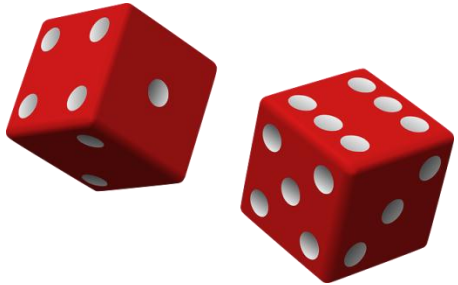
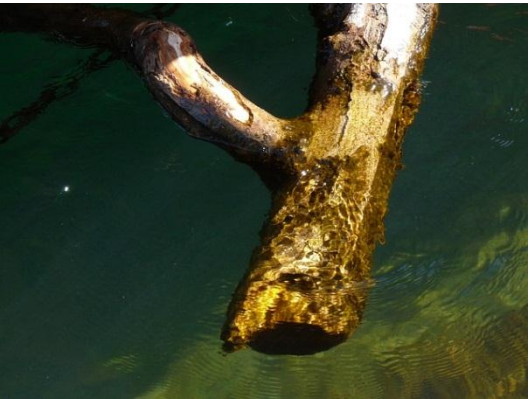
Duke
UNIVERSITY



GENETIC DRIFT:

Sampling error over single generations



Natural selection is *predictable*

- Some genotypes have a higher fitness
  - Higher fitness leads to more offspring
  - Genotypes become "overrepresented"

- If the fitness is known, then change by natural selection is *"predictable"*
- But not all evolutionary change is predictable...

Random chance matters

- Bag of many marbles
- Exactly half are brown
- Exactly half are blue

Random chance matters

- Bag of many marbles
- Exactly half are brown
- Exactly half are blue

- What if we picked exactly **4** marbles?
  - How many of each color would we get?

Random chance matters

- Bag of many marbles
- Exactly half are brown
- Exactly half are blue

- What if we picked exactly **4** marbles?
  - How many of each color would we get?

    ~5% chance we would get all
    4 marbles same color

Random chance matters

- Bag of many marbles
- Exactly half are brown
- Exactly half are blue

- What if we picked exactly **2** marbles?
  - How many of each color would we get?

Random chance matters

- Bag of many marbles
- Exactly half are brown
- Exactly half are blue

- What if we picked exactly **2** marbles?
  - How many of each color would we get?

    ~50% chance we would get
    both marbles same color

Sampling error!

- Previous example
  - Picking 4 is likely to get you roughly right proportions
  - Picking 2 is *not* likely to get you roughly right proportions
  - By picking MORE, you get **a more representative sample** of the original pool

Same principle applies in nature

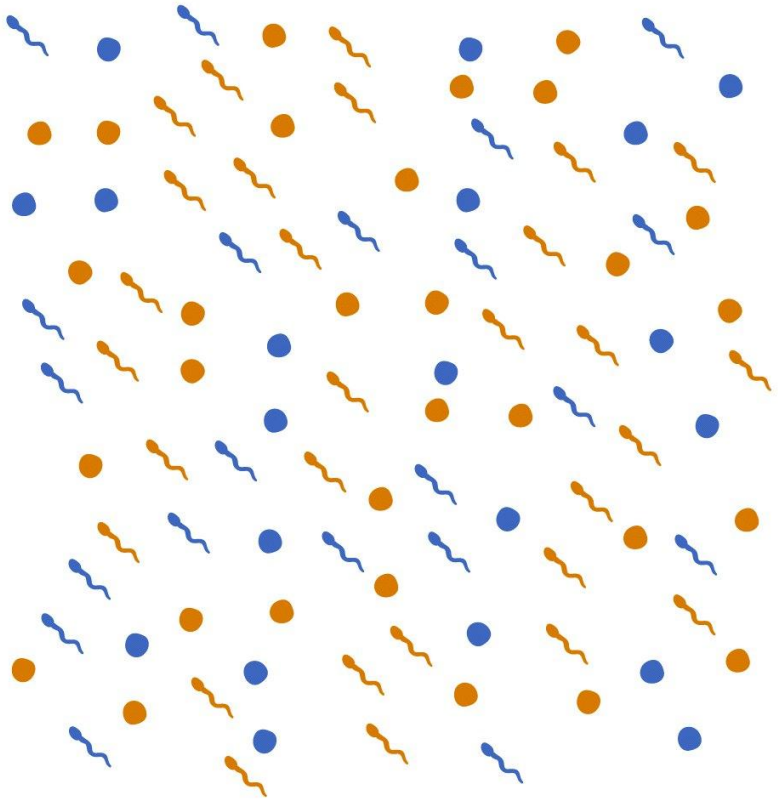


Figure 7-10b Evolutionary Analysis, 4/e

- Populations are **not** infinite
- Frequently, a small (not-perfectly-representative) sample of gametes form the next generation
  - Allele and genotype frequencies change
- Effect compounds over time

