UNIVERSITATEA DIN BUCUREȘTI

FACULTATEA DE MATEMATICĂ ȘI INFORMATICĂ

ȘCOALA DOCTORALĂ DE MATEMATICĂ

CURBE LORENZ ȘI APROXIMARE DE ORDIN SUPERIOR PENTRU DIVERGENȚE GENERALIZATE

REZUMAT

Autor:

Drd. **Carmen-Adriana GHEORGHE (Dragnea)**

Coordonator științific:

Prof. univ. dr. **Vasile PREDA**

BUCUREȘTI 2013

Cuprins

INTRODUCERE

1. Motivația alegerii temei

Curba Lorenz este una dintre cele mai importante modalități de analiză a distribuției venitului. Sunt două posibilități de a defini o curbă Lorenz. Prima este să specificăm o funcție de repartiție și apoi să găsim curba Lorenz în funcție de această repartiție, iar a doua metodă este prezentarea directă a funcției curbei Lorenz ce satisface anumite condiții pentru a fi bine definită.

Teza de doctorat *"Curbe Lorenz și aproximare de ordin superior pentru divergențe generalizate"* are ca **obiectiv general** realizarea unui studiu complex privind curbele Lorenz parametrice și legătura acestora cu divergențele generalizate.

Cercetarea urmărește prezentarea conceptelor teoretice, identificarea metodelor de studiu empiric privind definirea și aproximarea a unor clase noi de Curbe Lorenz, cu **scopul** dezvoltării unor posibile funcții de repartiție care să modeleze întregul set de date.

Motivația alegerii temei de cercetare o constituie noutatea, actualitatea și importanța Curbelor Lorenz. Lucrarea clasică a lui Atkinson (1970) a făcut să crească interesul pentru ordinea statistică referitoare la compararea distribuțiilor venitului, cum ar fi ordinea Lorenz.

Tipuri de cercetare și abordări noi

Originea repartiției Lorenz este într-un scurt articol intitulat "Metode de măsurare a concentrației veniturilor". Acesta propunea în 1905 o metodă simplă, denumită ulterior repartiția Lorenz, pentru vizualizarea repartiției veniturilor sau bunăstării în raport cu inegalitatea sau concentrația veniturilor dobândite. Conform Derobert și Thieriot (2003), termenul "repartiție Lorenz" apare pentru prima dată în King (1912), o lucrare de statistică scrisă pentru economiști și cercetători de științe sociale.

Indicele lui Gini în funcție de Curba Lorenz

Cea mai cunoscută modalitate de a măsura inegalitatea utilizând Curba Lorenz este **indicele Gini**. Acesta a fost introdus în articolul „Variability and Mutability" (1912) de către Corrado Gini (1884-1965), un ideolog italian, dar totodată un renumit sociolog și demograf.

Activitatea de debut a lui Gini a fost asociată cu problema modului de evaluare a inegalităților dintre venituri și avere în diferite țări. Gini a fost cercetător în cadrul Institutului Național de Statistică

Internațională și profesor la catedra de statistică a Universității din Padova. A predat, printre altele, demografie și statistică economică. Ulterior, ca profesor al Universității din Roma (după 1925) a făcut multe inovații semnificative în statistică. Cu toate aceste realizări, rămâne cunoscut posterității prin indicele ce-i poartă numele.

Indicele lui Gini pentru populația României în raport cu Europa

În acest paragraf am dorit evidențierea poziției României în Europa din perspectiva inegalității venitului, măsurată prin intermediul indicelui lui Gini.

2005

2011

Am reprezentat indicele lui Gini pentru anii 2005, respectiv 2011, astfel: prin galben am evidențiat țările cu inegalitatea veniturilor mai mică decât în România, iar prin albastru statele cu indicele lui Gini mai mare.

Inegalitatea a crescut în Bulgaria, Franța, Germania, dar în Cehia, Elveția veniturile au devenit mai puțin inegale. Pentru Figura 1.4.3 am selectat țările cu evoluția crescătoare a inegalității, exact ca și în România. Maximul este atins de Bulgaria, România situându-se printre țările cu o creștere moderată.

Cronologia evoluției măsurilor de inegalitate

În continuarea prezentării măsurilor inegalității am ilustrat o succintă cronologie a acestora în figura de mai jos. Observăm că abia după anii 1970 interesul asupra Curbelor Lorenz a crescut substanțial, fiind declanșat de lucrările lui Atkinson (1970) și Gastwirth (1971). Atkinson a prezentat importanța comparării Curbelor Lorenz în contextul bunăstării economice, iar Gastwirth a definit pentru prima dată Curba Lorenz prin intermediul unei funcții de repartiție.

Cercetările recente au avut ca scop dorința de a caracteriza ordinea Lorenz cu familii parametrice de distribuții ale veniturilor – Wang (2013), Krause (2013), Cowell (2007). Pentru modelele cu un parametru și cu doi parametri aceasta este directă, și este cunoscută liniaritatea în cazul familiilor Pareto și log-normale - Arnold (1987a). Pentru modelele cu trei și patru parametri, ordinea Lorenz corespunzătoare distribuțiilor Singh-Maddala și Dagum a fost redată de Wilfling și Kramer (1993) și Kleiber (1999). Ambele distribuții sunt cazuri speciale ale repartiției Generalizate Beta de ordinul al doilea (GB2), prezentată de McDonald (1984).

Familia de repartiții Weibull Exponențializată a fost introdusă de Mudholkar și Sarisastva (1993) ca fiind o generalizare a familiei de repartiții Weibull. Aplicațiile acestei repartiții au fost ilustrate de Mudholkar (1995).

Pronind de la dualitatea între funcțiile de repartiție și Curbele Lorenz am utilizat trei tipuri de cercetare: *generalizarea, extinderea* și *abordarea nouă*.

Cercetările pornind de la:

- o repartiție utilizând *generalizarea* repartițiilor exponențializate au condus la determinarea de noi curbe Lorenz;

- o curbă Lorenz au avut ca rezultate *extinderi* prin generări de curbe lorenz utilizând mecanismul Marshall-Olkin, dar și un alt mecanism: prin compunerea cu repartiția Gamma am introdus Curbe Lorenz mixte.

Revenind la măsuri ale inegalității, am abordat *noi cercetări* privind divergențele generalizate și aproximarea lor prin Curbe Lorenz.

2. Direcții de cercetare

Avantajele alegerii Curbelor Lorenz sunt: simplitatea reprezentării grafice și claritatea din punct de vedere vizual (câteva linii explică construcția graficului și încă de la prima vedere redă toate informațiile pe care le conține). Metoda permite o estimare numerică a concentrației bunăstării prin măsurarea ariei definită de Curba Lorenz și prima bisectoare. Mai mult, aceasta admite comparații între repartiții de venit diferite, cu excepția cazurilor în care curbele se intesectează. Aici intervin *limitările* și *deficiențele* care au condus la o extensie a modelelor existente în literatură. In primul rând întâlnim nepotrivirea în totalitate a curbelor empirice pe setul de date. Pentru o mai bună estimare a setului de date ne propunem construirea unor modele mai complexe ce combină o curbă parametrică Lorenz cu o densitate de repartiție cunoscută. Mai mult, utilizarea aproximării parametrice mixte aduce o concordanță mai mare prin introducerea de restricții stricte.

Ipoteza generală urmărită spre verificare a fost preluată în urma documentării preliminare din numeroasele studii realizate pe această temă de-a lungul timpului și anume posibilitatea îmbunătățirii familiilor de Curbe Lorenz prin crearea de noi conexiuni cu metode ce pot genera CL performante.

În realizarea scopului propus, este necesară rezolvarea următoarelor **obiective principale**:

- evidențierea importanței curbei Lorenz și a indicelui lui Gini în contextul dependenței acestuia de CL;
- dezvoltarea de noi clase de Curbe Lorenz, prin extinderea familiilor consacrate de curbe Lorenz;
- tratarea complexă a modelor de Curbe Lorenz mixte pentru familii parametrice de CL prin compunerea acestora cu repartiția Gamma;
- studiul intervalelor de încredere ale curbei Lorenz Generalizată;
- abordarea sistematică a intervalelor de încredere ale Curbelor Lorenz Generalizate pentru familia de repartiții Weibull Exponențializate;
- dezvoltarea de familii parametrice noi pentru Curbe Leimkuhler, precum și Curbe Leimkuhler mixte;
- elaborarea proprietăților asimptotice ale estimatorilor non-parametrici ai curbei Leimkuhler si Indicelui Kakwani;
- generalizarea familiilor parametrice noi de Curbe Leimkuhler în scopul obținerii unor Curbe Leimkuhler mixte;
- propunerea unor aproximări de ordin superior pentru divergențe generalizate.

Scopul și problemele cercetate au determinat structura tezei, care cuprinde, pe lângă introducere, alte șapte capitole, concluzii și propuneri, bibliografie și anexe.

3. Structura lucrării

Lucrarea este organizată în șapte capitole. După secțiunea introductivă ce își concentrează atenția asupra prezentării câtorva concepte de bază, în **primul capitol** definim curba Lorenz, indicele lui Gini în funcție de curba Lorenz și modele de Curbe Lorenz parametrice. Pentru exemplificare analizăm indicele lui Gini pentru populația României în raport cu Europa și Curbele Lorenz corespunzătoare veniturilor totale ale gospodăriilor românești în ultimele trei decenii. Acest capitol se bazează pe articolele [1], [2], [3], [14] și [15].

În **cel de-al doilea capitol** dezvoltăm noi clase de Curbe Lorenz, prin extinderea familiilor consacrate de curbe Lorenz, având ca bază lucrările [4], [10] și [11]. Propunem folosirea entropiei maximă Tsallis pentru Curba Lorenz. Prezentăm mecanismul Marshall-Olkin pentru a genera familii Lorenz și Leimkuhler de tip Marshall-Olkin.

Introducem în **capitolul al treilea** modele de Curbe Lorenz mixte pentru familii parametrice de CL prin compunerea acestora cu repartiția Gamma, rezultate publicate în [7].

Construim în **capitolul al patrulea** (conform [6]) intervale de încredere ale Curbelor Lorenz Generalizate pentru familia de repartiții Weibull Exponențializate și rata de hazard corespunzătoare. Dezvoltăm estimatorul de verosimilitate maximă pentru această familie și studiem intervalele de încredere ale curbei Lorenz Generalizată prin simulare.

Al cincilea capitol conține proprietățile asimptotice ale estimatorilor non-parametrici ai curbei Leimkuhler si Indicelui Kakwani prezentate în [9].

În strânsă legătură cu secțiunea anterioară **capitolul al șaselea** dezvoltă familii parametrice noi de Curbe Leimkuhler, precum și Curbe Leimkuhler mixte, având la bază articolul [8].

În **ultimul capitol** am prezentat bazându-ne pe [5] și [13] aproximări de ordin superior pentru divergențe generalizate. După realizarea unei sinteze asupra teoremelor Arnold și Thompson am aplicat formula lui Taylor în cazul AB-divergenței în vederea obținerii unor rezultate de aproximare pentru divergențe generalizate folosind curbe Lorenz.

Structura matematică a rezultatelor

Pe lângă împărțirea firescă a rezultatelor în capitole am urmărit și structurarea din punct de vedere matematic a rezultatelor în patru direcții principale:

a. Prezentarea unei forme analitică nouă a funcției din care se obține Curba Lorenz și validarea condițiilor de existență din definiție;

b. Obținerea estimatorului de verosimilitate maximă ai parametrilor unora dintre modelele parametrice prezentate și utilizarea aproximării normale pentru construirea de intervale de încredere;

c. Estimări neparametrice ale unora dintre modelele studiate;

d. Introducerea unei noi divergențe pentru două măsuri de probabilitate, stabilirea unor mărginiri superioare ale acesteia.

4. Contribuții autor

Rezultatele cercetărilor efectuate în cadrul tezei de doctorat au fost valorificate prin publicarea a cincisprezece articole științifice, dintre care

- *șapte* au fost prezentate cu ocazia unor conferințe internaționale, dintre care:

  (2012) *The estimate of generalized Lorenz confidence intervals using Exponentiated Weibull Distribution*, Conferința 15 SPSR, Universitatea din București (studiu cuprins în **capitolul 4**).

  (2013) *New parametric families of Leimkuhler Curves*, Conferința 16 SPSR, Bucuresti, (studiu cuprins în **capitolul 6**).

- *șase* sunt publicate în reviste de matematică recenzate:

  (2007) *Mortality modeling for Romanian population*. Proceedings of The 4-th International Colloquium "Mathematics in Engineering and Numerical Physics" October 6-8 , 2006, Bucharest, Romania, pp. 38-45. Balkan Society of Geometers, Geometry Balkan Press 2007, (studiu cuprins parțial în **capitolul 1**).

  (2007) *The Hill estimator for income of Romanian households*, Proceedings of International Conference Trends and Challenges in Applied Mathematics, ISBN 978-973-755-283-9/pbk, Ed. Matrix Rom, Bucuresti, p. 193- 196, (studiu cuprins parțial în **capitolul 1**).

  (2007) *Life tables for Romania. Survival function. Life expectancy for Romanian people,* U.P.B, Scientific Bulletin, Series A, vol. 69, no. 1, Bucuresti, (studiu cuprins parțial în **capitolul 1**).

(2009) *Estimation of weighted maximum entropy densities within the study upon Lorenz Curves for grouped data*, Proceedings of 16th European Young Statisticians Meetings, ASE, Bucuresti. (studiu cuprins parțial în **capitolul 2**)
(2009) *Approximation of Csiszar's f-Divergence using Kullback-Leibler Distance*, Proceedings of 9th Balkan Conference on Operational Reserch, Constanta, (studiu cuprins parțial în **capitolul 7**).
(2013) *Mixture Lorenz Curves. Three new models*, Proceedings of 18th European Young Statisticians Meetings, Osjek, Croatia, (studiu cuprins în **capitolul 3**)

- *șapte* se află în proces de publicare:

(2013) *Asymptotic properties of the nonparametric estimators of the Leimkuhler curve and Kakwani index*, U.P.B, Scientific Bulletin, Bucuresti, cotat ISI, factor impact *0.30* în 2012, (studiu cuprins în **capitolul 5**).
(2013) *Maximum Tsallis entropy and the Lorenz curve*(studiu cuprins în **capitolul 2**)
(2013) *A class of Lorenz Curves generated by the Marshall Olkin technique*, (studiu cuprins în **capitolul 2**).
(2013) *Some approximations for generalised divergences*, (studiu cuprins în **capitolul 7**).
(2013) *Bounds for some generalised divergences in trems of Lorenz curves*, (studiu cuprins în **capitolul 7**).
(2013) *Alternative measurements of the income inequality among the households* (studiu cuprins patrțial în **capitolul 1**)
(2013) *Analyzing the Romanian income inequality in the European context* (studiu cuprins patrțial în **capitolul 1**)

I
CURBE LORENZ

Noțiuni introductive. Definirea Curbei Lorenz

În 1905 a apărut în "Publicația Asociației de Statistică Americană" un scurt articol intitulat "Metode de măsurare a concentrației veniturilor". Acesta propunea o metodă simplă, denumită ulterior curba Lorenz, pentru vizualizarea distribuției veniturilor sau bunastării în raport cu inegalitatea sau concentrația veniturilor dobândite. Autorul Max Otto Lorenz și-a completat lucrarea de doctorat în 1906 la Universitatea din Wiscounsin, fără a menționa acest articol care dealtfel a fost și singura publicație a lui Lorenz într-o revistă științifică.

Curba Lorenz este un instrument statistic folosit pentru analiza distribuției veniturilor. Fie *f(t)* densitatea de repartiție, unde *t* reprezintă venitul și verifică o inegalitate de forma $\alpha \le t \le \beta$. Graficul repartiției Lorenz corespunzător acestei distribuții a veniturilor se obține prin trasarea punctelor $(x, y) = (x(t),\ y(t))$, unde $x = x(t)$ este procentul de populație cu venituri mai mici sau egale decât *t* și $y = y(t)$ este procentul veniturilor totale ale acestei populații.

Fie $X_1, X_2, ..., X_n$ un eșantion independent identic distribuit dintr-o funcție de repartiție F_X absolut continuă cu $F_X(0) = 0$.

Rezultă ecuațiile

$$x = x(t) = \int_{\alpha}^{t} f(s)ds \ ; \ y = y(t) = \frac{1}{m}\int_{\alpha}^{t} sf(s)ds$$

unde $m = \int_{\alpha}^{\beta} tf(t)dt$ este media lui X.

