Chapter 6

Moment Generating Functions

6.1 Definition and Properties

Our previous discussion of *probability generating functions* was in the context of discrete r.v.s. Now we introduce a more general form of generating function which can be used (though not exclusively so) for continuous r.v.s.

The *moment generating function* (MGF) of a random variable X is defined as

$$M_X(\theta) = \mathrm{E}(e^{\theta X}) = \begin{cases} \sum_x e^{\theta x}\mathrm{P}(X = x) & \text{if } X \text{ is discrete} \\ \int_{-\infty}^{\infty} e^{\theta x} f_X(x)dx & \text{if } X \text{ is continuous} \end{cases} \tag{6.1}$$

for all real θ for which the sum or integral converges absolutely. In some cases the existence of $M_X(\theta)$ can be a problem for non-zero θ: henceforth we assume that $M_X(\theta)$ exists in some neighbourhood of the origin, $|\theta| < \theta_0$. In this case the following can be proved:

(i) There is a *unique* distribution with MGF $M_X(\theta)$.

(ii) Moments about the origin may be found by power series expansion: thus we may write

$$\begin{aligned} M_X(\theta) &= \mathrm{E}(e^{\theta X}) \\ &= \mathrm{E}\left(\sum_{r=0}^{\infty} \frac{(\theta X)^r}{r!}\right) \\ &= \sum_{r=0}^{\infty} \frac{\theta^r}{r!}\mathrm{E}(X^r) \quad [\text{i.e. interchange of E and } \textstyle\sum \text{ valid}] \end{aligned}$$

i.e.

$$M_X(\theta) = \sum_{r=0}^{\infty} \mu'_r \frac{\theta^r}{r!} \qquad \text{where } \mu'_r = \mathrm{E}(X^r). \tag{6.2}$$

So, given a function which is known to be the MGF of a r.v. X, expansion of this function in a power series of θ gives μ'_r, the r^{th} moment about the origin, as the coefficient of $\theta^r/r!$.

(iii) Moments about the origin may also be found by differentiation: thus

$$\begin{aligned} \frac{d^r}{d\theta^r}\{M_X(\theta)\} &= \frac{d^r}{d\theta^r}\left\{\mathrm{E}(e^{\theta X})\right\} \\ &= \mathrm{E}\left\{\frac{d^r}{d\theta^r}(e^{\theta X})\right\} \\ &\qquad\qquad (\text{i.e. interchange of E and differentiation valid}) \\ &= \mathrm{E}\left(X^r e^{\theta X}\right). \end{aligned}$$

So

$$\left[\frac{d^r}{d\theta^r}\{M_X(\theta)\}\right]_{\theta=0} = \mathrm{E}(X^r) = \mu'_r. \tag{6.3}$$

(iv) If we require moments about the mean, $\mu_r = \mathrm{E}[(X-\mu)^r]$, we consider $M_{X-\mu}(\theta)$, which can be obtained from $M_X(\theta)$ as follows:

$$M_{X-\mu}(\theta) = \mathrm{E}\left(e^{\theta(X-\mu)}\right) = e^{-\mu\theta}\mathrm{E}(e^{\theta X}) = e^{-\mu\theta}M_X(\theta). \tag{6.4}$$

Then μ_r can be obtained as the coefficient of $\frac{\theta^r}{r!}$ in the expansion

$$M_{X-\mu}(\theta) = \sum_{r=0}^{\infty} \mu_r \frac{\theta^r}{r!} \tag{6.5}$$

or by differentiation:

$$\mu_r = \left[\frac{d^r}{d\theta^r}\{M_{X-\mu}(\theta)\}\right]_{\theta=0}. \tag{6.6}$$

(v) More generally:

$$M_{a+bX}(\theta) = \mathrm{E}\left(e^{\theta(a+bX)}\right) = e^{a\theta}M_X(b\theta). \tag{6.7}$$

Example

Find the MGF of the $N(0,1)$ distribution and hence of $N(\mu,\sigma^2)$. Find the moments about the mean of $N(\mu,\sigma^2)$.

