Census vs. **Sampling**

Parameter: a numerical measure (or characteristic) of the population, examples includes μ, σ^2, and π.

Statistic: a numerical measure (or characteristic) of a sample, examples include $\bar{X}$, s^2, and p.

Sampling error: the absolute difference between the parameter θ and its statistic $\hat{\theta}$, that is, $|\,\theta - \hat{\theta}|$.

Sampling distribution: the probability distribution of a statistic.

Standard error: the standard deviation of a statistic.

Sampling distribution of the sample mean $\bar{X}$

Suppose a random variable X of interest has mean μ_X and standard deviation σ_X and a random sample is taken from a(n infinite) population. Then the mean of $\bar{X}$ is μ_X and the variance of $\bar{X}$ is σ_X^2/n, that is,

$$E(\bar{X}) = \mu_X \text{ and } Var(\bar{X}) = \frac{\sigma_X^2}{n} \text{ (or } St.Dev(\bar{X}) = \frac{\sigma_X}{\sqrt{n}}\text{)}. \qquad (1)$$

Note $E(\bar{X}) = \mu_{\bar{X}}$ and $Var(\bar{X}) = \sigma_{\bar{X}}^2$. If a sample is from a finite population, $E(\bar{X}) = \mu_X$ and $Var(\bar{X}) = \frac{N-n}{N-1}\frac{\sigma_X^2}{n}$, and $\frac{N-n}{N-1}$ is called a finite population correction factor. $\sigma_{\bar{X}}$ is also called the **standard error of the sample mean**.

Furthermore,

- if X is a normal random variable, then $\bar{X}$ is also a normal random variable with mean and variance as in (1);
- if X is **not** a normal random variable, then $\bar{X}$ is approximately a normal random variable provided n is large, say, $n \geq 30$ according to the central limit theorem (see the definition of the CLT on page 257 of the textbook).

Example 1. Rotor bearings are produced with mean weight 1.64 and standard deviation 0.03 gram.

a) Find the mean and the standard deviation of the sample mean from a sample of size 5.
b) Find the sampling distribution of the sample mean from a sample of size 5 if the weight of a rotor bearing has a normal distribution.
c) Under the assumption of b), find the probability that the sample mean is between 1.635 and 1.645.

Example 2. Starting salaries of newly graduated accounting majors have mean \$25,000 and standard deviation \$2,000. If a random sample of 100 recently graduated accounting majors is taken, find the probability that the sample mean will be within \$300 of the population mean.

MSL Homework: 7.1, 7.6
HC Homework: 7.5, 7.8

Sampling distribution of the sample proportion *p*

Suppose a random sample of size n is drawn from a population. Let π denote the proportion of the population possessing some characteristic and p the sample proportion. Then p is approximately a normal random variable with mean π and variance $\pi(1-\pi)/n$, that is, $\mu_p = E(p) = \pi$ and $\sigma_p^2 = Var(p) = \pi(1-\pi)/n$, provided $n \geq 25$, $n\pi \geq 5$, and $n(1-\pi) \geq 5$ according to the central limit theorem. If a sample is from a finite population, $Var(p) = \frac{N-n}{N-1}\frac{\pi(1-\pi)}{n}$. Note that σ_p is the standard error of the sample proportion p.

Example 3. In an election 55% of the registered voters favor a certain candidate. If we take a random sample of 100 voters, what is the probability that based on the sample information we will predict the wrong winner, that is, the probability that the sample proportion of the voters who favor the candidate is less than 0.5.

MSL Homework: 7.9, 7.13
HC Homework: 7.12, 7.14