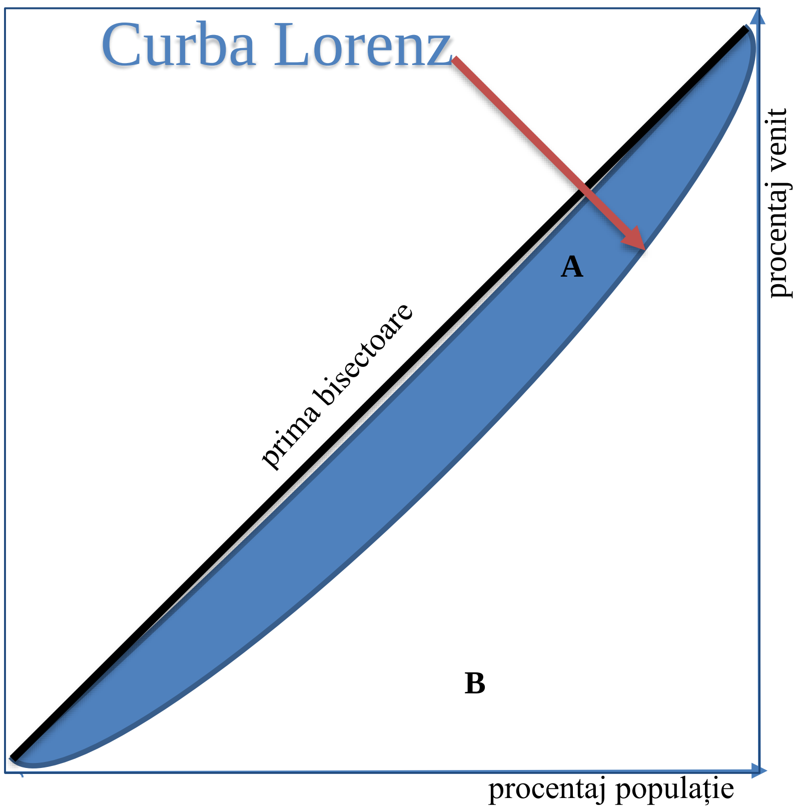

Figura 1.1. Curba Lorenz

Relațiile (1.1) sunt reprezentarea parametrică a repartiției Lorenz. Pentru a-l elimina pe ***t***, notăm *u=F(s)*, unde *F(t)* este funcția de repartiție a lui *f(t)*. Rezultă că *du=dF(s)=f(s)ds* și $s = F^{-1}(u)$.

Vom avea

$$x=F(t)\,;$$

$$y = \frac{1}{m}\int_{F(\alpha)}^{F(t)} F^{-1}(u)du = \frac{1}{m}\int_{0}^{x} F^{-1}(u)du =: l(x)$$

Notăm coada superioară a lui F_X cu

$$\bar{F}_X(x) = 1 - F_X(x),$$

iar statisticile de ordine prin

$$X_{(1)} \le \ldots \le X_{(n)}.$$

Presupunem că există media venitului și o notăm prin

$$\mu_X = \int_0^\infty yF^{-1}(y)dy.$$

Curba Lorenz este dată de

$$\{(p, L_X(p)), 0 \le p \le 1\}, \text{ cu}$$

$$L_X(p) = \int_{x\ge 0} I(x \le F_X^{-1}(p))xdF_X(x),$$

unde $\mathrm{I}(*)$ este funcție indicator.

Fie L clasa tuturor variabilelor aleatoare nenegative cu media finită și fie X din L, cu densitatea de repartiție $f(x)$. Atunci funcția de repartiție $F(x) = E(\mathrm{I}_{(0,x)})$ va fi procentul de unități ce au un venit mai mic sau egal cu x. Valorile pe care le poate lua $F(x)$ sunt între 0 și 1. Presupunem că există media veniturilor, aceasta fiind dată de $\mu = E(\mathrm{I}_{(0,\infty)})$. Atunci momentul de ordinul întâi al lui X va fi $F_1(x) = \mu^{-1}E(X \cdot I_{(0,x)})$ și reprezintă cota veniturilor din totalul câștigat de un individ ce are venitul mai mic sau egal cu x. Graficul trasat în pătratul-unitate ce conține $F(x)$ pe axa absciselor și $F_1(x)$ pe axa ordonatelor reprezintă Curba Lorenz, unde x are valori de la 0 la ∞.

Definiția 1.1. - ***Gastwirth (1971)*** - Fie $X \in L$ cu funcția de repartiție $F(\cdot)$ și inversa sa $F^{-1}(y) = \inf\{x: F(x) \geq y\}$. Curba Lorenz $L(\cdot)$ este definită prin

$$L(p) = \mu^{-1} \int_0^p F^{-1}(y)dy, \text{ pentru } 0 \leq p \leq 1.$$

De fapt, Curba Lorenz este corelația dintre procentul de populație și procentul de venituri care revin acesteia.

Definiția 1.2. - ***Kakwani (1980)*** O funcție $L_X(p)$, continuă pe $[0,1]$ este o curbă Lorenz dacă satisface condițiile următoare:

i) $L(p) = 0$, dacă $p = 0$;
ii) $L(p) = 1$, dacă $p = 1$;
iii) $L'(0^+) \geq 0$;
iv) $L''(p) \geq 0$, pentru orice $0 \leq p \leq 1$.

Shorrocks (1983) a construit **Curba Lorenz generalizată**:

$$GL_X(p) = \mu_X L_X(p).$$

Fie X și Y două variabile aleatoare cu funcțiile de repartiție F și G. Curba Lorenz Generalizată poate fi utilizată pentru a defini o relație de ordine parțială pe clasa de funcții de repartiție după cum urmează:

$$F \leq_{gl} G \Leftrightarrow GL_X(p) \leq GL_Y(p), \forall 0 \leq p \leq 1$$

În acest caz citim că Y este mai mic sau egal cu X în sensul Shorrocks.

Proprietăți ale Curbelor Lorenz

Kakwani (1981) a demonstrat că o Curbă Lorenz are următoarele proprietăți:

1. Este încadrată într-un pătrat $p \in [0,1]$, $L(p) \in [0,1]$.

2. Curba $L(p)$ nu este definită pentru $\mu_X = 0$ sau $\mu_X = \infty$.

3. Dacă variabila este pozitivă și are densitatea de repartiție continuă, atunci Curba Lorenz este funcție continuă și se află sub prima bisectoare sau este egală cu aceasta.

4. $L(p)$este o funcție crescătoare și convexă ($L'(0) > 0$, $L''(p) > 0$ pentru orice $0 \leq p \leq 1$).

5. Curba Lorenz este invariantă cu factor pozitiv de scalare: *X* și *cX* au aceeași Curbă Lorenz.

6. Media populației este obținută din cuantila $F(\mu)$ pentru care $L(p) = 1$.

Indicele lui Gini în funcție de Curba Lorenz

Cea mai *importantă* modalitate de a măsura inegalitatea utilizând Curba Lorenz este **indicele Gini**. Acesta a fost introdus în 1912 de către Corrado Gini.

Indicele lui Gini este reprezentat în *Figura 1.1.* ca fracția ce are la numărător (A) aria suprafeței situată între linia de egalitate și curba Lorenz, iar la numitor (A+B) toată suprafața de sub prima bisectoare:

$$G = \frac{A}{A+B}.$$

O valoare mică a indicelui lui Gini indică o repartiție mai uniformă a veniturilor. Deși în practică nu sunt niciodată atinse, valoarea *0* corespunde egalității între venituri, iar *1* inegalității totale a veniturilor.

Altă definiție a indicelui lui Gini (în funcție de Curba Lorenz) este

$$G = 1 - 2\int_0^1 L(p)dp.$$

O generalizare importantă a indicelui lui Gini a fost propusă de Yitzhaki (1983):

$$G_\nu = 1 - \nu(\nu - 1)\int_0^1 (1-p)^{\nu-2} L_X(p)dp; \nu > 1$$

Observăm că pentru $\nu = 2$ obținem indicele lui Gini.

II
NOI CLASE DE CURBE LORENZ

Familii de curbe Lorenz

Distribuția veniturilor constă în compararea veniturilor indivizilor dintr-un grup sau societate, și reprezintă unul dintre aspectele economiei și structurii sociale. Studiul statistic al distribuției veniturilor este util pentru a oferi informație, nu constă în recomandări de distribuție (în acest scop se utilizează “teoria redistribuirii veniturilor”).

Există multiple modalități prin care poate fi analizată distribuția veniturilor. Un exemplu ar fi compararea veniturilor celor mai bogați (10%) cu veniturile celor mai săraci 10%. În majoritatea societăților, cei mai bogați 10% controlează mai mult de jumătate din venitul total al societății.

Curba Lorenz este cea mai utilizată metodă de caracterizare a distribuției veniturilor.

Curbe Lorenz pentru repartiții clasice

Fie $X_1, X_2, ..., X_n$ un eșantion independent dintr-o funcție de repartiție F_X absolut continuă, cu $F_X(0) = 0$. Notăm coada superioară a lui F_X cu $\bar{F}_X(x) = 1 - F_X(x)$, și statisticile de ordine cu $X_{(1)} \leq \ldots \leq X_{(n)}$. Presupunem că există media venitului și o notăm cu

$$\mu_X = \int_0^\infty yF^{-1}(y)dy.$$

In 1976 Singh și Maddala au adăugat funcției de repartiție rata de hazard a datelor despre venit. Fie X o variabilă aleatoare dintr-o funcție de **repartiție Singh-Maddala** de forma:

$$F_X(x) = 1 - \left(1 + \left(\frac{x}{\sigma}\right)^a\right)^{-q}, x > 0; a, q, \sigma > 0,$$

Dacă $q > \frac{1}{a}$ atunci curba Lorenz pentru (2.1.10) este dată de Sarabia (2008):

$$L_X(p) = I_z\left(1 + \frac{1}{a}; -\frac{1}{a} + q\right); 0 \leq p \leq 1$$

unde $z = 1 - (1-p)^{\frac{1}{q}}$ și I_z este funcția Beta regularizată incompletă dată prin raportul dintre funcția Beta incompletă și funcția Beta: $I_z(\alpha, \beta) = \dfrac{B(z; \alpha, \beta)}{B(\alpha, \beta)}$.

Funcția de repartiție corespunzătoare repartiției Dagum de tipul I este dată de Dagum (1977) prin

$$F_X(x) = 1 - \left(1 + \left(\frac{x}{\sigma}\right)^{-a}\right)^{-q}, x > 0;$$

unde $a, q, \sigma > 0$. Curba Lorenz a fost dată de Dagum (1977) prin

$$L_X(p) = I_z\left(q + \frac{1}{a}; 1 - \frac{1}{a}\right); 0 \le p \le 1$$

unde $z = 1 - (1-p)^{\frac{1}{q}}$ și I_z este funcția Beta incompletă regularizată.

Fiind disponibil un număr foarte mare de forme funcționale, apare întrebarea cum să alegem între ele când încercăm să facem o estimare. O modalitate este să estimăm mai multe funcții și să o alegem pe aceea care se potrivește cel mai bine pe setul de date. Modelul cel mai potrivit ar putea fi cel care maximizează funcția de verosimilitate, introdus de Chotikapanich și Griffiths (2002).

Curbe Lorenz pentru funcții noi de repartiție de tip exponențializat

Prima repartiție exponențializată a fost introdusă de Gupta (1998) ca fiind o familie nouă de funcții pe care a denumit-o repartiție exponențială exponențializată (EE). Cei doi parametri ai acesteia sunt parametrul de formă și parametrul de scală. Când $\alpha = 1$ obținem familia exponențială clasică.

Gupta și Kundu (2001) au observat că majoritatea proprietăților familiei EE sunt comune cu cele ale familiilor de funcții de repartiție Weibull sau Gamma. Gupta (1998) a introdus o altă repartiție cu doi parametri, denumită **Pareto exponențializată II** (EP) care are funcția de repartiție de forma:

$$F_X(x;\theta,\lambda) = [1-(1+x)^{-\lambda}]^{\theta}, x > 0, \lambda > 0, \theta > 0.$$

Când $\theta = 1$, repartiția de mai sus corespunde repartiției Pareto de ordinul II. Ei au arătat că $EP(\theta,\lambda)$ poate fi foarte utilă în analiza mulțimilor ce conțin date despre supraviețuire.

Ali (2007) a introdus funcția de repartiție **Pareto exponențializată I**:

$$F_X(x,\alpha,\beta) = \left[1 - \left(\frac{x}{\sigma}\right)^{-\alpha}\right]^{\beta}; x \ge \sigma; \sigma > 0; \alpha > 0; \beta > 0.$$

Pentru a determina curba Lorenz avem nevoie de cuantila funcției de repartiție:

$$F_X^{-1}(y) = \sigma\left(1 - y^{\frac{1}{\beta}}\right)^{-\frac{1}{\alpha}}, 0 < y < 1$$

și de media μ_X:

$$\mu_X = \beta\sigma \cdot B\left(\beta, -\frac{1}{\alpha} + 1\right); \alpha > 1$$

unde $B(\cdot,\cdot)$ este funcția Beta. Pentru integrala din medie am utilizat $t = 1 - \left(\frac{x}{\sigma}\right)^{-\alpha}$.

Teorema 2.1

Fie $X \in L$ cu funcția de repartiție $F_X(\cdot)$ *Pareto exponențializată I* și funcția sa cuantilă $F^{-1}(y)$. Atunci funcția Lorenz corespunzătoare va fi

$$L_X(p) = I_t\left(\beta; -\frac{1}{\alpha} + 1\right); t = p^{\frac{1}{\beta}}, \text{ pentru } 0 \le p \le 1.$$

Introducem funcția de repartiție **Singh-Maddala Exponențializată** ESM$(\mathrm{a,q},\sigma,\alpha)$ ca o generalizare a repartiției Singh-Maddala. Funcția de repartiție pentru ESM va fi:

$$F_X(x)=\left(1-\left(1+\left(\frac{x}{\sigma}\right)^{a}\right)^{-q}\right)^{\alpha},\ x>0; a,q,\sigma,\alpha>0,$$

cu media:

$$\mu_X=\alpha\sigma q\binom{\alpha-1}{k}\sum_{k=0}^{\infty}(-1)^{k+\frac{1}{a}}B(-qk-q;\frac{1}{a}+1).$$

Teorema 2.2
Fie $X\in L$ cu funcția de repartiție *Singh Maddala Exponențializată* $F_X(\cdot)$. Atunci curba Lorenz este

$$L_X(p)=\frac{\sum_{k=0}^{\infty}(-1)^{k+\frac{1}{a}}\binom{\alpha-1}{k}I_{s_0}\left(-q(k+1);\frac{1}{a}+1\right)}{\sum_{k=0}^{\infty}(-1)^{k+\frac{1}{a}}\binom{\alpha-1}{k}B\left(-q(k+1);\frac{1}{a}+1\right)}-1,$$

pentru $0\le p\le 1$, unde $s_0=(1-p^{\frac{1}{\alpha}})^{-\frac{1}{q}}$.

În continuare definim funcția de repartiție **Dagum Exponențializată I** ED-I$(\mathrm{a,q},\sigma,\alpha)$ ca o generalizare a repartiției Dagum de tip I. Funcția de repartiție pentru ED-I va fi:

$$F_X(x)=\left(1-\left(1+\left(\frac{x}{\sigma}\right)^{-a}\right)^{-q}\right)^{\alpha},\ x>0; a,q,\sigma,\alpha>0,$$

cu limita la zero $\lim_{x\to 0}F_X(x)=0$, iar la infinit $\lim_{x\to\infty}F_X(x)=1$.

Teorema 2.3
Fie $X\in L$ cu funcția de repartiție $F_X(\cdot)$ *Dagum Exponențializată I.* Atunci curba Lorenz este

$$L_X(p)=\frac{\alpha(-1)^{\alpha-\frac{1}{a}}\left(I_{t_0}(-\alpha,\alpha)+\frac{1}{a}I_{t_0}(-\alpha+\frac{1}{q}-1,\alpha)+\frac{1+a}{2a^2}I_{t_0}(-\alpha+\frac{2}{q}-1,\alpha)+\ldots\right)}{1-\alpha B(1-\frac{1}{q};\alpha)},$$

pentru $0\le p\le 1$, unde $t_0=(1-p^{\frac{1}{\alpha}})^{-\frac{1}{q}}$.

Entropia maximă Tsallis și Curba Lorenz

Cercetările recente în statistică au intensificat interesul asupra entropiei Tsallis. Teoria a fost introdusă de Constantino Tsallis (1988) cu scopul de a generaliza entropia standard Boltzmann-Gibbs. În noua teorie este introdus parametrul real q pentru măsurarea gradului de nedeterminare.

Entropia Tsallis a unei variabile aleatoare X cu valori reale și densitatea de repartiție $f(x)$ este definită prin:

$$H_{Tsallis}(f)=\left[1-\int_{-\infty}^{\infty}f^{q}(x)dx\right]\frac{1}{q-1}$$

Holm (1993) a obținut o familie de densități de repartiție ale entropiei Shannon maximă sub restricțiile mediei și indicelui Gini.

Ryu (2008) a testat performanța și utilitatea unei măsuri de inegalitate în funcție de indicele Bonferroni prin metoda maximizării entropiei Shannon. Tot în același an Rhode (2008) a prezentat o legătură formală între două măsuri ale inegalității: entropia generalizată și curba Lorenz. Concluzia lui Rhode este una foarte importantă, și anume că putem considera curba Lorenz ca bază pentru majoritatea măsurilor de inegalitate.

