TALLINNA PEDAGOOGIKAÜLIKOOL
MATEMAATIKA – LOODUSTEADUSKOND
MATEMAATIKA – INFORMAATIKA OSAKOND

Birgit Bollverk

AEGRIDADE ANALÜÜS STATISTIKAPAKETIS SPSS

Proseminaritöö

Juhendaja: Katrin Niglas

Tallinn 2001

SISUKORD

SISSEJUHATUS

Nähtuste arvulist esitust võib leida juba varaorjanduslikust ühiskonnast. Kolmandal aastatuhandel e.m.a koguti demograafilisi andmeid Egiptuses ja Hiinas. Andmed riigi kodanike kohta olid vajalikud ülevaate saamiseks nõutavast tööjõust (püramiidide ehitamisel), sõjaväe moodustamisel ja maksude määramisel [5]. Põhjendatud otsuste tegemiseks, seaduspärasuste leidmiseks ja ülevaatliku informatsiooni esitamiseks toimus pidev statistika areng. Tänapäeval on statistikast saanud paljutahuline teadus, mis sisaldab endas mitmeid teisi alateadusi ja vastavaid õppeaineid.

Statistikaga puutume kokku ka pidevalt igapäevaelus, näiteks ajalehtedes näeme tihti mingite uuringute tulemusel saadud arvandmeid teatud ajahetkedel või kindlatel ajaperioodidel, mida on illustreeritud graafikute abil. Selles proseminaritöös vaatleme selliseid arvandmete ridu, mis on seotud kindla ajamomendiga - kuupäev, aasta algus jne. või mingi ajavahemikuga - kuu, kvartal, aasta jne.

Käesoleva proseminaritöö eesmärgiks on selgitada aegrea mõistet ning kirjeldada statistilisi meetodeid aegridade töötlemiseks ja analüüsimiseks ning järelduste ja üldistuste tegemiseks.

Proseminaritöö esimese peatükis uurime millised andmeread on aegread ja kuidas neid liigitatakse. Teises punktis vaatleme aegridade elementaar analüüsi, mille põhjal me saame iseloomustustada nähtuste üldist muutumist ning lihtsamate aegridade korral teha järeldusi ka üldise arengutendentsi kohta. Kui aegread on väga muutlikud nii, et on raske otsustada kas areng toimub tõusu või languse suunas, siis kasutatakse aegrea ühtlustamiseks aegridade tasandamist. Kolmandas peatükis uurimegi mitmeid erinevaid aegridade tasandamismeetodeid. Neljandas peatükis selgitame, mis on interpoleerimine, ekstrapoleerimine ja prognoosimine. Järgnevalt vaatleme millistest komponentidest aegread koosnevad ja kuidas teha kompleksanalüüsi. Lõpuks esitame aegridade analüüsi statistikapaketis SPSS 10.1.

Töö põhiosas käsitletud aegridade kirjeldamiseks on kasutatud põhiliselt raamutuid [1], [2], [3] ja [4] ning näidetes kasutavad aegread pärinevad Eesti Statistikaameti poolt väljaantud väljaannetest [6], [7], [8], [9].

1. AEGRIDADE MÕISTE JA LIIGID

Ühiskondlikud ja majanduslikud nähtused muutuvad ajaliselt, s.t. nad on dünaamilised. Nii näiteks materiaalseid väärtusi toodetakse ja kulutatakse pidevalt. Muutuvate nähtuste arengu ja muutumise jälgimine ning kirjeldamine eeldab nende kohta hangitud andmete koondamist aegridadesse.

Aegreaks nimetatakse nähtuste ajalist muutumist iseloomustavate arvandmete rida. *Aegrea elemendid* on nähtust iseloomustava tunnuse arvväärtused ning neile vastavad teatud ajamomendid või –perioodid.

Aegread liigitatakse *moment- ja perioodridadeks.* Nende range eristamine on oluline eeskätt keskmiste arvutamisel.

Momentrida on aegrida, mille iga element on seotud teatud ajamomendiga - kuupäevaga, mingi aasta alguse või lõpuga jne. Näiteks Tallinnas elavate inimeste arv iga aasta alguses. Momentrea olulisemaks iseärasuseks on asjaolu, et nähtust iseloomustava tunnuste arvväärtuste summal ei ole reaalset sisu .

Perioodrida on aegrida, mille iga element on seotud mingi ajavahemikuga, perioodiga (kuu, kvartal, aasta). Neid ridu nimetatakse mõnel juhul ka *intervallridadeks.* Perioodridu võib liigitada pidevateks ja sõredateks. *Pidevate perioodridade* perioodid järgnevad vahetult üksteisele. Pideva perioodrea näiteks on andmed ettevõtte kaubatoodangu kohta kvartalite lõikes ühel aastal. Pideva perioodrea elementide arvväärtuste summal on majanduslik sisu – nad väljendavad sama tunnuse väärtust pikema perioodi kohta. Nii väljendab nelja kvartali kaubatoodang ettevõtte aastast kaubatoodangut.

Sõredates perioodridades on tunnuse arvväärtused perioodide kohta, mis ei järgne vahetult üksteisele. Sõredate perioodridade perioodid võivad olla eraldatud üksteisest korrapäraste, s.t. võrdsete intervallidega või hoopis suvaliselt võetud intervallidega. Sõreda perioodrea näiteks võiks olla andmed Eesti riigieelarve tulude kohta, kus perioodidevahelised intervallid on ebakorrapärased. Sõreda perioodirea elementide arvväärtuse summa ei ole mõtestatud [1].

2. AEGRIDADE ELEMENTAARANALÜÜS

Ühiskondlike ja majanduslike nähtuste ajalise muutumise uurimisel võib eristada kahte põhilist astet:

1) aegridade elementaaranalüüs;
2) üldise arengutendentsi ehk trendi uurimine.

Elementaaranalüüs seisneb aegridade lihtsamate karakteristikute arvutamises. Sellised karakteristikud on järgmised: *absoluutne juurdekasv, juurdekasvutempo, kasvutempo, keskmine kasvutempo ja keskmised tasemed.* Nende lihtsamate aegrea karakteristikute abil on võimalik anda nähtuse muutumise üldine iseloomustus. Lihtsamate aegridade puhul on arvutatud elementaarkarakteristikute abil võimalik teha järeldusi ka üldise arengutendentsi kohta. Pikemate ja keerulisemate aegridade puhul tuleb aga üldise arengutendentsi hindamiseks kasutada spetsiaalseid meetodeid [1].

2.1. Aegridade kesmised tasemed

1) *Aritmeetiline kesmine.*

Aritmeetilise keskmisena leitakse perioodridade keskmisi tasemeid. Seejuures kasutatakse lihtsa aritmeetilise keskmise valemit

$$\bar{x} = \frac{\sum x}{n}, \quad (1)$$

kui üksteisele järgnevad sama suurusjärku perioodid on ühepikkused. Valemis (1) on x rea liikme väärtus ning n on rea liikmete arv. Nii leitakse näiteks keskmine elektrienergia toodang kümne üksteisele järgneva aasta jooksul: kümne aasta toodang liidetakse kokku ja jagatakse 10-ga.

2) *Kaalutud aritmeetiline keskmise*

$$\bar{x} = \frac{\sum xf}{\sum f} \quad (2)$$

järele tekib vajadus siis, kui perioodid ei ole ühepikkused. Aegrea eri liikmete kaalud peavad sel juhul olema võrdelised vastavate perioodide pikkustega. Valemis (2) on x aegrea elemendi väärtus ja f on tema vastav kaal.

Näide. On teada, et 2000. aasta kolmandas kvartalis sündinud laste arv oli 3304 last, 2000. aasta esimesel poolaastal sündinud laste arv oli 6714 last ja 1999. aastal sündinud sündinud laste arv oli 12544 last (Allikas: [8] lk.24). Leida keskmine laste sündimus kvartalis sel ajavahemikul.

Lahendus

Kuna ülejäänud perioodide pikkused on lühima kordsed, valime ühtlustatud perioodide pikkuseks kvartali. 2000. aasta esimese poolaasta kahes kvartalis oli laste keskmine sündimus 3357 last kvartalis ja 1999. aasta neljas kvartalis oli laste keskmine sündimus 3136 last kvartalis. Keskmine laste sündimus kvartalis on

$$\bar{x} = \frac{3304 \cdot 1 + 3357 \cdot 2 + 3136 \cdot 4}{1+2+4} = \frac{22562}{7} = 3223 \text{ last.}$$

3) *Kronoloogiline keskmine*.

Momentridade puhul kasutatakse aegrea keskmise taseme arvutamiseks kronoloogilist keskmist. Sellisel juhul saame momentrea keskmise arvutada järgmise valemi abil

$$\bar{x}_{kron.} = \frac{\frac{x_1}{2} + x_2 + \ldots + x_{n-1} + \frac{x_n}{2}}{n-1} = \left(\frac{x_1 + x_n}{2} + \sum_{i=1}^{n-1} x_i \right) \frac{1}{n-1},$$

kus x_i on mingi momentsuurus uuritavas aegreas. Sellisel kujul avaldatud keskmist nimetatakse kronoloogiliseks keskmiseks [1].

Näide. Arvuta keskmine rahvaarv Eestis 80-ndatel ja 90-ndatel aastatel, kui meil on teada rahvaarv iga aasta alguses 1980-2000.

Aasta	Rahvaarv aasta alguses
1980	1477219
1981	1487666
1982	1498414
1983	1508745
1984	1518617
1985	1528781
1986	1540190
1987	1552221
1988	1561900
1989	1568655

Aasta	Rahvaarv aasta alguses
1990	1571050
1991	1566334
1992	1544374
1993	1516728
1994	1499255
1995	1485645
1996	1476301
1997	1462130
1998	1453844
1999	1445580

Allikas: [6], lk 40 ja [8], lk 24.

Lahendus

Eespool toodud valemi põhjal saame keskmise rahvaarvu Eestis 80-ndatel aastatel arvutada järgmiselt

$$\bar{x} = \frac{\frac{1472190}{2} + 1482247 + 1493085 + ... + 1546304 + 1558137 + \frac{1565662}{2}}{9} =$$

$$= \frac{13673751}{9} = 1519306.$$

Keskmise rahvaarv Eestis 90-ndatel aastatel on

$$\bar{x} = \frac{\frac{1571648}{2} + 1570451 + 11562216 + ... + 1462130 + 1453844 + \frac{11445580}{2}}{9} =$$

$$= \frac{13512926}{9} = 1501436.$$

Järelikult keskmine rahvaarv Eestis on vähenenud peaaegu 20 000 inimese võrra.

4) *Keskmine kasvutempo geomeetrilise keskmisena.*

Kui on teada aegrida, mille liikmete absoluutsed väärtused on $y_1, y_2, ..., y_n$, kus y tähistab näiteks piimatoodangut mingil aastal, siis võib neil andmetel leida kokku (n - 1) ahelindeksit ehk aastast kasvutempot, mis näitavad, kui intensiivselt on toimunud y kasv aastast aastasse. Ahelindeksid on kordsed ehk multiplikatiivsed suurused. Nende liitmisel ja seega ka nende aritmeetilisel keskmisel pole praktiliselt mõtet. Et individuaalindeksite reast saab ühes arvus väljenduva üldistuse ainult korrutamise teel, siis tähendab see, et ka nende keskmise arvutamisel tuleb kasutada niisugust keskmise valemit, mis eeldab lähtesuuruste korrutamist.

Teisiti väljendatult tähendab see, et mistahes nähtuse keskmine kasvutempo (keskmine ahelindeks) tuleb arvutada individuaalsete kasvutempode geomeetrilise keskmisena. Valemit

$$\bar{x}_{geom.} = \sqrt{x_1 \cdot x_2 \cdot ... \cdot x_n} \qquad (3)$$

saab geomeetrilse keskmise arvutamiseks kasutada siis, kui indeksid on olemas. On aga lähteandmeteks ainult absoluutarvud kogu uuritava ajaperioodi alguses ja lõpus, siis seda kasutada ei saa. Viimasel juhul tuleb kasutada valemit

$$\bar{x}_{geom.} = \sqrt[n-1]{\frac{y_n}{y_1}}\,. \qquad (4)$$

Viimane valem on geomeetrilise keskmise arvutamiseks eriti sobiv, kui andmed on lünklikud. Valemi (4) kasutamine on otstarbekas ka siis, kui on teada aegrea kõik liikmed, sest jääb ära ahelindeksite arvutamine ja korrutamine. Üldse on valemit (3) geomeetrilse keskmise arvutamiseks mõtet kasutada ainult siis, kui ainsateks andmeteks on ahelindeksid, rea liikmete absoluutseid väärtusi aga pole teada [1].

