Methods of calculating N_e:

- Real populations may not meet the assumptions in the ideal population of the Wright and Fisher model
- Effective population size N_e: the size that would have an ideal population that gives rise to drift or inbreeding observed in a real population. In other words: it translates the size of a real population into the size of an idealized population that shows the same effects (drift or inbreeding) as the population under study).
- One method to calculate N_e: find the largest non-unit eigenvalue of a transition matrix P of a Markov chain. Calculate $N_e = 1/(2(1 - \lambda))$
- Can also use mutation effective size (defined from probability of identity in state of a gene under some mutation model with a defined rate)
- Most common method: coalescence effective size. Coalescence time of two alleles is the number of generations that must be traced back in the genealogy to the ancestor from which they come. If we can estimate the mean time of coalescence T in a real population we can predict effective size $N_e = T/2$

Prediction of N_e in populations with no selection:

- The absence of self-fertilization means that the appearance of inbreeding is delayed by one generation, reducing its' rate of increase, where $N_e = N+1/2$. This effect is negligible and will be ignored.
- Difference number of males and females: $N_e = \frac{4N_m N_f}{N_m + N_f}$, where N_m and N_f are the number of individuals that can intervene in reproduction
- The harmonic mean should be used whenever we average (effective) population sizes to predict their impact on dift or inbreeding

Variable population size across generations:

- In a population where all other assumptions of the ideal population by Wright and Fisher are met, if the population has a variable size: $1/N_e = 1/t \sum_{i=0}^{t-1} 1/2N_i$ (the effective size is the harmonic mean of population sizes in each generation)
- Example (population bottleneck): if a population size goes from 2500, 2000, 50, 1500 to 3000 then effects generated by inbreeding and drift are high because they correspond to that of a population of $N = 228$. This is why genetic diversity in a population can be lower than what is expected at its' current size.
- Variance and inbreeding effective sizes coincide but differ if population size changes
- Drift of allele frequencies depends on the number of offspring. If genealogical or molecular marker data is available, variance effective size should be estimated from rate of increase in coancestry
- Inbreeding in offspring depends on number of parents. If genealogical or marker data is available, inbreeding effective size should be estimated from rate of increase in inbreeding

Non-random contribution of parents to offspring:

- The probability that two random gametes taken at random in a population come from the same individual: $(\sum k_i^2 - \sum k_i)/(2N(2N - 1))$
- Using variance of samping k_i gametes, N_e can be calculated as $4N/(2 + S_K^2)$
- Under the ideal population, the distribution of the number of gametes contributed by the parents is Poisson
- These can be extended to where number of males and females differ

Partial mating among relatives:

- Hardy-Weinburg disequilibrium implies a decrease in the frequency of heterozygotes. In the previous calculation for N_e, this corresponds to the 2 term in the denominator

(representing variation from Mendelian segregation), which gets reduced by a fraction of α from chapter 4 since heterozygotes contribute 2 alleles to offspring generating less variation in allele frequencies
  - The variance term (representing variation from different contributions from individuals) in the previous equation for N_e is increased by a fraction of $alpha$, because homozygotes contribute a single allele to offspring generating more variation in allele frequencies
  - Culminating in $N_e = 4N/(2(1 - \alpha) + S_k^2(1 + \alpha))$

Overlapping generations:

  - Generations can become imbricated creating a complex structure of ages with differential survivals and fecundities within an age class.
  - With constant age structure, N_e can be approximated as $4N_a I_g/(2 + S_K^2)$ where N_a is the number of reproductive individuals that enter the population in each cohort and I_g is the generational interval
  - Without a constant age structure, then if genealogical relationships are available the effective size can be calculated from rate of increase in inbreeding $N_e = 1/(2\Delta \bar{F})$

Prediction of N_e in selected populations:

  - Genetic draft: frequency of a neutral allele carried by a successful breeder will increase in successive generations. These changes are more directional than those by drift.
  - Selection reduces N_e, translating to increases in genetic drift and inbreeding
  - Linkage: linked loci decrease N_e
  - Background selection
  - Deleterious mutations: natural selection compensates for declines in fitness
  - Neutral gene: depends on frequency of recombination. As linkage between neutral and selective alleles increases, reduction of effective size is greater

Prediction of N_e in subdivided populations:

  - GO BACK AND REVIEW

Applications of theory of N_e to conservation:

  - N_e provides a summary of the past history of a population in regards to inbreeding and drift
  - Mating procedures can be designed to eliminate the variance of parental contributions, leaving behind Mendelian segregation of heterozygotes as the only source of drift
  - Maker assisted selection can reduce variation from Mendelian segregation of heterozygotes too
  - Or, manipulate meiosis such that we use more than one gamete from a single meiosis

Problems

(1) In problem 4.1 we considered a set of ideal populations of census size N = 20 individuals each. Suppose now that the populations are maintained with 4 males and 16 females each generation. What would be the expected values of the heterozygosity, the variance in the allele frequencies and the inbreeding at generations $t = 5, 20$ and 100?
    Test answer+
(2) In problem 4.2 we considered a plant species whose natural reproduction is a combination of autogamy and allogamy in equal proportions. From a large population of this species a line with census size $N = 20$ was founded, which was maintained for 10 generations in the greenhouse by allogamous reproduction.

Self Assessment