

VII. KEYFITZ NATHAN: A NÉPESSÉG ELŐREJELZÉSÉNEK KORLÁTAI

(The limits of population forecasting) = Population and Development Review.
1981. 4. sz. 579-593. p.

A jövő népe sség számos - társadalmi-gazdasági-műszaki - tényezőtől függ, amelyek a népe sség előrejelzésénél nem vehető k figyelembe. Bármennyire interdisciplinári- sokká válunk is, mégis vannak bizonyos világos korlátai azon változók összefüggései ismeretének, amelyeknek a kombinált működése hozza létre a jövő népe sséget. A demográfusokat nem lehet nagyobb mértékben felelőssé tenni a népe sség 20 éves előrejelzésének pontatlanságáért, mint a geológusokat, meteorológusokat vagy a közgazdászokat, ha nem sikerül előre jelenteniük a földrengéseket, hideg teleket vagy gazdasági depressziókat 20 évvel előre.

Amiért mi felelőssé vagyunk tehető k az, hogy figyelmeztessük egymást és a mi közönségünket arra, hogy mi lehet a becsléseink valószínű hibája. A statisztika megkezdte a tudományos érettség megközelítését a valószínűségi hiba számítása terén (pl. az asztronómiai megfigyelésekben) és azután tovább ment a konfidencia intervallumokig és a szignifikancia tesztekig. A statisztikusok a számokkal dolgozó többi embertől a hibákat kezelő hatékony módszereik tekintetében különböznek. Az agronómus-statisztikusok nemcsak azt közlik a farmerekkel, hogy nagyobb terméshozamot takarítanak majd be egy új burgonyafajtából, hanem azt is, hogy ez a hozam milyen valószínűségi korlátok közé esik, mondjuk 19 az 1-hez eséllyel. A kísérleti próbákat egy valószínűségi modellnek megfelelően alakítják ki és a modellen egzakt konfidencia intervallumokat számítanak és ismertetnek.

Valamely népe sségi előrejelzés felhasználójának nincs kevésbé szüksége hibájának ismeretére, mint egy terménybecslés vagy egy munkanélküliségi becslés felhasználójának. Sajnálatos módon a népe sségi előrejelzéseket nem lehet egy valószínűségi modellbe foglalni az eljárások véletlen allokációja vagy az egységek véletlen kiválasztása következtében. A demográfusok azonban rendelkeznek egy eddig rendszeresen ki nem aknázott segédforrással: régebbi előrejelzésekkel és ezeknek a későbbi népe sségi eredményekkel való összehasonlításával. Az ilyen összehasonlítások felhasználása a jelen előrejelzés pontosságának megítélésére csupán attól függ, hogy találjunk egy sorozatot a múltból, amely valamiképpen megfelel a jelen előrejelzésnek. 1980-ban mi a 2000. év népe sségét becsüljük meg. A becslés pontosságának a megítéléséhez visszatekintünk az 1950-ben ugyanezzel a módszerrel készített 1970-es év becslésére, hogy lássuk, milyen szorosan közelítette meg az 1970-es megvalósulását. Ha a) az előrejelzésre használt módszer ugyanaz volt és

b) a népesség fluktuálása a jövőben a múltéhoz hasonló, akkor az 1950-1970-es ellenőrzés 1980-2000-re az információ egy egységét szolgáltatja.

Azt hihetnénk, hogy a népességelőrejelzőket megszállja a türelmetlenség azért, hogy lássák, hogy a múltban mennyire helyesen jártak el és hogy a felhasználók makacsul kérnek felvilágosítást a folyamatban lévő előrejelzések hibájáról. Sem ilyen türelmetlenség, sem ilyen igény nem tapasztalható. A nyomtatásban megjelent egyetlen ilyen elemzés, amit ismerek, Louis Henry (1977) tollából származik a francia megyékre vonatkozólag. A szerző úgy véli, hogy a hiba túl nagy és annál a szerény kijelentésnél marad, hogy az előrejelzések jobbak, mint egyszerűen annak a feltételezése, hogy a múltbeli népesség állandó marad.

Az össznépeesség a leggyakrabban használt, jövőre utaló egyetlen szám, de különleges felhasználásában néhány alcsoport rendszerint nagyobb érdeklődéssel bír. Különböző korú férfiak és nők, munkaerő, iskolás gyermekek, nyugdíjasok mindannyian számos esetben voltak az előrejelzések tárgyai. Mindezek vonatkozásában a múlt előrejelzések kiértékelésének az elvégzése, hasonlóképpen, mint az a következőkben történik, hasznos projekt lehet a jövő számára.

Ennek az össznépeesség előrejelzésével foglalkozó cikknek, az egyetlen számon kívül jelentősége van, ha tekintetbe vesszük, hogy az egyetlen szám hibája az összes életkorok hibáinak átlaga és hogy az idősebb életkorok pontosabban adóttak mint a fiatalabbak. Mi olyan módszert mutatunk be, ami más területen is jól felhasználható. A módszer alkalmazható a jövedelemre, árakra és egyéb esetekre. Nemcsak országokat és időintervallumokat tesz összehasonlíthatóvá, hanem lehetővé tesz megállapításokat arról, hogy vajon a népesség előrejelzése pontosabb-e mint a jövedelemé vagy a megélhetési költségek indexéé. Ezek a kiterjesztések egy későbbi tanulmány anyagai. Régebbi tapasztalatok aggregálása.

A régebbi előrejelzések hibáját a jelenlegire vonatkoztatva, szükség van néhány megelőző összehasonlítás aggregálására. Csupán az anekdota kedvéért : talán szörnyű történet kerekedhetik egy egyetlen régebbi hibából, de néhány tucat ezekből megbízható információt szolgáltatathat. A probléma tehát az, hogy a múlt tapasztalatát olyan formába hozzuk, amelyek aggregálhatók.

Mint más statisztikai elemzésben is, az első feladat a rendszeres feljegyzés készítése. Csupán a múltban készített előrejelzések gyűjtése és azt követőleg számbavett népességgel való összehasonlítás után merülnek fel a következtetés kérdései. Ez a tanulmány összegezni kívánja a múlt előrejelzéseket és az előrejelzésként használt előrebecsléseket, bizonyos átlagok kifejezéseiben. A népességelőrejelzések felhasználója számára hozzávetőleges becslést nyújt arról a fokról ameddig megbízhat a jövőt érintő vizsgált előrejelzésekben.

