

H U S D Y R A V L

Generell Avlslære I (HA 1)

Forelesninger ved Norges Landbrukshøgskole

Av

Ola Syrstad

Landbruksbokhandelen

ISBN 82-557-0054-4

Ås-NLH 1977

E. F.

HUSDYRAVL

Generell avislære I (HA1)

Forelesninger ved Norges Landbrukshøgskole

Av

Ola Syrstad

LANDBRUKSBOKHANDELEN

ISBN 82-557-0054-4

Ås-NLH 1977

FORORD

Kurset "Generell avlslære I" (HA1) er ført opp i undervisningsplanen med tre forelesninger pr. veke, dvs. ca. 40 forelesninger i alt. Kurset skal gi "ei kort innføring i de mest sentrale deler av det teoretiske grunnlaget for moderne husdyravl (populasjonsgenetikk, genetiske parametre, seleksjon, avlsmetoder)". Det har vært forutsetningen at kurset skal gi den teoretiske innføringa som er nødvendig for at de studentene som måtte ønske det, kan fortsette med ett eller flere kurs i spesiell avlslære (HA3, HA4, HA5).

Dette kompendiet er et konsentrat av forelesningene i kurset, og vil under kurset bli supplert med eksempler og kommentarer. Det vil bli lagt stor vekt på løsning av oppgaver i tilknytting til det stoffet som blir lagt fram i forelesningene. Etter hvert kapittel er det derfor tatt med noen øvingsoppgaver for gjennomgåing i timene eller til bruk for heimearbeid.

Som hjelpelitteratur til kurset vil disse bøkene høve:

Falconer, D.S. 1960. Introduction to Quantitative Genetics. The Ronald Company Press, New York, 365 s.

Johansson, I. & Rendel, J. 1963. Arftlighet och husdjursförädling. LT's förlag, Stockholm. 368 s. (Engelsk utgave i 1968).

Pirchner, F. 1968. Population Genetics in Animal Breeding. W.H. Freeman & Company, San Francisco, 274 s.

Alle de tre bøkene er å finne i biblioteket til Institutt for husdyravl, og kan studeres der, men må ikke fjernes fra biblioteket.

Ås-NLH, juni 1977.

Ola Syrstad

INNHOOLD

	Side
I. Innleiing.	1
II. De grunnleggende arvelovene.	3
A. Noen grunnbegreper i arvelära.	3
B. Gjennomsnittlig virkning av et gen. - Additiv arv.	5
C. Kjønnbestemmelse. - Kjønnbundet nedarving.	6
D. Spalting og nykombinasjon av arveanlegg.	7
E. Nedarving av kvantitative egenskaper.	9
III. Genfrekvens. - Genotypefrekvens.	12
IV. Endring av genfrekvensene.	16
A. Mutasjon.	16
B. Migrasjon.	18
C. Seleksjon.	18
1. Seleksjon mot et resessivt gen.	
2. Seleksjon mot et dominant gen.	
3. Seleksjon ved overdominans.	
D. Tilfeldige endringer i genfrekvensen (genetisk drift).	23
V. Innavl og slektskap.	26
A. Mål for innavlsgrad.	26
B. Mål for slektskap.	28
C. Innavl i små populasjoner.	29
D. Endring i genotypefrekvensene ved innavl.	32
VI. Beskrivelse av populasjoner med omsyn på kvantitative egenskaper.	34
A. Frekvensfordeling. - Gjennomsnitt og standardavvik.	34
B. Oppdeling av variasjonen. - Varianskomponenter.	36
C. Korrelasjon og regresjon.	39
D. Path-koeffisienter.	40
VII. Variasjon i en kvantitativ egenskap.	44
A. Genetisk og miljøbestemt variasjon.	44
B. Oppdeling av den genetiske variasjon.	45
C. Oppdeling av den miljøbestemte variasjon.	45
D. Gjentaksgard.	
E. Korrelasjon mellom genotype og miljø.	45
F. Samspill mellom genotype og miljø.	49

	Side
VIII. Genetisk andel av variasjonen. - Arvegrad.	51
A. Definisjoner.	51
B. Utrekning av arvegraden for en egenskap.	52
1. Likhhet foreldre - avkom.	
2. Likhhet mellom søsken (helsøsken og halvsøsken).	
C. Arvegrad for et gjennomsnitt av flere målinger.	57
IX. Korrelasjon mellom egenskaper.	60
A. Fenotypisk og genetisk korrelasjon.	60
B. Utrekning av den genetiske korrelasjon.	61
X. Seleksjon.	63
A. Seleksjonsdifferanse.	63
B. Genetisk virkning av seleksjon.	66
C. Indirekte seleksjon.	67
D. Seleksjon for flere egenskaper.	69
1. Tandem-metoden.	
2. Minstekrav-metoden.	
3. Indeksmetoden.	
XI. Utvalgsmetoder.	73
A. Masseutvalg.	73
B. Utvalg etter avstamning.	74
C. Utvalg etter avkomsgransking.	76
D. Utvalg etter slektninger i sideledd. - Familieutvalg.	79
1. Individet ikke medrekna i familiegjennomsnittet.	
2. Individet medrekna i familiegjennomsnittet.	
XII. Virkning av langvarig seleksjon.	84
A. Seleksjonsforsøk.	84
B. Gransking av genetisk framgang på grunnlag av data fra praksis.	85
XIII. Innavlsdepresjon og heterosis.	88
A. Innavlsdepresjon.	88
B. Heterosis.	89
C. Avlsmetoder som utnytter heterosis.	91
1. Rasekryssing.	
2. Kryssing av innavlsliner.	
3. Gjensidig seleksjon.	

I. INNLEIING

Moderne husdyravl bygger på arvelæra. De grunnleggende arvelovene ble riktignok utforma på grunnlag av forsøk med planter, men det ble snart klarlagt at nedarvingsmekanismen i prinsippet er lik for alle organismer med kjønna formering.

I den første tida etter at arvelovene ble oppdaga var arvelighetsforskerne i første rekke opptatt av å studere nedarvings-tilhøva for egenskaper der individa innafor ei art skilte seg fra hverandre på en slik måte at de kunne grupperes i noen få, klart avgrensa grupper (f.eks. blomsterfarge). Slike egenskaper blir kalt kvalitative egenskaper. Eksempler på kvalitative egenskaper hos husdyra er hårfarge, blodtype, enkelte former for misdanning o.l.

De fleste av de egenskapene som er av praktisk betydning i husdyrbruket er likevel av en slik natur at individa ikke lar seg gruppere i noen få avgrensa grupper. Det er tvert imot slik at en finner alle mulige mellomformer, slik at en får en kontinuerlig variasjon fra den ene ytterligheten til den andre. Disse egenskapene blir kalt kvantitative egenskaper. Kroppsstorleik, mjølkeavdrått og veksthastighet er eksempler på slike egenskaper. For å registrere kvantitative egenskaper må en i regelen foreta målinger, og de blir derfor også kalt metriske egenskaper.

Kvantitative egenskaper viser oftest ei fordeling som tilnærma svarer til normalfordelinga. I tilfelle der fordelinga er utprega skjev kan en ofte foreta enkle transformasjoner som gjør at en får data som er mere i samsvar med normalfordelinga.

I de første åra etter at de grunnleggende arvelovene var gjenoppdaga (i 1900) mente en at disse lovene bare gjaldt for kvalitative egenskaper. Men snart viste det seg at nedarving av kvantitative egenskaper kunne forklares ut fra de samme prinsippa, skilnaden er bare at en kvantitativ egenskap i regelen er bestemt av flere arveanlegg. Dessuten har det vist seg at de fleste kvantitative egenskapene er sterkt påvirka av de ytre tilhøva som individa lever under.

Enkelte egenskaper kan ikke uten videre klassifiseres som enten kvalitative eller kvantitative etter den omtalen som er gått ovafor. Dette gjelder bl.a. egenskaper der kjennetegnet er et antall, f.eks. antall avkom pr. fødsel hos flerfødende dyr. Slike egenskaper

kan etter sin natur ikke vise fullt ut kontinuerlig variasjon. Mangelen på kontinuitet blir mer markert jo lågere det gjennomsnittlige antall er. Det har likevel vist seg at slike "antalls-egenskaper" i regelen blir nedarva på samme måte som kvantitative egenskaper. Dette gjelder også for mange egenskaper der individa kan grupperes bare i to grupper, f.eks. friske eller sjuke. En tenker seg at resistensen mot sjukdom i virkeligheten viser en kontinuerlig variasjon. Når verdien av denne kontinuerlig variable kommer under et visst nivå (terskelverdi) blir dyra sjuke, dersom verdien ligger over dette nivået er de friske. Slike egenskaper blir gjerne kalt terskelegenskaper.

De arbeidsmetodene som en nytter ved studiet av kvantitative egenskaper skiller seg på minst to måter fra de som blir nytta i den kvalitative genetikken:

a) Fordi de kvantitative egenskapene ikke viser enkle spaltingshøve, slik som kvalitative egenskaper gjerne gjør, er interessen ikke i første rekke konsentrert om de enkelte individa, men om grupper av individer, det en gjerne kaller populasjoner. Læra om nedarvings-tilhøva for kvantitative egenskaper blir derfor også kalt populasjonsgenetikk.

b) Etersom de kvantitative egenskapene viser en kontinuerlig variasjon, vil en måtte bruke statistiske (biometriske) metoder for å analysere variasjonen. Dette har gitt opphav til nemninga biometrisk genetikk.

De tre nemningene (kvantitativ genetikk, populasjonsgenetikk, biometrisk genetikk) brukes ofte som fullstendig synonyme. Strengt tatt er det en viss skilnad, i det populasjonsgenetikk også kan anvendes for kvalitative egenskaper. Det samme gjelder i en viss grad for biometrisk genetikk.

II. DE GRUNNLEGGENDE ARVELOVENE.

A. Noen grunnbegreper i arvelæra.

Ifølge moderne arvelære blir arveanlegga overført fra foreldre til avkom i form av udelelige enheter. Disse enhetene blir kalt arveanlegg eller gener. Arveanlegga er plassert på kromosoma, som er noen trådforma legemer i cellekjernen.

I den frødde eggcella, zygoten, opptrer kromosoma parvis (diploid), i det ett medlem av hvert kromosompar er kommet fra hvert av foreldra. Kroppscellene er, arvemessig sett, kopier av den frødde eggcella. I kjønnscellene, gametene, som dannes ved ei spesiell form for celledeling (reduksjonsdeling) er kromosoma til stede i enkelt dose (haploid), dvs. ett kromosom fra hvert kromosompar.

Dette at kromosoma, og dermed også arveanlegga, forekommer parvis i alle celler unntatt kjønnscellene betyr ikke at de to arveanlegga som utgjør et par er like i sin virkning. Dersom det arveanlegget som er kommet fra faren har en annen virkning enn det som er kommet fra mora, sier en at individet er heterozygot med omsyn på dette anleggsparet, og at hvert av de to arveanlegga er til stede i heterozygot form. Er de to arveanlegga i anleggsparet identiske i sin virkning, sier en at individet er homozygot for dette arveanlegget (subst. heterozygoti - homozygoti).

Antall kromosompar, og forma på de ulike kromosompara, er det samme for alle individer innen samme art. De arveanlegga som kan forekomme på et bestemt sted på et bestemt kromosom blir kalt allelomorfe gener eller alleler. Allele gener er gener som hører til samme locus (flert. loci). Dersom et locus omfatter mere enn to ulike arveanlegg, snakker en om multiple alleler. Men et enkelt individ kan aldri ha mere enn to arveanlegg fra samme locus.

Som symbol for arveanlegga nytter en til vanlig bokstaver. Alle anlegg som hører til et bestemt locus får samme bokstav. For å skille dem fra hverandre bruker en indekser (f.eks. A_1 , A_2 osv.). Dersom det bare er tale om to alleler, kan en også bruke stor bokstav for det ene og liten bokstav for det andre (f.eks. A, a).

De arveanlegga som et individ har, utgjør tilsammen individets genotype. En kan også snakke om genotypen for en bestemt egenskap eller med omsyn på et bestemt genpar. Fenotypen er de personlige særdrag som individet har (framtoningspreget). Når en skal registrere en egenskap, er det alltid fenotypen en registrerer.

For egenskaper som ikke er påvirkta av ytre tilhøve vil det til hver genotype svare én bestemt fenotype. Dersom fenotypen for heterozygoter ligger midt mellom fenotypene for de tilsvarende to homozygotene, foreligger det intermediær nedarving. Er heterozygotene lik den ene av homozygotene, har en dominans. En sier at det genet som slår igjennom, er dominant eller at det dominerer over det allele genet, som er resessivt. Det forekommer også at en har partiell dominans, dvs. at heterozygotene ligger mellom de to homozygotene, men nærmere den ene enn den andre. Dersom de heterozygote individa er "bedre" enn noen av de to homozygotene, har en overdominans.

De ulike dominanstilhøva kan illustreres som vist i fig. 1. Her er de to homozygote typene gitt verdien a og $-a$, og d angir graden av dominans:

$d = 0$ Intermediær nedarving

$0 < d < a$ Partiell dominans

$d = a$ Fullstendig dominans

$d > a$ Overdominans.

Dersom det er tydelig dominans angir en ofte det dominante arveanlegget med stor bokstav og det resessive anlegget med liten bokstav.

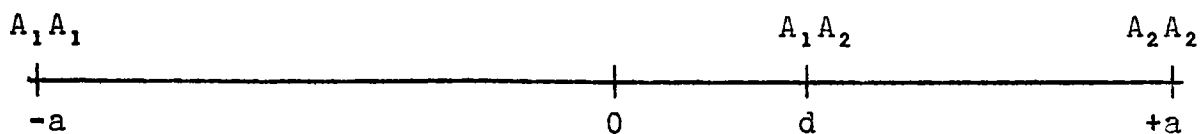


Fig. 1.

Når det foreligger dominans, vil virkningen av å skifte ut et arveanlegg med et anna anlegg fra samme locus ikke bare avhenge av forskjellen mellom de to arveanlegga, men også av hva for anlegg som er til stede på den tilsvarende plassen på det

andre kromosomet i dette kromosomparet. En kan si at "nyttene" av et arveanlegg avhenger av hvilken partner det har. Dette at virkningen av en bestemt faktor avhenger av hva for andre faktorer som er til stede, kalles samspill. Dominans er samspill mellom arveanlegg fra samme locus.

På samme måte er epistasi et samspill mellom arveanlegg fra ulike loci. Det foreligger mange eksempler på at et bestemt arveanlegg har virkning bare hvis det forekommer samtidig med et gitt arveanlegg fra et annet locus. En kan si at det er en viss kombinasjon av arveanlegg som virker, og ikke de enkelte anlegga hver for seg.

Ett og samme arveanlegg har ofte virkning på flere egenskaper hos de individa som fører anlegget. Dette kalles pleiotropi. Når det foreligger pleiotropi, vil egenskapene ikke nedarves uavhengig av hverandre, men ha en tendens til å følges at på en bestemt måte.

En annen årsak til at ulike egenskaper ikke blir nedarva helt uavhengig av hverandre er kopling. Når to arveanlegg er plassert på samme kromosompar, vil de ha en tendens til å følge hverandre under nedarvinga. Jo nærmere sammen de to arveanlegga ligger på kromosomet, jo tettere er koplinga. Dersom koplinga er tett kan det være vanskelig å avgjøre om arveanlegga hører til samme locus eller ulike loci.

Iblant inntreffer det at de to kromosoma i kromosomparet utveksler deler. Dette kalles overkryssing. Ved overkryssing kan ei kopling mellom to arveanlegg bli brutt.

B. Gjennomsnittlig virkning av et gen. - Additiv arv.

Det enkelte genet gir seg til kjenne gjennom den virkningen det har at en bytter inn dette genet i stedet for et annet gen på samme locus. Ved dominans (og epistasi) vil virkningen av ei slik innbytting være avhengig av hva for genotype innbyttinga foregår i.

La oss anta at vi har et locus med to gener, A_1 og A_2 , med partiell dominans for A_1 , og at verdien av de tre mulige genotypene er som angitt i fig. 1.

Innbytting av genet A_1 i stedet for A_2 vil gi dette resultatet:

$$\begin{array}{l} A_1A_2 \rightarrow A_1A_1 \quad a - d \quad = -a + d \\ A_2A_2 \rightarrow A_1A_2 \quad a + d \end{array}$$

Den gjennomsnittlige virkningen av å erstatte et A_2 -gen med A_1 avhenger av hvor stor del av A_2 -gena som forekommer i genotypen A_2A_2 og hvor stor del som forekommer i genotypen A_1A_2 .

Den gjennomsnittlige virkningen av et gen kan defineres som det gjennomsnittlige avvik fra populasjonsmidlet for de individa som har fått dette genet fra det ene av foreldra, når det genet som det har fått fra det andre av foreldra er tatt tilfeldig fra populasjonen. Dette blir også kalt den additive virkning av genet.

Den gjennomsnittlige virkningen av å skifte ut A_2 med A_1 er lik skilnaden i additiv verdi for de to gena.

C. Kjønnbestemmelse. - Kjønnbundet nedarving.

Regelen om at kromosoma i kroppsellene hos diploide organismer er parvis like, har ett viktig unntak. Hos individer av de ene kjønnnet består ett av kromosompara av to ulike kromosomer, mens individer av det motsatte kjønnnet har to like kromosomer også i dette paret. Kromosoma i dette kromosomet blir kalt kjønnskromosomer. Hos pattedyra har hannene to ulike og hoene to like kjønnskromosomer. En sier at hannene er det heterogametiske og hoene det homogametiske kjønnnet. Det ene av de to kjønnskromosoma hos hannene er likt med de to kjønnskromosoma hos hoene. Disse kromosoma blir kalt X-kromosomer, mens det andre kjønnskromosomet hos hannene blir kalt Y-kromosom.

Av dette følger at hannene produserer to typer av gameter med omsyn til hva slags kjønnskromosom de fører. Den ene typen fører et X-kromosom og den andre et Y-kromosom. De to gamettyperne blir til vanlig dannet i tilnærma samme antall. Alle gameter fra hoene fører et X-kromosom. Når ei sædcelle med Y-kromosom smelter sammen med ei eggcelle, blir det derfor dannet ei zygote med ett X-kromosom og ett Y-kromosom, og zygoten utvikler seg til et hannlig individ. Dersom sædcella fører et X-kromosom vil zygoten få to X-kromosomer, og individet blir ei hoe.

Hos fuglene er hoene det heterogametiske og hannene det homogametiske kjønnnet.

X-kromosoma er bærere av gener på samme måte som andre kromosomer, mens Y-kromosomet er ansett for å være praktisk talt tomt for gener. Egenskaper som er bestemt av gener på X-kromosomet vil derfor vise et nedarvingsmønster som avviker fra det en ellers

kjenner. Hos individer av det heterogametiske kjønnnet vil et resessivt gen gi seg utslag jamvel om det forekommer i enkelt dose, fordi det ikke blir dekket av en dominant partner. En egenskap som er bestemt av et resessivt gen på X-kromosomet vil derfor vise seg med langt større frekvens hos individer av det heterogametiske kjønnnet.

Et velkjent eksempel på en egenskap som viser kjønnsbundet nedarving er en viss form for fargeblindhet hos mennesker. Fargeblindhet er bestemt av et resessivt gen på X-kromosomet. La oss kalle dette genet A_2 , og det allele genet for normal fargesans A_1 . En mann som har genet A_2 på X-kromosomet sitt vil være fargeblind til tross for at genet forekommer i enkelt dose. Han vil gi X-kromosomet, og med det genet for fargeblindhet, til døtrene sine, mens sønnene, som får faren sitt Y-kromosom, går fri. Også døtrene vil i regelen få normal fargesans, fordi de har genet A_1 på det X-kromosomet de har fått fra morsida. Men de kan gi genet for fargeblindhet (A_2) videre til barna sine, og sønnene vil i så fall bli fargeblinde. Dette at egenskapen opptrer i annenhver generasjon er typisk for egenskaper som er bestemt av et resessivt gen på X-kromosomet.

D. Spalting og nykombinasjon av arveanlegg.

All sammenparing av individer med ulike arveanlegg blir i arvelæra betegnet som kryssing. Ved kryssing av to individer som hver for seg er homozygote i et bestemt locus, men for ulike arveanlegg, vil avkommet bli heterozygot. Avkom av denne første kryssingsgenerasjonen blir gjerne kalt F_1 . Dersom foreldrene er av genotypene A_1A_1 og A_2A_2 , får avkommet genotypen A_1A_2 . Fenotypisk kan avkommet være lik den ene av foreldretypene (dominans) eller være en mellomting mellom foreldra (intermediær nedarving eller partiell dominans).

F_1 -individene vil produsere to genetisk ulike typer av kjønns-celler (med arveanlegg A_1 eller A_2). Ved innbyrdes paring av F_1 vil en derfor få disse genotypene:

$$A_1 \text{ fra far} + A_1 \text{ fra mor} = A_1A_1$$

$$A_1 \text{ " " } + A_2 \text{ " " } = A_1A_2$$

$$A_2 \text{ " " } + A_1 \text{ " " } = A_2A_1$$

$$A_2 \text{ " " } + A_2 \text{ " " } = A_2A_2$$

De to midterste er genetisk like og like med F_1 , mens de to andre er like med hvert sitt av foreldra i den første kryssinga. Dersom det foreligger fullstendig dominans vil det fenotypiske spaltingshøvet bli 3:1, ellers 1:2:1. Disse spaltningshøva er typiske for monohybride spaltninger (ett anleggspaar). Avkom etter innbyrdes paring av F_1 -individer kalles F_2 .

La oss så tenke oss at vi parer sammen to individer som hver for seg er homozygote og som er ulike i to egenskaper. Dersom genotypen for de to individa er $A_1A_1B_1B_1$ og $A_2A_2B_2B_2$, vil avkommet bli av genotypen $A_1A_2B_1B_2$. Dersom de to loci (A og B) er lokalisert på ulike kromosomer, vil F_1 individa produsere fire genetisk ulike typer av gameter, nemlig A_1B_1 , A_1B_2 , A_2B_1 og A_2B_2 . Ved innbyrdes paring av F_1 vil disse fire typene kombinere seg på alle mulige måter. For å lette oversikten kan en sette opp et "sjakkbrett":

Kjønns-celler	A_1B_1	A_1B_2	A_2B_1	A_2B_2
A_1B_1	$A_1A_1B_1B_1$	$A_1A_1B_1B_2$	$A_1A_2B_1B_1$	$A_1A_2B_1B_2$
A_1B_2	$A_1A_1B_1B_2$	$A_1A_1B_2B_2$	$A_1A_2B_1B_2$	$A_1A_2B_2B_2$
A_2B_1	$A_1A_2B_1B_1$	$A_1A_2B_1B_2$	$A_2A_2B_1B_1$	$A_2A_2B_1B_2$
A_2B_2	$A_1A_2B_1B_2$	$A_1A_2B_2B_2$	$A_2A_2B_1B_2$	$A_2A_2B_2B_2$

Dette kalles ei dibybrid spalting (to anleggspaar). Av de 16 kombinasjonene er det fire av genotypen $A_1A_2B_1B_2$, og to av hver av genotypene $A_1A_2B_1B_1$, $A_1A_2B_2B_2$, $A_1A_1B_1B_2$ og $A_2A_2B_1B_2$. Det er derfor bare ni ulike genotyper. Blant disse er begge foreldre-

genotypene og genotypen for F_1 , men det har også kommet fram seks "nye" genotyper, og to av disse er homozygote for begge de to anleggspara ($A_1A_1B_2B_2$ og $A_2A_2B_1B_1$).

Ved paring av individer som er ulike i n arveanlegg som blir nedarva uavhengig av hverandre, og innbyrdes paring av F_1 -individua, er antall kombinasjonsmuligheter i F_2 4^n og antall mulige genotyper 3^n . Av disse vil 2^n være homozygote. Dersom foreldra er ulike i f.eks. 10 anleggspar som de hver for seg er homozygote for, vil det i F_2 være nesten 60 tusen mulige genotyper, men bare vel ett tusen av disse vil være homozygote for alle de ti anleggspara. Dette viser at det i en populasjon med stor grad av heterozygoti, slik tilfellet er i de fleste populasjoner av husdyr, vil være nesten en uendelighet av ulike genotyper. Og enda har en her rekna med bare to mulige arveanlegg for hvert locus. Dersom det forekommer multiple alleler, blir antall mulige genotyper enda langt større. I et locus med m alleler vil det være m mulige homozygote og $m(m-1)/2$ mulige heterozygote genotyper, dvs. i alt $m + m(m-1)/2 = m(m+1)/2$ ulike genotyper. Dersom en har n loci som nedarves uavhengig av hverandre, og hvert locus omfatter m arveanlegg, er antall mulige genotyper gitt av formelen $[m(m+1)/2]^n$

Iblant har en bruk for å undersøke om resultatene som er kommet fram blant avkommet etter ei paring er i samsvar med en bestemt hypotese for nedarvingstilhøva. En må da jamføre antall individer av ulike fenotyper med det antall en skulle vente etter hypotesen. Avvik mellom observerte og "ventede" antall kan prøves statistisk etter kji-kvadrat-metoden (χ^2).

