To get challenging sampling distributions (discussed in prior videos):

1. Randomization (AKA permutation, Chapter 13)
2. Simulation (Chapter 19)

To get an estimate of precision

(Note about a potential source of confusion: you are *approximating* the sampling distribution but not using it to test a hypothesis; just use it to get a SE)

1. Bootstrapping (Chapter 19)

Bootstrapping:

- 're-sampling' the actual data
  - **Sampling with replacement**
  - Pick an equal number of points for each group

- Approximates the *sampling distribution* of an estimate
  - ***But NOT*** *the null (sampling) distribution as we saw with simulation and randomization*

- Nonparametric and be applied to virtually any parameter – including means, proportions, correlations, linear model coefficients

- Used to find confidence interval and the bootstrap standard error
  - Precision method
  - Particularly useful when there is no ready formula for standard error (median, eigenvalue)

- Estimate uncertainty in phylogenies

Bootstrapping Method:

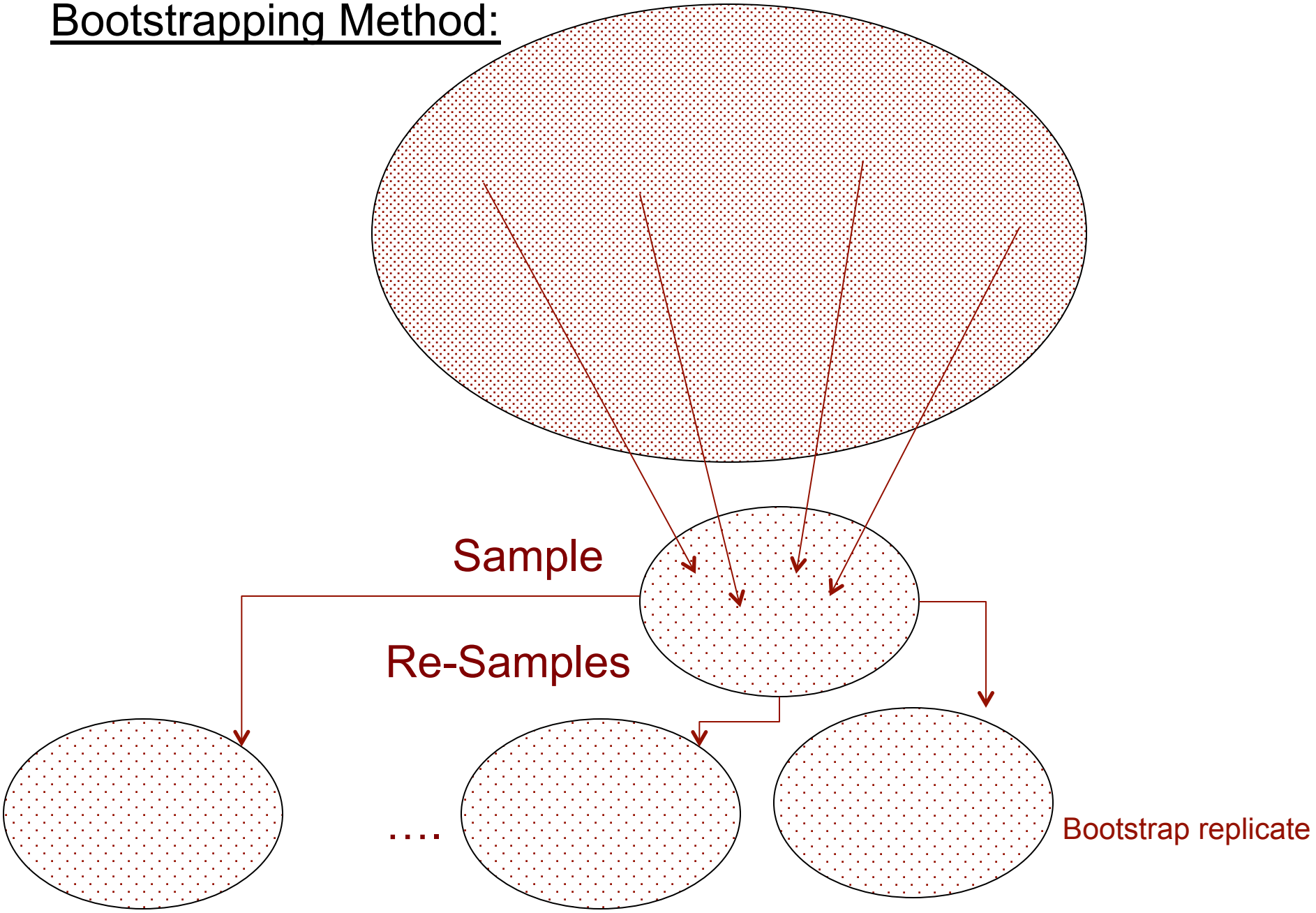
1. Use computer to take random individuals **with replacement** from the data

