SAMPLING DISTRIBUTION

SAMPLING DISTRIBUTION

- Let $X_1, X_2, \ldots, X_n$ be a r.s. of size n from a population and let $T(x_1, x_2, \ldots, x_n)$ be a real (or vector)-valued whose domain includes the sample space of $(X_1, X_2, \ldots, X_n)$. Then, the r.v. or a random vector $Y = T(X_1, X_2, \ldots, X_n)$ is called a ***statistic***. The probability distribution of a statistic Y is called the ***sampling distribution*** of Y.

SAMPLING DISTRIBUTION

- The ***sample mean*** is the arithmetic average of the values in a r.s.

$$\bar{X} = \frac{X_1 + X_2 + \cdots + X_n}{n} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

- The ***sample variance*** is the statistic defined by

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2$$

- The ***sample standard deviation*** is the statistic defined by S.

SAMPLING FROM THE NORMAL DISTRIBUTION

Properties of the Sample Mean and Sample Variance

- Let $X_1, X_2, \ldots, X_n$ be a r.s. of size n from a $N(\mu,\sigma^2)$ distribution. Then,

$a)\ \bar{X}$ and S^2 are independent rvs.

$$b)\ \bar{X} \sim N\left(\mu, \sigma^2 / n\right)$$

$$c)\ \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$$

SAMPLING FROM THE NORMAL DISTRIBUTION

- Let $X_1, X_2, \ldots, X_n$ be a r.s. of size n from a $N(\mu,\sigma^2)$ distribution. Then,

$$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

- Most of the time σ is unknown.

$$\frac{\bar{X}-\mu}{S/\sqrt{n}}.$$

SAMPLING FROM THE NORMAL DISTRIBUTION

$$\frac{\bar{X}-\mu}{S/\sqrt{n}}=\frac{(\bar{X}-\mu)/\left(\sigma/\sqrt{n}\right)}{\sqrt{S^2/\sigma^2}}=\frac{N(0,1)}{\sqrt{(n-1)\chi^2_{n-1}/(n-1)}}=t_{n-1}$$

In statistical inference, Student’s t distribution is very important.

SAMPLING FROM THE NORMAL DISTRIBUTION

- Let $X_1, X_2, \ldots, X_n$ be a r.s. of size n from a $N(\mu_X, \sigma_X^2)$ distribution and let $Y_1, Y_2, \ldots, Y_m$ be a r.s. of size m from an independent $N(\mu_Y, \sigma_Y^2)$.
- If we are interested in comparing the variability of the populations, one quantity of interest would be the ratio

$$\sigma_X^2 / \sigma_Y^2 \rightarrow S_X^2 / S_Y^2$$

SAMPLING FROM THE NORMAL DISTRIBUTION

- The F distribution allows us to compare these quantities by giving the distribution of

$$\frac{S_X^2 / S_Y^2}{\sigma_X^2 / \sigma_Y^2} = \frac{S_X^2 / \sigma_X^2}{S_Y^2 / \sigma_Y^2} \sim F_{n-1,m-1}$$

- If $X \sim F_{p,q}$, then $1/X \sim F_{q,p}$.
- If $X \sim t_q$, then $X^2 \sim F_{1,q}$.