

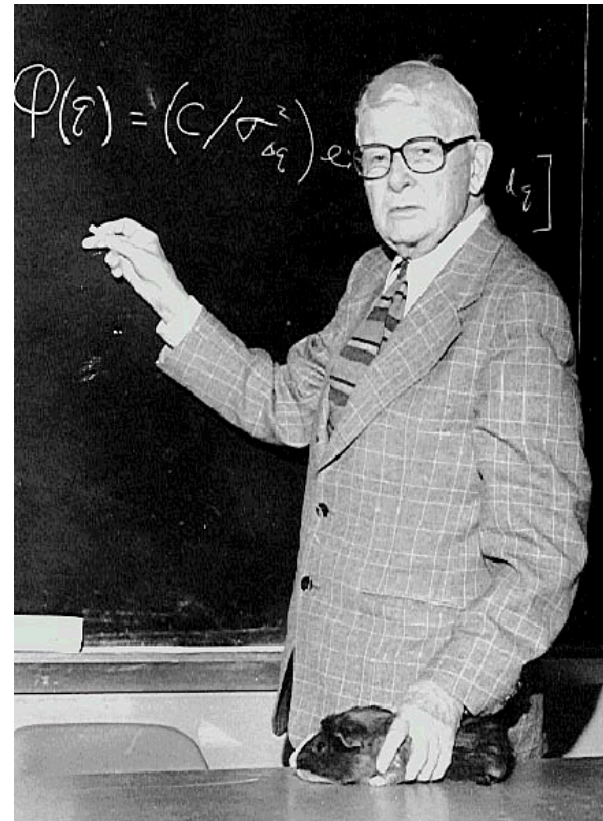
Week 6

Simulation and Modeling Evolution

Population genetics

Sewell Wright J.B.S. Haldane

- Microevolution
- Change in the genetic make up of populations $p_t \xrightarrow{\text{evolution}} p_{(t+1)}$
- Evolutionary forces
- Deterministic forces
 - Selection
 - Migration
 - [Mutation]
- Stochastic forces
 - Genetic drift
 - [Recombination]
- Population genetics as a balance of interacting forces



R. A. Fisher

What do Mendelian alleles do in a population?



Godfrey Harold Hardy
(1877 – 1947)

Hardy-Weinberg Equilibrium



Wilhelm Weinberg
(1862 — 1937)

These two gentlemen figured it out.

Minding Your Ps and Qs

Genotype:	A_1A_1	A_1A_2	A_2A_2
Relative Frequency:	x_{11}	x_{12}	x_{22}

Constraint:

$$x_{11} + x_{12} + x_{22} = 1$$

Treat genotype/allele frequencies as random variables

Minding Your Ps and Qs

First calculate alleles freqs

Freq A_1 allele

$$p = x_{11} + \frac{1}{2}x_{12}$$

Freq A_2 allele

$$q = 1 - p = x_{22} + \frac{1}{2}x_{12}$$

Where is this coming from?

Minding Your Ps and Qs

Numerical Example

Sickle Cell vs. Normal Hemoglobin in Sub Saharan Village

Genotype:	A_1A_1	A_1A_2	A_2A_2
Observed Numbers:	411	1404	185
Relative Frequencies:	0.2055	0.702	0.0925

Calculate the frequency of A_1 allele

Minding Your Ps and Qs

Numerical Example

Sickle Cell vs. Normal Hemoglobin in Sub Saharan Village

Genotype:	A_1A_1	A_1A_2	A_2A_2
Observed Numbers:	411	1404	185
Relative Frequencies:	0.2055	0.702	0.0925