Fie X o variabilă aleatoare în intervalul $[x_0, x_1]$. Atunci avem

$$H_{Tsallis}(f) = \left[1 - \int_{x_0}^{x_1} f^q(x)dx\right]\frac{1}{q-1}$$

Notăm $p = F(x)$, apoi entropia Tsallis va fi

$$H_{Tsallis}(f) = \left[1 - \int_0^1 \left(\frac{dF^{-1}}{dp}\right)^{1-q} dp\right]\frac{1}{q-1}$$

unde $F^{-1}(\cdot)$ este inversa funcției de repartiție $F(\cdot)$.

Căutăm maximul funcției descrisă în Holm (1993) pentru indicele Gini și media cunoscute.

Considerăm următoarea problemă de optimizare:

$$Max\left\{H_{Tsallis}(f) = \left[1 - \int_0^1 \left(\frac{dF^{-1}}{dp}\right)^{1-q} dp\right]\frac{1}{q-1}\right\}$$

astfel încât $\int_0^1 (1-p)\frac{dF^{-1}}{dp}dp = \mu - x_0$ și $\int_0^1 (1-p)\,\mathrm{p}\frac{dF^{-1}}{dp}dp = G\cdot\mu$.

Fie o familie de curbe Lorenz cu $L(0) = 0$ și $L(1) = 1$.

Pentru cazul $\lambda_1 \neq 0, \lambda_2 = 0$ funcția de repartiție va fi

$$F(x) = 1 - \left[\frac{3q-1}{2q-1} - \frac{(x-\mu)B\left(2;3-\frac{1}{q}\right)}{G\cdot\mu}\right]^{\frac{q}{q-1}},$$

iar funcția Lorenz corespunzătoare entropiei Tsallis:

$$L_X^{Tsallis}(p) = \frac{G}{B\left(2;3-\frac{1}{q}\right)}\frac{q}{2q-1}\left[\frac{q-1}{q}p - 1 + (1-\mathrm{p})^{2-\frac{1}{q}}\right] + p$$

În continuare analizăm cazul $\lambda_1 = 0, \lambda_2 \neq 0$:

Vom avea:

$$L_X^{Tsallis}(p) = p - \left(\frac{q-1}{3q-2}\right)^2\left\{Gp + \frac{\left[pB\left(p;1-\frac{1}{q},1-\frac{1}{q}\right) - B\left(p;2-\frac{1}{q},1-\frac{1}{q}\right)\right]2^{3-\frac{2}{q}}\Gamma\left(\frac{3}{2}-\frac{1}{q}\right)}{\sqrt{\pi}\,\Gamma\left(1-\frac{1}{q}\right)}\right\}$$

Cazul $\lambda_1 \neq 0, \lambda_2 \neq 0$ conduce la:

Folosind dezvoltarea

$$(1+\gamma p)^{-\frac{1}{q}} = \sum_{m\ge 0} A\left(-\frac{1}{q}, m\right)\gamma^m p^m$$

unde

$$A(s,m) = \frac{\Gamma(s+1)}{\Gamma(s+1-m)\Gamma(m+1)},$$

rezultă:

$$F^{-1}(p) = \lambda_1^{-\frac{1}{q}} \sum_{m\ge 0} A\left(-\frac{1}{q}, m\right)\gamma^m B(p, m+1, 1-\frac{1}{q}) + x_0.$$

Restricții de medie. Inegalități

În cele ce urmează considerăm că media se află într-un interval compact $[\mu', \mu'']$. Astfel restricțiile sunt de forma

$$G\mu' \le \int_0^1 (2p-1)F^{-1}(p)dp \le G\mu''$$

$$\mu' \le \int_0^1 F^{-1}(p)dp \le \mu''.$$

După transformări obținem restricțiile

$$\mu' - x_0 \le \int_0^1 (1-p)\frac{dF^{-1}(p)}{dp}dp \le \mu'' - x_0$$

$$G\mu' \le \int_0^1 (1-p)p\frac{dF^{-1}(p)}{dp}dp \le G\mu''$$

în care se maximizează entropia Tsallis.

Teorema 2.4
Soluția acestei probleme este dată de

$$\frac{dF^{-1}(p)}{dp} = \left[\lambda_1(1-p) + \lambda_2 p(1-p)\right]^{-1/q}$$

unde λ_1, λ_2 sunt multiplicatori Lagrange pentru care se pot obține informații din restricțiile de mai sus.
În continuare prezentăm soluțiile acestei probleme.

Teorema 2.5
Dacă $\lambda_1 \ne 0, \lambda_2 = 0$ atunci

$$\frac{dF^{-1}(p)}{dp} = \lambda_1^{-1/q}(1-p)^{-1/q} \text{ și}$$

$$F^{-1}(p) = \lambda_1^{-1/q}\left[1-(1-p)^{1-1/q}\right] + x_0,$$

unde

$$\lambda_1^{-1/q} \in \left[\max\left\{\frac{(\mu'-x_0)(2-q)}{q}, G\mu'\left[B(2,3-\frac{1}{q})\right]^{-1}\right\}, \min\left\{\frac{(\mu''-x_0)(2-q)}{q}, G\mu''\left[B(2,3-\frac{1}{q})\right]^{-1}\right\}\right]$$

În acest caz soluția problemei este curba Lorenz $L_X^{Tsallis}$ ca o funcție interval, adică

$$L_X^{Tsallis}(p) = \left[\frac{1}{\mu''}\left\{\lambda_1^{-\frac{1}{q}}\left(p-1+(1-p)^{2-\frac{1}{q}}\right) + px_0\right\}, \frac{1}{\mu'}\left\{\lambda_1^{-\frac{1}{q}}\left(p-1+(1-p)^{2-\frac{1}{q}}\right) + px_0\right\}\right].$$

Teorema 2.6

Dacă $\lambda_1 = 0, \lambda_2 \neq 0$ atunci

$$\frac{dF^{-1}(p)}{dp} = \lambda_2^{-1/q}(p(1-p))^{-1/q} \text{ și}$$

$$F^{-1}(p) = \lambda_2^{-1/q} B_p(1-\frac{1}{q}, 1-\frac{1}{q}) + x_0,$$

unde

$$\lambda_2^{-1/q} \in \left[\max\left\{ \frac{(\mu' - x_0)}{B(1-\frac{1}{q}, 2-\frac{1}{q})}, \frac{G\mu'}{B(2-\frac{1}{q}, 2-\frac{1}{q})} \right\}, \min\left\{ \frac{(\mu'' - x_0)}{B(1-\frac{1}{q}, 2-\frac{1}{q})}, \frac{G\mu''}{B(2-\frac{1}{q}, 2-\frac{1}{q})} \right\} \right]$$

În acest caz soluția optimă este curba Lorenz interval

$$L_X^{Tsallis}(p) = \left[\frac{\lambda_2^{-1/q}}{\mu''} \int_0^p B_t\left(1-\frac{1}{q}, 1-\frac{1}{q}\right) dt + \frac{px_0}{\mu''}, \frac{\lambda_2^{-1/q}}{\mu'} \int_0^p B_t\left(1-\frac{1}{q}, 1-\frac{1}{q}\right) dt + \frac{px_0}{\mu'} \right]$$

Teorema 2.7

Dacă $\lambda_1 \neq 0, \lambda_2 \neq 0$ atunci

$$\frac{dF^{-1}(p)}{dp} = \lambda_2^{-1/q}(1-p)^{-1/q}(1+\gamma p)^{-1/q} \text{ și}$$

$$F^{-1}(p) = \lambda_1^{-1/q} \sum_{m \geq 0} A(-\frac{1}{q}, m)\gamma^m B_p(m+1, 1-\frac{1}{q}) + x_0,$$

cu $A(s,m)$ dat de (2.3.30), unde λ_1, λ_2 verifică inegalitățile

$$\mu' - x_0 \leq \lambda_1^{-1/q} \sum_{m \geq 0} A(-\frac{1}{q}, m)\gamma^m \leq \mu'' - x_0$$

$$\mu' G \leq \lambda_1^{-1/q} \sum_{m \geq 0} A(-\frac{1}{q}, m)\gamma^m \leq \mu'' G.$$

În acest caz soluția optimă este curba Lorenz interval

$$L_X^{Tsallis}(p) = [\frac{\lambda_1^{-1/q}}{\mu''} \sum_{m \geq 0} A(-\frac{1}{q}, m)\gamma^m \int_0^p B_t\left(m+1, 1-\frac{1}{q}\right) dt + \frac{px_0}{\mu''},$$
$$\frac{\lambda_2^{-1/q}}{\mu'} \sum_{m \geq 0} A(-\frac{1}{q}, m)\gamma^m \int_0^p B_t\left(m+1, 1-\frac{1}{q}\right) dt + \frac{px_0}{\mu'}].$$

În finalul acestei secțiuni vom considera cazul când media este cunoscută dar indicele Gini aparține la interval dat $[G', G'']$, unde $G' < G''$. Relativ la astfel de restricții curba Lorenz care maximizează entropia Tsallis este dată în următoarea teoremă. Și în acest caz au loc relațiile (2.3.10) și (2.3.11):

Teorema 2.8

Dacă $\lambda_1 = 0$ și $\lambda_2 \neq 0$ atunci $\frac{dF^{-1}}{dp} = \lambda_2^{-\frac{1}{q}} [p(1-p)]^{-\frac{1}{q}}$, și

$F^{-1}(p)=\lambda_2^{-\frac{1}{q}}B\left(p;1-\frac{1}{q},1-\frac{1}{q}\right)+x_0$; pentru $1-\frac{1}{q}>0$, unde $\lambda_2^{-\frac{1}{q}}=\dfrac{x_0-\mu}{B\left(1-\frac{1}{q},2-\frac{1}{q}\right)}$; iar x_0 verifică inegalitatea

$$\frac{G'\mu B\left(1-\frac{1}{q},2-\frac{1}{q}\right)}{B\left(2-\frac{1}{q},2-\frac{1}{q}\right)}+\mu\le x_0\le\frac{G''\mu B\left(1-\frac{1}{q},2-\frac{1}{q}\right)}{B\left(2-\frac{1}{q},2-\frac{1}{q}\right)}+\mu .$$

În acest caz soluția este curba Lorenz

$$L_X^{Tsallis}(p)=\frac{G}{B\left(2;3-\frac{1}{q}\right)}\frac{q}{2q-1}\left[\frac{q-1}{q}p-1+(1-\mathrm{p})^{2-\frac{1}{q}}\right]+p$$

Mecanismul Marshall-Olkin pentru a obține curbe Lorenz

Marshall și Olkin (1997) au propus o metodă interesantă de a adăuga un parametru nou unei repartiții cunoscute. Repartiția rezultată, denumită repartiția extinsă Marshall-Olkin (MO), include repartiția originală ca fiind un caz particular și oferă o flexibilitate mare în modelarea diferitelor tipuri de date.

Prezentarea mecanismului Marshall-Olkin

ie $\bar{F}(x)=1-F(x)$ funcția de supraviețuire a unei variabile aleatoare continuă X ce depinde de un vector parametru $\beta=(\beta_1,...,\beta_q)^T$ de dimensiune q. În continuare fie $f(x)=\dfrac{dF(x)}{dx}$ densitatea de repartiție asociată funcției de repartiție $F(x)$. Atunci repartiția extinsă MO are funcția de supraviețuire pentru $x\in(-\infty,\infty),\alpha>0$ dată de

$$\bar{G}(x)=\frac{\alpha\bar{F}(x)}{1-\bar{\alpha}\bar{F}(x)}=\frac{\alpha\bar{F}(x)}{F(\mathrm{x})+\alpha\bar{F}(x)}$$

unde $\bar{\alpha}=1-\alpha$. În mod evident, ecuația (2.4.1) este o modalitate de a obține noi repartiții parametrice din cele existente. De exemplu, pentru $\alpha=1$ avem $\bar{G}(x)=\bar{F}(x)$. Densitatea de repartiție $g(x)$ ce corespunde repartiției (2.4.1) va avea forma:

$$g(x)=\frac{\alpha f(x)}{(1-\bar{\alpha}\,\bar{\mathrm{F}}(\mathrm{x}))^2}\text{, pentru } x\in(-\infty,\infty),\alpha>0$$

Câteva cazuri speciale ale ecuației (2.4.1) discutate în literatura recentă au considerat repartiția $F(x)$ ca fiind Pareto – în Ghitany (2005), Gamma – în Ristic (2007) sau Lomax în Ghitany (2007).

Marshall și Olkin (1997) au observat că repartițiile acestor forme extinse au o proprietate de stabilitate, în sensul că aplicând transformarea lui $\bar{G}$ nu obținem un rezultat nou, ci funcția rezultată conține același termen $\bar{F}(x)$, dar are o altă valoare a parametrului α. De aici Economu și Caroni (2007) au arătat că repartițiile extinse Marshall-Olkin au proprietatea de proporționalitate.

Utilizând această schemă de parametrizare (M-O), Preda V., Panaitescu E. și Ciumara R. (2011) au definit și analizat proprietățile asimptotice ale repartiției Poisson Exponențializată modificată. Această nouă repartiție este o generalizare a repartiției EP propusă de Kus (2007).

În acest an a apărut articolul "Asupra repartiției Weibull extinse Marshall-Olkin" în care Cordeiro și Lemonte (2013) au studiat repartiția Weibull extinsă Marshall-Olkin cu trei parametri (MOEW). Această repartiție a mai fost studiată de Ghitany (2005) și Zhang și Xie (2007). Ghitany a arătat că MOEW poate fi obținută ca o repartiție compusă prin mixarea repartiției exponențială, iar

Zhang și Xie au investigat caracteristicile modelului bazându-se pe graficul probabilităților funcției Weibull.

Familii Lorenz de tip Marshall-Olkin

Fie L_A o curbă Lorenz. Atunci din proprietățile Curbei Lorenz avem $L_A(0)=0$, $L_A(1)=1, L'_A(0)\geq 0, L''_A(\mathrm{p})\geq 0$.

Teorema 2.4.1

Dacă L_A este o curbă Lorenz iar $\beta\in[0,1]$ atunci $L_B:[0,1]\to[0,1]$, $L_B(\mathrm{p})=\dfrac{\beta L_A(p)}{1-\bar{\beta}L_A(p)}$, unde $\bar{\beta}=1-\beta$, este o curbă Lorenz.

Familii Lorenz generate de curbe Lorenz cunoscute utilizând principiul Marshall-Olkin

Corolarul 2.4.1.

Fie $L_A(p;\alpha)$ Curba Lorenz Exponențială. Atunci curba Lorenz generată prin mecanismul Marshall-Olkin va fi

$$L_B(p;\alpha;\beta)=1-\frac{\beta\left(e^{p\alpha}-1\right)}{e^{\alpha}-1+\beta\left(e^{p\alpha}-1\right)},\alpha>0,\beta\in[0,1].$$

Corolarul 2.4.2.

Fie $L_A(p;\alpha,\gamma)$ Curba Lorenz definită de **Ortega (1991)**. Atunci pentru orice $\beta\in[0,1],\alpha\geq 0,0<\gamma\leq 1$ curba Lorenz generată prin mecanismul Marshall- Olkin va fi

$$L_B(p;\alpha,\gamma,\beta)=1-\frac{1+p^{\alpha}(1-p)^{\gamma}-p^{\alpha}}{\beta p^{\alpha}\left[1-(1-p)^{\gamma}\right]+1+p^{\alpha}(1-p)^{\gamma}-p^{\alpha}}.$$

Corolarul 2.4.3.

Fie $L_A(p;\alpha,\gamma)$ Curba Lorenz definită de **Rasche (1980).** Atunci pentru orice $\beta\in[0,1],\gamma\geq 1,0<\alpha\leq 1$ curba Lorenz generată prin mecanismul Marshall- Olkin va fi

$$L_B(p;\alpha,\gamma,\beta)=\frac{\beta\left[1-(1-p)^{\alpha}\right]^{\gamma}}{2-\left[1-(1-p)^{\alpha}\right]^{\gamma}+\beta\left[1-(1-p)^{\alpha}\right]^{\gamma}}.$$

Corolarul 2.4.4.

Fie $L_A(p;\alpha,\gamma,\delta)$ din ecuația (2.1.15) Curba Lorenz de tip **Pareto Generalizată**. Atunci pentru orice $\beta\in[0,1],\alpha\geq 0,\gamma\geq 1,0<\delta\leq 1$ curba Lorenz generată prin mecanismul Marshall Olkin va fi

$$L_B(p;\alpha,\gamma,\delta,\beta)=\frac{\beta p^{\alpha}\left[1-(1-p)^{\delta}\right]^{\gamma}}{1-p^{\alpha}\left[1-(1-p)^{\delta}\right]^{\gamma}+\beta p^{\alpha}\left[1-(1-p)^{\delta}\right]^{\gamma}}.$$

Corolarul 2.4.5.

