Serik Sagitov, Chalmers Tekniska Högskola, November 23, 2005

5. Quantitative genetics

Types of quantitative traits

metric traits (cont. data): height, weight

meristic traits (discrete data): litter size, #bristles

threshold traits (categorical data): affected or not

5.1 broad sense heritabitily

5.2 narrow sense heritabitily

5.3 truncation selection

5.4 resemblance between relatives

5.1 Broad sense heritabitily

As a rule a QT is influenced by many genes (polygenes)

each gene exerting relatively small effect

considerable environmental variation

Ex 1: bristle number

Fig 9.1, p. 399: *D.melanogaster*

X = number of bristles on 5th abdominal sternite

histogram with $\mathrm{N}(\bar{X}, s^2)$ curve, $\bar{X} = 18.7$, $s = 2.1$

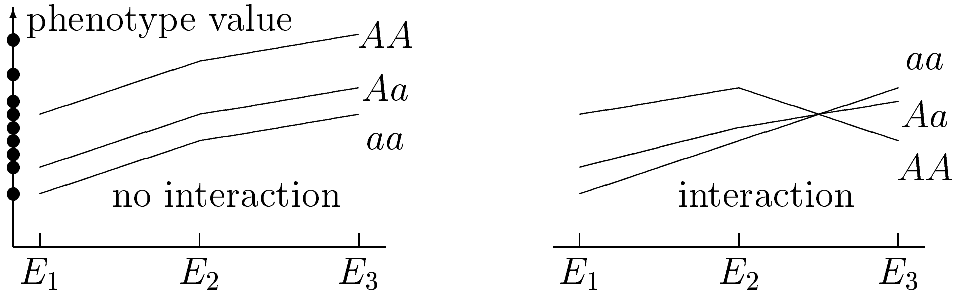
phenotypic variance $\sigma_p^2 = \mathrm{Var}(X) \approx (2.1)^2$

Phenotypic value of an individual: $P = M + E$
genotype value $M = \mu + G$ (mean μ)
environmental deviation E (mean 0)

Assumption: independent deviations G and E
$P = \mu + G + E$, variance decomposition $\sigma_p^2 = \sigma_g^2 + \sigma_e^2$

$$\boxed{\text{Broad sense heritabitily } H^2 = \sigma_g^2/\sigma_p^2}$$

Absence of genotype-environment interaction:



Examples of environmental effects:
nutritional and climatic factors, maternal effects
cultural environment

Ex 2: estimation of variance components
1) estimate σ_e^2 observing inbred populations
2) find $\sigma_g^2 + \sigma_e^2$ from random-bred populations

Thorax length in $D.melanogaster$ (in 0.01 mm)
inbred populations $\sigma_e^2 = 0.186$
random-bred populations $\sigma_p^2 = 0.366$ $(= \sigma_g^2 + \sigma_e^2)$
$\sigma_g^2 = \sigma_p^2 - \sigma_e^2 = 0.180$, $H^2 = 49.2\%$

Ex 3: LDL-cholesterol level

Human Apo-E locus binds LDL receptor protein

three common alleles e_2 (0.1), e_3 (0.7), e_4 (0.2)

LDL-cholesterol level in $\frac{\text{mg LDL}-\text{chol}}{\text{dL blood plasma}}$

genot	e_2e_2	e_2e_3	e_2e_4	e_3e_3	e_3e_4	e_4e_4	
freq	0.01	0.14	0.04	0.49	0.28	0.04	sum $=1$
M	76	90	115	100	110	106	$\mu=102$
G	-26	-12	13	-2	8	4	$\sigma_g^2=54.2$
A	-16.8	-9.4	-0.7	-2	6.7	15.4	$\sigma_a^2=39.2$
D	-9.2	-2.6	13.7	0	1.3	-11.4	$\sigma_d^2=15.0$

Broad sense heritability

given $\sigma_p^2 = 554.2$ compute $\sigma_e^2 = 500.0$, $H^2 = 0.098$

High LDL-chol increases risks of coronary heart desease

e_2e_4 genotype is at greatest risk

e_4e_4 genotype's children are at greatest risk

Breeding value of a genotype

$A =$ twice the average G across possible offspring

factor 2: only half of genes come from one parent

Dominance effect $D = G - A$, whithin locus interaction

$$\boxed{P = \mu + A + D + E}$$

5.2 Narrow sense heritabitily

One locus model

two alleles A_1 and A_2 with frequencies p and q

assuming random mating and HWE

genotype	A_1A_1	A_1A_2	A_2A_2
frequency	p^2	$2pq$	q^2
M	$\mu^* + a$	$\mu^* + d$	$\mu^* - a$
G	$2q\alpha - 2q^2d$	$(q-p)\alpha + 2pqd$	$-2p\alpha - 2p^2d$
A	$2q\alpha$	$(q-p)\alpha$	$-2p\alpha$
D	$-2q^2d$	$2pqd$	$-2p^2d$