Solution If $Z \sim N(0,1)$,

$$\begin{aligned} M_Z(\theta) &= \mathrm{E}(e^{\theta Z}) \\ &= \int_{-\infty}^{\infty} e^{\theta z}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}dz \\ &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \exp\{-\frac{1}{2}(z^2-2\theta z+\theta^2)+\frac{1}{2}\theta^2\}dz \\ &= \exp(\tfrac{1}{2}\theta^2)\frac{1}{\sqrt{2\pi}}\int_{\infty}^{\infty} \exp\{-\frac{1}{2}(z-\theta)^2\}dz. \end{aligned}$$

But here $\frac{1}{\sqrt{2\pi}}\exp\{...\}$ is the p.d.f. of $N(\theta,1)$), so

$$M_Z(\theta) = \exp(\frac{1}{2}\theta^2) \tag{6.8}$$

If $X = \mu+\sigma Z, X \sim N(\mu,\sigma^2)$, and

$$\begin{aligned} M_X(\theta) &= M_{\mu+\sigma Z}(\theta) \\ &= e^{\mu\theta}M_Z(\sigma\theta) \quad \text{by (6.7)} \\ &= \exp(\mu\theta + \tfrac{1}{2}\sigma^2\theta^2). \end{aligned}$$

Then

$$\begin{aligned} M_{X-\mu}(\theta) &= e^{-\mu\theta}M_X(\theta) = \exp(\tfrac{1}{2}\sigma^2\theta^2) \\ &= \sum_{r=0}^{\infty}\frac{(\frac{1}{2}\sigma^2\theta^2)^r}{r!} = \sum_{r=0}^{\infty}\frac{\sigma^{2r}}{2^r r!}\theta^{2r} \\ &= \sum_{r=0}^{\infty}\frac{\sigma^{2r}}{2^r}.\frac{(2r)!}{r!}.\frac{\theta^{2r}}{(2r)!}. \end{aligned}$$

Using property (iv) above, we obtain

$$\begin{aligned} \mu_{2r+1} &= 0, \quad r = 1, 2, ... \\ \mu_{2r} &= \frac{\sigma^{2r}(2r)!}{2^r r!}, \quad r = 0, 1, 2, ... \end{aligned} \tag{6.9}$$

e.g. $\mu_2 = \sigma^2; \quad \mu_4 = 3\sigma^4.$ ♢

6.2 Sum of independent variables

Theorem

Let X, Y be independent r.v.s with MGFs $M_X(\theta), M_Y(\theta)$ respectively. Then

$$M_{X+Y}(\theta) = M_X(\theta).M_Y(\theta). \tag{6.10}$$

Proof

$$\begin{aligned} M_{X+Y}(\theta) &= \mathrm{E}\left(e^{\theta(X+Y)}\right) \\ &= \mathrm{E}\left(e^{\theta X}.e^{\theta Y}\right) \\ &= \mathrm{E}(e^{\theta X}).E(e^{\theta Y}) \quad \text{[independence]} \\ &= M_X(\theta).M_Y(\theta). \end{aligned}$$

Corollary If $X_1, X_2, ..., X_n$ are independent r.v.s,

$$M_{X_1+X_2+\cdots+X_n}(\theta) = M_{X_1}(\theta).M_{X_2}(\theta)...M_{X_n}(\theta). \tag{6.11}$$

Note: If X is a *count* r.v. with PGF $G_X(s)$ and MGF $M_X(\theta)$,

$$M_X(\theta) = G_X(e^\theta): \quad G_X(s) = M_X(\log s). \tag{6.12}$$

Here the PGF is generally preferred, so we shall concentrate on the MGF applied to *continuous* r.v.s.

Example

Let $Z_1, ..., Z_n$ be independent $N(0, 1)$ r.v.s. Show that

$$V = Z_1^2 + \cdots + Z_n^2 \sim \chi_n^2. \tag{6.13}$$

Solution Let $Z \sim N(0, 1)$. Then

$$\begin{aligned} M_{Z^2}(\theta) = \mathrm{E}\left(e^{\theta Z^2}\right) &= \int_{-\infty}^{\infty} e^{\theta z^2}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}dz \\ &= \int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}\exp\{-\frac{1}{2}(1-2\theta)z^2\}dz. \end{aligned}$$

Assuming $\theta < \frac{1}{2}$, substitute $y = \sqrt{1-2\theta}z$. Then

$$M_{Z^2}(\theta) = \int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}y^2}.\frac{1}{\sqrt{1-2\theta}}dy = (1-2\theta)^{-\frac{1}{2}}, \quad \theta < \frac{1}{2}. \tag{6.14}$$

