Éltető Ödön—Vita László

Jövedelemeloszlások közelítése és prognosztizálása

Bevezetés

Cikkünkben a jövedelemeloszlások prognosztizálásának egy viszonylag egyszerű, s ugyanakkor hatékonynak tűnő módszerét írjuk le. A módszer azokon a kutatásokon alapszik, amelyek egyrészt az empirikus-jövedelemeloszlásokat leíró függvények típusának, másrészt a jövedelmi egyenlőtlenséget kialakító legfőbb tényezőknek az időbeli stabilitását vizsgálták. A stabilitás-vizsgálat az első esetben annak a kérdésnek az eldöntését jelenti, hogy a különböző évekre vonatkozó empirikus jövedelemeloszlások ugyanolyan típusú valószínűségeloszlással közelíthetők-e, a második esetben pedig arra irányul a vizsgálódás, hogy a jövedelemegyenlőtlenséget — pontosabban a jövedelmek relatív szórásnégyzetét — ugyanazok a tényezők, s ugyanolyan mértékben határozzák-e meg az egyes évek során.

Mindkét fajta stabilitás-vizsgálatra az adja meg a lehetőséget, hogy a Központi Statisztikai Hivatal 1962 óta ötévenként rendszeresen — lényegében hasonló technikával és hasonló fogalmi rendszerben — hajt végre országos reprezentatív jövedelmi felvételeket. Cikkünkben a három legutóbbi (az 1967., 1972. és 1977. évekre vonatkozó) jövedelmi felvétel adatainak elemzéséből indulunk ki. A továbbiakban mindig az egy főre jutó személyes jövedelmek eloszlását, illetve egyenlőtlenségét vizsgáljuk, majd e vizsgálataink eredményei alapján teszünk kísérletet a jövedelemeloszlások prognosztizálására. Ennek megfelelően vizsgálataink és prognózósaink természetesen csak a statisztikailag regisztrálható jövedelmekre vonatkoznak. A számba nem vehető jövedelmek arányáról és a jövedelmek egyenlőtlenségére gyakorolt hatásáról különböző — nagyságrendben eléggé eltérő — hipotézisek láttak már napvilágot. Mivel azonban ezek empirikus adatokkal nem verifikálhatók, célszerűbbnek látjuk, ha a jövedelmi felvételek tényleges adatait használjuk fel számításainkhoz.

A tanulmány első részében a jövedelemeloszlások közelítéséhez használt függvényeket, ezek néhány tulajdonságát, a paraméterbecslési eljárásokat és az illeszkedésvizsgálatok eredményeit ismertetjük. A második részben azt mutatjuk be, hogyan alkalmazhatók a jövedelemeloszlásokat leíró függvények típusának időbeli változására vagy stabilitására vonatkozó eredmények bizonyos hipotézisekre támaszkodva jövedelemeloszlások prognosztizálására prognózis változatok készítésére.

1. Jövedelemeloszlások közelítése eloszlásfüggvényekkel

1.1 *Az alkalmazott eloszlásfüggvények és néhány tulajdonságuk*

Vizsgálatainkban a *Pareto*-eloszlást (P), a két-, illetve háromparaméteres lognormális eloszlást (L-2, illetve L-3), gamma-eloszlást (G-2, illetve G-3), *Champernowne*-féle eloszlást (C-2, illetve C-3), valamint a *logisztikus* vagy más néven általánosított *sech*^{2} eloszlást (L) alkalmaztuk a magyar jövedelemeloszlások leírására. Lényegében ezek azok az eloszlásfüggvények, amelyek legtöbbször szerepelnek a nemzetközi irodalomban jövedelemeloszlások közelítésére. (L. pl. [1], [2], [14], [15].) Ezek közül a P és C-3 eloszlások egyáltalán nem bizonyultak alkalmasnak a magyar jövedelemeloszlások leírására, mert – egy-két esettől eltekintve – irreálisan magas alsó korlátokat adtak a legalacsonyabb jövedelmekre. Ezért velük a továbbiakban nem is foglalkozunk. A többi valószínűségeloszlás $f(x)$ sűrűségfüggvényét és $F(x)$ eloszlásfüggvényét az 1. táblázat tartalmazza.[1]

1. táblázat

A vizsgált valószínűségeloszlások sűrűség- és eloszlásfüggvényei

Eloszlás	Sűrűségfüggvény	Eloszlásfüggvény
L-2	$\frac{1}{x\sigma\sqrt{2\pi}}\exp\left\{-\frac{(\ln x-m)^2}{2\sigma^2}\right\},\quad x>0$	$\Phi(\ln x\mid m,\sigma^2)$
L-3	$\frac{1}{(x-\tau)\sigma\sqrt{2\pi}}\exp\left\{-\frac{[\ln(x-\tau)-m]^2}{2\sigma^2}\right\},\quad x>\tau$	$\Phi(\ln(x-\tau)\mid m,\sigma^2)$
G-2	$\frac{x^{\alpha-1}\exp\{-x/\beta\}}{\beta^\alpha\Gamma(\alpha)},\quad \alpha,\beta>0,\ x>0$	$\frac{\Gamma_{x/\beta}(\alpha)}{\Gamma(\alpha)}$
G-3	$\frac{(x-\gamma)^{\alpha-1}\exp\left\{-\frac{x-\gamma}{\beta}\right\}}{\beta^\alpha\Gamma(\alpha)},\quad \alpha,\beta>0\ \ x>\gamma$	$\frac{\Gamma_{\frac{x-\gamma}{\beta}}(\alpha)}{\Gamma(\alpha)}$
C-2	$\frac{2}{\pi}\frac{\alpha\left(\frac{m}{x}\right)^\alpha}{x\left[1+\left(\frac{m}{x}\right)^{2\alpha}\right]},\quad x>0$	$1-\frac{2}{\pi}\operatorname{arc\,tg}\left(\frac{m}{x}\right)^\alpha$
L	$\frac{\alpha\beta x^{\beta-1}}{(1+\alpha x^\beta)^2},\quad x>0$	$1-\frac{1}{1+\alpha x^\beta}$

[1] A táblázatokban szereplő Gamma, ill. Béta-függvények definíciója:

Teljes Gamma-függvény: $\Gamma(p)=\int_0^\infty x^{p-1}e^{-x}dx,\quad p>0$

Nem teljes Gamma-függvény: $\Gamma_x(p)=\int_0^x t^{p-1}e^{-t}dt$

Teljes Béta-függvény: $B(p,q)=\int_0^\infty x^{p-1}(1-x)^{q-1}dx,\quad p,q>0$

Nem teljes Béta-függvény: $B_x(p,q)=\int_0^x t^{p-1}(1-t)^{q-1}dt$

$\Phi(x\mid m,\sigma^2)$ az m, σ paraméterű normális eloszlásfüggvényt jelöli.

2. táblázat

A vizsgált valószínűségeloszlások jellemzői

Eloszlás	μ_r	M	D^2	Me	Mo
L-2	$e^{rm+\frac{1}{2}\tau^2\sigma^2}$	$e^{m+\frac{1}{2}\sigma^2}$	$M^2(e^{\sigma^2}-1)$	e^m	$e^{m-\sigma^2}$
L-3	$e^{rm+\frac{1}{2}r^2\sigma^2}+\tau$	$e^{m+\frac{1}{2}\sigma^2}+\tau$	$(M-\tau)^2(e^{\sigma^2}-1)$	$e^m+\tau$	$e^{m-\sigma^2}+\tau$
G-2	$\beta^r\dfrac{\Gamma(\alpha+r)}{\Gamma(\alpha)}$	$\alpha\beta$	$\dfrac{M^2}{\alpha}$	$\dfrac{3\alpha-1}{3}\beta$ *	$\beta(\alpha-1)$
G-3	$\beta^r\dfrac{\Gamma(\alpha+r)}{\Gamma(\alpha)}+\gamma$	$\alpha\beta+\gamma$	$\dfrac{(M-\gamma)^2}{\alpha}$	$\dfrac{3\alpha-1}{3}\beta+\gamma$ *	$\beta(\alpha-1)+\gamma$
C-2	$\dfrac{m^r}{\cos\dfrac{r\pi}{2\alpha}},\quad r<\alpha$	$\dfrac{m}{\cos\dfrac{\pi}{2\alpha}}$	$M^2\left(\dfrac{1}{2}\cdot\dfrac{\operatorname{tg}\dfrac{\pi}{\alpha}}{\operatorname{tg}\dfrac{\pi}{2\alpha}}-1\right)$	m	$m\left(\dfrac{\alpha-1}{\alpha+1}\right)^{\frac{1}{2\alpha}}$
L	$\alpha^{-\frac{r}{\beta}}\dfrac{r\dfrac{\pi}{\beta}}{\sin(r\pi/\beta)},\quad r<\beta$	$\alpha^{-\frac{1}{\beta}}\dfrac{\pi/\beta}{\sin(\pi/\beta)}$	$M^2\left(\dfrac{\operatorname{tg}(\pi/\beta)}{\pi/\beta}-1\right)$	$\alpha^{-\frac{1}{\beta}}$	$\left(\dfrac{\beta-1}{\beta+1}\right)^{\frac{1}{\beta}}\cdot\alpha^{-\frac{1}{\beta}}$

* Közelítő formulák

A vizsgált eloszlások jellemzői közül a 2. táblázatban az alábbi jellemzőket adjuk meg: a nem centrális

$$\mu_r = \int_0^\infty x^r(x)\,dx$$

momentumokat, külön feltüntetve az $M = \mu_1$ várható értéket; a

$$D^2 = \int_0^\infty (x - M)^2 f(x) dx = \mu_2 - \mu_1^2$$

szórásnégyzetet, ami egyben a második centrális momentum; továbbá az Me-vel jelölt mediánt, valamint az Mo-val jelölt móduszt.