Fie $L_A(p;a,b,d)$ Curba Lorenz de tipul **Beta Kawakani**. Atunci pentru orice $\beta\in[0,1],a>0,0<d\leq 1,0<b\leq 1$ curba Lorenz generată prin mecanismul Marshall-Olkin va fi

$$L_B(p;a,b,d,\beta)=1-\frac{p-1-ap^d(1-p)^b}{1-(1-\beta)\left(p-ap^d(1-p)^b\right)}.$$

Familii Leimkuhler de tip Marshall-Olkin

Fie K_A și K_B două curbe Leimkuhler definite prin: $K_A(p)=1-L_A(1-p)$ și $K_B(p)=1-L_B(1-p)$.

Teorema 2.4.2

Fie K_A o curbă Leimkuhler. Atunci pentru orice $p\in[0,1]$

$$K_B(p)=\frac{K_A(p)}{1-\left(\frac{1}{\beta}+1\right)K_A(p)}$$

este o curbă Leimkuhler, unde $\beta\in[0,1]$.

Familii Lorenz de tip Marshall-Olkin de ordin *n*

Teorema 2.4.3

Dacă L_{A_0} este o curbă Lorenz iar $\beta_1,...\beta_n\in[0,1]$ atunci $L_{A_n}:[0,1]\to[0,1]$,

$$L_{A_n}(\mathrm{p})=\frac{\beta_1...\beta_n L_{A_0}(p)}{1-(1-\beta_1...\beta_n)L_{A_0}(p)}$$

este o curbă Lorenz.

Familii Lorenz generate prin mecanismul Marshall Olkin din două curbe Lorenz cunoscute

Teorema 2.4.4

Fie L_A, L_B două curbe Lorenz și $\beta\in[0,1]$. Atunci $L_C:[0,1]\to[0,1]$,

$$L_C(\mathrm{p})=\frac{\beta L_A(p)}{1-\bar{\beta}L_B(p)}$$

este o curbă Lorenz.

III
MODELE DE CURBE LORENZ MIXTE

Introducere

Importanța Curbelor Lorenz în analiza statistică și economică a inegalității venitului ne motivează dorința de a găsi noi familii parametrice de Curbe Lorenz. Multitudinea de modele parametrice propuse în literatura de specialitate nu constituie un inconvenient, ci un considerent suplimentar dat de nepotrivirea în totalitate a curbelor empirice pe setul de venituri date.

Familii parametrice de Curbe Lorenz

Kakwani și Podder (1973, 1976) au propus primele modele parametrice de estimare a Curbelor Lorenz. În 1973 au introdus curba

$$L(p)=p^\gamma e^{-\eta(1-p)},\ \text{pentru } 0<p<1 \text{ și } \eta>0; 1<\gamma<2.$$

Utilizând sistemul de coordonate propus de Gini în 1932 de forma

$\eta = \frac{u+v}{\sqrt{2}}$ și $\pi = \frac{u-v}{\sqrt{2}}$, unde $0 < u < 1$, Kakwani și Podder au dat o altă definiție curbei $v = L(u)$, caracterizată prin

$$\eta = a\pi^{\alpha}(\sqrt{2} - \pi)^{\beta}, \text{ cu } a \geq 0,\ 0 \leq \alpha \leq 1 \text{ și } 0 < \beta \leq 1.$$

Propunem trei modele parametrice noi de Curbe Lorenz:

$$L_1(p;\theta,\nu) = \frac{1}{\ln p}\frac{p^{\nu} - p^{\theta}}{\nu - \theta}, \nu > \theta$$

$$L_2(p;\theta,k,\nu) = \frac{1}{\ln p}\frac{p^{\nu} - p^{\theta}}{\nu - \theta}[1 - (1-p)^k], \nu > \theta; \theta \geq 0; 0 < k \leq 1$$

$$L_3(p;\theta,\nu) = pe^{-\theta(2-p)}\frac{e^{\nu} - e^{\theta}}{\nu - \theta}, \nu > \theta; \theta \geq 0$$

Observație:
Aplicând limita după ν, cu $\nu \to \theta$ obținem
i) $\lim_{\nu \to \theta} L_1(p;\theta,\nu) = p^{\theta}$
ii) $\lim_{\nu \to \theta} L_2(p;\theta,\nu) = p^{\theta}[1 - (1-p)^k]$
iii) $\lim_{\nu \to \theta} L_3(p;\theta,\nu) = pe^{-\theta(1-p)}$

Teorema 3.1: Presupunem că $L_1(p;\theta,\nu)$ este definită și continuă pe [0,1], cu derivata de ordinul doi $L''_1(\cdot)$. Funcția $L_1(\cdot)$ este o Curbă Lorenz dacă și numai dacă

$$L_1(0;\theta,\nu) = 0,\ L'_1(0^+;\theta,\nu) \geq 0$$
$$L_1(1;\theta,\nu) = 1,\ L''_1(p;\theta,\nu) \geq 0, p \in (0,1).$$

Pentru curba L_3 indicele lui Gini va fi:

$$G_3 = 1 - 2\frac{1 - e^{\nu-\theta}}{\theta^2(\theta - \nu)}\left(\theta - 1 + e^{-\theta}\right)$$

Compunerea CL propuse cu repartiția Gamma

Curbele Lorenz compuse sunt o modalitate de a obține o potrivire mai bună (pe datele disponibile) prin construirea unor modele mai complexe ce combină o curbă parametrică Lorenz cu o densitate de repartiție cunoscută. Metoda de compunere a fost introdusă de Sarabia (2005).

Fie Curba Lorenz parametrică $L_1(p;\theta,\nu)$, unde θ,ν sunt vectori-parametru. De exemplu, θ poate fi un parametru scalar ce reprezintă un factor al omogenității populației.

Fie $\pi(\theta;\alpha,\lambda)$ o densitate de repartiție absolut continuă pe $\theta \subset \mathbb{R}$ unde α,λ sunt parametri reali.

Teorema 3.2: Expresia $\tilde{L}_1(p;\nu;\alpha,\lambda) = \int_{\Theta} L_1(p;\theta,\nu)\pi(\theta;\alpha,\lambda)d\theta$ definește o curbă Lorenz, unde $\pi(\theta;\alpha,\lambda) = \frac{\lambda^{\alpha}}{\Gamma(\alpha)}(\theta-1)^{\alpha-1}e^{-\lambda(\theta-1)}I(\theta > 1)$, oricare ar fi $\alpha,\lambda > 0$.

Teorema 3.3: Expresia $\tilde{L}_2(p;\alpha,\lambda,k,\nu) = \int_0^{\infty} L_2(p;\theta,k,\nu)\pi(\theta;\alpha,\lambda)d\theta$ definește o curbă Lorenz, unde $\pi(\theta;\alpha,\lambda) = \frac{\lambda^{\alpha}}{\Gamma(\alpha)}\theta^{\alpha-1}e^{-\lambda\theta}I(\theta > 0)$, oricare ar fi $\alpha,\lambda > 0$.

Teorema 3.4: Expresia $\tilde{L}_3(p;\alpha,\lambda,\nu) = \int_0^{\infty} L_3(p;\theta,\nu)\pi(\theta;\alpha,\lambda)d\theta$ definește o curbă Lorenz, unde $\pi(\theta;\alpha,\lambda) = \frac{\lambda^{\alpha}}{\Gamma(\alpha)}\theta^{\alpha-1}e^{-\lambda\theta}I(\theta > 0)$, oricare ar fi $\alpha,\lambda > 0$.

În continuare dorim să găsim indicele Gini pentru noua curbă Lorenz $\tilde{L}_3$. Pentru aceasta vom folosi Teorema 3 a lui Sarabia (2005) care ne indică următoarea proprietate:

$$G(\tilde{L}_3) = E_{\pi}[G(L_3)],$$

unde E_π este valoarea medie a lui $G(L_3)$ în raport cu funcția densitate de probabilitate
Forma finală a indicelui Gini va fi:

$$G_{\tilde{L}_3}(\lambda,\nu,\alpha)=1-2\lambda^{\alpha}e^{-\nu\lambda}(-\nu)^{\alpha-3}\frac{\Gamma(\alpha-2)}{\Gamma(\alpha)}\left[\Gamma(3-\alpha,-\nu\lambda)-\Gamma(3-\alpha,\nu(1+\lambda))\right].$$

Importanța curbelor Lorenz în analiza economică și statistică a inegalității veniturilor și bogăției motivează dorința de a găsi noi familii parametrice de curbe Lorenz. Concluzia noastră este că aproximarea parametrică mixtă aduce o concordanță mai mare prin introducerea de restricții stricte.

Prin utilizarea unor instrumente statistice corespunzătoare se pot face comparații între noile curbe mixte și cele clasice propuse de Sarabia (2005). Estimările parametrilor modelelor pot fi determinate prin utilizarea metodei a celor mai mici pătrate neliniară.

În acest capitol am introdus modele de Curbe Lorenz mixte pentru familii parametrice de CL prin compunerea acestora cu repartiția Gamma, rezultate publicate în [7].

IV
INTERVALE DE INCREDERE ALE CURBELOR LORENZ GENERALIZATE

Familii de repartiții Weibull Exponențializate și rata de hazard corespunzătoare

Repartiția Weibull este o funcție foarte bine cunoscută. Denumită după Walddi Weibull, un fizician suedez, a fost utilizată inițial în analiza rezistenței la rupere a materialelor.

Familia de repartiții Weibull Exponențializată a fost introdusă de Mudholkar și Sarisastva (1993) ca fiind o generalizare a familiei de repartiții Weibull. Aplicațiile acestei repartiții au fost ilustrate de Mudholkar (1995).

Aplicăm o extindere familiei repartiției exponențiale și obținem familia de patru parametri definită prin

$$F(x;\mu,k,\lambda;\theta)=\begin{cases}\left(1-e^{-\left(\frac{x-\mu}{\lambda}\right)^{k}}\right)^{\theta}, x\geq\mu\\ 0, x<\mu\end{cases}$$

unde $\mu \leq x$ este parametrul de locație, $k>0$ este parametrul de formă, $\lambda>0$ este parametrul de mărime și $\theta>0$ este al doilea parametru de mărime. Această familie se numește Repartiția Weibull Exponențializată.

Densitatea de repartiție corespunzătoare este

$$f(x;\mu,k,\lambda;\theta)=\frac{\theta k}{\lambda}\left(\frac{x-\mu}{\lambda}\right)^{k-1}e^{-\left(\frac{x-\mu}{k}\right)^{k}}\left[1-e^{-\left(\frac{x-\mu}{\lambda}\right)^{k}}\right]^{\theta-1}$$

Pentru câteva valori fixe ale parametrilor obținem repartiții diferite: dacă $\theta=1$ repartiția devine Weibull cu doi parametri. Pentru $k=1$ se reduce la familia exponențială, și dacă $\theta=k=1$ avem repartiția exponențială cu un parametru. Dacă $\theta=1; k=2$ obținem repartiția Rayleigh.

Avantajul principal al acestei repartiții este flexibilitatea funcției de hazard corespunzătoare:

$$h(x;\mu,k,\lambda;\theta)=\frac{\frac{\theta k}{\lambda}\left(\frac{x-\mu}{\lambda}\right)^{k-1}e^{-\left(\frac{x-\mu}{k}\right)^{k}}\left[1-e^{-\left(\frac{x-\mu}{\lambda}\right)^{k}}\right]^{\theta-1}}{1-\left(1-e^{-\left(\frac{x-\mu}{\lambda}\right)^{k}}\right)^{\theta}}$$

Curba Lorenz Generalizată este dată de formula (4.5):

$$GL_X(u)=\int_0^u\left(\lambda\left[\ln\left(\frac{1}{1-t^{\frac{1}{\theta}}}\right)\right]^{\frac{1}{k}}+\mu\right)dt, 0\le u\le 1$$

Estimatorul de verosimilitate maximă pentru repartiția Weibull Exponențializată

Fie $X_1, X_2,..., X_n$ un eșantion de variabile aleatoare independente dintr-o Repartiție Weibull Exponențializată F_X Atunci funcția de verosimilitate este dată de:

$$L(k,\lambda,\theta)=\prod_{i=1}^{n} f(x_i|k,\lambda,\theta)$$

$$=\left(\frac{\theta k}{\lambda}\right)^n\frac{1}{\lambda^{n(k-1)}}\prod_{i=1}^{n}x_i^{k-1}e^{-\sum_{i=1}^{n}\left(\frac{x_i}{\lambda}\right)^k}\prod_{i=1}^{n}\left(1-e^{-\sum_{i=1}^{n}\left(\frac{x_i}{\lambda}\right)^k}\right)^{\theta-1}. \qquad (4.6)$$

Logaritmând expresia funcției de verosimilitate (4.6) obținem:

$$\ln L(k,\lambda,\theta)=n\ln\left(\frac{\theta k}{\lambda}\right)-n(k-1)\ln\lambda+(k-1)\sum_{i=1}^{n}\ln x_i-\frac{1}{\lambda_k}\sum_{i=1}^{n}x_i^k+(\theta-1)\sum_{i=1}^{n}\ln(1-e^{-\left(\frac{x_i}{\lambda}\right)^k}).$$

Dorim maximizarea funcției de verosimilitate pentru fiecare parametru:

$$\frac{\ln L(k,\lambda,\theta)}{\partial\lambda}=n\ln\left(\frac{\theta k}{\lambda}\right)-n(k-1)\ln\lambda$$

$$+(k-1)\sum_{i=1}^{n}\ln x_i-\frac{1}{\lambda_k}\sum_{i=1}^{n}x_i^k+(\theta-1)\sum_{i=1}^{n}\ln(1-e^{-\left(\frac{x_i}{\lambda}\right)^k})$$

Teorema 3.1. Pentru $\theta>2$ estimatorii $(\hat{k},\hat{\lambda},\hat{\theta})$ parametrilor (k,λ,θ) sunt consistenți și $\sqrt{n}(\hat{k}-k,\hat{\lambda}-\lambda,\hat{\theta}-\theta)$ este repartizat asimptotic normal cu vectorul- medie nul și matricea de covarianță I^{-1}, unde

$$I=-\frac{1}{n}\begin{pmatrix} E\left(\frac{\partial^2\ln L}{\partial k^2}\right) & E\left(\frac{\partial^2\ln L}{\partial k\partial\lambda}\right) & E\left(\frac{\partial^2\ln L}{\partial k\partial\theta}\right)\\ E\left(\frac{\partial^2\ln L}{\partial\lambda\partial k}\right) & E\left(\frac{\partial^2\ln L}{\partial\lambda^2}\right) & E\left(\frac{\partial^2\ln L}{\partial\lambda\partial\theta}\right)\\ E\left(\frac{\partial^2\ln L}{\partial\theta\partial k}\right) & E\left(\frac{\partial^2\ln L}{\partial\theta\partial\lambda}\right) & E\left(\frac{\partial^2\ln L}{\partial\theta^2}\right)\end{pmatrix}$$

iar mediile derivatelor parțiale de ordinul doi ale funcției de verosimilitate sunt de forma:

Intervalele de încredere ale curbei Lorenz Generalizată utilizând aproximarea normală

În acest paragraf se va utiliza aproximarea normală pentru a construi intervalele de încredere ale Curbei Lorenz Generalizată (CLG). În primul rând căutăm un estimator pentru CLG.

Fie $X_1, X_2, ..., X_n$ un eșantion de variabile aleatoare independente din F_X. Un estimator consistent pentru CLG este:

$$GL_X(u) = \int_0^u F^{-1}(t)dt$$

unde $F^{-1}(p)$ este p-cuantila lui F .

În consecință, putem construi un interval de încredere bazat pe aproximarea normală pentru CLG astfel:

$$(\hat{\theta} - z_{1-\frac{\alpha}{2}}\hat{\sigma}_\nu / \sqrt{n}, \hat{\theta} + z_{1-\frac{\alpha}{2}}\hat{\sigma}_\nu / \sqrt{n})$$

Unde $z_{1-\frac{\alpha}{2}}$ este cuantila repartiției normale standard.

V
PROPRIETATI ASIMPTOTICE ALE ESTIMATORILOR NON-PARAMETRICI AI CURBEI LEIMKUHLER SI INDICELUI KAKWANI

Introducere

Fie $(\Omega, \mathscr{F}, P)$ un spațiu de probabilitate și X o variabilă aleatoare cu media $\mu_X < \infty$.

Notăm funcția de repartiție a lui X prin F_X și funcția cuantilă, funcția inversă continuă la stânga a lui F cu $Q(p) = \inf\{x \in \mathbb{R}; F(x) \geq p\}$. Atunci curba Lorenz este o funcție convexă crescătoare $L_F : [0,1] \to [0,1]$ definită de Gastwirth (1971) astfel:

$$L_F(p) = \mu_X^{-1} \int_0^p Q(y)dy, \text{ pentru } 0 \leq p \leq 1.$$

Presupunem că $X_1, X_2, ..., X_n$ sunt variabile aleatoare independente în spațiul de probabilitate $(\Omega, \mathscr{F}, P)$ pentru o repartiție F_X absolut continuă cu $F_X(0) = 0$.

