CHAPTER

4

Generating Functions

4.1. MOMENT GENERATING FUNCTION (AU 2018)

Definition. The moment generating function (m.g.f.) of a random variable X is denoted by $M_X(t)$ and is given by

$$M_X(t) = E(e^{tX}) = \begin{cases} \sum_x e^{tx} \cdot P(x), & \text{for discrete} \\ \int_{-\infty}^{\infty} e^{tx} \cdot f(x)\, dx, & \text{for continuous} \end{cases}$$

$M_X(t)$ exists if the series or integral is absolutely convergent

Generating moments through m.g.f.

$$M_X(t) = E(e^{tx}) = E\left[1 + tX + \frac{t^2X^2}{2!} + \ldots + \frac{t^rX^r}{r!} + \ldots\right]$$

$$= 1 + E(X) \cdot t + E(X^2) \cdot \frac{t^2}{2!} + \ldots + E(X^r) \cdot \frac{t^r}{r!} + \ldots$$

$$M_X(t) = 1 + \mu_1' t + \mu_2' \cdot \frac{t^2}{2!} + \ldots + \mu_r' \cdot \frac{t^r}{r!} + \ldots \quad \ldots(1)$$

$$\text{Here} \quad \mu_r' = E(X^r) = \begin{cases} \sum x^r \cdot P(x), & \text{for discrete} \\ \int_{-\infty}^{\infty} x^r \cdot f(x)\, dx, & \text{for continuous} \end{cases}$$

$\therefore$ We can observe that, the coefficient of $\frac{t^r}{r!}$ in $M_X(t)$ gives rth moment about origin. This mean $M_X(t)$ generate moments, hence it is called moment generating function. Differentiate (1) w.r.t. t r times and then putting $t = 0$, we get

$$\left[\frac{d^r}{dt^r} M_X(t)\right]_{t=0} = \left[\frac{\mu_r'}{r!} \cdot r! + \mu_{r+1}' t + \mu_{r+2}' \frac{t^2}{2!} + \ldots\right]_{t=0} = \mu_r'$$

$$\therefore \qquad \mu_r' = \left[\frac{d^r}{dt^r} \cdot M_X(t)\right]_{t=0}$$

Properties of moment generating function

1. $M_{CX} = M_X(ct)$, where C is a constant

   Proof : $M_{CX}(t) = E(e^{tCX}) = E[e^{X(Ct)}] = M_X^{(Ct)}$

2. ***Additive property of moment generating function*** **(AU 2017)**

   Statement. The moment generating function of the sum of n independent random variables is equal to the product of their respective m.g.fs.

   Symbolically, if $X_1, X_2, \ldots X_n$ are n independent random variables, then m.g.f. of their sum $X_1 + X_2 + \ldots + X_n$ is given by

   $$M_{X_1 + X_2 + \ldots + X_n}(t) = M_{X_1}(t) \,.\, M_{X_2}(t) \ldots M_{X_n}(t).$$

   Proof : Consider m.g.f. of $X_1 + X_2 + \ldots + X_n$ as

   $$M_{X_1 + X_2 + \ldots + X_n}(t) = E[e^{t(X_1+X_2+\ldots+X_n)}] \qquad (\because \text{ by the definition of m.g.f.})$$

   $$= E[e^{tX_1} . e^{tX2} \ldots e^{tXn}] = E(e^{tX_1}) \,.\, E(e^{tX2}) \ldots E(e^{tXn})$$

   $$(\because X_1, X_2, \ldots X_n \text{ are independent})$$

   $$= M_{X_1}(t) \,.\, M_{X_2}(t) \ldots M_{X_n}(t).$$

   $$\therefore \quad M_{X_1 + X_2 + \ldots + X_n}(t) = M_{X_1}(t) \,.\, M_{X_2}(t) \ldots M_{X_n}(t).$$

3. **Effect of change of origin and scale on m.g.f.** **(AU 2018)**

   Statement. The m.g.f. is not independent of shifting the origin scale.

   Proof : Let $U = \dfrac{X-a}{h}$

   Consider m.g.f. of U as

   $$M_U(t) = E(e^{tU}) = E\left[e^{t\left(\frac{X-a}{h}\right)}\right] = E\left[e^{\frac{tX}{h}} \,.\, e^{-\frac{ta}{h}}\right]$$

   $$= e^{-\frac{ta}{h}} E\left(e^{\frac{tX}{h}}\right) = e^{-\frac{ta}{h}} \,.\, M_X(t/h)$$

   $\therefore$ m.g.f. is affected by changing origin and scale *i.e.* it is not independent of changing the origin and scale.

