

## II

## TUNNUSE GENEETILISE DETERMINEERITUSE UURIMINE FENOTÜÜBI- JA PÕLVNEMISANDMETE ALUSEL

### 2.1 Klassikaline segregatsioonanalüüs

Diskreetse tunnuse (enamasti haiguse) geneetilise determineerituse uurimine baseerub paljudel juhtudel Gregor Mendeli juba 19. sajandi lõpul tehtud avastusel, et mingi diskreetse tunnuse poolest erinevate vanemate ristamisel saadud järglas põlvkonnas jaotuvad selle tunnuse väärtused kindlates proportsioonides. See ja teised Mendeli katsete põhjal sõnastatud järeldused on tänapäeval tuntud Mendeli seadustena ning need on aluseks, uurides vaid ühe alleelipaari pärandumist.

**Mendeli I seadus** e nn esimese põlvkonna ühtlikkuse seadus: homosügootsete isendite ristamisel saadakse esimeses põlvkonnas kõik ühetaolised järglased (vaadeldava tunnuse osas).

**Mendeli II seadus** e nn alleelide lahknemise seadus: heterosügootsete F1-hübriidide omavahelisel ristamisel toimub teises põlvkonnas (F2) tunnuste lahknemine kindlates sagedussuhetes (nn **lahknemissuhetes**).

**Klassikaline segregatsioon- e lahknemisanalüüs** püüab tuvastada diskreetse uuritava tunnuse määratust vaid ühes lookuses paikneva alleelidekomplekti (ehk siis ühe geeni) poolt ja põhineb uuringul, kas tunnuse väärtused lahknevad järglaste hulgas vastavalt Mendeli seaduspäradele? Jaatav vastus ongi tõend tunnuse väärtuste determineeritusest vaid ühe geeni poolt.

#### 2.1.1 Segregatsioonanalüüs dominantse tunnuse korral

Oletame, et uuritavat haigust põhjustab väga harva esinev geenmutatsioon. Tähistades haigust põhjustavat muteerunud alleeli  $D$  ja normaalset alleeli  $d$ , on oletusliku geeni suhtes võimalikud kolm erinevat genotüüpi:  $DD$ ,  $Dd$  ja  $dd$ . Kui haigus on dominantne, on haiged nii genotüübiga  $DD$  kui ka genotüübiga  $Dd$  indiviidid (inimestel on sellisteks haigusteks näiteks Huntingtoni tantstõbi ja *neurofibromatosis*). Selliste haiguste korral on genotüüp ja fenotüüp sedavõrd täpselt vastavuses, et sobivalt valitud vanemate korral on alleelide lahknemisseaduse täidetud selgelt tuvastatav.

Kolme võimaliku genotüübi baasil on moodustatavad 9 erinevat vanemate kombinatsiooni e paarumistüüpi. Kõigile nende vanempaaride järglastele on omane nii konkreetne genotüübijaotus kui ka fenotüübijaotus, mis kokkuvõtvalt on esitatud tabelis 2.1. Kuna dominantseid haigused on harva esinevad, on ilmselt ka haigust põhjustava alleeli  $D$  sagedus väike ning peaaegu kõik haiged indiviidid omavad genotüüpi  $Dd$ , mitte  $DD$ . Seega on ka enamus haige ja terve indiviidi järglaste genotüüpidest kujunenud genotüüpide  $Dd$  ja  $dd$  ristamisel. Tabelist on näha, et sellise paarumistüübi järglaste haigestumistõenäosus on  $1/2$ . Viimast kasutataksegi, kontrollimaks uuritava haiguse tingitust dominantsest alleelist.

Tabel 2.1 Lahknemissuhted dominantse haiguse korral

Paarumis- tüüp	Järglase genotüüp			Fenotüüp	
	$DD$	$Dd$	$dd$	Haige	Norm.
$DD \times DD$	1	0	0	1	0
$DD \times Dd$	0,5	0,5	0	1	0
$DD \times dd$	0	1	0	1	0
$Dd \times Dd$	0,25	0,5	0,25	0,75	0,25
$Dd \times dd$	0	0,5	0,5	0,5	0,5
$dd \times dd$	0	0	1	0	1

##### 2.1.1.1 Binoomjaotusel baseeruv test

Lihtsaimaks meetodiks, tuvastamiseks dominantset haigusealleeli, on testida, kas haigestunute proportsioon haige ja terve vanema järglaste hulgas on  $1/2$ , st kas haigestunute arv on binoomjaotusega  $B(n, 1/2)$ , kus  $n$  on kõigi valitud vanempaaride järglaste arv.

