Estimation of Mutation in Quantitative Traits:

- Mutation: the source of population genetic variation on which natural or aritifical selection acts (to produce genetic changes leading to adaptive evolution)
- Number of rounds of replicaiton from zygote to the formation of gametes varies between species and is why rates of mutations vary between species
- Men and women: the number of rounds of replication in ova stay at 22 while it continues to grow for spermatozoa in men
- However, chromosomal mutations increase in women because of cytoplasmic deterioration in oocytes over time
- Given the difficulty of identifying genotypes, it is only possible to obtain average estimates of the effect of mutations and their degree of dominance
- Mutational variance V_M is the increase in additive genetic variance per generation due to mutation
- And thus mutational heritability is the ratio of V_M/V_E
- Probability of fixation of a mutation:
  - In finite populations, genetic drift may cause beneficial alleles to be lost and deleterious alleles to be fixed
  - If the mutational selection coefficient is small ($N_e s/N < 1$)and $N_e s$ is high, then the probability of fixation can be approximated by the selective advantage s, so that most beneficial mutations of small effect are lost in a population of finite size
  - Neutral mutations have a P_f equal to their initial frequency $1/2N$
  - Deleterious mutations are destined to be eliminated by selection
  - Mutations where absolute value of N_e is substantially smaller than 1, the probability of fixation is the same for that of a neutral mutation. When mutation is advantageous or disadvantageous, the action of drift is more intense than selection and so the fate of the mutation is dependent on chance. This is referred to as quasi-neutral mutations.
  - Mutations can be beneficial, neutral or deleterious based on effective population size
- Estimating the rate of mutation and mutational effects
  - Balanced chromosome technique: cross an individual with another with a balanced chromosome (carrying inversions that inhibit recombinations along with visible markers for identification)
  - Cross wt male with If female to obtain copies of chromosome that are used to create multiple independent lines
  - Mutations accumulate in each line and disappear when deleterious heterozygous effects are sufficiently large
  - This approach has been applied to C. elegans, where lines are started by either self-fertilization or brother sister mating. Mutations accumulate in the genome, and because $N_e s << 1$, mutations behave as neutral and have a highly probability of fixation.
  - Mutations can be studied if they affect traits that are fitness components, either directly measured (productivity) or it may be in competition with the starting population (founder individuals in C. elegans can be frozen then thawed and compared with individuals from a line)
  - Accumulating mutations is expected to reduce the mean value of the trait will increasing the variance in that trait, and $\delta V/\delta M = \bar{s}(1+C^2)$, where U is the haploid genome mutation rate, and C is the coefficient of variation of mutational effects, and s is the average of selective effects of mutations weighted by their mutation rate (Bateman-Mukai method)

  - Bateman-Mukai enables the estimation of mutational variance (increase in additive genetic variance per generation due to new mutations) and mutational heritability
  - Past experiments typically show that most mutations are of small effects, and this is studied using the gamma distribution
  - U has been estimated for multiceullular eukaryotes (U = 0.08) and viruses (0.93).
  - Mean mutational coefficient of variation for fitness traits is $CV_M = 0.026$
  - The Bateman-Mukai method produces estimates with bias of unknown magnitude, because we do not know the distribution of mutational effects
  - rate of deleterious mutation can also be estimated from molecular data as $U = U_n\phi$, where U_n is the rate of nucleotide mutation per generation for the genome and phi is the fraction of mutations whose deleterious effect is large enough for selection to determine its elimination, calculated by $\phi = 1 - \pi_g/\pi_s$ where pi_s is the rate substitution rate of mutations that are assumed to be exclusively neutral, and pi_g is the total divergence observed between to twin species
  - Detecting mutation rates using molecular methods only discovers mutations of large effect, and those of small effect are not found because of insufficient power
- The dominance coefficient
  - Mutation accumulation experiments enable the estimates of average coefficient of dominance
  - First estimate is obtained as the ratio between the value of the trait from a cross between two lines along with the sum of the values of the trait in those lines. Mutations will be found in heterozygosis in the cross
  - The decline in mean of trait divded by the sum of declines occuring in lines makes up the ratio estimates, which is a measure of the average of dominance coefficients weighted by selection coefficients
  - Unbiased method for estimating average coefficient of dominance: regress the trait values at cross between lines of heterozygotes onto homozygotes (acting as the sum of trait values in parental lines)
  - Joint distribution of selection and dominance coefficients: hard to do because the effect of each mutation must be estimated individually
  - Large effect mutations in homozygosis are assumed to be recessive
  - Small-effect mutations are additive
  - Fisher hypothesis: a large amount of observed mutations are recessive. It is thought that they were initially additive and that dominance modifiers appeared at other loci that would cause the heterozygoye expression to become similar to wild homozygous
  - More common explanation: Wright's physiological theory of dominance. This is based on the idea that the enzymatic activity of heterozygotes is intermediate between homozygotes, but the flux of a metabolic pathway with several enzymes is non linear. But if the mutation has a high activity-reducing effect, then the heterozygote is more similar to the wild homozygote, so the mutation has recessive gene action. This explains why large-effect mutations tend to be recessive
- Beneficial mutations and summary of mutational parameters: there is a large variation in mutational parameters between estimates for different traits, species and estimation methods
- Somatic mutation rates may be 50x higher than those of the germline
- Mutations and environmental factors: can be additive or multiplicative. Can also be synergistic epistasis (greater than combining them) or antagonistic (lower than combining)
- Mutations can be conditionally deleterious (such as temperature dependent ones)
- Rate of mutation can also be contingent on the genetic background and whether a population is adapted to its environment. For instance, populations with low average fitness