4. **Uniqueness theorem of moment generating function.**

   The moment generating function of a distribution, if it exists, uniquely determines the distribution. This means corresponding to a given probability distribution, there is only one m.g.f. and corresponding to a given m.g.f., there is only one probability distribution.

   $\therefore \quad M_X(t) = M_Y(t) \;\Rightarrow\;$ X and Y are identically distributed.

Limitations of moment generating function

1. The moment generating function may exist but moments may not exist.

2. A random variable X can have m.g.f. and some or all moments, yet the m.g.f. does not generate the moments.
3. The m.g.f. may not exist but moments may exist.

4.2. CUMULANT GENERATING FUNCTION (AU 2018)

Definition. The cumulant generating function (c.g.f.) of a random variable X is denoted by $K_X(t)$ and is given by

$$K_X(t) = \log_e M_X(t)$$

Cumulants and moments through c.g.f.

Consider $K_1, K_2, \ldots K_r, \ldots$ are cumulants, then

$$K_X(t) = K_1.t + K_2 \cdot \frac{t^2}{2!} + \ldots + K_r \frac{t^r}{r!} + \ldots = \log M_X(t)$$

$$= \log \left[1 + \mu_1' t + \mu_2' \frac{t^2}{2!} + \ldots + \mu_r' \cdot \frac{t^r}{r!} + \ldots\right]$$

The rth cumulant K_r will be obtained by

$$K_r = \text{Coefficient of } \frac{t^r}{r!} \text{ in } K_X(t).$$

$$\therefore \quad K_1 t + K_2 \cdot \frac{t^2}{2!} + K_3 \frac{t^3}{3!} + K_4 \cdot \frac{t^4}{4!} + \ldots \ldots$$

$$= \left[\left(\mu_1' . t + \mu_2' \frac{t^2}{2!} + \mu_3' \frac{t^3}{3!} + \mu_4' \frac{t^4}{4!} + \ldots\right) - \frac{1}{2}\left(\mu_1' t + \mu_2' \frac{t^2}{2!} + \mu_3' \frac{t^3}{3!} + \ldots\right)^2 + \frac{1}{3}\left(\mu_1' t + \mu_2' \frac{t^2}{2!} + \ldots\right)^3 + \ldots\right] \quad \left(\because \log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - + \ldots\right)$$

By comparing the coefficients of powers of t on both sides, we get

$$K_1 = \text{Coefficient of } \frac{t^1}{1!} \text{ in } K_X(t) = \mu_1' = \text{mean} \qquad \therefore K_1 = \text{mean}$$

$$K_2 = \text{Coefficient of } \frac{t^2}{2!} \text{ in } K_X(t) = \mu_2' - \mu_1'^2 = \mu_2 = \text{variance} \qquad \therefore K_2 = \text{variance}$$

$$K_3 = \text{Coefficient of } \frac{t^3}{3!} \text{ in } K_X(t) = \mu_3' - 3\mu_2' \mu_1' + 2\mu_1'^3$$

$$= \mu_3 \text{. (Third Central moment)}$$

$$K_4 = \text{Coefficient of } \frac{t^4}{4!} \text{ in } K_X(t)$$

$$= \mu_4' - 3\mu_2'^2 + 4\mu_3' \mu_1' + 12 \mu_2' \mu_1'^2 - 6\mu_1'^4$$

$$= (\mu_4' - 4\mu_3' \mu_1' + 6\mu_2' \mu_1'^2 - 3\mu_1'^4) - 3(\mu_2'^2 - 2\mu_2' \mu_1'^2 + \mu_1'^4)$$

$$= \mu_4 - 3(\mu_2' - \mu_1'^2)^2 = \mu_4 - 3\mu_2^2$$

$$K_4 = \mu_4 - 3K_2^2 \qquad \therefore \mu_4 = K_4 + 3K_2^2$$

$\therefore$ **Moments from cumulants**

$$\text{Mean} = K_1$$

$$\text{Variance} = \mu_2 = K_2$$

$$\mu_3 = K_3$$

$$\mu_4 = K_4 + 3K_2^2$$

Properties of cumulant generating function

1. ***Additive property of cumulant generating function***

Statement. The cumulant generating function of sum of n independent random variables is equal to the sum of their respective cumulant generating functions.

