

Aretusväärtuste arvutamisest – kuidas seda tehakse ja milleks seda vaja on?

Tanel Kaart

EPKK Aretuskonverents 2022
08.04.2022, Pärnu



Euroopa Maaelu Arengu
Põllumajandusfond:
Euroopa investeringud
maapiirkondadesse

Olgu kohe alustuseks öeldud, et ...

- | Loomal ei ole vaid üks aretusväärtus. Aretusväärtuseid on palju: nii palju, kui on mõõdetud-registreeritud erinevaid näitajaid (piimatoodang, võõrutusmass, tiinestumine, haigestumine jne) – teoreetiliselt lõpmatu hulk.
- | Ka näiteks pullil on udarapõletiku aretusväärtus ja emisel sperma kvaliteedi aretusväärtus.
- | Aretusväärtused on ka taimedel, bakteritel, inimestel, ...
- | Aretusväärtus ei ole konstant – ta muutub ajas ja ruumis ehk tegu on suhtelise väärtusega, mis sõltub konkreetsest ajahetkest ja populatsioonist.

Definitsioon

- | Aretusväärtus – põllumajanduslooma geneetiline potentsiaal, võrreldes tõukaaslastega, arvestades majanduslikku tasuvust.
[Põllumajandusloomade tõuaretuse seadus, <https://www.riigiteataja.ee/akt/191855>]
- | Aretusväärtus (ingl. *breeding value*): kvantitatiivses geneetikas individuaalse fenotüübi kõrvalekalde osa populatsiooni keskmisest, mis on tingitud alleelide aditiivsest toimest. [<http://geneetika.ee/lexicon/>]
- | Aretusväärtus – vanematelt järglastele edasiantavate geenide väärtus, mis avaldub järglaste keskmises toodangus eakaaslastega võrreldes.
[https://www.epj.ee/assets/tekstid/sead/s_jk_kasiraamat_2022.pdf]

Geneetiline mudel

| Fenotüüp = Genotüüp + Keskkond

$$P = \underbrace{(\bar{G} + G)}_{\text{genotüübiväärtus}} + \underbrace{(\bar{E} + E)}_{\text{keskkonnaväärtus}} = \underbrace{(\bar{G} + \bar{E})}_{\text{keskmine fenotüübiväärtus}} + \underbrace{G}_{\text{genotüübiefekt}} + \underbrace{E}_{\text{keskkonnaefekt}} = \bar{P} + G + E$$

| Aga $G = A + D + I$

Üksikute alleelide summaarne efekt e **aretusväärtus** Dominantsi-efekt Epistaasi-efekt

Kui palju on loom keskmisest parem tänu oma geenidele

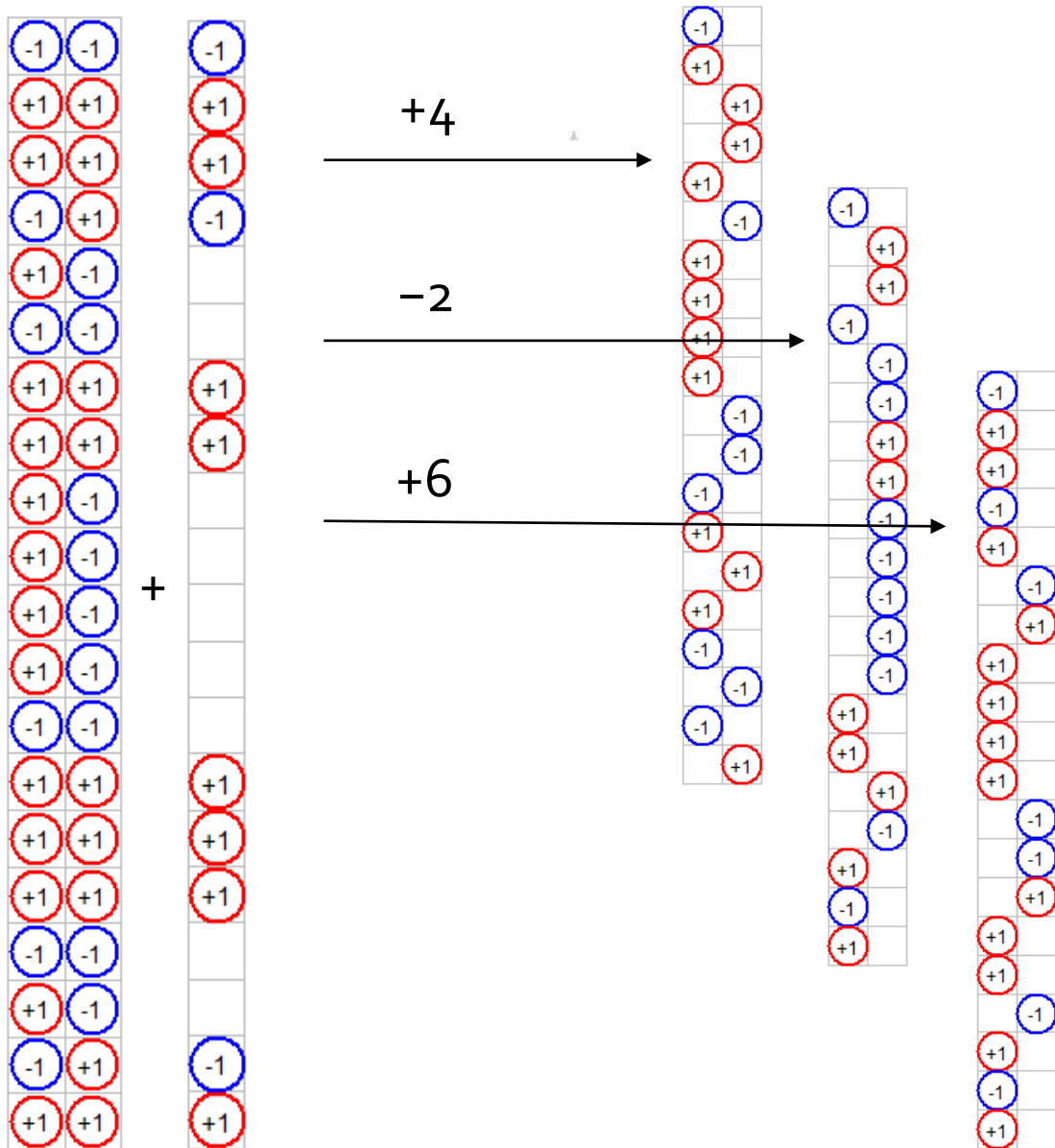
Kui palju on loom keskmisest parem tänu ümbritsevatele keskkonnale (söötmis-pidamistingimustele)

| Seega $P = \bar{P} + A + E^{(*)}$, kus $E^{(*)} (= E) = D + I + E$
hõlmab kõiki mitte-aditiivgeneetilisi efekte.

Miks vaid üksikute alleelide summaarne mõju?

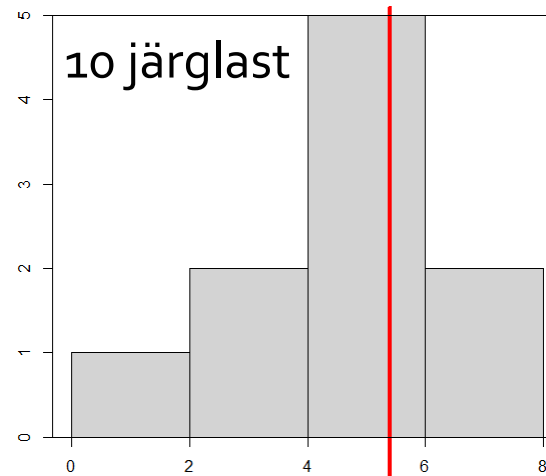
- | Sest vaid see mõjutab järglasi (kandub edasi järglastele).

Näide: 20 dialleelset lookust, igas lookuses alleeliefektid +1 või -1 pluss
dominantsiefektid.

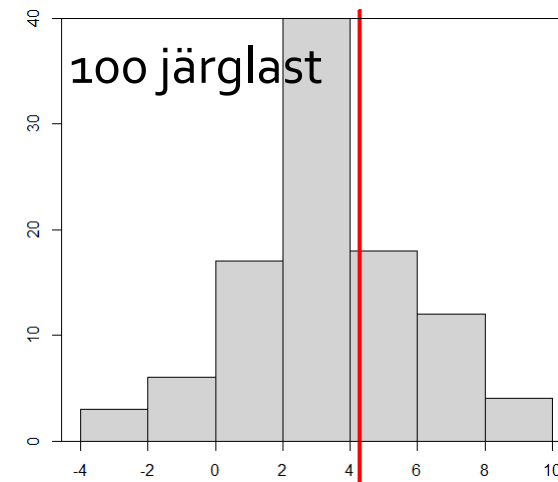


$$G = A + D = 8 + 5 = +13$$

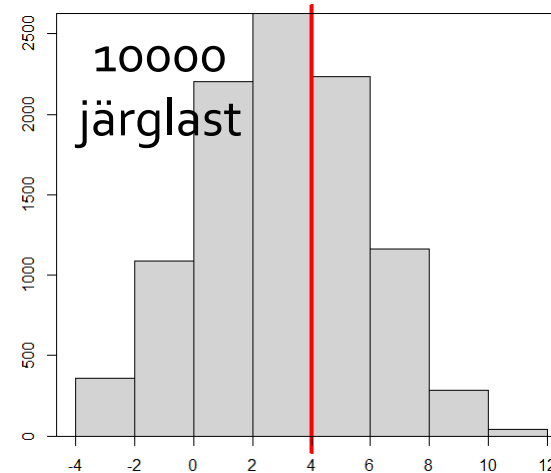
Keskmine = $5,4 \approx \frac{1}{2}A$?



Keskmine = $4,2 \approx \frac{1}{2}A$



Keskmine = $4,0 = \frac{1}{2}A$!



Geneetiline mudel

- | Aluseks on mudel (ettekujutus, visioon)

$$P = \bar{P} + A + E$$

- | Looma oodatav (keskmine) aretusväärtus avaldub kujul

$$E(A) = \frac{1}{2} A_{Isa} + \frac{1}{2} A_{Ema}$$

- | Looma tegelik aretusväärtus sõltub sellest, millised alleelid talle konkreetselt emalt ja isalt päranduvad ehk millistest meioosi protsessi käigus moodustunud sugurakkudest looma genotüüp kokku pannakse. Seda mõju nimetatakse Mendeli valiku (*MS*) mõjuks ning looma tegelik aretusväärtus avaldub kujul

$$A = \frac{1}{2} A_{Isa} + \frac{1}{2} A_{Ema} + MS$$

Geneetiline mudel

I Seega

$$\left. \begin{aligned} P &= \bar{P} + A + E \\ A &= \frac{1}{2} A_{I\text{sa}} + \frac{1}{2} A_{E\text{ma}} + MS \end{aligned} \right\} P_{\text{Järglane}} = \bar{P} + \frac{1}{2} A_{I\text{sa}} + \frac{1}{2} A_{E\text{ma}} + MS + E$$

I Kui aga isal on väga palju järglaseid, siis nende keskmine avaldub kujul

$$\bar{P}_{\text{Järglased}} = \bar{P} + \frac{1}{2} A_{I\text{sa}} + \frac{1}{2} \bar{A}_{E\text{ma}} + \bar{MS} + \bar{E} = \bar{P} + \frac{1}{2} A_{I\text{sa}} + 0$$

I Ja seega

$$A_{I\text{sa}} = 2 * (\bar{P}_{\text{Järglased}} - \bar{P})$$

Geneetiline varieeruvus ja päritavus

- | Aretusest ja aretusväärtusest on põhjust rääkida üksnes siis, kui loomade vahel eksisteerib nii fenotüübiline kui ka geneetiline varieeruvus (kui kõik loomad oleksid ühesugused, ei oleks ju kellegi vahel valida!).
- | Fenotüübilist varieeruvust mõõdab fenotüübidispersioon, mis on lahutatav aditiivgeneetiliseks ja nn keskkonnadispersiooniks (ingl. *variance*):

$$\text{var}(P) = \text{var}(A) + \text{var}(E)$$

- | **Päritavuskoeffitsient** või lihtsalt päritavus (ingl. *heritability*), näitab, kui suur osa populatsiooni fenotüübilisest erinevusest on tingitud päritavast geneetilisest (so aditiivgeneetilisest) erinevusest:

$$h^2 = \frac{\text{var}(A)}{\text{var}(P)}$$

Geneetiline varieeruvus ja päritavus

- Lihtsaim viis kasutada päritavuskoeffitsienti looma aretusväärtuse hindamisel on rakendada seda nn kaaluparameetrina, iseloomustamaks, millisel määral on looma fenotüübiväärtuse ja populatsiooni keskmise fenotüübiväärtuse erinevus tingitud üksikute alleelide summaarsest mõjust:

$$E(A) = h^2(P - \bar{P})$$

- Muide, valemist $A = \frac{1}{2}A_{Isa} + \frac{1}{2}A_{Ema} + MS$ järeldub ka, et

$$\text{var}(A) = \underbrace{\frac{1}{4}\text{var}(A_{Isa}) + \frac{1}{4}\text{var}(A_{Ema})}_{\frac{1}{2}\text{var}(A)} + \text{var}(MS),$$

ehk populatsiooni kogu aditiivgeneetilisest varieeruvusest on tervelt pool tingitud Mendeli valiku mõjust ja ei ole seega aretusega mõjutatav ...

