3.4

Cumulant Generating Function

Just as the moment generating function (m.g.f.) $M_X(t)$ or characteristic function (ch.f.) $\emptyset_X(t)$ of a r.v.X *generates its moments*, the logarithm of $M_X(t)$ or $\emptyset_X(t)$ generates a sequence of numbers called the **Cumulants of $\boldsymbol{X}$**. Cumulants are of interest for the following two reasons.

1. Moments in terms of cumulants can be obtained easily when compared to obtaining them from m.g.f. or ch.f.
2. j^{th} cumulant of a sum of independent r.vs is simply the sum of the j^{th} cumulants of the summand.

Since the ch.f. exists for every r.v(the m.g.f. need not exist for some r.vs), the cumulant generating function (c.g.f.) is defined as the logarithm of the ch.f.

Cumulant generating function: Let X be a r.v. with characteristic function $\emptyset_X(t) = E\left[e^{itX}\right]$. The cumulant generating function (c.g.f.) of X is defined by

$$\boldsymbol{K_X(t) = ln\left(\emptyset_X(t)\right)} \qquad \dots (1)$$

for all t in some open interval about 0 in $\boldsymbol{R}$, provided the RHS can be expanded as a convergent series in powers of t.

Thus,

$$K_X(t) = k_1(it) + k_2\frac{(it)^2}{2!} + \cdots + k_r\frac{(it)^r}{r!} \qquad \dots.(2)$$

Note that, $k_j =$ coef of $\frac{(it)^j}{j!}$ in $K_X(t)$ and it is called the $\boldsymbol{j^{th}}$ **Cumulant** of X

We have,

$$\emptyset_X(t) = 1 + \mu_1{}'(it) + \mu_2{}'\frac{(it)^2}{2!} + \cdots + \mu_r{}'\frac{(it)^r}{r!} \qquad \dots. (3)$$

From $(1), (2)$ and (3), we have

$$k_1(it) + k_2\frac{(it)^2}{2!} + \cdots = ln[1 + \mu_1'(it) + \cdots + \cdots]$$

$$= \left(\mu_1'(it) + \mu_2'\frac{(it)^2}{2!} + \cdots\right) - \frac{1}{2}\left(\mu_1'(it) + \mu_2'\frac{(it)^2}{2!} + \cdots\right)^2 + \frac{1}{3}\left(\mu_1'(it) + \mu_2'\frac{(it)^2}{2!} + \cdots\right)^3 - \cdots$$

$$(\because ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots)$$

Comparing the coefficients of like powers of t, we get the relationship between moments and cumulants. Hence, we have

$k_1 = \mu_1' =$ Mean $= \mu$ and $k_2 = \mu_2' - (\mu_1')^2 =$ variance $= \sigma^2$.
Thus, $\mu = k_1$ and $\sigma^2 = k_2$.

Note:

1. From (2), $K_X(t)$ can be written as

$$K_X(t) = \sum_{j=1}^{\infty} k_j \frac{(it)^j}{j!}$$

Thus j^{th} cumulant $= k_j =$ coef. of $\frac{(it)^j}{j!}$ in $K_X(t)$.

2. From (2), j^{th} cumulant is obtained as

$$k_j = (-i)^j \left.\frac{d^j K_X(t)}{dt^j}\right|_{t=0}$$

Example 1: If $X \sim B(n, p)$, then obtain the c.g.f. of X and hence obtain its mean and variance.

Solution: Since $X \sim B(n, p)$, its p.m.f. is given by

$$p(x) = \binom{n}{x} p^x q^{n-x}, x = 0,1,2,\dots,n$$

Then the characteristic function of X is given by

$$\emptyset_X(t) = E\left[e^{itX}\right] = \sum_{x=0}^{\infty} e^{itx}\binom{n}{x}p^x q^{n-x} = \sum_{x=0}^{\infty}\binom{n}{x}\left(pe^{it}\right)^x q^{n-x}$$

$$= \left(q + pe^{it}\right)^n$$

$$\Rightarrow \emptyset_X(t) = \left(q + pe^{it}\right)^n$$

Thus, the c.g.f. of X is given by

$$K_X(t) = ln\left(\emptyset_X(t)\right) = ln\left[q + pe^{it}\right]^n$$

$$\Rightarrow K_X(t) = n\, ln\left(q + pe^{it}\right)$$

$$= n\, ln\left[q + p\left(1 + (it) + \frac{(it)^2}{2!} + \cdots\right)\right]$$

$$= n\, ln\left[1 + (it)p + \frac{(it)^2}{2!}p + \cdots\right]$$

$$\Rightarrow K_X(t) = n\left[\left\{(it)p + \frac{(it)^2}{2!}p + \cdots\right\} - \frac{1}{2}\left\{(it)p + \frac{(it)^2}{2!}p + \cdots\right\}^2 + \cdots\right]$$

$$\Rightarrow K_X(t) = (it)(np) + \frac{(it)^2}{2!}(np - np^2) + \cdots$$

$\therefore k_1 =$ coef. of $(it) = np$ and $k_2 =$ coef. of $\frac{(it)^2}{2!} = np - np^2 = np(1-p) = npq$