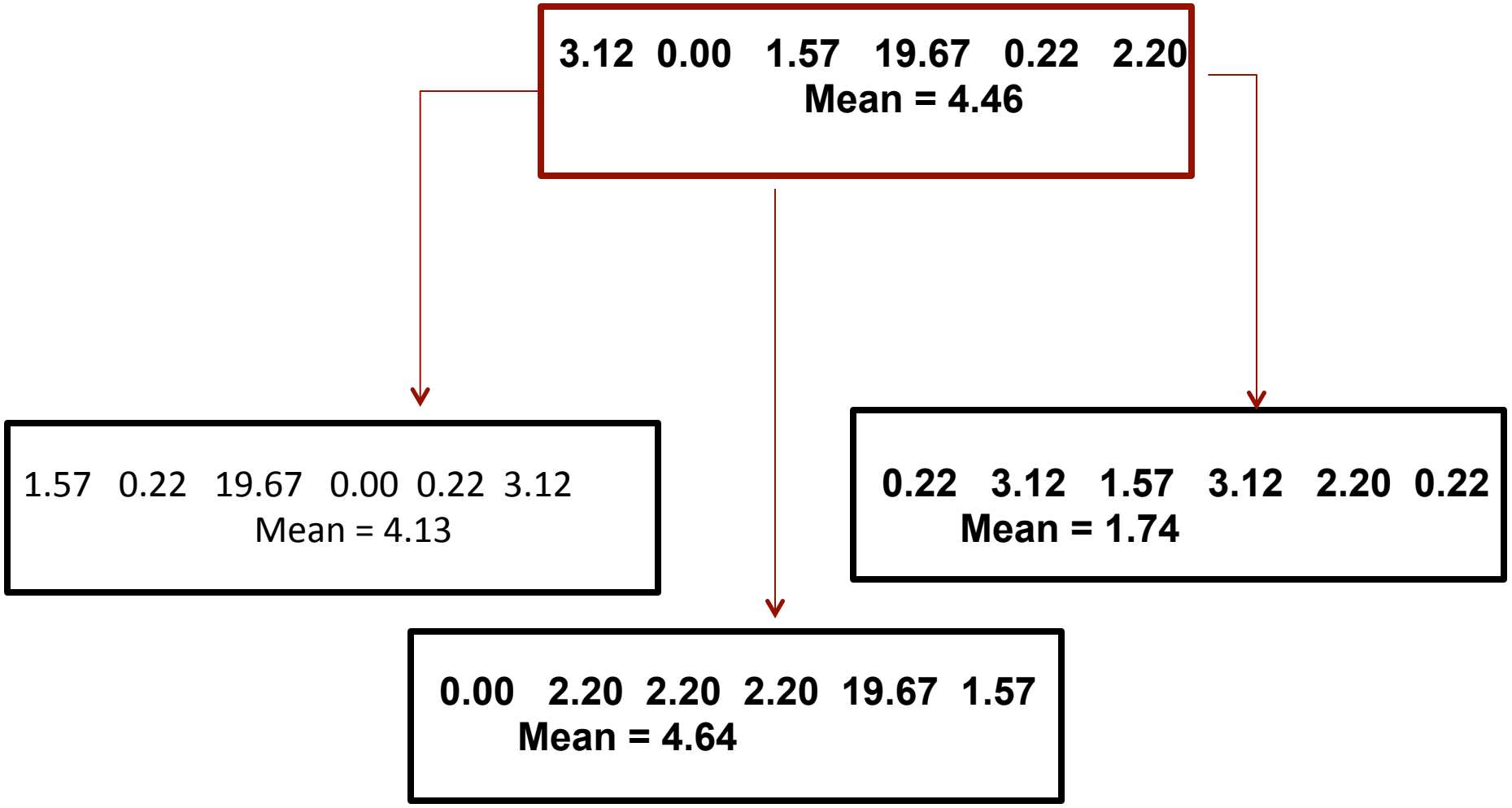
2. Calculate estimate from **bootstrap replicate**

