Addition Property of Some Known Probability distribution.

Result: Let $X_1, X_2, \cdots, X_K$ be independent and assume that $X_i \sim Bin(n_i, p)$ $i = 1, 2, \cdots, K$.

Consider the sum $S_K = \sum_{i=1}^{K} X_i = X_1 + X_2 + \cdots + X_K$.

Then probability distribution of S_K.

Proof: We use the moment generating function (MGF) technique to obtain the required result.

So let us compute MGF of S_K.

$$M_{S_K}(t) = E\left(e^{S_K t}\right) = E\left(e^{(X_1 + X_2 + \cdots + X_K)t}\right)$$

$$= E\left(e^{X_1 t} \cdot e^{X_2 t} \cdots e^{X_K t}\right)$$

$$= E\left(e^{X_1 t}\right) E\left(e^{X_2 t}\right) \cdots E\left(e^{X_K t}\right) \quad \left\{ \because X_i ; i = 1, 2, \cdots, K \text{ are independent variables} \right.$$

$$= M_{X_1}(t) \cdot M_{X_2}(t) \cdots M_{X_K}(t)$$

$$= (q + pe^t)^{n_1} (q + pe^t)^{n_2} \cdots (q + pe^t)^{n_K}$$

$$= (q + pe^t)^{n_1 + n_2 + \cdots + n_K} = (q + pe^t)^{\sum_{i=1}^{K} n_i} \quad \text{--- (1)}$$

Equation 1 is the MGF of a $Bin(\sum_{i=1}^{k} n_i, p)$ distribution. So what is the result we got?

If $X_1, X_2, \dots, X_k$ be indep. and $X_i \sim Bin(n_i, p), i=1,2,\dots,k,$ then $S_k = \sum_{i=1}^{k} X_i \sim Bin(\sum_{i=1}^{k} n_i, p)$ distn.

$$\boxed{S_k \sim Bin\left(\sum_{i=1}^{k} n_i, p\right)}$$

Remark: In the previous result suppose each of $n_i = 1$, $i=1,2,\dots,k$. That is we have situation like $X_1, X_2, \dots, X_k \overset{iid}{\sim} Bin(1,p) \cong Ber(p)$ then the sum $S_k = \sum_{i=1}^{k} X_i$ has $Bin(k, p)$ distribution.

$\Rightarrow$ Here $M_{S_k}(t) = \cancel{E e^{tX_1}}\ E(e^{S_k t})$

$$= E(e^{X_1 t + X_2 t + \dots + X_k t})$$

$$= E(e^{X_1 t} \cdot e^{X_2 t} \cdots e^{X_k t})$$

$$= E(e^{X_1 t})\, E(e^{X_2 t}) \cdots E(e^{X_k t})$$

$$= M_{X_1}(t)\, M_{X_2}(t) \cdots M_{X_k}(t) = (M_{X_1}(t))^k$$

$= (q + pe^t)^k$. this is the MGF of $Bin(k, p)$ distn.

Addition Property of Poisson Distribution:

Let $X_1, X_2, \ldots, X_K$ be independent Poisson r.v.s with $X_i \sim P(\lambda_i)$, $i=1,2,\ldots,K$. Then prob distribution of $S_K = \sum_{i=1}^{K} X_i$ is Poisson $P(\sum_{i=1}^{K} \lambda_i)$.

Note: If $X \sim P(\lambda)$ then $M_X(t) = e^{\lambda(e^t - 1)}$

Relation between Geometric & Negative Binomial Distn

Let $X_1, X_2, \ldots, X_K$ be iid $Geo(p)$ random variables. then show that $S_K = \sum_{i=1}^{K} X_i$ has negative Binomial $NB(K, p)$ distribution.

Relation between one parameter exponential and Gamma Distn:

Let $X_1, X_2, \ldots X_K$ iid $\exp(\lambda)$ then prob distn of $S_K = \sum_{i=1}^{K} X_i$ is given by gamma $G(K, \lambda)$.

Proof: If $X \sim \exp(\lambda)$ $f_X(x) = \lambda e^{-\lambda x}$, $x > 0$, $\lambda > 0$.

$$M_X(t) = \left(\frac{\lambda}{\lambda - t}\right) \quad t < \lambda.$$

Let us compute MGF of S_k:

$$M_{S_k}(t) = E(e^{S_k t}) = E[e^{(X_1+X_2+\cdots+X_k)t}]$$

$$= [M_{X_1}(t)]^k = \left[\frac{\lambda}{\lambda-t}\right]^k$$

this is MGF of a gamma $G(k,\lambda)$ variable. Hence result is proved

<u>Remark</u>: Let $X_1, X_2, \ldots, X_k$ be independent such that $X_i \sim G(r_i, \lambda)$ $i=1,2,\ldots,k$. Then find the pdf of $S_k = \sum_{i=1}^{k} X_i$.

<u>Result</u>:

* Let $X_1, X_2, \ldots, X_k$ be independent random variables such that $X_i \sim N(\mu_i, \sigma_i^2)$ $i=1,2,\ldots,k$.

Consider $Y = \sum_{i=1}^{k} (a_i X_i + b_i)$

$$= (a_1X_1+b_1) + (a_2X_2+b_2) + \cdots + (a_kX_k+b_k)$$

where a_i, b_i, $i=1,2,\ldots,k$ are given constants. Find prob distn of Y.