Minden előrejelzési időszakot egy egységnek vettem és több mint 1100 ilyen előrejelzést gyűjtöttem össze és a megfelelő megvalósulásokat. Például: India harmadik öt-éves terve, amelyet 1962-ben tettek közé, 1961-re 438 millió népességet adott meg és 1966-ra 492 milliót, ami 2,352 százalék átlagos évi növekedést jelent. A tényleges népesség, amint az a következő népszámlálásokból ismeretes, 447 millió volt 1961-ben és 502 millió 1966-ban, azaz a növekedés rátája 2,348 százalékos, az előrejelzésnél 0,004 százalékkal alacsonyabb volt. Ugyanebben a dokumentumban 1971-re 555 milliós előrejelzést közöltek és ez 2,396 százalék átlagos növekedést jelent, szemben a 2,352 százalékhoz képest mutatkozó 0,044 százalék hibával. Ezt külön megfigyelésként vesszük figyelembe.

Ami lehetővé teszi, hogy minden egyes a jövő időre vonatkozó előrejelzést elkülönített esetnek vegyünk és azután átlagoljunk az az, hogy az intervallumnak, amelyre az előrejelzés készült, úgy tűnik, hogy nincs hatása ennek a mértéknek a pontosságára, amint az a következőkből megállapítható. Ezzel szemben az az érv, hogy az 1961-1966 évek hibájának becslése és az 1961-1971-re vonatkozók nem függetlenek egymástól. Valaki, aki 1961-ben túl alacsonynak jelentett egy becslést, valószínűleg 1966-ban is túl alacsonyra jelenti 1971-re.

A függetlenség elérhetetlen az ehhez hasonló munkákban. Hasonló alacsony tendenciát mutat az országok és az 1940-es években működő szervezetek többsége; a fejlődő országokkal foglalkozó demográfusok által alkalmazott perspektívák általánosan alulbecsülték a halandóság csökkenését. A fejlődő országok tekintetében az 1940-es években készített előrejelzések alacsonyak voltak, az 1950-ben készített előrejelzések pedig magasak, az első esetben azért, mert a "baby boom" meglepetés volt, a másodikban pedig senki sem tudta, hogy a "baby boom" vége a küszöbön áll. Még abban az esetben is, amikor az előrejelzők teljesen el voltak szigetelve a közvéleménytől és egymástól, a szomszédos országok becslésének hibái nem voltak függetlenek; ha egy nyugat-európai országban a születési ráta váratlanul esik, akkor valószínűleg a többiben is esik.

A függetlenség hiánya nem hozott az összes torzulásokra ilyen eredményeket - senki sem állíthatja, hogy ezek magasabbak vagy alacsonyabbak lesznek, ha független megfigyelésekre szorítkozunk. De másodlagos kárt okoz a hiba megállapításának megakadályozásában a hibabecsléseinkben. Annak tisztázására, hogy ez további haladás lesz, sajnálatos módon, jelenleg nem állnak rendelkezésre adatok és alig is képzelhetők el.

1. Különbségek az előrejelzők között

Megkísérelhetjük az itt gyűjtött adatok felhasználását a különböző szervezetek és egyes szakemberek munkája minőségének a megállapítására, akik a jövőendő népszerűség becsléséről tanulmányokat készítenek. Ez nem része a jelen célkitűzésnek és ahogyan láttam, az egyes előrejelzők munkájának összevetése lényegében véve érdektelen. Gondolok itt a mesterlövészek nagy számára, akik mindannyian egyformán gyakorlottak, és akik egy össze-vissza mozgó célpontot vesznek célba. Egyesek ezt jobban csinálják mint mások, nem a gyakorlat különbségei következtében, hanem mert elég szerencsések voltak, hogy a célpont mozdulatlanul állt, amikor lőttek, míg másoknak az volt a balszerencsájük, hogy éppen akkor lőttek, amikor a célpont elmozdult. Eredménytelen lenne, számos mesterlövészt ilyen körülmények között, pontosságuk tekintetében osztályozni. (Az analógia Ronald Lee-től /1980/ származik.) Michael Stoto (1980) rámutatott, hogy az előrejelzés készítésének időpontja a hiba meghatározásának fő tényezője.

-Az előrejelzőket ezután úgy kívánjuk kezelni, mint az egy mozgó célpontra célzó mesterlövészeket és hibáikat nem mint az egyes mesterlövészek hibáit, hanem a célpont időben való eltalálásának valóságos nehézségét. Ami lényeges az az, vajon könnyebb-e a célpont eltalálása csupán 5 év vagy 20 év tekintetében, vajon ők közelebb kerülnek-e a lassan növekedő vagy a gyorsan növekedő országokban. Mi nem az előrejelzőket tanulmányozzuk, hanem a népszerűség előrejelezhetőségének a fokát.

Helyes a népszámlálás pontosságával kapcsolatos kérdéseket külön kezelni. Ez egy nagy tárgykör és itt nem kísérreljük meg az ezzel a kérdéssel való foglalkozást. A népszerűségmegvalósulását alkotó összes hivatalos számot megjegyzés nélkül elfogadjuk, még ha tudjuk is, hogy ezek maguk is hiba alanyok. Az összeírás bármiféle hibája (pl. az Egyesült Államokban nagyobb teljesség 1980-ban mint 1970-ben) bennefoglaltatik az előrejelzés hibájában, amikor a jelen számításokat elvégeztük; nem gondoljuk, hogy a hibát jobban lehetne csökkenteni, ha az ebből kivonható volna. Minden esetben a növekedési rátát ugyanabból a kiadványból számítottuk, úgy hogy legalább is védekeztünk az ellen, hogy hibás előrejelzési adatot vonjunk be az ellenőrzésbe.

