WORKING PAPERS ISSN 0781-4410

SUOMEN AKTUAARIYHDISTYS
The Actuarial Society of Finland

10

Kari Puustinen

ELÄKEMENON ENNUSTAMINEN (1984)

1 Abstract

This paper proposes a method to evaluate a long run pension expenditure in an employment pensions scheme. The method is based on computer simulation of pension events for every individual under the pensions scheme. The events are generated with Monte Carlo method from pseudo-random numbers according to given probabilities of events.

The initial population for simulation runs was converted from real register data of a Finnish pension fund. This data consists information of the active employees, former employees, invalids and other pensioners.

Every person in this model is in some state (active, invalid, retired, dead etc.) and once a year person has a possibility to transfer to some other state according to so-called transference functions (mortality, invalidisation, mortality of invalids, rehabilisation of invalids, resignation from membership, returnement of those resignated etc.). It is calculated the new pension expenditure in connection of every pension event.

The simulation was performed for years 1983-2015. The projections (time series of yearly expenditure) were registered into disc files. This realisation was repeated 50 times, and the results were then presented in graphic. The annual pension expenditure was also tested, and the conclusion was, that the normal approximation is applicable.

ELÄKEMENON ENNUSTAMINEN

SISÄLLYS

KARI PUUSTINEN

3.2.1984

1 Johdanto

Tässä tutkimuksessa ennustetaan erään kuvitteellisen eläkekassan eläkemenon kehitystä seuraavien 33 vuoden aikana, sekä arvioidaan saadun ennusteen luotettavuutta. Eläkekassa myöntää työntekijäin eläkelain mukaista eläketurvaa sekä rekisteröimätöntä lisäeläketurvaa.

Ennustamisessa käytetään yksilöpohjaista simulointitekniikkaa, missä vuotuinen eläkemeno muodostuu arpomalla erikseen kukin eläketapahtuma. Saatu eläkemeno on näin ollen eräs realisaatio, joka toteutuu annettujen todennäköisyyksien puitteissa. Suorittamalla simulointi useita kertoja pyritään löytämään eläkemenon kehitykselle "stokastinen hajontaviuhka".

Simuloinnin perustana ovat eläkekassan omat henkilöstöhallinnon tilastot ja yleiset kuolevuus- ,työkyvyttömyys-, ym. tilastot.

Lähtövuoden jäsenrekisteristä (sisältää kassan piiriin kuuluvat aktiivit työntekijät, työkyvyttömät ja eläkeläiset) edetään vuosi kerrallaan ja lasketaan kassan piirissä syntyvät eläketapahtumat muodostaen näin rekisteri seuraavaa vuotta varten. Simulointitekniikka perustuu siirtymätodennäköisyyksiin, minkä avulla rekisterin alkio siirtyy tilasta toiseen (tilat: aktiivit, työkyvyttömät,eläkeläiset,kuolleet).

Tutkimus on suoritettu Oy Porasto Ab:n HP3000 Series 44-tietokoneella, jolle simulointialgoritmi on ohjelmoitu. Käytetty tekniikka perustuu Martti Laihon pro gradu-tutkielmaan "Työeläkevakuutuksen yksilöpohjainen simulointimalli" (1977), missä tutkittiin vakuutusmaksutason riittävyyttä rahavirtausmallin avulla.

2 Lähtötilanne

2.1 Kassan säännöt

Eläkekassassa on kaksi osastoa: A-osasto ja B-osasto. B-osastosta myönnetään työntekijäin eläkelain (TEL) mukaista eläkettä, A-osastosta myönnetään työntekijöille lisäetuja.

A-osaston säännöt ovat eläkkeen laskemista koskevilta osiltaan seuraavat.

Jokainen kassan jäsen kuuluu myös A-osastoon.

Vuosityöansiona käytetään samaa ansiota, joka on myös työntekijäin eläkelaissa määrätty.

Vanhuuseläkkeen eläkeikänä on miehillä 65 vuotta ja naisilla 60 vuotta. Jokaisella kassan jäsenellä on myös mahdollisuus palvelusvuosiaan vastaavaan alennettuun eläkeikään palvelusta, joka ylittää 31 vuotta. Alhaisin mahdollinen eläkeikä on miehillä 62 vuotta ja naisilla 57 vuotta.

Eläkeikä määräytyy siis seuraavasti:

$$\max\{57, 60-(V60-TSALK-31)\}^{+} \quad \text{naiset}$$

$$\max\{62, 65-(V65-TSALK-31)\}^{+} \quad \text{miehet}$$

missä TSALK = työsuhteen alkamisvuosi
V60(vast. V65) = 60 vuoden (vast. 65) täyttämisvuosi

Työkyvyttömyyseläke määräytyy kuten vanhuuseläke. Ns. tuleva aika otetaan siis eläkettä määrättäessä huomioon.

Perhe-eläkettä voivat saada naislesket ja alle 20-vuotiaat lapset. Perusteina käytetään TEL-perusteita.

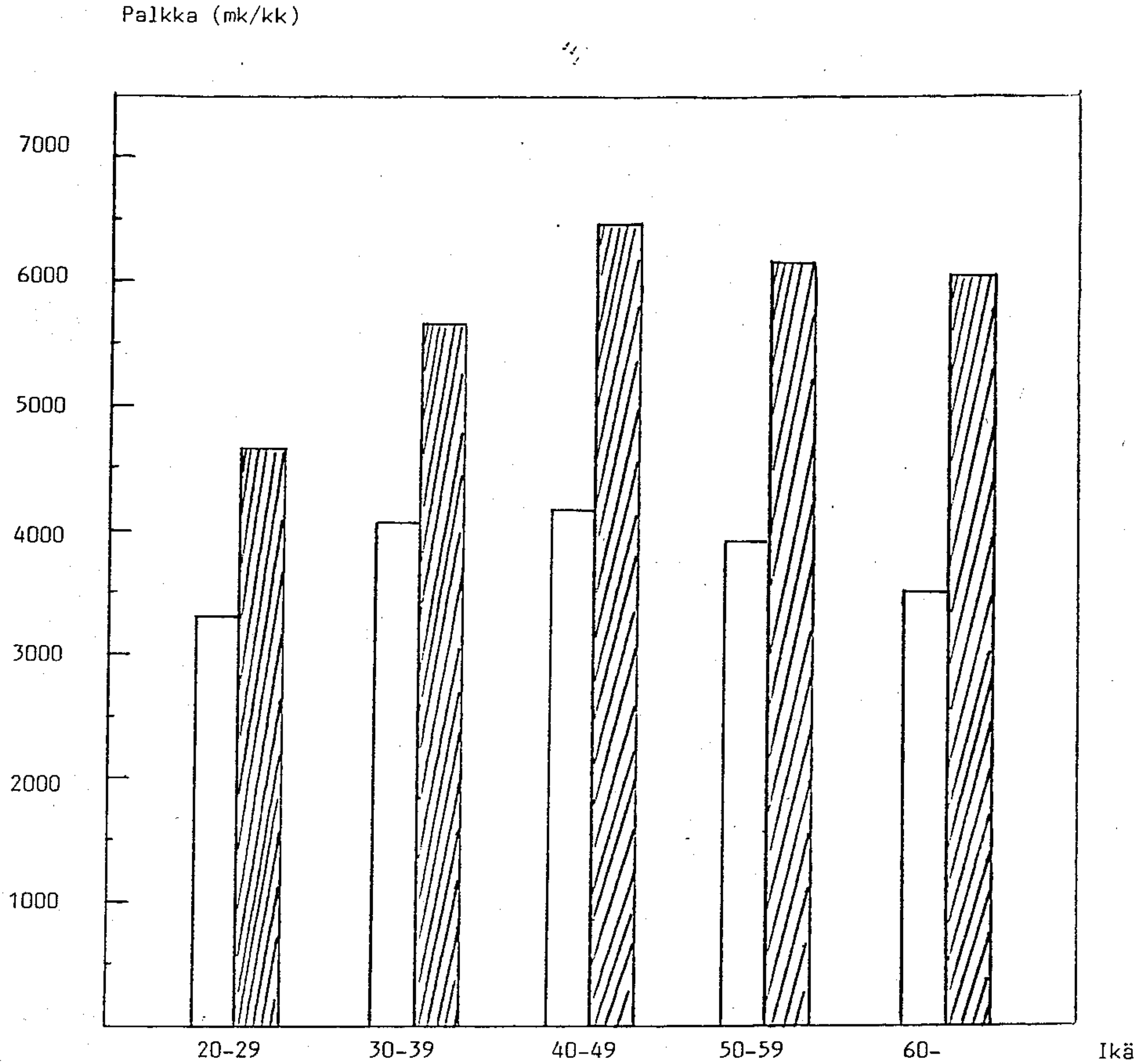
Vanhuuseläkkeen suuruus on 60 % eläkepalkasta.

2.2 Kassan rakenne

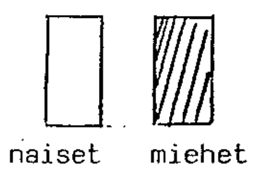
Tutkimusrekisteri sisältää tarvittavat tiedot eläkekassan piirissä olevista aktiiveista työsuhteista, entisistä työsuhteista eli vapaakirjoista sekä eläkeläisistä.

Rekisterissä on aktiiveja jäseniä 1696 kpl, joista 1010 miestä ja 686 naista. Vanhuuseläkeläisiä on 478 kpl, leskiä 252 kpl ja lapsen eläkkeitä 38 kpl. Työkyvyttömyyseläkeläisiä on 150 kpl.

Kassan palkkajakauma on esitetty alla olevassa kuvassa luokiteltuna 10 vuoden ryhmiin.



Kuva Eläkekassan palkkajakauma



2.3 Tilastot

Koska tutkimuksessa käytetty eläkekassa ei vastaa täydellisesti mitään todellista eläkelaitosta, vaan on muodostettu sopivasti yhdistelemällä Oy Porasto Ab:n asiakasrekisteritietoja, on tilastot laadittu yleisten, valtakunnallisten tilastotietojen perusteella ottaen tarvittaessa huomioon kassan säännöt.

Seuraavassa on lyhyt yhteenveto käytetyistä tilastoista.

Kuolevuutena käytettiin TEL-laskuperusteiden mukaista kuolevuutta ikäsiirtona miehille -2 ja naisille -9.

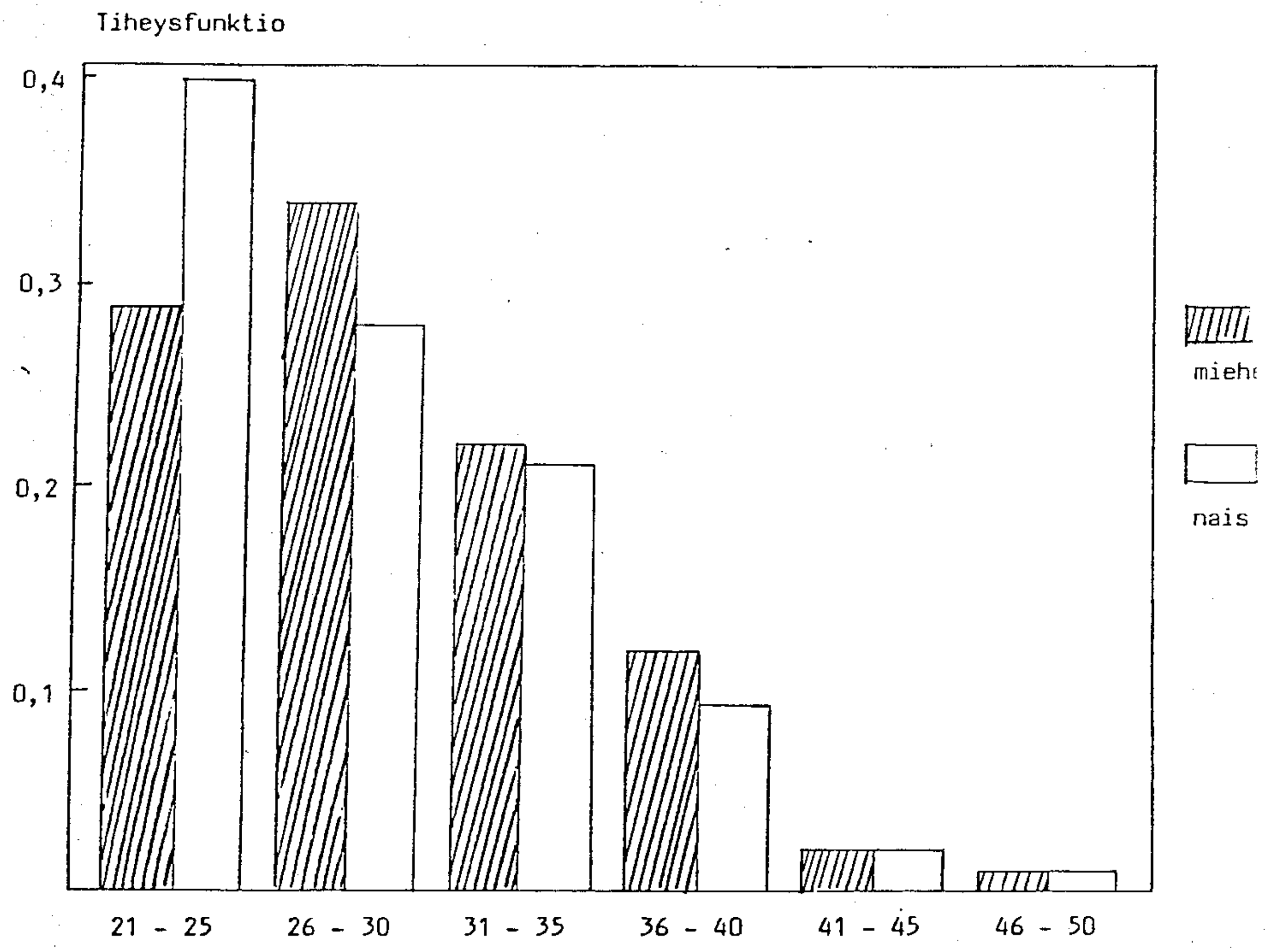
Invalidisoituvuus, eroavuus, keskipalkat sekä perheellisyysperusteet (avioisuus ja lasten syntyvyys) on laskettu eräiden Poraston asiakastietojen perusteella.

