

**TILASTOLLINEN TIETOJEN-  
KÄSITTELYJÄRJESTELMÄ  
SURVO 66**

**SEPPO MUSTONEN**

HELSINGIN YLIOPISTON  
TILASTOTIETEEN LAITOS  
STATISTISKA INSTITUTIONEN  
VID HELSINGFORS UNIVERSITET

**MONISTESARJA**

**TAMPEREEN YLIOPISTON TIETOKONEKESKUS**

**Moniste n:o 2**

**TAMPERE 1967**



SEPPÖ MUSTONEN 1967

# TILASTOLLINEN TIETOJEN- KÄSITTELYJÄRJESTELMÄ SURVO 66

SEPPÖ MUSTONEN

SEPPÖ MUSTONEN 1967

**MONISTESARJA**

TAMPEREEN YLIOPISTON TIETOKONEKESKUS

**Moniste n:o 2**

TAMPERE 1967

Copyright 1967 Seppo Mustonen

STATISTISKA INSTITUTIONEN

HELSINGFORS UNIVERSITET

SEPPÖ MUSTONEN 1967

Seppo Mustonen

# TILASTOLLINEN TIETOJEN- KÄSITTELYJÄRJESTELMÄ SURVO 66

SEPPÖ MUSTONEN

Seppo Mustonen

MONISTESARIA

TAMPEREEN YLIOPISTON TIETOKONEKESKUS

Moniste nro 2

TAMPERE 1967

## SISÄLLYSLUETTELO

	sivu
ESIPUHE	1
1. JOHDANTO	4
2. SURVO-OHJELMAN LAATIMINEN	21
2.1. Ohjelman rakenne	21
2.2. Sanat ja luvut	21
2.3. Muuttujat	22
2.4. SURVO-käskyt	25
2.5. Sääteleykäskyt	27
2.6. Muunnoskäskyt	29
2.7. Luokituskäskyt	37
2.8. Taulukointi	41
2.9. Tunnuslukujen laskeminen	47
2.10. Regressio- ja varianssianalyysi	50
2.11. Testit	52
3. SURVO-OHJELMAN KÄYTTÖ	53
3.1. Ohjelman lävistäminen	53
3.2. Havaintomatriisin lävistäminen	54
3.3. Ohjelman kääntäminen (T1-vaihe)	55
3.4. Ohjelman ajaminen (T2-vaihe)	56
3.5. Tulostus (T3-vaihe)	59
3.6. Virheilmoitukset	60
3.7. Muistitilan käyttö	62

## ESIPUHE

Ajatus tilastollisesta ohjelmointikielestä lienee syntynyt käymissäni keskusteluissa Martti Tienarin kanssa vuoden 1962 tienoilla. Tuolloin kehittelimme tilastollisten analyysimenetelmien ohjelmakirjastoja Suomen Kaapelitehtaan Tietokoneosastolla.

Huomion kohteena olivat olleet siihen asti etupäässä laskentatyöltään raskaat menetelmät, joissa tietokoneen tehokkuus on suurimmillaan. Yksinkertaisia menettelyjä varten ei ollut juuri laadittu yleisiä ohjelmia, koska tuntui siltä, ettei tietokonetta kannattanut käyttää yksittäisten, alkeistilastollisten tehtävien suorittamiseen, vaikka nämä ovat käytännössä yleisimpiä.

Ajatuksen kehittämisen kannalta oli ratkaisevaa havaita, että mikäli samanaikaisesti halutaan tehdä useita tilastollisia laskelmia eri menetelmillä samasta tilastoaineistosta, tietokoneen "hyötysuhde" paranee. Tärkeätä tässä on paitsi laskelmien samanaikaisuus myös eri menetelmien suuret rakenteelliset samankaltaisuudet tietojenkäsittelyn kannalta, koska nämä vasta antavat riittävät edellytykset löytää yhteinen ja yleinen kuvaustapa erityyppisille tehtäville. Erittäin tärkeältä tuntui tällöin myös tavoitella sellaista tehtävien kuvaustapaa, jonka tietokoneiden ohjelmointiin perehtymätönkin tutkija saattaisi vaivatta oppia.

Kehittelin ajatusta vuosina 1963-4 ja esitin ensimmäisen ehdotuksen tilastolliseksi ohjelmointijärjestelmäksi syksyllä 1964. Käytin siitä nimitystä SURVO 64; nimitys juontaa alkunsa sanasta "survey". Ehdotus jäi toteuttamatta siirryttyäni vuoden 1965 alussa Yhteiskunnallisen Korkeakoulun Tutkimuslaitokseen, mutta osatoteutuksena syntyi vuoden 1964 lopulla yleinen taulukointiohjelma.

Martti Tienari piti kuitenkin ajatusta vireillä ja niinpä vuonna 1965 Timo Alanko laati hänen kanssaan uuden ehdotuksen nimeltään SURVO 65, jonka toteuttamisesta ryhdyttiin keskuste-

lemaan vuoden 1966 syksyllä. Keskustelut johtivat siihen, että Suomen Kaapelitehtaan Tietokoneosaston ja Tampereen Yliopiston Tietokonekeskuksen välillä tehtiin sopimus, jossa Suomen Kaapelitehtaan Tietokoneosasto tilasi sopimuskumppaniltaan yleisen tilastollisen ohjelmointijärjestelmän SURVO 66 Elliott 803-tietokoneelle. Järjestelmän määrittelystä ja sen toteuttamisen valvonnasta vastaamaan asetettiin sopimuksessa suunnitteluryhmä, johon kuuluivat Kaapelitehtaalta Timo Alanko ja Martti Tienari sekä Tampereen Yliopistosta Pentti Kanerva ja allekirjoittanut projektin johtajana.

Ryhdyin välittömästi sopimuksen solmimisen jälkeen laatimaan uutta SURVO 66-suunnitelmaa. Lopullinen ehdotus, joka rakentuu jossain määrin SURVO 64:n pohjalle, tuli valmiiksi vuoden 1967 keväällä. Tässä suunnittelussa hyödyin paljon keskusteluista suunnitteluryhmän jäsenten kanssa. Olen erityisesti kiitollinen Timo Alangolle monista hyvistä ehdotuksista.

Keväällä 1967 muodostettiin Tampereen Yliopiston Tietokonekeskukseen työryhmä suunnitelman lopullista, käytännöllistä toteuttamista varten. Työryhmään kuuluivat Kari Kärkkäinen, Pentti Kanerva, Leena Lankinen, Tatu Kalin ja Matti Ylinen sekä allekirjoittanut. Ryhmä on pitänyt työnsä aikana viikottain neuvotteluja, joissa on tullut esille useita arvokkaita lisäehdotuksia. Näistä olen kiitollinen varsinkin Pentti Kanervalle ja Kari Kärkkäiselle.

Varsinainen ohjelmointityö on jaettu seuraavasti. Leena Lankinen on ohjelmoinut numeeriset muunnoskäskyt ja käskyn VARAN. Tatu Kalin on laatinut K-ohjelman, jonka avulla ylläpidetään järjestelmään liittyvää aliohjelmakirjastoa, ja ohjelmoinut käskyt MEAN, STDDEV ja CORREL. Matti Ylinen on puolestaan vastannut loogisista muunnoskäskyistä ja käskyistä REGRAN sekä TTEST. Tatu Kalin ja Matti Ylinen ovat lisäksi ansiokkaasti huolehtineet monista järjestelmän viimeistelyyn ja testaukseen liittyvistä toimenpiteistä. Allekirjoittaneen osalle on tullut järjestelmän yleissuunnittelun ja S-ohjelman laatimisen ohella käskyt DEF, CALL, CLASS, TABLE, TRANSF, FRACT, CHI2 ja PRINT.

SURVO-järjestelmä on suunniteltu ja toteutettu siten, että sitä on helppo edelleen kehittää ja laajentaa. Tarkoituksena onkin ryhtyä välittömästi saatujen käyttökokemusten jälkeen tähän kehitystyöhön ja siirtää järjestelmä nykyistä suuremmalle tietokoneelle.

Kiitokset tähänastisen suunnitelman toteutumisesta kohdistuvat paitsi suunnittelu- ja työryhmän jäsenille myös niille yhteisöille, joiden tuella työ on tapahtunut. Kiitän tästä myötävaikutuksesta erityisesti Oy Nokia Elekroniikkaa (ent. Suomen Kaapelitehtaan Tietokoneosasto), Tampereen Yliopiston Tietokonekeskusta, Tampereen Yliopiston Tutkimuslaitosta ja Helsingin yliopiston laskentakeskusta.

Tampereella, Helsingissä ja Kaivokselassa  
joulukuussa 1967

Seppo Mustonen

## 1. JOHDANTO

SURVO 66 on lähinnä alkeistilastollisiin tietojenkäsittelytehtäviin tarkoitettu ohjelmointijärjestelmä. Tavoitteena on ollut tätä järjestelmää suunniteltaessa, että tilastoaineiston käsittely voidaan kuvata ja toteuttaa helpommin kuin yleisen ohjelmointikielen tai valmiin erikoisohjelman avulla.

SURVO-järjestelmän avulla voidaan suorittaa annettuun tilastoaineistoon kuuluvilla muuttujilla ja näistä johdetuilla muuttujilla samanaikaisesti useita tilastollisia operaatioita. Tällaisia ovat:

1. aineiston laadun tarkkailu (muuttujien arvojen järkevyyden, havaintojen sisäinen ja keskinäinen ristiriidattomuus),
2. tunnuslukujen laskeminen (keski- ja hajaantoluvut, korrelaatiot),
3. taulukointi (frekvenssijakautumat, "ristiintaulukointi"),
4. tilastolliset analyysit (varianssi- ja regressioanalyysi),
5. testisuureiden laskeminen (t-testi,  $\chi^2$ -testi),
6. havaintoaineiston muuntaminen jatkotoimenpiteitä varten.

Operaatiot voidaan toteuttaa eri tavoin ehdollisina, mikä tarkoittaa, että operaatio kohdistuu vain halutut ehdot täyttyviin havaintoihin. Kaikille käsittelyn kohteena oleville objekteille, SURVO-objekteille (muuttujat, taulukot, muuttujien luokitustavat, luokat, ehdot jne.) voidaan antaa selväkieliset nimet, mikä helpottaa tehtävän esittämistä (ohjelmointia) ja tulosten tarkastelua. Tulokset saadaan aina SURVO-objektien nimillä ja muilla selväkielisillä tunnuksilla varustettuina kaavioina ja taulukkoina.

SURVO 66:lla käsiteltäväksi tarkoitettu tehtävä on kuvattava laatimalla tietokoneelle SURVO-ohjelma, joka koostuu SURVO-käskyistä ja sisältää kaikki tarpeelliset tehtävän määrittelyyn ja suorittamiseen tarvittavat ohjeet. Koska SURVO-järjestelmää laadittaessa tilastollisten tietojenkäsittelytehtävien ominaispiirteet on otettu jo etukäteen huomioon, ohjelmaa kirjoitettaessa päästään huomattavasti vähemmällä kuin yleistä ohjelmointikieltä (ALGOL, FORTRAN jne.) käytettäessä.



Esim.1. Olkoon tehtävänä laskea aineistosta, jossa on 20 muuttujaa ja 100 havaintoa muuttujien keskiarvot.

Äärimmilleen pelkistetty tämän tehtävän suorittava ALGOL-ohjelma kuuluu:

```
begin real x, sum(1:20); integer i,n;  
  for i:=1 step 1 until 20 do sum(i):=0;  
  for n:=1 step 1 until 100 do  
    for i:=1 step 1 until 20 do  
      begin read x(i); sum(i):=sum(i)+x(i) end;  
    for i:=1 step 1 until 20 do print sum(i)/100  
end
```

Vastaavan tehtävän selvittää SURVO-ohjelma

M@20 N@100

MEAN@X1-X20

END@

täydellisesti.

Yhdenmukaisuus, joka yksinkertaistaa ohjelmia, rajoittaa samalla käsiteltävien tehtävien luonnetta. SURVO-järjestelmä soveltuu käytettäväksi silloin, kun tilastoaineisto voidaan esittää ja on lävistetty reikäkorteille tai reikänauhalle havaintomatriisin muodossa, jossa kunkin havaintoyksikön (koehenkilö, haastateltava, otosyksikkö, tuoteyksikkö, koe-toisto) saamat arvot eri muuttujissa (havaintoyksikön tunnus- ja taustatiedot, testipistemäärät, vastaukset, mittausarvot) ilmaistaan etukäteen sovitussa järjestyksessä.

Nimitys havaintomatriisi seuraa siitä, että tällainen tilastoaineisto voidaan aina esittää taulukkona, matriisina, jossa kutakin havaintoyksikköä vastaa tietty vaakarivi ja kuttakin muuttujaa tietty pystyrivi.

SURVO-järjestelmä käsittelee (niinkuin yleensä muutkin tilastolliset ohjelmat) havaintomatriisia vaakariveittäin, siis havaintoyksiköittäin. Havaintomatriisin vaakarivejä (havaintoyksikön saamat arvot eri muuttujissa) sanotaan havaintovektoreiksi. Usein havaintovektorista käytetään myös nimitystä havainto.

Havaintomatriisia ei tavallisesti kokonaisuudessaan talleteta tietokoneen muistiin, vaan kustakin havaintovektorista

kootaan ainoastaan SURVO-ohjelman tarvitsemat tiedot, jotka suoraan yhdistetään aikaisemmista havaintovektoreista kerättyihin tietoihin.

Havaintomatriisiesitys ei mitenkään olennaisesti rajoita SURVO-järjestelmän käyttöä tilastollisissa tehtävissä, koska käytännössä kaikki järjestelmälliset tietojenkeruu- ja tilastointimenetelmät johtavat tulokseen, joka vastaa havaintomatriisia.

Havaintomatriisit lävistetään joko reikäkorteille tai reikänauhalle havaintovektoreittain siten, että kunkin muuttujan paikka lävistyksessä muiden tietojen suhteen on aina sama havaintoyksiköstä riippumatta. Vain numeeriset muuttujien arvot sallitaan; siis esim. kvalitatiiviset tiedot on koodattava numeerisiksi. Puuttuva havaintoarvo on samoin ilmaistava numeerisesti (muuttujan luonnollisista arvoista poikkeavasti, jolloin SURVO-ohjelmassa on mahdollista varautua tällaisiin tilanteisiin).

SURVO-ohjelma, jolla annetun havaintomatriisin käsittely kuvataan, ei kuitenkaan sisällä mitään tietoja havaintomatriisin lävistystavasta. Ainoa yhteys ohjelman kirjoittajan ja havaintomatriisin välillä ohjelmaa tehtäessä on, että tiedetään SURVO-järjestelmän antavan muuttujille nimet X1, X2, X3, .. siinä järjestyksessä, jossa ne esiintyvät havaintovektorissa.

Esim.2. Olkoon havaintomatriisi koottu Suomen kuntia ja vuoden 1964 kunnallisvaaleja koskevista tiedoista seuraavasti:  
(Lähde: Mitä-Missä-Milloin 1966)

Kunta	Kunta- muoto	Äänioik. (1000)	Äänestys- prosentti	Valtuutettuja yht. ei-sos.	Ennako- veroäyryn hintaa 1965	
Helsinki	1	328	77.2	77	40	12
Hanko	1	6	75.8	27	17	12.5
....						

Kuntamuoto on koodattu seuraavasti: kaupunki=1  
kauppala=2  
maalaiskunta=3

Havaintoyksikköinä ovat siis kunnat ja muuttujina

- X1 = kuntamuoto (1,2 tai 3),
- X2 = äänioikeutettujen lukumäärä tuhansina,
- X3 = äänestysprosentti,
- X4 = valtuutettujen kokonaismäärä,
- X5 = ei-sosialististen valtuutettujen lukumäärä,
- X6 = kunnallisveron ennakveroäyrin hinta v.1965.

Käytämme tätä havaintomatriisia kaikissa seuraavissa johdantoon kuuluvissa esimerkeissä. Näistä esimerkeistä koottu täydellinen SURVO-ohjelma tuloksineen esitetään johdannon lopussa.

Tarkastelemme nyt alustavasti SURVO-ohjelman laatimista. Ensimmäisenä toimenpiteenä on antaa tehtävälle (ohjelmalle) nimi ja ilmaista, montako muuttujaa sisältyy havaintomatriisiin. Tämän jälkeen muuttujien X-nimet (X1, X2, X3,...) ovat käytettävissä. Haluttaessa voidaan X-nimien rinnalle ottaa selväkielisiä muuttujien nimiä.

Esim.3.

VAALITUTKIMUS 1964

M06

CALL X1 KMUOTO  
X2 AANOIK  
X3 AAN%  
X4 VALT  
X5 EISOSV  
X6 AYRI65

on mahdollinen SURVO-ohjelman alku. Ohjelman nimenä on VAALITUTKIMUS 1964. SURVO-käskyllä M06 ilmaistaan, että havaintomatriisiin kuuluu 6 muuttujaa. Muuttujien nimeämiskäskyllä CALL annetaan tässä tapauksessa kaikille muuttujille selväkieliset nimet, jotka luetellaan X-nimien rinnalla.