2. Általános hiba

Egyetlen számban foglalhatjuk össze a mintegy 1100 előrejelzés összehasonlítását a megvalósulással: kerek számban kifejezve, az emlékezés és a magyarázat kedvéért, az elmúlt 30 év tapasztalata, a ma készített becslések 0,4 százalékos hibára mutatnak a négyzetes középhiba segítségével. A hibák számos más átlaga is felhasználható; előnyben részesítettem a négyzetek átlagát és abból a négyzetgyököt, mert ez nagyobb súlyt ad a nagyobb hibáknak, mint az egyszerű számtani középérték.

Ha valaki ennél többet nem tud és az az előrejelzés, hogy India népessége a következő 5 vagy 25 év folyamán 2,0 százalékkal fog növekedni, akkor kétharmados valószínűséggel az 1,6-2,4 százalékos variációs határa közrefogja a jelenleg fellépő növekedést - figyelembe véve, hogy tudatlanságunk normális eloszlású. Ha India jelenleg 690 milliós népességnél tart, az esély 2 az 1-hez, hogy népessége ettől kezdve 5 év múlva 747 és 777 millió között lesz és a 2001. évre 948 millió és 1100 milliárd között. Ha azt állítjuk, hogy India népessége 0,7 százalékkal fog növekedni, a következő század folyamán és eléri az 1,386 milliárdot, akkor 2 az 1-hez eséllyel lehet fogadni arra, hogy 931 és 2060 milliárd között lesz. A hatványozott fejlődés természeténél fogva az előrejelzés szórása ily módon számítva gyorsan nyílik szét; a hatványozott fejlődés ilyen használatát indokolja az a megállapítás, hogy a növekedés előrejelzése rátájának az eltérése a megvalósult rátától látszólag független attól az időtartamtól, amelyre az előrejelzés készült.

A gyakorlati következtetés az, hogy viszonylag rövid távú előrejelzések, mondjuk 10 vagy 20 évre, mondanak nekünk valamit, de egy negyedszázadon vagy hasonló időn túliak esetében egyszerűen nem tudjuk, hogy milyen lesz a népesség. Ez az exponenciális görbék által alkotott előrejelzések rohamos emelkedő formájának a következménye; ismételjük, hogy a kitevők használatának az érvényességét ezen a módon sugalmazza a százalékpontokban az állandó hiba 5-től 20 évig terjedően.

Az 1. tábla különféle becslések gyűjtését foglalja össze abban a vonatkozásban, ahogy egymás után elkészültek. Az 1950-es évet megelőző három gyűjtés kevésbé szoros, mint ami az Egyesült Nemzetek csoportjában is jelentkezett. Valamivel több súlyt kell adni ezeknek az újabb eseteknek és ez az oka, amiért a fenti illusztrációban nem használtam a 0,530 százalékpontot, hanem inkább a 0,4-et.

1. Az előrejelzés közepes négyzetes eltérése a rákövetkező megvalósulástól;
1939-1968-ban publikált becslések, százalékpontokban kifejezve

Ország	Előrejelzés időpontja	Előrejelzések évei	Előrejelzések száma	Közepes négyzetes hiba
Kanada ^a	1939-1954	1950-1970	14	1,123
Egyesült Államok ^b	- 1950 1950 +	1955-1975 1955-1980	15 92	0,907 0,340
Európa és a Szovjetúnió ^c	1944	1950-1970	96	0,949
Kelet-Európa 9 országa ^d	1965	1965-1980	96	0,219
Összes ország 1 000 000 ^e	1958-1968	1960-1975	810	0,476
Összes előrejelzés	1939-1968	1950-1980	1123	0,530

a Statistics Canada (1954)

b Egyesült Államok Népszámlálási Hivatala (1975)

c Notestein és mások (1944)

d Egyesült Államok Népszámlálási Hivatala (1965)

e Egyesült Nemzetek (1958,1966,1973,1979)

Abból a célból, hogy az előrejelzést részletesebben vizsgáljuk, előnyösebbnek tűnik egyetlen széles és viszonylag homogén sorozatot felhasználni. Az Egyesült Nemzetek középső variánsa 810 előrejelzést foglal magában, amelyet három különböző alkalommal végeztek el az egy milliós vagy ennél több lakosú országok tekintetében, mellőzve néhány kevés esetet, ahol a határok úgy változtak, hogy nem lehetett az egységet könnyen azonosítani. Ez az a sorozat, amelyen a legrészletesebb elemzést végeztem el. Nemcsak, hogy ez a legnagyobb és legváltozatosabb gyűjtés, hanem azt szakértő technikusok hosszú sora fejlesztette, akik az egész világ demográfusaival kapcsolatban voltak és akik módszereik tökéletesítésében érdekeltek. Erre gyakrabban történik hivatkozás, mint az előrejelzések bármelyik más sorozatára.

A számokat alátámasztó első sorozat a B10 ENSZ előrejelzés általános átlaga, amelyet az 1950-es évek végétől előrehaladva készítettek, nagyságban és növekedési rátában változó országokról, ezek némelyike fejlett ország és némelyike nem az, némelyik jó statisztikai alapadatokkal rendelkezik és némelyik nem. A minden esetben alkalmazott mérték volt az előrejelzésben foglalt növekedés évi rátája r , levonva a rákövetkező teljesítményt $R-t$. Az előrejelzés négy átlagos eltérése a rákövetkező megvalósulástól, százalékpontokban kifejezve a következő:

Átlagos eltérés, az előjel kitévésével	-0,070
Átlagos négyzetes középhiba	0,476
Az eltérés standard hibája	0,471
Az eltérés átlagos abszolút értéke	0,340

Az első számnál a mínusz jel azt jelenti, hogy az az 1950-es és az 1960-as évekről az összes előrejelzések kissé túl alacsonyak voltak, azaz egy lefelé irányuló torzulás volt. A négyzetes középhiba magában foglalja a teljes torzulás egy összetevőjét azon periódusok mindegyikében, amelyekben előrejelzéseket készítettek, plusz egy variancia összetevőt. A standard hiba csupán a variancia négyzetgyöke és mint minimum hiba alkalmazható, még ha nem is volt az összes országokban torzulás. Az abszolút közép eltérés itt is, mint mindenütt, a négyzetes középhiba gyökének mintegy kétharmada.