Eronneiden jäsenten palaavuus on oletettu nollaksi kaikille ikäryhmille.

Uusien jäsenten ikäjakauma on kuvan 1 tilaston mukainen.

Työkyvyttömien kuntoutuvuus on valtakunnallinen TEL-tilasto vuosilta 1966 - 1974.

Leskien kuolevuutena on käytetty Suomen tilastollisen vuosikirjan 1973 taulukkoa "Naisten väestökuolleisuus 1960-1970.



Kuva 1. Uusien jäsenten ikäjakauma

3 Simulointi

3.1 Simuloinnin periaate

Simulointitekniikka perustuu, kuten johdannossa todettiin, M Laihon tutkielmaan. Kyseisessä tutkimuksessa esitettiin teoreettinen malli, generoitiin pelkistetty rekisteri, jota taas käytettiin rahoitusmallin lähtökohtana. Käytetystä tekniikasta on esitetty yhteenveto liitteessä.

Tämän tutkimuksen osalta on simuloinnin periaate seuraava.

Rekisterin käsitteet aktiivit, vapaakirjat ja eläkeläiset toimivat systeemin "tiloina". Rekisterin henkilöt eli alkiot siirtyvät tilasta toiseen ns. siirtymätodennäköisyyksien avulla (kuva 2). Käyttäen em. todennäköisyyksiä arvotaan alkioiden liikkeet systeemin sisällä. Näin saadaan satunnainen, mutta mallin todennäköisyyksiin perustuva "polku", simulaatio. Simulointi suoritetaan useita kertoja, ja näin saadaan ennuste rekisterin kehityksestä tulevaisuudessa.

Uusi henkilö tulee kassan jäseneksi (sekä A- että B-osastoon) tilan (A,B) kautta (kuva 2). Vuoden kuluttua henkilö on joko edelleen aktiivijäsen, invalidisoitunut (joutunut tilaan (TA,B), eronnut (VPK eli vapaakirjalainen), päässyt A-osaston vanhuuseläkkeelle (VA), B-osaston vanhuuseläkkeelle (V) tai kuollut (K).

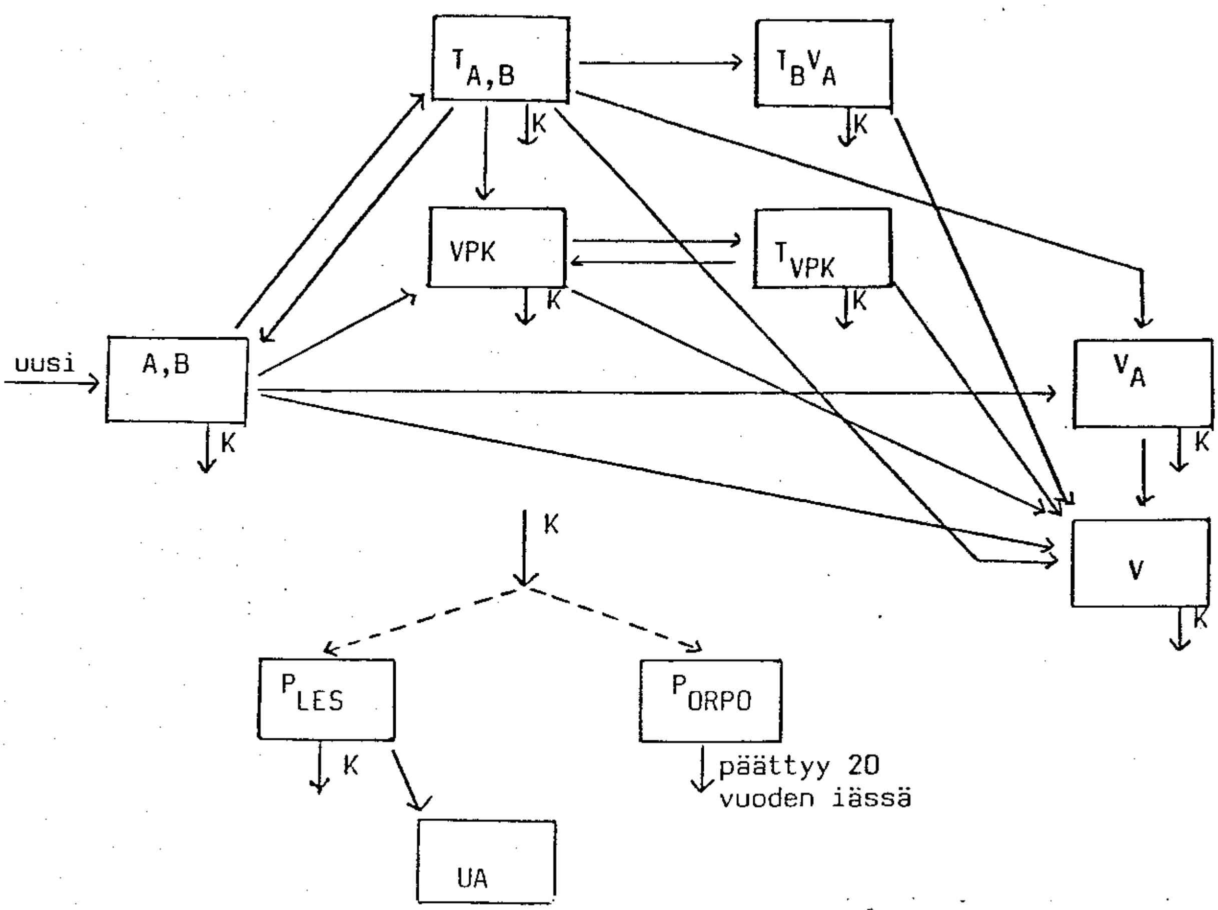
Työkyvytön voi mallin mukaan palata aktiiviksi, päästä vanhuuseläkkeelle A-osastosta, jolloin TEL:in mukainen työkyvyttömyys jatkuu (TB,VA), päästä TEL:in vanhuuseläkkeelle (VA-V), erota (VPK) tai kuolla. Myös vapaakirjalainen voi joutua työkyvyttömäksi (Tvpk).

Kun työntekijä kuolee, saattaa syntyä perhe-eläke lesken ja/tai lapsen osalle (Ples,Porpo). Perhe-eläke päättyy lesken uudelleenavioitumisen, kuolemaan tai lapsen täytettyä 20 vuoden iän.

Siirtymätodennäköisyydet syötetään ohjelmistolle parametritietoina. Todennäköisyydet ovat peräisin kassan omista tilastoista ja yleisistä, valtakunnallisista tilastoista.Todennäköisyydet arvotaan käyttäen satunnaisgeneraattoria. Arvonnassa käytettyä siemenlukua voidaan vaihdella, jolloin saadaan erilaisia satunnaispolkuja. Samalla siemenluvulla voidaan myös tietty simulaatio toistaa. Arvonta suoritetaan kerran jokaista vuotta kohden.

Systeemi generoi inputrekisterin perusteella uuden rekisterin, joka kuvaa kassan tilaa vuoden päästä lähtötilanteesta. Saatua rekisteriä käytetään taas seuraavan kierroksen inputrekisterinä.

Tarkempi selvitys simuloinnin periaatteesta on esitetty liitteessä.



Kuva 2.
Simulointimallin periaate

Tilat:

A,B	=	aktiivi (A-osasto ja B-osasto)
VPK	=	vapaakirja
T	=	työkyvytön
V	=	vanhuuseläke
P	=	perhe-eläke
UA	=	uudelleenavioituminen
K	=	kuollut

3.2 Simulaatiot

Jotta eri vuosien eläkemenot olisivat vertailukelpoisia keskenään, oletettiin inflaatioksi nolla prosenttia. Näin ollen aktiivien palkkataso muuttui ainoastaan väestön ikärakenteen muuttumisen myötä. Samalla perusteella pidettiin aktiivien lukumäärä muuttumattomana eli jokaisen poistuneen aktiivin tilalle otettiin uusi henkilö.

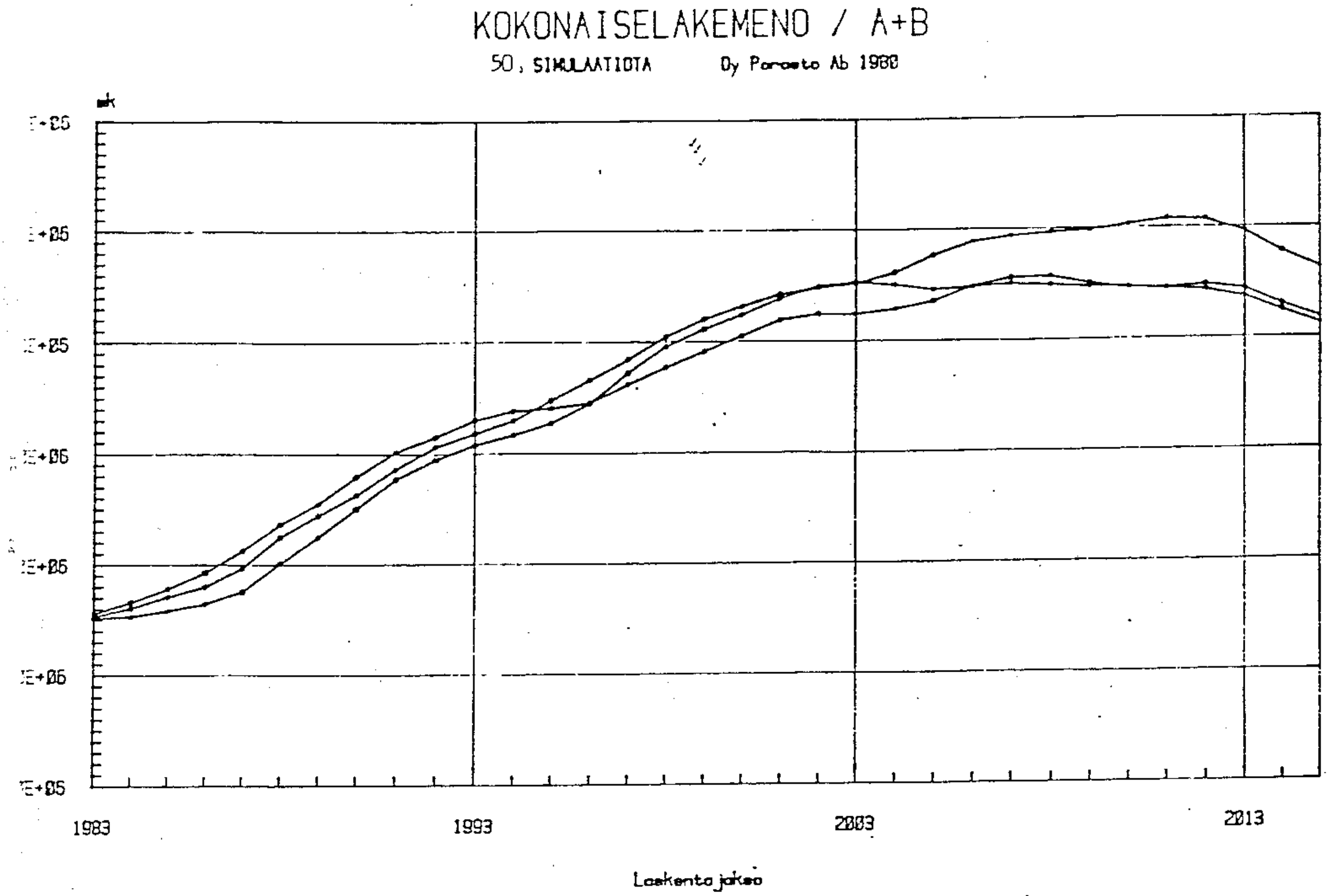
Simulaatiot suoritettiin 50 kertaa. Jokaisesta ajosta syntyivät aikasarjat kaikista eläkelajeista erikseen sekä kokonaiseläkemenosta. Ohjelma talletti nämä tietokoneen levymuistiin jatkokäsittelyä varten.

