Sampling Distributions (Ch. 6)

- The aim of the statistical analysis is inference
- Sample Statistics (mean, standard deviation, Proportion) are the main tools for making inference and decisions
- Probability distributions are models for Populations.

Two important measures

- Parameter: Numerical measure of a population. Its value is almost unknown.
- Statistic: Numerical measure of a sample. Its value can be calculated from observations.

	Parameter	Statistic
Mean	μ	$\bar{X}$
Variance	σ^2	S^2
Standard deviation	σ	S
Proportion	p	$\hat{p}$

Notes

$\boxed{1}$ The sample Statistics are random variables since they are computed from the sample.

Population: 5, 10, 20

Samples (size = n = 2): 5, 10 $\Rightarrow \bar{X} = \frac{5+10}{2} = 7.5$

5, 20 $\Rightarrow \bar{X} = \frac{5+20}{2} = 12.5$

10, 20 $\Rightarrow \bar{X} = \frac{10+20}{2} = 15$

[2] Since these sample statistics are random variables, there is a probability or <u>sampling</u> <u>distribution</u> of the statistic

[3] Let μ = mean of a population

$\bar{X}$ = Sample mean of a sample of size n

$$= \frac{\sum_{i=1}^{n} X_i}{n}$$

$\bar{X}$ is a <u>point</u> <u>estimator</u> of μ

<u>Sampling Distribution of $\bar{X}$</u>

- $\mu_{\bar{X}} = E(\bar{X}) = \mu$ [mean or average of $\bar{X} = \mu$]
- $\sigma_{\bar{X}} = \text{Std.}(\bar{X}) = \frac{\sigma}{\sqrt{n}}$ [Standard deviation of $\bar{X}$ is $\frac{\sigma}{\sqrt{n}}$]

<u>Result 1</u>: If a random sample of n observations is drawn from a normally distributed population, then the sampling distribution of $\bar{X}$ will be normally distributed.

<u>Result 2</u> (Central Limit Theorem): Based on a large sample of size n, the sampling distribution is approximately normal with

$$\mu_{\bar{X}} = \mu \text{ and } \sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$$

Example: The salaries of the persons working in computer company is normally distributed with mean 2000 KD and Std. 400 KD.

(a) What is the probability that a randomly selected person has a salary more than 2200 KD?

$$P(X > 2200) = P\left(Z > \frac{2200-2000}{400}\right)$$
$$= P(Z > 0.5)$$
$$= 1 - 0.6915$$
$$= \boxed{0.3085}$$

(b) What is the probability that a randomly selected person has a salary between 1900 KD and 2200 KD?

$$P(1900 < X < 2200) = P\left(\frac{1900-2000}{400} < Z < \frac{2200-2000}{400}\right)$$
$$= P(-0.25 < Z < 0.5)$$
$$= 0.6915 - 0.4013$$
$$= \boxed{0.2902}$$

(c) What is the probability that a randomly selected sample of 16 persons will get an average more than 2200 KD?

$$P(\bar{X} > 2200) = P\left(Z > \frac{2200-2000}{400/\sqrt{16}}\right) = P(Z > 2)$$
$$= 1 - 0.9772 = \boxed{0.0228}$$

(d) If a sample of size 16 is taken, what is the Prob. that its average is between 1800 and 2200?

$$P(1800 < \bar{X} < 2200) = P\left(\frac{1800-2000}{400/\sqrt{16}} < Z < \frac{2200-2000}{400/\sqrt{16}}\right)$$
$$= P(-2 < Z < 2) = 0.9772 - 0.0228 = \boxed{.9544}$$