E. Nedarving av kvantitative egenskaper.

Nedarvingstilhøva for kvantitative egenskaper kan forklares ved å anta at slike egenskaper er påvirka av flere arveanlegg, som hver for seg har liten virkning, men som adderer sin virkning. Slike arveanlegg kalles polymere gener eller polygener.

Et klassisk eksempel på kvantitativ nedarving er rau kornfarge hos kveite. Etter ei kryssing mellom en kveitesort med intens rau kornfarge og en sort med kvit farge ble det i F_2 funnet et spaltingshøve 63 raue:1 kvit. De raukorna typene varierte sterkt i fargeintensitet, fra intens rau (som den ene av foreldresortene)

til svært svak raufarge. For å forklare dette ble det antatt at kornfargen er bestemt av tre ulike anleggspaar. Hvert par består av et "kvitt" og et "raudt" arveanlegg. Den raue foreldresorten var homozygot for alle de tre "raue" arveanlegga og den kvite foreldresorten homozygot for alle de tre "kvite".

I F_2 vil antall "raue" arveanlegg variere fra 0 til 6. Dersom hvert arveanlegg har like stor verknad, og det er intermediær nedarving, vil en altså få sju klasser med varierende fargeintensitet. Fordelinga på de ulike klassene blir slik:

Antall "raue" arveanlegg	Antall indivi- vider (av 64)
(kvit) 0	1
1	6
2	15
3	20
4	15
5	6
6	<u>1</u>
Sum	64

Ved å tegne opp dette grafisk med antall "raue" arveanlegg som uavhengig og antall individer i ulike klasser som avhengig variabel vil en se at fordelingskurva nærmer seg den forma som gjelder for normalfordeling. Jamvel med et så beskjedent antall arveanlegg som i dette tilfellet vil det i praksis ofte være umulig å skille individer i de ulike intensitetsklassene fra hverandre, og en vil få en tilnærma kontinuerlig variasjon.

Det er ingen grunn til å anta at de ulike arveanlegga som påvirker en kvantitativ egenskap alle har like sterk virkning. En må vente at enkelte arveanlegg har en sterkere virkning enn andre. Heller ikke er det sikkert at nedarvinga på hvert enkelt locus er intermediær. Men jamvel om nedarvinga er dominant, vil en få varierende grader av egenskapen, fordi enkelte individer vil mangle ett eller flere av de arveanlegga som virker inn.

Øvingsoppgaver.

1. I et locus med to alleler, A_1 og A_2 , er verdien av de tre mulige genotypene:

$$A_1A_1 = 5$$

$$A_1A_2 = 4$$

$$A_2A_2 = 1$$

Hva er den additive virkningen av genet A_1 i en F -populasjon fra kryssinga $A_1A_1 \times A_2A_2$?

2. I den engelske ferasen korthorn forekommer det både raue, kvite og skimlete dyr (skimlet er ei blanding av raue og kvite hår). Blant avkommet etter paring av skimlet \times skimlet ble det i et materiale funnet 12 raue, 34 skimlete og 14 kvite dyr. Hvordan stemmer dette med hypotesen om monohybrid spalting, der skimlet er heterozygoter mellom raudt og kvitt (intermediær nedarving)?
3. I et forsøk der raue, hornete kyr av korthornrasen ble para med okser fra den svarte, kollete rasen aberdeen-angus ble alt avkommet i første generasjon (F_1) svart og kollet. Ved innbyrdes paring mellom F_1 -individene ble det produsert en F_2 -generasjon som omfatta 48 dyr. Hvor mange individer skulle en vente av hver type, dersom de to egenskapene (svart/raudt og hornet/kollet) nedarves uavhengig av hverandre?
4. I en populasjon forekommer det fire loci med følgende antall alleler:

Locus A: 3 alleler

Locus C: 2 alleler

" B: 4 "

" D: 3 "

Hvor mange mulige genotyper har en i denne populasjonen med omsyn på de fire loci? Hvor mange av disse genotypene er fullstendig homozygote?

III. GENOTYPEFREKVENNS - GENFREKVENNS.

Dersom en skal gi ei fullstendig genetisk beskrivelse av en populasjon, må en angi genotypen for hvert enkelt individ. For alle loci sett under ett vil det i regelen være like mange genotyper som individer. Men for hvert locus vil antall genotyper i reglen være ganske lite. Dersom en ser på ett enkelt locus, kan en derfor beskrive populasjonen ved å angi antall individer med en bestemt genotype, eller hvor stor del av individene som har en bestemt genotype. Dette siste tallet blir kalt genotypefrekvensen. Matematisk kan dette uttrykkes slik: Genotypefrekvensen P for genotypen A_1A_1 er lik N_P/N , der N er det totale antall individer i populasjonen og N_P er antall individer med genotypen A_1A_1 .

Ved overføring av arvestoffet fra en generasjon til den neste vil genotypene blir brutt ned, men genene vil bestå. I studiet av de genetiske endringene i en populasjon er en derfor mer interessert i frekvensen av de ulike genene enn i genotypefrekvensen. Med genfrekvens mener en den andelen som et bestemt gen utgjør av det totale antall gener i populasjonen (på et bestemt locus).

Når en kjenner genotypefrekvensene i en populasjon, kan en rekne ut genfrekvensene. La oss anta at det er to alleler, A_1 og A_2 , og at genotypefrekvensene for de tre mulige genotypene er P (for A_1A_1), H (for A_1A_2) og Q (for A_2A_2). Det totale antall gener i populasjonen er $2N$, og antall A_1 - og A_2 -gener er henholdsvis:

$$A_1: 2N \cdot P + N \cdot H = N(2P+H)$$

$$A_2: N \cdot H + 2N \cdot Q = N(H+2Q)$$

$$\text{Genfrekvensen for } A_1: p = \frac{N(2P+H)}{2N} = P + \frac{1}{2}H$$

$$\text{" " } A_2: q = \frac{N(H+2Q)}{2N} = \frac{1}{2}H + Q$$

En merker seg at summen av genfrekvensene for alle genene som tilhører ett og samme locus alltid vil være lik 1,0. Det samme gjelder for summen av genotypefrekvensene:

$$p + q = P + \frac{1}{2}H + \frac{1}{2}H + Q = P + H + Q$$

$$= \frac{N_P}{N} + \frac{N_H}{N} + \frac{N_Q}{N} = \frac{N_P + N_H + N_Q}{N} = \frac{N}{N} = 1$$

I en stor populasjon der det ikke forekommer mutasjon, migrasjon ("innvandring") eller seleksjon (utvalg), vil genfrekvensene være konstante fra generasjon til generasjon. Forutsetningen om at populasjonen må være stor, er nødvendig fordi en ellers kan få betydelige endringer i genfrekvensene på grunn av tilfeldigheter.

Dersom en i tillegg til disse vilkåra har tilfeldig paring, (panmixis) vil også genotypefrekvensene være konstante. Med tilfeldig paring forstår en at alle individa av et bestemt kjønn har samme sjanse til å bli partnere til et gitt individ av motsatt kjønn, slik at det ikke er noen tendens til sammenparing av individer som er genetisk like eller genetisk ulike.

Dette at genfrekvensene og genotypefrekvensene i en populasjon, under de forutsetningene som er angitt ovafor, er konstante fra generasjon til generasjon, er kjent som Hardy-Weinbergs lov. En slik populasjon sies å være i genetisk likevekt.

I en populasjon i likevekt er genotypefrekvensene fullstendig bestemt av genfrekvensene. La oss ta et locus med to alleler, A_1 og A_2 , der frekvensene i en gitt generasjon (foreldregenerasjonen, generasjon 0) er p (for A_1) og q (for A_2). Sannsynligheten for at en gamet (egg- eller sædcelle) fører genet A_1 er p . Sannsynligheten for at de to gametene som forener seg til en zygote begge fører genet A_1 , slik at en får et individ med genotypen A_1A_1 , er $p \cdot p = p^2$. På samme måte er sannsynligheten q^2 for at begge gametene fører genet A_2 (som gir genotypen A_2A_2). Sannsynligheten for at sædcella fører genet A_1 og eggcella A_2 er $p \cdot q$, og den samme sannsynligheten har en for at sædcella har genet A_2 og eggcella A_1 . En får da disse genotypefrekvensene:

Genotyper:	A_1A_1	A_1A_2	A_2A_2
Frekvenser:	p^2	$2pq$	q^2

På grunnlag av genotypefrekvensene rekker en ut genfrekvensene i avkomsgenerasjonen (generasjon 1) slik:

$$A_1: p' = p^2 + \frac{1}{2}(2pq) = p^2 + pq = p$$

$$A_2: q' = \frac{1}{2}(2pq) + q^2 = pq + q^2 = q$$

Dette er de samme genfrekvensene som i foreldregenerasjonen.

Den sammenhengen mellom genfrekvenser og genotypefrekvenser som kjennetegner en populasjon i likevekt inntreffer allerede etter én generasjon av tilfeldig paring, uansett hva for genotypefrekvenser en hadde i foreldregenerasjonen.

Hardy-Weinbergs lov kan nyttes for å granske om en populasjon blir reprodusert ved tilfeldig paring, ved at en jamfører de observerte genotypefrekvensene med de en skulle vente om paringa var tilfeldig. Om avvika er for store til at de kan skyldes tilfeldigheter, avgjøres ved hjelp av kji-kvadrat-metoden (χ^2).

Dersom det foreligger fullstendig dominans for det ene av allelene i et locus, slik at heterozygotene ikke kan skilles fra den ene av de to homozygote typene, er det ikke mulig å rekne ut genfrekvensene direkte. Under forutsetning av at populasjonen er i likevekt, slik at Hardy-Weinbergs lov gjelder, kan en rekne ut genfrekvensene indirekte, fordi frekvensen av det resessive genet er lik kvadratrot av genotypefrekvensen for den genotypen som er homozygot for det resessive genet.

Den sammenhengen mellom genfrekvenser og genotypefrekvenser som finnes i en populasjon i likevekt etter Hardy-Weinbergs lov, gjelder også for loci med mer enn to alleler. La oss anta at vi har tre alleler, A_1 , A_2 og A_3 , med frekvensene p , q og r . Frekvensen av de ulike genotypene er da gitt av dette uttrykket:

$$\begin{aligned} & [p(A_1) + q(A_2) + r(A_3)]^2 \\ &= p^2(A_1A_1) + 2pq(A_1A_2) + 2pr(A_1A_3) \\ & \quad + q^2(A_2A_2) + 2qr(A_2A_3) + r^2(A_3A_3) \end{aligned}$$

Regelen om at likevekt (dvs. konstante genotypefrekvenser fra generasjon til generasjon) inntreffer etter én generasjon med tilfeldig paring, gjelder for hvert locus sett isolert (uansett antall alleler), men ikke dersom en ser på flere loci i sammenheng. For å illustrere dette kan vi tenke oss at vi har en populasjon av individer med genotypene $A_1A_1B_1B_1$ og $A_2A_2B_2B_2$. Ved tilfeldig paring vil en få avkom av tre genotyper, de øvrige seks genotypene som vil være til stede når populasjonen er i likevekt, mangler. Disse genotypene vil opptre først i den påfølgende generasjon, men heller ikke nå med den frekvens som i likevektssituasjonen. Det viser seg at avviket fra likevekt halveres for hver generasjon (dersom de to loci er uavhengige, dvs. på ulike kromosompar).

Øvingsoppgaver

1. I en populasjon har en ved opptelling av 400 individer funnet dette antallet av ulike genotyper:
 A_1A_1 : 201 individer
 A_1A_2 : 158 "
 A_2A_2 : 41 "
 - a. Rekn ut genotypefrekvensene og genfrekvensene.
 - b. Er populasjonen i likevekt?
 - c. Rekn ut venta genotypefrekvenser i neste generasjon, under forutsetning av tilfeldig paring.

2. I en populasjon med tilfeldig paring og der det ikke forekommer mutasjon, migrasjon eller seleksjon, er A_1A_2 et genpar der A_1 dominerer fullstendig over A_2 . Ei gransking viser at 6,76 prosent av individa er av genotypen A_2A_2 .
 - a. Rekn ut frekvensen for gena A_1 og A_2 og for genotypene A_1A_1 og A_1A_2 .
 - b. Hvor stor er sannsynligheten for at et individ som viser en egenskap bestemt av genet A_1 er homozygot med omsyn på dette genet?
 - c. Hvor stor er sannsynligheten for at avkom etter paring mellom individer som begge viser denne egenskapen, skal få genotypen A_2A_2 ?
 - d. Hvor stor er sannsynligheten for at avkom etter paring mellom individer som fenotypisk er ulike, skal få genotypen A_2A_2 ?

IV. ENDRING AV GENFREKVENSENE.

De forutsetningene som må være til stede for at en populasjon skal være i genetisk likevekt, er sjelden fullt ut oppfylt. I de fleste populasjonene må en rekne med endringer i genfrekvensene som følge av mutasjon, migrasjon eller seleksjon. I tillegg til disse endringene vil en i alle populasjoner av endelig storleik ha forskyvninger i genfrekvensen på grunn av tilfeldigheter. Tilfeldige endringer i genfrekvensen blir kalt genetisk drift.

A. Mutasjon.

Med mutasjon forstår vi ei plutselig endring i arvestoffet. En skiller mellom genmutasjoner (punktmutasjoner), som er ei endring i strukturen av ett enkelt gen, og kromosommutasjoner, som omfatter større eller mindre deler av et kromosom. Det inntreffer f.eks. at et lite segment av et kromosom går tapt (delesjon), eller at to kromosomer bytter ut deler (translokasjon). En tredje type er genommutasjoner, som består i at ett eller flere kromosompar får et ekstra kromosom. Dersom alle kromosompar får ekstrakromosomer, slik at det blir mere enn to kromosomsett, oppstår polyploider. For husdyravlen er det mutasjoner i det enkelte genet som er av størst interesse, men også kromosommutasjoner er påvist.

Mutasjoner oppstår spontant, dvs. uten at det er mulig å påvise noen ytre årsak. Ved bestemte kjemiske eller fysiske påvirkninger (f.eks. temperatursjokk, kjemikalier, stråling) kan mutasjonshyppigheten aukes sterkt, og en får sannsynligvis også mutasjoner som ellers ikke forekommer.

Forekomsten av spontante mutasjoner er best undersøkt hos bananflua (*drosophila*), der en har bestemt mutasjonshyppigheten for ei rekke loci. I gjennomsnitt for de ulike loci er mutasjonsfrekvensen utrekna til å være av storleiksorden 10^{-5} (dvs. en mutasjon pr. 100 tusen gameter). Liknende mutasjonsfrekvenser er funnet for enkelte gener hos mennesker. Jamvel om mutasjonene er relativt sjeldne, vil det store antall loci gjøre at et stort antall individer vil føre ett eller anna gen som er oppstått ved mutasjon i siste generasjon.

De fleste mutasjonene er skadelige. Dette er hva en må vente, ettersom organismene er vel tilpassa til det miljøet de lever i, slik at ei tilfeldig endring av genotypen sjelden kan føre til noen forbedring. Enkelte mutasjoner er letale (dødelige) jamvel i heterozygotisk tilstand, men disse vil forsvinne like raskt som de oppstår og vil derfor ikke by på noe egentlig problem. Mere vanlig er det at virkningen er subletal, dvs. at den gir seg utslag i nedsatt livskraft uten å være direkte dødelig.

Den virkningen som en bestemt mutasjon har på genfrekvensen i vedkommende locus, vil være avhengig av mutasjonshyppigheten. En mutasjon som er så sjelden at den nærmest er en engangshending, har små sjanser til å holde seg i populasjonen, såfremt den ikke er fordelaktig for de individa som fører den. En slik mutasjon kan derfor ikke føre til noen varig endring av populasjonen.

En mutasjon som stadig gjentar seg kan derimot få en betydelig utbreiing. Frekvensen av det muterte genet vil aldri bli så låg at det går tapt på grunn av tilfeldigheter, og det vil oppstå et "mutasjonspress". Dersom genet er "nøytralt" (dvs. hverken skadelig eller fordelaktig), vil frekvensen av det muterte genet ha en tendens til å auke. Men de fleste mutasjonene ser ut til å være reversible, slik at det "nye" genet også kan mutere tilbake, til den opprinnelige forma. Før eller sia vil en derfor komme til et stadium da det oppstår en balanse mellom mutasjonene i de to retningene.

La oss anta at vi har et locus der gena A_1 og A_2 forekommer med genfrekvenser hhv. p og q . A_1 muterer til A_2 med en mutasjonshyppighet u slik at frekvensen av "nye" A_2 -gener er $u \cdot p$. Men samtidig vil enkelte A_2 -gener mutere tilbake til den opprinnelige typen. La oss anta at denne mutasjonen opptrer med frekvensen v . Frekvensen av A_2 -gener som går tapt pr. generasjon er da $q \cdot v$. Balanse mellom de to mutasjonene inntreer når avgangen på A_2 -gener er lik tilgangen, dvs. $u \cdot p = v \cdot q$. Ved å sette inn $p = (1-q)$ og løse med omsyn på q får en:

$$q = \frac{u}{u+v}$$

Mutasjonene har i første rekke betydning i sammenheng med evolusjonslæra, i det de bidrar til å holde den arvelige varia-

sjonen ved like og skaper grunnlag for et naturlig utvalg. I husdyravlen har de interesse særlig fordi en må søke å unngå at de gjør større skade enn nødvendig. En sjelden gang kan det nok inntruffe en mutasjon som kan komme til nytte, om ikke anna så fordi den representerer en "nyhet" (jfr. platinarev), men sjansene til dette er så små at det er lite å satse på.

B. Migrasjon.

Med migrasjon forstår en at en populasjon blir tilført arvestoff fra kilder utafør populasjonen. Den virkningen som ei slik tilføring vil ha på genfrekvensen i populasjonen er avhengig av hvor stort omfang migrasjonen har, og av hvor stor skilnad det er i genfrekvens mellom populasjonen og immigrantene.

Dersom genfrekvensen av et visst gen i populasjonen er q_0 og immigrantene er et tilfeldig utvalg av en populasjon der frekvensen av det samme genet er q_m , så vil genfrekvensen i populasjonen etter migrasjonen være:

$$q_1 = (1-m)q_0 + m \cdot q_m$$

der m angir proporsjonen av immigranter.

Endringa i genfrekvensen som følge av migrasjon blir da:

$$\Delta q = q_1 - q_0 = m(q_m - q_0)$$

Endring av genfrekvensen i en populasjon ved migrasjon er blitt mye nytta i husdyravlen. Omfanget kan variere fra en tilfeldig import av et enkelt avlsdyr og til et program der en tar sikte på å erstatte den gamle populasjonen med en ny. Jamvel om antall immigranter er lite, vil migrasjonen kunne få stor innvirkning på genfrekvensen, fordi immigrantene ofte blir brukt i avlen i langt større omfang enn deres relative andel av populasjonen skulle tilsi.

C. Seleksjon.

I en populasjon vil antall avkom variere fra individ til individ. Dersom individer som har en bestemt egenskap i gjennomsnitt får flere avkom enn individer som mangler denne egenskapen, foregår det en seleksjon for vedkommende egenskap. I populasjoner som er overlatt til seg sjøl, finner det sted en naturlig selek-

sjon for fruktbarhet og levedyktighet. Naturlig seleksjon vil også foregå i tamme populasjoner av f.eks. husdyr, men i dette tilfellet vil virkningene bli modifisert av den seleksjonen som husdyrbrukeren utfører.

Dersom skilnaden i antall avkom mellom ulike individer har sammenheng med individa sin genotype, vil seleksjonen føre til ei endring av genfrekvensen i populasjonen. La oss anta at vi har et genpar A_1A_2 som virker inn på hvor mange avkom et individ får, og slik at individer med genotypen A_1A_1 får flest avkom. Den relative reduksjonen i antall avkom for hver av de andre genotypene blir kalt seleksjonskoeffisienten for vedkommende genotype. Seleksjonskoeffisienten blir alltid angitt i høve til den genotypen som får det største gjennomsnittlige antall avkom. Dersom individa av genotypene A_1A_1 , A_1A_2 og A_2A_2 i gjennomsnitt etterlater seg hhv. 2,50, 2,25 og 1,50 avkom, så er seleksjonskoeffisientene for A_1A_2 og A_2A_2 :

$$A_1A_2: s_1 = \frac{2,50 - 2,25}{2,50} = 0,1$$

$$A_2A_2: s_2 = \frac{2,50 - 1,50}{2,50} = 0,4$$

1. Seleksjon mot et resessivt gen.

Dersom det foreligger fullstendig dominans, er det ikke mulig å skille heterozygotene fra den ene av de to homozygote typene. La oss anta at genet A_1 er dominant over A_2 , og at s er seleksjonskoeffisienten for genotypen A_2A_2 . I en populasjon med tilfeldig paring får en da denne fordelinga av de tre genotypene:

Genotype	A_1A_1	A_1A_2	A_2A_2
Genotypefrekvens før seleksjon	p^2	$2pq$	q^2
Selektiv verdi ("fitness")	1	1	$(1-s)$
Genotypefrekvens etter seleksjon	p^2	$2pq$	$q^2(1-s)$

Genfrekvensene i avkomspopulasjonen blir da:

$$p_1 = \frac{p^2 + pq}{p^2 + 2pq + q^2(1-s)} = \frac{p(p+q)}{p^2 + 2pq + q^2 - q^2s} = \frac{p}{1 - sq^2}$$

$$q_1 = 1 - p_1 = \frac{1 - sq^2 - p}{1 - sq^2} = \frac{q(1-sq)}{1 - sq^2}$$

og endringa i genfrekvensen:

$$\Delta p = -\Delta q = p_1 - p = \frac{p - (p - sq^2p)}{1 - sq^2} = \frac{s \cdot p \cdot q^2}{1 - sq^2}$$

Dersom ingen A_2A_2 -individer får avkom (f.eks. på grunn av at genet A_2 er letalt i homozygotisk form), dvs. $s = 1$, får en:

$$p_1 = \frac{p}{1 - q^2} = \frac{1}{1 + q}, \text{ og } q_1 = \frac{q(1-q)}{1 - q^2} = \frac{q}{1 + q}$$

$$\text{Videre: } \Delta p = -\Delta q = p_1 - p = \frac{1}{1 + q} - (1 - q) = \frac{q^2}{1 + q}$$

Formelen viser at endringa i genfrekvensen i høg grad er avhengig av hva genfrekvensen er på forhånd. Den absolutte endringa er størst når frekvensen av det resessive genet er ca. 2/3. Dersom frekvensen av dette genet allerede er låg, vil en ytterligere reduksjon foregå svært langsomt. Årsaken er at storparten av de resessive gena i dette tilfellet vil være å finne i kombinasjonen A_1A_2 , og dermed unndra seg seleksjon.

Et uønska resessivt gen vil i regelen forekomme med låg frekvens. Dette betyr at nevneren i formelen:

$$\Delta p = \frac{s \cdot p \cdot q^2}{1 - s \cdot q^2} \text{ vil være tilnærma lik } 1,0.$$

og en får: $\Delta p \approx s \cdot p \cdot q^2$,

som ved $s = 1,0$ gir: $\Delta p \approx p \cdot q^2$

Fra formelen $q_1 = \frac{q}{1 + q}$ kan en gå videre:

$$q_2 = \frac{q_1}{1 + q_1} = \frac{q}{1 + 2q},$$

$$q_3 = \frac{q_2}{1 + q_2} = \frac{q}{1 + 3q}$$

.....

$$q_t = \frac{q}{1 + tq}$$

Dette gir: $\frac{1}{q_t} = \frac{1 + tq}{q}$, eller $t = \frac{1}{q_t} - \frac{1}{q}$

Etter denne formelen kan en rekne ut hvor mange generasjoner det vil ta å få frekvensen av et uønska resessivt gen ned til en gitt verdi. Formelen gjelder under forutsetning av at ingen individer av den resessive typen (A_2A_2) får høve til å forplante seg.

Videre er det forutsatt at det ikke forekommer mutasjon fra A_1 til A_2 . Dersom slik mutasjon forekommer, vil det før eller sia oppstå likevekt mellom seleksjon og mutasjon. Denne situasjonen vil inntre når $s \cdot p \cdot q^2 = u \cdot p$, dvs. $q = \sqrt{\frac{u}{s}}$, der u er mutasjonshyppigheten (fra A_1 til A_2). For $s = 1$ gir dette: $q = \sqrt{u}$, dvs. $q^2 = u$. I en populasjon med tilfeldig paring vil likevekt inntre når frekvensen av den resessive typen er lik mutasjonshyppigheten.