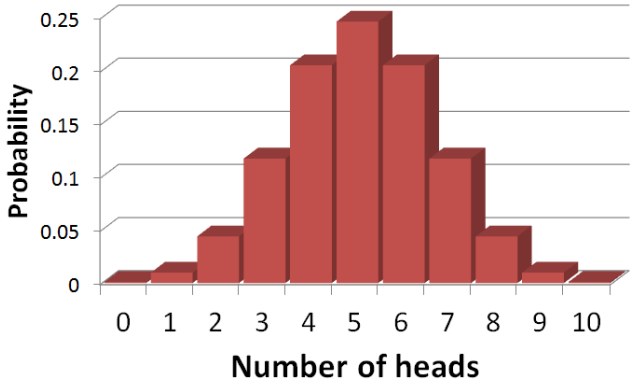
Sampling error is random in direction *over one generation*

- *Assuming there is more than one allele*, any allele is **about equally** likely to increase or decrease in frequency in one generation by sampling error
- If p=0.6, about equally likely to be p>0.6 or p<0.6 in next generation
  - But very unlikely to be EXACTLY p=0.6 again
- Allele frequency "drifts" due to sampling error: **genetic drift**

Small changes are likely.

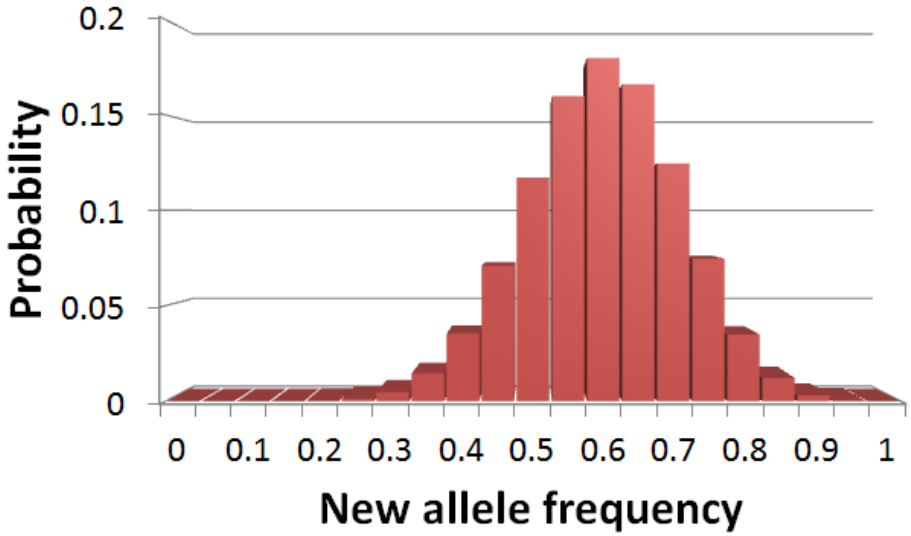
Big changes are possible but unlikely.

- Imagine tossing a coin 10 times (similar to p=0.5)
- May get 5 heads
  - Getting >5 heads or <5 heads equally likely
  - Getting 0-1 or 9-10 heads very unlikely
    - 10 heads: ~1/1000 chance



Same concept for populations

- Original population, p(A) = 0.6
- Shown below: probability of p(A) after one generation, if there are 10 (diploid) offspring



Magnitude of change compounds and relates to the population size

- Greater changes occur in the allele frequency if the sample (population) is smaller
- Population size = **400**
- A_1 starts at 0.5

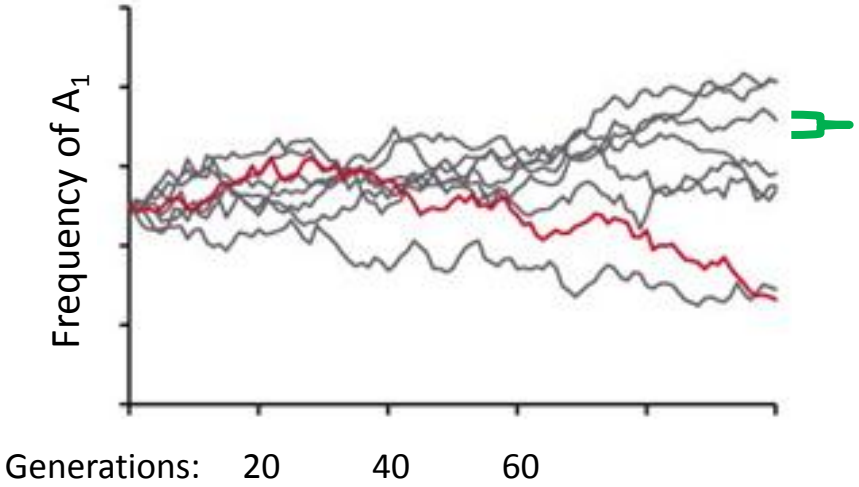


Figure 7.15 Evolutionary Analysis, 4/e

Magnitude of change compounds and relates to the population size

- Greater changes occur in the allele frequency if the sample (population) is smaller
- Population size = **40**
- A_1 starts at 0.5