Genotypic values M given around $\mu^* = \frac{M_{11}+M_{22}}{2}$
do not depend on the allele frequency p
Genotypic deviation $G = M - \mu$ depends on p
since population mean $\mu = \mu^* + (p-q)a + 2pqd$
Average effect of gene substitution
when a randomly chosen A_1 is replaced by A_2
$\alpha = p(a-d) + q(a+d)$
Breeding value of A_1A_1 genotype
$A_{11} = 2(pG_{11} + qG_{12}) = 2q\alpha$
random mate contributes allele A_1 with probability p
A and D values are uncorrelated
additive variance $\sigma_a^2 = 2pq\alpha^2$, additive gene effects
whithin locus interaction component $\sigma_d^2 = (2pqd)^2$
$\sigma_g^2 = \sigma_a^2 + \sigma_d^2$, so that $\sigma_p^2 = \sigma_a^2 + \sigma_d^2 + \sigma_e^2$

$$\boxed{\text{Narrow sense heritability } h^2 = \sigma_a^2/\sigma_p^2}$$

Narrow heritability h^2 is more important than broad H^2
since parents pass their genes not genotypes

Ex 3: LDL-cholesterol level

Narrow sense heritability

$554.2 = 500.0 + 39.2 + 15.0$, $h^2 = 0.071$

Ex 4: inheritance of rare diseases

autosomal disease allele A_2 of low frequency $q \approx 0$

assuming $\sigma_e^2 = 0$

Rare recessive disease: $d = a$

$\alpha = 2qa$, $\sigma_a^2 = 8pq^3a^2$, $\sigma_d^2 = 4p^2q^2a^2$

Low inheritance $h^2 = \frac{\sigma_a^2}{\sigma_a^2+\sigma_d^2} = \frac{q}{1+q} \approx 0$

affected offspring come usually from unaffected parents

Rare dominant disease: $d = -a$

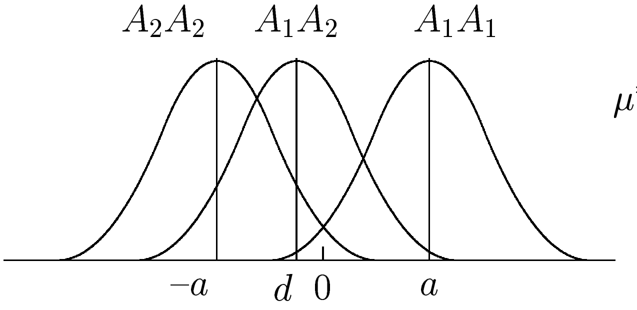
$\alpha = 2pa$, $\sigma_a^2 = 8p^3qa^2$, $\sigma_d^2 = 4p^2q^2a^2$

High inheritance $h^2 = 2\frac{p}{1+p} \approx 1$

affected offspring have one affected parent

Dependence on allele frequency

Phenotypic value distribution is a $(p^2, 2pq, q^2)$ mixture of three distributions for three genotypes



$\mu^* = 0$, $a > 0$

$d < 0$

Draw two pdf curves for phenotypic value

1) at $p = \frac{1}{2}$ with negative $\mu = \frac{d}{2}$

2) at $p = \frac{3}{4}$ with positive $\mu = \frac{a}{2} + \frac{3d}{8}$

Ex 5: coat coloration

$x =$ proportion of black color on the guinea pig coat

Normalizing transformation

if $0 \leq x \leq 1$, then

$P = \arcsin \sqrt{x}$ often has near normal distribution

Fixed environment: $P = M$ and $\sigma_p^2 = \sigma_g^2$

genotype	x	P	$P - \mu^*$
A_1A_1	0.87	68.87	$a = 68.87 - 61.60 = 7.27$
A_1A_2	0.76	60.67	$d = 60.67 - 61.60 = -0.93$
A_2A_2	0.66	54.33	$-a = 54.33 - 61.60 = -7.27$