2.2. Absoluutne juurdekasv

Nähtusi iseloomustavate tunnuste *absoluutseks juurdekasvuks* nimetatakse kahe elemendi väärtuste vahet. Absoluutne juurdekasv võib olla kas positiivne või negatiivne, mis iseloomustab seega tunnuse väärtuse vähenemist või suuremist [5]. Pikemate aegridade puhul võib absoluutse juurdekasvu leida eelneva või mingi teise aluseks võetud elemendi väärtusega võrreldes [1]. Aegrea elemendi arvväärtust nimetatakse majandusalases kirjanduses ka tasemeks ja absoluutset juurdekasvu iseloomustatakse kui kahe absoluutse taseme vahet. Absoluutne juurdekasv aegrea eelmise elemendi väärtusega võrreldes ehk nn. aheljuurdekasv *(d^a)* leitakse valemiga

$$d^a = y_t - y_{t-1}, \tag{5}$$

kus y_t on aegrea elemendi väärtus vaadeldaval ajamomendil ja y_{t-1} on aegrea elemendi väärtus eelmisel ajamomendil [5].

Absoluutne juurdekasv mingi varasema baasiks võetava väärtusega võrreldes *(d^b)* ehk nn. alusjuurdekasv leitakse valemiga

$$d^b = y_t - y_1, \tag{6}$$

kus y_1 on aegrea elemendi väärtus baasiks võetaval ajamomendil (-perioodil) [1].

Absoluutne nn. alusjuurdekasv mingi perioodi kohta on üksikute aheljuurdekasvude (5) summa, mis iseloomustab absoluutset juurdekasvu kogu vaadeldaval perioodil [5]. Absoluutsete juurdekasvude reast tavaliselt ei piisa ning tihtipeale on vaja uurida ka nähtuse suhtelist muutumist [1].

2.3. Kasvutempo

Kasvutempo on nähtust iseloomustava tunnuse vaadeldava ajamomendi (või -perioodi) arvväärtuse ja mingi eelmise ajamomendi (või -perioodi) arvväärtuse suhe. Kasvutempo on aegridade spetsiifiline näitarv, mida majandusalases kirjanduses nimetatakse ka indeksiks. Kasvutempo on lihtindeks, kirjanduses nimetatakse seda ka kasvukoefitsiendiks. Mõnede autorite määratluses on koefitsient vahetu suhtena arvutatud suhtarv ja indeks protsentides väljendatud suhtarv [5].

Kasvutempo võib leida kas eelmise ajamomendi (-perioodi) või mingi muu aluseks võetud ajamomendi (-perioodi) suhtes. Vastavalt sellele eristatakse ahel- ja aluskasvutempot.

Ahelindeks leitakse valemiga

$$i^a = \frac{y_t}{y_{t-1}}, \tag{7}$$

kus y_t on aegrea elemendi arvväärtus vaadeldaval ajamomendil (-perioodil) ja y_{t-1} on aegrea elemendi arvväärtus eelmisel ajamomendil (-perioodil).

Alusindeks leitakse valemiga

$$i^b = \frac{y_t}{y_1}, \tag{8}$$

kus y_1 on aegrea elemendi väärtus baasiks võetud ajamomendil või –perioodil.

Kui $i > 1$, siis on meil tegemist sõna otseses mõttes kasvu tempoga. Kui aga kasvutempo $i < 1$, siis nähtust iseloomustava tunnuse väärtus on vähenenud ning tegemist on kahanemise tempoga.

Kasvutempo (ehk indeks) võib olla väljendatud nii vahetu suhtena kui ka protsentides või promillides. Viimasel juhul korrutatakse vahetu suhtena leitud kasvutempo vastavalt 100 või 1000-ga. Ahelindeksite korrutis annab viimase alusindeksi. Lünklike aegridade täiendamiseks on oluline teada ka seda, et mistahes ajaperioodi (-momendi) ahelindeks võrdub sama perioodi (-momendi) alusindeksi ja eelmise perioodi alusindeksi jagatisega. Mistahes ajaperioodi (või -momendi) alusindeks võrdub kõigi ahelindeksite korrutisega alates baasiks võetud aastale järgnevast aastast kuni antud aastani [1].

2.4. Juurdekasvutempo

Nähtuste ajalise muutumise intensiivsust iseloomustab ka *juurdekasvutempo*, mis on absoluutse juurdekasvu ning selle arvutamisel aluseks võetud aegrea elemendi väärtuse suhe. Juurdekasvutempo võib leida kas aegrea eelmise elemendi või mingi teise aluseks võetud elemendi suhtes. Vastavalt sellele eristatakse ahel- ja alusjuurdekasvutempot [5].

Aheljuurdekasvutempo *(j^a)* leitakse valemiga:

$$j^a = \frac{d^a}{y_{t-1}} = \frac{y_t - y_{t-1}}{y_{t-1}} = i^a - 1, \qquad (9)$$

kus d^a on absoluutne juurdekasv aegrea eelmise elemendi väärtusega võrreldes ja y_{t-1} on aegrea elemendi arvväärtus eelmise ajamomendil või –perioodil.

Alusjuurdekasvutempo *(j^b)* leitakse valemiga:

$$j^b = \frac{d^b}{y_1} = \frac{y_t - y_1}{y_1} = i^b - 1, \qquad (10)$$

kus d^b on absoluutne juurdekasv aegrea baaselemendi arvväärtusega võrreldes ja y_1 on aegrea elemendi väärtus baasiks valitud ajamomendil (-perioodil).

Juurdekasvutempo võib nagu kasvutempogi väljenduda nii vahetu suhtena kui ka protsentides või promillides. Kui kasvutempo on juba leitud, siis võib juurdekasvutempo leida, lahutades kasvutempost vastavalt 1. Analoogselt võib juurdekasvutempo alusel leida kasvutempo, liites juurdekasvutempole vastavalt 1. Juurdekasvutempo võib olla ka negatiivne, kui absoluutne juurdekasv on negatiivne [1].

Näide . Olgu meil antud kahe viimase aasta keskmised brutopalgad kvartalite kaupa

	Keskmine brutopalk
1999	
I kvartal	4057
II kvartal	4554
III kvartal	4261
IV kvartal	4799
2000	
I kvartal	4501
II kvartal	5031
III kvartal	4694
IV kvartal	5279

Allikas: [7], lk.10.

Leia eelnevalt antud valemite põhjal absoluutne ahel- ja alusjuurdekasv, ahel- ja alusjuurdekasvutempo ning ahel- ja aluskasvutempo. Seejärel iseloomusta keskmise brutopalga muutust sellel ajavahemikul.

Võrdle oma vastuseid järgneva tabeliga.

	Keskmine brutopalk	Absoluutne ahel-juurdekasv	Absoluutne alus-juurde-kasv	Ahel-juurde-kasvu-tempo	Alus-juurde-kasvu-tempo	Ahel-kasvu-tempo	Alus-kasvu-tempo
1999							
I kvartal	4057						
II kvartal	4554	497	497	0,123	0,123	1,123	1,123
III kvartal	4261	-293	204	-0,064	0,050	0,936	1,050
IV kvartal	4799	538	742	0,126	0,183	1,126	1,183
2000							
I kvartal	4501	-298	444	-0,062	0,109	0,938	1,109
II kvartal	5031	530	974	0,118	0,240	1,118	1,240
III kvartal	4694	-337	637	-0,067	0,157	0,933	1,157
IV kvartal	5279	585	1222	0,125	0,301	1,125	1,301

Absoluutne aheljuurdekasv näitab kahe kõrvutipaikneva elemendi (hilisema ja varajasema) väärtuse vahet, s.t. näitab kahe üksteisele järgneva kvartali keskmise brutopalga vahet. Absoluutne aheljuurdekasv võib olla kas positiivne või negatiivne, näidates sel juhul kas keskmine brutopalk kasvab või kaheneb. Tabelist 1 näeme, et keskmine brutopalk vaheldumisi nii kasvab kui kahaneb.

Absoluutne alusjuurdekasv näitab andmete rea elemendi ja baasiks võetud elemendi vahet. Harilikult on baaselemendiks rea esimene liige. Siit näeme, et keskmine brutopalk 1999. aasta I kvartali palga suhtes on pidevalt kasvanud.

Ahelkasvutempo leitakse vaadeldava ajamomendi arvväärtuse ja mingi eelmise ajamomendi arvväärtuse suhtena. Kui ahelkasvutempo on suurem ühest, siis on tegemist kasvuga, kui aga väiksem ühest, siis kahanemisega. Tabelist näeme, et 1999. aasta ja 2000. aasta II ja IV kvartalis kasvas keskmine brutopalk eelneva kvartali keskmise brutopalga suhtes.

Aluskasvutempo saadakse vaadeldava ajamomendi arvväärtuse ja baasiks võetud arvväärtuse suhtena. Aluskasvutempo järgi näeme, et keskmine brutopalk kasvab 1999. aasta I kvartali palga suhtes.

Aheljuurdekasvutempo iseloomustab, kuivõrd on tunnuse väärtus vaadeldava perioodi jooksul muutunud võrreldes perioodi algusega. Andmete põhjal näeme, et näiteks 1999. aasta III kvartali keskmine brutopalga juurdekasv vähenes 6,4%, kuid 1999. aasta IV kvartalis oli keskmise brutopalga juurdekasv 12,6%.

Alusjuurdekasvutempo leitakse absoluutse alusjuurdekasvu ja baaselemendi suhtena. Siit näeme, et näiteks 2000. aasta IV kvartali palk on kasvanud 1999. aastaga I kvartali palgaga võrreldes 30,1%.

3. AEGRIDADE TASANDAMINE

Otseseid vaatlusandmeid kujutavad kogemuslikud ehk empiirilised aegread, mis võivad olla küllalt hüplevad. Sagedaste tõusude-languste tõttu võib olla puhuti isegi raske otsustada, kas areng toimub üldse tõusu või languse suunas, kõnelemata eri aegridades ilmnevate pikemaajaliste tõusu- või langustendentside täpsest kvantitatiivsest võrdlemisest [1]. Pikemaajaliste tendentside määramiseks kasutatakse *ridade nn. tasandamist*, millega elimineeritakse arengu üldsuunast nii ühele kui teisele poole toimunud kõikumised [2].

Ridade tasandamise moodustest on lihtsaim nn. *visuaalse trendi meetod*, s.o. üldist arengusuunda iseloomustava joone määramine silma järgi. See viis pole teaduslik ning see sisaldab liigselt subjektiivseid elemente. Eri uurijad võivad samu lähtekõveraid näha erinevalt ning saada seetõttu ka erinevaid tasandusjooned. Visuaalne trend on kasutatav tavaliselt ainult tasanduskõvera kuju leidmiseks. Teaduslike tasandamisviisidena kasutatakse *libiseva keskmise meetod* ja analüütiliste tasandusjoonte leidmiseks vähimruutude meetodit [1].

3.1. Libiseva keskmise meetod

Libiseva keskmise meetodit kasutatakse aegrea perioodilise komponendi kõrvaldamiseks [2]. Näiteks turgude ja kaupluste kaubakäibele avaldavad mõju aastaaegade vaheldumine, riiklikud pühad, palgapäevad, ent ka nädalapäevade vaheldumine. Soovides vabaneda ainult sellest mõjust, mida avaldab suhteliselt elavam müük nädala lõpupäevadel, s.o. nädalasisese perioodilisuse mõjust, kasutatakse seitsme päeva libisevat keskmist. Soovides vabaneda kuu- või aastasisesest perioodilisuse mõjust, leitakse kuu või kvartali libisev keskmine [1].