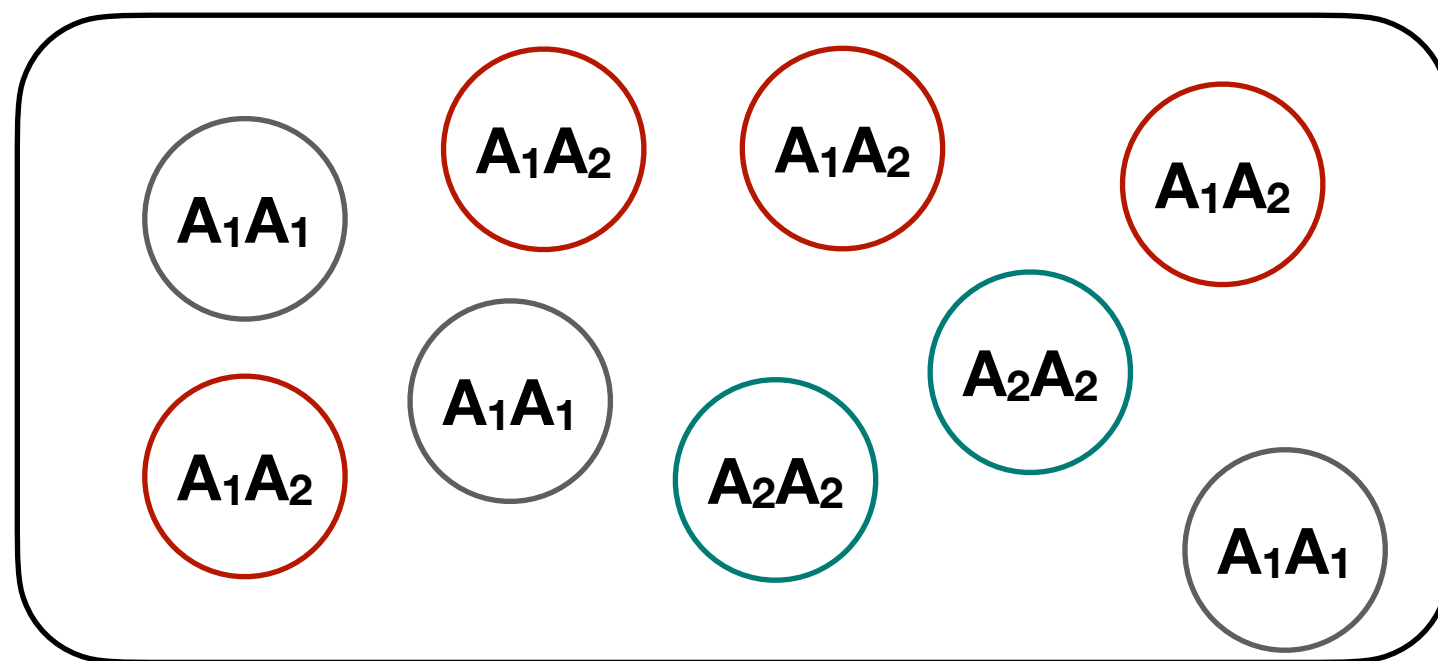
Calculate the frequency of A_1 allele

$$p = 0.2055 + \left(\frac{1}{2} \times 0.702\right) = 0.5565$$

Minding Your Ps and Qs

Consider this “experiment”:

- 1) Sample a genotype from popn
- 2) Sample an allele from that genotype



$$x_{11} = P(A_1A_1)$$

$$x_{12} = P(A_1A_2)$$

$$x_{22} = P(A_2A_2)$$

Probability of getting A_1 allele, p , is equal to

$$p = (x_{11} \times 1) + (x_{12} \times \frac{1}{2}) + (x_{22} \times 0)$$

The Hardy-Weinberg Law

Assumptions:

- Diploid sexual population
- Infinite population size
- Random mating
- no selection
- no migration
- no mutation

This is a Null model. Why is this useful?

Hardy-Weinberg

Write down genotype freqs in terms of allele freqs

Genotype:	A_1A_1	A_1A_2	A_2A_2
H-W Frequency:	p^2	$2pq$	q^2

$$p^2 + 2pq + q^2 = 1$$

Again ask yourself, where is this coming from?

Hardy-Weinberg

Write down genotype freqs in terms of allele freqs

Genotype:	A_1A_1	A_1A_2	A_2A_2
H-W Frequency:	p^2	$2pq$	q^2

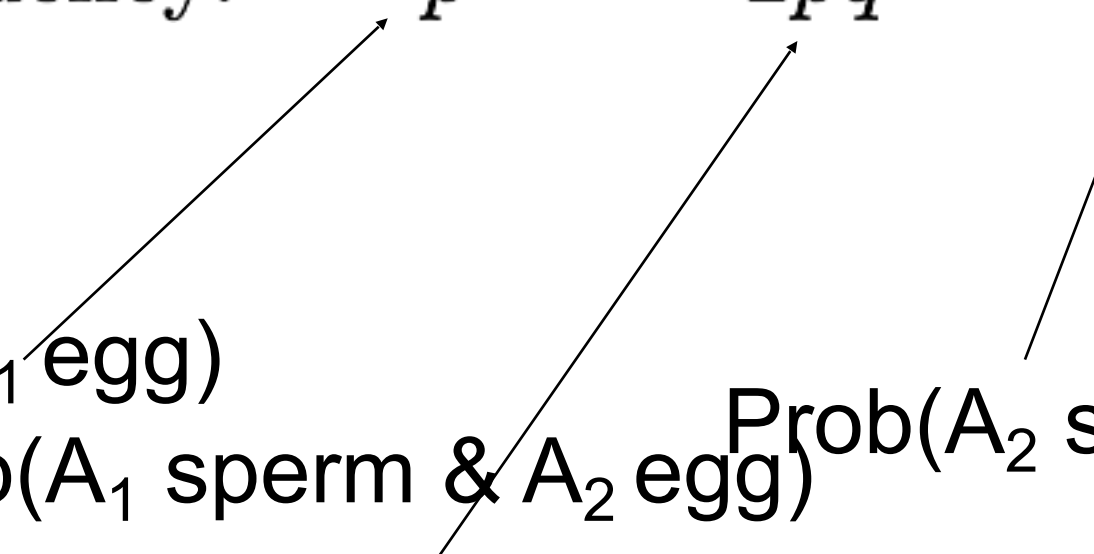
Prob(A_1 sperm & A_1 egg)

Prob(A_1 sperm & A_2 egg)

+

Prob(A_2 sperm & A_1 egg)

Prob(A_2 sperm & A_2 egg)



Hardy-Weinberg

Numerical Example

Back to sickle cell anemia...