Variances as functions of p

$\sigma_a^2 = -6.92p^4 - 40.25p^3 - 33.22p^2 + 80.39p$

$\sigma_d^2 = 3.46p^4 - 6.92p^3 + 3.46p^2$

$\sigma_g^2 = -3.46p^4 - 47.17p^3 - 29.76p^2 + 80.39p$

Fig 9.15A-C, p. 433: variance profiles

Ex 6: chromosome inversions

D. pseudoobscura: inversions in the 3rd chromosome

$A_1 =$ standard, $A_2 =$ arrowhead chromosomes

overdominance $w_{11} = 0.47$, $w_{12} = 1.00$, $w_{22} = 0.62$

Fig 9.15D, p. 433: variance profiles

p	0	0.2	0.418	0.8	1
$\alpha = 0.38 - 0.91p$	0.38	0.198	0	-0.348	-0.53
$\sigma_a^2 = 2pq\alpha^2$	0	0.0125	0	0.0388	0
$\sigma_d^2 = 0.83(pq)^2$	0	0.0212	0.049	0.0212	0
$\sigma_g^2 = \sigma_a^2 + \sigma_d^2$	0	0.0337	0.049	0.06	0
$h^2 = \sigma_a^2/\sigma_g^2$	–	37.1%	0	64.7%	–

equilibrium frequency $\hat{p} = \frac{1-0.62}{1-0.47+1-0.62} = 0.418$

> Fundamental theorem of natural selection:
> the increase in average fitness at any time is the additive genetic variance in fitness at that time

5.3 Truncation selection

Artificial selection aiming at a certain phenotypic value
use a truncation point T for parent selection
so that the offspring of selected parents have
phenotypic distribution with a desired bias

To estimate heritability compare phenotypic mean values
μ = parent mean before selection
μ_s = mean for selected parents
μ' = mean for the offspring of selected parents

$$\boxed{R/S = \text{realized heritability}}$$

$S = \mu_s - \mu$ selection differential
$R = \mu' - \mu$ response to selection

$$\boxed{\text{Prediction equation: } R = Sh^2 \text{ irrespective of } T}$$

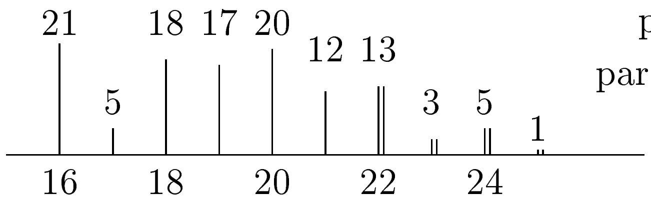
Ex 7: seed weight

Fig 9.6, p. 409: edible beans of the genus *Phaseolus*

P = weight of seed in mg, truncation point $T = 650$

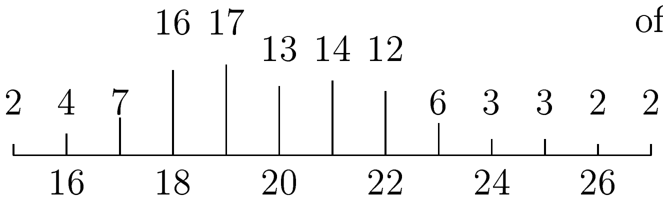
$\mu = 403.5$, $\mu_s = 691.7$, $\mu' = 609.1$, $\frac{R}{S} = \frac{205.6}{288.2} = 71.3\%$

Ex 8: drosophila bristles



parents total 115

parents selected 22



offspring total 101

Truncation selection with T=22

verify that $\mu = 19.304$, $\mu_s = 22.727$, $\mu' = 20.149$

realized heritability $h^2 = \frac{\mu' - \mu}{\mu_s - \mu} = \frac{0.845}{3.423} = 0.247$

Repeated truncation selection

Selection program over n generations with

new truncation points changing in certain direction

$\mu_0 \xrightarrow{T_0} \mu_{s0} \xrightarrow{h^2} \mu_1 \xrightarrow{T_1} \mu_{s1} \xrightarrow{h^2} \ldots \mu_{n-1} \xrightarrow{T_{n-1}} \mu_{s(n-1)} \xrightarrow{h^2} \mu_n$