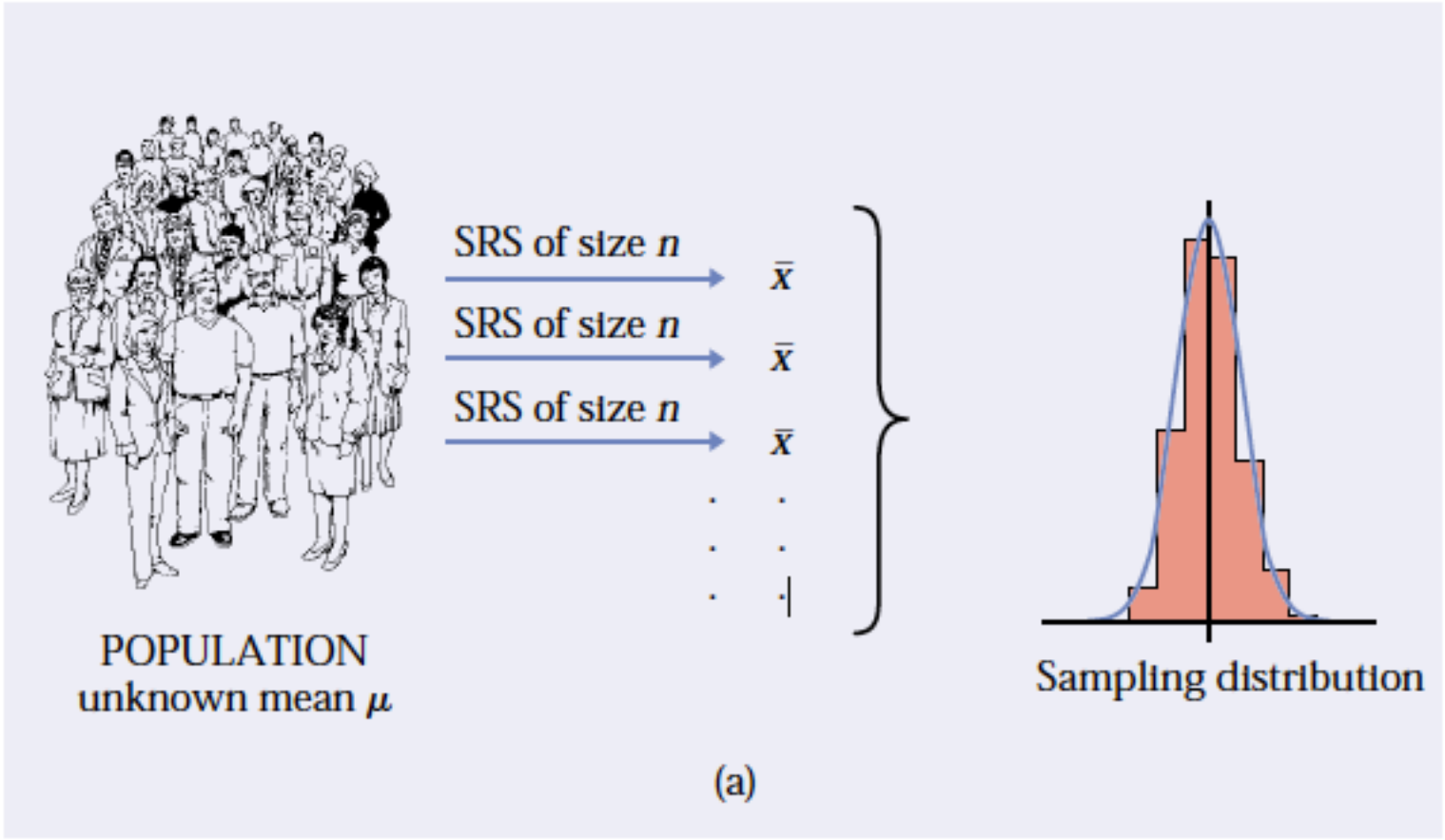
3. Repeat steps 1 and 2 **many times** to approximate the **sampling distribution of the estimate**
   - **not** the null distribution
   - Overall mean of the bootstrap replicate estimates are close to mean of original data
   - Similar- but not the same -concept as the Standard Error of the sampling mean

