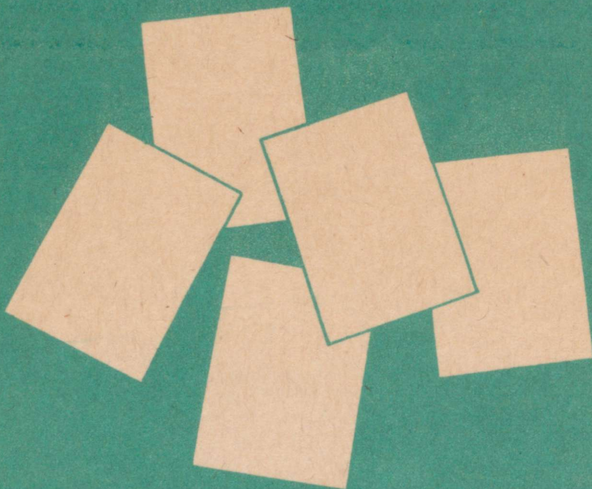

PROGRAMME KÕIGILE

I

TARTU
1971

TARTU RIIKLIK ÜLIKOOL
Arvutuskeskus

PROGRAMME KÕIGILE

I

Metoodiline materjal

Teine trükk

Tartu 1971

ПРОГРАММЫ ДЛЯ ВСЕХ
I
Методический материал
Издание второе
На эстонском языке
Тартуский государственный университет
ЭССР, г. Тарту, ул. Юликооли, 18

Vastutav toimetaja J. Torokoff
Korrektor R. Palm

TRÜ rotaprint 1971. Paljundamisele antud 18. III 1971.
Trükipoognaid 2,75. Tingtrükipoognaid 2,5. Arvestuspoognaid 2,46. Trükiarv 500. Paber 30 x 42. 1/4.
Tell. nr. 206.

Hind 10 kop.

Kogumikus "Programme kõigile" kavatseb TRÜ Arvutuskeskus avaldada niisuguste programmide lühikirjeldusi, mida saab laialdaselt kasutada TRÜ teiste teaduskondade ning laboratooriumite uurimistöös. Senised kogemused on nimelt näidanud, et paljud TRÜ töötajad ei kasuta elektronarvuti abi oma teaduslikus töös just seetõttu,et nad ei ole tuttavad TRÜ Arvutuskeskuses olemasolevate programmide võimalustega ja nende kasutamisega.

Käesolevasse esimesse kogumikku on koondatud siiani kõige laialdasemat kasutamist leidnud programmid matemaatilise planeerimise, statistika ja ankeetide töötlemise valdkonnast. Nende programmide autoriteks on Ülo Kaasik, Maie Viitso (matemaatiline planeerimine); Ain Iher, Airi Laumets Sven Veldre, Tiina Veldre (statistika) ja Ants Laumets (ankeetide töötlemine).

TRANSPORDIÜLESANDE LAHENDAMISE PROGRAMM

Transpordiülesandeks nimetatakse erikujulist lineaarset planeerimisülesannet: leida mittenegatiivsed suurused x_{ij} ($i = 1, 2, \ldots, m$; $j = 1, 2, \ldots, n$) nii, et oleks rahuldatud kitsendused

$$\sum_{j=1}^{n} x_{ij} = a_i \quad (\text{iga } i = 1,2,\ldots,m \text{ korral})$$

$$\sum_{i=1}^{m} x_{ij} = b_j \quad (\text{iga } j = 1,2,\ldots,n \text{ korral})$$

ja avaldis

$$\sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} x_{ij}$$

omandaks minimaalse väärtuse (vt. Ü.Kaasik, "Matemaatiline planeerimine", Tallinn, 1967, III ptk.).

Vaadeldavas programmis eeldatakse, et etteantud suurused a_i, b_j ja c_{ij} on kõik positiivsed täisarvud, ning mitte suuremad kui 99999. Ühtlasi nõutakse muidugi, et oleks täidetud tingimus

$$\sum_{i=1}^{m} a_i = \sum_{j=1}^{n} b_j \ .$$

Ladude arv m ja tarbijate arv n peavad rahuldama tingimusi:

$$m \leq 100; \quad n \leq 39; \quad m + n \leq 128; \quad n \cdot m \leq 2822 \ .$$

Seega suuremate ülesannete korral tuleb "tarbijateks" nimetada need punktid, mille arv on väiksem.

Algandmed tuleb perforeerimiseks esitada järgmise tabeli kujul:

Tarbijad / Laod	1	2	...	n	Varud
1	c_{11}	c_{12}	...	c_{1n}	a_1
2	c_{21}	c_{22}	...	c_{2n}	a_2
...	...	...	...	...	...
m	c_{m1}	c_{m2}	...	c_{mn}	a_m
Vajadused	b_1	b_2	...	b_n	$\sum a_i = \sum b_j$

SIMPLEKSMEETODI PROGRAMMID

Majandustegevuse planeerimisel ja juhtimise organiseerimisel tekkivad probleemid võib matemaatiliselt sõnastada lineaarsete planeerimisülesannetena[1]: leida lineaarvõrrandisüsteemi

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= b_2 \\ \dots\dots\dots\dots\dots\dots\dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n &= b_m \end{aligned}$$

mittenegatiivsete lahendite hulgast see, mis muudab minimaalseks lineaarse funktsiooni

$$c_1x_1 + c_2x_2 + \dots + c_nx_n ,$$

kus m on võrrandisüsteemi võrrandite arv ja n on tundmatute arv.

Kõiki sellisel kujul esitatavaid ülesandeid on põhimõtteliselt võimalik lahendada simpleksmeetodi abil. Selleks on arvutile "Ural-4" koostatud kaks simpleksmeetodi programmi[2].

---

[1] Neist vt. Ü.Kaasik, Lineaarsed planeerimisülesanded. Matemaatika ja kaasaeg, 2, Tartu, 1964, lk. 31 - 46.

[2] Nende programmide kasutamise täielik juhend on avaldatud väljaandes "Труды Вычислительного центра", 10, Тарту, 1967. (vene keeles).

Neist üks kasutab vaid arvuti sisemälu, võimaldades väiksemaid ülesandeid lahendada suhteliselt kiiresti. Nimelt peab ülesanne sel juhul rahuldama järgmisi kitsendusi:

$$(n + 3)(m + 2) \leq 1784,$$
$$n \leq 140 .$$

Teine programm kasutab peale sisemälu ka magnettrumleid, mistõttu on võimalik lahendada ülesandeid, mille mõõtmed rahuldavad tingimusi

$$(n + 1)(m + 2) \leq t \cdot 16384,$$
$$m \leq 450 ,$$

kus t on kasutatavate trumlite arv (TRÜ Arvutuskeskuses maksimaalselt t = 6). Selle programmi kasutamise puhul on ülesande lahendamise aeg võrdlemisi pikk (näit. ülesande puhul, kus n = 100 ja m = 100, umbes 3 tundi); lahendusaja pikkuse ja ka välisseadmete väikse töökindluse tõttu pole praktiliselt lahendatud ülesandeid, mille võrrandisüsteemi maatriksielementide arv on suurem kui 40000.

VÄHIMRUUTUDE MEETODI PROGRAMM

Antud programmi võib kasutada juhul, kui on tarvis leida katsetulemuste põhjal mingi suuruse (funktsiooni) A ja parameetrite x_j ($j = 1, 2, \ldots, n$) vaheline lineaarne seos.

