Course 11

Population Variance总体方差:

$$\sigma^2 = \frac{\sum(x_i - \mu)^2}{N}$$

Sampling Distribution样本分布: The probability distribution of **statistic统计量**

Random Sample随机样本
The rv's $X_1, X_2, \ldots, X_n$ are said to form a (simple) random sample of size n if

1. The X_i's are independent rv's.
2. Every X_i has the **same probability distribution**.

- Sampling with replacement or from an infinite population is **random sampling**.(放回抽样是随机抽样)
- Sampling without replacement from a finite population is generally considered **not random sampling**. However, if the sample size n is much smaller than the population size $N(n/N \leq 0.05)$, it is approximately random sampling.(不放回抽样不是随机抽样，但当抽取的样本占总体的比例很小时可以近似成随机抽样)
- That is, the $\bar{X}$ Sampling distribution is centered at the population mean μ.
- And the S^2 Sampling distribution is centered at the population variance σ^2

The Central Limit Theorem(CLT)中心极限定理
If n is sufficiently large, has approximately a normal distribution with

$\bar{X}$ is sample average

- $E(\bar{X}) = \mu_{\bar{X}} = \mu$
- $V(\bar{X}) = \sigma^2_{\bar{X}} = \sigma^2/n$
- $\sigma_{\bar{X}} = \sigma/\sqrt{n}$

T_0 is $T_0 = X_1 + X_2 + \cdots + X_n$

- $E(T_0) = \mu_{T_0} = n\mu$
- $V(T_0) = \sigma^2_{T_0} = n\sigma^2$

The larger the value of n, the better the approximation, and usually, if $n > 30$, the **Central Limit Theorem(CLT)** can be used.

The Distribution of a Linear Combination分布的线性组合

Given a collection of n random variables $X_1, \ldots, X_n$ and n numerical constants $a_1, \ldots, a_n$, they satisfy $Y = a_1X_1 + \cdots + a_nX_n = \sum_{i=1}^{n} a_iX_i$, the rv is called a linear combination of the X_i's.

Let $X_1, X_2, \ldots, X_n$ have mean values $\mu_1, \ldots, \mu_n$ respectively, and variances of $\sigma_1^2, \ldots, \sigma_n^2$, respectively.

- $E(\sum_{i=1}^{n} a_iX_i) = \sum_{i=1}^{n} a_iE(X_i) = \sum_{i=1}^{n} a_i\mu_i$
- If $X_1, X_2, \ldots, X_n$ are independent, $V(\sum_{i=1}^{n} a_iX_i) = \sum_{i=1}^{n} a_i^2V(X_i) = \sum_{i=1}^{n} a_i^2\sigma_i^2$
- For any $X_1, X_2, \ldots, X_n$, $V(\sum_{i=1}^{n} a_iX_i) = \sum_{i=1}^{n}\sum_{i=1}^{n} a_ia_j\mathrm{Cov}(X_i, X_j)$

Corollary推论:

- $E(X_1 - X_2) = E(X_1) - E(X_2)$
- If X_1 and X_2 are independent, $V(X_1 - X_2) = V(X_1) + V(X_2)$

Homework

Section 5.3 38, 41

Section 5.4 46, 51, 55

Section 5.5 58, 70, 73