STAT 100 Lecture 13: Sampling Distributions

Nate Strawn

http://www.math.umd.edu/∼nstrawn/

Last Time...

1. Calculations with Normal Distributions
2. The Normal Approximation to the Binomial Distribution
   - The (n, p) Binomial is approximated by $N(np, \sqrt{np(1-p)})$.
   - $P[a \leq X_{\text{binomial}} \leq b] \approx P[a - 0.5 \leq X_{\text{normal}} \leq b + 0.5]$.

Today's Agenda

1. Sampling Distribution of a Statistic (Section 7.1, 7.2)

2. The Central Limit Theorem (Section 7.3)

3. Examples

Parameters versus Statistics

Parameters are intrinsic features of a population, whereas Statistics are quantities inferred from samples from the population. For example, μ and σ are Parameters of $N(\mu, \sigma)$. If $x_1, x_2, \ldots, x_N$ are all generated from $N(\mu, \sigma)$ then

$$\overline{x} = \frac{x_1 + x_2 + \cdots + x_N}{N}$$

and

$$s = \sqrt{\frac{(x_1 - \overline{x})^2 + (x_2 - \overline{x})^2 + \cdots + (x_N - \overline{x})^2}{N - 1}}$$

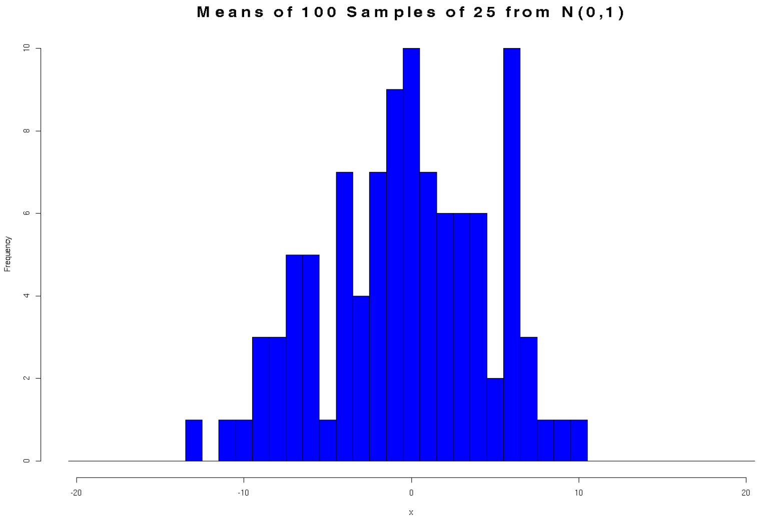
are Statistics.

Parameter or Statistic?

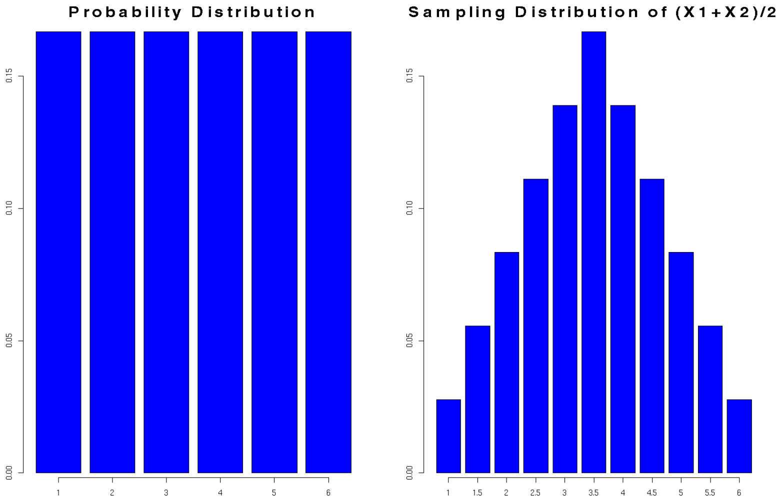
Example

1. *Median of a population.*
2. *Mean of an empirical distribution.*
3. *Sample variation.*

Variation of a Statistic



Sampling Distribution



Conditions for Randomness

Observations $X_1, X_2, \ldots, X_N$ are a Random Sample from a population if they are independent and they always have the same distribution.