Kui m katset on teostatud selliselt, et igas katses i on antud parameetritele x_j erinevaid väärtusi x_{ij}, ning alati on fikseeritud katsetulemused A_i ($i = 1, \ldots, m$), siis eeldades, et seosed suuruse A ja parameetrite x_j vahel on lineaarsed, saame järgmise võrrandite süsteemi:

$$\left.\begin{array}{l} A_1 = a_1x_{11} + a_2x_{12} + \ldots + a_jx_{1j} + \ldots + a_nx_{1n} \\ A_2 = a_1x_{21} + a_2x_{22} + \ldots + a_jx_{2j} + \ldots + a_nx_{2n} \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ A_i = a_1x_{i1} + a_2x_{i2} + \ldots + a_jx_{ij} + \ldots + a_nx_{in} \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ A_m = a_1x_{m1} + a_2x_{m2} + \ldots + a_mx_{mj} + \ldots + a_nx_{mn} \end{array}\right\} \quad (1)$$

Kui katsevigade piirides ühesuguste parameetrite x_j puhul on teostatud mitu katset (s.t. katset on korratud samade tingimuste juures), siis võib iga võrrandit võrrandisüsteemis (1) asendada võrranditega kujul

$$\bar{A}_o \pm \sigma_i = a_1\bar{x}_{i1} + a_2\bar{x}_{i2} + \ldots + a_n\bar{x}_{in} \, ,$$

kus $\bar{A}_i$ ja $\bar{x}_{ij}$ on vastavalt suuruste A_i ja parameetrite x_{ij}

keskväärtusteks samadel tingimustel korratud katsete juures (s.t. i-nda katseseeria puhul). Sel juhul on otstarbekas arvutada suuruse A_i ruutkeskmine viga σ_i.

Järgnev ülesanne seisneks kordajate (a_1, a_2, ..., a_n) määramises selliselt, et nad võimalikult hästi rahuldaksid kõiki võrrandeid süsteemis (1) ja seega uuritavat nähtust saaks iseloomustada ühe üldise lineaarse seosega kujul

$$A = a_1x_1 + a_2x_2 + \ldots + a_nx_n \ .$$

Antud programmi abil võib leida võrrandisüsteemi (1) ligikaudse lahendi (s.t. määrata kordajad a_1, a_2, ..., a_n) selliselt, et lahend (kordajad) muudab minimaalseks hälvete $\sum_{j=1}^{n} a_j x_{ij} - A_i$ ruutude summa, s.t. $\sum_{i=1}^{m}(\sum_{j=1}^{n} a_j x_{ij} - A_i)^2$ on minimaalne. (Sel nõudel baseeruvat lahendusmeetodit nimetatakse vähimruutude meetodiks).

Antud programmi abil on võimalik arvutada järgmisi suurusi:

1) kordajate a_j väärtused (väljundamisel tähistatakse X[M]);
2) kordajate a_j ruutkeskmised vead (tähistatakse V[X[M]]);
3) parameetrite x_j aritmeetilised keskmised (tähistatakse B[M]);
4) parameetrite x_j aritmeetiliste keskmiste hälbed (tähistatakse V[B[M]]);
5) hälbed $\sum_{j=1}^{n} a_j x_{ij} - A_i$ (tähistatakse E[K]);
6) mitmene korrelatsioonikordaja (tähistatakse KK);

7) (Kui peetakse vajalikuks)

parameetrite x_{ij} vahelised korrelatsioonikordajad. Väljundamisel trükitakse korrelatsioonimaatriks (tähistatakse KO[K, M]).

Algandmete hulk on piiratud järgmiselt:

1) parameetrite x_j arv $n \leq 32$;

2) võrrandite arv (s.t. teostatud katsete arv) $m \leq 653$;

3) alati peab olema $n \leq m$.

Arvutusteks vajalikud algandmed tuleb esitada järgmise vormi kohaselt:

ülesande nr. (suvaline kuuekohaline numbrite või vene tähestiku tähtede kombinatsioon).

m =

n =

i	x_{i1}	x_{i2}	....	x_{in}	A_i	σ_i
1						
2						
3						
⋮	⋮	⋮		⋮	⋮	⋮
m						

OBJEKTIDE RÜHMITUSTE LEIDMINE OBJEKTIDEVAHELISTE SARNASUSKORDAJATE MAATRIKSIST

Antud programmiga saab leida objektide rühmitusi, kui objektidevaheline sarnasus on väljendatud sarnasuskordajaga (suuremale sarnasusele vastab suurem kordaja väärtus; identsete objektide puhul on kordaja väärtuseks 1,0000). Üheks taoliseks sarnasuskordajaks on Q-tehnika lineaarne korrelatsioonikordaja r. Kõikide objektide omavahelised sarnasused olgu esitatud nelinurkse sarnasuskordajate maatriksina. Maatriks arvutatakse eelnevalt magnettrumlile salvestava programmiga või viiakse arvutisse perfosisendilt. Maatriksi elemendid (kahe objekti sarnasuskordaja väärtused) peavad väljenduma arvudena piirides +,9999 kuni -,9999.

Programm moodustab korraga ühe objektide rühma. Iga objekt võib kuuluda mitmesse rühma, kuid võib ainult ühele rühmale olla algobjektiks. Rühmitamise aluseks on suurimate sarnasuskordajate rida, milles igat objekti esindab sarnasuskordaja antud objekti ja talle kõige sarnasema objekti vahel. Rühma alustamiseks leitakse seni rühma mittealustanud objektide seast kaks sellist objekti, millel on ülalmainitud reas kõige suuremad ühised suurimad sarnasuskordajad, ja lõpetame, kui antud rühma koosseisu arvatud objektide rühmasiseste sarnasuskordajate aritmeetiline keskmine saab väiksemaks suurimate sarnasuskordajate rea vähimast liikmest. Rühmade leidmise lõpetame, kui seni rühma mittealustanud objektide hulgas pole enam ühise suurima sarnasusega paari.

Piiramised algandmete osas on järgmised:

1) Objektide arv $N \leq 181$; tunnuste arv M on piiratud sarnasuskordajate maatriksi arvutamise programmi nõuetega (vt. lk. 19).

2) Kui sarnasuskordajate maatriksit eelnevalt ei arvutata, vaid see viiakse arvutisse abiprogrammiga, siis olgu $N \leq 181$ ja maatriksi diagonaalist üleval pool asuv kolmnurk tuleb perforeerimiseks esitada järgmiselt:

$$
\begin{array}{cccc}
r_{1,2} & r_{1,3} & \cdots & r_{1,N} \\
 & r_{2,3} & \cdots & r_{2,N} \\
 & & \cdots\cdots\cdots & \\
 & & & r_{N1,N}
\end{array}
$$

Ülesande esitamisel tuleb ette anda:

1) maatriksi järk N;

2) millise programmiga on tarvis maatriks arvutada.