A 3. táblázat a szóban forgó eloszlások

$$F_1(x) = \frac{1}{\mu_1} \int_0^\infty t f(t)\, dt$$

ún. első momentum – eloszlásfüggvényét tartalmazza.[2]

3. táblázat

A vizsgált valószínűségeloszlások első momentum-eloszlásfüggvénye

Eloszlás	Első momentum-eloszlásfüggvény
L-2	$\Phi(\ln x \mid m + \sigma^2, \sigma^2)$
L-3	$\frac{\tau}{M}\,\Phi(\ln(x-\tau) \mid m, \sigma^2) + \frac{M-\tau}{M}\,\Phi(\ln(x-\tau) \mid m + \sigma^2, \sigma^2)$
G-2	$\frac{\Gamma_{x/\beta}(\alpha+1)}{\Gamma(\alpha+1)}$
G-3	$\frac{1}{\alpha\beta+\gamma}\;\frac{\beta\Gamma_{\frac{x-\gamma}{\beta}}(\alpha+1) + \gamma\Gamma_{\frac{x-\gamma}{\beta}}(\alpha)}{\Gamma(\alpha)}$
C-2	$\frac{1}{\pi}\cos\frac{\pi}{2\alpha}\,B_{\frac{1}{1+\left(\frac{m}{x}\right)^{2\alpha}}}\left(\frac{1}{2}+\frac{1}{2\alpha}, \frac{1}{2}-\frac{1}{2\alpha}\right)$
L	$\frac{1}{\pi/\beta}\sin\pi/\beta\left[B\left(1-\frac{1}{\beta}, 1+\frac{1}{\beta}\right) - B_{\frac{1}{1+\alpha x^\beta}}\left(1-\frac{1}{\beta}, 1+\frac{1}{\beta}\right)\right]$

[2] A 2. és 3. táblázatban található formulák többségének levezetése [8]-ban és [16]-ban található.

Az $F_1(x)$ első momentum-eloszlásfüggvény felhasználásával definiálható végül a

$$v = \frac{1 - F_1(M)}{1 - F(M)} \cdot \frac{F(M)}{F_1(M)}$$

egyenlőtlenségi mérőszám, melyet a hazai mikro jövedelemstatisztika rendszeresen használ és közöl. Belátható, hogy e mérőszám az M-nél magasabb, illetve alacsonyabb jövedelemmel rendelkező személyek átlagjövedelmének hányadosa. A részleteket illetően a [4] tanulmányra utalunk.

E jellemzők bemutatását azért tartjuk célszerűnek, mert egyrészt elég jól mutatják az eloszlás jellegét – a módusz pl. a sűrűségfüggvény maximum helyét adja –, másrészt mert az egyes eloszlásoknak az empirikus jövedelemeloszlásokhoz való illeszkedése értékelése során a jellemzők empirikus adatokból számított értékeit összevetettük az egyes eloszlások becsült paraméterei alapján meghatározható értékekkel.

A 2. táblázatból látható, hogy az m paraméter a C-2 eloszlásnál épp a medián, az L-2 eloszlásnál pedig a medián logaritmusa, továbbá L-2 σ paramétere, a G-2 és C-2 α paramétere és az L β paramétere egyértelműen meghatározzák a relatív szórást (illetve általánosabban az eloszlás egyenlőtlenségét).

A 2. táblázatban közölt jellemzők alapján egyszerűen meghatározhatók az egyes eloszlások aszimmetriáját, illetve csúcsosságát mérő

$$\gamma_1 = \frac{\mu_3 - 3M\mu_2 + 2M^3}{D^3},$$

illetve

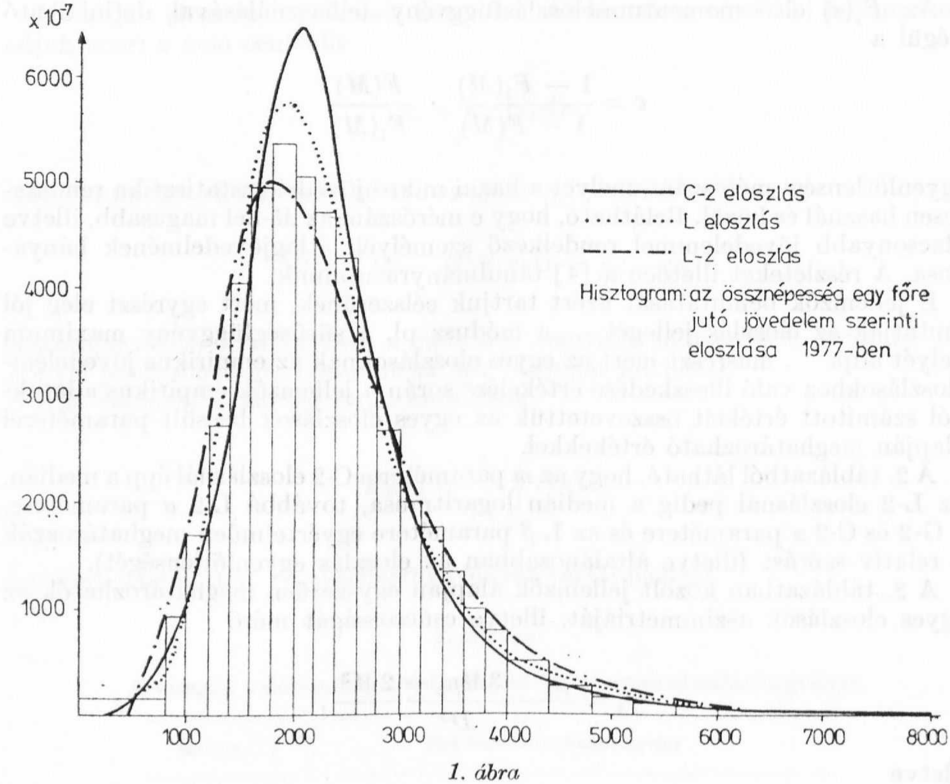
$$\gamma_2 = \frac{\mu_4 - 4M\mu_3 + 6M^2\mu_2 - 3M^4}{D^4} - 3$$

együtthatók is, amelyek jó szolgálatot nyújtanak az egyes sűrűségfüggvények alakjának jellemzéséhez. γ_1 pozitív értéke bal oldali, negatív értéke jobb oldali aszimmetriára utal, γ_2 pedig aszerint pozitív vagy negatív, hogy az eloszlás csúcsosabb vagy lapultabb-e a standard normális eloszlásnál.

Talán nem érdektelen ebből a szempontból összehasonlítani a jól ismert L-2 eloszlás, illetve a kevésbé ismert C-2 és L eloszlások sűrűségfüggvényeit, annál is inkább, mert az utóbbi a továbbiakban fontos szerepet játszik. Az 1. ábra mutatja a három sűrűségfüggvényt azonos – az össznépesség 1977. évi jövedelemeloszlásának megfelelő – várható érték és szórás mellett. Összehasonlításul a szóban forgó empirikus eloszlás hisztogramját is berajzoltuk az ábrába.

Látható, hogy a – bal oldali – aszimmetria, de méginkább a csúcsosság tekintetében jelentős különbség van a három eloszlás között. Leginkább aszimmetrikus és csúcsos a C-2 eloszlás, legkevésbé az L-2 eloszlás, az L eloszlás mindkét tekintetben közepes helyet foglal el. A 4. táblázatban megadjuk a három ábrázolt eloszlás néhány jellemzőjét.

Látható, hogy különösen γ_2 igen érzékeny mérőszám, a csúcsosságokban az ábrán mutatkozó nem túl nagy eltéréseket erősen felnagyítva mutatja



1. ábra

4. táblázat

Az ábrázolt három eloszlás néhány jellemzője

Eloszlás	γ_1	γ_2	Me	Mo
L-2	1,316	3,232	2154,8	1838,2
L	2,694	34,345	2170,1	1987,2
C-2	3,798	132,561	2178,5	2061,2

Általánosságban is kimutatható, hogy azonos várható értékű és szórású L-2 és L eloszlások közül mindig az utóbbi mutatja a nagyobb aszimmetriát és csúcsosságot. L és L-2 aszimmetriája közötti különbség ugyanis

$$\frac{M^3}{D^3}\cdot\frac{\operatorname{tg}^2(\pi/\beta)}{(\pi/\beta)^2}\left[\frac{1}{1-\frac{1}{3}\operatorname{tg}^2(\pi/\beta)}-\frac{\operatorname{tg}(\pi/\beta)}{\pi/\beta}\right]$$

alakban írható, ami $\beta > 3$ esetén mindig pozitív. Hasonlóképpen megmutatható az is, hogy a két eloszlás csúcsosságának a különbsége is mindig pozitív

$\beta > 4$ esetén. Mindkét különbség értéke annál nagyobb, minél közelebb van β 3-hoz, illetve 4-hez. $\beta \to \infty$ esetén pedig mindkét különbség 0-hoz tart, azaz az L eloszlás alakja annál jobban közelíti az ugyanolyan várható értékű és szórású L-2 eloszlását, minél nagyobb L β paraméterének értéke.

1.2 *A paraméterek becslése*

Mint a bevezetésben már jeleztük, vizsgálataink adatbázisát a KSH által 1963 óta rendszeresen végrehajtott reprezentatív jövedelmi felmérések 1967-re, 1972-re és 1977-re vonatkozó adatai képezték. Mint ismeretes, e felmérések adatai igen jól egyeznek a megfelelő makrostatisztikai adatokkal, s néhány kisebb módosítást, pontosítást leszámítva lényegében azonos módon és technikával készültek. Ily módon kitűnő adatbázist nyújtanak a jövedelemeloszlások dinamikus vizsgálatára.

Kiemelkedő fontossága miatt az egy főre jutó személyes rendelkezésű jövedelmek eloszlásainak időbeli alakulását tettük vizsgálat tárgyává a gazdaságilag aktív, az inaktív és a teljes népesség körében. E két népességcsoport jövedelemeloszlásainak egymástól elkülönített vizsgálatát az a tény teszi indokolttá, hogy nagymértékben különböznek egymástól az aktív és inaktív népesség jövedelmi forrásai és a jövedelmeik egyenlőtlenségét előidéző tényezők is, ami feltehetően eltérő jellegű jövedelemeloszlásokhoz vezet.

