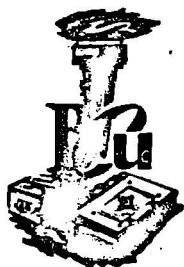


Poliana Ștefănescu

**MODELE
CAUZALE
ÎN
SOCIOLOGIE**

Editura Universității din București



BIBLIOTECA CENTRALĂ
UNIVERSITARĂ
București

Cota 1302941
C20000416
Inventar C2000416

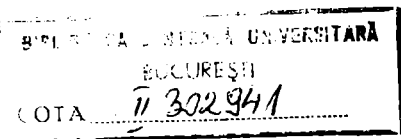
DOLIANA ȘTEFANESCU

MODELE CAUZALE ÎN SOCIOLOGIE

POLIANA ȘTEFĂNESCU

**MODELE CAUZALE
ÎN
SOCIOLOGIE**

**Editura Universității din București
– 1999 –**



69/100

© Editura Universității din București
Șos. Panduri, 90.92, București - 76235; Telefon/Fax: 410.23.84

B.C.U. București



C20000416

Tehnoredactare computerizată: Constanța Titu

Descrierea CIP a Bibliotecii Naționale

ȘTEFĂNESCU, POLIANA

Modele cauzale în sociologie / Poliana Ștefănescu

București; Editura Universității din București, 1999

156 p.; 21 cm

Bibliogr.

ISBN: 973-575-312-X

316

I. ANALIZA CAUZALĂ

I. 1 INTRODUCERE

Cauzalitatea are o importanță deosebită în sociologie. Ea oferă o relație simplă de determinare : un fapt social trebuie explicat printr-un alt fapt social, care îi este cauza (Durkheim, 1974).

Cauzalitatea socială este o categorie fundamentală a determinismului social care desemnează relațiile de generare sau producere dintre fenomene, fapte sau evenimente sociale. Specificarea unei relații cauzale este în strânsă legătură cu identificarea factorilor contingenți, care pot împieta asupra universalității relațiilor de cauzare socială.

Datorită dificultăților care însoțesc definirea și dezvoltarea acestui concept, se va folosi o definiție simplă:

cauzalitatea reprezintă o relație între doua fenomene, caracterizată prin aceea că unul dintre ele îl produce pe celălalt (Zamfir,1987).

Fenomenul produs este **efectul** iar fenomenul care îl produce este **cauza**.

Ideea de cauzalitate are la bază postularea unui proces ontologic în cadrul căruia fenomenul “**cauza**” reprezintă factorul activ, responsabil de producerea fenomenului “**efect**”.

Relația cauzală poate fi reprezentată grafic astfel :

C —→ **E**

unde C = cauza, E = efectul iar “ —→ ” indică procesul de producere a fenomenului “**efect**” de către fenomenul “**cauza**”.

În cercetarea socială, dar și în limbajul social comun formulările cauzale sunt frecvente. Relația cauzală poate asocia două tipuri de fenomene particulare sau complexe. De exemplu : incendiul produce panica, respectiv progresul economic determină structura politică.

I. 2 EXPLICAȚIA CAUZALĂ

Elaborarea unei explicații cauzale implică specificarea celor două componente : **enunțul cauzal și modelul explicativ.**

Enunțul cauzal postulează dependența unui fenomen (efectul), de alt fenomen (cauza) sub forma unei legi universale și necesare

“ ori de câte ori C atunci E “ sau

“ C este cauza lui E “

Din punct de vedere empiric, enunțul cauzal ia forma unui enunț de covariație :

“ variația lui C este însoțită de variația lui E “.

În funcție de particularitățile fenomenelor, covariația poate fi discretă sau continuă.

Covariația discretă se manifestă atunci când diferitele stări distincte ale lui C sunt însoțite de stări distincte ale lui E ; de exemplu, stratificarea (promovarea) personalului administrativ este determinată de caracteristici structurale (Heck & Johnsrud, 1994).

Covariația continuă (cantitativă) se explică prin enunțul: “ cu cât C este mai mare, cu atât E este mai mare “ ; de exemplu variația dintre performanța în muncă și satisfacție (Bagozzi, 1980).

Enunțul cauzal are funcție explicativă și predictivă a fenomenelor.

Modelul explicativ este o completarea a enunțului cauzal, oferind explicația procesului prin care C produce E.

Este necesar să se distingă explicarea unui fenomen

“ De ce E ? - Pentru că C”

de explicarea unei legi cauzale

“ De ce C este cauza lui E ? “ sau “ Cum C produce E ?”.

Orice lege cauzală exprimată prin enunț cauzal este testată empiric prin multiplicarea observațiilor directe, astfel încât să se poată demonstra că apariția fenomenului **cauză** favorizează sau se asociază cu producerea fenomenului **efect**. De exemplu se poate considera relația dintre performanța școlară a elevilor și nivelul de educație al părinților (Cohen,1983).

Dacă într-un eșantion reprezentativ, cea mai mare parte a elevilor cu note bune la învățătură au părinți cu un nivel superior de educație, atunci se poate afirma în termeni probabilistici că gradul de instruire al părinților este o **cauză** a rezultatelor școlare ale copiilor.

Atât enunțul cauzal cât și demonstrarea empirică sunt completate de modelul explicativ care expune mecanismele producerii cauzale.

Modelul explicativ este un instrument valoros pentru definirea unei legi cauzale. Funcțiile sale rezolvă cerințele teoretice referitoare la :

Emiterea de ipoteze asupra unor posibile relații cauzale pe baza cunoștințelor teoretice acumulate anterior.

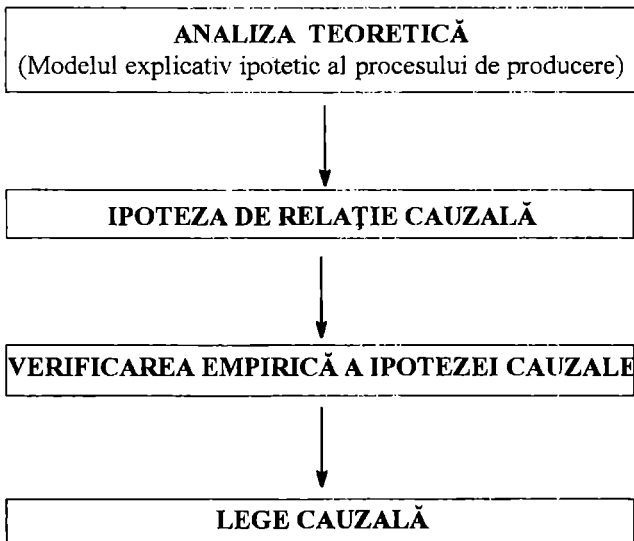
Precizarea unei relații cauzale prin cunoașterea mecanismelor de producere a fenomenelor sociale (identificarea entităților abstracte legate cauzal precum și contextul în care se manifestă relația cauzală).

Generalizarea unei relații cauzale prin asumarea faptului că o regularitate empirică este independentă de un context particular.

Verificarea ipotezelor cauzale prin selectarea unui enunț cauzal plauzibil și prin susținerea sau respingerea unei regularități empirice.

Schema grafică următoare expune etapele de stabilire a unei legi cauzale, conform Zamfir (1987).

ETAPELE DE STABILIRE A UNEI LEGI CAUZALE



I.3 STABILIREA CAUZALITĂȚII

Stabilirea cauzalității implică următoarele aspecte :

1. Stabilirea sensului cauzării

Prin obținerea unei covariații empirice a două variabile (fenomene) se poate presupune o posibilă relație cauzală.

Determinarea sensului cauzării parcurge următoarele etape:

Stabilirea succesiunii ia în considerare faptul că relația cauzală este în mod necesar temporală. Cele două fenomene: cauza și efectul sunt în mod necesar decalate în timp : cauza precede efectul (de exemplu incendiul precede panica). În sociologie însă, este dificil să se determine clar succesiunea temporală deoarece în multe situații sunt analizate procese sociale de lungă durată, în care cauza coexistă cu efectul. De exemplu, contextul economic coexistă cu cel politic.

Costner și Schoenberg (1973) în “Structural Equation Models in the Social Sciences” reiau analiza datelor lui Olsen (1968) referitoare la indicatorii de dezvoltare industrială și indicatorii de dezvoltare politică. Analiza începe cu un model simplu în care dezvoltarea industrială este considerată ca fiind cauza dezvoltării politice. Deoarece modelul nu are consistență din punct de vedere statistic s-a considerat că, în timp, și dezvoltarea politică influențează dezvoltarea economică, ceea ce a condus la formularea unui model complicat și dinamic în timp.

Presupunerea unui proces de producere implică imaginarea unui proces plauzibil de producere a variației unui fenomen de către altul. Dacă se reușește această supoziție, atunci este garantată adoptarea unui sens al cauzării.

Bagozzi (1980) a formulat un model causal pentru a studia relația între performanța profesională și satisfacția muncii. Modelul, într-o formă simplificată, arată astfel:

PERFORMANȚĂ \longrightarrow SATISFACTIE

Analiza cauzală a trebuit să răspundă unor întrebări de genul : “ Legătură între performanță și satisfacția în muncă este mit sau realitate ? Performanța influențează satisfacția, sau satisfacția influențează performanța ?”

Un caz particular îl reprezintă existența ambelor sensuri de cauzare între două fenomene care poartă numele de relație circulară de determinare reciprocă. Se poate presupune însă, că un sens este mai puternic decât celălalt :

C \longleftrightarrow E

Duncan, Haller & Portes (1968) au prezentat un model causal al influenței prietenilor asupra ambițiilor sociale ale subiecților, pe baza unui sondaj în rândul studenților.

Simplificat, modelul arată astfel :

AMBIȚII SOCIALE ALE SUBIECTULUI	\longleftrightarrow	AMBIȚII SOCIALE ALE PRIETENULUI	
------------------------------------	-----------------------	------------------------------------	--

În plus, relația causală este perfect simetrică deoarece s-a demonstrat empiric că efectele sunt identice.

În știință pot apare "răsturnări" în ceea ce privește sensul cauzării. În tradiția sociologică și psihologică se postula următoarea relație causală:

ATITUDINI → COMPORTAMENT

Empiric, există o corelație ridicată între aceste două variabile, iar sensul cauzării pare evident : atitudinile determină comportamentul. Deoarece stabilirea empirică a succesiunii este imposibilă în acest caz, noile abordări științifice au pus în discuție acest sens al determinării. Comportamentul uman, odata cristalizat, nu mai poate fi modificat în raport cu atitudinile. Acestea din urmă pot fi însă modificate în acord cu comportamentul. Se ajunge astfel la următoarea relație causală :

COMPORTAMENT → ATITUDINI

2. Stabilirea duratei și a distanței

Causalitatea presupune un decalaj temporal : efectul apare după o perioadă de timp de la activarea cauzei. Apare următoarea problemă : cât de mare poate fi decalajul temporal între cauză și efect ? Dacă se ia ca satisfăcător răspunsul:

"mai mare ca zero și mai mic ca infinitul", ce sens poate avea o relație causală între două evenimente puternic corelate empiric dar separate de o perioadă foarte mare de timp ?

La dificultatea duratei se adaugă și cea distanței. Efectul este separat de cauză și spațial ; dar cât de depărtat ? În acest caz, definirea causalității nu dă un răspuns.

Eliminarea acestor probleme trebuie realizată prin presuposițiile care decurg din modelul explicativ al procesului de producere causală.

3. Problema cauzelor aparente poate apare atunci când legile cauzale sunt induse din regularități empirice sau sunt tributare prejudecăților colective interpretate tendențios.

Un exemplu clasic îl reprezintă studiile legate de sărăcie. În 1976 Comunitatea Europeană a inițiat un sondaj în țările membre, asupra

percepției sărăciei (Sen, 1979). Una din întrebări se referea la cauzele percepute ale sărăciei. Astfel, în Marea Britanie, în 43 % din cazuri răspunsul a indicat că lenea și lipsa de voință constituie principala cauză; această proporție este exagerat de mare în comparație cu răspunsurile date în celelalte state, care indicau în principal cauze structurale.

Acest sondaj a evidențiat faptul că ideologiile asupra sărăciei sunt produsul unor forțe particulare în societăți particulare. Ele sunt, mai degrabă, produsul unui context istoric și al unor circumstanțe curente, decât adevăruri (sau neadevăruri) universale. Și așa cum cursul istoriei și contextul se schimbă, acest lucru se întâmplă cu ideologiile.

În 1989, un nou sondaj al Comunității Europene în Marea Britanie nu a obținut decât în 18 % din cazuri explicații individuale sau patologice asupra cauzelor sărăciei (Alcock, 1993).

Sărăcia nu este produsă de o slăbiciune sau un eșec individual; ea este rezultatul proceselor sociale sau structurale (Ferge & Millar, 1987). Starea de sărăcie nu se datora, deci, unor caracteristici personale, ci contextului social.

Este posibil ca cele două fenomene considerate apriori cauză și efect, să fie efecte independente ale aceleiași cauze.

De exemplu, variația concomitentă a cererii de mărfuri și a economiilor bănești ale populației constituie o regularitate empirică. Totuși, de aici nu se poate deduce o lege cauzală între cele două variabile, deoarece, la rândul lor, fiecare este legată de veniturile banesti ale populației.

O regularitatea empirică, oricât ar fi de puternică, nu constituie expresia unei legi cauzale decât atunci când este validată și teoretic, prin asocierea unei explicații a procesului de producere.

4. Riscul de a fi corect din rațiuni false apare în situația în care se pornește de la o presuposiție falsă asupra cauzelor, prezicându-se corect o corelație empirică.

Zamfir (1987) menționează un exemplu elocvent: în secolul trecut, în unele comunități americane, circula obiceiul de a purta la gât o punguță cu amfetamină pentru a alunga răceala.

Cele două fenomene erau corelate empiric, dar ulterior s-a dovedit că amfetamina nu are acțiune curativă. Datorită mirosului neplăcut, ea ținea lumea la distanță, diminuând astfel riscul contaminării.

Babbie (1992) discută următoarea situație :

În Statele Unite ale Americii exista convingerea că absolvenții Universităților din Liga Ivy (Harvard, Yale, Columbia, Princetown, etc.) obțin mai multe succese în cariera profesională decât absolvenții altor universități. Deși cele două fenomene sunt corelate empiric, cercetările sociologice efectuate încă din anii 50 (Haveman & West, " They Went to College " - 1952) au atras atenția asupra altor cauze ale succesului profesional :

Cine își putea permite să-și trimită copiii la universitățile din Liga Ivy ? - Bineînțeles că familiile bogate.

Cine dispunea de relațiile profesionale și de afaceri prin care să-și ajute copiii ? - Familiile bogate sau cu poziție socială.

Cu alte cuvinte, statusul socio - economic al familiilor studenților poate explica aparenta relație cauzală.

Disputa nu este însă încheiată, deoarece Universitățile în cauză se mândresc cu performanțele intelectuale ale studenților lor și acordă numeroase burse de merit celor foarte studioși, dar care nu dispun de mijloace materiale suficiente pentru a plăti școala.

5. Problema numărului observațiilor pe care se construiesc teoriile sociologice, este importantă.

În multe situații se dispune de un număr redus de cazuri, ceea ce nu permite separarea accidentalului, întâmplătorului, de ceea ce este general, necesar.

Tot în acest context trebuie menționată problema interdependenței și a interacțiunii fenomenelor. Acest aspect poate perturba un proces cauzal, în lipsa unui control teoretic și empiric.

Schema cauzală are o pronunțată deschidere empirică. Adeseori se procedează iluzoriu, formulând ipoteze cauzale numai pe baza procedurilor empirice inductive. În această situație, termenii relației cauzale, cauza și efectul, sunt dați în experiența înainte și independent de formularea

ipotezelor cauzale. De regulă, cauza și efectul sunt formulate mai întâi teoretic, ca entități abstracte, și apoi operaționalizate printr-un sistem de indicatori empirici.

Procesul prin care se trece de la conceptele abstracte la variabile măsurabile se numește **operaționalizare**. Această constă în specificarea operațiilor care trebuie efectuate pentru a observa sau măsura reperele concrete ale conceptelor abstracte.

6. Variația contextuală a cauzalității

Spre deosebire de omogenitatea realității fizice, realitatea socială are un caracter variabil în spațiu și timp.

Structura globală a unei societăți influențează puternic relațiile dintre fenomenele sociale, ajungând până la imprimarea unor orientări diferite. Din acest motiv, este riscantă generalizarea la toate societățile, a proceselor cauzale identificate într-un anumit tip de societate. Ceea ce într-un context reprezintă o cauză, în altul poate înceta să mai fie.

Primele cercetări asupra mobilității sociale au pus în evidență consensul privind rolul școlii ca factor de determinare a mobilității dintre statusul moștenit și statusul realizat. În 1927 Sorokin ("Social Mobility") a demonstrat rolul pe care îl au instituțiile de educație în distribuirea indivizilor pe diferite poziții profesionale, ținând seama de aspirațiile lor și de posibilitățile structurii sociale.

La începutul anilor '60 mai multe anchete efectuate în țările anglo-saxone arătau că dobândirea unei diplome de studii superioare celei a tatălui nu garantează neapărat o poziție socială mai bună (Paradoxul lui Anderson). Studiile de mobilitate socială efectuate în Franța începând cu 1970, confirmă rezultatul anterior.

Paradoxul se explică prin faptul că structura diplomelor nu este identică cu structura posturilor oferite pe piața muncii. Școala nu acționează ca o instituție autonomă care redistribuie pozițiile sociale în funcție de diplomele pe care le acordă.

I.4 MULTICAUZALITATE

Relația cauzală avută în discuție până acum considera cauza ca fiind identificată cu un singur fenomen. De cele mai multe ori însă, fenomenele sociale au o cauză multiplă.

În aceste situații explicația cauzală constă în identificarea complexului cauzal.

Aspectele specifice multicauzalității sunt următoarele :

4.1. *Structura listei cauzale :*

Lista cauzală cuprinde factori determinanți în explicarea diferitelor fenomene sociale. În analiza socială se face distincție între cauze și condițiile care contribuie la explicarea efectului (de exemplu, vântul nu este cauza unui incendiu, dar poate fi o condiție importantă).

Faptul că la alegerile prezidențiale din S.U.A., între anii 1940 - 1960, au câștigat candidații cu numele cel mai lung este o coincidență, nicidecum o cauză.

În studiile efectuate asupra calității vieții (Zamfir coord.1984), s-a pornit de la o teorie abstractă asupra calității vieții și a surselor sale de variație. Calitatea vieții a fost definită ca un concept evaluativ : rezultatul evaluării condițiilor de viață, a "stării vieții", prin prisma unor criterii de valoare (sistemul de necesități și aspirații ale fiecărei persoane). Pornind de la o asemenea definiție, apar distincte două categorii de variabile cauzale abstracte :

- variabile referitoare la starea vieții, la condițiile de viață ;
- variabile referitoare la necesitățile, aspirațiile persoanelor : configurația necesităților și nivelul de aspirații.

Pe această bază se pot formula diferite legi abstracte ale variației calității vieții :

- creșterea (îmbunătățirea) stării vieții și a condițiilor de viață determină o creștere a calității vieții;
- creșterea nivelului de aspirații determină o scădere relativă a calității vieții.

4.2. Mecanisme și circuite determinative

În explicarea cauzală nu este suficientă alcătuirea unei simple liste cu factori determinanți. Este nevoie să se dezvolte un model teoretic pentru fiecare factor, care să evidențieze mecanismele prin care se realizează contribuția sa la variația variabilei - efect. În acest mod s-a procedat pentru relația cauzală nivel școlar / calitatea vieții (C. Zamfir, coord., 1984).

4.3. Mărimea contribuției factorilor determinanți

Când lista cauzală conține mai mulți factori este important să se determine contribuția fiecăruia la explicarea efectului.

Tehnica cea mai simplă o constituie corelația sau covarianța. Calcularea coeficientului de corelație simplu sau multiplu oferă o măsură a covariației variabilelor cauză și efect

Pentru a măsura contribuția independentă a fiecărui factor în parte se folosesc tehnici statistice elaborate : analiza de regresie, analiza path, analiza structurala.

4.4. Lanțuri cauzale

Factorii cuprinși în lista cauzală nu acționează simultan ci sunt decalajați în timp, formând un lanț de cauze și efecte în care un factor determină efectul prin intermediul altor factori.

Astfel, nivelul școlar contribuie substanțial la variația calității vieții, dar el acționează prin tipul de profesie care, la rândul său, determină venitul, tipul de ocupație, statusul social și modul de viață (Zamfir,1984).

În viața socială lanțurile cauzale pot fi declanșate de producerea unui fenomen social important - o criză economică sau politică, de un proces social mai general - restructurare economică, reforma sistemului educațional, sau o caracteristică structurală profundă a societății - schimbarea de regim politic, tranziția la un alt tip de economie (privată sau statală).

Un exemplu îl constituie fenomenul de ereditate profesională analizat în studiile de mobilitate socială. Mediul de origine influențează evoluția profesională a individului. Intre două generații mobilitatea operează între două categorii vecine :

tata → fiu

S-a demonstrat însă, că efectul mediului de origine persistă și la a treia generație : bunic → fiu → nepot, de exemplu:

bunic/țăran → tata/cadru mediu → nepot/cadru superior.

Efectul de descendență este puternic, mediul de origine reprezentat de bunic acționând direct și indirect asupra poziției sociale dobândite de nepot. De exemplu, s-a constatat că la fiul unui tată cu studii superioare, un bunic cu studii superioare constituie un atuu ; pentru fiul unui muncitor, un bunic muncitor poate mări riscul de imobilitate (Pohl, 1982)

Un exemplu interesant oferă Spicker (1993), care în studiul său despre sărăcie și securitate socială comentează explicațiile familiale ale sărăciei. Un aspect particular îl constituie “ ciclul deprivării “ sau “ deprivarea transmisă “. Aceasta se referă la faptul că părinții săraci determină dezavantaje în dezvoltarea copiilor, dar, “ciclul deprivării” continuă : părinții care au fost ei înșiși deprivați într-un anumit mod în copilărie, devin la rândul lor părinții unei alte generații de copii deprivați.

I.5 STRUCTURA STANDARD A TEORIEI CAUZALE

O teorie cauzală completă cuprinde următoarele seturi de variabile (vezi și Zamfir, 1987):

1. Variabila de explicat (efectul)
2. Variabile cauzale abstracte (endogene)
3. Variabile contextuale (exogene)
4. Variabile acționale (pragmatice)

Variabila de explicat se referă la fenomenul efect.

În general, teoriile sociologice sunt grupate în două categorii: teorii centrate pe explicarea fenomenului efect și teorii centrate pe explicarea fenomenului cauză.

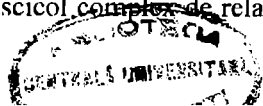
Variabilele cauzale abstracte sau endogene sunt acele variabile care au influență cauzală directă asupra variabilei de explicat.

Teoria abstractă generală a fenomenului studiat este formată din : variabila de explicat (efectul), variabilele endogene și enunțurile cauzale care descriu relațiile dintre ele. Această teorie abstractă nu poate explica prin ea însăși un fenomen concret. Ea reprezintă numai structura teoretică ce stă la baza unei analize explicative particulare care se realizează prin intermediul variabilelor contextuale.

Variabilele contextuale sau exogene descriu modul în care diferitele componente ale unui context social particular determină cauzal fenomenul de explicat (efectul), indirect prin intermediul variabilelor cauzale abstracte (endogene).

Teoria abstractă folosită în explicarea unui fenomen social concret trebuie să fie legată de contextul social al acestuia. Deci, variabilele contextuale sau exogene reprezintă punctele în care teoria abstractă (variabilele cauzale abstracte) se conectează la un context social concret. Ele reprezintă cauzele concrete ale respectivului fenomen în contextul social dat.

Contribuția globală a unei variabile contextuale la explicarea variabilei efect reprezintă rezultanta unui fascicol complex de relații de



determinare, și poate varia semnificativ de la un context social la altul (Zamfir, 1987).

În teoria sociologică a sărăciei studiile efectuate în paralel în diferite state au condus la metode diferite de măsurare și explicare a fenomenului. Un sărac în Marea Britanie este diferit de un sărac în India, de exemplu. După cum, noțiunea de sărăcie din secolul XVIII diferă de noțiune de sărăcie din secolul XX (Alcock, 1993).

Așadar, teoria cauzală abstractă trebuie completată cu o teorie contextuală concretă.

Variabilele acționale sau pragmatice se referă la acțiuni individuale și colective care afectează variabilele contextuale sau exogene.

Teoria pragmatică acțională a fenomenului studiat este reprezentată de conectarea variabilelor acționale la variabila efect prin intermediul variabilelor contextuale (exogene) și a variabilelor cauzale abstracte (endogene).

Teoriile contextuale oferă explicarea unui fenomen așa cum se manifestă el într-o societate oarecare.

Teoriile pragmatice analizează capacitatea diferitelor acțiuni de a modifica modul de manifestare a fenomenului studiat.

Figura I.1 reprezintă grafic structura standard a teoriei cauzale.

STRUCTURA STANDARD A TEORIEI SOCIALE

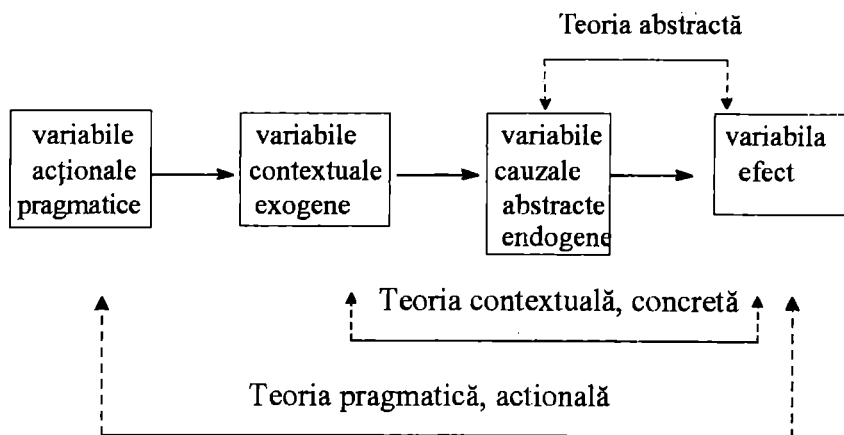


Figura I.1

I.6 MODELE

În mod evident modelele constituie simplificări ale realității. Ele există în diferite nivele de abstractizare, de la modele metafizice (meta-modele), sau paradigme, sau ipoteze asupra lumii, până la modele precise sau teorii științifice specifice.

Relațiile cauzale între fenomene constituie o preocupare de bază a științelor sociale. Modelele sunt exprimări explicite ale acestor relații cauzale. Modelele descriu aspectele sistematice ale unui fenomen, văzute într-o multitudine de circumstanțe.

Conform Popper (1981), principiul cauzalității este afirmația că orice eveniment poate fi explicat cauzal, adică prezis pe cale deductivă.

Principiul deducției presupune testarea unei teorii prin “aplicarea empirică” a consecințelor derivate din ea. Prin experimente științifice, observații sau aplicații se va testa dacă noile consecințe deduse din teorie se susțin și practic.

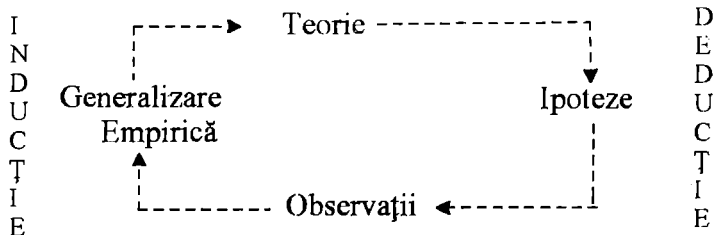
Principiul inducției se bazează pe inferența de la **enunțuri singulare** care descriu observații, experimente, la **enunțuri universale**, ipoteze sau teorii.

Sunt două abordări în construirea unui model :

Metoda deductivă - începe cu un set de axiome din care derivă un set de predicții asupra variației unor fenomene. De exemplu, teoria economică a raportului cerere / ofertă se poate folosi pentru studiul comportamentului cumpărătorilor față de un produs (Berry, 1984).

Metoda inductivă - se bazează pe observații empirice și pe presupuneri intuitive asupra variației așteptate în circumstanțe diferite. Relația dintre nivelul școlar și variația calității vieții nu este dedusă dintr-o axiomă, dar are la bază alte experiențe și observații (Zamfir, coord., 1984).

În practica actuală, teoria și cercetarea empirică interacționează într-o alternanță fără sfârșit de deducție, inducție, deducție, s.a.m.d. Wallace (1971) a reprezentat grafic acest proces circular :



În științele sociale modelele sunt în mare parte inductive. Ipotezele și predicțiile derivate din asemenea modele sunt componente importante ale științelor sociale aplicate.

Studiile empirice reprezintă o etapă importantă în dezvoltarea și rafinarea unui model. Observațiile empirice asupra fenomenelor studiate nu pot “demonstra” teorii sau ipoteze, dar contribuie la validarea explicațiilor cauzale.

În cadrul modelului probabilistic este util să se distingă două tipuri de cauze : cauze necesare și cauze suficiente.

O cauză necesară reprezintă o condiție care trebuie să fie prezentă pentru a produce “efectul”. De exemplu, pentru a obține o diplomă de licență, este necesar să urmezi cursuri universitare, dar trebuie trecute și examenele corespunzătoare.

O cauză suficientă reprezintă o condiție care, dacă este prezentă, garantează producerea “efectului”. De exemplu, neprezentarea la un examen reprezintă o cauză suficientă pentru a nu trece examenul.

Descoperirea unei cauze necesare și suficiente constituie rezultatul cel mai important al unei cercetări. Din nefericire, foarte rar se descoperă asemenea cauze. În științele sociale o relație cauzală rezultă din cauze fie necesare, fie suficiente (Babbie, 1992).

Obiectivul fundamental al analizei empirice este de a găsi contraexemple care să demonstreze inconsistența între valorile observate și predicțiile efectuate.

Dacă predicțiile modelului sunt dezmințite empiric, atunci este reluată analiza teoretică și formulat un nou model sau îmbunătățit modelul testat anterior.

Relațiile cauzale estimate într-un model pot determina variabilele specifice, interpretarea rezultatelor și direcția de dezvoltare a modelului.

Dar fără o analiză cauzală teoretică de bază rezultatele statistice vor reprezenta un set de cifre abstracte. Teoria conduce la interpretarea și atașarea unor semnificații concrete rezultatelor.

Un model conceptual redă o teorie formală sau o problemă complexă în formă verbală sau schematică. Modelul, ca și ipotezele pe care le conține, poate fi construit din observații acumulate, intuiție și analogie, sau dedus (derivat) din teorie.

Înainte de a stabili un model matematic sau statistic cercetătorul trebuie să se bazeze pe existența unui model semantic sau conceptual.

De exemplu, în educație, modelele conceptuale descriu teorii și ipoteze despre școală, învățare, schimbări în dezvoltare, efecte ale resurselor, într-o formă simplificată. Este de multe ori posibil să se translateze aceste modele formalizate în modele statistice.

În Rotariu (1991) se argumentează utilitatea introducerii metodelor matematice în cercetarea sociologică: Chiar dacă sociologia nu va realiza profituri explicative deosebite, se pot înregistra influențe pozitive, pe linia precizării conceptelor, datorită rigorii limbajului matematic.

Folosirea metodelor statistico-matematice în investigarea domeniului socio-uman datează de la începutul secolului. Reputați cercetători au investigat atât posibilitatea măsurării variabilelor cât și exprimarea relațiilor socio-umane în forma matematică.

De exemplu, măsurarea gradului de inteligență (Sperman, 1904), a atitudinii (Thurstone, 1928); reprezentarea matematică a principiilor învățării (Hull, 1943) și ale psihologie sociale (Lewin, 1936).

Începând cu anii 50, cercetarea în acest domeniu s-a intensificat, beneficiind de dezvoltarea metodelor matematice și de apariția calculatoarelor electronice.

În cercetarea sociologică românească au fost folosite metode automate de calcul și analiză statistică a datelor, începând din anii 70. Fără a fi exhaustivă, enumerarea poate începe cu Rotariu, 1974, Chelcea și colab., 1975, Zamfir și Ștefănescu, 1977, și continuă cu Zamfir, coord. 1995 și Sandu, 1996.

Reprezentarea fenomenelor sociologice și comportamentale nu este simplă. Aceasta presupune izolarea unui set particular de fenomene empirice, selectarea unui sistem particular formal și apoi stabilirea unei corespondențe între ele. Așa cum mai multe limbi pot fi folosite pentru a exprima aceeași idee, mai multe modele formale pot fi folosite pentru a reprezenta același fenomen empiric.

Caracteristica de bază a tuturor modelelor este reprezentarea unor aspecte ale lumii reale într-o formă abstractă ; astfel, procesul de modelare se mai numește și abstractizare.

În aplicarea unui model, investigatorul identifică obiecte și relații din lumea reală, cu anumite elemente și relații din sistemul formal.

Dându-se o interpretare specifică modelului, consecințele sale logice sunt comparate cu datele pentru a determina gradul de corespondență dintre model și realitate. Această etapă a procesului de cercetare implică o interpretare conceptuală dar și una statistică. În mod obișnuit, datele conduc la modificarea sau extinderea modelului original.

Procedurile statistice care vor fi abordate în capitolele următoare sunt proiectate pentru analiza relațiilor cauzale. Aceste tehnici testează consistența diferitelor modele și estimează mărimea efectelor cauzale.

I.7 VARIABILE LATENTE

Variabilele latente sunt caracteristici care pot fi definite conceptual dar nu pot fi măsurate direct. În științele sociale, elaborarea legilor cauzale se bazează deseori pe analizarea unor variabile teoretice care nu pot fi măsurate direct.

De exemplu, cum ar putea fi măsurată “ *ambiția* “ sau “ *nivelul aspirațiilor* “ ? Teoria abstractă a calității vieții a condus la acceptarea faptului că acest fenomen nu poate fi măsurat, cuantificat direct, ci prin intermediul unor variabile cauzale măsurabile.

Majoritatea teoriilor despre variabile latente pornesc de la presupunerea că variația sau comportamentul unui fenomen se poate măsura cantitativ, în una sau mai multe dimensiuni.

Astfel se definesc variabile manifeste măsurabile cu scopul de a crea măsuri pentru variabile latente.

În Cooley și Lohnes (1976) se expun dificultățile de a identifica și defini variabilele manifeste, și de a determina atât importanța lor relativă cât și relația dintre variabilele manifeste și cele latente. Aceste discuții au fost continuate de Bartholomew (1987), J. Scott Long (1983), Bolen (1992).

Premisa de bază în studiul variabilelor latente este : covariația empirică observată între variabilele manifeste (observate) se datorează legăturii dintre fiecare variabilă manifestă și variabila latentă - adică, variabila latentă "explică" relația dintre variabilele observate. Dacă variabila latentă există și poate fi caracterizată, atunci, controlând această variabilă va rezulta o diminuare a covariației între variabilele observate. În consecință, se spune că variabila latentă este "adeverata" sursă a covariațiilor observate.

De exemplu, este posibil să caracterizăm o variabilă latentă - "*convingeri religioase*", dacă ea explică relațiile observate între : frecventarea bisericii, confesarea la preot, afirmarea importanței religiei.

Nivelul de aspirații (Zamfir, coord, 1984), care este o variabilă cauzală abstractă importantă poate fi definit ca variabilă latentă. Ea nu poate fi măsurată direct, ci prin intermediul unor variabile contextuale. Putem considera că nivelul de aspirații explică relațiile observate între : nivelul școlar, mediul social de proveniență, statusul socio - profesional.

Este primul exemplu, în literatura română de specialitate, în care apare definită experimental și teoretic o asemenea variabilă, deși nu a fost folosit termenul de variabilă latentă. Pentru a preveni alte confuzii menționăm că în Zamfir, 1987, p93 apare termenul de cerință funcțională latentă cu o semnificație diferită de semnificația folosită în lucrarea de față.

Sandu, 1996, propune diferite modele cauzale ale atitudinii față de reforma - economică, socială, politică și culturală; el analizează concepte definite ca variabile latente, cum ar fi : "prodemocrație", "pro-economia de piață", "modernitate acțională", "modernitate identitară", "optimism".

Analiza variabilelor latente a debutat cu analiza factorială, care caracteriza variabilele latente continue prin analizarea indicatorilor observabili continui sau dihotomici.

Analiza regresiei a continuat examinarea legăturilor dintre indicatorii măsurabili și a pus în evidență aportul acestora la variația variabilei efect.

Metoda LISREL

Joreskog și Sorbom (1979, 1981) au pus bazele unei noi analize, concentrate pe definirea variabilelor latente și pe studierea relațiilor dintre ele. Analiza utilizează structuri de covarianță și modele cu ecuații structurale liniare. Tehnica respectiva poartă numele de LISREL (Analysis of Linear Structural Relationships by maximum likelihood and least squares method - Analiza relațiilor structurale liniare cu metoda verosimilității maxime și metoda celor mai mici pătrate).

Modelul LISREL, în forma sa generală are două părți :

Modelul de măsurare, și Modelul cu ecuații structurale.

Modelul de măsurare precizează modul în care variabilele latente sau construcțiile ipotetice, depind de variabilele observate.

Modelul cu ecuații structurale specifică relațiile cauzale între variabilele latente, descrie efectele cauzale și atașează variantele (dispersiile) explicate și neexplicate.

Metoda LISREL estimează coeficienții necunoscuți ai unui set de ecuații structurale liniare. Este special proiectată să adapteze modele care includ variabile latente, erori de măsurare atât pentru variabilele dependente cât și independente, cauzare reciprocă, simultaneitate și interdependență.

LISREL este și numele programului pe calculator care estimează coeficienții necunoscuți într-un set de ecuații structurale liniare. Programul poate analiza simultan date din mai multe populații.

Așa cum a fost implementată în programul LISREL, metoda include proceduri speciale ca : analiză factorială confirmatorie, regresie multiplă, analiză path, modele recursive și nerecursive, modele economice pentru date variabile în timp, și modele structurale de covarianță.

Un avantaj îl constituie modulul **PRELIS** cu care este pregătită structura datelor în vederea definirii variabilelor latente și a relațiilor cauzale. Acest program calculează covarianțe și corelații Pearson pentru datele continue, corelații polichorice pentru date ordinale, și corelații tetrachorice pentru datele dihotomice.

LISREL este o metodă foarte bună de analiză a modelelor cauzale în științele sociale, comportamentale sau ale educației.