2. Seleksjon mot et dominant gen.

Ved seleksjon mot et dominant gen får en disse genotypefrekvensene:

Genotyper	A_1A_1	A_1A_2	A_2A_2
Frekvenser før seleksjon	p^2	$2pq$	q^2
Selektiv verdi	$1-s$	$1-s$	1
Frekvenser etter seleksjon	$p^2(1-s)$	$2pq(1-s)$	q^2

Frekvensen av genet A_1 etter seleksjon blir:

$$p_1 = \frac{p^2(1-s) + pq(1-s)}{p^2(1-s) + 2pq(1-s) + q^2} = \frac{p(1-s)}{1 - s(1-q^2)}$$

Endringa i genfrekvens:

$$\Delta p = p_1 - p = \frac{p(1-s) - p + p \cdot s(1-q^2)}{1 - s(1-q^2)} = \frac{-s \cdot p \cdot q^2}{1 - s(1-q^2)}$$

Ved høge verdier av s vil seleksjon mot et dominant gen raskt føre til at genet blir utrydda. Dersom s er låg, dvs. genotypene A_1A_1 og A_1A_2 har nesten like høg selektiv verdi som A_2A_2 , vil nevneren i uttrykket for Δp være tilnærma 1,0, og en får

$$\Delta p \approx -s \cdot p \cdot q^2$$

dvs. tilnærma samme endring i genfrekvens som ved seleksjon mot det resessive genet.

3. Seleksjon ved overdominans.

Ved overdominans vil heterozygotene ha høgere selektiv verdi enn noen av de to homozygote typene. En kan sette opp dette skjemaet:

Genotyper	A_1A_1	A_1A_2	A_2A_2
Frekvenser før seleksjon	p^2	$2pq$	q^2
Selektiv verdi	$1-s_1$	1	$1-s_2$
Frekvenser etter seleksjon	$p^2(1-s_1)$	$2pq$	$q^2(1-s_2)$

Genfrekvensen i avkomsgenerasjonen blir:

$$p_1 = \frac{p^2(1-s_1) + pq}{1 - p^2s_1 - q^2s_2}$$

og endringa:

$$\begin{aligned} \Delta p = p_1 - p &= \frac{p^2(1-s_1) + pq - p(1-p^2s_1 - q^2s_2)}{1 - p^2s_1 - q^2s_2} \\ &= \frac{-pq(s_1 \cdot p - s_2 \cdot q)}{1 - p^2s_1 - q^2s_2} \end{aligned}$$

Likevekt inntreer når $s_1 \cdot p = s_2 \cdot q$.

Dersom $s_1 = s_2$ får en likevekt når $p = q = 0.5$. Ved overdominans vil seleksjonen ikke føre genfrekvensene mot ekstreme verdier, men tvert imot virke til at de blir opprettholdt på et midlere nivå. Mange granskinger fra de siste åra kan tyde på at overdominans ikke er uvanlig for egenskaper som bestemmer individa sin selektive verdi i populasjoner som er overlatt til seg sjøl.

D. Tilfeldige endringer i genfrekvensen (genetisk drift).

I en populasjon der det ikke forekommer mutasjon, migrasjon eller seleksjon vil de gametene som overfører arveanlegga fra en generasjon til den neste være en tilfeldig prøve av alle de gametene som individa i foreldregenerasjonen har produsert. Hver gamet fører ett gen fra hvert locus, og sannsynligheten for at en gamet fører et bestemt gen er lik frekvensen av dette genet i foreldregenerasjonen. Men tilfeldighetene kan gjøre at genfrekvensen blant de gametene som gir opphav til individene i neste generasjon vil avvike mere eller mindre fra det en skulle vente. Slike endringer i genfrekvensen blir kalt genetisk drift.

Hvor stor genetisk drift en må rekne med, er avhengig av størrelsen av populasjonen. I en populasjon med N individer pr. generasjon vil sambandet mellom generasjonene dannes av $2N$ gameter. La oss anta at frekvensen av genet A_1 i foreldregenerasjonen er p , dette vil da også være sannsynligheten for at en bestemt gamet skal være bærer av genet A_1 . Forekomsten av A_1 eller "ikke- A_1 " er en "enten-eller"-hending (binomial variasjon), og standardavviket (dvs. "feilen") på frekvensen av ei slik hending er gitt av formelen

$$\sigma_p = \sqrt{\frac{p(1-p)}{n}}$$

der p er sannsynligheten for ei viss hending og n er antall observasjoner. I vårt tilfelle er hver gamet én observasjon, og formelen blir:

$$\sigma_{\Delta p} = \sqrt{\frac{p(1-p)}{2N}}$$

Av formelen går det fram at endringer i genfrekvensen på grunn av tilfeldigheter er størst i små populasjoner, og avtar omvendt proporsjonalt med kvadratrota av antall individer pr. generasjon i populasjonen. Videre ser en at endringene avhenger av genfrekvensen i utgangspopulasjonen, de er størst ved midlere genfrekvenser ($p \approx 0.5$) og avtar når frekvensene går mot ekstreme verdier.

Det ligger i sakens natur at endring i genfrekvensen på grunn av genetisk drift kan gå snart i den ene og snart i den andre retningen. I enkelte høve vil det likevel inntreffe at endringa går i samme retning i flere generasjoner etter hverandre. I en liten populasjonen kan det derfor hende at et gen som opprinnelig forekom med en midlere frekvens, på grunn av genetisk drift har fått frekvensen 0 eller 1.0. I det siste tilfellet ($p = 1.0$) sier en at genet er festna i populasjonen. Når et gen først er festna, vil det ikke kunne forekomme noen genetisk drift på vedkommende locus. På dette viset vil tilfeldige endringer i genfrekvensen bære i seg en tendens til å føre genfrekvensene over fra midlere verdier ($p \approx 0.5$) til ekstreme verdier ($p \approx 0$ eller 1.0).

Dersom en populasjon blir oppdelt i mindre undergrupper, vil tilfeldighetene kunne gjøre at frekvensen av et bestemt gen beveger seg i forskjellig retning i de ulike undergruppene. Dette vil gjøre at den genetiske variasjonen innafor hver av undergruppene blir mindre enn i utgangspopulasjonen. Når genfrekvensene går mot ekstreme verdier, vil en automatisk få en auke i frekvensen av homozygote individer og en tilsvarende nedgang i frekvensen av heterozygoter (uttrykket $2pq$, som angir frekvensen av heterozygoter i en populasjon med tilfeldig paring, er størst for midlere verdier av p og q).

Øvingsoppgaver.

1. I en populasjon muterer genet A_1 til A_2 med en hyppighet på $4 \cdot 10^{-5}$. Frekvensen for mutasjoner i motsatt retning (fra A_2 til A_1) er $1 \cdot 10^{-5}$. Hva vil frekvensen av genet A_1 være når det er likevekt mellom de to mutasjonene?

2. I en populasjon A, der et bestemt gen har frekvensen 0.48 vil en auke denne frekvensen ved innkryssing av individer fra en annen populasjon, B, der frekvensen av det samme genet er 0.80. Hva vil venta frekvens av genet i populasjon A være etter at en i to generasjoner har tatt inn hanndyr fra populasjon B, og brukt disse til halvparten av hunndyra i populasjonen?

3. I en populasjon er genet A_2 (i genparet A_1A_2) er resessivt gen som i homozygot tilstand fører til nedsatt forplantingsevne. I en viss generasjon forekommer genet A_2 med frekvensen 0.2. En forutsetter at det ikke forekommer mutasjon eller migrasjon. Hva er venta frekvens av genet i neste generasjon, dersom:
 - a) A_2A_2 -individua i gjennomsnitt etterlater seg halvparten så mange avkom som individer av de andre genotypene?
 - b) Ingen av A_2A_2 -individua får høve til å gi avkom?
 - c) Hvor mange generasjoner vil det i det siste tilfellet (b) ta før frekvensen av genet er redusert til 0.1?

4. I en populasjon med 200 individer pr. generasjon er frekvensen av genet A_1 i en viss generasjon funnet å være 0.40. I neste generasjon er frekvensen 0,44. Er dette forenlig med hypotesen enn at det ikke har foregått mutasjon, migrasjon eller seleksjon i populasjonen?

V. INNAVL OG SLEKTSKAP.

I en populasjon der genfrekvensene er konstante fra generasjon til generasjon vil også genotypefrekvensene være konstante, under forutsetning av at vilkåret om tilfeldig paring er oppfylt. Ved avvik fra tilfeldig paring vil genotypefrekvensene kunne endre seg jamvel om genfrekvensene er konstante. Ei vanlig form for avvik fra tilfeldig paring er at paring mellom individer som er i slekt med hverandre forekommer ofte (eller sjeldnere) enn tilfeldighetene skulle tilsi. Paring mellom slektninger blir kalt slektskapsavl eller innavl.

A. Mål for innavlsgrad.

Individer som er i slekt med hverandre, har en eller flere felles aner. Det er derfor en mulighet for at to slike individer er bærere av gener som kan føres tilbake til ett og samme gen hos en felles ane. Avkom etter paring mellom slektninger kan følgelig på et gitt locus ha to gener som begge er kopier av ett og samme gen i en tidligere generasjon. Gener som er kopier av ett bestemt gen, sies å være identiske i opphav, til forskjell fra gener som bare er identiske i funksjon (dvs. har samme virkning).

Som mål for hvor sterkt et individ er innavla, bruker en innavlskoeffisienten. Innavlskoeffisienten for et individ kan defineres som sannsynligheten for at et tilfeldig valgt genpar hos individet består av to gener som er identiske i opphav. Med utgangspunkt i denne definisjonen kan en utlede formelen for utrekning av innavlskoeffisienten:

Individet Z er avkom etter paring mellom to individer, X og Y, som går tilbake til en felles ane, A, slik det er skissert på fig. 2.

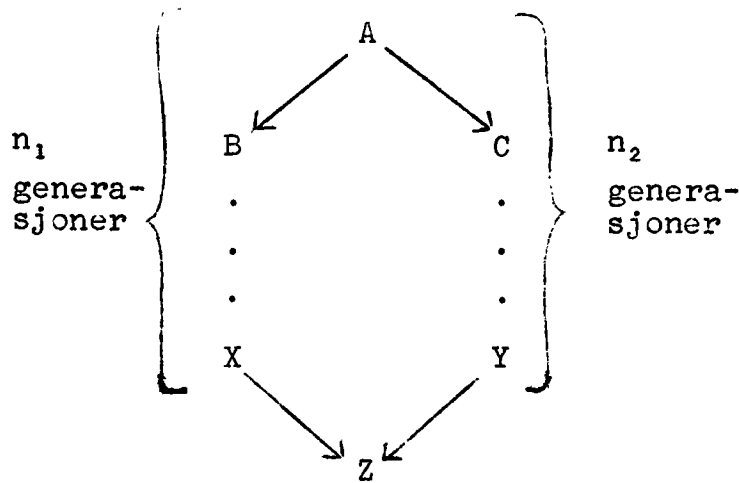


Fig. 2.

Sannsynligheten for at B og C har mottatt ett og samme gen fra A, er lik $\frac{1}{2}$. Sannsynligheten for at B og C har mottatt hvert sitt gen fra A, men at de to gena likevel er identiske i opphav (fra en tidligere generasjon) er lik $\frac{1}{2}F_A$, der F_A er innavlskoeffisienten for A. Sannsynligheten for at de gena (på et gitt locus) som B og C har fått fra A, er identiske i opphav, er derfor lik $\frac{1}{2} + \frac{1}{2}F_A = \frac{1}{2}(1+F_A)$.

Sannsynligheten for at det genet som B har fått fra A skal passere videre til Z er lik $\frac{1}{2}^{n_1}$, der n_1 er antall generasjoner fra B til Z (= antall generasjoner fra A til X). På samme måte er sannsynligheten for at det genet som C har fått fra A skal passere videre til Z lik $\frac{1}{2}^{n_2}$, der n_2 er antall generasjoner fra C til Z (= antall generasjoner fra A til Y).

Sannsynligheten for at Z på et gitt locus skal ha to gener som er identiske i opphav, får en ved å multiplisere sammen de tre sannsynlighetene som er angitt ovafor:

$$F_Z = \frac{1}{2}(1+F_A) \cdot \frac{1}{2}^{n_1} \cdot \frac{1}{2}^{n_2} = \frac{1}{2}^{(n_1+n_2+1)}(1+F_A)$$

der n_1 og n_2 er antall generasjoner fra hvert av foreldra og tilbake til deres felles ane, og F_A er innavlskoeffisienten for denne anen.

Dersom en "felles ane" forekommer mere enn én gang på den ene eller begge sider av anetavla, må en foreta utrekninga særskilt for hver av de sambandslinene som kan trekkes mellom de to foreldra gjennom den felles ane. En må bare passe på at hver slik line ikke passerer mere enn én gang gjennom ett og samme individ. Innavlskoeffisienten med omsyn på vedkommende ane får en ved å summere resultatene av utrekningene for de ulike sambandslinene. På samme måte går en fram dersom foreldra har flere felles aner. Den generelle formelen for innavlskoeffisienten blir derfor:

$$F_Z = \sum \left[\frac{1}{2}^{(n_1+n_2+1)} (1+F_A) \right]$$

der summasjontegnet angir at en skal summere både over alle sambandslinjer gjennom en felles ane og over alle felles aner.

Tabell 1. Innavlskoeffisienter for regelmessige paringsystemer (etter Pirchner, 1968).

Generasjon	Sjølbe-fruktning	Full-søsken	Foreldre x avkom		Halv-søsken
			I	II	
1	0.500	0.250	0.250	0.250	0.125
2	0.750	0.375	0.375	0.375	0.219
3	0.875	0.500	0.500	0.438	0.305
4	0.938	0.594	0.594	0.469	0.381
5	0.969	0.672	0.672	0.484	0.449
∞	1.000	1.000	1.000	0.500	1.000

- I) Paring til den yngste av foreldra.
 II) Paring til den eldste av foreldra.

B. Mål for slektskap.

Også målinga av graden for slektskap tar utgangspunkt i sannsynligheten for at gener er identiske i opphav. Slektskapskoeffisienten for slektskapet mellom to individer kan defineres som sannsynligheten for at et tilfeldig valgt gen hos det ene individet er identisk i opphav med et gen hos det andre individet. (Legg merke til at slektskapskoeffisienten ikke er sannsynligheten for at de to , individa har gener som er identiske i opphav på et tilfeldig valgt locus. Det er heller ikke sannsynligheten for at et tilfeldig valgt

gen hos ett individ er identisk i opphav med et tilfeldig valgt gen på samme locus hos det andre individet).

La oss utlede formelen for slektskapskoeffisienten ved å se på slektskapet mellom individa X og Y i fig. 2. Sannsynligheten for at et tilfeldig valgt gen hos X er kommet fra X og Y sin felles ane A er lik $\frac{1}{2}^{n_1}$ der n_1 er antall generasjoner fra X til A. Sannsynligheten for at det samme genet også har passert fra A til Y er lik $\frac{1}{2}^{n_1} \cdot \frac{1}{2}^{n_2} (1+F_A)$, der n_2 er antall generasjoner fra A til Y (faktoren $1+F_A$ tar omsyn til sannsynligheten for at de to gena i genparet hos A er identiske i opphav). Formelen for slektskapskoeffisienten blir derfor:

$$R_{XY} = \Sigma \left[\frac{1}{2}^{(n_1+n_2)} (1+F_A) \right]$$

Ved å jamføre denne formelen med formelen for innavlskoeffisienten for avkom etter de to individa finner en:

$$R_{XY}/F_Z = 2$$

dvs. innavlskoeffisienten for et individ er halvparten av koeffisienten for slektskap mellom de to foreldra.

Et spesialtilfelle av slektskap mellom to individer har en når det ene individet er ane til det andre. (f.eks. slektskapet mellom X og A i fig. 2). I dette tilfellet blir formelen forenkla, i det $n_2 = 0$, og en får:

$$R_{XA} = \Sigma \left[\frac{1}{2}^{n_1} (1+F_A) \right]$$

der n_1 er antall generasjoner som skiller de to individa (X og A).

C. Innavl i små populasjoner.

I en tallmessig liten populasjon som er stengt for immigrasjon, vil det i løpet av noen generasjoner bli slik at de fleste individa er mere eller mindre i slekt med hverandre. Jamvel om det er tilfeldig paring mellom individa i populasjonen, vil det derfor forekomme paring mellom slektninger, dvs. innavl.

Når en skal vise hvordan den innavlen som oppstår automatisk i en populasjon avhenger av storleiken på populasjonen, er det enklest å ta utgangspunkt i en diploid tvekjønn organisme, der det er fullstendig tilfeldig paring, sjølbefruktning medrekna, og der antall individer, N , er konstant fra generasjon til generasjon. Dersom en ser på ett enkelt locus, vil det da i utgangs-generasjonen (generasjon 0) bli danna $2N$ ulike slag av gameter. En forutsetter at hvert av disse slaga forekommer i samme antall. Sannsynligheten for at et individ i neste generasjon er oppstått etter sammensmelting av to gameter med gener som er identiske i opphav er da $1/2N$, som derfor blir gjennomsnittlig innavls-koeffisient for individa i generasjon 1:

$$F_1 = \frac{1}{2N}$$

La oss så gå videre til neste generasjon. Sannsynligheten for at et individ i denne generasjonen er bærer av to gener (i et gitt locus) som er identiske i opphav fra den nærmest foregående generasjonen, er igjen $1/2N$. Sannsynligheten for at de to gena ikke er identiske i opphav fra nærmest foregående generasjon er da $1-1/2N$. Men også i dette siste tilfellet er det en mulighet for at de kan være identiske i opphav fra en tidligere generasjon. Denne sannsynligheten er lik sannsynligheten for at to tilfeldig valgte gameter produsert av generasjon 0 er bærere av gener som er identiske i opphav, dvs. innavls-koeffisienten for generasjon 1. Innavlskoeffisienten for generasjon 2 blir da:

$$F_2 = 1/2N + (1-1/2N)F_1$$

eller generelt:

$$F_t = 1/2N + (1-1/2N)F_{t-1}$$

Det første leddet i formelen, $1/2N$, referer seg til den "nye" innavlen, og vil være konstant fra generasjon til generasjon. Men i tillegg drar en med seg storparten av innavlen fra tidligere generasjoner.

Ved omskriving av formelen for F_t kan en få et uttrykk for auken i innavlskoeffisient fra en generasjon til den neste:

$$F_t - F_{t-1} = 1/2N(1-F_{t-1})$$

Leddet $(1-F_{t-1})$ angir sannsynligheten for at et individ i generasjon $t-1$ på et gitt locus ikke fører gener som er identiske i opphav, dvs. hvor stor andel av individa som ikke er identisk homozygote på vedkommende locus. Formelen viser at denne andelen blir redusert med faktoren $1/2N$ for hver ny generasjon:

$$1 - F_t = (1-F_{t-1}) - 1/2N(1-F_{t-1}) = (1-1/2N) \cdot (1-F_{t-1})$$

Ved å sette inn for $1 - F_{t-1} = (1-1/2N)(1-F_{t-2})$ osv. vil en til slutt ende opp med:

$$1 - F_t = (1-1/2N)^t,$$

som gir $F_t = 1 - (1-1/2N)^t$

I en populasjon av en organisme som ikke er tvekjønna vil individa ikke kunne være bærere av to gener som er identiske i opphav fra den nærmest foregående generasjonen. Dette vil føre til at innavlen vil bli en generasjon forsinka i høve til det som formlene ovafor angir, ellers blir det nesten ingen skilnad.

En annen og langt mere alvorlig komplikasjon i praksis er at de ulike individa sjelden blir gitt de samme sjansene til å bidra med gameter for produksjon av neste generasjon. I husdyravlen er det f.eks. vanlig at antall hanndyr som brukes til avl er langt mindre enn antall hodyr. Ved å erstatte det virkelige antall avlsdyr, N , med det effektive antall, N_e , kan en likevel bruke de formlene som er utleda foran. Når en har ulikt antall hanndyr og hodyr, vil det effektive antall avlsdyr (med omsyn på auke av innavlsgraden) være det dobbelte av det harmoniske midlet av antall hanndyr (N_m) og antall hodyr (N_f):

$$N_e = 2H, \quad \text{der} \quad \frac{1}{H} = \left(\frac{1}{N_m} + \frac{1}{N_f} \right) / 2$$

som gir:

$$\frac{1}{N_e} = \frac{1}{4N_m} + \frac{1}{4N_f}$$

Også variasjon i antall avkom mellom individer av samme kjønn (ut over det som tilfeldighetene skulle tilsi) vil gjøre det effektive antall avlsdyr mindre enn det virkelige antall.

D. Endring i genotypfrekvensene ved innavl.

Individer som fører to gener av identisk opphav på et gitt locus, vil være homozygote på vedkommende locus. Denne forma for homozygoti vil komme i tillegg til den homozygotien som er til stede i en populasjon i likevekt, slik den er gitt av Hardy-Weinbergs lov. Innavl vil derfor automatisk føre til en auke av homozygotien i en populasjon, og en tilsvarende nedgang i heterozygotien.

For et locus med bare to alleler, A_1 og A_2 , med frekvensene p og q , får en disse genotypfrekvensene:

Genotype	Populasjon i likevekt	Innavl		Sum
		Genpar av ikke-identiske gener	Genpar av identiske gener	
A_1A_1	p^2	$p^2(1-F)$	$p \cdot F$	$p^2 + pqF$
A_1A_2	$2pq$	$2pq(1-F)$		$2pq - 2pqF$
A_2A_2	q^2	$q^2(1-F)$	$q \cdot F$	$q^2 + pqF$
Sum	1	$1-F$	F	1

Reduksjonen i frekvensen av heterozygoter er lik $2pqF$. Av dette følger at innavlskoeffisienten F angir den relative nedgangen i heterozygotien på grunn av innavl. Nedgangen blir fordelt likt på de to homozygote genotypene, som hver får en auke i frekvensen på $p \cdot q \cdot F$.

Øvingsoppgaver

- 1.a. Rekn ut koeffisienten for slektskap mellom to individer A og B som har felles far og felles morfar, og den felles morfar er resultatet av fullsøskenparing.
- b. Rekn ut innavlskoeffisienten for avkom C etter paring mellom disse to individene.

- c. I gjennomsnittet er individa i populasjonen heterozygote i 30 pst. av genpara. Rekn ut venta grad av heterozygoti hos individ C.
2. I en populasjon som formerer seg ved tilfeldig paring, består hver generasjon av 500 hodyr og 50 hanndyr, som alle bidrar likt til neste generasjon (bortsett fra tilfeldige variasjoner i antall avkom).
- Hva er det "effektive antall" avlsdyr i populasjonen?
 - Hvor stor blir innavlskoeffisienten etter to generasjoner?
 - Hvor mange generasjoner går det før innavlskoeffisienten kommer over 0,1?
 - På et gitt locus med to alleler, A_1 og A_2 , er genfrekvensen for A_1 i utgangspopulasjonen 0,3.
Rekn ut venta genotypefrekvenser i populasjonen på det tidspunktet da innavlskoeffisienten er 0,1.

VI. BESKRIVELSE AV POPULASJONER MED OMSYN PÅ
KVANTITATIVE EGENSKAPER.

A. Frekvensfordeling. - Gjennomsnitt og standardavvik.

Ved registrering av en kvantitativ egenskap må en i regelen bruke ei eller anna form for måling, og resultatet av målinga tjener som karakteristikk av det individet målinga gjelder for. Måling av individa i en populasjon vil gi et nytt resultat for hvert individ, slik at en ender opp med nesten like mange måleresultater som antall individer. Dersom en skal gi ei fullstendig beskrivelse av populasjonen, må en angi resultatet av målinga for hvert enkelt individ. For populasjoner med mange individer krever dette stor plass og blir lite oversiktlig. For å lette oversikten kan en slå sammen flere nærstående måleresultater i klasser, og angi antall individer som faller innfor hver klasse. En kan også angi antall individer i hver klasse som proporsjon av det totale antallet, slik at en får ei frekvensfordeling. Ved å framstille frekvensfordelinga grafisk vil en få en visuell beskrivelse av populasjonen med omsyn på vedkommende egenskap.

En kvantitativ egenskap vil i regelen vise ei fordeling som tilnærma svarer til normalfordelinga. I så fall vil populasjonen være fullstendig beskrevet av to parametre, gjennomsnittet og standardavviket.

Gjennomsnittet (dvs. det aritmetriske gjennomsnitt) er vel kjent som uttrykk for en kvantitativ egenskap i en populasjon. Det er gitt av formelen:

$$\bar{X} = \Sigma X_i / N$$

der X_i er resultatet av målinga for det enkelte (det i -te) individet, og N er antall individer. Dersom måleresultatet foreligger i form av ei frekvensfordeling, er gjennomsnittet:

$$\bar{X} = \Sigma f_i \cdot X_i$$

der X_i er gjennomsnittet (eller midtverdien) av den enkelte klassen og f_i er frekvensen for vedkommende klasse.

Gjennomsnittet er en særdeles viktig karakteristikk av populasjonen, men er likevel ufullstendig uten ett eller anna uttrykk for hvor stor variasjonen omkring gjennomsnittet er. Som mål for denne variasjonen bruker en standardavviket, som er definert ved formelen:

$$s = \sqrt{\Sigma(X_i - \bar{X})^2 / (N-1)}$$

Uttrykket $\Sigma(X_i - \bar{X})^2$ er kvadratsummen, dvs. summen av kvadrata på alle avvika mellom resultatet av den enkelte målinga og det aritmetriske gjennomsnittet. Ved omskriving av uttrykket kan en få en formel som er hendigere å bruke praksis:

$$\begin{aligned} (X_i - \bar{X})^2 &= \Sigma(X_i^2 - 2X_i\bar{X} + \bar{X}^2) \\ &= \Sigma X_i^2 - 2\bar{X} \cdot \Sigma X_i + N \cdot \bar{X}^2 \\ &= \Sigma X_i^2 - \frac{2(\Sigma X_i)^2}{N} + \frac{(\Sigma X_i)^2}{N} \\ &= \Sigma X_i^2 - (\Sigma X_i)^2 / N \end{aligned}$$

Når kvadratsummen blir dividert med antall frihetsgrader (N-1), får en middelkvadratet. Standardavviket er kvadratrotta av middelkvadratet.