OBJEKTIDE RÜHMITUSTE LEIDMINE OBJEKTIDEVAHELISTE KAUGUSTE MAATRIKSIST

Antud programmiga saab leida objektide rühmitusi, kui objektidevaheline sarnasus on väljendatud kaugustena (s.o. mida sarnasemad on objektid, seda väiksem on nendevaheline kaugus; identsete objektidevaheline kaugus võrdub nulliga). Üheks taoliseks sarnasuse karakteristikuks on näiteks eukleidiline kaugus M-mõõtmelises ruumis (kus M on tunnuste arv). Kõikide objektide omavahelised kaugused olgu esitatud nelinurkse maatriksina. Maatriks arvutatakse eelnevalt magnettrumlile salvestatava programmiga või viiakse arvutisse perfosisendilt. Maatriksi elemendid (kahe objekti vahelised kaugused) peavad väljenduma arvudena +0000 kuni +9999 suvalise (maatriksi piirides konstantse) järguga.

Programm moodustab korraga ühe objektide rühma. Iga objekt võib kuuluda vaid ühte rühma, seega uutesse rühmadesse valitakse vaid vabu objekte. Rühmitamise aluseks on vähimate kauguste rida, milles iga objekti esindab kaugus sellest objektist kuni kõige sarnasema objektini. Rühma (esimese või uue) alustamiseks leitakse sellised vabad objektid, millel on ülalmainitud reas kõige väiksemad ühised vähimad kaugused. Rühma järgmiseks potentsiaalseks objektiks valitakse selline vaba objekt, millel on vähim kauguste summa antud rühma juba valitud objektide suhtes. Seejärel arvutame rühma rühmaväliste kauguste (kaugused ülejäänud objektideni) aritmeetilise keskmise oodatava juurdekasvu ja antud rühma rühmasiseste (rühma objektide omavahelised kaugused) kauguste aritmeetilise keskmise oodatava juurdekasvu jagatis K. Kui jagatis K on suurem etteantud väärtusest (milleks tavaliselt valitakse konstandid 0,0000 või 1,0000), siis potentsiaalne element võetakse rühma liikmeks ja asutakse leidma rühma uut potentsiaalset liiget. Kui jagatis K on väiksem etteantud väärtusest, siis potentsiaalset elementi rühmaga ei liideta, rühma moodustamine loetakse lõpetatuks ja hakatakse koostama uut rühma. Rühmade leidmine lõpetatakse, kui vabade objektide hulgas pole enam ühise vähima kaugusega objektide paari.

Konkreetse rühma leidmise käigus trükitakse rühma koosseis, kusjuures iga rühmaga liidetud objekti kohta lisatakse järgmised andmed:

1) rühmasiseste kauguste aritmeetiline keskmine peale antud objekti liitmist rühmaga;
2) liidetud objekti number;
3) jagatis K objekti rühmaga liitmise momendil;
4) rühma antud koosseisu juures vähimat rühmasiseste kauguste summat omava objekti number.

Samad andmed esitatakse ka viimase potentsionaalse objekti kohta, mis rühma lõpetamisel jäi liitmata. Välja trükitakse ka kõikide rühmadevaheliste kauguste summad (summa ühe rühma kõikide objektide kaugustest teise rühma kõikide objektideni).

Algandmed on piiratud järgmiselt:

1) objektide arv $N \leq 256$; tunnuste arv M on piiratud kauguste maatriksi arvutamise programmi nõuetega (vt. lk. 19).

2) kui kauguste maatriksit eelnevalt ei arvutata, vaid see viiakse arvutisse abiprogrammiga, siis olgu $N \leq 181$ ja maatriksi diagonaalist ülevalpool asuv kolmnurk tuleb perforeerimiseks esitada järgmiselt:

$$
\begin{array}{cccc}
d_{1,2} & d_{1,3} & \cdots & d_{1,N} \\
 & d_{2,3} & \cdots & d_{1,N} \\
 & & \cdots\cdots\cdots & \\
 & & & d_{N-1,N}
\end{array}
$$

Ülesande esitamisel tuleb ette anda:

1) maatriksi järk (objektide arv) N ;

2) millise programmiga on tarvis maatriks arvutada.

SPEARMAN'i ASTAKKORRELATSIOONIKORDAJATE MAATRIKSI ARVUTAMISE PROGRAMM

Kui uuritavad tunnused on seotud mittelineaarselt, või kui tunnus(ed) ei väljendu arvudena, vaid tunnuse intensiivsuse astmetena (astakud), siis pole seoste ja mõjutuste statistiliseks määramiseks võimalik kasutada lineaarset korrelatsioonikordajat (r). Kui numbriliselt või astakutena väljendatud tunnused muutuvad teineteise suhtes lineaarselt, või mittelineaarselt, kuid monotoonselt, siis võib rakendada Spearman'i astakkorrelatsioonikordajat

$$\rho_{jk} = 1 - \frac{6\sum_{i=1}^{N} d_{ijk}^2}{N^3 - N},$$

kus N - vaatluspaaride arv;

d_{ijk} - i objekti j-nda ja k-nda tunnuse astakute erinevus.

See valem on kasutatav, kui ühegi tunnuse piirides ei esine samaväärseid objekte. Kui tunnustes esineb võrdse astakuga objekte, tuleb kasutada keerulisemat valemit (programmeerimiseks sobivale kujule teisendatult):

$$\rho_{jk} = \frac{A_j + A_k - 12\sum_{i=1}^{N} d_{ijk}^2}{2\sqrt{A_j A_k}},$$

kus A_j ja A_k leitakse valemiga

$$A = N^3 - N - \sum_{\sum t=N} (t^3 - t) ,$$

t - sama astakuga objektide kordus antud (j-nda või k-nda) tunnuse juures;

d_{ijk} - i objekti j-nda ja k-nda tunnuse astakute erinevus.

Programm võimaldab arvutada kuni 128. järku korrelatsioonikordajate maatriksit (režiim A) või arvutada selliste maatriksite rea, kasutades algandmete koguhulgast vaid N objekti iga maatriksi leidmisel (režiim B): esimese maatriksi arvutamisel objektid 1 kuni N, teise maatriksi arvutamisel objektid 1+k kuni N+k, kolmanda maatriksi arvutamisel objektid 1+2k kuni N+2k jne. (režiim B).

Laitrükkali abil trükitakse välja maatriksi alumine kolmnurk (25-veeruliste ribadena). Spearman'i astakkorrelatsioonikordajad trükitakse kordaja märgi ja nelja komajärgse kohaga. Kui mingi tunnus oli konstantne (selline olukord võib esineda režiim B puhul), siis trükitakse sellest tunnusest sõltuvate kordajate asemel sümbolid ****.

Algandmete hulk on piiratud järgmiselt:

Režiim A: Tunnuste arv $M \leq 128$,
objektide arv N peab olema selline, et
$M \times N \leq 3304$.

Režiim B: Tunnuste arv $M \leq 128$,
ühe maatriksi arvutamisel kasutatavate objektide arv N peab olema selline, et $N \times M \leq 256$,

objektide koguarv N_1 peab olema selline, et

$M \times N_1 \leq 3040.$

Algandmeteks võivad olla arvud vahemikus +0000 kuni +9999. Koma asukoht sama tunnuse piirides on konstantne, eri tunnustel suvaline.

Algandmed esitada tabelina:

Arvutamise režiim ja parameetrid

režiim A puhul: M ja N,

režiim B puhul: M, N, N_1 ja k.