LISREL a fost utilizat pentru validarea unor modele teoretice în domenii diferite : sociologie, educație și evaluare, psihologie, marketing și administrație, demografie, studii asupra personalității și inteligenței, medicină și biologie (Hayduk,1989; Bolen, 1993; Yli-Luoma,1990; Bagozzi,1994)

II. METODOLOGIA MODELELOR CAUZALE

În acest capitol vor fi tratate aspectele privind complexa problematică a modelelor cauzale, având ca suport lucrări de referință în domeniu : Joreskog & Sorbom (1977,1982,1990,1993), Scott Long (1983), Berry(1984), Bollen(1989), Hayduk (1989), Bollen și Long (1993), von Eye și Clogg (1994).

Stilul prezentării este adaptat științelor socio-umane utilizând un aparat matematic accesibil și un număr impresionant de cazuri explicative.

II.1. METODOLOGIA DE ANALIZĂ A DATELOR

1.1. Modele de analiză a datelor

În cercetările sociologice ne lovim adesea de o dificultate majoră, anume prelucrarea unui volum impresionant de informații.

În acest context se preferă aplicarea unor proceduri specifice statisticii matematice. Astfel sunt proiectate modele de analiză a datelor cu intenția de a surprinde posibilele interacțiuni dintre variabilele lor.

Printre procedeele statistice utilizate curent menționăm procedura de regresie multiplă, analiza de varianță, determinarea corelației canonice și a corelației parțiale, analiza de discriminare, analiza factorială și procedurile de clasificare.

În continuare, vor fi puse în evidență etapele ce trebuiesc de regulă parcurse în procesul necesar de cunoaștere a realității.

Cazul 1. Stabilirea interdependențelor dintre variabile

Să presupunem existența variabilelor X, U, V, W ce exercită, fiecare în parte, o influență asupra unei alte variabilei Y . Ne propunem să determinăm intensitățile acestor influențe.

Așadar modelul teoretic va cuprinde ansamblul de variabile X, U, V, W, Y împreună cu interacțiunile dintre ele.

Acest lucru poate fi reprezentat global printr-un graf orientat (Figura 1.1). Un arc orientat de la variabila W la variabila Y semnifică o influență pe care o exercită variabila W asupra variabilei Y.

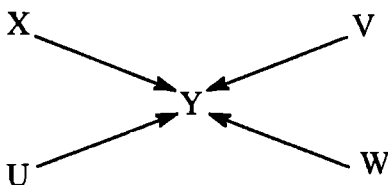


Figura 1.1.



Figura 1.2.

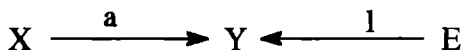


Figura 1.3.

Pentru modelul din Figura 1.1 este important de a stabili în ce măsură variabila Y este influențată de grupul de variabile X, U, V, W. Acest lucru este dificil de realizat în momentul în care nu deținem suficiente informații despre variabilele în cauză.

În plus dependențele descrise în Figura 1.1 pot îmbrăca forme variate, ca de pildă :

$$Y = h(X, U, V, W) \quad (1)$$

$$Y = h_1(X, U) + h_2(V, W)$$

$$Y = h_1(X) + h_2(U) + h_3(V) + h_4(W)$$

$$Y = h_1(X) \cdot h_2(U) + h_3(V) \cdot h_4(W)$$

unde de cele mai multe ori forma funcțiilor h, h₁, h₂, h₃, h₄ nu este cunoscută.

Surmontarea acestor dificultăți ar putea fi realizată prin impunerea unor restricții suplimentare. În acest caz, în locul modelului teoretic vom

avea un model empiric, simplificat, model care va trebui să surprindă esența raporturilor din modelul teoretic.

De regulă modelul teoretic are un grad mare de complexitate. Se va renunța la unele variabile ce nu sunt definitorii pentru el, precum și la acele relații dintre variabile ce nu prezintă o prea mare importanță. În acest mod, se obține un model simplificat, care să simuleze totuși realitatea: model empiric.

În cazul concret enunțat anterior să acceptăm că variabilele U, V, W nu au o prea mare influență asupra variabilei Y . În această nouă situație modelul empiric ar putea fi cel sugerat de figura 1.2.

Scopul esențial urmărit era de a determina gradul de influențare exercitat de ansamblul de variabile X, U, V, W asupra variabilei Y . Dar acest lucru este greu de prognozat chiar și în cazul modelului empiric din figura 1.2 deoarece nu știm forma concretă h după care variabila X acționează asupra variabilei Y ,

$$Y = h(X) \quad (2)$$

Suntem așadar constrânși în a propune noi restricții. Astfel putem presupune o formă liniară a funcției h , adică

$$Y = a \cdot X \quad (3)$$

valoarea coeficientului a semnificând puterea de influențare a variabilei X asupra variabilei Y . Din relația (3) rezultă că o creștere a lui X cu o unitate antrenează o variație a lui Y cu a unități. Parametrul a poate lua și valori negative.

Modelul empiric 1.2 va fi mai bine definit prin introducerea unei noi variabile E care să compenseze faptul că setul de variabile U, V, W a fost neglijat într-o primă variantă.

În noua sa variantă din figura 1.3, modelul empiric este definit prin ecuația

$$Y = a \cdot X + E \quad (4)$$

Cum eroarea E este, de regulă, independentă de variabila X , impunem condiția $\text{Cov}(X, E) = 0$. În plus, $\text{Med}(E) = 0$, erorile compensându-se. Modelul (4) cu aceste restricții conduce la modelul clasic al regresiei liniare cu o singură variabilă independentă X , variabila dependentă fiind Y .

Cazul 2. Reducerea dimensiunii modelului

Tehnicile de analiză factorială pot fi folosite pentru reducerea dimensiunii spațiului variabilelor analizate. Dacă într-un model se află un număr mare de variabile : X_1, X_2, \dots, X_n , acestea pot fi înlocuite cu un număr mai mic de factori F_1, F_2, \dots, F_m ($m < n$). Fiecare dintre variabilele inițiale va fi dependentă, mai mult sau mai puțin, de acești factori.

Pentru simplificare vom considera un număr de patru variabile X_1, X_2, X_3, X_4 pe care ne propunem să le exprimăm în funcție de numai doi factori F_1, F_2 .

Așadar avem următoarele relații :

$$\begin{aligned} X_1 &= h_1(F_1, F_2) \\ X_2 &= h_2(F_1, F_2) \\ X_3 &= h_3(F_1, F_2) \\ X_4 &= h_4(F_1, F_2) \end{aligned} \quad (5)$$

unde atât funcțiile h_1, h_2, h_3, h_4 cât și factorii F_1, F_2 trebuiesc determinați.

Admițând liniaritatea funcțiilor h_j , $1 \leq j \leq 4$, relațiile (5) devin :

$$\begin{aligned} X_1 &= a_{11} \cdot F_1 + a_{12} \cdot F_2 \\ X_2 &= a_{21} \cdot F_1 + a_{22} \cdot F_2 \\ X_3 &= a_{31} \cdot F_1 + a_{32} \cdot F_2 \\ X_4 &= a_{41} \cdot F_1 + a_{42} \cdot F_2 \end{aligned} \quad (6)$$

În formulele (6) va trebui să determinăm efectiv ponderile a_{ij} ale factorului F_j în variabila X_i , $1 \leq i \leq 4$, $1 \leq j \leq 2$.

Modelul (6) este reprezentat grafic în figura 1.4.

Egalitățile (6) nu sunt de cele mai multe ori satisfăcute, semnul “egal” fiind folosit în sensul de “aproximativ egal”.

În cazul în care se va utiliza în (6) semnul “egal” în adevăratul său înțeles, vom introduce variabilele E_j , $1 \leq j \leq 4$, ce semnifică erorile făcute în procesul de aproximare.

Așadar relațiile (6) se vor rescrie astfel :

$$\begin{aligned} X_1 &= a_{11} \cdot F_1 + a_{12} \cdot F_2 + E_1 \\ X_2 &= a_{21} \cdot F_1 + a_{22} \cdot F_2 + E_2 \\ X_3 &= a_{31} \cdot F_1 + a_{32} \cdot F_2 + E_3 \\ X_4 &= a_{41} \cdot F_1 + a_{42} \cdot F_2 + E_4 \end{aligned} \quad (7)$$

Reprezentarea grafică a modelului factorial (7) este dată în figura 1.5.

În modelul (7) factorii F_1 și F_2 sunt variabile fictive. Ele exprimă caracteristici comune prezente în variabilele X_1 - X_4 fără însă a avea întotdeauna posibilitatea de a le defini precis.

Factorii F_1, F_2 nu sunt direct măsurabili, fiind observabili numai în mod indirect prin intermediul măsurătorilor efectuate asupra variabilelor X_1, X_2, X_3, X_4 .

1.2. Soluționarea modelului empiric

Să încercăm să soluționăm modelul de regresie (4) (Figura 1.3) în ipoteza clasică $Med(E) = 0, Cov(X, E) = 0$.

Așadar vom estima valoarea parametrului a și varianța erorii E (adică $Var(E)$), aceștia fiind de fapt cei doi parametri necunoscuți ai modelului (4).

În vederea simplificării calculelor vom presupune

$$Med(X) = 0 \quad (8)$$

fapt ce va antrena obligatoriu, din egalitatea (4),

$$Med(Y) = 0 \quad (9)$$

aceasta deoarece s-a acceptat și ipoteza $Med(E) = 0$.

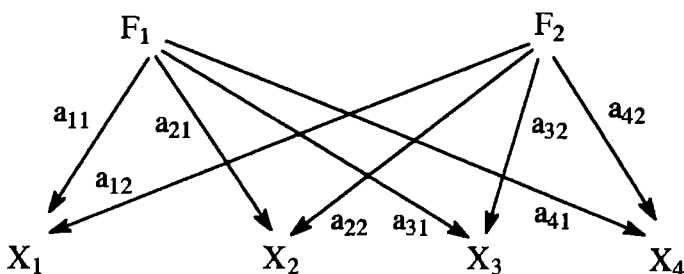


Figura 1.4.

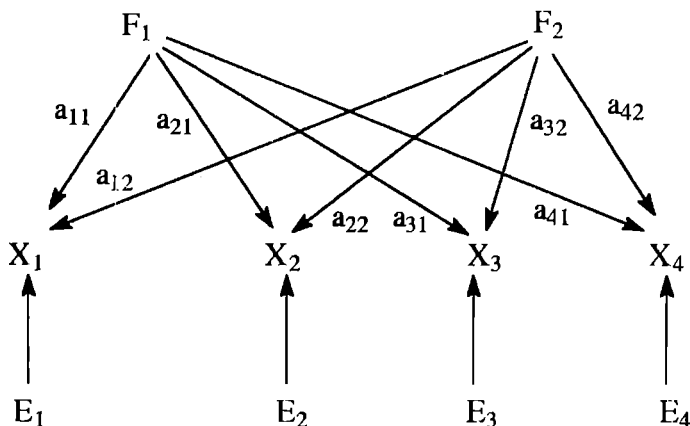


Figura 1.5.

Esențial în evaluarea modelului empiric (4) este de a obține o estimare a parametrului a , varianța erorilor E devenind în acest caz cunoscută, fiind estimată în funcție de a :

$$\text{Var}(E) = \text{Var}(Y) - a^2 \cdot \text{Var}(X) \quad (10)$$

aceasta deoarece variabilele X și E au fost presupuse de covarianță nulă.

Varianțele variabilelor X , respectiv Y vor fi deduse din datele experimentale x_1, x_2, \dots, x_k , respectiv y_1, y_2, \dots, y_k .

Tinând seamă de restricțiile (8) și (9), în locul observațiilor x_j, y_j , $1 \leq j \leq k$, se vor utiliza valorile $\underline{x}_j, \underline{y}_j$:

$$\underline{x}_j = x_j - x_+ \quad (11)$$

$$\underline{y}_j = y_j - y_+$$

unde

$$x_+ = [x_1 + x_2 + x_3 + \dots + x_k] / k \quad (12)$$

$$y_+ = [y_1 + y_2 + y_3 + \dots + y_k] / k$$

În acest mod mediile empirice ale seturilor de valori $\underline{x}_1, \underline{x}_2, \underline{x}_3, \dots, \underline{x}_k$, respectiv $\underline{y}_1, \underline{y}_2, \underline{y}_3, \dots, \underline{y}_k$ sunt nule, în concordanță cu restricțiile (8) și (9).

Operarea cu perechile de valori (x_j, y_j) , $1 \leq j \leq k$, în locul perechilor (x_j, y_j) nu reduce generalitatea problemei. În fond, s-au ales originile scalelor de măsurare pentru variabilele X și Y chiar în punctele x_+ , respectiv y_+ , definite de (12).

Modelul (4) rescris pentru fiecare individ j conduce la :

$$y_j = a \cdot x_j + e_j, \quad 1 \leq j \leq k \quad (13)$$

unde e_j este eroarea rezultată în cazul j , când în locul valorii reale y_j a variabilei Y este prezisă valoarea $a \cdot x_j$.

Parametrul a va fi determinat astfel încât suma pătratelor tuturor erorilor să fie minimă :

$$e_1^2 + e_2^2 + e_3^2 + \dots + e_k^2 = \text{minim} \quad (14)$$

fapt ce revine la minimizarea funcției

$$g(a) = [y_1 - a \cdot x_1]^2 + [y_2 - a \cdot x_2]^2 + \dots + [y_k - a \cdot x_k]^2 \quad (15)$$

Trinomul de gradul al doilea $g(a)$ se rescrie în forma

$$g(a) = C_{xx} \cdot a^2 - 2 \cdot C_{xy} \cdot a + C_{yy} \quad (16)$$

unde

$$\begin{aligned} C_{xx} &= [x_1 \cdot x_1 + x_2 \cdot x_2 + x_3 \cdot x_3 + \dots + x_k \cdot x_k] / k \\ C_{xy} &= [x_1 \cdot y_1 + x_2 \cdot y_2 + x_3 \cdot y_3 + \dots + x_k \cdot y_k] / k \\ C_{yy} &= [y_1 \cdot y_1 + y_2 \cdot y_2 + y_3 \cdot y_3 + \dots + y_k \cdot y_k] / k \end{aligned} \quad (17)$$

Din considerente elementare de statistică și ținând seama de formulele (11) deducem că valorile C_{xx} , C_{xy} , C_{yy} sunt estimații pentru $\text{Var}(X)$, $\text{Cov}(X, Y)$, respectiv $\text{Var}(Y)$.

Este știut faptul că valoarea minimă a trinomului de gradul al doilea $g(a)$ se realizează pentru

$$a = C_{xy} / C_{xx} \quad (18)$$

Având deja estimată valoarea parametrului a și aplicând formula (10) se obține imediat estimația celui alt parametru necunoscut al modelului (4), anume $\text{Var}(E)$. Variația variabilei E ne dă mărimea erorii pe care o acceptăm atunci când modelul teoretic din figura 1.1 este analizat prin intermediul modelului empiric dat de ecuația de regresie (4) (figura 1.3).

Menționăm complexitatea calculelor legate de estimarea parametrilor necunoscuți, fapt evidențiat chiar și în cazul celui mai simplu model, acela al ecuației de regresie liniară cu o singură variabilă independentă.

Sugerăm și o altă posibilitate de abordare, mai simplă și unitară, bazată pe utilizarea operatorului de covarianța.

O asemenea tehnică va fi curent folosită pentru modele mult mai generale.

Astfel, din relația $Cov(X, E) = 0$, aplicând (4) obținem

$$\begin{aligned} Cov(X, Y) &= Cov(X, a \cdot X + E) = a \cdot Cov(X, X) + Cov(X, E) = \\ &= a \cdot Cov(X, X) = a \cdot Var(X) \end{aligned} \quad (19)$$

egalitate ce ne conduce la o prognoză a valorii a ,

$$a = Cov(X, Y) / Var(X). \quad (20)$$

Folosind estimațiile C_{xy} și C_{xx} pentru $Cov(X, Y)$, respectiv $Var(X)$, din relația (20) obținem imediat estimația (18). Am ajuns astfel mult mai rapid și logic la estimația (18).

Un comentariu similar ar putea fi făcut în cazul modelului empiric de analiză factorială definit anterior.

În vederea soluționării modelului empiric (7) se introduc, de regulă, ipoteze suplimentare referitoare la factori (de exemplu independența lor), cât și ipoteze privind erorile E_1, E_2, E_3, E_4 (ca de pildă: media erorilor este nulă, erorile sunt independente între ele și sunt necorelate în raport cu factorii).

Remarca 1. Trebuie subliniat faptul că estimarea parametrilor unui model empiric se bazează pe datele colectate. Utilizarea de fiecare dată a acestora ne conduce atât la un volum uriaș de calcule cât și la repetarea unor proceduri de calcul similare. Pentru evitarea acestor neajunsuri se preferă folosirea datelor primare numai pentru deducerea valorilor elementelor matricii de covarianță-varianță atașată variabilelor modelului. Ulterior vor fi utilizate numai aceste matrici de covarianță, renunțându-se la folosirea directă a datelor primare x_j, y_j , din eșantionul de volum k .

În acest mod scade simțitor volumul de calcule. În plus această tehnică este în concordanță cu scopul urmărit, anume acela al determinării intensității dependențelor dintre variabilele modelului.

În acest context, folosirea cu prioritate a operatorului de covarianță este adecvată măsurării gradului de dependență liniară dintre oricare două variabile aleatoare.

Optăm astfel pentru o tratare calitativă a datelor, în concordanță cu scopul propus.

Remarca 2. Deși prin folosirea matricii de covarianță datele primare nu mai sunt direct utilizate ulterior, totuși numărul lor prezintă o importanță majoră în stabilirea acurateții estimărilor. Precizia rezultatelor crește odată cu mărirea volumului k al eșantionului studiat.

1.3. Validarea modelului empiric

De multe ori operaționalizarea modelului teoretic este imposibil de realizat datorită complexității sale. Prin urmare ne limităm la un model empiric, relativ simplu, care să surprindă trăsăturile esențiale ale modelului teoretic și totodată acesta să poată fi și efectiv soluționat.

Prin intermediul modelului empiric vom explica indirect modelul teoretic.

Apar în mod natural, cel puțin, următoarele întrebări :

- Modelul empiric aproximează “destul de bine” modelul teoretic?
- Cum putem fi “siguri” de acest lucru ?
- Care ar fi procedurile de validare ale modelului empiric ?

Vom căuta să sugerăm răspunsuri la întrebările enunțate, comentariile fiind făcute pentru modelele simple deja abordate.

Astfel în cazul 1, dacă varianța erorii E “este prea mare” atunci putem afirma cu siguranță că modelul empiric (4) (figura 1.3) nu este adecvat modelului teoretic din figura 1.2

În această situație, probabil modelul empiric ar trebui să mai conțină și alte variabile (U , V sau W) în afara variabilei X , luată inițial în considerare.

De asemenea pot fi imaginate și alte posibile abordări. Astfel introducerea unor noi variabile nu este întotdeauna necesară. O discuție separată ar trebui făcută privind forma concretă a funcției h ce apare în ecuația de structură (2). Nu este obligatoriu ca funcția h să aibă o formă liniară. În plus ar putea exista o corelație între variabila X și eroarea E .

Fiecare model empiric astfel propus va fi validat.

O discuție asemănătoare, dar cu mai multe variante de abordare, trebuie să fie făcută pentru Cazul 2.

În acest ultim caz suma d a varianțelor erorilor E_j corespunzătoare ecuațiilor de structură (7),

$$d = \text{Var}(E_1) + \text{Var}(E_2) + \text{Var}(E_3) + \text{Var}(E_4) \quad (21)$$

ar putea constitui o măsură a gradului de fidelitate a modelului experimental în încercarea de simulare a modelului real.

De fapt, folosirea unei măsuri de tip (21) la validarea modelului empiric ar trebui nuanțată. Astfel am putea accepta ponderile p_1, p_2, p_3, p_4 privind importanța specifică pe care o acordăm fiecărei ecuații de structură (7). În această situație măsura d a veridicității modelului experimental este dată de formula :

$$d = p_1 \cdot \text{Var}(E_1) + p_2 \cdot \text{Var}(E_2) + p_3 \cdot \text{Var}(E_3) + p_4 \cdot \text{Var}(E_4)$$

Atunci când d ia valori “mari” nu se poate spera la o bună modelare a realității. Aceasta s-ar datora, de exemplu, unui număr neadecvat de factori F , caz în care vom recurge la o mărire a numărului lor.

Este posibil ca factorii modelului teoretic să prezinte anumite caracteristici specifice care nu au fost sesizate în modelul experimental. Astfel, nu toți acești factori sunt factori generali, o parte dintre aceștia putând avea caracteristici proprii unui grup restrâns de variabile.

La aceste ipoteze se mai adaugă o posibilă corelare a factorilor (aparitia factorilor oblici). În plus, rămâne discutabilă acceptarea în modelul factorial din figura 1.5 a unor dependențe liniare ale variabilelor în funcție de presupușii factori. De fapt, prin ecuațiile de structură (7) s-a postulat existența unor dependențe liniare fără însă a efectua în prealabil teste confirmatorii în acest sens.

În încercarea de apreciere a raporturilor reale din cadrul modelului teoretic se va desfășura o muncă uriașă pentru propunerea mai multor modele empirice, soluționarea lor efectivă, analizarea lor separată și efectuarea unor studii comparative.

Este necesar de a evidenția și latura neplăcută a efortului deosebit ce trebuie depus în operațiunea de cercetare a realității. Astfel, toate

aceste operații complexe pot să ne conducă la unele soluții greșite. Ne putem înșela în selectarea unui model experimental. De pildă pot rezulta modele empirice “plauzibile”, modele ce sunt necontradictorii în raport cu datele experimentale culese, dar aberante din punctul de vedere al interpretării dependențelor cauzale pe care le prognozează.

1.4. Necesitatea unui cadru general

Cele doua modele statistice discutate în detaliu în secțiunea 1.1 se remarcă prin simplitatea lor. Modelele reale sunt complexe și dificil de soluționat efectiv prin utilizarea unui aparat matematic elementar.

Apare necesară abordarea acestor modele într-un cadru general care să permită o tratare unitară a lor.

Vom enumera pe scurt caracteristici metodologice ce se pretează la o astfel de generalizare.

Scopul principal urmărit este determinarea gradului de dependență dintre diferitele variabile ale modelului teoretic. În vederea rezolvării acestei probleme, modelul teoretic este simplificat propunându-se un model empiric. Modelul empiric va trebui să rețină caracteristicile esențiale ale modelului teoretic.

Soluționarea modelului empiric este, de regulă, dificil de realizat. Din această cauză se introduc ipoteze suplimentare (privind forma funcțiilor de dependență, restricții impuse variabilelor și erorilor de măsurare). Este esențial ca toate aceste noi ipoteze să nu contrazică perceperea noastră asupra modelului teoretic. Ipotezele introduse au ca scop principal tocmai simplificarea procedurilor de soluționare efectivă a modelului empiric.

Menționăm faptul că rezolvarea modelului empiric se realizează pornindu-se de la măsurătorile efectuate asupra indicatorilor. Soluționarea modelului empiric se reduce la estimarea coeficienților ecuațiilor structurale respective, fapt ce presupune un aparat matematic dezvoltat, precum și utilizarea calculatoarelor electronice.

Ca tehnică de abordare a modelului cauzal menționăm rolul prioritar ce trebuie acordat aplicării operatorului de covarianță.

Odată cu soluționarea modelului empiric se va trece la validarea sa. Se are în vedere o apropiere cât mai mare a modelului empiric de modelul real.

Toată această problematică este prezentă atât în analiza de regresie cât și în analiza factorială. Procedurile de soluționare sunt similare.

Prin urmare este necesară introducerea unui concept mai general care să cuprindă laolaltă cele două analize statistice. Această "unificare" este realizată de modelul causal structural abordabil prin una din cele două forme : ecuațiile structurale sau graful causal (cap. II.2).

Concluzii

Se impune o abordare unitară și comparativă a diferitelor procedee de analiză statistică a datelor experimentale.

Vom enumera câteva astfel de proceduri statistice : previziune și evaluarea tendinței de evoluție (ecuații de regresie), stabilirea unor indicatori pentru seturi de variabile (analiză factorială și analiză discriminantă, corelație canonică) determinarea gradului maxim de dependență liniară dintre grupuri specificate de variabile (corelație canonică), stabilirea intensității legăturilor dintre grupuri de variabile și în interiorul grupurilor (analiză de varianță), gruparea indivizilor unui eșantion dat în raport de anumite caracteristici comune, specificate (analiză de clasificare), stabilirea deosebirilor dintre diversele clase de variabile și selectarea acelor variabile ce sunt definitorii pentru fiecare clasă (analiză factorială, analiză discriminantă).

Metodologia generală (model teoretic, model experimental, soluționarea și validarea modelului empiric) este prezentă în toate procedeele statistice enumerate.

Introducerea unui concept mai general care să cuprindă mai multe tipuri de analize statistice va fi făcută în secțiunea următoare.

II.2. NOȚIUNEA DE MODEL CAUZAL

2.1. Reprezentarea modelelor cauzale

Modelul cauzal constă în precizarea unui sistem de ecuații structurale ce definesc legăturile dintre variabilele modelului.

Un model cauzal mai poate fi reprezentat printr-un graf ale cărui noduri sunt variabilele modelului. Un arc orientat de la nodul 1 la nodul 2 simbolizează o relație cauzală între aceste două noduri, cea de a doua variabilă constituind efectul influenței primei variabile. Evidențierea gradului de influență a primei variabile asupra celei de a doua variabile este realizată prin coeficienții ecuației structurale respective sau prin ponderea atribuită în graful cauzal arcului orientat ce leagă cele două variabile.

Precizăm faptul că o măsură a dependenței dintre cele două variabile (de exemplu coeficientul de corelație Pearson) dă intensitatea relației de dependență (liniară) dintre aceste variabile fără a indica însă aportul cauzal al fiecărei variabile tratate separat.

Ne putem imagina diferite relații cauzale între variabilele modelului urmând ca din datele experimentale să obținem o confirmare sau o negare a ipotezei făcute.

Metodologia de “diagnostic” va fi discutată pe larg în secțiunile următoare punându-se totodată în evidență aspecte ce contribuie la selectarea dintr-o multitudine de relații cauzale posibile a acelor relații ce sunt adecvate modelului teoretic.

În cadrul modelului cauzal (empiric sau teoretic) pot fi prezente trei tipuri de variabile: variabile latente, variabile observabile precum și variabile-eroare. O variabilă-eroare poate înlocui comportarea unui grup întreg de variabile, grup ce nu a fost inclus explicit în model. O variabilă de acest tip sugerează în ultimă instanță și nivelul de “credibilitate” al relației de dependență dintre variabilele definite ale modelului studiat. Variabila-eroare include o multitudine de alți factori ce au rămas neprecizați în model. În cazul în care variabilele eroare au o influență majoră se impune îmbunătățirea modelului propus inițial prin precizarea și a altor relații dintre variabilele modelului cât și prin introducerea

unor variabile suplimentare. În acest fel modelul va deveni “mai precis” diminuându-se totodată ponderea factorilor necunoscuți în explicarea dependențelor cauzale dintre variabile.

2.2. Model cu variabile latente

Variabila latentă reprezintă un concept abstract. O variabilă latentă este o variabilă neobservabilă direct (ci numai indirect prin intermediul altor variabile), adică o variabilă nemăsurabilă sau o variabilă ce presupune un grup de factori (de multe ori nu îndeajuns de clar definit). În cercetările sociologice întâlnim variabile ca: inteligența, clasa socială, conceptul de putere, perspective, industrializare, democratizare politică, variabile ce răspund caracterizării enunțate anterior. Aceste variabile sunt variabile latente comportând un grad mai mare sau mai mic de abstractizare.

Variabilele observabile pot fi măsurabile direct. Astfel venitul, vârsta, gradul de educație, numărul de indivizi ai unui grup sunt variabile ce se măsoară efectiv. Aceste variabile “observabile” reprezintă de fapt indicatori pentru variabilele latente, variabilele latente nefiind “direct observabile”.

Variabilele (latente și observabile) ale modelului cauzal mai pot fi clasificate în variabile endogene, proprii modelului studiat și variabile exogene, exterioare modelului (vezi și I.5). Această clasificare este relativă, specifică unui anumit model, rolul variabilelor în model fiind determinat în funcție de interpretările pe care aceste variabile le capătă în cadrul modelului. Variabile endogene într-un model pot deveni variabile exogene în alt model și reciproc. Pentru exemplificare, vom expune cazul următor:

Cazul 1. Să studiem un model, adesea abordat, în care intervin următoarele variabile latente:

W_1 = democrația politică în anul 1985;

W_2 = democrația politică în anul 1995;

U_1 = industrializarea în anul 1985.

Vom încerca să precizăm relațiile cauzale dintre aceste variabile latente adică ecuațiile structurale ale modelului.

Gradul de industrializare al anilor 1985 afectează democratizarea politică a acelor ani și va influența totodată nivelul democratizării politice a anilor 1995. Democratizarea politică din 1995 este, cel puțin parțial, un rezultat al gradului de democratizare al anilor 1985.

Intenționăm să studiem nivelul de dezvoltare al unei țări în raport cu democrația politică și industrializarea sa. Ținând cont de cele precizate putem accepta relațiile cauzale h_1 și h_2 între variabilele latente W_1, W_2, U_1 ale modelului, adică :

$$\begin{aligned} W_1 &= h_1(U_1) \\ W_2 &= h_2(U_1, W_1) \end{aligned} \quad (1)$$

Deoarece fenomenul studiat este evoluția democratizării politice prin intermediul variabilelor endogene W_1 și W_2 atunci variabila U_1 este o variabilă exogenă.

Modelul descris prin ecuațiile de structură (1) este un model simplificat. Variabilele endogene W_1, W_2 sunt numai parțial "explicate" de relațiile (1) ele fiind influențate de fapt și de alți factori ce nu au fost specificați efectiv în model. Vom nota prin V_1, V_2 ansamblul de factori neprecizați ce exercită o influență asupra variabilelor endogene W_1 , respectiv W_2 . În acest caz putem aduce următoarea îmbunătățire ecuațiilor structurale (1) ținând seamă și de contribuția factorilor perturbatori V_1, V_2 :

$$\begin{aligned} W_1 &= h_1(U_1, V_1) \\ W_2 &= h_2(U_1, W_1, V_2) \end{aligned} \quad (2)$$

Din punct de vedere statistic variabilele latente V_1, V_2 pot fi considerate variabile aleatoare de medie zero, ele fiind necorelate cu variabila exogenă U_1 . Variabilele V_1 și V_2 pot reprezenta un ansamblu de factori diferiți de U_1 ce influențează variabilele endogene ale modelului.

Soluționarea modelului (2) constă în determinarea intensității relațiilor cauzale h_1, h_2 (precizarea formei funcțiilor h_1, h_2). De regulă această problemă este dificil de abordat datorită necunoașterii formei concrete pentru funcțiile h_1, h_2 . Chiar în cazul selectării dintr-o infinitate de posibilități a tipului particular de funcție ce va fi luată în considerare, apar probleme legate de interpretarea concretă a parametrilor acestor funcții.

De exemplu pentru funcțiile h_1, h_2 de tip exponențial,

$$\begin{aligned} h_1(U_1, V_1) &= \exp(a_{11} \cdot U_1 + a_{12} \cdot V_1) \\ h_2(U_1, W_1, V_2) &= \exp(a_{21} \cdot U_1 + a_{22} \cdot V_2 + b_{21} \cdot V_1) \end{aligned} \quad (3)$$

este dificil de stabilit semnificația reală a parametrilor a_{ij}, b_{ik} . Din aceste considerente cât și din motive legate de simplificarea procedurilor de estimarea parametrilor (a_{ij}, b_{ik}) proprii funcțiilor de dependență cauzală h_1, h_2 , se va prefera în cele ce urmează o formă liniară pentru această dependență.

În această ipoteză ecuațiile structurale (2) din Cazul 1 vor avea următoarea formă:

$$\begin{aligned} W_1 &= a_{11} \cdot U_1 + V_1 \\ W_2 &= a_{21} \cdot U_1 + b_{21} \cdot W_1 + V_2 \end{aligned} \quad (4)$$

Acceptând convențiile operațiilor cu vectori și matrici, ecuațiile structurale (4) pot fi rescrise matriceal astfel:

$$W = A \cdot U + B \cdot W + V \quad (5)$$

În relația (5) prin W, U, V înțelegem vectorii coloană ce conțin cele m variabile latente endogene, cele n variabile latente exogene, respectiv cele m variabile latente perturbatoare, neprecizate (factori sau "erori de aproximare") ce completează "explicarea" variabilelor latente endogene. Deci:

$$\begin{aligned} W &= (W_1, W_2, W_3, \dots, W_m)' ; \quad U = (U_1, U_2, U_3, \dots, U_n)' ; \\ V &= (V_1, V_2, V_3, \dots, V_m)' ; \end{aligned}$$

Matricile $A = (a_{ij})_{i,j}, 1 \leq i \leq m, 1 \leq j \leq n$, $B = (b_{ik})_{i,k \leq m}$, semnifică intensitatea influenței cauzale a variabilelor latente exogene asupra variabilelor latente endogene, respectiv influența reciprocă dintre variabilele (latente) endogene.

În cazul concret al modelului liniar sugerat de Cazul 1 avem:

$$\begin{aligned} m &= 2 ; \quad n = 1 ; \quad W = (W_1, W_2)' ; \quad U = (U_1) ; \\ V &= (V_1, V_2)' ; \quad A = (a_{11}, a_{21})' ; \\ B &= \begin{vmatrix} 0 & 0 \\ b_{21} & 0 \end{vmatrix} \end{aligned}$$

Remarca 1. Reamintim că apostroful adăugat după un vector sau matrice semnifică transpunerea vectorului sau matricii respective (în tabloul respectiv liniile devin coloane iar coloanele linii). Vectorii vor fi considerați ca fiind vectori coloană. Prin transpunerea unui vector coloană obținem vectorul linie având aceleași componente ca și vectorul coloană. Atunci când dorim să punem în evidență componentele unui vector X (el fiind considerat de regulă un vector coloană) vom prefera, pentru economisire de spațiu, enumerarea componentelor vectorului $\text{transp } X'$ ce este de fapt un vector linie.

Odată precizat modelul (cu variabile latente), problema majoră ce va trebui abordată în continuare va consta în soluționarea modelului cauzal liniar (5).

Aceasta revine la determinarea efectivă a valorilor elementelor matricilor A și B , matrici ce semnifică intensitatea dependențelor cauzale dintre variabilele modelului. Menționăm dificultatea soluționării modelului cauzal (5), model ce conține în acest caz numai variabile latente, adică variabile ce nu pot fi măsurate direct.

2.3. Model cu variabile observabile

Datorită imposibilității măsurării directe a variabilelor latente ale modelului (5), vom fi nevoiți să apreciem indirect aceste variabile prin intermediul altor variabile ce sunt însă măsurabile.

Variabilele observabile, socotite variabile ce se pot manifesta, au o multitudine de denumiri, ca de exemplu : indicatori, variabile măsurabile sau variabile “împuternicite” (mandatate) de variabile latente pentru a le reprezenta.

Analog modelului cu variabile latente, poate fi conceput un model cu variabile măsurabile, în locul fiecărei variabile latente fiind trecuți indicatori. Astfel variabilele latente exogene devin variabile măsurabile independente X , iar locul variabilelor latente endogene este preluat de variabilele măsurabile dependente Y .

În urma discuției din secțiunea 2.2 am optat pentru o comportare liniară a relației de dependență dintre variabilele observabile.

În acest mod, variabila latentă exogenă U_1 (“grad de industrializare”) din Cazul 1 va putea fi mandatată (reprezentată) prin comportamentul uneia sau mai multor variabile independente măsurabile, ca de pildă :

X_1 = produsul brut pe cap de locuitor;

X_2 = consumul de energie convențională pe cap de locuitor;

X_3 = procentul forței de muncă angrenată în industrie, etc.

Pentru studierea comportamentului celor două variabile latente endogene W_1, W_2 (“democratizarea politică” în 1985 și cea din 1995) pot fi folosiți, de exemplu, patru indicatori :

- libertatea presei (Y_1 în 1985, Y_5 în 1995),
- libertatea de exprimare a opoziției politice (Y_2 și Y_6)
- corectitudinea alegerilor (Y_3 , respectiv Y_7),
- activitatea legislativă curentă (Y_4 și Y_8).

Modelul poate fi detaliat. Astfel fiecare indicator poate fi la rândul său exprimat în funcție de noi indicatori sau variabile măsurabile.

O variabilă latentă poate fi mandatată de una sau mai multe variabile măsurabile (indicatori). În unele cazuri se acceptă pentru o variabilă latentă un singur indicator. În această situație comportamentul variabilei latente se identifică cu acela al indicatorului atașat ei.

Va trebui să existe o corelație foarte puternică între variabila observabilă (indicatorul) și variabila latentă respectivă. Această restricție este însă înlăturată dacă pentru urmărirea comportamentului aceleiași variabile latente sunt folosiți mai mulți indicatori. În ultima variantă fiecare variabilă măsurabilă va prelua o parte din influența specifică variabilei latente și deci corelația dintre o variabilă observabilă mandatată de o variabilă latentă și variabila latentă respectivă nu va trebui să fie neapărat foarte puternică.

Cu cât avem mai mulți indicatori cu atât dependența acestora de variabila latentă respectivă poate să scadă ; deci și valoarea absolută a corelației dintre un indicator și variabila latentă asociată lui va scadea.

În concluzie, vom prefera utilizarea mai multor indicatori pentru “explicarea” unei singure variabile latente.

Ecuatiile de structură pentru modelul (mixt) cu variabile măsurabile atașat modelului inițial cu variabile latente (Cazul 1) generalizează setul de ecuații (4). Pe lângă vechile legături dintre variabilele latente de finite

în (4) se introduc noi raporturi de dependență dintre variabilele latente și indicatorii asociați lor. Astfel avem :

$$\begin{aligned} X_1 &= d_1 \cdot U_1 + Z_1 \\ X_2 &= d_2 \cdot U_1 + Z_2 \\ X_3 &= d_3 \cdot U_1 + Z_3 \end{aligned} \quad (6)$$

$$\begin{aligned} Y_1 &= d_4 \cdot W_1 + E_1 & Y_5 &= d_8 \cdot W_2 + E_5 \\ Y_2 &= d_5 \cdot W_1 + E_2 & Y_6 &= d_9 \cdot W_2 + E_6 \\ Y_3 &= d_6 \cdot W_1 + E_3 & Y_7 &= d_{10} \cdot W_2 + E_7 \\ Y_4 &= d_7 \cdot W_1 + E_4 & Y_8 &= d_{11} \cdot W_2 + E_8 \end{aligned} \quad (7)$$

Variabilele observabile independente X_1, X_2, X_3 sunt trei indicatori ai variabilei latente exogene U_1 ("industrializare").