Vahepala: aretaja võrrand

- I Aretaja võrrand (ingl. *breeders' equation*):

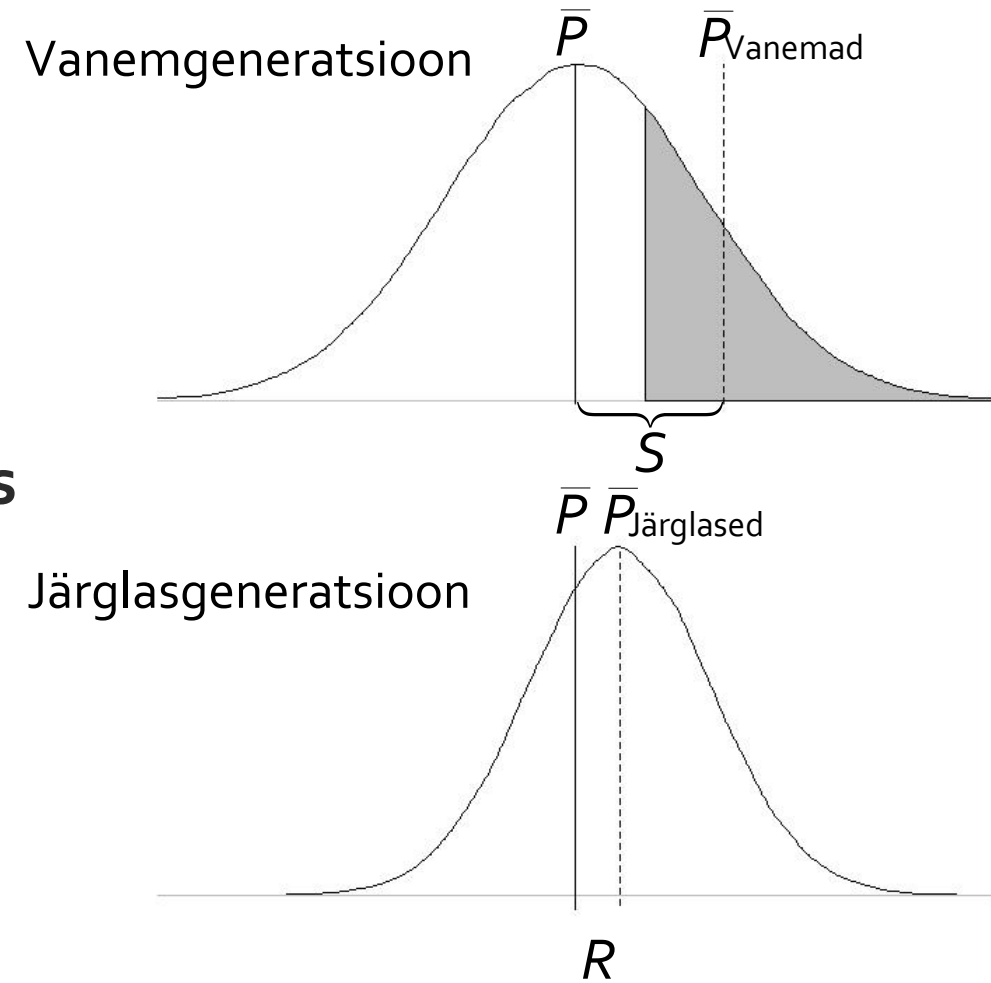
$$R = h^2 \times S,$$

kus $S = \bar{P}_{\text{Vanemad}} - \bar{P}$ on **selektsoonidiferents** ja $R = \bar{P}_{\text{Järglased}} - \bar{P}$ on **valikuedu** ehk **selektsooniefekt**.

- I Toodud võrrandi edasiarendus:

$$R = \frac{h^2 \times i \times \sqrt{\text{var}(P)}}{L},$$

kus i on **selektsooniintensiivsus** (normeeritud selektsoonidiferents), $\sqrt{\text{var}(P)}$ on fenotüübiline standardhälve ja L on generatsioonintervall.



Aretusväärtuse hindamine

- | Aluseks on mudel (ettekujutus, visioon)

$$P = \bar{P} + A + E$$

- | Definitsiooni kohaselt

$$A = 2 * (\bar{P}_{\text{Järglased}} - \bar{P})$$

- | Eeldused:

- | järglasi on hästi palju (teoreetiliselt lõpmatu palju);
- | teine vanem on valitud juhuslikult;
- | järglased elavad juhuslikes (st kõikvõimalikes) tingimustes – see garanteerib ka, et mistahes fenotüübil mõõdetavad keskmised erinevused saavad tuleneda üksnes geneetilistest erinevustest (erinevad söötmis-pidamistingimused rolli ei mängi);
- | järglastel on huvipakkuva tunnuse väärtused mõõdetud.

Aretusväärtuse hindamine: selektsiooniindeksid

- | Kui aga
 - | järglasi ei ole (mõõdetud) ja/või
 - | järglasi ei ole hästi palju?
- | Klassikaline lahendus: selektsiooniindeksid.

Aretusväärtuse hindamine: selektsiooniindeksid

- I **Selektsiooniindeks** on valem aretusväärtuse hinnangu $E(A)$ leidmiseks tavalise lineaarse regressioonivõrrandi kujul

$$E(A) = b * (P_x - \bar{P}),$$

kus

P_x tähistab indiviidi enese või tema sugulas(t)e (keskmist) fenotüübiväärtust, \bar{P} on populatsiooni keskmine ja

b on sobivalt valitud kaaluparameeter ehk regressioonikordaja (mis näitab aretusväärtuse muutust fenotüübilise erinevuse $P_x - \bar{P}$ muutumisel ühe ühiku võrra).

Aretusväärtuse hindamine: selektsiooniindeksid

- I Inglise keeles on selektsiooniindeksi kujul hinnatud aretusväärtuste tähistamiseks kasutusel lühend **BLP** (*Best Linear Predictor* ehk parim lineaarne prognoos):

$$\text{BLP}(A) = b * (P_x - \bar{P}).$$

- I Sõna „lineaarne“ märgib seda, et fenotüübilise erinevuse muutumisel muutub ka aretusväärtus ja seda lineaarselt.
- I Sõna „parim“ (*Best*) tuleneb aga saadud hinnangu matemaatilistest omadustest – tänu kordaja b hindamisele vähimruutude meetodil on saadud hinnangu varieeruvus vähim ja täpsus suurim.

Vahepala: regressioon

- Francis Galton uuris 19. sajandi lõpus laste ja nende vanemate pikkuse seost ja leidis, et

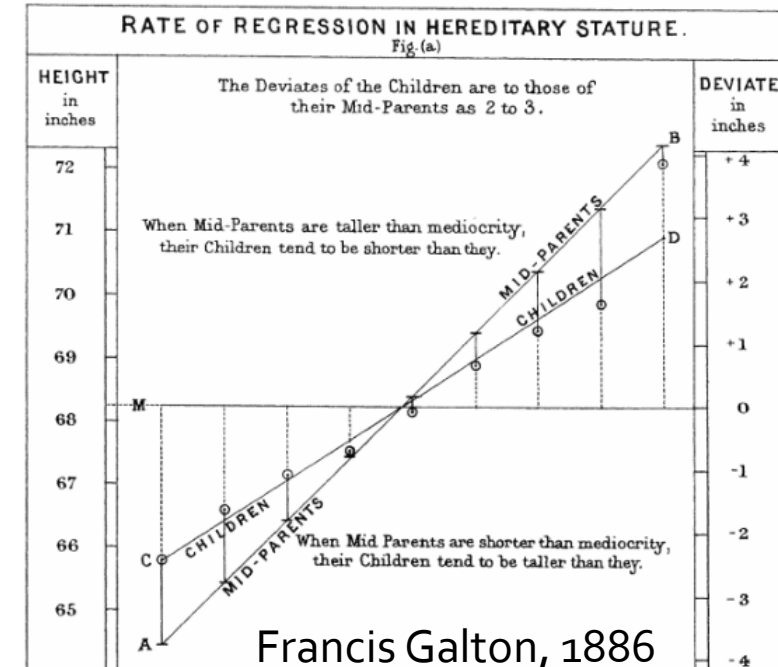
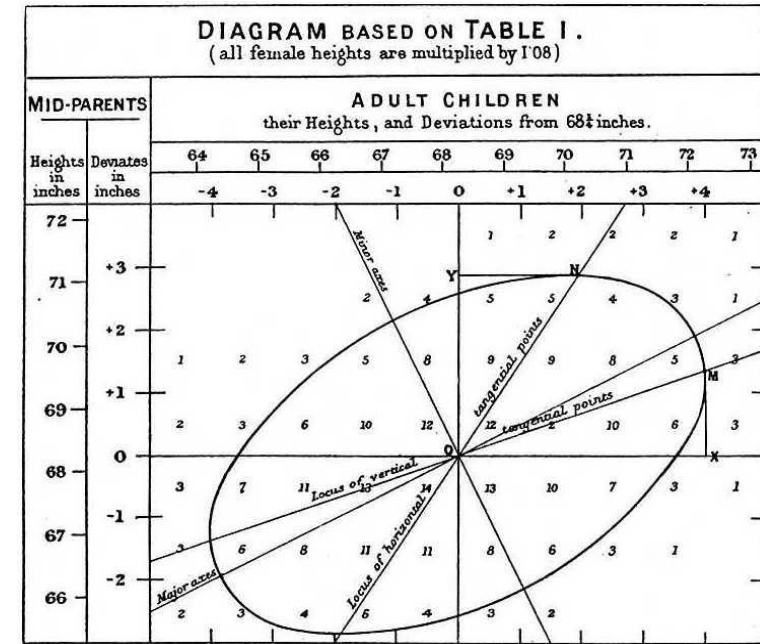
$$\bar{P}_{\text{Lapsed}} - \bar{P} = b * (\bar{P}_{\text{Vanemad}} - \bar{P}),$$

kus alati $b < 1$.

Ehk vanemate pikkuse erinevus keskmisest pikkusest ilmneb ka järglastes, aga see erinevus on väiksem – toimub järglaste keskmise pikkuse nõ tagasilangus (ingl. *regression*) üldkeskmise suunas.

- Tänapäeval me teame, et

- toodud valemis $b = h^2$,
- pikkuse erinevuse langus toimub põhjusel, et pikkus ei ole 100% ulatuses päritav,
- Galtoni tuletatud seose näol on tegu „aretaja võrrandiga“.

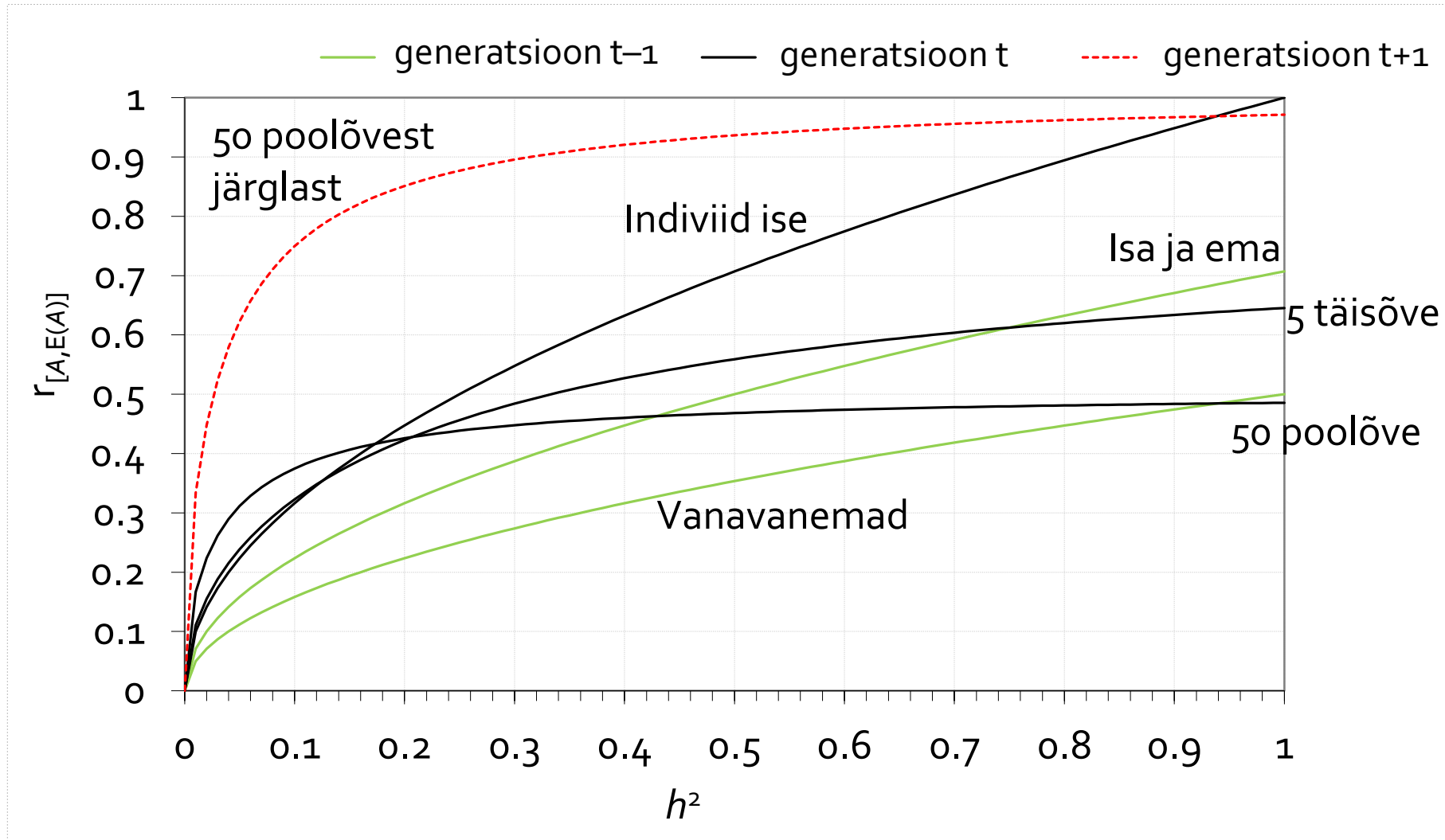


Aretusväärtuse hindamine: selektsiooniindeksid

$$E(A) = b * (P_X - \bar{P})$$

Informatsiooniallikas (kasutatav fenotüübiväärtus P_X)	Võrrandi kordaja b	AV-e hinnangu täpsus $r_{[A,E(A)]}$
Vanavanemad (4 väärtuse keskmine)	h^2	$\sqrt{\frac{1}{4}h^2}$
Vanemad (2 väärtuse keskmine)	h^2	$\sqrt{\frac{1}{2}h^2}$
Indiviid ise (1 väärtus)	h^2	$\sqrt{h^2}$
Indiviid ise (n korduva mõõtmise keskmine)	$\frac{nh^2}{1+(n-1)R}$	$\sqrt{\frac{nh^2}{1+(n-1)R}}$
Täisõved (n mõõtmise keskmine)	$\frac{\frac{1}{2}nh^2}{1+(n-1)(\frac{1}{2}h^2 + c_{FS}^2)}$	$\sqrt{\frac{\frac{1}{4}nh^2}{1+(n-1)(\frac{1}{2}h^2 + c_{FS}^2)}}$
Poolõved (n mõõtmise keskmine)	$\frac{\frac{1}{4}nh^2}{1+(n-1)(\frac{1}{4}h^2 + c_{HS}^2)}$	$\sqrt{\frac{\frac{1}{16}nh^2}{1+(n-1)(\frac{1}{4}h^2 + c_{HS}^2)}}$
Järglased (n mõõtmise keskmine)	$\frac{\frac{1}{2}nh^2}{1 + \frac{1}{4}h^2(n-1)}$	$\sqrt{\frac{\frac{1}{4}nh^2}{1 + \frac{1}{4}h^2(n-1)}}$