Symbolically, if $X_1, X_2 \ldots X_n$ are n independent random variables, then

$$K_{X_1+X_2+\ldots+X_n}(t) = K_{X_1}(t) + K_{X_2}(t) + \ldots + K_{X_n}(t).$$

Proof : If $X_1, X_2, \ldots X_n$ are n independent random variables, then from the additive property of m.g.f.,

$$M_{X_1+X_2+\ldots+X_n}(t) = M_{X_1}(t) \,.\, M_{X_2}(t) \ldots\ldots M_{X_n}(t)$$

Consider logarithms,

$$\log [M_{X_1+X_2+\ldots+X_n}(t)] = \log [M_{X_1}(t) \,.\, M_{X_2}(t) \ldots M_{X_n}(t)]$$

$$K_{X_1+X_2+\ldots+X_n}(t) = \log M_{X_1}(t) + \log M_{X_2}(t) + \ldots + \log M_{X_n}(t).$$

$$= K_{X_1}(t) + K_{X_2}(t) + \ldots + K_{X_n}(t).$$

2. ***Effect of change of origin and scale on c.g.f.***

Statement. The cumulant generating function is not independent of shifting the origin and scale.

Proof : Let $U = \dfrac{X-a}{h}$

Consider c.g.f. of U

$$K_U(t) = \log M_U(t) = \log \left[E\left(e^{t\left(\frac{X-a}{h}\right)} \right) \right] = \log \left[E\left(e^{\frac{tX}{h}} \,.\, e^{-\frac{ta}{h}} \right) \right]$$

$$= \log \left[e^{-\frac{ta}{h}} \,.\, E\left(e^{\frac{tX}{h}} \right) \right] = \log \left[e^{-\frac{ta}{h}} \,.\, M_X(t/h) \right] = \log e^{-\frac{ta}{h}} + \log M_X(t/h)$$

$$K_U(t) = -\frac{ta}{h} + K_X(t/h)$$

$\therefore$ Cumulant generating function affected by shifting the origin and scale *i.e.* not independent.

Properties of Cumulants

1. ***Additive property of cumulants***

Statement. The rth cumulant of the sum of n independent random variables is equal to the sum of the rth cumulants of the individual cumulants.

Symbolically, if $X_1, X_2 \dots X_n$ are n independent random variables, then

$$K_r(X_1 + X_2 + \dots + X_n) = K_r(X_1) + K_r(X_2) + \dots + K_r(X_n)$$

Proof : From the additive property of c.g.f.

$$K_{X_1+X_2+\dots+X_n}(t) = K_{X_1}(t) + K_{X_2}(t) + \dots + K_{X_n}(t)$$

Differentiating w.r.t. t, r times and putting $t = 0$, we get

$$\left[\frac{d^r}{dt^r} K_{X_1+X_2+\dots+X_n}(t)\right]_{t=0} = \left[\frac{d^r}{dt^r} K_{X_1}(t)\right]_{t=0} + \left[\frac{d^r}{dt^r} K_{X_2}(t)\right]_{t=0} + \dots + \left[\frac{d^r}{dt^r} K_{X_n}(t)\right]_{t=0}$$

$$\Rightarrow \quad K_r(X_1 + X_2 + \dots + X_n) = K_r(X_1) + K_r(X_2) + \dots + K_r(X_n)$$

2. ***Effect of change of origin and scale on cumulant*** **(AU 2018)**

Statement. Cumulants are independent of change of origin but not on scale except first cumulant.

Proof : Let $U = \dfrac{X-a}{h}$, from the property of c.g.f., we have

$$K_U(t) = -\frac{at}{h} + K_X\left(\frac{t}{h}\right)$$

$$\Rightarrow \quad K_1' \cdot t + K_2' \cdot \frac{t^2}{2!} + \dots + K_r' \cdot \frac{t^r}{r!} + \dots$$

$$= -\frac{at}{h} + K_1 \cdot \frac{t}{h} + K_2 \cdot \frac{(t/h)^2}{2!} + \dots + K_r \cdot \frac{(t/h)^r}{r!} + \dots$$

where K_r' and K_r are rth cumulants of U and X respectively.

Now comparing coefficients, we get

$$K_1' = \frac{K_1 - a}{h}, \quad K_r' = \frac{K_r}{h^r}, \quad r = 2, 3, \dots$$

First cumulant is not independent of origin and scale. But except first cumulant, the remaining all cumulants are independent of changing the origin but dependent on changing the scale.