Variabilele observabile Y_1, Y_2, Y_3, Y_4 , respectiv Y_5, Y_6, Y_7, Y_8 sunt indicatori prin intermediul cărora se va aprecia (indirect) comportamentul variabilelor latente endogene W_1 , respectiv W_2 (democrația politică în anul 1985, respectiv cea a anului 1995).

Variabilele $Z_i, 1 \leq i \leq 3, E_j, 1 \leq j \leq 8$, reprezintă "erorile de măsurare" pentru variabilele observabile X_i , respectiv Y_j .

Este normal să presupunem că aceste perturbații de măsurare Z_i, E_j ce afectează relația de dependență dintre variabilele latente și variabilele observabile au media zero și sunt necorelate cu toate variabilele latente (U_1, W_1, W_2, V_1, V_2) ale modelului cauzal.

Perturbațiile Z_i, E_j pot fi privite ca o influență a altor variabile asupra variabilelor observabile. Un studiu detaliat asupra acestui concept va fi făcut ulterior, urmărindu-se descompunerea acestor factori neprecizați ("erori") în mai multe componente (componente de grup și componente specifice).

Ecuțiile (6) și (7) pot fi rescrise în forma matricială

$$\begin{aligned} X &= D_X \cdot U + Z \\ Y &= D_Y \cdot W + E \end{aligned} \quad (8)$$

unde vectorii (coloană) X, Y, U, W, Z, E au componentele :

$$\begin{aligned} U &= (U_1) ; & W &= (W_1, W_2) ; & X &= (X_1, X_2, X_3)' ; \\ Y &= (Y_1, Y_2, Y_3, Y_4, Y_5, Y_6, Y_7, Y_8) ; & Z &= (Z_1, Z_2, Z_3)' ; \\ E &= (E_1, E_2, E_3, E_4, E_5, E_6, E_7, E_8) \end{aligned} \quad (9)$$

Transpusele D_X' , D_Y' ale matricilor D_X , D_Y au elementele:

$$D_X' = (d_1, d_2, d_3) ; \quad (10)$$

$$D_Y' = \begin{vmatrix} d_4 & d_5 & d_6 & d_7 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & d_8 & d_9 & d_{10} & d_{11} \end{vmatrix}$$

Dacă p, q reprezintă numărul maxim de indicatori acceptați pentru fiecare variabilă latentă endogenă, respectiv variabilă latentă exogenă a modelului inițial atunci matricile D_X , D_Y au dimensiunea $q \times n$, respectiv $p \times m$.

Pentru Cazul 1, cu ecuațiile de structură (6) și (7), avem $p = 4$ și $q = 3$.

2.4. Indicatori statistici utilizați

Cazul 2. Să presupunem o variabilă latentă U măsurată indirect prin intermediul a două variabile observabile X_1, X_2 . Variabilele observabile $X_i, i=1,2$, vor putea fi caracterizate (incomplet) prin k realizări ale lor, anume $x_{i1}, x_{i2}, \dots, x_{ik}$.

Modelul empiric va trebui să precizeze gradul de influență al variabilei latente U asupra fiecăreia dintre variabilele măsurabile X_i .

Variabila U determină cauzal variabilele X_1 și X_2 cu perturbațiile Z_1 , respectiv Z_2 . Așadar modelul empiric este definit de următoarele ecuații de structură:

$$X_1 = d_1 \cdot U + Z_1 ; \quad X_2 = d_2 \cdot U + Z_2 \quad (11)$$

În continuare vom căuta să sugerăm o posibilă analiză a modelului cauzal (11).

Coeșicientul $d_j, j=1,2$, dă intensitatea relației de dependență liniară dintre variabila latentă U și variabila observabilă X_j . Creșterea cu o unitate a valorii variabilei latente U antrenează o variație cu d_j unități pentru variabila măsurabilă X_j .

Dacă, în plus, presupunem independența perturbațiilor Z_j în raport cu variabila latentă U , adică $\text{Cov}(Z_j, U) = 0$, atunci gradul de dependență liniară $\text{Cov}(X_j, U)$ dintre X_j și U este proporțional cu d_j :

$$\begin{aligned} \text{Cov}(X_j, U) &= \text{Cov}(d_j \cdot U + Z_j, U) = d_j \cdot \text{Cov}(U, U) + \text{Cov}(Z_j, U) = \\ &= d_j \cdot \text{Cov}(U, U) = d_j \cdot \text{Var}(U) \end{aligned} \quad (12)$$

Să analizăm consecințele relației (12).

Astfel, în cazul în care avem o dependență liniară pozitivă între variabila latentă U și indicatorii ei X_1, X_2 (adică $d_1 > 0, d_2 > 0$) iar perturbațiile Z_1, Z_2 sunt necorelate, atunci rezultă o corelație pozitivă între variabilele măsurabile X_1, X_2 , deoarece :

$$\begin{aligned} \text{Cov}(X_1, X_2) &= \text{Cov}(d_1 \cdot U + Z_1, d_2 \cdot U + Z_2) = d_1 \cdot d_2 \cdot \text{Cov}(U, U) + \\ &+ d_1 \cdot \text{Cov}(U, Z_2) + d_2 \cdot \text{Cov}(Z_1, U) + \text{Cov}(Z_1, Z_2) = d_1 \cdot d_2 \cdot \text{Var}(U) \neq 0 \end{aligned} \quad (13)$$

În situația în care covarianța $\text{Cov}(Z_1, Z_2)$ este nenulă iar coeficienții de dependență liniară d_1, d_2 sunt pozitivi, nu putem afirma nimic în privința semnului indicatorului statistic $\text{Cov}(X_1, X_2)$. Afirmția se justifică interpretând identitatea :

$$\text{Cov}(X_1, X_2) = d_1 \cdot d_2 \cdot \text{Var}(U) + \text{Cov}(Z_1, Z_2) \quad (14)$$

În Cazul 2 am utilizat mai mulți operatori statistici : media (**Med**), covarianța (**Cov**), varianța (**Var**), corelația (**Cor**). Este util de enumerat mai multe proprietăți ale acestor operatori, proprietăți ce vor fi adeseori utilizate în continuare:

$$\begin{aligned} \text{Cov}(X_1, X_2) &= \text{Med}((X_1 - \text{Med}(X_1)) \cdot (X_2 - \text{Med}(X_2))) = \\ &= \text{Med}(X_1 \cdot X_2) - [\text{Med}(X_1)] \cdot [\text{Med}(X_2)] \end{aligned} \quad (15)$$

$$\text{Var}(X_1) = \text{Cov}(X_1, X_1) = \text{Med}(X_1^2) - [\text{Med}(X_1)]^2$$

$$\text{Var}(X_1) \neq 0 ; \quad \text{Var}(X_1) = 0 \text{ dacă și numai dacă } X_1 = \text{constant}$$

$$\text{Var}(c \cdot X_1) = c^2 \cdot \text{Var}(X_1)$$

$$\text{Var}(X_1 + X_2) = \text{Var}(X_1) + \text{Var}(X_2) + 2 \cdot \text{Cov}(X_1, X_2)$$

Menționăm că operatorul de covarianță este simetric, aditiv în raport cu fiecare dintre parametrii săi, multiplicativ cu constantele c , valoarea sa fiind nulă dacă cel puțin unul dintre parametrii săi este constant. Așadar:

$$\text{Cov}(X_1, X_2) = \text{Cov}(X_2, X_1) ; \quad \text{Cov}(c, X_1) = 0$$

$$\text{Cov}(c \cdot X_1, X_2) = c \cdot \text{Cov}(X_1, X_2)$$

$$\text{Cov}(X_1 + X_3, X_2) = \text{Cov}(X_1, X_2) + \text{Cov}(X_3, X_2) \quad (16)$$

Operatorul de covarianță statistică este adaptat măsurării gradului de dependență liniară dintre variabile.

Estimarea valorile rezultate prin aplicarea operatorilor statistici enumerați se va efectua utilizându-se realizări ale variabilelor aleatoare asupra cărora se acționează. Fiecărui operator “teoretic” ce acționează asupra unor variabile aleatoare i se asociază un operator “empiric” ce va acționa asupra mulțimilor de realizări ale variabilelor aleatoare respective.

Am pus astfel în evidență un paralelism între conceptele “teoretic” și “experimental”. Astfel o variabilă aleatoare se va putea manifesta printr-o multitudine de realizări ale sale.

Media experimentală x_{i+} și covarianța experimentală s_{12} vor estima media teoretică $Med(X_i)$ a variabilei X_i , respectiv covarianța teoretică $Cov(X_1, X_2)$ dintre variabilele X_1 și X_2 , atunci când se utilizează k realizări ale acestor variabile:

$$x_{i+} = (x_{i1} + x_{i2} + x_{i3} + \dots + x_{ik})/k \quad (17)$$

$$s_{12} = [(x_{11}-x_{1+}) \cdot (x_{21}-x_{2+}) + \dots + (x_{1k}-x_{1+}) \cdot (x_{2k}-x_{2+})]/[k - 1]$$

Covarianța va fi operatorul statistic utilizat cu precădere în vederea estimării relațiilor de dependență dintre variabilele modelelor cauzale propuse. Uneori acest indicator va fi înlocuit de “covarianță normalizată” (corelația).

Remarca 2. Dacă c este o constantă atunci

$$Cov(X + c, U) = Cov(X, U) + Cov(c, U) = Cov(X, U) \quad (18)$$

relație ce subliniază faptul că pentru determinarea covarianței dintre variabilele X și U nu are importanță stabilirea unei origini pentru scala de măsurare. Acest lucru este util mai ales în acele măsurători sociologice pentru care se pot aprecia diferențele dintre orice două observații dar este greu de stabilit o origine de referință comună a măsurătorilor.

Așadar operatorul de covarianță nu este afectat de o translație, cu o cantitate fixată (nu întotdeauna cunoscută), a valorilor măsurate.

2.5. Convenții privind grafurile cauzale

Modelul cauzal mai poate fi reprezentat și ca un graf orientat în care nodurile sale sunt variabilele modelului cauzal iar arcele (orientate) dintre noduri semnifică relația cauzală de dependență dintre variabilele asociate acelor noduri.

Un arc orientat (cu săgeată) de la nodul X_1 la nodul X_2 simbolizează o relație de “dominare” între aceste variabile, variabila X_1 constituind una dintre cauzele ce acționează asupra variabilei X_2 . Variabila X_2 constituie efectul cauzei X_1 .

Prin reprezentare grafică obținem o imagine globală asupra raporturilor cauzale dintre variabilele modelului studiat. În graful atașat modelului cauzal, de cele mai multe ori nodurile sunt reprezentate diferit în funcție de tipul de variabile asociate lor :

- nodurile atașate variabilelor latente se încercuiesc,
- nodurile corespunzătoare variabilelor observabile se încadrează într-un dreptunghi,
- nodurile afiliate variabilelor perturbatoare (erorilor) rămân nemodificate.

Definiții și convenții de scriere.

În textul expunerii și în grafulurile cauzale vom adopta următoarea convenție de scriere:

- variabilele latente vor fi reprezentate de majuscule subliniate (ele fiind variabile de importanță majoră pentru modelul cauzal),
- variabilele observabile, având o funcție de variabile intermediare, ajutoare, vor fi scrise obișnuit,
- variabilele perturbatoare (erorile), vor fi notate cu caractere italice.

De regulă, pentru variabile se vor folosi litere mari, literele mici fiind utilizate în special pentru a marca observațiile variabilelor respective. Cuvintele scrise cu caractere îngroșate (bold) vor reprezenta tablouri de variabile sau de valori având un nume comun și care se pretează la o tratare unitară (asemănătoare vectorilor și matricilor de date).

Urmând convențiile enumerate, modelul cauzal cu variabile latente definit prin ecuațiile structurale (4) va avea reprezentarea grafică din figura 2.1.

2.6. Măsurarea dependențelor cauzale

Intensitatea legăturilor cauzale ale modelului din figura 2.1 este dată de valorile parametrilor a_{11}, a_{21}, b_{21} din ecuația de structură (4) sau specificată prin ponderea arcelor orientate din diagrama cauzală respectivă.

Ecuatiile de structură (6) cuprind atât variabile latente cât și variabile observabile. Diagrama cauzală corespunzătoare este prezentată în figura 2.2.

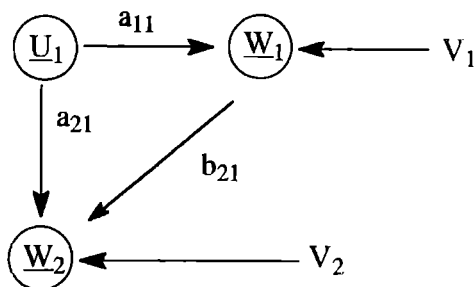


Figura 2.1.

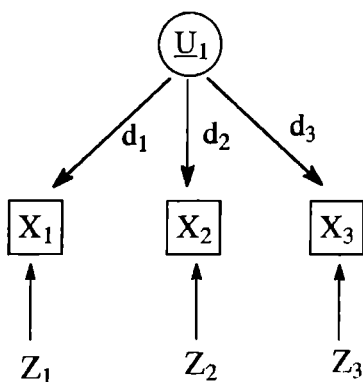


Figura 2.2.

Dacă, în diagrama cauzală din figura 2.2, între perturbațiile Z_i și Z_j nu există arce dublu orientate atunci vom subînțelege o covarianță nulă între aceste variabile, $Cov(Z_i, Z_j) = 0$.

Așadar existența unui arc dublu orientat între două variabile perturbatoare presupune o covarianță nenulă între acestea.

În diagrama cauzală efectul total dintre două variabile este determinat prin sumarea efectului direct cu efectul indirect dintre variabilele respective, adică :

$$\text{Efect_total} = \text{Efect_direct} + \text{Efect_indirect} \quad (19)$$

Astfel efectul direct măsurat între variabilele \underline{U}_1 și \underline{W}_2 din diagrama 2.1 este dat de valoarea coeficientului a_{21} , pe când efectul indirect între aceleași variabile se realizează prin intermediul variabilei latente \underline{W}_1 .

Substituind prima relație (4) în cea de a doua obținem :

$$\underline{W}_2 = a_{11} \cdot b_{21} \cdot \underline{U}_1 + a_{21} \cdot \underline{U}_1 + (b_{21} \cdot V_1 + V_2) \quad (20)$$

Intensitatea efectului indirect a variabilei latente \underline{U}_1 asupra variabilei latente \underline{W}_2 este măsurată prin valoarea produsului $a_{11} \cdot b_{21}$.

Prin urmare valoarea efectului direct dintre oricare două noduri ale unei diagrame cauzale este dată de ponderea arcului ce leagă (direct) aceste noduri (adică lungimea drumului direct dintre noduri).

Într-un graf, un drum (orientat) de la nodul X la nodul Y este definit de șirul de arce orientate prin intermediul cărora se poate ajunge (direct sau indirect) de la X la Y. "Lungimea" (ponderea) unui astfel de drum este calculată ca fiind produsul ponderilor arcelor componente drumului.

Valoarea efectului indirect de la nodul X la nodul Y va fi dată de suma "lungimilor" tuturor drumurilor indirecte ce duc de la X la Y.

Astfel în diagrama cauzală din figura 2.1 există un singur drum indirect de la \underline{U}_1 la \underline{W}_2 , anume $\underline{U}_1 - \underline{W}_1 - \underline{W}_2$, aceasta prin utilizarea nodului intermediar \underline{W}_1 . Ponderea acestui unic drum indirect dintre \underline{U}_1 și \underline{W}_2 este $a_{11} \cdot b_{21}$, cantitate ce măsoară de fapt dependența "indirectă" dintre aceste variabile latente.

Prin gruparea termenilor asemenea, ecuația (20) se poate rescrie în forma :

$$\underline{W}_2 = (a_{11} \cdot b_{21} + a_{21}) \cdot \underline{U}_1 + (b_{21} \cdot V_1 + V_2) \quad (21)$$

Coeficientul a_{21} măsoară efectul direct al variabilei \underline{U}_1 în explicitarea lui \underline{W}_2 . În schimb produsul $a_{11} \cdot b_{21}$ semnifică efectul indirect

pe care îl exercită variabila \underline{U}_1 asupra aceleiași variabile \underline{W}_2 prin intermediul variabilei latente \underline{W}_1 .

Coefficientul variabilei \underline{U}_1 în ecuația (21) are valoarea $a_{21} + a_{11} \cdot b_{21}$ și reprezintă de fapt efectul total ce acționează asupra variabilei \underline{W}_2 și care este datorat cauzei \underline{U}_1 .

Concluzii

În acest paragraf am sugerat o dublă reprezentare a unui model cauzal, atât prin ecuațiile sale de structură cât și prin graful orientat atașat acestor ecuații. Nodurile grafului cauzal sunt variabilele ale modelului.

Raportul cauză-efect este dictat de sensul arcului orientat dintre nodurile respective. Variabila origine a unui arc constituie "cauza", iar variabila țintă a aceluiași arc semnifică "efectul". Intensitatea legăturii dintre două noduri ale grafului cauzal este dată de "ponderea" ce îi este atribuită arcului ce leagă direct acele noduri.

În procesul de explicitare a unui model cauzal este necesară utilizarea unor indicatori statistici (medie, varianță, covarianță, corelație).

Au fost puse în evidență atât modele cu variabile latente cât și modele mixte (cu variabile latente și observabile). Menționăm grija deosebită ce trebuie acordată definirii corecte a indicatorilor (alegerii variabilelor măsurabile) în încercarea de apreciere justă a comportamentului variabilelor latente.

O importanță aparte trebuie acordată măsurării influențelor cauzale indirecte cât și a efectului total dintre oricare două variabile ale modelului empiric.

II.3. VALIDAREA RELAȚIEI DE CAUZALITATE

3.1. Introducere

În procesul de evidențiere a relației de cauzalitate directă dintre două variabile apare necesară parcurgerea mai multor etape.

În primul rând, se va încerca confirmarea sau infirmarea unei posibile legături (directe sau indirecte) dintre două variabile oarecare ale modelului conceput. Tehnica adoptată în această situație constă în izolarea perechii de variabile studiate în raport cu toate celelalte variabile ale modelului.

Dacă se dovedește că, într-adevăr, între cele două variabile astfel separate există vreo legătură, atunci se va analiza în continuare în ce măsură această legătură ar putea fi o legătură cauzală directă. Mai precis, se va aprecia dacă cele două variabile selectate se pot asocia din punct de vedere cauzal sau acest fapt conduce la o contradicție.

Urmatoarea etapă după faza de asociere constă în stabilirea sensului cauzalității în cadrul perechii (X, Y) de variabile selectate. Se va preciza dacă avem un raport cauzal de la X la Y , sau a un raport cauzal în sens invers, de la Y la X .

Există oricând posibilitatea unei relații cauzale duble, reciproce, între variabilele X și Y .

3.2. Tehnici de izolare a variabilelor

Intenționăm să stabilim dacă în cadrul unei perechi fixate de variabile X, Y poate fi acceptată sau infirmată o relație cauzală între componente. În acest sens se va folosi o procedură de izolare a perechii (X, Y) de restul variabilelor modelului.

Variabilele X și Y sunt într-o relație de dependență, dacă, de exemplu, modificări ale valorilor lui X ar antrena modificări ale valorilor variabilei Y , aceasta atunci când toate celelalte variabile rămân nemodificate.

Un astfel de procedeu este ideal. Dar de cele mai multe ori nu poate fi aplicat în practică, cel puțin în forma prezentă. Există situații în care nu se pot păstra vechile valori ale variabilelor modelului atunci când, de exemplu, se operează o modificare asupra variabilei X .

În acest caz nu se poate afirma (cu certitudine) că eventuala variație a valorilor variabilei Y se datorează exclusiv fluctuațiilor variabilei X . Prin urmare suntem puși în imposibilitatea de a confirma o posibilă relație de asociere între componentele perechii (X, Y) .

În acest raționament de “izolare” a perechii de variabile (X, Y) s-au acceptat implicit trei ipoteze :

– Dacă variabilele X și Y sunt într-o relație cauzală directă (de la X la Y) atunci Y nu poate fi modificat fără “acordul” lui X . De fapt, relația cauză-efect a fost interpretată în mod determinist, în forma $Y = f(X)$.

– Efectul este perfect predictibil atunci când se cunoaște cauza.

– Unei cauze unice i se asociază un efect unic.

Oricând se poate renunța, cel puțin parțial, la astfel de presupuneri sau se pot propune alte noi ipoteze.

De exemplu, admiterea unei relații cauzale dintre Y și X ce nu este “deterministă” impune introducerea unei variabile perturbatoare E , relația cauzală fiind definită de :

$$Y = f(X) + E \quad (1)$$

Funcția f joacă rolul componentei deterministe de influență a cauzei X asupra variabilei-efect Y . Această componentă deterministă este de cele mai multe ori aproximată cu o funcție liniară, fapt infirmat nu de puține ori în practică.

În realitate nu se poate realiza o izolare perfectă a grupului de variabile (X, Y) în raport cu restul variabilelor modelului.

În modelul (1) eroarea E poate fi interpretată (din punct de vedere determinist) și ca o manifestare a altor variabile $Z_1, Z_2, Z_3, \dots, Z_n$ ce nu au fost cuprinse în model, adică :

$$E = h(Z_1, Z_2, Z_3, \dots, Z_n) \quad (2)$$

Ținând cont de condiția de “izolare”, și de faptul că modificările variabilei Y sunt datorate exclusiv variațiilor variabilei X , atunci din modelul (1) cu precizarea (2) rezultă inexistența unei dependențe între variabilele X și Z_j , $1 \leq j \leq n$, și deci între X și $h(Z_1, Z_2, \dots, Z_n)$.

Așadar prin impunerea condiției de “izolare” nu va fi posibilă existența unei dependențe între variabila X și variabila “perturbatoare” E .

În aplicațiile sociologice se utilizează deseori coeficientul de corelație Pearson drept măsură a dependenței (liniare) dintre două variabile. În această situație o valoare nenulă a corelației (Pearson) $Cor(X,E)$ violează condiția de izolare a perechii (X,Y) din modelul (1).

Trebuie subliniat însă faptul că fenomenul de cauzalitate este mult mai profund, corelația neînsemnând neapărat cauzalitate. Renunțarea în cadrul modelului propus la unele variabile “intermediare” distorsionează de multe ori rezultatele. Acest aspect ar putea fi pus în evidență prin neglijarea condiției de pseudo-izolare, caz ce va fi discutat în continuare.

Exemplul 1. Să presupunem că modelul real (figura 3.1) este definit de următoarele ecuații de structură :

$$\begin{aligned} Y_1 &= a_{11} \cdot X_1 + Z_1 \\ Y_2 &= b_{21} \cdot Y_1 + a_{21} \cdot X_1 + Z_2 \end{aligned} \quad (3)$$

Variabila Y_2 este influențată atât direct cât și în mod indirect (prin intermediul lui Y_1) de către variabila X_1 .

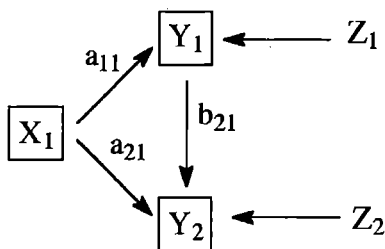


Figura 3.1.

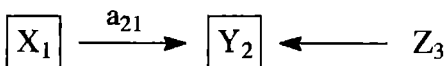


Figura 3.2.

Condiția de izolare presupune:

$$\text{Cov}(Z_1, Z_2) = 0, \text{Cov}(X_1, Z_1) = 0, \text{Cov}(X_1, Z_2) = 0 \quad (4)$$

Neglijarea în modelul (3) a variabilei intermediare Y_1 conduce la următorul model simplificat (figura 3.2) :

$$Y_2 = a_{21} \cdot X_1 + Z_3 \quad (5)$$

unde de fapt

$$Z_3 = b_{21} \cdot Y_1 + Z_2 \quad (6)$$

Absența din modelul (3) a variabilei intermediare Y_1 ar putea fi sesizată practic prin nerespectarea condiției de pseudo-izolare.

Într-adevăr coeficientul de corelație Pearson $\text{Cor}(X_1, Z_3)$ dintre variabilele X_1 și Z_3 este nenul, deoarece :

$$\begin{aligned} \text{Cov}(X_1, Z_3) &= \text{Cov}(X_1, b_{21} \cdot Y_1 + Z_2) = b_{21} \cdot \text{Cov}(X_1, Y_1) + \text{Cov}(X_1, Z_2) = \\ &= b_{21} \cdot \text{Cov}(X_1, a_{11} \cdot X_1 + Z_1) = b_{21} \cdot a_{11} \cdot \text{Var}(X_1) + b_{21} \cdot \text{Cov}(X_1, Z_1) = \\ &= a_{11} \cdot b_{21} \cdot \text{Var}(X_1) \end{aligned} \quad (6)$$

Menționăm că valorile a_{11} și b_{21} sunt, de regulă, nenule iar dispersia variabilei aleatoare X_1 este strict pozitivă dacă X_1 nu este o constantă.

O sursă permanentă de confuzie este generată de diferitele interpretări pe care le acordăm în mod curent coeficienților de dependență, ca o influența directă, indirectă, sau printr-un efect cumulativ.

De multe ori corelația dintre două variabile este dictată de prezența unei a treia variabile ce le influențează în parte pe primele două.

Renunțarea la unele variabile sau indicarea unor false dependențe poate denatura grav rezultatele. Modelele empirice ce urmăresc o “simplificare” a realității prin omiterea unor relații cauzale esențiale sau a unor variabile “importante” constituie surse permanente de erori.

Aceste afirmații vor fi evidențiate prin mai multe exemple.

Exemplul 2. În cazul în care dorim să determinăm influența “gradului de inteligență” X a subiectului asupra calității Y_1 a produsului pe care acesta îl realizează, va trebui neapărat să studiem pe lângă legătura cauzală directă de la X la Y_1 și o legătură cauzală indirectă ce ia în considerare tipul de activitate Y_2 pe care individul o desfășoară.

Astfel în modelul experimental propus vor fi neapărat prezente și relațiile cauzale $X \longrightarrow Y_2, Y_2 \longrightarrow Y_1$.

Specificarea unui asemenea tip de relație este necesară în special în “activitățile plictisitoare”, atunci când indivizi inteligenți “obosiți” de o muncă intelectuală neinteresantă pentru ei realizează produse de o calitate inferioară.

Neluarea în considerare a dependenței cauzale indirecte

$X \longrightarrow Y_2 \longrightarrow Y_1$ se bazează pe falsă percepție că întotdeauna variabila latentă “inteligenta” influențează numai în sens pozitiv procesul de obținere a produselor de calitate. Dar “gradul de inteligență” poate acționa indirect cu efect negativ asupra calității produsului rezultat. Acest fenomen este adesea prezent în cazul unor activități intelectuale “neinteresante”, de rutină.

Nu de puține ori, în modelele cauzale propuse apar relații cauzale false generate printr-o concepere inițială neadecvată a indicatorilor variabilelor latente.

Exemplul 3. În modelul definit în figura 3.3 variabilele latente $\underline{U}, \underline{W}$ sunt studiate indirect prin intermediul indicatorilor $I_1=(X_1, X_2)$, $I_2=(X_2, X_3)$. Prezența variabilei observabile X_2 în componența celor doi indicatori I_1 și I_2 va antrena, de regulă, un raport de dependență, nu neapărat real, între variabilele latente corespunzătoare $\underline{U}, \underline{W}$.

Exemplul 4. Dacă în modelul descris în figura 3.4 sunt identificați greșit indicatorii variabilelor latente $\underline{U}, \underline{W}$ (a se vedea figura 3.5), atunci evident intensitățile raporturilor cauzale dintre aceste variabile latente vor fi distorsionate.

Este important de semnalat faptul că o lipsă de corelare dintre variabile ale modelului nu implică neapărat absența cauzalității dintre acele variabile. Astfel, un coeficient de corelație nenul este o condiție necesară a cauzalității în cazul izolării perechii de variabile respective.

Atragem însă atenția că procedura de izolare a perechilor de variabile nu este întotdeauna aplicabilă practic.

În plus, există posibilitatea ca între variabilele X și Y să existe o dependență funcțională, sau o dependență stohastică ce nu este neapărat “liniară”. Vom comenta acest lucru în exemplul următor.

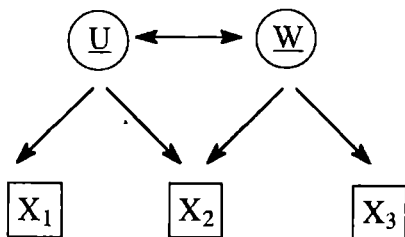


Figura 3.3.

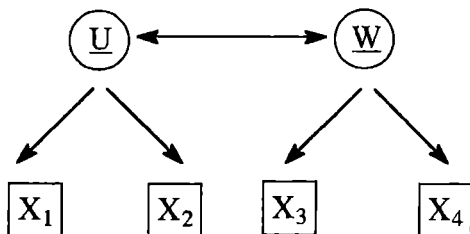


Figura 3.4.

Exemplul 5. Să presupunem o dependență funcțională de tipul $Y = a.X^2$. Dacă variabila observabilă, independentă, X ia în egală măsură atât valori pozitive cât și valori negative atunci $\text{Cor}(X, Y) = 0$, fapt ce ar sugera absența unei cauzalități.

Această concluzie este însă infirmată de model deoarece există o dependență funcțională neliniară clar definită.

În cazul în care între variabilele modelului sunt legături de dependență neliniare (ce nu pot fi "bine aproximate" prin acceptarea unor dependențe liniare), se aplică transformări adecvate (logaritmare, exponentiere, etc.). În final, s-ar putea ca între noile variabile astfel rezultate să existe o puternică dependență de tip liniar.

Astfel modelul neliniar descris anterior prin ecuația $Y = a.X^2$ este redus la modelul liniar $Y = a.X_1$ prin utilizarea transformării $X_1 = X^2$.

O selectare neadecvată a variabilelor modelului la care se poate adăuga și o interpretare greșită a tipurilor variabilelor implicate (endogene sau exogene) conduce de nenumărate ori la erori majore mai ales în faza de soluționare efectivă a modelului empiric.

Exemplul 6. Fie modelul real ale cărui relații cauzale sunt descrise în figura 3.6, în care variabila X_3 este interpretată greșit drept o variabilă exogenă, toate celelalte variabile fiind considerate endogene.

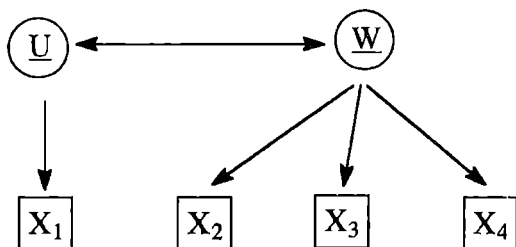


Figura 3.5.

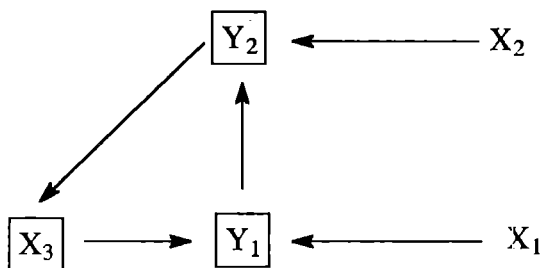


Figura 3.6.

Menționăm că în domeniul socio-uman se admite principiul: “variabilele endogene nu pot influența variabilele exogene”.

Acest lucru antrenează obligatoriu neglijarea influenței variabilei endogene Y_2 asupra presupusei variabile exogene X_3 , fapt ce va denatura în mod evident percepția privind raporturile reale din modelului teoretic.

Un asemenea aspect este des semnalat în cazul sistemelor evolutive în timp atunci când, de exemplu, variabila dependentă Y_2 nu influențează comportarea prezentă a variabilei X_3 ci determină modificări ulterioare ale acesteia (o nouă “versiune temporală” a lui X_3).

În figura 3.6, o interpretare adecvată a tipului variabilei X_3 , anume ca variabilă endogenă a modelului, va elimina eroarea făcută. În noua situație astfel creată, variabila X_3 devine modificabilă prin acțiunea oricărei alte variabile endogene a modelului empiric.

Ipoteza de pseudo-izolare ar putea să nu fie verificată și datorită corelărilor dintre variabilele perturbatoare. Vom sugera un exemplu în acest sens.

Exemplul 7. În modelul cu variabile observabile din figura 3.7 presupunem o corelație nenulă între variabilele perturbatoare Z_1 și Z_2 . Modelul este definit prin ecuațiile de structură :

$$\begin{aligned} Y_2 &= b_{21} \cdot Y_1 + Z_2 \\ Y_1 &= a_{11} \cdot X_1 + Z_1 \end{aligned} \quad (7)$$

unde

$$\begin{aligned} \text{Med}(Z_1) &= \text{Med}(Z_2) = 0 \\ \text{Cov}(X_1, Z_1) &= \text{Cov}(X_1, Z_2) = 0 \end{aligned} \quad (8)$$

În cazul perechii de variabile (X_1, Y_1) efectul de pseudo-izolare se va traduce prin condiția $\text{Cov}(X_1, Z_1) = 0$. Nu același lucru se poate afirma despre perechea (Y_1, Y_2) . Într-adevăr

$$\begin{aligned} \text{Cov}(Y_1, Z_2) &= \text{Cov}(a_{11} \cdot X_1 + Z_1, Z_2) = \\ &= a_{11} \cdot \text{Cov}(X_1, Z_2) + \text{Cov}(Z_1, Z_2) \neq 0 \end{aligned} \quad (9)$$

Menționăm faptul că fenomenul de corelare a factorilor perturbatori apare mai ales în cadrul modelele dinamice.

O autocorelare în timp a erorilor va antrena nerespectarea condiției de pseudo-izolare, condiție operantă pentru evidențierea unei posibile cauzalități.

3.3. Relația de asociere

În urma efectuării unei analize adecvate a modelului empiric în care s-a utilizat și tehnica pseudo-izolării, se vor propune în final perechi de variabile (X, Y) ce se “asociază”.

Problema fundamentală ce apare într-o astfel de situație este dacă între variabilele perechii selecționate există o dependență cauzală directă (și nu una indirectă) și în plus, dacă este posibilă precizarea sensului acestei dependențe directe. Este util de a studia eventualitatea unor reacții reciproce dintre variabilele perechii analizate.

Menționăm că la definirea relației de asociere dintre variabile trebuie ținut seama și de următoarele două aspecte :

- Toți coeficienții (parametri) modelului să poată fi identificați.
- Sensul raporturilor cauzale dintre variabila latentă și variabilele observabile atașate acesteia este uneori postulat.

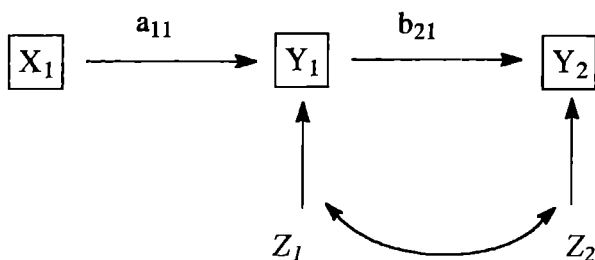


Figura 3.7.

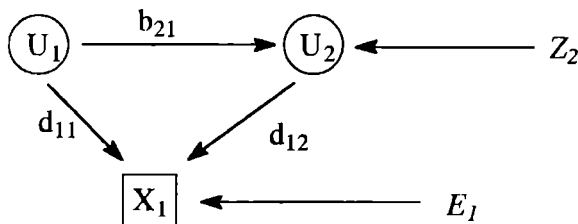


Figura 3.8.

Să comentăm aceste aspecte :

I. Prima cerință nu este satisfăcută în mod automat.

Astfel modelul 3.8 cu variabila observabilă X_1 și variabilele latente U_1, U_2 descrise de ecuațiile de structură

$$\begin{aligned} \underline{U}_2 &= b_{21} \cdot \underline{U}_1 + Z_2 \\ X_1 &= d_{11} \cdot \underline{U}_1 + d_{12} \cdot \underline{U}_2 + E_1 \end{aligned} \quad (10)$$

nu poate fi efectiv soluționat.

Dupa efectuarea calculelor rezultă :

$$X_1 = (d_{11} + d_{12} \cdot b_{21}) \cdot U_1 + (d_{12} \cdot Z_2 + E_1) = d \cdot U + E_2 \quad (11)$$

unde

$$\begin{aligned} d &= d_{11} + d_{12} \cdot b_{21} \\ E_2 &= d_{12} \cdot Z_2 + E_1 \end{aligned} \quad (12)$$

Utilizând relațiile (11), printr-o procedură de regresie simplă, vom putea eventual estima valoarea coeficientului d care verifică (12) dar nu vom putea să precizăm valorile parametrilor d_{11}, d_{12}, b_{21} .

Nu același lucru se poate spune în cazul în care renunțăm la existența unei dependențe cauzale între variabilele latente U_1 și U_2 (cazul $b_{21} = 0$).

II. În modelul cauzal empiric, raporturile cauzale admise între variabilele latente și variabilele observabile asociate lor sunt strâns legate de domeniul concret abordat. Astfel, în cercetările sociologice sau psihologice se acceptă faptul că o variabilă latentă U constituie o cauză pentru variabila observabilă X . În acest caz, variabila observabilă este privită ca un efect al variabilei latente.

Asupra variabilei observabile mai poate acționa și un factor perturbator E , factor interpretat drept eroare de măsurare.

Vom avea așadar ecuația de structură de forma

$$X = d_{11} \cdot U + E \quad (13)$$

cu condițiile suplimentare $Cov(U, E) = 0$ (variabilele latente și erorile de măsurare sunt necorelate), $Med(E) = 0$ (o compensare a erorilor).

Vom semnala în continuare și alte aspecte interesante a căror soluționare, însă, depășește cadrul lucrării.

O problemă majoră ce ține de filosofia dependențelor cauzale este aceea de a preciza în ce măsură este posibil să existe o relație cauzală reciprocă între două variabile.

Un loc aparte îl ocupă studiul comportării perturbațiilor (factorilor aleatori). Asupra acestora se vor efectua teste statistice pentru a decide dacă poate fi pusă în evidență o autocorelare în timp sau dacă are loc o modificare în timp a varianțelor factorilor.

Un fenomen interesant este acela al extinderii relației de dependență (liniară) la o “multicoliniaritate”, variabilele observabile fiind la rândul lor explicate (printr-o relație liniară) de alte noi variabile observabile. În situația nou creată, variabilele observabile influențate de factori comuni vor fi corelate.