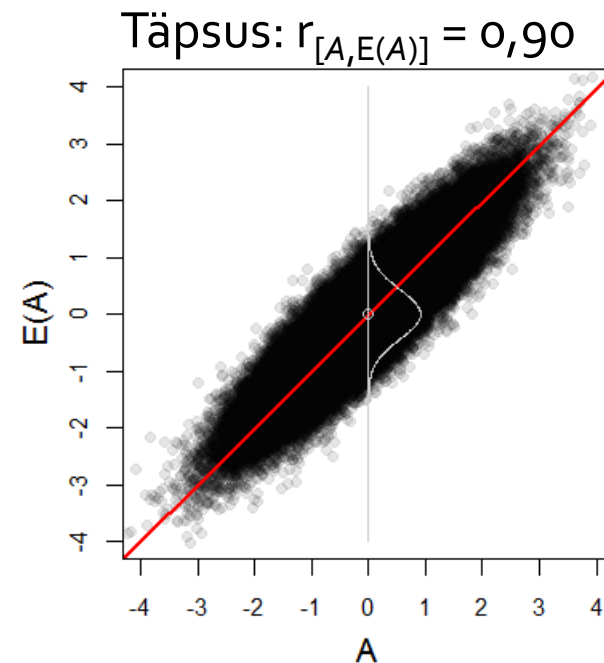
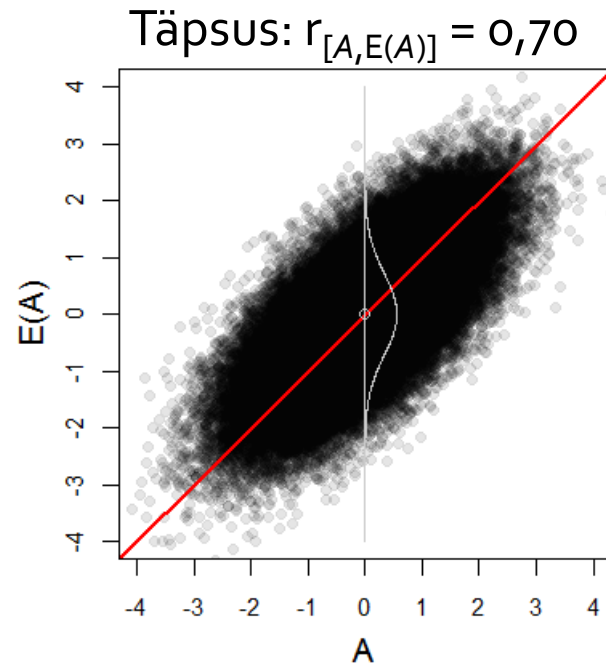
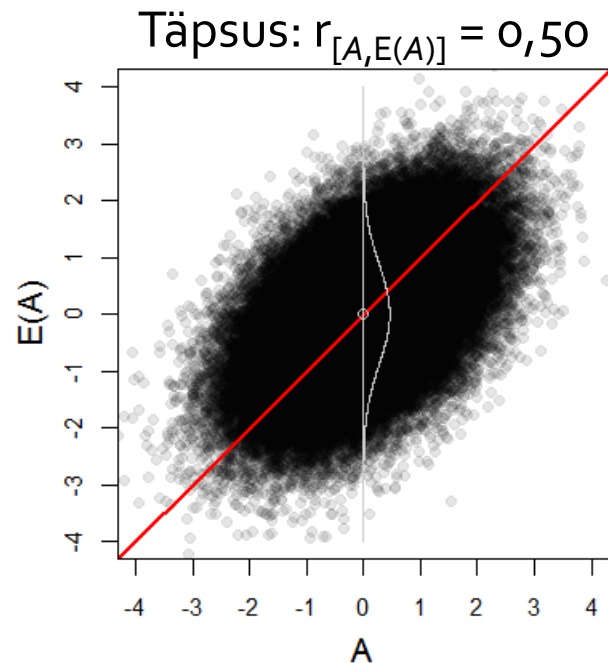
Aretusväärtuse hinnangu täpsus



$$E(A) = b * (P_x - \bar{P})$$

Vahepala: hinnangu täpsuse tähendus

- Mida see hinnangu täpsus, ehk korrelatsioon tegeliku ja hinnatud aretusväärtuse vahel endast kujutab?



Vahepala: kui järglasi on palju ...

- l Muide, ka valemist hindamaks aretusväärtust vanemale tema järglaste järgi:

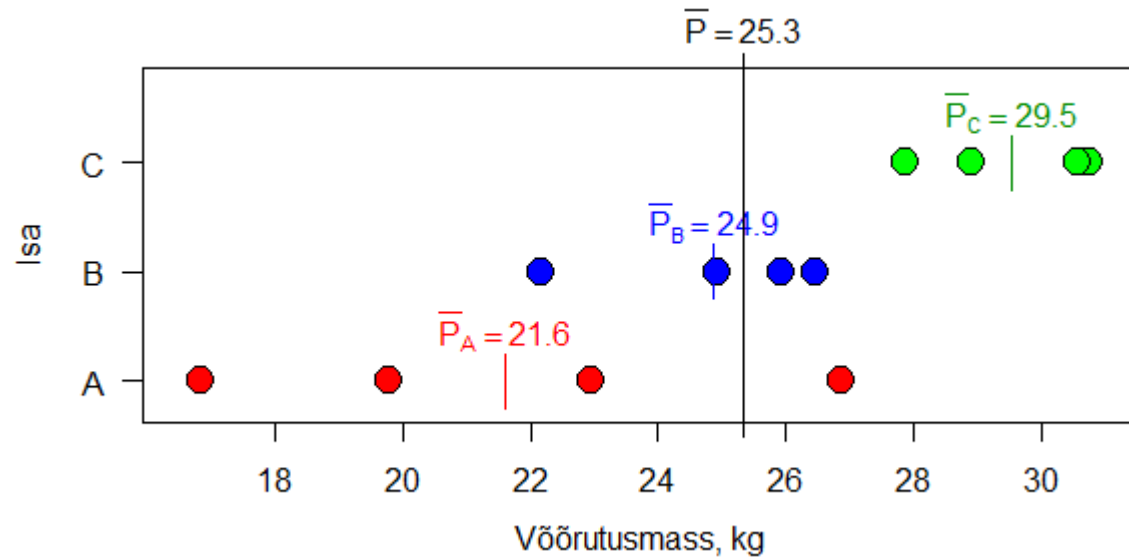
$$E(A) = \frac{\frac{1}{2}nh^2}{1 + \frac{1}{4}h^2(n-1)} * (\bar{P}_{\text{Järglased}} - \bar{P}),$$

järeldub järglaste arvu lõpmatusete lähenedes juba tuttav seos

$$A = 2 * (\bar{P}_{\text{Järglased}} - \bar{P}).$$

$$\text{Sest } \lim_{n \rightarrow \infty} \left[\frac{\frac{1}{2}nh^2}{1 + \frac{1}{4}h^2(n-1)} \right] = \frac{\frac{1}{2}h^2}{\frac{1}{4}h^2} = 2.$$

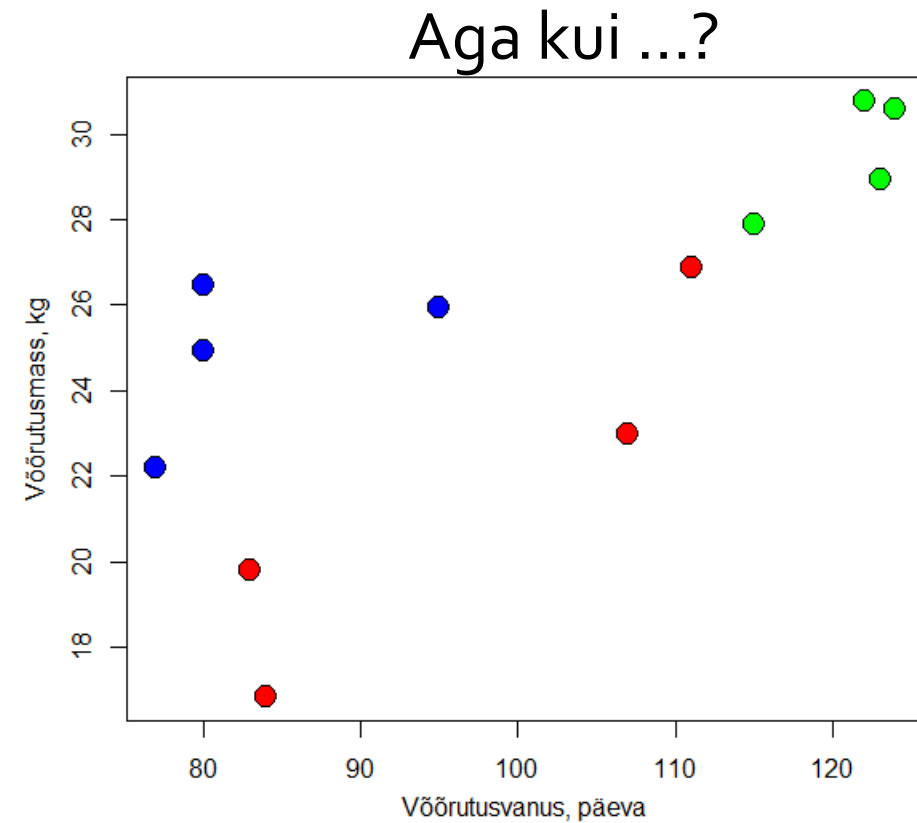
I Näide. 12 talle; uuritav tunnus: võõrutusmass; 3 isa järglased.



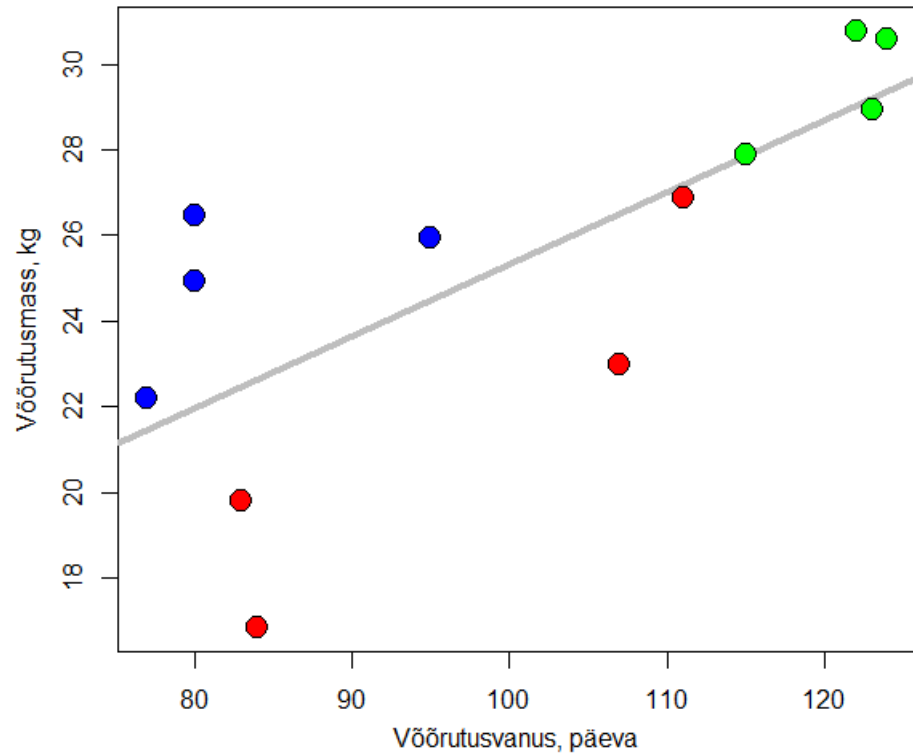
$$E(A_A) = 2 * (\bar{P}_A - \bar{P}) = 2 * (21,6 - 25,3) = 2 * (-3,7) = -7,4 \text{ kg}$$

$$E(A_B) = 2 * (\bar{P}_B - \bar{P}) = 2 * (24,9 - 25,3) = 2 * (-0,4) = -0,8 \text{ kg}$$