Libisev keskmine on fikseeritud arvu naabervaatluste (näiteks kolme või viie) aritmeetiline keskmine, mis liigub läbi kogu rea aja kasvamise suunas. Olgu aegreas n elementi $y_1, y_2, ..., y_n$. Võtame neist m esimest elementi ning leiame nende aritmeetilise keskmise $\bar{y}_1^{(m)} = \frac{1}{m}\sum_{i=1}^{m} y_i$. Edasi leitakse elementide $y_2, y_3, ..., y_{m+1}$

aritmeetiline keskmine $\bar{y}_2^{(m)} = \frac{1}{m}\sum_{i=2}^{m+1} y_i$ ning analoogilisel viisil jätkates liigutakse rea lõpuni [2]. Libiseva keskmise $\bar{y}^{(m)}$ arvväärtused leitakse valemiga

$$\bar{y}_k^{(m)} = \frac{1}{m}\sum_{i=k-\frac{m-1}{2}}^{k+\frac{m-1}{2}} y_i \,, \tag{11}$$

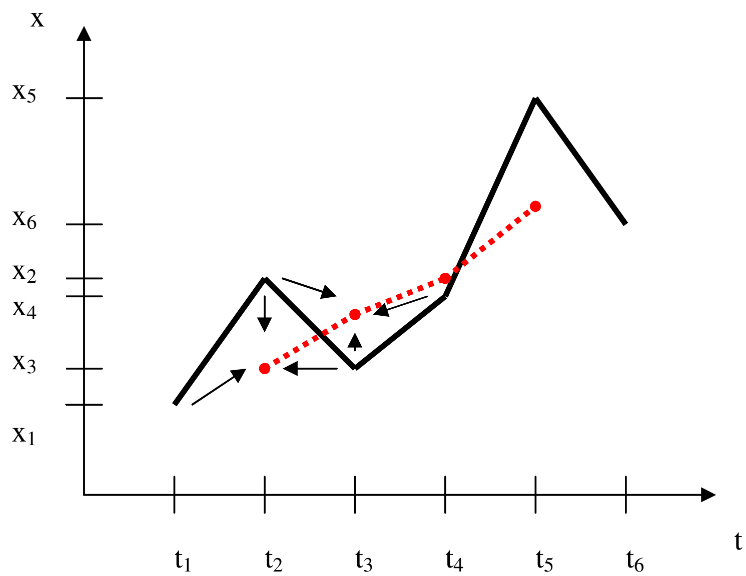
kus *m* on libisemissammu pikkus. Murru lugejas olevat suurust nimetatakse libisevaks summaks. Tavaliselt võetakse libisemissammu pikkuseks mingi paaritu arv osaperioode (päevi, kuid, aastaid) [1].

Nii saadud uus rida $\bar{y}_1^{(m)}, \bar{y}_2^{(m)}, \ldots, \bar{y}_{n-(m-1)}^{(m)}$ on *m - 1* elemendi võrra lühem kui lähterida [2]. Kui *i = 1, 2, …, n*, siis saab tuletada lähtereast kokku *n - (m - 1)* elemendist koosneva libiseva keskmise arvväärtuse rea. Libiseva keskmise saab leida kõigi osaperioodide kohta, mille järjenumber *k* mahub vahemikku

$$\frac{m+1}{2} \le k \le n - \frac{m-1}{2} \qquad [1].$$

Illustreerime toimuvat joonisel 1 kolmeelemendilise libisemissammu korral. Pidev joon näitab aegrea tegelikke väärtusi ja nende muutumist. Libiseva keskmise leidmiseks võetakse kolm esimest rea liiget ning arvutatakse nende keskmine. Ajaliselt paigutub see keskmine teise rea liikme kohale. Järgmise keskmise arvutamiseks jäetakse kolmikust välja ajahetkele 1 vastav rea liige ja lisatakse sellele ajahetkele 4 vastav rea liige. Sellest kolmikust arvutatud keskmine paigutub ajaliselt rea kolmanda liikme kohale. Analoogiliselt jätkatakse, kuni jõutakse rea kõige hilisema liikmeni.

Rea liikmed, millest keskmine arvutatakse, ja libisevaid keskmisi iseloomustavad punktid on ühendatud nooltega. Saadud keskmised on ühendatud katkendjoonega. Joonte võrdlemisel on näha, et tunnuse väärtused on pärast tasandamist ühendatud märksa siledama joonega, kui nad olid enne seda. Kõrvaldatud on suur osa üksikutele rea liikmetele iseloomulikke kõrvalekaldeid üldtendentsist [3].



Joonis 1. Aegrea tasandamine libiseva keskmisega

Libiseva keskmise puhul tuleb silmas pidada, et tunnuse väärtuse kasvamise korral kalduvad libisevad keskmised tunnuse väärtusi alla hindama ja kahanemise korral üle hindama.

Tunnuse tegelikke väärtusi läbiva joone ja libisevaid keskmisi läbiva joone lõikepunkti loetakse sageli ajahetkeks, millest alates võib uut arengusuunda väljakujunenuks pidada [3].

Näide. Olgu antud Eestimaa rahvastiku loomulik iive kuude kaupa 1999-2000. Tasanda loomulik iive libiseva keskmise meetodiga, kui libisemissamm on 3. Tulemused on esitatud järgmises tabelis.

Aeg	Sündimus	Suremus	Loomulik iive	Kolme kuu libisev keskmine
1999				
Jaanuar	952	1670	-718	
Veebruar	965	1552	-587	-630
Märts	1106	1690	-584	-543
Aprill	1100	1559	-459	-460
Mai	1200	1538	-338	-411
Juuni	1062	1497	-435	-393
Juuli	1069	1474	-405	-392
August	1099	1435	-336	-356
September	1102	1429	-327	-397
Oktoober	991	1520	-529	-480
November	921	1504	-583	-574
Detsember	978	1587	-609	-666
2000				
Jaanuar	1081	1886	-805	-755
Veebruar	1031	1881	-850	-695
Märts	1164	1594	-430	-523
Aprill	1055	1344	-289	-333
Mai	1246	1527	-281	-293
Juuni	1136	1446	-310	-256
Juuli	1144	1322	-178	-233
August	1190	1401	-211	-257
September	970	1352	-382	-342
Oktoober	1069	1502	-433	-443
November	980	1495	-515	-535
Detsember	915	1571	-656	

Allikas: [6], lk. 24.

Antud aegrida ja libiseva keskisega tasandatud aegrida graafiliselt on lisas 3.

3.2. Analüütiline tasandamine

Aegrea analüütilise tasandamise all mõistetakse selle taandamist mingile geomeetrilisele joonele (sirge, parabool vm.) ja vastava funktsiooni kasutamist aegrea esitamisel. Tasandusjoone kuju valik on oluline, et õigesti peegeldada aegrea suundumusi. Tavaliselt leitakse tasandusjoone kuju empiirilise rea vaatlemisel intuitiivselt nn. visuaalsel meetodil. Lisaks visuaalsele vaatlemisele kasutatakse abimeetodeid, mis võimaldavad leida taandava joone funktsiooni [5]. Kui aegrida tasandatakse vähimruutude meetodil, läbitakse järgmised 3 tööetappi:

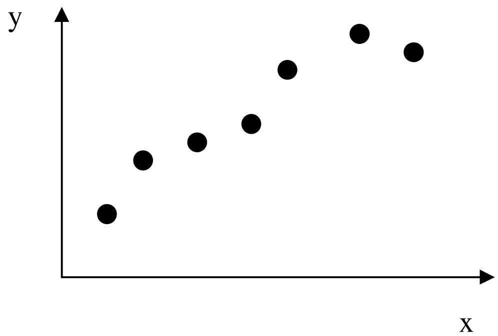
1) Valitakse sobiv tasadusjoon;
2) Nn. normaalvõrrandite süsteemi või sellest tuletatud lihtsustatud võrrandisüsteemi abil leitakse empiirilist kõverat tasandava teoreetilise joone parameetrilised hinnangud;
3) Leitakse teoreetilise joone punktide arvväärtused ning konstrueeritakse tasandusjoon.

Kõige sagedamini kasutatakse tasandusjooneks sirget, suhteliselt tihti ka teist järku parabooli. Perioodiliste võngete tasandamiseks sobib vahel sinusoid. Põhimõtteliselt võib kasutada kõiki, sealhulgas ka suletud geomeetrilisi jooni. Mõnegi protsessi kulgu võib hästi tasandada ringjoone või ellipsoidi lõik [1].

3.3. Vähimruutude meetodi olemus ja omadused

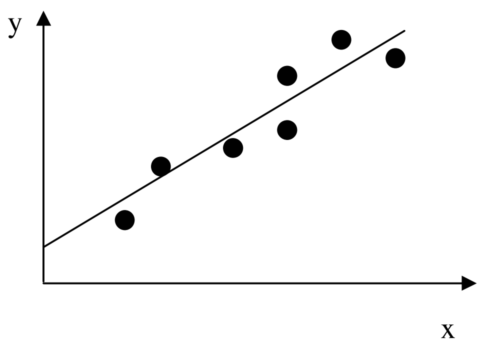
Olgu meil teada kahe tunnuse väärtused paaridega $x_1, y_1; x_2, y_2; ...; x_n, y_n$, kus x_i *(i = 1 ,…, n)* on aeg ja y_i *(i = 1 ,…, n)* on ajas muutuv suurus. Meie eesmärgiks on leida selline funktsioon, mis kõige paremini iseloomustab meile teada olevaid väärtusi.

Kui punktid $x_1, y_1; x_2, y_2; ...; x_n, y_n$ asetsevad graafikul vastavalt allpool toodud joonisele,



Joonis 1. Antud punktide asend.

siis iseloomustab antud väärtuste parve kõige paremini sirge.



Joonis 3. Antud punktid ja neid kõige paremini iseloomustav joon.

Järgnevalt otsime väärtuste parvega kõige paremini sobivat sirget, mille võrrand $y = a + bx$ aitaks iseloomustada nähtuste x ja y vahelist seost.

Kuna x_i ja y_i $(i = 1,2,\ldots,n)$ on meil teada, siis tuleb meil leida parameetrite a ja b väärtused. Parameetrite a ja b leidmiseks kasutatakse kõige sagedamini vähimruutude meetodit.

Vähimruutude meetodil leitakse sirge, mille puhul vaatlusel saadud punktide ja seost kirjeldava sirge vaheliste hälvete ruutude summa on minimaalne

$$S = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 , \tag{12}$$

kus $\hat{y}_i$ on trendifunktsiooni abil leitud teoreetilised väärtused ning y_i on aegrea liikmete tegelikud väärtused.

Arvutamise lihtsustamiseks on eelmises valemis (12) tehtud asendus, mis võtab arvesse otsitava trendifunktsiooni kuju

$$\hat{y}_i = a + bx_i .$$

Asenduse tulemusel saame

$$S = \sum_{i=1}^{n} (y_i - bx_i - a)^2 .$$

Meie ülesandeks on leida funktsiooni *S* miinimumile vastavad *a* ja *b* väärtused. Seega otsime kahe muutuja funktsiooni miinimumi. Selleks peavad funktsiooni *S* osatuletised *a* ja *b* järgi võrduma nulliga, s.t.

$$\frac{\partial S}{\partial a} = 0 \qquad \text{ja} \qquad \frac{\partial S}{\partial b} = 0 .$$

Pärast tuletise võtmist saame kahe tundmatuga kahest võrrandist koosneva lineaarse võrrandisüsteemi

$$\sum_{i=1}^{n} y_i = na + b\sum_{i=1}^{n} x_i$$
$$\sum_{i=1}^{n} x_i y_i = a\sum_{i=1}^{n} x_i + b\sum_{i=1}^{n} x_i^2$$

mille lahendamine annab *a* ja *b* väärtused. Seda lineaarset võrrandisüsteemi nimetatakse ka normaalvõrrandisüsteemiks.

Sageli uuritakse keerulisi nähtusi, mida iseloomustatakse rohkem kui ühe või kahe tunnusega. Üldjuhul on nad kõik üksteisega seotud.

Selle seose kirjeldamiseks tuleb kasutada mitme muutuja funktsioone, kõige sagedamini kasutatakse lineaarset mitme muutuja funktsiooni

$$y_t = a + b_1 x_{1i} + b_2 x_{2i} + ... + \varepsilon_i . \qquad (13)$$

Aegridade tasandamiseks kasutatakse tihti ka erinevat järku polünoome. Sel juhul püütakse arvutamistööde lihtsustamiseks piirduda võimalikult madalat järku polünoomidega. Teiseks tuleb funktsiooni valikul silmas pidada asjaolu, et kuigi kõrgemat järku polünoomide abil on võimalik rida kirjeldada täpsemalt, kuid see ei tähenda, et leitud trendifunktsiooni prognoosivõime suureneks [3].

