Theorem

If $X_1, X_2, X_3, \ldots, X_n$ are $n$ independent random variables with means $\mu_1, \mu_2, \mu_3, \ldots, \mu_n$, variances $\sigma^2_1, \sigma^2_2, \sigma^2_3, \ldots, \sigma^2_n$, and moment generating functions $M_{X_i}(t)$ for $i = 1, 2, \ldots, n$, respectively, then if

$$Y = \sum_{i=1}^{n} a_i X_i$$

$$\mu_Y = E[Y] = E\left[\sum_{i=1}^{n} a_i X_i\right] = \sum_{i=1}^{n} a_i E[X_i] = \sum_{i=1}^{n} a_i \mu_i$$

$$\sigma^2_Y = Var[Y] = Var\left[\sum_{i=1}^{n} a_i X_i\right] = \sum_{i=1}^{n} a_i^2 Var[X_i] = \sum_{i=1}^{n} a_i^2 \sigma^2_i$$

$$M_Y(t) = \prod_{i=1}^{n} M_{X_i}(a_i t)$$

Corollary

If $X_1, X_2, X_3, \ldots, X_n$ is a random sample from a distribution with mean $\mu$, variance $\sigma^2$, and moment generating functions $M(t)$ then if

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

$$\mu_{\bar{X}} = E[\bar{X}] = E\left[\frac{1}{n} \sum_{i=1}^{n} X_i\right] = \frac{1}{n} \sum_{i=1}^{n} E[X_i] = \frac{1}{n} \sum_{i=1}^{n} \mu = \frac{n\mu}{n} = \mu$$

$$\sigma^2_{\bar{X}} = Var[\bar{X}] = Var\left[\frac{1}{n} \sum_{i=1}^{n} X_i\right] = \frac{1}{n} \sum_{i=1}^{n} Var[X_i] = \frac{1}{n} \sum_{i=1}^{n} \sigma^2 = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}$$

$$M_{\bar{X}}(t) = \prod_{i=1}^{n} M\left(\frac{t}{n}\right) = \left[M\left(\frac{t}{n}\right)\right]^n$$

Theorem

If $X_1, X_2, X_3, \ldots, X_n$ is a random sample from a distribution with mean $\mu$, variance $\sigma^2$, from a normal distribution, $N(\mu, \sigma^2)$, then the distribution of the sample mean, $\bar{X}$, has a normal distribution $N(\mu, \frac{\sigma^2}{n})$.

Proof:

$$M_{\bar{X}}(t) = \prod_{i=1}^{n} M\left(\frac{t}{n}\right) = \left[M\left(\frac{t}{n}\right)\right]^n = \prod_{i=1}^{n} e^{\frac{t^2}{2} \left(\frac{1}{n}\right)^2 \frac{\sigma^2}{n}} = e^{\frac{t^2}{2} \left(\frac{1}{n}\right)^2 \frac{\sigma^2}{n}}$$

It is the moment generating function of a random variable having $N(\mu, \frac{\sigma^2}{n})$.

Theorem: $Y = \sum_{i=1}^{n} c_i X_i \sim N(\sum_{i=1}^{n} c_i \mu_i, \sum_{i=1}^{n} c_i^2 \sigma^2_i)$
STAT 5844, Review of Sampling Distribution

**Theorem**

Let the distributions of independent random variables $X_1, X_2, X_3, \ldots, X_n$, be $\chi^2(r_1),\chi^2(r_2),\ldots,\chi^2(r_n)$, respectively, then $Y = \sum_{i=1}^{n} X_i$ has a Chi-square distribution, $\chi^2(r_1 + r_2 + \ldots + r_n)$.

**Theorem**

Let the mutually independent random variables, $Z_1, Z_2, Z_3, \ldots, Z_n$, be normally distributed with mean 0 and variance 1, then $W = Z_1^2 + Z_2^2 + \ldots + Z_n^2$, has a Chi-square distribution, $\chi^2(n)$.

**Corollary**

If $X_1, X_2, X_3, \ldots, X_n$, are mutually independent normally distributed random variables with means with means $\mu_1, \mu_2, \mu_3, \ldots, \mu_n$, variances $\sigma_1^2, \sigma_2^2, \sigma_3^2, \ldots, \sigma_n^2$, respectively, then the distribution of

$$W = \sum_{i=1}^{n} \frac{(X_i - \mu_i)^2}{\sigma_i^2}$$

is $\chi^2(n)$.

**Theorem**

If $X_1, X_2, X_3, \ldots, X_n$, is a random sample from the normal distribution $N(\mu, \sigma^2)$, and

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

then

1) $\bar{X}$ and $S^2$ are independent,

2) $\frac{nS^2}{\sigma^2}$ is $\chi^2(n - 1)$.

**Theorem (Central Limit Theorem)**

If $\bar{X}$ is the mean of a random sample of size $n$, $X_1, X_2, X_3, \ldots, X_n$, from a distribution with finite mean, $\mu$, and variance, $\sigma^2$, then

$$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} = \frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n} \cdot \sigma}$$

has approximately a **standard normal** distribution as $n \to \infty$.

This implies that sample mean is approximately normally distributed for large sample.

If $X_i$’s are symmetric unimodal and continuous random variables then the approximation is good for $n \geq 4$. 
Normal approximation to Binomial Distribution

Let $X_1, X_2, X_3, \ldots, X_n$, be a random sample from a Bernoulli distribution with a mean, $\mu = p$, and variance, $\sigma^2 = p(1-p)$, $0 < p < 1$. Then $Y = \sum_{i=1}^{n} X_i$ is Binomial, $b(n, p)$, and by the Central Limit Theorem, the random variable $W = \frac{Y - np}{\sqrt{np(1-p)}} = \frac{X - p}{\sqrt{p(1-p)/n}}$ is $N(0, 1)$.

This implies that Binomial distribution can be approximated by normal distribution for large sample. Usually, the approximation is good if $np > 10$ and $n(1-p) > 10$. Some people use $np > 5$, $n(1-p) > 5$.

Theorem

If random variable $Z$ has a standard normal distribution and $U$ has a Chi-square distribution with $r$ degrees of freedom, and $Z$ and $U$ are independent, then

$$T = \frac{Z}{\sqrt{U/r}}$$

follows a $t$-distribution with $r$ degrees of freedom.

Its p.d.f. is

$$g(t) = \frac{\Gamma\left(\frac{r+1}{2}\right)}{\sqrt{\pi} \cdot r \cdot \Gamma\left(\frac{r}{2}\right) \left(1 + \frac{t^2}{r}\right)^{\frac{r+1}{2}}}$$

, $-\infty < t < \infty$.

When $r = 1$, $t$-distribution is the Cauchy distribution.

$E[T] = 0$, if $r \geq 2$

$Var[T] = r/(r - 2)$, if $r \geq 3$, $(\sigma^2$ does not exist when $r = 1, 2)$

Corollary

If $X_1, X_2, X_3, \ldots, X_n$, is a random sample from the normal distribution $N(\mu, \sigma^2)$, then

$$T = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} = \frac{\overline{X} - \mu}{\sqrt{nS/(n-1)}}$$

has a $t$-distribution with degrees of freedom $n - 1$. 