Olgu haigestunud järglasi kokku  $r$ . Kontrollitav hüpoteesidepaar on siis kujul

$$\begin{array}{ll} H_0: p = 1/2, & \\ H_1: p \neq 1/2; & \Leftrightarrow \\ & H_0: r = n/2, \\ & H_1: r \neq n/2. \end{array}$$

Et binoomjaotuse tõenäosusfunktsioon on kujul  $P(x=k) = C_n^k p^k (1-p)^{n-k}$ , kus  $X \sim B(n, p)$ , siis võrdub olulisuse tõenäosus, mille alusel nullhüpootees kas kummutatakse või mitte, kahekordse tõenäosusega, et binoomjaotusega  $B(n, 1/2)$  juhusliku suuruse absoluutväärtus on väiksem või võrdne suurusest  $n/2 - |n/2 - r|$ :

$$2P(x \leq n/2 - |n/2 - r|) = 2 \times \sum_{x=0}^{n/2 - |n/2 - r|} C_n^x \underbrace{\left(\frac{1}{2}\right)^x \left(1 - \frac{1}{2}\right)^{n-x}}_{(1/2)^n} = \left(\frac{1}{2}\right)^{n-1} \times \sum_{x=0}^{n/2 - |n/2 - r|} C_n^x \quad (2.1)$$

kus  $C_n^x$  märgib kombinatsioonide arvu  $n$  elemendist  $x$  kaupa.

---

**Näide.** Opalestsentse dentiini uuringuil (Neel ja Schull, 1954) koguti andmed 112 haigestunud vanema järglaste kohta ja leiti sama haigus 52-l. Kas on põhjust lükata ümber väide uuritava haiguse dominantsest geneetilistest tingitusest?

Vastavalt valemile (2.1) on lahknemissuhte  $1/2$ -ga võrdumise hüpoteesi kohta käiv olulisustõenäosus

$$\left(\frac{1}{2}\right)^{111} \sum_{x=0}^{52} C_{112}^x = 0,5058,$$

mistõttu ei ole mingit alust lükata ümber hüpoteesi lahknemissuhte võrdumisest  $1/2$ -ga ja seega on andmed kooskõlas ka hüpoteesiga uuritud haiguse ühe dominantse alleeli poolt määratusest.

---

### 2.1.1.2 Normaaljaotusel baseeruv test

Kuna binoomjaotus on vaatluste arvu ehk siis järglaste arvu suurenedes lähendatav normaaljaotusega, on viimane kasutatav ka testimaks haigestunute proportsiooni  $p$  erinevust  $1/2$ -st terve ja haige vanema järglaste hulgas. Hüpooteeside kontrollimiseks vajalik teststatistik, mis nullhüpoteesi kehtides on standardse normaaljaotusega, on kujul

$$z = \frac{|r - n/2|}{(n/4)^{1/2}},$$

sest teatavasti  $X \sim N(\mu, \sigma^2)$  korral  $(X - \mu)/\sigma \sim N(0, 1)$ , kus  $\hat{\mu} = np$  ja  $\hat{\sigma} = np(1-p)$ .

---

**Näide.** Eelnevas näites kirjeldatud opalestsentse dentiini andmete alusel leitud  $z$ -statistik on

$$z = (52 - 56)/5,29 = -0,756.$$

Eksimise tõenäosus lahknemissuhte  $1/2$ -ga võrdumise hüpoteesi kummutamisel on võrdne tõenäosusega, et standardse normaaljaotusega juhuslik suurus on väiksem kui  $-0,756$  või suurem kui  $0,756$ . Et vastav tõenäosus on  $0,4496$ , on järeldus sama, mis binoomjaotuse kasutamise korralgi – andmed on kooskõlas hüpoteesiga opalestsentse dentiini määratusest ühe dominantse alleeli poolt.