Genotype:	A_1A_1	A_1A_2	A_2A_2
H-W Frequency:	p^2	$2pq$	q^2
Observed Frequencies:	0.2055	0.702	0.0925
H-W Expected Frequencies:	0.309	0.494	0.197

Hardy-Weinberg

Numerical Example

Back to sickle cell anemia...

Genotype:	A_1A_1	A_1A_2	A_2A_2
H-W Frequency:	p^2	$2pq$	q^2
Observed Frequencies:	0.2055	0.702	0.0925
H-W Expected Frequencies:	0.309	0.494	0.197

Pretty big deviations, could assess how big using statistics!

Heterozygosity

“temperature of a population”

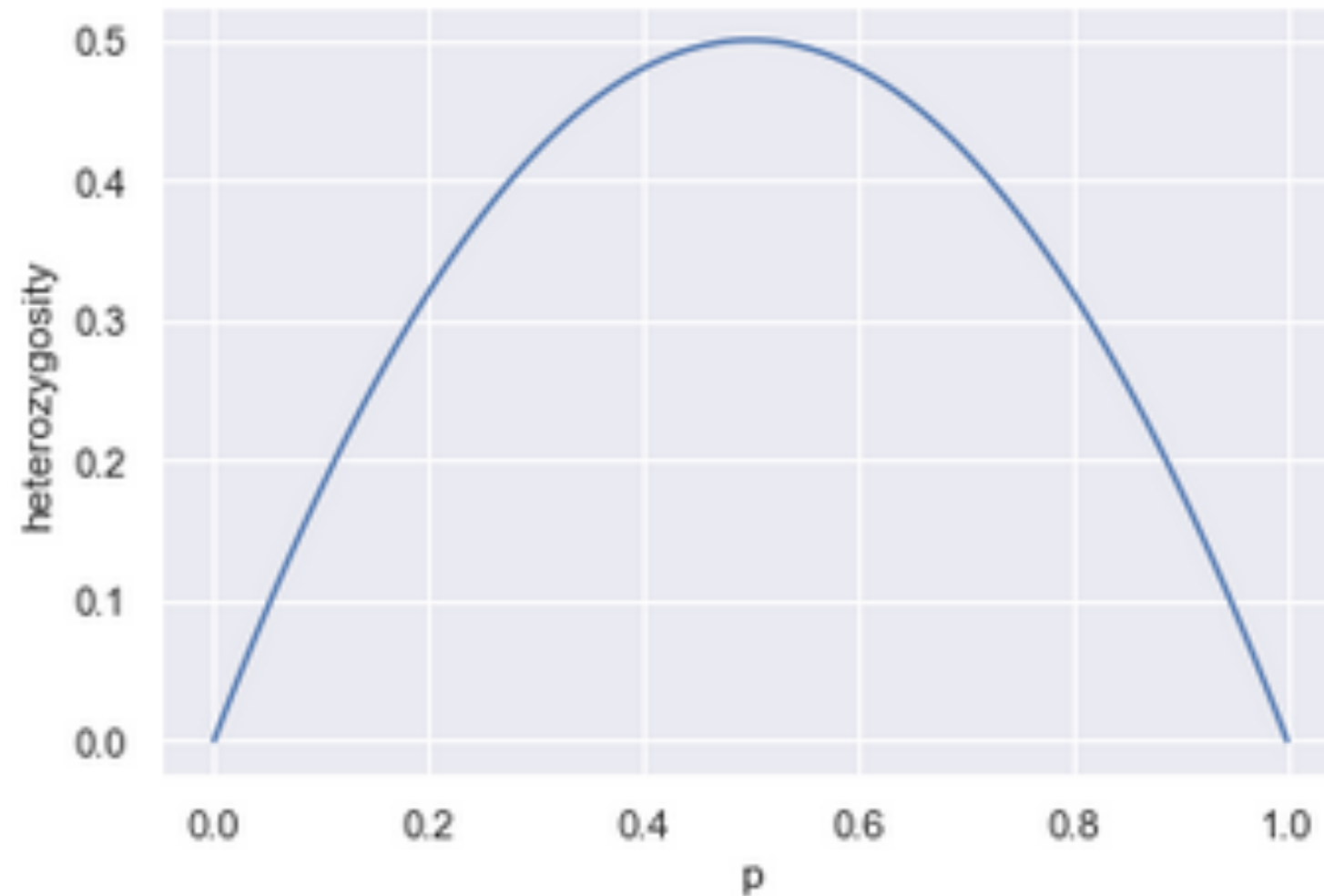
Genotype:	A_1A_1	A_1A_2	A_2A_2
H-W Frequency:	p^2	$2pq$	q^2