Yleensä on aiheellista ja joskus suorastaan välttämätöntäkin kuvata muuttujien luonnetta, jotta tietyt toimenpiteet, ennen kaikkea erilaiset tarkistukset, saataisiin ohjelmassa suoritetuksi.

Esim.4. Esimerkkiaineistossamme muuttujat AAN% ja AYRI65 saattavat olla yksidesimaalisia lukuja muiden muuttujien ollessa kokonaislukuja. Mikäli halutaan säilyttää yhden desimaalin tarkkuus näitä muuttujia käsiteltäessä, on annettava seuraavanlainen käsky SURVO-ohjelmassa,

DEF@AAN% AYRI65 S:1 ,

jolloin SURVO-järjestelmä saa tietoonsa käskyssä esiintyvän skaalausparametrin (S:1) avulla tämän muuttujien AAN% ja AYRI65 erikoisominaisuuden.

Esim.5. Jos on aihetta epäillä, että aineiston lävistyksessä on virheitä, jotka aiheuttavat muuttujille mielettömiä arvoja, on syytä tarkkailla muuttujia seuraavasti:

Käsky

DEF@AAN% L:0 U:100

tarkastaa, etteivät muuttujan AAN% arvot ole alle 0 (L:0) eivätkä yli 100 (U:100).

Käsky

DEF@AYRI65 L:8

vaatii, että AYRI65 on vähintään 8.

Haluttaessa johtaa havaintomatriisiin kuuluvista muuttujista uusia, on käytettävissä erilaisia muunnoskäskyjä.

Esim.6. Käsky

SUB@SOSV VALT EISOSV

laskee erotuksen VALT-EISOSV ja asettaa tämän uuden muuttujan SOSV/(=sosialististen valtuutettujen lukumäärä) arvoksi.

Muuttujaa SOSV voidaan tämän jälkeen käyttää ohjelmassa täysin alkuperäisten muuttujien veroisena.

Käskyllä

MULT@AANEST 0.01 AAN% AANOIK S:0

saadaan äänestäneiden lukumäärä AANEST täysinä tuhansina.

(AANEST = 0.01 x AAN% x AANOIK)

Esim.7. Käsky

EQUAL@SOPU AAN% 0

määrittelee uuden loogisen muuttujan SOPU, jolla on joko arvo tosi tai epätosi. SOPU on tosi, jos AAN% = 0 (=vaaleja ei pidetty). SOPU on epätosi, jos AAN% ≠ 0. Kunnat, joissa vaaleja ei pidetty, voidaan tämän jälkeen karsia tarkastelusta kirjoittamalla käsky

STOP@ IF:SOPU .



Edellinen esimerkki kuvasi mahdollisuutta järjestää SURVO-ohjelman toiminta ehdolliseksi. Käskyissä, joissa ehdollinen toiminta on mahdollista, ehto ilmaistaan IF-sanalla, johon on liitetty jokin looginen muuttuja. Käsky toteutetaan vain niiden havaintojen osalta, joissa IF-ehto on voimassa eli looginen muuttuja on arvoltaan tosi.

Esim.8. Lasketaan keskimääräinen äänestysprosentti maalaiskunnissa siten, että kunkin kunnan äänestysprosentti otetaan huomioon äänioikeutettujen lukumäärän osoittamalla painolla. Tämä tapahtuu käskyillä

```
EQUAL@MLK KMUOTO 3  
MEAN@ AAN% W:AANOIK IF:MLK .
```

EQUAL-käskyn määrittelemä looginen muuttuja MLK on tosi vain, jos KMUOTO=3. MEAN-käsky laskee tällä ehdolla (IF:MLK) muuttujan AAN% keskiarvon painottamalla sen arvot muuttujan AANOIK (W:AANOIK) arvoilla.

MEAN-käskyn antama tulos näyttää seuraavalta:

```
IF: MLK N= 466
```

```
WEIGHT VARIABLE: AANOIK  
SUM OF WEIGHTS 1487.00
```

```
VARIABLE MEAN
```

```
AAN% 79.89099
```

MEAN-käsky edellisessä esimerkissä oli ensimmäinen varsinainen tilastointikäsky, joka hoitaa paitsi tuloksen laskemisen myös sen painatuksen tietyssä vakio muodossa. Sama käytäntö koskee kaikkia tilastointikäskyjä.

Esim.9. (Taulukointi)

Muodostetaan kaksiulotteinen frekvenssitaulukko, joka kuvaa kuntamuodoittain äänestysaktiivisuutta äänestysprosentin mukaan luokitettuna. Äänestysprosentteissa käytetään seuraavaa luokkajakoa

luokka äänestysprosentti

ALLE75		-	74.9
75/80	75.0	-	79.9
80/85	80.0	-	84.9
YLI85	85.0	-	

Määritellään ensin vaadittu luokitustapa CLASS-käskyllä

```

CLASS@ AANAKT S:1
      ALLE75 0 74.9
      75/80 75 79.9
      80/85 80 84.9
      YLI85 85 100

```

Tämän jälkeen voidaan antaa taulukointikäsky TABLE, jossa juuri määritelty luokitustapa AANAKT tulee käyttöön muuttujan AAN% arvoja luokiteltaessa:

```

TABLE@ KMUOTO -
      TAULU1 AAN% AANAKT
      M:C%

```

TABLE-käskyn tulostus näyttää seuraavalta:

N= 535

TABLE:TAULU1

```

COLUMN VARIABLE:KMUOTO
ROW VARIABLE: AAN% CLASSIFICATION:AANAKT

```

FREQUENCIES

	1	2	3	TOTAL
ALLE75	3	0	68	71
75/80	24	10	143	177
80/85	16	6	175	197
YLI85	3	7	80	90
TOTAL	46	23	466	535

PER CENT BY COLUMNS

	1	2	3	TOTAL
ALLE75	6.52	-	14.59	13.27
75/80	52.17	43.48	30.69	33.08
80/85	34.78	26.09	37.55	36.82
YLI85	6.52	30.43	17.17	16.82
TOTAL	100.00	100.00	100.00	100.00

Prosenttitaulukon painatus aiheutuu parametrasta M:C% .

Tilastointikäskyjen tuloksiin voidaan usein kohdistaa uusia SURVO-käskyjä.

Esim.10. Edellisessä esimerkissä lasketun frekvenssitaulukon TAULU1  $\chi^2$ -arvo saadaan käskyllä  
CHI2@ TAULU1 .

Esim.11. Suoritetaan viivallinen regressioanalyysi, jossa selitettävänä muuttujana on AYRI65 ja selittävinä muuttujina AANOIK, AAN%, SOS% (=sosialististen valtuutettujen prosentuaalinen osuus) ja MLK (loogisen muuttujan arvoa tosi vastaa numeerinen arvo 1 ja arvoa epätosi numeerinen arvo 0). Suoritus tapahtuu vaiheittain siten, että ensin muodostetaan uusi muuttuja SOS% , sitten korrelaatiomatriisi (+keskiarvot ja -hajonnat), jossa kaikki regressiomalliin tulevat muuttujat ovat mukana ja lopuksi annetaan käsky REGRAN, joka kohdistetaan tähän korrelaatiomatriisiin:

```
MULT@A 100 SOSV
DIV@ SOS% A VALT
CORREL@ AYRI65 AANOIK AAN% SOS% N:KORR
REGRAN@ KORR
        AYRI65 (=SELITETTAVA)
        AANOIK AAN% SOS% MLK (=SELITTAJAT) .
```

SURVO-ohjelma päättyy aina käskyyn END@ . Seuraavassa esitetään edellä olevista esimerkeistä laajentaen koottu täydellinen SURVO-ohjelma ja sen antamat tulokset.

VAALITUTKIMUS 1964

M@6

CALL@X1 KMUOTO

X2 AANOIK

X3 AAN%

X4 VALT

X5 EISOSV

X6 AYRI65

DEF@ AAN% AYRI65 S:1

EQUAL@ SOPU AAN% 0 (EI VAALEJA)

STOP@ IF:SOPU

EQUAL@ MLK KMUOTO 3

NOT@ URB MLK

MEAN@ AAN% W:AANOIK IF:MLK

@ AAN% W:AANOIK IF:URB

CLASS@ AANAKT S:1

ALLE75 0 74.9

75/80 75 79.9

80/85 80 84.9

YLI85 85 100

DEF@ KMUOTO L:1 U:3

TABLE@KMUOTO - T:AYRI65

TAULU1 AAN% AANAKT M:CR%V

CHI2@ TAULU1

TABLE@ TAULU2 KMUOTO - T:AYRI65

VARAN@ TAULU2

SUB@ SOSV VALT EISOSV

MULT@ A 100 SOSV

DIV@ SOS% A VALT

CORREL@ X2-X6 SOSV SOS% MLK N:KORR8

REGRAN@ KORR8

AYRI65 (=SELITETTAVA)

AANOIK AAN% SOS% MLK (=SELITTAJAT)

REGRAN@ KORR8

SOS%

AAN% VALT MLK

CORREL@ AAN% SOS% AYRI65 IF:MLK N:R/MLK

@ AAN% SOS% AYRI65 IF:URB N:R/URB

TTEST@ R/MLK R/URB

END@



VAALITUTKIMUS 1964

CLASSIFICATION: AANAKT  
CLASS LIMITS

ALLE75 .000000 74.90000  
75/80 75.00000 79.90000  
80/85 80.00000 84.90000  
YL185 85.00000 100.00000

VARIABLES

NO.	NAME	SCALE
1	KMUOTO	0
2	AANOIK	0
3	AAN%	1
4	VALT	0
5	EISOSV	0
6	AYR165	1
7	SOSV	0
8	A	0
9	SOS%	0
96	URB	
98	MLK	
99	SOPU	

FREE STORE: 2137

VAALITUTKIMUS 1964

TRANSLATED PROGRAM=

VAALITUTKIMUS 1964

TRANSLATED PROGRAM+DATA 545

VAALITUTKIMUS 1964

N= 545

VAALITUTKIMUS 1964

IF: MLK N= 466

WEIGHT VARIABLE: AANOIK  
SUM OF WEIGHTS 1487.00

VARIABLE MEAN

AAN% 79.89099

VAALITUTKIMUS 1964

IF: URB N= 69

WEIGHT VARIABLE: AANOIK  
SUM OF WEIGHTS 1222.00

VARIABLE MEAN

AAN% 78.77152

VAALITUTKIMUS 1964  
N= 535

TABLE:TAULU1

COLUMN VARIABLE:KMUOTO  
ROW VARIABLE: AAN% CLASSIFICATION:AANAKT

FREQUENCIES

	1	2	3	TOTAL
ALLE75	3	0	68	71
75/80	24	10	143	177
80/85	16	6	175	197
YL185	3	7	80	90
TOTAL	46	23	466	535

PER CENT BY COLUMNS

	1	2	3	TOTAL
ALLE75	6.52	-	14.59	13.27
75/80	52.17	43.48	30.69	33.08
80/85	34.78	26.09	37.55	36.82
YL185	6.52	30.43	17.17	16.82
TOTAL	100.00	100.00	100.00	100.00

PER CENT BY ROWS

	1	2	3	TOTAL
ALLE75	4.23	-	95.77	100.00
75/80	13.56	5.65	80.79	100.00
80/85	8.12	3.05	88.83	100.00
YL185	3.33	7.78	88.89	100.00
TOTAL	8.60	4.30	87.10	100.00

MEANS OF AYR165

	1	2	3	TOTAL
ALLE75	11.667	-	11.956	11.944
75/80	12.875	12.100	12.329	12.390
80/85	12.813	12.667	12.494	12.525
YL185	12.667	12.214	12.825	12.772
TOTAL	12.761	12.283	12.422	12.445

STANDARD DEVIATIONS OF AYR165

	1	2	3	TOTAL
ALLE75	.57735	-	1.2950	1.2721
75/80	.79741	.77460	1.6616	1.5431
80/85	.96393	.51640	1.6153	1.5500
YL185	2.0817	.95119	1.5813	1.5509
TOTAL	.96459	.78032	1.5970	1.5279

VAALITUTKIMUS 1964

CHI2@TAULU1= 18.179063 DF= 6 FREQUENCY MIN.= 3.052

VAALITUTKIMUS 1964  
N= 535

TABLE:TAULU2

ROW VARIABLE:KMUOTO

FREQUENCIES

1	46
2	23
3	466
TOTAL	535

MEANS OF AYRI65

1	12.761
2	12.283
3	12.422
TOTAL	12.445

VAALITUTKIMUS 1964

ANALYSIS OF VARIANCE OF TABLE: TAULU2

SOURCE	SSD	DF	VARIANCE	F
MEAN	82857.877	1		
A	5.4496875	2	2.7248437	1.1679405
RESIDUAL	1241.1736	532	2.3330331	
TOTAL	84104.500	535		

A=KMUOTO



VAALITUTKIMUS 1964  
N= 535

KORR8

VARIABLE	MEAN	STDDEV
AANOIK	5.063551	15.71146
AAN%	79.27645	8.353936
VALT	22.82056	7.681728
EISOSV	12.96075	3.995825
AYR165	12.44486	1.527907
SOSV	9.859813	6.524584
SOS%	40.37009	17.75471
MLK	.8710280	.3354826

CORRELATION MATRIX: KORR8

	AANOIK	AAN%	VALT	EISOSV	AYR165	SOSV	SOS%	MLK
AANOIK	1.000	.0042	.5951	.3930	.0260	.4599	.1650	-.3100
AAN%	.0042	1.000	.1389	-.1356	.1737	.2466	.4129	-.0252
VALT	.5951	.1389	1.000	.5279	.1490	.8541	.4742	-.5017
EISOSV	.3930	-.1356	.5279	1.000	.2088	.0091	-.4399	-.0820
AYR165	.0260	.1737	.1490	.2088	1.000	.0475	-.0362	-.0395
SOSV	.4599	.2466	.8541	.0091	.0475	1.000	.8278	-.5404
SOS%	.1650	.4129	.4742	-.4399	-.0362	.8278	1.000	-.3416
MLK	-.3100	-.0252	-.5017	-.0820	-.0395	-.5404	-.3416	1.000

VAALITUTKIMUS 1964

REGRESSION ANALYSIS

CORRELATION MATRIX: KORR8

VARIANCE OF DEPENDENT VARIABLE AYR165	2.3345
RESIDUAL VARIANCE	2.2300
MULTIPLE CORRELATION	.21155

REGRESSION COEFFICIENTS AND STANDARD DEVIATIONS:

VARIABLE	COEFF	STDDEV	T
CONSTANT	9.8484	.65088	15.131
AANOIK	.00264	.00434	.60845
AAN%	.04409	.00858	5.1400
SOS%	-.01449	.00430	-3.3677
MLK	-.37570	.21466	-1.7502

VAALITUTKIMUS 1964

REGRESSION ANALYSIS

CORRELATION MATRIX: KORR8

VARIANCE OF DEPENDENT VARIABLE SOS%	315.23
RESIDUAL VARIANCE	200.74
MULTIPLE CORRELATION	.60265

REGRESSION COEFFICIENTS AND STANDARD DEVIATIONS:

VARIABLE	COEFF	STDDEV	T
CONSTANT	-31.225	6.4867	-4.8137
AAN%	.76743	.07421	10.341
VALT	.79457	.09327	8.5193
MLK	-8.4684	2.1155	-4.0030

VAALITUTKIMUS 1964

IF: MLK N= 466

R/MLK

VARIABLE	MEAN	STDDEV
AAN%	79.19549	8.605026
SOS%	38.03863	16.84677
AYRI65	12.42167	1.596967

CORRELATION MATRIX: R/MLK

	AAN%	SOS%	AYRI65
AAN%	1.000	.4251	.1729
SOS%	.4251	1.000	-.0839
AYRI65	.1729	-.0839	1.000

VAALITUTKIMUS 1964

IF: URB N= 69

R/URB

VARIABLE	MEAN	STDDEV
AAN%	79.82319	6.430217
SOS%	56.11594	15.68020
AYRI65	12.60145	.9298152

CORRELATION MATRIX: R/URB

	AAN%	SOS%	AYRI65
AAN%	1.000	.4906	.1793
SOS%	.4906	1.000	.3330
AYRI65	.1793	.3330	1.000

VAALITUTKIMUS 1964

T TEST

R/MLK	N=	466	R/URB	N=	69	DF=	533
VARIABLE	MEAN	STDDEV	VARIABLE	MEAN	STDDEV	T	
AAN%	79.195	8.6050	AAN%	79.823	6.4302	-.58214	
SOS%	38.039	16.847	SOS%	56.116	15.680	-8.3906	
AYRI65	12.422	1.5970	AYRI65	12.601	.92982	-.91202	

VAALITUTKIMUS 1964

Kirjoitetun SURVO-ohjelman käsittely tapahtuu pääpiirteittäin seuraavasti.