---

### 2.1.1.3 $\chi^2$ -test

Et standardse normaaljaotusega juhusliku suuruse ruut on  $\chi^2$ -jaotusega vabadusastmete arvuga 1, on ka eelnevalt defineeritud  $z$ -statistiku ruut nullhüpoteesi  $p = 1/2$  kehtides jaotusega  $\chi_1^2$ . Olulisustõenäosus, võimaldamaks kummutada nullhüpoteesi, on võrdne tõenäosusega, et ühe vabadusastmega  $\chi^2$ -jaotusega juhusliku suuruse väärtus on suurem teststatistiku  $z^2$  väärtusest, kus

$$z^2 = \frac{(r - n/2)^2}{n/4}. \quad (2.2)$$

Analoogse statistikuni viib ka andmetest hinnatud sageduste võrdlemine teoreetiliste eelduste ( $p = 1/2$ ) täidetuse korral leitud. St, et juhuslik suurus

$$Z^2 = \frac{(X - np)^2}{np} + \frac{((n - X) - n(1 - p))^2}{n(1 - p)} \quad (2.3)$$

on nullhüpoteesi korral  $\chi^2$ -jaotusega vabadusastmete arvuga 1.

Asendades juhusliku suuruse  $Z^2$  avaldises (2.3) suuruse  $X$  tema andmete alusel leitud väärtusega  $r$  ning võttes nullhüpoteesi järgselt tõenäosuse  $p$  võrdseks  $1/2$ -ga, saamegi teststatistiku  $z^2$  valemi (2.2):

$$z^2 = \frac{(r - np)^2}{np} + \frac{((n - r) - n(1 - p))^2}{n(1 - p)} = \frac{(r - np)^2}{np} + \frac{(np - r)^2}{n(1 - p)} = \frac{(r - np)^2}{np(1 - p)} = \frac{(r - n/2)^2}{n/4}.$$

$\chi^2$ -testi eeliseks võrreldes binoom- ja normaaljaotusel baseeruvate testidega on selle rakendatavus ka enam kui kahe indiviidide grupi korral.

**Näide.** Vastavalt valemile (2.2) saame teststatistiku väärtuseks

$$z^2 = \frac{(52 - 112/2)^2}{112/4} = \frac{16}{28} = \frac{4}{7} = 0,5714.$$

Analoogselt tulemusele jõuame ka, kasutades traditsioonilist  $\chi^2$ -statistiku arvutamise eeskirja empiiriliste ja teoreetiliste sageduste baasil (vajalikud arvutused on mugav koondada tabelisse).

	Haigeid	Terveid	Kokku
Empiiriline ( $E$ )	52	60	112
Teoreetiline ( $T$ )	56	56	112
$(E - T)^2 / T$	0,2857	0,2857	0,5714

Tõenäosus, et jaotusega  $\chi^2$  juhusliku suuruse väärtus on suurem kui 0,5714 on 0,4497, mis on identne normaaljaotusel baseeruva testi korral leituga.

### 2.1.1.4 Suurima tõepära meetod

Lisaks eelnevatele meetoditele on hüpotees  $p = 1/2$  testitav parameetri  $p$  funktsioonina esitatud tõepärafunktsiooni abil. Juhul, kui vaatluse all on kokku  $n$  haige ja terve vanema järglast, kellest  $r$  on haiged, avaldub tõepärafunktsioon kujul

$$L(p) = C_n^r p^r (1 - p)^{n-r}.$$

Tähistades tõepärafunktsiooni maksimumi  $0 \leq p \leq 1$  korral  $L_1$  ja  $p = 1/2$  korral  $L_0$ , on tõepärasuhte statistik defineeritud seosega

$$2 \ln(L_1/L_0) = 2(\ln L_1 - \ln L_0).$$

Nullhüpoteesi kehtides on see statistik asümptootiliselt  $\chi^2$ -jaotusega vabadusastmete arvuga 1.