For en egenskap med normalfordeling vil den proporsjon av populasjonen som avviker mindre enn et gitt antall standardavvik fra gjennomsnittet være:

<1 standardavvik:	68,3	prosent
<2	"	: 95,4 "
<3	"	: 99,7 "

Standardavviket har samme nemning (måleenhet) som gjennomsnittet. Standardavviket er m.a. utgangspunktet når en skal rekne ut standardfeilen på gjennomsnittet. Denne standardfeilen er gitt av formelen: $s_{\bar{X}} = s/\sqrt{N}$. Kvadratet på standardfeilen, $s_{\bar{X}}^2/N$, blir ofte kalt variansen på gjennomsnittet.

I enkelte høve kan det være nyttig å ha et relativt mål for variasjonen, f.eks. dersom en skal jamføre variasjonen for ulike egenskaper. Et slikt mål er variasjonskoeffisienten, som angir

standardavviket som proporsjon av gjennomsnittet (som desimalbrøk eller i prosent).

B. Oppdeling av variasjonen. - Varianskomponenter.

Variasjonen mellom populasjoner og mellom individer av samme populasjon er sjølve grunnlaget for avlsarbeidet. En nærmere analyse av variasjonen er derfor ofte nødvendig.

I mange tilfelle kan en populasjon tenkes delt opp i undergrupper (underpopulasjoner). Slike grupper kan f.eks. være ulike buskaper, avkom etter ulike hanndyr e.l. Den totale variasjonen i populasjonen kan da deles opp i én del som refererer seg til forskjell mellom de ulike gruppene, og én del som angir variasjonen innafor grupper. Oppdelinga foregår ved hjelp av en variansanalyse.

La oss anta at vi har delt inn i k grupper med $n_1, n_2 \dots n_k$ individer, og at observasjonene for individa i de ulike gruppene er symbolisert med X_i , der $i = 1, 2 \dots k$. En har da:

$$\begin{aligned} \text{Total kvadratsum:} \quad KST &= \sum X^2 - (\sum X)^2/N \\ \text{Kvadratsum innen grupper:} KSI &= \sum X_1^2 - (\sum X_1)^2/n_1 \\ &+ \sum X_2^2 - (\sum X_2)^2/n_1 \\ &+ \dots \\ &+ \underline{\sum X_k^2 - (\sum X_k)^2/n_k} \\ KSI &= \sum X^2 - \sum (\sum X_i)^2/n_i \end{aligned}$$

Kvadratsummen "mellom grupper" kommer fram som differansen mellom total kvadratsum og kvadratsum "innen grupper":

$$KSM = KST - KSI = \sum (\sum X_i)^2/n_i - (\sum X)^2/N$$

Det antall frihetsgrader som svarer til hver av de to "del"-kvadratsummene (KSI og KSM) er:

$$\text{Innen grupper: } (n_1-1) + (n_2-1) + \dots + (n_k-1) = N - k$$

$$\text{Mellom grupper: } k - 1$$

Det neste steget i variansanalysen er å rekne ut middelkvadrata:

$$\text{Middelkvadrat "mellom grupper": } s_m^2 = KSM/(k-1)$$

$$\text{Middelkvadrat "innen grupper": } s_i^2 = KSI/(N-k)$$

Formålet med variansanalysen er i regelen å finne ut om det er noen reell forskjell mellom de ulike gruppene. En slik forskjell vil gi seg utslag i at middelkvadratet "mellom grupper" er større enn middelkvadratet "innen grupper". Den statistiske sikkerheten av forskjellen kan undersøkes ved at en rekker ut $F = s_m^2/s_i^2$ og jamfører resultatet med de verdiene som er angitt for det tilsvarende antall frihetsgrader i statistiske tabeller.

Her slutter den vanlige variansanalysen. Men i avlslæra er det ofte nødvendig å gå et steg videre. For å forstå dette må en først få klarhet i hva middelkvadratet "mellom grupper" står for. Dette middelkvadratet er bestemt ikke bare av variasjonen mellom de ulike grupper, men også av antall observasjoner i hver gruppe. Dersom en tenker seg at en auker datamengda ved å måle flere individer i hver gruppe, vil også alle kvadratsummene auke. Men antall grupper er uforandra, og middelkvadratet "mellom grupper" vil derfor auke i samme grad som kvadratsummen.

Matematisk kommer dette best fram dersom formelen for utrekning av kvadratsummen "mellom grupper" skrives om slik:

$$\begin{aligned} \text{KSM} &= \Sigma(\Sigma X_i)^2/n_i - (\Sigma X)^2/N \\ &= \Sigma n_i \cdot \bar{X}_i^2 - (\Sigma n_i \cdot \bar{X}_i)^2/N \end{aligned}$$

Dersom antallet er det samme i alle grupper, dvs. $n_1 = n_2 \dots = n_k = n$ får en:

$$\begin{aligned} \text{KSM} &= n \cdot \Sigma \bar{X}_i^2 - n^2 (\Sigma \bar{X}_i)^2 / n \cdot k \\ &= n [\Sigma \bar{X}_i^2 - (\Sigma \bar{X}_i)^2 / k] \end{aligned}$$

Uttrykket inne i den store parentesen er kvadratsummen på gjennomsnitta for de enkelte gruppene. Middelkvadratet "mellom grupper" er derfor lik n ganger middelkvadratet på disse gjennomsnitta:

$$s_m^2 = n \cdot s_{\bar{X}_i}^2$$

Dette siste middelkvadratet er igjen sammensatt av to deler. Den ene delen er uttrykk for den virkelige variasjonen "mellom grupper", dvs. variansen på de "sanne" gjennomsnitta for hver gruppe. La oss angi denne delen med σ_m^2 . Den andre delen er den

tilleggsvariasjon som en har fordi det observerte gjennomsnitt bare er et tilnærma riktig uttrykk for det "sanne" gjennomsnitt så lenge det bygger på et avgrensa antall observasjoner. Dette tillegget, variansen på et gjennomsnitt, er lik σ_i^2/n , der σ_i^2 er variansen "innen grupper". En får da:

$$s_{\bar{X}_i}^2 = \sigma_m^2 + \frac{\sigma_i^2}{n}, \text{ og}$$

$$s_m^2 = n \cdot s_{\bar{X}_i}^2 = n \cdot \sigma_m^2 + \sigma_i^2$$

Det som vi i første rekke er interessert i, er variansen på de "sanne" gjennomsnitta for hver gruppe, σ_m^2 . Dette blir kalt for varianskomponenten "mellom grupper". En løser med omsyn på σ_m^2 , og får:

$$\sigma_m^2 = \frac{s_m^2 - \sigma_i^2}{n}$$

Det beste estimatet av σ_i^2 (varianskomponenten "innen grupper") er s_i^2 (middelkvadratet innen grupper) fra variansanalysen. En erstatter derfor σ_i^2 i formelen med s_i^2 , og får

$$\sigma_m^2 = \frac{s_m^2 - s_i^2}{n}$$

Den totale variansen i populasjonen er lik summen av varianskomponenten "mellom grupper" og varianskomponenten "innen grupper":

$$\sigma_T^2 = \sigma_m^2 + \sigma_i^2$$

Den andelen som σ_m^2 utgjør av σ_T^2 er et relativt mål for graden av forskjell "mellom grupper", dvs. for graden av likhet mellom observasjoner som hører til i samme gruppe. Denne andelen blir kalt koeffisienten for intraklassekorrelasjon:

$$t = \frac{\sigma_m^2}{\sigma_m^2 + \sigma_i^2}$$

Den har 0 som sin lågste og 1,0 som sin høgste verdi.

Når antall observasjoner varierer fra gruppe til gruppe blir utrekninga noe mere komplisert. Det er utleda formler som gir det "effektive" antall i slike høve. Dersom variasjonen i antall mellom ulike grupper ikke er alt for ekstrem, vil det aritmetriske gjennomsnitt av antalla i de ulike gruppene være ei god tilnærming.

C. Korrelasjon og regresjon.

I avlslæra har en ofte bruk for å studere sammenhengen mellom to ulike egenskaper. Det statistiske målet for graden av samvariasjon (kovariasjon) er korrelasjonskoeffisienten. Den er definert ved formelen:

$$r_{XY} = \frac{\text{Cov}(X \cdot Y)}{\sqrt{\text{Var}(X) \cdot \text{Var}(Y)}}$$

Kovariansen mellom X og Y er lik produktsummen for XY dividert med antall frihetsgrader. Produktsummen er summen av produkta av avvik fra gjennomsnittet for sammenhørende (parvise) verdier av X og Y:

$$\text{Produktsum}(X \cdot Y) = \sum (X_i - \bar{X})(Y_i - \bar{Y})$$

På samme måte som for kvadratsummen kan også denne formelen skrives om til ei form som gir enklere reknearbeid:

$$\sum (X_i - \bar{X})(Y_i - \bar{Y}) = \sum (X_i \cdot Y_i) - \frac{\sum X_i \cdot \sum Y_i}{N}$$

Ved å multiplisere teller og nevner i formelen for r_{XY} med antall frihetsgrader, får en:

$$r_{XY} = \frac{\text{prod.sum}(X \cdot Y)}{\sqrt{\text{kv.sum}(X) \cdot \text{kv.sum}(Y)}}$$

og det er denne forma som blir mest nytta i praksis.

Kvadratet av korrelasjonskoeffisienten angir hvor stor del av variasjonen i den ene av de to variable som kan forklares av variasjonen i den andre. Korrelasjonskoeffisienten kan ha verdier mellom 1,0 og -1,0. Verdier <0 angir negativ korrelasjon, og forteller at de to variablene har en tendens til å svinge i motsatt retning.

Et anna mål for sammenhengen mellom parvise observasjoner er regresjonskoeffisienten. Den angir hvor mye den ene (avhengige) variable endrer seg når den andre (uavhengige) variable blir endra med én enhet, og blir utrekna etter formelen:

$$b_{Y/X} = \frac{\text{Cov}(X \cdot Y)}{\text{Var}(X)} = \frac{\text{prod.sum}(X \cdot Y)}{\text{kv.sum}(X)}$$

Ved å jamføre med formelen for r_{XY} ser en at

$$\begin{aligned} b_{Y/X} &= r_{XY} \cdot \frac{\sqrt{\text{kv.sum}(Y)}}{\sqrt{\text{kv.sum}(X)}} \\ &= r_{XY} \cdot s_Y/s_X \end{aligned}$$

I motsetning til korrelasjonskoeffisienten, som er uten nemning, må regresjonskoeffisienten gis opp med nemning (måleenhet) både for den avhengige og for den uavhengige variable.

I mange høve er det naturleg å se på en variabel som avhengig av to eller flere uavhengige variable samtidig. Dette blir kalt multipel regresjon. For to uavhengige variable får funksjonen denne forma:

$$Y = a + b_1 \cdot X_1 + b_2 \cdot X_2$$

der X_1 og X_2 angir de to uavhengige variable og b_1 og b_2 er partielle regresjonskoeffisienter for Y og på hhv. X_1 og X_2 . Den partielle regresjonskoeffisienten angir regresjonen av den avhengige variable på den ene av de uavhengige variable når de andre uavhengige variable blir holdt konstante.

D. Path-koeffisienter.

Formelen $b_{Y/X} = r_{XY} \cdot s_Y/s_X$ kan skrives om til $r_{XY} = b_{Y/X} \cdot s_X/s_Y$. Av dette går det fram at korrelasjonskoeffisienten angir auken i Y, angitt i antall standardavvik, når X auker med ett standardavvik. En kan si at korrelasjonskoeffisienten er en standardisert regresjonskoeffisient. På tilsvarende måte kan en standardisere partielle regresjonskoeffisienter. En standard partiell regresjonskoeffisient blir ofte kalt path-koeffisient (path= sti, smal vei).

Fordi den refererer seg til ett standardavvik av de variable og ikke til vanlige måleenheter er den uten nemning. På denne måten har path-koeffisienten mye til felles med korrelasjonskoeffisienten. Forskjellen er i første rekke at path-koeffisienten har en bestemt retning, dvs. den skiller mellom årsak og virkning. I grafiske framstillinger blir retningen av path-koeffisienten angitt med en pil, slik det er vist i fig. 3.

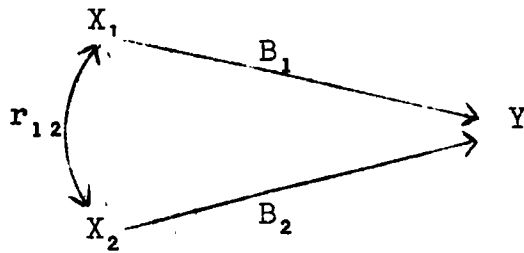


Fig. 3.

Figuren skal forståes slik: Y er en variabel som er "bestemt" av to andre variable X_1 og X_2 . Mellom X_1 og X_2 er det en korrelasjon som er angitt ved korrelasjonskoeffisienten r_{12} . B_1 er en path-koeffisient som angir hvor mye (angitt i standardavvik) Y endrer seg når X_1 endres med ett standardavvik mens X_2 blir holdt konstant. Kvadratet på path-koeffisienten fra X_1 til Y angir hvor stor del av variasjonen i Y som skyldes variasjon i X_1 (med X_2 holdt konstant). På tilsvarende måte er det med kvadratet på path-koeffisienten fra X_2 til Y. Endelig er det en del av variasjonen i Y som ikke kan tilskrives enten X_1 eller X_2 , men de to variablene i fellesskap. Denne delen, uttrykt som proporsjon av den totale variasjon i Y, er lik det dobbelte av path-kjeden gjennom X_1 og X_2 til Y. Den andelen av variasjonen i Y som er bestemt av X_1 og X_2 , hver for seg eller i fellesskap, er da:

$$B_1^2 + B_2^2 + 2B_1 \cdot B_2 \cdot r_{12}$$

Dersom Y er fullstendig bestemt av X_1 og X_2 , vil denne summen være lik 1:

$$B_1^2 + B_2^2 + 2B_1 \cdot B_2 \cdot r_{12} = 1$$

Korrelasjonen mellom X_1 og Y er lik summen av den direkte path fra X_1 til Y (B_1) og pathkjeden gjennom X_2 til Y ($r_{12} \cdot B_2$). Generelt gjelder at når to variable er knytta sammen med ei kjede av pathkoeffisienter og korrelasjonskoeffisienter, kan en rekne ut korrelasjonskoeffisienten mellom de to variable ved å multiplisere sammen de koeffisientene som går inn i kjeden. Dersom det er flere slike kjeder må en foreta utrekninga særskilt for hver kjede og summere resultatata. Det eneste en må passe på er at hver kjede ikke må inneholde mere enn én korrelasjonskoeffisient, og at alle pathkoeffisientene må ha retning bort fra denne korrelasjonskoeffisienten.

Øvingsoppgaver.

1. Data for kullavdrått for fem purker som hver har hatt fire kull viser følgende antall griser:

Purke nr.	Antall griser			
1	8	10	11	9
2	7	11	10	8
3	6	9	8	7
4	10	11	11	12
5	7	6	8	11

- Rekn ut standardavviket for antall griser i kullet, når en ser alle kulla under ett.
- Utfør en variansanalyse av materialet, der variasjonen blir fordelt på "mellom purker" og "innen purker".
- Rekn ut varianskomponenten "mellom purker".
- Rekn ut intraklassekorrelasjonen "innen purker".

2. For et materiale som omfatter 800 avkom etter 40 ulike fedre, hver med 20 avkom, har en beregnet følgende kvadratsummer:

Kvadratsum mellom avkomsgrupper: 11.700

" innen avkomsgrupper: 76.000

- Rekn ut middelkvadrater og varianskomponenter mellom og innen avkomsgrupper.
- Rekn ut intraklassekorrelasjonen innen avkomsgrupper.

3. For et materiale som omfatter 600 parvise observasjoner (x og y) har en oppgitt:

$$\Sigma x = 1200$$

$$\Sigma x^2 = 3000$$

$$\Sigma y = 1320$$

$$\Sigma y^2 = 3600$$

$$\Sigma xy = 2700$$

- a. Rekn ut korrelasjonskoeffisienten mellom x og y, og regresjonskoeffisienten for y på x.
- b. Sett opp regresjonslikninga for y som funksjon av x.

VII. VARIASJON I EN KVANTITATIV EGENSskap.

A. Genetisk og miljøbestemt variasjon.

Ettersom de kvantitative egenskapene i regelen er bestemt av arveanlegg på et stort antall loci, er det ikke mulig å foreta arveanalyser som fører fram til en genetisk formel for det enkelte individet, slik en ofte gjør for kvalitative egenskaper. Slike analyser blir ytterligere vanskeliggjort av at de kvantitative egenskapene i de fleste høve er sterkt påvirka av ei rekke ikke-genetiske tilhøve. For disse ikke-genetiske tilhøva som virker inn på en egenskap nytter en i populasjonsgenetikken samlenemninga "miljø", jamvel om enkelte av dem faller utafor det som en ville kalle "miljø" i daglig tale.

Dette at fenotypen er et resultat av genotype og miljø kan i si mest almenyldige form skrives slik:

$$P = f(G, E)$$

der P, G og E symboliserer hhv. fenotype, genotype og miljø. Formelen uttrykker at fenotypen er en funksjon av genotype og miljø, uten å angi nærmere om G og E virker hver for seg eller i fellesskap.

Det enkleste tilfellet en kan tenke seg er at G og E utøver sin virkning uavhengig av hverandre. I dette tilfellet kan en skrive formelen slik:

$$P = G + E$$

i det en tenker seg at E angir miljøvirkningen som avvik fra gjennomsnittlig miljø for populasjonen. Med denne definisjonen vil den gjennomsnittlige E for populasjonen bli 0, og gjennomsnittlig fenotypeverdi lik den gjennomsnittlige genotypeverdi.

Variasjon i G og E vil hver for seg føre til en like stor variasjon i P. Dersom det ikke foreligger noen korrelasjon mellom genotype og miljø, kan en derfor skrive variansen på fenotypen slik:

$$\sigma_P^2 = \sigma_G^2 + \sigma_E^2$$

En sier at σ_G^2 er den genetiske og σ_E^2 den miljøbestemte (eller miljømessige) varians. Det er grunn til å merke seg at begge de to komponentene blir målt ved den virkningen de har på fenotypen.

B. Oppdeling av den genetiske variasjon.

I et tidligere avsnitt (s. 5) er det forklart at virkningen av et gen kan spaltes opp i én del som skyldes den additive (dvs. den gjennomsnittlige) virkningen av genet, og én del som skyldes særvirkningen i en gitt genkombinasjon. Denne siste delen kan igjen deles opp i dominansvirkningen og epistasivirkningen. Ved overføring av genene fra én generasjon til den neste vil genkombinasjonene bli brutt opp, slik at et gen vil forekomme i en annen kombinasjon hos avkommet enn hos foreldra. Ved de vanlige seleksjonsmetodene er det stort sett bare den additive virkningen av genet som en kan gjøre seg nytte av. Den additive virkningen blir derfor også kalt avlsverdien.

På tilsvarende måte kan en tenke seg den genetiske varians delt opp i komponenter som svarer til den additive virkningen, dominansvirkningen og epistasivirkningen:

$$\sigma_G^2 = \sigma_A^2 + \sigma_D^2 + \sigma_I^2, \text{ der}$$

σ_A^2 = additiv genetisk varians.

σ_D^2 = varians på grunn av dominans, og

σ_I^2 = varians på grunn av epistasi.

C. Oppdeling av den miljøbestemte variasjon.

I avlsmessig sammenheng er den miljøbestemte variasjon å se på som en feilkilde, som virker til å skjule den genetiske forskjellen mellom individa. En er derfor interessert i at denne variasjonen skal være så liten som mulig.

Også den miljøbestemte variansen kan spaltes opp i mindre komponenter. Ved denne oppdelinga kan en legge flere ulike kjennetegn til grunn. En kan f.eks. skille mellom den varians som skyldes påviselige, vel definerte miljøfaktorer (f.eks. fóring, alder etc.), og den delen som ikke kan føres tilbake til noen bestemt årsak. Variasjon av den første typen blir ofte kalt systematisk miljøvariasjon (i motsetning til tilfeldig miljøvariasjon). De systematiske miljøfaktorene kan igjen deles inn i faktorer som er knytta til det "ytre miljø" (f.eks. fóring, klima, hus), og faktorer som gjelder det "indre miljø", dvs. den

fysiologiske tilstanden individet er i (f.eks. alder). Dersom en kjenner virkningen av en bestemt systematisk miljøfaktor kan en fjerne, eller iallfall redusere, virkningen av den gjennom korrigeringer. Slike korrigeringer er mye nytta i det praktiske avlsarbeidet.

Ei form for miljøvariasjon som har stor betydning i dyreriket, og i særlig grad for pattedyr, er den som skyldes maternal effekt eller morsvirkning. Denne miljøvariasjonen gjør seg gjeldende allerede i fosterstadiet, fordi de ulike mødrene ikke gir like gode muligheter for utvikling i denne perioden (prenatalt miljø). Også i den første tida etter fødsel vil en ofte ha en morsvirkning (postnatalt miljø), m.a. på grunn av forskjell i mjølkeevne mellom mødrene. Morsvirkningen gjør at individer med samme mor blir mere like enn slektskapet mellom dem skulle tilsi. Dette er særlig utpreget for kulløsken.

Også andre miljøårsaker kan ha en tendens til å virke likt på bestemte typer av slektninger. Individer med samme far vil f.eks. ofte være født innafor et sterkt avgrensa tidsrom, og periodiske miljøsvingninger vil derfor slå ut i samme retning for alle. En miljøvirkning som er felles for bestemte grupper av slektninger, blir i avlslæra gjerne kalt C-effekt.

D. Gjentaksgard.

For egenskaper som kommer til uttrykk flere ganger i løpet av individet sitt liv (f.eks. mjølkeavdrått hos kyr, kullstorleik hos purker) er det mulig å dele variansen i én del som angir variasjon mellom individer, og én del som angir variasjon fra måling til måling for samme individ (dvs. variasjon "innen individer"). Denne siste variasjonen vil i sin helhet være miljøbestemt, og skyldes miljøvariasjoner av avgrensa varighet (temperære variasjoner). Variasjonen "mellom individer" er dels genetisk og dels miljøbestemt, og den miljøbestemte delen skyldes permanent miljøforskjell mellom ulike individer. Et eksempel på en slik permanent miljøforskjell er den virkningen som miljøet i oppdrettstida har på produksjonsevna senere i livet.

Forholdet mellom variasjonen "mellom individer" og variasjonen "innen individer" bestemmer graden av likhet mellom resultatene av ulike målinger av samme egenskap hos ett og samme individ.

Denne likheten blir oftest uttrykt som intraklassekorrelasjon:

$$r = \frac{\sigma_m^2}{\sigma_m^2 + \sigma_i^2}$$

der σ_m^2 og σ_i^2 angir variansen hhv. mellom og innen individer. Her er σ_i^2 lik den temperære miljøvarians, $\sigma_{E_t}^2$, mens σ_m^2 er satt sammen av den permanente miljøvarians (mellom individer) og den genetiske varians:

$$\sigma_m^2 = \sigma_G^2 + \sigma_{E_p}^2$$

Intraklassekorrelasjonen mellom ulike registreringer av samme egenskap på samme individ blir kalt gjentaksgarden, fordi den angir i hvor stor grad resultatet av ei registrering gjentar seg ved senere registreringer av egenskapen på dette individet. (I engelsk litteratur bruker en uttrykket "repeatability", som ofte er oversatt med "reproduserbarhet").

Den temperære miljøvariasjon vil etter sin definisjon slå ut snart i den ene og snart i den andre retningen for samme individ. For gjennomsnittet av flere målinger vil den tendere til å jamne seg ut, i samsvar med den vanlige formelen for variansen på et gjennomsnitt:

$$\sigma_e^2 = \sigma^2/n$$

Regresjonen av framtidige målinger på gjennomsnittet av n målinger er derfor:

$$b_{Y/\bar{X}} = \frac{\text{Cov}(\bar{X}Y)}{\sigma_{\bar{X}}^2} = \frac{\sigma_m^2}{\sigma_m^2 + \sigma_i^2/n} = \frac{n \cdot r}{1 + (n-1)r}$$

Her er \bar{X} gjennomsnittet av n målinger for samme individ og Y resultatet av ei senere måling for dette individet.

Denne formelen kan nyttes når en vil jamføre individer med ulike antall målinger. Venta framtidig avdrått, Y, reknes ut slik:

$$Y = \bar{P} + b_{Y\bar{X}}(\bar{X} - \bar{P}).$$

der \bar{P} symboliserer gjennomsnittet for populasjonen. Y er også et estimat av individet sin avdråttsevne, dvs. summen av genotype og permanent miljøinnvirkning.