Objekti nr. \ Tunnuse nr.	1	2	3 ...		... M
1	xxxx	xxxx	xxxx	...	xxxx
2	xxxx	xxxx	xxxx	...	xxxx
...	....	....	....	...	....
N (režiim A) või N_1(režiim B)	xxxx	xxxx	xxxx	...	xxxx

MITMEMÕÕTMELISE NORMEERITUD RUUMI EUKLEIDILISTE KAUGUSTE JA R- JA Q-TEHNIKA LINEAARSETE KORRELATSIOONIKORDAJATE MAATRIKSITE ARVUTAMISE PROGRAMM

Algandmed normeeritakse tunnuste kaupa ($j = 1,2,3,\ldots,M$) üle kõikide objektide (N objekti), kasutades valemit

$$x'_{ij} = \frac{x_{ij} - \bar{x}_j}{S_{x_j}} ,$$

kus

$$\bar{x}_j = \frac{\sum_{i=1}^{N} x_{ij}}{N} , \quad S_{x_j} = \sqrt{\frac{\sum_{i=1}^{N} (x_{ij} - \bar{x}_j)^2}{N - 1}} .$$

Programm võimaldab normeerimise ajal leida ka veel

variatsioonikordajaid $c_j = \dfrac{S_{x_j}}{\bar{x}}$,

aritmeetiliste keskmiste standardhälbeid $S_{\bar{x}_j} = \dfrac{S_{x_j}}{\sqrt{N}}$,

aritmeetiliste keskmiste 95% usalduspiire

$U_j^{(95)} = t_{95,\, N-1} \cdot S_{\bar{x}_j}$

ja ka 99% usalduspiire $U_j^{(99)} = t_{95,\, N-1} \cdot S_{\bar{x}_j}$.

Tunnuste j ja k vaheline lineaarne (R-tehnika) korrelatsioonikordaja arvutatakse valemiga

$$r_{jk} = \frac{\sum_{i=1}^{N} x'_{ij} \cdot x'_{ik}}{N - 1} .$$

Objektide k ja ℓ vaheline eukleidiline kaugus M-mõõtmelises ruumis leitakse valemiga

$$\frac{d}{M} = \sqrt{\frac{\sum_{i=1}^{M} (x'_{ik} - x'_{i\ell})^2}{M}} .$$

Q-tehnika lineaarsete korrelatsioonikordajate (r^{*}) saamiseks teostatakse veel teistkordne normeerimine objektide kaupa (j = 1,2,3,...,N) üle kõigi M mõõtme

$$x''_{ij} = \frac{x'_{ij} - \bar{x}'_j}{S_{x'_j}} ;$$

$$\bar{x}'_j = \frac{\sum_{i=1}^{M} x'_{ij}}{M} ; \quad S_{x'_j} = \sqrt{\frac{\sum_{i=1}^{M} (x'_{ij} - \bar{x}'_j)^2}{M - 1}} .$$

r^{*}_{jk} arvutatakse valemiga

$$r^{*}_{jk} = \frac{\sum_{i=1}^{M} x''_{ij} \cdot x''_{ik}}{M - 1}$$

Lisaks põhivariantidele r, r^{*}, $\frac{d}{10M}$ on võimalik leida veel erikujulisi variante: $|r|$, r^2, sign $r(r^2)$, $\frac{d^2}{100M}$, $|r^{*}|$, $(r^{*})^2$, sign $r^{*}((r^{*})^2)$. Soovi korral võib kõiki neid variante teisendada valemiga $D' = 0{,}5 - 0{,}5D$. Teisendatud kuju on sobiv edasiseks töötlemiseks rühmitamisprogrammidega.

Arvutatava maatriksi alumine kolmnurk trükitakse laitrükkaliga välja 25-veeruliste ribadena. Korrelatsioonikordajad ja kaugused trükitakse märgi ja nelja õige tüvekohaga. Olenevalt programmi iseärasustest trükitakse allpool diagonaali asuvate korrelatsioonikordajate puhul +1,0000 asemel -,0000, ja -1,0000 asemel +,0000 (magnettrumlile salvestatakse õiged väärtused). Kuna +,0000 ja -,0000 võivad tekkida ka loomulikul teel (korrelatsiooni puudumisel), tuleb nende juhtude eristamiseks võrrelda vastavate tunnuste või objektide korrelatsioone ülejäänud tunnuste või objektidega. Segaduse ärahoidmiseks on parem samas ülesandes mitte kasutada identseid tunnuseid või objekte. Ka pole soovitav töödelda väga ebaühtlase jaotusega tunnuseid, sest normeerimisel võrdsustatakse arvu $\mp 9,999$ absoluutselt ületavad normeeritud hälbed selle väärtusega. Ülisuurte hälvete asendamine väärtusega $\mp 9,999$ toob kaasa korrelatsioonikordajate väärtuste muutumise, sest diagonaalil asuv kordaja ei saavuta enam väärtust ,9999.

Soovi korral on programmi abil võimalik välja trükkida algandmeid, ühekordselt normeeritud algandmeid, kahekordselt normeeritud algandmeid - märgi ja nelja õige tüvekohaga. Märgi ja kaheksa tüvekohaga (neist kuus õiget) on võimalik välja trükkida tunnuste parameetreid $\bar{x}$, s, $s_{\bar{x}}$, c ja aritmeetilise keskmise usalduspiire $U^{(95)}$, $U^{(99)}$.

Edasiseks töötlemiseks on võimalik salvestada arvutatavat maatriksit nelinurksel kujul magnettrumli suvalisse ossa, kusjuures maatriksi diagonaalile kantakse soovitud konstant.

Kuna programm koosneb standardsetest osadest, on neid võimalik kasutada ka muu informatsiooni töötlemiseks:

a) neljakohaliste kümnendarvude (kaks arvu pikas pesas) teisendamiseks kahendsüsteemi ja viimiseks magnettrumlile (massiividena, mille pikkus ei ületa 183 perfokaarti);

b) magnettrumlil asuva informatsiooni normeerimiseks rida või veergu pidi;

c) magnettrumlil asuva informatsiooni (lühikese pesa fikseeritud komaga kahendsüsteemi arvud) väljatrükkimiseks nelinurkse maatriksina (trükitakse 25-veerulised ribad; igal arvul on märk ja neli kümnendkohta);

d) magnettrumlil asuva nelinurkse maatriksi transponeerimiseks magnettrumli mingisse teise ossa.

Ülaltoodud operatsioonide korral võib informatsioon paikneda magnettrumli suvalises kohas ja hõivata ühe või mitu trumlit.

Algandmete hulk on piiratud järgmiselt:

1) algandmete väljatrükkimine pole kitsendatud;

2) maatriksi transponeerimisel ja normeerimisel ei tohi veergude või ridade arv olla suurem kui 1770;

3) korrelatsiooni- ja kaugustemaatriksite arvutamisel on soovitav, et r arvutamisel $26N \leq 3546$ ja $M \leq 1770$, ning r^* või d arvutamisel $26M \leq 3546$ ja $N \leq 1770$. Ülesannet on võimalik lahendada ka siis, kui N või M on suurem kui $3546 : 26 \approx 136$, kuid siis toimub maatriksi

arvutamine ja trükkimine kitsamate kui 25-veeruliste ribadena, kusjuures arvutamiskiirus väheneb. Võimalik maksimaalne veergude arv trükitavas ribas w_{max} ja objektide arv (r) või tunnuste arv (d ja r^{*}) J on seotud järgmiselt /$(w + 1) \cdot J \leq 3546$/:

J	≤	136	141	147	154	161	168	177	186	197	208	221
w_{max}	=	25	24	23	22	21	20	19	19	18	17	16	15
J	≤	236	253	272	295	322	354	394	443	506	591	
w_{max}	=	14	13	12	11	10	9	8	7	6	5	

Algandmed peavad olema täielikud, s.t. igal objektil olgu kõikide tunnuste mõõteandmed.