see higher rates of beneficial mutation compared to those with normal average fitness (though these are compensatory)
- Mutational parameters for quasi-neutral traits: can be studied using mutation accumulation or, starting with a population lacking genetic variability, artifical selection
- If the trait is neutral, then variability is determined exclusively by mutation (increase) or drift (reduction)
- Under the neutral model, the heterozygosity in the mutation-drift equilibrium is a function of N_e and mutation rate per locus
- Alternatively, knowing heterozygosity and mutation rate we can estimate effective population size
- Additive genetic variance in each generation is the result of the decrease in drift and the increase by new mutation, so equilibrium will be reached
- Though traits are not generally neutral, the following expression can be used if N_e is sufficiently small: $\hat{H} = \frac{4N_e u}{4N_e u+1}$
- The genetic variance between mutation accumulation lines increases at a rate $2\delta F V_A$
- The response to artificial selection due to mutation can be expressed as $R_t = t\hat{V}i/\sqrt{V_P}$ where i is the intensity with which selection is applied and V_P is the phenotypic variance in that trait
- Genes with low effect have effect sizes that are indistinguishable from additivity, but those of large effect are usually recessive and have deleterious pleiotropic effects on fitness
- Implications of deleterious mutations in populations of large sizes: the maximum possible frequency of a deleterious allele at the mutation selection balance is $\hat{q} = u/sh$
- Small effect of alleles on heterozygosis implies a substantial reduction of equilibrium frequency
- Mutation load: decrease in mean fitness of a population compared to that with no segregating deleterious mutations
- The mutation load for a recessive allele is equal to u which implies that in deleterious mutation selection equilibrium, the load only depends on mutation rate regardless of frequency
- Joint effect of all loci with presumed deleterious alleles: $L \approx 1 - e^{-2U}$
- The mean number of mutations eliminated per individual by selection is L(z-y) where L is the load, z is the mean number of del. mutations in individuals and the y is the mean number of mutations carried by survivors
- The number of mutations eliminated by selection may be greater than expected due to synergistic epistasis
- Non-panmictic populations: equilibrium load decreases as the inbreeding coefficient increases (showing how inbreeding purges deleterious mutations)
- Hard selection: fitness directly translates into survival probability, quite unrealistic in terms of the amount of progeny needed for survival
- But mutation load refers to population fitness relative to the mutation-free genotype which is often non-existent in natural populations so the issue with hard selection is reduced/disappears (so long as fitness is considered in relative terms)
- Relaxation of selection in human population: possible accumulation of deleterious mutations
- Attempted eugenics in 1970s over the world was a failure due to the ineffective elimination of recessive alleles which are masked in heterozygotes
- Consider loads outside of mutation: load due to segregation of heterozygotes in overdominant fitness models, recombination, advantageous allelic recombination, or environemntal heterogeneity or inbreeding
- Estimating the average dominance coefficient in mutation-selection equilibrium: using balanced chromosome designs as mentioned before. Homozygous individuals can be obtained for a chromosome, and carriers of chromosomes heterozygous for all loci. Then

regress the heterozygotes on homozygotes to estimate the harmonic mean of h weighted by s, given by: $\hat{b} = \frac{\bar{s}}{\overline{s/h}}$

- Estimating mutational parameters in equilibrium from panmictic and inbred populations: if data on means and variances of fitness are available, then we can compute the mutation rate, mean selection coefficient of mutations and dominance coefficients
- Mutation and recombination: sex is one of the most important evolutionary aspects for the fixation of beneficial (or deleterious) mutations
- Sex carries a two-fold cost assuming an equal number of descendants per individual
- Fisher and Muller proposed the evolutionary advantage of recombination: the probability of fixing beneficial mutations at different loci will be independent with recombination
- Fixation of beneficial mutations is much faster in populations with sex and mutation, providing a net selective advantage
- Muller expanded on deleterious mutations. They can be assembled inthe same genome and thus simultaneously eliminated. In asexual species, mutations accumulate without genome reconstruction, so the minimum number of mutations carried by a given genome grows irreversibly (Muller's ratchet)
- The fixation probability of mutations for any rate of recombination can be obtained from an extension of the formula for predicting N_e under linkage, which is $N_e \approx \frac{-U}{s+(L/2)}$
- Hill-Robertson effect: when two alleles of different loci segregate, interference can occur so their probability of fixation is smaller than that if they did not segregate simultaenously (occurs in the absence of LD or through interaction between loci)
- In chapter 5 we saw that selecting for a loci increases the magnitude of drift on an unselected, neutral locus due to random associations, reducing N_e ascribed to the neutral locus
- Hill-Robertson is an extension of this when two loci are linked. The association between loci will persist over more generations and will be diluted by recombination, progressivly reducing N_e
- If two linked genes increased in frequency due to selection N_e decreases but the effect is reduced if recombination is present
- Hill-Robertson effect also predicts an increase in the rate of fixating deleterious mutations if no recombination is present