Mínthogy ezek egyike sem teljesen érdektelen, az alaperedmény a négyzetes középhiba gyökének megfelelő eltérés. A legközelebbi 0,1 százalékhöz való pontos kerekítéssel 0,5 százalékpontot létesít, de lekerekítve 0,4-ig, indokoltnak látszik a jelenlegi felhasználás számára, mínthogy az újabb előrejelzések valamivel kisebb eltérést mutatnak. Egy második ok az, hogy a növekedés alacsonyabb rátájú előrejelzései kisebb hibát mutatnak és a növekedés rátái valószínűleg elmaradnak a jelenlegi szintektől.

3. Nyilvánvalóan javuló előrejelzések

A 2. tábla mutatja a mi négy mértékünket, az előrejelzéseknek a megvalósulástól való eltéréseiről, mint az három különböző időszakban közzétett munkában megjelent. Az átlagos négyzetes középhiba, a standard hiba és az átlagos abszolút hiba, mind állandó csökkenést mutat. Az 1960-as évek végének az előrejelzései az 1950-es évek vége hibáinak csupán mintegy kétharmadát mutatják.

2. Négy hibamérték az ENSZ három időszakra vonatkozó előrejelzésénél

Időszak	Átlagos eltérés (1)	Négyzetes középérték (2)	Az előrejelzés standard hibája (3)	Átlagos abszolút hiba (4)
---------	------------------------	-----------------------------	---------------------------------------	------------------------------

Az előrejelzés hibái 4 egymást követő időszakra

1958	-0,266	0,555	0,487	0,417
------	--------	-------	-------	-------

Az előrejelzés hibái 3 egymást követő időszakra

1963	0,123	0,429	0,411	0,293
------	-------	-------	-------	-------

Az előrejelzés hibái 2 egymást követő időszakra

1968	0,159	0,357	0,319	0,257
------	-------	-------	-------	-------

Összes előrejelzés átlaga

	-0,070	0,476	0,471	0,340
--	--------	-------	-------	-------

Szerettük volna azt mondani, hogy a módszerek tökéletesedtek és ez nyilvánvalóan valószínű. Egy alternatív lehetőség az, hogy hirtelen változások következtek be az 1960-as évek folyamán, beleértve a "baby boom" végét Európában és a halandóság csökkenését másutt, valamint hogy az 1970-es évek kevésbé váratlan változásokat jeleztek. Ezekről eltekintve, az utóbbi években a népesség növekedése jelenlegi szintjének és rátainak ismerete is tökéletesedett. A tökéletesedés három forrása közötti megkülönböztetés és annak az ismerete, hogy mi tulajdonítható ezek mindegyikének, a létező adatok mellett, nem lehetséges.

Ilyen bizonytalanságok mellett végül is három úton lehet a tábla eredményeit arra felhasználni, hogy jelenleg mi a jövő, azaz milyen legyen az 1981-ben végzendő előrejelzés:

1. Azzal lehetne érvelni, hogy a jövő éppen úgy változhat mint a múlt és a jelenlegi demográfiai módszerek egészükben véve sem jobbák, sem rosszabbak, mint a régebbiek és átlagosan eléri az egész beszámolási időszak alatt a négyzetes középhiba mintegy 0,5 százalékos értékét.

2. Azt is lehet mondani, hogy a régebbi munkák tökéletesedtek és hogy a jelenlegi előrejelzések legfeljebb az 1960-as évek végén mutatkozó hibákat tartalmazhatják. Ennek a véleménynek megfelelően a négyzetes középhibát 0,00357-re vagy durván 0,4 százalékpontonra lehet felvenni.

3. Azt is lehetne mondani, hogy a késői 1960-as évek hibái kétharmadát teszik ki az 1950-es évek hibáinak és ezért az 1970-es évek végének a hibái az 1960-as évek végén mutatkozó hibáknak kétharmadára tehetők, a négyzetes középhiba mintegy 0,24 százalékpontos eltéréssel. Az ebből a megfontolásból levezetett szűk terjedelemmel való számítás azonban ésszerűtlen lenne.

Ezenfelül egy vagy két dekádot igényel az ezen értelmezések közötti döntés, mindamellett helyesnek tűnik a legújabb adatokra nagyobb súlyt fektetni.

4. Állandóság a különböző előrejelzési időközök tekintetében

A 3. tábla összesíti a 810 ENSZ előrejelzést, összevetve az azt követő megvalósulással, bemutatva az 5 évtől 20 évig terjedő előrejelzési időközöket. Az előrejelzések eltérései a megvalósulástól viszonylag egyenlők az egyes időközökben. Nézzük például a táblában a 4. oszlopot. Az első sorozatban a legalacsonyabb négyzetes középhiba 0,536, a legmagasabb 0,589; a második sorozatban az ingadozás 0,408-tól 0,465-ig terjed. Néhány variáció jelentkezik az idényváltási pont vonatkozásában, amint azt az 1. táblából elvárhattuk, nemkülönben a növekedési ráta tekintetében.

A 3. táblából az a fontos következtetés, hogy a növekedés várható rátája becslésében a hiba közel változatlan, tekintettel arra az időszakra, amelyre a becslés készült. A 2. tábla felső részében extrapolálva, kissé kevesebb hiba fog mutatkozni az előrejelzés 100 éves időközében, de a tábla második részének ilyen extrapolálása hosszabb időszakra nagyobb hibát fog jelezni. A csökkenés vagy a növekedés kisebb tendenciája figyelmen kívül hagyható.

Bizonyos különbségek vannak az ezen mértéken való pontosításban a gazdan országoknak a szegény országokhoz való viszonyításában, a nagy országoknak a kis országokkal való összevetésében és a különböző kontinensek között. De mindezek másodlagos jelentőségűnek tűnnek és végső soron a növekedési rátával társult különbségekből erednek. A próbák világosan jelezték, hogy a használt mértéken (az előrejelzésnek a megvalósulástól való százalékos eltérése négyzetes középérték hibája) a lassan növekedő országok pontosabban adottak mint a gyorsan növekedők.