Algandmeteks võivad olla arvud vahemikus -9999 kuni +9999 (välja arvatud -0000); koma asukoht eri tunnustel on suvaline ja samal tunnusel konstantne. Arv -0000 ei tohi algandmena esineda; seda kasutatakse algandmete massiivi täiendamiseks, et tunnuse või objekti andmete lõpp langeks kokku perfokaardi lõpuga. Algandmed ei tohi sisaldada konstantseid tunnuseid (s = 0). Kui mingi tunnuse keskmine võib osutuda täpselt nulliks, siis tuleb see asjaolu operaatori töö kergendamiseks eraldi ära märkida.

Algandmed esitada tabelina:

1) ülesande parameetrid M ja N;
2) milliste näitajate maatrikseid on tarvis arvutada ja välja trükkida;
3) millised maatriksid tuleb salvestada magnettrumlil, milline konstant viia diagonaalile,

millise programmiga töödelda salvestatud maatrikseid;

4) millised lisaandmed välja trükkida;

5) kas algandmeid tuleb perforeerida tunnuse või objekti kaupa (olenevalt sellest, kas tunnuste või objektide hulka muudetakse järgnevate ülesannete lahendamisel).

Algandmete tabel (objekti kaupa perforeerimiseks) on järgmine:

Tunnuse nr. \ Objekti nr.	1	2 ...		... M
1	xxxx	xxxx	. . .	xxxx
2	xxxx	xxxx	. . .	xxxx
...	....	....	. . .	....
M	xxxx	xxxx		xxxx

Algandmete perforeerimisel tunnuse kaupa tuleb tunnused paigutada tabeli veergudena.

PROGRAMM LINEAARSETE KORRELATSIOONIKORDAJATE JA MÕNEDE STATISTILISTE PARAMEETRITE ARVUTAMISEKS LÜNKLIKU ALGANDMESTIKU KORRAL

Lünkliku algandmestiku korral (s.o. juhul kui mõnel objektil pole iga tunnus mõõdetud) võib osutuda, et erinevate tunnuste paari (i,j) juures on erinev ka nende objektide arv (N_{ij}), millistel mõlemad tunnused on samaaegselt mõõdetud. Allpool antud valemite järgi summeerides jäetakse N objektist välja need, millistel tunnused i ja j ei ole samaaegselt mõõdetud.

Iga tunnuste paari i ja j ($j \leq i \leq M$; M-tunnuste arv) jaoks arvutatakse järgmised statistilised parameetrid

aritmeetilised keskmised

$$\bar{x}_{i\cdot j} = \frac{\sum\limits_{N_{ij}} x_i}{N_{ij}}\ , \quad x_{j\cdot i} = \frac{\sum\limits_{N_{ij}} x_j}{N_{ij}}\ ;$$

standardhälbed

$$S_{x_{i\cdot j}} = \frac{\sqrt{\sum\limits_{N_{ij}} (x_i - \bar{x}_{i\cdot j})^2}}{\sqrt{N_{ij} - 1}}\ , \quad S_{x_{j\cdot i}} = \frac{\sqrt{\sum\limits_{N_{ij}} (x_j - \bar{x}_{j\cdot i})^2}}{\sqrt{N_{ij} - 1}}\ ;$$

korrelatsioonikordaja

$$r_{ij} = \frac{\sum\limits_{N_{ij}} (x_i - \bar{x}_{i\cdot j})(x_j - \bar{x}_{j\cdot i})}{\sqrt{\sum\limits_{N_{ij}} (x_i - \bar{x}_{i\cdot j})^2 \sum\limits_{N_{ij}} (x_j - \bar{x}_{j\cdot i})^2}} ;$$

variatsioonikordajad

$$c_{i\cdot j} = \frac{s_{x_{i\cdot j}}}{\bar{x}_{i\cdot j}} , \quad c_{j\cdot i} = \frac{s_{x_{j\cdot i}}}{\bar{x}_{j\cdot i}} ;$$

aritmeetiliste keskmiste standardhälbed

$$S_{\bar{x}_{i\cdot j}} = \frac{s_{x_{i\cdot j}}}{\sqrt{N_{ij}}} , \quad S_{\bar{x}_{j\cdot i}} = \frac{s_{x_{j\cdot i}}}{\sqrt{N_{ij}}} ;$$

aritmeetiliste keskmiste 95% ja 99% usalduspiirid

$$U^{(95)}_{i\cdot j} = t_{95;N_{ij}-1} \cdot S_{\bar{x}_{i\cdot j}} , \quad U^{(99)}_{i\cdot j} = t_{95;N_{ij}-1} \cdot S_{\bar{x}_{i\cdot j}} ,$$

$$U^{(95)}_{j\cdot i} = t_{95;N_{ij}-1} \cdot S_{\bar{x}_{j\cdot i}} , \quad U^{(99)}_{j\cdot i} = t_{95;N_{ij}-1} \cdot S_{\bar{x}_{j\cdot i}} ,$$

kus t_{95} ja t_{99} on Student'i t-jaotuse vastavad kvantiilid.

Arvutatud tulemustest trükitakse alati välja r_{ij} ja N_{ij}, soovi korral ka $\bar{x}$, s_x, c, $S_{\bar{x}}$, 95^U ja 99^U.

Ka algandmete maatriksi väljatrükk on võimalik. Konkreetse tunnuste kombinatsiooni puhul teostatud arvutuste tulemused trükitakse 4-8 kümnendkohaga arvude tulpadena (q). Laitrükkali abil trükitakse 1 - 25-veeruliste ribadena (w) välja mainitud tulpadest koosneva maatriksi alumine kolmnurk.

Algandmete hulk on piiratud järgmiselt:

1) tunnuste arv $M \leq 1640$, objektide arv $N \leq 1640$.

2) objektide arv N peab praktiliselt olema märksa väiksem lubatust, sest muidu tuleb tulemusi väljastada liiga kitsaste ribadena. Ühes ribas korraga trükitavate (tulpadest koosnevate) veergude arv sõltub soovitud kümnendkohtade arvust q (kui q=4 $w_{max} \leq 25$, kui q=5 $w_{max} \leq 21$, kui q=6 $w_{max} \leq 18$) ja objektide arvust N. Võimalik maksimaalne veergude arv ribas w_{max} ja objektide arv N on seotud järgmiselt $/(w + 1)\, N \leq 3284/$:

N	$\leq$	126	131	136	142	149	156	164	172	182	193	205
w_{max}	=	25	24	23	22	21	20	19	18	17	16	15
N	$\leq$	217	234	252	273	298	328	364	410	469	547	656
w_{max}	=	14	13	12	11	10	9	8	7	6	5	4

Algandmed võivad olla lünklikud. Kui $N_{ij} \geq 3$ teostatakse kõik arvutused. Kui $N_{ij} = 2$, siis leitakse r_{ij}, N_{ij}, $\bar{x}_{i \cdot j}$, $\bar{x}_{j \cdot i}$, $S_{x_{i \cdot j}}$, $S_{x_{j \cdot i}}$, $c_{i \cdot j}$, $c_{j \cdot i}$, $S_{\bar{x}_{i \cdot j}}$, $S_{\bar{x}_{j \cdot i}}$.