Problems

(1) In a population with $N = 100$ individuals and $N_e = 50$, an additive mutation occurs with favourable effect $s = 0.1$ in the homozygote. (a) What is the probability that the mutation will end up being fixed in the population? (b) What would be that probability if the mutation were recessive? (c) If the mutation were deleterious and additive, what would be the mean persistance time and the average number of copies present in the population prior to elimination?

$$P_f = \frac{1 - e^{-s(N_e/N}}{1 - e^{-2N_e s}} = \frac{1 - e^{-0.1(50/100)}}{1 - e^{-2(50)(0.1)}} = 0.049$$

$$P_f \approx \frac{\sqrt{2N_e s/\pi}}{N} = \frac{\sqrt{2(50)0.1/\pi}}{100} = 0.018$$

$$N_e s = 50(-0.1) = -5 << -1; P_f = 0; h = -0.5$$

$$t = 2[ln(-1/sh) + 1 - \gamma = 2[ln(-1/0.1(-0.5) + 1 - 0.5772] = 6.84$$

$$-1/sh = -1/(0.1(-0.5)) = 20$$

(2) In a mutation accumulation experiment where 100 crossbred brother-sister lines of Drosophila melanogaster were used, the mean viabilities initially and in geneation 100 were 0.8 and 0.35, and the corresponding variances between lines 0.05 and 0.14, respectively. If the distribution of effects of mutations is assumed to be exponential, what are the estimates

that can be obtained for the genomic mutation rate, U, the mean effect of mutations $\bar{s}$ and the mutational variance V_M?

$$\triangle M = (v_0 - v_f)/t = (0.8 - 0.35)/100 = 0.0045$$
$$\triangle V = (V_0 - V_f)/t = (0.14 - 0.05)/100 = 0.0009$$
$$C^2/ = 1/\beta = 1/1$$
$$\frac{\triangle V}{\triangle M} = \bar{s}(1 + C^2)$$
$$\frac{\triangle V}{2\triangle M} = \bar{s} = 0.0009/(2(0.0045)) = 0.1$$
$$U = \frac{(1 + C^2)\triangle M^2}{\triangle V} = \frac{2(0.0045)^2}{0.0009} = 0.045$$
$$V_M = \triangle V/2 = 0.0009/2 = 0.000045$$

(3) A strain of Drosophila melanogaster devoid of genetic variation was created by means of crosses using balanced chromosomes. With this strain, lines were founded with effective size $N_e = 20$ and maintained for 100 generations in panmixia. We then evaluate a neutral quantitative trait whose mutational variance has been estimated to be $V_M = 0.0005$. What would be the approximate expected value of the variance between lines and the additive variance of the trait in the last generation?

$$V_{B,t} = 2tV_M = 2(100)(0.0005) = 0.1$$
$$V_A \approx 2N_eV_M = 2(20)0.0005 = 0.02$$
$$V_{A(t)} = V_{A(t-1)}[1 - 1/2(N_e)] + V_M$$

See ch7.r for the code to implement the recursive calculation for $V_{A(t)}$

(4) The rate of mutation of lethal recessives is $2.6 * 10^{-6}$ per gene and generation and the frequency of such alleles in a population is $1.4 * 10^{-4}$. Are these estimates compatible with the complete recessivity of lethal alleles? If not, what would be the most likely value of the dominance coefficient h?

The maximum possible frequency of a deleterious allele at the deleterious mutation selection balance can be approximated by $\hat{q}^2 = u/s$. For recessive lethals, $s = 1$ and $u = \sqrt{q} = \sqrt{2.6 * 10^{-6}} = 0.00161$, so this estimate is not compatable with complete recessivity of lethal alleles because there is about an order of magnitude difference between expected and observed allele frequencies. To calculate the dominance coefficient, we use $h = u/q = 0.00161/1.4 * 10^{-4} = 0.0186$

(5) The following data correspond to fitness value of six individuals of a population of large size and the mean fitness of their progeny by self-fertilization. Is it possible to estimate the deleterious genomic mutation rate, U, the mean effect of the mutations $\bar{s}$ and its average dominance coefficient, $\bar{h}$?

Yes, if we assume that the population is at the deleterious mutation-selection equilibrium. See chp7.r for the code to answer this question.

Self Assessment

(1) The estimate of heritability obtained from the regression of offspring values on that of their parents will necessarily be biased if the latter are not a random sample of the population.

False, if the group of parents were selected, the regerssion estimates would not be affected if the bias would occur to the same extent in the numerator and the denominator of the equation $b_{OP} = \frac{cov(O,P)}{\sigma_P^2}$