După înlocuirea lui F_X în (5.1) cu funcția de repartiție empirică continuă la dreapta F_n, Gastwirth a construit curba Lorenz empirică $L_n : [0,1] \to [0,1]$, ca fiind:

$$L_n(p) = \mu_n^{-1} \int_0^p Q_n(y)dy, \text{ pentru } 0 \leq p \leq 1,$$

unde $\mu_n = (x_1 + ... + x_n)/n$.

Problema găsirii unei estimări neparametrice a verosimilității maxime (NPMLE) pentru funcția de repartiție F pe baza a două eșantioane independente: un eșantion $\{X_i; i = 1,..,m\}$ pornind de la F și un eșantion $\{Y_i; i = 1,..,n\}$ pornind de la G a fost abordată de Vardi (1982). Estimatorul empiric al lui F este

$F_n(t) = \nu_n^{-1}\int_0^t y^{-1}dG_n(y)$, unde $\nu_n = \int_0^\infty y^{-1}dG_n(y)$ și G_n, estimatorul empiric al lui G, este dat de $G_n(t) = 1/n\sum_{i=1}^{n} I(Y_i \le t)$, iar $I(A)$ descrie indicatorul evenimentului A.

Presupunem că procesul empiric este dat de

$$\alpha_n(t) = n^{1/2}[F_n(t) - F(t)], t \ge 0 .$$

Sen (1984) a demonstrat convergența slabă a procesului α_n la un proces Gaussian în ipoteza $EY^{-2} < \infty$.

Fie $Y_{n:1} \le ... \le Y_{n:n}$ o statistică ordonată corespunzătoare lui $Y_1,...,Y_n$. Atunci estimatorul eșantionului ce corespunde cuantilei $Q(p)$ este definit de $Q_n(p)(=Y_{n:k})$ unde $k(=k_n)$ este un întreg aleator, bine ales, ce depinde de statisticile de ordine și e definit prin

$$k_n = \max\left\{k; \sum_{i=1}^{k} Y_{n:i}^{-1} \le p\left(\sum_{i=1}^{n} Y_{n:i}^{-1}\right)\right\}; 0 < p < 1.$$

Vom considera în continuare ipotezele lui Fakoor (2011):

(A1) Presupunem că funcția de repartiție a lui F are funcția de densitate de repartiție continuă f într-o vecinătate a lui $Q(p)$ și $f(Q(p))$ este strict pozitivă și finită pentru orice $0 < p < 1$.

(A2) Presupunem că $E(Y^{-2-\delta}) < \infty$, pentru anumiți $\delta > 0$, și

$$\sup\{|f'(x)|; x \in \mathbb{R}^+\} < \infty .$$

(A3) Avem $\int_0^\infty (G(y))^{1/r} y^{-2}dy < \infty$, pentru anumiți parametri r>2.

Folosind aproximările procesului empiric

$$\gamma_n(t) = \sqrt{n}\left(G_n(t) - G(t)\right), t \ge 0$$

se poate obține o aproximare tare pentru α_n.

Horvath (1985) definește procesul $\Gamma(t,n)$ pentru aproximarea α_n astfel încât

$$\Gamma(t,n) = \nu^{-1}\int_0^t y^{-1}dK(y,n) - \nu^{-1}F(t)\int_0^\infty y^{-1}dK(y,n) .$$

Atunci procesul Gaussian cu medie zero va avea covarianța

$$E(\Gamma(x,n)\Gamma(y,m)) = (mn)^{-\frac{1}{2}}(m \wedge n)\left(\sigma(x \wedge y) - F(x)\sigma(y) - F(y)\sigma(x) + F(x)F(y)\sigma\right) \quad (5.3)$$

unde $\sigma(t) = \nu^{-2}\int_0^t y^{-2}dG(y)$, și $\sigma = \lim_{t\to\infty}\sigma(t) = \nu^{-2}\int_0^\infty y^{-2}dG(y)$.

Notăm procesul Lorenz

$$l_n(p) = \sqrt{n}\{L_n(p) - L_F(p)\},\ 0 \le p \le 1.$$

Goldie (1977) a demonstrat consistența uniformă a lui L_n către L_F și a determinat convergența slabă a procesului Lorenz către un proces Gaussian în anumite condiții.

Curba Leimkuhler

Familia de curbe Lorenz poate modela corect curba de început a venitului (de jos, corespunzătoare persoanelor cu venituri mici), dar în unele cazuri nu suficient de relevant curba de sus a observațiilor veniturilor. Sarabia (2008) a introdus o curbă care modelează partea de sus a repartiției.

Fie $K_n(p)$ statisticile Leimkuhler definite de

$$K_n(p) = 1 - \mu_n^{-1}\int_0^{1-p} Q_n(u)du$$

Comportamentul asimptotic al statisticii Leimkuhler

Arătăm în continuare consistența tare uniformă și aproximația tare Gaussiană pentru $K_n(p)$. Pentru asta vom folosi rezultatele corespondente pentru statisticile Lorenz și rezultate clasice de aproximare tare.

Teorema 5.3.1 *(Consistența tare uniformă)*

Presupunem $\delta > 0$ astfel încât $F' = f$ este continuă și strict pozitivă pe $[Q(p)-\delta; Q(p)+\delta]$. Atunci

$$\sup_{0<p<1} |K_n(p) - K_F(p)| = O\left(\sqrt{\frac{\log\log n}{n}}\right) \text{ P-a.s.}$$

Definim în continuare procesul Leimkuhler

$$k_n(p) = \sqrt{n}\left[K_n(p) - K_F(p)\right], p \in [0,1]$$

și procesul Gaussian $\{\Gamma(t,n), t \geq 0\}$.

Teorema 5.3.2 *(Aproximarea tare Gaussiană)*

Presupunem că sunt îndeplinite (A1)-(A3). Atunci pentru un spațiu de probabilitate potrivit de mare există un proces Gaussian cu doi parametri, de medie zero $\{\Gamma(t,n), t \geq 0\}$ și există r>2 astfel încât

i1) $\sup\limits_{0\leq p\leq 1}\left|k_n(1-p) - \frac{1}{\mu}\left(K_F(1-p)G(n) - H(n,p)\right)\right| = O\left((n^{-\frac{1}{4}}\log n) \vee (n^{-\lambda})\right)$ a.p.t,

pentru $0 < \lambda < \dfrac{r-2}{2r}$; unde $G(n) = \int_0^1 \dfrac{\Gamma(Q(y),n)}{f(Q(y))}dy$ și $\mathrm{H}(n,p) = \int_{1-p}^1 \dfrac{\Gamma(Q(y),n)}{f(Q(y))}dy$.

i2) există un șir de procese Gaussiene de medie zero $\{\Gamma_n(t), t \geq 0\}$ și există $r > 2$ pentru orice λ cu $0 < \lambda < \dfrac{r-2}{2r}$, atunci avem:

$$\sup_{0\leq p\leq 1}\left|k_n(1-p) - \frac{1}{\mu}\left(K_F(1-p)G_n(n) - H_n(n,p)\right)\right| = O\left((n^{-\frac{1}{4}}\log n) \vee (n^{-\lambda})\right) \text{ a.p.t,}$$

unde $G_n(n) = \int_0^1 \dfrac{\Gamma_n(Q(y))}{f(Q(y))}dy$ și $H_n(n,p) = \int_{1-p}^1 \dfrac{\Gamma_n(Q(y))}{f(Q(y))}dy$.

Einmahl (2007) reafirmă că legea Strassen a logaritmului iterat se poate aplica pentru orice funcțională continuă $\psi : C[0,1] \to \mathbb{R}$ cu probabilitate unu și $\left\{\psi(S_{(n)} / \sqrt{2nLLn}; n \geq 1\right\}$ este mărginită.

Luăm în considerare motivația lui Fakoor (2011). Fie $C(0,1)$ spațiul de funcții continue pe [0;1] și S mulțimea de funcții absolut continue Strassen:

$$S = \{\varphi \mid \varphi : [0,1] \to \mathbb{R}, \varphi(0) = 0, \int_0^1 (\varphi'(x))^2 dx \leq 1\}$$

Dacă $S_0 = \{\varphi : [0,1] \to \mathbb{R} \mid \exists \varphi_0 \in S, \varphi(n) = \varphi_0(\sigma(Q(u))) - F(Q(u))\sqrt{\sigma}\}$ atunci pentru $\varphi_0 \in S_0$ relativ la $K_n(p)$ noi definim

$$\psi_{\varphi_0}(p) = \frac{1}{\mu}\left(K_F(1-p)\int_0^1 \frac{\varphi_0(y)}{f(Q(y))}dy - \int_{1-p}^1 \frac{\varphi_0(y)}{f(Q(y))}dy\right), \; 0 < p < 1.$$

Fie $S_1 = \left\{\psi_{\varphi_0}\right\}_{\varphi_0 \in S_0}$. Atunci avem

Teorema 5.3.3 *(Funcționala LIL)*

Presupunem că sunt îndeplinite (A1)-(A3). Atunci pentru spațiul de probabilitate $(\Omega, \mathscr{F}, P)$ șirul de aplicații $k_n(\cdot) / \sqrt{2\log\log n}$ este P-aproape sigur relativ compact în $C(0,1)$ în raport cu norma supremum. În plus, mulțimea de puncte limită este S_1.

Curba Kakwani

Kakwani (1980) a introdus măsura de inegalitate K_r definită de

$$K_r = 1 - r(r+1)\int_0^1 L_X(p)(1-p)^{r-1}dp$$

ca o generalizare parametrică a indicelui Gini (G). Pentru $\mathrm{r} = 1$ este clar că $K_1 = G$.

Noi vom introduce Curba Kakwani corespunzătoare ca fiind

$$K_r = 1 - r(r+1)L_X(p)(1-p)^{r-1},\ r \geq 1$$

Introducem statisticile Kakwani prin

$$K_{r,n}(p) = 1 - r(r+1)(1-p)^{r-1}\mu_n^{-1}\int_0^p Q_n(u)du$$

Comportamentul asimptotic al statisticii curbei Kakwani

În această parte arătăm consistența uniformă tare și aproximarea Gaussiană tare pentru $K_{r,n}(p)$.

Pornind de la Sen (1984) și *Lema 1* a lui Fakoor (2011) obținem următoarele rezultate.

Teorema 5.5.1 *(Consistența uniformă tare)*

Presupunem că au loc (A1), (A2) și $r \geq 1$ atunci

$$\sup_{0<p<1} |K_{r,n}(p) - K_{r,F}(p)| = O\left(\sqrt{\frac{\log\log n}{n}}\right) \text{ P-a.s.}$$

Teorema 5.5.2 *(Aproximarea Gaussiană tare)*

Dacă presupunem că au loc (A1)-(A3) atunci există un spațiu de probabilitate $(\Omega, \mathcal{F}, P)$ pe care putem construi un proces Gaussian cu doi parametri de medie zero $\{\Gamma(t,n), t \geq 0\}$ și există r>2 astfel încât

i1) $$\sup_{0\leq p\leq 1}\left\{\left|k_{r,n}(1-p) - \frac{1}{\mu}\left(r(r+1)(1-p)^{r-1}G(\mathrm{n}) - (1-K_{r,F}(p))\,\mathrm{H}(\mathrm{n},\mathrm{p})\right)\right|(1-p)^{1-r}\right\} =$$

$$= O\left((n^{-\frac{1}{4}}\log n) \vee (n^{-\lambda})\right) \text{ P-a.s., pentru } 0 < \lambda < \frac{r-2}{2r};$$

Dacă $0 < p_1 < 1$ și λ este definit ca mai sus, atunci

$$\sup_{0\leq p\leq p_1}\left\{\left|k_{r,n}(1-p) - \frac{1}{\mu}\left(r(r+1)(1-p)^{r-1}G(n) - (1-K_{r,F}(p))\,\mathrm{H}(\mathrm{n},\mathrm{p})\right)\right|(1-p)^{1-r}\right\} =$$

$$= O\left((n^{-\frac{1}{4}}\log n) \vee (n^{-\lambda})\right) \text{ P-a.s.}$$

i2) De asemeni putem construi un proces Gaussian $\{\Gamma_n(t), t \geq 0\}$ cu media zero și există r>2 astfel ca

$$\sup_{0\leq p\leq 1}\left\{\left|k_{r,n}(1-p) - \frac{1}{\mu}\left(r(r+1)(1-p)^{r-1}G_n(n) - (1-K_{r,F}(p))\,\mathrm{H}_n(\mathrm{n},\mathrm{p})\right)\right|(1-p)^{1-r}\right\} =$$

$$= O\left((n^{-\frac{1}{4}}\log n) \vee (n^{-\lambda})\right) \text{ pentru orice } \lambda \in \left(0, \frac{r-2}{2r}\right), \text{ și}$$

$$\sup_{0\leq p\leq p_1}\left\{\left|k_{r,n}(1-p) - \frac{1}{\mu}\left(r(r+1)(1-p)^{r-1}G_n(n) - (1-K_{r,F}(p))\,\mathrm{H}_n(\mathrm{n},\mathrm{p})\right)\right|(1-p)^{1-r}\right\} =$$

$$= O\left((n^{-\frac{1}{4}}\log n) \vee (n^{-\lambda})\right) \text{ pentru orice } \lambda \in \left(0, \frac{r-2}{2r}\right), \text{ cu } 0 < p_1 < 1.$$

Fie $S_1 = \left\{\hat{\psi}_{\varphi_0}\right\}_{\varphi_0 \in S_0}$. Atunci avem

Teorema 5.3 *(Funcționala LIL)* Dacă (A1)-(A3) au loc atunci există un spațiu de probabilitate $(\Omega, \mathscr{F}, P)$ astfel încât șirul de funcții $-\dfrac{k_{r,n}(t)(1-t)^{1-r}}{r(r+1)\sqrt{2\log\log n}}, t \in (0,1)$, pentru $r > 1$ este P-aproape sigur relativ compact în $C(0,1)$ relativ la norma supremum. În plus, mulțimea de puncte limită este S_1.

Comportamentul asimptotic al indicelui Kakwani

Teorema 5.4.
Dacă (A1) și (A2) au loc atunci

$$|K_{r,n} - K_{r,F}| = O\left(\sqrt{\frac{\log\log n}{n}}\right) \text{ P-a.s.}$$

Concluzii

După un secol de analiză intensă a familiilor de curbe Lorenz în domeniul veniturilor și a distribuțiilor de avere, a fost introdusă recent curba Leimkuhler pentru o potrivire corectă a datelor din partea superioară a valorilor observate. Am definit rezultate asimptotice pentru estimatori neparametrici ai curbei Leimkuhler și pentru indeicele Kakwani. În secțiunea 5.3 și 5.5 am obținut aproximarea lor Gaussiană tare pentru un eșantion care corespunde unei distribuții dependentă de lungime. S-a determinat o lege a procesului logaritmului iterat pentru acești estimatori.

VI
FAMILII PARAMETRICE NOI DE CURBE LEIMKUHLER

Introducere

Definiția curbei Leimkuhler propusă de Sarabia (2008) a fost punctul de plecare al acestui capitol. Utilizând această abordare recentă propunem o nouă familie de Curbe Leimkuhler parametrice. Vom analiza o mare varietate de proprietăți, inclusiv măsuri ale inegalității.

Curba Leimkuhler

Familiile de curbe Lorenz pot modela corect valorile mici observate, dar uneori incorect datele din partea superioară. În informetrică, de exemplu, studiile analizează sursele cele mai productive.

Curbe Leimkuhler mixte

Fie $L_X(p,\theta,\nu)$ o curbă Lorenz parametrică, $\pi(\theta,\alpha,\lambda)$ densitatea de repartiție Gamma pe $(0,\infty)$ și $K(p,\theta,\nu)$ o curbă Leimkuhler parametrică.

Definim **curba Leimkuhler mixtă** prin $\tilde{K}(p,\theta,\nu,\lambda) = \mathrm{E}_{\pi(\theta,\alpha,\lambda)}(K(p,\theta,\nu))$, unde

$$\pi(\theta,\alpha,\lambda) = \frac{\lambda^a}{\Gamma(\alpha)}\theta^{\alpha-1}e^{-\lambda\theta}I(\theta > 0).$$

Teorema 6.1.

Fie $\tilde{K}(p,\theta,\nu,\lambda)$ o curbă Leimkuhler mixtă. Atunci are loc relația:

$$\tilde{K}(p,\theta,\nu,\lambda)=1-\tilde{\mathrm{L}}_X(1-\mathrm{p},\alpha,\lambda,\nu)$$

Teorema 6.2.

Dacă $L_X(p,\theta)$ este curba Lorenz exponențială atunci funcția corespunzătoare **curbei Leimkuhler Gamma-Exponențială** este

$$\tilde{K}(p,\alpha,\lambda)=1-\frac{1-p}{\left[1-\lambda^{-1}\ln(1-p)\right]^{\alpha}}$$

În strânsă legătură cu capitolul cinci în **capitolul al șaselea** am dezvoltat familii parametrice noi de Curbe Leimkuhler, precum și Curbe Leimkuhler mixte, având la bază articolul [8].