Suurima tõepära hinnang parameetritele  $p$ , mis maksimeerib  $L_1$ , on leitav standardsel viisil, võrdsustades tõepärafunktsiooni tuletise  $p$  järgi nulliga:

$$\frac{\partial \ln L(p)}{\partial p} = C_n^r [r p^{r-1} (1 - p)^{n-r} - (n - r) p^r (1 - p)^{n-r-1}] = 0,$$

millest

$$p^{r-1} (1 - p)^{n-r-1} [r(1 - p) - (n - r)p] = p^{r-1} (1 - p)^{n-r-1} (r - np) = 0$$

ja  $\hat{p} = r/n$ .

Tõepärasuhte statistik esitub siis kujul

$$\begin{aligned} 2(\ln L_1 - \ln L_0) &= 2 \left\{ \ln \left[ C_n^r \left( \frac{r}{n} \right)^r \left( 1 - \frac{r}{n} \right)^{n-r} \right] - \ln \left[ C_n^r \left( \frac{1}{2} \right)^r \left( 1 - \frac{1}{2} \right)^{n-r} \right] \right\} \\ &= 2 \left[ r \ln \left( \frac{r}{n} \right) + (n - r) \ln \left( \frac{n - r}{n} \right) - n \ln \left( \frac{1}{2} \right) \right]. \end{aligned}$$

Otsuse vastuvõtmiseks vajalik olulisustõenäosus on siis võrdne tõenäosusega, et  $\chi^2$ -jaotusega juhuslik suurus omandab väärtuse, mis on suurem leitud tõepärasuhte statistiku väärtusest.

**Näide.** Opalestsentse dentiini andmete vastav logaritmitud tõepärafunktsioon (konstant kombinatsioonide arvu näol jäetakse enamasti kirjutamata, sest nii suurima tõepära hinnangu kui ka tõepärasuhte konstrueerimisel taandub see nagunii välja) on

proportsionaalne suurusega

$$\ln L(p) \propto 52 \ln(p) + 60 \ln(1-p).$$

Parameetri  $p$  suurima tõepära hinnang on  $52/112$ . Tõepärasuhte statistiku leidmiseks vajalikud arvutused on mugav koondada tabelisse.

Mudel	$p$	$L(p)$	$2 \ln L(p)$
Üldine	$52/112$	0,07541	-5,1695
Dominantne	$1/2$	0,05666	-5,7414

Tõepärasuhte statistik on siis 0,5719. Olulisustõenäosus, saadud statistiku võrdlemisel  $\chi^2$ -jaotusega juhusliku suurusega, on 0,4495, mis on peaaegu identne eelnevalt  $\chi^2$ -testil või normaaljaotusel baseerual testil leituga.

### 2.1.2 Segregatsioonanalüüs kodominantse tunnuse korral

Kodominantsus mingi lookuse (geeni) suhtes tähendab, et antud lookuses heterosügootsed indiviidid erinevad oma fenotüübi poolest sama lookuse suhtes homosügootsetest indiviididest.

Näiteks dialleelse lookuse korral on kõik kolm võimalikku genotüüpi üksteisest fenotüübi alusel eristatavad. Seetõttu on ka erinevalt dominantsest lookusest eristatavad kõik kuus võimalikku paarumistüüpi ja iga ühe puhul neist võib kontrollida järglaste lahknemist vastavalt Mendeli seaduspäradele. Et osade vanempaaride järglaste hulgas on eristatavad kolm erinevat gruppi, ei ole kirjeldatud binoom- ja normaaljaotusel baseeruvad testid neil juhtudel rakendatavad. Küll aga on enam kui kahe võimaliku genotüübi juhule üldistatavad  $\chi^2$ -test ja tõepärasuhte test.