Hence

$$\begin{aligned} M_V(\theta) &= (1-2\theta)^{-\frac{1}{2}}.(1-2\theta)^{-\frac{1}{2}}...(1-2\theta)^{-\frac{1}{2}} \\ &= (1-2\theta)^{-n/2}, \quad \theta < \tfrac{1}{2}. \end{aligned}$$

Now χ_n^2 has the p.d.f.

$$\frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} w^{\frac{n}{2}-1} e^{-\frac{1}{2}w}, \quad w \geq 0; n \text{ a positive integer.}$$

Its MGF is

$$\begin{aligned} & \int_0^\infty e^{\theta w} \frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} w^{\frac{n}{2}-1} e^{-\frac{1}{2}w} dw \\ &= \int_0^\infty \frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} w^{\frac{n}{2}-1} \exp\{-\frac{1}{2}w(1-2\theta)\} dw \\ & \qquad\qquad (t = \tfrac{1}{2}(1-2\theta) \quad (\theta < \tfrac{1}{2})) \\ &= (1-2\theta)^{-\frac{n}{2}} \frac{1}{\Gamma(\frac{n}{2})} \int_0^\infty t^{\frac{n}{2}-1} e^{-t} dt \\ &= (1-2\theta)^{-\frac{n}{2}}, \qquad \theta < \tfrac{1}{2} \\ &= M_V(\theta). \end{aligned}$$

So we deduce that $V \sim \chi_n^2$. Also, from $M_{Z^2}(\theta)$ we deduce that $Z^2 \sim \chi_1^2$.

If $V_1 \sim \chi_{n_1}^2, V_2 \sim \chi_{n_2}^2$ and V_1, V_2 are independent, then

$$\begin{aligned} M_{V_1+V_2}(\theta) = M_{V_1}(\theta).M_{V_2}(\theta) &= (1-2\theta)^{-\frac{n_1}{2}}(1-2\theta)^{-\frac{n_2}{2}} \qquad (\theta < \tfrac{1}{2}) \\ &= (1-2\theta)^{-(n_1+n_2)/2}. \end{aligned}$$

So $V_1 + V_2 \sim \chi_{n_1+n_2}^2$. [This was also shown in Example 3, §5.8.2.]

6.3 Bivariate MGF

The bivariate MGF (or *joint* MGF) of the continuous r.v.s (X, Y) with joint p.d.f. $f(x, y), \quad -\infty < x, y < \infty$ is defined as

$$M_{X,Y}(\theta_1, \theta_2) = \mathrm{E}\left(e^{\theta_1 X + \theta_2 Y}\right) = \int_{-\infty}^\infty \int_{-\infty}^\infty e^{\theta_1 x + \theta_2 y} f(x, y) dx dy, \tag{6.15}$$

provided the integral converges absolutely (there is a similar definition for the discrete case). If $M_{X,Y}(\theta_1, \theta_2)$ exists near the origin, for $|\theta_1| < \theta_{10}, |\theta_2| < \theta_{20}$ say, then it can be shown that

$$\left[\frac{\partial^{r+s} M_{X,Y}(\theta_1, \theta_2)}{\partial\theta_1^r \partial\theta_2^s}\right]_{\theta_1=\theta_2=0} = \mathrm{E}(X^r Y^s). \tag{6.16}$$

The bivariate MGF can also be used to find the MGF of $aX + bY$, since

$$M_{aX+bY}(\theta) = \mathrm{E}\left(e^{(aX+bY)\theta}\right) = \mathrm{E}\left(e^{(a\theta)X + (b\theta)Y}\right) = M_{X+Y}(a\theta, b\theta). \tag{6.17}$$

Example *Bivariate Normal distribution*

Using MGFs:

(i) show that if $(U,V) \sim N(0,0;1,1;\rho)$, then $\rho(U,V) = \rho$, and deduce $\rho(X,Y)$, where $(X,Y) \sim N(\mu_x,\mu_y;\sigma_x^2,\sigma_y^2;\rho)$;

(ii) for the variables (X,Y) in (i), find the distribution of a linear combination $aX+bY$, and generalise the result obtained to the multivariate Normal case.