Systeemin toimintaa havainnollistaa alla oleva taulukko. Taulukosta nähdään, kuinka systeemi generoi eri eläketapahtumia. Vanhuuseläkkeiden lukumäärä (VA) näyttää hitaasti kasvavan, samoin työkyvyttömyyseläkkeiden lukumäärä (TK). Eroamisia (ER) ja kuolleita (K) näyttää sen sijaan syntyvän melko satunnaisesti. Viimeisessä sarakkeessa ovat kuntoutuneet työkyvyttömät (KUNT).

VUOSI	KANTA	K	VA	TK	ER	UUDET	KUNT
1983	1696	2	22	6	18	46	2
1984		4	20	8	30	62	0
1985		12	14	16	28	70	0
1986		4	30	6	34	74	0
1987		4	34	12	22	70	2
1988		8	46	18	16	88	0
1989		10	42	20	20	92	0
1990		4	48	20	26	98	0
1991		6	48	20	14	86	2
1992		0	42	12	40	92	2

Taulukko 1. Aktiiviväestön muutokset eräässä simulaatiossa 10 vuoden ajalta.

Kuvasta 3 nähdään, kuinka ennusteet poikkeavat toisistaan käyrien leikatessa toisensa monta kertaa.

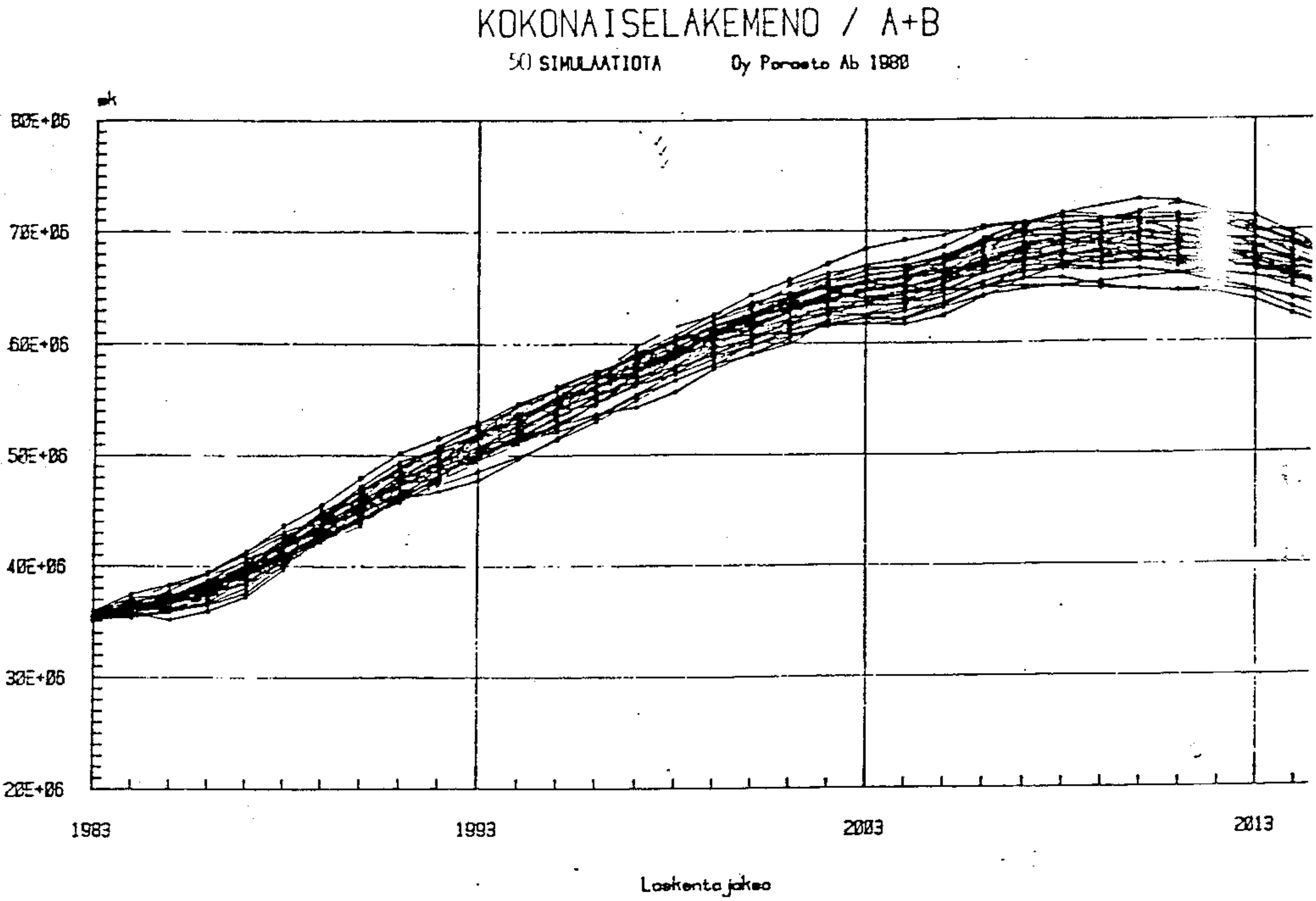


Kuva 3. Simulointimallin käyttäytyminen 3 simulaation perusteella

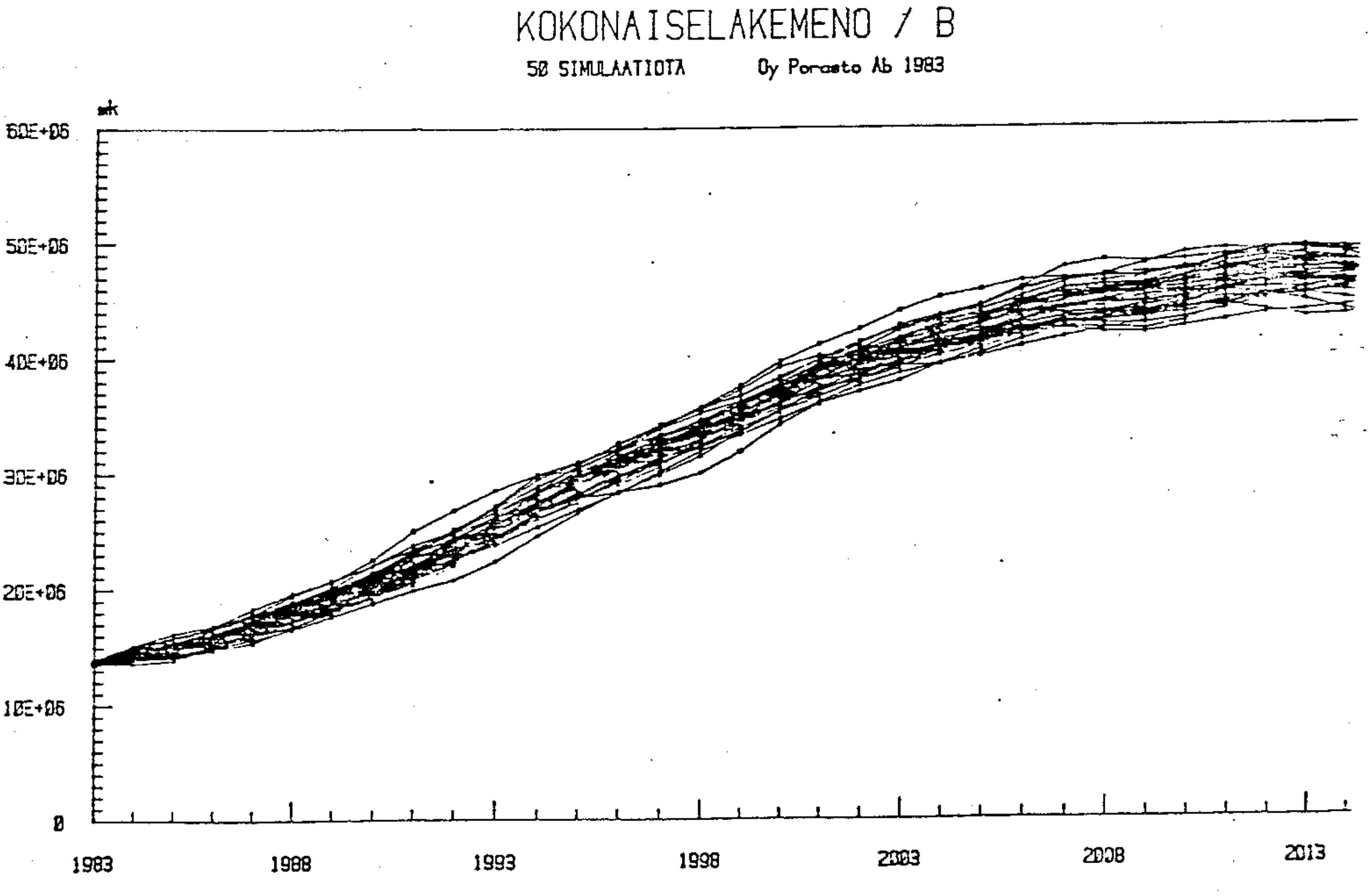
Eläkemenoennusteet on esitetty graafisesti kuvissa 4-9. Käyrät on piirretty kokonaiseläkemenosta, vanhuuseläkemenosta ja työkyvyttömyyseläkemenosta. Perhe-eläke jätettiin erittelemättä. Todettakoon, että perhe-eläkkeen merkitys on samaa suuruusluokkaa kuin työkyvyttömyyseläkemenon eli lähtövuonna noin 20 % kokonaiseläkemenosta. Tulostus suoritettiin sekä koko eläkejärjestelmästä (A+B) että TEL-osastosta (B).

Kuvista 4 ja 5 voidaan havaita, kuinka kokonaiseläkemenoennusteen ja vanhuuseläkemenoennusteen viuhkan leveys kasvaa ennustejakson pidentyessä. Koska kyseessä on kuvitteellinen eläkekassa, ei kovin syvällisiä analyysejä kannata tehdä käyrien käyttäytymisestä. On kuitenkin mielenkiintoista havaita, että kassan kokonaiseläkemeno saavuttaa huippunsa vuoden 2010 paikkeilla, kun taas TEL-eläkemeno jatkaa kasvamistaan, tosin hidastuneena.

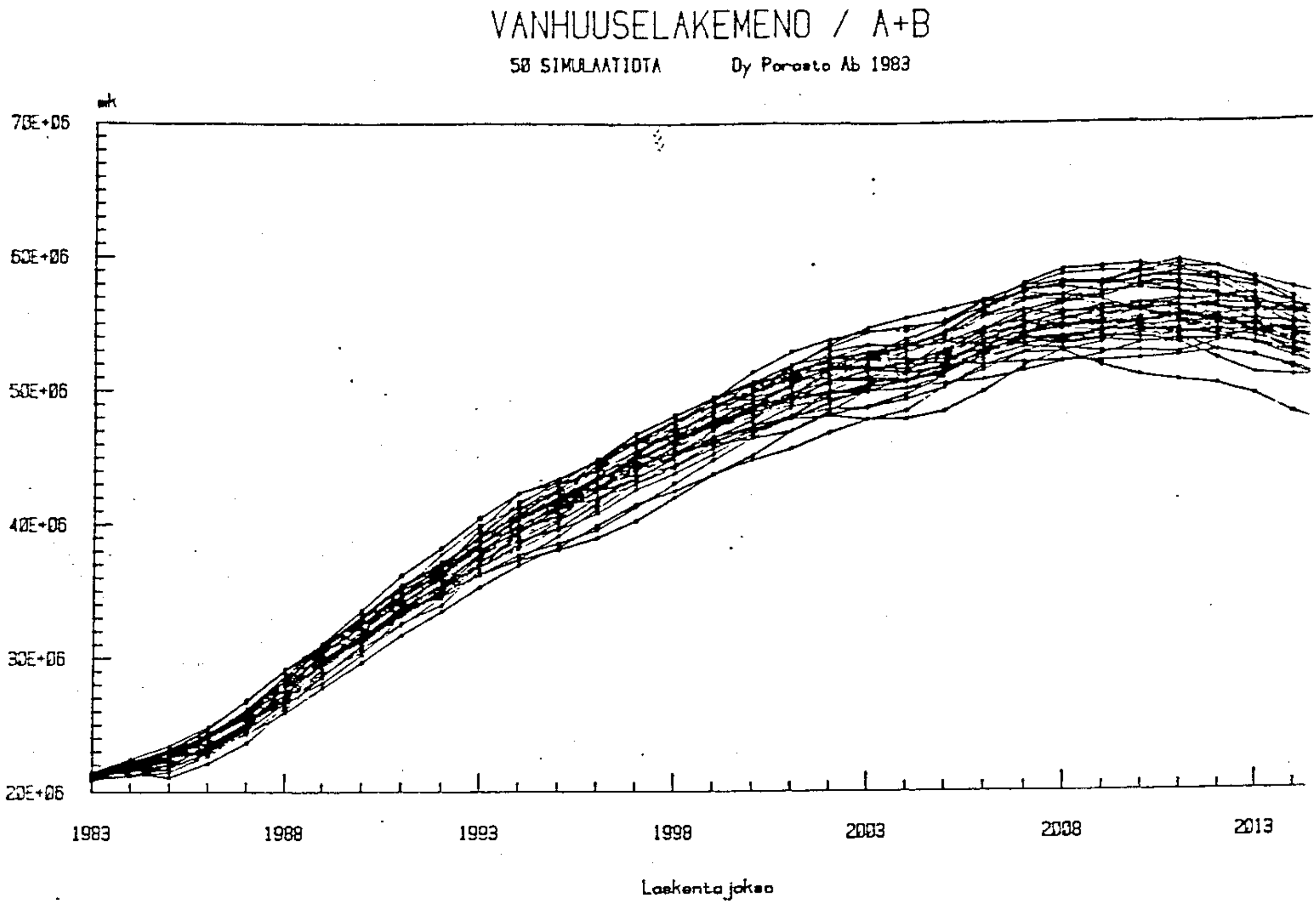
Työkyvyttömyyseläkemenolla näyttää olevan pitkällä tähtäimellä hieman aleneva trendi. Tämä johtunee väestön ikärakenteen muuttumisesta. Eroamistodennäköisyys on mallissa melko suuri, jolloin kassan uusien henkilöiden "vinosta" ikärakenteesta johtuen kassan keski-ikä alenee.