Det er vanlig at variasjon som skyldes påviselige miljøårsaker (f.eks. forskjell mellom buskaper, virkning av alder osv.) såvidt mulig blir fjerna før en rekner ut gjentaksgarden. Komponentene $\sigma_{E_p}^2$ og $\sigma_{E_t}^2$ vil da inneholde bare tilfeldig (ikke identifisert) varians.

I virkeligheten er skillet mellom permanent og temporær miljøvariasjon ikke så skarpt som det er framstilt her. En miljøfaktor kan f.eks. ha en virkning som strekker seg over flere målinger (f.eks. flere laktasjoner hos samme ku), men virkningen er ikke permanent. I slike tilfelle vil korrelasjonen mellom resultatene av målinger som ligger nær hverandre i tid være høgere enn korrelasjonen mellom målinger foretatt med større tidsavstand.

I variansen "innen individer" vil en, i tillegg til den egentlige miljøvariens, også få med den variansen som skyldes at en ikke er i stand til å registrere egenskapen helt nøyaktig. Gjentaksgarden vil derfor være avhengig av hvor nøyaktige målinger en har å bygge på.

E. Korrelasjon mellom genotype og miljø.

Hittil har en forutsatt at genotype og miljø varierer uavhengig av hverandre, og at den fenotypiske variens derfor er lik summen av den genetiske og den miljøbestemte variens. Dersom denne forutsetninga ikke er oppfylt, dvs. dersom det foreligger en korrelasjon mellom genotype og miljø, vil det inngå enda en komponent i den fenotypiske variens. Denne siste komponenten er lik det dobbelte av kovariansen mellom genotype og miljø. En får da:

$$\sigma_P^2 = \sigma_{(G+E)}^2 = \sigma_G^2 + \sigma_E^2 + 2 \text{Cov}(G,E)$$

Det siste leddet, $2 \text{Cov}(G,E)$ kan også skrives som $2 \cdot \sigma_G \cdot \sigma_E \cdot r_{GE}$, der r_{GE} angir korrelasjonskoeffisienten mellom genotype og miljø.

I analyser av data vil det ikke være mulig å isolere den delen av den fenotypiske variens som skyldes korrelasjon mellom genotype og miljø. I de fleste høve vil den havne sammen med den genetiske variens. Dette er heller ikke så bakvendt som det i første omgang kan se ut til. Korrelasjon mellom genotype og miljø vil si at individer med en bestemt genotype bevisst blir gitt et miljø som er bedre eller dårligere enn gjennomsnittet.

Kovarianskomponenten er derfor på en måte avledet av den genetiske varians, og kan ses på som en del av denne.

Ei vanlig form for korrelasjon mellom genotype og miljø i husdyrbruket oppstår ved at fórmengd (og førsammensetning) blir tilpassa etter dyra sitt næringskrav. Ettersom næringskravet i høg grad er bestemt av produksjonsevna, som igjen til dels er genetisk fastlagt, vil det i gjennomsnitt være slik at dyr med en god genotype også får den beste fóringa, dvs. en positiv korrelasjon mellom genotype og miljø. Men en kan også tenke seg situasjoner der denne korrelasjon er negativ.

F. Samspill mellom genotype og miljø.

En annen forutsetning som vi har gjort, er at genotype og miljø adderer sin virkning, dvs. at virkningen av en bestemt miljøforskjell er konstant, uavhengig av hva for genotype den virker sammen med. Dersom dette ikke er tilfelle, slik at den fenotypiske virkningen av ei gitt miljøendring varierer etter genotypen, foreligger det et samspill mellom genotype og miljø. Et slikt samspill kan være av minst to ulike slag. Den mest vanlige forma er at ei viss endring i en miljøfaktor ikke gir like stort utslag for de ulike genotypene, uten at dette fører til noen endring i den innbyrdes rangeringa av genotypene. Det har f.eks. vist seg at virkningen av ei miljøendring ofte er proposjonal heller enn additiv, dvs. at den gir større utslag på individer med "god" enn på individer med "dårlig" genotype. I dette tilfellet foreligger det et samspill i statistisk, men ikke i biologisk forstand. Et virkelig biologisk samspill har en først når den fenotypiske rangeringa av de ulike genotypene er forskjellig fra det ene miljøet til det andre.

De ulike formene for samspill mellom genotype og miljø kan illustreres skjematisk ved disse tabellene, der talla i tabellen angir fenotypen:

a) Ikke samspill

G \ E	I	II
A	1	2
B	3	4

b) Samspill i statistisk,
men ikke biologisk
forstand.

G \ E	I	II
A	1	2
B	3	6

c) Samspill i biologisk
forstand (omrangering
av genotypene).

G \ E	I	II
A	1	4
B	2	3

VIII. GENETISK ANDEL AV VARIASJONEN. - ARVEGRAD.

A. Definisjoner.

Mulighetene for å forbedre en egenskap ved seleksjon er i høg grad bestemt av hvor stor genetisk variasjon det er for egenskapen, og hvor stor del den genetiske variasjon utgjør av den totale (fenotypiske) variasjon. Jo større den genetiske variasjon er i høve til den fenotypiske, jo sikrere kan en dømme om det enkelte individet sin genotype ut fra fenotypen.

Den genetiske andel av den fenotypiske variasjon blir uttrykt ved arvegraden (eng. "heritability," ofte oversatt med "arvbarhet"). Arvegraden er høvet mellom den genetiske og den fenotypiske varians:

$$h^2 = \sigma_G^2 / \sigma_P^2 = \sigma_G^2 / (\sigma_G^2 + \sigma_E^2)$$

Den har 0 som sin lågste og 1,0 som sin høgste verdi. En arvegrad på 0 betyr at all fenotypisk variasjon er miljøbestemt ($\sigma_G^2 = 0$), mens 1,0 i arvegrad sier at det ikke er noen miljøbestemt variasjon ($\sigma_E^2 = 0$).

Dersom det ikke foreligger noen korrelasjon mellom genotype og miljø (jfr. foreg. kapitel) vil den genetiske varians også være kovariansen mellom genotype og fenotype:

$$\text{Cov}(G,P) = \text{Cov } G(G+E) = \text{Cov } (G,G) = \sigma_G^2$$

Arvegraden kan derfor også defineres som regresjonen av genotype på fenotype:

$$h^2 = \sigma_G^2 / \sigma_P^2 = \text{Cov}(GP) / \sigma_P^2 = b_{G/P}$$

Endelig kan en definere arvegraden ved hjelp av korrelasjonen mellom fenotype og genotype:

$$h^2 = b_{G/P} = r_{PG} \cdot \frac{\sigma_G}{\sigma_P} = r_{PG} \cdot h$$

$$h = r_{PG} \quad , \quad h^2 = r_{PG}^2$$

dvs. arvegraden er kvadratet av korrelasjonskoeffisienten mellom fenotype og genotype. Her går det klart fram at arvegraden er et

mål for hvor nøyaktig en kan dømme om genotypen til et individ på grunnlag av fenotypen.

σ_G^2 omfatter her all genetisk varians, enten den skyldes additiv arv, dominans eller epistasi. Formelen angir derfor arvegraden i videste forstand. Men ved de vanlige formene for seleksjon er det bare den additive genetiske variasjon som kan utnyttes. Det er derfor denne delen av den genetiske variasjon som i første rekke har interesse. Arvegraden i snevere forstand er definert som høvet mellom den additive genetiske varians og den fenotypiske varians:

$$h^2 = \sigma_A^2 / \sigma_P^2$$

Når en snakker om arvegraden for en egenskap er det i regelen denne siste definisjonen som er lagt til grunn.

B. Utrekning av arvegraden for en egenskap.

Når en skal rekne ut arvegraden for en egenskap, må en kjenne både den fenotypiske og den (additive) genetiske variansen for egenskapen i den populasjonen en arbeider med. Den fenotypiske variansen kan reknes ut direkte på grunnlag av måleresultata. Men den genetiske varians lar seg ikke rekne ut på samme måten, etter som en aldri kjenner genotypen for det enkelte individet fullt ut (for kvantitative egenskaper). Dersom en kan måle den miljøbestemte varians, kan den genetiske varians reknes ut som differansen mellom den fenotypiske og den miljømessige varians. Men for å kunne måle den miljømessige varians må en ha en populasjon som er genetisk helt ensarta, og en slik populasjon vil sjelden være tilgjengelig. En må derfor nøye seg med indirekte utrekningsmetoder, dvs. metoder som tar utgangspunkt i andre definisjoner av arvegraden enn den som er mest vanlig ($h^2 = \sigma_G^2 / \sigma_P^2$).

Tidligere er vist at arvegraden kan defineres som kvadratet av korrelasjonskoeffisienten mellom fenotype og genotype. Denne korrelasjonskoeffisienten er også pathkoeffisienten fra genotype til fenotype. Den fenotypiske korrelasjonen mellom to individer som er i slekt, er derfor bestemt av arvegraden for den egenskapen det gjelder, og av korrelasjonen mellom genotypene for de to individene (se fig. 4) (under forutsetning av at det ikke foreligger noen miljøkorrelasjon).

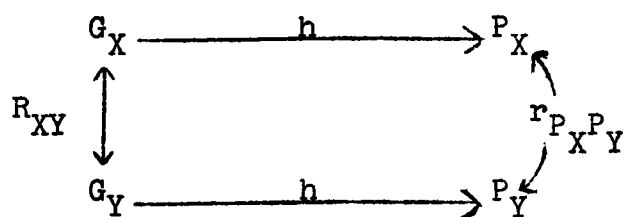


Fig. 4.

Tidligere har vi definert slektskapskoeffisienten mellom to individ som sannsynligheten for at et tilfeldig valgt gen hos det ene individet er identisk i opphav med et gen hos det andre. En annen og like vanlig definisjon er at slektskapskoeffisienten mellom to individ er lik korrelasjonskoeffisienten mellom genotypene deres. En får derfor:

$$r_{P_X P_Y} = h \cdot R_{XY} \cdot h, \text{ som gir}$$

$$h^2 = r_{P_X P_Y} / R_{XY}$$

Den fenotypiske korrelasjon mellom slektninger kan reknes ut direkte. Når en så også kjenner slektskapskoeffisienten, kan en finne arvegraden etter denne formelen.

Ved utrekning av arvegraden er det mest vanlig å bygge på likheten mellom foreldre og avkom eller på likheten mellom søsken (helsøsken eller halvøsken).

1. Likhet foreldre - avkom.

Koeffisienten for slektskap mellom et individ og det ene av foreldra er lik 0,5. Arvegraden kan derfor reknes ut som to ganger den fenotypiske korrelasjon mellom foreldre og avkom.

Denne metoden er grei å nytte i de fleste tilfelle. Men en må være på vakt mot mulige feilkilder. De viktigste av disse er:

a. Metoden forutsetter at det ikke er noen genetisk likhet mellom de to foreldra (dvs. at det er tilfeldig paring). Dersom det foreligger en slik likhet, enten på grunn av slektskap eller på grunn av gjensidig seleksjon av paringspartnere (paring av "de beste

med de beste"), vil koeffisienten for slektskap mellom avkommet og det ene av foreldra være større enn 0,5. For å unngå denne feilen kan en rekne ut korrelasjonen mellom foreldre og avkom som partiell korrelasjon (dvs. korrelasjonen når fenotypen for det andre av foreldra blir holdt konstant). Mere vanlig er det å dele materialet i grupper etter den andre av foreldra, foreta utrekningene for hver gruppe, og etterpå dra sammen resultatene for ulike grupper. Dersom en f.eks. ønsker å rekne ut arvegraden for mjølkeavdrått på grunnlag av likheten mellom avdråttsoppgaver for mødre og døtre, grupperer en de ulike mor-datter-para i grupper der en i hver gruppe har døtre etter en bestemt okse. En rekner ut kvadratsummer og produktsummer for hver gruppe, adderer summene for de ulike gruppene, og rekner ut korrelasjonskoeffisienten på grunnlag av de sammenslåtte kvadrat- og produktsummene.

b. Metoden forutsetter at det ikke foreligger miljøkorrelasjon mellom foreldre og avkom. Denne forutsetninga er ikke alltid oppfylt. Foreldre og avkom vil f.eks. ofte ha stått i samme buskap. Dersom det er en forskjell i miljø mellom ulike buskaper, vil en få en miljølikhet mellom foreldre og avkom, og med det en sterkere fenotypisk korrelasjon enn den som skyldes den genetiske likheten mellom dem.

En kan unngå denne feilen ved å foreta utrekninga særskilt for hver buskap, og etterpå slå sammen resultatene for ulike buskaper. En annen utvei er å uttrykke fenotypen for hvert individ som avvik fra midlet for samtidige individer i samme buskap (buskapsmiddel).

c. I et selektert materiale vil korrelasjonen mellom fenotype og genotype være mindre enn i et materiale der det ikke har foregått noen seleksjon. Foreldra vil i de fleste tilfelle være mere eller mindre sterkt selekterte individer, og den fenotypiske korrelasjonen mellom foreldre og avkom vil i så fall være mindre enn halvparten av arvegraden.

Dersom foreldra utgjør en uselektert gruppe, vil standardavviket "mellom foreldre" være like stort som standardavviket "mellom avkom", og regresjonskoeffisienten av avkom på foreldre er derfor lik korrelasjonskoeffisienten mellom dem. Men regresjonskoeffisienten er ikke påvirket av seleksjon for den uavhengig variable, som i dette tilfelle er fenotypen for foreldra. Ved å erstatte korrelasjonskoeffisienten (i formelen for arvegraden) med regresjons-

koeffisienten av avkom på foreldre vil en derfor kunne unngå at seleksjon blant foreldre virker inn på resultatet.

d. I tillegg til de nevnte feilkildene (og eventuelle andre) vil en alltid måtte rekne med en tilfeldig feil, som snart vil gå i den ene og snart i den andre retningen. Storleiken av den tilfeldige feilen (standardfeilen) på en regresjonskoeffisient er gitt av formelen:

$$s_b = \sqrt{s_{Y \cdot X}^2 / \Sigma(X - \bar{X})^2}$$

der X angir den uavhengige og Y den avhengige variable, og $s_{Y \cdot X}^2$ er variansen i Y når X er holdt konstant (dvs. variansen omkring regresjonslina). Kvadratsummen i X, $\Sigma(X - \bar{X})^2$, kan også skrives som $s_X^2(N-1)$, der N er antall par observasjoner. Videre kan $s_{Y \cdot X}^2$ skrives som $s_Y^2(1-r_{XY}^2)$, fordi r_{XY}^2 angir den del av variansen i Y som skyldes variasjonen i X. En får da:

$$s_b = \sqrt{s_Y^2(1-r_{XY}^2) / s_X^2(N-1)}$$

Ettersom variansen i regelen vil være tilnærma lik for foreldre og avkom ($s_Y^2 \approx s_X^2$), og r_{XY}^2 i de fleste høve vil være liten (sjelden over ca. 0,1), kan en med god tilnærming sette:

$$s_b \approx 1/\sqrt{N}$$

Ved utrekning av arvegraden blir feilen på regresjonskoeffisienten fordobla:

$$s_{h^2} \approx 2/\sqrt{N}$$

Formelen viser at for å få en standardfeil mindre enn f.eks. 0,1 på den utrekna arvegraden må en ha mer enn 400 foreldre-avkom-par.

Den arvegraden som blir utrekna på grunnlag av likheten mellom foreldre og avkom (regresjon eller korrelasjon) vil tilnærma svare til arvegraden i snevreste forstand, dvs. høvet mellom den additive genetiske varians og den fenotypiske varians. Dominans vil ikke bidra til likhet mellom foreldre og avkom, fordi hvert avkom alltid får bare ett av de to gena som utgjør et genpar hos hvert av foreldra. Av epistasivirkningen vil bare en mindre del være felles for foreldre og avkom.

2. Likhet mellom søsken (helsøsken eller halvøsken).

Koeffisienten for slektskap mellom søsken er 0,5 (helsøsken) og 0,25 (halvøsken). Arvegraden kan derfor reknes ut som to ganger den fenotypiske korrelasjon mellom helsøsken, eller som fire ganger den fenotypiske korrelasjon mellom halvøsken. Da antall helsøsken eller halvøsken sjelden vil være avgrensa til to, er det mest praktisk å rekne ut korrelasjonen som intraklassekorrelasjon, der hver hel- eller halvøsken-gruppe utgjør én klasse (variansanalyse):

$$t = \sigma_m^2 / (\sigma_m^2 + \sigma_i^2)$$

Det mest vanlige er å nytte halvøskengrupper, og fortrinnsvis halvøsken med samme far (paternale halvøsken). En unngår da den feilkilden som kan ligge i at individer som har samme mor i noen grad er utsatt for lik maternal effekt, og derfor vil være mere fenotypisk like enn det som skyldes den genetiske likheten mellom dem. Men det kan være grunn til å være på vakt mot andre feilkilder:

a. Metoden forutsetter at det ikke er noen genetisk forskjell mellom mødrene til avkom etter ulike fedre (ut over det som skyldes tilfeldigheter). Dersom det skulle være en slik forskjell, vil likheten mellom halvøsken i regelen bli større enn det som kan tilskrives slektskapet gjennom den felles far. Størst blir feilen dersom det er en tendens til at de beste fedrene er nytta til det beste mor-materialet og omvendt, men jamvel om det ikke er noen slik tendens, vil genetisk forskjell mellom mødre i de ulike gruppene føre til en auke i intraklassekorrelasjonen.

b. Metoden er svært følsom for en eventuell miljøkorrelasjon mellom avkom etter samme far. En slik miljøkorrelasjon kan oppstå dels fordi avkom etter samme far ofte vil være å finne i samme buskap (eller iallfall innafor et geografisk avgrensa område), og dels fordi de vil være samtidige i større grad enn tilfeldig valgte individer i populasjonen (C-effekt). Virkningen av buskapen kan fjernes i variansanalysen ved at en foretar analysen "innen buskaper" (med buskap som overordna klassifisering). Det samme oppnår en ved å uttrykke fenotypen for hvert enkelt individ som avvik fra buskapsmidlet.

c. Seleksjon blant fedre fører til redusert genetisk variasjon mellom fedrene og dermed også mindre varians mellom avkomsgruppene, slik at intraklassekorrelasjonen blir mindre enn det en ville få om det ikke var noen seleksjon. Men seleksjonen er sjelden så nøyaktig at den reduserer den genetiske variasjon i vesentlig grad.

Seleksjon blant avkom vil også i regelen virke til å minske forskjellen mellom de ulike avkomsgruppene, og dermed til lågere intraklassekorrelasjon. Dersom seleksjonen er sterkere i enkelte grupper enn i andre og uavhengig av gjennomsnittet for gruppen, vil seleksjon kunne resultere i auka intraklassekorrelasjon, men det er vanskelig å tenke seg at dette skal kunne inntreffe i praksis.

d. Den tilfeldige feilen på koeffisienten for intraklassekorrelasjon avhenger ikke bare av det totale antall observasjoner (N), men også av hvor mange klasser (k) en har. Det kan vises at en for et gitt antall observasjoner får den minste standardfeil på intraklassekorrelasjonen når antall observasjoner pr. klasse ($n = N/k$) er lik $1/t$. I dette høvet er:

$$s_t \approx t\sqrt{8/k} = \sqrt{8t/N}$$

For å få en standardfeil som er lik $\frac{1}{t}$ eller mindre må en etter dette ha minst 32 klasser (avkomsgrupper) med $1/t$ avkom pr. klasse. Ved utrekning av arvegraden etter formelen $h^2 = 4t$ (halvsøsken-grupper), blir også feilen firedobla.

Ved planlegging av et forsøk eller ei gransking for å rekne ut arvegraden for en egenskap, vil en ikke kjenne t , men kan kanskje på grunnlag av tidligere granskinger gjøre et overslag over hva for område den vil ligge i. En kan da planlegge granskinga slik at en får best mulig utnytting av et gitt antall individer.

C. Arvegrad for et gjennomsnitt av flere målinger.

Tidligere er det forklart at den miljøbestemte varians kan tenkes delt opp i en permanent og en temporær del:

$$\sigma_E^2 = \sigma_{E_p}^2 + \sigma_{E_t}^2$$

Den permanente miljøvirkningen er konstant gjennom individet sitt liv, mens den temporære varierer fra tid til tid. Dersom en har

flere målinger for samme egenskap hos et individ, men til ulik tid (f.eks. flere avdråttår hos kyr) vil den temporære miljøvariansen være lik variansen mellom ulike målinger for samme individ. Når en tar gjennomsnittet av flere målinger, blir denne miljøvariansen redusert, slik at den for et gjennomsnitt av n målinger blir $1/n$ av komponenten for ei måling. Dermed blir også den totale (fenotypiske) varians redusert, og den genetiske andel av variansen auker:

$$h_n^2 = \frac{\sigma_G^2}{\sigma_G^2 + \sigma_{E_p}^2 + \sigma_{E_t}^2/n}$$

Vi merker oss at $(\sigma_G^2 + \sigma_{E_p}^2)$ er variansen "mellom individer" ($=\sigma_m^2$) og $\sigma_{E_t}^2$ variansen "innen individer" ($=\sigma_i^2$).

Gjentaksgraden, som er definert ved formelen

$$r = \sigma_m^2 / (\sigma_m^2 + \sigma_i^2)$$

kan også skrives som:

$$r = \frac{\sigma_G^2 + \sigma_{E_p}^2}{\sigma_G^2 + \sigma_{E_p}^2 + \sigma_{E_t}^2} \equiv \frac{\sigma_G^2 + \sigma_{E_p}^2}{\sigma_P^2}$$

Dette gir: $\sigma_G^2 + \sigma_{E_p}^2 = r \cdot \sigma_P^2$

og: $\sigma_{E_t}^2 = (1-r)\sigma_P^2$

som innsatt i formelen for h_n^2 gir:

$$\begin{aligned} h_n^2 &= \frac{\sigma_G^2}{r \cdot \sigma_P^2 + (1-r)\sigma_P^2/n} \\ &= \frac{\sigma_G^2}{\sigma_P^2[r + (1-r)/n]} = h^2 \frac{n}{1 + (n-1)r} \end{aligned}$$

Av formelen ser en at auken i arvegrad når en har et gjennomsnitt av flere målinger i stedet for ei enkelt måling er størst når verdien av r er låg.

Øvingsoppgaver.

1. Arvegrad for kullstorleik hos purker målt som antall griser i første kull, skal reknes ut på grunnlag av data for mor-datter-par fra tre store buskaper. En har disse opplysningene ($X =$ mødre, $Y =$ døtre).

Buskap:	Antall par (n)	Σ X	Σ X^2	Σ Y	Σ Y^2	Σ $X \cdot Y$
A	200	2600	34600	2800	40270	36500
B	300	3300	37272	3360	39460	37400
C	400	4000	46444	4200	45864	42100

Rekn ut arvegraden ved hjelp av

a. Korrelasjon mødre - døtre.

b. Regresjon døtre/mødre.

Gjør utrekningene først for hele materialet under ett, og deretter ved å rekne ut kvadratsummer og produktsummer for hver buskap og slå sammen resultatata.

c. Rekn ut (tilnærma) hvor stor tilfeldig feil (standardfeil) en må rekne med på den utrekna verdien av arvegraden.

2. En variansanalyse av avdråttsoppgaver (kg mjølkefeitt) for døtre etter 50 okser med 20-30 døtre pr. okse, i alt 1250 døtre, har gitt disse kvadratsummene:

Kvadratsum "mellom okser": 39886

" "innen okser" : 346800

a. Fullfør variansanalysen.

b. Rekn ut arvegraden for mjølkeavdrått, når en forutsetter at all likhet mellom døtrene etter hver okse skyldes slektskap gjennom deres felles far.

c. Hva blir arvegraden for et gjennomsnitt av tre avdråttsår, dersom gjentaksgraden er 0,4?

3. Det skal utføres et forsøk for å finne arvegraden for førutnyttning hos voksende storfe. Tidligere granskinger tyder på at en kan vente en arvegrad på ca. 0,2. En vil bruke halvsøsken-grupper, og disponerer i alt 800 forsøksplasser. Hva er det optimale antall individer (avkom) pr. gruppe?

IX. KORRELASJON MELLOM EGENSKAPER.

A. Fenotypisk og genetisk korrelasjon.

Individa i en populasjon viser i regelen variasjon i et stort antall egenskaper. I enkelte høve opptre egenskapene helt uavhengig av hverandre, men svært ofte er det en sammenheng mellom dem, slik at individer som viser høge verdier av en bestemt egenskap tenderer til å ligge over (eller under) gjennomsnittet også for en annen egenskap. I mange høve ligger denne sammenhengen klart i dagen og er nokså sjølsagt (f.eks. sammenhengen mellom kroppsvekt og visse kroppsmål), i andre høve kan det være vanskelig å finne noen nærliggende biologisk forklaring. Graden av sammenheng blir uttrykt ved hjelp av korrelasjonskoeffisienten mellom resultatene av målinger av de to egenskapene på samme individ. Korrelasjonskoeffisienten refererer seg til fenotypene og blir derfor kalt fenotypisk korrelasjon.