Ohjelma lävistetään reikänauhalle ja sen tarvitsema tilastoaineisto joko reikänauhalle tai reikäkortteille.

SURVO-ohjelman käyttöä valvoo tietokoneessa SURVO-järjestelmän systeemiohjelma, S-ohjelma, jonka ohjaamana toiminta jakautuu kolmeen päävaiheeseen:

T1-vaihe: SURVO-ohjelman kääntäminen,

T2-vaihe: tilastoaineiston (havaintomatriisin) käsittely,

T3-vaihe: laskennan viimeistely ja tulostus.

T1-vaiheessa S-ohjelma lukee, tarkastaa ja varastoi reikänauhalle lävistetyn SURVO-ohjelman. Samalla kukin SURVO-käskey varaa työskentelyynsä tarvitsemat tilat koneen muistista ja asettaa nämä tilat alkuarvoihinsa (summauspaikat nolllataan jne.)

T2-vaiheessa S-ohjelma lukee havaintomatriisin laajuutta ja lävistystapaa koskevat parametritiedot sekä tämän jälkeen itse havaintomatriisin havaintovektori kerrallaan. Kunkin havainnon kohdalla S-ohjelma käy läpi SURVO-ohjelman käskyt ja jokainen SURVO-käskey kerää tällöin tarvitsemansa tiedot havaintovektorista. Esim. käskey CORREL kerää IF-ehdon havainnoista frekvenssiä ja käskeyn parametreina olevien muuttujien summia, neliösummia ja tulosummia, joista haluttu korrelaatiomatriisi myöhemmin (T3-vaiheessa) muodostetaan.

T3-vaiheeseen siirrytään kaikkien havaintojen tultua käsitellyiksi T2-vaiheessa. T3-vaiheessa S-ohjelma käy vielä kerran SURVO-ohjelman läpi käskey käskeyltä, ja tällöin jokainen käskey hoitaa tarvittavat viimeistelytoimenpiteet ja antaa selväkielisin selityksin varustettuina omat tuloksensa.



## 2. SURVO-OHJELMAN LAATIMINEN

### 2.1. OHJELMAN RAKENNE

SURVO-ohjelma koostuu ohjelman nimestä ja SURVO-käskyistä. Ohjelman nimenä saa olla mikä tahansa merkkijono, joka käsittää numeroita, kirjaimia ja Elliott-reikänauhakoodin erikoismerkkejä merkkiä @ lukuunottamatta. Ohjelman nimi ei ole pakollinen, mutta sen käyttö on kuitenkin suositeltavaa, koska mm. ohjelmalla saadut tulokset varustetaan tällä nimellä ja ovat silloin alkuperältään välittömästi tunnistettavissa.

SURVO-käskyt ovat kaikki muotoa

<SURVO-operaattori>@<parametrilista> .

SURVO-operaattori eli SURVO-käskyn nimi ilmaisee käskyn tehtävän ja siihen liittyy merkin @ välittämänä parametrilista, jolla täsmennetään käskyn suoritustapa. Parametrilistan rakenne riippuu käskyn luonteesta; jopa saman käskyn parametrilistan laajuus vaihtelee käskyn käyttötilanteesta riippuen.

Tämä johtuu siitä, että SURVO-järjestelmässä jokaisella käskyllä on etukäteen sovittu vakiotoimintatapa, josta sallitut poikkeamat ja laajennukset ilmoitetaan lisäparametrien (täsmennysparametrit) avulla. Näin saavutetaan eräitä etuja (turhia parametreja ei tarvitse mainita) käskyjä kirjoitettaessa.

Ohjelma päättyy aina käskyyn END@ . Käskyt toteutetaan siinä järjestyksessä, jossa ne on ohjelmaan kirjoitettu. Vaikka SURVO-käskyt ovat toiminnoiltaan suuressa määrin toisistaan riippumattomia, jokaiselta käskyltä kuitenkin vaaditaan, että ne SURVO-objektit (muuttujat, taulukot, ehdot jne.), joita käskyssä ei uusina määritellä, on määriteltä aikaisemmissä käskyissä.

### 2.2. SANAT JA LUVUT

SURVO-käskyn parametrilistassa on tavallisesti useita käskyn suoritusta tarkentavia parametreja, jotka ovat joko sanoja tai lukuja.

Sanat ovat joko erilaisten SURVO-objektien nimiä tai käskyn suoritusta tarkentavia yleissanoja. Sanalla tarkoitetaan tässä yhteydessä ainakin yhden kirjaimen sisältävää merkkijonoa, joka päättyy joko välilyöntiin tai rivinsiirtoon. Kaikki muut merkit paitsi @ : - ( ) joiden käyttö on sallittu vain

tietyissä erikoistarkoituksissa, ovat mahdollisia sanoissa. Esim. A AAA A12 H2SO4 V1967 PAINO PAINO% 3.KERTA ovat sallittuja sanoja. Esimerkkeinä kielletyistä sanoista mainittakoon

1967 (ei sisällä yhtään kirjainta),

X-Y (- on muuttujaluettelon tunnus, kts.2.3.1),

A:B (: on täsmennysparametrin tunnus, kts.2.4.2).

Edellä esitetyn vastaisesti sanana voi esiintyä pelkästään merkki - , mutta tällä sanalla on aina jokin erikoismerkitys, joka riippuu ao. SURVO-käskyn ominaisuuksista.

Yleensä SURVO-objektien nimiksi tarkoitettut sanat saa valita vapaasti. Poikkeuksina ovat operaattorit (käskyjen nimet) ja muuttujien X-nimet X1, X2, X3,... (kts.2.3.1). Samaa sanaa ei tietenkään saa samassa ohjelmassa käyttää eri tarkoituksiin.

Sanojen pituutta ei ole rajoitettu, mutta SURVO 66 ottaa vain 6 ensimmäistä merkkiä huomioon. Esim. sana ASUKASLUKU tulkitaan samaksi kuin ASUKASTIHEYS, koska molemmat ymmärretään sanana ASUKAS. Sekaannus vältetään esim. valitsemalla lyhenteet ASLUKU ja ASTIH.

Ohjelmassa tarvittavat luvut, jotka ovat tavallisesti parametreina esiintyviä vakioita, kirjoitetaan normaaleja lukujen esitystapoja käyttäen. Sallittuja ovat siis esim. 2 -534 0.25 -75.4 . Luku päättyy samoin kuin sana joko yhteen tai useampaan välilyöntiin tai rivinsiirtoon.

### 2.3. MUUTTUJAT

SURVO-järjestelmässä pystytään käsittelemään numeerisia ja loogisia muuttujia. Seuraavassa käsitellään vain numeerisia muuttujia; loogiset muuttujat esitellään kohdassa 2.6.3.

#### 2.3.1. Muuttujien nimet

Havaintomatriisissa olevien muuttujien lukumäärä  $m$  ilmaistaan SURVO-ohjelmassa käskyllä  $M@m$  , missä vakion  $m$  on oltava positiivinen kokonaisluku. Tämän käskyn jälkeen käsiteltäville muuttujille on käytettävissä X-nimet X1, X2,..., Xm . Nimet valitaan siinä järjestyksessä, jossa muuttujat ovat havaintomatriisissa.

Muuttujan X-nimen saa vaihtaa uudeksi nimeksi, jollaiseksi kelpaa mikä tahansa ohjelmassa ennen käyttämätön sana. Uusien muuttujanimien antamiseen on erityisesti tarkoitettu CALL-käskey (kts. 2.6.1). Esim. käsky

```
CALL@ X3 PAINO
      X7 PITUUS
```

antaa muuttujille X3 ja X7 uudet nimet PAINO ja PITUUS. Muuttujien nimeäminen voi tapahtua myös yleisemmällä määrittelykäskeyllä DEF (kts. 2.6.1) ja johdettujen muuttujien nimet määräytyvät välttämättä niitä määriteltäessä muunnoskäskeyillä (kts. 2.6.2). Esim. käsky

```
SUB@ IKA 1967 SYNT.V
```

määrittelee muuttujan IKA erotuksena 1967 - SYNT.V .

Usein käskeyn peräkkäisinä parametreina joudutaan mainitsemaan useita peräkkäin määriteltyjä (peräkkäiset X-nimet esimerkiksi) muuttujia. Tällainen muuttujaluettelo merkitään kätevimmin SURVO-ohjelmassa muodossa A-B , missä A on muuttujista ensimmäinen (ensimmäisenä määritelty) ja B viimeinen. Esim. X2-X5 on muuttujaluettelon X2 X3 X4 X5 lyhennetty esitys.

### 2.3.2. Muuttujien arvot

Muuttujien arvot varastoidaan SURVO 66:ssa joko kokonaislukuina tai 10:n potensseilla skaalattuina (so. sopivalla vakioluvulla kerrottuina) desimaalilukuina kiinteän desimaalipilkun muodossa (so. desimaaliluvut kokonaisluvuiksi muunnettuina).

Ellei ohjelmassa toisin määrätä, kaikkia muuttujia käsitellään kokonaislukuarvoisina. Täten, jos kaikki havaintomatriisiin sisältyvät muuttujat ja näistä muunnoskäskeyillä johdetut muuttujat ovat kokonaislukuarvoisia, ohjelman kirjoittajan ei tarvitse tietää skaalauksesta mitään. (Skaalauksesta erikseen kohdassa 2.3.3.)

On suositeltavaa, vaikka ei välttämätöntä, että muuttujien arvoja tarkkaillaan SURVO-ohjelmassa, etteivät karkeat esikäsitteily- ja lävistysvirheet pääse häiritsemään tuloksia. Tarkkailuun on käytettävissä paitsi loogiset muuttujat (kts. 2.6.3) muuttujien yleinen määrittelykäskey DEF (kts. 2.6.1), jolla voidaan ilmoittaa mm. muuttujan pienin ja/tai suurin sallittu arvo. Muuttujan rajojen ilmoittaminen ei ole pakollista muulloin kuin käytettäessä muuttujaa luokittelumuuttujana taulukoinnissa ilman luokitustapaa (kts. TABLE-käskey), jolloin SURVO-järjestelmä tarvitsee rajat varatakseen sopivasti tilaa taulukkoja varten.

Jotta laskutarkkuus säilyisi hyvänä ja vältettäisiin ns. ylivuodot laskutoimituksissa, on syytä välttää turhan moninumeroisia muuttujia. Esim. muuttujaa X, jonka alkuperäiset arvot ovat 1950, 1951, ..., 1967 on parasta yksinkertaistaa siirtymällä vaikkapa muuttujaan  $X - 1900$  tai  $X - 1950$  muunnoskäsken SUB avulla; tämä suositus koskee erityisesti tilanteita, joissa muuttujan arvon neliö tai muu vastaava muunnos joudutaan laskemaan (keskihajonnat, korrelaatiot, varianssi- ja regressioanalyysi).

Siis vakiona pysyttelevät etunumerot ja merkityksettömät desimaalit on hyvä poistaa.

### 2.3.3. Muuttujien skaalaus

Muuttujat, joiden arvot eivät ole kokonaislukuja, varastoidaan SURVO-ohjelmassa ilmoitetulla tarkkuudella skaalattuina. Skaalaus tapahtuu kertomalla muuttujan arvot sopivalla  $10^n$  kokonaislukupotenssilla ja varastoimalla nämä tulot muuttujan skaalatuiksi arvoiksi kiinteään desimaalipilkun muodossa (kokonaislukuina). Jos kerroin on  $10^s$ , sanotaan kokonaislukua s muuttujan skaalaksi. Kun  $s > 0$ , skaala s ilmaisee, monenko "desimaalin tarkkuudella" muuttujan arvoja käsitellään SURVO-ohjelmassa. Esim. muuttujan arvo 123.456 varastoidaan

skaalassa	muodossa
4	1234560
3	123456
2	12346 (huom. pyöristys)
1	1235
0	123
-1	12
-2	1
-3	0 .

Tulostettaessa skaalattuja arvoja SURVO-järjestelmä palauttaa luvut oikeaan suuruuteen (edellyttäen, että on käytetty järkevää skaalaa) jakamalla ao. skaalaustekijällä skaalatun arvon. Esim. luvun 123.456 tulostus eri skaaloissa varastoinnin jälkeen näyttää seuraavalta:

skaala	tulostus
4	123.456
3	123.456
2	123.460
1	123.500
0	123.000
-1	120.000
-2	100.000
-3	0.000 .



Muuttujan skaala on silloin, kun kokonaislukuesitykseen eli perusskaalaan, 0-skaalaan ei voida tyytyä, syytä valita huolella, jotta ei menetettäisi liikaa tarkkuutta (liian pieni skaala) tai jouduttaisi varastoimaan älyttömän suuria lukuja (liian suuri skaala).

Muuttujan skaala ilmoitetaan tavallisesti DEF-käskyn skaalausparametrin (S-parametrin) avulla (kts. 2.6.1). Muunnoskäskyillä johdettujen uusien muuttujien skaala määräytyy ensisijaisesti muunnoskäskyssä olevan S-parametrin avulla tai, ellei S-parametria anneta, automaattisesti (kts. 2.6.2).

#### 2.4. SURVO-KÄSKYT

Muuttujiin ja muihin SURVO-objekteihin kohdistuvat toimenpiteet ilmaistaan SURVO-ohjelmassa käskyillä, jotka ovat muotoa

<SURVO-operaattori>@<parametrilista> .

SURVO-operaattori on samalla käskyn nimi ja ilmaisee käskyn tehtävän. Parametrilistan avulla tarkennetaan käskyn suoritusta.

##### 2.4.1. Operaattorit

SURVO-käskyt voidaan jakaa pääryhmiin säätelykäskyt, muunnoskäskyt, luokituskäskyt ja tilastointikäskyt. Jako on kuitenkin suhteellisen mielivaltaisen; muodollisesti kaikki käskyt noudattavat samoja sääntöjä.

SURVO-operaattoreina olevien sanojen käyttö muihin tarkoituksiin on kielletty. Merkki @, joka seuraa aina operaattoria, on kirjoitettava operaattorin perään ilman välilyöntiä.

##### 2.4.2. Parametrit

Parametrilista koostuu parametreista, jotka ovat joko sanoja (muuttujien ja muiden SURVO-objektien nimet, käskyn suoritusta tarkentavat yleissanat) tai lukuja (vakiot). Lisäksi parametrilistaan voi kirjoittaa sulkujen () välissä olevia huomautuksia ja selityksiä, joilla sulkuineen ei ole mitään vaikutusta. Tällaiseksi huomautukseksi kelpaa mikä tahansa merkkijono, joka ei sisällä sulkumerkkejä eikä rivinsiirtoa.

Parametrit jaetaan niiden käskyssä esiintyvän tehtävän mukaisesti pakollisiin parametreihin ja täsmennysparametreihin.

Pakollisia parametreja ei koskaan saa jättää pois parametrilistasta ja ne on mainittava määrättyssä järjestyksessä. Pakollisten parametrien lukumäärä ei kuitenkaan ole aina vakio, vaan sama käsky saattaa eri kerroilla kohdistua eri määrään muuttujia tai muita SURVO-objekteja. Pakollisen parametrin



sijainti muiden pakollisten parametrien suhteen määrää sen tehtävän käskyä toteutettaessa.

Täsmennysparametrit ovat kaksiosaisia ja muotoa

<täsmennyssana>:<parametrisana> .

Täsmennyssana ilmoittaa täsmennysparametrin tehtävän ja parametrinaan voi olla sana tai luku. Täsmennysparametrin saa kirjoittaa mihin kohtaan tahansa parametrilistassa, koska merkkiin : päättyvä täsmennyssana riittää kertomaan, mistä parametrilla on kysymys.

Mahdollisia täsmennyssanoja ovat N, S, L, U, IF, M, T, W ja vastaavia parametrin sanoja sanotaan seuraavassa N-parametriksi, S-parametriksi jne. Käskyssä sallitaan luonnollisesti vain yksi kunkin lajin täsmennysparametri.

Täsmennysparametrin tehtävä riippuu viime kädessä käskystä, jonka yhteydessä sitä käytetään. Täsmennysparametrin poisjättäminen merkitsee yleensä, että ao. täsmennysparametrin arvoksi valitaan käskyn normaalia suoritustapaa vastaava arvo, parametrin alkuarvo.