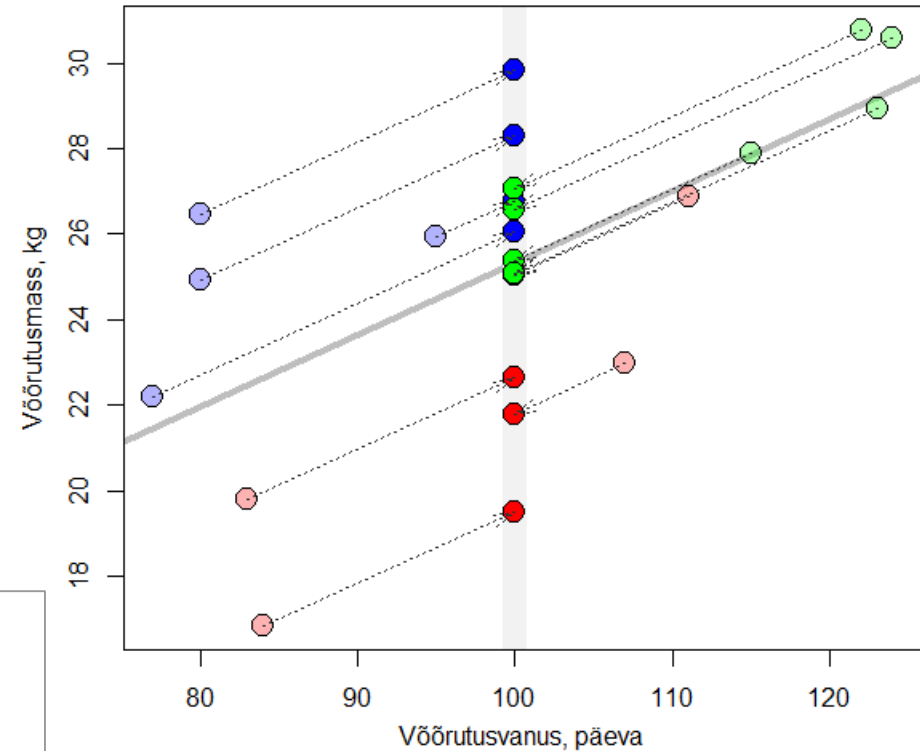
$$E(A_C) = 2 * (\bar{P}_C - \bar{P}) = 2 * (29,5 - 25,3) = 2 * 4,2 = 8,4 \text{ kg}$$



Hindame võõrutusmassi ja -vanuse vahelise seose



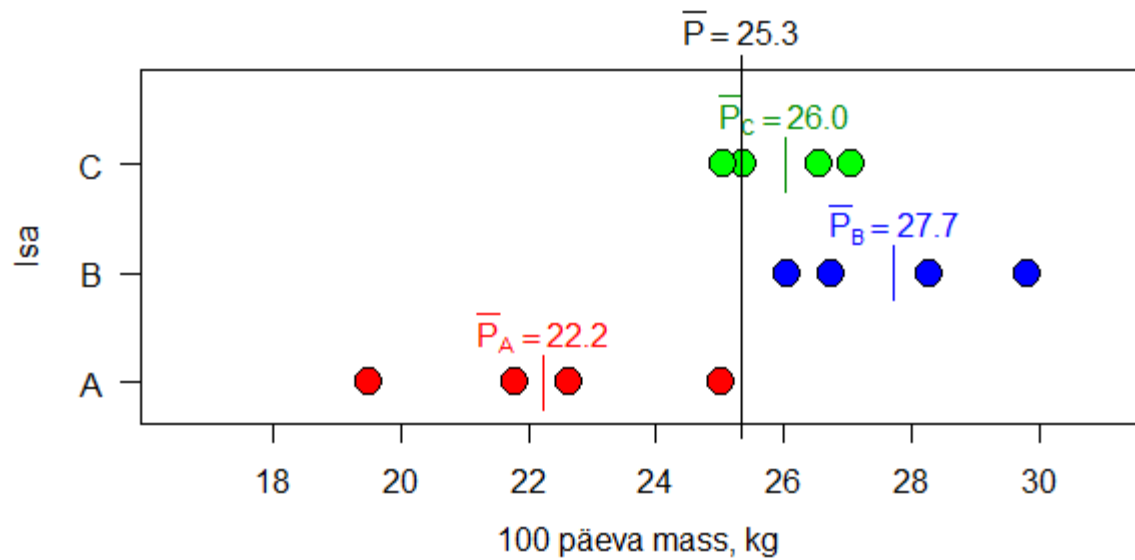
ja leiame igale tallele hinnangulise 100 päeva massi:
100 päeva mass
= $Võõrutusmass + 0,168 * (100 - Võõrutusvanus)$



```
> summary(lm(vmass ~ vvanus, data=vm.data))
```

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	8.54197	4.89488	1.745	0.11156	
vvanus	0.16791	0.04814	3.488	0.00584	**

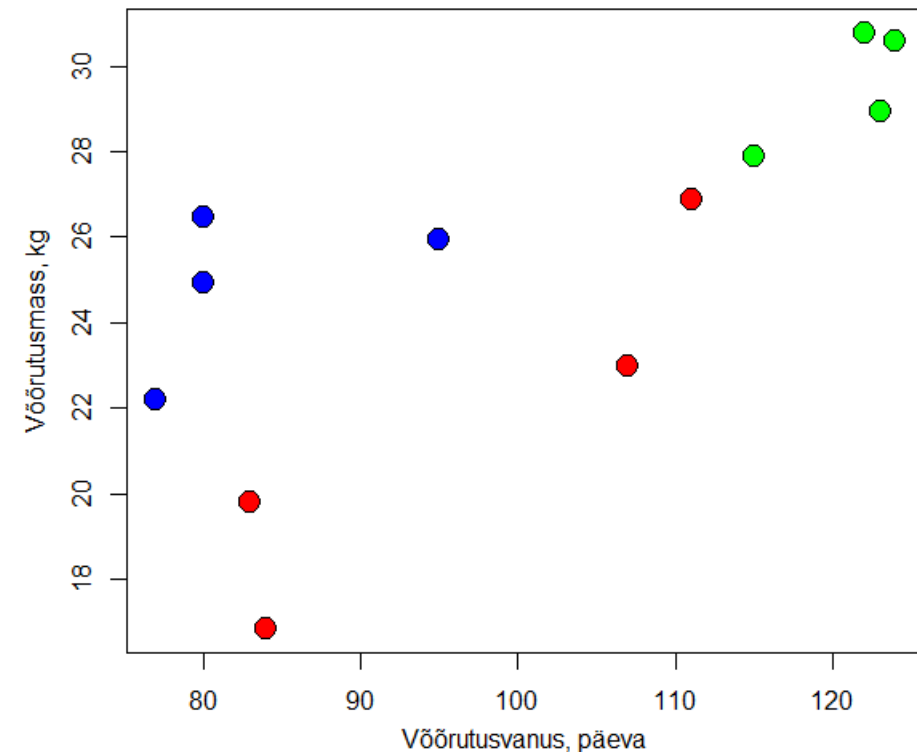


$$E(A_A) = 2 * (\bar{P}_A - \bar{P}) = 2 * (22,2 - 25,3) = 2 * (-3,1) = -6,2 \text{ kg}$$

$$E(A_B) = 2 * (\bar{P}_B - \bar{P}) = 2 * (27,7 - 25,3) = 2 * (2,4) = 4,8 \text{ kg}$$

$$E(A_C) = 2 * (\bar{P}_C - \bar{P}) = 2 * (26,0 - 25,3) = 2 * 0,7 = 1,4 \text{ kg}$$

Aga ka võõrutusvanus võib sõltuda isast ...



Lahenduseks on statistiline mudel, mis võtab samaaegselt arvesse nii võõrutusvanuste kui ka isade erinevuse.

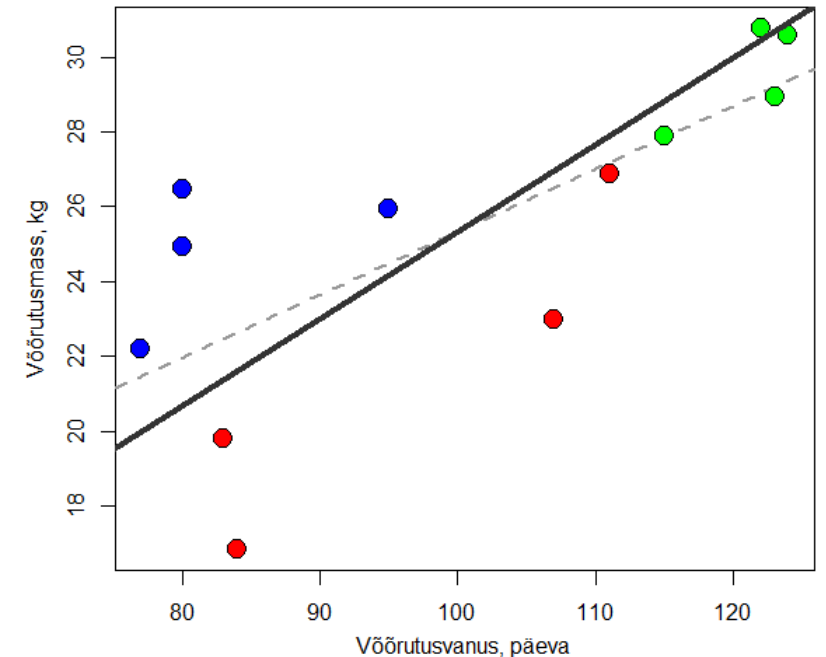
```
> summary(lm(vmass ~ vvanus + isa, data=vm.data))
```

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-0.85181	5.47278	-0.156	0.88017
vvanus	0.23340	0.05618	4.155	0.00319 **
isaB	6.36143	1.40884	4.515	0.00196 **
isaC	2.15619	1.83410	1.176	0.27355

```
> emmeans(lm(vmass ~ vvanus + isa, data=vm.data), ~isa,
at=list(vvanus=100))
```

isa	emmean	SE	df	lower.CL	upper.CL
A	22.5	0.872	8	20.5	24.5
B	28.8	1.276	8	25.9	31.8
C	24.6	1.452	8	21.3	28.0



Muutuvad nii võõrutusvanuse mõju kui ka isade hinnangulised keskmised.

$$E(A_A) = 2 * (22,5 - 25,3) = 2 * (-2,8) = -5,6 \text{ kg}$$

$$E(A_B) = 2 * (28,8 - 25,3) = 2 * (3,5) = 7,0 \text{ kg}$$

$$E(A_C) = 2 * (24,6 - 25,3) = 2 * (-0,7) = -1,4 \text{ kg}$$

Selektsiooniindeksi puudused

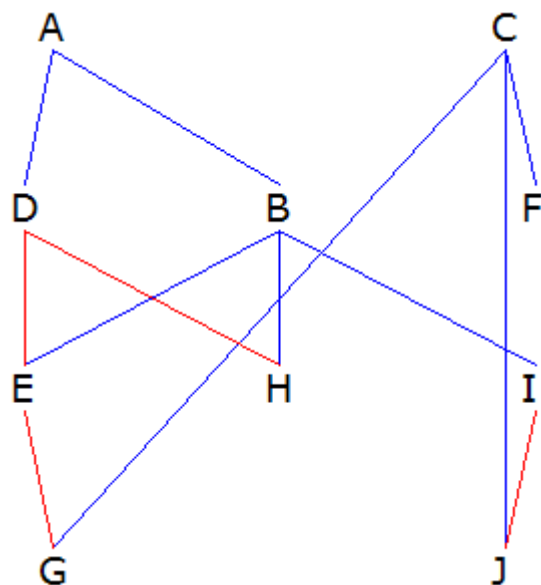
- | Selektiooniindeksite rakendamine suurtes populatsioonides on sageli pärsitud mitmel põhjusel.
 - | Esiteks on keerulisemate põlvnemisskeemide puhul tülikas ja vägagi töömahukas kõikvõimalikke sugulussidemeid arvestavate indeksite välja töötamine.
 - | Teiseks eeldab selektsiooniindeksite teooria kõigi mitte-aditiivgeneetiliste faktorite mõjude täpset teadmist ja mõõtmistulemuste nende suhtes korrigeerimist, mida on aga suuremahuliste, mitmeid aastaid ja loomakasvatustevõteteid hõlmavate uuringute korral pea võimatu teostada.
 - | Kolmandaks ei saa siiski alati lugeda selektsiooniindeksites kasutatavaid populatsiooni iseloomustavaid geneetilisi parameetreid teadaolevaiks, nende väärtuste hindamiseks selektsiooniindeksid aga enamasti ei sobi.

BLUP-mudel

- | **BLUP-mudel** (ka BLUP-meetod) on **üldine lineaarne segamudel** (ingl. *General Linear Mixed Model*, GLMM), mis võimaldab samaaegselt leida nii
 - | keskkonnamõjude parimad lineaarsed nihketa hinnangud (*Best Linear Unbiased Estimate*, BLUE) kui ka
 - | aretusväärtuste parimad lineaarsed nihketa prognoosid (*Best Linear Unbiased Predictor*, BLUP).
- | Sõna „parim“ (*Best*) tähendab nagu selektsiooniindeksite puhulgi, et saadud hinnangud ja prognoosid täpseimad (vähima varieeruvusega).
- | Sõna „nihketa“ (*Unbiased*) märgib aga seda, et tulemused ei ole korrektse mudeli puhul kuhugi suunas kallutatud (on keskmiselt õiged).

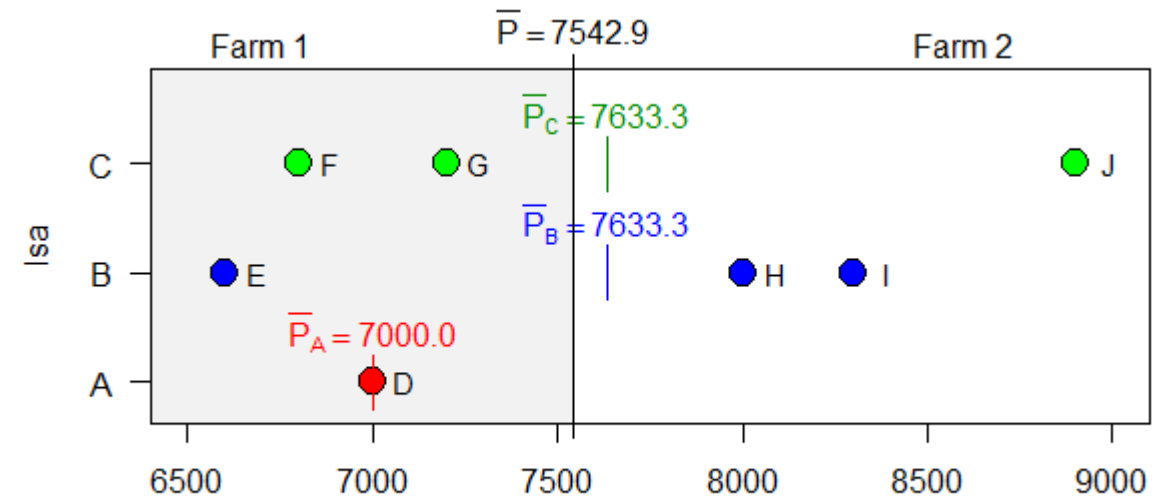
BLUP looma mudel

- BLUP looma mudel võimaldab võtta arvesse kõigi loomade omavahelise geneetilise suguluse.



