să colectez în secțiunea *Links* o serie de materiale (cărți, articole, aplicații statistice) pentru estimarea modelelor ierarhice liniare, în mod special în R.

point-and-click ar permite învățarea comenzilor în maximum 1-2 zile. Alte programe, care se bazează exclusiv pe sintaxă (R, Mplus, sau HLM) necesită o investiție mai mare de timp pentru începători.

Bibliografie

Săptămâna 1 (2–8 Iunie): Introducere în modele ierarhice liniare

Capitolele selectate pentru această săptămână oferă o perspectivă generală asupra situațiilor în care este necesară folosirea modelelor ierarhice.

Obligatoriu

- Luke, Douglas A. (2004). *Multilevel Modeling*. Thousand Oaks, CA: SAGE Publications. Capitolul I (pp. 1–9). URL: <http://www.u.arizona.edu/~mishler/MultilevelModelingBook.pdf>.
- Bickel, Robert. (2007). *Multilevel Analysis for Applied Research: It's Just Regression!* New York: Guilford Press. Capitolul I (pp. 1–16).
- Snijders, Tom A. B., și Roel J. Bosker. (1999). *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling*. Thousand Oaks, CA: SAGE Publications. Capitolul II (pp. 6–12).

Opțional

- Goldstein, Harvey. (2011). *Multilevel Statistical Models (4th edition)*. London: Wiley. Capitolul I (pp. 1–14).
- Raudenbush, Stephen W., și Anthony S. Bryk. (2002). *Hierarchical Linear Models: Applications and Data Analysis Methods (2nd edition)*. Thousand Oaks, CA: SAGE Publications. Capitolul I (pp. 3–15).

Săptămâna 2 (9–15 Iunie): Modele cu 2 niveluri

Săptămâna aceasta va fi alocată familiarizării cu notația folosită în analizele care folosesc modele ierarhice. Tratatamentul oferit de Raudenbush și Bryk (2002) este puțin mai tehnic, însă reprezintă un sistem de notație foarte des folosit în literatură.

Obligatoriu

- Raudenbush, Stephen W., și Anthony S. Bryk. (2002). *Hierarchical Linear Models: Applications and Data Analysis Methods (2nd edition)*. Thousand Oaks, CA: SAGE Publications. Capitolul II (pp. 16–37).

Opțional

- Bickel, Robert. (2007). *Multilevel Analysis for Applied Research: It's Just Regression!* New York: Guilford Press. Capitolul V (pp. 150–182).

- Steenbergen, Marco R., și Jones, Bradford S. (2002). Modeling Multilevel Data Structures. *American Journal of Political Science*, **46**(1), 218–237. URL: http://www.kenbenoit.net/courses/mlm/SteenbergenJones_AJPS_2002.pdf.

Săptămâna 3 (16–22 Iunie): Aplicații practice ale modelelor

Vom face o pauză de la prezentările matematice pentru a vedea câteva aplicații practice ale modelelor ierarhice. Participanții ar trebui să aibă în acest moment cunoștințele de bază necesare pentru a înțelege 70–80% din procedurile folosite în aceste analize.

Obligatoriu

- Wells, Jason M., și Kriekhaus, Jonathan. (2006). Does National Context Influence Democratic Satisfaction? A Multi-Level Analysis. *Political Research Quarterly*, **59**(4), 569–578.
- Cho, Wendy K. Tam, James G. Gimpel, și Joshua J. Dyck. (2006). Residential Concentration, Political Socialization, and Voter Turnout. *The Journal of Politics*, **68**(1), 156–167. URL: http://www.polsci.buffalo.edu/contrib/faculty_staff/documents/dyck/dyckResConcentration.pdf.