Täsmennysparametrien tavalliset tehtävät ovat:

täsmennys- sana	tehtävä	parametri- sana	puuttumisen seuraus
N	<u>nimiparametri</u> , nimen antaminen SURVO-objektille tietyissä käskyissä	uusi nimi	SURVO-objekti jää nimettömäksi
S	<u>skaalausparametri</u> , ilmoittaa uuden SURVO-objektin skaalan	kokonais- lukuvakio	riippuu käskystä, yleensä perusskaala
L	<u>muuttujan alaraja</u>	vakio	ei alarajaa
U	<u>muuttujan yläraja</u>	vakio	ei ylärajaa
IF	<u>ehto</u> , jolla käsky toteutetaan T2-vaiheessa	looginen muuttuja	käsky toteutetaan aina
M	<u>menetelmäparametri</u>		normaali toteutustapa
T	taulukoitava muuttuja TABLE-käskyssä	muuttuja	ei taulukoitavaa muuttujaa
W	painomuuttuja MEAN-,STDDEV- ja CORREL-käskyssä	muuttuja	ei painomuuttujaa

Esim. käsky

DEF@ X3 N:PAINO L:20.5 U:99 S:1

antaa muuttujalle X3 nimen PAINO, asettaa sen pienimmäksi sallituksi arvoksi 20.5 ja suurimmaksi 99 sekä ilmoittaa S-parametrin välityksellä, että muuttujan arvot varastoidaan skaalassa 1 (yhden desimaalin tarkkuudella). (Kts. 2.6.1.)

## 2.5. SÄÄTELYKÄSKYT

Säätelykäskyjä, joilla yleisesti vaikutetaan SURVO-ohjelman toimintaan, ovat END, WAIT, STOP, M, N, SPACES, COMMENT ja NAME.

### END-käsky

END@ päättää kirjoitetun ohjelman.

### WAIT-käsky

WAIT@ IF:<ehto>

keskeyttää havaintovektorin käsittelyn T2-vaiheessa, mikäli ehto on voimassa. Jatkettaessa normaalisti (kts. 3.4) jatketaan saman havaintovektorin käsittelyä seuraavasta SURVO-käskystä.

### STOP-käsky

STOP@ IF:<ehto>

päättää havaintovektorin käsittelyn T2-vaiheessa, mikäli ehto on voimassa ja siirrytään seuraavan havaintovektorin käsittelyyn. STOP- ja END-käskyn välillä olevat käskyt jäävät silloin toteutumatta ac. havaintovektorin osalta. STOP-käsky tarjoaa yksinkertaisen keinon hylätä tietyn ehdon täyttävät havainnot. Normaalisti STOP-käsky ilmoittaa teleprintterin välityksellä näin hylättyjen havaintojen tunnukset (järjestysnumerot) ja hylkäämisen syyn (ehdon). (Kts. 3.4.)

### M-käsky

M@m ,

missä m on positiivinen kokonaisluku, asettaa havaintomatriisiin kuuluvien muuttujien lukumääräksi m. M-käsky on annettava aina SURVO-ohjelmassa ja ennen ensimmäistä muuttujien nimiä parametreinaan tarvitsevaa käskyä.

### N-käsky

`N@n` ,

missä n on positiivinen kokonaisluku, ilmoittaa SURVO-ohjelmalla käsiteltävien havaintovektoreiden (havaintojen) lukumäärän n. N-käskyn saa kirjoittaa mihin kohtaan ohjelmaa tahansa; sen saa jättää jopa pois, jolloin havaintojen lukumäärä todetaan T2-vaiheessa. (Kts. 3.4.)

### SPACES-käsky

`SPACES@ k` ,

missä k on positiivinen kokonaisluku, ilmoittaa, montako merkkiä mahtuu riville tulostuksessa käytettävässä teleprinterissä. Käskyllä on merkitystä taulukkomuotoisissa tulostuksissa, joita antavat CORREL- ja TABLE-käskyt. Ellei SPACES-käskyä ole SURVO-ohjelmassa, toiminta vastaa käskyä `SPACES@ 63` , jolloin normaali (kapein) teleprinterin telanleveys riittää. SPACES-käskyllä ilmoitetun telanleveyden ylittävät taulukot painetaan SURVO-järjestelmässä automaattisesti sopiviin lohkoihin ositettuina.

### COMMENT-käsky

`COMMENT@ <merkkijono, joka ei sisällä merkkiä @>`

on tyhjä käsky, jolla voidaan antaa ohjelmaan liittyviä selityksiä.

### NAME-käsky

`NAME@ <nimi> <merkkijono, joka ei sisällä merkkiä @>`

antaa merkkijonolle nimen, jota voidaan myöhemmissä käskyissä käyttää tämän merkkijonon asemasta. Näin saadaan lyhennetyksi ja selvennetyksi ohjelmassa usein toistuvia parametrilistan osia. NAME-käskyllä annettu merkkijonon nimi saa olla uusilla NAME-käskyillä nimettävien merkkijonojen osana.

Esim.

```
NAME@ OSA1 X1 X2 X5 X6 X9
      @ OSA2 X3 X4 X7 X8 X10
      @ KAIKKI OSA1 OSA2
DEF@ KAIKKI L:0 U:9
CORREL@ OSA1 IF:MIES
        @ OSA1
        @ OSA2 IF:MIES
        @ OSA2
```

Tässä ohjelman osassa KAIKKI merkitsee samaa kuin X1 X2 X5 X6 X9 X3 X4 X7 X8 X10 . Huomattakoon, että esimerkissä on käytetty myös toista ohjelman kirjoittamista lyhentävää keinoa; saman SURVO-operaattorin toistuessa monta kertaa peräkkäin ei toistoissa tarvitse mainita muuta kuin pelkkä `@` .

## 2.6. MUUNNOSKÄSKYT

SURVO-järjestelmään kuuluu useita eri muunnoskäskyjä uusien, alkuperäisistä muuttujista johdettavien muuttujien määrittelymiseksi. Muunnoskäskyillä johdetut muuttujat ovat täysin alkuperäisten muuttujien veroisia määrittelyn tapahtuttua ja niitä voi esim. toistuvasti muuntaa edelleen.

Muunnoskäskyihin luetaan tässä yhteydessä myös muuttujien ominaisuuksia kuvaavat määrittelykäskyt. Varsinaisia muunnoskäskyjä on kahta tyyppiä: numeeriset muunnoskäskyt ja loogiset muunnoskäskyt.

### 2.6.1. Määrittelykäskyt

Käytössä olevien muuttujien ominaisuuksia määritellään käskyillä DEF ja CALL.

#### DEF-käsky

Tällä käskyllä voidaan antaa samanaikaisesti usean muuttujan  $u_1, u_2, \dots, u_r$  yhteinen alaraja  $a$ , yläraja  $b$  ja skaala  $s$  sekä yhden muuttujan ( $r=1$ ) tapauksessa myös muuttujan uusi, selväkielinen nimi. Käskyn rakenne on

```
DEF@ u1 u2 ... ur  
      L: {alaraja a}  
      U: {yläraja b}  
      S: {skaala s}  
      N: {uusi nimi} .
```

DEF-käsky aiheuttaa seuraavat toimenpiteet:

1<sup>o</sup> Muuttujien  $u_1, u_2, \dots, u_r$  skaalaksi valitaan  $s$ . Jos  $S$ -parametria ei anneta, muuttujien skaaloihin ei vaikuteta.

SURVO-ohjelman alkaessa kaikkien muuttujien oletetaan olevan perusskaalassa ( $s=0$ ).

2<sup>o</sup> Mikäli käskyssä mainitaan vain yksi muuttuja ( $r=1$ ), tämä saa  $N$ -parametrin mukaisen nimen, edellyttäen, että  $N$ -parametria on käytetty. Usean muuttujan ( $r>1$ ) tapauksessa  $N$ -parametrilla ei ole mitään merkitystä.

3<sup>o</sup> Jos  $L$ - ja/tai  $U$ -parametria on käytetty, tarkastetaan kunkin havaintovektorin osalta, että

$$a \leq u_i \leq b, \quad i=1,2,\dots,r .$$

Puuttuva  $L$ -parametri tulkitaan  $-\infty$ :ksi (alarajaa ei tarkasteta) ja puuttuva  $U$ -parametri  $+\infty$ :ksi (ylärajaa ei tarkasteta). Rajan ylittävästä muuttujan arvosta annetaan virheilmoitus (kts. 3.4) ja virheellinen arvo korjataan lähimpään sallittuun arvoon.

Esimerkkejä DEF-käskyn käytöstä:

Esim.1. DEF@ X12-X15 L:1 U:5 asettaa muuttujien X12, X13, X14 ja X15 alarajaksi 1 ja ylärajaksi 5.

Esim.2. Olkoot X1, X2 ja X3 muuttujia, joiden mahdolliset arvot ovat 0 ja 1. Tiedetään lisäksi, että jokaisessa havainnossa vain yksi niistä saattaa olla arvoltaan 1 eli  $X1+X2+X3$ , joka seuraavassa muunnoskäskyllä ADD määritellään muuttujaksi SUM3, on  $\leq 1$ . Seuraava ohjelman osa tarkastaa nämä seikat.

```
DEF@ X1-X3 L:0 U:1
ADD@ SUM3 X1-X3
DEF@ SUM3 U:1
```

Esim.3. DEF@ X7 S:1 N:PITUUS antaa muuttujalle X7 nimen PITUUS ja ilmoittaa, että muuttujan arvot on varastoitava skaalassa 1 (yhden desimaalin tarkkuudella).

### CALL-käsky

Käskyä käytetään pelkästään muuttujien selväkielisten nimien antamiseen. CALL-käsky on tarpeen vain, ellei muilla käskyillä (DEF ja muunnoskäskyt) ole annettu nyt muuttujalle haluttavaa nimeä. CALL-käskyn yleinen muoto on

```
CALL@ u1 nimi1
      u2 nimi2
      .. ..
      ur nimir ,
```

missä  $u_1, u_2, \dots, u_r$  ovat tunnettuja muuttujanimeä (tavallisesti X-nimeä) ja  $nimi_1, nimi_2, \dots, nimi_r$  näiden uudet nimet. CALL-käskyyn kirjoitetuilla täsmennysparametreilla ei ole vaikutusta.

Uusia nimiä annettaessa on kiellettyä valita sellaisiksi X-nimeä. Havaintomatriisiin kuuluvan muuttujan X-nimi on uuden nimen ohella käytettävissä samanveroisena kaikissa käskyissä. Tulostuksessa SURVO-järjestelmä käyttää kuitenkin aina ensisijaisesti selväkielistä nimeä. Muuttujalla ei voi olla yhdellä kertaa yhtä useampaa selväkielistä nimeä. CALL-käskyllä saa kuitenkin nimen vaihtaa missä kohdassa ohjelmaa tahansa, jolloin vain viimeksi määriteltä nimi on käytettävissä X-nimen ohella.

Esim. Ohjelma on kirjoitettu käyttäen englanninkielisiä muuttujanimeä, jotka on määriteltä ohjelman alussa vaikkapa CALL-käskyllä

```
CALL@ X2 LENGTH
      X3 WEIGHT
      X4 RADIUS
      .. ..
```



Ohjelmassa on käytetty näitä selväkielisiä nimiä useissa käskyissä. Haluttaessa "suomentaa" muuttujien nimet tulostusta varten päästään vähimmällä (muuttamatta ainoatakaan käskyä) lisäämällä ohjelman loppuun sopivasti ennen END-käskyä joko

```
CALL@ X2 PITUUS
      X3 PAINO
      X4 SAEDE
```

tai

```
CALL@ LENGTH PITUUS
      WEIGHT PAINO
      RADIUS SAEDE
```

### 2.6.2. Numeeriset muunnoskäskyt

Numeeriset muunnoskäskyt on tarkoitettu uusien numeeristen muuttujien muodostamiseen tai entisten muuttujien arvojen muokkaamiseen. Käskyjen rakenne noudattaa yleensä seuraavaa kaaviota:

```
<operaattori f>@<tuloksena oleva muuttuja u>
      <operandiluettelo u1 u2 ... ur>
      S:<skaala s>
      IF:<ehto>
```

Muunnoskäsky f määrää kussakin havaintovektorissa muuttujan u arvon operandien  $u_1, u_2, \dots, u_r$  funktiona eli

$$f(u_1, u_2, \dots, u_r) \rightarrow u$$

Esim.1. Käsky ADD suorittaa operaation  $u_1 + u_2 + \dots + u_r \rightarrow u$  kirjoittamalla

```
ADD@ u u1 u2 ... ur
```

Tuloksena oleva muuttuja u saa olla joko aikaisemmin määritetty (nimeltään tunnettu) tai uusi muuttuja. Muuttujan u ollessa uusi ei ole sallittua käyttää siitä X-nimeä. Mikäli ohjelman kirjoittaja ei halua "selväkielisiä" nimiä, on yksinkertaisinta ottaa johdetuille muuttujille esim. tyyppiä Y1, Y2, Y3, ... olevia nimiä.

Parametrit (operandit)  $u_1, u_2, \dots, u_r$  saavat joko kokonaan tai osittain olla vakioita.

Muunnoskäsky toteutuu, mikäli käsiteltävänä oleva havainto täyttää IF-ehdon. Ellei IF-ehto ole voimassa, käsky ei muuta muuttujan u arvoa. S-parametrin merkitystä kuvataan jäljempänä.

Numeerisia muunnoskäskyjä ovat SET, ADD, SUB, MULT, DIV, MOD, SQRT, LOG, EXP, MAX, MIN, ORDER, ja LAG. Erikseen käsi-

tällä muuttujien yleistä muunnoskäskeyä TRANSF (kts. 2.7.2).  
Numeeriset muunnoskäskyt toimivat seuraavasti:

käskey	toiminta
SET@ u u <sub>1</sub>	$u_1 \rightarrow u$
ADD@ u u <sub>1</sub> u <sub>2</sub> ... u <sub>r</sub>	$u_1 + u_2 + \dots + u_r \rightarrow u$
SUB@ u u <sub>1</sub> u <sub>2</sub>	$u_1 - u_2 \rightarrow u$
MULT@ u u <sub>1</sub> u <sub>2</sub> ... u <sub>r</sub>	$u_1 u_2 \dots u_r \rightarrow u$
DIV@ u u <sub>1</sub> u <sub>2</sub>	$u_1 / u_2 \rightarrow u$
MOD@ u u <sub>1</sub>	$ u_1  \rightarrow u$
SQRT@ u u <sub>1</sub>	$\sqrt{u_1} \rightarrow u$
LOG@ u u <sub>1</sub>	$\log_e u_1 \rightarrow u$
EXP@ u u <sub>1</sub>	$\exp u_1 \rightarrow u$
MAX@ u u <sub>1</sub> u <sub>2</sub> ... u <sub>r</sub>	$\max(u_1, u_2, \dots, u_r) \rightarrow u$
MIN@ u u <sub>1</sub> u <sub>2</sub> ... u <sub>r</sub>	$\min(u_1, u_2, \dots, u_r) \rightarrow u$
ORDER@ u	havaintovektorin järjestysnumero $\rightarrow u$
LAG@ u u <sub>1</sub> k	u <sub>1</sub> :n arvo k havaintoa aikaisemmin $\rightarrow u$

Seuraavat tilanteet muunnoskäskyyä toteutettaessa aiheuttavat virheilmoituksen (kts. 3.6).

käskey	tilanne	virhenumero
DIV	$u_2 = 0$	29
SQRT	$u_1 < 0$	30
LOG	$u_1 \leq 0$	31

Virheilmoituksen jälkeen u asetetaan arvoon 0. On hyvä tarkkailla DEF-käskyn avulla näiden muunnoskäskyyjen operandien arvoja, jotta turhat virheilmoitukset vältettäisiin.

Esim.2.

```
SET@ X 1
@ X 2 IF:LAJI2
```

Muuttujalle X annetaan arvo 2, mikäli ehto LAJI2 on voimassa ja arvo 1, mikäli ehto LAJI2 ei ole voimassa.

Esim.3.

```
MULT@ TILAVUUS PITUUS LEVEYS KORKEUS
```

Esim.4. Lauseke  $Y = \sqrt{X_1^2 + X_2^2}$  lasketaan käskyillä

```
MULT@ Y X1 X1
      @ A X2 X2
ADD@ Y Y A
SQRT@ Y Y
```

Tässä A on käytössä pelkkänä apumuuttujana. Ellei esim. muuttujaa X2 tarvita enää myöhemmissä käskyissä, A:n paikalle voidaan kirjoittaa X2.

Esim.5. Käskyparilla

```
MAX@ X7 X7 1
MIN@ X7 X7 5
```

on sama vaikutus muuttujan X7 arvoihin kuin käskyllä

```
DEF@ X7 L:1 U:5 .
```

LAG-käskyllä siirretään informaatiota havaintovektorista toiseen ja näin tulee mahdolliseksi tutkia havaintojen keskinäisiä riippuvuussuhteita. LAG-käsky on ainoa käsky ORDER-käskyn ohella koko SURVO-järjestelmässä, joka edellyttää, että havaintovektorit käsitellään määrättyssä järjestyksessä.

Annettaessa käsky LAG@ u u<sub>1</sub> k valitaan muuttujan u arvoksi k ensimmäisen havaintovektorin kohdalla muuttujan u<sub>1</sub> arvo ensimmäisessä havaintovektorissa. Käyttäen ORDER-käskyä ja muita muunnoskäskyjä ehdollisina saadaan tämä sopimus muutetuksi halutunlaiseksi.