Lineaarse mitme muutuja funktsiooni ja erinevat järku polünoomi korral leitakse tundamatute parameetrite väärtused vähimruutude meetodiga analoogiliselt, kui sirge juhtumil. Tundmatute parameetrite leidmiseks peab antud väärtuste y_t ja trendifunktsiooni vastavate väärtuste vahede ruutude summa olema minimaalne. Selleks võetakse osatuletised tundmatute parameetrite järgi ja pannakse nulliga võrduma. Saadud normaalvõrrandisüsteemist leitakse otsitavad tundmatud parameetrid.

3.4. Muude trendifunktsioonide kasutamine

Trendifunktsioonide parameetrite hindamine ja aegrea tasandamine nende abil sõltub palju sellest, kas trendifunktsioonid on lineaarsed parameetrite suhtes, või saab funktsiooni kuidagi selliselt teisendada, et viia ta parameetrite suhtes lineaarsele kujule VRMi kasutamiseks.

1) *Parameetrite suhtes lineaarsed funktsioonid.*

Siia klassi kuuluvad eeskätt kõik polünoomid. Kõrgemat järku polünoomidena avalduvate trendifunktsioonide parameetrite hinnangute arvutamine toimub põhimõtteliselt analoogiliselt juba käsitletule (punkt 3.3.).

2) *Parameetrite suhtes mittelineaarsed funktsioonid, mis lihtsate teisenduste (tavaliselt logaritmimine) abil on lineariseeritavad.*

Sellesse trendifunktsioonide klassi kuuluvad eeskätt mitmesugused eksponentfunktsioonid, nagu näiteks

$$y_t = a_0 a_1^t \,, \tag{14}$$

$$y_t = a_0 a_1^t a_2^{t^2} \,, \tag{15}$$

ning astmefunktsioonid, nagu näiteks

$$y_t = a_0 t^{a_1} \text{, jt.} \tag{16}$$

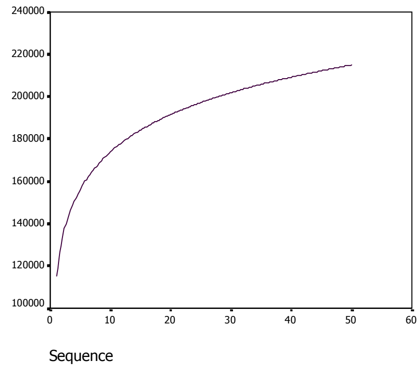
Viimati toodud kolme trendifunktsiooni (14) – (16) logaritmimisel saame parameetrite suhtes lineaarsed funktasioonid

$$\ln y_t = \ln a_0 + t \ln a_1 \,, \tag{17}$$

$$\ln y_t = \ln a_0 + t \ln a_1 + t^2 \ln a_2 \,, \tag{18}$$

$$\ln y_t = \ln a_0 + a_1 \ln t \,. \tag{19}$$

Graafiliselt on need funktsioonid järgmised



Joonis 4.

Trendifunktsioonide (17) – (19) parameetrite hinnangute arvutamiseks vajalike võrrandisüsteemide koostamine toimub analoogiliselt eespool käsitletule.

Analoogiliselt toimub ka mitmesuguste poollogaritmiliste trendifunktsioonide, nagu näiteks

$$y_t = a_0 + a_1 \ln t \tag{20}$$

parameetrite hinnangute arvutamine.

3) *Mittelineariseeritavad kõverjoonelised trendifunktsioonid.*

On olemas terve hulk huvitava kujuga trendifunktsioone, mis on aga parameetrite suhtes mittelineaarsed ning mida ka lihtsate teisenduste abil ei saa lineaarseks teisendada. Taoliste funktsioonide seas on siiski ka selliseid, millede puhul teatud tingimustel on võimalik parameetrite hinnangute leidmiseks kasutada tavalist VRMi. Need on nn. asümptoodiga (teatud piiriga) funktsioonid, millede puhul on võimalik VRMi kasutada juhul, kui asümptoot on eelnevalt teada või ette antud.

Sellesse trendifunktsioonide klassi kuuluvad näiteks modifitseeritud eksponentfunktsioon

$$y_t = k + a_0 a_1^t,$$

logistiline kõver

$$y_t = \frac{k}{1 + a_0 e^{-a_1 t}},$$

Gomperzi funktsioon

$$y_t = k a_0^{a_1 t},$$

kus kõigis k-asümtoodi (funktsiooni piiri) arvväärtus.

Kui asümptoodi k arvväärtus on eelnevalt teada, viime selle aegrea tasemetega ühele poolele vastavas trendifunktsiooni valemis, teeme vastavad teisendused, ning saamegi parameetrite suhtes lineaarsed funktsioonid

$$\ln(y_t - k) = \ln a_0 + t \ln a_1 ,$$

$$\ln(\frac{k}{y_t} - 1) = \ln a_0 + a_1 t ,$$

$$\ln(\frac{y_t}{k}) = \ln(\ln a_0) + t \ln a_1 .$$

Seejärel koostame vastavad võrrandisüsteemid, millede lahendamisel saame vastavate trendifunktsioonide parameetrite VRM hinnangud. Kui asümptoodi k arvväärtus ei ole teada ning seda ei saa ka ligilähedaselt hinnata, tuleb taoliste trendifunktsioonide parameetrite hindamiseks kasutada muid meetodeid nagu näiteks kolme summa meetod, kolme punkti meetod, mitmesugused iteratsioonimeetodid VRMi kasutamisega jne [1].

3.5. Trendifunktsiooni statistilise usaldatavuse hindamine

Aegrea tasandamiseks ning trendi kindlakstegemiseks on kõige olulisem probleem sobiva funktsiooni valimine, eriti veel siis, kui trendifunktsiooni kasutatakse uuritava nähtuse prognoosimudelina.

Trendi valiku probleemi lahendamine on otstarbekas seostada ka formaalse statistilise analüüsiga, eeskätt trendifunktsiooni statistilise usaldatavuse hindamisega. Trendifunktsiooni valiku formaalse kriteeriumina saab kasutada selle ruutkeskmist viga ehk standardviga

$$s = \sqrt{\frac{\sum (y_t - \hat{y}_t)^2}{n - p - 1}} .$$

Sobivamaks (vähemalt formaalsest statistilisest kriteeriumist lähtudes) on trendifunktsioon, mille puhul standardviga on väiksem (eeldusel, et parameetrite arv p on sama).

Trendi- ehk tasandusfunktsiooni vastavuse kontrollimiseks empiirilise aegreaga kasutatakse ka nn. aproksimeerimisviga

$$V_a = \frac{1}{n} \sum \frac{|y_t - \hat{y}_t|}{y_t} \times 100 ,$$

mis tavaliselt esitatakse protsentides ning hea vastavuse korral V_a ei tohiks olla suurem kui 5%, rahuldava vastavuse korral mitte suurem kui 10%.

Polünoomidena avalduvate (või polünoomseteks teisenduvate) trendifunktsioonide kasutamise korral oleks mõistlik trendifunktsiooni statistilise usaldatavuse kontroll läbi viia, kasutades selleks nullhüpoteesi (trendifunktsiooni tegelikud parameetrid on võrdsed nulliga, s.t. uuritavas reas ei olegi trendi) paikapidavuse kontrollimist.

Dispersioonanalüüsi saame kompaktselt esitada dispersioonanalüüsi tabelina.

Varieerumise allikas	Ruutude summa	Vabadus astmete arv	Ruutude keskmine	Empiiriline *F*
Trend (a_i)	$S_1 = \sum(\hat{y}_t - \bar{y})^2$	p	$\frac{S_1}{p}$	$F = \frac{S_1(n-p-1)}{S_2 p}$
Jääkliikmed	$S_2 = \sum(y_t - \hat{y}_t)^2$	$n-p$ - *1*	$\frac{S_2}{n-p-1}$	
Üldine	$S = \sum(y_t - \bar{y})^2$	*n - 1*		

Kus S on uuritava nähtuse üldist varieerumist iseloomustav hälvete ruutude summa, S_1 on uuritava nähtuse varieerumise seda osa, mis on tingitud trendi olemasolust aegreas, peegeldav hälvete ruutude summa, S_2 on jääkvarieerumist iseloomustav hälvete ruutude summa, milles kajastub hälbimine trendifunktsiooni suhtes paljude juhuslike tegurite toimel ja F on kontrollsuurus, mille kontrollimisel Fischeri F-jaotuse tabeliandmetega F_α usaldatavusnivoo $(1-\alpha)\times 100\%$ juures saame teha järeldused kas nullhüpoteesi paikapidavuse ($F < F_\alpha$) või selle ümberlükkamise ($F > F_\alpha$) kohta.

Samasugusele tulemusele jõuaksime ka lineaarse trendi parameetri hinnangu a_1 statistilist usaldatavust kontrollides. Hinnangu a_1 ruutkeskmine ehk standardviga arvutatakse valemiga

$$s_{a_1} = \sqrt{\frac{\sum(y_t - \hat{y}_t)^2}{(n-2)\sum t^2}}$$

ning tõenäosusega $(1-\alpha)\times 100\%$ lineaarse trendi tegelik parameeter α_1 paikneb intervallis

$$\alpha_1 = a_1 \pm t_\alpha s_{a_1},$$

kus t_α on Studenti *t*-jaotuse tabeliväärtus usaldusnivool $(1-\alpha)\times 100\%$. Hinnangu a_1 statistilise usaldatavuse kontrolliks saab kasutada kontrollsuurust

$$t_{a_1} = \frac{a_1}{s_{a_1}},$$

mida tuleb võrrelda *t*-jaotuse tabeliväärtusega etteantud usaldusnivool $(1-\alpha)\times 100\%$ [1].

3.6. Eksponenttasandamine

Eksponenttasandamise eeliseks on see, et eksponentkeskmised arvutatakse iga ajamomendi (perioodi) jaoks sellele momendile (perioodile) vastava tunnuse tegeliku väärtuse ja sellele vahetult eelnevale ajamonedile vastava eksponentkeskmise kaalutud keskmisena

$$S_t = \alpha y_t + (1-\alpha)S_{t-1} = S_{t-1} + \alpha(y_t - S_{t-1}), \tag{21}$$

kus S_t on eksponentkeskmise arvväärtus ajamomendil (-perioodil) *t*, α on tasandusparameeter, mis iseloomustab aegrea tegeliku taseme osakaalu eksponentkeskmise arvutamisel $(0 < \alpha < 1; \alpha = const.)$ [3].

Eksponenttasandamise mudel (21) on juba ise teatud määral isereguleeruv, sest kogu aeg võetakse arvesse aegrea tegeliku taseme ning eelmise ajamomendi (-perioodi) eksponentkeskmise taseme vahe. Mudeli (21) käivitamisel tuleb lahendada kaks olulist probleemi:

(a) algtingimuste, s.t. sisuliselt S_0, valik – kasutatakse kas eelmise perioodide keskmist taset või alustatakse rea algtasemest, s.t. $S_0 = y_1$;

(b) tasandusparameetri α valik, millest oleneb mudeli reageerimiskiirus muutustele uuritavas reas – suurema α arvväärtuse korral reageerimiskiirus suurem, ent siis langeb mudeli abil saavutatav juhuslike hälbimiste filtreerimisvõime [1]. Kui α = 1, siis mingit tasandamist ei toimu (valemiga (21) arvutatud rea liikmed võrduvad originaalrea elementidega). Kui α = 0, siis jäävad kõik rea liikmed konstantseteks ning võrduvad elemendiga S_0 [2].

Näide. Järgmises tabelis on toodud ühelt aksialt saadav tulu (dollarites) 17 aasta jooksul ning kolm erineva parameetriga α (0,6, 0,3, 0,8) tasandatud rida.