	A	B	C	D	E	F	G	H	I	J
A	1	0.5	0	0.5	0.5	0	0.25	0.5	0.25	0.25
B	0.5	1	0	0.25	0.625	0	0.3125	0.625	0.5	0.3125
C	0	0	1	0	0	0.5	0.5	0	0	0.5
D	0.5	0.25	0	1	0.625	0	0.3125	0.625	0.125	0.3125
E	0.5	0.625	0	0.625	1.125	0	0.5625	0.625	0.3125	0.3125
F	0	0	0.5	0	0	1	0.25	0	0	0.25
G	0.25	0.3125	0.5	0.3125	0.5625	0.25	1	0.3125	0.15625	0.40625
H	0.5	0.625	0	0.625	0.625	0	0.3125	1.125	0.3125	0.5625
I	0.25	0.5	0	0.125	0.3125	0	0.15625	0.3125	1	0.15625
J	0.25	0.3125	0.5	0.3125	0.3125	0.25	0.40625	0.5625	0.15625	1

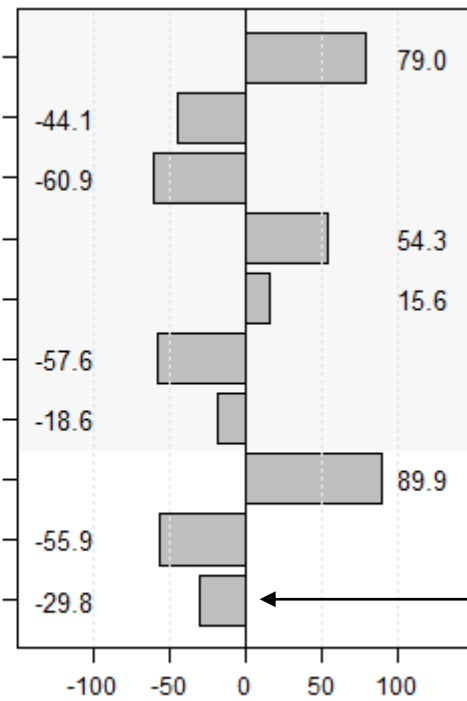
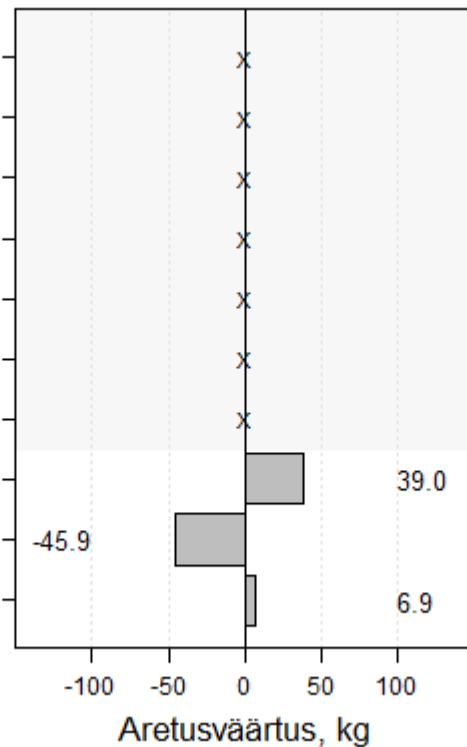
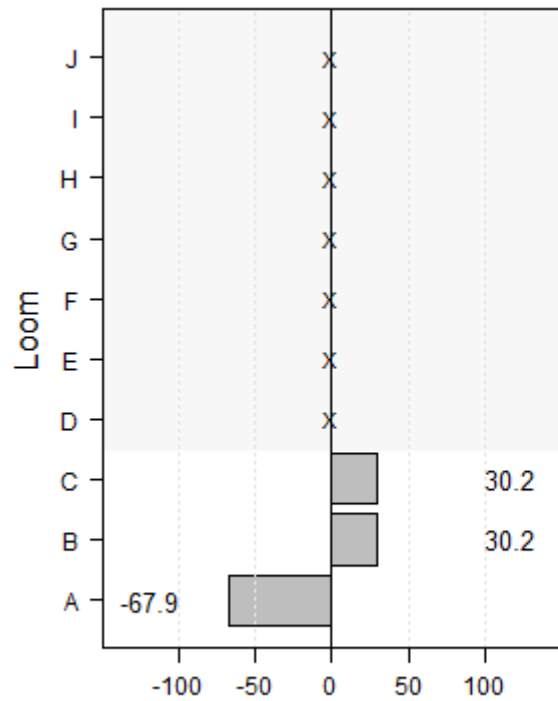
Näide. Uuritav tunnus: piimatoodang, 2 farmi, 7 lehma (3 isa järglased + sugupuud).



Selektsiooniindeks

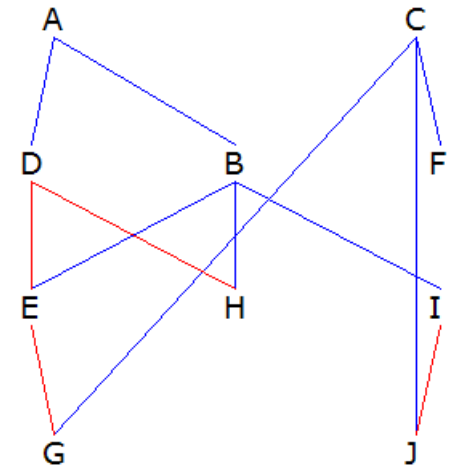
BLUP isa mudel

BLUP looma mudel



$$h^2 = 0,25$$

Piim, kg



Võetakse näiteks arvesse seda, et A on B isa.

BLUP-mudel – mittegeneetilised tegurid

- | Võimalik on arvestada vaid nende teguritega, mis on registreeritud.
- | Mõtet on arvestada vaid nende teguritega, mis uuritavat tunnust mõjutavad.
- | Klassika: *HYS*-efekt (*Herd-Year-Season*, karja-aasta-sesooni efekt) + igale tunnusele tema olemusest tulenevad täiendavad mõjutegurid.

6.5 Sigivustunnuste geneetiline hindamine (lehma sigivusandmete alusel)

Sigivustunnuste geneetisel hindamisel kasutatakse mitme tunnusega BLUP-loomamudelit. Hindamismudelis sisalduvad lisaks geneetilisele efektile ja lehma alalisele keskkonnaefektile järgmised efektid ja koosmõjud:

- sünniaasta
- tõug
- poegimise aasta*kuu
- kari*seemendusaasta
- laktatsioon*taastumisperioodi pikkus
- laktatsioon*poegimisvanus
- seemendustehnik



Aretusväärtuste skaala ja ulatus

- | Aretusväärtused hinnatakse mõõdetud tunnuse ühikutes ja skaalal.
- | Keskmise aretusväärtus on 0 (null) ja varieeruvus sõltub nii mõõdetud tunnuse väärtuste varieeruvusest kui ka geneetilisest determineeritusest.
- | See, kas „hea“ on nullist suurem, nullist väiksem või hoopis nulliga võrduv aretusväärtus, sõltub mõõdetud tunnusest:
 - | piimatoodangu (kg), rasvaprotsendi (%), kasvukiiruse (kg) jne puhul on „mida suurem, seda parem“, ehk hea on positiivne aretusväärtus,
 - | surnultsünni või mõne haiguse riski puhul aga „mida väiksem, seda parem“, ehk hea on negatiivne aretusväärtus,
 - | paljude välimikutunnuste puhul on aga „parim“ (optimaalseim) keskmine väärtus, ehk nulliline aretusväärtus.

Suhtelised aretusväärtused

- | **Suhtelised aretusväärtused** on ühesugusele (sama keskmise, varieeruvuse ja suunaga) punktiskaalale teisendatud aretusväärtused.
- | Nad võimaldavad lihtsamalt kõrvutada erinevate loomade geneetilist potentsiaali erinevate tunnuste osas.
- | Samas on üksnes suhteliste aretusväärtuste alusel sageli raske otsustada looma sisulise paremuse või halvemuse üle.
- | Suhtelised aretusväärtused (SAV) leitakse valemiga

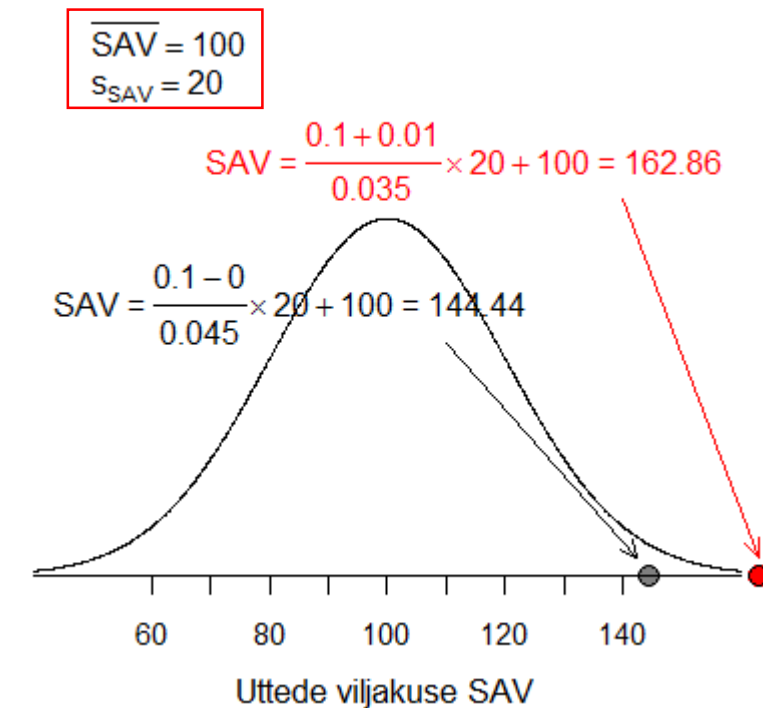
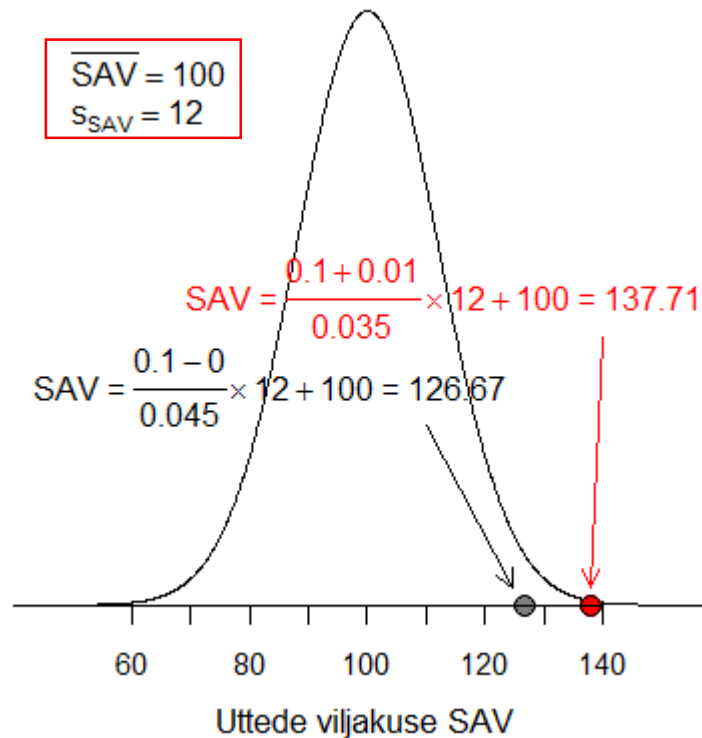
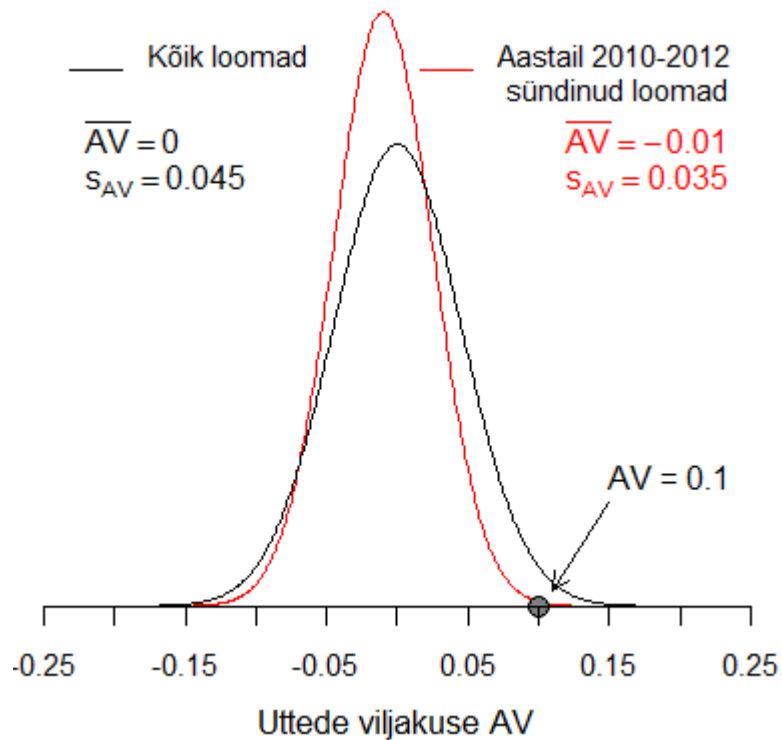
$$SAV = \frac{AV - \overline{AV}}{s_{SAV}} \times s_{SAV} + \overline{SAV}$$

kus AV , \overline{AV} ja s_{SAV} on vastavalt algne aretusväärtus ning selle keskmine ja standardhälve (**baas**)populatsioonis, ning SAV , \overline{SAV} ja s_{SAV} on suhteline aretusväärtus ning selle ette antud keskmine ja standardhälve (baas)populatsioonis.