Ismeretes, hogy egy empirikus eloszláshoz illeszkedő valószínűségeloszlás paramétereit többféle eljárással is lehet becsülni, részben attól függően, hogy milyen kritérium alapján bíráljuk el az illeszkedés jóságát. Ha ugyanahhoz az empirikus eloszláshoz többféle eloszlásfüggvényt illesztünk, elvileg olyan paraméterbecslő eljárást kell alkalmazni, amely ugyanazon kritérium alapján ad optimális becslést a vizsgált eloszlásfüggvények ismeretlen paramétereire. Ugyanis csak ezúton válik a különböző eloszlások illeszkedése összemérhetővé. Ideális az lett volna, ha minden esetben a paraméterek ún. *maximum likelihood becslését* tudtuk volna meghatározni. A vizsgált eloszlásfüggvények többségénél azonban a maximum likelihood becslés meghatározása még egyedi adatok használata esetén is elég bonyolult, csoportosított adatokra pedig — a jövedelemeloszlások osztályközös gyakoriságok formájában álltak rendelkezésünkre — elméletileg sem teljesen megoldott. Egyedül az L-2 eloszlásra volt meg — 1972-re és 1977-re — a paraméterek maximum likelihood becslése, ami ebben az esetben az egyedi jövedelmek logaritmusainak átlaga, illetve szórásnégyzete.

Ezért általában más becslési eljárásokat alkalmaztunk, így pl. esetenként az ún. *momentum módszert*. Ennek során annyi

$$m_r = \frac{1}{n}\sum_{i=1}^{k} n_i x_i^r, \qquad n = \sum_{i=1}^{k} n_i, \qquad r = 1, 2, \ldots$$

mintabeli momentumot, ahány becsülendő paraméter van, egyenlővé teszünk a μ_r elméleti momentumokkal, s az így adódó egyenleteket megoldva a μ_r-ekben szereplő paraméterekre, kapjuk a paraméterek momentum-becsléseit. m_r formulájában x_i az i-edik háztartás egy főre jutó jövedelme, n_i pedig a taglétszáma

Gyakran alkalmazzák a paraméterek becslésére a *minimális χ^2 módszert* is amely azon Θ- értékeket tekinti az ismeretlen Θ paraméter-vektor becslésének amelyre a

$$\chi^2(\Theta) = n \sum_{i=1}^{k} \frac{\left[\frac{n_i}{n} - p_i(\Theta)\right]^2}{p_i(\Theta)} \tag{1.1}$$

kifejezés minimális Itt

$\frac{n_i}{n}$ — az i-edik jövedelmi kategóriába tartozó személyek relatív gyakorisága

$p_i(\Theta)$ — az i-edik jövedelmi kategóriába esés valószínűsége az illesztett eloszlásfüggvény alapján.

De becsülhetők az ismeretlen paraméterek a hisztogramhoz, illetve az empirikus eloszlásfüggvényhez a *legkisebb négyzetek módszerével* illesztett sűrűség-, illetve eloszlásfüggvény alapján is. Előbbi esetben — egyenletes osztályközöket feltételezve — a

$$\Delta(\Theta) = \sum_{i=1}^{k} \left[\frac{n_i}{n} - p_i(\Theta)\right]^2 \tag{1.2}$$

négyzetösszeget, az utóbbi esetben pedig az

$$S(\Theta) = \sum_{i=1}^{k-1} [F_n(x^i) - F(x^i, \Theta)]^2 \tag{1.3}$$

négyzetösszeget kell minimalizálni, ahol

$$F_n(x^i) = \sum_{j=1}^{i} \frac{n_j}{n},$$

x^i pedig az i-edik jövedelem kategória felső határa.

Megjegyzendő, hogy a különböző becslési eljárások másképp illesztik a szóban forgó valószínűségeloszlást az empirikus adatokhoz, s így természetesen más-más becsléseket adnak a paraméterekre. Így pl. a minimális χ^2 módszer — mivel relatív eltéréseket tekint — elsősorban az eloszlás szélein, a kis relatív gyakoriságú kategóriákhoz igyekszik jó illeszkedést biztosítani. Ezzel szemben $S(\Theta)$ minimalizálásakor inkább az eloszlás közepén — ott, ahol az eloszlásfüggvény meredeken emelkedik — lesz jobb az illeszkedés, $\Delta(\Theta)$ minimalizálása során ugyanakkor egyformán értékelődnek a sűrűségfüggvénynek a hisztogramtól vett eltérései az eloszlás különböző részein.

Az esetek túlnyomó többségében mindegyik felsorolt becslési eljárás nemlineáris, explicite nem megoldható egyenletrendszerekhez vezet. Legegyszerűbb a momentum-becslések meghatározása, különösen akkor, ha csak két paramétert kell becsülni. Ez utóbbi esetben, ha egyedi — és nem csoportosított — adatokból számított mintabeli átlag és szórásnégyzet áll rendelkezésre, a momentum módszer elég megbízható becsléseket szolgáltat a keresett két paraméterre. Ezt tapasztalataink is alátámasztották. 1972-re és 1977-re ugyanis meg tudtuk határozni mindhárom vizsgált eloszlás esetén az egyedi adatokból számolt momentumokon alapuló becsléseket az összes tekintett két-paraméteres eloszlásfüggvényre, és ezek — összevetve más módszerekkel nyert becslésekkel — elég hatásosaknak bizonyultak, különösen az L eloszlásoknál. A háromparaméteres eloszlások (L-3, G-3) esetén ezzel szemben a momentum módszer meglehetősen rossz hatásfokú becsléseket eredményezett, nem elsősorban azért, mert a harmadik mintabeli momentumot csak csoportosított adatokból tudtuk számítani, hanem inkább amiatt, mert a harmadik paraméternek — küszöb paraméter lévén — nem sok köze van a harmadik momentumhoz.

Mint említettük, az elbírálás egységességét, az eredmények összehasonlíthatóságát az biztosítja, ha mindegyik évre, mindhárom rétegre és valamennyi figyelembe vett eloszlásfüggvényre azonos módszerrel történik a paraméterek becslése. Ilyen egységes módszernek kínálkozott az empirikus eloszlásokhoz legkisebb négyzetek módszerével illesztett eloszlásfüggvények alkalmazása, mivel erre, mint nem lineáris regresszióra kész programcsomag állt rendelkezésre. Az iterációhoz a kezdő értékeket általában a momentum módszerrel nyert becslések szolgáltatták. Kivételt képezett az L eloszlás, amelynél — mint erről könnyűszerrel meggyőződhetünk — az

$$\ln \frac{F(x)}{1 - F(x)} = \ln \alpha + \beta \ln x \tag{1.4}$$

összefüggés alapján egyszerűen nyerhetők regressziós becslések az α és β paraméterekre.

Az (1.3) kifejezés minimálizálása néhány esetben nem szolgáltatott elfogadható becslést a keresett paraméterekre. Ez volt a helyzet az L-eloszlással 1967-re és 1972-re, feltehetően több lokális szélsőérték létezése miatt. A G-3 eloszlás esetén viszont 1972-re és 1977-re nem kaptunk megfelelő becsléseket, itt elsősorban a túl magasra becsült küszöb-paraméter okozott problémát. Az L-eloszlás paramétereire ezekben az esetekben a fenti (1.4) összefüggésből adódó lineáris regressziós becsléseket a G-3 eloszlás esetén pedig a momentum becsléseket fogadtuk el.

Az egyes eloszlások paramétereire végül elfogadott becsléseket az 5. táblázat tartalmazza. Megjegyezzük, hogy az illeszkedés-vizsgálatokhoz a táblázatban közöltnél több tizedesjegyet használtunk.

1.3 *A vizsgált valószínűségeloszlások illeszkedése*

A vizsgált valószínűségeloszlások becsült paramétereit felhasználva minden esetben meghatároztuk az egyes empirikus jövedelemeloszlásokat közelítő elméleti valószínűségeloszlásokat, s összevetettük azokat a megfelelő empirikus jövedelemeloszlásokkal. Ez az összevetés többféleképpen is megtehető.

Elvben természetesen ugyanazt a kritériumot kellene alapul venni az összehasonlítások során is, mint amelyek szerint optimalizáltunk akkor, amikor a vizsgált valószínűségeloszlások ismeretlen paramétereit becsültük. A jelen esetben tehát elméletileg az (1.3) mérőszám szerinti összehasonlítás volna indokolt, mert a paraméterek becslése a legtöbb esetben $S(\Theta)$ minimalizálása útján történt. Úgy véljük azonban, hogy gyakorlati szempontból elősorban az illeszkedés relatív mértéke bír jelentőséggel, ami például az (1.1)-ben definiált χ^2-értékek alapján történő összehasonlítást indokolná. A szokatlanul nagy mintaelemszámok miatt viszont a szokásos χ^2-próbás illeszkedésvizsgálat még viszonylag kis eltérések esetén is általában szignifikáns eltérést jelzett volna az empirikus és az azokhoz illesztett eloszlások között. Ezért inkább az leíró mérőszámok használata mellett döntöttünk. Az (1.5) mérőszámok ugyanis