Thus mean and variance are given by $\mu = k_1 = np$ and $\sigma^2 = k_2 = npq$

Example 2: If $X\sim$ Poisson $P(\lambda)$, then find the c.g.f. of X and hence obtain its mean and variance.

Solution: Since $X\sim P(\lambda)$, , its p.m.f. is given by

$$p(x) = \frac{e^{-\lambda}\lambda^x}{x!}\ , x = 0,1,2, \ldots$$

The characteristic function of X is given by

$$\emptyset_X(t) = E\left[e^{itX}\right] = \sum_{x=0}^{\infty} e^{itx} p(x) = \sum_{x=0}^{\infty} e^{itx} \frac{e^{-\lambda}\lambda^x}{x!} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{\left(\lambda e^{it}\right)^x}{x!}$$

$$= e^{-\lambda} e^{\lambda e^{it}} = e^{\lambda(e^{it}-1)}$$

Thus, $\emptyset_X(t) = e^{\lambda(e^{it}-1)}$

The c.g.f. of X is given by

$$K_X(t) = ln\left(\emptyset_X(t)\right) = ln\left[e^{\lambda(e^{it}-1)}\right] = \lambda\left(e^{it} - 1\right)$$

$$= \lambda\left[1 + (it) + \frac{(it)^2}{2!} + \frac{(it)^3}{3!} + \cdots - 1\right]$$

$$\Longrightarrow K_X(t) = (it)\lambda + \frac{(it)^2}{2!}\lambda + \cdots$$

Thus, $k_1 =$ coef. of (it) in $K_X(t) = \lambda$ and $k_2 =$ coef. of $\frac{(it)^2}{2!}$ in $K_X(t) = \lambda$

Hence, mean $=$ variance $= \lambda$

Example 3: If $X \sim NB(r, p)$, then find the c.g.f. of X and hence obtain its mean and variance.

Solution: Since $X \sim NB(r, p)$, its p.m.f is given by

$$p(x) = \binom{-r}{x} p^r (-q)^x, x = 0,1,2, \ldots$$

The characteristic function of X is given by

$$\emptyset_X(t) = E\left[e^{itX}\right] = \sum_{x=0}^{\infty} e^{itx} p(x) = \sum_{x=0}^{\infty} e^{itx} \binom{-r}{x} p^r (-q)^x$$

$$= p^r \sum_{x=0}^{\infty} \binom{-r}{x} \left(-qe^{it}\right)^x = p^r\left(1 - qe^{it}\right)^{-r}$$

$$\Longrightarrow \emptyset_X(t) = p^r\left(1 - qe^{it}\right)^{-r}$$

The c.g.f. is given by

$$K_X(t) = ln\left(\emptyset_X(t)\right) = ln\left(p^r\left(1 - qe^{it}\right)^{-r}\right)$$

$$= r\, ln\, p - r\, ln\left(1 - qe^{it}\right)$$

Now, $\frac{d}{dt}\left(K_X(t)\right) = (-r)\frac{-iqe^{it}}{1-qe^{it}} = \frac{irqe^{it}}{1-qe^{it}}$

$$\therefore k_1 = (-i)\frac{d}{dt}\left(K_X(t)\right)\bigg|_{t=0} = (-i)\frac{(irq)}{1-q} = \frac{rq}{p}$$

And $\frac{d^2}{dt^2}\left(K_X(t)\right) = irq\frac{d}{dt}\left[\frac{e^{it}}{1-qe^{it}}\right] = irq\left[\frac{(1-qe^{it})ie^{it}+e^{it}qie^{it}}{\left(1-qe^{it}\right)^2}\right]$

$$\therefore k_2 = (-i)^2\frac{d^2}{dt^2}\left(K_X(t)\right)\bigg|_{t=0} = (-i)^2(irq)(i)\left[\frac{p+q}{p^2}\right]$$

$$\Longrightarrow k_2 = \frac{rq}{p^2}$$

Thus mean and variance are given by

$\mu = k_1 = \frac{rq}{p}$ and $\sigma^2 = k_2 = \frac{rq}{p^2}$ respectively.

Example 4: If $X \sim N(\mu, \sigma^2)$, then obtain the c.g.f. of X and hence find its mean and variance.