Este de preferat ca în modele cauzale să fie alese variabile observabile necorelate. Corelarea dintre variabilele observabile ce caracterizează un indicator prezintă inconvenientul major al creșterii, de cele mai multe ori, a varianțelor (fluctuațiilor) ponderilor acestor variabile în cadrul indicatorului respectiv.

Acest fenomen se va manifesta concret printr-o diminuare a preciziei în determinarea ponderilor variabilelor corelate din cadrul indicatorului respectiv.

Vom exemplifica acest aspect în continuare.

Exemplu 8. Fie variabilele observabile Y_1, X_1, X_2 . Dorim să stabilim pentru variabila Y_1 un indicator de tipul

$$Y_2 = a_1 \cdot X_1 + a_2 \cdot X_2 \quad (14)$$

Rămâne să precizăm ponderile a_1, a_2 ale variabilelor X_1 , respectiv X_2 în indicatorul Y_2 .

Eroarea pe care o facem aproximând variabila Y_1 prin intermediul indicatorului asociat Y_2 , este egală cu E , unde $E = Y_1 - Y_2$. Eroarea E nu este corelată cu variabilele X_1 și X_2 , adică

$$\text{Cov}(E, X_1) = 0 \quad ; \quad \text{Cov}(E, X_2) = 0 \quad (15)$$

Cum $E = Y_1 - a_1 \cdot X_1 - a_2 \cdot X_2$, din condițiile (15) deducem următorul sistem liniar de două ecuații în necunoscutele a_1, a_2

$$\begin{aligned} a_1 \cdot \text{Var}(X_1) + a_2 \cdot \text{Cov}(X_1, X_2) &= \text{Cov}(Y_1, X_1) \\ a_1 \cdot \text{Cov}(X_1, X_2) + a_2 \cdot \text{Var}(X_2) &= \text{Cov}(Y_1, X_2) \end{aligned} \quad (16)$$

ale cărui soluții sunt de fapt :

$$a_1 = [\text{Cov}(Y_1, X_1) \cdot \text{Var}(X_2) - \text{Cov}(Y_1, X_2) \cdot \text{Cov}(X_1, X_2)] / d$$

$$a_2 = [\text{Cov}(Y_1, X_2) \cdot \text{Var}(X_1) - \text{Cov}(Y_1, X_1) \cdot \text{Cov}(X_1, X_2)] / d \quad (17)$$

unde d este chiar determinantul sistemului (16), adică

$$d = \text{Var}(X_1) \cdot \text{Var}(X_2) - [\text{Cov}(X_1, X_2)]^2 \quad (18)$$

În cazul în care variabilele X_1 și X_2 sunt puternic corelate (patratul corelației este aproape de 1), atunci determinantul d al sistemului (16) este aproximativ nul. Așadar sistemul (16) va fi prost condiționat. Soluțiile sale a_1, a_2 date de relațiile (17) vor prezenta fluctuații mari în cazul unor erori de calcul acceptabile. Acest fenomen este datorat împărțirii, în expresiile (17), cu o cantitate d apropiată de zero.

Remarci similare pot fi enunțate și în cazul unor modele cu variabile latente și observabile dar în care indicatorii asociați diverselor variabile latente sunt corelați între ei (ca și în modelul din Exemplul 3, Figura 3.3).

În plus, o influență cauzală reciprocă între variabilele U și W din Exemplul 3 va face dificilă includerea variabilei X_2 numai în componența unuia dintre indicatorii variabilelor latente U sau W .

Indicatorii variabilelor U, W “aproximează” comportamentul variabilelor latente respective. O corelare între U și W va antrena o foarte posibilă corelare între indicatorii acestor variabile. Așadar cei doi indicatori vor trebui să aibă o “componentă comună”.

3.4. Stabilirea sensului cauzalității

Odată presupusă o relație de asociere dintre două variabile ne mai rămâne să precizăm sensul acesteia (de la cauză la efect) și eventual acceptarea unui dublu sens cauzal. Ultima situație este echivalentă cu acceptarea unor relații de influențare reciprocă între variabile în care nu este obligatoriu să avem aceleași ponderi de influență în ambele sensuri.

O problemă foarte importantă la determinarea ponderilor de influență cauzală dintre variabilele modelului constă în stabilirea corectă a duratei de întârziere de la aplicarea “cauzei” până la manifestarea “efectului” ei.

În modelele cauzale din domeniul sociouman se admite implicit ipoteza că un eveniment viitor nu poate fi o cauză a unui eveniment prezent sau trecut. Așadar modelul nu va permite ca un eveniment să fie o cauză pentru un alt eveniment “anterior” lui.

De multe erori neinterpretarea corectă a acestui deziderat constituie o permanentă sursă de erori, lucru ce va fi evidențiat printr-un exemplu.

Exemplul 9. Presupunem că în modelul teoretic cu variabile observabile cauza este determinată de variabila X (variabila independentă) efectul ei fiind sesizat de variabila Y (o variabilă dependentă).

Modelul are ecuația structurală

$$Y = a \cdot X + Z \quad (19)$$

unde perturbația Z verifică restricția $\text{Med}(Z) = 0$.

Modelul (19) va fi însă privit ca un model dinamic.

Deci la momente discrete t de timp, $t \in \{1, 2, 3, \dots\}$, sunt satisfăcute ecuațiile de structură

$$Y_t = a \cdot X_t + Z_t \quad (20)$$

De cele mai multe ori, în evoluția dinamică a sistemelor, este prezent un fenomen de “continuitate” ce se va manifesta în șirul valorilor observabile. Acest lucru ar putea fi sugerat prin relațiile :

$$X_t = b \cdot X_{t-1} + E_t \quad (21)$$

Valoarea coeficientului b caracterizează intensitatea fenomenului de “continuitate”.

Utilizând formulele (20) și (21) se stabilesc coeficienții ecuației de “continuitate” pentru șirul de valori $\{Y_t\}_{t \in \mathbb{N}}$:

$$\begin{aligned} Y_t &= a \cdot X_t + Z_t = a \cdot (b \cdot X_{t-1} + E_t) + Z_t = \\ &= b \cdot [a \cdot X_{t-1} + Z_{t-1} - Z_{t-1}] + a \cdot E_t + Z_t = \\ &= b \cdot Y_{t-1} + [a \cdot E_t + Z_t - b \cdot Z_{t-1}] \end{aligned} \quad (22)$$

În cadrul modelului dinamic vom presupune independența între erorile de măsurare precum și independența între “cauze” și erori, fapt exprimat prin relațiile :

$$\begin{aligned} \text{Cov}(X_t, Z_t) &= 0 ; \quad \text{Cov}(X_{t-1}, Z_t) = 0 ; \quad \text{Cov}(X_{t-1}, E_t) = 0 \\ \text{Cov}(Z_t, E_j) &= 0 \end{aligned} \quad (23)$$

Să admitem că dintr-o greșeală de interpretare, acceptăm variabila X ca fiind efectul acțiunii variabilei Y , adică

$$X_t = a_1 \cdot Y_{t-1} + V_t \quad (24)$$

coeficientul a_1 precizând gradul de influență al variabilei Y asupra variabilei X .

În acest context o estimare a valorii a_1 este dată de expresia

$$\begin{aligned} a_1 &= \text{cov}(X_t, Y_{t-1}) / \text{Var}(Y_{t-1}) = \\ &= \text{Cov}(b \cdot X_{t-1} + E_t, a \cdot X_{t-1} + Z_{t-1}) / \text{Var}(Y_{t-1}) = \\ &= a \cdot b \cdot \text{Var}(X_{t-1}) / \text{Var}(Y_{t-1}) \end{aligned} \quad (25)$$

Deoarece coeficienții a și b sunt nenuli atunci din (25) rezultă o valoare nenulă pentru a_1 . Acest fapt contrazice însă bunul simț deoarece variabila dependentă Y_{t-1} la momentul $t-1$ nu poate influența în nici un fel variabila independentă X_t la momentul t .

O greșeală asemănătoare, frecventă în domeniul sociouman, se manifestă atunci când în modelul empiric variabila independentă X nu poate fi măsurată chiar la momentul t , ci la momentul imediat următor, $t+1$.

În acest caz, acceptarea ecuației de structură

$$Y_t = a_2 \cdot X_{t+1} + Z_t \quad (26)$$

conduce la următoarea estimare a intensității a_2 a dependenței dintre variabilele X și Y :

$$\begin{aligned} a_2 &= \text{Cov}(X_{t+1}, Y_t) / \text{Var}(X_{t+1}) = \\ &= \text{Cov}(b \cdot X_t + E_t, a \cdot X_t + Z_t) / \text{Var}(X_{t+1}) = \\ &= a \cdot b \cdot \text{Var}(X_t) / \text{Var}(X_{t+1}) = a \cdot b \end{aligned} \quad (27)$$

unde am utilizat relațiile (20) și (21) și s-a presupus o nealterare în timp a varianței variabilelor X_t , în sensul $\text{Var}(X_t) = \text{Var}(X_{t+1})$.

Coeficientul $a_2 = a \cdot b$ astfel obținut diferă față de valoarea reală a , ce caracterizează influența variabilei independente X asupra variabilei dependente Y (formula (19)).

În modelele sociologice și psihologice sunt adeseori prezente generalizări ale modelului discutat anterior. Astfel, variabila Y la momentul t este dependentă atât de trecutul său cât și de variabilele independente X observabile până la momentul t , fiind satisfăcută o ecuație structurală de forma :

$$Y_t = b_1 \cdot Y_{t-1} + b_2 \cdot Y_{t-2} + \dots + a_0 \cdot X_t + a_1 \cdot X_{t-1} + a_2 \cdot X_{t-2} + \dots + Z_t \quad (28)$$

Menționăm că pentru domeniul socio-uman s-a acceptat tacit ipoteza că indicatorii construiți din variabile observabile sunt un “efect” al variabilei latente U pe care intenționează să o “reprezinte” (variabila U fiind, în fond, adevărata cauză).

În acest context, sensul relațiilor cauzale dintre o variabilă latentă U și mulțimea variabilelor observabile X asociate ei este descris de următoarea ecuație de structură:

$$X = d \cdot U + Z \quad (29)$$

variabila Z semnificând factorul perturbator.

Modelele cu variabile latente și observabile ce vor fi abordate în secțiunile următoare se vor baza, de cele mai multe ori, pe ecuații structurale de tipul (29).

Semnalăm însă faptul că există cercetări sociologice sau psihologice în care ipoteza enunțată nu este justificată.

Astfel, variabilele latente ale modelului empiric pot constitui “efecte” ale unor “cauze” determinate de variabilele măsurabile.

În această nouă accepțiune ecuațiile de structură ale modelului empiric sunt de forma:

$$U = d \cdot X + Z \quad (30)$$

variabila Z semnificând eroarea de aproximare.

Concluzii

Prin numeroase exemple s-a pus în evidență metodologia de stabilire a unei posibile relații de asociere, condițiile ca această asociere să devină raport de cauzalitate și eventuala precizare a sensului cauzalității. În acest context vom trata diferențiat tipurile de dependență (directă sau indirectă).

Postularea sensului raporturilor cauzale dintre variabilele latente și cele observabile este cerută într-o măsură însemnată de domeniul de cercetare abordat.

Raporturile cauzale dintre variabilele modelului empiric pot fi evidențiate mai ușor dacă se vor aplica transformări adecvate acestor variabile.

O importanță aparte trebuie acordată modelelor dinamice unde “efectul” acțiunii unei cauze se poate manifesta cu o întârziere esențială în interpretare. Asemenea modele dinamice sunt frecvente în domeniul sociouman.

II.4. TEHNICI DE SOLUȚIONARE A MODELELOR CAUZALE

4.1. Principiul statistic fundamental

În această secțiune vom urmări pe exemple modul de aplicare a procedurii generale de soluționare efectivă a unui model cauzal. Altfel spus, vom căuta să estimăm parametri necunoscuți a_1, a_2, \dots, a_k ai modelului empiric astfel încât modelul cauzal particular rezultat prin precizarea acestor parametri să fie "optim" din punctul de vedere al concordanței sale cu datele experimentale colectate.

Modelul empiric va conține variabilele observabile X_1, X_2, \dots, X_m și variabilele latente V_1, V_2, \dots, V_q , fiecare dintre acestea putând fi la rândul ei o variabilă endogenă sau exogenă.

În urma operațiunii de colectare a datelor $x_{i1}, x_{i2}, \dots, x_{in}$ privind variabila observabilă X_i , $1 \leq i \leq m$, se poate calcula matricea de covarianță empirică $C = (c_{ij})_{1 \leq i, j \leq m}$, unde :

$$\begin{aligned} c_{ij} &= (x_{i1} \cdot x_{j1} + x_{i2} \cdot x_{j2} + \dots + x_{in} \cdot x_{jn})/n - x_{i+} \cdot x_{j+} \\ x_{i+} &= (x_{i1} + x_{i2} + x_{i3} + \dots + x_{in})/n \end{aligned} \quad (1)$$

Menționăm faptul că matricea C a covarianțelor empirice dintre variabilele observabile nu depinde de parametri a_1, a_2, \dots, a_k ai modelului cauzal ci numai de măsurătorile concrete efectuate asupra variabilelor observabile respective.

Optarea pentru un anumit model empiric, adică propunerea anumitor relații de cauzalitate între variabilele acestuia, implică introducerea unor parametri necunoscuți a_1, a_2, \dots, a_k .

Parametrii a_t , $1 \leq t \leq k$, au în modelul experimental diferite semnificații ca de pildă : intensitatea legăturilor cauzale directe dintre variabilele modelului, mărimea erorilor de măsurare, indicatori privind gradul de fidelitate al modelului propus, etc.

În funcție de relațiile cauzale postulate între variabilele modelului se obțin ecuațiile de structură respective ce depind de parametrii a_t , $1 \leq t \leq k$.

În mod evident matricea de covarianță teoretică $S_0 = (s_{ij})_{1 \leq i, j \leq m+q}$ dintre variabilele modelului va depinde de parametrii acestuia a_1, a_2, \dots, a_k .

Fie $S = (s_{ij})_{1 \leq i, j \leq m}$ partea din matricea simetrică S_0 ce se referă numai la covarianțele teoretice s_{ij} dintre variabilele observabile, $s_{ij} = \text{Cov}(X_i, X_j)$, $1 \leq i, j \leq m$.

Estimarea parametrilor a_1, a_2, \dots, a_k se va face bazându-ne pe următorul principiu statistic fundamental, anume matricea covarianțelor teoretice $S(a_1, a_2, \dots, a_k)$ să fie "cât mai apropiată" de matricea covarianțelor experimentale C .

Dacă cele două matrici S și C "nu concordă" atunci putem spune cu certitudine că modelul empiric ales nu este adecvat, el fiind infirmat de datele experimentale. Atragem atenția asupra faptului că pot exista mai multe modele empirice având aceeași matrice S a covarianțelor teoretice.

Termenul "cât mai apropiat" folosit la compararea matricilor S și C nu a fost îndeajuns precizat, el putând avea, într-un anumit context, și accepțiunea de "identic". În această ultimă situație determinarea parametrilor a_t , $1 \leq t \leq k$, se va face astfel încât matricile S și C să aibă aceleași elemente. Menționăm faptul că o asemenea abordare nu este însă întotdeauna posibilă.

Exemplul 1. Fie modelul empiric cu variabile observabilele X_1, X_2, X_3 descris prin graful cauzal din figura 4.1.

Modelul are ecuațiile de structură :

$$\begin{aligned} X_2 &= a_1 \cdot X_1 + E_2 \\ X_3 &= a_3 \cdot X_1 + a_2 \cdot X_2 + E_3 \end{aligned} \quad (2)$$

unde E_1, E_2, E_3 sunt variabile-eroare.

Modelul (2) este un model ierarhic.

Variabila X_1 din modelul din figura 4.1 va fi interpretată ca variabilă exogenă, ea fiind în acest caz neafectată de erori. Această variabilă va influența variabilele endogene X_2 și X_3 .

Asupra variabilelor endogene X_2 și X_3 acționează și alți factori a căror comportare a fost sugerată prin prezența variabilelor “eroare” E_2 , respectiv E_3 .

Având în vedere presupunerile făcute este normal să acceptăm restricțiile :

$$\begin{aligned} \text{Cov}(X_1, E_2) = 0 & ; & \text{Cov}(X_1, E_3) = 0 \\ \text{Cov}(X_2, E_3) = 0 & ; & \text{Cov}(E_2, E_3) = 0 \end{aligned} \quad (3)$$

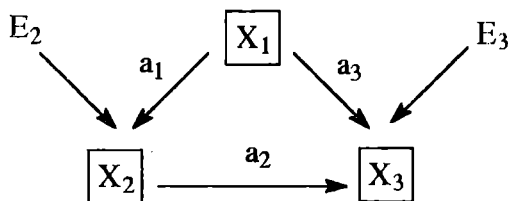


Figura 4.1.

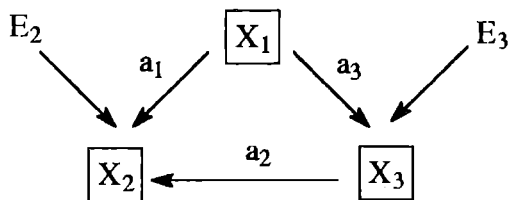


Figura 4.2.

Parametrii necunoscuți ai modelului din figura 4.1 sunt în număr de cinci. Astfel parametri a_1, a_2, a_3 sunt definiți în ecuațiile de structură (2) iar a_4 și a_5 reprezintă varianțele “erorilor” E_2 , respectiv E_3 ,

$$a_4 = \text{Var}(E_2) ; \quad a_5 = \text{Var}(E_3) \quad (4)$$

Valorile a_4 și a_5 sunt importante deoarece precizează gradul de influență asupra variabilelor endogene X_2, X_3 a altor factori E_2, E_3 ce acționează în afara factorului exogen X_1 .

Din datele experimentale $x_{i1}, x_{i2}, \dots, x_{in}$ referitoare la variabilele observabile $X_i, 1 \leq i \leq 3$, aplicând formulele (1) se obține efectiv matricea $C = (c_{ij})_{1 \leq i, j \leq n}$ ale cărei elemente sunt covarianțele empirice dintre variabilele măsurabile.

Utilizând proprietățile operatorilor de medie și covarianță precizate în secțiunea II.2.4, precum și ecuațiile de structură (2) cu restricțiile (3), rezultă egalitățile următoare :

$$\begin{aligned} \text{Var}(X_2) &= a_1^2 \cdot \text{Var}(X_1) + a_4 \\ \text{Cov}(X_1, X_2) &= a_1 \cdot \text{Var}(X_1) \\ \text{Cov}(X_1, X_3) &= [a_3 + a_1 \cdot a_2] \cdot \text{Var}(X_1) \\ \text{Cov}(X_2, X_3) &= a_1 \cdot [a_3 + a_1 \cdot a_2] \cdot \text{Var}(X_1) + a_2 \cdot a_4 \\ \text{Var}(E_3) &= a_3^2 \cdot \text{Var}(X_1) + a_2^2 \cdot \text{Var}(X_2) + a_5 + 2 \cdot a_2 \cdot a_3 \cdot \text{Cov}(X_1, X_2) \end{aligned} \quad (5)$$

Ținând seama de notațiile făcute, $s_{ij} = \text{Cov}(X_i, X_j)$, $s_{ii} = \text{Var}(X_i)$, $1 \leq i, j \leq 3$, identitățile (5) conduc la definirea matricii simetrice $S = (s_{ij})_{1 \leq i, j \leq 3}$, ale cărei elemente sunt covarianțelor teoretice dintre variabilele modelului. Astfel :

$$\begin{aligned} s_{12} = s_{21} &= a_1 \cdot s_{11} ; \quad s_{13} = s_{31} = (a_3 + a_1 \cdot a_2) \cdot s_{11} \\ s_{22} &= a_1^2 \cdot s_{11} + a_4 ; \quad s_{23} = s_{32} = a_1 \cdot (a_3 + a_1 \cdot a_2) \cdot s_{11} + a_2 \cdot a_4 \\ s_{33} &= a_3^2 \cdot s_{11} + a_2^2 + s_{22} + a_5 + 2 \cdot a_2 \cdot a_3 \cdot s_{12} \end{aligned} \quad (6)$$

Identificând matricea covarianțelor teoretice S cu matricea C a covarianțelor empirice obținem un sistem neliniar de șase ecuații în necunoscutele $s_{11}, a_1, a_2, a_3, a_4, a_5$:

$$\begin{aligned} s_{11} &= c_{11} \\ a_1 \cdot s_{11} &= c_{12} \\ [a_3 + a_1 \cdot a_2] \cdot s_{11} &= c_{13} \\ a_1^2 \cdot s_{11} + a_4 &= c_{22} \\ a_1 \cdot [a_3 + a_1 \cdot a_2] \cdot s_{11} + a_2 \cdot a_4 &= c_{23} \\ a_3^2 \cdot s_{11} + a_2^2 \cdot [a_1^2 \cdot s_{11} + a_4] + a_5 + 2 \cdot a_1 \cdot a_2 \cdot a_3 \cdot s_{11} &= c_{33} \end{aligned} \quad (7)$$

Utilizând procedeul substituției se obțin următoarele estimări a_1, a_2, a_3, a_4, a_5 ale parametrilor modelului cauzal din figura 4.1 :

$$\begin{aligned}
 a_1 &= c_{12}/c_{11} \\
 a_2 &= [c_{11} \cdot c_{23} - c_{12} \cdot c_{13}] / [c_{11} \cdot c_{22} - c_{12}^2] \\
 a_3 &= [c_{13} \cdot c_{22} - c_{12} \cdot c_{23}] / [c_{11} \cdot c_{22} - c_{12}^2] \\
 a_4 &= c_{22} - c_{12}^2 / c_{11} \\
 a_5 &= c_{33} - a_3^2 \cdot c_{11} - a_2^2 \cdot c_{22} - 2 \cdot a_2 \cdot a_3 \cdot c_{12}
 \end{aligned} \tag{8}$$

Menționăm faptul că modelul cauzal din figura 4.1 poate fi soluționat întotdeauna. Într-adevăr, cu valorile $c_{ij}, 1 \leq i, j \leq 3$, deduse din măsurătorile $x_{it}, 1 \leq i \leq 3, 1 \leq t \leq n$, efectuate asupra variabilelor observabile X_1, X_2, X_3 obținem estimările parametrilor necunoscuți a_1, a_2, a_3, a_4, a_5 prin aplicarea formulelor (8).

Exemplul 2. Fie modelul cauzal, ierarhic, cu variabile observabile, descris în figura 4.2 și având ecuațiile de structură :

$$X_2 = a_1 \cdot X_1 + a_2 \cdot X_3 + E_2 \tag{9}$$

$$X_3 = a_3 \cdot X_1 + E_3$$

îndeplinind condițiile

$$\text{Cov}(X_1, E_2) = 0 \quad ; \quad \text{Cov}(X_1, E_3) = 0$$

$$\text{Cov}(X_3, E_2) = 0 \quad ; \quad \text{Cov}(E_2, E_3) = 0 \tag{10}$$

Modelul (9) se deosebește de modelul (2) (figurile 4.1-4.2) prin inversarea sensului relației cauzale dintre variabilele măsurabile X_2 și X_3 . În fond din modelul (2) se obține modelul (9) efectuând permutările de variabile : X_2 cu X_3 și E_2 cu E_3 .

Modelul din figura 4.2 este, de asemenea, întotdeauna soluționabil. Un calcul analog celui din exemplul 1 ne conduce la următoarele estimări ale parametrilor modelului (9) :

$$\begin{aligned}
 a_1 &= [c_{12} \cdot c_{33} - c_{13} \cdot c_{23}] / [c_{11} \cdot c_{33} - c_{13}^2] \\
 a_2 &= [c_{11} \cdot c_{23} - c_{13} \cdot c_{12}] / [c_{11} \cdot c_{33} - c_{13}^2] \\
 a_3 &= c_{13} / c_{11}
 \end{aligned} \tag{11}$$

$$a_4 = c_{22} - a_1^2 \cdot c_{11} - a_2^2 \cdot c_{33} - 2 \cdot a_1 \cdot a_2 \cdot c_{13}$$

$$a_5 = c_{33} - c_{13}^2 / c_{11}$$

unde evident $a_4 = \text{Var}(E_2)$, $a_5 = \text{Var}(E_3)$.

Estimarea parametrilor a_1, a_2, \dots, a_k ai modelului cauzal prin procedeul de identificare a matricii covarianțelor teoretice S cu matricea C a covarianțelor empirice nu este întotdeauna posibilă. Vom da un exemplu în acest sens.

Exemplul 3. Fie modelul cauzal cu variabile observabile descris prin graful din figura 4.3, cu ecuațiile structurale:

$$X_2 = a_2 \cdot X_1 + E_2$$

$$X_3 = a_3 \cdot X_1 + E_3 \quad (12)$$

Variabilele X_2, X_3 vor fi considerate drept variabile endogene, ele fiind influențate de variabila exogenă X_1 precum și de alți factori ce sunt grupați în variabilele E_2 , respectiv E_3 .

Presupunerea făcută impun următoarele restricții :

$$\text{Cov}(X_1, E_2) = 0 ; \text{Cov}(X_1, E_3) = 0 ; \text{Cov}(E_2, E_3) = 0 \quad (13)$$

Să determinăm varianțele și covarianțele teoretice dintre variabilele modelului analizat în funcție de parametrii necunoscuți a_2, a_3, a_4, a_5 , unde :

$$a_4 = \text{Var}(E_2) ; a_5 = \text{Var}(E_3) \quad (14)$$

Ținând seama de (12) și (13) deducem următoarele expresii pentru elementele s_{ij} , $1 \leq i, j \leq 3$, ale matricii S a covarianțelor teoretice :

$$s_{11} = \text{Var}(X_1)$$

$$s_{22} = a_2^2 \cdot s_{11} + a_4$$

$$s_{33} = a_3^2 \cdot s_{11} + a_5 \quad (15)$$

$$s_{12} = a_2 \cdot s_{11}$$

$$s_{13} = a_3 \cdot s_{11}$$

$$s_{23} = a_2 \cdot a_3 \cdot s_{11}$$

Identificând matricea S cu matricea C obținem în final următorul sistem neliniar de 5 ecuații în necunoscutele a_j :

$$\begin{aligned} a_2^2 \cdot c_{11} + a_4 &= c_{22} \\ a_3^2 \cdot c_{11} + a_5 &= c_{33} \\ a_2 \cdot c_{11} &= c_{12} \\ a_3 \cdot c_{11} &= c_{13} \\ a_2 \cdot a_3 \cdot c_{11} &= c_{23} \end{aligned} \tag{16}$$

Sistemul (16) nu are întotdeauna soluție deoarece valorile a_2 și a_3 obținute din ecuațiile 3 și 4 nu verifică neaparat și ecuația 5. Într-adevăr relația:

$$c_{12} \cdot c_{13} = c_{11} \cdot c_{23} \tag{17}$$

ramâne de cele mai multe ori nesatisfăcută pentru valorile c_{ij} deduse din datele experimentale x_{it} (formula (1)).

4.2. Criterii concrete

Așadar nu întotdeauna este posibil de a estima valorile parametrilor a_1, a_2, \dots, a_k din condiția de identificare a matricelor de covarianță, teoretică și empirică, $S(a_1, \dots, a_k) = C$

Suntem nevoiți să rezolvăm un sistem cu $m \cdot (m+1)/2$ relații neliniare și având k necunoscute a_1, a_2, \dots, a_k ,

$$s_{ij}(a_1, a_2, \dots, a_k) = c_{ij} \quad ; \quad 1 \leq i \leq j \leq m \tag{18}$$

unde m este numărul variabilelor măsurabile ale modelului cauzal.

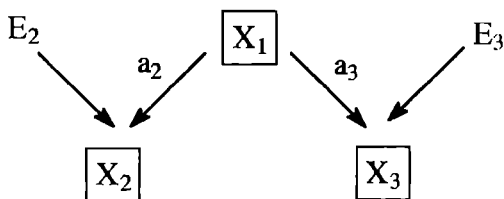


Figura 4.3

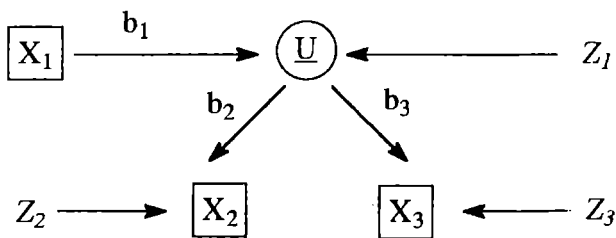


Figura 4.4

De multe ori sistemul(18) este supradimensionat rezultând mai multe restricții decât numărul k de parametri necunoscuți.

Astfel în cazul

$$m \cdot (m + 1) > 2 \cdot k \quad (19)$$

sistemul (18) nu are în general “soluție exactă”.

În această situație estimarea parametrilor a_1, a_2, \dots, a_k se va face astfel încât matricile S și C să fie “cât mai apropiate”. Parametrii a_1, a_2, \dots, a_k se vor determina astfel încât “distanța” $d(S,C)$ dintre matricile S și C să fie minimă.

Precizăm faptul că există o multitudine de posibilități de a defini “distanțe” dintre două matrici.

Cea mai des folosită distanță este distanța euclidiană d_1 ,

$$d_1^2(S,C) = (s_{11} - c_{11})^2 + \dots + (s_{ij} - c_{ij})^2 + \dots + (s_{mm} - c_{mm})^2 \quad (20)$$

Putem însă opta și pentru alte “distanțe” ca de pildă d_2 ,

$$d_2(S,C) = |s_{11} - c_{11}| + \dots + |s_{ij} - c_{ij}| + \dots + |s_{mm} - c_{mm}| \quad (21)$$

Deoarece într-un model causal precizia de măsurare a variabilelor observabile nu este, de regulă, aceeași, sugerăm utilizarea unor distanțe “ponderate”. Astfel, se vor acorda ponderi distincte pentru diverse variabile observabile, perechi sau grupuri de variabile.

O dificultate majoră a acestui mod de abordare constă în stabilirea criteriilor de definire concretă a ponderilor.

În literatura de specialitate, în cadrul studiului modelelor cauzale, sunt utilizate și alte tipuri de “distanțe” selecționate în urma unor considerente statistice. Înțelegerea semnificației acestora necesită o pregătire matematică adecvată.

Menționăm faptul că, prin folosirea diverselor distanțe, obținem estimări diferite pentru parametrii modelului cauzal analizat. Prin urmare, în final, este necesară aplicarea unor teste statistice (de exemplu testul hi-patrat pentru a alege sau infirma anumite variante. Nu vom intra însă în detalii tehnice.

Pe baza unui exemplu simplu vom căuta să sugerăm complexitatea unei astfel de probleme.

Exemplu 4. Dorim să soluționăm modelul cauzal din figura 4.3 (vezi exemplul 3) știind că din datele experimentale au fost deduse următoarele valori (formula 1) :

$$\begin{aligned} c_{11} = 1 ; \quad c_{22} = 2 ; \quad c_{33} = 3 ; \\ c_{12} = 1 ; \quad c_{13} = 3/2 ; \quad c_{23} = 2 ; \end{aligned} \quad (22)$$

Conform exemplului 3, soluționarea modelului cauzal se reduce la rezolvarea “aproximativă” a sistemului supradimensionat (16), în necunoscutele a_2, a_3, a_4, a_5 . În acest sens vom opta pentru aplicarea “distanței” d_1 dată de formula (20).

Așadar vom determina acele valori a_2, a_3, a_4, a_5 care asigură minimul funcției h_1 ,

$$\begin{aligned} h_1(a_2, a_3, a_4, a_5) = (s_{11} - c_{11})^2 + (s_{12} - c_{12})^2 + (s_{13} - c_{13})^2 + \\ + (s_{22} - c_{22})^2 + (s_{23} - c_{23})^2 + (s_{33} - c_{33})^2 \end{aligned} \quad (23)$$

unde constantele c_{ij} sunt cele specificate în (22) iar expresiile s_{ij} sunt calculate prin aplicarea formulelor (15).

Minimizarea funcției h_1 este în fapt o optimizare cu restricții, valorile parametrilor a_4 și a_5 trebuind să fie pozitive (ele semnifică varianțe ale erorilor, formula (14)).

În vederea simplificării, din punct de vedere matematic, a expunerii vom accepta condițiile restrictive:

$$s_{11} = c_{11} ; \quad s_{22} = c_{22} ; \quad s_{33} = c_{33} \quad (24)$$

Pentru rezolvarea sistemului (16) în condițiile (22) este suficient să găsim o “soluție aproximativă” a sistemului derivat

$$\begin{aligned} a_2 &= 1 \\ a_3 &= 3/2 \\ a_2 \cdot a_3 &= 2 \end{aligned} \quad (25)$$

Rezolvarea sistemului (25) se reduce la minimizarea funcției

$$h_2(a_2, a_3) = (a_2 - 1)^2 + (a_3 - 3/2)^2 + (a_2 \cdot a_3 - 2)^2 \quad (26)$$

Exemplul 5. Considerăm modelul cauzal din figura 4.4, cu variabila latentă \underline{U} și variabilele observabile X_1, X_2, X_3 .

Variabila X_1 va fi privită ca variabilă exogenă, restul variabilelor fiind endogene. Variabilele eroare au fost desemnate prin Z_1, Z_2, Z_3 .

Un asemenea tip de model apare frecvent în cercetările socioumane atunci când suntem nevoiți să măsurăm influența b_1 , a unei variabile observabile X_1 , asupra variabilei latente \underline{U} . Efectul variabilei latente este determinat indirect prin intermediul variabilelor măsurabile X_2 și X_3 .

Așadar modelul cauzal descris în figura 4.4 are ecuațiile de structură următoare :

$$\begin{aligned} \underline{U} &= b_1 \cdot X_1 + Z_1 \\ X_2 &= b_2 \cdot \underline{U} + Z_2 \\ X_3 &= b_3 \cdot \underline{U} + Z_3 \end{aligned} \quad (29)$$

și sunt indeplinite restricțiile :

$$\text{Cov}(X_i, Z_i) = 0 ; \quad \text{Cov}(Z_i, Z_j) = 0 ; \quad 1 \leq i, j \leq 3 \quad (30)$$

Substituind prima ecuație de structură (29) în următoarele două, rezultă egalitățile :

$$\begin{aligned} X_2 &= b_1 \cdot b_2 \cdot X_1 + (b_2 \cdot Z_1 + Z_2) \\ X_3 &= b_1 \cdot b_3 \cdot X_1 + (b_3 \cdot Z_1 + Z_3) \end{aligned} \quad (31)$$

Cu notațiile

$$\begin{aligned} a_2 &= b_1 \cdot b_2 ; & E_2 &= b_2 \cdot Z_1 + Z_2 \\ a_3 &= b_1 \cdot b_3 ; & E_3 &= b_3 \cdot Z_1 + Z_3 \end{aligned} \quad (32)$$

ecuațiile (31) se rescriu în forma :

$$\begin{aligned} X_2 &= a_2 \cdot X_1 + E_2 \\ X_3 &= a_3 \cdot X_1 + E_3 \end{aligned} \quad (33)$$

adică tocmai ecuațiile de structură (12) ale modelului 4.3.

Suntem astfel tentați să reducem modelul causal din figura 4.4 la modelul causal descris în figura 4.3, model deja studiat.

O asemenea abordare este eronată, în primul rând datorită faptului că nu sunt respectate restricțiile (13) din exemplul 3.

Într-adevăr, folosind condițiile (30) obținem :

$$\text{Cov}(X_1, E_2) = \text{Cov}(X_1, b_2 \cdot Z_1 + Z_2) = b_2 \cdot \text{Cov}(X_1, Z_1) + \text{Cov}(X_1, Z_2) = 0 \quad (34)$$

$$\text{Cov}(X_1, E_3) = \text{Cov}(X_1, b_3 \cdot Z_1 + Z_3) = b_3 \cdot \text{Cov}(X_1, Z_1) + \text{Cov}(X_1, Z_3) = 0$$

dar

$$\text{Cov}(E_2, E_3) = \text{Cov}(b_2 \cdot Z_1 + Z_2, b_3 \cdot Z_1 + Z_3) = b_2 \cdot b_3 \cdot \text{Var}(Z_1) \quad (35)$$

expresie ce este în general nenulă ($\text{Var}(Z_1) \leq 0$).

Cu notațiile :

$$s_{ij} = \text{Cov}(X_i, X_j) ; \quad 1 \leq i, j \leq 3 \quad (36)$$

$$b_4 = \text{Var}(Z_2) ; \quad b_5 = \text{Var}(Z_3) ; \quad b_6 = \text{Var}(Z_1)$$

un calcul similar celui din (15) ne conduce la aflarea matricii S a covarianțelor teoretice, $S = (s_{ij})_{1 \leq i, j \leq 3}$.

Soluționarea modelului se reduce la găsirea unei "soluții aproximative" $b_1, b_2, b_3, b_4, b_5, b_6$ a sistemului neliniar :

$$s_{11} = c_{11}$$

$$b_1^2 \cdot b_2^2 \cdot s_{11} + b_2^2 \cdot b_6 + b_4 = c_{22}$$

$$b_1^2 \cdot b_3^2 \cdot s_{11} + b_3^2 \cdot b_6 + b_5 = c_{33}$$

$$b_1 \cdot b_2 \cdot s_{11} = c_{12} \quad (37)$$

$$b_1 \cdot b_3 \cdot s_{11} = c_{13}$$

$$b_1^2 \cdot b_2 \cdot b_3 \cdot s_{11} + b_2 \cdot b_3 \cdot b_6 = c_{23}$$

Rezolvarea sistemului (37) se va face analog modului de abordare din exemplul 4, prin utilizarea, de exemplu, a distanței d_1 dată de formula (20).

4.3. Soluționarea unor modele clasice

În această secțiune intenționăm să abordăm modele statistice clasice utilizând metodologia pentru soluționarea modelelor cauzale. Ne vom referi la modele de regresie liniară și la modele de analiză factorială. Analizarea acestor modele a fost făcută pe unele cazuri particulare, simplificate, urmărindu-se astfel cu ușurință calculele matematice și principiile aplicate.

Exemplul 6. În secțiunea II.1.2 (formula (4)) am rezolvat, printr-un procedeu standard, modelul liniar de regresie cu o variabilă independentă,

$$Y = a_1 \cdot X + E \quad (38)$$

în ipoteza independenței erorii de măsurare E de variabila X ,

$$\text{Cov}(X, E) = 0 \quad (39)$$

Modelul cauzal a fost de fapt descris în secțiunea II.1.2, figura 1.3 și are trei parametri : a_1 , $a_2 = \text{Var}(X)$, $a_3 = \text{Var}(E)$.