Opțional

- Solt, Frederick. (2008). Economic Inequality and Democratic Political Engagement. *American Journal of Political Science*, **52**(1), 48–60. URL: http://www.kean.edu/~jkeil/Welcome_files/Inequality_and_Political_Engagement.pdf.
- Pittau, M. Grazia, Roberto Zelli, și Andrew Gelman. (2010). Economic Disparities and Life Satisfaction in European Regions. *Social Indicators Research*, **96**(2), 339–361. URL: <http://stat.columbia.edu/~gelman/research/published/pittauzelligelman.pdf>.

Săptămâna 4 (23–29 Iunie): Eșantioane, centrare, ICC

Săptămâna aceasta ne vom aborda câteva întrebări esențiale pentru modelele ierarhice: De câte cazuri avem nevoie pentru a aplica aceste modele? Cum trebuie să fie pregătite datele pentru a estima aceste modele? Cum hotărâm dacă este justificat să folosim un model ierarhic în loc de o regresie OLS, sau o altă metodă statistică? Deși uneori ar putea fi considerate aspecte de importanță secundară, ele sunt totuși frecvent neînțelese în analizele practice.

Obligatoriu

- Bickel, Robert. (2007). *Multilevel Analysis for Applied Research: It's Just Regression!* New York: Guilford Press. Capitolul X (pp. 266–284).
- Enders, Craig K., și Davood Tofghi. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods*, **12**(2), 121–138. URL: <http://www-bcf.usc.edu/~sping/PM-515/references/Enders%202007%20centering%20in%20multilevel%20analysis.pdf>.
- Snijders, Tom A. B., și Roel J. Bosker. (1999). *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling*. Thousand Oaks, CA: SAGE Publications. Capitolul III (pp. 16–22).

Opțional

- Stegmueller, Daniel. (2013). How Many Countries for Multilevel Modeling? A Comparison of Frequentist and Bayesian Approaches. *American Journal of Political Science*, **57**(3), 748–761. URL: http://www.daniel-stegmueller.com/files/Stegmueller_AJPS2013.pdf.
- Paccagnella, Omar. (2006). Centering or not centering in multilevel models? The role of the group mean and the assessment of group effects. *Evaluation Review*, **30**(1), 66–85. URL: ftp://131.252.97.79/Transfer/ES_Pubs/ESVal/multilevel_analysis/centering/Paccagnella_06_multilevel_centering.pdf.
- Kreft, Ita G. G. (1995). The Effects of Centering in Multilevel Analysis: Is the Public School the Loser or the Winner? Lucrare prezentată la Întâlnirea Anuală a *American Educational Research Association*. San Francisco, CA: 18–22 Aprilie. URL: <http://files.eric.ed.gov/fulltext/ED392837.pdf>.

Săptămâna 5 (30 Iunie – 6 Iulie): Estimarea modelelor ierarhice liniare

Acesta va fi cel mai tehnic capitol din întregul modul, prezentând modalitatea prin care o mare parte a modelelor ierarhice sunt estimate. Recomand participanților să înceapă cu capitolul lui Eliason (1993), care oferă o prezentare generală a principiului *Verosimilității Maxime*, iar apoi să continue cu capitolul lui Hox (2010). Acesta din urmă se referă exclusiv la modele ierarhice, și include câteva metode suplimentare de estimare, prezentate pe scurt.

Obligativu

- Hox, Joop J. (2010). *Multilevel Analysis: Techniques and Applications (2nd edition)*. New York: Routledge. Capitolul III (pp. 40–53).
- Eliason, Scott R. (1993). *Maximum Likelihood Estimation: Logic and Practice*. Thousand Oaks, CA: SAGE Publications. Capitolul I (pp. 1–21).

Opțional

- Raudenbush, Stephen W., și Anthony S. Bryk. (2002). *Hierarchical Linear Models: Applications and Data Analysis Methods (2nd edition)*. Thousand Oaks, CA: SAGE Publications. Capitolul III (pp. 38–67).