4.3. CHARACTERISTIC FUNCTION **(AU 2018)**

Sometimes m.g.f. does not exist, since the integral $\int_{-\infty}^{\infty} e^{tx} \cdot f(x)\,dx$ or the series $\sum_x e^{tx}$. P(x) does not converge absolutely for real values of t for some distributions (for example Cauchy distribution). In such cases, the characteristic function is used to calculate moments instead of m.g.f.

Definition. The characteristic function (c.f.) of a random variable X is denoted by $\phi_X(t)$ and is given by

$$\phi_X(t) = E(e^{itX}) = \begin{cases} \sum_x e^{itx} \, . \, P(x), & \text{for discrete} \\ \int_{-\infty}^{\infty} e^{itx} \, . \, f(x) \, dx, & \text{for continuous} \end{cases}$$

Properties of characteristic function

1. $\phi_X(0) = 1$

   Proof : $\phi_X(t) = E(^{itX})$

   at $t = 0$, $\phi_X(0) = E(e^0) = E(1) = 1$

2. $| \phi_X(t) | \leq 1$ *i.e.* characteristic function lies inbetween -1 and $+1$.

   Proof : Let X be a continuous random variable, then

$$\phi_X(t) = \int_{-\infty}^{\infty} e^{itx} f(x) \, dx$$

$$| \phi_X(t) | = \left| \int_{-\infty}^{\infty} e^{itx} f(x) \, dx \right| \leq \int_{-\infty}^{\infty} | e^{itx} f(x) | \, dx$$

$$= \int_{-\infty}^{\infty} | e^{itx} | \, . \, | f(x) | \, dx = \int_{-\infty}^{\infty} | e^{itx} | \, . \, f(x) \, dx \ (\because \ f(x) \geq 0)$$

$$= \int_{-\infty}^{\infty} f(x) \, dx \qquad (\because | e^{itx} | = \sqrt{\cos^2 tx + \sin^2 tx} = 1)$$

$$= 1$$

$\therefore \quad | \phi_X(t) | \leq 1.$

3. $\phi_{CX}(t) = \phi_X(ct)$, where c is constant

   Proof : $\phi_{CX}(t) = E(e^{itCX})$ (By the definition)

$$= E(e^{ict.X}) = M_X(Ct)$$

4. Additive property of characteristic function

   Statement. The characteristic function of the sum of n independent random variables is equal to the product of their characteristic functions.

   Symbolically, if $X_1, X_2, \ldots X_n$ are n independent random variables, then

$$\phi_{X_1 + X_2 + \ldots + X_n}(t) = \phi_{X_1}(t) \, . \, \phi_{X_2}(t) \ldots\ldots \phi_{X_n}(t).$$

   Proof : Consider characteristic function of $X_1 + X_2 + \ldots + X_n$

$$\phi_{X_1 + X_2 + \ldots + X_n}(t) = E[e^{it(X_1 + X_2 + \ldots + X_n)}]$$

$$= E(e^{itX_1} \cdot e^{itX_2} \ldots\ldots e^{itX_n}) = E(e^{itX_1}) \cdot E(e^{itX_2}) \ldots\ldots E(e^{itX_n})$$

($\because$ $X_1, X_2, \ldots X_n$ are independent)

$$= \phi_{X_1}(t) \cdot \phi_{X_2}(t) \ldots\ldots \phi_{X_n}(t).$$

5. ***Effect of charge of origin and scale on c.f.***

Statement. Characteristic function is not independent of changing the origin and scale.

Proof : Let $U = \dfrac{X-a}{h}$.

Consider the characteristic function of U

$$\phi_U(t) = E(e^{itU}) = E\left[e^{it\left(\frac{X-a}{h}\right)}\right] = E\left[e^{\frac{itX}{h}} \cdot e^{-\frac{ita}{h}}\right] = e^{-\frac{ita}{h}} \cdot E\left(e^{\frac{itX}{h}}\right)$$

$$\phi_U(t) = e^{-\frac{ita}{h}} \cdot \phi_X\left(\frac{it}{h}\right)$$

$\therefore$ The characteristic function is affected by shifting both origin and scale, *i.e.*, c.f. is not independent of changing the origin and scale.