Variabilele X și Y sunt variabile endogene, ambele observabile.

Un calcul elementar și utilizarea relațiilor (38)-(39) ne conduc la determinarea matricei S a covarianțelor teoretice:

$$S = \begin{vmatrix} \text{Var}(X) & \text{Cov}(X, Y) \\ \text{Cov}(X, Y) & \text{Var}(Y) \end{vmatrix} = \begin{vmatrix} a_2 & a_1 \cdot a_2 \\ a_1 \cdot a_2 & a_1^2 \cdot a_2 + a_3 \end{vmatrix} \quad (40)$$

Identificând matricea S cu matricea covarianțelor empirice C se obține un sistem neliniar de 3 ecuații cu 3 necunoscute :

$$\begin{aligned} a_2 &= c_{11} \\ a_1 \cdot a_2 &= c_{12} \\ a_1^2 \cdot a_2 + a_3 &= c_{22} \end{aligned} \quad (41)$$

a cărui soluție este:

$$a_1 = c_{12}/c_{11} ; \quad a_2 = c_{11} ; \quad a_3 = c_{22} - c_{12}^2/c_{11} \quad (42)$$

Estimația parametrului a_3 este importantă în evaluarea credibilității unei dependențe liniare de forma (38).

Valoarea estimată a parametrului a_1 din ecuația de structură (38) este egală cu c_{12}/c_{11} , aceeași valoare fiind obținută și în cazul unei abordări statistice clasice (secțiunea II.1.2, formula (18)).

Extinderea acestui model în varianta cu m variabile independente nu implică dificultăți majore. Vom justifica aceasta pentru $m = 2$.

Exemplul 7. Modelul de regresie liniară cu două variabile independente X_1, X_2 :

$$Y = a_1 \cdot X_1 + a_2 \cdot X_2 + E \quad (43)$$

este dat de graful cauzal descris în figura 4.5, cu restricțiile :

$$\text{Cov}(X_1, E) = 0 ; \quad \text{Cov}(X_2, E) = 0 \quad (44)$$

Variabilele X_1, X_2, Y sunt variabile endogene și măsurabile, E fiind variabila-eroare.

Identificând matricea S a covarianțelor teoretice

$$S = \begin{pmatrix} \text{Var}(X_1) & \text{Cov}(X_1, X_2) & \text{Cov}(X_1, Y) \\ \text{Cov}(X_2, X_1) & \text{Var}(X_2) & \text{Cov}(X_2, Y) \\ \text{Cov}(Y, X_1) & \text{Cov}(Y, X_2) & \text{Var}(Y) \end{pmatrix} \quad (45)$$

cu matricea covarianțelor experimentale C obținem un sistem neliniar de 6 ecuații cu 6 necunoscute:

$$\begin{aligned} a_3 &= c_{11} \\ a_4 &= c_{22} \\ a_1^2 \cdot a_3 + a_2^2 \cdot a_4 + a_5 + 2 \cdot a_1 \cdot a_2 \cdot a_6 &= c_{33} \\ a_6 &= c_{12} \\ a_1 \cdot a_3 + a_2 \cdot a_6 &= c_{13} \\ a_1 \cdot a_6 + a_2 \cdot a_4 &= c_{23} \end{aligned} \quad (46)$$

unde noii parametri introduși au semnificațiile : $a_3 = \text{Var}(X_1)$, $a_4 = \text{Var}(X_2)$, $a_5 = \text{Var}(X_3)$, $a_6 = \text{Cov}(X_1, X_2)$.

Prin rezolvarea sistemului neliniar (46), deducem valorile coeficienților a_1 și a_2 ale ecuației de structură (43) precum și varianța a_5 a erorii comise :

$$\begin{aligned}
 a_1 &= [c_{13} \cdot c_{22} - c_{23} \cdot c_{12}] / [c_{11} \cdot c_{22} - c_{12}^2] \\
 a_2 &= [c_{11} \cdot c_{23} - c_{13} \cdot c_{12}] / [c_{11} \cdot c_{22} - c_{12}^2] \\
 a_5 &= c_{33} - c_{11} \cdot a_1^2 - c_{22} \cdot a_2^2 - 2 \cdot c_{12} \cdot a_1 \cdot a_2
 \end{aligned}
 \tag{47}$$

În cercetările socio-umane este des folosită analiza factorială atât pentru reducerea numărului efectiv de indicatori cu care se lucrează cât și pentru evidențierea unor factori ce nu pot fi măsurați în mod direct. Detalii privind aceste aspecte pot fi obținute în J. Torrens - Ibern(Modèles et méthodes de l'analyse factorielle, Edition Dunod, Paris, 1972).

Vom încerca să soluționăm un model clasic de analiză factorială prin tehnica variabilelor latente.

Exemplul 8. Considerăm variabilele măsurabile X_1, X_2, \dots, X_m și intenționăm să determinăm ponderile a_1, a_2, \dots, a_m cu care un factor comun \underline{U} influențează fiecare dintre aceste variabile.

Avem în fond modelul cauzal descris în figura 4.6, unde X_1, X_2, \dots, X_m sunt variabile observabile iar factorul \underline{U} neprecizat și imposibil de măsurat în mod direct, este variabila latentă.

Toate aceste variabile sunt endogene iar erorile E_j satisfac restricțiile:

$$\begin{aligned}
 \text{Cov}(E_i, E_j) &= 0, \quad 1 \leq i < j \leq m \\
 \text{Cov}(\underline{U}, E_j) &= 0, \quad 1 \leq j \leq m
 \end{aligned}
 \tag{48}$$

Variabila latentă \underline{U} reprezintă “partea comună” prezentă în variabilele X_1, X_2, \dots, X_m și care acționează din “umbră” dictând, mai mult sau mai puțin, un comportament unitar, de grup.

Ecuatiile structurale ale modelului sunt:

$$X_j = a_j \cdot \underline{U} + E_j ; \quad 1 \leq j \leq m \tag{49}$$

Introducând parametrii suplimentari:

$$\begin{aligned}
 b_j &= \text{Var}(E_j) ; \quad 1 \leq j \leq m \\
 a_0 &= \text{Var}(\underline{U})
 \end{aligned}
 \tag{50}$$

aplicând proprietățile operatorului de covarianța și utilizând relațiile (49)-(50) obținem:

$$s_{ij} = \text{Cov}(X_i, X_j) = \text{Cov}(a_i \cdot \underline{U} + E_i, a_j \cdot \underline{U} + E_j) = a_i \cdot a_j \cdot a_0$$

$$s_{jj} = \text{Var}(X_j) = a_j^2 \cdot \text{Var}(\underline{U}) + \text{Var}(E_j) = a_j^2 \cdot a_0 + b_j \quad (51)$$

Așadar identificarea matricilor de covarianță teoretică și empirică conduce la soluționarea următorului sistem nelinier având $m \cdot (m + 1) / 2$ ecuații cu $2 \cdot m + 1$ necunoscute :

$$a_j^2 \cdot a_0 + b_j = c_{jj} \quad ; \quad 1 \leq j \leq m$$

$$a_i \cdot a_j \cdot a_0 = c_{ij} \quad ; \quad 1 \leq i < j \leq m \quad (52)$$

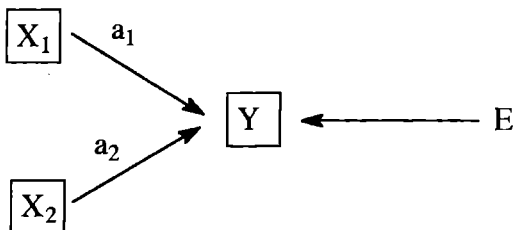


Figura 4.5.

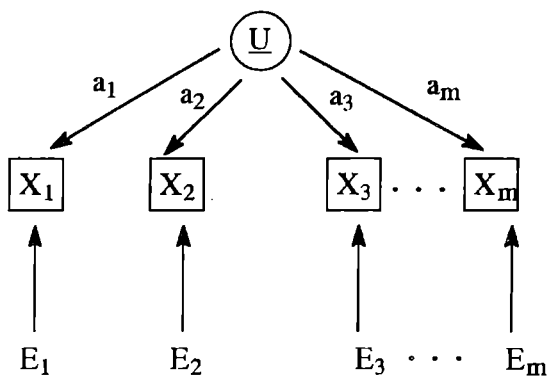


Figura 4.6.

Pentru rezolvarea “aproximativă” a sistemului nelinier (52) se utilizează de regulă distanța euclidiană dată de formula (20).

Prin introducerea mai multor variabile latente în modelul cauzal din figura 4.6 se obțin modele de analiză factorială cu mai mulți factori regăsindu-se rezultate clasice.

Concluzii

Estimarea parametrilor modelelor cauzale nu se face la întâmplare ci se bazează pe respectarea unor principii.

Formularea unor astfel de principii este legată de filosofia modelului în cauză. În acest context, pentru modelele dinamice evolutive am propus “principiul minimizării erorilor”, inspirat din mecanică pentru caracterizarea stărilor de echilibru.

În urma aplicării acestor principii, soluționarea modelului cauzal se reduce la o problemă tehnică de matematică, în general la o problemă de optimizare cu restricții. În acest sens sunt, de regulă, utilizați algoritmi iterativi, programați pe calculator.

Exemplele prezentate au fost alese astfel încât rezolvarea lor să fie facilă din punct de vedere matematic, dar cu intenția vădită de a pune în evidență principiile aplicate, precum și dificultățile conceptuale ce pot, în orice moment, să apară.

Prin aplicarea metodologiei modelelor cauzale au fost regăsite rezultate clasice. Este cazul analizei de regresie și analizei factoriale. Avantajul acestor procedee constă în tratarea unitară a tuturor variabilelor modelului studiat, obținerea soluțiilor fiind elegantă și de multe ori mai facilă.

Utilizarea unor produse soft pentru estimarea parametrilor modelului cauzal nu trebuie să se desfășoare în mod automat, fără discernământ. Devine obligatorie testarea mai multor variante, pentru care este necesară clarificarea principiile aplicate precum și a ipotezelor de lucru folosite.

Trebuie avut permanent în minte caracterul relativ al rezultatelor obținute, rezultate ce ar putea fi modificate în momentul obținerii de informații suplimentare privind variabilele modelului sau diversele raporturi dintre acestea.

Estimații ale parametrilor, care sunt comparativ apropiate și care au fost deduse prin aplicarea sistematică a unor procedee de lucru diferite, vin în sprijinul consolidării unei anumite imagini privind interdependențele dintre variabilele modelului cauzal teoretic.

II.5. RAPORTUL DINTRE VARIABILELE LATENTE ȘI VARIABILELE MĂSURABILE

Relațiile dintre variabilele latente și variabilele observabile (măsurabile) sunt descrise în cadrul unui model causal prin intermediul ecuațiilor de structură. În acest proces are o foarte mare importanță modul în care este conceput modelul, selectarea variabilelor observabile și interdependențele cu variabilele latente, modul de validare al măsurătorilor precum și gradul lor de credibilitate.

5.1. Conceperea modelelor cauzale

În secțiunile anterioare au fost evidențiate probleme ce apar în cazul în care trebuie soluționat efectiv un model causal deja precizat : optarea pentru anumite tehnici de estimare a parametrilor, impunerea unor ipoteze suplimentare în cazul în care nu este posibilă identificarea parametrilor modelului, reevaluarea legăturilor cauzale dintre variabilele admise (variabile observabile și variabile latente) precum și tipul lor în cadrul modelului (variabile endogene și variabile exogene).

Ideea primordială urmărită este de a soluționa modelul causal prin “ajustarea” anumitor raporturi dintre variabilele sale (se vor propune noi ipoteze sau se va renunța succesiv la alte ipoteze “neconcludente”).

În continuare, vom pune în evidență aspecte premergătoare acestei faze, strâns legate de modul de concepere al modelului causal.

În această perspectivă este obligatorie parcurgerea următorilor pași:

- 1. Definirea teoretică a conceptului ce urmează a fi studiat ;**
- 2. Precizarea dimensiunilor sale ;**
- 3. Efectuarea de observații ;**
- 4. Propunerea unui model.**

Vom detalia aceasta problematică.

1) Procesul de măsurare are la bază un concept teoretic cu un grad de abstractizare mai mare sau mai mic. Astfel, conceptul de inteligență (sau cel de dezvoltare economică) presupune un grad foarte

ridicat de abstractizare, pe când conceptul de pregătire școlară este cu mult mai concret.

Fiecare model causal este axat pe studierea unui anumit fenomen, bazându-se pe un concept teoretic. Conceptul identifică aspecte caracteristice fenomenului studiat. Este discutabil însă modul de delimitare a unui concept în vederea caracterizării unui fenomen.

Odată definit conceptul ce va sta la baza modelului, se vor preciza caracteristici de operativitate, detalii specifice, semnificația unor termeni.

De exemplu, putem accepta următoarea definiție teoretică a conceptului de "sărăcie" (Dicționar sociologic, coordonat de Cătălin Zamfir, Lazăr Vlăsceanu, București, 1993): "Situatie caracterizată prin lipsa (insuficiența) mijloacelor materiale necesare vieții".

În cadrul acestui concept pot fi evidențiate următoarele aspecte:

- Grade diferite de sărăcie în funcție de natura mijloacelor de subsistență.

- Sărăcie individuală și sărăcie colectivă.

- Factorii răspunzători de amploarea sărăciei: nivelul de dezvoltare economică, situația pieței forței de muncă, gradul de ocupare a forței de muncă și șomajul, caracteristicile locurilor de muncă, tipul de control asupra resurselor materiale, inflația, modul de distribuție a veniturilor, caracteristici demografice, sociale și culturale, stilul de viață.

- Pragul de sărăcie și forme agravante de manifestare (subnutriția și foamea), nivel minim de trai decent (formule de calcul).

- Standarde utilizate pentru trasarea pragului de sărăcie: nivelul general al costului vieții, programe prioritare de sprijin.

- Sărăcie primară și secundară; nevoi de subsistență fizică, sărăcia stilului de viață, sărăcie culturală.

- Segmente ale populației afectate de sărăcie.

2) În funcție de problematica abordată ne vom opri la un număr mai mare sau mai mic de caracteristici ale conceptului teoretic enunțat. Numărul acelor caracteristici majore certifică dimensiunea conceptului, adică numărul de aspecte distincte în cadrul conceptului.

Există și posibilitatea divizării unei componente majore în mai multe subcomponente, fapt ce va conduce la o dimensiune variabilă a aceleiași concept, în funcție de natura concretă a studiului de cercetare întreprins.

Fiecărei caracteristici (dimensiuni) astfel precizată i se atașează în modelul cauzal câte o variabilă latentă.

Astfel, în cadrul studierii conceptului de “sărăcie” având ca scop estimarea nivelului de sărăcie și a gradului de afectare a diverselor segmente din populație, au fost construite următoarele variabile latente: status socio-economic, aprecierea subiectivă a standardului de viață, aspirații materiale. Evidențierea și măsurarea unor raporturi cauzale dintre variabilele modelului sugerează dinamica și complexitatea fenomenului respectiv, moduri de abordare concretă, constructivă.

3) Pasul următor definirii variabilelor latente ale modelului constă în a preciza modalitățile prin care aceste variabile vor fi “apreciate indirect”, prin propunerea unor variabile măsurabile asociate variabilelor latente.

Variabilele observabile pot fi foarte diverse: răspunsuri la chestionare, rapoarte, rezultate ale recensămintelor, date furnizate periodic de anumite aparate de măsură, reportaje mass-media, etc.

Accastă multitudine de măsurători presupune scale diferite, de multe ori necomparabile (uneori este necesară ajustarea lor), tipuri diverse de date (cantitative sau calitative, discrete sau continue).

Alegerea adecvata a variabilelor observabile și a modului lor de măsurare este de o importanță majoră în micșorarea dispersiilor variabilelor latente.

4) Ultima etapă a acestui proces de definitivare a modelului cauzal constă în propunerea unor indicatori și formularea relațiilor de structură dintre acești indicatori și variabilele latente asociate caracteristicilor dominante ale conceptului teoretic studiat.

Remarcăm impunerea unor ipoteze suplimentare în concordanță cu fenomenul sociologic studiat. Astfel, în cazul unui model evolutiv în timp, “cauzele” vor fi întotdeauna variabile (latente sau măsurabile) anterioare din punct de vedere temporal, viitorul neputând influența trecutul.

În multe modele sociologice anumite dependențe între variabile nu pot exista în realitate, chiar dacă măsurătorile efectuate sugerează o foarte puternică corelație între datele experimentale respective. În acest sens menționăm exemplul clasic al unui studiu britanic în care s-a determinat o foarte puternică, corelare între numărul C de copii născuți în mediul

rural și numărul B al cuiburilor de berze din comunitățile respective. Această corelare semnificativă nu poate sugera un model causal de tipul $B \longrightarrow C$ sau $C \longrightarrow B$ deoarece este unanim acceptat faptul că berzele nu contribuie în mod determinant la nașterea mai multor copii (și reciproc, nașterea unui număr mare de copii nu influențează decisiv creșterea numărului cuiburilor de barză).

5.2. Tehnici de validare

Variabilele latente nu pot fi măsurate direct, ci indirect prin intermediul variabilelor observabile. Așadar unei variabile latente i se asociază una sau mai multe variabile observabile cu ajutorul cărora se va studia comportamentul variabilei latente.

În acest proces, datorită în primul rând unei mari doze de subiectivism la alegerea variabilelor observabile atașate variabilelor latente studiate, se pune cu prioritate întrebarea dacă variabilele observabile selectate măsoară într-adevăr caracteristici ale variabilelor latente. Oare un indicator a fost bine ales pentru o anumită variabilă latentă? Ar trebui el modificat sau se impune chiar renunțarea la un astfel de indicator? La acest gen de întrebări vom răspunde în continuare.

Stabilirea veridicității măsurătorilor presupune aplicarea mai multor tehnici:

1. Verificarea dimensiunilor conceptului teoretic ;
2. Folosirea unor variabile de control ;
3. Utilizarea constructivă a unor predicții teoretice;
4. Evidențierea unor similarități și contraste.

În cele ce urmează vom exemplifica tehnicile mai sus menționate și vom pune în evidență avantajele și dezavantajele aplicării lor.

1) Dimensiunea conceptului teoretic.

O procedura teoretică calitativă în procesul de validare a măsurătorilor constă în a stabili dacă variabilele observabile utilizate în modelul causal studiat cuprind, în totalitate, conceptul teoretic abordat sau se limitează numai la anumite aspecte ale sale (nu întotdeauna esențiale).

Acest deziderat este relativ ușor de satisfăcut. Astfel, urmând definiția teoretică a conceptului sociologic studiat, se determină numărul componentelor sale esențiale (dimensiunea modelului causal). Fiecărei componente i se va atribui câte o variabilă latentă.

În final, trebuie să urmărim dacă fiecărei variabile latente astfel selectată i se asociază cel puțin o variabilă măsurabilă. În caz contrar, anumite aspecte (esențiale) ale conceptului teoretic vor fi neglijate, neputând fi puse în evidență de modelul causal.

Este evidentă dependența acestei proceduri de validare a măsurătorilor de o “bună definiție teoretică” a conceptului analizat, definiție acceptată de către analistul modelului.

Menționăm doza mare de subiectivism prezentă în analiza caracteristicilor conceptului cât și importanța relativă a acestor caracteristici.

Oricând vom putea opta pentru o dimensiune mai mare sau mai mică a modelului. În funcție de acest lucru, va trebui să efectuăm noi măsurători sau vom fi nevoiți să renunțăm la altele (nu la întâmplare, ci în urma unui studiu temeinic).

Prin acest procedeu teoretic vom reduce problema validării măsurătorilor la o abordare locală, privind fiecare variabilă latentă în mod separat.

2) Variabile de control.

Utilizarea unei variabile de control C este eficientă în procesul de verificare a veridicității măsurătorilor numai atunci când unei variabile latente U i se asociază o singură variabilă observabilă.

În anumite condiții restrictive, manifestarea variabilei latente U poate fi “aproximată” prin comportamentul unei variabile de control C .

Intenționăm să apreciem dacă variabila observabilă X este sau nu adecvată pentru măsurarea indirectă a comportamentului variabilei latente U .

O validare în acest sens ar putea fi dictată de o corelație puternică dintre variabila observabilă X și variabila de control C . Variabila de control constituie, în anumite condiții, un “inlocuitor” al variabilei latente. În cazul în care variabila X măsoară “adecvat” comportamentul variabilei latente U , atunci valoarea corelației $Cor(X,U)$ este mare (aproape de 1). Cum în anumite condiții particulare, variabila de control C poate constitui

un bun substitut al variabilei \underline{U} , vom obține de asemenea o valoare ridicată a corelației $\text{Cor}(X,C)$ dintre două variabile măsurabile.

Vom pune în evidență câteva greșeli ce pot apare într-un astfel de raționament.

Cazul 1. Vom considera modelul din figura 5.1. având următoarele ecuații de structură:

$$\begin{aligned} X &= d \cdot \underline{U} + E \\ C &= d_0 \cdot \underline{U} + E_0 \end{aligned} \quad (1)$$

unde erorile E_0 și E sunt necorelate între ele și de asemenea necorelate cu variabila latentă \underline{U} . În plus,

$$\text{Med}(E) = \text{Med}(E_0) = 0$$

Condițiile $d_0 = 1$ și $E_0 = 0$ semnifică o identificare “perfectă” a variabilei latente \underline{U} prin variabila de control C .

Din ecuațiile de structură (1), ținând seama de restricțiile impuse erorilor, deducem

$$\text{Cor}(X,C) = d \cdot d_0 \cdot \text{Var}(\underline{U}) / [\text{Var}(X) \cdot \text{Var}(C)]^{1/2} \quad (2)$$

În cazul în care vom standardiza variabilele observabile X și C (adică $\text{Var}(X) = \text{Var}(C) = 1$), atunci valoarea corelației dintre aceste variabile se va calcula după formula :

$$\text{Cor}(X,C) = d \cdot d_0 \cdot \text{Var}(\underline{U}) \quad (3)$$

Presupunând, pentru simplificarea raționamentelor, că și variabila latentă \underline{U} este standardizată ($\text{Var}(\underline{U}) = 1$) deducem că parametrii modelului (1) iau valorile :

$$d = \text{Cor}(X,\underline{U}),$$

$$d_0 = \text{Cor}(C,\underline{U}),$$

și deci formula (3) devine:

$$\text{Cor}(X,C) = \text{Cor}(X,\underline{U}) \cdot \text{Cor}(C,\underline{U}) \quad (4)$$

În cazul în care variabila de control C nu este puternic corelată cu variabila latentă \underline{U} (variabila C “nu simulează bine” comportarea lui \underline{U}) atunci valoarea absolută a corelației $\text{Cor}(C,\underline{U})$ este mică. Din formula

(4) deducem că și valoarea absolută a corelației $Cor(X,C)$ nu poate fi prea mare chiar dacă corelația $Cor(X,U)$ este foarte mare.

Așadar, aplicând criteriul variabilei de control, cum valoarea absolută a corelației $Cor(X,C)$ nu este prea ridicată, suntem conduși la invalidarea posibilității ca variabila X să măsoare “acceptabil” comportamentul variabilei latente U . Acest lucru este însă infirmat de ipoteza unei puternice corelații între X și U .

Prin urmare, aplicarea procedurii variabilelor de control poate conduce la erori, deoarece valoarea corelației dintre o variabilă observabilă X și o variabilă latentă U nu depinde numai de intensitatea corelației dintre variabila X și variabila de control C atașată lui U . Intervin în acest caz și alți factori (formulele (2)-(4)). O valoare scăzută a corelației $Cor(X,C)$ se poate datora și faptului că variabila de control C nu se adaptează suficient comportamentului variabilei latente ($Cor(C,U)$ ia valori mici în modul).

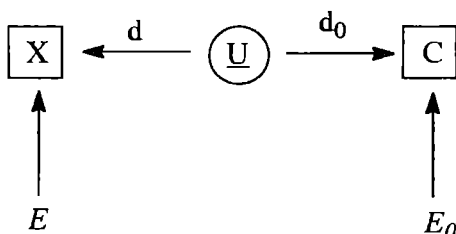


Figura 5.1.

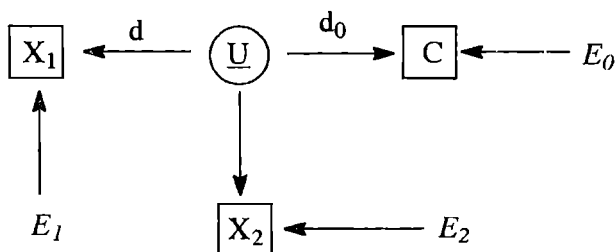


Figura 5.2.

Drept variabilă de control poate fi acceptată orice variabilă măsurabilă care, în anumite condiții particulare, are o manifestare similară cu variabila latentă.

Astfel putem considera drept variabile de control unele caracteristici esențiale ale variabilei latente, caracteristici ce sunt măsurabile cu anumite restricții (de exemplu: pentru unele segmente din întreaga populație, măsurători într-un interval de timp limitat, aprecieri asupra unui fenomen similar care s-a desfășurat deja, o aproximare a fenomenului printr-o singură componentă dominantă, etc.).

În situația în care o variabilă latentă este reprezentată prin mai multe variabile măsurabile asociate ei, atunci validarea veridicității observațiilor nu poate fi efectuată utilizându-se doar o singură variabilă de control.

Intuitiv, o singură variabilă de control nu poate lua în considerare o multitudine de aspecte specifice variabilei latente, respectând totodată și raporturile de dependență dintre aceste caracteristici.

În momentul în care în variabila de control C se acordă o mai mare importanță unei caracteristici măsurabile X_1 a variabilei latente \underline{U} , atunci acest lucru se va resimți în detrimentul unei alte caracteristici X_2 . Astfel va fi imposibil să se obțină valori ridicate pentru ambele corelații $\text{Cor}(X_1, C)$ și $\text{Cor}(X_2, C)$ și deci procedura variabilei de control este neaplicabilă.

Cazul 2. Vom considera variabila latentă \underline{U} drept un "amestec" de parametru a , $0 < a < 1$, inițial neprecizat, dintre două variabile standardizate, necorelate, X_1 și X_2 , adică :

$$\underline{U} = a^{1/2} \cdot X_1 + (1 - a)^{1/2} \cdot X_2 \quad (5)$$

unde $0 < a < 1$.

Indicatorul \underline{U} va fi astfel caracterizat în totalitate numai de variabilele măsurabile X_1 și X_2 . Deci, în formula (5), eroarea de aproximare este nulă. În această situație suntem convinși de necesitatea acceptării ambilor factori X_1 și X_2 pentru studierea comportamentului variabilei \underline{U} .

Considerăm drept variabilă de control variabila C ,

$$C = b^{1/2} \cdot X_1 + (1 - b)^{1/2} \cdot X_2 \quad (6)$$

cu $0 < b < 1$, fixat.

Comportamentul variabilei de control C este identic cu cel al variabilei \underline{U} pentru $b = a$. În plus, variabila C are aceeași structură ca a variabilei \underline{U} , proporția variabilelor măsurabile din “amestec” putând fi însă diferită.

Modelul causal este definit de ecuațiile de structură:

$$\begin{aligned} X_1 &= d_1 \cdot \underline{U} + E_1 \\ X_2 &= d_2 \cdot \underline{U} + E_2 \\ C &= d_0 \cdot \underline{U} + E_0 \end{aligned} \quad (7)$$

unde $\text{Cov}(E_j, \underline{U}) = 0$ pentru $j=0,1,2$ (figura 5.2).

Din ipotezele făcute rezultă standardizarea variabilelor \underline{U} și C . Ținând seama de independența dintre X_1 și X_2 și de forma variabilei de control C obținem :

$$\begin{aligned} \text{Cor}(X_1, C) &= \text{Cor}(X_1, b^{1/2} \cdot X_1 + (1-b)^{1/2} \cdot X_2) = b^{1/2} \\ \text{Cor}(X_2, C) &= \text{Cor}(X_2, b^{1/2} \cdot X_1 + (1-b)^{1/2} \cdot X_2) = (1-b)^{1/2} \end{aligned} \quad (8)$$

În mod evident, corelațiile $\text{Cor}(X_1, C)$, $\text{Cor}(X_2, C)$ definite prin relațiile (8), nu pot lua simultan valori mari în modul atunci când $0 < b < 1$.

Deci variabila de control C nu este în stare să valideze necesitatea utilizării variabilelor măsurabile X_1 și X_2 pentru explicitarea lui \underline{U} .

Putem argumenta că acest fapt s-ar datora “nepotrivirii” dintre variabila latentă \underline{U} și variabila de control C . O astfel de obiecție este falsă deoarece pentru $b = a$ variabila de control C se identifică cu variabila \underline{U} , caz în care vom avea, de regulă, valori mici în modul pentru unul dintre coeficienții de corelație $\text{Cor}(X_1, C)$ sau $\text{Cor}(X_2, C)$.

Cazul 3. Să considerăm variabila \underline{U} definită prin

$$\underline{U} = a_1 \cdot X_1 + a_2 \cdot X_2 + a_3 \cdot X_3 + \dots + a_m \cdot X_m \quad (9)$$

unde X_1, X_2, \dots, X_m sunt variabile aleatoare independente și standardizate ($\text{Var}(X_j) = 1$).

Impunem condiția:

$$a_1^2 + a_2^2 + a_3^2 + \dots + a_{m-1}^2 + a_m^2 = 1 \quad (10)$$

pentru ca și variabila \underline{U} să fie standardizată ($\text{Var}(\underline{U}) = 1$).

Prin urmare, pentru $j=1,2,\dots,n$ avem:

$$\text{Cor}(X_j, \underline{U}) = \text{Cov}(X_j, \underline{U}) = a_j \quad (11)$$

Din formulele (10) și (11) rezultă faptul că există cel puțin un indice j , $1 \leq j \leq m$, astfel încât valoarea absolută a corelației $\text{Cor}(X_j, \underline{U})$ să nu depășească $m^{-1/2}$.

Asfel, pentru $n = 4$ există cel puțin o variabilă X_j , $j=1,2,3,4$, pentru care $\text{Cor}(X_j, \underline{U})$ să nu depășească valoarea 0.5

Devine clară neaplicativitatea procedurii variabilei de control C în cazul unui număr mare m de variabile observabile și independente, X_1, X_2, \dots, X_m .

Oricât de bine am alege variabila de control (de exemplu $C = \underline{U}$) pot rezulta corelații $\text{Cor}(C, X_j)$ a căror valoare absolută este mică, nedepășind $m^{-1/2}$.

Deși validitatea observațiilor X_1, X_2, \dots, X_n nu poate fi pusă la îndoială, totuși procedura propusă infirmă veridicitatea cel puțin a unei măsurători X_j , $1 \leq j \leq m$, aceasta datorită unei valori absolute mici a corelației $\text{Cor}(X_j, C)$.

3) Predicții teoretice și modele ajutătoare

De multe ori se folosesc anumite rezultate teoretice pentru a confirma veridicitatea variabilelor observabile. Această metodă este de regulă aplicată succesiv în mai mulți pași, fiecare etapă validând sau invalidând grupurile de variabile ce caracterizează în parte variabilele latente analizate. În acest proces se pot concepe și folosi modele cauzale ajutătoare.

Cazul 4. Vom încerca să exemplificăm un astfel de procedeu în cazul a patru variabile măsurabile $X_{11}, X_{12}, X_{21}, X_{22}$ asociate variabilelor latente \underline{U}_1 , respectiv \underline{U}_2 (Figura 5.3.).

Modelul cauzal este definit de ecuațiile de structură:

$$X_{ij} = d_{ij} \cdot \underline{U}_i + E_{ij}, \quad 1 \leq i, j \leq 2 \quad (12)$$

cu restricțiile:

$$\text{Cov}(E_{ij}, \underline{U}_k) = 0, \quad \text{Med}(E_{ij}) = 0, \quad \text{Cov}(E_{ij}, E_{st}) = 0 \quad (13)$$

unde $1 \leq i, j, k, s, t \leq 2$.

În aceste condiții avem :

$$\begin{aligned} \text{Cov}(X_{1j}, X_{1k}) &= d_{1j} \cdot d_{2k} \cdot \text{Cov}(\underline{U}_1, \underline{U}_2) \\ d_{ij} &= \text{Cov}(X_{ij}, \underline{U}_i) / \text{Var}(\underline{U}_i) \end{aligned} \quad (14)$$

adică

$$\text{Cov}(X_{1j}, X_{2k}) = \text{Cor}(X_{1j}, \underline{U}_1) \cdot \text{Cor}(X_{2k}, \underline{U}_2) \cdot \text{Cor}(\underline{U}_1, \underline{U}_2) \quad (15)$$

În cazul în care știm că, din punct de vedere teoretic, avem o corelație puternică între variabilele latente \underline{U}_1 și \underline{U}_2 atunci acest efect se va propaga și asupra corelației dintre variabilele măsurabile X_{1j} , X_{2k} (formula (15)).

Așadar vom valida indirect corelația dintre variabilele latente \underline{U}_1 și \underline{U}_2 prin calcularea efectivă, din datele experimentale, a coeficientului de corelație $\text{Cor}(X_{1j}, X_{2k})$.

Procedeul propus își are limitele sale. Într-adevăr $\text{Cor}(\underline{U}_1, \underline{U}_2)$ nu depinde numai de $\text{Cor}(X_{1j}, X_{2k})$ ci și de alți factori (ca de exemplu $\text{Cor}(X_{1j}, \underline{U}_1)$, respectiv $\text{Cor}(X_{2k}, \underline{U}_2)$, formula (15)).

4) Similarități și contraste

O comportare adecvată a unui model cauzal poate fi apreciată în mod indirect, în comparație cu comportamentul altor modele cauzale. Ideea de similaritate și contrast va fi cu succes utilizată în acest context.

Astfel doua modele cauzale “*asemănătoare*” vor avea o comportare relativ similară. O abatere de la această regulă este posibilă, dar întotdeauna va naște o îndoială referitoare la veridicitatea modelului empiric propus. Spulberarea acestei îndoieli va presupune efectuarea unor teste suplimentare asupra modelului în cauză.

Drept “*model similar*” ar putea fi considerată o parte esențială a modelului studiat, submodel ce poate fi soluționat independent de modelul din care derivă.

Există modele cauzale, care, din punct de vedere teoretic, se manifestă în “*opoziție*”. Dacă acest aspect nu se menține și la nivelul modelelor empirice corespunzătoare atunci vom fi nevoiți să invalidăm cel puțin unul dintre cele două modele experimentale ce au fost propuse.

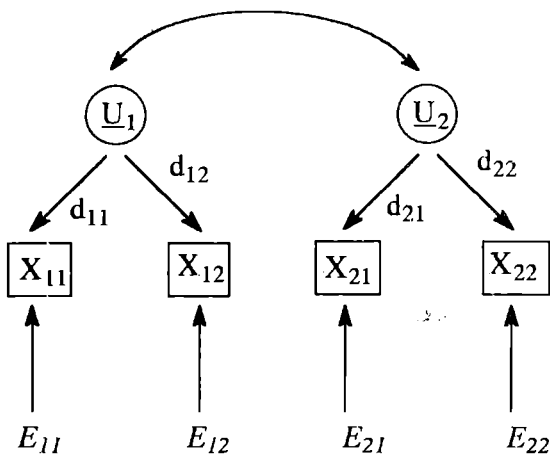


Figura 5.3

Concluzii

În această secțiune s-a discutat modul concret de concepere a unor modele cauzale. S-au evidențiat tehnicile de definire a variabilelor latente ale modelului, precum și modul în care urmează să se efectueze măsurători asupra acestora.

Scopul final al acestei etape este propunerea unui model empiric al cărui comportament să fie asemănător cu acela al modelului teoretic.

Următoarea etapă principală o constituie validarea modelului empiric. Acest lucru se va desfășura pe mai multe planuri.

Vom avea o validare a dimensiunii modelului, fiind astfel siguri că modelul conține acele variabile latente care dau caracteristicile esențiale ale conceptului teoretic abordat. Urmează apoi o validare a variabilelor observabile care caracterizează indirect variabilele latente. Variabilele observabile trebuie să "exprime" preponderent influența specifică a variabilelor latente corespunzătoare și nu acțiunea altor factori. În acest sens pot fi utilizate variabile de control, procedeu ce are însă restricții de aplicare.

Pe lângă validarea variabilelor se va face și o validare indirectă a întregului model empiric propus, prin compararea cu alte modele "similare" sau cu modele cu un comportament "opus".

În acest context vor fi folosiți și indicatorii de siguranță.

III. APLICAȚII ÎN SOCIOLOGIE : SĂRĂCIA

III. 1 DEFINIREA CONCEPTULUI DE SĂRĂCIE

1.1 Definiție empirică

În cercetările asupra sărăciei, o problemă importantă o constituie definirea acestui concept. Există multe abordări legate de definirea și măsurarea sărăciei care însă nu sunt încadrate într-un cadru teoretic fundamental (Townsend 1979). Înainte de 1960 s-au publicat puține lucrări pe această temă.

Începutul a fost făcut în 1901 de Rowntree într-un studiu despre sărăcie în comitatul York, Anglia. Acest studiu definea sărăcia primară ca fiind venitul inferior celui necesar la satisfacerea necesităților minime pentru menținerea eficienței fizice. Aceeași definiție a fost folosită și în raportul Beveridge (Parlamentul Regatului Unit, 1942).

Aceste studii empirice asupra sărăciei se concentrau pe evidențierea numărului de săraci și nu pe măsurarea gradului lor de sărăcie.

Problema esențială în definirea sărăciei este existența unei definiții obiective unice pe baza căreia să se poată stabili exact când o persoana sau o familie se află în stare de sărăcie.

➤ Conceptul de sărăcie a evoluat de la o interpretare absolută și obiectivă către o abordare de natură subiectivă și relativă.

De la abordarea “absolută” de la începutul secolului, în ultimile 2 decenii s-au definit concepte de sărăcie relativă sau subiectivă (Sen, 1979 ; Van Praag 1984).

În definirea sărăciei au intervenit, pe lângă elementele legate de supraviețuire, și standardul corespunzător al locuinței sau participarea la viața socială.

Fiecare concept de sărăcie are avantajele și dezavantajele sale. Acestea vor fi clarificate cu ajutorul a trei întrebări fundamentale:

- Sărăcia este relativă sau absolută ?
- Sărăcia este obiectivă sau subiectivă ?
- Sărăcia este un concept material sau ne-material ?