VII
APROXIMARE DE ORDIN SUPERIOR PENTRU DIVERGENTE GENERALIZATE

Noțiuni preliminare

Problema găsirii și estimării distanței corecte (diferență sau discriminare) între două densități de probabilitate este printre cele mai importante în Teoria Probabilităților. φ-divergența Csiszar are aplicații în antropologie, genetică, finanțe, economie, științe politice, biologie și procesarea semnalelor.

Aproximări de ordin superior
Aproximări folosind formula lui Taylor

Definim (φ,a)**-divergența generalizată** dintre măsurile de probabilitate ξ_1,ξ_2 ca fiind

$$\Gamma_{\varphi,a}(\xi_1,\xi_2)=\int_X \mathrm{g}^a(x)\varphi\left(\frac{\mathrm{f}(x)}{\mathrm{g}(x)}\right)d\mu(x).$$

Pentru $a=1$ obținem φ-divergența Csiszar din (7.1).

Teorema 7.1.

Fie $0<\alpha<1<\beta$, funcția φ conform ipotezelor din paragraful anterior, și $\varphi\in C^n([\alpha,\beta]),n\geq 1$ cu $|\varphi^{(n)}(t)-\varphi^{(n)}(1)|$ o funcție convexă în t. Fie $0<y<min\left(1-\alpha,\beta-1\right)$ fixat. Atunci

$$\Gamma_{\varphi,a}(\xi_1,\xi_2)\leq\sum_{j=1}^{n}\frac{|\varphi^{(\mathrm{j})}(1)|}{\mathrm{j}!}\left|\int_X g^{a-j}(\mathrm{x})(\mathrm{f}(\mathrm{x})-g(x))^j\,d\mu\right|$$

$$+\frac{\omega(\varphi^{(n)},\mathrm{y})}{(n+1)!\mathrm{y}}\int_X g^{a-1-n}(\mathrm{x})\,|\,\mathrm{f}(\mathrm{x})-g(x)\,|^{n+1}\,d\mu.$$

Dacă $\varphi^{(\mathrm{j})}(1)=0,\ \mathrm{j}=1,\dots,n,$ atunci

$$\Gamma_{\varphi,a}(\xi_1,\xi_2)\leq\frac{\omega(\varphi^{(n)},\mathrm{y})}{(n+1)!\mathrm{y}}\int_X g^{a-1-n}(\mathrm{x})\,|\,\mathrm{f}(\mathrm{x})-g(x)\,|^{n+1}\,d\mu\,,$$

unde $\omega(\varphi^{(n)}, y) := \sup_x |\Delta_y(\varphi^{(n)}; x)|$ este modulul de continuitate de ordin *1*, iar $\Delta_y(\varphi^{(n)}; x) := \varphi^{(n)}(x+y) - \varphi^{(n)}(x)$ este diferența finită de ordinul *I*. Pentru *n* par sunt obținute prin

$$\tilde{\varphi}(t) := \frac{|t-1|^{n+1}}{(n+1)!}, \alpha \le t \le \beta$$

Teorema 7.2

Fie $0 < \alpha < 1 < \beta$, funcția φ conform ipotezelor din paragraful anterior, și $\varphi \in C^n([\alpha, \beta]), n \ge 1$. Presupunem $\omega(\varphi^{(n)}, \tau) \le \omega_0$, unde $0 < \tau \le \beta - \alpha, \omega > 0$. Fie $x \in \mathbb{R}$ și notăm prin

$$\chi_n(x) := \int_0^{|x|} \left\lceil \frac{p}{\tau} \right\rceil \frac{(|x| - p)^{n-1}}{(n-1)!} dp$$

unde $\lceil m \rceil$ este partea întreagă superioară a numărului m. Atunci

$$\Gamma_{\varphi,a}(\xi_1, \xi_2) \le \sum_{j=1}^{n} \frac{|\varphi^{(j)}(1)|}{j!} \left| \int_X g^{a-j}(x)(f(x) - g(x))^j d\mu \right| + \omega_0 \int_X g^a(x) \chi_n\left(\frac{f(x) - g(x)}{g(x)} \right) d\mu$$

Egalitatea în (7.20) este obținută pentru funcția

$$\tilde{\varphi}_n(t) := \omega_0 \chi_n(p-1), \alpha \le p \le \beta,$$

atunci când n este par.

Corolarul 7.1.

Avem pentru $0 < \tau \le \beta - \alpha$

$$\begin{aligned}\Gamma_{\varphi,a}(\xi_1, \xi_2) \le & \sum_{j=1}^{n} \frac{|\varphi^{(j)}(1)|}{j!} \left| \int_X g^{a-j}(x)(f(x) - g(x))^j d\mu \right| + \\ & + \omega(\varphi^{(n)}, \tau)\{ \frac{1}{(n+1)!\tau} \int_X g^{a-1-n}(x) |f(x) - g(x)|^{n+1} d\mu + \\ & + \frac{1}{2n!} \int_X g^{a-n}(x) |f(x) - g(x)|^n d\mu + \\ & \frac{\tau}{8(n-1)!} \int_X g^{a+1-n}(x) |f(x) - g(x)|^{n-1} d\mu \}\end{aligned}$$

și

$$\begin{aligned}\Gamma_{\varphi,a}(\xi_1, \xi_2) \le & \sum_{j=1}^{n} \frac{|\varphi^{(j)}(1)|}{j!} \left| \int_X g^{a-j}(x)(f(x) - g(x))^j d\mu \right| + \\ & + \omega(\varphi^{(n)}, \tau)\{ \frac{1}{n!} \int_X g^{a-n}(x) |f(x) - g(x)|^n d\mu + \\ & + \frac{1}{n!(n+1)\tau} \int_X g^{a-1-n}(x) |f(x) - g(x)|^{n+1} d\mu \}\end{aligned}$$

În particular vom avea

$$\begin{aligned}\Gamma_{\varphi,a}(\xi_1, \xi_2) \le & \varphi'(1) \left| \int_X g^{a-1}(x)(f(x) - g(x)) d\mu \right| + \\ & + \omega(\varphi^{(n)}, \tau)\{ \frac{1}{2\tau} \int_X g^{a-2}(x) |f(x) - g(x)|^2 d\mu + \\ & + \frac{1}{2} \int_X g^{a-1}(x) |f(x) - g(x)| d\mu \quad + \frac{\delta}{8} \}\end{aligned}$$

și

$$\Gamma_{\varphi,a}(\xi_1,\xi_2)\le\varphi'(1)\left|\int_X g^{a-1}(\mathrm{x})(\mathrm{f}(\mathrm{x})-g(x))d\mu\right|+$$
$$+\omega(\varphi',\tau)\{\int_X g^{a-1}(\mathrm{x})\left|\mathrm{f}(\mathrm{x})-g(x)\right|d\mu+$$
$$+\frac{1}{2\tau}\int_X g^{a-2}(\mathrm{x})\left(\mathrm{f}(\mathrm{x})-g(x)\right)^2 d\mu\}$$

Rezultate bazate pe formula generalizată Taylor-Widder

Teorema 7.3

Pentru toți $x\in[\alpha,\beta]$ și $n\geq 0$ avem

$$\Gamma_{\varphi,a}(\xi_1,\xi_2)\le\sum_{j=1}^{n}\left|B_j\varphi(1)\right|\int_X g^a(x)h_i\left(\frac{f(x)}{g(x)},1\right)d\mu+\left|B_{n+1}\varphi(1)\right|\left|\int_X g^a(x)\,\mathrm{N}_n\left(\frac{f(x)}{g(x)},1\right)d\mu\right|+$$
$$+\omega_0\int_X\left(g^a(x)+\frac{\mathrm{g}^{a-1}(x)\left|f(x)-g(x)\right|}{y}\right)\left|\mathrm{N}_n\left(\frac{f(x)}{g(x)},1\right)\right|d\mu\,.$$

Teorema 7.4.

Pentru toți $x\in[\alpha,\beta]$ și $n\geq 0$ avem

$$\Gamma_{\varphi,a}(\xi_1,\xi_2)\le\sum_{j=1}^{n}\left|B_j\varphi(1)\right|\int_X g^a(x)h_i\left(\frac{f(x)}{g(x)},1\right)d\mu+\left|B_{n+1}\varphi(1)\right|\left|\int_X g^a(x)\,\mathrm{N}_n\left(\frac{f(x)}{g(x)},1\right)d\mu\right|+$$
$$+A\int_X\left(g^{a-\alpha}(x)\left|f(x)-g(x)\right|\right)\left|\mathrm{N}_n\left(\frac{f(x)}{g(x)},1\right)\right|d\mu\,.$$

Teorema 7.5

Pentru toți $x\in[\alpha,\beta]$ și $n\geq 0$ avem

$$\Gamma_{\varphi,a}(\xi_1,\xi_2)\le\sum_{j=1}^{n}\left|B_j\varphi(1)\right|\int_X g^a(x)h_i\left(\frac{f(x)}{g(x)},1\right)d\mu+\left|B_{n+1}\varphi(1)\right|\left|\int_X g^a(x)\,\mathrm{N}_n\left(\frac{f(x)}{g(x)},1\right)d\mu\right|+$$
$$+\frac{\omega_0}{p}\int_X\left(g^{a-1}(x)\left|f(x)-g(x)\right|\right)\left|\mathrm{N}_n\left(\frac{f(x)}{g(x)},1\right)\right|d\mu\,.$$

Teorema 7.6

Fie $p=\int_X\left(g^{a-1}(x)\left|f(x)-g(x)\right|\right)\left|\mathrm{N}_n\left(\frac{f(x)}{g(x)},1\right)\right|d\mu$ astfel încât $0<p\le\min(1-a,b-1)$.

Atunci

$$\Gamma_{\varphi,a}(\xi_1,\xi_2)\le\sum_{j=1}^{n}\left|B_j\varphi(1)\right|\int_X g^a(x)h_i\left(\frac{f(x)}{g(x)},1\right)d\mu+\left|B_{n+1}\varphi(1)\right|\left|\int_X g^a(x)\,\mathrm{N}_n\left(\frac{f(x)}{g(x)},1\right)d\mu\right|+$$
$$+\omega\left(B_{n+1}\varphi;\int_X\left(g^{a-1}(x)\left|f(x)-g(x)\right|\right)\left|\mathrm{N}_n\left(\frac{f(x)}{g(x)},1\right)\right|d\mu\right).$$

Teorema 7.7

Pentru toți $x\in[\alpha,\beta]$ și $n\geq 0$ avem

$$\Gamma_{\varphi,a}(\xi_1,\xi_2)\le\sum_{j=1}^{n}\left|B_j\varphi(1)\right|\int_X g^a(x)h_i\left(\frac{f(x)}{g(x)},1\right)d\mu+$$

$$+\left\|B_{n+1}\varphi\right\|_{\infty,[\alpha,\beta]}\int_X g^{a}(x)\left|\mathrm{N}_n^*\left(\frac{f(x)}{g(x)},1\right)\right|d\mu$$

Aproximarea de ordin superior pentru divergențe generalizate folosind curbe Lorenz
Cazul AB-divergenței

Fie f,g două densități de repartiție. Atunci AB – divergența definită recent de Cichocki (2011) între f și g este $D_{\alpha,\beta}(f\,\|\,g)$

$$D_{\alpha,\beta}(f\,\|\,g)=-\frac{1}{\alpha\beta}\int_X (f^{\alpha}(x)g^{\beta}(x)-\frac{\alpha}{\alpha+\beta}f^{\alpha+\beta}(x)-\frac{\beta}{\alpha+\beta}g^{\alpha+\beta}(x))dx,$$

unde $\alpha,\beta>0$ cu $\alpha+\beta\neq 0$.

Prin ajustarea parametrilor α și β Chichocki (2011) a arătat că pot fi obținute un număr mare de divergențe noi sau standard. Astfel divergența este exprimată într-o formă explicită ca fiind:

$$D_{\alpha,\beta}(f\,\|\,g)=\int_X d_{\alpha,\beta}(f(x),g(x))dx,$$

unde

$$d_{\alpha,\beta}(f(x),g(x))=\begin{cases}-\dfrac{1}{\alpha\beta}\left(f^{\alpha}(x)g^{\beta}(x)-\dfrac{\alpha}{\alpha+\beta}f^{\alpha+\beta}(x)-\dfrac{\beta}{\alpha+\beta}g^{\alpha+\beta}(x)\right);\alpha,\beta,\alpha+\beta\neq 0\\ \dfrac{1}{\alpha^2}\left(f^{\alpha}(\mathrm{x})\ln\left(\dfrac{f(x)}{g(x)}\right)^{\alpha}-f^{\alpha}(x)+g^{\alpha}(x)\right);\alpha\neq 0,\beta=0\\ \dfrac{1}{\alpha^2}\left(\ln\left(\dfrac{g(x)}{f(x)}\right)^{\alpha}+\left(\dfrac{g(x)}{f(x)}\right)^{-\alpha}-1\right);\alpha=-\beta\neq 0\\ \dfrac{1}{\beta^2}\left(g^{\beta}(\mathrm{x})\ln\left(\dfrac{g(x)}{f(x)}\right)^{\beta}-g^{\beta}(x)+f^{\beta}(x)\right);\alpha=0,\beta\neq 0\\ \dfrac{1}{2}\left(\ln f(x)-\ln g(x)\right)^2;\alpha,\beta=0\end{cases}$$

Astfel rezultă cazurile ce conduc la divergențe importante: Alpha-divergența pentru $\alpha+\beta=1$ (Amari (2007); Cichocki (2011)); Beta-divergența atunci când $\alpha=1$ (Basu (1998), Mihoko (2002)) și divergența Kullback-Leibler pentru $\alpha=1$ și $\beta=0$.

O proprietate importantă a divergenței este aceea că $D_{\alpha,\beta}(f\,\|\,g)$ este nenegativă pentru toți f și g și este nulă dacă și numai dacă $f\equiv g$ aproape peste tot (Cichocki (2011)).

În continuare, rescriem AB-divergența pentru $\alpha+\beta=a$

$$D_{\alpha,\beta}(f\,\|\,g)=\int_x d_{\alpha,\beta}(f(x),g(x))dx=\int_x \varphi(t)dt$$

Astfel considerăm

$$\varphi(t)=-\frac{1}{\alpha\beta}t^{\alpha}+\frac{1}{\beta(\alpha+\beta)}t^{\alpha+\beta}+\frac{1}{\alpha(\alpha+\beta)}$$

Derivata de ordin *k* va fi de forma

$$\varphi^{(k)}(t)=\frac{1}{\beta}\Big[-(\alpha-1)(\alpha-2)...(\alpha-k)t^{\alpha-k}+(a-1)...(a-k)t^{a-k}\Big]$$

Observăm că $\varphi^{(k)}(1)=0, k=1,...,n$.

Teorema 7.4.3.

Fie $0<\alpha<1<\beta$, funcția φ de forma (7.4.6) și $0<y<min(1-\alpha,\beta-1)$ fixat. Atunci, sub ipotezele Teoremei 7.4.2, avem

$$\Gamma_{\varphi,a}(\xi_1,\xi_2)\le\frac{\omega(\varphi^{(n)},\mathrm{y})}{(n+1)!\mathrm{y}}\int_X g^{a-1-n}(\mathrm{x})\,|\,\mathrm{f(x)}-g(x)\,|^{n+1}\,d\mu\,,$$

unde $\omega(\varphi^{(n)},\mathrm{y}):=\sup_x|\Delta_y(\varphi^{(n)};x)|$ este modulul de continuitate de ordin *1*, iar $\Delta_y(\varphi^{(n)};x):=\varphi^{(n)}(x+y)-\varphi^{(n)}(x)$ este diferența finită de ordinul *I*.

Rezultate de aproximare pentru divergențe generalizate folosind Curbe Lorenz

Teorema 7.4.4.