Esim.6. Tutkitaan muotoa

$$x_t = ax_{t-1} + bx_{t-2} + c \quad , \quad t=1,2,\dots,n$$

olevaa, aikasarjaa  $x_t$  liittyvää regressiomallia, jossa a, b ja c ovat estimoitavia parametreja. Oletetaan, että  $x_1 = x_0 = x_{-1}$ . Olkoon muuttujan  $x_t$  nimenä XT. Regressiomallissa tarvittavat selittäjät  $x_{t-1} = XT,1$  ja  $x_{t-2} = XT,2$  saadaan käyttöön käskyillä

```
LAG@ XT,1 XT 1
      @ XT,2 XT 2 .
```

Regressiomalli saadaan lasketuksi tämän jälkeen seuraavasti:

```
CORREL@ XT XT,1 XT,2 N:SARJA2
REGRAN@ SARJA2
      XT
      XT,1 XT,2
```

Muunnoskäskeyn tulos u varastoidaan S-parametrin osoittamassa skaalassa. Alkuperäiseen havaintovektoriin kuuluvilla muuttujilla skaalaa ei kuitenkaan muuteta.

Jos S-parametria ei mainita käskeyssä, skaala valitaan muuttujalle u seuraavasti:

Merkitään  $s_i$  = operandin  $u_i$  skaala,  $i=1,2,\dots,r$   
 $s$  = tuloksen u skaala

käskey	skaala s
SET@ u $u_1$	$s = s_1$
ADD@ u $u_1 u_2 \dots u_r$	$s = \max(s_1, s_2, \dots, s_r)$
SUB@ u $u_1 u_2$	$s = \max(s_1, s_2)$
MULT@ u $u_1 u_2 \dots u_r$	$s = s_1 + s_2 + \dots + s_r$
DIV@ u $u_1 u_2$	$s = s_1 - s_2$
MOD@ u $u_1$	$s = s_1$
SQRT@ u $u_1$	$s = s_1$
LOG@ u $u_1$	$s = s_1$
EXP@ u $u_1$	$s = s_1$
MAX@ u $u_1 u_2 \dots u_r$	$s = \max(s_1, s_2, \dots, s_r)$
MIN@ u $u_1 u_2 \dots u_r$	$s = \max(s_1, s_2, \dots, s_r)$
ORDER@ u	$s = 0$
LAG@ u $u_1 k$	$s = s_1$

Yleinen periaate skaalan automaattisessa valinnassa on, että tulos saatetaan tarkimpaan järkevään skaalaan.

### 2.6.3. Loogiset muuttujat ja muunnoskäskeyt

Ehtoa ilmaisevan IF-parametrin parametrinaan on aina looginen muuttuja (Boolean muuttuja), jonka ainoat mahdolliset arvot kunkin havainnon kohdalla ovat tosi ja epätosi. Jos IF-parametrina oleva looginen muuttuja on tosi, sanotaan, että IF-ehto on voimassa ja havainto kuuluu käskeyn toimintapiiriin. Jos looginen muuttuja on epätosi, IF-ehto ei ole voimassa, eikä käskeyä toteuteta ao. havainnon kohdalla.

Loogisia muuttujia saa käyttää myös numeerisina muuttujina, jolloin arvoa tosi vastaa 1 ja arvoa epätosi 0 (skaalana peruskaala). Myös tavallista numeerista muuttujaa voi tällöin käyttää loogisena muuttujana, jolloin arvoa tosi vastaavat kaikki muuttujan arvot  $\neq 0$  ja arvoa epätosi muuttujan arvo 0.

Loogisten muuttujien määrittelyyn ja muuntamiseen on käytettävissä loogiset muunnoskäskyt EQUAL, LESS, LESSEQ, BETWEEN, OR, AND, NOT.

Käskyjen esittelyssä käytetään merkintöjä:

$e$  = looginen muuttuja, jonka arvon muunnoskäsky määrää,  
 $e_1, e_2, \dots, e_r$  = loogisia muuttujia, joiden arvot tunnetaan,  
 $u_1, u_2, u_3$  = numeerisia muuttujia tai vakioita.

käsky	ehto, jolla e=tosi
EQUAL@ e $u_1$ $u_2$	$u_1 = u_2$
LESS@ e $u_1$ $u_2$	$u_1 < u_2$
LESSEQ@e $u_1$ $u_2$	$u_1 \leq u_2$
BETWEEN@ e $u_1$ $u_2$ $u_3$	$u_1 \leq u_2 \leq u_3$
OR@ e $e_1$ $e_2$ ... $e_r$	ainakin yksi muuttujista $e_1, e_2, \dots, e_r$ on tosi
AND@ e $e_1$ $e_2$ ... $e_r$	kaikki muuttujat $e_1, e_2, \dots, e_r$ ovat tosia
NOT@ e $e_1$	$e_1$ =epätosi

Esim.1.

EQUAL@ MIES X1 1  
 @ NAINEN X1 2

MIES on tosi, jos X1=1 ja NAINEN on tosi, jos X1=2. Jälkimmäisen käskyn asemasta voitaisiin NAINEN määritellä myös käskyllä

NOT@ NAINEN MIES .

Esim.2.

LESS@ NUORI IKA 18  
 AND@ POIKA NUORI MIES

Käskyihin EQUAL, LESS, LESSEQ ja BETWEEN voi kirjoittaa myös S-parametrin. Se määrää tarkkuuden, jolla loogisen muuttujan arvon antavaa epäyhtälöä tarkastellaan. Tämä tapahtuu esim. EQUAL-käskyssä laskemalla erotus  $u_1 - u_2$  tarkimmassa järkevässä skaalassa, siirtämällä erotus S-parametrin osoittamaan skaalaan ja tutkimalla, onko erotus tämän jälkeen =0. Muissa käskyissä menettely on täysin vastaava.

Esim.3. Käskyssä

LESS@ ALLEL X2 1 S:1

ALLEL saa arvon tosi muuttujan X2 skaalasta riippuen seuraavasti:

X2:n skaala	ALLEL=tosi, jos
0	$X2 \leq 0$
1	$X2 \leq 0.9$
2,3,4,...	$X2 \leq 0.95$

#### 2.6.4. PRINT-käsky

Muunnoskäskyjä täydentää PRINT-käsky, jonka avulla tulostetaan T2-vaiheen aikana havainnoittain alkuperäisten ja johdettujen muuttujien arvoja havaintomatriisin muodossa reikänauhalle. PRINT-käsky tekee täten mahdolliseksi muokata muunnoskäskyjen avulla (jopa ehdollisesti) havaintomatriisi uuteen muotoon, jota voidaan käyttää myöhemmissä SURVO-ohjelmissa ja muissa tilastollisissa ohjelmissa. PRINT-käskyn yleinen rakenne on

```
PRINT@ u1 u2 ... ur  
      M: <tulostuskanavan numero>  
      IF: <ehto> .
```

Käsky tulostaa havainnoittain summatarkistettuna muuttujien  $u_1, u_2, \dots, u_r$  arvot M-parametrin ilmoittaman tulostuslaitteen kautta. M-parametrin mahdolliset arvot ovat 1,2,3, joista 1 ja 2 tarkoittavat reikänauhanlävistäimiä sekä 3 suoraankytkettyä teleprinteriä. Korkeintaan 2 erilaista PRINT-käskyllä tulostettavaa havaintomatriisia voidaan siis ottaa samanaikaisesti reikänauhalle.



## 2.7. LUOKITUSKÄSKYT

Paitsi muunnoskäskyillä myös luokituskäskyillä muunnetaan muuttujien arvoja ryhmittämällä halutut muuttujan arvoalueen osat omiksi luokikseen.

Muuttujien arvojen luokittelussa on tärkein tehtävä CLASS-käskyllä, jolla määritellään luokitustavat eli ne säännöt, joilla muuttujan arvot kuvataan luokkiin (luokkanimiksi, luokkanumeroiksi). Kullakin CLASS-käskyllä määritellyllä luokitustavalla on oma nimensä, johon viittaamalla luokitustapa saadaan käyttöön eri muuttujien luokituksessa. Luokitustapoja käytetään käskyissä TRANSF ja TABLE.

Tässä esitellään CLASS-käskyn ohella TRANSF-käsky, joka on CLASS-käskyyn perustuva yleinen muuttujien muunnoskäsky.

### 2.7.1. CLASS-käsky

CLASS-käskyllä määritellään muuttujille tarkoitettu luokitustapa. Tämä tapahtuu kirjoittamalla käskyn parametreiksi luokitustavan nimi ja luokkien määritykset, joista jokainen käsittää luokan nimen sekä muuttujan alimman ja ylimmän tähän luokkaan kuuluvan arvon. Käskyn rakenne on siis seuraava:

```
CLASS@ <luokitustavan nimi>  
      <luokan nimi  $n_1$ > <alaraja  $a_1$ > <yläraja  $b_1$ >  
      <luokan nimi  $n_2$ > <alaraja  $a_2$ > <yläraja  $b_2$ >  
      ...  
      <luokan nimi  $n_r$ > <alaraja  $a_r$ > <yläraja  $b_r$ >  
      M:<luokitusmenetelmä>  
      S:<skaala s>
```

Näin määritelty luokitustapa on käytettävissä kaikille skaalassa  $s$  (ellei S-parametria anneta,  $s=0$ ) varastoiduille muuttujille. Luokkaan  $n_i$  ( $i=1,2,\dots,r$ ) sijoitetaan sellaiset havaintoarvot  $x$ , joille pätee  $a_i \leq x \leq b_i$ . Luokkanimet  $n_i$  saavat olla osittain samoja; luokat siis voivat koostua useista erillisistä väleistä. Kaikki luokkanimet  $n_i$  ovat joko ei-negatiivisia kokonaislukuja tai kaikki ovat sanoja; luokitustapaa, jossa luokkaniminä on osittain sanoja, osittain kokonaislukuja, ei sallita.

Esim.1. Olkoot mielipidettä mittaavat muuttujat koodattu muodossa

- 1= täysin eri mieltä
- 2= osittain eri mieltä
- 3= melkein samaa mieltä
- 4= täysin samaa mieltä
- 5= ei osaa sanoa .

Muodostetaan uusi luokitustapa

- 1= samaa mieltä (3,4)
- 2= ei osaa sanoa (5)
- 3= eri mieltä (1,2) .

Tämä luokitustapa (MIELI) syntyy CLASS-käskyllä

```
CLASS@ MIELI
  1  3  4
  2  5  5
  3  1  2 .
```

Esim.2. Edellisessä esimerkissä oleva luokitustapa voidaan määritellä myös "selväkielisin" luokkanimien muodossa

```
CLASS@ MIELI
  SAMA  3  4
  EISANO 5  5
  ERI    1  2 .
```

Jokaisella luokalla on luokkanumero, joka luokkanimien ollessa kokonaislukuja on sama kuin luokkanimi. Jos luokkanimet ovat sanoja, luokat numeroidaan 1,2,3,... luokkien esiintymisjärjestyksessä. Esimerkissä 2 luokkia SAMA, EISANO ja ERI vastaavat tässä järjestyksessä luokkanumerot 1,2 ja 3.

Käytettäessä luokkaniminä kokonaislukuja ja sovellettaessa luokitustapaa taulukoinnissa (kts. TABLE-käsky) varataan myös pienimmän ja suurimman luokkanumeron välisille tyhjille (käyttämättä jääneille) luokkanumeroille omat taulukkorivinsä, jolloin tietokoneen muistitilaa saattaa tuhlautua aiheetta. Esim. vuosiluvut 1950, 1960 ja 1965 luokitustavan ainoana luokkaniminä on tällöin syytä korvata vaikkapa nimillä V1950, V1960 ja V1965, joita vastaavat luokkanumerot ovat 1,2 ja 3. Jos luokitustapaa käytettäisiin jollekin muuttujalle luokkanimien 1950, 1960 ja 1965 taulukoinnissa, varattaisiin tilaa 13 turhalle luokalle (luokat 1951-1959, 1961-1964).

TRANSF-käskyssä sen sijaan "vaihtelevien" luokkanumeroiden käytöstä ei ole lainkaan haittaa vaan päinvastoin usein hyötyä.

Pienin luokitettava arvo ( $=\min a_i$ ) ja suurin luokitettava arvo ( $=\max b_i$ ) rajoittavat luokitusvälin, jolla luokitettavan muuttujan arvot tavallisesti sijaitsevat. On hyvä pitää huolta joko loogisten muuttujien tai DEF-käskyn avulla, että muuttujan arvot lankeavat luokitusvälille. Luokitustapaa käytettäessä luokitusvälin ulkopuoliset arvot aiheuttavat virheilmoituksen ja ne sijoitetaan ylimääräiseen luokkaan, johon joutuminen tulkitaan luokitustapoja käytävissä käskyissä (TRANSF, TABLE) havainnon puuttumiseksi. (Kts. 3.4.)

Kaikille luokitusvälin arvoille ei ole pakko määritellä luokkaa. Vailla luokkaa olevista luokitusvälin arvoista ei seuraa virheilmoitusta, mutta ne tulkitaan kuten luokitusvälin ulkopuoliset arvotkin puuttuviksi.

Esim.3. Käsky

```
CLASS@ JAKO1
  A 5 6
  B 1 2
  C 7 7
  A 9 10
```

määrittelee luokitustavan JAKO1, jossa luokitusvälin alaraja on 1 ja yläraja 10. Luokitusvälin arvot luokitetaan seuraavasti:

arvo	luokka
1	B
2	B
3	- (ei luokkaa)
4	-
5	A
6	A
7	C
8	-
9	A
10	A

Jos luokkien määritykset peittävät toisiaan (tilanne, jota on syytä välttää), ristiriita selvitetään aina ensimmäisenä annetun määrityksen hyväksi.

Esim.4. Käskyllä

```
CLASS@ JAKO2
  1 1 3
  2 5 6
  3 2 5
```

luokitus tapahtuu seuraavasti:

arvo	luokka
1	1
2	1
3	1
4	3
5	2
6	2

CLASS-käskyyn kirjoitetulla M-parametrilla, jonka mahdolliset arvot ovat FAST ja SHORT, valitaan luokituksessa käytettävä menetelmä. Arvolla FAST käytetään nopeaa, mutta muistitilaa usein paljon kuluttavaa "epäsuorien osoitteiden" menetelmää. (CLASS-käsky rakentaa taulukon, josta oikea luokkanumero löytyy suoraan muuttujan arvon määräämästä paikasta.) Arvolla SHORT valitaan käytössä 3-10 kertaa hitaampi, mutta muistin käytön suhteen säästeliäs hakumenettely. (CLASS-käsky varastoi luokkanumerot ja -rajat sellaisenaan ja vertaa muuttujan arvoa luokkien määrittelyjärjestyksessä rajoihin, kunnes oikea luokka löydetään.) Jos M-parametria ei käytetä, valinta menetelmien FAST ja SHORT välillä tapahtuu automaattisesti.

### 2.7.2. TRANSF-käsky

TRANSF-käsky (TRANSFORM) on yleinen muuttujan muunnoskäsky, jolla muuttujan arvoksi asetetaan toisen muuttujan luokitustavalla luokitettu arvo (luokkanumero). Tarvittu luokitustapa on määriteltävä ennen TRANSF-käskyä CLASS-käskyllä. Käskyllä on rakenne

```
TRANSF@ u u1 c
      M: m
      IF:<ehto> ,
```

missä

u = uusi luokitettu muuttuja,  
u<sub>1</sub> = luokitettava muuttuja,  
c = CLASS-käskyllä määriteltä luokitustapa,  
m = puuttuvan (luokittamattoman) arvon muunnettu arvo

Muuttujan u arvo määräytyy kunkin IF-ehdon täyttävän havainnon kohdalla luokittamalla muuttujan u<sub>1</sub> arvo luokitustapaa c käyttäen ja sijoittamalla saatu luokkanumero muuttujan u arvoksi. Jos muuttujan u<sub>1</sub> arvoa ei saada luokitetuksi luokitustavalla c, muuttujan u arvoksi tulee M-parametrin arvo m; M-parametrin puuttuessa käskystä oletetaan, että m=0.

## 2.8. TAULUKOINTI

Yksi- ja kaksiulotteisten frekvenssijakautumien laskemiseen, "ristiintaulukointiin" ja näiden pohjalla tapahtuviin muihin tavanomaisiin taulukointitehtäviin ("prosenttitaulukot", ehdolliset keskiarvot ja -hajonnat) on käytettävissä TABLE-käsky. Ennen käskyn esittelyä on tarpeen sopia tärkeimmistä taulukointiin liittyvistä nimityksistä.