A 810 becslésre az alapvető különbség nyers összegzése nyerhető, három csoportra osztva ezt. A 30 leghalványabban növekvő ország évi 1,8 százalékos növekedésig 0,29 százalékpont négyzetes középhibát mutatott. A középső 30 ország évi 1,8-2,6 százalékos növekedéssel, 0,48 százalékpont négyzetes középhibát mutatott. A 30 leggyorsabban növekvő ország évi 2,6 százaléknál több növekedéssel, 0,60 százalékos pontos középértéket mutatott. Ezek a különbségek kétségtelenül jelentősek. Lássuk most, hogy mit jelentenek ezek az alkalmazásban.

3. Az ENSZ három, ütemváltási évre (jumping-off years) vonatkozó előrejelzésének négyzetes középértékű hibája, százalékpontokban

Ütemváltási év	Előrejelzési időszak	Növekedési ráta			Növekedés összes rátája (4)
		Lassú (1)	Közepes (2)	Gyors (3)	
1958	1955-60	0,320	0,555	0,793	0,589
	1955-65	0,338	0,545	0,713	0,553
	1955-70	0,323	0,534	0,698	0,541
	1955-75	0,310	0,538	0,691	0,536
	Összesen	0,323	0,543	0,725	0,555
1963	1960-65	0,236	0,421	0,527	0,413
	1960-70	0,248	0,447	0,487	0,408
	1960-75	0,250	0,481	0,596	0,465
	Összesen	0,245	0,450	0,539	0,429
1968	1965-70	0,298	0,359	0,398	0,354
	1965-75	0,245	0,418	0,390	0,359
	Összesen	0,273	0,390	0,394	0,357
Mindösszesen		0,288	0,478	0,604	0,476

Az Egyesült Államok hivatalos előrejelzése (1980. évi Statisztikai Évkönyv) az 1980. évi 222 millió népesség növekedését 2000-re 260 millióra teszi, ami magában foglal egy 0,79 százalékos átlagos évi növekedési rátát. Ezen a kétharmad konfidencia sávok $0,79 \pm 0,29$ vagy 0,50-től 1,08 lesznek. Ez azt jelenti, hogy 2 az 1 esélyünk arra, hogy a népesség a 2000. évben 222 (1,0050)²⁰ és 222 (1,0108)²⁰ között lesz, azaz 245 és 275 millió között. Ugyanebben a hivatalosan közreadott táblában változatok is szerepelnek, 246 milliós alacsony és 280 milliós magas szinten. Ekként megmutattuk, hogy a hivatalosan publikált alacsony és magas változatok némiképpen szélesebbek mint a kétharmad konfidencia-intervallum - a valószínűség, hogy a hivatalos változók közrefogják a való helyzetet, kétharmadnál valamivel több.

Eredményeinket gondosabban kívánjuk kifejtteni. Ha a jelenleg készített előrejelzések éppen olyan jók mint a régebbiek és nem jobbak azoknál és ha a népesség összetevőinek előre nem látott kanyarai és fordulatai éppen olyan nagyok mint a régebbiek és nem nagyobbak azoknál, akkor úgy vehetjük, hogy a most elkövetett hibát ugyanabból az elosztásból vettük, mint a régebbi hibát; így az esély, hogy a 2000. év népessége 245 és 275 millió között lesz, kétharmad és az esély, hogy a hivatalos terjedelem közrefogja, kevéssel több mint kétharmad. Az esély egyharmad arra, hogy a 2000. év népessége kevesebb lesz, mint a publikált 245 millió alacsony változat vagy nagyobb mint a 280 milliós magas változat.

Ez az eredmény csalódást fog okozni azoknak, akik úgy vélik, hogy a népesség tökéletesen megjósolható, de aligha fogja meglepni azokat, akik jól emlékeznek a múltbeli hibáikra, lévén ez valójában sem több sem kevesebb mint a múlt hibái átlagolásának egyik módja. Azt tudatja velünk, hogy van bizonyos ismeretünk 20 évre előre a népességről, de ez az ismeretsáv széles.

Egy hasonló számítás azt mutatja, hogy valójában semmit sem tudunk a népességről, mostantól számítva 50 évre. Nem tudunk jobban kockáztatni mint 2-1 arányban, ha a távolság kisebb mint 285-től 380 milliós, a 2030. év tekintetében.

A gyorsabban növekedő országok vonatkozásában ne az Egyesült Államokra megfelelő 0,29-et használjuk, hanem 0,48-at, vagy a fejlődő országok legtöbbször tekintetében a 0,59-et.

5. A hiba eloszlása

Aligha lehet azt elvárni, hogy minden előrejelzés 0,4 százalékpontos vagy egyéb számú hibának legyen alávetve; a hibát úgy kell felfogni mint egy véletlen változót. Néhány előrejelzés az eseményt követően úgy tekinthető, mint a jövő népesség pontos előrelátása, míg mások nagyon távol vannak ettől.

Az 1958-ban készített 20 éves előrejelzés eloszlása 90 ország vonatkozásában a következőképpen jelenik meg:

<u>Az előrejelzés eltérése a megvalósulástól</u>	<u>Esetek száma</u>
Több mint 1,5 százalékponttal alatta	1
1,5 és 1,0 közötti százalékponttal alatta	6
1,0 és 0,5 közötti százalékponttal alatta	15
0,5 és 0 közötti százalékponttal alatta	35
0 és 0,5 közötti százalékponttal fölötte	31
Több mint 0,5 százalékponttal fölötte	2

Az eloszlásnak egyetlen kielégítő csúcsa van; az egész lefelé ferdül annak a ténynek a következtében, hogy a fejlődő országok számára történő előrejelzések alacsonyak a halandóság javulásának alulbecslései miatt.

6. Előrejelzések az országoknál nagyobb vagy kisebb területek számára

Az országon belüli adatok - államoké, tartományoké, vidékeké, megyéké, városoké - sokkal teljesebben kerülnek megvitatásra, mint az eddigiekben előadott országos adatok és az eddig bemutatott számítások könnyebben végezhetők el, mihelyt az adatok gyűjtése megtörtént.