4. Calculate **bootstrap standard error**
   - sample standard deviation of all bootstrap replicate estimates

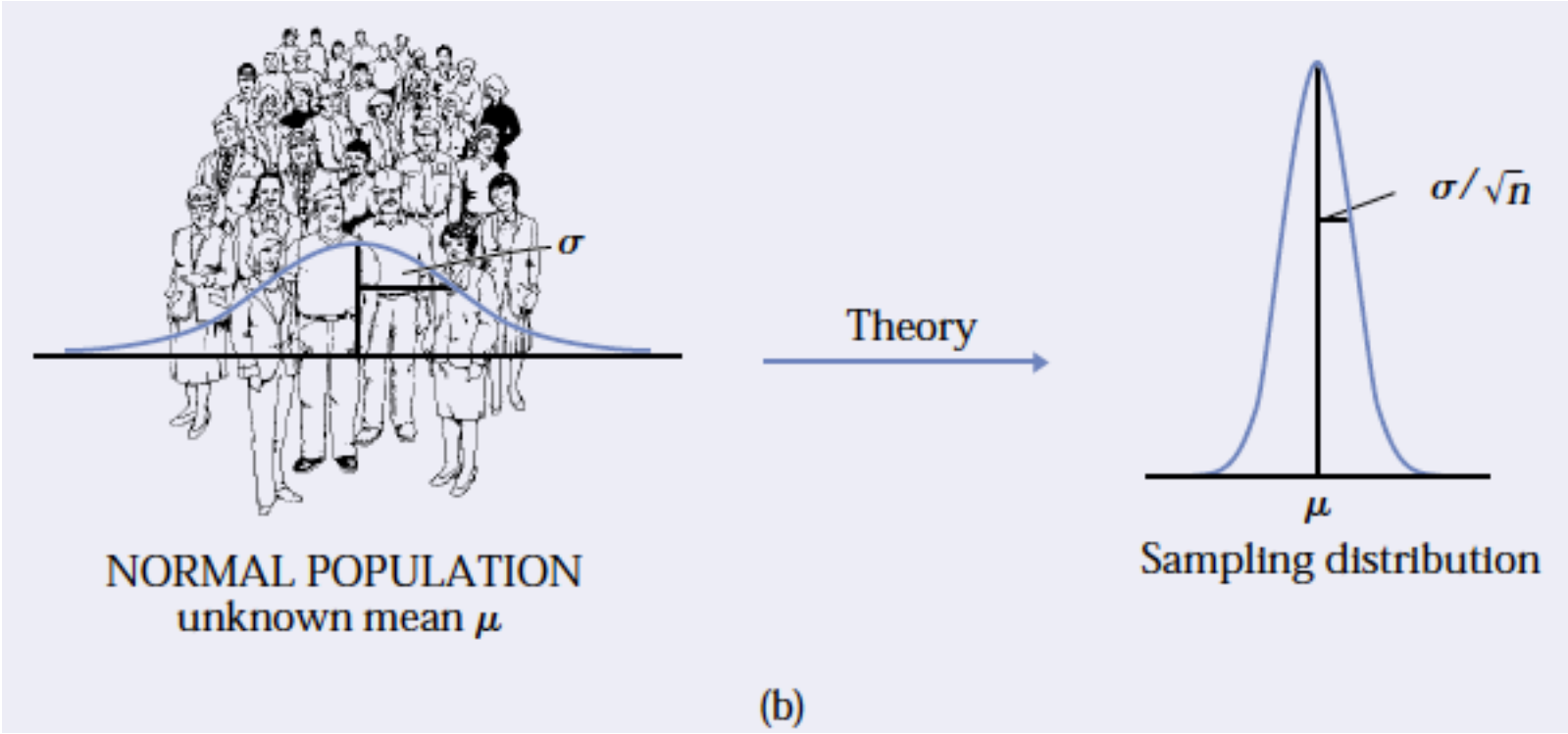
Bootstrapping Method:




3.12 0.00 1.57 19.67 0.22 2.20
Mean = 4.46
1.57 0.22 19.67 0.00 0.22 3.12
Mean = 4.13
0.22 3.12 1.57 3.12 2.20 0.22
Mean = 1.74
0.00 2.20 2.20 2.20 19.67 1.57
Mean = 4.64