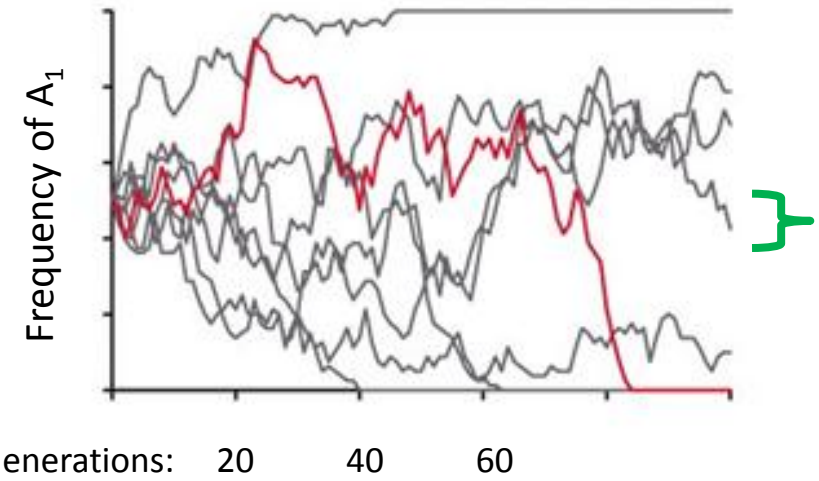


Figure 7.15 Evolutionary Analysis, 4/e

Magnitude of change compounds and relates to the population size

- Greater changes occur in the allele frequency if the sample (population) is smaller
- Population size = **4**
- A_1 starts at 0.5

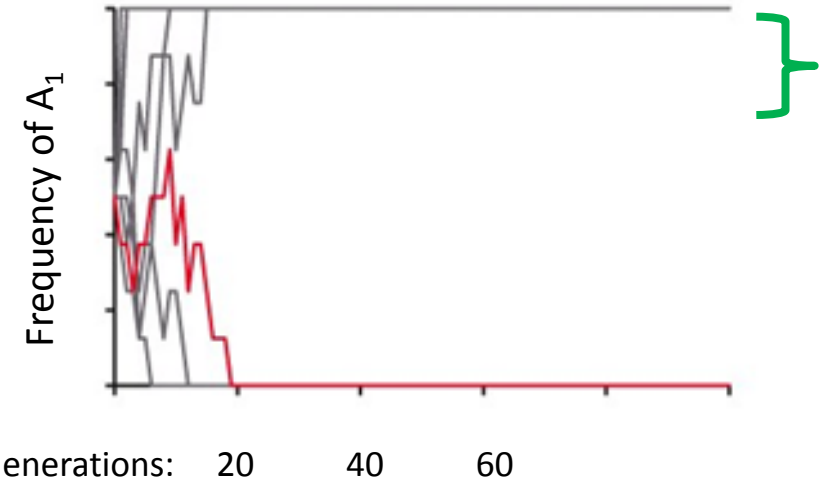


Figure 7.15 Evolutionary Analysis, 4/e

How big are the individual steps (on average)?

- Variance in allele frequency due to one generation of drift

  $= (pq)/(2N)$

  – p & q are allele frequencies, N is the population size
- Standard deviation is (slight over-)estimate of average allele frequency change in one generation: $\sqrt{(pq)/(2N)}$

- For N=4, p=0.5, q=0.5, average change estimate ~ 0.18
  – Likely to go to p=0.68 or p=0.32, *on average* – could be more or less
- For N=40, p=0.5, q=0.5, average change estimate ~ 0.06
- For N=400, p=0.5, q=0.5, average change estimate ~ 0.02

Take-home messages

- Drift is **strongest in small populations**
  - Variance due to one generation of drift = (pq)/(2N)
- Drift is neither ***predictable in direction*** in one generation nor ***exactly replicable in degree***
  - Under exact same conditions, get different results from genetic drift
- Drift can cause ***big changes in allele frequency over time***

(next video)

Image Credits, Unit 13-1

- Coin flip, © Nicolai Sorokin, all rights reserved, www.photoxpress.com
- Bag of marbles, © Ivonne Wierink, all rights reserved, www.photoxpress.com
- Genetic drift, © 2007 Pearson Education Inc., all rights reserved, Figure 7.15, Evolutionary Analysis 4/e by Freeman & Herron
- Outdoor steps, © JoAnn Hess, all rights reserved