<u>note</u> $\boxed{\text{If } X \sim N(\mu,\sigma^2) \text{ then } M_X(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2}}$

Solution: Let compute MGF of Y as follows:

$$M_Y(t) = E(e^{tY}) = E\left[e^{t\sum_{i=1}^{k}(a_i x_i + b_i)}\right]$$

$$= E\left[e^{t\cdot\sum_{i=1}^{k} b_i} \cdot e^{t\sum_{i=1}^{k} a_i x_i}\right]$$

$$= \left[e^{t\cdot\sum_{i=1}^{k} b_i}\right] \cdot E\left[e^{t\cdot\sum_{i=1}^{k} a_i x_i}\right]$$

$$= \left[e^{t\cdot\sum_{i=1}^{k} b_i}\right] E\left[e^{t(a_1 x_1 + a_2 x_2 + \cdots + a_k x_k)}\right]$$

$$= \left(e^{t\cdot\sum_{i=1}^{k} b_i}\right) E\left[e^{a_1 t x_1} \cdot e^{a_2 t x_2} \cdots e^{a_k t x_k}\right]$$

$$= \left[e^{t\cdot\sum_{i=1}^{k} b_i}\right] E\left(e^{a_1 t x_1}\right) E\left(e^{a_2 t x_2}\right) \cdots E\left(e^{a_k t x_k}\right)$$

($\because X_i, i = 1, 2, \ldots, k$ are independent)

$$= \left(e^{t\cdot\sum_{i=1}^{k} b_i}\right) M_{X_1}(a_1 t)\, M_{X_2}(a_2 t) \cdots M_{X_k}(a_k t)$$

$$= e^{t\cdot\sum_{i=1}^{k} b_i}\, e^{\mu_1 a_1 t + \frac{1}{2} a_1^2 t^2 \sigma_1^2} \cdot e^{\mu_2 a_2 t + \frac{1}{2} a_2^2 t^2 \sigma_2^2} \cdots e^{\mu_k a_k t + \frac{1}{2} a_k^2 t^2 \sigma_k^2}$$

$$= e^{t\cdot\sum_{i=1}^{k} b_i} \cdot e^{t(\mu_1 a_1 + \mu_2 a_2 + \cdots + \mu_k a_k)}\, e^{\frac{1}{2} t^2 (a_1^2 \sigma_1^2 + a_2^2 \sigma_2^2 + \cdots + a_k^2 \sigma_k^2)}$$

$$M_Y(t) = e^{t \cdot \sum_{i=1}^{K} b_i} \, e^{t \cdot \sum_{i=1}^{K} a_i \mu_i + \frac{1}{2} t^2 \sum_{i=1}^{K} \sigma_i^2 a_i^2}$$

$$= e^{t\left[\sum_{i=1}^{K} (a_i \mu_i + b_i)\right] + \frac{1}{2} t^2 \sum_{i=1}^{K} a_i^2 \sigma_i^2}$$

this is the MGF of a normal variable with appropriate mean and variance resp.

That is

$$\boxed{Y \sim N\left(\sum_{i=1}^{K} (a_i \mu_i + b_i), \sum_{i=1}^{K} a_i^2 \sigma_i^2\right)}$$

Dear students: we started with one dimensional random variables and studied several probabilistic properties of different types of random variables. These concepts were extended to two-dimensional random variables as well.

In todays lecture we observed some of these properties for n-dimensional rvs as well.

Next we try to see a result when sample size $X_1, X_2, \ldots, X_n$ becomes extensively large i.e, as $n \to \infty$. Before that kindly note some definition regarding convergence of r.v's.

Convergence in Probability:

A sequence of random variables $X_1, X_2, \cdots X_n, \cdots$ is said to converge in Probability to a rv X if for every $\epsilon > 0$ we have

$$\lim_{n\to\infty} P\left[|X_n - X| > \epsilon\right] \longrightarrow 0 \quad \text{as } n \to \infty.$$

Law of Large Number (weak):

Let $X_1, X_2, \cdots, X_n$ be iid random variables with $E(X_i) = \mu$ and $V(X_i) = \sigma^2 < \infty$. Define $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$.

Then for every $\epsilon > 0$

$$\lim_{n\to\infty} P(|\bar{X}_n - \mu| > \epsilon) = 0 \quad \text{--- (1)}$$

that is $\bar{X}_n$ converges in Probability μ.

Proof: Proof follows from Chebyshev Inequality.

we have

$$P(|\bar{X}_n - \mu| > \epsilon) = P((\bar{X}_n - \mu)^2 \geq \epsilon^2)$$

$$\leq \frac{E(\bar{X}_n - \mu)^2}{\epsilon^2} = \frac{V(\bar{X}_n)}{\epsilon^2} = \frac{\sigma^2}{n\epsilon^2}$$

$$\longrightarrow 0 \quad \text{as } n \to \infty$$

So $\bar{X}_n \xrightarrow{\text{Converges}} \mu$ in probability as $n \to \infty$.

<u>Ex</u> : Let $X_1, X_2, \cdots, X_n \overset{iid}{\sim} Ber(p)$ then

$$\bar{X}_n \xrightarrow{Prob.} p \quad \text{as } n \to \infty$$

<u>Ex</u> : Let $X_1, X_2, \cdots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$

then $\bar{X}_n \xrightarrow{Prob.} \mu$ as $n \to \infty$

One of the most beautiful application of the concept of Convergence will be discussed in the next Class.

Ex : Let $X_1, X_2, \cdots, X_n \overset{iid}{\sim} Ber(p)$ then

$$\bar{X}_n \xrightarrow{Prob.} p \quad \text{as } n \to \infty$$

Ex : Let $X_1, X_2, \cdots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$

then $\bar{X}_n \xrightarrow{Prob.} \mu$ as $n \to \infty$

One of the most beautiful application of the concept of convergence will be discussed in the next class.