6. ***Uniquess theorem of characteristic function***

Characteristic function uniquely determines the distribution.

i.e. if $\phi_X(t) = \phi_Y(t)$, then $f(x) = f(y)$ means X and Y are identically distributed.

4.4. PROBABILITY GENERATING FUNCTION

Definition. The probability generating function (p.g.f.) of a discrete random variable X is denoted by $P_X(S)$ and is given by

$$P_X(S) = E(S^X) = \sum_{x=0}^{\infty} S^x P(x).$$

Explanation : Let us suppose X is a discrete random variable taking the values 0, 1, 2, ... with respective probabilities $P_0, P_1, P_2, \ldots$ Then expectation of the function S^x of the random variable X is

$$P_X(S) = E(S^X) = P_0 \cdot S^0 + P_1 \cdot S^1 + P_2 \cdot S^2 + \ldots$$

$$= \sum_{x=0}^{\infty} S^x \cdot P_x = \sum_{x=0}^{\infty} S^x \cdot P(X = x) \quad \text{or} \quad \sum_{x=0}^{\infty} S^x \cdot P(x)$$

where $P(x)$ is p.m.f. of X.

Properties of Probability general function

1. ***Additive property of p.g.f.***

Statement. The probability generating function of the sum of 'n' independent random variables is equal to product of their p.g.fs.

Symbolically, if $X_1, X_2, \ldots X_n$ are 'n' independent random variables, then

$$P_{X_1 + X_2 + \ldots + X_n}(S) = P_{X_1}(S) \cdot P_{X_2}(S) \ldots\ldots P_{X_n}(S).$$

Proof : Consider the probability generating function of $X_1 + X_2 + \ldots + X_n$ as

$$P_{X_1 + X_2 + \ldots + X_n}(S) = E(S^{X_1 + X_2 + \ldots + X_n})$$
$$= E(S^{X_1} \,.\, S^{X_2} \ldots S^{X_n}) = E(S^{X_1}) \,.\, E(S^{X_2}) \ldots E(S^{X_n})$$
$(\because X_1, X_2, \ldots X_n$ are independent)
$$= P_{X_1}(S) \,.\, P_{X_2}(S) \ldots P_{X_n}(S).$$

2. ***Effect of change of origin and sale on p.g.f.***

Statement. The probability generating function is not independent of changing the origin and scale.

Proof : Let $U = \dfrac{X - a}{h}$

Consider probability generating function of U as

$$P_U(S) = E(S^U)$$
$$= E\left(S^{\frac{X-a}{h}}\right) = E(S^{X/h} \,.\, S^{-a/h}) = S^{-a/h} \,.\, E(S^{X/h})$$
$$= S^{-a/h} \,.\, [E[S^{1/h})^X]] = S^{-a/h} \,.\, P_X(S^{1/h})$$

$\therefore$ The probability generating function is affected by shifting the origin and scale *i.e.* p.g.f. is not independent of changing the origin and scale.

4.5. WEAK LAW OF LARGE NUMBERS

Let $X_1, X_2, \ldots X_n$ be a sequence of random variables and $\mu_1, \mu_2, \ldots \mu_n$ be their respective expectations (means) and let

$$B_n = \text{Var}(X_1 + X_2 + \ldots + X_n) < \infty, \text{ then}$$

$$P\left\{\left| \frac{X_1 + X_2 + \ldots + X_n}{n} - \frac{\mu_1 + \mu_2 + \ldots + \mu_n}{n} \right| < \varepsilon\right\} \geq 1 - \eta$$

$\forall\ n > n_0$, ε and η are arbitrary small positive numbers, provided $\lim\limits_{n \to \infty} \dfrac{B_n}{n^2} \to 0$.

Weak Law of Large numbers for i.i.d. random variables

If the variables $X_1, X_2, \ldots X_n$ are independently and identically distributed (i.i.d.) *i.e.* $E(X_i) = \mu$ and $\text{Var}(X_i) = \sigma^2 \forall\ i = 1, 2, \ldots n$, then $B_n = \text{Var}(X_1 + X_2 + \ldots + X_n) = \sum\limits_{i=1}^{n} \text{Var}(X_i) = n\sigma^2$. (Covariance term vanish since variables are independent).

$$\therefore \lim_{n \to \infty} \frac{B_n}{n^2} = \lim_{n \to \infty} \frac{n\sigma^2}{n^2} = \lim_{n \to \infty} \frac{\sigma^2}{n} = 0$$