Call heterozygosity the probability of sampling an individual that is heterozygous

$$\hat{H} = 2pq = 2p(1 - p)$$

Heterozygosity

“temperature of a population”



$$\hat{H} = 2pq = 2p(1 - p)$$

The Hardy-Weinberg Law

Assumptions:

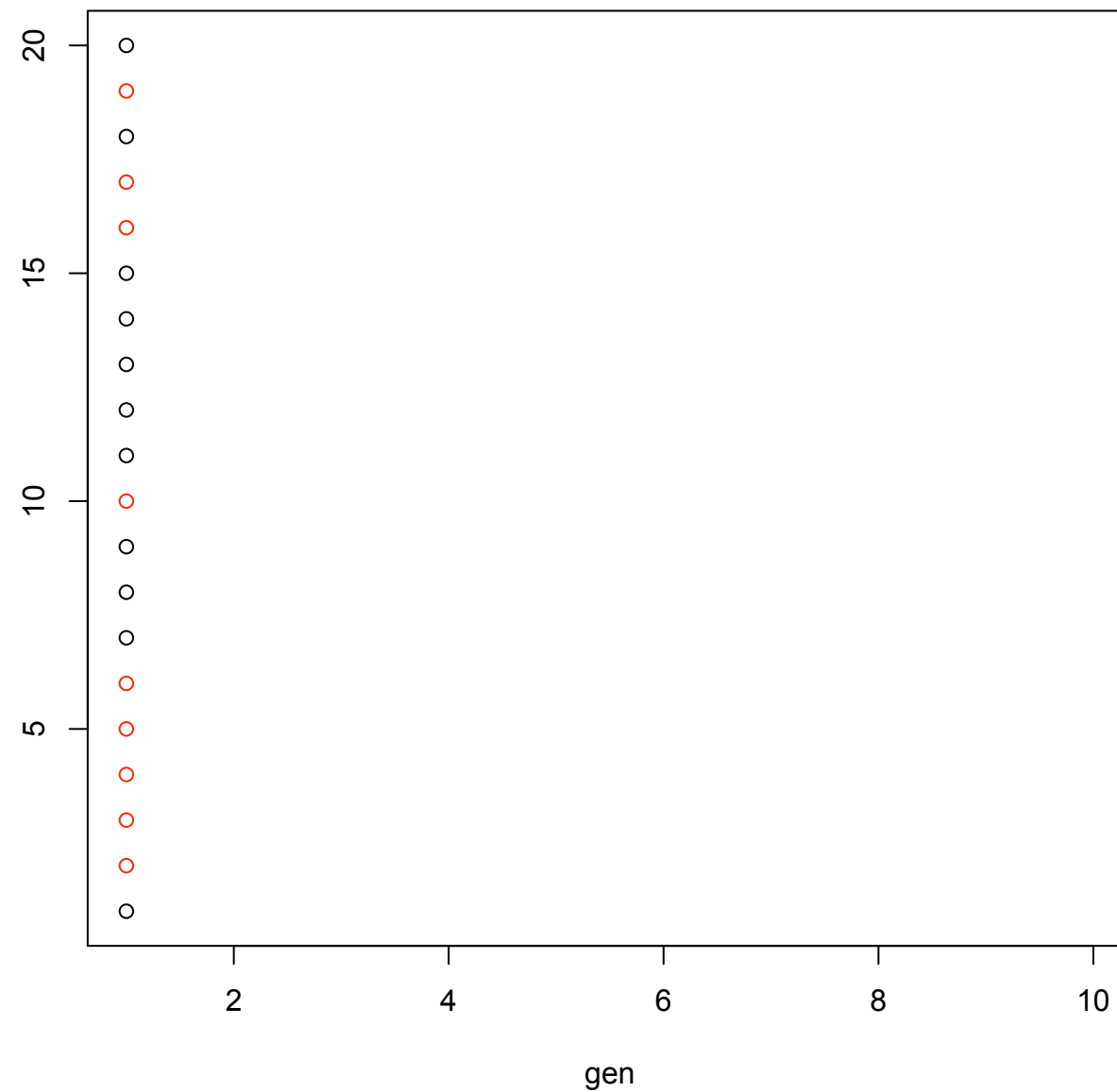
- Diploid sexual population
- Infinite population size
- Random mating
- no selection
- no migration
- no mutation

This is a Null model. Why is this useful?