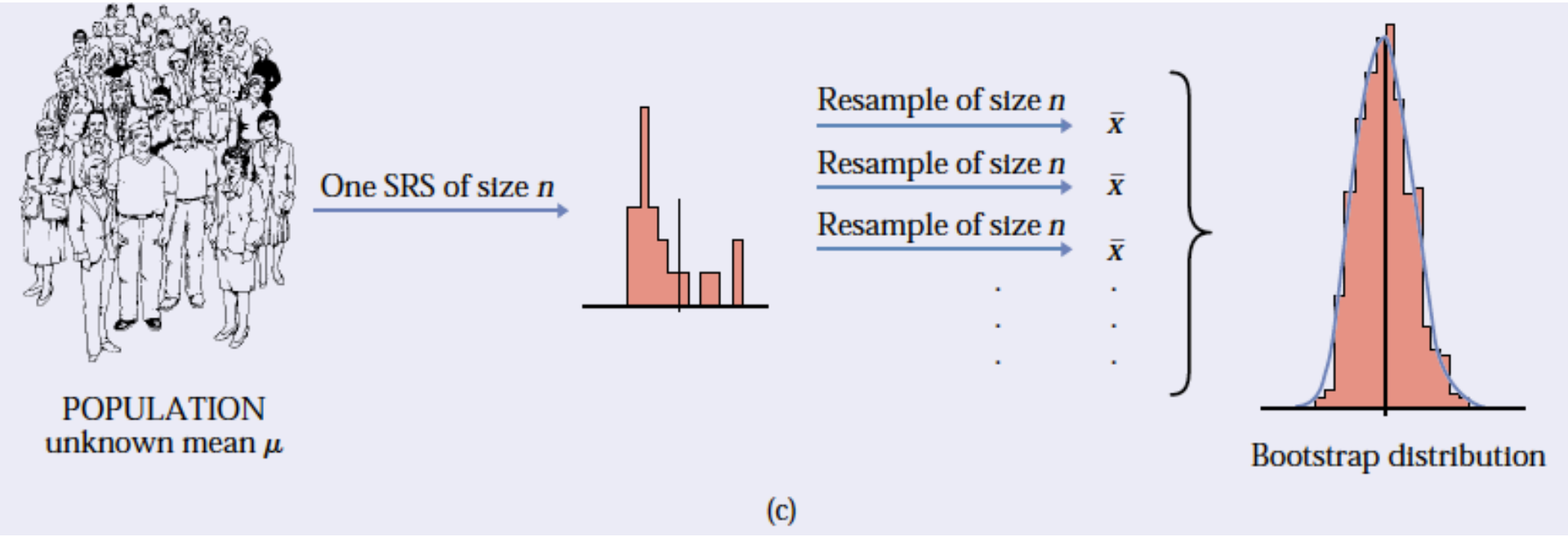


http://bcs.whfreeman.com/ips5e/content/cat_080/pdf/moore14.pdf



http://bcs.whfreeman.com/ips5e/content/cat_080/pdf/moore14.pdf

Bootstrapping Examples:



http://bcs.whfreeman.com/ips5e/content/cat_080/pdf/moore14.pdf

Providing an estimate of uncertainty along branches:

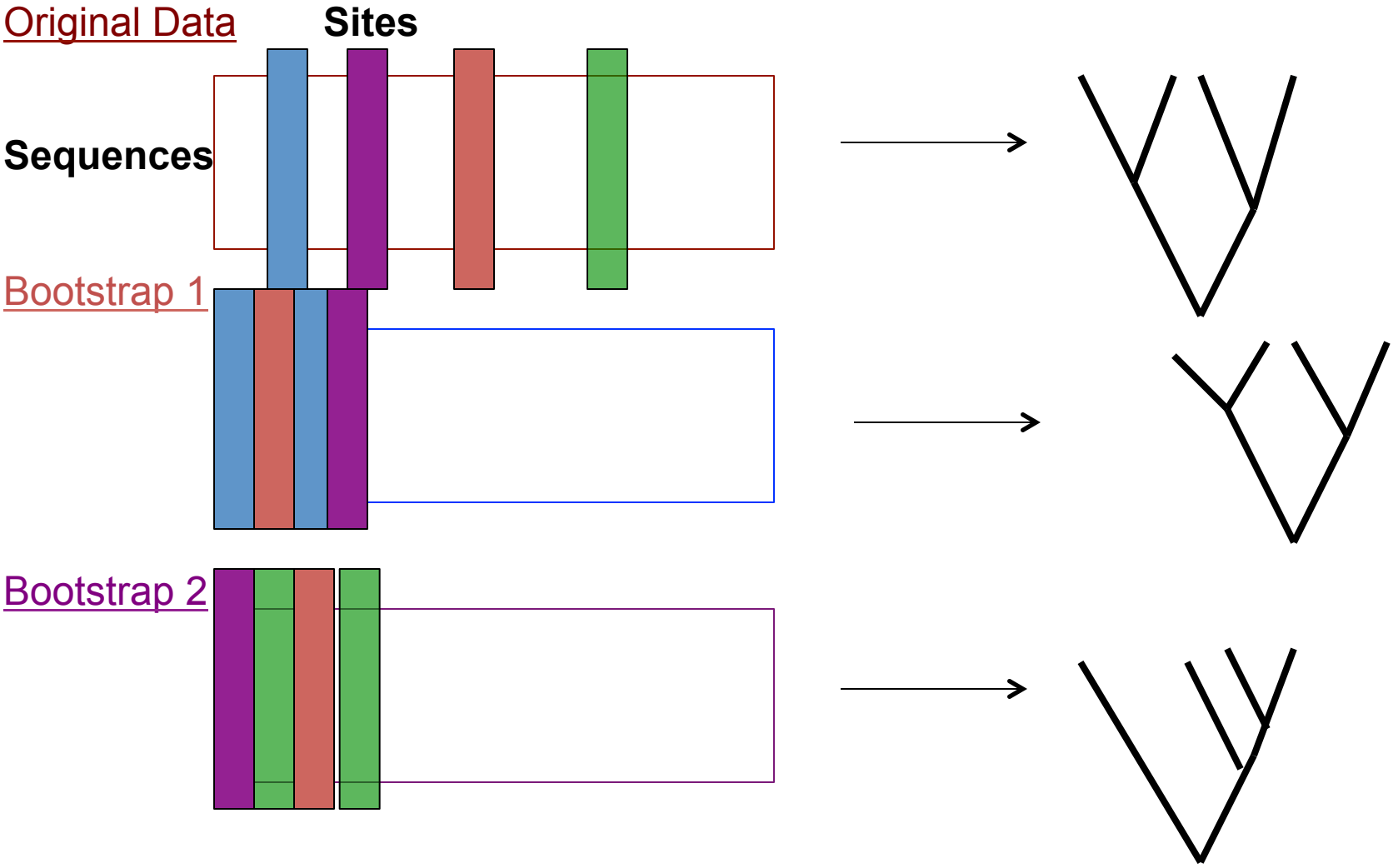
- Bootstrap doesn't rely on parametric assumptions
  - Allows us to infer variability in parameters in models that are too complex for easy calculations of their variances

