



UNIVERSITATEA BABEȘ-BOLYAI
TRADIȚIE ȘI EXCELENȚĂ



UBBFSEGA
Universitatea Babeș-Bolyai | Facultatea de Științe Economice și Gestionarea Afacerilor



Str. Teodor Mihali nr. 58-60
Cluj-Napoca, RO-400591
Tel.: 0264-41.86.52-5
Fax: 0264-41.25.70
econ@econ.ubbcluj.ro
www.econ.ubbcluj.ro

TEZA DE ABILITARE

Contribuții teoretice privind curba ROC și aplicații econometrice în asigurări și scientometrie

Cristian Mihai DRAGOȘ

Cluj Napoca
2019

REZUMAT

Această teză de abilitare sintetizează activitatea mea de cercetare după susținerea tezei de doctorat în anul 2004 la Universitatea Babeș-Bolyai din Cluj Napoca, în domeniul Cibernetică, Statistică și Informatică Economică.

Prima parte a tezei cuprinde contribuții teoretice și aplicații, grupate în trei secțiuni:

Secțiunea A: Contribuții teoretice și aplicații ale curbei ROC pentru modele multinomiale de decizie discretă; **Secțiunea B:** Modele multinomiale de decizie discretă aplicate în asigurări; **Secțiunea C:** Abordare econometrică a factorilor de influență a productivității științifice a țărilor.

Secțiunea A include contribuții teoretice privind curba ROC (Receiver Operating Characteristic) pentru modelele multinomiale de decizie discretă și aplicații ale acesteia în scoringul financiar și industria auto.

Prima parte a secțiunii A este intitulată *Aplicații ale curbei ROC binare în scoringul financiar*. Este realizată o trecere în revistă a metodelor statistice din scoringul financiar în scopul de a compara eficiența lor relativă. Sunt evaluate analiza discriminantă, econometria regresiiilor discrete (modele logit și probit), metodele neparametrice și algoritmi de inteligență artificială (rețele neuronale și algoritmi genetici). Un studiu experimental realizat prin previziuni urmărește selecția modelului cu cea mai bună acuratețe a previziunii. Rezultatele, evaluate prin curba ROC arată că creșterea numărului de variabile din modelul de scoring are un impact semnificativ asupra predictibilității modelului și asupra reducerii varianței erorii de previziune în eșantioanele de testare.

A doua parte a secțiunii A este intitulată *Definirea unei curbe ROC bidimensionale pentru modele categoriale de decizie discretă*. Studiul face o recenzie a aplicațiilor care utilizează modele categoriale de decizie discretă. Necesitatea unui instrument adecvat pentru evaluarea acurateții previzionale a unui model este pusă în evidență. Definim curba ROC pentru modele categoriale multinomiale, atât pentru o alternativă de alegere, cât și global, pentru întregul set de alternative. Senzitivitatea globală poate

fi exprimată ca media ponderată a sensibilităților celor M alternative. În eșantion, ponderile sunt frecvențele relative ale indivizilor care aleg alternativele respective. Specificitatea globală poate fi exprimată de asemenea ca media ponderată a specificităților individuale, ponderile fiind frecvențele relative din eșantion aferente indivizilor care nu aleg alternativele respective. O aplicație asupra pieței românești a automobilelor pune în evidență diferențele dintre două modele prin prisma curbei ROC nou definite.

A treia parte a secțiunii A este intitulată *Definirea unei curbe ROC bidimensionale pentru modele ordinale de decizie discretă*. Curba ROC este de mai mult de șapte decenii un instrument foarte popular de evaluare a acurateții previziunii pentru modelele binare de decizie discretă. Câțeva studii au generalizat acest instrument în spații multidimensionale pentru modele ordonate și categoriale. Pornind de la un set de principii, definim o curbă ROC bidimensională pentru modelele ordonate de decizie discretă: 2D-ROC-ODCM. Pentru a defini specificitatea, noul instrument ia în considerare distanța dintre fiecare alternativă și alternativa în care se află observația. Folosind simulări, evaluăm comportamentul 2D-ROC-ODCM în anumite situații: corelații de intensitate diferită între variabila endogenă și regresori, variația numărului de regresori și relația dintre varianța empirică a ariei de sub curbă și volumul eșantionului de testare.

Secțiunea B include aplicații ale modelelor categoriale de decizie discretă și curba ROC corespunzătoare în estimarea comportamentului uman în asigurările de viață și auto.

Prima parte a secțiunii B este intitulată *Estimarea comportamentului consumatorilor de asigurări auto facultative*. Aplicația modelează comportamentul persoanelor în raport cu cumpărarea sau nu a polițelor de asigurări facultative auto, folosind modele de decizie discretă, presupunând atât decizii de tip ierarhic, cât și neierarhic. Eșantionul este format din 311 proprietari de autoturisme din județul Cluj, România. Estimațiile econometrice utilizează modelul logit binar, logit multinomial și nested logit. Puterea predictivă a modelelor este evaluată prin curba ROC pentru modele multinomiale. Rezultatele arată că principalii factori care influențează cumpărarea unei polițe de asigurări auto facultative sunt preferința/aversiunea la risc, distanța anuală parcursă cu mașina, nivelul de educație al șoferului și raportul dintre venitul

conducătorului auto și prețul mașinii. Spre deosebire de studii anterioare care estimează profilul de risc al individului doar prin variabile proxy care nu țin cont de aspecte comportamentale, studiul nostru a integrat cu succes profilul de risc ca o variabilă specifică.