Suhtelised aretusväärtused

$$SAV = \frac{AV - \overline{AV}}{S_{SAV}} \times S_{SAV} + \overline{SAV}$$

- | Suhtelise aretusväärtuse väärtus sõltub baaspopulatsiooni valikust!
- | Suhtelise aretusväärtuse väärtus sõltub ette antud keskmisest ja standardhälbest!



Suhtelised üldaretusväärtused e aretusindeksid

- | **Aretusindeksid** leitakse summeerimaks üksikuid (suhtelisi) aretusväärtusi.
- | See, milliste (suhteliste) aretusväärtuste baasil üldine indeks arvutada, nagu ka see, milliste kaaludega üldises indeksis üksikuid aretusväärtuseid arvestada, on kokkuleppe küsimus.
- | Indeksisse kuuluvate aretusväärtuste ja nende kaalude üle otsustamisel võib lähtuda nii sisulisest sarnasusest ja olulisusest kui ka majanduslikust väärtusest.
- | Majanduslikel väärtustel (1 kg piimarasva, 1 kg võõrutusmassi, ühe seemenduse, ühe talle jne hind) baseeruvate kaaludega indekseid nimetatakse ka **kasumiindeksiteks** ja nende väärtuste arvutatakse rahalistes ühikutes.

Suhtelised üldaretusväärtused e aretusindeksid

- | Aretusindeksid muutuvad ajas:
 - | lisandub uusi mõõdetavaid tunnuseid,
 - | aretuseesmärgid teisenevad,

| ning on erinevates populatsioonides erinevad.

Tunnused / indeksid Eestis	Nimetus	Tunnused / indeksid Põhjamaades
* Suhteline koguaretusväärtus ($SKAV = 0,5 * SPAV + 0,25 * SSAV + 0,25 * SVAV$)	Net Total Merit Index, NTM	$0,63 * \text{Toodang} + 0,29 * \text{Funktsionaalsus} + 0,06 * \text{Kehaehitus} + 0,02 * \text{Söödakasutus}$ (valemis olevad tunnustegruppide aretusväärtused arvutatakse kokku enam kui 15 üksiktunnuse baasil)
Toodang		
* Piimatoodang, kg	Milk Index	$0,30 * EBV_{\text{milk1}} + 0,25 * EBV_{\text{milk2}} + 0,45 * EBV_{\text{milk3}}$
* Rasvatoodang, kg	Fat Index	$0,30 * EBV_{\text{fat1}} + 0,25 * EBV_{\text{fat2}} + 0,45 * EBV_{\text{fat3}}$
* Valgutoodang, kg	Protein Index	$0,30 * EBV_{\text{protein1}} + 0,25 * EBV_{\text{protein2}} + 0,45 * EBV_{\text{protein3}}$ (alaindeksid näitavaid erinevate laktatsioonide aretusväärtuseid)
* Suhteline piimajõudluse üldaretusväärtus ($SPAV = 0,1 * AV_{\text{Rasv, kg}} + 0,4 * AV_{\text{Valk, kg}}$)	Yield	$-0,25 * \text{Milk Index} + 0,55 * \text{Fat Index} + 0,70 * \text{Protein Index}$
Somaatiliste rakkude arv / udara tervis		
* Suhteline udara tervise üldaretusväärtus (SSAV)	Udder health	$0,15 * CM11 + 0,15 * CM12 + 0,25 * CM2 + 0,45 * CM3$; CM11 ja CM12 – kliinilise mastiidi esinemine -15. kuni 50. päeval ja 51. kuni 305. päeval esimesest poegimisest, CM2 ja CM3 – kliinilise mastiidi esinemine -15. kuni 150. päeval teisest ja kolmandast poegimisest.

Vahepala: standardhälve

- | Mõõdab varieeruvust, erinevust, mitmekesisust.
- | Näitab, kui palju keskmiselt erinevad üksikud väärtused keskmisest.
- | Normaaljaotuse järgi jaotuvate suuruste (kõik aretusväärtused ja -indeksid, paljud jõudlus- ja välimikunäitajad) korral
 - | ~68,3% väärtustest jäävad vahemikku keskmine \pm standardhälve;
 - | ~95,4% vahemikku keskmine ± 2 *standardhälve;
 - | ~99,7% vahemikku keskmine ± 3 *standardhälve.

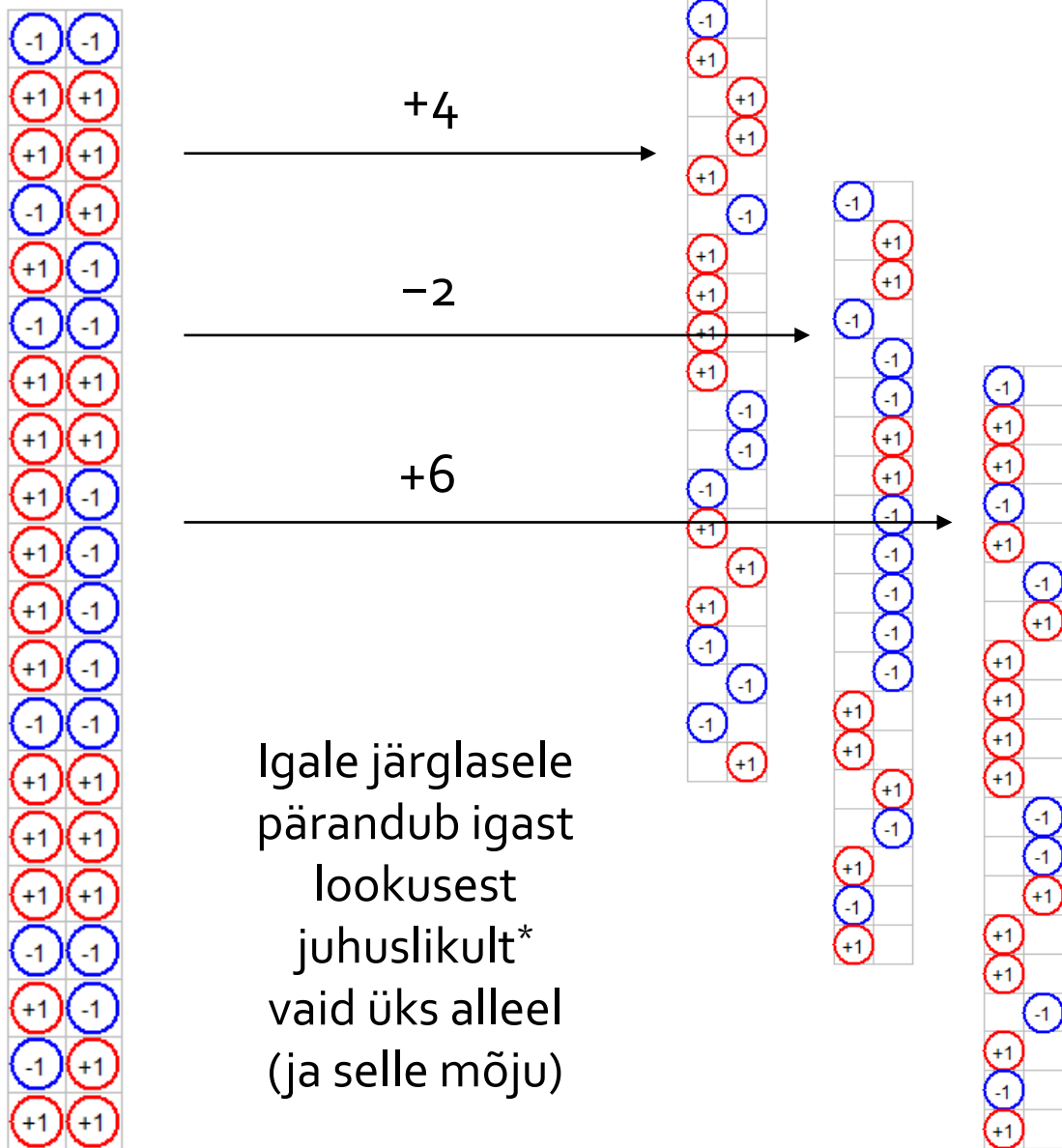
Näiteks SPAV $\sim N(100, 12)$. Seetõttu jääb

- 68,3% loomade SPAV vahemikku 88-112 punkti ja
- 95,4% loomade SPAV vahemikku 76-124 punkti.

Vahepala: normaaljaotus

- I Tekib siis, kui mõõdetava-hinnatava suuruse väärtused on kujunenud väga suure hulga väikeste tegurite summeerumise tulemusena (tsentraalne piirteoreem ehk Ljapunovi teoreem).

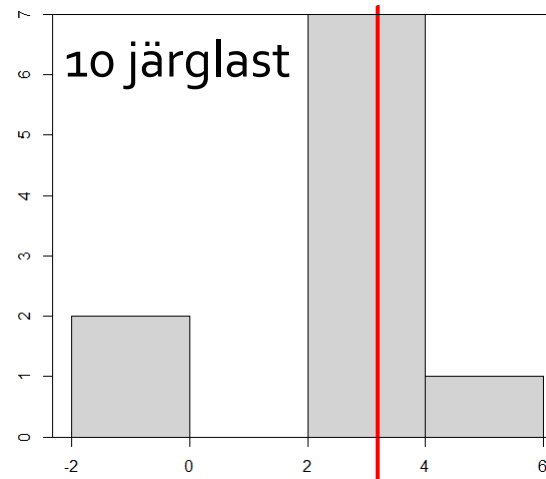
Näide: 20 dialleelset lookust, igas lookuses vaid alleeliefektid +1 või -1.



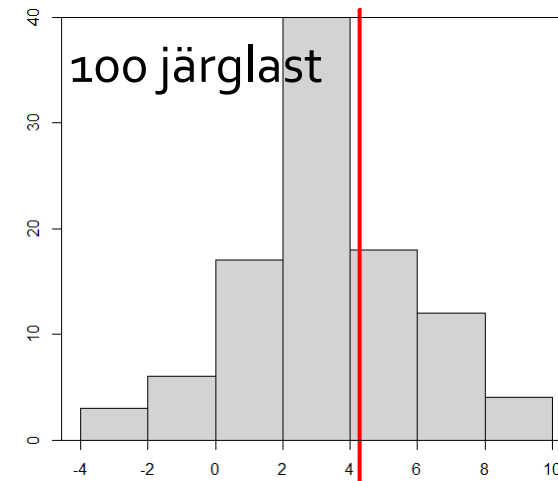
Igale järglasele pärandub igast lookusest juhuslikult* vaid üks alleel (ja selle mõju)

$$G = A = +8$$

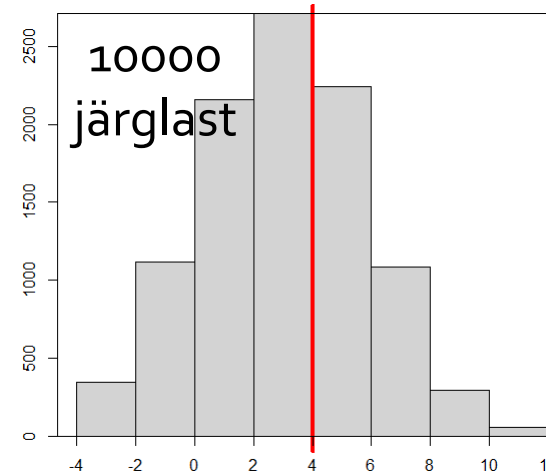
Keskmine = 3,2 \approx $\frac{1}{2}A$?



Keskmine = 4,2 \approx $\frac{1}{2}A$



Keskmine = 4,0 = $\frac{1}{2}A$!



Vahepala: normaaljaotus

- | Tekib siis, kui mõõdetava-hinnatava suuruse väärtused on kujunenud väga suure hulga väikeste tegurite summeerumise tulemusena (tsentraalne piirteoreem ehk Ljapunovi teoreem).
- | Kõik aretusväärtused ja -indeksid, geneetilised riskiskoorid ja majanduslikud indeksid jaotuvad normaaljaotuse järgi.
- | Sellest lähtuvalt saab isegi Exceli abil leida mitmesuguseid piirväärtuseid ja/või tõenäosuseid.

Vahepala: standardhälve ja normaaljaotus

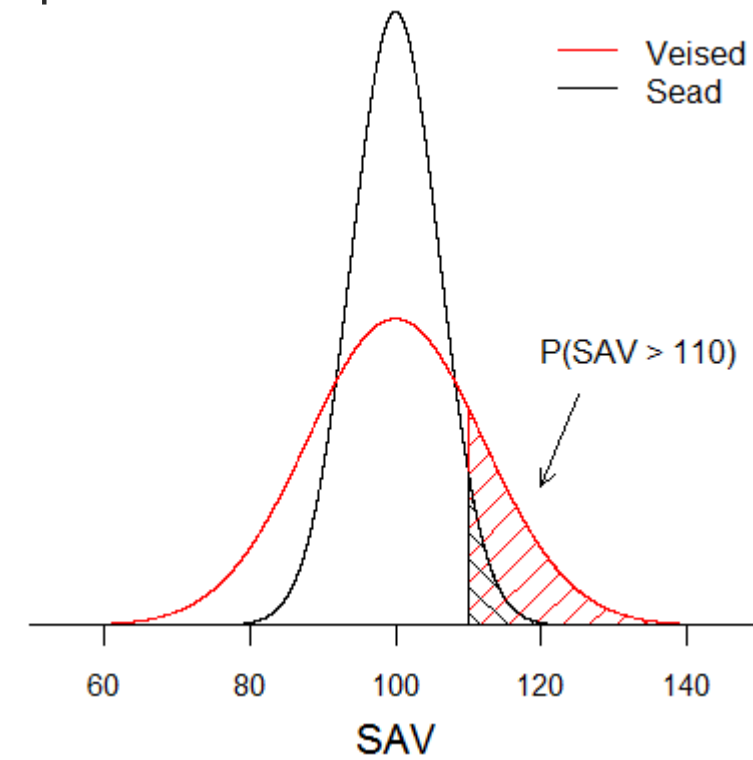
- | Erinevate liikide-tõugude-aretusorganisatsioonide puhul võivad aretusindeksite keskmised ja standardhälbed olla erinevad.
- | Näiteks Eestis on
 - | piima- ja lihaveistel aretusindeksite keskmiseks 100 ja standardhällbeks 12 punkti,
 - | sigade aretusindeksite keskmiseks 100 ja standardhällbeks 6 punkti.