1.2 Concepte alternative de sărăcie

1.2.1 Sărăcia absolută

Acest concept măsoară sărăcia ca fiind venitul sau consumul inferior unui nivel absolut care reprezintă un minim “absolut”.

Această definiție a fost utilizată pe scară largă, dar prezintă dezavantajul că este dificil de a stabili un minim “obiectiv” valabil pe o perioadă lungă de timp sau în cadrul unor grupuri eterogene ale populației.

În timp sau spațiu este posibil ca bunurile pe care le consumă oamenii să difere, și, în acest caz, și definiția “minimului” se va schimba.

Metoda bugetului care a fost dezvoltată în Marea Britanie de Booth și Rowntree se referă la definirea unui “coș de bunuri”, considerat a fi absolut minimal pentru ca o familie să poată supraviețui. Suma de bani corespunzătoare acestui coș de bunuri este considerată a fi nivelul de sărăcie.

O a doua variantă, dezvoltată de Orshansky (1968), a atașat coșului de bunuri anumite standarde de dietă, care să asigure o existență sănătoasă. Conform acesteia, o familie trăiește în sărăcie dacă cheltuiește peste 30 % din venitul său pentru hrană.

1.2.2 Sărăcia relativă

Măsurarea acestui concept s-a efectuat în mai multe moduri:

- a) S-au considerat ca fiind săraci subiecții cuprinși în decila 1 sau 2 a distribuției veniturilor.
- b) S-a calculat jumătate din mediana veniturilor pentru a desemna nivelul de sărăcie.

Un asemenea standard va crește rapid în perioadele de avânt economic, iar familiile sărace anterior, care beneficiază de o creștere a veniturilor, vor fi considerate tot sărace, deoarece mediana veniturilor va crește corespunzător.

- c) Stabilirea scalelor de echivalență ;

În țările vest europene a fost calculată o linie de sărăcie pe baza noțiunii de venit echivalent și analize comparative de la o societate la alta. Pe baza studiilor efectuate de OECD s-a stabilit o scală de echivalență a venitului care conține indicele 1 pentru o persoană singură, indicele 0,7 pentru fiecare persoană adultă din familie, exclusiv susținătorul, și 0,5 pentru fiecare copil din familie.

Ca orice încercare arbitrară, această scală a fost criticată de experți deoarece conduce la comparații de utilitate interpersonală.

În România, Institutul pentru Cercetarea Calității Vieții a stabilit o scală de echivalență, rezultată din analiza coșului de bunuri și servicii al familiei cu 4 persoane. Astfel :

Pentru calculul minimumului de decență scala cuprinde indicele 1 pentru capul de familie, 0,8 pentru fiecare adult , 0,6 pentru fiecare din primii 2 copii și 0,5 pentru următorii copii.

Pentru calculul minimumului de subzistență scala cuprinde indicele 1 pentru capul de familie, 0,7 pentru fiecare adult și primii 2 copii, iar indicele 0,6 începând cu al 3-lea copil.

O altă metodă de calcul a nivelului de sărăcie se bazează pe sumele minime alocate de stat pentru securitate socială. Familiile al căror venit este inferior acestor standarde oficiale sunt considerate sărace.

Folosirea acestei definiții poate conduce la concluzii paradoxale, în sensul următor : dacă factorul politic determină scăderea nivelului ajutoarelor sociale, atunci numărul săracilor va scade. Explicația este următoarea :

Persoanele care nu beneficiază de ajutor social ci trăiesc dintr-un venit scăzut obținut din muncă se află în situația în care venitul rămâne constant în timp, iar minimul social legal scade : deci aceste persoane vor depăși linia de sărăcie legală și vor trece în categoria non-săracilor, deși situația materială a acestora nu s-a schimbat, în fond.

Pe de altă parte, pentru săraci, cei care trăiesc din ajutoarele sociale, scăderea nivelului legal și deci a unicului lor venit nu are nici un impact asupra stării lor de sărăcie, deoarece ei vor rămâne tot săraci.

În acest context, scăderea venitului va avea un impact negativ asupra percepției venitului și asupra stării subiective de deprivare. De aceea este important ca în cercetarea asupra sărăciei să se ia în considerare și aprecierile subiective asupra sărăciei.

1.2.3 Sărăcia subiectivă

Definițiile subiective ale sărăciei sunt abordări relativ noi în domeniul cercetării sărăciei. Aceste definiții se bazează pe investigații care folosesc aprecierea proprie a familiei asupra venitului minim sau suficient necesar traiului.

Prin aceasta se presupune că subiecții înșiși sunt experții care estimează cel mai bine nevoile familiei lor. Inițiativa românească în abordarea acestui concept (Zamfir, coord. 1995) va fi prezentată în secțiunea III.2.

Începutul în acest sens a fost făcut de articolul “The Poverty Line: Concept and Measurement”, publicat de Goedhart și colaboratorii în 1977. Van Praag de la Universitatea Leyden (Olanda) a dezvoltat acest concept în Van Praag și colab. (1980, 1982).

Ideea de bază derivă din faptul că, una din modalitățile de a afla nivelul minim de venit sau de consum necesar oamenilor pentru a menține ceea ce ei consideră a fi un nivel “decent” de viață sau un nivel minim, este ca oamenii să fie întrebați direct.

De exemplu, în sondajul efectuat de Goedhart și colab. (1977) s-a pus întrebarea :

“Ce venit considerați a fi suficient pentru familie?”

Într-un context asemănător, Danziger și colab. (1984) au analizat următoarea întrebare:

“Trăind în aceleași condiții ca acum și efectuând cheltuielile pe care le considerați necesare, care ar fi cel mai mic venit cu care s-ar putea descurca familia dumneavoastră până la sfârșitul lunii?”

Mai precis, Van Praag a aplicat următoarea întrebare de evaluare a venitului :

În situația dvs. ce venit	f. rău	\$...
considerați a fi :foarte rău?	rău	\$...
, rău ?, insuficient ?,	insuficient	\$...
suficient?	suficient	\$...
, bun ?, foarte bun?	bun	\$...
	foarte bun	\$...

Așa cum este de așteptat, răspunsurile la asemenea întrebări variază sistematic în funcție de venitul și structura familiei respective.

“Liniile de sărăcie subiectivă” au fost calculate găsim punctul în care distribuția venitului actual se intersectează cu venitul “minimal necesar”.

În realitate, sărăcia este o stare determinată social, iar pragurile oficiale scad către ceea ce programele politice și administrative consideră a fi un nivel corespunzător de resurse pentru a sprijini viața într-o comunitate.

Pare mult mai realist faptul ca înșiși membrii comunității respective să fie întrebați direct asupra nivelului de venit minim necesar. Și, deoarece răspunsurile la asemenea întrebări riscă să supraevalueze suma (dar nu exagerat), această abordare poate reprezenta un compromis între măsurile absolute și relative.

Abordarea “sărăciei subiective” conduce la următoarele considerații: întrebările care cer oamenilor să-și exprime o opinie primesc adeseori răspunsuri extrem de diferite, chiar dacă enunțul întrebării se schimbă aparent foarte puțin.

Dacă aplicăm întrebarea folosită de Danziger (1984) în care sunt incluse expresiile :

“trăind în aceleași condiții ca până acum” și
“cheltuielile pe care le considerați necesare”

sondajul va ocoli răspunsuri care să includă referiri la condițiile de viață, și va reflecta nivelul cheltuielilor necesare pe o perioadă scurtă de timp.

În ceea ce privește perioada de timp pentru care se cere aprecierea venitului, se consideră că o săptămână este un termen prea scurt, iar un an constituie un interval în care situația unei familii se poate modifica radical. De aceea, pentru majoritatea studiilor, intervalul de timp pentru care se cere aprecierea subiectivă a venitului este de o lună (se consideră că situația familiei este relativ stabilă pe o astfel de perioadă).

1.2.4 Indicele de deprivare relativă

Această metodă a fost privită ca o metodă de “consens social”, fiind dezvoltată și folosită de Townsend (1970, 1979, 1987).

Acest indice definește deprivarea ca fiind “starea de dezavantajare relativă care poate fi observată și demonstrată în raport cu comunitatea locală, societatea sau națiunea căreia aparține un individ, o familie sau un grup”, Townsend (1979).

În această definiție sunt importante două aspecte :

- a) Sărăcia este o stare de dezavantaj, ceea ce înseamnă că nu numai venitul sau situația materială sunt importante ci și condițiile de trai ne-materiale. Conceptul de “sărăcie” este deci diferit de celelalte concepte (definiții) prezentate anterior, în sensul că sărăcia este considerată a fi mai mult decât o lipsă de venituri - ea este considerată a fi o stare de deprivare multiplă în care sunt incluse aspecte complexe ale existenței.
- b) Sărăcia este un concept relativ. Se definește întotdeauna ca termen de comparație cu ceilalți membri din societate sau comunitatea locală. Caracterul relativ derivă din faptul că acest concept este determinat de situația concretă din societatea actuală. Așadar, conceptul de sărăcie este relativ atât în timp cât și în spațiu.

Politicile sociale europene se concentrază pe definirea sărăciei atât ca “inegalitate a veniturilor” dar și ca “inegalitate a oportunităților” (Muffels, 1992 ; Zamfir coord.,1995).

Townsend a dezvoltat această metodă utilizând o lista de 66 itemi care acoperă multiple forme ale deprivării :

deprivarea dietei, îmbrăcămintei, comoditățile locuinței, combustibil și lumină, condiții de muncă, sănătate, educație , mediu, contacte sociale, deprivare recreațională și socială.

Subiecților li s-a cerut să indice care din aceste elemente nu pot fi realizate cu resursele de care dispun. Suma tuturor deprivărilor a fost utilizată ca indicator al deprivării, și prin aceasta al sărăciei.

Se poate afirma că starea de sărăcie are un caracter “cumulativ” în sensul că cel mai mare risc de a trăi în sărăcie îl au familiile care probează o combinație de caracteristici de deprivare.

Fiecare metodă de definire, și deci de măsurare, a sărăciei a condus la construirea unui prag specific care este cunoscut și sub numele " linia de sărăcie". În studiile comparative asupra sărăciei se disting așadar linii de sărăcie: absolută, relativă, subiectivă, care sunt specifice unui anumit context economic și social.

În Gustafsson și Lindblon (1993) sunt analizate simultan zece linii de sărăcie corespunzătoare celor mai dezvoltate state din Europa, America și Australia. Se remarcă particularitățile de definire și măsurare întâlnite în fiecare caz. Deși, aparent, statele respective par să aibă standarde economice similare, liniile de sărăcie diferă foarte mult, în special acelea între continentul european și cel american.

Studierea fenomenului de sărăcie a condus la studii pe termen lung dintre care cel mai important este LIS - Luxemburg Income Study, care este o bază de date internațională. Aceasta conține standarde de venit și praguri de sărăcie calculate pe baza sondajelor efectuate în 17 țări industrializate. Setul de date specific unei țări conține informații bogate asupra venitului familiei și asupra structurii familiei. Se depun eforturi pentru a se asigura comparabilitatea datelor din țări diferite. Aceasta permite cercetătorilor și politicianilor să compare nivelul și distribuția sărăciei dintre diferite națiuni și să evalueze impactul pe care îl au diversele politici anti-sărăcie în diferite contexte sociale.

În România, cercetările asupra sărăciei au fost lansate de Institutul pentru Cercetarea Calității Vieții (ICCV) din București, începând din 1991 (Mărginean, 1991) studiile repetându-se periodic. Aceste cercetări au ca scop investigarea realității socio-economice a populației și definirea obiectivă a diferitelor praguri de sărăcie (Zamfir, coord. 1995).

Seturile de date sunt prezervate pentru a face posibile analiza secundară, analize comparative în timp, dar și pentru a servi diferitelor interpretări teoretice legate de fenomenul sărăciei.

III. 2 SĂRĂCIA ÎN ROMÂNIA

2.1 Caracteristici ale eșantionului

Măsurarea sărăciei în România a făcut obiectul unor importante cercetări ale Institutului pentru Cercetarea Calității Vieții (ICCV), începând cu 1991 (Mărginean & Socol, 1991) și continuând în 1993, 1994 (Zamfir, coord, 1995 b). Pe baza datelor primare obținute din sondajele efectuate în 1993 și 1994 am abordat definirea conceptului de sărăcie subiectivă și relațiile cauzale care se manifestă direct sau latent în contextul social prezent. Pe parcursul expunerii vom căuta să păstrăm denumirea și definirea unor concepte studiate și în Zamfir, coord. 1995, care se referă la același set de date.

Așa cum s-a menționat în paragraful 1.2.3, conceptul de sărăcie subiectivă implică o măsurare directă, bazată fie pe venit (Muffels, 1990), fie pe o apreciere calitativ - ordinală, conform Praag, 1982.

Cercetarea efectuată la ICCV a utilizat o întrebare bazată pe venit :
“ Pentru a-și asigura strictul necesar, de ce venit lunar ar avea nevoie - în prezent - familia dvs.? “

Răspunsul la această întrebare a fost numit : **Minimul Subiectiv de Viață - MSV.**

La nivel european, cercetările efectuate începând cu 1990, au aplicat cu precădere, măsurători de acest tip. De exemplu, Muffels, 1992 :

“ Ce venit lunar considerați a fi absolut minim pentru familia dvs.?
Cu alte cuvinte, dacă ați avea mai puțin nu v-ați putea descurca. “

S-a constatat că răspunsul la asemenea întrebare este în funcție de componența și venitul real al familiei. Există tendința ca familiile cu venituri totale mari să declare venituri minime subiective și mai mari.

Se impune analizarea minimului subiectiv de viață, pentru a decide dacă această măsură a sărăciei poate constitui un indicator pe baza căruia să fie enunțată o definiție obiectivă.

O altă apreciere subiectivă a venitului în cercetarea ICCV este dată de întrebarea :

“ Pentru a trăi bine, de ce venit lunar ar avea nevoie - în prezent - familia dvs.? “

Răspunsul la această întrebare va fi numit : **Minimul Subiectiv pentru o Viață Bună - MSVB.**

Cele două praguri de sărăcie, pe tipuri de familie, calculate la ICCV prin metoda normativă (Zamfir, coord.1995) sunt :

Nivelul minim de trai decent - MD , care asigură acoperirea la minimum atât a necesităților de alimentație pe termen lung, cât și satisfacerea la minimum a celorlalte necesități materiale și spirituale.

Nivelul minim de subsistență - MS, asigură numai cheltuielile pentru servicii și bunuri de primă necesitate, pe termen scurt fără a lua în calcul îmbrăcămintea și încălțăminte persoanelor, pe considerentul că ele există deja.

O familie este considerată săracă, dacă realizează un venit total inferior pragului de sărăcie aplicat.

Venitul total al familiei - VENTOT, a fost calculat pe baza informațiilor conținute în chestionar, informații care includ sume obținute din salariu, precum și din activități suplimentare, burse, proprietăți, agricultură, etc (Zamfir, coord. 1995).

Este interesant a observa care sunt similitudinile și diferențele dintre nivelele de venit subiective și obiective (analitice), corespunzătoare aceleiași familii. Este util a investiga dacă pragurile “legale” sau “oficiale” de sărăcie se pot baza pe informații subiective.

Pe de o parte, venitul subiectiv este o estimare directă efectuată de persoana vizată, ceea ce ar trebui să-i confere credibilitate. Pe de altă parte, subiecții au tendința de a supraevalua sumele minime necesare ; deci, n-ar putea fi folosite direct, ca praguri de sărăcie obiectivă.

Analizând datele de la ICCV (1994) am studiat indicatorii obiectivi și subiectivi ai sărăciei, în contextul românesc :

Față de **pragurile de sărăcie obiectivă**, situația se prezintă astfel : 71 % din familii au un venit total sub minimul de decență (MD) ; 29 %

din familii au venitul mai mare sau egal cu MD. Venitul a 31 % din familii se situează sub minimul de subzistență (MS), iar 69 % au venitul mai mare sau egal ca MS.

Față de aceste praguri, se dezvoltă o stare cronică de sărăcie pentru 71 % din populație, dintre care, aproape jumătate (31%), reprezintă situații grave, în care se pune problema supraviețuirii pe termen scurt.

Față de **minimele subiective de venit**, situația este următoarea: 69 % din familii au venitul total sub minimul subiectiv de viață - MSV ; 31 % din cazuri au venitul mai mare sau egal cu MSV. 90,5 % din familii au un venit inferior minimului subiectiv pentru o viață bună - MSVB, 1,9 % au venituri egale cu MSVB iar 7,6 % realizează venituri superioare MSVB.

Deci, din cele 31 % familii cu venituri care depășesc dezideratele subiective, aproximativ o treime (9,5 %) declară o situație în care trăiesc bine.

După cum se remarcă, există doi poli de concentrare, două extreme în societatea românească (1994):

- a) Un grup de 30 % care se află la limita supraviețuirii ; aceștia au venituri sub nivelul minim de subzistență - MS.
- b) Un grup de 7,6 % care trăiesc bine și realizează venituri peste nivelul subiectiv pentru o viață bună - MSVB.

Primul grup este format din familii cu mai mulți copii (22,6 % din familii au 2 copii - față de procentul de 15,4 % în tot eșantionul sau 8,7 % din familii au 3 copii - față de 4,4 în eșantion) al căror susținător are un nivel scăzut de educație (72 % din cazuri au cel mult școala profesională) și al cărui statut ocupațional este : 32 % sunt muncitori, 27,9 % pensionari, 10 % tehnicieni și 7 % șomeri.

O caracteristică importantă a acestui grup o constituie mediul de reședință : 63,4 % familii din acest grup trăiesc în orașe. Acest subgrup din mediul urban este format în majoritate din muncitori care au în întreținere mai mulți copii, comparativ chiar cu restul familiilor defavorizate : 33,2 % au 2 copii, 11,1 % au 3 copii, etc.

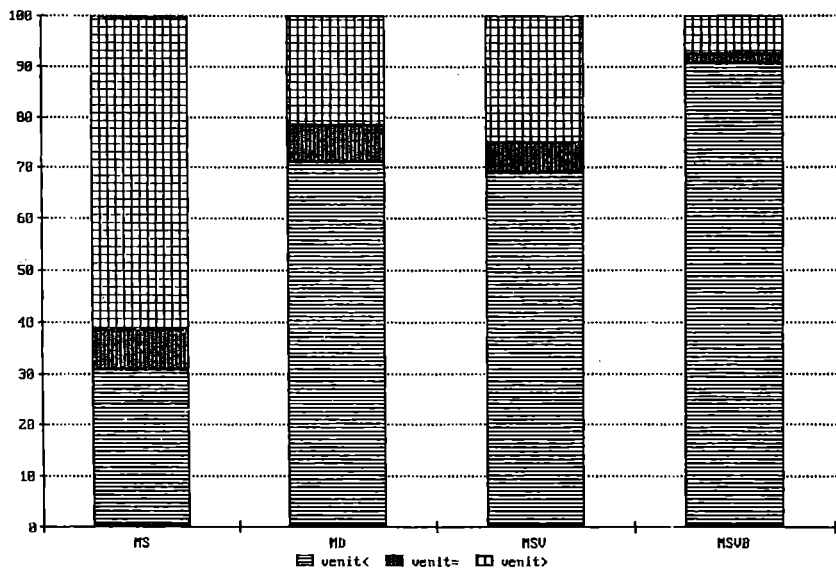
Așadar, sărăcia se caracterizează nu numai prin venituri scăzute, ci și printr-un cumul de factori de deprivare.

Cel de al doilea grup (b), acei 7,6 % “favorizați” ai societății îl constituie : 3 % familii din mediul rural și 4,6 % din mediul urban, având în întreținere mai puțini copii decât media din eșanțion. Peste jumătate din aceste familii nu au copii în întreținere (57,1 %). Proporția familiilor cu un copil rămâne constantă față de eșanțion, iar familiile cu 2 sau mai mulți copii sunt în scădere. Aceste familii realizează venituri din activități suplimentare sau dispun de teren agricol și diferite unelte, iar produsele din gospodărie le asigură consumul curent.

Graficul III.1 reprezintă proporțional situația venitului total față de cei doi indicatori subiectivi (MSV, MSVB) și cei doi obiectivi (MD, MS).

Se remarcă o similitudine de comportament între un indicator obiectiv : minimul de decență - MD, și un indicator subiectiv: minimul subiectiv de viață - MSV.

Situația veniturilor în raport cu pragurile de sărăcie: MS,MD, MSV, MSVB



GRAFIC III.1

2.2 Praguri de sărăcie

În continuare vom stabili două praguri de sărăcie, cărora le vom atașa numele de "linie de sărăcie" datorită analogiei cu reprezentarea grafică a acestor valori.

Se definește **linia de sărăcie subiectivă** punctul în care venitul total al familiei - **VENTOT**, este egal cu minimul subiectiv de viață - **MSV**.

Linia de sărăcie obiectivă este punctul în care venitul total al familiei - **VENTOT**, este egal cu minimul obiectiv de decență - **MD**.

Linia de sărăcie este o noțiune dinamică; ea evoluează în concordanță cu situația socio-economică (a individului și a familiei). De aceea pragul de sărăcie (linia de sărăcie) se recalculează periodic pe baza cercetărilor sociologice, corelate cu situația economică a momentului. În acest mod se poate identifica sărăcia sub cele două aspecte ale sale : numărul de săraci și amploarea sărăciei.

Conform graficului III.1, linia de sărăcie subiectivă departajează populația în proporții similare cu linia de sărăcie obiectivă. Vom analiza în continuare, dacă cele două praguri de sărăcie coincid. În mod natural apare întrebarea :

Există o diferență semnificativă între minimul subiectiv - **MSV** și minimul de decență - **MD** ?

Am constatat că testul t de egalitatea a mediilor, confirmă cu o eroare de 5 % faptul că nu există o diferență semnificativă între **MD** și **MSV**.

Deci, veniturile pe care le reclamă o familie, pentru a-și asigura strictul necesar în condițiile actuale, se situează în apropierea nivelului obiectiv (analitic), care la rândul său este calculat ținând cont de structura familiei.

Analiza de egalitate a mediilor pentru cei doi indicatori va fi reluată pe variabile de mediu și pe grupe socio - demografice.

Tabelul 1. Mediile indicatorilor MSV și MD calculate pe caracteristici de rezidență și ocupaționale (1994)

	EȘANTION		URBAN		RURAL	
	activi	pens	activi	pens	activi	pens
MSV	357202	227210	385342	267102	307646	184708
MD	347763	221347	371248	246289	306405	194774

Am introdus în analiză și factorul de mediu : urban - rural, fapt ce a condus la următoarele concluzii :

În mediul rural, testul t pentru egalitatea mediilor confirmă cu o eroare de 5 %, că nu există o diferență semnificativă între media minimumului de decență și media minimumului subiectiv. În mediul rural minimumul subiectiv al unei familii - MSV are o medie apropiată (ușor inferioară) de media minimumului de decență - MD, calculat obiectiv.

În mediul urban, testul t pentru egalitatea mediilor confirmă următoarea afirmație : există o diferență semnificativă între minimumul de decență și minimumul subiectiv de viață.

Ipoteza alternativă, acceptată cu o eroare de 5 % este :

“ În mediul urban, media minimumului subiectiv (MSV) este mai mare decât media minimumului de decență (MD).

Deci, familiile din mediul urban declară un minim subiectiv superior minimumului de decență, calculat pe baza coșului de produse și servicii. Aceasta se datorează atât stilului de viață din mediul urban, cât și lipsei de resurse suplimentare. Și în Zamfir, 1995, se pune în evidență această tendință de supraevaluare, explicată prin faptul că la sat o parte din consum este acoperit din producția proprie, veniturile bănești propriu-zise fiind astfel necesare într-o proporție ceva mai redusă.

O treime din familiile cuprinse în eșantion sunt formate din pensionari (28,2 % în mediul urban și 39,4 % în mediul rural).

În consecință, în compararea minimumului subiectiv cu minimumul obiectiv de decență am introdus un factor “ economico - demografic “,

conform căruia populația va fi departajată în două grupe : pensionari și persoane active economic.

Analiza statistică pe care am efectuat-o răspunde la următoarea întrebare :

“ Există o diferență semnificativă între minimumul de decență - MD și minimumul subiectiv de viață - MSV, în cadrul grupului activ economic ? “

Testul t de egalitate a mediilor , aplicat și pe factori de mediu, confirmă, cu o eroare de 5 % :

Nu există o diferență semnificativă între minimumul de decență și minimumul subiectiv de viață al familiilor întreținute de o persoană activă economic, indiferent de mediul de reședință (urban, rural, sau total)

Întrebarea anterioară se va aplica și pensionarilor. Pentru aceștia, la nivelul eșantionului, am constatat că nu există diferențe semnificative între standardul subiectiv și cel obiectiv. Diferențele apar pe subdomeniile de reședință, conform următoarelor precizări :

În mediul rural, pentru familiile de pensionari, nu există diferențe semnificative între minimumul de decență și minimumul subiectiv de viață.

În mediul urban, însă, pentru familiile de pensionari, există diferențe semnificative și acestea se manifestă astfel :

“media minimumului subiectiv de viață este mai mare decât media minimumului de decență “.

Dacă în mediul rural populația nu declară necesități subiective materiale foarte mari (comparativ cu standardul obiectiv precalculat), în schimb în mediul urban, categoria pensionarilor se comportă diferit. Aceștia declară un minim subiectiv mai mare decât minimumul de decență. Deci, pentru această categorie demografică un venit egal cu minimumul de decență este insuficient, deoarece aceste persoane nu au venituri suplimentare sau autoconsum.

2.3 Concluzii și precizări

În Mărginean & Socol, 1991 se remarcă diferențieri în configurația profilelor diagnozei calității vieții, pe criteriul profesional și de rezidență.

În general, aprecierea directă a oamenilor asupra venitului minim necesar familiei lor este realistă și nu se remarcă estimări exagerate (Zamfir coord, 1995).

Testele statistice aplicate conduc la concluzia că metodologia folosită de ICCV la calcularea minimumului de decență este adecvată atât pe ansamblu, cât și pentru grupa persoanelor active, indiferent de rezidență.

Minimumul de decență calculat de ICCV poate constitui un prag de sărăcie obiectiv pentru majoritatea populației. Excepție fac pensionarii din mediul urban, care nu se încadrează în afirmația anterioară și pentru care metodologia de calcul a minimumului de decență trebuie adaptată. Rezultatele prezentate anterior confirmă aceste aspecte.

III.3 DINAMICA LINIEI DE SĂRĂCIE SUBIECTIVĂ

3.1. Linia de sărăcie subiectivă (LSS)

Așa cum s-a precizat în secțiunea anterioară, pragul subiectiv de sărăcie (linia de sărăcie) constituie unul din indicatorii alternativi de constatare a sărăciei.

În continuare, vom studia linia de sărăcie definită, de **Minimul Subiectiv de Viață - MSV** ; acest indicator este bazat pe răspunsul la întrebarea :

“ Pentru a-și asigura strictul necesar, de ce venit lunar ar avea nevoie - în prezent - familia dvs.? “

MSV depinde de venitul actual al familiei precum și de alți factori, dintre care remarcăm compoziția familiei.

Așadar MSV este dat de funcția :

$$MSV = f(V, X) \quad (III.1)$$

unde V este un indicator al veniturii actual iar X este un vector care conține alți factori.

Linia de sărăcie subiectivă va fi acel nivel V^* pentru care :

$$V^* = f(V^*, X) \quad (III.2)$$

Cu alte cuvinte, **Linia de Sărăcie Subiectivă** trece prin punctul în care :

minimul subiectiv de viață este egal cu venitul total al familiei:
MSV = VENTOT.

Dacă venitul total al familiei se situează sub linia de sărăcie, atunci vom considera că familia respectivă este săracă.

Dacă venitul total depășește minimul subiectiv, putem afirma că familia nu se află în “ starea de sărăcie “.

Funcția propusă pentru minimul subiectiv - MSV depinde de un indicator al venitului total al familiei, de compoziția familiei și de categoria demografică în care se încadrează susținătorul familiei.

Astfel, pentru susținătorul de familie am definit trei categorii demografice :

- Tânăr sub 30 ani
- Adult, apt de muncă, peste 30 ani
- Pensionar

În grupa pensionarilor intră persoanele care și-au declarat acest statut - 32,7 % din eșantion.

Tinerii sub 30 ani , care locuiesc singuri și, eventual, au întemeiat o familie , constituie o categorie demografică restrânsă : 6,8 % din eșantion. Deși, în România, în majoritatea cazurilor, tinerii locuiesc cu părinții până când se căsătoresc, sunt prezente situațiile în care tinerii studiază și muncesc în alte localități decât părinții lor.

Adulții peste 30 ani, apti de muncă, constituie majoritatea cu 60,5 %.

Compoziția familiei este pusă în evidență de factorii următori : numărul de adulți (exclusiv capul familiei) și numărul de copii (sub 16 ani) în întreținere.

Venitul total al familiei poate fi exprimat direct sau prin intermediul unui indicator de venit. În acest studiu, am preferat folosirea următorului indicator :

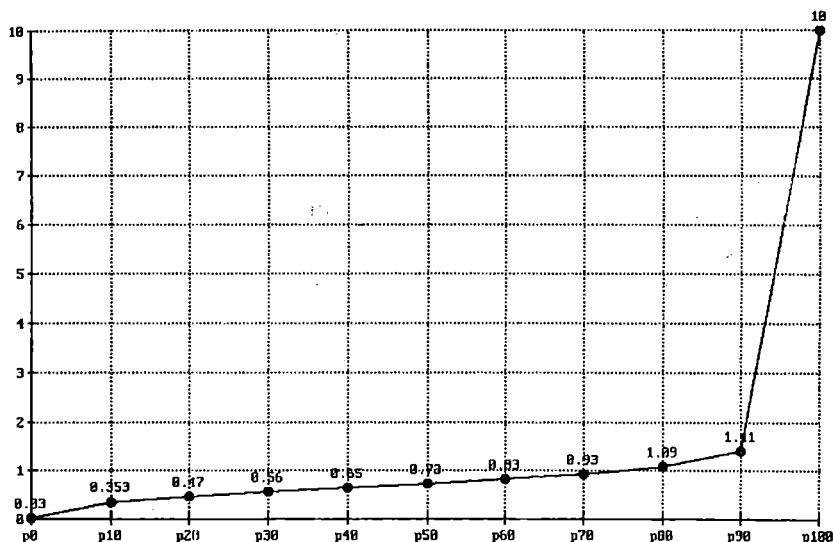
raportul dintre venitul total și minimul de decență

$$\text{RAPORT} = \text{VENTOT} / \text{MD} \quad (\text{III.3})$$

Evident, un indicator de venit poate fi calculat în diferite moduri. O motivație privind alegerea indicatorului de venit este bazată pe posibilitatea utilizării variabilei RAPORT în a studia influența venitului actual în contextul în care se pune problema dacă acest venit poate asigura un trai decent. Acest indicator variază de la 0,03 la 10,2.

Graficul din figura III.2 redă distribuția valorilor indicatorului RAPORT în cadrul decilelor.

Evoluția variabilei RAPORT în cadrul decilelor



GRAFIC III.2

Regăsim, sub o altă formă, un rezultat discutat anterior și anume că, în decila a 7- a se află veniturile mai mici decât nivelul de decență (peste 70 % din populație). Utima decilă prezintă o creștere bruscă, de la valoarea 1,41 la 10. Această imagine corespunde acelor aproximativ 10 % indivizi din populație, care depășesc cu 150 % minimul de decență.

Cu variabilele PENS, TANAR, ADULT, NRCOPII, RAPORT menționate anterior propunem un model de regresie multiplă pentru explicarea minimului subiectiv de viață. Acest model va include atât variabile continue cât și variabile dummy (variabile discrete, cu valori 0 sau 1) :

$$Y = b_0 + b_1 * PENS + b_2 * TÂNĂR + b_3 * ADULT + b_4 * NRCOPII + b_5 * RAPORT \quad (III.4)$$

Explicarea notațiilor din formula (III.4) :

Y este **variabila dependentă** (endogenă). Ea va fi asociată minimului subiectiv.

Variabilele independente (exogene) din regresie sunt :

a) Variabile demografice

Acestea au fost definite ca variabile 'dummy' și iau valoarea 1 sau 0 în funcție de prezența sau absența caracteristicii pe care o reprezintă. În analiza statistică, variabilele dummy se utilizează pentru a defini categorii ale unei variabile nominale.

Categoria de referință este categoria adulților peste 30 ani. Ea este prezentă implicit, atunci când celelalte două categorii sunt absente.

Categoriile definite explicit sunt :

PENS = 1, dacă susținătorul familiei este pensionar.

În caz contrar, **PENS = 0**.

TÂNĂR = 1, dacă susținătorul familiei este un tânăr în vârstă sub 30 ani;

În caz contrar **TÂNĂR = 0**.

b) Variabile referitoare la structura familiei :

ADULT reprezintă numărul persoanelor adulte din familie exclusiv capul familiei ;

NRCOPII reprezintă numărul de copii în întreținere ;

c) Indicatorul de venit

RAPORT este indicatorul de venit calculat după formula :

$$\mathbf{RAPORT = VENTOT / MD.}$$

Cu aceste notații, odată determinați coeficienții ecuației de regresie III.4, se pot face predicții pentru diverse situații particulare :

$Y = b_0$, reprezintă minimul subiectiv (MSV) pentru un adult, apt de muncă, singur, fără copii, fără venituri.

$Y = b_0 + b_1 + b_3$, reprezintă MSV pentru o familie de 2 persoane, în care susținătorul este pensionar (s-a ignorat venitul familiei).

$Y = b_0 + b_2 + b_5$, reprezintă MSV pentru un tânăr sub 30 de ani, care are un venit apropiat de nivelul obiectiv de decență.

$Y = b_0 + b_3 + b_4 + b_5$, reprezintă minimumul subiectiv pentru o familie formată din doi adulți și un copil, pentru care venitul actual VENTOT este identic cu minimumul de decență MD.

Precizări și concluzii

Ecuția de regresie III.4 poate fi utilizată în predicția minimumului subiectiv de viață și pentru alte situații.

În continuare, printr-o analiză statistică a datelor se va determina dacă modelul empiric propus în acest paragraf (formula III.4) este viabil.

Se va studia dacă toate variabilele din model contribuie semnificativ la explicarea minimumului subiectiv de viață, și în caz afirmativ, se va discuta ponderea fiecăreia.

În cazul în care modelul III.4 nu este validat de analiza statistică a datelor atunci este necesar a se propune un nou model.

3.2. Estimarea coeficienților ecuației de regresie

Analiza statistică a datelor s-a efectuat pe un număr de 2318 cazuri semnificative din sondajul anului 1994 (volumul total al eșantionului este de 2504, dar au fost reținute numai cazurile cu răspunsuri complete).

Pentru estimarea coeficienților de regresie parțiali și standardizați s-a folosit metoda stepwise, care garantează obținerea unui rezultat consistent din punct de vedere statistic, deoarece introduce în ecuație numai acele variabile ce sunt semnificative statistic.

Ecuția estimată conform modelului III.4 conducea la reziduuri mari a căror repartiție avea o formă mult mai ascuțită decât repartiția normală. Un test statistic de concordanță a confirmat “nenormalitatea” reziduurilor ecuației de regresie III.4. În această situație am optat pentru logaritmul variabilei dependente - MSV (trecerea la o scală logaritmică pentru indicatorul “minimum subiectiv de viață”).

Așadar, în noul model pe care îl propunem va fi studiat comportamentul variabilei transformate :

$$Y = \text{Ln} (\text{MSV}) \quad (\text{III.5})$$

Cu alte cuvinte, modelul teoretic propus va trebui să țină cont de particularitățile datelor observate. În cazul nostru se va determina dinamica “logaritmului” minimumului subiectiv.

Menționăm că prin această transformare de logaritmare a variabilei MSV ne apropiem de modelul utilizat de Muffels et al. (1990) pentru contextul vest european. Aceștia au propus o funcție de regresie în vederea studierii logaritmului indicatorului obținut din aplicarea MIQ - minimum income question (întrebarea despre venitul minim).

Pentru cercetarea noastră, parametrii estimați ai modelului III.4 cu precizarea III.5 sunt dați de :

$$Y = 12,095 - 0,365 * PENS + 0,209 * ADULT + 0,108 * \\ * NRCOPII + 0,122 * RAPORT \quad (III.6)$$

Aspecte privind modelul III.6

Precizăm că, în această formă, ecuația de regresie III.6 este “consistentă”, iar coeficientul multiplu de corelație este $R = 0,560$.

Așadar, modelul teoretic se reformulează astfel :

“Minimul subiectiv de viață este o funcție care depinde de factorul de venit, structura familiei și categoria socio-demografică a capului de familie.”

Se observă că variabila TÂNĂR nu a intrat în regresie, fapt explicat și de coeficientul de corelație slab cu variabila dependentă ($r = -0,0002$).

Această absență indică faptul că, la nivelul eșantionului categoria tinerilor sub 30 ani, care sunt și susținători de familie, nu aduce o influență suplimentară asupra evaluării minimului subiectiv de viață. Așadar, această categorie demografică (a tinerilor sub 30 de ani) poate fi inclusă în grupul persoanelor active, apte de muncă (categoria demografică de referință).

Variabilele ce caracterizează venitul și structura familiei au o influență pozitivă asupra minimului subiectiv de viață.

Deci evaluarea subiectivă a venitului minim va fi cu atât mai mare cu cât indicatorul de venit este mai mare, sau familia are mai mulți copii sau mai mulți adulți în componența sa.

Un caz aparte îl constituie pensionarii - întreținători de familie : aceștia au tendința de a diminua evaluarea minimului subiectiv.

Tabelul următor conține evoluția minimumului subiectiv de viață MSV pentru diferite categorii de familii, în ipoteza că acestea realizează un venit total egal cu minimumul de decență.

TABELUL 2. ESTIMAREA MSV PE CATEGORII DE FAMILII CARE AU VENITUL TOTAL APROXIMATIV EGAL CU MINIMUL DE DECENȚĂ (anul 1994)

Alți membri ai familiei	Capul de familie	
	Adult activ	Pensionar
0	202197	140365
1 Adult	249197	172992
1 Copil	225258	156373
1 Adult + 1 Copil	277618	192721
1 Adult + 2 Copii	309279	214701

Concluzii

Deoarece s-a demonstrat anterior că media MSV este egală cu media MD, se poate considera că în figura III.3 graficul trasat conform tabelului 2, exprimă estimarea liniei de saracie subiectivă pentru diferitele categorii de familii (pensionari și adulți activi).