A doua parte a secțiunii B este intitulată *Estimarea deciziei probabiliste a alegerii dintre produsele de asigurări de viață utilizând modele de decizie discretă*. Studiul vizează piața de asigurări de viață din România care este în plină ascensiune. Competiția este în creștere, atât între companii, cât și între produsele aceleiași companii. Utilizând date colectate de la clienții unei companii, s-au luat în considerare două tipuri de variabile: atribute ale produselor de asigurare (ex.: riscul, profitabilitatea) și caracteristici ale indivizilor (ex.: genul, vârsta, venitul). Pe baza elementelor de teorie economică și a modelului logit multinomial se explică comportamentul asiguraților. Se estimează de asemenea variațiile cotelor de piață datorate schimbărilor posibile în valorile unor atribute sau caracteristici.

A treia parte a secțiunii B este intitulată *Estimarea econometrică a subscrierii pentru asigurările de viață sau pensii private*. Sunt testate ipoteze privind influența factorilor comportamentali și socio-demografici asupra subscrierii pentru polițe de asigurări de viață sau pensii private. Datele sunt obținute dintr-un eșantion de 1579 de indivizi, reprezentativ pentru consumatorii români. Chestionarul este construit pentru a observa patru categorii de determinanți: factori socio-demografici, factori comportamentali financiari generali și specifici și un indice propriu construit pentru a educația financiară. Prin regresii de tip logit sunt puse în evidență mecanismele diferite care disting între decizie și intenție privind cumpărarea asigurărilor de viață și a pensiilor private. Se arată că factorii comportamentali specifici și educația financiară sunt foarte semnificativi pentru decizie, dar nu și pentru intenția de a cumpăra a celor două produse financiare.

Secțiunea C include aplicații econometrice în scientometrie, care pun în evidență factorii determinanți ai productivității științifice a țărilor.

Prima parte a secțiunii C este intitulată *Abordare econometrică a factorilor care influențează productivitatea științifică a țărilor în științele mediului și ecologie*. Diferite studii academice au măsurat influența factorilor economici, politici și lingvistici asupra producției științifice globale a țărilor. Analize separate, pe domenii, pot releva factori determinați specifici. Studiul nostru demonstrează că Indicele Performanței de Mediu, calculat de universitatea din Yale este influențată foarte semnificativ productivitatea activității de cercetare în științele mediului și ecologie. Variabilele de control ca finanțarea educației, limba engleză sau publicarea în reviste indexate ISI Thomson din propria țară sunt de asemenea semnificative. Metodologia utilizează regresii multiple OLS, cu rezultate foarte convingătoare ($R^2 = 0.752$). Poziția relativă a celor 92 de țări din eșantion este de asemenea discutată. Este construit și un clasament al țărilor, luând în considerare atât productivitatea științifică, cât și calitatea mediului. Se remarcă diferențe uriașe între țările care domină clasamentul și cele care ocupă ultimele poziții.

A doua parte a secțiunii C este intitulată *Abordare econometrică a factorilor care influențează productivitatea științifică în domeniul economiei și afacerilor*. Studiul pune în evidență principalii determinanți ai productivității științifice în economie și afaceri, respectiv finanțarea educației și cercetării, mărimea populației, numărul de reviste proprii și engleza ca limbă oficială. Sunt utilizate regresii multiple OLS și date extrase din *Web of Knowledge* și *World Bank* pentru 56 de țări. Se subliniază și relația care există între producția științifică și eficiența utilizării fondurilor de cercetare. Clasamentele țărilor din eșantion arată că există procese de învățare la nivel național, producția fiind dublată de eficiență.

A treia parte a secțiunii C este intitulată *Productivitate științifică versus eficiența finanțării activităților de cercetare-dezvoltare analiză bibliometrică a țărilor africane*. Uzual, analizele bibliometrice la nivel de țară măsoară performanța științifică prin producția totală de articole și citări și prin productivitate, raportând la numărul de locuitori. Studiul nostru evaluează și eficiența, prin raportarea outputului științific la finanțarea disponibilă. Rezultatele arată că doar două țări africane (Africa de Sud și Tunisia) au parcurs un proces de învățare și au devenit entități mature în procesele de cercetare dezvoltare, productivitatea fiind acompaniată de eficiență.

A doua parte a tezei prezintă planul de dezvoltare a carierei, accentuând orientările în activitățile de cercetare. Principalele mele preocupări de cercetare în econometrie și statistică pentru economie și afaceri se referă la următoarele axe de expertiză: (1) metode de evaluare a acurateții previziunii pentru modelele multinomiale de decizie discretă; (2) metode statistice specifice aplicate în scientometrie; (3) aplicații ale regresiiilor discrete în economie și afaceri. Cercetările viitoare vor include în principal dezvoltări teoretice și aplicații pe aceste trei axe menționate. Voi insista asupra identificării unor metode de inferență statistică pentru modelele teoretice existente, asupra evoluțiilor econometrice ale metodelor statistice deja utilizate în scientometrie și asupra identificării unor noi variabile explicative (chiar provenind din psihologie sau alte laturi ale comportamentului uman) pentru unele fenomene economice deja studiate.

A treia parte a tezei include referințele bibliografice.