Now, if we sample without replacement, then the population changes for each X_i and they are also dependent. However, the sample can be considered Random if the population is much larger than N.

Example: Sampling Distribution

Example

Consider the set $S = \{1, 2, 3, 4\}$. Two samples will be drawn from S without replacement.

1. *List all possible samples and evaluate $\overline{x}$ for each sample.*
2. *Determine the sampling distribution of $\overline{X}$.*

The Sampling Distribution of $\overline{X}$ for a Normal Population

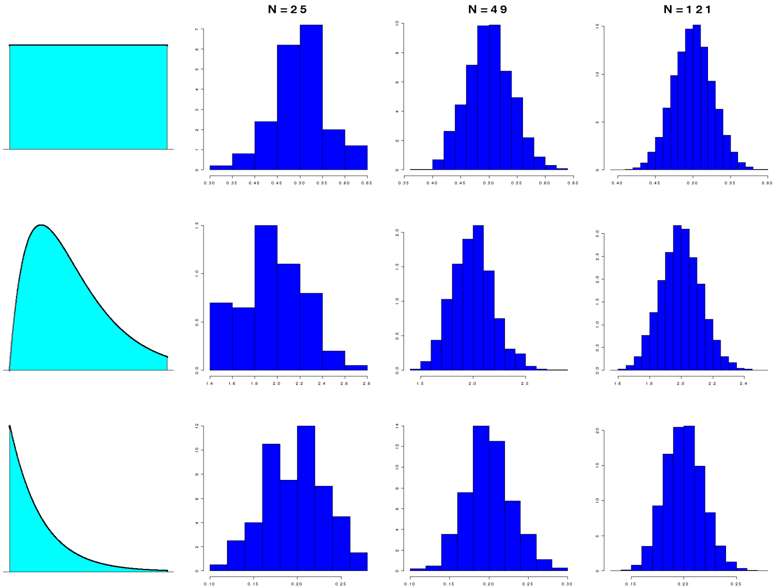
Fact

If the population is distributed as $N(\mu, \sigma)$, *then* $\overline{X}$ *is distributed as* $N\left(\mu, \frac{\sigma}{\sqrt{N}}\right)$.

Example

The number of Skittles in a bag has a binomial distribution with $n = 1000$ *and* $p = 0.9$. *What is the probability that the average number of Skittles in 9 bags is greater than* 800*?*

Sample Distributions for $\overline{X}$ as N gets Large



Parameters of the Sampling Distribution of $\overline{X}$

Fact

For a population with mean μ *and standard deviation* σ*, the sampling distribution of* $\overline{X}$ *based on samples of size* N *satisfies*

$$E(\overline{X}) = \mu,$$

$$Var(\overline{X}) = \frac{\sigma^2}{N},$$

and

$$sd(\overline{X}) = \frac{\sigma}{\sqrt{N}}.$$

The Central Limit Theorem

For ANY population, the distribution of $\overline{X}$ is approximately normal when N is large. Moreover, the distribution of

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{N}}$$

is well approximated by $N(0,1)$ if μ and σ are the population mean and population standard deviation, respectively.

Example: Application of the Central Limit Theorem

Example

Consider a population with $\mu = 5$ and $\sigma = 2$.

1. *If a random sample of size $N = 49$ is selected, what is the probability that the sample mean is between 4.5 and 5.5?*
2. *If a random sample of size $N = 25$ is selected, what is the probability that the sample mean is less than 7?*

For Next Time

- MINITAB Project 03! Due Monday!
- Group Problems:

Group	1	2	3	4	5
Problem	7.1	7.3	7.5	7.7	7.9

Group	6	7	8	9	10
Problem	7.11	7.13	7.15	7.21	7.23