Algandmeteks võivad olla arvud -9998 kuni -0000 ja +0001 kuni +9999; koma asukoht eri tunnustel on suvaline, samal tunnusel konstantne. Kui mingi mõõtmistulemus on täpselt null, siis tuleb see algandmete tabelis märkida kujul -0000. Puuduva arvu tunnuseks on +0000; algandmete tabelis tuleb lünga asemel kirjutada 0000 või veel parem ——. Arv -9999 ei tohi algandmena esineda, seda kasutatakse algandmete massiivi täiendamiseks, et tunnuse või objekti lõpp langeks kokku perfokaardi lõpuga.

Algandmed esitada tabelina:

1) ülesande number - suvaline kuue sümboli kombinatsioon (numbrid ja vene tähestiku tähed);
2) millised näitajad on tarvis arvutada ja välja trükkida;
3) ülesande parameetrid N, M, q;
4) kas algandmed tuleb perforeerida tunnuse või objekti kaupa (olenevalt sellest, kas järgnevates ülesannetes muudetakse tunnuste või objektide hulka).

Algandmete tabel (tunnuste kaupa perforeerides)

Objekti nr. \ Tunnuse nr.	1	2	...	...	M
1	xxxx	xxxx	...	...	xxxx
2	xxxx	xxxx	...	...	xxxx
...	....	....	...	...	....
N	xxxx	xxxx	...	...	xxxx

Algandmete perforeerimisel objektide kaupa tuleb tunnused paigutada tabeli ridadena, objektid tabeli veergudena.

FAKTORANALÜÜSI PROGRAMM ELEKTRONARVUTILE "URAL-4"

1. Faktoranalüüsi rakendamise võimalustest.

Nähtuse üksikute omaduste ja eri külgede vaheliste seoste uurimine annab võimaluse saada informatsiooni ka selle nähtuse kui terviku kohta. Kasutades nähtust iseloomustavaid ja kirjeldavaid (nii kvantitatiivseid kui kvalitatiivseid) tunnuseid, luuakse uuritava nähtuse teatav mudel. Mudeli koostamiseks on vaja määrata, millised tunnused on olulised ja annavad küllalt täpse kirjelduse uuritavast nähtusest. Samuti on tarvis eelnevalt otsustada, kas selliseid olulisi tunnuseid on üldse teada või kas neid saab moodustada antud tunnuste lineaarsete kombinatsioonidena. Praktika seisukohalt on tähtis, et oluliste tunnuste arv oleks võimalikult väike, ent siiski küllalt suur selleks, et nähtus nende kaudu osutuks nõutava täpsusega modelleerituks. Mõnede sellist tüüpi probleemide lahendamisel saab kasutada faktoranalüüsi.

Faktoranalüüsis lähtutakse korrelatsioonimaatriksist, mis teatavasti esitab lineaarsed seosed kõigi tunnuste vahel. Faktoranalüüsi tulemusena saadav faktorite maatriks esitab samuti lineaarsed seosed kõigi tunnuste vahel, kuid mitte enam tunnustevaheliste korrelatsioonidena, vaid leitud korrelatsioonidena faktorite vahel. Kui nõutava täpsuse tagamiseks vajalike faktorite arv on väiksem tunnuste arvust ja faktoreid on võimalik mõistlikul viisil tõlgendada, siis oleme saavutanud kokkuhoiu nähtuse kirjeldamiseks tarvilike näitajate arvus ja määranud need näitajad faktorite näol (kus iga faktor kujutab endast tunnuste lineaarset kombinatsiooni). Täiesti uute ja uurimata nähtuste puhul annab faktoranalüüs võimaluse püstitada hüpoteese nähtuse tekkemehhanismi kohta, klassifitseerida tunnuseid või indiviide ja kontrollida olemasolevaid hüpoteese. Eriti huvitav on võimalus faktoranalüüsi abil hüpoteese püstitada, sest enamik teisi meetodeid võimaldab ainult olemasolevaid hüpoteese kontrollida.

Kuna faktoranalüüsi tulemuste tõlgendamine on küllaltki keeruline, siis on meetodi praktilisel kasutamisel tingimata tarvis tutvuda vastava kirjandusega. Selleks võib soovitada:

1. P. Lorenz, "Anschauungsunterricht in Mathematischer Statistik", Band III, Leipzig, 1961.
2. R.B. Cattell, "Factor Analysis", New York, 1952.
3. T. Veldre, "Ülevaade faktoranalüüsist", käsikiri TRÜ Arvutuskeskuses.

2. Programmi kasutamine.

Programmi kasutamiseks tuleb ette anda:

1) soovitav täpsus faktori koordinaatide leidmisel. Maksimaalne võimalik täpsus on 2^{-18}. Mida suurem on soovitav täpsus, seda kauem läheb aega arvutusteks;

2) ülesande nimetus viietähelise väljendina (vene tähestiku tähed ja numbrid);

3) korrelatsioonimaatriksi järk (s.o. tunnuste arv). Maksimaalne võimalik järk on 40. Leitavate faktorite

maksimaalne arv on 9. Mida suurem on korrelatsioonimaatriksi järk, seda kauem läheb aega faktormaatriksi saamiseks. Faktormaatriks arvutatakse faktorite (s.o. faktormaatriksi veergude) kaupa. Mida rohkem faktoreid leitakse, seda suurem on ka ajakulu;

4) vaatluste või katsete arv, mille põhjal on korrelatsioonid leitud. Seda arvu on tarvis siis, kui soovitakse χ^2-testi abil kontrollida, kas faktoreid on piisavalt või tuleks nende arvu suurendada;

5) korrelatsioonimaatriksi ülemine kolmnurk ilma diagonaali elementideta. Maatriksi ülemise kolmnurga elemendid kirjutatakse järjestikku (esimesest reast alates) üheteistkümne kaupa gruppidesse. Seejuures märgitakse ainult neli kümnendkohta pärast koma. Kui korrelatsioonikordaja on negatiivne, siis tuleb numbrite ette kirjutada miinusmärk (plussmärki pole tarvis kirjutada). Andmete koosseisu kuuluvad nullid tuleb asendada numbrite kombinatsiooniga "0001";

6) võtmete numbrid, mis peavad ülesande lahendamise kestel olema sisse lülitatud. Kui sisse on lülitatud

a) võti nr. 1, siis trükitakse korrelatsioonimaatriksi ülemine kolmnurk, tunnuste ja katsete arv (s.o. punktides 3, 4 ja 5 nõutud andmed);

b) võti nr. 3, siis arvutatakse ja trükitakse χ^2 väärtus ning vabadusastmete arv pärast iga faktori arvutamist;

c) võti nr. 5, siis trükitakse jääkmaatriks nelinurksel kujul pärast iga faktori arvutamist;

d) võti nr. 6, siis trükitakse iga faktori korral iteratsioonide arv ja faktormaatriks pärast soovitud täpsuse saavutamist;

e) võti nr. 7, siis trükitakse iga faktori korral alglähend ja igal iteratsioonisammul faktormaatriks.