Aasta	Aksialt saadav tulu	Tasandatud $\alpha = 0{,}6$	Tasandatud $\alpha = 0{,}3$	Tasandatud $\alpha = 0{,}8$
1	0,54	0,54	0,54	0,54
2	1,15	0,54	0,54	0,54
3	1,13	0,906	0,723	1,028
4	1,69	1,040	0,845	1,110
5	1,53	1,430	1,099	1,547
6	1,79	1,490	1,228	1,539
7	1,34	1,670	1,397	1,740
8	1,20	1,472	1,380	1,420
9	2,02	1,301	1,326	1.244
10	2,38	1,732	1,534	1,865
11	2,97	2,121	1,788	2,277
12	1,82	2,630	2,143	2,831
13	2,07	2,144	2,046	2,022
14	3,15	2,100	2,053	2,060
15	2,92	2,730	2,382	2,932
16	2,73	2,844	2,543	2,922
17	3,00	2,776	2,599	2,768
Prognoos		**2,910**	**2,719**	**2,954**

Tabelis olevatest arvutustulemustest võib tähele panna, et suurema $\alpha = 0{,}8$ puhul on rea tasandamisefekt tunduvalt nõrgem kui väiksema $\alpha = 0{,}3$ korral.Tabeli viimases reas olevad arvud on prognoosid 18. aasta eeldavate tulude kohta. Väiksem tasandusparameeter $\alpha = 0{,}3$ annab konkreetses ülesandes tagasihoidlikuma prognoosi [2].

4. AEGRIDADE INTERPOLEERIMINE, EKSTRAPOLEERIMINE JA PROGNOOSIMINE

Aegridade tasandamine ning üldise arengutrendi uurimine võimaldab teha oletusi vahepealsete puuduvate elementide ehk tasemete arvväärtuste kohta. Aegrea puuduvate elementide arvväärtuste leidmist nimetatakse selle *interpoleerimiseks*, mis sisuliselt tähendab funktsiooni vahepealsete väärtuste arvutamist teadaolevate arvväärtuste järgi. Aegrea interpoleerimine võib toimuda nii aegrea elementaaranalüüsi näitajate kui ka tasandusjoone järgi, kusjuures eelistatavam ja täpsem on viimane variant. Kasutada saab ka graafilist interpoleerimist, mille korral aegrea vahepealsed puuduolevad elemendid leitaksed aegrea muutumist peegeldava arvjoonise järgi.

Ekstrapoleerimiseks nimetatakse sotsiaal- ja majandusnähtuste muutumist iseloomustava vaatlusega hõlmatud osa põhjal tehtud järelduste üldistamist selle vaatlusega mittehõlmatud osale. Ekstrapoleerimine võib olla nii retrospektiivne (ajas tagasi vaatav) kui ka perspektiivne (suunatud eelseisvatele ajaperioodidele või -momentidele). Uuritava nähtuse ajalise muutumise ekstrapoleerimise teoreetiliseks aluseks on sotsiaal- ja majandusnähtuste ajalist muutumist iseloomustavates aegridades peituv inertsus, s.t. nende arengu üldtendentsid säiluvad mõnda aega tõenäoliselt ka tulevikus [1].

Sotsiaal- ja majandusnähtustes täheldatakse, et esiteks on arenguprotsessid dünaamilsed ja pidevalt muutuvad, teisalt sisaldab areng inertsust. On täheldatav, et mida suurem on uuritava sotsiaal- või majandusnähtuste maht, seda suurem on ka inertsus. Seega on inertsus määrava tähtsusega eriti suurte, globaalsete süsteemide uurimisel. Võib öelda, et ekstrapoleerimine on statistiliste prognoosimeetodite ja – mudelite konstrueerimise aluseks.

Lihtasamad prognoosimudelid tuginevad tavaliselt eelneva arengu uurimisele, trendi kindlaksmääramisele ja selle ekstrapoleerimisele tulevikku. Uuritava nähtuse areng ei pea kulgema ühe täpselt kindlaks määratud trendifunktsiooni järgi. Korrektsel prognoosimisel leitakse tavaliselt mitmed prognoosimudelid ja prognoosihinnangud, mida analüüsivad eksperdid, kasutades tasandusfunktsioonide sobivuse ja statistilise usaldatavuse hindamist, empiirilisete andmete hälbimise uurimist tasandusjoonte suhtes jm [5].

Üldiselt levinud soovituseks on, et empiiriline aegrida, mille alusel tehakse kindlaks ekstrapoleerimiseks sobiv trendifunktsioon, peaks olema võimalikult pikk. Kuigi teoreetiliselt on see seisukoht õige, selgub praktilist laadi ülesannete lahendamisel sageli, et sotsiaal- ja majandusnähtuste arengut iseloomustavad pikad aegread on kohati väga ebaühtlase arenguga. Kui taoline ebaühtlane areng oli tingitud mingitest minevikus toimunud spetsiifilistest tingimustest, siis ei oleks mõistlik nende mõju tulevikku üle kanda. Sellistel juhtudel tuleks piirduda uuritava aegrea selle osaga, milles kajastub üldisele arengule omane ühtlane areng.

Ekstrapoleerimist saame kirjeldada mudeliga

$$y_{t+L} = f(y_t^*, L),$$

kus y_{t+L} on aegrea taseme ekstrapoleerimise teel saadud arvväärtus, L on prognoosperioodi pikkus ja y_t^* on ekstrapoleerimise aluseks võetud aegrea tase [1].

5. AEGREA KOMPONENDID JA KOMPLEKSANALÜÜS

Sotsiaal- ja majandusnähtusi peegeldavad aegread sisldavad üldiselt kolme komponenti:

1) arengutendents, s.t. uuritava nähtuse pikaajaline arengusuund, mis iseloomustab analüüsitava nähtuse muutumist;
2) lühiajalised süstemaatilised perioodilised võnked, mis on tingitud sesoonsuse, rütmilisuse, arengu tsüklilisuse mõjust;
3) juhuslik komponent, milles kajastub paljude juhuslikku mõju avalduvate tegurite koondmõju [3].

Eeldades aditiivset seost nende komponentide vahel, võime aegrea avaldada kujul

$$y_t = f(t) + g(t) + \varepsilon_t ,$$

kus *f(t)* on trend, mis avaldub mingi konkreetse trendifunktsioonina, *g(t)* on perioodilised võnked, ε_t on aegrea juhuslik komponent.

Aegrea kompleksanalüüsi all mõistetakse rea sellist käsitlemist, millega empiirilisest reast tuuakse eraldi esile ja uuritakse neid komponente. Seejuures tuleks läbida järgmised analüüsietapid:

1) uuritava aegrea üksikute komponentide kindlakstegemine ja rea lahutamine osadeks nende komponentide järgi;
2) üksikute komponentide põhjalik ja detailne statistiline analüüs;
3) komponentide taasühendamine peale eelnevaid kaht analüüsietappi.

Siinjuures peatume veel lühidalt perioodiliste, eeskätt sesoonsete ehk hooajaliste võngete kindlakstegemisel.

Aegridade iseloomustamiseks võib kasutada vastavaid indekseid, mida on kokku neli:

1) uuritava aegrea üldist varieerumist iseloomustavad indeksid aegrea tasemete y_t suhtena üldkeskmisse, s.t.

$$i_{t,K} = \frac{y_t}{\bar{y}};$$

2) trendi indeksid aegrea tasemete tasandatud arvväärtuse $\hat{y}_t$ suhtena üldkeskmisse $\bar{y}$

$$i_{t,T} = \frac{\hat{y}_t}{\bar{y}};$$

3) hooajalisuse ehk sesoonsuse indeksid aegrea ühenimeliste osaperioodide (kuud, kvartalid jt.) keskmiste $\bar{y}_t$ suhtena üldkeskmisse $\bar{y}$

$$i_{t,S} = \frac{\bar{y}_t}{\bar{y}};$$

4) jääkkomponendi (kui aegrea juhusliku liikme hinnangu) indeks, mis leitakse eelmiste kolme komponendi indeksite kaudu

$$i_{t,J} = \frac{i_{t,K}}{i_{t,T} \times i_{t,S}}.$$

Näide. Leiame regiooni elektrienergia tarbimise hooajalisuse indeksid (absoluutarvud, kwh).

Kuu / Aasta	I	II	III	IV	V	VI	VII
Esimene	69,4	56,3	38,3	31,4	26,7	28,3	21,5
Teine	97,4	78,0	50,2	33,6	42,6	42,9	41,8
Kolmas	111,6	98,6	92,6	61,0	76,5	80,2	82,3
Kokku	278,4	282,9	181,3	126,0	145,8	151,4	145,6
Kuu keskmine	92,8	77,6	60,4	42,0	48,6	50,5	48,5
Hooaja-lisuse indeks (%)	128,8	107,6	83,8	58,3	67,4	70,0	67,4

Kuu / Aasta	VIII	IX	X	XI	XII	Kokku
Esimene	19,9	25,7	47,7	76,9	116,4	558,5
Teine	45,3	53,6	62,8	88,4	123,7	760
Kolmas	84,6	93,6	121,2	149,4	225,2	1277,0
Kokku	149,8	172,9	231,7	314,7	465,3	2595,8
Kuu keskmine	49,9	57,6	77,2	104,9	155,1	72,1
Hooaja-lisuse indeks (%)	69,2	79,9	107,1	145,5	215,1	

Hooajaliste võngete all mõistetakse aegrea elementide arvväärtuste aastasiseseid võnkumisi (suurenemist ja vähenemist), mis korduvad reeglipäraselt aastast aastasse. Hooajalised võnked on näiteks turistide arvu kasv suvel, müüdava puuvilja koguste suurenemine sügisel jne. Hooajalise võnked mõjuvad sageli majanduslikule tegevusele negatiivselt ning neid püütakse vältida või nende mõju vähendada. Hooajalisuse indeksit väljendatakse tavaliselt protsentides [1].

6. AEGRIDADE ANALÜÜS STATISTIKAPAKETIS SPSS

SPSS for Windows on andmetöötlussüsteem, mis pakub väga häid vahendeid andmete haldamiseks ning statistiliseks analüüsiks graafilises keskkonnas. SPSS for Windows sisaldab

- Windows'ile omaseid menüüsid ning dialoogiaknaid.
- Andmeohje süsteemi (Data Editor), mis põhineb töölehel ning pakub võimalusi andmete defineerimiseks, sisestamiseks, redigeerimiseks ning kuvamiseks.
- Süstematiseeritud tulemilehte (Output Navigator), kuhu SPSS paigutab kõik tellitud (analüüsi) tulemused. Sellelt lehelt on lihtne leida huvitav tulem, selektiivselt kuvada või peita tulemit, muuta tulemuste kuvamise järjekorda, lisada kommentaare jne.
- Kõrgkvaliteetseid jooniseid (High-resolution graphics), mida on kasutajal võimalik oma soovi kohaselt redigeerida.

SPSS'i andmeohje süsteemis olevate andmete statistiline iseloom on järgmine

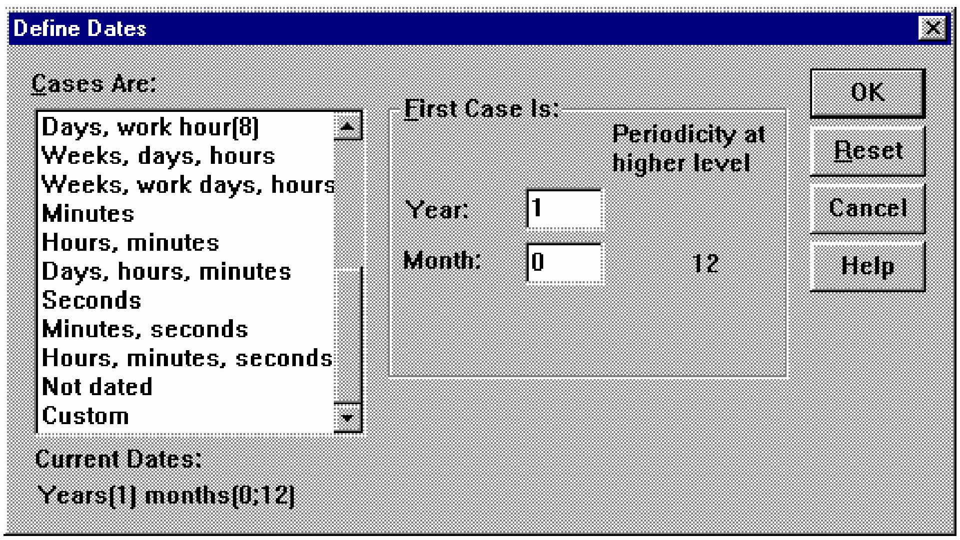
- iga rida töölehel vastab ühele objektile (*case*).
- iga veerg vastab ühele tunnusele (*variable*).
- iga lahter sisaldab ühe tunnuse väärtust konkreetse objekti jaoks. Lahter võib sisaldada ainult andmeid.
- andmetabel on alati nelinurkne. Andmetabeli suurus on määratud tunnuste ja objektide arvuga. Andmetabelis ei ole tühjasid lahtreid: kõiki tühju lahtreid tõlgendab SPSS kui puuduvaid väärtusi (*system-missing values*) ning tähistab need punkti või komaga [4].