KOKONAISELAKEMENO / A+B
50 SIMULAATIOTA
Oy Porasto Ab
mk
80E+06
70E+06
60E+06
50E+06
40E+06
30E+06
20E+06
1983
1993
2003
2013
Laskentajakso

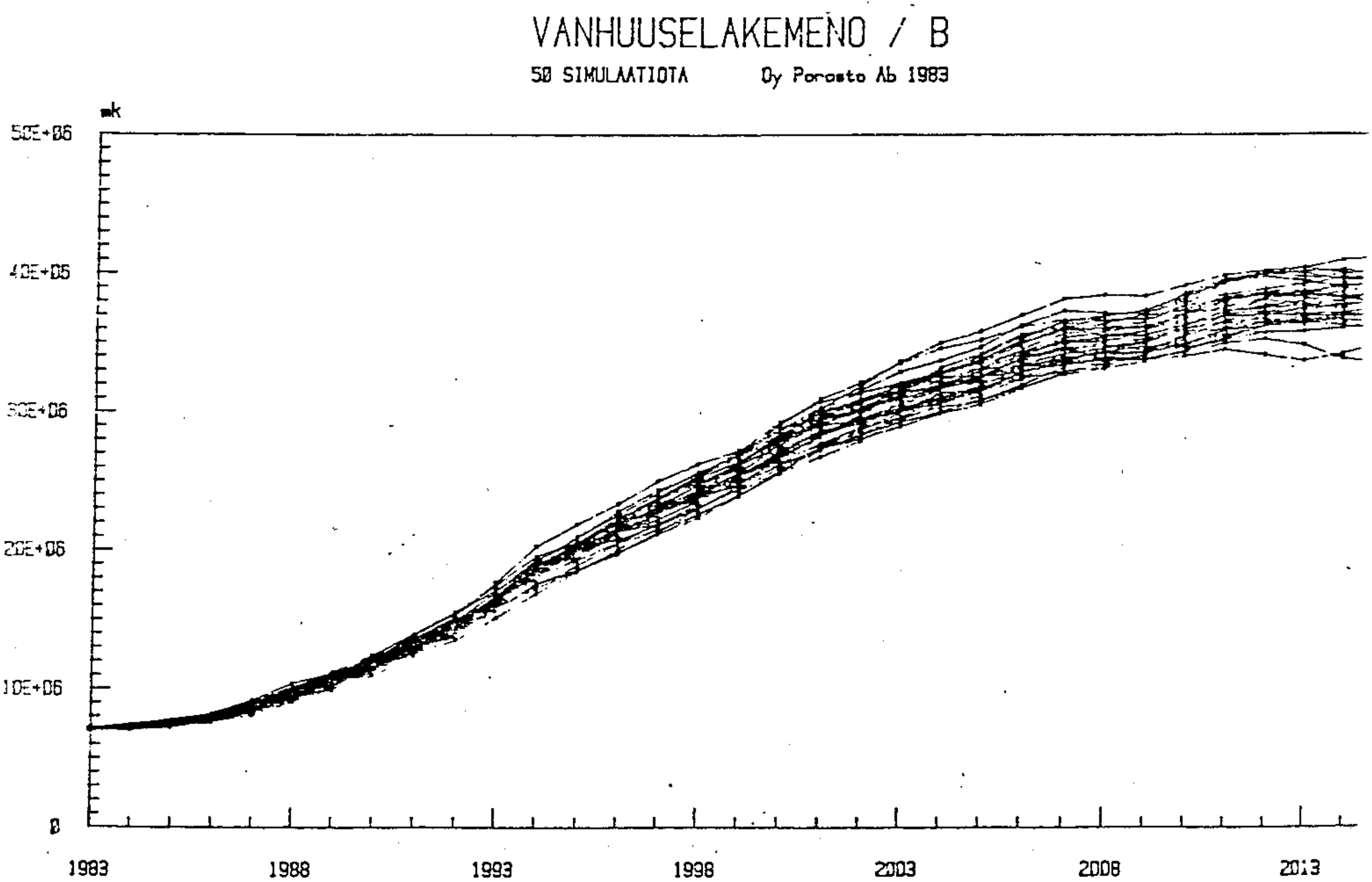
Kuva 4.


KOKONAISELAKEMENO / B
50 SIMULAATIOTA
Oy Porasto Ab 1983
mk
60E+06
50E+06
40E+06
30E+06
20E+06
10E+06
0
1983
1988
1993
1998
2003
2008
2013
Laskentajakso

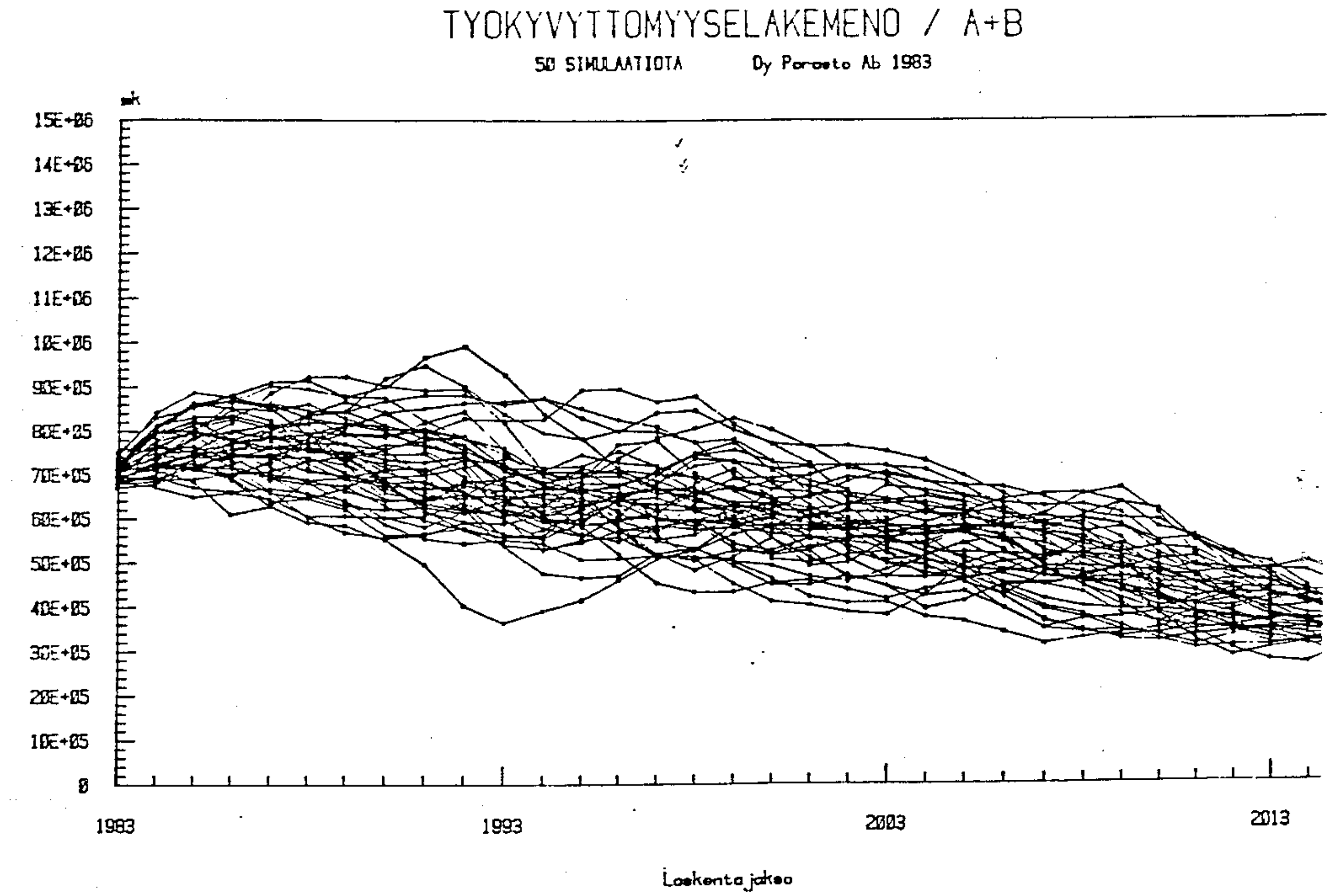
Kuva 5.


VANHUUSELAKEMENO / A+B
50 SIMULAATIOTA
Oy Porasto Ab 1983
mk
70E+06
60E+06
50E+06
40E+06
30E+06
20E+06
1983
1988
1993
1998
2003
2008
2013
Laskentajakso

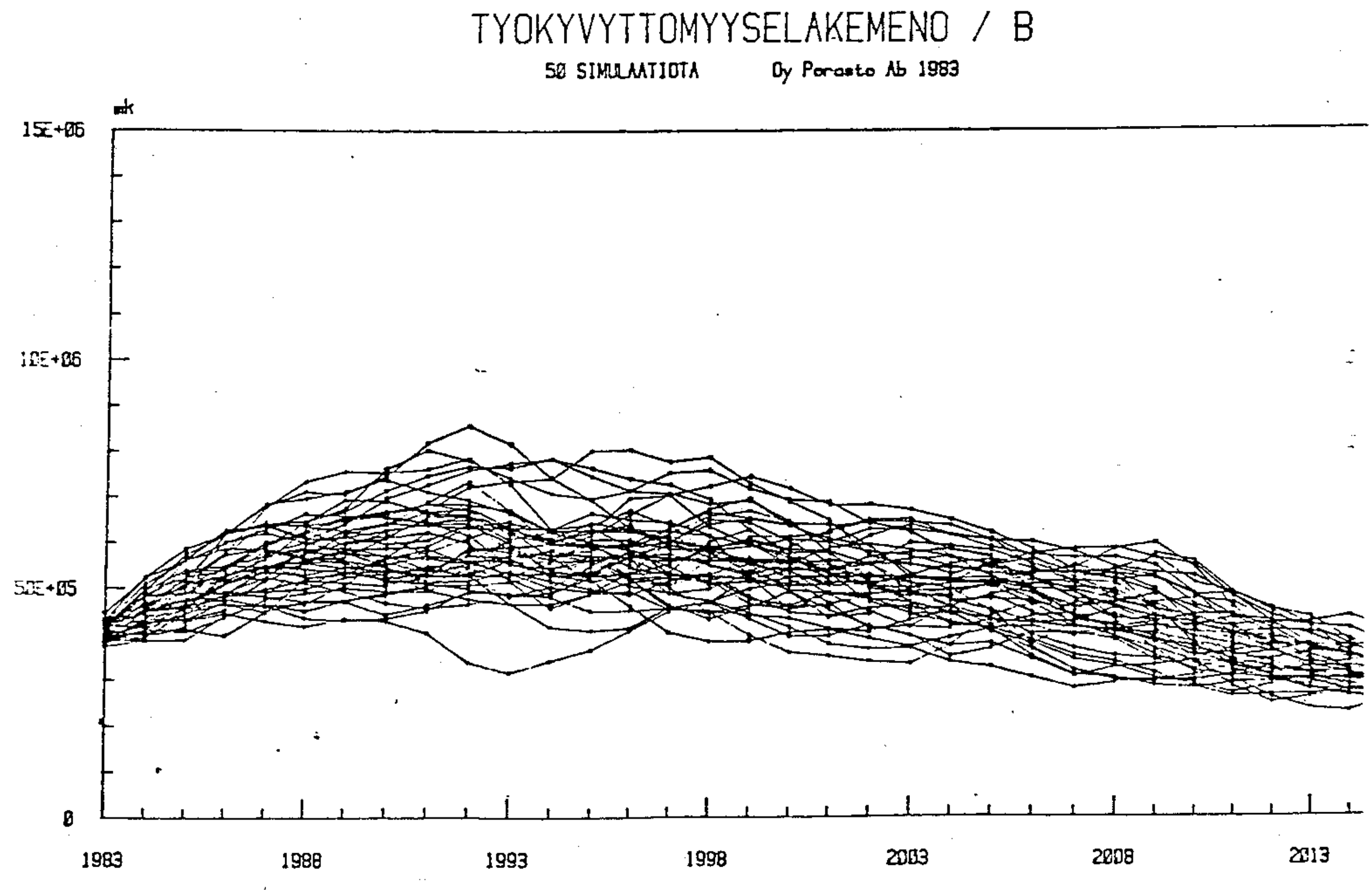
Kuva 6.