5. *táblázat*

A vizsgált valószínűségeloszlások becsült paraméterei

Eloszlás-paraméter	1967			1972			1977		
	Aktív	Inaktív	Összes	Aktív	Inaktív	Összes	Aktív	Inaktív	Összes
	népesség								
L-2 eloszlás									
$\hat{m}$	6,9934	6,6247	6,9625	7,3249	6,8671	7,2864	7,7198	7,4493	7,6831
$\hat{\sigma}$	0,3853	0,3967	0,3983	0,3866	0,4648	0,4102	0,3584	0,3994	0,3734
L-3 eloszlás									
$\hat{m}$	7,2238	6,7261	7,2354	7,4606	6,8328	7,5588	7,7444	7,4107	7,7964
$\hat{\sigma}$	0,3106	0,3638	0,3092	0,3383	0,4792	0,3143	0,3501	0,4139	0,3352
$\hat{\tau}$	−275,1	−79,3	−323,1	−215,1	31,3	−446,4	−54,5	62,6	−253,6
G-2 eloszlás									
$\hat{\alpha}$	6,7123	5,3833	6,0745	5,7973	4,5045	5,4093	6,9099	5,8450	6,4523
$\hat{\beta}$	172,6	148,9	185,8	280,2	232,3	290,4	345,7	315,0	357,5
G-3 eloszlás									
$\hat{\alpha}$	3,8070	2,7862	3,7263	2,693	0,958	2,823	1,782	1,387	1,868
$\hat{\beta}$	232,7	195,2	236,5	426,1	584,1	420,1	715,5	697,2	696,5
$\hat{\gamma}$	282,4	272,4	256,2	495,3	522,5	400,3	1137,4	906,3	1031,0
C-2 eloszlás									
$\hat{m}$	1090,2	753,6	1057,0	1519,1	962,8	1462,3	2253,7	1719,3	2173,1
$\hat{\alpha}$	3,6403	3,5664	3,5193	3,5565	3,0259	3,3441	3,9048	3,5361	3,7532
L-eloszlás									
$10^{14} \cdot \hat{\alpha}$	0,6242	1,5095	1,3697	5,8949	414,8642	28,7873	0,0146	1,6272	0,0722
$\hat{\beta}$	4,6859	4,7941	4,5931	4,1671	3,8118	3,9743	4,7229	4,2620	4,5376

$$\overline{RH}_a = \frac{1}{k_1}\sum_{j=1}^{k_1}\frac{\left|\frac{n_j}{n} - p_j\right|}{\frac{n_j}{n}}; \quad \overline{RH}_k = \frac{1}{k_2}\sum_{j=k_1+1}^{k_1+k_2}\frac{\left|\frac{n_j}{n} - p_j\right|}{\frac{n_j}{n}}$$

$$\overline{RH}_m = \frac{1}{k_3}\sum_{j=k_1+k_2+1}^{n}\frac{\left|\frac{n_j}{n} - p_j\right|}{\frac{n_j}{n}}; \quad \overline{RH} = \frac{1}{k}\sum_{j=1}^{k}\frac{\left|\frac{n_j}{n} - p_j\right|}{\frac{n_j}{n}} \tag{1.5}$$

azzal az igen előnyös tulajdonsággal rendelkeznek, hogy azt az átlagos relatív becslési hibát mutatják az eloszlás alsó, középső és felső szakaszára, illetve egészére nézve, amit a tényleges jövedelemeloszlásnak egy becsült eloszlással való közelítésekor követünk el. Ennek megfelelően k_1 az eloszlás alsó, k_2 az eloszlás középső, k_3 pedig az eloszlás felső szakaszába tartozó jövedelemkategóriák számát jelöli. Az eloszlások három szakaszát oly módon definiáltuk, hogy a személyek aránya minden esetben megközelítően 30, 60, illetve 10% legyen. Ez a szakaszokra bontás önkényes ugyan, de a jövedelemkategóriák adott száma és határai mellett nem találtunk jobb megoldást. Az (1.5) mutatók értékeit a 6. táblázat tartalmazza az egyes eloszlások alsó és középső szakaszának felső határaival együtt.

Az illeszkedés három szakaszon történő vizsgálatát azért tartottuk indokoltnak, mert kérdéses volt, hogy az eloszlások egyes szakaszainak közelítésére ugyanaz a valószínűségeloszlás alkalmas-e, illetve az egyes szakaszok egyformán közelíthetők-e. Az eredmények igazolták eljárásunkat, ugyanis az eloszlások alsó és középső szakasza általában a felső szakasznál lényegesen jobban közelíthető, s a különböző szakaszok közelítésére általában nem ugyanazok az eloszlások a legalkalmasabbak.

A teljes eloszláshoz való illeszkedést vizsgálva jól látható, hogy az utóbbi tíz év folyamán az *aktív* és *összes* népesség jövedelemeloszlásának típusában érdekes változás állt be. Míg ugyanis 1967-ben e két népességcsoport jövedelemeloszlása az L-3 eloszlással volt a legjobban közelíthető, addig 1977-ben a két szóban forgó jövedelemeloszlást már az L eloszlás írta le a legjobban. Az 1972-es év e szempontból érdekes átmenetet képez 1967 és 1977 között, mert ekkor az aktív népesség esetén a C-2 eloszlás tűnik a legalkalmasabbnak a jövedelemeloszlás közelítésére, az össznépességnél pedig az L-3 és C-2 eloszlások lényegében azonos mértékben a legmegfelelőbbek. Ez annál is inkább érdekes, mert a C-2 eloszlás 1967-ben még egyértelműen a legkevésbé használhatónak bizonyult. Az *inaktív* népesség jövedelemeloszlásának típusa ezzel szemben időben jóval stabilabbnak tűnik, mert az L-2 eloszlás mindhárom évben igen jól közelíti a szóban forgó tényleges eloszlást. Igaz ugyan, hogy 1972-ben az L-3 és L eloszlás, 1977-ben pedig az L-3 eloszlás az L-2-nél valamivel jobb eredményt ad, de a közöttük fennálló különbség minimális. Említést érdemel még az is, hogy a más vizsgálatok [14] során jól használhatónak tűnő G-2 és G-3 eloszlás nálunk 1972-ben és 1977-ben a legkevésbé alkalmasnak tűnt. Ez valószínűleg a G-3 eloszlás küszöb-paraméterének nem kielégítő becslésével is kapcsolatos.

6. táblázat

A vizsgált valószínűségeloszlások illeszkedése az empirikus jövedelemeloszlások különböző szakaszain

Eloszlás	1967			1972			1977		
	Aktív	Inaktív	Összes	Aktív	Inaktív	Összes	Aktív	Inaktív	Összes
	népesség			népesség			népesség		
	Alacsony jövedelmek								
Felső határ	900	700	900	1200	800	1200	1800	1400	1800
L-2	0,1291	0,0268	0,1334	0,2780	0,0179	0,2718	0,2602	0,1170	0,2186
L-3	0,0454	0,0107	0,0332	0,2173	0,0581	0,1320	0,2562	0,1310	0,1640
G-2	0,0258	0,0870	0,0309	0,1905	0,3026	0,1001	0,2672	0,2238	0,1752
G-3	0,1259	0,0287	0,1103	0,4018	0,4961	0,3832	0,6333	0,4642	0,5189
C-2	0,0647	0,0853	0,0746	0,0508	0,2060	0,0513	0,1067	0,1097	0,0205
L	0,1040	0,1575	0,1311	0,1291	0,0744	0,1331	0,1212	0,0656	0,0663
	Közepes jövedelmek								
Felső határ	1800	1300	1800	2400	1700	2400	3600	3000	3600
L-2	0,0469	0,0951	0,0413	0,0256	0,0645	0,0467	0,0320	0,0860	0,0267
L-3	0,0312	0,0932	0,0307	0,0268	0,0598	0,0303	0,0315	0,0845	0,0379
G-2	0,0454	0,1129	0,0477	0,0917	0,0973	0,0791	0,0973	0,1350	0,0921
G-3	0,0370	0,0943	0,0311	0,0903	0,1402	0,0913	0,0826	0,1195	0,0787
C-2	0,1057	0,1412	0,1146	0,0524	0,0858	0,0740	0,0666	0,0578	0,0822
L	0,1116	0,1336	0,1228	0,0561	0,0808	0,0786	0,0386	0,0504	0,0489
	Magas jövedelmek								
L-2	0,1521	0,2217	0,1721	0,0690	0,1865	0,0739	0,1366	0,2061	0,1343
L-3	0,1173	0,3104	0,1231	0,0914	0,1579	0,0993	0,1366	0,1978	0,1489
G-2	0,1629	0,4077	0,1934	0,1846	0,3256	0,1712	0,2226	0,4300	0,2708
G-3	0,1815	0,3473	0,1463	0,1636	0,2340	0,0575	0,2272	0,1914	0,2036
C-2	0,3479	0,9541	0,4103	0,1345	0,1859	0,1671	0,2827	0,4801	0,3520
L	0,1855	0,3585	0,1598	0,0998	0,1362	0,1222	0,1577	0,3415	0,1657
	Összesen								
L-2	0,1049	0,1670	0,1115	0,1180	0,1101	0,1269	0,1279	0,1488	0,1142
L-3	0,0690	0,2198	0,0690	0,1026	0,1016	0,0774	0,1267	0,1468	0,1094
G-2	0,0899	0,2930	0,1042	0,1423	0,2245	0,1037	0,1852	0,2916	0,1759
G-3	0,1123	0,2447	0,0926	0,2081	0,2312	0,1822	0,2766	0,2138	0,2370
C-2	0,1973	0,6391	0,2283	0,0675	0,1459	0,0841	0,1522	0,2688	0,1600
L	0,1404	0,2751	0,1394	0,0888	0,1036	0,1051	0,1016	0,1923	0,0941

Ami a szakaszonkénti illeszkedést illeti, csak egy olyan eset van (az összes népesség 1967. évi jövedelemeloszlása), amikor ugyanaz az eloszlástípus bizonyul a legjobbnak a tényleges eloszlás mindhárom szakaszán, mint a teljes eloszlásra nézve. Általánosságban annyit mondhatunk, hogy az eloszlások *középső-* és *felső* szakaszának közelítésére legtöbbször az L-2 és L-3 eloszlás bizonyult a legjobbnak. Az *alsó* szakasz tekintetében már nem ilyen egyöntetű a kép. Az aktív és összes népesség esetében 1967-ben a G-2 eloszlás, 1972-ben és 1977-ben pedig a C-2 eloszlás bizonyult a legjobbnak e szakaszon. Az inaktív népességnél ezzel szemben 1967-ben és 1972-ben az L-2 és L-3 eloszlás, 1977-ben pedig az L eloszlás.