En fenotypisk korrelasjon mellom to egenskaper kan være enten av genetisk eller av miljøbestemt opphav (eller begge deler). Korrelasjon mellom de to genotypene som ligger til grunn for to ulike egenskaper skyldes som regel pleiotropi, dvs. ett og samme gen virker inn på begge de to egenskapene. I enkelte situasjoner kan også kopling av gener føre til liknende resultat. Hver egenskap er oftest påvirka av et stort antall gener, og enkelte av disse kan virke på bare én enkelt egenskap, mens andre har virkning på flere egenskaper. En kan derfor få alle grader av korrelasjon mellom genotypene. Korrelasjon mellom genotypene for to ulike egenskaper blir kalt genetisk korrelasjon.

Hvor sterkt en genetisk korrelasjon mellom to egenskaper skal komme til uttrykk i en fenotypisk korrelasjon er avhengig av i hvor høg grad hver av de to egenskapene er arvelig bestemt, dvs. av arvegraden. Dette er vist i pathdiagrammet i fig. 5. Egenskapene (1) og (2) har arvegraden hhv. h_1^2 og h_2^2 , og path-koeffisientene fra genotype til fenotype blir da h_1 og h_2 . Den genetiske korrelasjon, r_G , vil føre til en fenotypisk korrelasjon som er lik produktet av koeffisientene langs path-kjeden ($h_1 \cdot r_G \cdot h_2$). Dette er den genetiske delen av den fenotypiske korrelasjon.

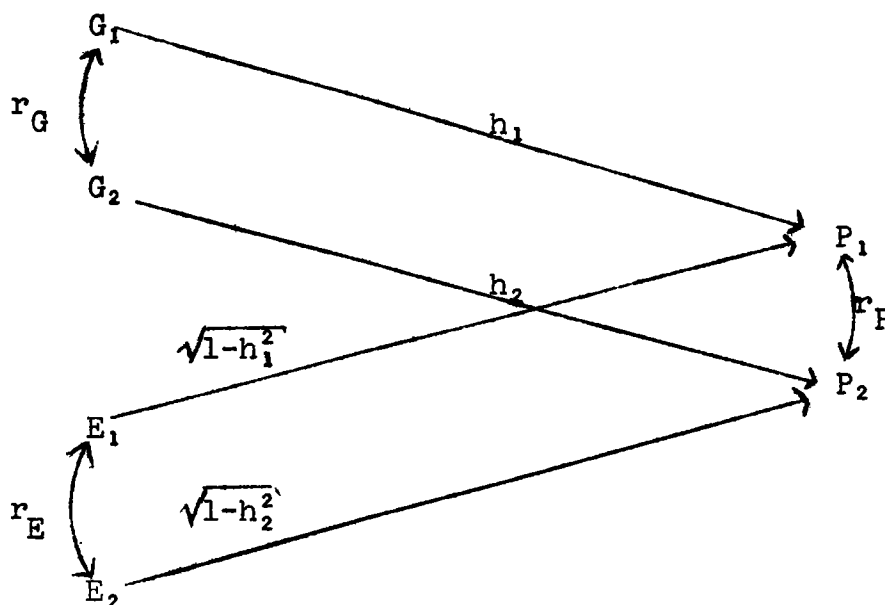


Fig. 5.

En miljøbestemt korrelasjon mellom to egenskaper oppstår dersom miljøtilhøva påvirker de to egenskapene i samme eller i motsatt retning. Path-koeffisienten fra "miljø" til fenotype er lik $\sqrt{1-h^2}$, og den miljøbestemte delen av den fenotypiske korrelasjon blir da $r_E \sqrt{(1-h_1^2)(1-h_2^2)}$. De to delene kan ha samme eller motsatt fortegn. Dersom arvegraden for egenskapene er låg, vil den fenotypiske korrelasjonen være dominert av den miljøbestemte delen.

B. Utrekning av den genetiske korrelasjon.

Den fenotypiske korrelasjonen mellom to egenskaper kan reknes ut direkte på grunnlag av målinger av de to egenskapene hos et større antall individer. Metodene for utrekning av den genetiske korrelasjonen er parallele til de som er omtalt for utrekning av arvegraden.

Den biometriske sammenhengen mellom en egenskap (1) hos et individ (X) og en annen egenskap (2) hos et avkom (Y) er vist i fig. 6.

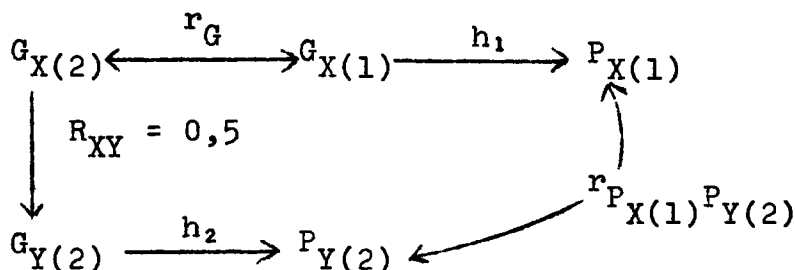


Fig. 6.

Det er her forutsatt at det ikke forekommer noen miljøkorrelasjon mellom de to individa. Korrelasjonen mellom fenotypen for egenskap (1) hos X og egenskap (2) hos Y er da gitt av pathkjeden som knytter dem sammen:

$$r_{P_{X(1)P_{Y(2)}}} = h_1 \cdot r_G \cdot R_{XY} \cdot h_2, \text{ som gir}$$

$$r_G = \frac{r_{P_{X(1)P_{Y(2)}}}}{h_1 \cdot R_{XY} \cdot h_2}$$

Et anna estimat av r_G får en ved å jamføre egenskap (2) hos X og egenskap (1) hos Y.

$$r_G = \frac{r_{P_{X(2)P_{Y(1)}}}}{h_2 \cdot R_{XY} \cdot h_1}$$

I praksis pleier en ofte å nytte det geometriske eller det aritmetriske gjennomsnittet av de to resultatata.

Øvingsoppgaver.

- Den fenotypiske korrelasjon mellom to egenskaper hos samme individ er utrekna til $r_P = 0,18$. Arvegraden for de to egenskapene er hhv. 0,36 (egenskap 1) og 0,16 (egenskap 2). Rekn ut den genetiske korrelasjonen (r_G) mellom egenskapene, når en forutsetter:
 - Ingen miljøkorrelasjon mellom dem.
 - En miljøkorrelasjon på $r_E = 0,2$.
- For to egenskaper med arvegrad hhv. 0,25 (egenskap 1) og 0,36 (egenskap 2) er korrelasjonen mellom egenskap (1) hos foreldre og egenskap (2) hos avkom utrekna til 0,06. En forutsetter at det ikke foreligger noen miljøkorrelasjon mellom foreldre og avkom. Rekn ut den genetiske korrelasjonen mellom egenskapene.

X. SELEKSJON.

Tidligere er det vist at genfrekvensen, og med det også genotypefrekvensen, i en populasjon kan endres gjennom seleksjon. Ved å selektere de individene som ligger nærmest opp til det avlsmålet en har satt seg, og sørge for at de selekterte individene får et større antall avkom enn resten av populasjonen, vil en få en auke i frekvensen av de gunstige genotypene. Seleksjon er det mest effektive hjelpemidlet en har for å forbedre den genetiske kvaliteten av en populasjon.

A. Seleksjonsdifferanse.

Seleksjon foregår på grunnlag av individene sin fenotype. Forskjellen mellom gjennomsnittlig fenotype for de individene som er selektert til avl (\bar{P}_s) og gjennomsnittet for hele populasjonen (\bar{P}) blir kalt seleksjonsdifferansen:

$$S = \bar{P}_s - \bar{P}$$

Når en kjenner fenotypen for de selekterte individene kan en etter dette rekne ut seleksjonsdifferansen.

Som nevnt tidligere vil de fleste kvantitative egenskapene ha ei fordeling som tilnærma svarer til normalfordelinga. Dersom seleksjonen blir gjennomført slik at det er ei skarp grense mellom selekterte og ikke selekterte individer (alle individer som ligger over denne grensa blir selektert til avl mens alle de som ligger under blir vraka), vil seleksjonsdifferansen være bestemt av hvor stor andel som blir selektert (b), og av hvor stort det fenotypiske standardavviket (σ_p) for vedkommende egenskap er. Ved en gitt andel selekterte vil seleksjonsdifferansen være direkte proporsjonalt med standardavviket. Seleksjonsdifferansen kan derfor skrives som:

$$S = i \cdot \sigma_p$$

der i er et mål for seleksjonsintensiteten. Ved å skrive om formelen til

$$i = S/\sigma_p$$

ser en at i er seleksjonsdifferansen angitt i antall standardavvik. En kan derfor si at i er en standardisert seleksjonsdifferanse.

Det kan vises at $i = z/b$, der z er ordinaten til normalkurven i det punktet som angir skillet mellom den selekterte og den ikke-selekterte delen av populasjonen (jfr. fig. 7). Tabeller over ordinatene til ulike punkter på normalkurven er å finne i de fleste lærebøkene i statistikk og i statistiske tabellverk.

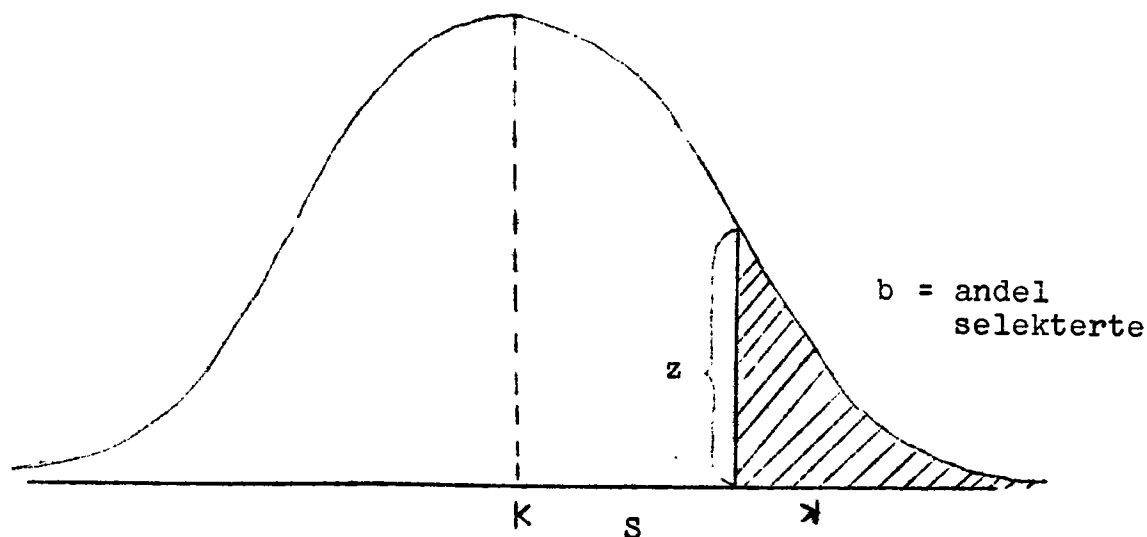


Fig. 7.

Men det er også utarbeidd tabeller som angir i direkte for ulike verdier av b . Et utdrag av en slik tabell er gjengitt på s.65 (tabell 2). Denne tabellen viser også skillet mellom selekterte og ikke-selekterte individer angitt i antall standardavvik over eller under populasjonsgjennomsnittet, og ordinaten til normalkurven på dette punktet.

Ved å framstille i som funksjon av b grafisk vil en se at sammenhengen er tilnærma rettlina for verdier av b mellom 0,2 og 0,9.

De standardiserte seleksjonsdifferansene som er angitt i denne tabellen gjelder bare dersom seleksjonen er foretatt blant et relativt stort antall individer. Dersom dette antallet er lite ($< \approx 50$), vil verdien bli noe mindre enn det tabellen viser. Tabell 3 angir standardiserte seleksjonsdifferanser når seleksjonen foregår blant hhv. 10, 20 og 40 individer.

Tabell 2. Seleksjonsdifferanse (i standardavvik) ved seleksjon i en stor populasjon med normalfordeling.

Andel selekterte	Abscisse (i standardavvik) til skillet mellom selekterte og ikke-selekterte	Ordinat til normalkurven i skillet mellom selekterte og ikke-selekterte	Seleksjonsdifferanse (i standardavvik)
b	d	z	i = z/b
0.01	2.33	0.027	2.66
0.02	2.05	0.048	2.42
0.03	1.88	0.068	2.27
0.04	1.75	0.086	2.15
0.05	1.65	0.103	2.06
0.10	1.28	0.176	1.76
0.15	1.04	0.233	1.56
0.20	0.84	0.280	1.40
0.30	0.52	0.348	1.16
0.40	0.25	0.386	0.96
0.50	0.00	0.399	0.80
0.60	-0.25	0.386	0.64
0.70	-0.52	0.348	0.50
0.80	-0.84	0.280	0.35
0.90	-1.28	0.176	0.20

Tabell 3. Seleksjonsdifferanse (i standardavvik) ved seleksjon blant hhv. 10, 20 og 40 individer fra en populasjon med normalfordeling (delvis etter Pirchner, 1967).

Andel selekterte	Seleksjonsdifferanse (i standardavvik)		
	N = 10	N = 20	N = 40
0,05		1,87	1,96
0,10	1,54	1,64	1,69
0,20	1,27	1,33	1,36
0,30	1,06	1,11	1,13
0,40	0,89	0,93	0,95
0,50	0,74	0,77	0,78
0,60	0,60	0,62	0,63
0,70	0,45	0,48	0,49
0,80	0,32	0,34	0,34
0,90	0,17	0,18	0,19

B. Genetisk virkning av seleksjon.

Når en skal rekne ut hvor stor genetisk endring en kan vente ved seleksjon for en kvantitativ egenskap, tar en utgangspunkt i at arvegraden kan defineres som regresjonen av genotype på fenotype. Den gjennomsnittlige genotypeverdi for individer som er selektert for avl kan derfor estimeres som:

$$\bar{G}_s = \bar{G} + (\bar{P}_s - \bar{P}) \cdot h^2$$

der \bar{G} er gjennomsnittlig genotypeverdi for populasjonen. Dette gir:

$$\Delta G = \bar{G}_s - \bar{G} = (\bar{P}_s - \bar{P}) \cdot h^2 = S \cdot h^2$$

Genfrekvensen blant avkom er den samme som i den selekterte delen av foreldrepopulasjonen, og den additive genetiske verdi blir derfor også den samme. Dersom h^2 angir arvegraden i snevreste forstand (bare additiv genetisk varians inkludert), vil formelen ovafor gi venta endring i avlsverdi (dvs. additiv genetisk verdi) av populasjonen som følge av seleksjon i én generasjon.

Ved seleksjon i en normalfordelt populasjon med skarpt skille mellom selekterte og ikke-selekterte, dvs. $S = i \cdot \sigma_p$, får en:

$$\Delta G = i \cdot \sigma_p \cdot h^2$$

Dette kan skrives om til:

$$\Delta G = i \cdot \sigma_p \cdot h \cdot \frac{\sigma_G}{\sigma_p} = i \cdot \sigma_G \cdot h$$

Tidligere er vist at kvadratrota av arvegraden, h , er lik korrelasjonen mellom fenotype og genotype. Dette leder fram til den generelle formelen:

$$\Delta G = i \cdot \sigma_G \cdot r_{PG}$$

der r_{PG} er korrelasjonen mellom den genotypen en ønsker å forbedre og det fenotypiske uttrykket som seleksjonen blir bygd på (seleksjonskriteriet). I denne formen er formelen gyldig også i de tilfella da seleksjonen ikke foregår på grunnlag av fenotypen for det individet en ønsker å vurdere, men på grunnlag av fenotypen for en slektning (eller slektninger). Dette kommer en tilbake til i neste kapitel.

Formelen angir venta genetisk framgang i løpet av én generasjon. Fra et praktisk synspunkt er det ofte mere interessant å kunne rekne ut venta framgang pr. tidsenhet. Dette får en ved å dividere med den gjennomsnittlige tidsavstand mellom foreldre og avkom, det en gjerne kaller generasjonsintervallet. En får da:

$$G/L = \frac{i \cdot \sigma_G \cdot r_{PG}}{L}$$

der L er generasjonsintervallet.

Formelen viser at den genetiske endringa pr. tidsenhet er bestemt av:

- a. Intensiteten av seleksjonen (representert ved i).
- b. Storleiken av den genetiske variasjon (σ_G).
- c. Nøyaktigheten av seleksjonen (r_{PG}).
- d. Lengda av generasjonsintervallet (L).

Ved planlegging av avlsarbeidet er oppgaven ofte å finne fram til en optimal kombinasjon av disse fire faktorene. Den øvre grensa for seleksjonsintensiteten er bestemt av formeringsevnen, dvs. antall avkom pr. individ. Denne formeringsevnen varierer sterkt fra art til art, og kan i enkelte høve aukes ved tekniske hjelpemidler (f.eks. kunstig sædovertføring.) Ved å la avlsdyra leve så lenge som mulig kan en få flere avkom etter hvert dyr, og dermed auke seleksjonsintensiteten. Storleiken av den genetiske variasjonen vil i regelen være gitt. Nøyaktigheten av seleksjonen kan aukes ved at en reduserer den miljøbestemte variasjonen, enten ved å standardisere miljøet eller ved å korrigere for virkningen av kjente miljøfaktorer. For enkelte egenskaper kan en auke nøyaktigheten ved å nytte gjennomsnittet av flere målinger, men dette kommer ofte i konflikt med ønsket om å holde generasjonsintervallet så kort som mulig.

C. Indirekte seleksjon.

Enkelte egenskaper er av en slik natur at seleksjon på grunnlag av fenotypen for vedkommende egenskap vanskelig lar seg praktisere. Dette gjelder bl.a. kjønnsavgrensa egenskaper (for individer av det kjønnnet som ikke viser egenskapen) og egenskaper som først kommer til uttrykk på et sent stadium i individa sitt liv (holdbarhet)

eller etter at de er slakta (slaktekvalitet). For noen egenskaper er registreringa så komplisert og arbeidskrevende at den ikke er praktisk gjennomførlig i det omfang som ville være nødvendig (f.eks. forutnytting hos mjølkekyr). Endelig har en de egenskapene som har så låg arvegrad at seleksjon på grunnlag av fenotypen er lite effektivt. I slike høve kan det være aktuelt å gå en omvei, og å bygge seleksjonen på en annen egenskap, som i seg sjøl kan være av liten interesse, men som viser sammenheng med den egenskapen en ønsker å forbedre. Dette blir kalt indirekte seleksjon.

Nøyaktigheten av indirekte seleksjon er avhengig av arvegraden for den egenskapen som danner grunnlaget for seleksjonen (seleksjonskriteriet) og av den genetiske korrelasjonen mellom denne egenskapen (2) og den egenskapen en ønsker å forbedre (1). Dette er illustrert i fig. 8.

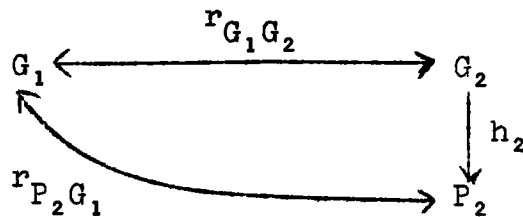


Fig. 8.

Nøyaktigheten, dvs. korrelasjonen mellom P₂ og G₁, er lik produktet av koeffisientene langs pathkjeden som knyttet P₂ til G₁:

$$r_{P_2G_1} = h_2 \cdot r_{G_1G_2}$$

Indirekte seleksjon vil være mere nøyaktig enn direkte seleksjon når $h_2 \cdot r_{G_1G_2} > h_1$. Dette kan inntreffe når arvegraden for egenskap (2) er vesentlig høgere enn for egenskap (1), samtidig som den genetiske korrelasjon mellom de to egenskapene er høg. Dersom arvegraden er 0,04 for egenskap (1) og 0,25 for egenskap (2), vil indirekte seleksjon for å forbedre egenskap (1) være mere nøyaktig enn direkte seleksjon om den genetiske korrelasjonen er høgere enn 0,40.

D. Seleksjon for flere egenskaper.

Den økonomiske verdien av individa i en populasjon vil i regelen være bestemt av mange egenskaper. Anstrengelsene for å forbedre populasjonen gjennom seleksjon må derfor være retta mot alle disse egenskapene.

Når en skal forbedre flere egenskaper i en populasjon, kan en i prinsippet gå fram på en av tre måter:

1. Tandem-metoden.

Denne metoden går ut på at en forsøker å forbedre de ulike egenskapene i tur og orden, dvs. bare én egenskap i hver generasjon. I denne generasjonen vil framgangen i vedkommende egenskap bli den samme som om dette var den eneste egenskapen en ønska å forbedre. Men i de andre generasjonene får en ingen framgang for denne egenskapen (forutsatt at det ikke foreligger noen genetisk korrelasjon mellom egenskapene). Dersom en har n like viktige og innbyrdes uavhengige egenskaper med samme arvegrad, vil framgangen pr. generasjon for hver egenskap bli $1/n$ av den framgangen en ville få om en hadde bare én egenskap å ta omsyn til.

2. Minstekrav-metoden.

Etter denne metoden søker en å forbedre alle egenskapene samtidig, ved at en for hver egenskap fastsetter et visst minstekrav som individa må fylle for å bli selektert til avl. Dette minstekravet må settes så høgt som mulig, men likevel ikke høgere enn at det blir igjen et tilstrekkelig antall individer til å holde populasjonen ved like. Dersom b er den andelen av populasjonen som må brukes til avl, og en ønsker å forbedre n like viktige og innbyrdes uavhengige egenskaper samtidig, så er den andelen som må oppfylle minstekravet for hver egenskap:

$$b' = \sqrt[n]{b}$$

En kan da rekne ut venta genetisk framgang ved å finne hva for seleksjonsdifferanse (angitt i standardavvik, tab. 2) som svarer til de aktuelle verdiene av b' . For en populasjon der 25 prosent av individa trengs til avl ($b = 0,25$) får en f.eks.:

Antall egenskaper	Andel selekterte for hver egenskap	Seleksjonsdifferanse (i standardavvik)
n	$b' = \frac{1}{\sqrt{n}} \sqrt{0,25}$	i
1	0,25	1,27
2	0,50	0,80
3	0,63	0,60
5	0,76	0,41
10	0,87	0,24

Minste-krav-metoden vil alltid være mere effektiv enn tandem-metoden, og skilnaden vil være større jo flere egenskaper en ønsker å forbedre. Men også minstekravmetoden har sine åpenbare svakheter. Den fører bl.a. til at et individ som er framifrå i mange egenskaper kan bli vraka som avlsdyr dersom det ligger litt under minstekravet i én eneste egenskap.

3. Indeksmetoden.

Det prinsippet som ligger til grunn for denne metoden er at svakheter med omsyn til enkelte egenskaper skal kunne oppveges av fordeler i andre egenskaper. En forsøker derfor å se alle egenskapene under ett, og gjøre seg opp ei mening om totalverdien av individet. Den genetiske totalverdien kan skrives slik:

$$H = a_1 \cdot G_1 + a_2 \cdot G_2 + \dots + a_k \cdot G_k = \sum a_i \cdot G_i$$

der G_i er genotypen for en bestemt egenskap (uttrykt som avvik fra gjennomsnittet for populasjonen), og a_i er den økonomiske nettoverdien av å forbedre egenskapen med én enhet. Dersom egenskapene er innbyrdes uavhengige, kan G_i best estimeres som $G_i = P_i \cdot h_i^2$, og seleksjonsindeksen blir:

$$I = a_1 \cdot h_1^2 \cdot P_1 + a_2 \cdot h_2^2 \cdot P_2 + \dots + a_n \cdot h_n^2 \cdot P_n$$

dvs. den enkelte egenskapen skal ha vekt etter produktet av den økonomiske verdi og arvegraden.

Dersom egenskapene er like viktige og har samme arvegrad, vil framgangen pr. generasjon for hver egenskap bli $1/\sqrt{n}$ ganger den framgangen en ville ha fått om en hadde bare en egenskap å ta omsyn til. Indeksmetoden er derfor \sqrt{n} ganger så effektiv som tandem-metoden. Den vil også alltid være mere effektiv enn

minstekrav-metoden. Praktiske omsyn kan likevel gjøre at en må foretrekke minstekravmetoden, f.eks. når de ulike egenskapene kommer til uttrykk på ulik tid eller når enkelte egenskaper er så kostbare å registrere at registreringa må avgrenses til et utvalg av individer. I slike høve er det aktuelt å foreta en foreløpig seleksjon på grunnlag av en enkelt egenskap etter minstekrav-metoden, mens den endelige seleksjonen bygger på en seleksjonsindeks.

Dersom det foreligger en fenotypisk eller genetisk korrelasjon mellom de egenskapene som en ønsker å endre gjennom seleksjon, vil disse korrelasjonene virke inn på koeffisientene i seleksjonsindeksen. Når en kjenner korrelasjonskoeffisientene mellom ulike egenskaper, er det mulig å utlede seleksjonsindeksen matematisk også i slike høve. Virkningen av å ta omsyn til korrelasjoner mellom egenskapene ved utledning av seleksjonsindeksen vil være at egenskaper som er positivt genetisk korrelerte får større vekt enn de ellers ville ha fått. Egenskaper som viser positiv miljøkorrelasjon vil derimot i regelen få mindre vekt. Positiv miljøkorrelasjon vil videre tendere til å auke vekta på de viktigste og redusere vekta på de mindre viktige egenskapene. I enkelte høve kan korrelasjoner mellom egenskapene jamvel føre til at vekta for en egenskap får et anna fortegn i indeksen enn det den økonomiske verdien skulle tilsi.