Az Egyesült Államok példát szolgáltatnak erre. Ezeknek az államoknak 1980-ra készített megelőző népességbecslése szembeállítható az 1980-as népszámlálással. Ha ez elkészült, a négyzetes középhiba 50-es szabadságfokon 1,64 százalékpont, ami mintegy kétszerese az országra megállapított hibának. Ez így van, annak a ténynek ellenére, hogy ebben az esetben az előrejelzés alkalmazta az államokból származó jelzőszámokat 1979-ig. Kétséggkívül, az előre nem látott belső vándorlás nagyobb mint a termékenység és a nemzetközi vándorlás ismeretlen eleme, de az, hogy mennyire nagyobb, még vizsgálatot igényel nagyobb részletességben és egyéb vidékek tekintetében.

A magasabb szintű aggregátumok - kontinensek vagy a világ - vonatkozásában lényeges adathiány áll fenn. Thomas Frejka (1981) legújabban a világnépesség előrejelzésének hasznos történetével szolgált, ez a gyűjtés azonban azt mutatja, hogy a rendelkezésre álló anyag nem elegendő valószínűségi megállapítások számára. Feltételezzük, hogy a sávok szorosabbak, mint az aggregáció növekedések, így az egész világ tekintetében a sávok szorosabbak lesznek mint a 0,4 százalékpontos négyzetes középhiba, amelyet az összes ország vonatkozásában javasoltunk, azt azonban, hogy mennyire szorosabbak, aligha tudjuk megmondani.

7. Naív előrejelzések

Valamely szofisztikus módszer megítéléséhez többre van szükségünk, mint hibáinak ismeretére. Egy szakmai előrejelzés hatékonyságának ki kell állnia az összehasonlítást egy naív módszerrel, amely például azt tételezi fel, hogy az elmúlt év növekedése folytatódni fog a jövőben. Ha a hiba nagyobb volt mint egy naív módszer hibája, arra a következtetésre kell jutnunk, hogy a demográfiai módszernek nincs értéke az előrejelzés számára. Általánosságban elvárjuk, hogy a szofisztikus módszer kevesebb hibát mutasson mint a naív módszer és azt is szeretnénk megtudni, hogy mennyivel kevesebbet. Ronald Lee (1978) rámutatott arra, hogy az Egyesült Államok Népszámlálási Hivatalának termékenységi előrejelzései milyen közel vannak az előrejelzést megelőző évek tényleges termékenységéhez; ez a megállapítás a 4. táblában bemutatott kísérletet sugalmazza.

Az 1950-es évek végén készített előrejelzések értékelésére az 1950. és 1955. évek megfigyelt népességét vehetjük és feltételezhetjük, hogy az a növekedési ráta, amelyet mutattak a jövőben is folytatódik. A 4. tábla (1) oszlopa mutatja, hogy mennyire jó a keletkező előrejelzés; átlagos négyzetes középhibája 0,90 százalékpontos, míg a megfelelő szám a (3) oszlopban 0,48. Tehát a szakmai munka, amely kifejezett ítéletet alkot a halandóságról és a termékenységről, a bizonytalanságnak egy szűkebb terjedelmét adja. A négyzetek arányát véve, azt mondhatjuk, hogy a szakmai előrejelzés háromszor annyi információt tartalmaz, mint az előrejelzett mértani növekedés.

A 4. tábla az országokat növekedési rátájuk szerint öt csoportba sorolva mutatja be. A múltból levezetett mértani növekedés zömmel mindenütt alacsony. A (2) oszlop be- számol egy valamivel kevésbé naív módszerről, demográfiai becslést végezve 5 évre előre és feltételezve, hogy a növekedés rátája folytatódik. Sokkal jobb mint az (1) oszlop és valóban megközelíti a publikált előrejelzések pontosságát. Megmutatja,

hogya a kezdeti kiindulásnak milyen jelentősége van; az előrejelző nem eszközöl további tökéletesítéseket kifejezett előrejelzési változásokra ehhez a kezdeti irányhoz képest öt, tíz vagy több évre előre a jövőben. Ez megegyezik azzal a ténnyel, hogy csupán nagyon kevés esetben keresztezi az előrejelzési görbe a megvalósulási görbét; sokkal gyakrabban a kettő egyszerűen divergál az indulástól kezdve. Röviden szólva: az előrejelző valami nagyon ahhoz hasonlót tett, mint amit a (2) oszlop mutat.

4. Hiba négyzetes középértéke százalékpontokban az 1955-1975 előrejelzésekre vonatkozólag, két egyszerű módszer szerint, valamint a jelenleg készített előrejelzésre: 18 népesség a növekedési ráták öt csoportjának mindegyikében

Növekedés	Mértani növekedéssel a tényleges 1950-1955-ből kiindulva (1)	Mértani növekedéssel a tervezett 1955-1960-ból kiindulva (2)	A publikált ENSZ előrejelzésekben ténylegesen alkalmazott módszerrel (3)
Leglassabb növekedés	0,58	0,32	0,26
↑	1,03	0,41	0,36
↑	0,93	0,73	0,63
↑	0,76	0,68	0,58
↓	1,18	0,87	0,71
Leggyorsabb növekedés			
Átlag	0,90	0,60	0,48

8. Az előrejelzés korlátai

Ennek a tanulmánynak a motivációja az az igény, hogy az előrejelzések felhasználóival közöljük azt, hogy a jövő népesség publikált előrejelzéseire milyen mértékben lehet támaszkodni. Megcáfolja azt a vádat, hogy a demográfusok semmit sem tudnak a jövőről, valamint azt a reményt, hogy pontosan meg tudják mondani azt, hogy mi fog történni.

A pontosság leírásának legtömörebb módja az elmúlt előrejelzések átlagos hibájának a bemutatása. Először a csillagászatban, majd az ipari feldolgozásban, a mezőgazdasági üzemtanban, a gyógyszerészetben, a humán népességi mintavételnél és számos egyéb területen alkalmazva, a statisztikai hagyomány a különleges tárgynak megfelelő alakot ölti fel, de mindig legfőbb gondja a következtetés és a hiba. A jelen tanulmány ezt a hagyományt a népesség előrejelzésére alkalmazza.