### 2.1.3 Segregatsioonanalüüs retsessiivse tunnuse korral

Retsessiivse tunnuse, mis fenotüübis avaldub vaid homosügootse genotüübi (näiteks  $dd$ ) korral, uurimisel ei ole võimalik täpselt tuvastada ühtki huvipakkuvat paarumistüüpi. Ainuke üheselt määratav genotüüp on  $dd$ , kuid kahe taolise homosügootse indiviidi omavaheline ristamine ei anna statistiliste analüüsides tarvis informatiivseid järglasi (kõik järglased on ühesugused). Sobivaimad vanemad, kelle järglasi uurida, on heterosügootid – peaks ju juhul, kui uuritav tunnus (haigus) on määratud ühe retsessiivse geeni poolt, paarumistüübi  $DdxDd$  järglastest vastavalt Mendeli seadustele  $1/4$  olema genotüübiga  $dd$  ja seega ka haiged. Probleem on aga selles, et retsessiivse tunnuse puhul on indiviidid genotüüpidega  $Dd$  ja  $dd$  fenotüübi järgi eristamatud ja seega ei ole võimalik vaid väliste tunnuste (haigestumine) alusel võtta vaatluse alla üksnes paarumistüübiga  $DdxDd$  perekondi. Et tervete vanemate järglane saab retsessiivse haiguse puhul olla haige vaid siis, kui mõlemad vanemad on haigust põhjustava lookuse suhtes heterosügootid, sisaldab parim katseplaan vaid haige(te) järglas(te)ga perekondade analüüsi. Samas jätab selline uuring kõrvale need heterosügootsete vanemate paarid kellel juhuse läbi haiged järglased puuduvad, tuues nii kaasa haigestumise sageduse ülehindamise.

Hoolimata võimatusest kaasata uuringusse kõiki retsessiivseid vanemate paare on just haige järglase järgi tuvastatud paarumistüübiga  $DdxDd$  perekondade analüüs peamine vahend, kontrollimaks kahtlusi mingi haiguse ühe retsessiivse alleeli poolt määratusest, kasutades üksnes fenotüübi ja põlvnemisandmeid.

Lahknemissuhte (retsessiivse haiguse alleeli korral võrdub see haigestumise tõenäosusega) ülehindamise vältimiseks püütakse lihtsamal juhtudel korrigeerida sobivalt nii haigete järglaste arvu kui ka kõigi järglaste arvu, keerulisematel juhtudel aga panna kirja erinevaid tõepärafunktsioone ja kasutada hindamisel numbrilisi meetodeid. Selgituseks veel nii palju, et kõik need meetodid uurivad perekondade valiku efekti arvestamiseks esmalt, kui suur osa mingi perekonna haigetest järglastest esmalt juhuslikult kogu populatsiooni haigetest moodustatud valimisse sattus. Defineeritakse nn tuvastamistõenäosus, mis näitab, kui suur tõenäosus on haigel indiviidil sõltumata teistest esmasesse juhuslikku valimisse sattuda; laiendatakse seda tõenäosust igale haige(te) järglas(te)ga perekonnale; jne.

Vajadusel vt: Sham, P. (1998) Statistics in Human Genetics. Arnold, London. Lk. 22-38.

## 2.2 Genotüübi- ja fenotüübisagedused populatsioonis, Hardy-Weinbergi seadus

Kui eelnevalt vaadeldud segregatsioonanalüüs baseerus kindlat tüüpi vanempaaride järglaste analüüsil, siis kuna kogu populatsioon on erinevate vanempaaride järglaste segu, on erinevate genotüüpide sagedus populatsioonis esitatav erinevatele paarumistüüpidele vastavate lahknemissuhete kaalutud keskmisena, kus kaaludeks on paarumistüüpide esinemissagedused populatsioonis. Juhul, kui vanempaarid moodustuvad juhuslikult, järgivad genotüüpide sagedused populatsioonis kindlat matemaatilist seaduspära. Selle seaduspära sõnastasid 1908. aastal üksteisest sõltumatult Inglise matemaatik G. H. Hardy ja Saksa arst W. Weinberg.

**Hardy-Weinbergi seadus** e populatsiooni geneetilise tasakaalu seadus e populatsiooni-geneetika põhiseadus: suures isoleeritud panmiktilises populatsioonis säilivad alleeli- ja genotüübisagedused põlvkonniti muutumatuks.

Matemaatiliselt väljendatuna tähendab see, et kui meil on kahealleelne lookus alleelidega  $A$  ja  $a$  ning genotüüpide  $AA$ ,  $Aa$  ja  $aa$  sagedused on vastavalt  $P$ ,  $2Q$  ja  $R$ , kus  $P + 2Q + R = 1$ , siis indiviidide juhusliku paaritumise korral kehtib võrdus

$$Q^2 = PR.$$

Hardy-Weinbergi seaduse kehtides on lihtsalt esitatavad ka genotüübi- ja alleelisageduste vahelised seosed. Nimelt, kui alleelide  $A$  ja  $a$  sagedused on vastavalt  $p$  ja  $q$ , kus  $p + q = 1$ , siis eelnevalt defineeritud genotüübisagedused on avaldatavad alleelisageduste kaudu kujul  $P = p^2$ ,  $R = q^2$  ja  $Q = pq$ . Juhul kui uuritava tunnuse geneetilise määratuse kontrollimiseks ei ole kasutada perekondade andmeid, sobib kasutada just Hardy-Weinbergi seadust – tunnuse jaotumine vastavalt sellele seadusele on ilmne tõend tunnuse tugevast geneetilisest determiinatsioonist.