7. táblázat

A jövedelemeloszlások tényleges és az illesztett valószínűségeloszlások becsült paraméterei alapján számított jellemzői

Eloszlás	1967			1972			1977		
	Aktív	Inaktív	Összes	Aktív	Inaktív	Összes	Aktív	Inaktív	Összes
	népesség			népesség			népesség		
					Medián				
Egyedi tényleges	—	760,0	1069,0	1525,0	965,0	1476,0	2252,0	1733,0	2186,0
Csoportosított tényleges	1102,0	760,0	1069,0	1525,0	965,0	1476,0	2256,6	1722,6	2180,2
L-2	1089,4	753,5	1056,3	1517,6	960,1	1460,3	2252,6	1718,6	2171,2
L-3	1096,6	754,6	1064,7	1523,1	959,0	1471,2	2254,1	1716,2	2178,4
G-2	1100,7	752,1	1066,4	1530,8	968,9	1474,0	2273,4	1736,2	2187,8
G-3	1090,8	751,2	1058,5	1500,9	887,4	1446,1	2174,0	1640,9	2100,0
C-2	1090,2	753,6	1057,0	1519,1	962,8	1462,3	2253,7	1719,3	2173,1
L	1074,9	763,7	1043,0	1495,4	968,3	1430,4	2253,2	1719,0	2172,4
					Relatív szórás				
Egyedi tényleges	—	—	—	0,455	0,547	0,474	0,397	0,482	0,415
Csoportosított tényleges	0,394	0,425	0,414	0,426	0,528	0,445	0,396	0,438	0,408
L-2	0,400	0,413	0,414	0,402	0,491	0,428	0,370	0,416	0,387
L-3	0,393	0,413	0,407	0,394	0,493	0,414	0,369	0,418	0,383
G-2	0,386	0,431	0,406	0,415	0,471	0,430	0,380	0,414	0,394
G-3	0,389	0,399	0,401	0,426	0,528	0,445	0,396	0,438	0,408
C-2	0,519	0,534	0,545	0,537	0,696	0,589	0,470	0,541	0,497
L	0,428	0,416	0,438	0,495	0,558	0,527	0,423	0,481	0,445
					v egyenlőtlenségi mutató				
Egyedi tényleges	1,877	1,929	1,916	1,892	2,142	1,958	1,792	1,923	1,835
L-2	1,851	1,886	1,891	1,855	2,104	1,927	1,773	1,894	1,816
L-3	1,851	1,895	1,892	1,847	2,104	1,912	1,773	1,895	1,813
G-2	1,857	1,999	1,925	1,954	2,093	1,996	1,835	1,949	1,885
G-3	1,834	1,860	1,866	1,942	2,200	2,006	1,838	1,949	1,875
C-2	1,937	1,959	1,973	1,962	2,119	2,025	1,858	1,968	1,903
L	1,801	1,826	1,834	1,960	2,105	2,020	1,812	1,947	1,866

Végül megjegyezzük még, hogy az aktív és összes népesség jövedelemeloszlása 1967-ben és 1972-ben kissé jobban volt közelíthető, mint 1977-ben.

Érdemes ezután még az empirikus eloszlásokból közvetlenül, illetve az azok közelítésére használt eloszlások paramétereinek felhasználásával számított néhány jellemzőt összevetni egymással. Választásunk a mediánra, relatív szórásra és a v egyenlőtlenségi mutatóra — mint a jövedelemeloszlások fontos jellemzőire — esett.

Az eredményeket a 7. táblázat tartalmazza. A táblázat adataiból látható, hogy a mediánt egy-két esettől eltekintve kielégítően közelítik a valószínűségeloszlások becsült paramétereiből számítható medián-értékek. Nem ez a helyzet viszont a relatív szórással és a v egyenlőtlenségi mutatóval. A vizsgált valószínűségeloszlások többsége ugyanis általában alulbecsli még a csoportosított adatok alapján becsült relatív szórást is, ami köztudottan alacsonyabb az

8. táblázat

A tényleges jövedelemeloszlások és az azokhoz legjobban illeszkedő valószínűségeloszlások, 1977

Egy főre jutó havi jövedelem, Ft	Aktív		Inaktív		Összes	
	népesség					
	tényleges	becsült[a]	tényleges	becsült[b]	tényleges	becsült[c]
– 800	1,2	0,8	3,5	2,8	1,5	1,1
800 – 1000	1,3	1,4	5,1	6,0	1,9	1,8
1000 – 1200	2,3	2,7	9,9	9,7	3,4	3,5
1200 – 1400	4,4	4,7	11,2	12,0	5,4	5,6
1400 – 1600	7,4	7,0	12,7	12,5	8,1	8,0
1600 – 1800	9,4	9,2	12,4	11,7	9,9	9,9
1800 – 2000	10,8	10,6	10,5	10,2	10,7	10,9
2000 – 2200	10,4	10,9	8,5	8,4	10,1	10,7
2200 – 2400	9,9	10,2	6,1	6,7	9,4	9,7
2400 – 2600	9,0	8,9	5,9	5,1	8,6	8,2
2600 – 2800	7,2	7,3	3,3	3,9	6,7	6,6
2800 – 3000	5,8	5,8	2,5	2,9	5,4	5,3
3000 – 3200	4,4	4,5	1,7	2,2	4,0	4,1
3200 – 3400	3,9	3,5	1,6	1,6	3,5	3,1
3400 – 3600	3,0	2,7	1,2	1,2	2,7	2,4
3600 – 3800	2,2	2,0	1,1	0,8	2,0	1,9
3800 – 4000	1,8	1,6	0,6	0,6	1,6	1,4
4000 – 4400	2,1	2,2	0,8	0,8	2,0	2,0
4400 – 4800	1,3	1,3	0,5	0,4	1,1	1,2
4800 – 5200	0,8	0,8	0,3	0,2	0,7	0,8
5200 – 5600	0,4	0,5	0,3	0,1	0,4	0,5
5600 – 6000	0,3	0,4	0,1	0,1	0,3	0,3
6000 –	0,7	1,0	0,2	0,1	0,6	1,0
Összesen	100,0	100,0	100,0	100,0	100,0	100,0
A megfigyelt személyek száma	41 421		6940		48 361	

a) L-eloszlás: $\hat{\alpha} = 1{,}4626419 \cdot 10^{-16}$, $\hat{\beta} = 4{,}7228700$
b) L-2 eloszlás: $\hat{m} = 7{,}4492505$, $\hat{\sigma} = 0{,}3993919$
c) L-eloszlás: $\hat{\alpha} = 7{,}2191982 \cdot 10^{-16}$, $\hat{\beta} = 4{,}5375546$

egyedi adatokból számított relatív szórásnál. Ezzel szemben az L, de különösen a C-2 eloszlás legtöbbször még az egyedi adatokból számított relatív szórást is felülbecsli, a C-2 eloszlás nemegyszer igen lényegesen. Ami a v egyenlőtlenségi mutatót illeti, lényegében itt is hasonló a helyzet. Azzal az eltéréssel, hogy 1967-ben az L-eloszlás feltételezése is a v egyenlőtlenségi mutató alulbecsléséhez vezet, 1972-ben és 1977-ben viszont a G-2 és G-3 eloszlás feltételezése is felülbecsli v-t. Ebből adódóan aztán nem mindig azoknál a valószínűségeloszlásoknál egyeznek a legjobban a becsült paraméterek alapján számított jellemzők a közvetlenül számítottakkal, amelyek az (1.5) mutató szerint a legjobban illeszkednek az empirikus eloszlásokhoz.

Mindent egybevetve úgy tűnik, hogy a jövedelemeloszlásokat közelítő valószínűségeloszlások típusában az utóbbi évtizedben változás állt be. Ez a változás azonban csak az aktív és összes népesség jövedelemeloszlásának típusát érintette, az inaktív népességét nem. *Eszerint az aktív és összes népesség jövedelemeloszlásának típusa az utóbbi tíz évben L-3-ról L-re változott, míg az inaktív népességé továbbra is jól közelíthető L-2 eloszlással.* Érdemes megjegyezni azonban, hogy az L-3, ill. L-2 eloszlás feltételezése az aktív és összes népesség esetében továbbra is megengedhető, mert e két eloszlás csak a jövedelemeloszlás *alsó* részében illeszkedik lényegesen rosszabbul az empirikus adatokhoz, mint az L eloszlás. Ennek hatása azonban az eloszlás egészére vonatkozó $\overline{RH}$ mérőszám értékére nem túl nagy.

Azt a tényt, hogy a vizsgált jövedelemeloszlások mindegyike jól leírható valamelyik tekintett kétparaméteres valószínűségeloszlással, a 2. részben fogjuk kihasználni.

Végül a 8. táblázatban megadjuk az 1977. évi empirikus eloszlásokat, és az azokhoz legjobban illeszkedő valószínűségeloszlásokat.

2. A jövedelemeloszlások prognosztizálása

Ideálisan a jövedelemeloszlások prognosztizálása olyan szimulációs modellek kidolgozását igényelné, amelyek nemcsak egyszerűen a jövedelemeloszlások előrejelzésére alkalmasak, hanem felhasználhatók annak meghatározására is, hogy egy adott jövedelemeloszlás elérése érdekében milyen konkrét gazdaságpolitikai vagy társadalompolitikai intézkedések meghozatalára van szükség. Az ilyen modellek kidolgozása azonban igen sok adatot, időt és munkát igényel, ezért gyakorlati megvalósításukra — tudomásunk szerint — eddig még nem került sor.

Úgy véljük azonban, hogy ha a jövedelmek adott típusú eloszlását tételezzük fel a jövőben, s az eloszlásfüggvényben szereplő ismeretlen paramétereket nem mechanikusan extrapoláljuk, hanem különféle gazdaság- és társadalompolitikai feltételezések alapján becsüljük, akkor az ideálishoz közel eső eredményekre juthatunk. Cikkünk e részében egy ilyen prognosztizálási eljárást ismertetünk és alkalmazunk a gyakorlatban.

2.1 *A prognosztizálás módszere*

Az eljárás alkalmazásának az az előfeltétele, hogy a jövedelmek eloszlása a prognosztizált időszakban *kétparaméteres valószínűségeloszlás*sal legyen közelíthető. Az 1. pontban ismertetett vizsgálataink alapján ez a feltevés rövid távon mindenképpen, de esetleg még hosszú távon is reálisnak tűnik.

Az a tény, hogy a jövedelmek eloszlása két paramétert tartalmazó eloszlás segítségével kielégítően közelíthető, egyben azt is jelenti, hogy a tervezett átlagos jövedelem (M_t) és a *jövedelmek relatív szórása* prognosztizált értékének (V_p) ismeretében már egyértelműen meghatározhatók az alapulvett valószínűségeloszlás ismeretlen paraméterei, majd ezek alapján az egész jövedelemeloszlás is.