Solution

(i) We have

$$\begin{aligned} M_{U,V}(\theta_1,\theta_2) &= \mathrm{E}(e^{\theta_1 U+\theta_2 V}) \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{\theta_1 u+\theta_2 v}\frac{1}{2\pi\sqrt{1-\rho^2}}\exp\left\{-\frac{1}{2(1-\rho^2)}[u^2-2\rho uv+v^2]\right\}dudv \\ &= \frac{1}{2\pi\sqrt{1-\rho^2}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\exp\{......\}dudv \\ &= .... = \exp\{\tfrac{1}{2}(\theta_1^2+2\rho\theta_1\theta_2+\theta_2^2)\}. \end{aligned}$$

Then

$$\begin{aligned} \frac{\partial M_{U,V}(\theta_1,\theta_2)}{\partial\theta_1} &= \exp\{.....\}(\theta_1+\rho\theta_2) \\ \frac{\partial^2 M_{U,V}(\theta_1,\theta_2)}{\partial\theta_1\partial\theta_2} &= \exp\{....\}(\rho\theta_1+\theta_2)(\theta_1+\rho\theta_2)+\exp\{....\}\rho. \end{aligned}$$

So

$$\mathrm{E}(UV) = \left[\frac{\partial^2 M_{U,V}(\theta_1,\theta_2)}{\partial\theta_1\partial\theta_2}\right]_{\theta_1=\theta_2=0} = \rho.$$

Since $\mathrm{E}(U)=\mathrm{E}(V)=0$ and $\mathrm{Var}(U)=\mathrm{Var}(V)=1$, we have that the correlation coefficient of U,V is

$$\rho(U,V) = \frac{\mathrm{Cov}(U,V)}{\sqrt{\mathrm{Var}(U).\mathrm{Var}(V)}} = \frac{\mathrm{E}(UV)-\mathrm{E}(U)\mathrm{E}(V)}{1} = \rho.$$

Now let

$$X=\mu_x+\sigma_x U, \quad Y=\mu_y+\sigma_y V.$$

Then, as we have seen in Example 1, §5.8.2,

$$(U,V)\sim N(0,0;1,1;\rho) \Longleftrightarrow (X,Y)\sim N(\mu_x,\mu_y;\sigma_x^2,\sigma_y^2;\rho).$$

It is readily shown that a correlation coefficient remains unchanged under a linear transformation of variables, so $\rho(X,Y)=\rho(U,V)=\rho$.

(ii) We have that

$$\begin{aligned} M_{X,Y}(\theta_1,\theta_2) &= \mathrm{E}\left[e^{\theta_1(\mu_x+\sigma_x U)+\theta_2(\mu_y+\sigma_y V)}\right] \\ &= e^{(\theta_1\mu_x+\theta_2\mu_y)}M_{U,V}(\theta_1\sigma_x,\theta_2\sigma_y) \\ &= \exp\{(\theta_1\mu_x+\theta_2\mu_y)+\tfrac{1}{2}(\theta_1^2\sigma_x^2+2\theta_1\theta_2\rho\sigma_x\sigma_y+\theta_2^2\sigma_y^2)]. \end{aligned}$$

So, for a linear combination of X and Y,

$$\begin{aligned} M_{aX+bY}(\theta)=M_{X,Y}(a\theta,b\theta) &= \exp\{(a\mu_x+b\mu_y)\theta+\tfrac{1}{2}(a^2\sigma_x^2+2ab\mathrm{Cov}(X,Y)+b^2\sigma_y^2)\theta^2\} \\ &= \text{MGF of } N(a\mu_x+b\mu_y, a^2\sigma_x^2+2ab\mathrm{Cov}(X,Y)+b^2\sigma_y^2)\theta^2), \end{aligned}$$

i.e.

$$aX+bY \sim N(a\mathrm{E}(X)+b\mathrm{E}(Y), a^2\mathrm{Var}(X)+2ab\mathrm{Cov}(X,Y)+b^2\mathrm{Var}(Y)). \qquad (6.18)$$

More generally, let $(X_1, ..., X_n)$ be multivariate normally distributed. Then, by induction,

$$\sum_{i=1}^{n} a_i X_i \sim N\left(\sum_{i=1}^{n} a_i \mathrm{E}(X_i), \sum_{i=1}^{n} a_i^2 \mathrm{Var}(X_i) + 2\sum_{i<j} a_i a_j \mathrm{Cov}(X_i, X_j)\right). \tag{6.19}$$