Øvingsoppgaver

1. I en populasjon kjenner en disse parametrene for en bestemt egenskap:

Gjennomsnitt:	105 enheter
Fenotypisk standardavvik:	12 enheter
Genetisk standardavvik:	6 enheter

 - a. Beregn den genetiske endring i løpet av én generasjon ved å selekterte avlsdyr blant den beste halvpart av populasjonen?
 - b. Hvor strengt må en selektere (i prosent) for å oppnå en genetisk framgang på 1 pst. pr. år, dersom generasjonsintervallet er 4 år?

2. Mellom to egenskaper, A og B, er det en genetisk korrelasjon på 0,7. Arvegraden er 0,36 for A og 0,16 for B, og de fenotypiske standardavvik hhv. 15 enheter (A) og 8 enheter (B).
- Hva er mest effektivt for å forbedre egenskap B, enten å selektere på grunnlag av A eller på grunnlag av B?
 - Hvor stor genetisk framgang pr. generasjon kan en vente i A og B ved å selektere de beste 30% etter egenskap A?
3. I en svinepopulasjon vil en selektere for rask vekst og tynt ryggspekk. La oss gå ut fra at det har like stor økonomisk betydning å redusere ryggspekktjukna med 1 mm som å auke veksthastigheten (tilvekst pr. dag) med 10 g. Arvegraden er 0,3 for veksthastighet og 0,5 for ryggspekktjukn. Vi forutsetter at de to egenskapene er fenotypisk og genetisk uavhengige. Konstruer en seleksjonsindeks for veksthastighet og ryggspekktjukn.

XI. UTVALGSMETODER.

A. Masseutvalg.

Med masseutvalg forstås en at avlsverdien til det enkelte individet blir vurdert på grunnlag av individet sin egen fenotype. Uttrykket fenotypeutvalg blir ofte nytta i samme betydning som masseutvalg.

Som vist tidligere kan en estimere genotypen (eller avlsverdien) for et individ ut fra fenotypen etter denne formelen:

$$G = \bar{P} + (P - \bar{P})h^2$$

der \bar{P} er gjennomsnittlig fenotype for populasjonen, P er fenotypen for det individet vurderinga gjelder, og h^2 er arvegraden for vedkommende egenskap. Dersom egenskapen er registrert flere ganger for samme individ, bør P være gjennomsnittet for alle disse målingene, etter at en først har fjerna virkningen av kjente miljøfaktorer. Ved jamføring av individer med ulikt antall målinger må en ved utrekning av sannsynlig avlsverdi etter formelen ovafor nytte den verdi av arvegraden som gjelder for det antall målinger individet har.

Tidligere er vist at virkningen av seleksjon er proporsjonal med nøyaktigheten, uttrykt som korrelasjonen mellom genotypen til individet og den fenotypen som danner grunnlaget for seleksjonen. Ved masseutvalg er denne korrelasjonen lik kvadratrot av arvegraden. Når arvegraden auker med f.eks. 20 prosent som følge av flere målinger pr. individ, mere nøyaktige målemetoder eller standardisering av miljøet, vil effektiviteten av seleksjonen (ved masseutvalg) ikke auke med 20 prosent, men med ca. 10 prosent. På bakgrunn av formelen ovafor kan dette synes merkelig. Forklaringa er at redusert miljøvariasjon fører med seg en reduksjon av den fenotypiske variasjonen, dvs. lågere verdi av differansen $(P - \bar{P})$.

Masseutvalg er en enkel og grei utvalgsmetode, og den gir ofte muligheter for intensiv seleksjon. Men for egenskaper med låg arvegrad er den lite nøyaktig. Den kan ikke nyttes for egenskaper som først lar seg registrere etter slakting, og for kjønnsavgrensa egenskaper kan den nyttes bare for individer av det ene kjønnet. For egenskaper som kommer til uttrykk på et så sent stadium at individa alt er nytta til avl, vil masseutvalg ofte ha liten virkning.

B. Utvalg etter avstamning.

Avstamning er ei samlenemning for alle anene til et individ. Anetavle er ei skjematisk oppstilling av avstamningen (anene).

Alle de arveanlegga et individ har, er kommet fra foreldra, som i sin tur har fått dem fra sine foreldre. Nøyaktigheten av utvalg etter avstamning vil derfor i høg grad være bestemt av hvor nøyaktig en kjenner arveanlegga til foreldra. For kvantitative egenskaper kommer dette til uttrykk gjennom arvegraden. Men etter som hvert avkom tar imot en tilfeldig halvpart av arveanlegga til hvert av foreldra (ett gen fra hvert genpar), vil jamvel et fullstendig kjennskap til genene hos foreldra ikke kunne gi noen helt sikker opplysning om genotypen til et enkelt avkom.

Ved å tegne opp et path-diagram over sammenhengen mellom genotypene for foreldre og avkom vil en se at 25 prosent av variasjonen i genotypen for avkommet er "bestemt" av genotypen for hvert av foreldra ($0,5^2 = 0,25$). Tilsammen svarer altså genotypene til foreldra for 50 prosent av variasjonen i genotype for avkommet. De resterende 50 prosent skyldes tilfeldigheter som følge av spaltinga av arveanlegg (dvs. variasjon mellom avkom etter det samme foreldreparet). Den øvre grensa for nøyaktighet ved utvalg etter avstamning er derfor lik $\sqrt{0,5} \approx 0,7$. Dette er en teoretisk grense, som en når dersom en kjenner genotypen til foreldra fullt ut (dvs. $h^2 = 1,0$ for kvantitative egenskaper). I praksis vil dette aldri være tilfelle, og nøyaktigheten vil i regelen være vesentlig mindre enn 0,7 (jfr. tab. 4).

Tabell 4. Nøyaktigheten av utvalg etter avstamning.

(Omarbeidd etter Skjervold, 1966).

Aner	$h^2 = 0,1$	$h^2 = 0,3$	$h^2 = 0,5$	$h^2 = 0,7$
Foreldre	0,22	0,39	0,50	0,59
Foreldre + besteforeldre	0,27	0,43	0,53	0,61
Mor, farmor, mormor	0,19	0,32	0,41	0,47

Et viktig spørsmål i samband med utvalg etter avstamning er hvor stor vekt en skal legge på de ulike anene. Den vekta som en ane bør ha, er bestemt av:

- a) Slektskapet mellom vedkommende ane og det individet vurderinga gjelder. Dette slektskapet, og dermed sannsynligheten for "felles arveanlegg", blir halvert for hver generasjon en går bakover i anetavla. En bør derfor alltid legge hovedvekta på de nærmeste anene, og en har som regel lite igjen for å gå lenger tilbake enn til besteforeldre.
- b) Hvor godt kjennskap en har til avlsverdien for vedkommende ane. Dess bedre kjennskap en har til en ane, dess mere vekt bør en legge på denne anen.
- c) Kjennskapet til mellomliggende aner. Dersom en ane forekommer i anetavla bak en ane som en kjenner avlsverdien for, skal den tillegges mindre vekt enn om avlsverdien til den mellomliggende anen er mindre godt kjent. Dette fører til at det kan være riktig å legge ulik vekt på to aner som forekommer i samme anerekke og som en har like sikre opplysninger om.

Dersom en kjenner avlsverdien til foreldra fullt ut, gir resten av anetavla ingen tilleggsinformasjon, og skal derfor heller ikke ha noen vekt ved vurderinga. Jo mindre kjennskap en har til foreldra, jo mere vekt er det riktig å legge på fjernere aner. Men dersom en har like sikre opplysninger om alle anene, skal hvert av besteforeldra alltid ha mindre enn halvparten av den vekta som en legger på hvert av foreldra. For egenskaper med middels høg arvegrad ($h^2 = 0,2-0,4$) høver det å gi hvert av foreldra om lag tre ganger så stor vekt som hvert av besteforeldra.

Dersom en har like sikre målinger (av fenotypen) for individet sjøl som for hver av anene, vil utvalg etter avstamningen alltid være mindre nøyaktig enn masseutvalg. Skilnaden i nøyaktighet er minst for egenskaper med låg arvegrad.

Utvalg etter avstamning er en utvalgsmetode som kan nyttes allerede på et tidlig stadium av individet sitt liv. Metoden er brukbar også for kjønnsavgrensa egenskaper, og i mange høve også for egenskaper som blir registrert først etter slakting. Som grunnlag for et foreløpig utvalg av avlsdyr (og bruksdyr) vil vurdering av avstamningen alltid være et nyttig hjelpemiddel, jamvel om det senere blir supplert med andre metoder.

I mange høve er det aktuelt å kombinere opplysninger om individets egen fenotype med opplysninger om anene. På denne måten vil en kunne auke nøyaktigheten av seleksjonen med opptil 20-25 prosent jamført med utvalg der en bygger bare på egen fenotype (masseutvalg). Vinninga er størst for egenskaper med låg arvegrad. Men en må være varsom så en ikke legger for stor vekt på anene (og dermed for lita vekt på egen fenotype).

C. Utvalg etter avkomsgransking.

Avkomsgransking er ei vurdering av avlsverdien til et individ på grunnlag av fenotypen for et større eller mindre antall avkom. Slektskapet mellom et individ og hvert enkelt avkom er det samme som mellom individet og hvert av foreldra. Men et individ har bare to foreldre, mens det kan ha mange avkom, og de feilkildene som en må rekne med for hvert enkelt avkom, vil derfor kunne jamne seg ut på gjennomsnittet.

Nøyaktigheten av avkomsgransking kan utledes av et diagram over den biometriske sammenhengen mellom genotypen (G) for et individ og den gjennomsnittlige fenotype (\bar{P}_n) for n avkom. I fig. 9 angir $G_1, G_2 \dots G_n$ og $P_1, P_2 \dots P_n$ henholdsvis genotypen og fenotypen for hvert enkelt avkom, h^2 er arvegraden for egenskapen, og t er intraklassekorrelasjonen mellom individer som har det ene av foreldra felles.

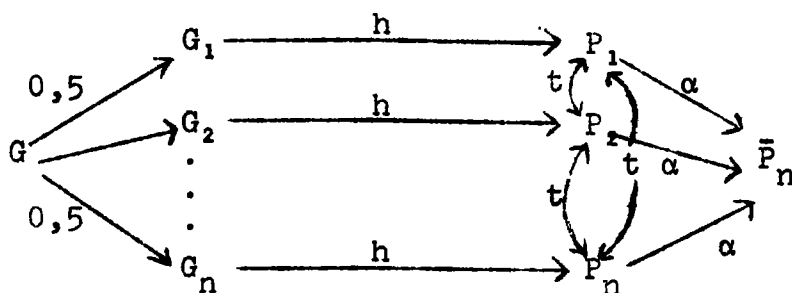


Fig. 9.

\bar{P}_n vil være fullstendig bestemt av $P_1, P_2 \dots P_n$. En har derfor:

$$n \cdot \alpha^2 + n(n-1) \alpha \cdot t \cdot \alpha = 1$$

$$n \cdot \alpha^2 [1 + (n-1)t] = 1$$

$$\alpha = \frac{1}{\sqrt{n[1 + (n-1)t]}}$$

$R_{\bar{P}_n G}$ er summen av produkta langs de path-kjedene som forbinder G med \bar{P}_n . Hver av disse kjedene er $0,5 \cdot h \cdot \alpha$, og summen blir:

$$R_{\bar{P}_n G} = n \cdot 0,5h \cdot \alpha = \frac{0,5 \cdot n \cdot h}{\sqrt{n[1 + (n-1)t]}} = \sqrt{\frac{0,25 \cdot n \cdot h^2}{1 + (n-1)t}}$$

Dersom det ikke foreligger noen likhet mellom avkomma ut over den som skyldes den genetiske likheten mellom halvsøsken, er $t = 0,25h^2$, og formelen blir:

$$R_{\bar{P}_n G} = \sqrt{\frac{0,25 \cdot n \cdot h^2}{1 + (n-1)0,25h^2}} = \sqrt{\frac{n}{n + 4/h^2 - 1}}$$

Av formelen går det fram at nøyaktigheten auker med stigende antall avkom og med stigende verdi av arvegraden. Når arvegraden går mot 1,0, går verdien av $R_{\bar{P}_n G}$ mot $\sqrt{n/(n+3)}$. Når n går mot uendelig, går $R_{\bar{P}_n G}$ mot 1,0 uavhengig av arvegraden.

Tabell 5. Nøyaktigheten av utvalg etter avkomsgransking.

Antall avkom	$t = 0,25h^2$		$t = 0,25h^2 + 0,1$	
	$h^2 = 0,2$	$h^2 = 0,6$	$h^2 = 0,2$	$h^2 = 0,6$
1	0,22	0,39	0,22	0,39
2	0,31	0,51	0,28	0,49
5	0,46	0,68	0,40	0,61
10	0,59	0,80	0,46	0,68
25	0,75	0,90	0,52	0,73
100	0,92	0,97	0,56	0,76
∞	1,00	1,00	0,58	0,77

Det kan iblant være av interesse å finne ut hvor mange avkom en må ha for at avkomsgransking skal være nøyaktigere enn egen fenotype som grunnlag for seleksjon. Utvalg etter avkomsgransking

er sikrere enn fenotypeutvalg (masseutvalg) når $R_{\bar{P}_n G} > h$, dvs.

$$\sqrt{\frac{0,25 \cdot n \cdot h^2}{1 + (n-1) \cdot 0,25h^2}} > h^2$$

som løst med omsyn på n gir:

$$n > \frac{4 - h^2}{1 - h^2}$$

Denne brøken er størst ved høye verdier av arvegraden (h^2). Den går mot 4 når h^2 går mot 0. Når h^2 går mot 1,0, går brøken mot uendelig. For midlere verdier av arvegraden ($h^2 = 0,2-0,4$) vil en måtte ha 5-6 avkom før avkomsgranskning er et mere nøyaktig mål for individets avlsverdi enn det dets egen fenotype er.

I mange høve kan det foreligge en korrelasjon mellom medlemmer av en avkomsgruppe ut over den som skyldes slektskapet ved at de har det ene av foreldra felles. Årsaken kan være miljølikhet av ulike slag (f.eks. morsvirkning for maternale søsken), eller det kan være slektskap også gjennom det av foreldra som ikke er felles for hele gruppa (f.eks. ved at gruppa er ei blanding av helsøsken og halvsøsken). Dette gir seg utslag i en auke i verdien av t ($t > 0,25h^2$) og dermed en auke i nevneren i den første formelen for $R_{\bar{P}_n G}$, dvs. en reduksjon av $R_{\bar{P}_n G}$. Dersom t er vesentlig større enn $0,25h^2$, hjelper det lite på nøyaktigheten å auke antall avkom.

En kan også se på $R_{\bar{P}_n G}$ som en pathkoeffisient fra G til \bar{P}_n . Korrelasjonen mellom gjennomsnittlig fenotype for to uavhengige og like store grupper av avkom kan derfor uttrykkes som:

$$R_{\bar{P}_n \bar{P}'_n} = R_{\bar{P}_n G} \cdot R_{\bar{P}'_n G} = R_{\bar{P}_n G}^2 = \frac{0,25 \cdot n \cdot h^2}{1 + (n-1)t}$$

Dette er også koeffisienten for regresjon av \bar{P}'_n på \bar{P}_n (fordi \bar{P}_n og \bar{P}'_n har samme standardavvik). Generelt gjelder at regresjonen av framtidige avkom (uavhengig av antall) på gjennomsnittet av n tidligere avkom er:

$$b = \frac{0,25 \cdot n \cdot h^2}{1 + (n-1)t}$$

På grunnlag av gjennomsnittet for n avkom kan en derfor rekne ut venta gjennomsnitt for framtidig avkom som:

$$\bar{P}' = \bar{P} + b(\bar{P}_n - \bar{P})$$

der \bar{P} er gjennomsnittet for populasjonen.

Avlsverdien for det individet som avkomsgranskinga gjelder for, angitt som avvik fra populasjonsgjennomsnittet, vil være det dobbelte av det tilsvarende avvik for framtidige avkom. Denne avlsverdien kan derfor estimeres slik:

$$G = \bar{P} + 2 \cdot b(\bar{P}_n - \bar{P})$$

Avkomsgransking er, iallfall i teorien, den utvalgsmetoden som gir mulighet for den sikreste vurderinga av individet sin avlsverdi. De viktigste forutsetningene for et sikkert resultat er at en har et rimelig stort antall avkom, og at det ikke foreligger miljøforskjeller mellom de ulike avkomsgruppene. Videre bør en passe på å få med alle avkom, ikke bare de beste. Ulike sterk seleksjon i de ulike gruppene forut for granskinga kan føre til nokså misvisende resultat.

Avkomsgransking er mest aktuelt for handyr, fordi disse får et større antall avkom over ei avgrensa tid. Særlig nyttig er avkomsgransking som grunnlag for utvalg for kjønnsavgrensa egenskaper som bare kommer til uttrykk hos hodyra, og for egenskaper som først kan registreres etter slakting. Ellers har denne utvalgsmetoden størst interesse for egenskaper med låg arvegrad.

D. Utvalg på grunnlag av slektninger i sideledd. - Familieutvalg.

Med familie forstår en i populasjonsgenetikken ei samling individ som hører til samme generasjon og har samme innbyrdes slektskap. I husdyravlen blir nemninga oftast nytta om helsøsken og halvsøsken.

Når en skal drøfte nøyaktigheten av utvalg etter slektninger i sideledd, må en ta omsyn til om det individet en ønsker å vurdere, er medrekna i gjennomsnittet for familien eller ikke.

1. Individet ikke medrekna i familiegjennomsnittet.

Denne situasjonen er illustrert i fig. 10. G_x er genotypen for det individet som skal vurderes, \bar{P}_n den gjennomsnittlige fenotype

for n familiemedlemmer (helsøsken eller halvøsken), og g slektskapskoeffisienten mellom familiemedlemmene, de andre symbolene er som i fig. 9. G_x og \bar{P}_n er knytta sammen med n pathkjeder som hver består av $g \cdot h \cdot \alpha$. Fra drøftinga av fig. 9 har en:

$$\alpha = 1/\sqrt{n[1+(n-1)t]}$$

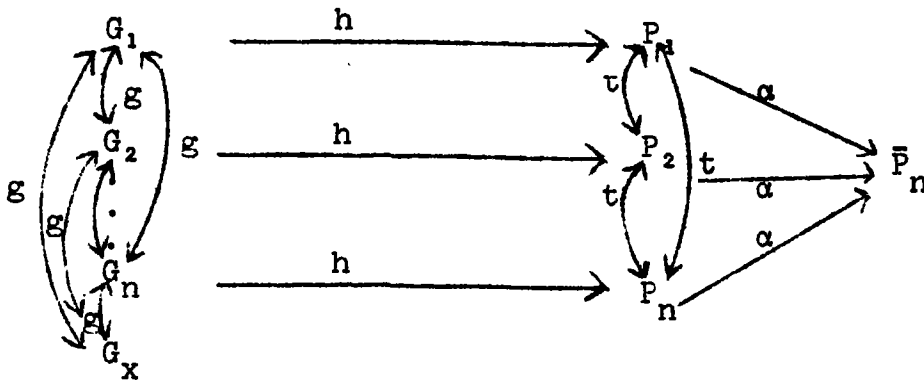


Fig. 10.

For korrelasjonen mellom G_x og \bar{P}_n får en da:

$$R_{G_x \bar{P}_n} = n \cdot g \cdot h \cdot \frac{1}{\sqrt{n \cdot [1+(n-1)t]}} = g \sqrt{\frac{n \cdot h^2}{1+(n-1)t}}$$

Tabell 6. Nøyaktigheten ved utvalg etter slektninger i sideledd (helsøsken og halvøsken).

Antall søsken	Helsøsken		Halvøsken	
	$h^2 = 0,2$	$h^2 = 0,6$	$h^2 = 0,2$	$h^2 = 0,6$
n				
1	0,22	0,39	0,11	0,19
2	0,30	0,48	0,15	0,26
5	0,42	0,58	0,23	0,34
10	0,51	0,64	0,29	0,40
100	0,68	0,70	0,46	0,49
∞	0,71	0,71	0,50	0,50

Nøyaktigheten auker med aukende verdi av g (slektskapskoeffisienten), n (antall familiemedlemmer) og h^2 (arvegraden), men avtar med

aukende verdi av t (den fenotypiske korrelasjonen mellom familie-medlemmene). Når n går mot uendelig og det ikke foreligger noen miljøkorrelasjon mellom medlemmer av samme familie (dvs. $t = g \cdot h^2$), vil $R_{G_x \bar{P}_n}$ gå mot \sqrt{g} , dvs. mot $\approx 0,7$ for helsøsken og $0,5$ for halv-søsken (jfr. tab. 6). Dette er de øvre grensene for nøyaktigheten av utvalg etter slektninger i sideledd når det individet som skal vurderes ikke er medrekna i gjennomsnittet for slektingene. Med en miljøkorrelasjon mellom familiemedlemmene vil nøyaktigheten gå raskt nedover.

2. Individet medrekna i familiegjennomsnittet.

Også dette tilfellet kan illustreres av fig. 10, dersom en ser bort fra G_x . La genotypen for det n -te medlemmet av familien (G_n) være den vi skal vurdere. G_n er knytta til \bar{P}_n med n path-kjeder. Den ene av disse kjedene består av $h \cdot \alpha$, de $(n-1)$ andre av $g \cdot h \cdot \alpha$. En får derfor:

$$R_{G_n \bar{P}_n} = h \cdot \alpha + (n-1) \cdot g \cdot h \cdot \alpha = h \cdot \alpha [1 + (n-1)g]$$

Ved å sette inn $\alpha = 1/\sqrt{n[1+(n-1)t]}$ får en:

$$R_{G_n \bar{P}_n} = h \frac{1 + (n-1)g}{\sqrt{n[1+(n-1)t]}}$$

Nøyaktigheten auker med aukende verdi av g (slektskapskoeffisienten) og h^2 (arvegraden), og avtar med aukende verdi av t . Hvordan ulike verdier av n (antall familiemedlemmer) virker, avhenger av verdien av h^2 , g og t . Når n er uendelig stor, spiller det ingen rolle om individet er medrekna i familiegjennomsnittet eller ikke, og grenseverdiene blir derfor de samme for de to alternativa (\sqrt{g}). For $n = 1$, dvs. ingen familiemedlemmer i tillegg til individet, faller familieutvalg sammen med masseutvalg, dvs. $R_{G_n \bar{P}_n} = h$. For egenskaper med høg arvegrad vil derfor familieutvalg være mere nøyaktig ved små enn ved store familier (jfr. tab. 7).

Tabell 7. Nøyaktigheten av familieutvalg (dvs. korrelasjonen mellom familiegjennomsnittet og genotypen til ett av familiemedlemmene).

Antall søsken	Helsøsken		Halvsøsken	
	$h^2 = 0,2$	$h^2 = 0,6$	$h^2 = 0,2$	$h^2 = 0,6$
n				
1	0,45	0,78	0,45	0,78
2	0,46	0,72	0,35	0,64
5	0,51	0,70	0,37	0,55
10	0,62	0,70	0,38	0,52
100	0,69	0,71	0,47	0,50
∞	0,71	0,71	0,50	0,50

Utvalg etter slektninger i sideledd høver godt for egenskaper som er kjønnsavgrensa og egenskaper som først kan registreres etter slakting. Ellers er det mest aktuelt i de tilfelle da det lar seg gjøre å registrere gjennomsnittlig fenotype for en familie uten å måtte gå vegen om hvert enkelt familiemedlem (f.eks. eggproduksjon hos høns).

Øvingsoppgaver.

1. For tre kyr i samme buskap har en disse avdråttsoppgavene (korrigert for kjente miljøfaktorer):

A: Gj.snitt for 2 år: 6000 kg

B: " " 3 år: 6500 kg

C: " " 6 år: 4500 kg

Gjennomsnittsavdråtten for buskapen er 5000 kg.

a. Rekn ut venta avlsverdi for hver av de tre kyrne, når arvegraden for mjølkeavdrått for ett enkelt år er 0,20. og gjentaksgarden er 0,40.

b. Kyrne A, B og C er hhv. mor, farmor og mormor til et individ, og avdråttsoppgavene deres skal nyttas for vurdering av avlsverdien til dette individet. Angi (tilnærma) hvor stor relativ vekt en bør legge på hver av de tre anemødrene. Grunngi svaret.

2. a. Hvor mange avkom må en ha i hver avkomsgruppe for at utvalg etter avkomsgransking skal være sikrere enn masseutvalg, dersom arvegraden er 0,50?
b. Hvor høg må arvegraden være for at masseutvalg skal være sikrere enn utvalg på grunnlag av 10 avkom?