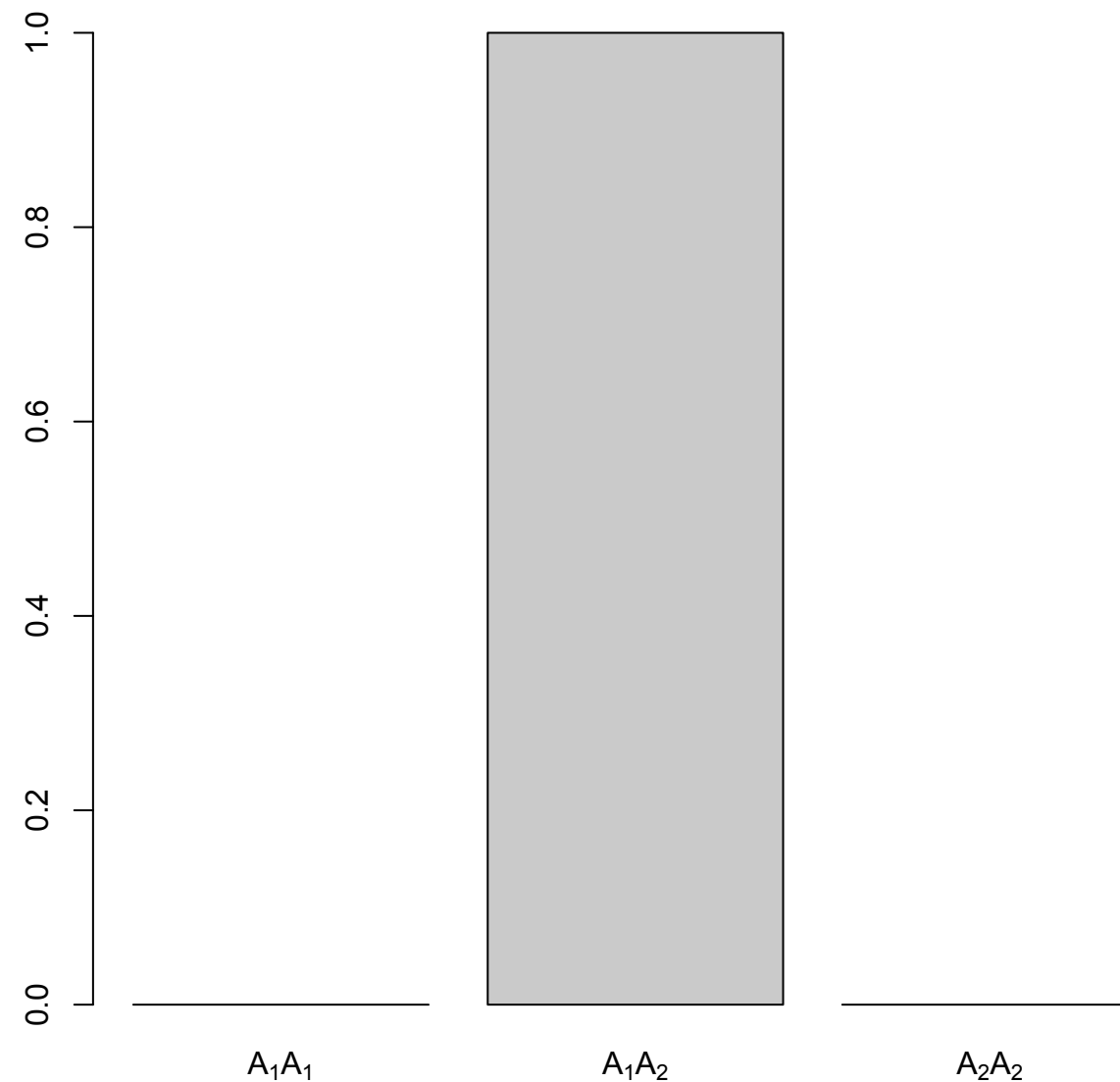
Genetic Drift: the algorithm

1. Choose an allele at random from the $2N$ alleles in the parental generation
2. Make an exact copy of the allele
3. Place the copy of the allele in the next generation
4. Go back to 1 until next generation has $2N$ alleles

Genetic Drift: the algorithm



Drift in population of size 1



What is probability of being heterozygote in next generation?