- Uncertainty in phylogeny construction

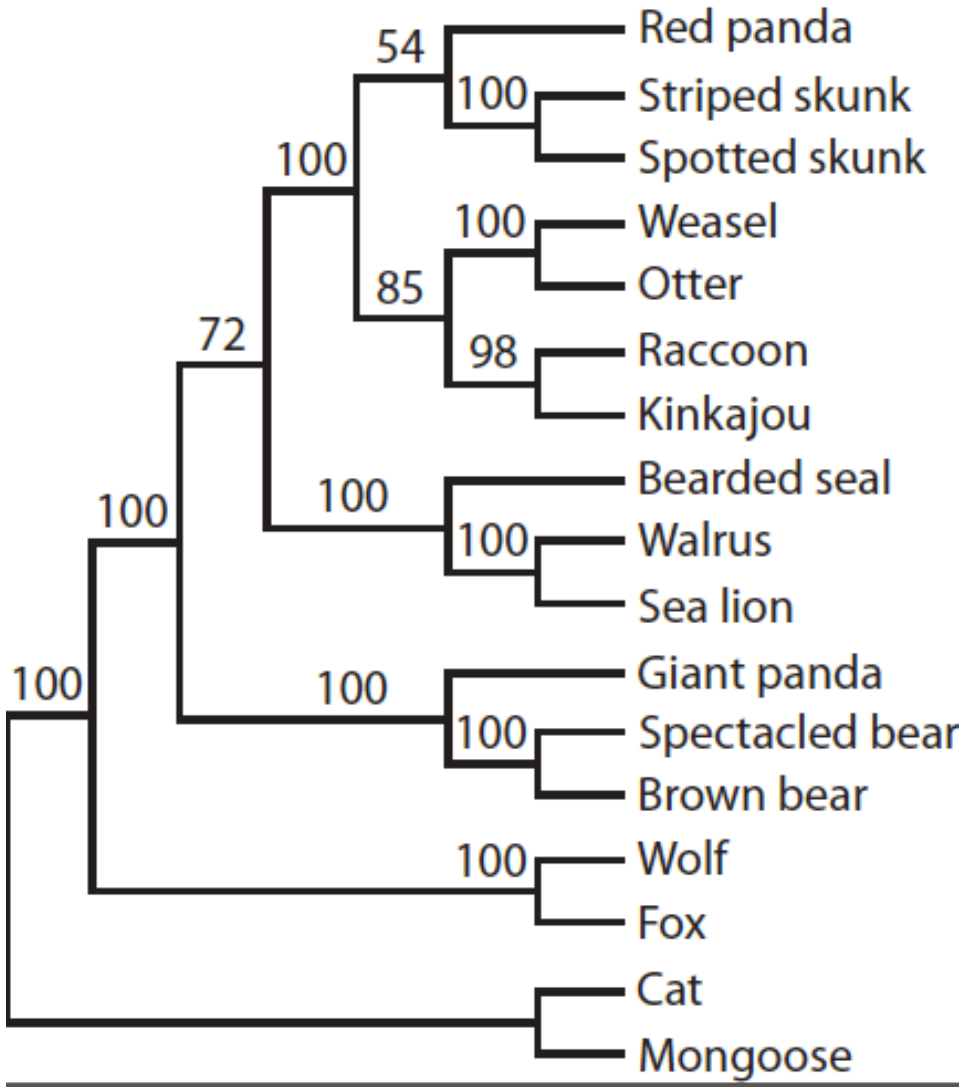
- To use the bootstrap method, data needs to be **independent points**. For phylogenetic construction, we typically have a matrix of dependent variables (species) and possibly independent characters (sites?)