3. En okse er avkomsgranska på grunnlag av avdråttsoppgaver for 30 døtre, og gjennomsnittsavdråtten er utrekna til 5600 kg mjølk, mens populasjonsmidlet er 5300 kg. Arvegraden for mjølke-mengd settes til 0,25 og fenotypisk standardavvik til 800 kg.
 - a. Hva er venta avdrått for framtidige døtre etter denne oxen, under forutsetning av konstant miljø?
 - b. Hva er sannsynlig avlsverdi for oxen i høve til populasjonsmidlet?
 - c. Hvor stor genetisk seleksjonsdifferanse vil en få ved å selektere de beste 20 prosent blant okser som er avkomsgranska på 30 døtre?

XII. VIRKNING AV LANGVARIG SELEKSJON.

Tidligere er vist hvordan venta genetisk virkning av seleksjon kan reknes ut på grunnlag av seleksjonsintensitet, nøyaktighet og genetisk variasjon. Dette er utleda rent teoretisk og det er nærliggende å stille spørsmålet om det har vist seg å holde stikk, også når seleksjonen strekker seg over flere generasjoner. Dette kan en få svar på gjennom seleksjonsforsøk eller ved å granske virkingen av seleksjon i data fra vanlig praksis.

A. Seleksjonsforsøk.

Et seleksjonsforsøk er et forsøk som blir utført for å studere virkingen av et visst seleksjonsopplegg. Forsøket kan være lagt opp med to grupper (liner) der det blir selektert i motsatt retning, eller det kan være ei gruppe der det blir foretatt seleksjon, og ei kontrollgruppe (uten seleksjon). Det kan også være aktuelt å utføre seleksjonsforsøk for å jamføre virkingen av ulike seleksjonsmåter. For å gi mulighet for å vurdere hvor stor del av et eventuelt utslag som kan skyldes tilfeldigheter, bør det være minst to grupper (paralleller) for hvert forsøksledd. Samtidig bør antall individer pr. gruppe i hver generasjon være såvidt stort at en i rimelig grad er sikra mot tilfeldige feil (genetisk drift).

De fleste seleksjonsforsøk er utført med insekter (bananfluer, tribolium) som dyremateriale, de har stor formeringsevne og kort generasjonsintervall, samtidig som de krever lite plass. Det er også utført mange seleksjonsforsøk med mus. De fleste husdyrarter egner seg dårlig til dette formålet, særlig på grunn av langt generasjonsintervall og fordi kostnadene med å holde et tilstrekkelig stort antall individer under forsøksstilhøve er uoverkommelige. I det siste er det likevel satt i gang flere seleksjonsforsøk med husdyr, i første rekke fjørfe og svin.

Et spørsmål som ofte blir stilt i samband med seleksjon, er dette: Hvor lenge kan seleksjonen fortsette før en når et punkt da det ikke er mulig å komme videre, eller da det iallfall vil gå langsommere? Resultata fra seleksjonsforsøka gir ikke noe entydig svar på dette, bl.a. fordi svaret sannsynligvis i høg grad avhenger av hvor intensiv og hvor nøyaktig seleksjonen er. I seleksjonsforsøk med bl.a. bananflue og mus har det vært liten

tendens til nedgang i virkningen av seleksjon før det har gått 15-20 generasjoner eller mer. Enkelte forsøk har fortsatt i over 50 generasjoner uten at mulighetene for seleksjon har vært uttømt, i andre høve har en nådd et "platå" etter 25-30 generasjoner. I husdyravlen, der en i regelen selekterer for flere egenskaper samtidig, og derfor vil ha relativt svak seleksjon for hver enkelt egenskap, er det neppe noen grunn til å frykte for noen nedgang i virkningen av seleksjon på lang tid.

Enkelte forsøk med seleksjon i to retninger har gitt tydelig usymmetrisk utslag for seleksjon. I de fleste tilfelle har utslaget vært mindre i positiv enn i negativ retning, men det fins også døme på det motsatte. Usymmetrisk utslag kan forklares ved at genfrekvensene i utgangspopulasjonen avviker fra 0,5, dersom de "positive" genene alt fra starten er i overvekt, må en vente mindre utslag for seleksjon i positiv enn i negativ retning. Fullstendig eller partiell dominans for "positive" gener kan virke på samme måte. Det er også mulig å forklare usymmetrisk utslag som et resultat av naturlig seleksjon. Flere seleksjonsforsøk har vist at sterk seleksjon i det lange løp kan føre til nedsatt forplantingsevne og levedyktighet. Dersom forsøket er utført uten kontrollgruppe kan ei usymmetrisk endring ganske enkelt skyldes at miljøet har endra seg i løpet av forsøkestida. Det er derfor viktig at en alltid har med ei kontrollgruppe.

I seleksjonsforsøk er det funnet at når seleksjonen opphører i en gruppe som har vært selektert i ei bestemt lei over flere generasjoner, får en ofte et markert tilbakeslag. Dette er forklart med at seleksjonen i det lange løp virker til å bygge opp gunstige genkombinasjoner (epistasi), og at disse etterhvert går i oppløsning når seleksjonspresset blir borte.

B. Gransking av genetisk framgang på grunnlag av data fra vanlig praksis.

Seleksjonsforsøk har vist at jamvel om en gjør sitt ytterste for å holde miljøtilhøva i forsøket konstante, er det ikke til å unngå at en får både periodiske svingninger og trendmessige endringer. I praksis, der en har langt mindre muligheter for å kontrollere miljøfaktorene, vil dette gjøre seg enda sterkere gjeldende. For å kunne studere virkningen av seleksjon i data fra

det praktiske husdyrbruket må en derfor først finne en metode som gjør det mulig å skille mellom miljøendringer og genetiske endringer.

Flere ulike metoder for å rekne ut den virkelige genetiske endringa i en populasjon har vært foreslått. De fleste av disse metodene er varianter av ett og samme prinsipp, og bygger på at avlsdyr som har gitt avkom til ulik tid utgjør en form for "genetisk kontroll" over det tidsrommet de har vært nytta. For å unngå feil på grunn av en mulig morsvirkning som har sammenheng med mora sin alder når avkommet blir født, er det tryggest å holde seg til avkomsgrupper med samme far. Endringa i gjennomsnittlig fenotype for avkom etter ett og samme hanndyr, men født til ulik tid (T_1 og T_2) kan skrives som:

$$\bar{S}_2 - \bar{S}_1 = \frac{1}{2}\Delta G + \Delta E$$

der ΔG og ΔE angir hhv. genetisk og miljøbestemt endring fra T_1 til T_2 . Leddet $\frac{1}{2}\Delta G$ kommer fram fordi avkom etter samme far bare bare viser halvparten av den genetiske framgangen i populasjonen som helhet. Endringa i gjennomsnittlig fenotype for hele populasjonen i samme tidsrom er:

$$\bar{P}_2 - \bar{P}_1 = \Delta G + \Delta E$$

Ved subtraksjon får en:

$$(\bar{P}_2 - \bar{P}_1) - (\bar{S}_2 - \bar{S}_1) = \frac{1}{2}\Delta G$$

$$\Delta G = 2[(\bar{P}_2 - \bar{P}_1) - (\bar{S}_2 - \bar{S}_1)] = 2[(\bar{P}_2 - \bar{S}_2) - (\bar{P}_1 - \bar{S}_1)]$$

Den genetiske framgangen pr. tidsenhet blir da:

$$\Delta G/T = 2[(\bar{P}_2 - \bar{S}_2) - (\bar{P}_1 - \bar{S}_1)] / (T_2 - T_1) = 2b_{(\bar{P}-\bar{S})/T}$$

der $b_{(\bar{P}-\bar{S})/T}$ er regresjonskoeffisienten for regresjonen av $(\bar{P}-\bar{S})$ på T .

Metoden forutsetter at det ikke er noen forskjell i alder eller i genetisk kvalitet mellom partnere til yngre og eldre hanndyr. En annen forutsetning er at de hanndyra som har avkom på tidspunkt (2) ikke er selektert, men utgjør et tilfeldig utvalg av de som er

representert med avkom på tidspunkt (1). Metoden krever store data-mengder for å gi pålitelig resultat.

Utrekning av genetisk framgang i husdyrpopulasjoner etter denne og andre liknende metoder har gitt resultater som samsvarer bra med det en skulle vente ut fra seleksjonsdifferansene. For f.eks. mjølkeavdrått er det i flere populasjoner funnet en framgang på snautt 1 prosent pr. år. Dette er likevel vesentlig mindre enn det som skulle være oppnåelig med et optimalt avlsprogram.

Øvingsoppgaver.

1. I et seleksjonsforsøk blir to grupper fra samme utgangsmateriale selektert i motsatt retning. Fem prosent av handyra og 20 prosent av hodyra i hver generasjon blir nytta til avl (masseseleksjon). Arvegraden for den egenskapen det selekteres for er utrekna til 0,16 og det fenotypiske standardavviket til 10 enheter. I løpet av 10 generasjoner er det blitt en forskjell på 12 enheter mellom gruppene. Hvordan samsvarer dette med det en skulle vente ut fra teorien?
2. Gjennomsnittsavdrått for en populasjon av mjølkefe auka fra 5100 kg i 1970 til 5700 kg i 1975. For tre tilfeldig valgte okser som hadde et stort antall døtre (med 1. laktasjonsår) både i 1970 og 1975 var gjennomsnittet for hvert av de to åra:

Okse	Gj.snitt for døtre	
	1970	1975
A	5310	5830
B	4900	5310
C	5400	5850

Bruk dette til å rekne ut den gjennomsnittlige genetiske framgang pr. år fra 1970 til 1975.

XIII. INNAVLSDEPRESJON OG HETEROSIS.

A. Innavlsdepresjon.

Innavl er tidligere definert som sammenparing av individer som er i slekt med hverandre. I det praktiske avlsarbeidet blir nemninga i regelen nytta bare ved paring av nære slektninger.

Den genetiske virkningen av innavl er en auke av homozygotien og en reduksjon av heterozygotien i populasjonen. Innavlskoeffisienten, F , angir den relative reduksjon i frekvensen av heterozygoter.

Det er ei gammel røynsle at innavl oftest fører med seg en auke i forekomsten av individer med mere eller mindre alvorlige arvefeil. Dette er lettest å oppdage for kvalitative egenskaper, og er en naturlig følge av auka homozygoti. De fleste arvefeil skyldes resessive gener (de som skyldes dominante gener vil utrydde seg sjølv) og innavl resulterer i en auke i frekvensen av individer som fører de uheldige resessive genene i homozygot form, slik at de gir seg til kjenne. Den relative auken i forekomsten av resessive homozygoter er:

$$\Delta Q/Q = pqF/q^2 = pF/q$$

der q er frekvensen av det resessive genet, og $p = 1 - q$.

De fleste resessive gener som i homozygot tilstand er letale eller fører til alvorlige arvefeil forekommer med låg frekvens. Jamvel en beskjedne innavlsgrad vil derfor føre til en sterk relativ auke i frekvensen av resessive homozygoter.

Men innavl har vist seg å gi en mere generell tilbakegang også for mange kvantitative egenskaper. For å forklare dette har en antatt at det også i disse høve foreligger en viss grad av dominans for "positive" gener (retningsbestemt dominans). For et locus med to alleler, A_1 og A_2 , som har frekvensene p og q , er endringa i genotypefrekvensen på grunn av innavl:

$$A_1 A_1 : +pqF$$

$$A_1 A_2 : -2pqF$$

$$A_2 A_2 : +pqF$$

der F er innavlskoeffisienten.

Dersom en gir de tre genotypene verdiene a , d og $-a$, blir endringa i gjennomsnittlig verdi:

$$\begin{aligned}\Delta G &= pqF \cdot a + (-2pqF) \cdot d + pqF(-a) \\ &= -2pqFd\end{aligned}$$

Endringa er negativ for alle positive verdier av d , dvs. når heterozygotene ligger nærmere den beste av de to homozygote typene. Den er proporsjonal med innavlskoeffisienten (F) og dominansgraden (d), og er størst ved midlere genfrekvenser ($p \approx q \approx 0,5$).

For en egenskap som er bestemt av gener på flere loci blir totalvirkningen lik summen av virkningene på de enkelte loci:

$$\Delta G = \Sigma - 2pqFd$$

Dette at innavl fører til auka frekvens av individer med arvefeil og en generell tilbakegang i mange kvantitative egenskaper blir kalt innavlsdepresjon. Det har vist seg at innavlsdepresjonen gjør seg sterkest gjeldende for egenskaper som har betydning for artens evne til å overleve (fruktbarhet og levedyktighet), men den er merkbar også for mange av de egentlige produksjonsegenskapene. Granskinger tyder på at en auke i innavlsgraden på én prosent-enhet fører til en tilbakegang på 0,5-1 prosent i de fleste av de egenskapene en legger vekt på i husdyrbruket.

B. Heterosis.

Heterosis eller krystingsfrodighet (eng. "hybrid vigour") er et fenomen som viser seg hos avkom etter paring mellom individer fra to ulike populasjoner. Avkom etter ei slik kryssing er ofte overlegne over foreldretypene, først og fremst i egenskaper som er knytta til forplanting og levedyktighet, men også i andre egenskaper.

Heterosis er et motstykke til innavlsdepresjon, og skyldes at kryssing fører til en auke i frekvensen av heterozygoter. La oss ta for oss et locus med to alleler, A_1 og A_2 . I to populasjoner er frekvensen av A_1 hhv. p og p' og av A_2 q og q' ($p + q = p' + q' = 1,0$). Ved å sette $p - p' = y$ får en:

$$p' = p - y \text{ og } q' = q + y$$

Frekvensen av genotypene A_1A_1 , A_1A_2 og A_2A_2 i hver av de to populasjonene og i gjennomsnitt for dem blir:

	I	II	(I+II)/2
A_1A_1	p^2	$(p-y)^2$	$\frac{1}{2}p^2 + \frac{1}{2}(p-y)^2 = p^2 - p \cdot y + \frac{1}{2}y^2$
A_1A_2	$2pq$	$2(p-y)(q+y)$	$p \cdot q + (p-y)(q+y) = 2pq + (p-q)y - y^2$
A_2A_2	q^2	$(q+y)^2$	$\frac{1}{2}q^2 + \frac{1}{2}(q+y)^2 = q^2 + q \cdot y + \frac{1}{2}y^2$

Ved kryssing av de to populasjonene får en disse genotypefrekvensene blant avkommet:

$$\begin{aligned}
 A_1A_1: p \cdot p' &= p(p-y) &= p^2 - p \cdot y \\
 A_1A_2: p \cdot q' + p' \cdot q &= p(q+y) + (p-y)q &= 2pq + (p-y)y \\
 A_2A_2: q \cdot q' &= q(q+y) &= q^2 + q \cdot y
 \end{aligned}$$

Forskjellen i genotypefrekvenser mellom avkomspopulasjonen og gjennomsnittet av de to foreldrepopulasjonene er:

$$\begin{aligned}
 A_1A_1 & - \frac{1}{2}y^2 \\
 A_1A_2 & + y^2 \\
 A_2A_2 & - \frac{1}{2}y^2
 \end{aligned}$$

Dersom de tre genotypene har verdien a , d og $-a$, blir den gjennomsnittlige forskjellen i verdi:

$$\Delta G = -\frac{1}{2}y^2 \cdot a + y^2 \cdot d - \frac{1}{2}y^2(-a) = d \cdot y^2,$$

eller, summert over alle loci: $\Delta G = \Sigma d \cdot y^2$

Heterosisvirkningen er proporsjonal med dominansgraden (d) og med kvadratet på forskjellen i genfrekvens mellom de to foreldrepopulasjonene (y).

Enkelte forskere hevder at heterosis av den storleiksorden som den en har funnet i f.eks. mais, ikke lar seg forklare ved hjelp av hypotesen om fullstendig eller partiell dominans for positive gener. De mener derfor at det til dels må foreligge overdominans, dvs. at heterozygotene er bedre enn noen av de homozygote typene. Formelene ovafor gjelder også i denne situasjonen ($d > a$).

Ettersom heterosis skyldes samme genetiske mekanisme som innavlsdepresjon, må en vente at de egenskapene som viser størst depresjon ved innavl, også er de som viser mest heterosis. Mange forsøk og granskinger har stadfesta at dette stort sett er riktig. Den største heterosisvirkningen har en for forplantingsegenskaper og levedyktighet. Dette er egenskaper som har låg arvegrad, dvs. liten additiv genetisk variasjon. En kan tenke seg at naturlig seleksjon gjennom tidene har ført til en reduksjon i den additive genetiske variasjonen, og at storparten av den genetiske variasjonen som er igjen, skyldes dominans (og epistasi). Det er derfor rimelig at nettopp disse egenskapene viser en sterk grad av heterosis. Men en har tydelig heterosis også for mange andre egenskaper (veksthastighet, mjølkeavdrått osv.).

C. Avlsmetoder som utnytter heterosis.

I en populasjon med tilfeldig paring kan proporsjonen av heterozygoter for et locus med to alleler maksimalt være 0,5. Denne proporsjonen har en når de to allelene har samme frekvens ($p=q=0,5$). Seleksjon vil i de fleste høve ha en tendens til å forskyve genfrekvensen i retning av ekstreme verdier. For å auke heterosisvirkningen må en nytte ei eller anna form for kryssing.

1. Rasekryssing.

Kryssing av to ulike raser er en gammel og velprøvd avlsmetode. Formålet kan ofte være et helt anna enn å utnytte heterosis:

a. Gjennomført kryssing blir nytta som et hjelpemiddel til å føre inn en ny rase i et distrikt. En kjøper inn hanndyr av den nye rasen og bruker disse til paring av egne hodyr. Arvestoffet fra den "gamle" rasen blir da halvert for hver generasjon, og etter noen generasjoner vil en i praksis ha skifta rase. En vil få fullt utslag for heterosis i første generasjon, men denne vil forsvinne samtidig med genene fra den "gamle" rasen.

b. Kombinasjonskryssing kaller en det når en krysser for å kombinere gode egenskaper fra to (eller flere) raser i en ny rase. Avkommet etter den første kryssinga (F_1) blir para innbyrdes, og danner grunnlaget for den nye rasen, som så blir utvikla vidare

gjennom seleksjon. En rase som er blitt til på denne måten blir ofte kalt en syntetisk rase.

Også i dette tilfellet får en full heterosisvirkning i første generasjon. I denne generasjonen blir genfrekvensene $p-\frac{1}{2}y$ og $q+\frac{1}{2}y$ (se foran). Frekvensen av heterozygoter i neste generasjon (F_1) blir derfor:

$$H = 2(p-\frac{1}{2}y)(q+\frac{1}{2}y) = 2pq + (p-q)y - \frac{1}{2}y^2$$

Jamført med gjennomsnittet av de to utgangsrasene er dette en auke i frekvensen av heterozygoter på $\frac{1}{2}y^2$, dvs. halvparten så mye som auken i første generasjon. Halvparten av heterosisvirkningen går altså tapt fra F_1 til F_2 . Men fra F_2 og utover vil den holde seg uforandret (bortsett fra eventuell virkning av seleksjon).

c. Bruksdyrkryssing (eller brukskryssing) er kryssing med sikte på å produsere avkom som bare skal tjene som bruksdyr. Hovedformålet med kryssing i dette høvet er ofte å utnytte de fordelene en annen rase har for en bestemt produksjonsretning, som f.eks. når kyr av mjølkerase blir para med okser av kjøtttrase for å skaffe dyr som egner seg for kjøttproduksjon. Men også utnytting av heterosis er viktig i denne sammenhengen.

I en rase som bare skal nyttes som far-rase trenger en ikke sette så store krav til de spesielle morsegenskapene (fruktbarhet, mjølkeevne osv.), som en ellers ville måtte gjøre. Ved bruksdyrkryssing vil det derfor være aktuelt å utvikle spesialiserte mor- og far-raser (eller liner). Ved seleksjon i far-rasen kan en legge ensidig vekt på de egentlige produksjonsegenskapene, mens mor-rasen må foredles med omsyn på både produksjonsegenskaper og morsegenskaper.

Ofte er heterosis vel så viktig for morsegenskapene som for produksjonsegenskapene. Det kan derfor være fornuftig å nytte rasekryssing for produksjon av mordyra, og pare disse med handyr enten fra en av foreldrerasesene eller fra en tredje rase. I siste høve vil en få full utnytting av heterosis både for morsegenskaper og for produksjonsegenskaper. Et avlsopplegg der både mødre og avkom er resultatet av kryssing blir iblant kalt bruksdyrkryssing over to generasjoner. Det blir bl.a. nytta i svineavl.

d. Rotasjonskryssing. I dyrearter der hodyra har liten forplantingsevne (f.eks. storfe og småfe) vil de fleste hodyra måtte tjene både som avlsdyr og som bruksdyr. Det kan derfor være aktuelt å legge opp til et system der en skifter mellom hanndyr av ulike raser, men nytter kryssings-hodyr til avl. Dette blir kalt alternerende kryssing (to raser) eller rotasjonskryssing (tre eller flere raser). Ved alternerende kryssing vil avkommet (når systemet er kommet i jamvekt) ha $2/3$ av genene sine fra den ene og $1/3$ fra den andre rasen. Heterosisvirkningen blir $2/3$ av den en får i første kryssingsgenerasjon etter de samme rasene. For rotasjonskryssing med tre raser blir proporsjonen av gener fra de tre rasene hhv. $4/7$, $2/7$ og $1/7$.

2. Kryssing av innavlsliner.

Ved innavl vil en populasjon bli delt opp i flere genetisk ulike liner. Innafor hver line vil en få en stor grad av homozygoti og liten genetisk variasjon. Når en krysser disse linene, alle med alle (diallel kryssing) får en, for populasjonen under ett, de samme genotyperefrekvensene som i utgangspopulasjonen, det en har oppnådd er å oppheve innavlsdepresjonen. Men det kan være stor forskjell mellom de ulike kryssingskombinasjonene. I denne sammenhengen skiller en mellom generell og spesiell kombinasjonsevne (avlsverdi) (eng. "general and specific combining ability"). Med "generell kombinasjonsevne" forstår en den evne som ei line har til å gi godt avkom i kryssing med andre liner generelt. Denne evnen (generell avlsverdi) kommer til uttrykk i den gjennomsnittlige fenotype for alt avkom som har individer fra denne lina som ett av foreldra. Den generelle kombinasjonsevne er i første rekke knytta til den additive genetiske verdi av lina. Spesiell kombinasjonsevne er den særvirkning som en får ved kryssing av to bestemte liner, ut over det som den generelle kombinasjonsevne for de to linene skulle tilsi. Verdien av avkommet etter ei kryssing mellom to liner er bestemt av den generelle kombinasjonsevne for hver av de to linene og den spesielle kombinasjonsevne som de har i høve til hverandre.

Ved å selektere de linene som har gitt best resultat ved kryssing og gjenta denne bestemte kryssinga kan en utnytte både

den generelle og den spesielle kombinasjonsevne. Denne metoden er mye nytta i maisforedlinga, og har gitt utslag i avling på opptil 25 prosent i høve til utgangsmaterialet. For fjørfe har en i stor utstrekning brukt samme framgangsmåte, men avslutta med ei "dobbel kryssing" (dvs. kryssing av avkom fra to ulike enkle kryssinger). Formålet med dette er bl.a. å unngå innavlsdepresjon, og dermed nedsatt fruktbarhet, i avlsdyrgenerasjonen (ved produksjon av rugeegg).

3. Kryssing med gjensidig seleksjon.

Når en har funnet fram til to populasjoner (raser eller liner) som gir godt resultat i kryssing, kan en søke å utvikle denne evnen videre gjennom seleksjon. Utvalg av avlsdyr til videreføring av hver populasjon må da foregå på grunnlag av den avlsverdien som individene viser ved kryssing med individer fra den andre populasjonen. Når dette blir gjentatt over flere generasjoner, vil de to populasjonene etterhvert tilpasse seg hverandre. Seleksjonen vil være både for generell og for spesiell kombinasjonsevne. I den grad den spesielle kombinasjonsevna skyldes heterozygoti vil seleksjonen føre genfrekvensene i de to populasjonene i motsatt retning.

I engelsk-språklig litteratur blir denne metoden kalt "reciprocal recurrent selection" eller bare "reciprocal selection". I mangel av et bedre uttrykk er det her oversatt til "gjensidig seleksjon". Metoden er utvikla på rent teoretisk grunnlag, og har vært tatt med til utprøving i flere seleksjonsforsøk, med noe skiftende resultat. I praktisk husdyravl er den hittil blitt lite nytta.

Øvingsoppgaver.

1. I en populasjon er genet A_1 fullstendig dominant over det allelle genet A_2 . Verdien av de tre genotypene er:

A_1A_1 og A_1A_2 : 10 enheter

A_2A_2 : 2 "

Genet A_1 har frekvensen 0,6.

Rekn ut venta innavlsdepresjon i prosent ved en innavlskoeffisient på 0,1.

2. Individua fra populasjonen i oppgave 1 blir kryssa med individ fra en annen populasjon, der frekvensen av A_1 er 0,3.
- Rekn ut venta heterosis i F_1 -generasjonen.
 - Jamfør kryssingsavkommet med den beste av de to foreldrepopulasjonene.
 - Hva er venta gjennomsnittlig verdi av avkom etter innbyrdes paring av F_1 -individua?