Săptămâna 6 (7–13 Iulie): Indici de potrivire

Săptămâna aceasta vom discuta modalitatea prin care putem alege modelul cel mai potrivit datelor de care dispunem din perspectiva potrivirii cu datelor. Modele ierarhice folosesc o serie de indicatori (AIC, BIC, devianță, logaritmul probabilității) care diferă în interpretare de R^2 din regresia OLS, cu care participanții sunt mai familiarizați.

Obligativu

- Luke, Douglas A. (2004). *Multilevel Modeling*. Thousand Oaks, CA: SAGE Publications. Capitolul II (pp. 33–37).

- Hox, Joop J. (2010). *Multilevel Analysis: Techniques and Applications (2nd edition)*. New York: Routledge. Capitolul III (pp. 47–51).²
- Ripley, Brian D. (2004). Selecting Amongst Large Classes of Models. În Adams, Niall, Martin Crowder, David J. Hand și David Stephens (eds.), *Methods and Models in Statistics: In Honour of Professor John Nelder, FRS*. London: Imperial College Press, pp. 155–170. URL: <http://www.stats.ox.ac.uk/~ripley/ModelChoice/Nelder80paper.pdf>.

Opțional

- Burnham, Kenneth P., și David R. Anderson. (2004). Multimodel Inference: Understanding AIC and BIC in Model Selection. *Sociological Methods & Research*, **33**(2), 261–304. URL: http://faculty.washington.edu/skalski/classes/QERM597/papers_xtra/Burnham%20and%20Anderson.pdf.
- Kuha, Jouni. (2004). AIC and BIC: Comparisons of Assumptions and Performance. *Sociological Methods & Research*, **33**(2), 188–229. URL: <http://boss.web.elte.hu/grb/aicbic/188.pdf>.

Săptămâna 7 (14–20 Iulie): Modele cu 3 niveluri

În ultima săptămână din acest modul vom vedea cum putem extinde cadrul statistic învățat până acum la situația în care avem date pentru 3 niveluri de analiză (e.g. studenți, grupați în clase, la rândul lor grupate în școli). Deși notația modelelor va fi mai complexă, toate aspectele statistice discutate până acum în cadrul cursului pot fi extinse ușor la situația modelelor cu 3 niveluri.

Obligatoriu

- Bickel, Robert. (2007). *Multilevel Analysis for Applied Research: It's Just Regression!* New York: Guilford Press. Capitolul VIII (pp. 218–247).

Opțional

- Raudenbush, Stephen W., și Anthony S. Bryk. (2002). *Hierarchical Linear Models: Applications and Data Analysis Methods (2nd edition)*. Thousand Oaks, CA: SAGE Publications. Capitolul VIII (pp. 228–251).

Examinare

Dacă cerințele de lectură pentru modul par a fi prea înalte, sper ca această impresie să fie contrabalansată de formatul examinării. Modulul nu va necesita scrierea unei lucrări la final, sau un examen scris. Participanții sunt rugați să se folosească de articolul lui Wells și Kriekhaus (2006), alocat pentru săptămâna 3, și a scrie o critică de maximum o pagină (A4, corp de literă Georgia 11, spațiere la un rând, format Word sau PDF) la adresa articolului din perspectiva modelelor analizate. Identificați 2 aspecte pozitive și 2 aspecte negative ale analizelor făcute de autori. Nu este nevoie de o prezentare teoretică a literaturii, sau de citate din capitolele alocate pentru acest modul: sunt interesat de părerile proprii și argumentate ale participanților privind calitatea analizelor din articolul alocat. Termenul limită de trimitere a lucrării prin email este 20 Iulie, 23:59 (ora României).

²Capitolul a fost deja alocat pentru o săptămână precedentă; doar aceste 5 pagini sunt necesare pentru această săptămână.

PS: Mult succes cu lecturile, și sper să ne vedem în Cluj Napoca!

PPS: Dacă pot fi de ajutor cu informații privind lecturile alocate sau modele ierarhice până când ne vom vedea în Cluj Napoca, nu ezitați să-mi trimiteți un email!