Sestap on näiteks 110-st punktist kõrgema suhtelise aretusväärtusega loomade hulk sigade ja veiste puhul erinev:

- sigadel $P(\text{SAV} > 110) = 0,048$ (so 4,8%)
- veistel $P(\text{SAV} > 110) = 0,202$ (so 20,2%)

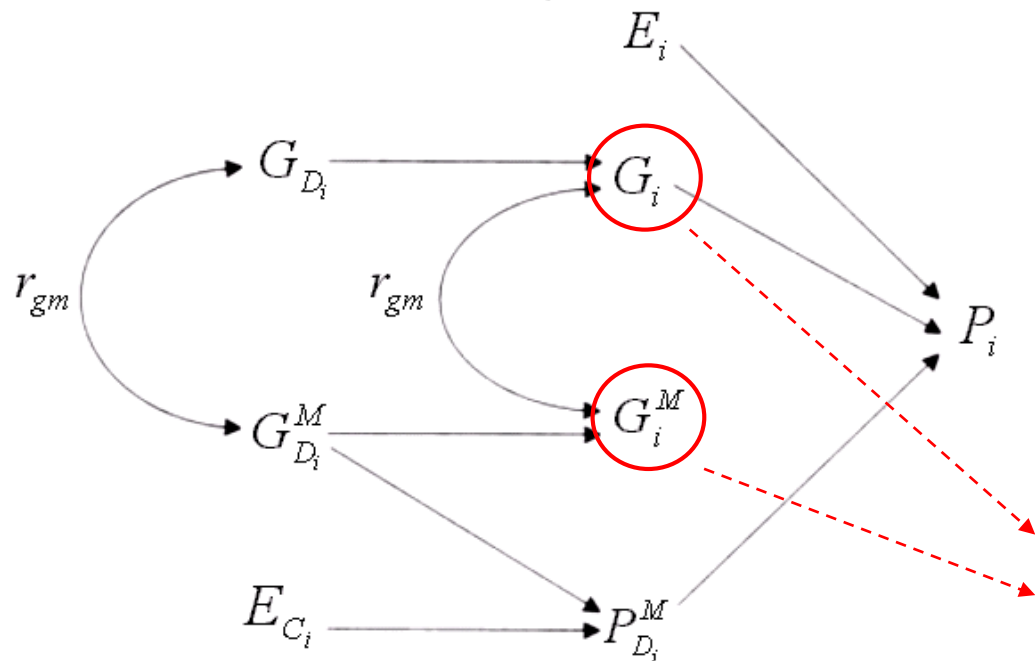
PS. Need tõenäosused on leitavad ka Excelis, viimane näiteks valemiga

$$= 1 - \text{NORM.DIST}(110, 100, 12, 1)$$



„Üks tunnus“, mitu aretusväärtust

- ▮ Näiteks **emapoolse efektiga** mudelid – enamuse noorloomadel mõõdetud tunnuste puhul mängib rolli nii
 - ▮ noorlooma enese geneetiline potentsiaal – näiteks võime kiirelt kasvada – kui ka
 - ▮ tema ema geneetiline potentsiaal kasvu mõjutada (näiteks läbi piimakuse).



 Eesti
Põllumajandusloomade
Jõudluskontroll

Poegimine

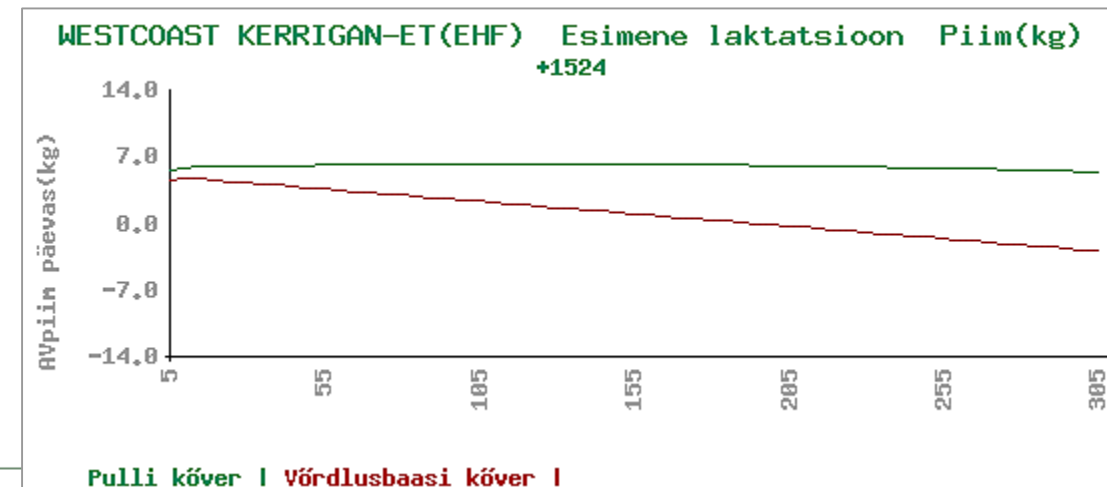
Pull	Loomade arv	Karjade arv	Usalduskoef.	EHF	
				Poegimiskergus	Elusünd
Vasika isana	193	14	0.93	93	93
Lehma isana	74	10	0.78	145	133

„Üks tunnus“, mitu aretusväärtust

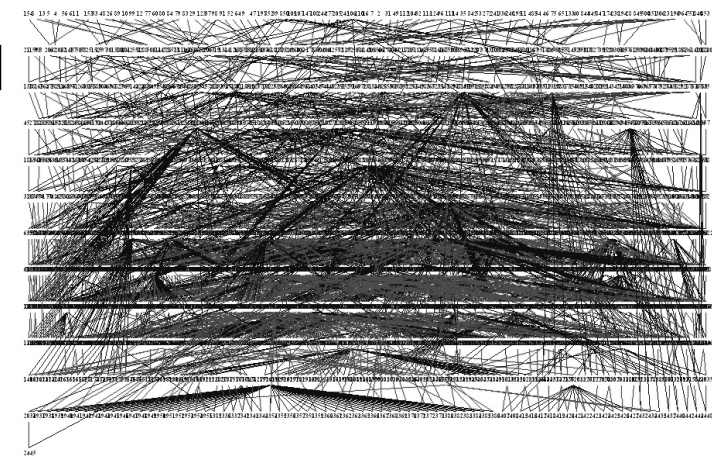
- | Korduvalt mõõdetud tunnused, näiteks piimatoodang, kehamass jne
- | Kas hinnata aretusväärtust
 - | summaarsele/keskmisele väärtusele või
 - | väärtuste dünaamikale-muutumisele?

Piimajoudlus (EHF)

SPAV	152	Aretusväärtused				
Usalduskoefitsient	0.90	Piim +1809	Rasva% -0.09	Rasv +57	Valgu% + 0.07	Valk +70
Tütarde arv	76	70 tütre 1. laktatsiooni keskmine toodang				
Karjade arv	8	Piim 11396	Rasva% 3.65	Rasv 416	Valgu% 3.31	Valk 377



„Random regression“ mudel

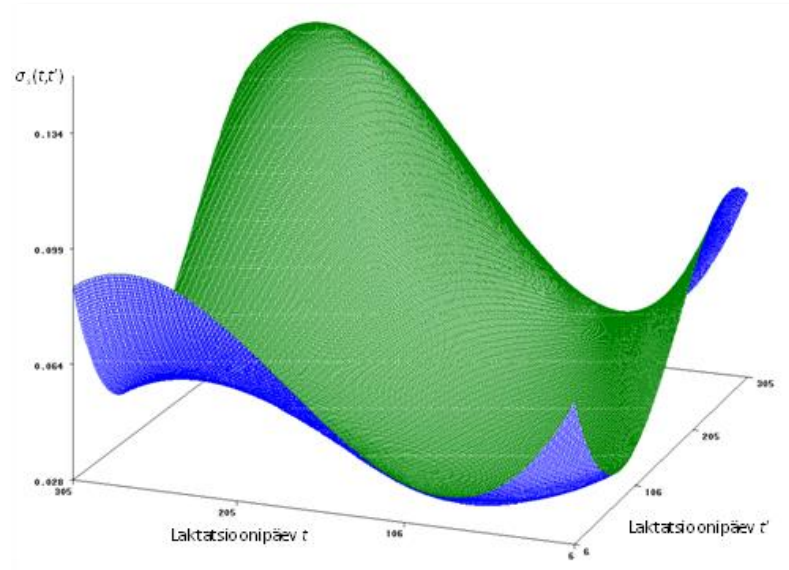
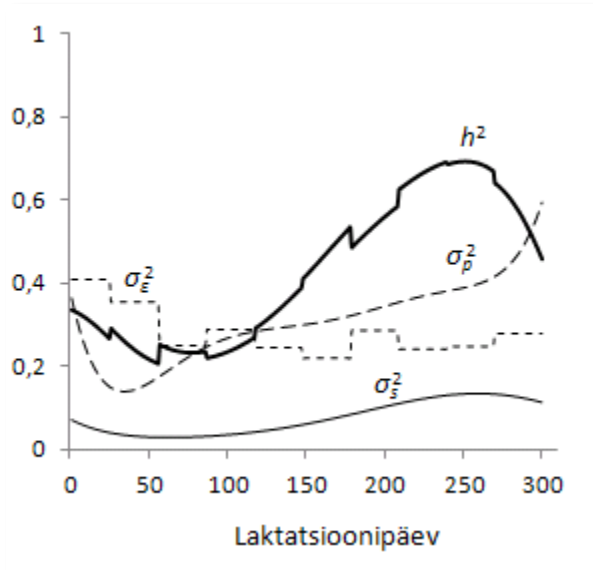


Kontroll-lüpside andmed ----->

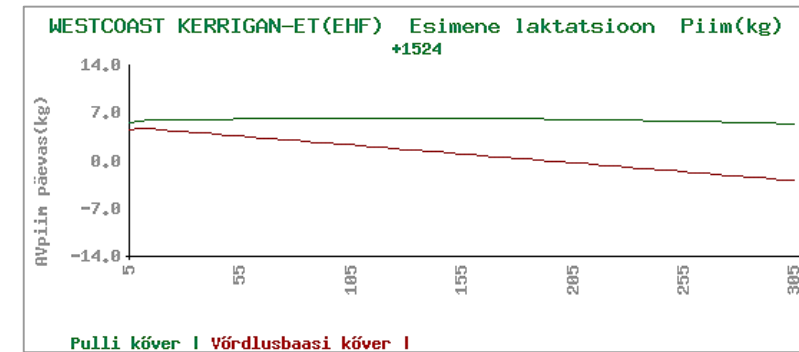
```

Ka={ 3.41982118 0.13816859 -0.42588716,
      0.07357666 0.79272650 0.07917200,
      -0.41973488 0.04679669 0.26542990};
Kp={ 2.09347304 -0.11831324 -0.12844095,
      0.22607809 0.46480827 0.05246716,
      -0.20639162 0.04065760 0.15288249};
ve=5.706000;
    
```

$$Cov[y_{ijkl}(t), y_{ijkl}(t')] = \sum_{m=0}^{np-1} \sum_{n=0}^{np-1} \varphi_m(t) \varphi_n(t') Cov(a_{km}, a_{kn}) + \sum_{m=0}^{np-1} \sum_{n=0}^{np-1} \varphi_m(t) \varphi_n(t') Cov(p_{lm}, p_{ln}) + Cov[\varepsilon_{ijkl}(t), \varepsilon_{ijkl}(t')]$$