**Näide.** Islandil juhuslikult valitud inimeste veregruppide  $MM$ ,  $MN$  ja  $NN$  sagedused saadi vastavalt 233, 385 ja 129 (Mourant, 1954). On need andmed vastavuses Hardy-Weinbergi seadusega?

Leiame vaadeldavate veregruppide oodatavad sagedused eeldusel, et nende kohta kehtib Hardy-Weinbergi seadus.

Kokku osales uuringus 747 inimest ja vaatluse all oli seega  $2 \times 747 = 1494$  geeni. Alleel  $M$  esines  $(2 \times 233) + 385 = 851$  korda, järelikult on alleeli  $M$  empiiriline sagedus  $851/1494 = 0,5696$ . Analoogselt saame alleeli  $N$  sageduseks 0,4304.

Genotüüpide  $MM$ ,  $MN$  ja  $NN$  Hardy-Weinbergi seadusele vastavad sagedused on siis  $(0,5696)^2 = 0,3245$ ,  $2(0,5696)(0,4304) = 0,4903$  ja  $(0,4304)^2 = 0,1852$ . Korrutades saadud sagedused inimeste arvuga, saame Hardy-Weinbergi seadusele vastavad erinevate veregruppidega inimeste arvud: 242,37 (veregrupiga  $MM$ ), 366,26 ( $MN$ ) ja 138,37 ( $NN$ ).

Empiiriliste andmete võrdlemiseks teoreetiliselt tuletatutega sobiva  $\chi^2$ -statistiku leidmiseks vajalikud arvutused on koondatud tabelisse.

	$MM$	$MN$	$NN$	Kokku
Empiiriline ( $E$ )	233	385	129	747
Teoreetiline ( $T$ )	242,37	366,26	138,37	747
$(E - T)^2 / T$	0,3622	0,9588	0,6345	1,9555

Leitud statistiku väärtusele 1,9555 vastav olulisustõenäosus ühe vabadusastme korral (lisaks tavapärasele ühe vabadusastme kaole, mis kaasneb teoreetiliste absoluutsete sageduste arvutamisel kitsendava tingimusena nende summa fikseerimisega, on alleelisageduste hindamisel tarvilik fikseerida ka üks neist – et alleelisageduste summa võrduks ühega – põhjustades sellega ka teise vabadusastme kao) on 0,1620. Seega ei ole alust ümber lükata hüpoteesi veregruppide  $MN$ -süsteemi geneetilisest Mendeli seadustele vastavast determiinatsioonist.

Suurima tõepära meetodi korral tuleb arvutada kaks tõepärafunktsiooni väärtust – üks neist on tõepärafunktsiooni maksimaalne väärtus üle kõigi võimalike kolme genotüübigrupi sageduste populatsioonis, teine aga tõepärafunktsiooni väärtus Hardy-Weinbergi seadusega kooskõlas olevate genotüübisageduste korral.

Genotüübisageduste suurima tõepära hinnanguteks populatsioonis on lihtsalt erinevate veregruppide suhtelised sagedused valimis:  $\hat{p}_{MM} = 233/747 = 0,3119$  (veregrupp *MM*),  $\hat{p}_{MN} = 385/747 = 0,5154$  ja  $\hat{p}_{NN} = 129/747 = 0,1727$ . Tõepärafunktsiooni maksimaalne väärtus on seega leitav korrutisest

$$L(\hat{p}_{MM}, \hat{p}_{MN}, \hat{p}_{NN}) = \prod_{i \in \{MM, MN, NN\}} (\hat{p}_i)^{n_i} = (0,3119)^{233} \times (0,5154)^{385} \times (0,1727)^{129}.$$

Hardy-Weinbergi tasakaalule vastavad genotüüpide *MM*, *MN* ja *NN* sagedused said leitud juba  $\chi^2$ -testi tarvis ning on vastavalt 0,3245, 0,4903 ja 0,1852. Tõepärafunktsioon avaldub teoreetiliste väärtuste puhul kujul  $L_0 = (0,3245)^{233} \times (0,4903)^{385} \times (0,1852)^{129}$ .