Studiind Tabelul 2 sau Figura III.3 se constată ca MSV minim corespunde unui pensionar singur, cu venituri sub minimumul de decență. MSV maxim va corespunde unei familii numeroase al cărei susținător este un adult activ, sau unei familii mai puțin numeroase, dar care realizează venituri mari raportate la minimumul de decență.

În ceea ce privește contribuția indicatorului de venit (variabila RAPORT), se poate afirma că acesta determină o majorare a MSV, care va fi cu atât mai consistentă cu cât venitul real al familiei este mai mare comparativ cu minimumul de decență.

Este necesar să analizăm și cele două subdomenii ale mediului de reședință : urban și rural.

Linia de sărăcie subiectivă estimată pe total eşantion

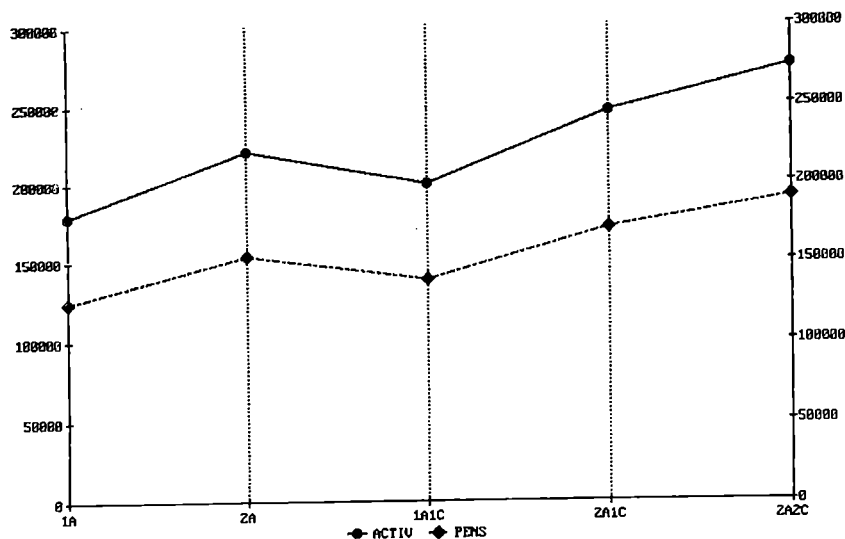


Figura III.3

3.3 Evoluția minimumului subiectiv de viață în mediul urban

Subdomeniul urban conține 60 % din eşantion. Ecuația de regresie estimată este următoarea :

$$Y = 12,225 - 0,291 * PENS + 0,171 * ADULT + 0,110 * NRCOPII + 0,195 * RAPORT - 0,109 * TÂNĂR \quad (III.7)$$

unde $Y = \text{Ln}(\text{MSV})$.

Se observă prezența semnificativă a variabilei TÂNĂR în ecuația de regresie (coeficientul de corelație cu Y este $r = -0,0423$).

Tabelul 3 conține evoluția minimumului subiectiv pentru diferite categorii de familii.

TABELUL 3. ESTIMAREA MSV ÎN MEDIUL URBAN, PENTRU DIFERITE TIPURI DE FAMILII, CU VENITUL EGAL CU MINIMUL DE DECENTĂ

Alți membri ai familiei	Capul de familie		
	Tânăr	Activ	Pensionar
0	222126	247707	185165
1 Adult	263550	293902	219696
1 Copil	247954	276509	206695
1 Adult + 1 Copil	294196	327748	245242
1 Adult + 2 Copii	328404	365858	273758

Discutarea modelului III.7

Se remarcă minimul subiectiv al unei familii întreținute de un pensionar, valoare care este inferioară în comparație cu celelalte tipuri de familii.

Categoria tinerilor declară minime subiective inferioare, dar destul de apropiate de cele ale familiilor întreținute de un adult activ cu vârsta peste 30 ani.

Ca și în cazul modelului general III.6, parametrii familiei și ai venitului contribuie la majorarea estimării minimului subiectiv.

Figura III.4 conține cele trei linii de sărăcie, corespunzătoare celor trei coloane din tabelul anterior (tânăr, adult activ, pensionar). Punctele marcate pe axa orizontală corespund componenței familiei : 1A înseamnă o persoană adultă (tânăr, activ sau pensionar) iar 1C - un copil.

3.4. Evoluția minimului subiectiv de viață în mediul rural

În mediul rural sunt 40 % din cazuri. Ecuația de regresie estimată în acest domeniu este :

$$Y = 11,76 - 0,328 * PENS + 0,257 * ADULT * 0,121 * NRCOPII + 0,119 * RAPORT + 0,192 * TÂNĂR \quad (III.8)$$

unde $Y = \text{Ln}(\text{MSV})$.

LINIA DE SĂRĂCIE SUBIECTIVĂ ESTIMATĂ PENTRU MEDIUL URBAN

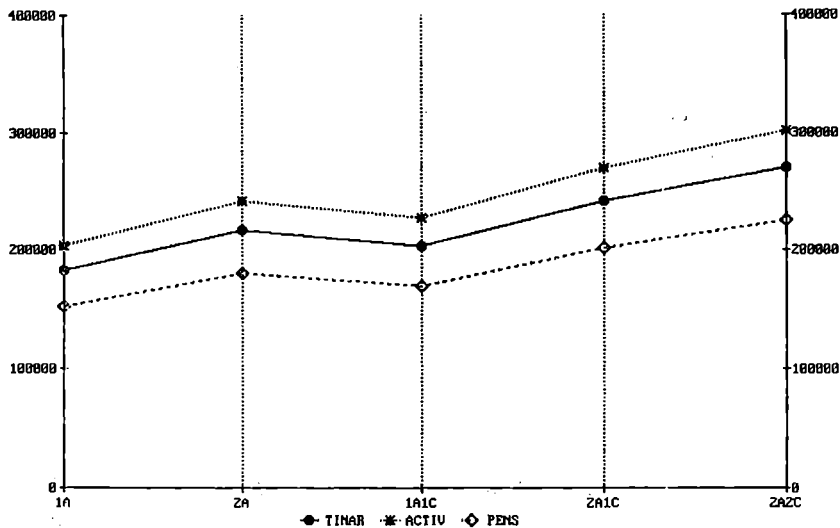


FIG III.4

LINIA DE SĂRĂCIE SUBIECTIVĂ ESTIMATĂ PENTRU MEDIUL RURAL

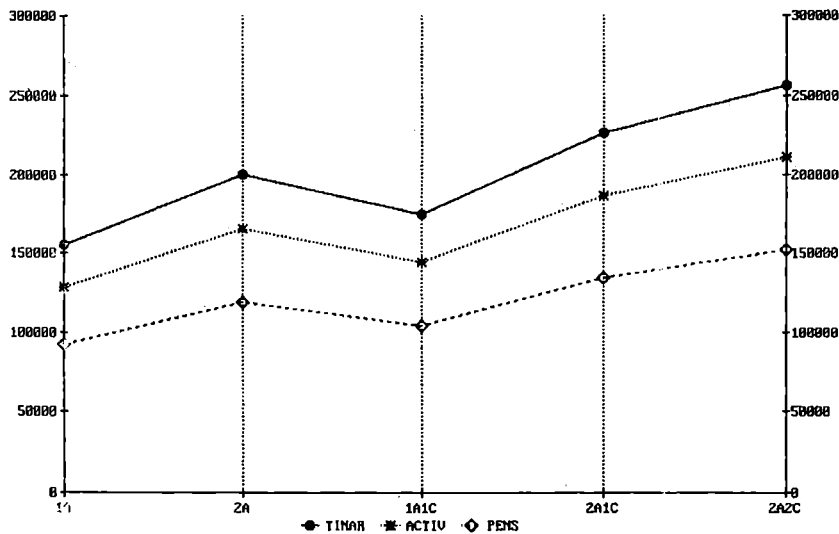


FIG III.5

Ecuția de regresie este consistentă, iar coeficientul de corelație multiplă este $R = 0,595$.

Tabelul 4 conține valori ale venitului subiectiv minim, pentru diferite tipuri de familii, calculate după predicția dată de ecuația III.8.

TABELUL 4. MSV PENTRU MEDIUL RURAL, PENTRU DIFERITE TIPURI DE FAMILII, CU VENITURI EGALE CU MINIMUL DE DECENTĂ

Alți membri ai familiei	Capul de familie		
	Tânăr	Activ	Pensionar
0	174731	144206	103881
1 Adult	225934	186465	134323
1 Copil	197205	162755	117243
1 Adult + 1 Copil	254995	210449	151600
1 Adult + 2 Copii	318062	237518	171099

Precizări privind ecuația de predicție III.8

Și în acest caz pensionarii estimează minime subiective inferioare față de celelalte categorii demografice.

Ceilalți parametri care intervin în regresie contribuie pozitiv la estimarea minimului subiectiv de viață.

Se remarcă o situație deosebită în categoria susținătorilor de familie tineri sub 30 ani. Aceștia estimează veniturile subiective cele mai ridicate, astfel : cu 68 % mai mari decât pensionarii și cu 36 % mai mult decât adulții peste 30 ani.

În figurile III.6 și III.7 sunt trasate liniile de sărăcie corespunzătoare tabelului 4, pentru familii ce au o compoziție diferită.

3.5. Concluzii privind dinamica liniei de sărăcie subiectivă

Din tabelele 2,3 și 4 se remarcă faptul că minimul subiectiv de viață - MSV este mai mare în mediul urban decât în mediul rural, pentru toate tipurile de familii și indicatori de venit.

Este relevantă analizarea raportului dintre veniturile subiective din mediul urban și rural, valori prezise de ecuațiile de regresie III.7, III.8, cu precizarea III.5, aceasta pentru fiecare categorie a capului de familie, anume:

- Pentru tinerii sub 30 ani valorile subiective sunt foarte apropiate. Raportul lor este 1,27 ;
- Pentru adulții peste 30 ani, raportul este 1,71 ;
- Pentru pensionari, raportul este 1,78.

Se constată că MSV estimat de tinerii din mediul urban este apropiat de MSV-ul tinerilor din mediul rural, ceea ce poate indica stiluri de viață și aspirații similare.

Pentru celelalte categorii demografice situația este diferită. Adulții peste 30 ani și pensionarii din mediul urban vor indica minime subiective mult mai mari decât cei din mediul rural (valori cu peste 50 %).

Se poate avansa următoarea ipoteză :

“Minimul subiectiv de viață din mediul urban se apropie de minimul subiectiv din mediul rural pe măsură ce vârsta susținătorilor de familie scade “.

Figurile III.6 și III.7 exemplifică și susțin această afirmație.

Pentru a urmări contribuția variabilelor exogene la evoluția MSV, se consultă tabelul 5 cu coeficienții de regresie standardizați - beta din modelele de regresie III.4, III.6, III.7.

Tabelul 5. Valorile coeficienților beta

variabila	tot eșantionul	urban	rural
ADULT	0,340	0,313	0,413
PENS	-0,272	-0,249	-0,232
NRCOPII	0,171	0,195	0,189
RAPORT	0,142	0,224	0,152
TÂNĂR	-	-0,053	0,070

LINII DE SĂRĂCIE SUBIECTIVĂ ESTIMATĂ
T - tineri; Ac - activi; P - pensionari

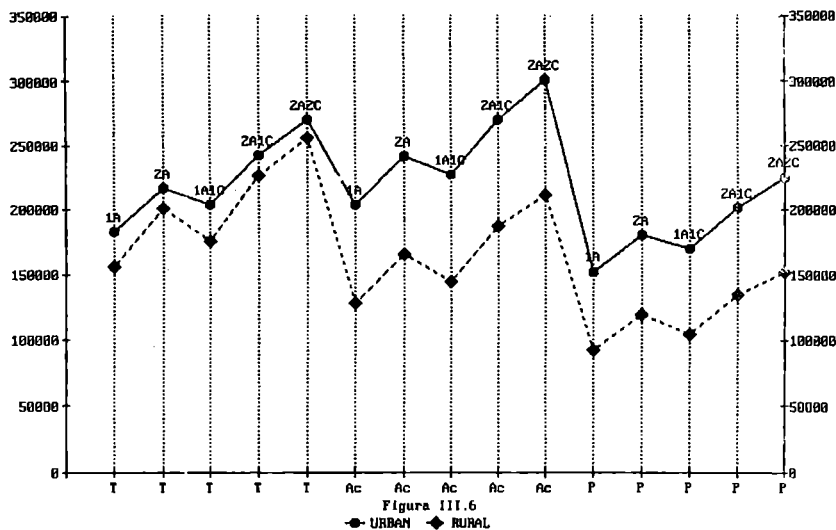


FIG III.6

VSM - DIFERENȚE PE TIPURI DE FAMILII
URBAN - RURAL

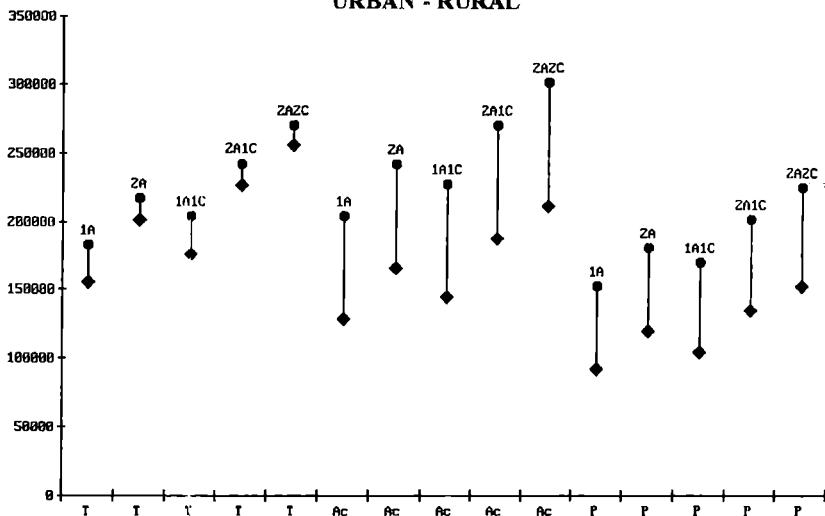


FIG III.7

Se remarcă faptul că pensionarii își reduc minimul subiectiv de viață comparativ cu celelalte categorii (coeficienții beta sunt negativi și sunt semnificativi din punct de vedere statistic).

Compoziția familiei: numărul adulților (ADULT) și numărul copiilor (NRCOPII) contribuie la majorarea estimării minimului subiectiv de viață.

Factorul de venit are o influență pozitivă, care devine maximă în mediul urban : coeficientul beta = 0,224 în mediul urban, față de beta = 0,152 în mediul rural.

Structura familiei și indicatorul de venit contribuie cu prioritate la declararea minimului subiectiv , deși nu se poate nega existența și a altor influențe (coeficientul de corelație multiplă pentru modelul II.6 este 0,560 fapt ce sugerează posibilitatea existenței și a altor factori de influență pentru MSV).

Subliniem faptul că toate aceste aspecte au fost puse în evidență prin utilizarea unor proceduri statistice clasice : teste statistice, analizarea unor tabele sau grafice, utilizarea caracteristicii predictive a modelului de regresie sau folosirea valorilor (absolute) ale coeficienților beta pentru ierarhizarea variabilelor independente ce explică comportamentul unei variabile dependente (de exemplu MSV).

În continuare intenționăm să aducem noi precizări și nuanțe rezultatelor deja enumerate, prin utilizarea unui aparat statistic mult mai puternic, bazat pe analiză cauzală.

Așa cum am menționat în capitolul II, analiza cauzală cu variabile latente generalizează modelul de regresie și modelul de analiză factorială. În plus, o asemenea analiză este adecvată domeniului socio-uman permițând evaluarea raportului dintre două variabile nemăsurabile direct (variabile latente) prin utilizarea efectului acestora asupra altor variabile ce pot fi observate (măsurate). Menționăm că analiza cauzală permite ca variabilele observabile să fie de tipuri diferite : continue, ordinale, nominale.

Pentru studierea conceptului de “ sărăcie subiectivă “ propunem două tipuri de modele cauzale : modele cauzale cu două, respectiv trei variabile latente. Analizarea acestor tipuri de modele va fi efectuată în paragrafele următoare.

III. 4. MODELE CAUZALE CU VARIABILE LATENTE IN STUDIUL SĂRĂCIEI SUBIECTIVE

4.1 Introducere

Starea de sărăcie subiectivă poate fi pusă în evidență prin metode subiective de estimare a standardului de viață. În Zamfir, coord. 1995, sunt menționate următoarele măsurători :

- Autoetichetarea globala - “În momentul actual vă considerați sărac? da sau nu “
- Standardul subiectiv de viață (SSV) - “ Cum apreciați veniturile actuale ale familiei dumneavoastră ?”
- Satisfacția cu veniturile (SV) - “Cât de mulțumit sunteți de veniturile actuale ale gospodăriei ?”
- Indicatorul subiectiv de sărăcie (ISS) a fost construit pe baza celor trei estimări subiective anterioare ale sărăciei, dar nu poate fi considerat “factor latent”, în accepțiunea dată în această lucrare.

În continuare, vom prezenta o abordare metodologică nouă asupra conceptului de “sărăcie subiectivă”. Analiza va folosi metoda LISREL, menționată în cap. II, și va expune atât definirea variabilelor latente cât și relațiile cauzale dintre acestea. Pentru unificarea conceptelor vom păstra denumirile atașate variabilelor măsurabile, în studiul ICCV (Zamfir, coord. 1995). Propunem două tipuri de modele cauzale ce vor fi discutate pe larg în cele ce urmează.

4.2. Tipul I. Model cauzal cu două variabile latente

4.2.1 *Definirea și selectarea unor modele cauzale*

În contextul cercetării despre sărăcia subiectivă , pot fi puse în evidență următoarele concepte teoretice :

“ **necesități materiale** “ și
“ **estimarea subiectivă a standardului de viață** “ .

“ Necesitățile materiale “ ale unei persoane sau familii se concretizează prin exprimarea dorinței de a poseda anumite valori : bani, casă, autoturism, diferite obiecte de folosință îndelungată. O asemenea listă conține, parțial, valori materiale de primă necesitate, care lipsesc la un moment dat.

Dar factorul uman, subiectiv, va determina necesități și aspirații specifice, de la caz la caz. Deci lista de valori materiale va avea un conținut determinat atât de criterii obiective cât și subiective.

“ Estimarea subiectivă a standardului de viață “ este un concept care determină modul în care individul își apreciază condițiile de trai, sub aspect cantitativ și calitativ.

Acest concept îl conduce pe individ să se autodefinească, ca fiind sărac sau nu, să-și exprime satisfacția sau dezaprobarea față de nivelul de trai sub multiple aspecte. Modul în care o persoană sau familie își evaluează standardul de viață depinde de particularitățile individuale și sociale ale momentului.

Într-o țară în tranziție, cum este România, este util de analizat cele două concepte teoretice care, cu siguranță, poartă amprenta epocii pe care o traversăm. Se poate enunța următorul model teoretic :

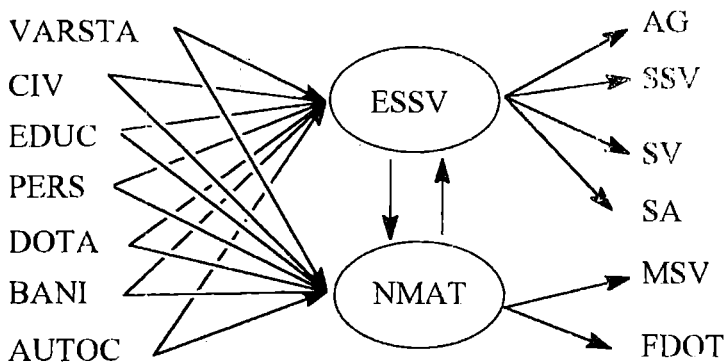
“ Necesitățile materiale și estimarea subiectivă a standardului de viață sunt influențate de indicatori sociali și caracteristici de familie ; cele două concepte teoretice se află într-o relație cauzală reciprocă “.

Modelul teoretic pe care îl propunem se reprezintă grafic prin diagrama din figura III. 8.

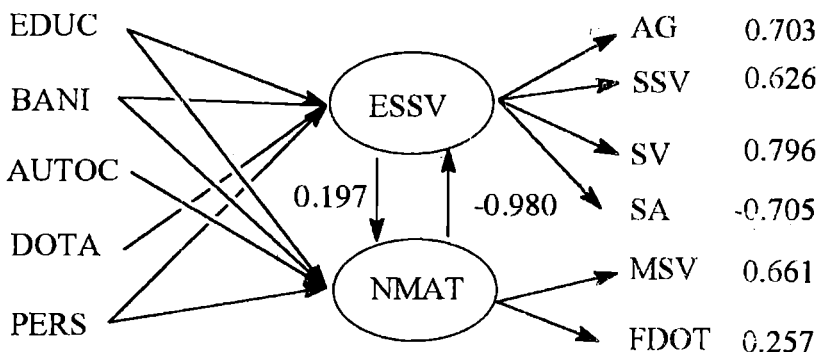
Cele doua concepte teoretice (necesități materiale și estimarea subiectivă a standardului de viață) pot fi definite ca variabile latente care determină comportamentul unor indicatori măsurabili.

Conceptul de “ **Necesități Materiale** “ - **NMAT** este definit prin intermediul a doi indicatori observabili (ce reprezintă “efectul “ măsurabil al acestui concept) :

– **Minimul subiectiv de viață - MSV**. Acest indicator corespunde definiției date în secțiunea III.2.



Sărăcie Subiectivă
Model causal cu 2 variabile latente (faza exploratorie)
FIG. III.8



Sărăcie Subiectivă
Model causal cu 2 variabile latente (faza confirmatorie)
FIG III. 9

TABELUL 6
DESCRIEREA INDICATORILOR

DENUMIRE	ENUNȚUL ÎNTREBĂRII	SCALA
VARSTA	Vârsta capului de familie	numar de ani împliniți
CIV	Starea civilă a capului de familie	1. căsătorit 2. văduv 3. divorțat 4. conviețuire 5. necăsătorit 6. separat
EDUC	Pregătirea școlară a capului de familie	1. fără școala 2. șc. primară 3. șc. generală 4. șc. profes. 5. liceu 6. șc. maiștri 7. șc. postlic. 8. inv. superior
PERS	Numărul de persoane din familie	
BANI	Venitul bănesc realizat din salariu, burse, alocații, etc	...lei **
AUTOC	Autoconsumul familiei (venituri provenite din producția proprie)	...lei **
VENTOT	Venitul total al familiei	BANI+AUTOC **
DOTA	"Pe care din următoarele bunuri le aveți în dotarea gospodăriei?"	1. frigider 2. aragaz 3. congelator 4. mașină de spălat 5. aspirator 6. aparat radio 7. TV alb-negru 8. Tv color 9. video 10. casetofon, magnetofon, pick-up 11. bicicleta / motocicletă 12. autoturism

DENUMIRE	ENUNȚUL ÎNTREBĂRII	SCALA
AG	Autoetichetarea globală: "În momentul actual vă considerați sărac ?"	1. Da 2. Nu
SSV	Standardul Subiectiv de Viață "Cum apreciați veniturile actuale ale familiei dvs.?"	1. nu ne ajung pt strict necesar 2. ne ajung pt strict necesar 3. ne ajung pt. trai decent 4. facem chelt. mai deosebite cu eforturi 5. reusim să avem tot ce este necesar fara mari eforturi
SV	Satisfacția cu veniturile : "Cât de mulțumit sunteți de veniturile actuale ale gospodăriei dvs. ? "	1. f.nemulțumit 2. nemulțumit 3. nici mulțumit nici nemulțumit 4. mulțumit 5. f.mulțumit
SA	Satisfacția cu alimentatia "Există produse alimentare care considerați că ar trebui să facă parte, frecvent din alimentația fam. dvs. iar pe care nu le puteți cumpăra în cantitatea dorită? "	1.nu 2 da, care ? (menționați max trei) : 1. carne 2. pește 3. mezeluri 14. băuturi alcoolice
MSV	Minimul Subiectiv de Viață "Pentru a-și asigura strictul necesar, de cât venit lunar ar avea nevoie -în prezent - familia dvs.?" lei
FDOT	Frustrări în dotarea gospodăriei : " Vă rugăm să precizați care sunt bunurile din tabel pe care nu le aveți, care sunt nesatisfăcătoare și a căror lipsă constituie o sursă de nemulțumire ?"	1. locuința dvs. 2. locuința copiii 3. mobila 4. aragaz, frig. 5. TV 6. autoturism 7. utilaj agr. 8. îmbrăcăminte 9. rechizite 10. mâncare bună

** - Metodologia de calcul este prezentată în Zamfir, coord.1995 (pag. 31-52)

- **Frustrări în ceea ce privește dotarea cu bunuri - FDOT** ;
Acest indicator cuantifică acele bunuri pe care subiecții le doresc începând cu mâncarea mai bună și sfârșind cu locuința

Efectul variabilei latente “ **Estimarea Subiectivă a Standardului de Viață** ” - **ESSV** este observat prin intermediul următorilor indicatori măsurabili :

- **Autoetichetarea globală - AG**
- **Standardul subiectiv de viață - SSV**
- **Satisfacția cu veniturile - SV**
- **Satisfacția cu alimentația - SA**

Tabelul 6 conține definițiile variabilelor măsurabile din această analiză.

Cele două concepte latente sunt determinate cauzal de un set de variabile exogene măsurabile (figura III.8) care se pot grupa în două categorii :

1. Variabile referitoare la susținătorul de familie :

- **vârsta - VÂRSTA**
- **starea civilă - CIV**
- **pregătirea școlară (educația) - EDUC**

2. Variabile referitoare la caracteristici de familie :

- **numărul de persoane din familie - PERS** ;
- **venitul bănesc al familiei - BANI** ;
- **autoconsumul - AUTOC** ; acest indicator exprimă veniturile provenite din producția proprie a familiei ;
- **dotarea gospodariei - DOTA** ; acest indicator exprimă gradul de înzestrare a gospodariei prin însumarea articolelor de folosință îndelungată pe care le deține familia.

Analiza cauzală

Am efectuat o analiză cauzală asupra modelului III.8.

În urma utilizării metodei LISREL s-au pus în evidență următoarele aspecte :

- a) - **variabilele măsurabile exogene semnificative sunt :**
- **pregătirea școlară a capului de familie**

- venitul bănesc al familiei
- autoconsumul familiei
- dotarea gospodăriei
- numărul de membri ai familiei

b) - nu toate variabilele exogene influențează simultan cele două variabile latente

c) - între cele două variabile latente există relații cauzale reciproce, cu intensități diferite și sensuri opuse.

Având în vedere aspectele menționate anterior, au fost eliminate din modelul III.8 variabilele exogene VARSTA și CIV.

Variabilele exogene : pregătire școlară (EDUC), venitul bănesc al familiei (BANI), numărul de persoane din familie (PERS) vor influența ambele variabile latente : ESSV (estimarea subiectivă a standardului de viață) și NMAT (necesități materiale).

În schimb, variabila exogenă AUTO (autoconsumul) influențează numai variabila nemăsurabilă direct NMAT (necesități materiale), neavând o influență directă (ci cel mult indirectă) asupra ESSV (estimarea subiectivă asupra standardului de viață).

O situație similară este realizată de variabila exogenă DOTA (dotarea gospodăriei), variabilă ce are o influență directă asupra conceptului de “estimare subiectivă a standardului de viață” (ESSV). Nu putem vorbi de o influență directă a variabilei exogene DOTA asupra conceptului “necesități materiale” (NMAT); acest lucru se poate realiza însă indirect prin intermediul variabilei latente ESSV.

Cu aceste modificări privind modelul cauzal din figura III.8 rezultă modelul cauzal reprezentat de diagrama III.9.

Așadar, prin aplicarea unei tehnici exploratorii (cap.7, Bollen,1989), am reușit să simplific modelul III.8, ajungând astfel la varianta III.9.

În final, prin utilizarea unei proceduri confirmatorii (cap.7, Bollen,1989) a fost acceptat modelul III.9 (noul model III.9 a fost validat prin aplicarea programului LISREL).

4.2.2 Discutarea modelului cauzal din figura III.9

Modelul cauzal descris în figura III.9 are două variabile latente. Discutarea acestui model poate fi făcută din mai multe puncte de vedere : indicatori ce caracterizează fiecare variabilă latentă în parte și ponderea lor în “explicarea “ variabilei latente respective , legăturile cauzale dintre variabilele latente, intensitatea relațiilor cauzale dintre variabilele exogene modelului și variabilele latente ale modelului, determinarea efectelor cauzale directe, indirecte și totale dintre diferitele variabile ale modelului propus. În continuare, vom detalia toate aceste aspecte.

4.2.2.1. Definirea variabilelor latente

Conceptul “ Necesități materiale “ nemăsurabil direct este definit indirect de indicatorii MVS (minimumul subiectiv de viață) cu coeficientul =0,661 și FDOT (frustrări de dotare) cu coeficientul = 0,257. Se remarcă ponderea substanțială a indicatorului MSV și contribuția mai slabă a frustrării datorită dotării materiale - FDOT.

Așadar, populația reprezentată prin eșantionul studiat în 1994 își concentrează necesitățile materiale asupra veniturilor bănești deoarece aproximativ 70 % din cazuri se aflau sub pragul de sărăcie, fie analitic fie subiectiv.

În ceea ce privește bunurile de folosință îndelungată, lipsa a mai mult de trei obiecte indică o frustrare marcantă și afectează 28 % din gospodăria (vezi și Zamfir coord.1995 p.99).

Din punct de vedere statistic, variabilă latentă “ necesități materiale “ este susceptibilă de îmbunătățiri. Astfel, s-ar putea acorda o nouă dimensiune conceptului de “necesități materiale “ studiindu-se și alte caracteristici ale sale. Din punct de vedere practic aceasta revine la definirea unor noi indicatori (în afara variabilelor observabile MSV și FDOT) care să măsoare efectul indirect al variabilei latente NMAT.

“ Estimarea subiectivă a standardului de viață “ - ESSV este determinată consistent de cei patru indicatori :

- Satisfacția cu veniturile - SV ; coeficient = 0,796
- Satisfacția cu alimentația - SA ; coeficient = - 0,705
- Autoetichetarea globală - AG ; coeficient = 0,703
- Standardul subiectiv de viață - SSV? ; coeficient = 0,626

Ținând cont de scalele de măsurare ale indicatorilor (a se vedea tabelul 6) se poate deduce că acest concept teoretic reflectă o caracteristică subiectivă care poate determina aprecieri favorabile sau defavorabile asupra nivelului de trai, a venitului, a calității alimentației și chiar asupra declarării stării de sărăcie.

4.2.2.2. *Relația cauzală dintre variabilele latente*

Coeficienții structurali între cele două variabile latente pun în evidență următoarea situație :

- “ Necesitățile materiale “ influențează puternic și negativ estimarea subiectivă a standardului de viață cu coeficientul = -0,980. Așadar, cu cât aspirațiile materiale sunt mai mari, cu atât aprecierea standardului de viață scade. Deci, cu cât o persoană are mai multe necesități materiale, în bani sau obiecte, ea își va aprecia defavorabil standardul de viață (se consideră sărac, venitul este insuficient, nivelul de trai este deteriorat iar alimentația este necorespunzătoare).
- “Estimarea subiectivă a standardului de viață “ are o slabă influență pozitivă asupra necesităților materiale; deși ia o valoare mică, coeficientul = 0,197, este însă semnificativ statistic.

O apreciere pozitivă a standardului de viață determină și o creștere a necesităților materiale, fapt explicat parțial de caracteristicile individuale, subiective.

4.2.2.3. *Relațiile cauzale între indicatorii exogeni și variabilele latente.*

Așa cum se observă în diagrama path (figura III.9), cele două variabile latente sunt influențate în mod direct de un set de variabile măsurabile, exogene.

Variabilele educația (EDUC), venitul bănesc (BANI), dotarea (DOTA) și numărul persoanelor din familie (PERS) influențează estimarea subiectivă a standardului de viață (ESSV); influența cea mai mare o are variabila BANI cu coeficientul = 0,617 (tabelul 7).

Variabilele EDUC, BANI, AUTOC, și PERS influențează necesitățile materiale ; influența cea mai mare este generată de variabila PERS (numărul de persoane din familie) cu ponderea = 0,658.

Se remarcă variabila autoconsum (AUTO) care are o influență negativă : atunci când autoconsumul crește, necesitățile materiale scad. Deoarece autoconsumul este o caracteristică principală a familiilor din mediul rural, se poate deduce, implicit, că veniturile proprii ale gospodăriei acoperă o bună parte din necesitățile traiului.

Menționăm că acest fenomen a fost evidențiat și de analiza de regresie privind dinamica liniei de sărăcie subiectivă (secțiunea III.3).

Tabelul 7 conține coeficienții cauzali standardizați ai variabilelor exogene EDUC, BANI, AUTO, DOTA, PERS ce acționează asupra variabilelor latente ESSV, respectiv NMAT (modelul cauzal din figura III.9).

Tabelul 7. Influența directă a factorilor exogeni asupra variabilelor latente (model cauzal III.9)

	ESSV	NMAT
EDUC	0,357	0,149
BANI	0,617	0,291
AUTO	-	-0,246
DOTA	0,449	-
PERS	0,355	0,658

4.2.2.4. Efecte directe, indirecte și totale

Într-un model cauzal reprezentat ca diagramă path se pun în evidență efectele directe și indirecte care se manifestă în grupul variabilelor endogene și exogene.

Nivelul de educație al capului de familie (EDUC) are efecte directe, pozitive asupra celor două variabile latente (0,357 respectiv, 0,149)

Dar educația mai are și un efect indirect, negativ, asupra estimării subiective a traiului (ESSV) : o pondere de -0,146. Urmărind diagrama modelului (Fig. III.9), explicația acestui fenomen este următoarea:

- a) – **efectul direct** : cu cât gradul de educație este mai mare, crește estimarea subiectivă asupra standardului de viață.

Din tabelul 7 rezultă că intensitatea efectului direct al variabilei exogene EDUC asupra variabilei latente ESSV este de 0,357.

- b) – **efectul indirect** : cu cât nivelul de educație este mai mare, va crește nivelul necesităților materiale și, pe această cale, va scădea estimarea subiectivă a standardului de viață. Din graful III.9 reiese că variabila exogenă EDUC influențează în mod indirect variabila latentă ESSV prin intermediul variabilei latente NMAT. Cum intensitatea influenței directe a variabilei EDUC asupra lui NMAT este de 0,149 (Tabelul 7) iar NMAT influențează negativ variabila ESSV cu ponderea -0,980 (figura III.9), rezultă o influență indirectă a variabilei exogene EDUC asupra variabilei latente ESSV cu o intensitate egală cu $0,149 * (-0,980) = -0,146$.

Un comentariu asemănător se poate face și pentru variabilele : BANI (venitul bănesc al familiei) și PERS (numărul de persoane din familie).

Prezența autoconsumului (AUTO) are ca efect direct scăderea necesităților materiale (în principal, scăderea minimumului subiectiv de viață) iar ca efect indirect : creșterea estimării subiective a standardului de viață.

Dotarea gospodăriei (DOTA), influențează direct estimarea standardului de viață (0,449) și indirect necesitățile materiale (0,088). În modelul cauzal III.9 variabila exogenă DOTA nu influențează direct variabila latentă „necesități materiale“ (NMAT). Asistăm numai la o influență indirectă prin intermediul variabilei latente ESSV. Având coeficienții de influențare directă DOTA - ESSV (egal cu 0,449, tabelul 7), respectiv ESSV - NMAT (egal cu 0,197, figura III.9) obținem o pondere pozitivă de $0,449 * 0,197 = 0,088$ a influenței indirecte DOTA - NMAT. Deci, dacă înzestrarea gospodăriei crește, atunci estimarea standardului de viață devine pozitivă (efect direct) iar nivelul necesităților materiale are o ușoară creștere (efect indirect).

Numărul persoanelor din familie (PERS) are atât efect direct cât și indirect (ambele pozitive) asupra necesităților materiale. Deci, dacă într-o familie sunt mai multe persoane, atunci nivelul necesităților materiale crește, fenomen absolut normal.

Mai deosebite sunt efectele (direct și indirect) pe care le are variabila exogenă PERS asupra estimării subiective a standardului de viață ESSV. Astfel, PERS are un efect direct pozitiv (0,355, tabelul 7) și un

efect indirect negativ, prin intermediul variabilei NMAT (coeficient : $-0,645 = 0,658 * (-0,980)$, tabelul 7 + graful III.9).

Efectul indirect PERS-ESSV este perfect explicabil ținând cont de nivelul aspirațiilor materiale NMAT : dacă numărul persoanelor PERS dintr-o familie crește, atunci necesitățile materiale NMAT se majorează și, în consecință, estimarea subiectivă a standardului de viață ESSV scade.

Explicația efectului direct PERS-ESSV este mai complicată.

În acest caz ar trebui investigată structura familiei, deoarece numai în familiile cu multe persoane care aduc un venit concret se poate explica faptul că o familie mai numeroasă influențează favorabil aprecierea subiectivă a standardului de viață ESSV.

Într-o astfel de situație contradictorie, se va lua în considerare efectul total al variabilei exogene PERS asupra variabilei latente ESSV (pondere $-0,290$), efect care este negativ. Menționăm că efectul total PERS - ESSV se obține adunând efectul direct PERS-ESSV (egal cu $0,355$, tabelul 7) cu efectul indirect PERS - ESSV, via NMAT. Efectul indirect PERS - ESSV este egal cu produsul dintre efectul direct PERS - NMAT și efectul direct NMAT - ESSV, adică : $0,658 * (-0,980) = -0,645$ (tabelul 7 graful III.9). Așadar, între variabila exogenă PERS și variabila latentă ESSV avem un efect total de:

$$0,355 - 0,645 = -0,290.$$

Se poate afirma că, în general, cu cât numărul membrilor PERS a unei familii este mai mare, estimarea subiectivă asupra standardului de viață ESSV se diminuează, ținând cont și de legătura acestei variabile cu necesitățile materiale.

O particularitate a modelului III.9 o reprezintă efectele indirecte și totale ale variabilelor latente asupra lor însele ; aceste efecte sunt negative și egale cu $-0,193$.

Explicația acestor efecte este dată de evoluția în timp a variabilelor latente (Joreskog, 1990). Într-adevăr, în timp, necesitățile materiale pot scădea (efect negativ) sau crește (efect pozitiv). De asemenea, estimarea subiectivă a standardului de viață poate evolua, în sens pozitiv sau negativ.

În contextul tipului I de model cauzal abordat, cele două variabile latente ESSV și NMAT vor avea o evoluție descendentă, în timp.