Solution: If $X \sim N(\mu, \sigma^2)$, then its characteristic function can be shown that

$$\emptyset_X(t) = e^{it\mu - \frac{1}{2}\sigma^2 t^2} \qquad \text{(recall!)}$$

Hence, the c.g.f. is given by

$$K_X(t) = ln\left[e^{it\mu - \frac{1}{2}\sigma^2 t^2}\right] = it\mu - \frac{1}{2}\sigma^2 t^2$$

$$\Longrightarrow K_X(t) = (it)\mu + \frac{(it)^2}{2!}\sigma^2$$

$\therefore k_1 =$ coef. of (it) in $K_X(t) = \mu$ and $k_2 =$ coef. of $\frac{(it)^2}{2!}$ in $K_X(t) = \sigma^2$

Thus, mean $= \mu$ and variance $= \sigma^2$.

Example 5: If $X \sim E(\lambda)$, then obtain the c.g.f. of X and hence obtain its mean and variance.

Solution: Since $X \sim E(\lambda)$, its ch.f. can be shown that

$$\emptyset_X(t) = \frac{\lambda}{\lambda - it} = \left(1 - \frac{it}{\lambda}\right)^{-1} \qquad \text{(recall!)}$$

The c.g.f. of X is given by

$$K_X(t) = ln(\emptyset_X(t) = (-1)\, ln\left(1 - \frac{it}{\lambda}\right) = (-1)\left[-\frac{it}{\lambda} - \frac{1}{2}\left(\frac{it}{\lambda}\right)^2 - \cdots\right]$$

$$\Longrightarrow K_X(t) = \left[(it)\frac{1}{\lambda} + \frac{(it)^2}{2!}\frac{1}{\lambda^2} + \cdots\right]$$

Thus, $k_1 =$ coef. of (it) in $K_X(t) = \frac{1}{\lambda}$ and $k_2 =$ coef. of $\frac{(it)^2}{2!}$ in $K_X(t) = \frac{1}{\lambda^2}$.
Thus, the mean and variance are given by $\mu = \frac{1}{\lambda}$ and $\sigma^2 = \frac{1}{\lambda^2}$ respectively.

Properties of Cumulants: Here we develop some useful properties of cumulants. Let $\boldsymbol{k_n(X)}$ be the $\boldsymbol{n}^{\text{th}}$ **cumulant** of a r.v. X.

Theorem 1: $\boldsymbol{k_n(cX) = c^n k_n(X)}$ for some real constant $\boldsymbol{c}$.

Proof: Consider $\emptyset_{cX}(t) = E\left[e^{itcX}\right] = E\left[e^{i(tc)X}\right] = \emptyset_X(tc)$

$$\Longrightarrow \emptyset_{cX}(t) = \emptyset_X(ct)$$

$$\Longrightarrow ln\,\emptyset_{cX}(t) = ln\,\emptyset_X(ct)$$

$$\Longrightarrow K_{cX}(t) = K_X(ct)$$

Then, $\frac{d^n}{dt^n}\left(K_{cX}(t)\right)\Big|_{t=0} = \frac{d^n}{dt^n}\left(K_X(ct)\right)\Big|_{t=0} = \frac{d^n}{ds^n}\left(K_X(s)\right)\Big|_{s=0} c^n$,where $ct = s$

Therefore, $(-i)^n \frac{d^n}{dt^n}\left(K_{cX}(t)\right)\Big|_{t=0} = (-i)^n c^n \frac{d^n}{dt^n}\left(K_X(t)\right)\Big|_{t=0}$

$\Longrightarrow k_n(cX) = c^n k_n(X)$

Theorem 2: $\boldsymbol{k_n(X+b) = \begin{cases} k_n(X) + b & , \ if \quad n = 1 \\ k_n(X) & , \ if \quad n > 1 \end{cases}}$

Proof: $\emptyset_{X+b}(t) = E\left[e^{it(X+b)}\right] = e^{itb} E\left[e^{itX}\right] = e^{itb} \emptyset_X(t)$

$\Longrightarrow \emptyset_{X+b}(t) = e^{itb} \emptyset_X(t)$

$\Longrightarrow ln[\emptyset_{X+b}(t)] = itb + ln[\emptyset_X(t)]$

$\Longrightarrow K_{X+b}(t) = itb + K_X(t)$

$$\Longrightarrow (-i)^n \frac{d^n}{dt^n}\left(K_{X+b}(t)\right)\Big|_{t=0} = (-i)^n \frac{d^n}{dt^n}\left((itb)\right)\Big|_{t=0} + (-i)^n \frac{d^n}{dt^n}\left(K_X(t)\right)\Big|_{t=0}$$

If $n = 1$, then $K_n(X+b) = b + K_n(X)$.