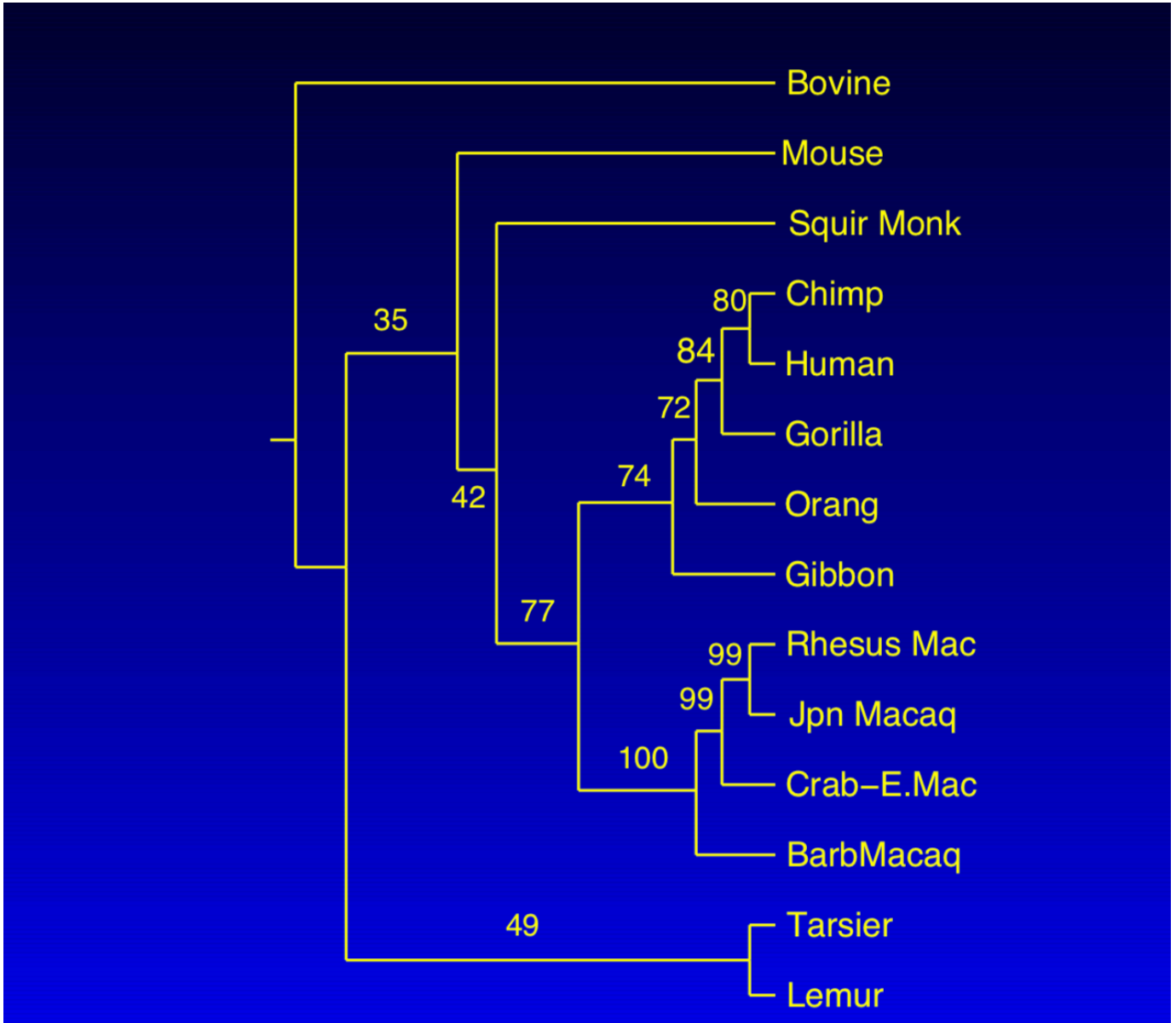
- The trees play the role of the parameter. We ask: how often does a particular branch occur among all the possible tree shapes (topologies, we don't usually care about branch lengths for this): Majority rule consensus

Providing an estimate of uncertainty along branches:



Providing an estimate of uncertainty along branches:





232 nucleotide, 14-species mitochondrial D-loop analyzed by parsimony, 100 bootstrap replicates

Bootstrapping Assumptions:

- Random sample
- Large enough so that frequency distribution of sample is reasonable approximation of frequency distribution of population
  - Too small samples, results in standard errors that are too small and confidence errors are that are too narrow --> overestimate precision

Bootstrapping Examples (PP 9):

In a hypothetical study of a predatory fish, the bootstrap was used to help generate confidence intervals for the mean waiting time between meals. A waiting time measurement for 79 fish in a random sample from the predatory fish population. The mean waiting time was 109 seconds. Ten thousand bootstrap replicates were used to calculate a sampling distribution for the estimate. The mean of the bootstrap replicates was 108.6 seconds, and the standard deviation of the bootstrap replicate estimates was 10.4 seconds. What is the bootstrap standard error of the estimated mean waiting time?

a. 10.4 seconds
b. 12.22 seconds
c. 1.086 seconds

Bootstrapping example:

5th instar *Manduca sexta* caterpillars trained to associate a mild electric shock with a specific odor (ethyl acetate; EA). Then assayed for learning in a Y-choice apparatus as larvae and again as adult moths, after metamorphosis. Can a moth remember what it learned as a caterpillar?