### 2.8.1. Käsitteitä ja nimityksiä

Taulukolla tarkoitetaan seuraavassa lukujen muodostamaa, suorakaiteen muotoista kaaviota, joka koostuu yhdestä tai useammasta näiden lukujen muodostamasta vaaka- ja pystyrivistä. Taulukko on kaksiulotteinen, jos vaaka- ja pystyrivejä on kumpiakin ainakin 2. Muussa tapauksessa taulukko on yksiulotteinen (vain 1 vaaka- tai pystyrivi). Taulukkoon kuuluvien lukujen paikkoja sanotaan ruuduiksi; ruutu tulee määrätyksi yksikäsitteisesti, jos tunnetaan, millä vaaka- ja millä pystyrivillä se sijaitsee.

Taulukoinnilla tarkoitetaan niitä toimenpiteitä, joilla rekisteröidään havaintovektoreiden vaikutus taulukon ruutujen sisältöön. Taulukoinnin alkaessa (SURVO-ohjelman alkaessa) taulukko on tyhjä; kaikki taulukon luvut ovat nolliä.

On mahdollista ajatella, että havaintovektorilla saattaisi olla vaikutusta useaan taulukon ruutuun samanaikaisesti. SURVO 66:ssa rajoitutaan kuitenkin taulukointiin, jolle on ominaista se, että yksi havaintovektori voi vaikuttaa korkeintaan yhteen taulukon ruutuun. Havainto siis kuuluu johonkin taulukon ruuduista tai ei kuulu niistä mihinkään.

Ruudun, johon havainto kuuluu, määrää yksi tai useampi muuttuja. Tällaista muuttujaa sanotaan luokittelumuuttujaksi. Luokittelumuuttuja, joka määrää vaakarivin, on rivimuuttuja ja luokittelumuuttuja, joka määrää pystyrivin, on sarakemuuttuja.

Muuttujaa, jonka arvo vaikuttaa luokittelumuuttujien määrämien taulukon ruudun sisältöön, sanotaan taulukoitavaksi muuttujaksi. Riippuen taulukoitavasta muuttujasta ja sen halutusta vaikutuksesta saadaan erityyppisiä taulukkoja.

Jos taulukoitavan muuttujan arvo lisätään aina ruudun entiseen arvoon, saadaan summataulukko. Taulukoitavan muuttujan ollessa vakio=1 summataulukko on frekvenssitaulukko, joka kertoo, montako havaintoa kuuluu kuhunkin taulukon ruutuun.



Lisättäessä taulukoitavan muuttujan arvon neliö aina ruudun entiseen arvoon saadaan neliösummataulukko. Samalla kertaa muodostettujen (tämä tehdään yhdellä TABLE-käskyllä) frekvenssi-, summa- ja neliösummataulukoiden perusteella saadaan helposti taulukoitavan muuttujan (ehdolliset) keskiarvo- ja keskihajontataulukot, so. taulukot, joiden ruuduissa ovat taulukoitavan muuttujan ehdolliset keskiarvot ja keskihajonnat.

Taulukkoon saattaa liittyä ylimääräisiä vaaka- ja pystyrivejä. Tällaisia ovat (taulukon ensimmäisinä) tunnusrivit, joilla nimetään vaaka- ja pystyrivit sekä (taulukon viimeisinä) marginaalit. Marginaalit ovat yksiulotteisia taulukkoja, joiden ruutuina ovat alkuperäisen taulukon ruudut vaakariveittäin ja pystyriveittäin yhdistettyinä. Havainto kuuluu siis oman taulukkoruutunsa ohella kahteen (yksiulotteisen taulukon tapauksessa yhteen) marginaaliruutuun. Taulukointi tapahtuu marginaaleissa samalla periaatteella kuin varsinaisessa taulukossa. Siten frekvenssitaulukon marginaalien luvut ovat vaaka- ja pystyrivien lukujen summia. Keskiarvotaulukoiden (keskihajontataulukoiden) kussakin marginaaliruudussa on taulukoitavan muuttujan keskiarvo (keskihajonta) laskettuna tähän marginaaliruutuun kuuluvista havainnoista.

Marginaalien yhtymäkohtana on yleisruutu, joka on kaikkien alkuperäisen taulukon ruutujen yhdelmä. Esim. frekvenssitauluko yleisruudussa on kokonaisfrekvenssi ja keskiarvotaulukon yleisruudussa taulukoitavan muuttujan yleiskeskiarvo.

Frekvenssitaulukon välittömiä johdannaisia ovat prosenttitaulukot. SURVO-järjestelmässä prosenttitaulukot saatetaan muodostaa kolmella eri tavalla muuntamalla frekvenssit prosentteiksi 1) vastaavasta vaakarivisummasta, 2) vastaavasta pystyrivisummasta, 3) kokonaisfrekvenssistä.



### 2.8.2. TABLE-käsky

TABLE-käskyllä muodostetaan yksi- ja kaksiulotteisia frekvenssi-, keskiarvo- ja keskihajontataulukkoita. Useampiulotteiset taulukot saadaan toistamalla TABLE-käskyä erilaisin IF-ehdoin.

TABLE-käskyillä muodostettuihin, nimettyihin taulukkoihin voidaan kohdistaa eräitä muita SURVO-käskyjä. CHI2-käsky laskee frekvenssitaulukon  $\chi^2$ -arvon ja VARAN-käsky tekee varianssi-analyysin taulukoitavasta muuttujasta keskiarvo- ja frekvenssitaulukon avulla.

TABLE-käsky noudattaa seuraavaa rakennetta:

```
TABLE@<sarakemuuttuja u> <luokitustapa c>
<taulukon nimi n1> <rivimuuttuja u1> <luokitustapa c1>
<taulukon nimi n2> <rivimuuttuja u2> <luokitustapa c2>
...
<taulukon nimi nr> <rivimuuttuja ur> <luokitustapa cr>
T:<taulukoitava muuttuja t>
M:<tulostuksen laajuus>
IF:<ehto>
```

Käsky suorittaa r taulukointitehtävää, joissa kaikissa on sama sarakemuuttuja u, mutta eri rivimuuttujat  $u_1, u_2, \dots, u_r$ . T- ja M-parametrit säätelevät myöhemmin kuvattavalla tavalla, mitkä taulukot näissä taulukoinneissa lasketaan. Ainoastaan IF-ehdon täyttävät havainnot taulukoidaan.

Kussakin taulukossa on CLASS-käskyllä määritellyn luokitustavan c mukainen määrä pystyrivejä ja i. taulukoinnissa ( $i=1,2,\dots,r$ ) rivimuuttujan  $u_i$  CLASS-käskyllä määritellyn luokitustavan  $c_i$  mukainen määrä vaakarivejä. Jokaista kunkin luokitustavan luokkanumeroa kohti varataan yksi taulukkorivi järjestyksessä pienimmästä luokkanumerosta suurimpaan. Huomattakoon erityisesti, että myös käyttämättömille pienimmän ja suurimman luokkanumeron välisille numeroille varataan omat rivinsä, jotka jäävät taulukoissa tyhjiksi. (Kts. 2.7.1.)

Esim.1. Käskyjen

```
CLASS@ KOOT
      PIENET  24  37
      KESKI  38  43
      SUURET  44  50
```

```
CLASS@ IAT
      NUORET   7  29
      VANHAT  30  39
```

```
TABLE@ KOKO  KOOT
      TAULU1  IKA  IAT
```

tuloksena saadaan TABLE-käskyn osalta frekvenssitaulukko, joka näyttää seuraavalta:

N= 307

TABLE:TAULU1

```
COLUMN VARIABLE:  KOKO  CLASSIFICATION:  KOOT
ROW VARIABLE:     IKA  CLASSIFICATION:   IAT
```

FREQUENCIES

	PIENET	KESKI	SUURET	TOTAL
NUORET	36	100	20	156
VANHAT	14	88	49	151
TOTAL	50	188	69	307

Muuttujia, joiden arvot ovat ei-negatiivisia kokonaislukuja (ja jotka on varastoitu perusskaalassa), saa käyttää luokittelumuuttujina ilman luokitustapaa. Tällöin luokitustavan paikalle on parametrilistaan kirjoitettava - . Jokaiselle luokittelumuuttujan mahdolliselle arvolle varataan silloin taulukoissa oma rivinsä, jonka tunnuksena on muuttujan arvo. Näitten luokittelumuuttujien ala- ja yläraja on ehdottomasti määriteltävä DEF-käskyllä ennen TABLE-käskyä, jotta tiedetään rivien oikea lukumäärä.

Esim.2. Olkoon muuttujan X1 arvot rajattu käskyllä

```
DEF@ X1 L:4 U:10 .
```

Tällöin taulukointikäsky, joka alkaa muodossa

```
TABLE@ X1 - (+ muita parametreja)
```

laskee taulukkoja, joiden kaikkien pystyrivit on varustettu tunnuksin 4,5,6,7,8,9,10.

Sarakemuuttujaa koskevat parametrit u ja c saa kokonaan jättää pois, jolloin muodostetaan yksiulotteisia taulukkoja erikseen kullakin rivimuuttujalla. Tällaisten taulukkojen ainoa pystyrivi jää vaille tunnusta.

Rivimuuttujia on sen sijaan oltava aina TABLE-käskyssä ainakin 1 (r2l).

Taulukon nimen asemasta parametrina voi olla pelkkä - , jolloin ko. rivimuuttujaa koskevat taulukot jäävät nimeämättä. Tällaisiin taulukkoihin ei ole mahdollista kohdistaa muita käskyjä.

Esim.3. Välillä 0:sta 9:ään vaihtelevien muuttujien X2, X3, X4 ja X5 ("suorat") frekvenssijakautumat saadaan käskyillä

```
DEF@ X2-X5 L:0 U:9
TABLE@ - X2 -
        - X3 -
        - X4 -
        - X5 - .
```

Jos luokittelumuuttujan luokka ei ole tietyillä muuttujan arvoilla määriteltä käytetyssä luokitustavassa, havainto jää taulukoimatta. "Puuttuvien" havaintojen lukumäärä näkyy tuloksesta vertaamalla taulukoiden kokonaisfrekvenssiä IF-ehdon täyttävien havaintojen määrään.

Esim.4.

```
CLASS@ ASENNE
      MYONT 1 2
      KIELT 4 5
EQUAL@ MIES X1 1
DEF@ X3 L:1 U:5
TABLE@ TAULU3 X3 ASENNE IF:MIES .
```

TABLE-käskyn tulostus näyttää seuraavalta:

```
IF: MIES N= 153
```

```
TABLE:TAULU3
```

```
ROW VARIABLE: X3 CLASSIFICATION:ASENNE
```

```
FREQUENCIES
```

```
MYONT      88
KIELT      45
TOTAL     133
```

Puuttuvia havaintoja (,joissa X3=3) on siis 153-133=20.

Jos M- ja T-parametreja ei käytetä TABLE-käskyssä, laske-  
taan pelkkiä frekvenssitaulukkoja.

Taulukoitava muuttuja t ilmoitetaan T-parametrilla ja täl-  
löin lasketaan muuttujan t keskiarvotaulukot.

M-parametrilla voidaan valita lisäksi seuraavat tulostukset:

M-parametri	tulostus
1) V	taulukoitavan muuttujan keskihajonnat
2) C%	frekvenssit prosentteina pystyrivisummista
3) C%V	1)+2)
4) R%	frekvenssit prosentteina vaakarivisummista
5) R%V	1)+4)
6) CR%	2)+4)
7) CR%V	1)+2)+4)
8) T%	frekvenssit prosentteina kokonaisfrekvenssistä
9) T%V	1)+8)
10) CT%	2)+8)
11) CT%V	1)+2)+8)
12) RT%	4)+8)
13) RT%V	1)+4)+8)
14) CRT%	2)+4)+8)
15) CRT%V	1)+2)+4)+8)

Jos T-parametri puuttuu TABLE-käskystä, kirjaimella V ei ole vaikutusta M-parametrissa.

Kaikki taulukot tulostetaan yhtenäisinä, ellei niiden leveys ylitä tiettyä rajaa. Tämä raja (63 merkkiä/rivi) on asetettu siten, että normaali teleprinterin telanleveys riittää ja vastaa sellaisen taulukon leveyttä, jossa on 7 pystyriviä. Mikäli raja ylitetään, TABLE-käsky painaa taulukot sopivasti ositetuina.

Rajaa voidaan muuttaa SPACES-käskyllä (kts. 2.5). Tällöin on syytä tietää, että kaikissa TABLE-käskyn painamissa taulukoissa yhden rivin leveys on 7 merkkiä. Esim. antamalla käsky SPACES@ 105 saadaan tulostetuiksi yhtenäisinä taulukot, joissa on korkeintaan 13 pystyriviä,  $(13+2) \cdot 7 = 105$ , sillä tilaa on varattava myös vaakarivien tunnuksille ja marginaalille. Tällöin on oltava käytössä teleprinteri, jonka telanleveys on ainakin 105 merkkiä.

## 2.9. TUNNUSLUKUJEN LASKEMINEN

Tunnuslukujen laskemiseen tarkoitettuja SURVO-käskyjä ovat:  
MEAN (keskiarvot),  
STDDEV (keskiarvot, keskihajonnat),  
CORREL (keskiarvot, keskihajonnat, kovarianssit, korrelaatiot)  
FRACT (mediaani, kvartiilit, desiilit, prosenttipisteet).  
Lisäksi huomattakoon mahdollisuus käyttää TABLE-käskyä muuttujan ehdollisten keskiarvojen ja keskihajontojen järjestelmälliseen taulukointiin.

### 2.9.1. Käskyt MEAN, STDDEV, CORREL

Näiden käskyjen yleinen rakenne on

```
<operaattori>@ u1 u2 ... ur  
IF:<ehto>  
N: <momenttitaulukon nimi>  
W: <painomuuttuja>  
M: <tulostuksen laajuus (vain CORREL)>  
T: <tulostuksen laajuus (vain CORREL)>
```

missä  $u_1, u_2, \dots, u_r$  ovat muuttujia. IF-ehdon täyttävistä havainnoista lasketaan operaattorin määräämät muuttujien  $u_1, u_2, \dots, u_r$  momentit. Momenttien laskemiseen tarvittavat summat säilytetään ja näiden summien muodostamiin momenttitaulukoihin on mahdollista viitata eräissä muissa käskyissä (REGRAN, TTEST), mikäli momenttitaulukko on nimetty N-parametrin avulla. M- ja T-parametreilla on merkitystä vain CORREL-käskyssä.

Tarvittaessa havainnot saadaan myös painotetuiksi käyttäen painomuuttujana W-parametrilla nimettyä muuttujaa. Tunnuslukuja laskettaessa painojen summa (painomuuttujan arvojen summa) normeerataan samaksi kuin havaintojen lukumäärä. W-parametrin puuttuessa kaikki havainnot ovat samanarvoisia (painomuuttujana vakio=1).

MEAN-käsky laskee pelkät keskiarvot. STDDEV-käsky laskee keskiarvojen lisäksi keskihajonnat.

CORREL-käsky laskee keskiarvojen ja keskihajontojen lisäksi muuttujien  $u_1, u_2, \dots, u_r$  korrelaatiomatriisin sekä vielä M-parametrin arvolla COVCOR kovarianssimatriisin. M-parametrin ollessa COV korrelaatiomatriisi jää pois, mutta kovarianssimatriisi tulostetaan. T-parametrilla J\*CODE korrelaatiomatriisi saadaan tulostetuksi reikänauhalle matriisimuotoisen tulostuksen lisäksi ns. J\*-koodissa, jossa muodossa se kelpaa kaikille korrelaatioesitykseen perustuville ELLIOTT 803:n tilastomaattisille kirjasto-ohjelmille. M-parametrin ollessa COV



T-parametrilla ei ole vaikutusta.

Korrelaatio- ja kovarianssimatriisien taulukkomuotoiseen tulostukseen pätee samat säännöt kuin TABLE-käskyllä tulostettuihin taulukkoihin.

Muodostettaessa momenttitaulukoita jatkotoimenpiteitä varten on tärkeätä tietää, että TTEST-käsky (t-testi) soveltuu vain käskyillä STDDEV ja CORREL muodostettuihin taulukkoihin ja REGRAN-käsky (regressioanalyysi) pelkästään CORREL-käskyn tuloksiin.

Esim.

```
NAME@ TAUSTA X3-X8 X12 X15-X20
STDDEV@ TAUSTA IF:MIES N:MIEHET
      @ TAUSTA IF:NAINEN N:NAISET
TTEST@ MIEHET NAISET
(Kts. TTEST).
```

### 2.9.2. FRACT-käsky

FRACT-käskyn tehtävänä on muuttujan fraktiilien (mediaanin, kvartiilien, desiilien ja prosenttipisteiden) laskeminen. Ennen FRACT-käskyn kirjoittamista SURVO-ohjelmaan on TABLE-käskyllä määriteltävä taulukko (frekvenssitaulukko), jossa tutkittava muuttuja on rivimuuttujana. Viittaamalla FRACT-käskyssä pelkästään tähän taulukkoon fraktiilit lasketaan rivimuuttujan marginaalijakautumasta. Taulukon sarakemuuttujalla on vaikutusta vain siinä tapauksessa, että joitain havaintoja ei voida luokitella sarakemuuttujan suhteen, jolloin havainto jää pois taulukosta. FRACT-käskyn yleinen rakenne on

```
FRACT@ <taulukon nimi> q r s ,
```

missä q, r ja s ovat ei-negatiivisia kokonaislukuja ja määrittelevät käskyn toimintalaajuuden.