Drift in population of size 1

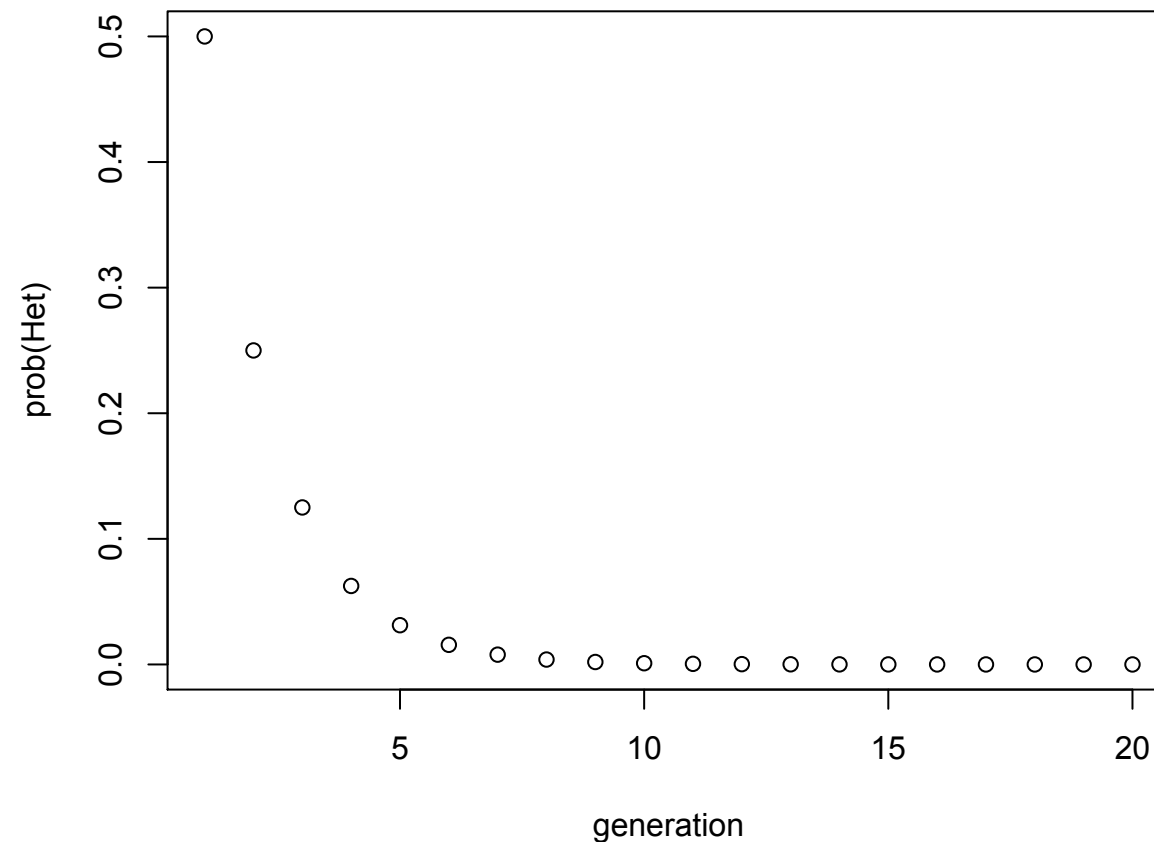
What is probability of being heterozygote in next generation?

$$Prob(Heterozygote) = \frac{1}{2}$$

$$Prob(Het|t) = \left(\frac{1}{2}\right)^t$$

Decay in heterozygosity due to drift

$$Prob(Het|t) = \left(\frac{1}{2}\right)^t$$



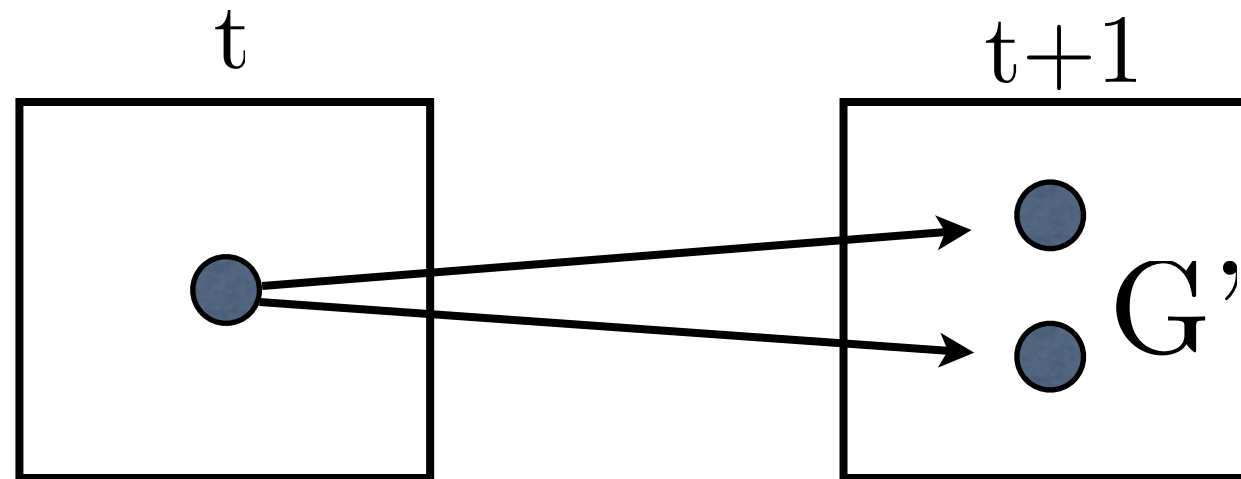
Question: How many generations until fixation on average?