Tõepärasuhte statistiku leidmiseks vajalikud arvutused on mugav koondada tabelisse.

Mudel	Parameetrite väärtused	-2 ln L
Üldine	0,3119; 0,5154; 0,1727	-1506,38
Hardy-Weinberg	0,3245; 0,4903; 0,1852	-1508,34

Tõepärasuhte statistik on siis 1,9592. Tõenäosus, et  $\chi^2_1$ -jaotusega juhusliku suuruse väärtus on sellest suurem, on 0,1616, mis on peaaegu identne eelnevalt  $\chi^2$ -testil leitud ega anna alust nullhüpoteesi kummutamiseks.

### Ülesanne 1.

Peatükis 2.1 kirjeldatud lahknemissuhete analüüs töötab ka siis, kui hüpoteetilisi genee on enam kui 1 – siis tuleb lihtsalt lisaks Mendeli I ja II seadusele rakendada teoreetiliste sageduste arvutamisel ka Mendeli III seadust. Alljärgnevalt üks näide taolisest analüüsist iseseisvaks läbilahendamiseks.

Mesilaste vastsed arenevad kärjekannudes ning neid hooldavad töomesilased. Töomesilaste ülesanne on haiged või surnud vastsed kõrvaldada, et tagada hügieeniline keskkond.

Selles tegevuses on kaks geneetilist kontrollitud etappi:

- (1) haiget vastset sisaldava kärjekannu avamine;
- (2) vastse eemaldamine kärjekannust.

Osa mesilasi on aga nn ebahügieenilised ning jätavad haiged või surnud vastsed kärke.

Kui ristata ebahügieenilisi mesilasi hügieenilistega (mõeldud on puhtaid liine, st et nii hügieenilised kui ka ebahügieenilised mesilased on saadud paljude põlvkondade vastavate omadustega mesilaste järjestikuse ristamise tulemusel ja on seega homosügootsed), on kõik järglased ebahügieenilised, mis näitab, et ebahügieeniline käitumine on dominantne tunnus.

Selgitamiseks täpsemalt mesilaste hügieenilisuse geneetilisi tagamaid, ristati saadud järglasi uuesti hügieeniliste mesilastega.

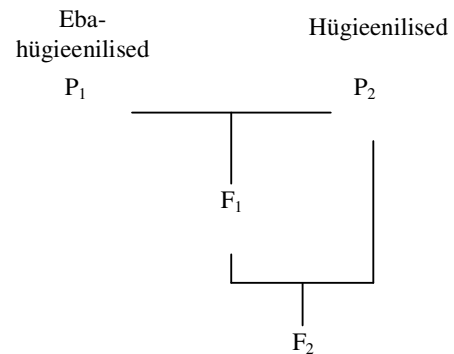
Katse skeem on kujutatud kõrvaloleval joonisel ja tulemused on kokku võetud järgmises tabelis.

	Ristandmesilaste (F2) tüüp				Kokku
	Hügieen.	Ebahüg. <sup>1</sup>	Ebahüg. <sup>2</sup>	Ebahüg. <sup>3</sup>	
$n_i$	110	98	87	105	400

<sup>1</sup> avavad kärjekannud, kuid ei kõrvalda sealt nakatunud vastseid,

<sup>2</sup> ei ava kärjekanne, kuid kõrvaldavad nakatunud vastseid avatud kärjekannudest,

<sup>3</sup> ei ava kärjekanne ega kõrvalda vastseid.



Kas antud katse põhjal saab ümber lükata teadlaste peades tekkinud hüpoteesi, et mesilaste hügieenilisus on määratud kahe dialleelse lookuse poolt, kus üks paar alleele kontrollib kärjekannu avamist, teine aga vastsete kõrvaldamist?