If $n > 1$, $K_n(X+b) = K_n(X)$.

Theorem3: If $\boldsymbol{X}$ and $\boldsymbol{Y}$ are independent random variables and $\boldsymbol{S = X + Y}$, then

$$\boldsymbol{k_n(S) = k_n(X) + k_n(Y).}$$

Proof: Since X and Y are independent,

$$\emptyset_S(t) = \emptyset_X(t)\emptyset_Y(t)$$

$\Longrightarrow ln[\emptyset_S(t)] = ln[\emptyset_X(t)] + ln[\emptyset_Y(t)] \Longrightarrow K_S(t) = K_X(t) + K_Y(t)$

$$\Longrightarrow (-i)^n \frac{d^n}{dt^n}\left(K_S(t)\right)\Big|_{t=0} = (-i)^n \frac{d^n}{dt^n}\left(K_X(t)\right)\Big|_{t=0} + (-i)^n \frac{d^n}{dt^n}\left(K_Y(t)\right)\Big|_{t=0}$$

$\Longrightarrow k_n(S) = k_n(X) + k_n(Y)$

Generalization: If $\boldsymbol{S = X_1 + \cdots + X_m}$ where $\boldsymbol{X_1, X_2, \dots, X_m}$ are independent random variables, then

$$\boldsymbol{k_n(S) = k_n(X_1) + k_n(X_2) + \cdots + k_n(X_m)}$$

Theorem 4: Let $\mu_j' = E(X^j)$ be the j^{th} moment of X about zero for $j = 1, 2, 3, \ldots, n$ where $\mu_0' = 1$. Let $k_1, k_2, \ldots, k_n$ be the n cumulants of X. Then

$$\mu_{r+1}' = \sum_{j=0}^{r} \binom{r}{j} \mu_j' k_{(r+1-j)} \qquad \text{.... (1)}$$

for $r = 0,1, \ldots, n-1$.

Proof: For $j = 0,1,2, \ldots, n$, we have

$$\mu_j' = \frac{d^j}{dt^j}\left(\emptyset_X(t)\right)\Big|_{t=0} \text{ and } k_j = (-i)^j \frac{d^j}{dt^j}\left(K_X(t)\right)\Big|_{t=0}$$

where $\emptyset_X(t) = E\left[e^{itX}\right]$ and $K_X(t) = ln[\emptyset(t)]$ or equivalently, $\emptyset_X(t) = e^{K_X(t)}$. Differentiating this last identify w.r.t. t gives

$$\emptyset_X'(t) = e^{K_X(t)} K_X'(t) \qquad \text{...(2)}$$

and evaluating this at $t = 0$ gives $i\mu_1' = ik_1 \Longrightarrow \mu_1' = k_1$ holds for $r = 0$.

Differentiating (2) for r times, it gives

$$\emptyset_X{}^{(r+1)}(t) = \sum_{j=0}^{r} \binom{r}{j} \emptyset_X{}^{(j)}(t) K_X{}^{(r+1-j)}(t)$$

(Use Leibnitz theorem for the n^{th} derivative of the product of two functions)

and evaluating this at $t = 0$ gives

$$\mu_{r+1}' = \sum_{j=0}^{r} \binom{r}{j} \mu_j' k_{(r+1-j)} \quad \text{for } r = 0,1, \ldots, n-1$$

Note: Taking $r = 0,1,2,3$ in (1) produces

$$\left.\begin{aligned} \mu_1' &= k_1 \\ \mu_2' &= k_2 + \mu_1' k_1 \\ \mu_3' &= k_3 + 2\mu_1' k_2 + \mu_2' k_1 \\ \mu_4' &= k_4 + 3\mu_1' k_3 + 3\mu_2' k_2 + \mu_3' k_1 \end{aligned}\right\} \qquad \text{.....(3)}$$

These recursive formulae can be used to calculate the (μ')s efficiently from ks and vice versa.

Let $\mu_j = E\left[\left(X - E(X)\right)^j\right] = E\left[(X - \mu_1')^j\right]$ for $j = 1,2,\dots$ are unknown as **central moments.**

Then formulae (3) simplify to

$\mu_2 = k_2, \mu_3 = k_3,\ \mu_4 = k_4 + 3{k_2}^2$and $k_2 = \mu_2\ \ , k_3 = \mu_3\ , k_4 = \mu_4 - 3{\mu_2}^2$

Note: Mean $= \boldsymbol{\mu} = \boldsymbol{k_1}$ **and variance** $= \boldsymbol{\mu_2} = \boldsymbol{\sigma^2} = \boldsymbol{k_2}$**.**