Kui ühtegi võtit pole sisse lülitatud, siis trükitakse vastus pärast üheksanda faktori leidmist. Soovitav on alati nõuda, et võti nr. 6 oleks sisse lülitatud.

N ä i d e:

Olgu tarvis teostada faktoranalüüs järgmise korrelatsioonimaatriksi põhjal:

1,0000	-0,2459	0,5436	-0,2115	0,0000	0,7436	0,2424
-0,2456	1,0000	0,2000	-0,1574	-0,5563	0,8432	0,4515
0,5436	0,2000	1,0000	-0,5421	-0,2113	0,2481	0,6213
-0,2115	-0,1574	-0,5421	1,0000	0,0000	0,9874	0,5321
0,0000	-0,5563	-0,2113	0,0000	1,0000	0,2528	0,6174
0,7436	0,8432	0,2481	0,9874	0,2528	1,0000	0,2217
0,2424	0,4515	0,6213	0,5321	0,6174	0,2217	1,0000

Elektronarvutil lahendamiseks esitame selle ülesande järgmiselt:

Tellija perekonna- ja eesnimi: K a l j u s t e, E d a
töökoht ja amet arstiteadusk. V k. üliõp.
telefon 41 - 20/300

Orienteeruv tähtaeg lahenduse saamiseks 20. 11. 67.

1) soovitav täpsus $E = 2^{-10}$

2) ülesande nimetus КАЛЮЕ

3) korrelatsioonimaatriksi järk $n = 7$

4) vaatluste arv $N = 515$

5) korrelatsioonimaatriks

R ⟶

-2459 5436 -2115 0001 7436 2424 2000 -1574 -5563 8432 4515

-5421 -2113 2481 6213 0001 9874 5321 2528 6174 2217

ANKEETIDE TÖÖTLEMISEST

Ankeetide matemaatiliseks töötlemiseks on TRÜ Arvutuskeskuses koostatud programm elektronarvutile "Ural-4". Kuna programm on koostatud standardsena (sellega on võimalik töödelda paljusid erinevaid ankeete), siis peavad ankeedid rahuldama teatud tingimusi, mida tuleb arvestada juba ankeedi koostamisel. Järgnevalt kirjeldatavad nõuded on põhiliselt seotud elektronarvuti "Ural-4" tehniliste võimalustega. Kaks põhilist nõuet, mida tuleb ankeedi koostamisel silmas pidada, on:

1) ankeet võib sisaldada kuni 440 küsimust;
2) igale küsimusele võib esitada kuni 15 vastusevarianti või alajaotust.

Vastusevariante või alajaotusi kodeeritakse numbritega 1, 2, 3, 4, ..., 15.

Näiteks võib vanuse järgi moodustada rühmad, mida kodeerime järgmiselt:

1. 15 - 20 a.
2. 21 - 25 a.
3. 26 - 30 a.
4. 31 - 35 a.
5. 36 - 40 a.
6. 41 - 45 a.
7. 46 - 50 a.

8. 51 - 55 a.
9. 56 - 60 a.
10. 61 - 65 a.
11. 66 - 70 a.
12. 71 ja vanemad.

Vastaja kas tõmbab kriipsu alla sellele klassile, kuhu tema vanus kuulub või ümbritseb klassi numbri ringiga.

Arvamused mingi nähtuse või objekti meeldivuse kohta võime kodeerida järgmiselt:

1. Meeldib.
2. Ükskõik.
3. Ei meeldi.

või

1. Väga meeldib.
2. Meeldib.
3. Ükskõik.
4. Ei meeldi.
5. Täiesti vastumeelne.

Küsimused, kus on võimalik vastata kas "jah" või "ei", kodeeritakse järgmiselt:

1. Jah.
2. Ei.

Oluline on, et vastuse variandid üksteist välistaksid. Juhul, kui pole võimalik vastuse variante nii koostada, et nad üksteist välistaksid, on soovitav juurde märkida, et vastaja kriipsutaks alla ainult ühe vastuse, sest programm on koostatud eeldusel, et igale küsimusele on antud ainult üks vastus.

Elektronarvutil töötlemiseks tuleb ankeetides leiduv materjal kanda perfokaartidele. Perforeerimisel kantakse ühele kaardile 44 küsimust, seega saab ühe ankeedi maksimaalseks pikkuseks olla 10 kaarti. Perfokaardil on 11 rida ja igale reale kantakse nelja küsimuse vastused järgmiselt: iga vastuse jaoks (arvud 0-st kuni 15-ni; "0" perforeeritakse sel juhul, kui küsimusele on jäetud vastamata) on kaks kümnendkohta, mis on skeemil

0 0	0 0	0 0	0 0
I	II	III	IV

näidatud ringikestena kastis.

Järjestikku nummerdatud küsimused paigutatakse perfokaardile järgmiselt:

1	2	3	4
5	6	7	8
9	10	11	12
...	...	...	...
41	42	43	44

45., 46., 47., 48. küsimuse vastused asuvad järgmise kaardi esimeses reas. Iga küsimus omab seega kaardil kindla koha. Et vältida vigu perforeerimisel, on otstarbekas tähistada ankeedis iga neljas küsimus (millega lõpeb rida kaardil) näiteks ringiga, iga 44-s küsimus aga ruuduga (tähistab kaardi lõppu).

Kogu töö, mida masin peab sooritama antud materjaliga, tuleb masinale esitada käsu kujul. Pole mõeldav, et masin koostaks kõikvõimalikud sagedustabelid, see võtaks väga palju masina töö aega ja ka uurijal oleks raske üles leida talle vajalikke tabeleid.

Masin koostab kahe turnuse vahel sagedustabeli.

Tabel 1.

		1. küsimuse vastusevariandid						
		0	1	2	3	...	n	
2. küsimuse vastusevariandid	0	N_{oo}	N_{01}	N_{02}	N_{03}	...	N_{0n}	$\sum_{k=0}^{n} N_{0k}$
	1	N_{10}	N_{11}	N_{12}	N_{13}	...	N_{1n}	$\sum_{k=0}^{n} N_{1k}$
	2	N_{20}	N_{21}	N_{22}	N_{23}	...	N_{2n}	$\sum_{k=0}^{n} N_{2k}$
	..	...	...	...	...	...	...	....
	m	N_{m0}	N_{m1}	N_{m2}	N_{m3}	...	N_{mn}	$\sum_{k=0}^{n} N_{mk}$
		$\sum_{i=0}^{m} N_{10}$	$\sum_{i=0}^{m} N_{11}$	$\sum_{i=0}^{m} N_{12}$	$\sum_{i=0}^{m} N_{13}$	...	$\sum_{i=0}^{m} N_{1k}$	$\sum_{k=0}^{n} \sum_{i=0}^{m} N_{1k}$

Nimetame esimeseks küsimuseks seda, mille vastusevariandid asetsevad erinevates veergudes ja teiseks küsimuseks seda, mille vastusevariandid asetsevad erinevatel ridadel.