4.2.3 Concluzii și comentarii

Rezultatele obținute prin soluționarea unui model cauzal utilizând programul LISREL sunt calitativ superioare celor din capitolul III.3, în care s-a folosit mecanismul regresiei liniare multiple precum și interpretarea adecvată a unor tabele și grafice.

În modelul cauzal descris de graful III.9 sunt puse în evidență influențe ale unor variabile exogene asupra variabilelor latente care nu sunt direct măsurabile. În plus, între variabilele latente ale modelului III.9 sunt specificate influențe cauzale reciproce, influențe ce nu au aceeași intensitate și semn.

Important este faptul că modelul III.9 a fost validat din punct de vedere statistic iar rezultatele parțiale obținute au o semnificație teoretică. Aceste rezultate confirmă intuiția noastră asupra mecanismului cauzal privind conceptul sărăciei subiective. Avem astfel și un argument empiric pentru validarea modelului III.9.

Am pus în evidență analiza complexă ce a condus în final la obținerea modelului III.9 : definirea conceptelor, detectarea variabilelor latente și a indicatorilor ce le explică indirect, precizarea variabilelor exogene și sugerarea sensului influenței lor asupra variabilelor endogene ale modelului, propunerea unor relații cauzale directe de dependență dintre variabilele exogene și endogene ale modelului. În acest mod, s-a ajuns la un model intermediar (graful cauzal III.8). Pasul decisiv îl constituie eliminarea variabilelor și legăturilor cauzale “ neconcludente “ ale modelului III.8. În acest sens s-a desfășurat un demers consistent de tip exploratoriu. De fiecare dată noul model rezultat din “ simplificarea “ modelului III.8 trebuia validat din punct de vedere statistic. În plus, consecințele deduse din interpretarea modelului cauzal derivat trebuiau să confirme aspectele teoretice și intuitive privind dependențele cauzale reale. Printr-o tehnică confirmatorie am ajuns în final să propunem modelul cauzal III.9. Nu excludem posibilitatea îmbunătățirii modelului cauzal III.9. Acest lucru se poate realiza prin introducerea unor noi variabile observabile care să caracterizeze mai precis variabilele latente ale modelului. De asemenea, se poate studia influența unor noi variabile exogene asupra variabilelor modelului. Extinderea numărului de variabile endogene ale modelului cauzal poate conduce la o rafinare a rezultatelor și va da posibilitatea unor noi interpretări.

În acest sens vom studia, în continuare, un model cauzal, al sărăciei subiective, cu trei variabile latente.

4.3. Tipul II. Model cauzal cu trei variabile latente, pentru studiul sărăciei subiective

4.3.1 *Sărăcia subiectivă și statusul socio-economic*

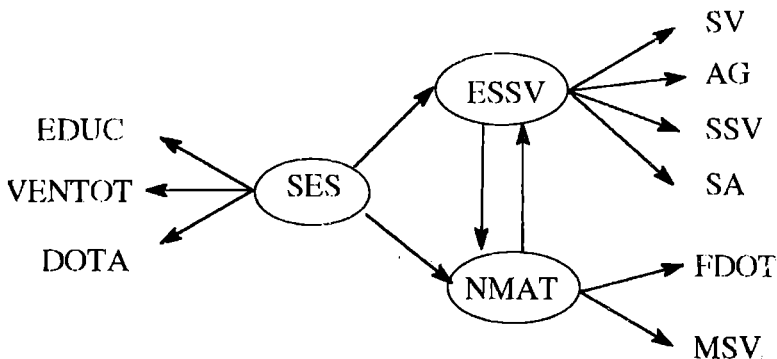
Așa cum s-a discutat în cazul I (modelul cauzal III.9), necesitățile materiale și estimarea subiectivă a standardului de viață sunt concepte teoretice, influențate de caracteristicile individuale dar și sociale. Aceste caracteristici sunt înglobate în conceptul de **status socio-economic**. În general, poziția în societate și potențialul economic al individului sau familiei dețin statusul socio-economic, evident cu deosebiri de la o societate la alta. În Dicționarul de sociologie (1993) **statusul socio-economic** are următoarea definiție :

“Poziția socială a unei persoane sau grup uman în cadrul societății determinată de apartenența de clasă, poziția în sistemul de stratificare socială, ocupația, nivelul de pregătire, venitul, participarea la conducerea vieții social-economice și politice, stilul de viață, prestigiul social etc. S.s.-e. poate fi analizat ca un indicator general (sintetic), un indice constituit din agregarea valorilor elementelor componente, în calitate de statusuri particulare : de clasă, ocupație, educație, venit, etc.”

În contextul actual al societății românești, statusul socio-economic poate fi definit de nivelul de educație, categoria ocupațională și acumulările materiale.

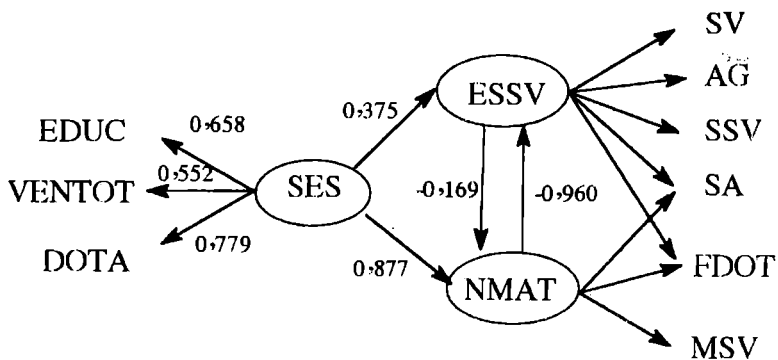
Statusul socio-economic influențează necesitățile materiale ale individului. În funcție de status, individul va majora sau diminua necesitățile materiale. Intervine, în acest caz, și un aspect subiectiv, pe lângă cel strict obiectiv, determinist.

Dar, statusul socio-economic determină și modificări ale estimării subiective a standardului de viață. Un status socio-economic superior va determina, în general, o apreciere pozitivă a nivelului de trai, deoarece o asemenea persoană nu va afirma că este săracă sau că îi lipsesc obiecte absolut necesare în gospodărie. Nici nu va declara că nivelul de trai s-a deteriorat sau că banii îi sunt insuficienți.



Model causal cu 3 variabile latente pentru sărăcia subiectivă (faza exploratorie)

Figura III.10



Model causal cu 3 variabile latente pentru sărăcia subiectivă (faza confirmatorie)

Figura III.11

Dacă, însă, familia are un status inferior, aprecierile subiective ale standardului de viață devin sumbre : individul se va considera sărac și dezavantajat din punct de vedere material.

În mod natural apar întrebările :

- Este statusul socio-economic factorul care influențează necesitățile materiale și estimarea standardului de viață ?
- Cum se manifestă această influență : direct sau indirect ?
- Care este intensitatea acestor influențe ?

Vom răspunde la aceste întrebări prin analizarea următorului model conceptual :

“ Statusul socio-economic are efecte directe și indirecte asupra necesităților materiale și asupra estimării subiective a standardului de viață “.

4.3.2 Influența cauzală a statusului socio-economic

Modelul cauzal următor va analiza trei variabile latente corespunzătoare conceptelor teoretice discutate anterior și anume :

1) Statusul Socio-Economic al familiei (SES), definit de trei indicatori măsurabili :

- **nivelul de educație al capului de familie - EDUC**
- **venitul total al familiei - VENTOT**
- **dotarea gospodăriei - DOTA**

Menționăm că în alte studii de specialitate SES este determinat și de ocupație. Nu am optat pentru asemenea variantă. Explicația este dată de faptul că ocupația, în contextul actual, este încă foarte corelată cu nivelul de educație.

2) - Necesități materiale - NMAT

3) - Estimarea subiectivă a standardului de viață - ESSV

Definirea ultimelor două variabile latente NMAT și ESSV s-a făcut în paragraful III.4.2 - Modelul I. Explicarea variabilelor măsurabile este dată de tabelul 6, discutat deja.

Reprezentarea grafică a modelului cauzal este dată în figura III. 10. Deși diagrama path III.10 este asemănătoare cu cea a modelului cauzal din figura III.8, menționăm că noul model cauzal este diferit. Astfel influența indicatorilor măsurabili exogeni ai modelului III.8 este preluată în modelul cauzal III.10 de variabila latentă SES (statusul socio-economic).

4.3.3 Analiza exploratorie și confirmatorie

În continuare, va fi expusă o analiză cauzală cu variabile latente a modelului III.10, utilizând metoda LISREL.

Procedura automată LISREL a condus la modificări în schema grafică III.10. În urma aplicării unei proceduri exploratorii s-a ajuns la modelul cauzal din figura III.11. Astfel, prezența variabilelor observabile SA (satisfacția cu alimentația) și FDOT (frustrări în dotarea gospodăriei) în explicarea comună a variabilelor latente ESSV (estimarea subiectivă a standardului de viață) și NMAT (necesități materiale) a fost impusă de îmbunătățirea stabilității statistice a modelului cauzal inițial III.10. Discuția teoretică va justifica modificările din modelul conceptual III.10. În urma aplicării programului LISREL s-au obținut coeficienții structurali ai modelului cauzal III.11 (tabelul 8).

Tabelul 8. Coeficienții structurali standardizați ai modelului empiric III.11

indicatori	variabile latente		
	SES	ESSV	NMAT
EDUC	0,658		
VENTOT	0,552		
DOTA	0,779		
AG		0,730	
SSV		0,638	
SV		0,796	
SA		-0,643	-0,291
FDOT		-0,336	0,217
MSV			0,596

4.3.4. *Discutarea modelului cauzal III.11*

Vom discuta pe larg modelul cauzal III.11 din punctul de vedere al procedurii de identificare a variabilelor latente, interdependențele dintre acestea, precum și calculul efectelor directe, indirecte și totale.

4.3.4.1. *Definirea variabilelor latente*

Variabila latentă, status socio-economic SES, este măsurată indirect prin efectul pe care îl are asupra următorilor indicatori :

- educația - EDUC - coeficient = 0,658
- venitul total - VENTOT - coeficient = 0,552
- dotarea gospodăriei - DOTA - coeficient = 0,779

Se remarcă indicatorul DOTA (înzestrarea gospodăriei), care este definitoriu pentru statusul socio-economic, ponderea sa în “explicarea” variabilei latente SES fiind cea mai mare (= 0,779).

Dar și nivelul de educație (EDUC) are o pondere semnificativă în definirea statusului socio-economic, ceea ce echilibrează aspectele materiale ale conceptului latent SES.

Celelalte variabilele latente au suferit următoarele modificări:

1. **Necesități materiale - N_{MAT}** ; acest concept este măsurat prin intermediul a 3 indicatori :
 - minimul subiectiv de viață- MSV - coeficient = 0,596
 - frustrări în dotarea gospodăriei - FDOT-coefic. = 0,217
 - satisfacția cu alimentația - SA - coeficient = - 0,291

Se remarcă faptul că minimul subiectiv de viață - MSV este indicatorul de referință pentru conceptul de “necesități materiale”, ponderea sa fiind cea mai mare în explicarea conceptului N_{MAT}. Deși acest concept influențează foarte mult estimarea venitului subiectiv, nu trebuie neglijat că el determină și aprecierea alimentației și lista bunurilor materiale care lipsesc sau sunt insuficiente în gospodărie

Deci, dacă nivelul aspirațiilor materiale crește, atunci aceasta va antrena o majorare a necesităților materiale și financiare.

Menționăm că analiza cauzală se bazează pe ipoteza enunțată în cap. II.3.4 conform căreia indicatorii sunt efecte ale variabilelor latente.

2. **Estimarea subiectivă a standardului de viață - ESSV ;**

este definită de 5 indicatori :

- satisfacția cu veniturile - SV ; coef. = 0,796
- autoetichetarea globală - AG ; coef. = 0,730
- standardul subiectiv de viață - SSV ; coef. = 0,638
- satisfacția cu alimentația - SA ; coef. = -0,643
- frustrări în dotarea gospodăriei - FDOT ; coef. = -0,336

Indicatorul de referință pentru această variabilă latentă este satisfacția cu veniturile - SV (ponderea 0,796), urmată descendent de variabilele observabile : AG, SA, SSV, FDOT.

Aprecierca subiectivă nefavorabilă asupra standardului de viață va influența negativ aprecierea venitului și a calității alimentației. În același context, subiectul se va considera sărac, va aprecia că nivelul de trai a fost afectat negativ, și va majora lista bunurilor materiale care îi lipsesc sau sunt insuficiente în gospodărie.

Situația descrisă mai sus este aceea a existenței unei stări latente de sărăcie subiectivă care nu poate fi explicată exhaustiv pe baza unor măsurători cantitative.

O caracteristică a modelului cauzal III.11 o constituie prezența variabilelor SA - satisfacția cu alimentația și FDOT - frustrări în dotarea gospodăriei, în definirea ambelor variabile latente ESSV și NMAT. Această "simultaneitate" de influență a fost impusă de procedura confirmatorie LISREL, pentru a îmbunătăți stabilitatea statistică a modelului structural III.10. Procedura este justificabilă din punct de vedere teoretic, astfel:

- Lista bunurilor care lipsesc sau sunt insuficiente a fost obținută cu întrebarea :

“ Indicați lucrurile pe care nu le aveți sau sunt insuficiente/ învechite și a căror lipsă constituie o sursă de nemulțumire pentru dumneavoastră “.

Răspunsul la această întrebare nu este numai o reflectare a necesităților materiale dar și a aprecierii standardului de viață. Este normal ca o persoană cu o percepție pozitivă asupra condițiilor de viață, nu va indica un număr mare de obiecte care să constituie o sursă de nemulțumire.

– Aspectele cantitative și calitative ale alimentației sunt puse în evidență de aprecierea standardului de viață, dar și de necesitățile materiale. Este evident că nivelul necesităților materiale va determina variații ale dietei alimentare.

4.3.4.2. *Analiza cauzală a variabilelor latente*

Așa cum s-a discutat și în cap. II, paragraful 2.2, clasificarea variabilelor latente din modelul cauzal în variabile endogene și variabile exogene este relativă și depinde de interpretările date acestor variabile în cadrul modelului.

În contextul acestui model s-a considerat că statusul socio-economic este variabila exogenă care influențează direct cele două variabile latente endogene : estimarea standardului subiectiv de viață și necesitățile materiale.

Statusul socio - economic influențează direct și pozitiv necesitățile materiale - coeficientul = 0,877 și estimarea subiectivă a standardului de viață - coeficient = 0,375.

Cele două variabile latente endogene se influențează reciproc astfel:

- Necesitățile materiale influențează puternic negativ estimarea standardului de viață - coeficient = -0,960;
Deci, cu cât nivelul aspirațiilor materiale este mai ridicat, aprecierea standardului subiectiv de viață scade
- Estimarea subiectivă asupra standardului de viață influențează slab și negativ necesitățile materiale - cu coeficientul = -0,169. Dacă estimarea subiectivă asupra standardului de viață ESSV devine favorabilă, atunci aspirațiile materiale vor scade puțin. Totuși legătura cauzală dintre conceptele ESSV și NEMAT este slabă.

4.3.4.3. Efecte directe, indirecte și totale

Datorită determinărilor cauzale prezentate anterior, variabila latentă exogenă SES - statusul socio-economic - are atât efecte directe (menționate mai sus) cât și efecte indirecte asupra variabilelor endogene.

Dacă efectele **directe** sunt pozitive (0,375, respectiv 0,877, figura III.11) cele **indirecte** sunt negative : un status socio-economic ridicat determină, în mod indirect o ușoară scădere a aspirațiilor materiale - efectul= - 0,063 (0,375 * (-0,169) = -0,063, figura III.11).

Acest efect este posibil datorită influenței factorului subiectiv care este ESSV. Explicația ia în considerare faptul că statusul socio-economic influențează pozitiv estimarea subiectivă a standardului de viață (ESSV), care la rândul său determină o diminuare a necesităților materiale.

SES determină, în mod indirect, și scăderea aprecierii subiective a standardului de viață ESSV - efectul= - 0,842 (0,877 * (-0,960) = -0,842, figura III.11). Efectul indirect este consistent, deoarece în fluxul cauzal apare medierea adusă de conceptul “necesități materiale”.

Deci, dacă statusul socio - economic crește, el va determina, în mod indirect o apreciere nefavorabilă a standardului de viață, deoarece necesitățile materiale vor crește semnificativ. Asadar, se confirmă enunțul cauzal formulat anterior.

Dar, în afara efectelor directe și indirecte, în analiza cauzală sunt puse în evidență și efectele totale ale variabilelor exogene asupra celor endogene.

Analiza **efectelor totale** ale statusului socio - economic SES conduce la următoarele influențe asupra variabilelor latente endogene ale modelului III.11 :

$$0,375 + 0,877 * (-0,960) = -0,467 \quad \text{pentru ESSV}$$

$$0,877 + 0,375 * (-0,169) = 0,814 \quad \text{pentru NMAT}$$

Deci, se poate afirma că, în timp, statusul socio - economic va exercita o influență semnificativă asupra necesităților materiale (sens pozitiv) și asupra estimării subiective a standardului de viață (cu o pondere mai mică și sens negativ).

Prin situația descrisă, modelul empiric și modelul conceptual se validează și se completează reciproc.

4.3.5 Comentarii

Modelele cauzale III.9 și III.11 au fost definite printr-o tehnică exploratorie-confirmatorie. În esență, efectul tuturor variabilelor observabile exogene din modelul III.9 este înlocuit în modelul III.11 printr-o singură variabilă latentă exogenă, anume statusul socio-economic SES.

Modelele cauzale III.9 și III.11 pun în evidență o influență directă puternică (sens negativ) a variabilei latente „necesități materiale“ (NMAT) asupra variabilei latente „estimarea subiectivă a standardului de viață“ ESSV (ponderea : -0,980, respectiv -0,960).

Influența directă a variabilei ESSV asupra variabilei NMAT este mult mai slabă în cele două modele. Această influență diferă însă ca semn în modelele cauzale III.9 și III.11.

Menționăm că cele două modele cauzale nu sunt “identice”, variabilele măsurabile ce explică statusul socio-economic SES (modelul III.11) nefiind aceleași cu variabilele observabile exogene ale modelului cauzal III.9.

Modelele III.9 și III.11 conduc la concluzii comune și nu neagă rezultatele teoretice și empirice cunoscute.

III.5 CONCLUZII FINALE

Am analizat conceptul de sărăcie subiectivă, folosind un aparat statistic divers. Am început cu interpretarea unor tabele și grafice (capitolul III.2), s-a continuat cu utilizarea ecuațiilor liniare de regresie multiplă pentru predicție și ierarhizarea unor variabile (capitolul III.3), studiul statistic încheindu-se prin soluționarea unor modele cauzale cu ajutorul metodei LISREL (capitolul III.4).

În cadrul fiecărei etape s-au discutat interpretările, din punct de vedere teoretic și empiric, ale rezultatelor obținute. Metode diferite de analiză statistică a datelor și modele cauzale diferite au pus în evidență puncte de convergență comune, rezultatele fiind comparabile.

Din aceste prelucrări statistice a reieșit net caracterul calitativ superior al procedurii LISREL de analiză cauzală.

Modelele de regresie folosesc numai variabile observabile continue sau dihotomice ("dummy"), și de cele mai multe ori au o singură variabilă dependentă.

Analiza factorială presupune implicit noțiunea de variabilă latentă (factorii), determină chiar intensitățile corelărilor dintre factori (analiza factorilor "oblici") dar nu reușește să stabilească sensul cauzal dintre doi factori și să precizeze intensitatea diferită de influențare reciprocă dintre aceștia.

Toate aceste neajunsuri nu apar la soluționarea unui model cauzal prin metoda LISREL.

Definirea unui model cauzal presupune și existența unor variabile (concepte, variabile latente) ce nu pot fi măsurate în mod direct, ci apreciate prin efectul pe care îl au asupra unor variabile observabile. Determinarea intensităților inter-actiunilor cauzale dintre diferitele tipuri de variabile (masurabile și latente) precum și stabilirea sensului cauză-efect devine

posibilă numai printr-o analiză cauzală complexă de tipul LISREL. Celelalte proceduri statistice enumerate soluționează numai parțial problema.

Un avantaj major al procedurii LISREL îl constituie faptul că variabilele observabile pot fi de tipuri diferite : nominale, ordinale, continue.

Pentru soluționarea modelului cauzal, metoda LISREL utilizează “coeficienții de dependență” dintre variabilele observabile ale modelului. Acești coeficienți sunt deduși din datele experimentale și pot avea semnificații diferite : măsuri ale corelării dintre variabile (coeficientul Pearson, Kendall, Spearman) sau, mai general, măsuri de asociere dintre diversele tipuri de variabile observabile (nominale, ordinale, continue). Menționăm în acest sens și lucrarea lui Albert M.Liebtrau, „Measures of associations“, Sage,London,1990.

În concluzie, procedura LISREL constituie un puternic instrument adecvat unei analize cauzale complexe.

Nu putem încheia fără a pune în evidență și uriașa muncă de analiză ce conduce în final la modelul cauzal propus. Aceasta cuprinde : operațiunea de culegere a datelor, eliminarea valorilor aberante, stabilirea conceptelor teoretice și a variabilelor observabile ale modelului, stabilirea ipotezelor privind natura variabilelor și relațiile cauzale dintre acestea, numeroase analize exploratorii și validări statistice ale modelelor selecționate.

BIBLIOGRAFIE

- Alcock P. (1993) - **Understanding Poverty**, Macmillan
- Angel R. & Gronfein, (1988) - **The use of subjective information in statistical models**, American Sociological Review, vol. 53
- Asher H. B. (1976) - **Causal Modeling**, Sage Publications
- Babbie, E. (1992) - **The Practice of Social Research**, Wadsworth Publishing Company, U.S.A.
- Bagozzi R. P (1980) - **Performance and Satisfaction in an industrial sale force : An examination of their antecedents and simultaneity** - Journal of Marketing,44, 65-72.
- Bagozzi R. P. & Heatherton T.F. (1994) - **A general Approach to Representing Multifaced Personality Constructs : Application to State Self-Esteem**, in Structural Equation Modeling vol 1,nr. 1
- Bartholomew D.J. (1987) - **Latent variables models and factor analysis**, New York, Oxford University Press
- Berry W.D. (1984) - **Nonrecursive Causal Models**, Sage Publ.
- Byrne B. M. (1989) - **A Primer of LISREL. Basic applications**, Springer -Verlag, New York
- Blalock H.M. (1968) - **Methods in Social Research**, MacMillan
- Blalock H. M. (1971) - **Causal Models in The Social Sciences** ,MacMillan
- Blalock H. M. jr. (1975) - **Measurements in the Social Sciences -Theories and Strategies**, MacMillan
- Bollen K. A. (1989) - **Structural equations with latent variables**, New York John Wiley
- Bollen K. A. Long S.J (1993) - **Testing Structural Equation Models**, Sage Publications

- Carley M. (1981) - **Social Measurement and Social Indicators**, London, George Allen & Unwin
- Casper L.M., McLanahan S.S., Garfinkel I. (1994) - **The Gender - Poverty Gap : What We Can Learn from other Countries**, American Sociological Review, vol 59, 594-605
- Chelcea S., Constantinescu I., Damian N., Ștefănescu St.(1975) - **Cerințe sociale ale culturii de masă (studiu statistic)** - Viitorul Social, an 4,nr. 3, București
- Chelcea S. (1992) - **Metode și tehnici de cercetare sociologică**, Universitatea București
- Clogg C.C. (1992) - **The Impact of Sociological Methodology on Statistical Methodology**, Statistical Science, vol 7, no 2, p. 183 - 207
- Cohen J. (1983) - **Peer influence on college aspirations with initial aspirations controlled**, American Sociological Review,vol. 84, p. 728 -734
- Coombs, C.H., Dawes R.M., Tversky A. (1970) - **Mathematical Psychology**, Prentice Hall
- Costner L. C. & Schoenberg R. (1973) - **“ Indicator Ills in Multiple Indicator Model “ in Structural Equation Models for Social Sciences**, editori : A. S. Goldberger și O.D. Duncan
- Cuttance P. & Ecob R. coord (1987) - **Structural Modeling by Example**, Cambridge University Press
- Draper N. R. & Smith H. (1981) - **Applied Regression Analysis**, New York, John Wiley
- Duncan O.D., Haller A.O. & Portes A. (1968) - **Peer influence on aspiration : A reinterpretation** - American Journal of Sociology, 74, p. 119-137
- Durkheim, E. (1974) - **Regulile metodei sociologice**, Editura Științifică, București
- Everitt B. S.(1987) -**Introduction to Optimization Methods and their Application in Statistics**, Chapman and Hall
- Ferge Z., Millar S.M. (1987) - **Dynamics of Deprivation**, Gower Publishing House
- Gustafsson B. & Lindblon M. (1993) - **Poverty Lines in Seven European Countries, Australia, Canada and U.S.A.**, Journal of European Social Policy, 3 (1)

- Hagenaars J. A. (1993) - **Loglinear Models with Latent Variables** - Sage Publications, Inc
- Hanushek E. & Jackson J. (1981) - **Statistical Methods for Social Scientist**, Academic Press
- Hayduk A. L. (1989) - **Structural Equation Modeling with LISREL**, The John Hopkins University Press
- Heck R. H & Johnsrud L. K. (1994) - **Workplace Stratification in Higher Education Administration : Proposing and testing a Structural Model**, in Structural Equation Modeling, vol1, nr1
- Holland P.W. (1992) - **Comment : It's the Interplay that's Important**, Statistical Science, vol. 7, nr 2, 198 - 201
- Hull C. L. (1943) - **Principle of Behaviour**, Appleton, N.Y.
- Joreskog K.G. (1977) - **Structural Equation models in the social sciences ; Specifications, estimation and testing**, in Applications of Statistics, Amsterdam
- Joreskog K.G. & Sorbom D. (1982) - **Recent development in structural equation modeling** - Journal of Marketing Research
- Joreskog K. G & Sorbom (1990) - **New developments in LISREL : Analysis of ordinal variables using polychoric correlations and weighted least squares** , Quality and Quantity 24, 387-404.
- Joreskog K. G & Sorbom D. (1993) - **LISREL 8 : Structural Equation Modeling with SIMPLIS Command Language**, Scientific Software International, Inc. Chicago
- Kurtz, N. (1983) - **Introduction to social statistics**, Mc Graw - Hill Book Company
- Linden L. (1986) - **Developmental Change in Linear Structural Equations**, Almqvist & Wiskell International, Stockholm
- Loehlin, J. (1987) - **Latent Variables Models. An introductory to factor, path and structure Analysis**, New Jersey
- McCutcheon A.L. (1987) - **Latent Class Analysis**, Sage Publication
- Mărginean I. & Socol Gh. (1991) - **Calitatea vieții în perioada de tranziție**, Sociologia Românească, serie noua, nr.3-4
- Mărginean I. (1994) - **Politica socială și economia de piață în România**, ICCV, Centrul de Informare și documentare, București

- Mărginean I.(1995) - **Ruralul și reforma socială în România**, Sociologia Românească, serie nouă, nr. 3-4
- Marshall G. (editor) 1994 - **The concise Oxford dictionary of Sociology**, Oxford University Press, U. K.
- Muffels R., Kapteyn J. A., Berghman J. (1990) - **Poverty in The Netherlands** - Report Dutch Ministry of Social Affaires
- Muffels R., Berghman J., Dirven H.J. (1992) - **A Multy Method Approach to Monitor the Evolution of Poverty**, Journal of European Social Policy, 2(3), 192-213
- Norusis M. SPSS (1985) - **Advanced Statistics Guide**, SPSS Inc.
- Pohl R., Soleilhavoup J. (1982) - **La Transmission du statut social sur deux ou trois generations**, Economie et Statistique, nr 144
- Popper, Karl R. (1981) - **Logica cercetarii**, Ed. Științifică și Tehnică, Bucuresti
- Praag B.M.S. (1971) - **The Poverty Line** - 's-Gravenhage
- Praag B.M.S., Hagenaars A.J.M., Weren H. (1980) - **Poverty in Europe**, University of Leiden, COEPS
- Rotariu, T. (1974) - **Influența factorilor mediu de proveniență, origine socială și pregătire școlară asupra reușitei la admiterea în învățământul superior**, Studia, Series Sociologia, Cluj-Napoca
- Rotariu, T. (1991) - **Curs de metode și tehnici de cercetare sociologică**, Univesitatea Babeș-Bolyai
- Rowntree B.S. (1901) - **Poverty - A Study of Town Life**, London, Macmillan
- Ruggles P. (1992) - **Drawing the Line**, The Urban Institute Press, Washington D. C.
- Sandu D. (1996) - **Sociologia tranziției : Valori și tipuri sociale în România**, Ed. Staff, București
- Schmidt M. & Steier R. (1993) - **A Latent State-Trait Model (not only) for Social Desirability**, Person. Individ. Diff. vol 14,no4
- Schmidt J. M. & Steier R. (1990) - **Latent state-trait models in attitude research**, Quality and Quantity,24, 427-445

- Scott Long J. (1983) - **Confirmatory Factor Analysis - A preface to Lisrel**, Sage Publications
- Scott Long J. (1983) - **Covariance Structure Models - An Introduction to LISREL**, Sage Publications
- Sen A. K. (1979) - **Issues in the Measurement of Poverty**, Scandinavian Journal of Economics, vol 81, 285-307
- Spearman C.E. (1904) - **General intelligence - objectively determined and measured**, American Journal of Psychology
- Spicker P. (1993) - **Poverty and Social Security**, Routledge, London and New York
- Ștefănescu St., Zamfir C. (1977) - **Structura stilului de conducere**, Viitorul Social, an. 6, nr.4, București
- Ștefănescu P. & Ștefănescu St. (1983) - **Probleme de măsurare în modele cauzale** - Lucrările celui de al IV-lea Simpozion - Modelarea cibernetică a proceselor de producție, București
- Ștefănescu P. (1995) - **Using a computer network for social work training**, în "Computers in Human Services", vol 12, no.1/2, p.75-82, Haworth Press, New York
- Ștefănescu P. (1996) - **Latent variables in the study of subjective poverty - the Romanian context**, lucrare prezentată la a IV - a Conferință internațională de metodologie în științele sociale, Universitatea Essex, U. K.
- Ștefănescu P. (1996) - **Analiza cauzala cu variabile latente. Aplicații în științele sociale**, ICCV, București
- Ștefănescu P. (1997) - **Studiul conceptului de sărăcie subiectivă în contextul actual al societății românești**, Revista de cercetări sociale, an 4, nr. 1
- Steyaert J., Colomby D., Rafferty J. (editori) ; (1996), **Human services and information technology : An international perspective**, Arena Publishing (U. K.) & Stakes (Finland)
- Torrens - Ibern, J. (1972) - **Modeles et methodes de l'analyse factorielle**, Edition Dunod, Paris

- Thurstone, L.L. (1928) - **Attitudes can be measured**, American Journal of Sociology
- Townsend P. (1974) - **The Concept of Poverty**, Heineman Educational Books, London
- Townsend P. (1979) - **Poverty in the U. K.**, Penguin Books, London
- Townsend P. (1986) - **Deprivation**, Journal of Social Policy, vol 16 125-146, Cambridge, University Press
- Townsend P. (1993) - **The International Analysis of Poverty**, Harvester, Wheatsheaf, Great Britain
- Vlăsceanu M. (1985) - **Conștiință și cauzalitate**, Editura Științifică și Enciclopedică, București
- Vlăsceanu L. (1986) - **Metodologia cercetării sociologice**, Ed. Științifică și Enciclopedică, București
- Von Eye A., Clogg C.(editori) 1994 - **Latent variable analysis; Applications for developmental research**,SAGE Publications London
- Wallace W. (1971) - **The Logic of Science in Sociology**, Aldine - Atherton, Chicago
- Weakliem D.I.. (1990) - **Relative Wages and the Radical Theory of Economic Segmentation**, American Sociological Review vol 5,574-590
- Wright S. (1960) - **The Treatment of Reciprocal Interaction, with or without Lag, in Path Analysis**, Biometrics, Sept. 1960,vol.16, no.3
- Yli-Luoma, P.V.J. (1991) - **Predictors of Critical Thinking Abilities**, Almqvist & Wiskell International, Stockholm
- Zamfir C. coord. (1980) - **Dezvoltarea umană a întreprinderii**, Editura Academiei, București
- Zamfir C. coord. (1984) - **Indicatori și surse de variație a calității vieții**, Ed. Academiei, București
- Zamfir C. (1987) - **Structurile gândirii sociologice**, Ed. Politică, București

Zamfir C. & Vlăsceanu L. coord (1993) - **Dicționar de Sociologie**, Editura Babel, București

Zamfir E. & Zamfir C. coord (1995a) - **Politici Sociale - România în Context European**, Ed. Alternative, București

Zamfir C. coord. (1995b) - **Dimensiuni ale sărăciei -1994**, Ed. Expert, București

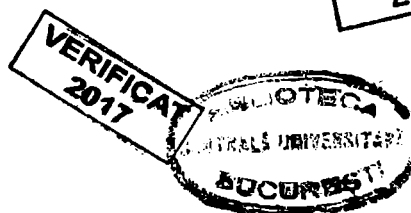
CUPRINS

I. ANALIZA CAUZALĂ	5
I.1 Introducere	5
I.2 Explicația cauzală	6
I.3 Stabilirea cauzalității	8
1. Stabilirea sensului cauzării	8
2. Stabilirea duratei și a distanței	10
3. Problema cauzelor aparente	10
4. Riscul de a fi corect din rațiuni false	11
5. Problema numărului observațiilor	12
6. Variația contextuală a cauzalității	13
I.4 Multicauzalitate	14
4.1. Structura listei cauzale	14
4.2. Mecanisme și circuite determinative	15
4.3. Mărimea contribuției factorilor determinanți	15
4.4. Lanțuri cauzale	15
I.5 Structura standard a teoriei cauzale	17
I.6 Modele	19
I.7 Variabile latente	22
II. METODOLOGIA MODELELOR CAUZALE	26
II.1. Metodologia de analiză a datelor	26
1.1. Modele de analiză a datelor	26
1.2. Soluționarea modelului empiric	30
1.3. Validarea modelului empiric	34
1.4. Necesitatea unui cadru general	36
Concluzii	37
II.2. Noțiunea de model cauzal	38
2.1. Reprezentarea modelelor cauzale	38
2.2. Model cu variabile latente	39
2.3. Model cu variabile observabile	42

2.4. Indicatori statistici utilizați	45
2.5. Convenții privind grafurile cauzale	47
2.6. Măsurarea dependențelor cauzale	48
Concluzii	51
II.3. Validarea relației de cauzalitate	52
3.1. Introducere	52
3.2. Tehnici de izolare a variabilelor	52
3.3. Relația de asociere	60
3.4. Stabilirea sensului cauzalității	63
Concluzii	67
II.4. Tehnici de soluționare a modelelor cauzale	68
4.1. Principiul statistic fundamental	68
4.2. Criterii concrete	74
4.3. Soluționarea unor modele clasice	79
Concluzii	83
II.5. Raportul dintre variabilele latente și variabilele măsurabile ...	84
5.1. Conceperea modelelor cauzale	84
5.2. Tehnici de validare	87
Concluzii	95
III. APLICAȚII ÎN SOCIOLOGIE : SĂRĂCIA	96
III.1 Definirea conceptului de sărăcie	96
1.1. Definiție empirică	96
1.2. Concepte alternative de sărăcie	97
1.2.1 Sărăcia absolută	97
1.2.2 Sărăcia relativă	97
1.2.3 Sărăcia subiectivă	99
1.2.4 Indicele de deprivare relativă	100
III.2 Sărăcia în România	103
2.1 Caracteristici ale eșantionului	103
2.2 Praguri de sărăcie	107
2.3 Concluzii și precizări	110

III.3 Dinamica liniei de sărăcie	111
3.1 Linia de sărăcie subiectivă (LSS)	111
Precizări și concluzii	115
3.2 Estimarea coeficienților ecuației de regresie	115
Concluzii	117
3.3 Evoluția minimumului subiectiv de viață în mediul urban	118
3.4 Evoluția minimumului subiectiv de viață în mediul rural	119
3.5 Concluzii privind dinamica liniei de sărăcie subiectivă	121
III.4 Modele cauzale cu variabile latente în studiul sărăciei subiective	
4.1 Introducere	125
4.2 Tipul I. Model cauzal cu două variabile latente	125
4.2.1 Definierea și selectarea unor modele cauzale	125
4.2.2 Discutarea modelului cauzal III.9	132
4.2.2.1 Definierea variabilelor latente	132
4.2.2.2 Relația cauzală dintre variabilele latente	133
4.2.2.3 Relațiile cauzale între indicatorii exogeni și variabilele latente	133
4.2.2.4 Efecte directe ,indirecte și totale	134
4.2.3 Concluzii și comentarii	137
4.3 Tipul II. Model cauzal cu trei variabile latente pentru studiul sărăciei subiective	138
4.3.1 Sărăcia subiectivă și statusul socio-economic	138
4.3.2 Influența cauzală a statusului socio-economic	140
4.3.3 Analiza exploratorie și confirmatorie	141
4.3.4 Discutarea modelului cauzal III.11	142
4.3.4.1 Definierea variabilelor latente	142
4.3.4.2 Analiza cauzală a variabilelor latente	144
4.3.4.3 Efecte directe, indirecte și totale	145
4.3.5 Comentarii	146
III.5 Concluzii finale	147
BIBLIOGRAFIA	149

**VERIFICAT
2007**



Tiparul s-a executat sub cda 551 la
Tipografia Editurii Universității din București

DATA RESTITUIRII

4. FEB. 2002	11 MAR 2003	3. MAR 2005
<u> </u>	31 OCT 2003	19. MAI. 2006
25 FEB. 2002	20 NOV 2003	<u> </u>
7. MAI. 2002	10. DEC 2003	<u> </u>
8. IAN. 2003	9. IAN. 2004	
11. IAN 2003	9. IAN. 2004	
	15. APR. 2004	
14 IAN 2003	25. MAI. 2004	
16. IAN. 2003	<u> </u>	
	8. DEC. 2004	
25. FEB. 2003	10. FEB. 2005	

18750

Modelele constituie simplificări ale realității. Ele există în diferite nivele de abstractizare, de la modele metafizice (metamodele) sau paradigme, sau ipoteze asupra lumii, până la modele precise sau teorii științifice specifice.

Relațiile cauzale între fenomene constituie o preocupare de bază a științelor sociale. Modelele exprimă explicit aceste relații cauzale.

Studiile empirice asupra fenomenelor studiate nu pot “demonstra” teorii sau ipoteze, dar contribuie la validarea explicațiilor cauzale.