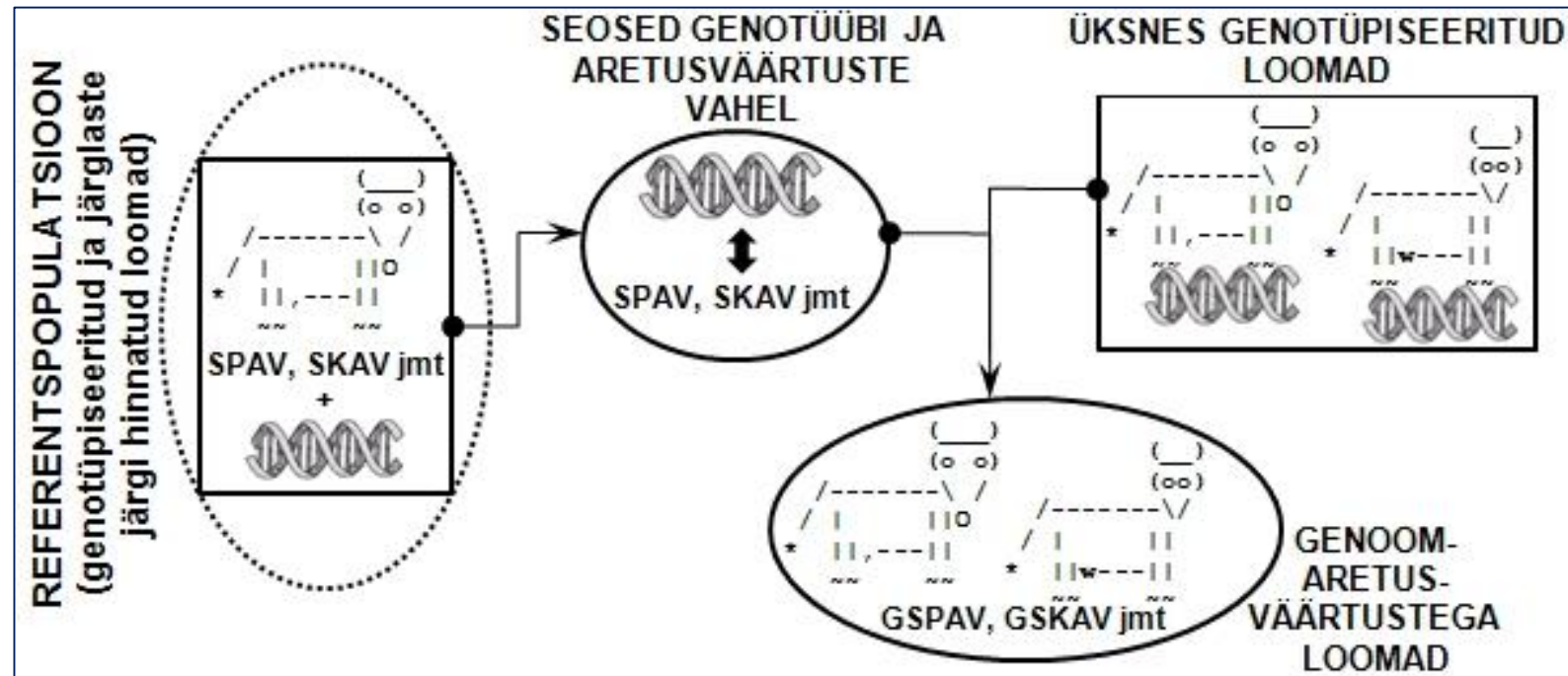


Aretusväärtused

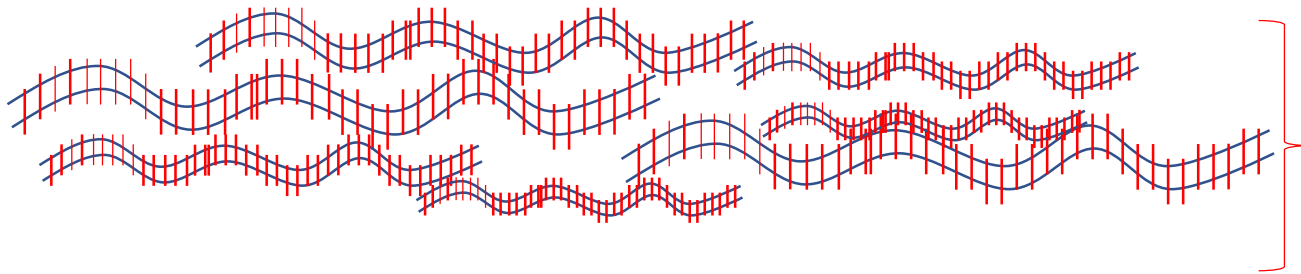


Genoomaretusväärtused ja -seleksioon

- Genoomaretusväärtus on genotüpiseeritud DNA-piirkondadele hinnatud arvuliste mõjude summa.



Genoomaretusväärtused ja -seleksioon



- * 50 000 geneetilist markerit,
- * igas positsioonis kaks võimalikku varianti,
- * igal variandil oma **arvuline efekt** x_i

Referentspopulatsioon

$$\left\{ \begin{array}{l} AV_1 \\ AV_2 \\ \dots \\ AV_{20\,000} \end{array} \right. = x_1 + x_2 + x_3 + \dots + x_{1\,000} + x_{1\,001} + \dots + x_{50\,000} + x_{50\,001} + \dots + x_{99\,999} + x_{100\,000}$$

↑
20 000 pulli aretusväärtused

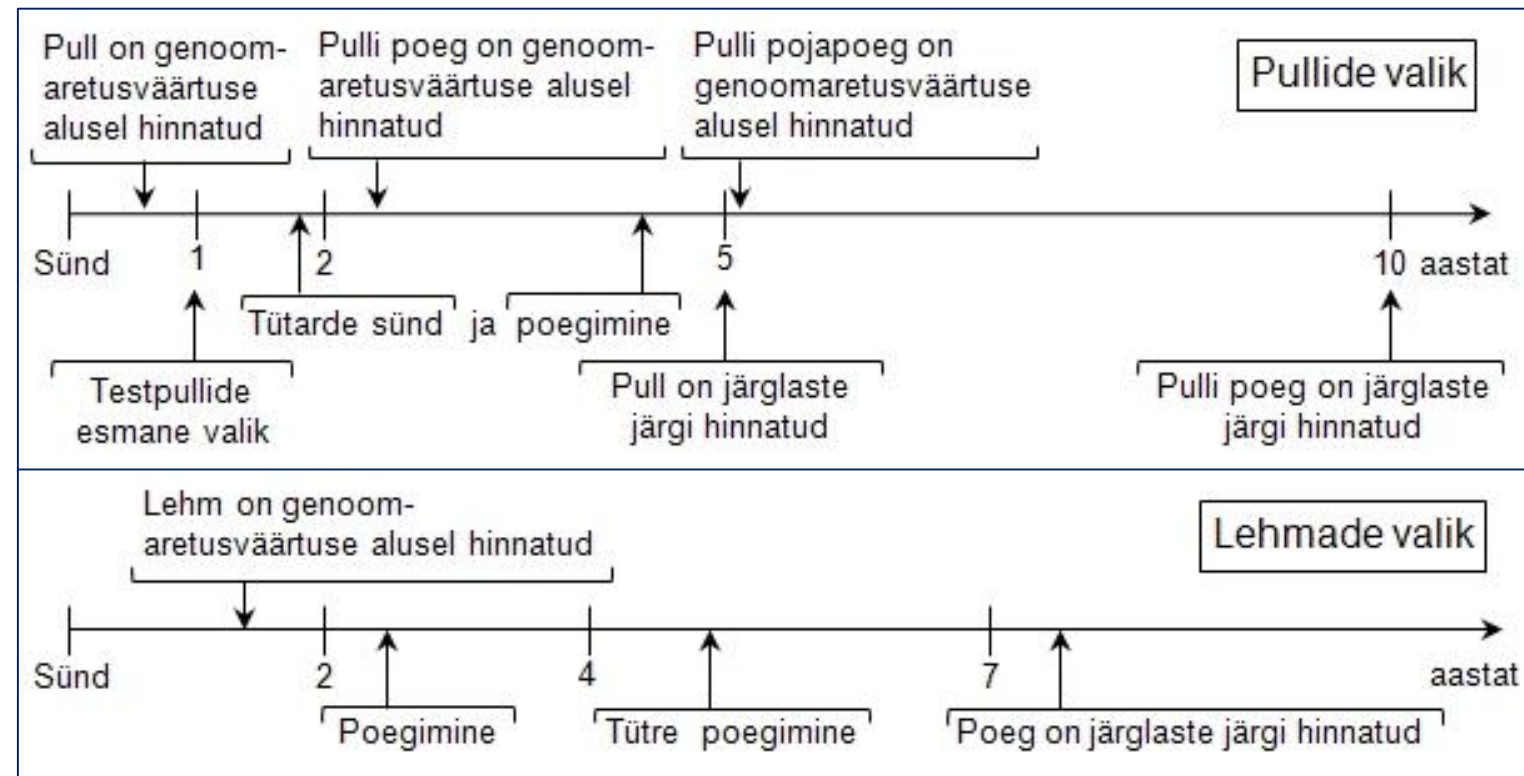
↑
100 000 tundmatut arvulist efekti

Tegu on umbes sama üheselt lahenduva ülesandega kui järgmise võrrandisüsteemi puhul:

$$\begin{cases} 110 = x + y + z + w + a + b + c + d + e + f \\ 98 = x + y + z + w + a + b + c + d + e + f \end{cases}$$

Genoomaretusväärtused ja -seleksioon

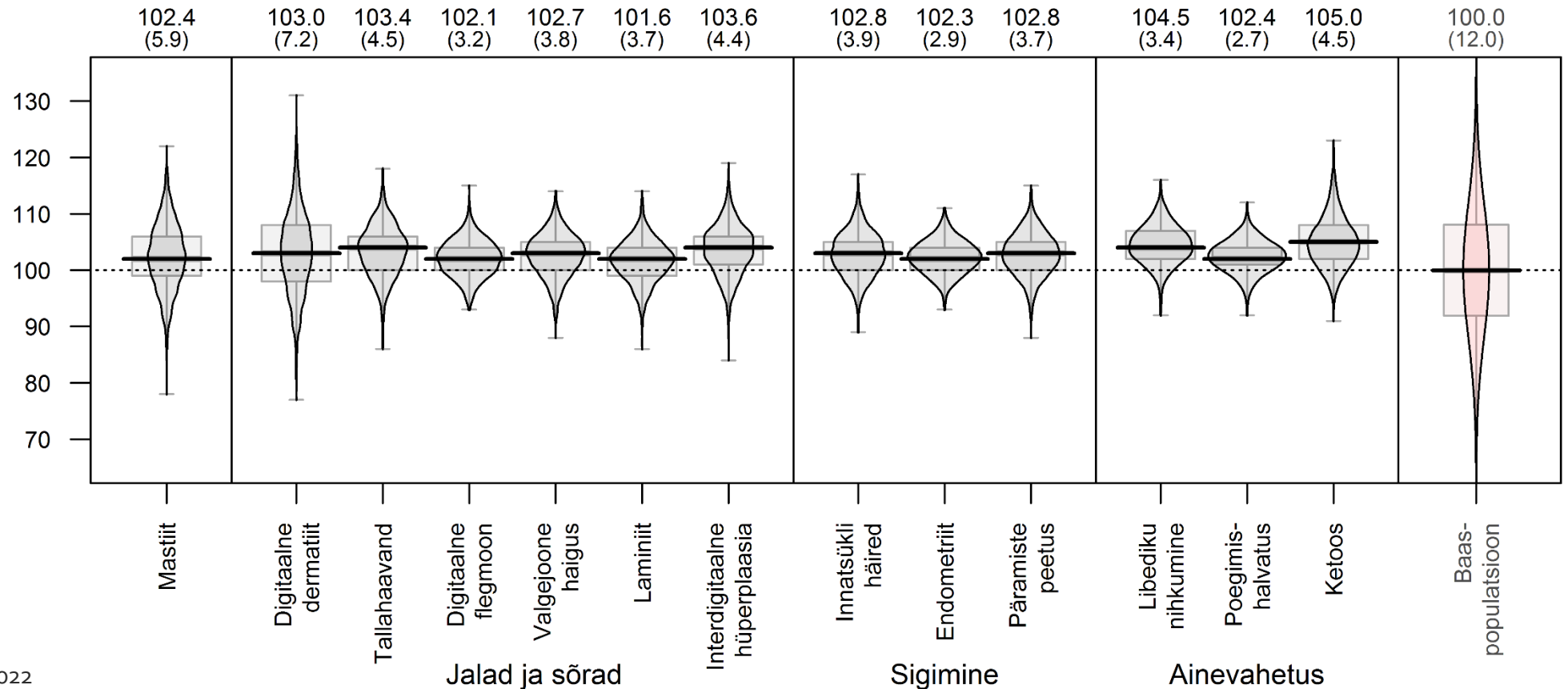
- | Generatsiooniintervall väheneb enam kui kaks korda – aretuseesmärk saavutatakse kiiremini.
- | Aretuses kasutamiseväärseteks mitteosutuvad individid on varem tuvastatavad.
- | Aretusväärtuse hinnangu täpsus on küll ~10% madalam võrreldes järglaste järgi hindamisega, aga ~poole kõrgem, kui põlvnemisindeksil.



Genoomaretusväärtused ja -selektioon

- Võimalus hinnata aretusväärtuseid tunnustele, mida ei olegi mõõdetud-registreeritud.

EHF lehmikute tervisetunnuste suhtelised genoomaretusväärtused Saksamaa skaalal



Aga kui loomi on vähe ja erinevaid söötmis- pidamis- jne tingimusi palju?

- | Robustsed, keskkonnast mittesõltuvad tunnused – näiteks välimiku lineaarsed hinded.
 - | Puuduseks ebatäpsus või isegi suutmatus peegeldada varieeruvust reaalselt huvi pakkuvates tunnustes.
- | Genoomaretusväärtused
 - | Tuleb tellida välismaalt (kallis), ei pruugi vastata siinsetele tingimustele.

Kokkuvõte

- | Aretusväärtuseid on palju.
- | Aretusväärtuste arvutusmeetodeid on palju – see, mis on sobilik, sõltub andmete/populatsiooni hulgast/suurusest ja struktuurist.
- | Õigete (st nihketa) hinnangute saamiseks on vajalik geneetilise materjali kasutamine erinevates keskkonnatingimustes ning nende keskkonnatingimuste registreerimine.
- | Suhtelised aretusindeksid on kokkuleppe ja tulevikuvisioni küsimus.

Ja lõpetuseks – millal on looma aretusväärtus null?

- | Loom vastab oma (pärandatavalt) geneetiliselt potentsiaalilt (baas)populatsiooni keskmisele.
- | Päritavus on null (tunnuse väärtus ei sõltu geenidest).
- | Uuritava tunnuse geneetiline varieeruvus populatsioonis on null.
- | Loomal ei ole uuritava tunnuse väärtust mõõdetud ning tal ei ole ka ühtki teadaolevat sugulast (ja loom kuulub baaspopulatsiooni).

(Pisut aegunud) lisamaterjalid

- | Selektsooniindeksid: http://ph.emu.ee/~ktanel/sel_indeks/
- | Aretusväärtuste hindamine isa ja looma mudelist:
http://ph.emu.ee/~ktanel/isa_ja_looma_mudel/,
 - | sh Exceli fail isa ja looma mudeli rakendamiseks:
http://ph.emu.ee/~ktanel/VL_0192/SAMod_1c.xls
- | Genoomselektsoon: <http://ph.emu.ee/~ktanel/genoomselektsoon/>