Decay in heterozygosity due to drift

Identity by origin: two alleles that are from same locus on the **same** chromosome

Identity by state: two alleles that are of the same form (e.g. same DNA sequence)

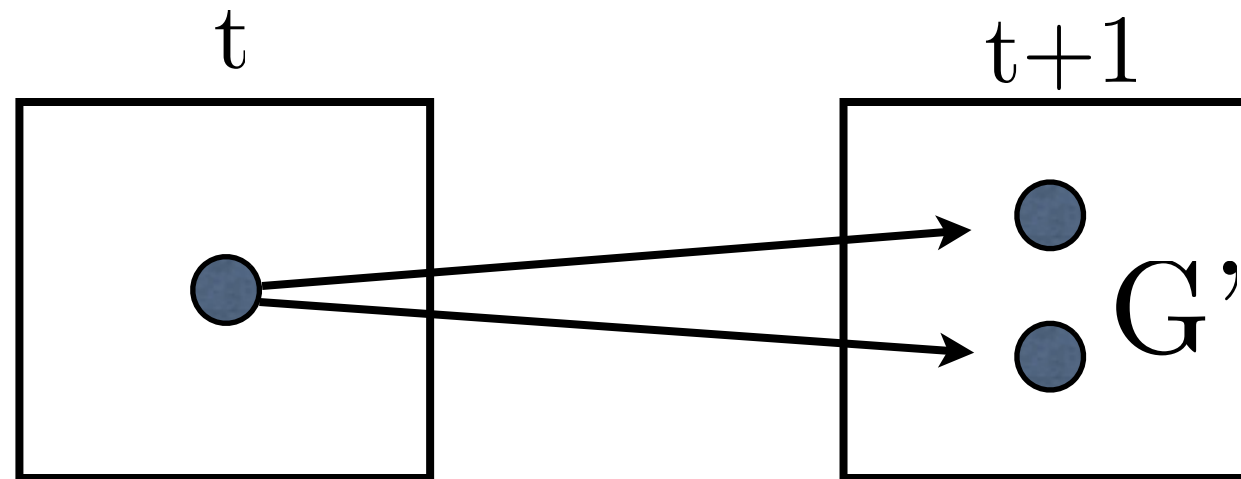
Decay in heterozygosity due to drift



Probability that two alleles chosen at random
are identical by state

No variation $G = 1$
Every allele unique $G = 0$

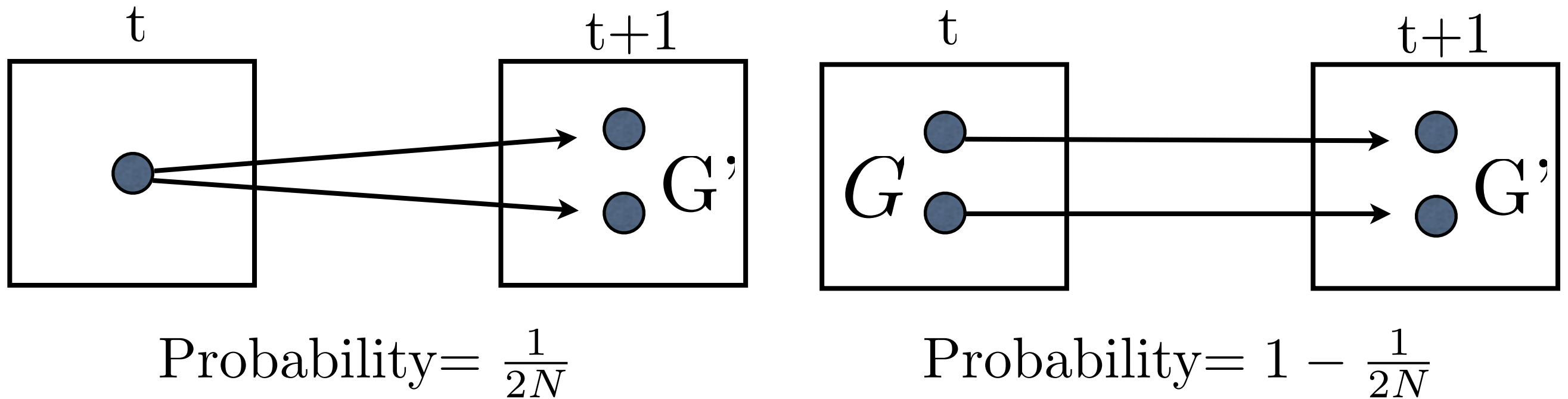
Decay in heterozygosity due to drift



Probability that two alleles chosen at random
are identical by state

$$G' = \frac{1}{2N} + \left(1 - \frac{1}{2N}\right) G$$

Decay in heterozygosity due to drift