Käskyllä saadaan määrätyksi mitkä tahansa taulukon rivimuuttujan prosenttipisteistä

$$P_0, P_1, P_2, \dots, P_{99} ,$$

missä  $P_i$  on se muuttujan arvo (viivallisesti interpoloituina), jonka alittaa i prosenttia taulukoiduista havaintoarvoista.

Esim.  $P_{50}$ =mediaani,  $P_{25}$ =alakvartiili ja  $P_{75}$ =yläkvartiili. Prosenttipisteistä lasketaan ja tulostetaan apuparametrien q, r ja s säätämänä

$$P_q, P_{q+r}, P_{q+2r}, \dots, P_s ,$$

jolloin vaaditaan, että  $0 \leq q \leq s \leq 99$  ja s on muotoa  $s = q+kr$ , missä k on ei-negatiivinen kokonaisluku. Esim. valitsemalla  $q=25$ ,  $r=25$  ja  $s=75$  lasketaan alakvartiili, mediaani

ja yläkvartiili tai valitsemalla  $q=50$ ,  $r=0$  ja  $s=50$  lasketaan vain mediaani. Valinnalla  $q=0$ ,  $r=1$  ja  $s=99$  saadaan kaikki prosenttipisteet.

FRACT-käskyn tulostuksessa käytetään käskyn kohteena olevan taulukon vaakarivien luokkanumeroiden asteikkoa olettaen, että luokkanumerot ovat luokkavälien alarajoja. (Kts. CLASS- ja TABLE-käsky.)

Pyrittäessä muuttujan alkuperäisen asteikon mukaisiin tuloksiin joudutaan FRACT-käskyn tulostamissa fraktiileissa tekemään seuraavanlaisia korjauksia.

1) Jos muuttujan arvot sellaisenaan ovat luokkanumeroita (luokitustapaa ei siis ole käytetty) ja luokkanumerot ovat luokkakeskusten arvoja, tulostetuista fraktiilien arvoista on vähennettävä 0.5 .

Yleinen sääntö on:

2) Jos muuttujan arvot ovat luokitettuja, menetellään seuraavasti. Olkoon FRACT-käskyn tulostama fraktiilin arvo  $f = c + e$  , missä  $c$  on luvun  $f$  kokonaisosa ja  $e$  sen desimaaliosa. Katsotaan luokitustavan määrittelystä, mitä voidaan pitää luokkanumeroa  $c$  vastaavan luokkavälin ala- ja ylärajana. Olkoon alaraja  $a$  ja ylärajana  $b$ . Lopullinen fraktiilin arvo on tällöin  $f' = a + e(b-a)$  .

Esim.

```
CLASS@ C      S:1
      I      10.5  16
      II     16.1  18
      III    18.1  20
TABLE@ Y/C    Y  C  M:C%
FRACT@ Y/C    50  0  50
```

Näillä käskyillä lasketaan skaalassa 1 varastoidun muuttujan Y mediaani jakautumasta Y/C, jossa Y on luokitettu luokitustapaa C käyttäen. FRACT-käskyn tulos voi näyttää seuraavalta:

```
FRACTILES OF      Y IN TABLE      Y/C
```

RESULTS ARE GIVEN IN TERMS OF CLASS VALUES

```
FRACT%      VALUE
      50      2.4246575
```

Mediaaniluokan numero on siis 2 (luokka II). Tulkittaessa tämän luokan alarajaksi 16.05 ja ylärajaksi 18.05 saadaan medianin lopulliseksi arvoksi  $16.05 + 0.425(18.05 - 16.05) = 16.90$  .

## 2.10. REGRESSIO- JA VARIANSSIANALYYSI

### 2.10.1. REGRAN-käsky

Käsky REGRAN laskee viivallisen regressiomallin

$$y = a_0 + a_1x_1 + a_2x_2 + \dots + a_rx_r$$

parametrit pienimmän neliösumman keinolla. Käskyn tuloksina saadaan paitsi regressiokertoimet myös niiden keskihajonnat, kertoimien ja hajontojen suhteet, jäännösvarianssi ja yhteiskorrelaatiokerroin.

REGRAN-käsky ei kohdistu suoraan muuttujiin vaan CORREL-käskyllä muodostettuun korrelaatiomatriisiin (, joka sisältää myös keskiarvot ja -hajonnat). Tähän korrelaatiomatriisiin tulee kuulua jokaisen malliin sisältyvän muuttujan (myös selitettävän muuttujan y) ja sillä tulee olla CORREL-käskyn N-parametrilla annettu nimi.

Korrelaatiomatriisi saa sisältää muitakin kuin malliin kuuluvia muuttujia. On täten mahdollista kohdistaa REGRAN-käsky toistuvasti samaan korrelaatiomatriisiin ja laskea useita regressiomalleja selitettävää ja selittäjiä vaihdellen. REGRAN-käskyn yleinen rakenne on

REGRAN@ <korrelaatiomatriisin nimi>

y  
x<sub>1</sub> x<sub>2</sub> ... x<sub>r</sub> .

Koska REGRAN-käsky toimii valmiin korrelaatiomatriisin pohjalta ja tämä voi olla ehdollisesti muodostettu, REGRAN-käsky ei saa sisältää IF-parametria.

### 2.10.2. VARAN-käsky

Käsky VARAN laskee yksi- tai kaksisuuntaisen varianssianalyysin tulokset TABLE-käskyllä saadusta taulukoitavan muuttujan (T-parametri TABLE-käskyssä) keskiarvotaulukosta. Käskyn rakenne on pelkästään

VARAN@ <taulukon nimi> .

Koska VARAN-käsky kohdistuu valmiiseen keskiarvotaulukkoon, joka voi olla ehdollisesti muodostettu, VARAN-käsky ei saa sisältää enää IF-parametria.

Varianssianalyysin suoritustavan (tavallinen yksi- tai kaksisuuntainen varianssianalyysi) määrää VARAN-käskyn kohteena olevan taulukon ulotteisuusluku. Analyysi kohdistuu TABLE-käskyn T-parametrilla ilmaistuun taulukoitavaan muuttujaan, jota

tarkastellaan luokittelumuuttujien eri tasoilla. Merkittäköön taulukoitavaa muuttujaa seuraavassa  $x$ :llä, taulukon rivimuuttujaa  $A$ :lla (vaakarivejä olkoon  $a$ ) ja sarakemuuttujaa  $B$ :llä (pystyivejä olkoon  $b$ ). Jos sarakemuuttuja  $B$  puuttuu (eli taulukko on yksiulotteinen), tehdään muuttujan  $x$  yksiulotteinen varianssianalyysi rivimuuttujan  $A$  suhteen. Kaksiulotteisen taulukon tapauksessa myös varianssianalyysi toteutetaan kaksisuuntaisena.

Yksiulotteisen varianssianalyysin suoritukselle ei ole rajoituksia. Keskiarvotaulukkoa vastaavassa frekvenssitaulukossa saa olla vaihteleva määrä havaintoja. Jos jokin frekvensseistä on nolla, vastaavaa  $A$ -muuttujan tasoa ei varianssianalyysissä oteta lainkaan huomioon.

Kaksisuuntaisen varianssianalyysin yhteydessä frekvenssitaulukon odotetaan olevan tasaisesti (tai ainakin lähes tasaisesti) täytetty. Analyysin tulos on tarkalleen oikea vain tasaisesti täytetyn taulukon tapauksessa.

Epätasaisesti täytetyn taulukon ollessa kysymyksessä käytetään karkeata likimääräismenettelyä, jossa frekvenssitaulukko "tasoitetaan" korvaamalla kunkin ruudun frekvenssi ( $n_{ij}$ ) keskimääräisellä frekvenssillä

$$\bar{n} = \frac{1}{ab} \sum_{i=1}^a \sum_{j=1}^b n_{ij}$$

ja summataulukon lukuja ( $s_{ij} = n_{ij} \bar{x}_{ij}$ , missä  $\bar{x}_{ij}$  on vastaava keskiarvotaulukon luku) "korjataan" vastaavasti kertomalla ne luvulla  $\bar{n}/n_{ij}$ . Mikäli ruudussa ei ole yhtään havaintoa ( $n_{ij}=0$ ), asetetaan

$$s_{ij} = \frac{as_{.j} + bs_{i.} - s_{..}}{(a-1)(b-1)},$$

missä  $s_{i.} = \sum_{j=1}^b s_{ij}$ ,  $s_{.j} = \sum_{i=1}^a s_{ij}$  ja  $s_{..} = \sum_{i=1}^a \sum_{j=1}^b s_{ij}$ .

(kts. Cochran, Cox: Experimental Designs, ss. 110-112).

Jos kuitenkin kokonainen rivi tai sarake on tyhjä, VARAN-käskyä ei toteuteta.

Epätasaisen taulukon tapauksessa ja erikseen joidenkin ruutujen ollessa tyhjiä varianssianalyysin tuloksiin ei pidä liikaa luottaa, koska käytetty menettely ei ole "oikea".

VARAN-käskyn tulostus on normaalin varianssianalyysin tuloskaavion mukainen.

## 2.11. TESTIT

### 2.11.1. CHI2-käsky

Käskyllä

CHI2@ <frekvenssitaulukon nimi>

lasketaan frekvenssitaulukon  $\chi^2$ -arvo. Käsky kohdistuu aina frekvenssitaulukkoon, vaikka samassa taulukoinnissa olisi muodostettu muitakin taulukkoja.

CHI2-käsky tulostaa paitsi  $\chi^2$ -arvon ja vapausasteet myös pienimmän odotetuista frekvensseistä =  $\min_{i,j} n_{i,j}/n$ , missä  $n_{i,j}$  on i.rivin marginaalifrekvenssi,  $n_{.j}$  on j.sarakkeen marginaalifrekvenssi ja n on kokonaisfrekvenssi.

Jos taulukossa on tyhjiä rivejä ja/tai sarakkeita, nämä poistetaan ennen  $\chi^2$ -arvon laskemista.

### 2.11.2. TTEST-käsky

TTEST-käsky vertailee kahden havaintoryhmän vastinmuuttujien keskiarvoja t-testin avulla. Käsky kohdistetaan ao. ryhmiin ja muuttujiin STDDEV- tai CORREL-käskyllä muodostettujen momenttitaulukoiden kautta. TTEST-käskyn voi kirjoittaa jommassakummassa seuraavista muodoista:

- 1) TTEST@ <taulukko 1> <taulukko 2> ,
- 2) TTEST@ <taulukko 1> <muuttuja 1>  
<taulukko 2> <muuttuja 2> .

Tapauksessa 1) on vaatimuksena, että molemmissa momenttitaulukoissa on samat (tai ainakin toisiaan vastaavat) muuttujat samassa järjestyksessä ja vertailu tapahtuu erikseen kaikilla vastinmuuttujapareilla. Tapauksessa 2) taulukoiden samanmuotoisuutta ei vaadita, mutta käsky vertaa tällöin keskenään vain parametreina mainittuja kahta muuttujaa, joista muuttujan 1 tulee olla taulukossa 1 ja muuttujan 2 vastaavasti taulukossa 2.

TTEST-käsky tulostaa kummastakin vertailtavasta ryhmästä erikseen havaintojen lukumäärän, muuttujien keskiarvot ja hajonnat sekä yhteisesti muuttujien keskiarvojen eroa kuvastavat t-arvot vapausasteineen.

Edellytyksenä TTEST-käskyn järkevälle käytölle on, että vertailtavissa ryhmissä ei ole yhteisiä havaintoja. Ryhmien tulee olla toisistaan riippumattomia.



### 3. SURVO-OHJELMAN KÄYTTÖ

SURVO 66-järjestelmän perustana on S-ohjelma (systeemiohjelma), joka normaalikäytössä on jaettu suoritusvaiheita T1, T2 ja T3 vastaten kolmeen osaan, käännösohjelma (SURVO 66 TAPE 1), ajo-ohjelma (SURVO 66 TAPE 2) ja tulostusohjelma (SURVO 66 TAPE 3). S-ohjelma on laadittu käyttäen Elliott 803- autokoodia ja -konekoodia.

Käännösohjelmalla muunnetaan reikänauhalle lävistetty SURVO-ohjelma ajo-ohjelman tarvitsemaan muotoon käännetyksi ohjelmaksi (TRANSLATED PROGRAM), jonka kone käännöksen jälkeen lävistää reikänauhalle. Käännetyt ohjelman ajaminen tapahtuu ajo-ohjelman valvonnassa, joka lukee aluksi käännetyt ohjelman ja tämän jälkeen reikäkortteille tai reikänauhalle lävistetyt havaintomatriisin. Lopuksi ajo-ohjelma tulostaa käännetyt ohjelman täydennettynä havaintomatriisista kerätyillä, SURVO-ope-raatioiden tarvitsemilla tiedoilla (TRANSLATED PROGRAM + DATA n , missä n = luettujen havaintovektoreiden lukumäärä). SURVO-ohjelman käyttö päättyy ajoon tulostusohjelman valvonnassa, joka lukee aluksi ajo-ohjelman tulostamat tiedot (TRANSLATED PROGRAM + DATA n ) ja antaa lopulliset tulokset.

#### 3.1. OHJELMAN LÄVISTÄMINEN

SURVO-ohjelma lävistetään reikänauhalle täsmälleen siinä muodossa, jossa se on kirjoitettu. Lävistyksessä on huomattava erityisesti seuraavat seikat:

1) Operaattorin (käskyn nimen) perässä oleva @ merkki on lävistettävä välittömästi operaattorin jälkeen ilman välilyöntiä. Vastaava sääntö koskee täsmennyssanan päätteenä olevaa : merkkiä ja muuttujaluettelon keskellä olevaa - merkkiä.

2) Parametreina olevat sanat ja luvut on erotettava toisistaan yhdellä tai useammalla välilyönnillä tai rivinsiirroilla.

3) Parametrisanassa tai luvussa oleva ? merkki poistaa saman sanan tai luvun tätä edellä olevien merkkien vaikutuksen.

4) Suluilla ( ) ja niiden välissä olevilla merkeillä ei ole vaikutusta parametrilistassa.

5) Jos operaattorin paikalle kirjoitetaan pelkkä @ merkki, tämä tarkoittaa samaa kuin edellisen SURVO-käskyn operaattori.

### 3.2. HAVAIMTOMATRIISIN LÄVISTÄMINEN

Havaintomatriisi voidaan lävistää joko reikäkorteille tai reikänauhalle.

Reikäkorttien sisäänluvussa on käytössä Elliott 803:n CINPUT-KINPUT-järjestelmä, jonka käytöstä on laadittu erilliselostuksia. Havaintomatriisi ja sen alussa olevat muuttujien sijaintia ja lävistystapaa korteilla ilmaisevat parametrikortit on lävistettävä tämän järjestelmän vaatimusten mukaisesti. Huomattakoon, että CINPUT-KINPUT-järjestelmä sallii kaikki numeeriset lävistystavat, joissa yhtä useampaa muuttujaa ei ole lävistetty samaan sarakkeeseen ja joissa muuttujat voivat olla usealla peräkkäisellä kortilla. Havaintojen lukumäärää ei voida ilmoittaa parametrikorteilla, vaan se tehdään joko N-käskyn avulla (kts. 2.5) tai havaintomatriisikorttien perään asetetulla loppukortilla, jonka ensimmäinen sarake on täyteen lävistetty.

Reikänauhalle havaintomatriisi lävistetään seuraavassa muodossa:

$$\begin{array}{l}
= \langle \text{havaintomatriisin nimi} \rangle \quad b1 \\
\langle \text{muuttujien lukumäärä } m \rangle \\
\langle \text{havaintojen lukumäärä } n \rangle \\
\begin{array}{llll}
x_{11} & x_{21} & x_{31} & \cdots & x_{m1} & \sum_i x_{i1} & (1. \text{ havaintovektori}) \\
x_{12} & x_{22} & x_{32} & \cdots & x_{m2} & \sum_i x_{i2} & (2. \text{ havaintovektori}) \\
\cdots & & & & & \cdots & \cdots \\
x_{1n} & x_{2n} & x_{3n} & \cdots & x_{mn} & \sum_i x_{in} & (n. \text{ havaintovektori})
\end{array}
\end{array}$$

Luvut on erotettava toisistaan yhdellä tai useammalla välilyönillä tai rivinsiirroilla.

Havaintomatriisi saa kummassakin lävistystavassa sisältää kokonaislukujen ohella myös desimaalilukuja.