VANHUUSELAKEMENO / B
50 SIMULAATIOTA
Oy Porasto Ab 1983
mk
50E+06
40E+06
30E+06
20E+06
10E+06
0
1983
1988
1993
1998
2003
2008
2013

Kuva 7.


TYOKYVYTTOMYYSELAKEMENO / A+B
50 SIMULAATIOTA Oy Poraste Ab 1983
mk
15E+06
14E+06
13E+06
12E+06
11E+06
10E+06
90E+05
80E+05
70E+05
60E+05
50E+05
40E+05
30E+05
20E+05
10E+05
0
1983
1993
2003
2013
Laskentajakso

Kuva 8.


TYOKYVYTTOMYYSELAKEMENO / B
50 SIMULAATIOTA Oy Poraste Ab 1983
mk
15E+06
10E+06
50E+05
0
1983
1988
1993
1998
2003
2008
2013

Kuva 9.

Laskentajakso

Simulaatioiden antama informaatio eläkemenon kehityksestä on riittävä monenlaisten johtopäätösten tekemiseen. Usein on kuitenkin mukavampaa käsitellä yhtä ennustekäyrää kuin koko simulaatiorealisaatioiden joukkoa. Eräs ennuste keskiarvokäyräksi ja hajontakäyräksi saadaan seuraavasti.

Tarkastellaan aluksi vuoden t eläkemenoennusteiden muodostamaa joukkoa $X = (x_1, x_2, \ldots x_n)$. Olkoon $\bar{x}$ otoksesta X laskettu keskiarvo, s vastaava keskihajonta ja m_k k. keskeismomentti:

$$\bar{x} = \frac{1}{n}\sum_{1}^{n} x_i$$

$$s = \sqrt{\frac{1}{n-1}\sum_{1}^{n}(x_i - \bar{x})^2}$$

$$m_k = \frac{1}{n}\sum_{1}^{n}(x_i - \bar{x})^k$$

Tällöin voidaan määritellä jakauman (joukko X) vinous (g1) ja huipukkuus (g2) seuraavasti:

$$g1 = \frac{m_3}{s^3}$$

$$g2 = \frac{m_4}{s^4} - 3$$

Taulukkoon 2 on laskettu vinous ja huipukkuus eläkelajeittain. Koska molemmat tunnusluvut ovat kaikkina vuosina varsin lähellä nollaa, on perusteltua olettaa otoksen X jakautuvan likimain normaalisti $N(\mu,\sigma^2)$ kaikkina vuosina $t \in (1983,\ldots,2015)$. Tällöin μ:n ja σ:n estimaatit $\mu = \bar{x}$ ja $\sigma = s$.

Koska huipukkuusluvut ovat pääosin negatiivisia, on jakauma hieman "leveä".

Vuosi	Kokonais-eläkemeno g1	Kokonais-eläkemeno g2	Vanhuus-eläkemeno g1	Vanhuus-eläkemeno g2	Työkyvyttömyys-eläkemeno g1	Työkyvyttömyys-eläkemeno g2
1983	.2333	-.6614	-.5544	.0615	.3341	-.0918
1984	.3609	-.4602	-.2865	-.3507	.2349	-.8536
1985	.3692	-.1905	-.0180	-.7822	.2083	-.7878
1986	.0876	-.5426	.0088	-.8668	.1921	-.6426
1987	-.0505	-.5233	.1329	-.1534	.0970	-.7415
1988	-.0552	-.2452	.2095	.0234	-.0241	-.1910
1989	.1022	-.5798	-.0986	-.1712	.0073	-.4529
1990	.4159	-.4447	.0894	-.2129	.1778	-.7633
1991	.5729	.4016	.3234	-.1167	.3120	-.5588
1992	.3746	.9495	.4031	.6328	.1687	.2327
1993	.2453	1.2925	.2409	.9235	.0618	.6798
1994	.3460	.8630	.0067	1.5198	.3516	.8211
1995	-.1104	-.2937	-.1619	.8369	.2769	.9291
1996	-.3207	-.3447	-.2234	-.0318	.3747	.5050
1997	-.3897	.3825	-.0512	-.3449	.3600	-.2416
1998	-.4217	1.1434	-.0289	-.6026	.3206	-.0705
1999	-.1308	.7509	-.0917	-.7559	.2318	-.3794
2000	.1587	.4672	-.1118	-.7790	-.0513	-.2656
2001	.0452	.0837	-.0905	-.6986	-.0808	-.2878
2002	.2509	.2529	-.0302	-.5575	-.0632	-.3287
2003	.4615	.4463	.1900	-.1816	.0713	-.3912
2004	.5769	.4827	.4325	.2545	-.0540	-.5399
2005	.5271	-.0738	.4534	.0821	-.2185	-.8285
2006	.3625	-.5386	.4395	-.2754	-.3799	-.8125
2007	.2691	-.6730	.5435	-.3357	-.3328	-.6001
2008	.1972	-.5912	.6049	-.3788	-.1346	-.7178
2009	.0703	-.7146	.4188	-.9075	.1601	-.6785
2010	.0575	-.7518	.3063	-.9329	.3778	-.6467
2011	-.0301	-.6724	.3035	-.8098	.3036	-.9620
2012	-.0711	-.5491	.0717	-.5120	.1194	-.9773
2013	-.2225	-.3257	-.3358	.0327	.1257	-.8925
2014	-.3589	-.1182	-.5010	.5478	.3030	-.3976
2015	-.3566	-.0097	-.4762	.7597	.3800	-.2157

Taulukko 2. Vinous (g1) ja huipukkuus (g2) eläkelajeittain (TEL-osasto)

Testataan vielä jakauman normaalisuutta Kolmogorovin yhteensopivuustestillä. Tämän testin määrittelemä testisuure D on

$$D = \sup_x | F_0(x) - F_n^*(x) |$$

missä

$F_0(x)$ on oletettu teoreettinen jakauma ,

$F_n^*(x) = \frac{i}{n}$,kun $x_{(i)} < x < x_{(i+1)}$, $i=0,1,\ldots,n$
ja $x_{(i)}$ on joukon X järjestetty otos

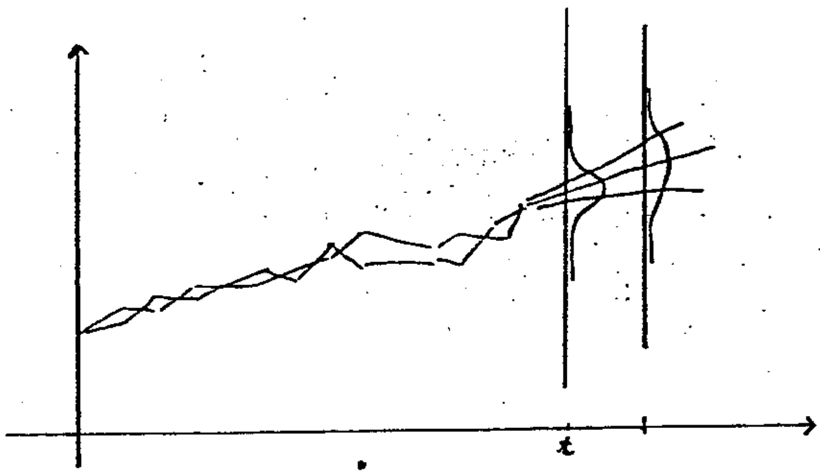
Testi suoritettiin eläkelajeittain TEL-eläkemenosta 49 simulaatiolle vuosina 1983 ja 1999. Tuloksena saatiin alla olevassa taulukossa näkyvät D-arvot.

laji	vuosi	
	1983	1999
0	0.0736	0.0859
2	0.0534	0.0921
3	0.0725	0.0654

Taulukko. Kolmogorov-testi,TEL-eläkkeiden D-arvot
(0=kokonaiseläkemeno,2=tk-eläkemeno
3=vanhuuseläkemeno)

Kriittinen piste D =0.1984 5%:n merkitsevyystasolla, kun n=47. Koska kaikki yllä olevat D-arvot ovat tätä pienempiä, voidaan nollahypoteesi "otos jakaantuu normaalisti $N(\mu,\sigma^2)$ " hyväksyä eli jakaumat ovat normaalisia.

Simulaatioiden normaalisuus on esitetty graafisesti kuvassa 10.



Kuva 10. Simulaatioiden otosarvot vuonna t oletetaan normaalisti jakautuneiksi

Seuraavaksi on tutkittu odotusarvon μ luotettavuutta $\bar{X}$:n arvioidun keskivirheen sekä luottamusvälin avulla. $\bar{x}$:n keskivirhe saadaan kaavasta $s/\sqrt{n}$. Kun tuntemattomat parametrit μ ja σ^2 on estimoitu kuten yllä ($\mu=\bar{x}, \sigma^2=s^2$), voidaan laskea μ:n luottamusvälit seuraavasti.

Muuttuja $t=\frac{\bar{x}-\mu}{s/\sqrt{n}}$ noudattaa Studentin jakaumaa vapausastein n-1. Tällöin voidaan konstruoida μ:lle luottamusväli:

$$\left(\bar{x}-t_{\xi}\frac{s}{\sqrt{n}}\ ,\ \bar{x}+t_{\xi}\frac{s}{\sqrt{n}}\right) .$$

Edellä t_{ξ} on t-jakauman kaksisuuntainen kriittinen arvo vapausastein f=n-1. Taulukoissa 3-5 on laskettu eläkelajeittain μ:lle 95 %:n luottamusvälit.

Hajonnan estimaattorin s luotettavuutta on mitattu variaatiokertoimella $s/\bar{x}$, mikä siis ilmoittaa , montako prosenttia hajonta on keskiarvosta (taulukot 3-5).

Kuvaan 11 on piirretty TEL-osaston kokonaiseläkemenon keskiarvo- ja hajontakäyrät vuosittain.

Vaihtoehtoinen tapa ja mikäli jakauma on tuntematon, ainoa tapa on ns. ei-parametristen menetelmien käyttö. Taulukoissa 6 ja 7 on laskettu TEL-osaston kokonaiseläkemenosta ja vanhuuseläkemenosta mediaani (keskimmäinen havainto) sekä pienin ja suurin havainto (MIN , MAX).

Koska työkyvyttömyyseläkemenon hajonta on melko suuri, ei kyseistä ennustetta ole tarkasteltu lainkaan keskiarvon ja hajonnan avulla (vaikka normaalisuusolettamus pätisikin, ovathan g1 ja g2 melko pieniä myös tässä tapauksessa). Kuvassa 12 on työkyvyttömyyseläkemenon ennusteista piirretty mediaani, pienin ja suurin arvo sekä ala- ja yläkvartiilit.