Ha például a jövedelmek eloszlása a jövőben L-2 eloszlással közelíthető, akkor a

$$\sigma_p = \sqrt{\ln (V_p^2 + 1)} \quad \text{és} \quad m_p = \ln M_t - \frac{1}{2}\sigma_p^2 \tag{2.1}$$

formulákból, L eloszlás feltételezése esetén pedig az

$$\alpha_p = \left(\frac{\pi/\beta_p}{M_t \sin(\pi/\beta_p)} \right)^{\beta_p} \tag{2.2}$$

összefüggés alapján adódnak az ismeretlen paraméterek, ahol β_p a

$$\frac{\operatorname{tg}(\pi/\beta_p)}{\pi/\beta_p} = V_p^2 + 1 \tag{2.3}$$

egyenlet megoldása.[3]

A tervezett M_t átlagos jövedelmi szint az egy főre jutó reáljövedelem tervezett növekedési üteme és bázisidőszaki színvonala alapján egyértelműen meghatározható. A relatív szórás V_p prognosztizált értéke ezzel szemben a szórásnégyzet-felbontás néven ismert eljárással kapott eredmények bizonyos időbeli stabilitása alapján tűnik meghatározhatónak.

Induljunk ki ezért a szórásnégyzet jól ismert

$$S^2 = S_b^2 + S_k^2 \tag{2.4}$$

felbontásából, ahol

$S^2 = \frac{1}{n} \sum_{=1}^{n} (x_i - \overline{x})^2$ — a jövedelmek *teljes* szórásnégyzete,

$S_b^2 = \frac{1}{n} \sum_{l=1}^{L} n_l S_l^2$ — a jövedelmek *belső* szórásnégyzete,

$S_k^2 = \frac{1}{n} \sum_{l=1}^{L} n_l (\overline{x}_l - \overline{x})^2$ — a jövedelmek *külső* szórásnégyzete,

$\overline{x}_l$, illetve $\overline{x}$ — az l-edik részsokaságra, ill. az egész sokaságra vonatkozó átlagos egy főre jutó jövedelem,

S_l^2 — az egy főre jutó jövedelmek szórásnégyzete az l-edik részsokaságban,

n_l — az l-edik részsokaságba tartozó személyek száma,

$n = \sum_{l=1}^{L} n_l$ — az összes személyek száma.

A részsokaságokat egy vagy több olyan ismérv alapján konstruáljuk meg, hogy azok a jövedelemegyenlőtlenségnek önálló közgazdasági tartalommal rendelkező, fontos tényezői legyenek. Erre általában mindig megvan a lehetőség.

Ha a részsokaságok képzésére használt ismérvek valóban lényegesek a jövedelemegyenlőtlenség alakulása szempontjából, akkor arra lehet számítani, hogy a jövedelmek szórásnégyzetének a figyelembe vett ismérvek által együt-

[3] A (2.3) egyenlet V_p ismeretében iterációval könnyen megoldható.

tesen megmagyarázott

$$H^2 = \frac{S_k^2}{S^2} \tag{2.5}$$

hányada időben több-kevesebb stabilitást mutat. E feltevés természetesen annál inkább helytálló, minél rövidebb távra szól a prognózis. Hosszú távon ugyanis – legalábbis elvileg – nemcsak H^2 nagysága, hanem még a jövedelemegyenlőtlenséget meghatározó tényezők azonossága is kérdésessé válhat.

Végezzük el ezután S_k^2-nek az egyes ismérveknek tulajdonítható

$$\tilde{S}_k^2(r) = \lambda S_k^2(r)$$

ún. korrigált hozzájárulásokra való

$$S_k^2 = \sum_{r=1}^{m} \tilde{S}_k^2(r) \tag{2.6}$$

felbontását, ahol

$S_k^2(r)$ – a jövedelmek teljes szórásnégyzetének az r-edik részsokaságképző ismérv által ténylegesen megmagyarázott része[4]

m – a részsokaságok képzésére használt ismérvek száma

$\lambda = \dfrac{S_k^2}{\sum_{r=1}^{m} S_k^2(r)}$ – korrekciós tényező.

Erre a korrekcióra azért van szükség, mert a gyakorlatban a részsokaságok képzésére használt ismérvek általában nem teljesen függetlenek egymástól, s következésképpen általában

$$S_k^2 \neq \sum_{r=1}^{m} S_k^2(r).$$

A (2.4) és (2.6) felhasználásával a relatív szórásnégyzet

$$V^2 = \frac{S^2}{\bar{x}^2} = V_b^2 + \sum_{r=1}^{m} \tilde{V}_k^2(r) \tag{2.7}$$

felbontásához jutunk, ahol

$$V_b^2 = \frac{S_b^2}{\bar{x}^2} \quad \text{és} \quad \tilde{V}_k^2(r) = \frac{\tilde{S}_k^2(r)}{\bar{x}^2}.$$

Ekkor azonban (2.5) és (2.6) következtében nyilvánvaló, hogy H^2 és a $\tilde{V}_k^2(r)$ mennyiségek ismeretében

$$V^2 = \frac{\sum_{r=1}^{m} \tilde{V}_k^2(r)}{H^2} \tag{2.8}$$

adja meg a relatív szórás négyzetét.

[4] Ezt az S_k^2 általános formulája alapján számítjuk úgy, hogy a részsokaságokat kizárólag az r-edik ismérv alapján képezzük.

Tegyük fel ezután, hogy H^2 értéke időben stabil, illetve meghatározott mértékben változik,[5] valamint azt, hogy a jövőben az r-edik ismérv jövedelemegyenlőtlenségre gyakorolt – $\tilde{V}_k^2(r)$-rel mért – hatásának t_r-szeresre való változtatása kívánatos. Ekkor a relatív szórásnégyzet V_p prongosztizált értéke (2.8) felhasználásával a következő:

$$V_p^2 = \frac{\sum_{r=1}^{m} t_r \tilde{V}_k^2(r)}{H^2} . \tag{2.9}$$

A (2.9)ben szereplő t_r együtthatók

$$t' = [t_1, t_2, \ldots, t_m] \tag{2.10}$$

vektora ezek szerint mindig valamilyen jövedelempolitikát reprezentál. E jövedelempolitika jellegét a bázisidőszakhoz viszonyítva mindig a figyelembe vett ismérvek, ill. a t_r együtthatók 1-től vett eltérése határozza meg, hiszen a

$$t_0' = [1, 1, \ldots, 1]$$

vektor reprezentálja a bázisidőszak jövedelempolitikáját. Mint a továbbiakban látni fogjuk, bizonyos korlátok között még arra is lehetőség van, hogy egy-egy t vektorhoz – s ily módon az annak megfelelő jövedelemeloszláshoz – konkrét cselekvési-intézkedési programot rendeljünk hozzá. Ez pedig már igen közel esik a jóval bonyolultabb szimulációs eljárásokkal elérhető eredményekhez.

Hangsúlyozzuk, hogy ezt a viszonylag egyszerű eljárást az teszi lehetővé, hogy kizárólag két ismeretlen paramétertől függő jövedelemeloszlást tételezzünk fel a jövőre nézve.

Végül meg kívánjuk még jegyezni, hogy a vázolt eljárás a jövedelmek relatív szórásnégyzete helyett közvetlenül a logaritmikus szórásnégyzetükre is alkalmazható, ha a jövedelmek L-2 típusú eloszlását tételezzük fel a jövőben (l. [5]).

2.2 *Egy illusztratív alkalmazás*

Az előző pontban leírt eljárás illusztrálására bemutatjuk az aktív népesség jövedelemeloszlása prognosztizálásának menetét. Ehhez mindenekelőtt azt tételezzük fel – az 1. pontban bemutatott eredményekkel összhangban –, hogy az aktív népességnek az egy főre jutó jövedelem szerinti eloszlása a jövőben L-2 vagy L eloszlással lesz közelíthető. A feladat az L-2 eloszlás σ, illetve az L eloszlás β paraméterének prognosztizálása.

Ehhez legelőször megadjuk a 9. táblázatban, hogy milyen mértékben járultak hozzá a jövedelmek relatív szórásnégyzetéhez 1977-ben:

– a háztartásban eltartott 19 éven aluli gyermekek száma,
– a háztartás keresőinek havi átlagos keresete,

[5] Természetesen az is feltételezhető, hogy H^2 a jövőben a H_1^2 és H_2^2 határok között mozog. Ilyenkor azonban V^2-re is két határ adódik, ami megnöveli a prognózisváltozatok számát.

— a háztartás keresőinek száma és
— a háztartásfő beosztása[6]
szerinti jövedelmi különbségek.

9. táblázat

Az aktív népesség 1977. évi relatív szórásnégyzetének felbontása

Ismérv	Csoportok száma	$V_k^2(r)$	$\tilde{V}_k^2(r)$
Eltartott gyermekek száma	5	0,05289	0,05498
Keresők átlagos keresete	11	0,01737	0,01805
Keresők száma	4	0,00758	0,00788
Háztartásfő beosztása	9	0,01405	0,01461
Összeg	—	0,09189	0,09552
Együttesen	—	0,09552	0,09552

A felsorolt négy tényező a relatív szórásnégyzetnek (0,15785) 60,5%-át magyarázta meg együttesen 1977-ben, λ értéke pedig 1,03952 volt.

A 10. táblázatban néhány konkrét prognózis-változatot generáló t vektort adunk meg. Feltételezve továbbá, hogy a jövőben a négy ismérv a relatív szórásnégyzetnek együttesen 57,5; 60,5; illetve 63,5%-át fogja meghatározni, megadjuk a relatív szórás ezekhez tartozó prognosztizált értékét, valamint a σ és β paraméterek értékeit (négy tizedesjegyre kerekítve).

Példaképpen: a IV. változat relatív szórásnégyzete a $H^2 = 0{,}605$ esetben a következőképpen adódik a (2.9) formula, valamint a 9. és 10. táblázatok megfelelő adatai alapján:

$$V_p^2 = \frac{0{,}6 \cdot 0{,}05498 + 2{,}00 \cdot 0{,}01805 + 1{,}05 \cdot 0{,}00788 + 1{,}30 \cdot 0{,}01461}{0{,}605}$$

$$= 0{,}15926, \text{ amiből } V_p \approx 0{,}3991.$$

Ebből végül (2.1), illetve (2.3) alapján a

$$\sigma_p = 0{,}3844 \quad \text{és} \quad \beta_p = 4{,}9610$$

prognosztizált paraméter értékek adódnak.