6.1. Ajaühikute defineerimine aegridades

Define Dates genereerib tunnuse, mida saab kasutada ajaühiku määramiseks. Näiteks kui teil on vaja sisestada objektidena järjestikused aastad, kuud või nädalapäevad, siis on vaja tunnust, kus iga objekti jaoks on sisestatud konkreetse ajahetke või –perioodi nimetus. SPSS annab uutele tunnustele automaatselt uued nimetused. Uue tunnuse nimetuseks on vastava komponendi nimetus koos alakriipsuga, näiteks YEAR_, QUARTER_, MONTH_, WEEK_, DAY_, HOUR_, MINUTE_, SECOND_ ja DATE_.

Selle tunnuse loomiseks

1) Vali menüü *Data/Define Dates.*



2) Vali *Case Are* loetelust ajaühikud, mida tahad kasutada..

<u>*Cases Are*</u> loetelus

Not dated eemaldab mingi eelnevalt defineeritud aja tunnused.

Custom võimaldab endapoolt loodud käsuvalemiga määratata vajalikud ajaühikud.

3) Sisesta väljale *First Case Is* väärtus(ed), mis on vaadeldava aja alguspunkt. .Järgmised väärtused saavad oma vastava ajalise väärtuse eelnevalt defineeritud perioodi järgi.

Näide 1. Meil oli antud 2000 aasta iga kuine sündimus, suremus ja iive ([8], lk 24) ning tahame neile andmetele genereerida vastava ajaväärtuse. Selleks defineerime dialoogiaknast *Define Dates* rippmenüüst *Year, month*, siis saame oma eelnevale tabelile kolm uut tunnust YEAR_, MONTH_ ja DATE_. Saadud tabel on järgmine

Year_	Month_	Date_	Sündimus	Suremus	Iive
2000	1	JAN 2000	1081	1886	-805
2000	2	FEB 2000	1031	1881	-850
2000	3	MAR 2000	1164	1594	-430
2000	4	APR 2000	1055	1344	-289
2000	5	MAY 2000	1246	1527	-281
2000	6	JUN 2000	1136	1446	-310
2000	7	JUL 2000	1144	1322	-178
2000	8	AUG 2000	1190	1401	-211
2000	9	SEP 2000	970	1352	-282
2000	10	OCT 2000	1069	1502	-433
2000	11	NOV 2000	980	1495	-515
2000	12	DEC 2000	915	1571	-656

Antud tabeli perioodiks on 12 kuud.

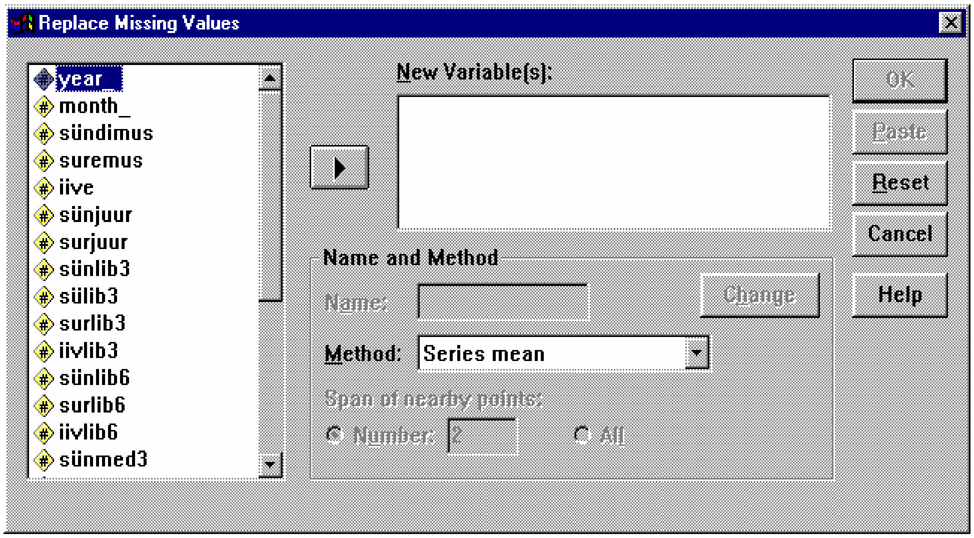
6.2. Puuduvate väärtuste asendamine

Puuduvad vaatlustulemused võivad olla probleemiks aegridade analüüsimisel, sest mõningad aegrea arvnäitajad ei ole sel juhul arvutatavad. *Replace Missing Value* loob olemasoleva aegrea põhjal uue tunnuse, kus puuduvad väärtused asendatakse mingil kindlal meetodil arvutatud hinnanguga.

Aegrea kuuele esimesele uuele tunnusele pannakse automaatselt uus tunnuse nimi, mida on endal võimalik muuta. Uue tunnuse nime kuju on järgmine: algse aegrea nimi alakriips ja järjekorra number. Näiteks kui tunnuse nimi on PRICE, siis esimese uue tunnuse nimi oleks PRICE_1.

Puuduvate väärtuste asendamiseks

1) Vali menüü *Transform/Replace Missing Value*.



2) Vali tunnuste loetelust tunnus, milles sa tahad puuduvad väärtused asendada.

3) Vali rippmenüüst *Method* hindamismeetod, mida tahad kasutada puuduvate väärtuste leidmiseks.

   Puuduvate väärtuste asendamise hindamismeetodid on järgmised

   Series means. Puuduvad väärtused asendatakse andmerea keskmisega.

   Mean of nearby points. Puuduvad väärtused asendatakse ümritsevate väärtuste keskmisega. *Span of nearby points* on puuduvast väärtusest üleval ja allpool asetsevate väärtuste arv, mida kasutatakse keskmise arvutamisel.

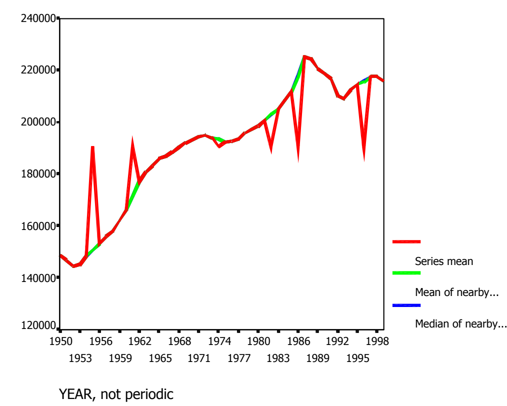
   Median of nearby points. Puuduvad väärtused asendatakse ümritsevate väärtuste mediaaniga. *Span of nearby points* on puuduvast väärtusest üleval ja allpool asetsevate väärtuste arv, mida kasutatakse mediaani arvutamisel.

   Linear interpolation. Puuduvate väärtuste asendamiseks kasutatakse lineaarset interpoleerimist. Viimast olemasolevat väärtust enne puuduvat väärtust ja esimest olemasolevat väärtust pärast puuduvat väärtust kasutatakse interpoleerimiseks. Kui puuduv väärtus on esimene või viimane, siis ei ole võimalik puuduvat väärtust asendada lineaarse interpoleerimise abil.

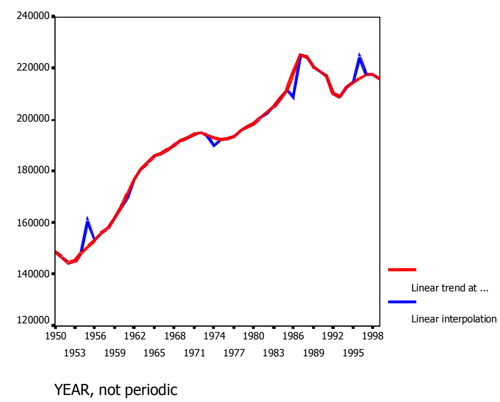
   Linear trend at point. Puuduvad väärtused asendatakse lineaarse trendi punktide abil. Selleks leitakse olemasolevat rida kõige paremini kirjeldav sirge ning puuduvad väärtused asendatakse lineaarse trendi järgi prognoositavate väärtustega.

4) Soovi korral sisesta väljale *Name and Method* uue tunnuse nimi või mõni muu meetod, millega muudad eelnevalt pakutud uue tunnuse nime või puuduvate väärtuste leidmiseks kasutatava meetodi ning vajuta nuppu *Change.*

Näide 1. Meil on antud Eestimaa koolides õppivate õpilaste koguarv aastast 1950 ([9], lk 35), kuid mingitel teadmata asjaoludel puudvad andmereast mõnede aastate õpilaste koguarvud. Nende andmete leidmiseks me kasutame käsku *Replace Missing Value* ja saame tabeli, kus oleme kasutanud andmete asendamiseks kõikvõimalikke meetodeid (vt. Lisa 1). Lisas 1 on puuduvad väärtused asendatud järgmiste meetoditega: tunnus smen on *series means* meetodiga, tunnus mean on *mean of nearby points* meetodiga, tunnus median on *median of nearby points* meetodiga, tunnus interpol on *linear interpolation* meetodiga ja tunnus lineart on *linear trend at point* meetodiga. Graafiliselt on uued andmeread järgmised



Joonis 1.



Joonis 2.

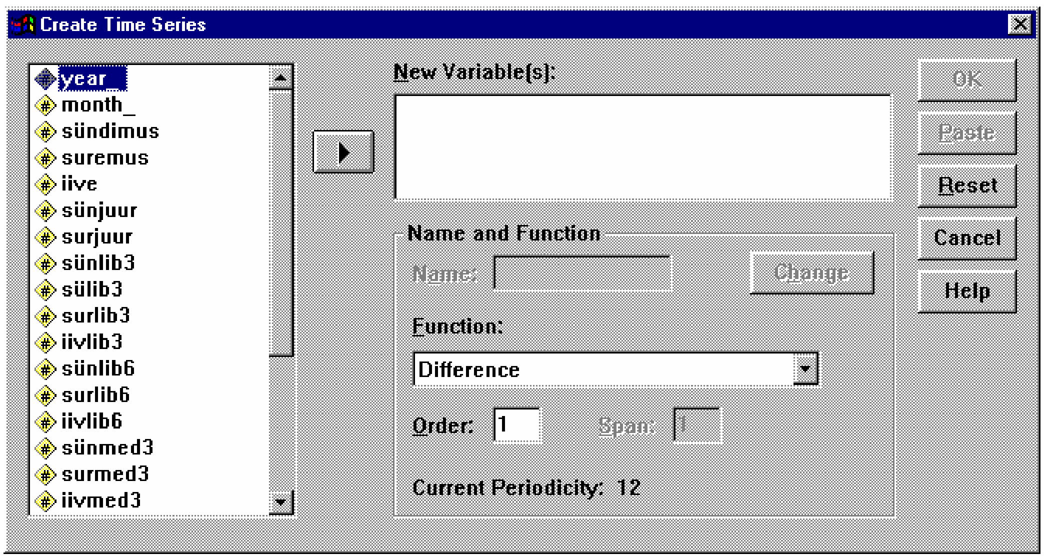
Graafilisest esitusest näeme, et kõige paremini asendasid puuduvaid väärtusi hindamismeetodid *mean of nearby points, median of nearby points* ja *linear trend at point.* Jooniselt 1 näeme, et sinine joonega tähistatud *median of nearby points* meetodiga leitud puuduvad väärtusted ühtivad rohelise joonega tähistatud *mean of nearby points* meetodiga leitud puuduvate väärtustega. Teiste hindamismeetodite korral olid uued väärtused väga erinevad ennustatavatest väärtustest.

6.3. Aegridade teisendamine

Käsk *Create Time Series* on statistikapaketi SPSS üks põhi võimalusi, mille abil on võimalik aegridasid kujundada. Lisaks sellele on võimalik veel kasutada mõne teise trendi dialoogi aknaid ajutiseks aegridade teisendamiseks enne oma andmete analüüsi. *Create Time Series* eesmärgiks on arvutada olemasoleva aegrea põhjal uus aegrida. *Create Time Series* funktsioonidega on võimalik aegridasid siluda liikuva keskmise või –mediaani abil ja leida objekti väärtuste vahelisi erinevusi.