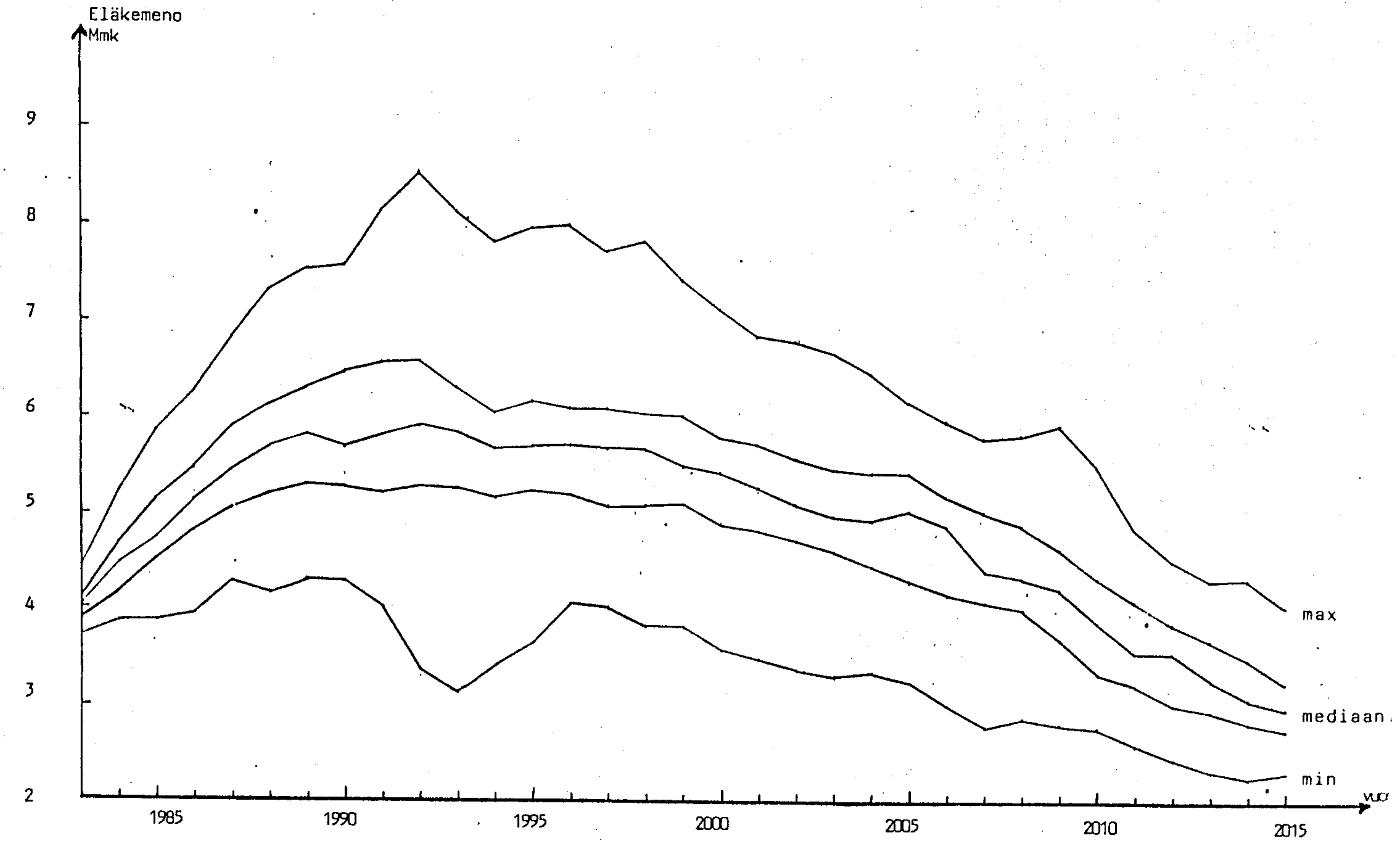
VUOSI	$\bar{X}$	$\bar{X}$:N KVIR ($S/\sqrt{n}$)	$\bar{X}-T_{.05}*S/IN$	$\bar{X}+T_{.05}*S/IN$	S	VARIAATIO ($100*S/\bar{X}$)
1983	13688	20	13648	13728	141	1.0
1984	14314	52	14211	14418	365	2.6
1985	14933	71	14790	15076	502	3.4
1986	15745	78	15588	15901	552	3.5
1987	16847	89	16668	17026	631	3.7
1988	17990	96	17797	18182	677	3.8
1989	19098	98	18900	19295	695	3.6
1990	20377	117	20141	20613	830	4.1
1991	21926	138	21650	22203	974	4.4
1992	23497	146	23205	23790	1029	4.4
1993	25354	145	25063	25645	1024	4.0
1994	27295	142	27009	27580	1005	3.7
1995	28938	139	28658	29218	985	3.4
1996	30565	145	30274	30857	1025	3.4
1997	32079	151	31775	32383	1071	3.3
1998	33521	154	33212	33829	1086	3.2
1999	34975	148	34678	35272	1045	3.0
2000	36694	147	36399	36989	1038	2.8
2001	38210	145	37918	38502	1028	2.7
2002	39348	146	39054	39641	1034	2.6
2003	40473	164	40143	40803	1162	2.9
2004	41455	167	41118	41791	1184	2.9
2005	42309	168	41971	42647	1189	2.8
2006	43425	180	43064	43786	1271	2.9
2007	44296	193	43909	44683	1363	3.1
2008	44614	198	44215	45012	1402	3.1
2009	44922	204	44513	45331	1441	3.2
2010	45562	206	45149	45975	1453	3.2
2011	46280	197	45883	46676	1396	3.0
2012	46742	195	46351	47134	1377	2.9
2013	46923	201	46519	47326	1421	3.0
2014	46912	195	46519	47304	1381	2.9
2015	46780	187	46405	47156	1322	2.8

Taulukko 3. Tunnuslukuja
TEL-eläkkeet, kokonaiseläkemeno

[illegible]
TUNNUSLUKUJA
SIMULOINTEJA 50 KPL
SKAALAUS 1000 MK

VUOSI	X	S	MEDIAANI	MIN	MAX
1983	7099	59	7101	6936	7193
1984	7283	127	7281	6967	7531
1985	7501	165	7516	7166	7858
1986	7888	202	7879	7494	8271
1987	8646	246	8633	8055	9209
1988	9618	282	9621	8974	10393
1989	10558	285	10556	9889	11182
1990	11683	315	11632	10911	12430
1991	13082	367	13030	12319	13954
1992	14458	410	14440	13466	15499
1993	16379	472	16379	15092	17638
1994	18424	570	18403	16834	20203
1995	19882	653	19915	18390	21813
1996	21392	762	21468	19668	23288
1997	22835	846	22844	21115	24958
1998	24173	879	24131	22307	26200
1999	25575	863	25544	23855	27179
2000	27302	936	27267	25481	29184
2001	28808	991	28963	26681	30897
2002	29882	1001	30014	27836	32061
2003	30948	1036	31086	28850	33541
2004	31855	1072	31812	29779	34918
2005	32610	1113	32520	30467	35744
2006	33764	1189	33721	31606	36911
2007	34710	1225	34454	32653	38117
2008	34960	1219	34617	33077	38407
2009	35312	1221	35063	33591	38332
2010	36146	1312	35986	33954	39090
2011	37028	1374	36927	34436	39835
2012	37558	1402	37442	34042	40197
2013	37778	1430	37756	33618	40378
2014	37880	1431	37957	33734	40869
2015	37889	1403	37978	33472	41148

Taulukko 7 – TEL-osaston vanhuuseläkemeno: tunnuslukuja
(X = odotusarvon estimaatti,
S = hajonnan estimaatti)



Kuva 12. TEL-osaston työkyvyttömyyseläkemenoennuste: mediaani, ala- ja yläkvartiilit sekä kunkin vuoden pienin ja suurin ennustettu luku

4 Yhteenveto

Edellä esitellyn menetelmän mukaan eläkemeno voidaan ennustaa pitkälle tulevaisuuteen. Mikäli simulointeja suoritetaan tarpeeksi paljon, saadaan melko luotettava kuva ennusteesta. Edellytyksenä luotettavuudelle on kuitenkin se, että käytetyt tilastot vastaavat todellisuutta. Tämän takia onkin kiinnitettävä erityistä huomiota tutkimusta valmisteleviin töihin: tilastojen keruuseen ja testaukseen.

Simulointitekniikasta johtuva atk-prosessin hitaus asettaa omat rajoituksensa menetelmän käytölle. Toisaalta varsinkin kokonaiseläkemenon kehitykselle saavutetaan melko tarkka ennuste jo varsin vähäisillä simulaatioiden määrillä.

Eläkemeno on merkittävin menoerä eläkekassan tilinpäätöksessä. Tämän tutkimuksen pääasiallinen tulos on eläkemenoennusteet ja niiden luotettavuus. Lisäksi syntyneet rekisterit ovat myös käytettävissä mahdollisia jatkotutkimuksia varten, esimerkiksi erilaisten rahoitusvaihtoehtojen määräämisen. Tällöin on vain pystyttävä ennustamaan rahastojen kehitys.

Liite: Simulointimallin periaate
(Otteita Martti Laihon pro gradu-tutkielmasta Työeläkevakuutuksen yksilöpohjainen simulointimalli)

4 SIMULOINTIMALLI

4.1 Mallin valintaperusteet

Tarkastelemme eläkelaitoksen toimintapiiriä ja vakuutuskantaa vakuutettujen henkilöiden ja (kuolleiden vakuutettujen) edunsaajien dynaamisina joukkoina. Molempien joukkojen alkiot vanhenevat jatkuvasti ajan suhteen, vaihtavat tilaa tai poistuvat joukosta jonkin satunnaisen tai ennalta määrätyn tapahtuman johdosta. Joukkoihin tulee satunnaisesti uusia alkioita, edunsaaja-alkioita kuitenkin vain vakuutetun alkion poistuttua (kuoltua).

Luvun 3 virtausmallissa on kuvattu eläkevakuutuksen kannalta laadittu vakuutettujen tilaluokitus, jonka mukaan vakuutettujen joukko voidaan jokaisena hetken jakaa yksikäsitteisesti tiloja vastaaviin osajoukkoihin, ja henkilövirrat, joiden mukaan vakuutetut voivat vaihtaa tilaansa. Jatkuva-aikainen virtausmalli ei kuitenkaan ole käyttökelpoinen kuvaamaan henkilöiden diskreettejä varantoja ja edestakaisia virtauksia luvussa 3 tarkasteltujen selitysmallie mukaan. Diskreetti demografisten kohorttimallien pohjalta kehitetty virtausmalli puolestaan vaatisi liian suuret taulukot ja olisi raskas muutettava uusien vaatimusten noustessa esiin.

Joustavimmin kaikki yksilölliset käsittelyvaatimukse kuten esim. liukuvan eläkeiän, perus- ja lisäeläkevakuutuksen vähimmäis- ja yhteensovitussäännökset, sääntömuutoksista johtuvat voimaantulo- ja vanhojen jäsenten etujen siirtymäsäännökset jne ovat toteutettavissa yksilöpohjaisella diskreetillä simulointi mallilla, jossa kutakin vakuutettujen tai edunsaajie joukkoon kuuluvaa henkilöä vastaa erillinen laskenta alkio, tietue, tarpeellisine ominaisuustietoineen.

Diskreetit simulointimallit ovat yleensä tapahtumaohjattuja eli aika etenee niissä satunnaisin välein tapahtumasta toiseen /11. s 17/. Eläkevakuutuksen piirissä vakuutustapahtumia, joiden johdosta vakuutettu siirtyy toiseen tilaan, sattuu kuitenkin vain harvoin saman henkilön kohdalla. Toisaalta jokainer laskenta-alkio (eronneita vakuutettuja lukuunottamat on käsiteltävä palkka- ja eläkesummien keräämistä varten kunkin laskentavuoden osalta erikseen ja laskenta-alkioita käsitellään toisistaan riippumattc mina, joten laskenta-alkiot on yksinkertaisinta käyc läpi peräkkäisjärjestyksessä kunkin laskentavuoden kohdalla ja käsitellä alkiot intervalliohjattuina kohorttimallien tapaan. Tällöin simulointimallin ajoitusmekanismi ja seuraavan toteuttavan tapahtuman etsintätekniikka /11. ss 31-33/ jäävät niin yksinkertaisiksi, että malli voidaan ohjelmoida yleisillä ohjelmointikielillä.

Näin päädytään simulointimalliin, joka koostuu dynaamisesta joukosta kohdassa 4.2 kuvattavia henkilömalleja.

4.2 Henkilömalli

Yksittäisen eläkevakuutuksen eri vaiheita voidaan kuvata kuvion 4.2.1 mukaisella kaaviolla, jota tässä kutsutaan henkilömalliksi. Oikeampi nimi olisi ehkä vakuutuksen vaihemalli, mutta nimi henkilömalli korostaa enemmän sitä, että tarkastelun kohteena on yksittäinen vakuutettu ja kukin hänen mahdollinen edunsaajansa.