$S_0 = \mu_{s0} - \mu_0$, $R_0 = \mu_1 - \mu_0$, $R_0 = S_0 h^2$

$S_1 = \mu_{s1} - \mu_1$, $R_1 = \mu_2 - \mu_1$, $R_1 = S_1 h^2, \ldots$

Total response to selection assuming constant h^2

$\mu_n - \mu_0 = R_0 + R_1 + \ldots + R_{n-1} = (S_0 + \ldots + S_{n-1})h^2$

cumulative selection differential $C_n = S_0 + \ldots + S_{n-1}$

Ex 9: body weight in mice

Fig 9.19, p. 445

body weight in mice plotted against C_t

linearity supports the assumption of constant h^2

which is usally true for at least ten first generations

Ex 10: oil content in corn

Fig 9.4, p. 407: selection for high oil content in corn seeds

over 76 generations, $\mu_0 = 4.8\%$, $\mu_{76} = 18.8\%$

Given that C_t increased by 1.1% per generation

estimate $h^2 = \frac{18.8-4.8}{1.1\times76} = 0.168$

5.4 Resemblance between relatives

Another characterisation of h^2 via comparison of

$P_o =$ male offspring's phenotypic values

$P_f =$ father's phenotypic values

Regression line

$P_o = \mu_o + b(P_f - \mu_f)$ with the slope $b = \frac{\mathrm{Cov}(P_o,P_f)}{\mathrm{Var}(P_f)}$

Diallelic model neglecting the environmental component

$\mathrm{Cov}(P_o, P_f) = \mathrm{E}(P_o \cdot P_f) - \mu^2 = pq\alpha^2 = \frac{1}{2}\sigma_a^2$

$b = \frac{\sigma_a^2}{2\sigma_p^2} = \frac{h^2}{2}$

joint distribution	$O=a$	$O=d$	$O=-a$	total
$P=a,\quad A_1A_1$	p^3	p^2q	0	p^2
$P=d,\quad A_1A_2$	p^2q	pq	pq^2	$2pq$
$P=-a,\ A_2A_2$	0	pq^2	q^3	q^2
total	p^2	$2pq$	q^2	1

Offspring and midparent value

$P_h = \frac{1}{2}(P_m + P_f)$

$\mathrm{Cov}(P_o, P_h) = \frac{1}{2}\sigma_a^2$, $\mathrm{Var}(P_h) = \frac{1}{2}\sigma_p^2$, $b = h^2$

Ex 11: pupa weight

Fig 9.2, p.401: pupa weight of flour beetle

(son, father) regression line slope $b = 0.11$, $\hat{h}^2 = 0.22$

Ex 12: shell breadth in mm in 119 sibships of snail

Observed frequencies in 119 sibships of snail

shell breadth in mm (midparent value, offspring mean)

22 (16.25, 17.73)	11 (23.75, 22.84)
31 (18.75, 19.15)	4 (26.25, 23.75)
48 (21.25, 20.73)	3 (28.75, 25.42)

Sample means and sample variances

$\bar{P}_h = 20.26$, $\bar{P}_o = 20.18$, $s_h^2 = 8.18$, $s_o^2 = 3.31$

Sample covariance

$$= \frac{1}{118}[22 \cdot (16.25 - 20.26) \cdot (17.73 - 20.18) + \ldots + 3 \cdot (28.75 - 20.26) \cdot (25.42 - 20.18)] = 5.18$$

Estimated narrow-sense heritability $\hat{h}^2 = \frac{5.18}{8.18} = 0.634$

Observed heritabilities

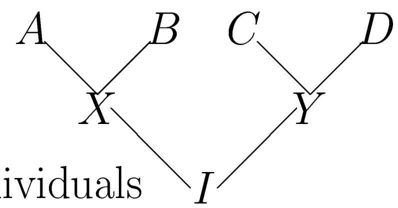
Fig 9.17-18, p. 438-9: animal and plant h^2, human H^2
low heritabilities of fitness related traits

General covariance and slope

Table 9.7 p. 436: covariances between close relatives

$$\boxed{\mathrm{Cov}(X,Y) = r\sigma_a^2 + u\sigma_d^2} \quad \boxed{b = (r-u)h^2 + uH^2}$$

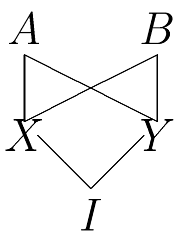
$r = 2F_{XY}$
$u = F_{AC}F_{BD} + F_{AD}F_{BC}$



Coefficient of coancestry for two individuals

$$\boxed{F_{XY} = F_I = \mathrm{P}(\text{IBD genes of hypothetical offspring } I)}$$

Ex 13: full siblings



Two genes in I are IBD if they both come
1. from the same grandparent
2. from the same chromosome of that grandparent

$F_{XY} = 0.5 \cdot 0.5 = 0.25$
$r = 2 \cdot 0.25 = 0.5$
$u = F_{AA}F_{BB} + F_{AB}F_{BA} = 0.5 \cdot 0.5 + 0 \cdot 0 = 0.25$

Covariance and slope
$\mathrm{Cov}(X,Y) = \frac{\sigma_a^2}{2} + \frac{\sigma_d^2}{4}$, $b = \frac{h^2}{4} + \frac{H^2}{4}$