(If the Xs are also independent, the covariance terms vanish – but then there is a simpler derivation (see HW 8).) ♢

6.4 Sequences of r.v.s

6.4.1 Continuity theorem

First we state (without proof) the following:

Theorem

Let $X_1, X_2, ...$ be a sequence of r.v.s (discrete or continuous) with c.d.f.s $F_{X_1}(x), F_{X_2}(x), ...$ and MGFs $M_{X_1}(\theta), M_{X_2}(\theta), ...$, and suppose that, as $n \to \infty$,

$$M_{X_n}(\theta) \to M_X(\theta) \qquad \text{for all } \theta,$$

where $M_X(\theta)$ is the MGF of some r.v. X with c.d.f. $F_X(x)$. Then

$$F_{X_n}(x) \to F_X(x) \qquad \text{as } n \to \infty$$

at each x where $F_X(x)$ is continuous.

Example

Using MGFs, discuss the limit of $\mathrm{Bin}(n, p)$ as $n \to \infty, p \to 0$ with $np = \lambda > 0$ fixed.

Solution Let $X_n \sim \mathrm{Bin}(n, p)$, with PGF $G_X(s) = (ps + q)^n$. Then

$$M_{X_n}(\theta) = G_{X_n}(e^\theta) = (pe^\theta + q)^n = \{1 + \frac{\lambda}{n}(e^\theta - 1)\}^n \quad \text{where } \lambda = np.$$

Let $n \to \infty, p \to 0$ in such a way that λ remains fixed. Then

$$M_{X_n}(\theta) \to \exp\{\lambda(e^\theta - 1)\} \quad \text{as } n \to \infty,$$

since

$$\left(1 + \frac{a}{n}\right)^n \to e^a \quad \text{as } n \to \infty, a \text{ constant}, \tag{6.20}$$

i.e.

$$M_{X_n}(\theta) \to \text{MGF of Poisson}(\lambda) \tag{6.21}$$

(use (6.12), replacing s by e^θ in the Poisson PGF (3.7)). So, invoking the above continuity theorem,

$$\mathrm{Bin}(n, p) \to \mathrm{Poisson}(\lambda) \tag{6.22}$$

as $n \to \infty, p \to 0$ with $np = \lambda > 0$ fixed. Hence in large samples, the binomial distribution can be approximated by the Poisson distribution. As a rule of thumb: the approximation is acceptable when n is large, p small, and $\lambda = np \leq 5$.

6.4.2 Asymptotic normality

Let $\{X_n\}$ be a sequence of r.v.s (discrete or continuous). If two quantities a and b can be found such that

$$\text{c.d.f. of } \frac{(X_n - a)}{b} \to \text{c.d.f. of } N(0,1) \quad \text{as } n \to \infty, \tag{6.23}$$

X_n is said to be *asymptotically normally distributed* with mean a and variance b^2, and we write

$$\frac{X_n - a}{b} \overset{a}{\sim} N(0,1) \quad \text{or} \quad X_n \overset{a}{\sim} N(a, b^2). \tag{6.24}$$

Notes: (i) a and b need not be functions of n; but often a and b^2 are the mean and variance of X_n (and so are functions of n).

(ii) In large samples we use $N(a, b^2)$ as an approximation to the distribution of X_n.

6.5 Central limit theorem

A restricted form of this celebrated theorem will now be stated and proved.

Theorem

Let $X_1, X_2, ...$ be a sequence of independent identically distributed r.v.s, each with mean μ and variance σ^2. Let

$$S_n = X_1 + X_2 + \cdots + X_n, \qquad Z_n = \frac{(S_n - n\mu)}{\sqrt{n}\sigma}.$$

Then

$$\begin{aligned} Z_n &\overset{a}{\sim} N(0,1) \quad \text{or} \quad \mathrm{P}(Z_n \le z) \to \mathrm{P}(Z \le z) \quad \text{as } n \to \infty, \text{ where } Z \sim N(0,1), \\ \text{and } S_n &\overset{a}{\sim} N(n\mu, n\sigma^2). \end{aligned}$$