Az előrejelzés elég nehéz ahhoz, hogy elkedvetlenítse még a legbátrabbakat is és hosszú ideig a demográfusok ezt az előrevetítés fejezete alatt végezték, azt állítva, hogy csupán a változó feltételek sorozatának következményeit dolgozzák ki. Mégis, ha a demográfusok nem tudják megítélni, hogy a halandóság vagy a termékenység melyik szint-változata valósulhat meg valószínűleg, ezt akkor senki sem tudja megmondani és a gyakorlatban a felhasználók az ilyen megítélés tekintetében a demográfusoktól függenek. A középső változatú előrevetítések mint előrejelzések értelmezése közel általános a felhasználók között.

Valamely előrejelzés minőségének a leírásánál gondoskodni lehet egy olyan érték-közzől, amelyről fel lehet tételni, hogy a valóság abba bizonyos valószínűséggel bele esik. Egyetlen számban összegeztük a jelen vizsgálatot, mondván, hogy az esély 2 az 1-hez, hogy a növekedés \pm 0,4 százalékpontos előrejelző rátája a jövő időszak vonatkozásában közrefogja a megvalósult növekedési rátát. A gyakorlatban ez használható előrejelzéseket jelent az elkövetkező 5 vagy 20 esztendőre, gyakorlatilag azonban egyáltalán semmi információt a népességről mától kezdődő 100 évre előre.

Az 1. és 2. táblák azt mutatják, hogy a hiba nyilvánvalóan csökkenő jellegű az idők folyamán, vagy azért, mert a népesség változik rendszeresebb módon, vagy mert az előrejelző szakemberek szakképzettebbek lesznek. Az 1950-es évek végén készített ENSZ előrejelzések négyzetes középhibája 0,555 volt; az 1960-as évek kezdetére 0,429, az 1960-as évek végére 0,357. A különbség nem tulajdonítható annak a ténynek, hogy rövidebb idő állt rendelkezésre a legutóbbi munka ellenőrzése számára, mert akkor is így jelenik meg, ha korlátozzuk az ütemváltási pontok összehasonlítását 10 éves előrejelzésekre. Annak a lehetőségnek a megengedésével, hogy itt egy trend volt, javasoltam, hogy kerekítsük le az egyetlen számos átlagot a kezdeti 0,530-ról 0,4-re.

Eltekintve az előrejelzési módszerek abszolút kifejezésekben való értékelésétől, azt is vizsgáltam, hogy hogyan vethetők össze a naív módszerekkel. A végrehajtás egyik módja egyszerűen a megelőző öt év növekedési rátájának a megállapítása és a népesség átvitele a jövőbe a növekedésnek ezzel a rátájával. Ennek a rátának az eltérése a megvalósult rátától a négyzetes középhiba szerint 0,90 hibát mutat (4. tábla), a tényleges előrejelzés 0,48-ával szemben. A megítélés és a demográfiai módszerek valóban csökkentik az előrejelzési hibákat.

Az a felhasználó, aki kifogásolja, hogy a valószínűségi sorozat, amely a hiba becslésére adatokat szolgáltat, nem ugyanaz, mint amihez alkalmazni kívánják, abszolút helytálló megállapítást tesz. De hasonlóképpen a farmer, aki bizalommal ültet burgonyát, olyan fajtákat használva, amelyek egy kísérlet során jól szerepeltek, nem biztosítható a kísérlet urnájából való véletlen mintavétellel; későbbi, olyan időben dolgozik, amelyben a külső körülmények számos vonatkozásban különböznek. A logikus ugrás fajtája, amelyet megteszünk az előrejelzés sikerének régebbi fokáról a most megteendő egyik várható sikerhez, mindenütt megkívánja a statisztikai adatok felhasználását. Az agrónomus megtervezhet valamely kísérletet a felhasználás hibájának csökkentésére, de nem tudja azt kiküszöbölni.

Megjegyzendő, hogy ez alapvető tekintetben különbözik számos becslés bemutatásának attól a jelenlegi gyakorlatától, amely egyiket magasnak, a másikat alacsonynak, megint másikat közepesnek nevez. Valamilyen valószínűségi megállapítás nélkül, a magas és alacsony becslések haszon nélküliek annak a jelzésére, hogy milyen mértékben lehet megbízni a közepes számban és hogy mikor kell erre használni az alacsony vagy a magasat. Ugyancsak nem nyerünk valamiféle segítséget abból az ismeretből, hogy az előrejelzések mindegyike megfelel a feltételezések különböző sorának és hogy a felhasználótól függ, hogy figyelembe vegye a feltételezések három sorozatát, eldöntse, hogy melyik a reális és azt az egyet válassza ki. Ha ez számára valóban gondot jelent és ha megvan a gyakorlata a feltételezések alternatív sorozatának a megfontolásában és annak eldöntésében, hogy melyik ezek közül a legreálisabb, akkor ezenfelül számításokat is végezhet - ez már viszonylag könnyű kérdés, mihelyt a feltételezések részletesen ismertek. Ha másrésről, ami gyakrabban előfordul, a felhasználó az eredményeket tekinti és azt veszi három előrejelzés közül, amelyik a legmegfelelőbbnek tűnik neki, akkor a demográfus egyáltalán semmit sem adott a számára - a felhasználó, akitől elvárják, hogy azon az alapon válasszon, hogy az eredmények közül melyik látszik a legjobbnak, választhat akár a véletlen számok sorozata között is.

9. Számítsunk-e hibát?

Mint hogy adott a bizonyosság, hogy a jövő különbözni fog a múlttól, helytelennek tűnik azt feltételezni, hogy a jövőbeli fluktuációkat ugyanabból a valószínűségi sorozatból következtessek ki, mint az elmúlt fluktuációkat. Még maguk az előrejelzések is szemben találják magukat ezzel a nehézséggel; bármennyire bonyolultak lehetnek is, nem kerülhetik el annak a megállapítását, hogy a jövő lényeges vonatkozásokban hasonlít a múlthoz. Valaki, aki ezt a megállapítást kifogásolja, nem készíthet előrejelzéseket. Tiszteletben tartható az a nézet, hogy egyáltalán semmi sem mondható a jövő népeiségről, hogy tudományunk kevésbé fejlett ahhoz, hogy az előrejelzés számára alapul szolgáljon. De ha elegendő jónak vélik az előrejelzés számára, akkor elegendő jó az előrejelzések hibáinak a becslésére is.