Fie $0<\alpha<1<\beta$, funcția φ conform ipotezelor din paragraful 7.1, și $\varphi\in C^n([\alpha,\beta]),n\ge1$ cu $|\,\varphi^{(n)}(t)-\varphi^{(n)}(1)\,|$ o funcție convexă în t. Fie $0<y<min(1-\alpha,\beta-1)$ fixat. Atunci

$$\Gamma_{\varphi,a}(\xi_1,\xi_2)\le\sum_{j=1}^{n}\frac{|\,\varphi^{(\mathrm{j})}(1)\,|}{\mathrm{j}!}\left|\int_X\sum_{k=0}^{j}C_j^k\mu_F^{j-k}\mu_G^{j-a-k}\left[L_F''(\mathrm{F}(x))\right]^{k-j}\left[L_G''(\mathrm{G}(x))\right]^{j-a-k}d\mu\right|$$

$$+\frac{\omega(\varphi^{(n)},\mathrm{y})}{(n+1)!\mathrm{y}}\int_X\sum_{k=0}^{n+1}C_j^k\mu_F^{n+1-k}\mu_G^{n+1-a-k}\left[L_F''(\mathrm{F}(x))\right]^{k-n-1}\left[L_G''(\mathrm{G}(x))\right]^{n+1-a-k}d\mu\,.$$

Dacă $\varphi^{(\mathrm{j})}(1)=0,\ \mathrm{j}=1,...,n,$ atunci

$$\Gamma_{\varphi,a}(\xi_1,\xi_2)\le\frac{\omega(\varphi^{(n)},\mathrm{y})}{(n+1)!\mathrm{y}}\int_X\sum_{k=0}^{n+1}C_j^k\mu_F^{n+1-k}\mu_G^{n+1-a-k}\left[L_F''(\mathrm{F}(x))\right]^{k-n-1}\left[L_G''(\mathrm{G}(x))\right]^{n+1-a-k}d\mu\,,$$

unde $\omega(\varphi^{(n)},\mathrm{y}):=\sup_x|\Delta_y(\varphi^{(n)};x)|$ este modulul de continuitate de ordin *1*, iar $\Delta_y(\varphi^{(n)};x):=\varphi^{(n)}(x+y)-\varphi^{(n)}(x)$ este diferența finită de ordinul *I*.

Teorema 7.4.5.

Fie $0<\alpha<1<\beta$, funcția φ conform ipotezelor din paragraful 7.1, și $\varphi\in C^n([\alpha,\beta]),n\ge1$ cu $|\,\varphi^{(n)}(t)-\varphi^{(n)}(1)\,|$ o funcție convexă în t.

Atunci

$$\Gamma_{\varphi,a}(\xi_1,\xi_2)\le\sum_{j=1}^{n}|\,\mathrm{B}_i\,\varphi(1)\,|\left|\int_X\mu_G^{-a}\left[L_G''(\mathrm{G}(x))\right]^{-a}h_j\left(\mu_G\mu_F^{-1}\frac{L_G''(G(x))}{L_F''(\mathrm{F}(x))},1\right)d\mu\right|$$

$$+\left\|\mathrm{B}_{n+1}\,\varphi\right\|_{\infty,[\alpha,\beta]}\int_X\mu_G^{-a}\left|N_n^*\left(\mu_G\mu_F^{-1}\frac{L_G''(G(x))}{L_F''(\mathrm{F}(x))},1\right)\right|\left[L_G''(\mathrm{G}(x))\right]^{-a}d\mu$$

În acest ultim capitol am prezentat bazându-ne pe [5], [12] și [13] aproximări de ordin superior pentru divergențe generalizate. După realizarea unei sinteze asupra teoremelor Arnold și Thompson am aplicat formula lui Taylor în cazul AB-divergenței în vederea obținerii unor rezultate de aproximare pentru divergențe generalizate folosind curbe Lorenz.

Concluzii, contribuții personale și direcții viitoare de cercetare

§ 1. Concluzii

Din multitudinea perspectivelor prin care pot fi abordate Curbele Lorenz, ne-am concentrat atenția asupra extinderii claselor de Curbe Lorenz parametrice și aproximarea de ordin superior pentru divergențe generalizate. După o analiză riguroasă a literaturii de specialitate ne-am propus prin lucrarea de față crearea unor familii noi de curbe Lorenz și realizarea de noi conexiuni ale acestora cu mecanismul Marshall-Olkin, entropia Tsallis, curbe Leimkuhler și divergențe generalizate.

De obicei datele pe care le analizăm sunt disponibile sub formă de venit, caz în care coeficientul lui Gini poate fi estimat prin aproximarea CL ca o serie de segmente. Pentru a stabili gradul de inegalitate nu este suficientă o asemenea abordare deoarece aceasta nu ține cont de inegalitatea în cadrul claselor de venit. Ca o metodă alternativă, în primul rând găsim din datele pe grupuri o funcție pentru Curba Lorenz, și apoi calculăm indicele lui Gini ca fiind de două ori aria dintre CL estimată și prima bisectoare. Plecând de la acest raționament au fost introduse de-a lungul timpului diverse funcții pentru curbele Lorenz. De exemplu, Kakwani (1980), Rasche (1980), Ortega (1991), Sarabia (1999), etc…

În această lucrare am abordat multiple modalități de a obține o curbă Lorenz. Ca o primă metodă am definit funcții noi de repartiție de tip exponențializat (§2.2) și apoi am determinat curba Lorenz în funcție de aceaste repartiții, iar ca o altă metodă am generat direct funcția curbei Lorenz prin intermediul mecanismului Mrashall-Olkin (§2.4) și prin compunerea cu repartiția Gamma (§3.3).

Mai mult, colateral cu aceste obiective s-au realizat și legăturile inter-disciplinare ale statisticii matematice cu teoria informației (Curba Leimkuhler), analiză matematică (dezvoltări de tip Taylor), teoria măsurii (divergențe) și econometrie. Pentru atingerea acestor obiective, cercetarea întreprinsă a avut trei coordonate: o cercetare bibliografică, una analitică și una de tip empiric.

§2. Contribuții personale

Cercetarea a atins obiectivele enunțate în introducere, și anume în primul capitol am evidențiat importanța curbei Lorenz și a indicelui lui Gini în contextul dependenței acestuia de CL, am dezvoltat în al doilea capitol noi clase de Curbe Lorenz, prin extinderea familiilor consacrate de curbe Lorenz, iar în al treilea capitol am modelat Curbe Lorenz mixte pentru familii parametrice de CL prin compunerea acestora cu repartiția Gamma.

În continuare, în capitolul al patrulea, am studiat intervalele de încredere ale curbei Lorenz Generalizată.

În al cincilea capitol am abordat proprietățile asimptotice ale estimatorilor non-parametrici ai curbei Leimkuhler si Indicelui Kakwani.

În strânsă legătură cu secțiunea anterioară capitolul al șaselea am dezvoltat familii parametrice noi de Curbe Leimkuhler, precum și Curbe Leimkuhler mixte.

În ultimul capitol am prezentat aproximări de ordin superior pentru divergențe generalizate. După realizarea unei sinteze asupra teoremelor Arnold și Thompson am aplicat formula lui Taylor în cazul AB-divergenței în vederea obținerii unor rezultate de aproximare pentru divergențe generalizate folosind curbe Lorenz.

§3. Limitele cercetării

Oricât de complexă și fundamentată ar fi o cercetare științifică, pe lângă rezultatele obținute și contribuțiile semnificative aduse, există și unele limite care sunt inerente, dar care oferă posibilitatea continuării și aprofundării domeniului studiat.

§4. Perspective de cercetare

Importanța curbelor Lorenz în analiza economică și statistică a inegalității veniturilor și bogăției motivează dorința de a găsi noi familii parametrice de curbe Lorenz. Multitudinea de modele parametrice propuse în literatura de specialitate nu este un inconvenient, dar un motiv în plus dat de neconcordanța totală a curbelor empirice pe datele de venituri stabilite. Concluzia noastră este că aproximarea parametrică mixtă aduce o concordanță mai mare prin introducerea de restricții stricte.

Prin utilizarea unor instrumente statistice corespunzătoare ne propunem pe viitor efectuarea de comparații între noile curbe mixte și cele clasice propuse de Sarabia (2005). Estimările parametrilor modelelor pot fi determinate, de exemplu, prin utilizarea metodei celor mai mici pătrate neliniară.

§5. Etica cercetării

Lucrarea de față a urmărit liniile generale ale eticii cercetării științifice, prin utilizarea citărilor în conținutul său și alcătuirea unei bibliografii corespunzătoare.

Lucrări publicate

[1] **Gheorghe** Carmen-Adriana; Cipu Corina Elena **(2007)** *Mortality modeling for Romanian population*. Proceedings of The 4-th International Colloquium "Mathematics in Engineering and Numerical Physics" October 6-8 , 2006, Bucharest, Romania, pp. 38-45. Balkan Society of Geometers, Geometry Balkan Press 2007.

[2] **Gheorghe** Carmen-Adriana; Cipu Corina Elena **(2007)** *The Hill estimator for income of Romanian households*, Proceedings of International Conference Trends and Challenges in Applied Mathematics, ISBN 978-973-755-283-9/pbk, Ed. Matrix Rom, Bucuresti, p. 193- 196.

[3] Cipu Corina Elena; **Gheorghe** Carmen-Adriana **(2007)** *Life tables for Romania. Survival function. Life expectancy for Romanian people,* U.P.B, Scientific Bulletin, Series A, vol. 69, no. 1, Bucuresti, p. 37-36.

[4] **Gheorghe** Carmen-Adriana **(2009)** *Estimation of weighted maximum entropy densities within the study upon Lorenz Curves for grouped data*, Proceedings of 16th European Young Statisticians Meetings, ASE, Bucuresti.

[5] **Gheorghe** Carmen-Adriana **(2009)** *Approximation of Csiszar's f-Divergence using Kullback-Leibler Distance*, Proceedings of 9th Balkan Conference on Operational Reserch, Constanta, Romania, *http://civile.utcb.ro/balcor*.

[6] **Gheorghe** Carmen-Adriana; Preda Vasile; Cipu Corina Elena **(2012)** *The estimate of generalized Lorenz confidence intervals using Exponentiated Weibull Distribution*, Conferința 15 SPSR, Universitatea din București

[7] Drăgulin Mircea; **Gheorghe** Carmen-Adriana **(2013)** *Mixture Lorenz Curves. Three new models*, Proceedings of 18th European Young Statisticians Meetings, Osjek, Croatia.

[8] **Gheorghe** Carmen-Adriana **(2013)** *New parametric families of Leimkuhler Curves*, Conferința 16 SPSR, ASE, Bucuresti.

[9] **Gheorghe** Carmen-Adriana **(2013)** *Asymptotic properties of the nonparametric estimators of the Leimkuhler curve and Kakwani index*, U.P.B, Scientific Bulletin, Bucuresti. In curs de publicare.

[10] Preda, Vasile; **Gheorghe** Carmen-Adriana **(2013)** *Maximum Tsallis entropy and the Lorenz curve.* În proces de publicare.

[11] Preda, Vasile; **Gheorghe** Carmen-Adriana **(2013)** *A class of Lorenz Curves generated by the Marshall Olkin technique.* În proces de publicare.

[12] Preda, Vasile; **Gheorghe** Carmen-Adriana **(2013)** *Some approximations for generalised divergences.* În proces de publicare.

[13] Preda, Vasile; **Gheorghe** Carmen-Adriana **(2013)** *Bounds for some generalised divergences in trems of Lorenz curves.* În proces de publicare.

[14] Cipu, Corina Elena; **Gheorghe** Carmen Adriana **(2013)** *Different measurement tools of the income inequality among the Romanian households and their limit.* Romanian Journal of Economic Forecasting, Romania. Trimis spre publicare.

[15] Cipu, Corina Elena; **Gheorghe** Carmen Adriana **(2013)** *Analyzing the Romanian income inequality in the European context.* Studii Economice, National Institute of Economic Research. Trimis spre publicare.

Referințe bibliografice

(1) Ali, S.M.; Silvey, S.D. (1996). A general class of coefficients of divergence of one distribution from another. J. Roy. Stat. Soc.

(2) Amari S.I. (2007). Integration of stochastic models by minimizing α-divergence. Neural Comput., 19, 2780–2796.

(3) Anastassiou, G.A., (1993) Moments in Probability and Approximation Theory, Pitman, England

(4) Anastassiou, G.A., (2001) Taylor integral remainders and moduli of smoothness, Korean Proceedings,Vol.1, Inequalities Theory and Applications, Eds.Y.Cho,J.K.Kim,and S.Dragomir, pp.1-31, Nova Publ., NY.

(5) Anastassiou, G.A., (2005) Higher order optimal approximation of Csiszar's f-divergence, Nonlinear Analysis, 61, 309-339.

(6) Arnold, B. C. (1987a). Bivariate distributions with Pareto conditionals, *Statistics & ProbabilityLetters, 5, 263-266*

(7) Arnold, B. C. (1987b) . Majorization and the Lorenz Order: A Brief Introduction, Springer-Verlag, Berlin

(8) Arnold, B. C. ; P. L. Brockett; C. A. Robertson; B. Y. Shu (1987c). Generating ordered families of Lorenz curves by strongly unimodal distributions, J. Bus. Econ. Statist. 5, 305-308

(9) Arora, S. ; Jain, K. (2005 a). Testing for Generalised Lorenz Dominance , Statistical Methods and Applications.

(10) Arora, S., Jain, K. ; Pundir, S. (2005 b). On Cumulated Mean Income Curve, Model Assisted Statistics and Applications.

(11) Asimit, Alexandru V. and Furman, Edward and Vernic, Raluca (2009) On a Multivariate Pareto Distribution. Insurance: Mathematics and Economics, Vol. 46, No. 2.

(12) Atkinson, A.B. (1970) On the measurement of inequality, *J. Economic Theory 2, 244-263.*

(13) Basu, A.; Harris, I.R.; Hjort, N.L.; Jones, M. (1998) Robust and efficient estimation by minimising a density power divergence. Biometrika, 85, 549–559.

(14) Basseville, M. (1989). Distance measures for signal processing and pattern recognition. Signal Processing.

(15) Beach, C., Davidson, R., (1983). Distribution-free statistical inference with Lorenz curves and income shares. Review of Economic Studies 50, 723–735.

(16) Beach, C., Richmond, J., (1985). Joint confidence intervals for income shares and Lorenz curves. International Economic Review 26, 439–450.

(17) Bennett, N.G.; S. Horiuchi. (1981). Estimating the completeness of death registration in a closed population. Population Index, 47, 207-21.

(18) Birkhoff, G. ; S. MacLane. (1959). A survey of modern algebra (rev. ed). New York: Macmillan.

(19) Bishop, J. A. Chakraborti, S., Thistle, P. D. (1989). Asymptotically distribution free statistical inference for Generalised Lorenz Curve , Review of Economic Statistics, 71, 725-727.

(20) Bongaarts, J. ; R.A. Bulatao. (1999). Completing the demographic transition. Population and Development Review, 25, 515-29.

(21) Burnham, K. P., Anderson D. R. (2002). Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach, Second Edition. Springer Science, New York

(22) Caswell, H. (2001). Matrix population models (2d ed). Sunderland MA: Sinauer.

(23) Cencov, N.N.(1982). StatisticalDecision Rules and Optimal Inference, volume 14 of Translations in Mathematics. American Mathematical Society, Providence, RI.

(24) Chakraborti, S., (1994). Asymptotic distribution-free joint confidence intervals for Generalised Lorenz Curves based on Complete data . Statistics and Probability Letters, 21, 229-235.

(25) Chernoff, H. (1952). Measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. Ann. Math. Stat..

(26) Chernoff, H. (1956). Large-sample theory: Parametric case. Ann. Math. Stat.

(27) Chong, K. M. (1974), Some extensions of a theorem of Hardy, Littlewood and Polya and their applications, Can. J. Math. 26, 1321-1340.

(28) Cipu Corina Elena; **Gheorghe Carmen-Adriana** (2007) Life tables for Romania. Survival function. Life expectancy for Romanian people, U.P.B, Scientific Bulletin, Series A, vol. 69, no. 1, Bucuresti, p. 37-36.

(29) Cipu, Corina Elena; **Gheorghe Carmen Adriana** (2013a) Different measurement tools of the income inequality among the households and their limit. Romanian Journal of Economic Forecasting, Romania. Trimis spre publicare.

(30) Cipu, Corina Elena; **Gheorghe Carmen Adriana** (2013b) Analyzing the Romanian income inequality in the European context. Studii Economice, National Institute of Economic Research. Trimis spre publicare.

(31) Cichocki, A.; Cruces, S.; Amari, S.I. (2011) Generalized alpha-beta divergences and their application to robust nonnegative matrix factorization. Entropy, 13, 134–170.

(32) Ciucu, G., Craiu, V., Săcuiu I. (1974) Probleme de statistică matematică, Editura tehnică

(33) Coale, A.J. (1972). The growth and structure of human populations. Princeton: Princeton University Press.

(34) Cordeiro, G., Lemonte, A. (2013) On the Marshall–Olkin extended Weibull distribution Statistical Papers, vol. 54, issue 2, pages 333-353

(35) Cover, T.M.; Thomas, J. A. (1991). Elements of Information Theory. John Wiley & Sons, Inc.

(36) Craiu, M. (2006) Statistica matematica. Teorie si probleme, Ed. MatrixRom

(37) Craiu, V. (1972) Verificarea ipotezelor statistice, Editura didactică și pedagogică

(38) Dasgupta, P.; Sen, A. K.; Starrett, D. (1973), Notes on the measurement of inequality, J. Econ Theory 180-187.

(39) David, H., 1981. Order Statistics, 2nd Edition. Wiley, New York.

(40) Davidson, R., Duclos, J.-Y. (1997). Statistical inference for the measurement of the incidence of taxes andtransfers. Econometrica 65, 1453–1465.

(41) de Haan, L., Peng, L. (1999). Exact rates of convergence to a stable law. Journal of the London Mathematical Society 59, 1134–1152.