#### 3.2.1. Desimaaliluvut reikäkorttilävistyksessä

Muuttujat, joiden arvot ovat desimaalilukuja, lävistetään kokonailukuina jättämällä desimaalipiste pois. Tällaisen muuttujan kortilta varaaman kentän viimeiset sarakkeet vastaavat desimaaleja. DEF-käskyn S-parametrin avulla on ilmoitettava muuttujan skaala eli monenko desimaalin tarkkudella muuttuja on lävistetty. Muuttujat, joihin ei ole kohdistettu mainitunlaista DEF-käskyä, käsitellään kokonaislukuarvoisina perusskaalassa.

### 3.2.2. Desimaaliluvut reikänauhalävistyksessä

Desimaaliluvut lävistetään tavalliseen tapaan käyttäen desimaalipistettä. Havaintomatriisia luettaessa muuttujien arvot muunnetaan siihen skaalaan, jonka SURVO-ohjelma kullekin muuttujalle määrää.

### 3.3. OHJELMAN KÄÄNTÄMINEN (T1-VAIHE)

SURVO-ohjelman kääntäminen tapahtuu käännösohjelman (SURVO 66 TAPE 1) avulla seuraavasti:

T1.1. Aloituksella 40 0 luetaan käännösohjelma SURVO 66 TAPE 1 koneen muistiin.

T1.2. Asetetaan reikänauhalle lävistetty SURVO-ohjelma lukijaan ja annetaan aloitus 40 16, jolloin kääntäminen alkaa. Kun SURVO-ohjelma on saatu luetuksi, käännösohjelma tulostaa (kanava 1) käännetyn SURVO-ohjelman (TRANSLATED PROGRAM) reikänauhalle.

Vaihe T1.2 voidaan toistaa ja kääntää useita SURVO-ohjelmia peräkkäin.

Käännösvaiheen aikana tulostetaan (kanava 1) myös luokitus-tapojen määrittelyt ja sen päätyttyä numeeristen ja loogisten muuttujien luettelo.

Jos käännettävässä SURVO-ohjelmassa on virheitä, tulostetaan (kanava 3) kunkin havaitun virheen jälkeen sen laatua ilmaiseva virhenumero ( ERROR <virhenumero> ) ja virheellinen käsky. Kohdasta 3.6 ilmenevät eri virhenumeroiden merkitykset. Virheilmoituksen jälkeen käännösohjelma jää Elliott-autokoodijärjestelmän WAIT-käskyyn. Uusien virheiden löytämiseksi voidaan yrittää jatkaa suoraan WAIT-käskystä muuttamalla numerogeneraattorin viimeisen N2-napin asento. Käännös päättyy virheellisen SURVO-ohjelman tapauksessa WAIT-käskyyn, josta jatkamalla saadaan käännetty ohjelma tulostetuksi. Tällaisella käännetyllä ohjelmalla ei kuitenkaan yleensä ole merkitystä, vaan alkuperäinen SURVO-ohjelma on korjattava ja käännettävä uudelleen alusta.

### 3.4. OHJELMAN AJAMINEN (T2-VAIHE)

Käännösohjelmalla käännetyn SURVO-ohjelman ajaminen tapahtuu ajo-ohjelman (SURVO 66 TAPE 2) valvonnassa seuraavasti:

Jos havaintomatriisi on reikänauhalla, B-napin on oltava alaspainettuna koko T2-vaiheen ajan.

- T2.1. Aloituksella 40 0 luetaan ajo-ohjelma SURVO 66 TAPE 2 koneen muistiin.
- T2.2. Aloituksella 40 16 luetaan käännetty SURVO-ohjelma (TRANSLATED PROGRAM). Ajo-ohjelma jää lukemisen jälkeen kaksiaäniseen (oktaavi-intervalli) WAIT-käskyyn (DATA WAIT). Käännetyn ohjelman sisäänluku on summatarkistettu.
- T2.3. Havaintomatriisin lukeminen alkaa jatkamalla DATA WAIT-käskystä (muuttamalla numerogeneraattorin viimeisen N2-napin asento) ja SURVO-ohjelman T2-vaihe, havaintovektoreitten käsittely toteutetaan.  
Tämä vaihe päättyy, kun N-käskyn mukainen määrä havaintoja on luettu ja käsitelty (tai havaintomatriisin ollessa reikäkorteilla loppukortti on luettu) kaksiaäniseen (kvintti-intervalli) WAIT-käskyyn (RESULTS WAIT).  
Jos havaintomatriisi on reikänauhalla, N-käskyllä ei ole vaikutusta, vaan luettavien havaintojen määrä saadaan havaintomatriisin alusta.
- T2.4. Jatkettaessa RESULTS WAIT-käskystä tulostetaan käännetty SURVO-ohjelma täydennettynä havaintomatriisista kerätyillä tiedoilla (TRANSLATED PROGRAM + DATA n , missä n on luettujen havaintojen lukumäärä) reikänauhalle (kanava 1).

Toistamalla vaiheet T2.2 - T2.4 voidaan ajaa useita SURVO-ohjelmia peräkkäin.

Vaiheen T2.3 päätyttyä voidaan jatkaa ilmoitettu havaintojen määrä ylittäen havaintovektoreiden lukemista aloituksella 40 16:00 4 ( 40 16/00 4 havaintomatriisin ollessa reikänauhalla).

Poistamalla havaintomatriisi lukijasta vaihe T2.3 saadaan keskeytetyksi ja antamalla tämän jälkeen aloitus 40 16:00 5 siirrytään vaiheeseen T2.4 käyttäen havaintojen lukumääränä tähän asti luettujen havaintojen määrää.

Vaiheessa T2.4 saatua nauhaa (TRANSLATED PROGRAM + DATA n) saatetaan käyttää paitsi tulostusohjelmassa (kts. 3.5) myös uudelleen ajo-ohjelmassa, mikäli halutaan kerätä lisää havain-toja aikaisempien täydennykseksi. Tällöin käydään läpi vaiheet T2.1 - T2.4 kuten edellä sillä erotuksella, että vaiheessa T2.2 luetaan käännetty SURVO-ohjelma aikaisemmin havainnoin täydennetyssä muodossa (TRANSLATED PROGRAM + DATA n). Tätä mahdollisuutta kannattaa käyttää tilanteessa, jossa havain-toja on jouduttu hylkäämään esim. lävistysvirheistä johtuen (ehdol-liset STOP-käskyt SURVO-ohjelmassa, kts. 2.5 ja 3.4.1). Hylä-tyt havaintovektorit voidaan tällöin korjata ja ottaa myöhem-min mukaan ilman, että koko havaintomatriisi jouduttaisiin lu-kemaan uudelleen. Tämä toimenpide ei kuitenkaan ole yleensä silloin mahdollinen, jos SURVO-ohjelmassa käytetään ORDER- ja LAG-käskyjä, joiden toiminta riippuu havaintovektoreiden kes-kinäisestä järjestyksestä.

#### 3.4.1. Ajon keskeytyminen T2-vaiheessa

Havaintomatriisin käsittely voi keskeytyä useasta eri syys-tä.

Lävistysvirheet. Havaintomatriisin ollessa reikänauhalla kiellettyjä tietonauhan merkkejä koskevat virheilmoitukset ovat Elliott 803-autokoodijärjestelmän mukaiset. Havaintovek-toreiden tarkastussummavirheistä ilmoitetaan tulostamalla (ka-nava 3) virheen ollessa suurempi kuin 0.00001

SUM ERROR <havainnon järjestysnumero> <summavirhe>

ja ajo-ohjelma jää WAIT-käskyyn, josta normaalisti jatkamalla havaintovektori hyväksytään summavirheestä huolimatta. Havain-tovektorin sisäänluku uusitaan aloituksella 40 16/00 4 .

Reikäkorteilla olevan havaintomatriisin sisältämistä vir-heistä annettavat ilmoitukset ovat CINPUT-KINPUT-järjestelmän mukaiset. Havaintovektoreiden sisäänluku uusitaan aloituksella 40 16:00 4 .



Muuttujien rajojen ylitykset. Jos muuttujan arvo rikkoo DEF-käskyllä ilmoitettuja rajoja, seuraa virheilmoitus (kanava 3).

<havainnon järjestysnumero><muuttujan nimi>=<muuttujan arvo>  
ERROR 32

ja ohjelma jää WAIT-käskyyn, josta jatkettaessa normaalisti virheellinen arvo muutetaan lähimpään sallittuun arvoon. Virheilmoitusta ei kuitenkaan anneta (, mutta korjaus tehdään), jos F2=20 ohjauspöydässä.

Ellei muuttujan arvo lankea käytetyn luokitustavan luokitusvälille (kts. 2.7.1), seuraa virheilmoitus (kanava 3)

<havainnon järjestysnumero><luokitustapa><muuttujan arvo>  
ERROR 40

ja ohjelma jää WAIT-käskyyn, josta jatkettaessa normaalisti havaintoarvo jää luokittamatta. Tämä virheilmoitus jätetään antamatta, jos F2=02 ohjauspöydässä.

STOP- ja WAIT-käskyt. Toteutettu STOP-käsky aiheuttaa ilmoituksen (kanava 3)

STOP <havainnon järjestysnumero> <ehto, jolla STOP-käsky on toteutunut> ,

jonka jälkeen SURVO-ohjelma jatkuu normaalisti uuden havaintovektorin lukemisella.

SURVO-ohjelman WAIT-käskyn toteutuessa ajo-ohjelma jää normaaliin WAIT-käskyyn, josta jatkettaessa ohjelman toteutus jatkuu normaalisti.

Käskyt DIV, SQRT ja LOG antavat niille sopimattomien operandin arvojen tapauksessa omat virheilmoituksensa (kts. 2.6.2) ja ohjelma jää WAIT-käskyyn.

### 3.5. TULOSTUS (T3-VAIHE)

Ajo-ohjelmalla käsitellyn SURVO-ohjelman lopulliset tulokset saadaan tulostusohjelman (SURVO 66 TAPE 3) avulla seuraavasti:

T3.1. Aloituksella 40 0 luetaan tulostusohjelma SURVO 66 TAPE 3 koneen muistiin.

T3.2. Aloituksella 40 16 luetaan ajo-ohjelman tulosnauha TRANSLATED PROGRAM + DATA n , jolloin T3-vaihe, viimeistely- ja tulostusvaihe toteutetaan. Tulostusohjelman ajo päättyy ilmoitukseen (kanava 3) END OF PROGRAM .

Toistamalla vaihe T3.2 voidaan saada usean SURVO-ohjelman tulokset peräkkäin.

Ennen kunkin SURVO-käskyn omaa osuutta tulostetaan (kanava 1) SURVO-ohjelman nimi ja käskyn suorituksen edellytyksenä ollut ehto muodossa IF:<ehto> . Ohjelman nimeä ei kuitenkaan tulosteta joka käskyn kohdalla, jos F2=40 ohjauspöydässä.

Toteutumassa olevan SURVO-käskyn tulostus voidaan (esim. reikänauhan loppuessa) uusia aloituksella 40 16:00 11 . Jos SURVO-käskyn tulostus jostain syystä jää kesken tai keskeytetään, seuraavan käskyn tulostukseen päästään aloituksella 40 16:00 6 .

### 3.6. VIRHEILMOITUKSET

S-ohjelma ja SURVO-käskyt ilmoittavat SURVO-ohjelmaa käännettäessä ja ajettaessa havaituista virheistä tulostamalla (kanava 3)

ERROR <virhenumero> .

Käännösvaiheessa tulostetaan tämän jälkeen vielä mahdollisen virheen sisältänyt SURVO-käsky. Joissakin virheilmoituksissa virhenumeroa saattaa edeltää muita virheen laatua tarkentavia tietoja.

Virheilmoituksen jälkeen ohjelma jää aina WAIT-käskyyn. Virheen luonteesta riippuu, jatketaanko ajoa vai joudutaanko se keskeyttämään. Käännösvaiheessa virheellinen käsky yleensä siivutetaan ja jää siis kääntämättä, mistä jatkettaessa voi koitua uusia, aiheettomia virheilmoituksia.

Alla on lueteltu kaikki käytössä olevat virhenumerot ja niiden merkitykset.

- 1 a)liian suuri muuttujien lukumäärä,  
b)M-käsky puuttuu,  
c)liian suuri indeksi X-nimessä
- 2 käskystä puuttuu parametreja
- 3 vakio muuntyyppisen parametrin tilalla
- 4 tuntematon tai kielletty parametri
- 5 a)sama nimi jo aikaisemmin,  
b)vakio nimen asemasta
- 6 parametrien lukumäärä väärä
- 7 a)kielletty luokan numero,  
b)luokkanimet eivät kaikki samaa tyyppiä
- 8 vakion tilalla nimi tai kielletty sana
- 9 luokan alaraja suurempi kuin yläraja
- 12 skaalausvirhe, liian suuri skaalaustekijä
- 13 a)liian pitkä merkkijono parametrilistana (yli 600 merkkiä),  
b)liian pitkä sana (yli 50 merkkiä)
- 14 @ merkki liikaa?
- 15 määrittelemätön muuttuja
- 16 NAME-käskyillä nimettyjen merkkijonojen tila täynnä
- 17 tuntematon operaattori
- 18 käskyllä liikaa parametreja (yli 100)
- 19 tuntematon täsmennyssana
- 20 virhe muuttujaluettelon lyhennysmerkinnässä

- 21 virhe IF-parametrissa
- 22 muuttujien lukumäärä ilmoitettu ohjelmassa eri tavalla kuin tietonauhalla/tietokorteilla
- 23 parametrin tyyppi väärä
- 24 täsmennysparametrin tyyppi väärä
- 25 skaalausvirhe
- 26 työtila täynnä. Jatkettaessa työtilan ylittäneen käskyn tilalle asetetaan END-käsky ja kääntäminen loppuu. "Supistetulla" ohjelmalla voi ajaa.
- 27 muuttujille ja vakioille varattu tila täynnä. Jatko sama kuin virheen 26 jälkeen.
- 28 käsky ei ole mukana käytetyssä SURVO-versiossa
- 29 nolla jakajana jakolaskussa (tulos=0)
- 30 negatiivinen juurrettava (tulos=0)
- 31 nolla tai negatiivinen luku logaritmoitavana (tulos=0)
- 32 muuttujan arvo asetettujen rajojen ulkopuolella. Jatkettaessa muutetaan lähimpään sallittuun arvoon.
- 33 luokitustavan skaala ei ole sama kuin muuttujan skaala
- 34 luokitustapoja liikaa
- 35 taulukkonimiä liikaa (nimi jää antamatta)
- 36 ylivuoto laskutoimituksessa
- 37 summavirhe ohjelmanauhan sisäänluvussa (laskettu summa tulostettu). Jatkettaessa toimitaan kuin ei summavirhettä olisikaan.
- 38 regressiomallissa muuttujia, jotka eivät sisälly korrelaatiomatriisiin
- 39 momenttimatriisi singulaarinen
- 40 arvo, jota ei voida luokitella. Jatkettaessa katsotaan tuntemattomaksi.
- 41 sulkujen sisällä oleva selitys jatkuu uudelle riville. Jatkettaessa etsitään edelleen loppusulkua.
- 42 taulukossa korkeintaan yksi "ei-tyhjä" vaaka- tai pystyrivi,  $\chi^2$ -arvo mahdoton laskea
- 43 a) skaalattu muuttuja luokittelumuuttujana ilman luokitus-  
tapaa,  
b) ilman luokitustapaa oleva luokittelumuuttuja, jonka ala- ja/tai ylärajaa ei ole määritelty DEF-käskyllä. Jatkettaessa L:0 U:9.
- 44 vaaka- tai pystyrivi tyhjä varianssianalyysissä
- 45 t-testissä yhteensopimattomat momenttitaulukot

### 3.7. MUISTITILAN KÄYTTÖ

SURVO-järjestelmä tarvitsee muistista tilaa S-ohjelman lisäksi SURVO-ohjelman käskyille, käskyjen työtiloille ja niiden käyttämille SURVO-objekteille.

Käskyjen tilankäytöstä (ja myös ajankäytöstä) annetaan tietoja kustakin käskystä tehdyssä erilliselostuksessa. Kaikkiin on SURVO-ohjelman käytössä noin 3000 muistipaikkaa. Ohjelman tarvitsema tila saadaan tietää käännosvaiheen lopussa, jolloin tulostetaan SURVO-ohjelmalta käyttämättä jääneiden muistipaikkojen lukumäärä (FREE STORE: ). Esim. johdantoluvussa oleva ohjelma tarvitsee  $2500 - 2137 = 323$  muistipaikkaa (ohjelma on ajettu S-ohjelmalla, jossa SURVO-ohjelman käytössä on 2500 muistipaikkaa).

SURVO-ohjelman tilankäytöstä riippumatta siinä esiintyvien SURVO-objektien lukumäärille on seuraavat rajoitukset:

numeerisia ja loogisia muuttujia sekä (muunnoskäskyjen) vakioita korkeintaan	100	200
luokitustapoja korkeintaan	30	50
nimettyjä taulukoita korkeintaan	30	100
NAME-käskyllä nimettyjä merkkijonoja korkeintaan	10	20

Elint 503