Tietue

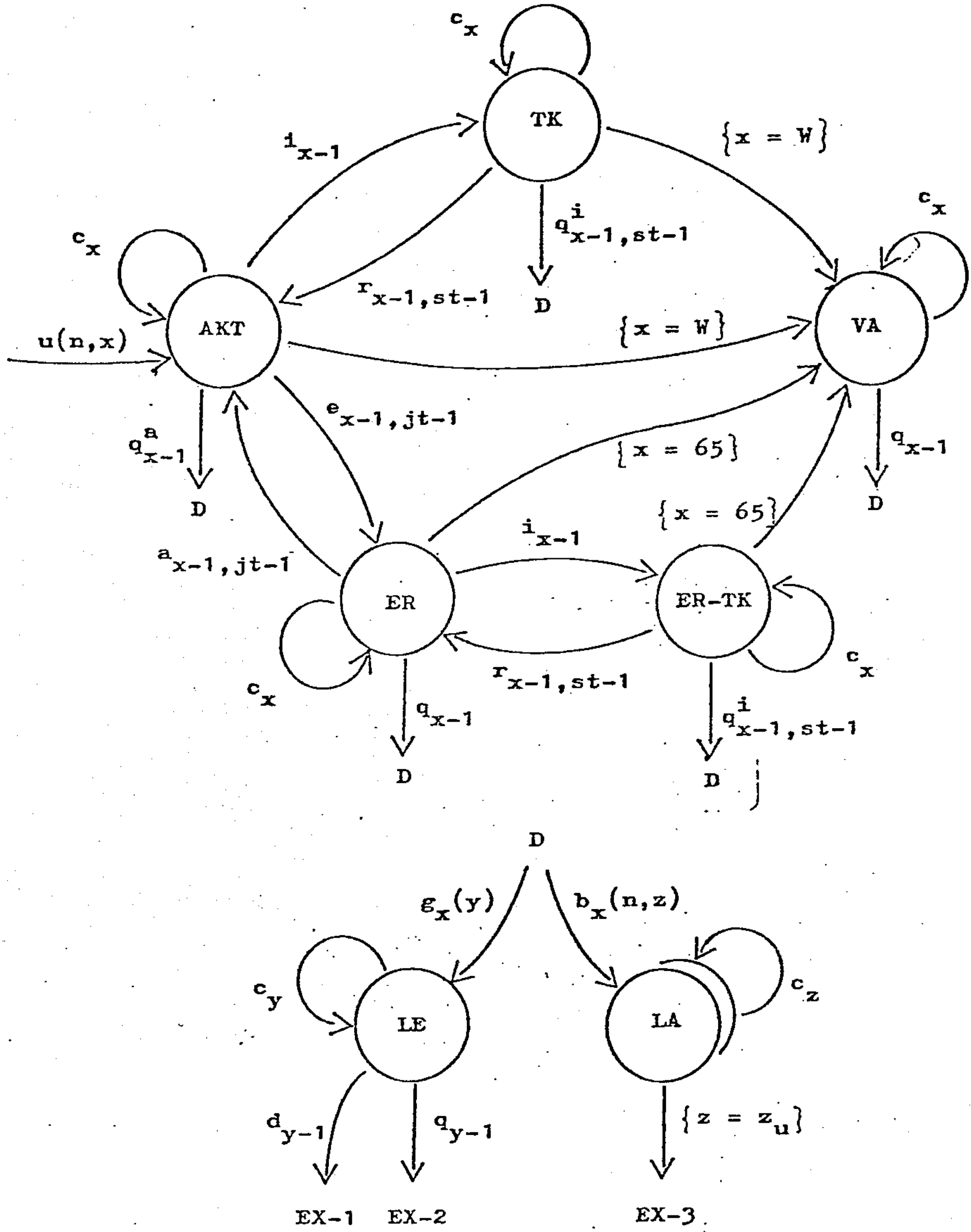
Laskenta-alkio kuvataan taulukossa 4.2.1 luetelluill tiedoilla, joita voidaan lisätä tarpeen mukaan. Näistä palkkatietoa muutetaan aktiiveilla vuosittain palkkakehitysmallin mukaan ja eläke- ja vapaakirjatietoa korjataan eläkeindeksisarjan mukaan. Jäsenyyden muutosvuodeksi Vo asetetaan uudelle jäsenelle sekä jäsenyyteen palaavalle työhöntulovuosi ja jäsenyydestä eroavalle eroamisvuosi. Viimeinen tila muutosvuosi Po asetetaan aina tilan vaihtuessa.

Taulukko 4.2.1

tunnus	selitys
POS	henkilöryhmän numero
JAS	jäsenyyskoodi
TIL	tilakoodi, joka yhdessä jäsenyyskoodin kanssa yksilöi kaaviossa 4.2.1 kuvatun tilan
X_o	syntymävuosi
V_o	viimeinen jäsenyydenmuutosvuosi
P_o	viimeinen tilanvaihdosvuosi
SUK	sukupuolikoodi
S	palkka mk/kk
E	eläke mk/kk
. . .	... muita laskennassa tarvittavia tietoja

Laskenta-alkion perusteet

Kuvio 4.2.1 Henkilömallin tila- ja siirtymäkaavio



Laskenta-alkion perustiedoista lasketaan kunakin laskentavuonna v ennen alkion käsittelyä taulukossa 4.2.2 luetellut suureet.

Taulukko 4.2.2

tunnus	laskusääntö	tiedon selitys
x	$v - X_0$	ikä vuonna v
jt	$v - V_0$	nykyisen jäsenyyden tai poissaolon kesto
st	$v - P_0$	nykyisen tilan kesto (lähinnä työkyvyttömillä työkyvyttömyyden kesto)

Laskenta-alkion laskettavat ominaisuustiedot

Tilat

Tässä tutkielmassa rajoitumme tarkastelemaan taulukossa 4.2.3 lueteltuja vakuutetun ja edunsaajie tiloja. Tilajoukon valitseminen näin ei ole kuvatta van simulointitekniikan kannalta oleellista, vaan tilajoukko voidaan välita kulloinkin tarkasteltavien eläke- tai avustusetujen ja käsittelytarkkuuden muka

Kaaviossa 4.2.1 kuvattu tila D on poistumatila, joho siirtynyt laskenta-alkio poistetaan vakuutettujen joukosta ja alkiolle perustetaan edunsaajien joukkoo leski- ja lapsialkiot luvussa 3 kaavailtujen f_x- ja ikäjakaumaperusteiden mukaan.

Jos laskenta-alkion tila laajennetaan tarkoittamaan myös tiettyä henkilöryhmää, ikää, työsuhteen (jäsenyyden) tai poissaolon kestoa sekä taulukossa 4.2.3 tarkoitetun nykyisen tilan kestoa, joilla tilasiirtymät luvussa 3 pääasiassa selitettiin, ja vanhuuseläkkeelle siirtyminen käsitellään myös todennäköisyysfunktiona, voidaan henkilömallin vakuutettua koskeva osa ja vastaavasti leski- ja lapsialkiot tulkita ensimmäisen asteen Markov-malliksi. Tällä on kuitenkin merkitystä vain mallin perusteita ja tulkintaa ajatellen. Rajoittumalla käsittelemään vain taulukossa 4.2.3 lueteltuja tiloja säilytämme mallin havainnollisuuden ja kunkin laskenta-alkion osalta ratkaistavien siirtymien määrä tulee Markov-mallin siirtymämatriisiin verrattuna hallittavammaks vaikkakin siirtymämallit monimutkaistuvat.

Taulukko 4.2.3

laskenta-alkio	tilan tunnus	tilan selitys
vakuutettu	AKT	aktiivi eli maksava jäsen
	TK	työkyvytön jäsen, saa yleensä tavoitevanhuuseläkkeensä suuruista työkyvyttömyyseläkettä, ansaits samalla vakuutusmaksuista vapautettuna vanhuuseläkettä
	ER	eronnut jäsen, on eläketapahtuman sattuessa oikeutettu vapaakirjojaan vastaavaan eläkkeeseen
	ER-TK	eronnut jäsen, joka on tullut työkyvyttömäk ja saa vapaakirjojaan vastaa työkyvyttömyyseläkettä
	VA	vanhuuseläkeläinen, saa vapaakirjojaan ja mahdollista eläkkeellesiirtymistä välittömästi edeltäneenä jäsenyysjaksona ansaittua eläkettä
	D	kuollut vakuutettu
leski	LE	edunsaajaleski, saa eläkettä, joka on puolet siitä, mihin edunjättäjä oli oikeutettu
	EX-1	uudelleenavioitunut leski, saa kertasuorituksena kahden vuoden eläkkeen, jonka jälkeen poistetaan edunsaaji joukosta
	EX-2	kuollut leski
lapsi	LA	lapsieläkkeensaaja
	EX-2	määräiän (18 vuotta) täyttän lapsieläkkeensaaja, poisteta edunsaajien joukosta

Laskenta-alkioiden tilat

Siirtymät

Intervalliohjatussa mallissa voidaan kullekin laskenta-alkiolle sallia vain yksi siirtymä eli tilanvaihdos laskennan aikaväliä kohden. Jos näitä siirtymiä tapahtuu todellisuudessa useammin merkittävässä määrin, on laskennan aikaväliä lyhennettävä ja lasketut siirtymätodennäköisyydet korjattava tätä vastaaviksi.

Teoriassa intervallipohjaisen mallin siirtymätodennäköisyydet tulee estimoida tarkkailemalla kohdejoukkoa hetkinä

$$t,\ t+1,\ t+2,\ \ldots$$

aina valitun aikavälin (mallissamme vuoden) kuluttua edellisestä tarkasteluhetkestä ja rekisteröiden vain näin havaitut siirtymät. Näin lasketut siirtymä todennäköisyydet vastaavat kohdejoukon "kehittymisnopeutta". Tätä siirtymätulkintaa vastaavia tilastoja ei kuitenkaan ole eläkevakuutuksen alalta saatavissa muista ilmiöistä kuin kuolleisuudesta. Muiden siirtymäilmiöiden tilastoja on sen vuoksi esim. tilastoimatta jääneiden kuolleiden osalta korjattava puolivuotiskuolevuudella.

Kaaviossa 4.2.1 ei esitetä eräitä merkitykseltään vähäisinä pidettyjä siirtymiä, joita ovat esimerkiksi siirtymät:

TK -> ER, AKT -> ER-TK, jne.

Kaaviossa 4.2.1 siirtymäfunktioiden argumenteissa merkintä "-1" osoittaa, että käytettävät yksivuotiset siirtymätodennäköisyydet lasketaan väestötieteessä ja vakuutusmatematiikassa omaksutun käytännön mukaan siten, että esimerkiksi kuolevuus q_x ilmoittaa x-ikäisen kuolintodennäköisyyden seuraavan ikävuoden kuluessa /12. s 9/.

Tilanvaihdoksen hetken t tilasta s_i hetken t+1 tilaan s_j käsittelemme tapahtuneeksi hetkellä t+1/2 siten, että laskenta-alkion katsotaan olleen tilassa s_i ajan (t, t+1/2] laskentavuodesta v ja tilassa s_j ajan (t+1/2, t+1] , missä t tarkoittaa laskentavuotta edeltäneen vuoden loppua ja t+1 laskentavuoden loppua.

Vakuutusmatematiikassa käytetään jatkuville todennäköisyysfunktioille tunnuksina kreikkalaisia kirjaimia ja diskreeteille yksivuotistodennäköisyyksille latinalaisia kirjaimia. Tämän kanssa yhdenmukaisesti rajoitumme käyttämään todennäköisyysfunktioiden merkinnöissä tavallisia pieniä kirjaimia. Funktiomerkintöjen alaviitteet ovat argumentteja, yläviitteet osoittavat tietylle tilalle eriytettyä todennäköisyyt ja sulkeisiin merkitään Eriksenin /8./ esimerkin muka funktion tulosparametrit.

Vakuutetun (diskreettiä) ikää merkitään tässä tutkielmassa x:llä, lesken ikää y:llä ja lapsen ikää z:lla Nämä lasketaan taulukossa 4.2.2 x:lle annetun säännön mukaan. Merkintä W tarkoittaa eläkeikää ja Z_u lapsieläkkeen päättymisikää.

Kaavion 4.2.1 yksittäiset funktiomerkinnät on lueteltu selityksineen taulukossa 4.2.4. Yläviite a tarkoittaa tässä työkykyisille ja i työkyvyttömille eriytettyä funktioa. Funktiomerkintä q on vakiintunut väestötieteeseen /21./ ja vakuutusmatematiikkaan /27, 29, 14 ja 12 — tarkoittamaan yksivuotista kuolevuutta tai -kuolleisuutta, mutta muille funktioille en ole löytänyt vakiintuneita yksivuotismerkintöjä.
Merkinnät e (tosin maastamuuttavuuden asemesta eroavuutena) ja r sekä perheellisyysperustemerkinnät b ja olen lainannut Eriksenilta /8./. Loput siirtymäfunktiomerkinnät olen kehittänyt kohdetilojen (a ja i) sekä sopivien jäljellä olevien kirjainten (c ja d) mukaan.

Kaarisulkeisiin merkitty ehto kaaviossa 4.2.1 tarkoittaa muut siirtymät, paitsi kuoleman, voittavaa siirtymää silloin, kun ehto on tosi.

Funktiot määritetään yleensä taulukoina, joita on mahdollista muuntaa tai vaihtaa kesken simuloinnin. Siirtymäfunktioiden taulukot samoin kuin palkkakehitysmallin perustetaulukko laaditaan erikseen jokaiselle henkilöryhmälle.