Az a lehet érvelni, hogy az előrejelzés szakmai munka, ami analóg az orvos munkájához. A páciens kevésbé érdekelt a konfidencia intervallumok tekintetében egy operáció kimenetelével kapcsolatban, mint az orvos, aki a lehető legjobb munkát végzi. Éppen úgy, ahogyan az orvos nem tárgyalja meg a hibát pácienseivel, ugyanúgy a kliens, aki fizet az előrejelzésért, nem kíván elrettentő történeteket hallani arról, hogy milyen rosszak voltak az előrejelzések a múltban; az amit igényel: a demográfusban való bizalom.

Ez nem az ebben a tanulmányban kifejtett vélemény. A kliens, akire gondolok, nem egy ideges páciens, akinek nyugalmát és bizalmát meg kell óvni, hanem egy intelligens döntéshozó, akinek ismernie kell az információ feltételezett pontosságát, mielőtt hozzákezd döntése meghozatalához. Vannak nagyobb nehézségek az előrejelzések hibáinak becslésénél, de ugyanúgy vannak nehézségek magukban az előrejelzésekben is. Ha hasznosan tudunk extrapolálni népeiséget a múlt alapján, akkor felhasználhatjuk a múltat, hogy határokat szabjunk az extrapolált népeiségnek. A demográfusok tudatában vannak annak, hogy van határa a pontosságnak, amellyel a jövő megismerhető a népeiség vonatkozásában, éppen úgy mint más kérdésekben is. A múlt tapasztalata mond nekünk valamit erről a határról.

10. Függelék: a számítás módszere

Abból a célból, hogy a különböző egységeket közös alapra helyezzük, minden előrejelzést értékelni kell azon különbség alapján, amelyet az általa képviselt évi növekedési ráta és a sorban következő népességi megvalósulás felmutat. A jegyzetekben felsorolt ENSZ kiadványok anyagot szolgáltatnak a szükséges számítások illusztrálásához. Etiópia tekintetében r -nek az előrejelzési ráta növekedésének R -től, a megvalósult rátától való eltérése $-0,01283$ százalékpont az 1955-1975 időközre. Kelet-Afrika többi három állama tekintetében ezek a számok: $-0,00588$, $-0,00908$ és $-0,01255$.

A négy előrejelzés átlagos hibája:

$$\frac{-0,01283 \quad -0,00588 \quad -0,00908 \quad -0,01255}{4} = -0,01008$$

A négyzetes középérték alapján számított hiba:

$$\sqrt{\frac{(-0,01283)^2 + (-0,00588)^2 + (-0,00908)^2 + (-0,01255)^2}{4}}$$

A hiba standard eltérése:

$$\sqrt{\frac{(-0,01283+0,01008)^2 + (-0,00588 + 0,01008)^2 + (\dots)^2 + (\dots)^2}{4}}$$

és az átlagos abszolút hiba:

$$\frac{0,01283+0,00588+0,00908+0,01255}{4} = 0,01008$$

J e g y z e t e k

- Frejka, Thomas (1981): "World population projections: A concise history". (Világ-népeségi előrejelzések: rövid történet.) Center for Policy Studies Working Paper No.66. New York: The Population Council
- Henry, Louis és Gutierrez, Hector (1977): "Qualité des prévisions démographiques á court terme: Étude de l'extrapolation de la population totale des départements et villes de France, 1821-1975" (A rövid távú demográfiai előrejelzések minősége: tanulmány Franciaország megyéi és városai össznépességének extrapolálásáról, 1821-1975.) Population. 32. No.3. 625-647.p.
- Lee, Ronald D. (1978): "Appraisal of the fertility assumptions employed in the social security projections" (A társadalombiztosítási előrejelzésekben alkalmazott és a termékenységre vonatkozó feltevések értékelése.) Unpublished ms. Kiadatlan kézirat.
- (1980) "Aiming at a moving target: Period fertility and changing reproductive goals." (Törekvés az elérendő célra: termékenységi periódus és változó reprodukciós célok.) Population Studies. 206-226.p.
- Notestein, Frank W. és mások (1944): "The future population of Europe and the Soviet Union: population projections 1940-1970." (Európa és a Szovjetunió jövő népessége: népesség előrejelzések 1940-1970.) Geneva. League of Nations.
- Statistics Canada (1954): "Memorandum on the projection of population statistics, 1954" (Memorandum a népességstatisztikai előrejelzésről.) Ottawa: Dominion Bureau of Statistics.
- Soto, Michel A. (1979): "The accuracy of population projections." (A népesség előrejelzések pontossága.) Working Paper 79-75. Laxenburg Austria: International Institute for Applied Systems Analysis.
- United Nations (1958): "The future growth of world population" (A világ népességének jövő növekedése.) Population Studies No.28. New York. United Nations.
- (1966) "World population prospects as assessed in 1963." (A világ népességének kilátásai, amint azt 1963-ban megállapították.) Population Studies No.41. New York: United Nations.
- (1973) "World population prospects as assessed in 1968.*" (A világ népességének kilátásai, amint azt 1968-ban megállapították.) Population Studies. No.53. New York. United Nations.
- (1979) "World population trends and policies." 1977 Monitoring Report. Vol.1. Population trends. (A világnépesség trendjei.) Population Studies. No.62. New York. United Nations.

US Bureau of the Census (1965): "Projections of the population of the Communist Countries of Eastern Europa, by age and sex, 1965-1985" (Kelet-Európa kommunista országai népességének kor és nem szerinti előrejelzése 1965-1985.) by James L. Scott International Population Reports. Series P-91.No.14. Washington.D.C. (1975) "Population estimates and projections: Projections of the population of the United States: 1975 to 2050" (Népességelőrejelzések és becslések; az Egyesült Államok népességének előrejelzése: 1975-től 2050-ig.) Current Population Reports. Series P-25. No.601.