Chapter 8. Introduction to Statistics.

8.1

Population : consists of all units of interest.

Sample : consists of observed units collected from the population.
It is used to make statements about the population.

Statistics : Any function of a sample.

Simple random sampling : is a sampling design where units are collected from the entire population independently of each other, all being equally likely to be sampled.

8.2. Simple descriptive statistics.

Mean : $\bar{X} = \dfrac{X_1 + X_2 + \cdots + X_n}{n} = u = E(X)$.

Unbiased : A statistic is unbiased if its expected value is the parameter which it estimates.

Notation : u = population mean.

$\bar{X}$ = sample mean, estimator of u.

δ = population standard deviation.

s = sample standard deviation, estimator of δ.

δ^2 = population variance.

s^2 = sample variance, estimator of δ.

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2 = \frac{\sum_{i=1}^{n} (X_i)^2 - n \cdot \bar{X}^2}{n-1}$$

$s = \sqrt{s^2}$.

Standard error of any statistic is its standard deviation, and that we can often only estimate.

$$Var(\bar{X}) = Var\left(\frac{X_1 + X_2 + \cdots + X_n}{n}\right)$$

$$= \frac{1}{n^2} \cdot (Var(X_1) + Var(X_2) + \cdots + Var(X_n))$$

$$= \frac{1}{n^2} \cdot (\underbrace{\delta^2 + \cdots + \delta^2}_{n})$$

$$= \frac{\delta^2}{n}$$

Therefore, $\delta(\bar{X}) = \sqrt{Var(\bar{X})} = \frac{\delta}{\sqrt{n}}$ → if X is Bernoulli, then $\delta = \sqrt{p(1-p)}$

We can only estimate δ with s.