Uudet vakuutetut

Laskentavuonna ensikertalaiset jäsenet generoidaan vakuutettujen kantaan sen jälkeen, kun muut laskenta-alkiot on ensin käsitelty ko. vuoden osalta.
Mallille annetaan henkilöryhmäkohtaisesti aktiivien jäsenten tavoitemiehityksen aikasarja. Simulointiohjelma pitää kirjaa kunkin henkilöryhmän jäsenten määrästä ja uusia vakuutettuja generoidaan ryhmään sen verran, että ryhmän tavoitemiehitys tulee täytettyä. Uuden laskenta-alkion syntymävuosi ratkaistaan uusille jäsenille henkilöryhmittäin annettavan ikäjakauman mukaan ja palkaksi asetetaan ikää vastaava ko. henkilöryhmän keskipalkka.

Taulukko 4.2.4

funktio	selitys
q_x	todennäköisyys, että x-ikäinen henkilö kuolee seuraavan ikävuoden aikana
$e_{x,jt}$	todennäköisyys, että x-ikäinen ja viimeiset jt vuotta yhtäjaksoisesti jäsenenä ollut henkilö on eronnut jäsenyydestä seuraavan vuoden lopussa
$a_{x,jt}$	todennäköisyys, että x-ikäinen ja viimeiset jt vuotta yhtäjaksoisesti poissa ollut entinen jäsen on jäsen seuraavan vuoden lopussa
i_x	todennäköisyys, että x-ikäinen työkykyinen (aktiivi) jäsen tai entinen jäsen on työkyvytön seuraavan vuoden lopus
$r_{x,st}$	todennäköisyys, että x-ikäinen ja viimeiset st vuotta yhtäjaksoisesti työkyvyttömänä ollut henkilö on kuntoutunut seuraavan vuoden lopussa
d_y	todennäköisyys, että y-ikäinen leski on seuraavan vuoden lopussa uudelleen avioitunut
c_x	tilan säilyttävä siirtymä
	<u>yhdistetyt funktiot:</u>
$u(n,x)$	merkintä, joka kuvaa n:n x-ikäisen uuden jäsenen eläkelaitoksen piiriin tulemisen todennäköisyyttä; tämän algoritminen käsittely on kuvattu alaluvun 4.2 kohdassa "uudet vakuutetut,
$g_x(y)$	todennäköisyys, että x-ikäisenä kuolleelta vakuutetulta jää y-ikäinen edunsaajaleski
$b_x(n,z)$	todennäköisyys, että x-ikäisenä kuolleelta vakuutetulta jää n z-ikäistä edunsaajalasta

Henkilömallin kaavion 4.2.1 funktiomerkinnät

4.3 Simulointitekniikasta

Valintaongelma

Simulointiohjelmaa voidaan verrata tavalliseen vakuutusmatemaattiseen tilastointiohjelmaan, joka käy läpi eläkelaitoksen rekistereitä kooten näistä vuositilastoja, sillä erotuksella, että simulointiohjelman on ensin ratkaistava itse mitä vakuutettuje ja eläkkeensaajien kannassa on laskentavuoden aikana tapahtunut. Kunkin laskenta-alkion, tietueen, osalt on ratkaistava, mikä siirtymä kulloinkin todennäköisyyskertoimiltaan erilaisista vaihtoehdoista valitaa Vakuutetun kuoltua on ratkaistava, mikä edunsaajakombinaatio vakuutetun jälkeen jää, kun tiedetään eri kombinaatioiden todennäköisyydet. Vastaavasti on aina uutta laskenta-alkiota perustettaessa ratkaitava henkilön syntymävuosi, kun uusien ko. laskenta-alkioiden ikäjakauma on annettu.

Nämä kaikki ratkaistavat ongelmat ovat samaa tyyppiä

> On valittava jokin vaihtoehtoisista tapahtumista $E_1, .., E_k$, joilla on todennäköisyydet $p_1, .., p_k$ siten, että $\sum p_i = 1$.
>
> Valintaproseduurilta vaaditaan, että valittujen tapahtumien suhteelliset frekvenssit f_i/n lähenevät vastaavia todennäköisyyksiä p_i $(i=1,..,k)$, kun valintojen lukumäärä n kasvaa rajatta, eli proseduuri toistaa tilastollisen stabiliteetin, joka on kuvattu todennäköisyyksillä p_i.

Valintaongelma ratkaistaan ns. Monte Carlo metodilla, jota ovat kuvanneet mm. Beard et al /5. ss 91-97/, Ackoff /2. ss 350-354/ ja Martin /15. ss 31-33/.

Monte Carlo metodi

Monte Carlo metodi on numerinen, epäanalyyttinen tekniikka, jolla voidaan ratkaista approksimatiivisesti erilaisia funktionaalisia yhtälöitä. Erityisesti korkeamman asteen osittaisia differentiaaliyhtälöitä ei käytännössä juuri muuten voida ratkaistakaan /3. s 33/.

Monte Carlo metodin toinen tärkeä sovellutusalue, josta tässä varsinaisesti olemme kiinostuneita, on stokastisten mallien simulointitekniikka. Kohdesysteemin prosessien deterministisiä ja stokastisia piirteitä jäljitellään askel askeleelta. Kukin satunnaismuuttaja x kuvataan todennäköisyysjakaumalla, josta muuttujan arvot kulloinkin "poimitaan" Monte Carlo proseduurilla /15. s 33/.

Todennäköisyysjakauman kertymäfunktiolle $y=F(x)$ muodostetaan käänteisfunktio $x=F^{-1}(y)$ ja jos analyyttistä käänteisfunktiota ei ole, käytetään askeleittain lineaarista interpolaatiota /11. ss 118-121, 2. s 354/. Proseduuri edellyttää, että on käytettävissä sarja tasaisesti välille (0, 1) jakautuvia satunnaislukuja $\{U\}$. Aina uutta satunnaismuuttajan x arvoa tarvittaessa otetaan seuraava satunnaisluku U ja ratkaistaan

$$x = F^{-1}(U)$$

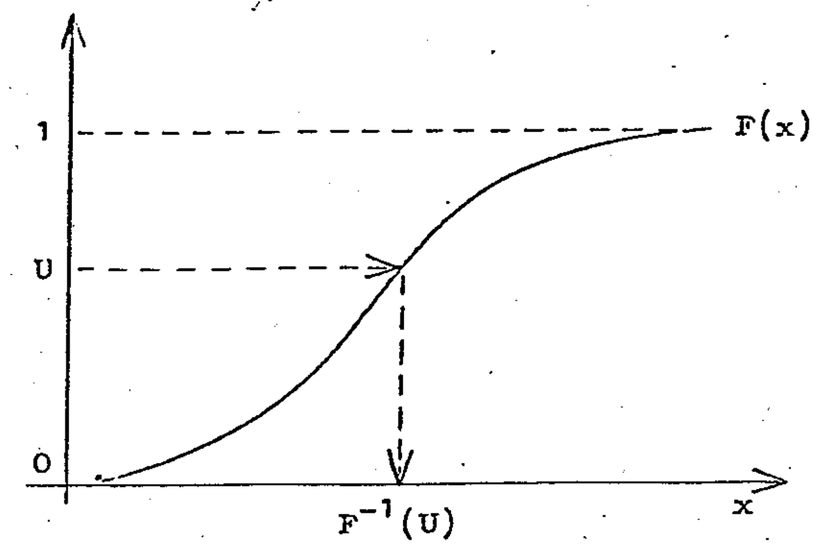
eli yhtälön $U = F(x)$ juuri x /5. s 91/ kuvion 4.3.1 mukaisesti.

Diskreettien tapahtumien $E_1,\ldots,E_k$ valintaongelmaan sovellettuna todennäköisyyksien p_i $(i=1,\ldots,k)$ kertymäfunktio on kuvion 4.3.2 mukainen porrasfunktio, josta uusi tapaus E ratkaistaan seuraavasti /15. s 33/:

jos $U \leq p_1$, niin valitaan E_1,
muuten valitaan sellainen E_j, että

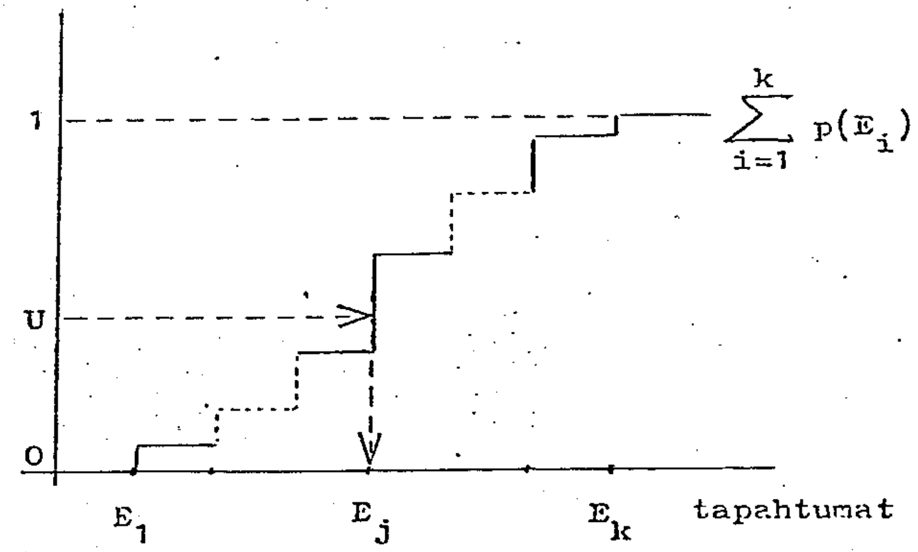
$$\sum_{i=1}^{j-1} p_i < U \leq \sum_{i=1}^{j} p_i \quad (j=2,\ldots,k).$$

Kuvio 4.3.1



Käänteisfunktiomenetelmä

Kuvio 4.3.2



Diskreetti tapahtumavalitsin

4.4 Satunnaisluvuista

Monte Carlo metodin tarvitsema satunnaislukusarja korvataan tietokonesimuloinnissa yleensä matemaattisella algoritmilla, jonka tuottamalla lukusarjalla on satunnaislukujen ominaisuudet muuten paitsi, että lukusarja on täysin määräytynyt, kun laskenta-algoritmille simulointiajon alussa annettava ns. siemenluku l. satunnaislukusiemen (engl. random number seed) tunnetaan, ja että lukusarjassa on vain äärellinen määrä eri lukuja ja lukusarja on periodinen ts. sama lukusarja toistuu aina tietyn jakson jälkeen /11. s 112/.

Lukusarjan täydellisen ennaltamääräytyvyyden vuoksi algoritmilla muodostettavia satunnaislukuja kutsutaan pseudosatunnaisluvuiksi ja niiden laskenta-algoritmia pseudosatunnaislukugeneraattoriksi erotuksena varsinaisista satunnaisluvuista, joiden arvoja ei edeltäkäsin voida tietää /11. s 109/.

Käyttämällä pseudosatunnaislukugeneraattorille samaa siemenlukua simulointiajo voidaan samoilla lähtötiedoilla toistaa täydellisesti. Tämä pseudosatunnaislukujen piirre on tarpeellinen mallia testattaessa, mutta siitä voi olla hyötyä varsinaisen simulointitutkimuksen aikana haluttaessa selvittää jonkin satunnaiskäyttäytymiseen vaikuttamattoman parametrin muutoksen (mallissamme esimerkiksi eläkelaskusäännön muutoksen) vaikutusta mahdollisimman vähillä simulointikerroilla.

Erilaisia pseudosatunnaislukugeneraattoreita on käsitelty runsaasti kirjallisuudessa /esim. 11. ss 112-114 ja 3. s 273-277/. Muodostetun pseudosatunnaislukugeneraattorin generoimien lukusarjojen tilastollisia ominaisuuksia on testattava ennen kuin generaattori voidaan ottaa käyttöön. Lukusarjan on noudatettava tarkoitettua jakaumaa ja sen lukujen on aina yhden jakson sisällä oltava autokorreloimattomia. Jakson tulisi olla mahdollisimman pitkä. Yleensä satunnaislukugeneraattoreina käytetään valmiiksi testattuja yleisiä kirjasto-ohjelmia.

Kuvattavana olevan simulointimallin satunnaislukugeneraattorina käytetään kielivalinnan luonnollisena seurauksena Extended ALGOLin perusfunktiota RANDOM. Tämä generoi käyttökelpoisia tasaisesti välille [0, 1) jakautuvia pseudosatunnaislukuja ja muuttaa nimiparametrina käsittelemäänsä argumenttia jokaisella funktion kutsulla /33. s 25/.

Ennen funktion RANDOM (arg) ensimmäistä kutsua argumentin arvoksi asetetaan satunnaislukusarjan siemenluku, jonka arvoksi Ruokosalmi /23. s 25/ suosittelee lukua 2^{19}, 2^{20} lähellä olevaa paritonta kokonaislukua. Simulointiohjelmassamme käytetään siemenlukuasetusta

arg:= 2 ** 20 + NO * 1000 - 1

missä NO on ohjelmalle parametrina annettava kokonaisluku. Lukua NO olemme käyttäneet samalla ko. simulointiajon yksilöivänä tunnuksena simulointitulosten dokumentoinnissa.