$$G' = \frac{1}{2N} + \left(1 - \frac{1}{2N}\right) G$$

Decay in heterozygosity due to drift

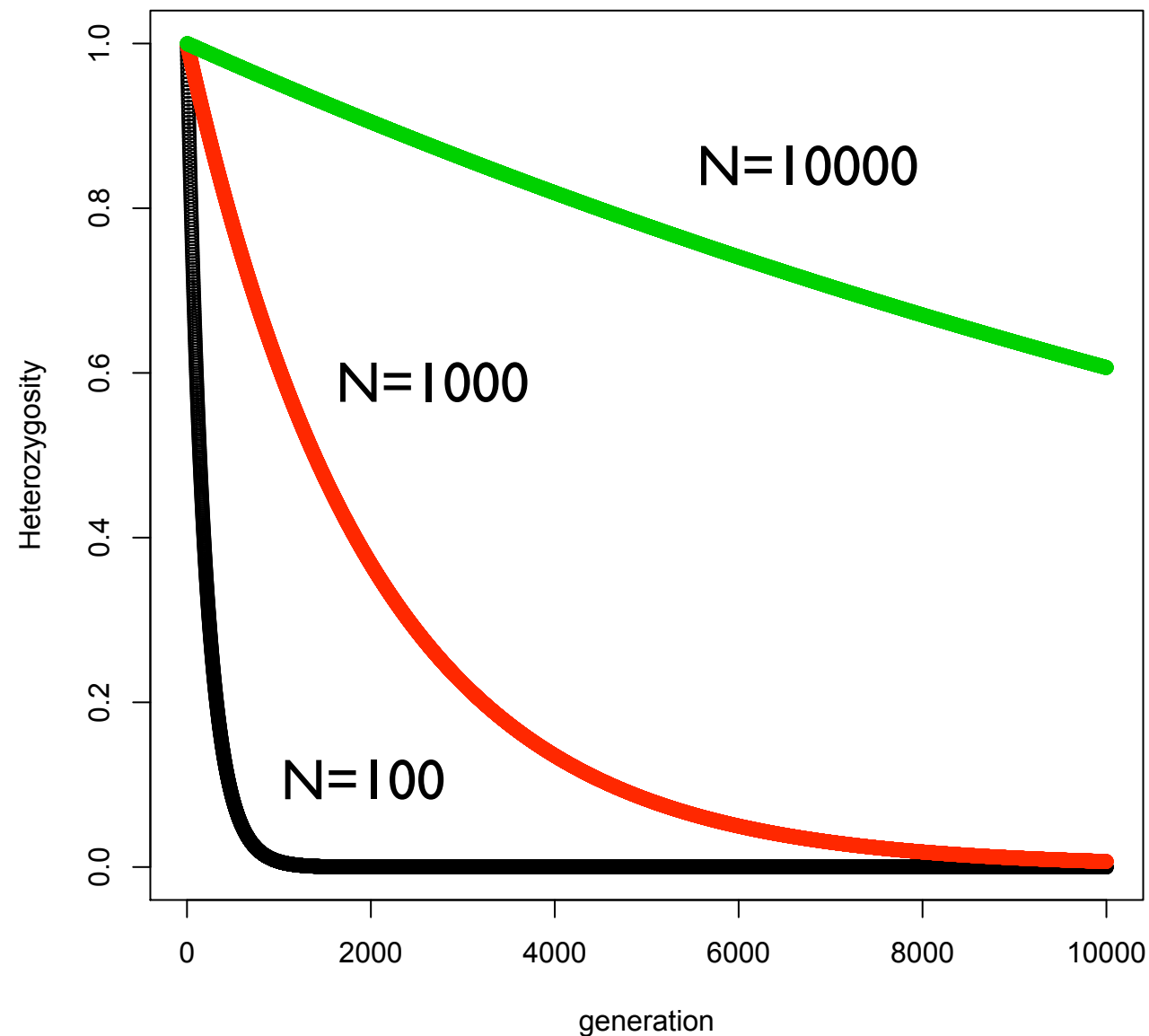
Probability that two alleles chosen at random
are different by state

$$H = 1 - G$$

$$H' = 1 - G' = \left(1 - \frac{1}{2N}\right) H$$

$$\begin{aligned}\Delta_N H &= H' - H \\ &= \left(1 - \frac{1}{2N}\right) H - H \\ &= -\frac{1}{2N} H\end{aligned}$$

Decay in heterozygosity due to drift



Difference Equation

$$H_t = H_0 \left(1 - \frac{1}{2N} \right)^t$$

Ultimately all Heterozygosity is lost!

So who fixes?

2N different alleles: $Prob(fix) = \frac{1}{2N}$

i copies of an allele: $Prob(fix) = \frac{i}{2N}$

So for neutral alleles: $Prob(fix|p) = p$