Then weak law of large numbers for sequence of $\{X_n\}$ of i.i.d. random variables is

$$P\left\{\left| \frac{X_1 + X_2 + \ldots + X_n}{n} - \mu \right| < \varepsilon\right\} > 1 - \eta \quad \forall\ n > n_0$$

$$\Rightarrow \quad P\{| \overline{X}_n - \mu | < \varepsilon\} > 1 - \eta \quad \forall\ n > n_0$$

$$P\{|\overline{X}_n - \mu| < \varepsilon\} \to 1 \quad \text{as } n \to \infty$$

i.e.

$$P\{|\overline{X}_n - \mu| \geq \varepsilon\} \to 0 \quad \text{as } n \to \infty$$

(or)

This means $\overline{X}_n$ convergence in probability to μ, *i.e.* $\overline{X}_n \xrightarrow{P} \mu$.

4.6. STRONG LAW OF LARGE NUMBERS

The strong law of large numbers state that the sample mean (average) converges almost surely to the population mean.

Let $X_1, X_2, \ldots, X_n$ be independently and identically distributed (*i.i.d.*) *i.e.*, $E(X_i) = \mu$ and Var $(X_i) = \sigma^2$ then the sample mean $\overline{X}_n = \dfrac{X_1 + X_2 + \ldots + X_n}{n}$ converges to the population mean μ with probability unity.

PROBLEMS

CHAPTER

6

Central Limit Theorem and Its Applications

6.1. CONVERGENCE IN LAW (OR) CONVERGENCE IN DISTRIBUTION

A sequence of random variables $X_1, X_2, \ldots, X_n$ is said to be convergence is law or convergence in distribution to a random variable X if

$$\lim_{n \to \infty} F_n(x) = F(x)$$

for all x at which F is continuous. $F_n(x)$ and $F(x)$ are distribution functions of random variables X_n and X respectively.

6.2. CONVERGENCE IN PROBABILITY

A sequence of random variables $X_1, X_2, \ldots, X_n$ is said to be convergence in probability to a constant a if for any $\varepsilon > 0$,

$$\lim_{n \to \infty} P\{| X_n - a | < \varepsilon\} = 1$$

or

$$\lim_{n \to \infty} P\{| X_n - a | \geq \varepsilon\} = 0$$

It is written as $X_n \xrightarrow{p} a$ as $n \to \infty$.

6.4. CENTRAL LIMIT THEOREM (CLT)

The central limit theorem was first stated by Laplace in 1812 and the proof was given by Liapounoff in 1901.

Statement. If $X_1, X_2, \ldots X_n$ be n independent random variables with mean $E(X_i) = \mu_i$ and $\text{Var}(X_i) = \sigma_i^2$, $i = 1, 2, \ldots n$, then it can be proved that under certain very general conditions, the random variable $S_n = X_1 + X_2 + \ldots + X_n$ is asymptotically normal distribution with mean $\mu = \sum_{i=1}^{n} \mu_i$ and Variance $\sigma^2 = \sum_{i=1}^{n} \sigma_i^2$

(With the standard deviation σ)

6.5. LINDEBERG—LEVY CENTRAL LIMIT THEOREM (CLT) : (CENTRAL LIMIT THEOREM FOR I.I.D. RANDOM VARIABLES)

According to Lindeberg—Levy, if $X_1, X_2, \ldots X_n$ are independently and identically distributed (*i.i.d*) random variables with $E(X_l) = \mu_1$, $\text{Var}(X_i) = \sigma_1^2$, $i = 1, 2, \ldots, n$, then the sum $S_n = X_1 + X_2 + \ldots + X_n$ is asymptotically normally distributed with mean $\mu = n\,\mu_1$ and variance $\sigma^2 = n\,\sigma_1^2$.

The Lindeberg-Levy Central limit theorem can also be stated in the following three ways.

(*i*) If $X_1, X_2, \ldots X_n$ are *i.i.d.* variables with mean μ_1 and variance σ_1^2 (finite) and $S_n = X_1 + X_2 + \ldots + X_n$, then

$$\lim_{n \to \infty} P\left\{a \le \frac{S_n - n\,\mu_1}{\sigma_1/\sqrt{n}} \le b\right\} = \phi(b) - \phi(a) = \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx$$

for $-\infty < a < b < \infty$ and $\phi(-\infty) = 0$, $\phi(\infty) = 1$.

where $\phi(x)$ is a distribution function.

