1. Goal: to make inferences about the parameter from the statistic

2. Randomness: a model of how the data were produced

3. The i.i.d. (or random sample) assumption

4. Big idea: Every statistic has a sampling distribution

5. Example:

   ```
   > inc = families$INCOME
   > hist(inc,main='Population Distribution',xlab='Income')
   > r5= replicate( 10000, mean( sample(inc,5,replace=F) ) )
   > hist(r5,main='10,000 samples of size 5',xlab='Sample mean of Income')
   ```

6. Notation:

   - \( n \) sample size
   - \( x_1, \ldots, x_n \) result of sample
   - \( \bar{x} \) sample mean
   - \( \bar{X} \) sample mean as a random variable
   - \( \mu, \sigma \) values of population parameters
   - \( \mu_{\bar{X}}, \sigma_{\bar{X}} \) values of parameters of distribution of \( \bar{X} \)

7. The distribution of \( \bar{X} \):

   (a) \( \mu_{\bar{X}} = \mu \).
   (b) \( \sigma^2_{\bar{X}} = \sigma^2/n \)
   (c) If \( n \) is large, the distribution of \( \bar{X} \) is approximately normal. (Central Limit Theorem.)
   (d) If the \( x_i \) have a normal distribution, then the distribution of \( \bar{X} \) is exactly normal.

---

**Homework**

2. Do problems 5.6.50,52
Population Distribution

10,000 samples of size 5

10,000 samples of size 10

10,000 samples of size 30

10,000 samples of size 100