Tabelis nr.1 arv N_{11} on nende vastajate arv, kes esimesele küsimusele vastasid vastusevariandiga 1 ja teisele küsimusele samuti vastusevariandiga 1. N_{21} on nende vastajate arv, kes esimesele küsimusele vastasid vastusevariandiga 1 ja teisele küsimusele vastusevariandiga 2. N_{13} on nende vastajate arv, kes vastasid esimesele küsimusele vastusevariandiga 3 ja teisele küsimusele vastusevariandiga 1. N_{10} on nende vastajate arv, kes jätsid esimesele küsimusele vastamata ja teisele küsimusele vastasid vastusevariandiga 1. N_{00} on nende vastajate arv, kes jätsid mõlemale küsimusele vastamata.

Niisiis, iga veerg tabelis tähendab teatud kindlat esimese küsimuse vastusevarianti (0, 1, 2, 3, ..., 15 vasakult paremale) ja iga rida teise küsimuse vastusevarianti (0, 1, 2, ..., 15 ülalt alla). Paremalt äärmises veerus on reasummad ja alumises reas veerusummad. Nimetatud rea ja veeru ristumiskohal on kogu tabeli summa. Kõik arvutused sooritatakse selle tabeli alusel.

Järgnevalt vaatleme käsu formuleerimist. See tähendab eeskirja andmist masinale, missuguste küsimuste vahel tuleb koostada sagedustabel.

Käsklus masinale kahe tunnuse vahelise (näit. 5 ja 12) sagedustabeli koostamiseks antakse kujul 5 - 12 (1). On võimalik anda veel lisatingimusi, s.t. sorteerida ankeete tunnuste kaupa ja koostada sagedustabel väljaeraldatud ankeetide põhjal.

Käsk

5 (7) 6 (1,2) 3 (1,4,5) 2 (4,5) 1 - 7 (2)

tähendab, et ankeetide hulgast on välja valitud need, kus 5. küsimusele on antud vastus 7; 6. küsimusele kas vastus 1 või vastus 2; 3. küsimusele kas vastus 1 või vastus 4 või vastus 5; 2. küsimusele kas vastus 4 või vastus 5, ja nende ankeetide järgi on koostatud sagedustabel 1. ja 7.küsimuse vahel.

Antud käsul on neli eeltingimust. Üldse võib eeltingimusi olla maksimaalselt 9. Sulgudes esinevaid (komaga eraldatud) arve võib olla maksimaalselt 3. Toome veel mõned näited käskudest:

2 (1) 3 - 5 (3)
2 (2) 3 - 5 (4)
2 (3) 3 - 5 (5)
2 (1,2,3) 3 - 5 (6)
12 (14,15) 20 (1,2,3) 32 (7,8,11) 16 - 24 (7)

Käskudes (3), (4), (5), (6) on üks eeltingimus, käskudes (6) aga kolm eeltingimust. Käsuga (6) saadud sagedustabel on sama, mis käskudega (3), (4) ja (5) saadud tabelid üheskoos. Oluline on,et iga käsk lõpeks sagedustabeli koostamisega (s.o. kahe küsimuse numbriga ja nende vahel oleva miinusmärgiga).

Käskude masinasse viimiseks tuleb need arvud teisendada kaheksandsüsteemi ja kirjutada programmiblanketile kahte veergu järgmiselt (näiteks käsk (2)):

5	07
6	0102
3	010405
2	0405
1 -	7

Selles näites on arvud valitud nii, et nende esitused kaheksand- ja kümnendsüsteemis ühtuksid.

Teised eespool toodud näited:

2		01	(3)
3	-	5	
2		02	(4)
3	-	5	
2		03	(5)
3	-	5	
2		010203	(6)
3	-	5	
14		1617	
24		010203	
40		071013	
20	-	30	(7)

Vasakpoolses tulbas on küsimuste numbrid, parempoolses tulbas on kas miinusmärgiga teise küsimuse number või eeltingimuse vastusevariantide väärtused.

Käsu jaoks kasutatakse perfokaardil kümmet rida. Üheteistkümnendasse ritta, või (kui perfokaardil on käske vähem kui kümme rida, siis) viimasele käsule järgnevasse ritta kirjutame vasakpoolse tulba alla -0 (miinus nulli), mis on tunnuseks, et kaardil enam rohkem käskusid pole.

Näide:

5	07
6	0102
3	010405
2	0405
1	- 7
2	01
3	- 5
2	02
3	- 5
5	- 14
-0	

Numbrid sulgudes, mis on omavahel eraldatud komaga, tuleb kaardile kanda kahekohaliste arvudena. See on vajalik selleks, et perforeerija teaks missugusest veerust perforeerida.

Kui on vaja esitada pikemat käsku, kui kaardile mahub, siis kirjutame selle uuele kaardile, eelmisele aga jätame tühja ruumi või kirjutame hiljem sinna mõne niisuguse käsu, mis sinna mahub. Tuleb meeles pidada, et käsk lõpeb miinusmärgiga 2. küsimuse numbri ees ja et kaart lõpeb miinus nulliga käsule järgnevas esimeses tühja rea vasakus veerus.

Tulemustest trükib masin kõigepealt käsu järgmisel kujul:

ПРИЗНАК $\longrightarrow \downarrow$ 002 (010203) 003 - 005

mis on kujult samane käsu (6) kirjutusviisiga. Nooled näitavad siin küsimuse muutumise suunda sagedustabelis. Nool paremale näitab rea esimese küsimuse (003) muutumise suunda, nool alla teise küsimuse (005) muutumise suunda (need on küsimused, mille vahel koostatakse sagedustabel, mitte eeltingimused).

Edasi trükitakse sagedustabel sellisel kujul nagu on näidatud tabelis 1.

Teises tabelis esitatakse samad andmed aga protsentides kogu tabeli summast. Kolmas ja neljas tabel on sagedustabeli protsendid ridade ja veergude järgi.

Peale nende trükitakse välja järgmised suurused:

Ж = X^2 üle kogu tabeli;

H = vabadusastmete arv;

Ф = Tšuprovi koefitsient.

Järgnevate suuruste arvutamisel on 0-rida ja 0-veerg välja jäetud:

KK = korrelatsioonikordaja;

M = tabeli summa (kus 0-veergu ja 0-rida pole arvestatud);

Д2 = teise tunnuse standardhälve;

Д1 = esimese tunnuse standardhälve;

C2 = teise tunnuse keskmine;

C1 = esimese tunnuse keskmine;

B2 = teise tunnuse variatsioonikordaja;

B1 = esimese kordaja variatsioonikordaja;

C→ = üksikute ridade keskmised tabelis (tinglikud keskmised);

C↓ = üksikute veergude keskmised tabelis (tinglikud keskmised).

Oluline on märkida, et korrelatsioonikordajat, standardhälvet ja keskmisi võib kasutada ainult nende küsimuste puhul, mida saab kvantitatiivselt mõõta, või siis, kui näitajat ei saa küll kvantitatiivselt mõõta, kuid ta muutub kindlas suunas. Viimaste kasutamisel tuleb olla ettevaatlik.

Oluline on märkida, et juba ankeedi koostamise ajal peab olema teada, kus ja missuguse programmi järgi seda töötlema hakatakse, et oleks võimalik arvestada vastavaid nõudeid ankeedi koostamisel. Kui sellele aga pole möeldud, siis võib ankeedi töötlemisega tekkida tõsiseid raskusi.

S I S U K O R D

Hind 10 kop.