Sampling distributions

- The probability distribution of a statistic is called a sampling distribution.
- $\overline{X}$: the sampling distribution of the mean

$$\overline{X} = \frac{X_1 + X_2 + \ldots + X_n}{n}$$

$$\mu_{\overline{X}} = \frac{\mu + \mu + \ldots + \mu}{n} = \mu$$

$$\sigma^2_{\overline{X}} = \frac{\sigma^2 + \sigma^2 + \ldots + \sigma^2}{n} = \frac{\sigma^2}{n}$$

The Central Limit Theorem

- When *n* is sufficiently large (i.e. greater than 15), the sample mean follows approximately a normal distribution:

$$\overline{X} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$

Example 1

- An electronics company manufactures resistors that have a mean resistance of 100Ω and a standard deviation of 10Ω. The distribution of resistance is normal. Find the probability that a random sample of n=25 resistors will have an average resistance less than 25Ω .

Example 2

- Suppose that a random variable X has a continuous uniform distribution

$$f(x)=\begin{cases} 1/2, & 4 \le x \le 6 \\ 0, & otherwise \end{cases}$$

Find the distribution of the sample mean of a random sample of size n=40.

Sampling distribution

- If we have two independent populations with means μ_1 and μ_2 and variance σ_1^2 and σ_2^2, and if $\overline{X}_1$ and $\overline{X}_2$ are the sample means of two independent random samples of sized n_1 and n_2 from these population, then the sampling distribution of

$$Z = \frac{\overline{X}_1 - \overline{X}_2 - (\mu_1 - \mu_2)}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}}$$

is approximately standard normal, if the conditions of the central limit theorem apply. If the two populations are normal, then the sampling distribution of Z is exactly standard normal.

Example 3

- The effective life of a component used in a jet-turbine aircraft engine is a random variable with mean 5000 hr and standard deviation 40hr. The distribution of effective life is fairly close to a normal distribution. The engine manufacturer increases the mean life to 5050hr and decreases the standard deviation to 30hr. Suppose that a random sample of n1=16 components is selected from the old process and a random sample of n2=25 components is selected from the improved process. What is the probability that the difference in the two sample means is at least 25 hr?