Odds ratio:

Adult response	**learned**	**control**
Choose Clean Air	**32**	**25**
Choose EA	**9**	**21**
	41	**46**

5th instar *Manduca sexta* caterpillars trained to associate a mild electric shock with a specific odor (ethyl acetate; EA). Then assayed for learning in a Y-choice apparatus as larvae and again as adult moths, after metamorphosis. Can a moth remember what it learned as a caterpillar?

Odds ratio:

Adult response	**learned**	**control**
Choose Clean Air	**32**	**25**
Choose EA	**9**	**21**
	41	**46**

Learned:

$p_1 = 32/41 = 0.78$

$O_1 = 0.78/0.22 = 3.56$

control:

$p_2 = 25/46 = 0.54$

$O_2 = 0.54/0.46 = 1.19$

$O_1/O_2 = 3.56/1.19 = 2.99$

The odds of choosing the clean air in a trail are about three times greater in the treatment group (learned) than in the control group.

5th instar *Manduca sexta* caterpillars trained to associate a mild electric shock with a specific odor (ethyl acetate; EA). Then assayed for learning in a Y-choice apparatus as larvae and again as adult moths, after metamorphosis. Can a moth remember what it learned as a caterpillar?

We want a standard error for this estimate and a 95% confidence interval for the true population OR

let 1 represent choice air and 0 represent choose EA

Learn group : 32 X 1 and 9X0

Control group : 25X1 and 21X0

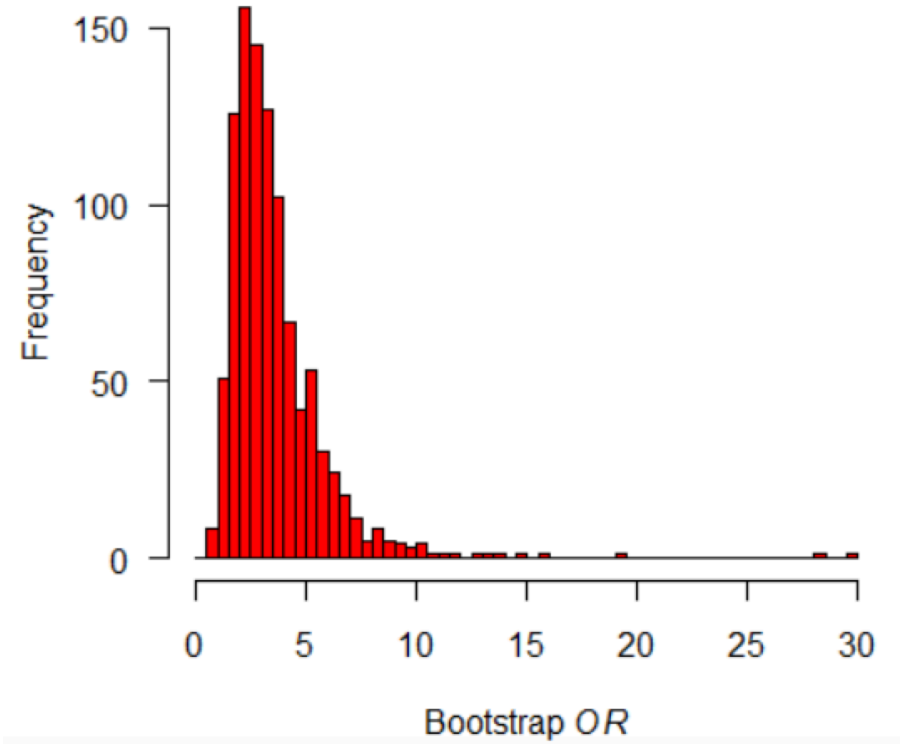
Process:

1. Take a random sample with replacement from each group
2. Calculate OR
3. Repeat 10,000 times

5th instar *Manduca sexta* caterpillars trained to associate a mild electric shock with a specific odor (ethyl acetate; EA). Then assayed for learning in a Y-choice apparatus as larvae and again as adult moths, after metamorphosis. Can a moth remember what it learned as a caterpillar?

We want a standard error for this estimate and a 95% confidence interval for the true population OR

Bootstrap SE = 2.26
(using R) we can get 95% CI
(1.21, 8.67)

If we had calculated the 95% CI for OR as in chapter 9, we would have gotten:

(1.17, 7.65)