A 10. táblázatban szereplő t vektorok t_r elemeinek nagyságát alapvetően két tényező határozza meg: egyfelől a vizsgált sokaság t_r-nek megfelelő ismérv szerinti megoszlásának, másfelől pedig az adott ismérv szerinti jövedelmi arányoknak a jövőbeni alakulása. A jövedelmi arányok azt mutatják, hogy a szóban forgó ismérv szerint képzett részsokaságok egy főre jutó átlag jövedelmei hogyan aránylanak egymáshoz. Ragadjuk ki példaképpen a napjainkban legfontosabb jövedelemdifferenciáló ismérvet, az eltartott gyermekek számát.

[6] A beosztás a végzett munkát minősíti aszerint, hogy
— szellemi vagy fizikai
— vezető, irányító vagy irányított
— szakképzett vagy szakképzetlen
tevékenység-e. A háztartásfő beosztása szerint az aktív keresős háztartásokat 9 csoportba sorolták.

10. táblázat

A négy ismérv relatív szórásnégyzethez való abszolút hozzájárulásának feltételezett változása 1977-hez képest és ennek hatása

Tényező	I.	II.	III.	IV.	V.	VI.
	t értéke az ... változat esetén					
Eltartott gyermekek száma	1,00	1,00	1,00	0,60	0,40	0,20
Keresők átlagkeresete	1,50	2,00	3,00	2,0	2,50	3,00
Keresők száma	1,05	1,05	1,05	1,05	1,05	1,05
Háztartásfő beosztása	1,20	1,30	1,50	1,30	1,40	1,50
	$H^2 = 0{,}575$					
Relatív szórás (V_p)	0,4331	0,4537	0,4922	0,4094	0,4083	0,4072
σ_p	0,4146	0,4326	0,4658	0,3936	0,3926	0,3917
β_p	4,6362	4,4655	4,1876	4,8567	4,8674	4,8782
	$H^2 = 0{,}605$					
Relatív szórás (V_p)	0,4222	0,4423	0,4799	0,3991	0,3980	0,3970
σ_p	0,4050	0,4227	0,4552	0,3844	0,3835	0,3825
β_p	4,7338	4,5579	4,2713	4,9610	4,9720	4,9832
	$H^2 = 0{,}635$					
Relatív szórás (V_p)	0,4121	0,4317	0,4684	0,3895	0,3885	0,3875
σ_p	0,3961	0,4134	0,4454	0,3759	0,3749	0,3740
β_p	4,8295	4,6485	4,3534	5,0632	5,0745	5,0860

A 11. táblázat tartalmazza a különböző gyermekszámú háztartások népességének 1977. évi tényleges és jövőben feltételezett megoszlását, valamint a jövedelmi arányok három különböző változatát az azoknak megfelelő t_1 együtthatókkal együtt.

11. táblázat

Megoszlás és jövedelmi arányok az eltartott gyermekek száma szerint

Eltartott gyermekek száma	1977		Feltételezett			
	megoszlás	arányok	megoszlás	arányok A	arányok B	arányok C
0	32,9	100	32,1	100	100	100
1	28,3	77	28,5	80	85	90
2	29,0	64	31,0	70	75	80
3	6,6	52	6,6	60	65	70
4 és több	3,2	36	1,8	40	45	55
t_1*	1,00		—	0,611	0,419	0,264

* A t_1 együtthatók az 1977-re vonatkozó tényleges és a táblázatba foglalt feltételezésekből adódó külső szórásnégyzetek hányadosai.

Látható, hogy az A jelű eset a IV. változathoz, a B jelű eset pedig az V. változathoz közelálló t_1 értékeket eredményez.

A jövedelem források szerinti összetételének bázis évi ismeretében még annak vizsgálatára is lehetőség nyílik, hogy milyen konkrét jövedelempolitikai intézkedések meghozatalára van szükség a kívánt jövedelmi arányok eléréséhez. Ennek részleteibe azonban itt nem áll módunkban belemenni.

A többi három ismérv szerint is hasonló módon lehetne eljárni, és igen nagy számú prognózis-változatot lehetne létrehozni. Ez azonban természetesen azzal jár, hogy egyre több feltevéssel kell élnünk a változatok kidolgozása során. Egy erre alkalmas gépi programmal azonban, véleményünk szerint, viszonylag egyszerűen megoldható a különféle változatok generálása, s a megfelelő jövedelemeloszlásoknak az összehasonlítása egymással és a bázisidőszak jövedelemeloszlásával. Az egyedüli problémát az okozhatja, hogy a valóságban a figyelembe vett jövedelmi arányok nem egymástól függetlenül alakulnak, bár λ-nak 1-hez közeli értéke arra utal, hogy a négy szóban forgó ismérv kölcsönös függősége nem túl nagy, s így talán megengedhető, hogy jövedelemdifferenciáló hatásukat egymástól függetlennek tekintsük.

Visszatérve a 10. táblázatra, látható, hogy a benne szereplő t vektorok egymástól elég lényegesen eltérő jövedelempolitikákat képviselnek. Míg ugyanis az I., II. és III. változat az eltartott gyermekek száma szerint mutatkozó jövedelmi arányok fenntartása mellett kívánja több-kevesebb mértékben fokozni a kereseteknek és a háztartásfő beosztásának a jövedelem-differenciáló hatását, addig a másik három változat esetében a gyermekszám szerinti jövedelmi különbségek különböző mértékű mérséklése mellett kerül sor ugyanerre. A relatív szórásnégyzethez való abszolút hozzájárulásoknak az eltérő nagyságrendje miatt ennek az a következménye, hogy az első három változat megvalósulása a jövedelmek relatív szórásának elég jelentős mértékű fokozódásával járna, míg a második három változat esetében a gyermekszám szerinti különbségek mérséklődése általában ellensúlyozná a többi tényező jövedelemdifferenciáló hatásának fokozódását. Ezért úgy véljük, hogy a gyakorlatban csak a IV., V. és VI. változat jöhet szóba tényleges alternativaként, mert különben az egyre nagyobbá váló abszolút jövedelmi különbségek igen könnyen társadalmi feszültségek okozójává válhatnának.

Végül a 12. táblázatban azt közöljük, hogy a 10. táblázatban felsorolt hat változat esetében — $H^2 = 0{,}605$ feltételezése mellett — hogyan alakul a négy vizsgált ismérv százalékos hozzájárulása a relatív szórásnégyzethez.

12. táblázat

A négy ismérv százalékos hozzájárulása a relatív szórásnégyzethez

Ismérv	I.	II.	III.	IV.	V.	VI.
	változat					
Eltartott gyermekek száma	30,9	28,1	23,9	20,7	13,9	7,0
Keresők átlagkeresete	15,2	18,5	23,5	22,7	28,5	34,4
Keresők száma	4,6	4,2	3,6	5,2	5,2	5,2
Háztartásfő beosztása	9,8	9,7	9,5	11,9	12,9	13,9
Egyéb tényezők	39,5	39,5	39,5	39,5	39,5	39,5
Együtt	100,0	100,0	100,0	100,0	100,0	100,0

Látható, hogy a t_2 együttható növelése még a t_1 együttható állandó szinten tartása mellett is csökkenti az eltartott gyermekek számának a jövedelmek relatív szórásnégyzetéhez való százalékos hozzájárulását (I–III. változat), de a két érintett tényező relatív súlyának lényeges módosítása csak t_1 és t_2 egymással ellentétes irányú változtatása útján érhető el.

Hátra van még a különböző változatokhoz tartozó jövedelemeloszlások meghatározása. Ehhez – mint láttuk – a σ_p és β_p paraméterek mellett még az m_p és α_p paraméterek meghatározására van szükség a (2,1) és (2,2) összefüggések alapján.

Illusztrációképpen a 10. táblázatban szereplő IV. változathoz tartozó jövedelemeloszlásokat határozzuk meg az L-2, illetve L eloszlás feltételezése mellett. Tételezzük fel, hogy az aktív népesség egy főre jutó tervezett jövedelme a prognosztizált időszakban 3200 Ft lesz.[7] Az m_p és α_p paraméterek prognosztizált értékei a fenti feltétel és a IV. változathoz tartozó σ_p, illetve β_p értékek mellett az alábbiak:

H^2	m_p	α_p
0,575	7,99344	$1{,}33563 \cdot 10^{-17}$
0,605	7,99701	$5{,}71343 \cdot 10^{-18}$
0,635	8,00027	$2{,}48786 \cdot 10^{-18}$

Az itt és a 10. táblázatban található paramétereknek megfelelő jövedelemeloszlásokat a 13. táblázat mutatja.

Itt is látható a két eloszlásnak az 1.1 pontban említett jellegzetes eltérése, az ti., hogy az L eloszlás csúcsosabb, jobban tömörül a módusz körül. Így pl. ha a prognosztizált időszakban a jövedelmek eloszlása L-2 eloszlást fog követni,

13. táblázat

A IV. változatnak megfelelő prognosztizált jövedelemeloszlások

Egy főre jutó havi jövedelem (1980-as árakon)	$H^2 = 0{,}575$		$H^2 = 0{,}605$		$H^2 = 0{,}635$	
	L-2 eloszlás	L eloszlás	L-2 eloszlás	L eloszlás	L-2 eloszlás	L eloszlás
–1500	4,2	3,4	3,8	3,1	3,4	2,9
1500–2000	11,7	9,1	11,4	8,8	11,0	8,5
2000–2500	17,4	17,3	17,5	17,2	17,5	17,1
2500–3000	18,0	20,9	18,3	21,3	18,7	21,6
3000–3500	15,1	17,8	15,5	18,2	15,9	18,6
3500–4000	11,3	12,1	11,5	12,3	11,8	12,5
4000–4500	7,9	7,4	8,0	7,5	8,0	7,5
4500–5000	5,2	4,4	5,2	4,4	5,2	4,4
5000–5500	3,4	2,7	3,3	2,6	3,3	2,6
5500–6000	2,2	1,6	2,1	1,6	2,0	1,5
6000–7000	2,2	1,7	2,1	1,6	2,0	1,5
7000–	1,4	1,6	1,3	1,4	1,2	1,3
Összesen	100,0	100,0	100,0	100,0	100,0	100,0

[7] Ez hozzávetőlegesen az aktív népesség 1985-re tervezett egy főre jutó személyes jövedelmének felel meg 1980-as árszinten kifejezve.

akkor H^2-nek és a relatív szórásnak a feltételezett értékétől függően a személyek 33—35%-a fog 2500 és 3500 Ft közötti jövedelemmel rendelkezni, míg ha a jövedelemeloszlás L eloszlás szerint alakul, akkor ez az arány 39—40% körül lesz. Az eltérés az eloszlás típusában természetesen a jövedelem egyenlőtlenségben is megmutatkozik: a relatív szórás azonos mértéke mellett L-2 eloszlás esetén valamivel nagyobb egyenlőtlenséget jeleznek a különböző egyenlőtlenségi mutatók, mint L eloszlás esetén.