Uue aegrea tunnuse loomiseks

1) Vali menüü *Transform/Create Time Series*



2) Vali väljale *New Variable(s)* tunnus, mida soovid teisendada uueks aegreaks. Kasutada on võimalik ainult arvtunnuseid.

3) Vali funktsioon rippmenüüst *Function*, mida tahad kasutada aegrea teisendamisel. Aegrea teisendamise funktsioonid on järgmised

   Difference on aegreas üksteisele järgnevate väärtuste vaheline mittesesoonne erinevus. Järk *(order)* on eelmiste väärtuste arv, mida kasutadakse erinevuse leidmiseks. Kuna üks vaatlustulemus kaob iga erinevuse järguga, siis uue aegrea alguses asendatakse need vastava puuduva väärtuse tähisega. Näiteks, kui järk on üks, siis leitakse üksteisele järgnevate aegrea liikmete erinevus ning esimesel objektil on uueks väärtuseks süsteemi puuduv väärtus.

   Seasonal difference. Olgu meil eelnevalt ajaühiku määramiseks defineeritud vastav tunnus. Näiteks, kui aegrea objektideks on kuud, siis perioodiks on 12.

Sellisel juhul kui järk (*oder*) on üks ja haare (*span*) on vastav periood, siis arvutatakse iga aegrea elemendi jaoks tema erinevus ühe perioodi võrra eespool olevast aegrea elemendist. Seega järk on perioodide arv, mille tagant aegrea elementide erinevust arvutatakse ning haare on perioodide arv. Puuduvate väärtuste arv rea alguses on võrdne haare korda järk. Näiteks, kui haare on 12 ja järk on 2, siis esimesed 24 liiget uues aegreas tähistatakse puuduva väärtuse tähisega.

Centered moving average on antud objekti ja teda ümbritsevate objektide väärtuste keskmine. Haare (*span*) on aegrea väärtuste hulk, mida kasutatakse keskmise arvutamisel. Kui haare on paarituarv, siis uue aegrea arvväärtused saadakse eelnevalt määratud arvu aegrea järjestikuste liikmete keskmistest, mille leidmisel liigutakse edasi nõnda, et iga uue keskmise arvutamisel jäetakse välja ajaliselt kõige varasem ning lisatakse järelejäänutele ajaliselt vahetult järgnev uus aegrea liige. Kui aga haare on paarisarv, siis leitakse uue aegrea liikmed samuti kui haare on paarituarv, kuid erinevuseks on vaid see, et esialgsetest keskmistest leitakse uued nii, et haare on kaks. Puuduvate väärtuste arv aegrea alguses ja lõpus on võrdne *n/2*, kus *n* on haare. Näiteks kui haare on viis, siis puuduvate väärtuste arv nii aegrea alguses kui lõpus on võrdne kahega.

Prior moving averange on antud objekti väärtuse ja temale eelnenud objektide väärtuste keskmine. Haare (*span*) on eelnevate aegrea väärtuste arv, mida kasutatakse keskmise leidmiseks. Puuduvate väärtuste arv aegrea alguses on võrdne vastava haarde väärtusega.

Running median on antud objekti ja teda ümbritsevate objektide väärtuste mediaan. Haare (*span*) on aegrea väärtuste hulk, mida kasutatakse mediaani arvutamisel. Kui haare on paarituarv, siis uue aegrea arvväärtused saadakse eelnevalt määratud arvu aegrea järjestikuste liikmete mediaanina, mille leidmisel liigutakse edasi nõnda, et iga uue mediaani arvutamisel jäetakse välja ajaliselt kõige varasem ning lisatakse järelejäänutele ajaliselt vahetult järgnev uus aegrea liige. Kui aga haare on paarisarv, siis leitakse uue aegrea liikmed samuti kui haare on paarituarv, kuid erinevuseks on vaid see, et esialgsetest mediaanidest leitakse uued nii, et haare on kaks. Puuduvate väärtuste arv aegrea alguses ja lõpus on võrdne *n/2*, kus *n* on haare. Näiteks kui haare on viis, siis puuduvate väärtuste arv nii aegrea alguses kui lõpus on võrdne kahega.

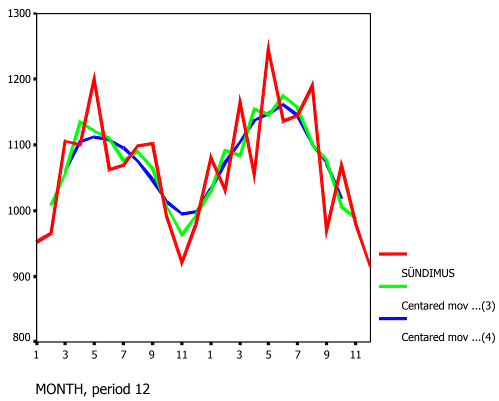
Cumulative sum on andmerea väärtuste kogusumma kuni aegrea lõpuni. Seda ei ole mõtet arvutada momentridade (vt. peatükk 1) korral, kuna sellel ei ole mingit reaalset sisu.

Smoothing. Uue andmerea leidmine baseerub esialgsete andmete silumisele.

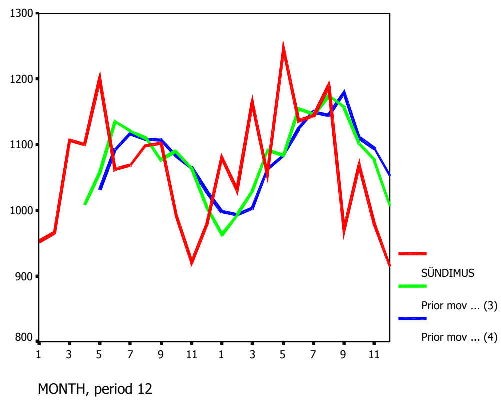
4) Määra vajadusel kas *order* või *span.*

5) Soovi korral võid panna uuele tunnusele nime või valida uue teisendusfunktsiooni, selleks siseta väljale *Name and Function* nimi või vali rippmenüüst uus funktsioon: ning vajuta nuppu *Change.*

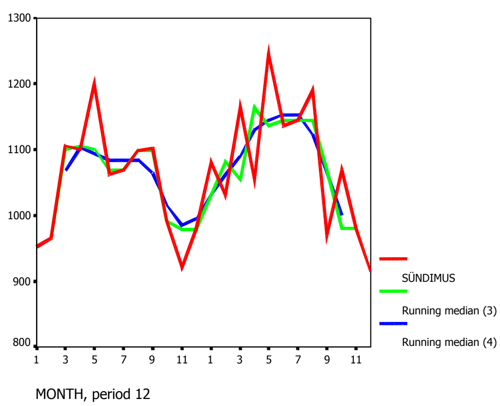
Näide 1. Meil antud kahe aasta sündimus Eestis kuude kaupa ([8], lk 24). Teisendame antud andmeid käsuga *Create Time Series.* Selle tulemusel saame tabeli uute teisendatud andmeridadega (vt Lisa 2). Graafiliselt on uued tasandatud andmeread järgmised



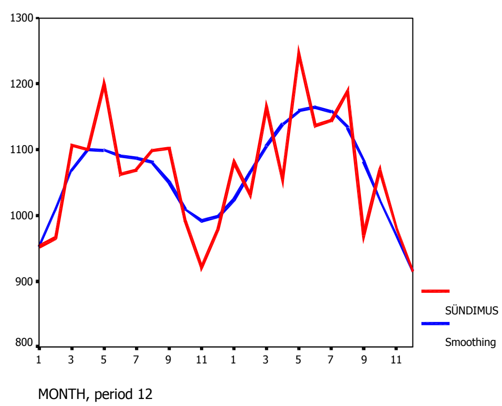
Joonis 1.



Joonis 2.



Joonis 3.



Joonis 4.

Eelnevatelt joonistelt 1 - 4 näeme, et kõik tasandamisfunktsioonid on ühtemoodi efektiivsed. Mida väiksem on keskmise ja mediaani leidmisel haare (vt punkt 3.1.)

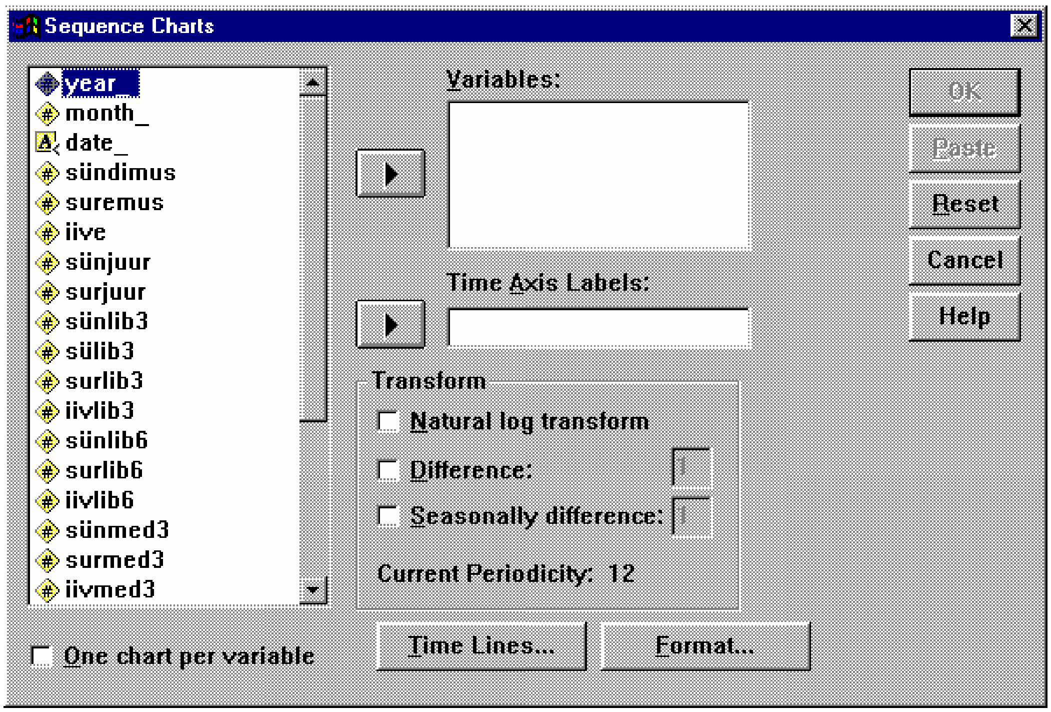
seda sarnasem on tasandatud joon esialgsele. Haarde suurenemisel muutub tasandatud joon järjest siledamaks.

6.4. Aegridade kujutamine graafiliselt

Sequence Charts joonistab graafikuid aegrea objektide jadadest. See meetod nõuab, et meil oleks kas aegread või mõned teised andmed, kus objektid on tähendusrikkas järjestuses.

Aegridade kujutamiseks graafiliselt

1) Vali menüü *Graphs/Sequence*



2) Vali üks või mitu arvtunnust ja vii need väljale *Variable.*

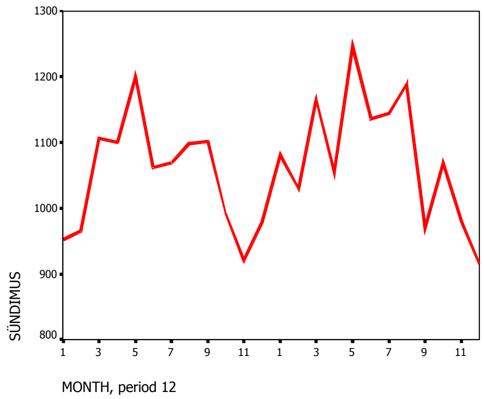
3) Paiguta tunnus, mis on ajatelje ühikute kirjeldus väljale *Time Axis Labels.* Selle tunnuse väärtusi kasutatakse *x*-telje väärtustena. Valitav tunnus võib olla arvtunnus, tekst või pikk tekst.

Valikuliselt on võimalik valida järgmisi andmete teisendamise viise

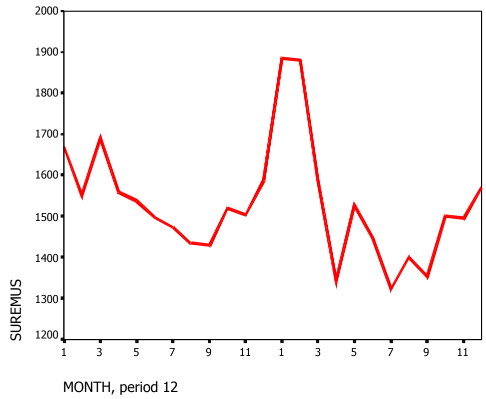
- *Transform* võimaldab valida andmete teisendamiseks kolme teisendusfunktsiooni vahel: *natural log transform, difference* ja *seasonal difference.* Kui valid teisendusfunktsiooni *difference* või *seasonal difference*, siis tuleb sisestada erinevuse järk vastavasse teksti lahtrisse. See number peab olema positiivne täisarv.