(42) Drăgulin Mircea; **Gheorghe Carmen-Adriana** (2013) Mixture Lorenz Curves. Three new models, Proceedings of 18th European Young Statisticians Meetings, Osjek, Croatia.

(43) Drees, H., de Haan, L., Resnick, S., (2000). How to make a Hill plot. Annals of Statistics 28, 254–274.

(44) Dumitrescu M. (2000) Survey Sampling and Applications, Editura Tehnica, Bucuresti

(45) Dumitrescu, Monica, Bătătorescu, Anton (2006) Applied Statistics using the R-Sytem, Editura Universitatii din Bucuresti

(46) Economou P, Caroni C (2007) Parametric proportional odds frailty models. Commun Stat Simul Comput 36:579–592

(47) Einmahl, U. (2007). A generalization of Strassen's functional LIL. J Theor Probab (2007) 20: 901–915

(48) Embrechts, P., Kl W uppelberg, C., Mikosch, T., (1997). Modelling Extremal Events. Springer, Berlin.
(49) Fakoor, V., Ghalibaf, M., Azarnoosh, H.A. (2011). Asymptotic behaviors of the Lorenz curve and Gini index in sampling from length-biased distribution. Statistics and probability letters 81, 1425-1435.
(50) Fraundorf, P. (2007). Thermal roots of correlation-based complexity
(51) Gail, M. H.; Gastwirth, J. L. (1971), A scale-free goodness-of-fit test for the exponențial distribution based on the Lorenz curve, J. Amer. Statist. Ass. 73 (1978), 786-793.
(52) Gantmacher, F.R. (1959). Matrix theory. New York: Chelsea.
(53) Gastwirth, J. L. (1971), A general definition of the Lorenz curve, Econometrica 39 1037-1039.
(54) Gastwirth, J. L. (1971). A General Definition of the Lorenz Curve. Econometrica 39, 1037–1039.
(55) **Gheorghe Carmen-Adriana** (2009a) Approximation of Csiszar's f-Divergence using Kullback-Leibler Distance, Proceedings of 9th Balkan Conference on Operational Reserch, Constanta, Romania, http://civile.utcb.ro/balcor.
(56) **Gheorghe Carmen-Adriana** (2009b) Estimation of weighted maximum entropy densities within the study upon Lorenz Curves for grouped data, Proceedings of 16th European Young Statisticians Meetings, ASE, Bucuresti.
(57) **Gheorghe Carmen-Adriana** (2013a) Asymptotic properties of the nonparametric estimators of the Leimkuhler curve and Kakwani index, U.P.B, Scientific Bulletin, Bucuresti, în proces de publicare.
(58) **Gheorghe Carmen-Adriana** (2013b) New parametric families of Leimkuhler Curves, Conferința 16 SPSR, ASE, Bucuresti.
(59) **Gheorghe Carmen-Adriana**; Cipu Corina Elena (2007a) Mortality modeling for Romanian population. Proceedings of The 4-th International Colloquium "Mathematics in Engineering and Numerical Physics" October 6-8 , 2006, Bucharest, Romania, pp. 38-45. Balkan Society of Geometers, Geometry Balkan Press 2007.
(60) **Gheorghe Carmen-Adriana**; Cipu Corina Elena (2007b) The Hill estimator for income of Romanian households, Proceedings of International Conference Trends and Challenges in Applied Mathematics, ISBN 978-973-755-283-9/pbk, Ed. Matrix Rom, Bucuresti, p. 193- 196.
(61) **Gheorghe Carmen-Adriana**; Preda Vasile; Cipu Corina Elena (2012) The estimate of generalized Lorenz confidence intervals using Exponentiated Weibull Distribution, Conferința 15 SPSR, Universitatea din București
(62) Ghitany, M.E, AL-Awadhi, F.A and Alkhalfan,L.A.(2007) Marshall-Olkin extended Lomax distribution and its applications to censored data, Communications in statistics-Theory and Methods, 36, 1855-1866.
(63) Ghitany, M.E.(2005) Marshall-Olkin extended Pareto distribution and its application, International Journal of Applied Mathematics, 18, No.1, 17-31.
(64) Gini, C. (1912). 'Variability and Mutability', C. Cuppini, Bologna, 156 pages. Reprinted in Memorie di metodologica statistica (Ed. Pizetti E, Salvemini, T).
(65) Goldie, C. (1977). Convergence theorems for empirical Lorenz curves and their inverses.Advances in Applied Probability, 9, 765–791.
(66) Goldstein, J.R. (2002). Population momentum for gradual demographic transitions: An alternative approach. Demography, 39, 65-73.
(67) Goldstein, J.R. ; G. Stecklov. (2002). Long-range population projections made simple. Population and Development Review, 28, 121-41.
(68) Gruner, C. M. (1998). Quantifying Information Coding Limits in Sensory Systems. PhD thesis, Dept. Electrical & Computer Engineering, Rice University, Houston, Texas, 1998.
(69) Gupta, M.R. (1984) Functional form for estimating the Lorenz curve, Econometrica, 52, 1313-1314.
(70) Gupta, R. C., Gupta, R.D. , Gupta, P.L. (1998) Modelling failure time data by Lehman alternatives. Communications in statistics- Theory and Methods, 27, p 887-904.
(71) Gupta, R. D., Kundu, D. (2001a). Exponentiated exponential distribution, an alternative to gamma and Weibull distributions. Biometrical Journal, 43(1), 117-130.
(72) Haeusler, E., Teugels, J., (1985). On asymptotic normality of Hill's estimator for the exponent of regular variation. Annals of Statistics 13, 743–756.
(73) Hall, P., (1982). On some simple estimates of an exponent of regular variation. Journal of the Royal Statistical Society B 44, 37–42.
(74) Hartman, P, Winter, A. (1941) On the law of iterated logarithm, Amer. J. Math. 63, p 169-176.
(75) Hill, B.M., (1975). A simple general approach to inference about the tail of a distribution. Annals of Statistics 3, 1163–1174.
(76) Hochberg Y., A. C. Tamhane (1987). Multiple Comparison procedures, Wiley, New York.
(77) Hochberg, Y.; Tamhane, A. (1987) Multiple Comparison Procedures. New York: Wiley.
(78) Horvath, L. (1985). Estimation from a length-biased distribution. Statistics and Decisions, 3, 91/113.
(79) Iosifescu, M., S. Grigorescu, Gh. Oprişan, Gh. Popescu, (1984), Elemente de modelare stohastică, Editura Tehnică, Bucureşti.
(80) Jeffreys, H. (1946). An invariant form for the prior probability in estimation problems. Proc. Roy. Soc. A.

(81) Johnson, D. H. ; Orsak, G. C. (1993). Relation of signal set choice to the performance of optimal non-Gaussian detectors. IEEE Trans. Comm.
(82) Johnson, D. H.; Gruner, C. M. (1998). Information-theoretic analysis of neural coding. ICASSP, Seattle,WA.
(83) Johnson, D. H.; Gruner, C. M. (1998).. Information-theoretic analysis of signal processing systems: Application to neural coding. In Inter. Sympos. Information Theory,MIT,Cambridge,MA.
(84) Kailath, T. (1967). The divergence and Bhattacharyya distance measures in signal selection. IEEE Trans. on Comm. Tech.
(85) Kakwani, N. C. (1980). Income Inequality and Poverty: Methods of Estimation and Policy Applications. Oxford University Press.
(86) Kenichi Hirose (2004): Topics in Quantitative Analysis of Social Protection Systems, *Social Security Department*, International Labour Office, Geneva.
(87) Keyfitz, N. (1968). Introduction to the mathematics of population. Reading MA: Addison-Wesley.
(88) Keyfitz, N. (1971). On the momentum of population growth. Demography, 8, 71-80.
(89) Kim, Y.J. (1987). Dynamics of populations with changing rates: Generalization of stable population theory. Theoretical Population Biology, 31, 306-22.
(90) Kim, Y.J. ; R. Schoen. (1993). On the intrinsic force of convergence to stability. Mathematical Population Studies, 4, 89-102.
(91) Kim, Y.J. ; R. Schoen. (1996). Populations with quadratic exponential growth. Mathematical Population Studies, 6, 19-33.
(92) Kim, Y.J. ; R. Schoen. (1997). Population momentum expresses population aging. Demography, 34, 421-27.
(93) King, W.I. (1912). The Elements of Statistical Method. New York: Macmillan
(94) Kleiber C. (2005). Lorenz Curve in Economy and Econometry, Gini-Lorenz Continental Conference, Siena
(95) Kleiber C., Kotz S. (2000): Statistical Size Distributions in Economics and Actuarial Sciences, John Wiley
(96) Kleiber, C., (1999). Tail dominance of income distributions, working paper.
(97) Kullback, S. (1959). Information Theory and Statistics. Wiley, New York.
(98) Kullback, S.; Leibler, R.A. (1951). On information and sufficiency. Ann. Math. Stat.
(99) Land, K.C. ; A. Rogers. (1982). Multidimensional mathematical demography. New York: Academic Press.
(100) Lee, R. (1974). The formal dynamics of controlled populations and the echo, the boom, and the bust. Demography, 11, 563-85.
(101) Leslie, P.H. (1945). On the use of matrices in certain population mathematics. Biometrika, 33, 183-212.
(102) Levy, M., Solomon, S. (1997). New Evidence for the Power-Law Distribution of
(103) Li, N. S. Tuljapurkar. (2000). The solution of time-dependent population models. Mathematical Population Studies, 7, 311-29.
(104) Li, N. ; S. Tuljapurkar. (1999). Population momentum for gradual demographic transitions. Population Studies, 53, 255-62.
(105) Lorenz, M. O. (1905). Methods of Measuring the Concentration of Wealth. Publications of the American Statistical Association 9, 209–219.
(106) Lotka, A.J. (1939). Theorie analytique des associations biologiques. Paris: Hermann.
(107) Marshall,A. N, Olkin, I.(1997) A new method for adding a parameter to a family of distributions with applications to the exponential and Weibull families, Biometrica 84, 641-652.
(108) McDonald, J., 1984. Some generalized functions for the size distribution of income. Econometrica 52, 647–663.
(109) McDonald, J., Xu, Y., 1995. A generalization of the beta distribution with applications. Journal of Econometrics 66, 133–152.
(110) Mihoko, M.; Eguchi, S. Robust blind source separation by beta divergence. Neural comput. 2002, 14, 1859–1886.
(111) Ogwang, T., Rao, U., (1996). A new functional form for approximating the Lorenz curve. Economics Letters 52, 21–29.
(112) Pollard, J.H. (1973). Mathematical models for the growth of human populations. Cambridge: Cambridge University Press.
(113) Preda, V. (1992a) Inferență statistică pentru medii condiționate, Editura Universității București
(114) Preda, V. (1992b) Probleme de statistică matematică. Estimări, Editura Universității București
(115) Preda, V. (1992c) Teoria deciziilor statistice, Editura Academiei Romane, București
(116) Preda, V. ; Craiu V. (1980). Probleme de decizie multiplă, Editura Universității București
(117) Preda, V. ; Craiu V. (1981). Verificarea normalității datelor, Editura Universității București
(118) Preda, V., E Panaitescu, R Ciumara (2011b) The Modified Exponential-Poisson Distribution. Proceedings of The Romanian Academy 12 (1), 22-29
(119) Preda, V., I Stancu-Minasian, M Beldiman, AM Stancu (2011a) On a general duality model in multiobjective fractional programming with n-set functions. Mathematical and Computer Modelling 54 (1-2), 490-496

(120) Preda, Vasile; **Gheorghe Carmen-Adriana** (2013a) A class of Lorenz Curves generated by the Marshall Olkin technique. În proces de publicare.
(121) Preda, Vasile; **Gheorghe Carmen-Adriana** (2013b) Bounds for some generalised divergences in trems of Lorenz curves. În proces de publicare.
(122) Preda, Vasile; **Gheorghe Carmen-Adriana** (2013c) Maximum Tsallis entropy and the Lorenz curve. În proces de publicare.
(123) Preda, Vasile; **Gheorghe Carmen-Adriana** (2013d) Some approximations for generalised divergences . În proces de publicare.
(124) Preston, S.H. (1986). The relation between actual and intrinsic growth rates. Population Studies, 40, 495-501.
(125) Preston, S.H. ; A.J. Coale. (1982). Age structure, growth, attrition, and accession: A new synthesis. Population Index, 48, 217-59.
(126) Richmond, J. (1982). A general method for constructing simultaneous confidence intervals , Journal of American statistical Association, Vol.77, No. 378, 455-460.
(127) Rogers, A. (1975). Introduction to multiregional mathematical demography. New York: Wiley.
(128) Rohde N. (2009), An alternative functional form for estimating the Lorenz curve, Economics Letters, 105, 61-63.
(129) Sallila, S., Hiilamo, H. (2004). Rethinking relative measures of poverty , Working paper no. 368, Luxembourg Income Study Working Paper Series.
(130) Sarabia, J. M. (2008). Parametric Lorenz Curves: Models and Applications. In D. Chotikapanich (Ed.), Modeling Income Distributions and Lorenz Curves. Economic Studies in Inequality, pp. 167–190. Springer.
(131) Sarabia, J. M., Prieto, F., Sarabia, M. (2010). Revisiting a functional form for the Lorenz curve. Economics Letters, Elsevier, vol. 107(2), pages 249-252, May.
(132) Savin, N., (1993). Multiple hypothesis testing. In: Griliches, Z., Intriligator, M.D. (Eds.), Handbook of Econometrics, Chapter 14. North-Holland, Amsterdam.
(133) Schluter C., Trede M. (2001) Tails of Lorenz Curves, TMR Network, Living Standards.
(134) Schoen, R. (1988). Modeling multigroup populations. New York: Plenum.
(135) Schoen, R. (2003). Dynamic populations with uniform natural increase across states. Mathematical Population Studies, 10, 195-210.
(136) Schoen, R. ; S.H. Jonsson. (2003). Modeling momentum in gradual demographic transitions. Demography, 40, 621-35.
(137) Schoen, R. ; Y.J. Kim. (1994a). Cyclically stable populations. Mathematical Population Studies, 4, 283-95.
(138) Schoen, R. ; Y.J. Kim. (1994b). Hyperstability. Paper presented at the Annual Meeting of the Population Association of America, Miami, May 5-7.
(139) Schoen, R. ; Y.J. Kim. (1996). Stabilization, birth waves, and the surge in the elderly. Mathematical Population Studies, 6, 35-53.
(140) Schoen, R. ; Y.J. Kim. (1997). Exploring cyclic net reproduction. Mathematical Population Studies, 6, 277-90.
(141) Schoen, R. ; Y.J. Kim. (1998). Momentum under a gradual approach to zero growth. Population Studies, 52, 295-99.
(142) Sen, P. K. (1984). On asymptotic representations for reduced quantiles in sampling from a length-biased distribution. Calcutta Statistical Assosiation Bulletin, 33, 59-67.
(143) Shorrocks, A. F. (1983). Ranking Income Distributions , Econometrica, 50, 3-17.
(144) Singh, S., Maddala, G., (1976). A function for size distribution of incomes. Econometrica 44, 963–970.
(145) Temkin, L. S. (1993). Inequality. Oxford: Oxford University Press.
(146) Topsøe, F. (2000). Some inequalities for information divergence and related measures of discrimination. IEEE Trans. Info. Th.
(147) Tsallis, C. (1988). "Possible generalization of Boltzmann-Gibbs statistics". Journal of Statistical Physics 52: 479–487.
(148) Tuljapurkar, S. (1990). Population dynamics in variable environments (Vol. 85). New York: Springer.
(149) Tuljapurkar, S.D. (1993). Entropy and convergence in dynamics and demography. Journal of Mathematical Biology, 31, 253-71.
(150) Vardi, Y. (1982). Nonparametric estimation in the presence of length-bias. The Annals of Statistics, 10, 616/620.
(151) Vernic R., Bjoern Sundt (2009) Recursions for Convolutions and Compound Distributions with Insurance Applications, EAA Lecture Notes, Springer Berlin Heidelberg.
(152) von Collani E., Dumitrescu M., (2004) Stochastic methods for production processes, Economic Quality Control, 19
(153) Weissman, I., (1978). Estimation of parameters and large quantiles based on the k largest observations. Journal of the American Statistical Association 73, 812–815.

(154) Widder, D.V. (1928) A generalisation of Taylor's formula, Trans.AMS 30, 126-154.
(155) Wilfling, B., W. Kramer (1993). The Lorenz-ordering of Singh–Maddala income distributions. Economic Letters 43, 53–57.
(156) Yitzhaki, Shlomo (1983) "On an Extension of the Gini Inequality Index," International Economic Review, Department of Economics, University of Pennsylvania and Osaka University Institute of Social and Economic Research Association, vol. 24(3), pages 617-28.
(157) Zhang TL, Min Xie.(2007) Failure data analysis with extended Weibull distribution, Communications in Statistics-Simulation and computation, 36, 579-592.