(*ii*) $$\lim_{n \to \infty} P\left\{a \le \frac{S_n - E(S_n)}{\sqrt{\text{Var}(S_n)}} \le b\right\} = \phi(b) - \phi(a)$$

(*iii*) $$\lim_{n \to \infty} P\left\{a \le \frac{\bar{X}_n - E(\bar{X}_n)}{\sqrt{\text{Var}(\bar{X}_n)}} \le b\right\} = \phi(b) - \phi(a)$$

i.e., $$\lim_{n \to \infty} P\left\{a \le \frac{\bar{X}_n - \mu_1}{\sigma_1/\sqrt{n}} \le b\right\} = \phi(b) - \phi(a)$$

6.6. APPLICATIONS OF CENTRAL LIMIT THEOREM

PROBLEM 1. Discuss the distribution of sum of n *i.i.d.* binomial variates.

SOLUTION. If $X_1, X_2, \ldots, X_n$ are *i.i.d* $B(r, p)$ (consider r, p are parameters of binomial distribution) *i.e.*, $E(X_i) = rp$, $\text{Var}(X_i) = rpq$.

If we consider sum of n *i.i.d.* binomial variates as S_n

$$S_n = X_1 + X_2 + \dots + X_n = \sum_{i=1}^{n} X_i \text{ then}$$

$$E(S_n) = E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i) = \sum_{i=1}^{n} rp = nrp$$

$$\text{Var}(S_n) = \text{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \text{Var}(X_i) = \sum_{i=1}^{n} rpq = nrpq$$

By using Lindeberg-Levy central limit theorem,

$$\lim_{n \to \infty} P\left\{a \le \frac{S_n - E(S_n)}{\sqrt{\text{Var}(S_n)}} \le b\right\} = \phi(b) - \phi(a)$$

$$\Rightarrow \quad \lim_{n \to \infty} P\left\{a \le \frac{S_n - nrp}{\sqrt{nrpq}} \le b\right\} = \phi(b) - \phi(a),\ 0 \le p < 1.$$

2. Prove that if Y_n is a binomial variate with parameters n and p, then

$$\lim_{n \to \infty} P\left\{a \le \frac{Y_n - np}{\sqrt{np(1-p)}} \le b\right\} = \phi(b) - \phi(a),\ 0 < p < 1.$$

Proof: Let us consider $X_1, X_2, \dots, X_n$ be n *i.i.d.* Bernoullian variates *i.e.*, $X_i \sim$ i.i.d (1, p) if we let $S_n = X_1 + X_2 + \dots + X_n$, then $S_n \sim B(n, p)$.

Given $Y_n \sim B(n, p)$, hence by using Lindeberg-Levy central limit theorem for Y_n,

$$\lim_{n \to \infty} P\left\{a \le \frac{Y_n - E(Y_n)}{\sqrt{\text{Var}(Y_n)}} \le b\right\} = \phi(b) - \phi(a)$$

Since $Y_n \sim B(n, p)$, $E(Y_n) = np$ and $\text{Var}(Y_n) = npq = np(1-p)$

$$\therefore \quad \lim_{n \to \infty} P\left\{a \le \frac{Y_n - np}{\sqrt{np(1-p)}} \le b\right\} = \phi(b) - \phi(a),\ 0 < p < 1.$$

3. If Y_n is distributed as poisson P(n), then

$$\lim_{n \to \infty} P\left\{a \le \frac{Y_n - n}{\sqrt{n}} \le b\right\} = \phi(b) - \phi(a), \text{ then prove that}$$

$$\lim_{n \to \infty} P(Y_n \le n) = \frac{1}{2} \text{ i.e., } \sum_{k=0}^{n} \frac{e^{-n} n^k}{k!} = \frac{1}{2} \text{ as } n \to \infty$$

Proof: Let us consider $X_1, X_2, \dots, X_n$ be *i.i.d.* P(1), then

$$S_n = X_1 + X_2 + \dots + X_n \ (= Y_n) \sim P(n)$$

This means $S_n = Y_n$, by using Lindeberg-Levy Central limit theorem,

$$\lim_{n \to \infty} P\left\{a \le \frac{Y_n - E(Y_n)}{\sqrt{\text{Var}(Y_n)}} \le b\right\} = \phi(b) - \phi(a)$$