2.3 *Néhány záró megjegyzés*

Cikkünkben érthető módon a jövedelemeloszlások prognosztizálására helyeztük a hangsúlyt. A jövedelemeloszlások különféle valószínűségeloszlásokkal való közelítése azonban önmagában is fontos kérdés, s az empirikus ökonometria régi törekvései közé tartozik. Az ezt célzó legismertebb kísérletek PARETO [13], GIBRAT [7] és CHAMPERNOWNE [2] nevéhez fűződnek, de a téma kutatása napjainkban is változatlan intenzitással folyik ([9], [12], [14] és [15]). Hazánkban az első ilyen vizsgálatokra a hatvanas évek elején került sor [3]. Cikkünk első része azt bizonyítja, hogy az ilyen vizsgálatokat időről-időre mindenképpen célszerű megismételni, mert a jövedelemeloszlásokat közelítő valószínűségeloszlások típusában az idő folyamán változások állhatnak be. Természetesen igen izgalmas kérdés volna annak vizsgálata is, hogy miért következnek be e változások, s hogy mennyire nemzetközi érvényűek a változási tendenciák. E kérdések megválaszolása azonban még további széles körű vizsgálódást igényel. Ugyanígy adós még az elmélet annak vizsgálatával is, hogy társadalompolitikai szempontból a jövedelmek melyik valószínűségeloszlás szerinti megoszlása lenne kívánatosabb.

Az a tény, hogy az eloszlások minden lényeges jellegzetessége kevés számú paraméterbe sűríthető, önmagában is nagy jelentőségű, de — mint láttuk — jól felhasználható a jövedelemeloszlások prognosztizálására is. A 2.2 pontban az általunk javasolt eljárást egy adott népességcsoportra, az aktív népességre alkalmaztuk illusztratív módon. A módszer természetesen más népességcsoportokra is alkalmazható. Ennek lényegében csak az az előfeltétele, hogy előzetesen feltárjuk az adott népességcsoportok jövedelemegyenlőtlenségére ható legfőbb tényezőket. A különböző népességcsoportok jövőbeni súlyának és jövedelemeloszlásának ismeretében ezután igen könnyen előállítható az össznépesség jövedelemeloszlása is. Mivel már az egyes népességcsoportokra is nagy számú változat állítható elő — a változatok közé beleértve a jövedelmek eltérő típusú eloszlásának feltételezését is —, még fokozottabban igaz ez az össznépesség jövedelemeloszlására.

A 2.1 pontban leírt módszernek igen előnyös tulajdonsága az, hogy nemcsak az abszolút jövedelemkategóriák szerinti jövedelemeloszlás előállítására alkalmas, hanem az ún. kvantilis-eloszlások (pl. decilis-eloszlások) meghatározására is. A kvantilis eloszlások az L-2, illetve L eloszlás esetében csak a σ, ill. β paramétertől függenek.

További előnyös tulajdonság az is, hogy a prognosztizált eloszlások paramétereinek ismeretében igen egyszerűen megoldható az abszolút kategóriák szerinti jövedelemeloszlás átszámítása tetszőleges árszínvonalra. Ha ugyanis a fogyasztói árindex értéke nem nagyon függ az egy főre jutó jövedelem nagyságától — mint például az 1970-es évtizedben —, akkor az új árszínvonalon zámított jövedelemeloszlás paraméterei L-2 eloszlást alapul véve

$$m'_p = m_p + \ln I_p \quad \text{és} \quad \sigma'_p = \sigma_p.$$

L eloszlást feltételezve pedig

$$\alpha'_p = \alpha_p \cdot I_p^{-\beta_p} \quad \text{és} \quad \beta'_p = \beta_p,$$

ahol a vessző nélküli paraméterek egy adott év árszínvonalán meghatározott értékek, a vesszővel jelölt értékek egy ettől eltérő árszínvonalon meghatározott paraméterértékek, I_p pedig a két árszínvonal arányát kifejező fogyasztói árindex.

A javasolt módszer leggyengébb pontja a t vektor elemeinek meghatározása, amennyiben egy-egy konkrét t-hez nem rendelhető hozzá teljes egyértelműséggel egy társadalom- és jövedelempolitikai intézkedés-sorozat. Úgy véljük azonban, hogy e részleges hiányosságért bőséges kárpótlást nyújt a módszer igen egyszerű és könnyen áttekinthető volta.

(*Beérkezett: 1981. március 28-án*)

IRODALOM

[1] Aitchison, J.—Brown, J. A. C.: The Lognormal Distribution, Cambridge University Press, 1969.
[2] Champernowne, D. G.: The Graduation of Income Distributions, *Econometria*, Vol. 20. (1952): pp. 591—615.
[3] Éltető Ö.: Jövedelem-eloszlások jellegének és tulajdonságainak vizsgálata. (Az életszínvonal elemzésének és nemzetközi összehasonlításának kérdései.) Akadémiai Kiadó, Budapest, 1962. 191—206. old.
[4] Éltető Ö.—Frigyes E.: Új jövedelemegyenlőtlenségi mutatók tulajdonságai és hasznosítási lehetőségeik. *Szigma*, I. évf. (1968.) 17—28. old.
[5] Éltető Ö.: Kísérlet a jövedelemeloszlás prognosztizálására, a családi jövedelemkülönbségekre ható tényezők alapján. *Gazdaság*, XV. évf. 4. szám.
[6] Fisk, D. R.: The Graduation of Income Distributions, *Econometrica*, Vol. 29. (1961.) pp. 171—185.
[7] Gibrat, R.: Les Inegalités Economiques, Paris, Recueil Sirey, 1931. 296.p.
[8] Johnson, N. L.—Kotz, S.: Continuous Univariate Distributions Volumes 1 and 2, New York, Wiley 1970.
[9] Kloek, T.—H. K. Van Dijk: Efficient Estimation of Income Distribution Parameters, *Journal of Econometrics*, Vol. 8. (1978) pp. 61—74.
[10] Központi Statisztikai Hivatal: A családi jövedelmek színvonala és szóródása 1972-ben. Statisztikai Időszaki Közlemények. 351. kötet, Budapest, 1975.
[11] Központi Statisztikai Hivatal: A családi jövedelmek színvonala és szóródása 1977-ben. Statisztikai Időszaki Közlemények, 462. kötet, Budapest, 1980.
[12] McDonald, J. B.—Ransom, M. R.: Functional Forms, Estimation Techniques and the Distribution of Incomes, *Econometrica*, Vol. 47 (1979) pp. 1513—1525.
[13] Pareto, V.: Cours d'Economie Politique, Lausanne and Paris: Rouge & Cie, 1897.
[14] Salem, A. B. Z.—Mount, T. D.: A Convenient Descriptive Model of Income Distribution: The Gamma Density, *Econometrica*, Vol. 42 (1974), pp. 1115—1128.
[15] Singh, S. K.—Maddala, G. S.: A Function for Size Distributions of Incomes, *Econometrica*, Vol. 44. (1976), pp. 963—970.
[16] Vita L.: Jövedelemeloszlások prognosztizálása a jövedelemeloszlásokat leíró függvények típusának időbeli stabilitás-vizsgálata alapján. Kézirat, OT Tervgazdasági Intézet, 1981.

APPROXIMATION AND FORECASTING OF INCOME DISTRIBUTIONS

The study consists of two main parts. In the first part the authors examine temporal changes in the type of functions describing Hungarian income distributions on the basis of grouped income distribution data (with 26—30 income brackets) of three conse-

cutive income surveys (in the years of 1967, 1972 and 1977). They come to the conclusion that the type of distribution of the active and the entire population according to per capita income changed from L-3 to L between 1967 and 1977, while income distribution of the inactive population may be satisfactorily approximated also further on by the L-2 distribution.

In the second part a simple forecasting method is presented and illustrated based on the assumption that the income distribution of a population group may be approximated also in the future with some two-parameter probability distribution. The authors determine future values of parameters of the supposed distribution on the basis of planned average income and the decomposition of relative variance of incomes as well as on that of changing the contribution of some most important factors to relative variance in the desired direction and to the desired extent.

ПОДХОД К РАСПРЕДЕЛЕНИЮ ДОХОДОВ И ИХ ПРОГНОЗИРОВАНИЕ

Данная работа складывается из двух частей. В первой авторы рассматривают изменения по времени в типах зависимостей, описывающих распределение доходов в Венгрии по данным следующих друг за другом обследований доходов (1967, 1972 и 1977 гг.), сгруппированных соответственно распределению доходов (26—30 классы являются общими). Они приходят к такому выводу, что в период с 1967 по 1977 гг. тип распределения доходов активного и всего населения из расчета на душу изменился с L_3 до L в то время как распределение доходов по неактивному населению и в последующем может удовлетворительно устанавливаться с помощью распределения по L_2.

Во второй части приводится и иллюстрируется такой простой метод прогнозирования базирующийся на такой предпосылке, в соответствии с которой распределение доходов по отдельным группам населения и впредь может устанавливаться посредством какого-то вероятного распределения с использованием двух параметров. Значения параметров предполагаемого распределения в будущем авторы устанавливают на основании дифференциации квадрата относительного рассеивания доходов и планируемых средних доходов, т. е. желаемого направления и величины участия некоторых наиболее существенных факторов.