- *One Chart per variable* tähistamisel joonistatakse iga tunnus väljalt *Variable* erinevale graafikule, muidu joonistatakse kõik valitud tunnused ühele graafikule.

Näide 1. Kui kasutame käsku *One Chart per variable*, siis saame tabeliandmetest (vt Lisa 3) järgmised graafikud

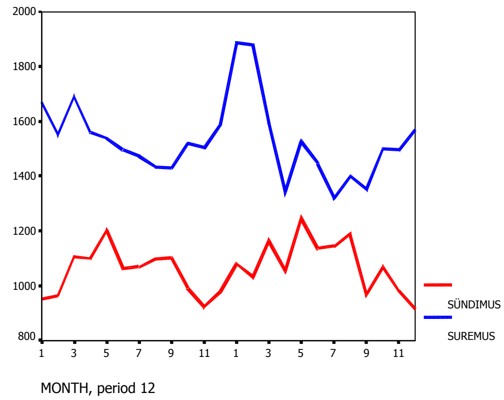


Joonis 1.



Joonis 2.

Kui me ei kasuta käsku *One Chart per variable,* siis saame järgmise graafiku



Joonis 3.

- *Time Lines…* defineerib aja teljega ristuva objekti väärtusi eraldava joone.

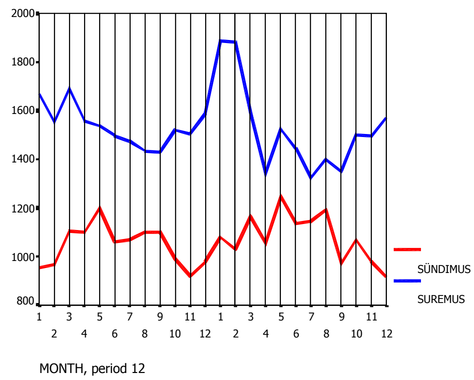
  Dialoogiaknas *Time Axis Reference Lines* on võimalik valikuliselt valida kas

  1) *No reference line*
  2) *Lines at each change of*
  3) *Line at date*

  Kui valida *No reference line,* siis jooniselt puudub objektide väärtusi eraldav joon. Näiteks joonised 1-3.

Kui valida *Lines at each change of* , siis peab valima tunnuse loetelust väljale *Reference Variable* tunnuse, mille alusel need eraldavad jooned joonistatakse.

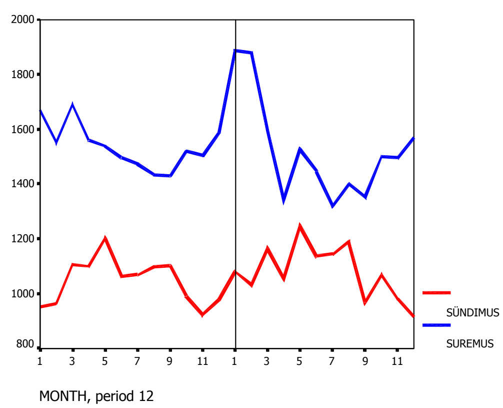
Näide 2. Me tahame leida sündimuse ja suremuse graafikule iga kuu järelt eraldavaid jooni, siis kasutame käsku *Lines at each change of* ning saame järgmise graafiku



Joonis 4.

Line at date joonistab ühe eraldusjoone meie poolt defineeritud ajahetkel või kindla objekti numbri kohas. Kui meil on defineeritud ajaühikud menüüs *Define Dates,* siis tuleb meil sisetada eelnevalt defineeritutud aja komponentidele kindlad arvud, millisel ajahetkel me seda eraldavat joont tahame joonisele saada. Kui meil ei ole defineeritud ajaühikuid menüüs *Define Dates,* siis tuleb sisetada vastava objekti number, kuhu joonisel eraldusjoon tõmmatakse.

Näide 3. Me tahame joonisele saada eraldusjoont kahe aasta vahele, siis peame väljalt *Line at date* märkima, et tahame eraldusjoont 2000 aasta esimese kuu juurde ning saame järgmise graafiku



Joonis 5.

- *Format* dialoogiaknas on võimalik määrata ajatelje paiknevust, valida joon- või pinddiagrammi vahel ning joonisele tuua aegrea keskmist kujutav joon.

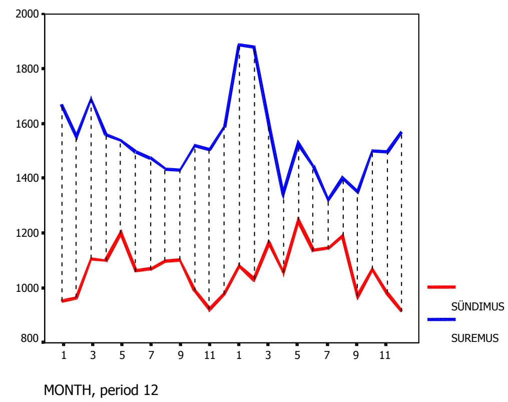
  Kui valida *Time on horizontal axis*, siis asetsevad aega tähistavad väärtused horisontaalteljel ning aegrea väärtused vertikaalteljel. Kui seda ei valita, siis asetsevad aja väärtused vertikaalteljel ning andmerea väärtused horisontaalteljel.

  Kui olete eelnevalt valinud ainult ühe tunnuse või *One chart per variable*, siis saame väljalt *Single Variable Chart(s)* valida joon- või pinddiagrammi.

  Valides *Reference line at mean of series*, siis joonistatakse graafikule rea keskmist kujutav joon.

  Kui sa oled eelnevalt selekteerinud mitu tunnust, siis *Connect cases between variable* annab võimaluse rõhutada kahe tunnuse vahelist erinevust teatud ajahetkel.

  Näide 4. Järgneval graafikul tahame näidata sündimuse ja suremuse vahelist erinevust iga kuu. Käsuga *Connect cases between variable* saame

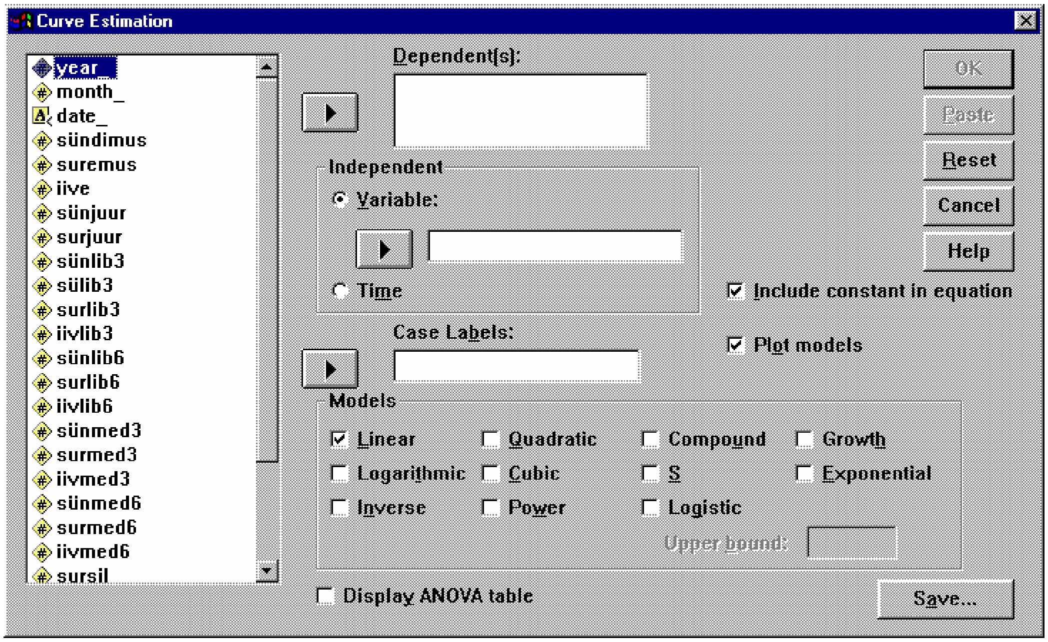


Joonis 6.

6.5. Regressioon

Kõvera hinnangu leidmiseks

1) Vali menüü *Analyze/Regression/Curve Estimation.*

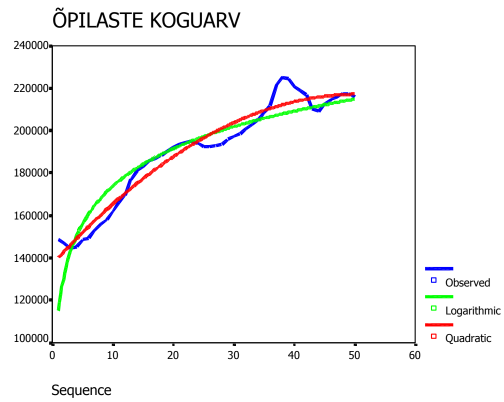


2) Vali tunnuste loetelust üks või mitu sõltuvat tunnust väljale *Dependent(s)*. Igale sõltuvale tunnusele leitakse eraldi vastav lähendusjoon.
3) Vali tunnuste loetelust sõltumatu tunnus väljale *Independent* või märgi *Time.*
4) Vali hindamismudel või -mudelid väljalt *Model.* Sul on võimalik valida 11 erineva hindamismudeli hulgast omale sobiv mudel.
5) Soovi korral on võimalik salvestada prognoositavad väärtused, jääkliikmed (sõltuva tunnuse objekti väärtus miinus mudeli poolt prognoositav väärtus) ja uue tunnuse prognoositavad intervallid (ülemine ja alumine piir). Selleks vajutada nuppu *Save.*

Näide 1. Leiame erinevate hindamismudelite abil lähendjoone Eestimaa koolis õppivate laste koguarvule (vt Lisa 1).

Esimesel joonisel lähendame aegrida logaritm- ja ruutfunktsioonile. Tulemused on järgmised

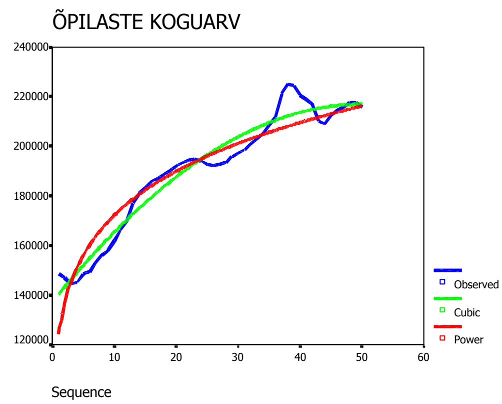
Dependent	Mth	Rsq	d.f.	F	Sigf	b0	b1	b2
KOGUARV	LOG	.874	48	332.98	.000	115055	25507.0	
KOGUARV	QUA	.952	47	466.76	.000	136914	3161.77	-31.146



Joonis 1.

Teisel joonisel lähendame aegrida astme- ja kuupfunktsioonile. Tulemused on järgmised

Dependent	Mth	Rsq	d.f.	F	Sigf	b0	b1	b2	b3
KOGUARV	CUB	.952	46	304.63	.000	137153	3108.03	-28.538	-.0341
KOGUARV	POW	.892	48	394.91	.000	124181	.1418		



Joonis 2.

Kirjandus

[1] Aarma, A., Vensel, V. Statistika teooria põhikursus. Tln., “Külim”, 1996

[2] Käerdi, H. Statistika ja tõenäosusteooria alused.Tln., ERA, 1999

[3] Listra, E. Äristatistika I. Tln., TTÜ, 1998

[4] Niglas, K. Statistilise andmetöötluse pakett SPSS 10.0. Tln., TPÜ, 2001

[5] Roomets, S. Statistika algkursus. Tln.,TPÜ, 1999

[6] Eesti statistika aastaraamat. Tln., ESA, 1995

[7] Eesti Statistika. 2000. Kuukiri nr.10

[8] Eesti Statistika. 2001. Kuukiri nr.1

[9] Haridus 1999/2000. ESA. 2000