Since Y_n follows poisson distribution,

$$E(Y_n) = n,\ \text{Var}(Y_n) = n.$$

$$\therefore \quad P\left\{a \le \frac{Y_n - n}{\sqrt{n}} \le b\right\} = \phi(b) - \phi(a) \text{ as } n \to \infty \quad \dots(1$$

Now let us take $a = -\infty$ and $b = 0$ in (1), we have

$$P\left\{a \le \frac{Y_n - n}{\sqrt{n}} \le b\right\} = P\left\{-\infty \le \frac{Y_n - n}{\sqrt{n}} \le 0\right\}$$

$$= P\left\{\frac{Y_n - n}{\sqrt{n}} \le 0\right\} = P\{Y_n - n \le 0\} = P\{Y_n \le n\} \quad \text{...(2)}$$

and we have,

$$\phi(b) - \phi(a) = \phi(0) - \phi(-\infty) = \frac{1}{2} \quad \text{...(3()}$$

From (2) and (3), we get

$$P(Y_n \le n) = \frac{1}{2} \text{ as } n \to \infty$$

$$\Rightarrow \quad \sum_{k=0}^{n} \frac{e^{-n} n^k}{k\,!} = \frac{1}{2} \text{ as } n \to \infty.$$

6.7. RELATION BETWEEN CLT AND WLLN

The two important laws, central limit theorem (CLT) and weak law of large numbers (WLLN) hold for a sequence $\{X_n\}$ of *i.i.d.* random variables with finite mean μ and variance σ^2.

The central limit theorem CLT is so stronger than the WLLN.

The WLLN for the sequence $\{X_n\}$ of *i.i.d.* random variables provide an estimate of the $P\left\{\frac{|S_n - n\mu|}{n} \ge \varepsilon\right\}$.

From the WLLN, $P\left\{\left|\frac{X_1 + X_2 + \dots + X_n}{n} - \mu\right| \ge \varepsilon\right\}$

$$= P\left\{|\bar{X}_n - \mu| \ge \varepsilon\right\} = P\left\{\left|\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}}\right| \ge \frac{\varepsilon}{\sigma/\sqrt{n}}\right\} = P\left\{|z| \ge \frac{\varepsilon\sqrt{n}}{\sigma}\right\}$$

where $z \sim N(0, 1)$ standard normal variate.

$$= 1 - P\left\{|z| \le \frac{\varepsilon\sqrt{n}}{\sigma}\right\}$$

$$= 1 - P\left[-\frac{\varepsilon\sqrt{n}}{\sigma} \le z \le \frac{\varepsilon\sqrt{n}}{\sigma}\right] = 1 - \left\{\phi\left(\frac{\varepsilon\sqrt{n}}{\sigma}\right) - \phi\left(\frac{-\varepsilon\sqrt{n}}{\sigma}\right)\right\}$$

where $\phi(\cdot)$ is distribution function of standard normal variate.

The relation between CLT and WLLN also states that

(1) For the sequence $\{X_n\}$ WLLN holds for independent and uniformly bounded variables and CLT holds good provided by

$$B_n = \text{Var}(X_1 + X_2 + \dots + X_n) = \sigma_1^2 + \sigma_2^2 + \dots + \sigma_n^2 \to \infty \text{ as } n \to \infty$$

(2) For the sequence $\{X_n\}$ of independent random variables, CLT may hold but the WLLN may not hold.

6.8. LIAPOUNOFF'S CENTRAL LIMIT THEOREM (CLT)

Liapounoff was generalised the central limit theorem (CLT) when the variables are not identically distributed. $X_1, X_2, \ldots, X_n$ be n independent random variables with the means $E(X_i) = \mu_i$ and variances $Var(X_i) = \sigma_i^2$, $i = 1, 2, \ldots, n$ respectively. And let us consider third absolute moment about mean of X_i be

$$\rho_i^3 = E\{|X_i - \mu_i|^3\},\ i = 1, 2, \ldots n \text{ is finite, and let } \rho^3 = \sum_{i=1}^{n} \rho_i^3.$$

Liapounoff stated that CLT, if $\lim_{n \to \infty} \frac{\rho}{\sigma} = 0$, then the distribution of the sum

$$S_n = X = X_1 + X_2 + \ldots + X_n \text{ is asymptotically } N(\mu, \sigma^2)$$

where $\mu = \sum_{i=1}^{n} \mu_i$ and $\sigma^2 = \sum_{i=1}^{n} \sigma_i^2$.

• • •