Proof Let $Y_i = X_i - \mu \quad (i = 1, 2, ...)$. Then $Y_1, Y_2, ...$ are i.i.d. r.v.s, and

$$S_n - n\mu = X_1 + \cdots + X_n - n\mu = Y_1 + \cdots + Y_n.$$

So

$$M_{S_n - n\mu}(\theta) = M_{Y_1}(\theta).M_{Y_2}(\theta)....M_{Y_n}(\theta) = \{M_Y(\theta)\}^n,$$

and

$$\begin{aligned} M_{Z_n}(\theta) &= M_{\frac{S_n - n\mu}{\sqrt{n}\sigma}}(\theta) = \mathrm{E}\left[\exp\left(\frac{S_n - n\mu}{\sqrt{n}\sigma}\theta\right)\right] \\ &= \mathrm{E}\left[\exp\left\{(S_n - n\mu)\left(\frac{\theta}{\sqrt{n}\sigma}\right)\right\}\right] \\ &= M_{S_n - n\mu}\left(\frac{\theta}{\sqrt{n}\sigma}\right) = \left\{M_Y\left(\frac{\theta}{\sqrt{n}\sigma}\right)\right\}^n. \end{aligned}$$

Note that

$$\mathrm{E}(Y) = \mathrm{E}(X - \mu) = 0: \qquad \mathrm{E}(Y^2) = \mathrm{E}\{(X - \mu)^2\} = \sigma^2.$$

Then

$$\begin{aligned} M_Y(\theta) &= 1 + \mathrm{E}(Y)\frac{\theta}{1!} + \mathrm{E}(Y^2)\frac{\theta^2}{2!} + \mathrm{E}(Y^3)\frac{\theta^3}{3!} + \cdots \\ &= 1 + \tfrac{1}{2}\sigma^2\theta^2 + o(\theta^2) \end{aligned}$$

(where $o(\theta^2)$ denotes a function $g(\theta)$ such that $\frac{g(\theta)}{\theta^2} \to 0$ as $\theta \to 0$). So

$$M_{Z_n}(\theta) = \{1 + \tfrac{1}{2}\sigma^2(\tfrac{\theta^2}{n\sigma^2}) + o(\tfrac{1}{n})\}^n = \{1 + \tfrac{1}{2}\theta^2.\tfrac{1}{n} + o(\tfrac{1}{n})\}^n$$

(where $o(\frac{1}{n})$ denotes a function $h(n)$ such that $\frac{h(n)}{1/n} \to 0$ as $n \to \infty$).

Using the standard result (6.20), we deduce that

$$M_{Z_n}(\theta) \to \exp(\tfrac{1}{2}\theta^2) \quad \text{as } n \to \infty$$

– which is the MGF of N(0,1).

So

$$\text{c.d.f. of } Z_n = \frac{S_n - n\mu}{\sqrt{n}\sigma} \to \text{c.d.f. of } N(0,1) \quad \text{as } n \to \infty,$$

i.e.

$$Z_n \overset{a}{\sim} N(0,1) \quad \text{or} \quad S_n \overset{a}{\sim} N(n\mu, n\sigma^2). \tag{6.25}$$

□

Corollary

Let $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$. Then $\overline{X}_n \overset{a}{\sim} N(\mu, \frac{\sigma^2}{n})$. (6.26)

Proof $\overline{X}_n = W_1 + \cdots + W_n$ where $W_i = \frac{1}{n}X_i$ and $W_1, ..., W_n$ are i.i.d. with mean $\frac{\mu}{n}$ and variance $\frac{\sigma^2}{n^2}$. So

$$\overline{X}_n \overset{a}{\sim} N(n.\frac{\mu}{n}, n.\frac{\sigma^2}{n^2}) = N(\mu, \frac{\sigma^2}{n}).$$ □

(*Note*: The theorem can be generalised to

independent r.v.s with different means & variances
dependent r.v.s

–but extra conditions on the distributions are required.

Example 1

Using the central limit theorem, obtain an approximation to Bin(n, p) for large n.

Solution Let $S_n \sim$ Bin(n, p). Then

$$S_n = X_1 + X_2 + \cdots + X_n,$$

where

$$X_i = \begin{cases} 1, & \text{if the } i\text{th trial yields a success} \\ 0, & \text{if the } i\text{th trial yields a failure.} \end{cases}$$

Also, $X_1, X_2, ..., X_n$ are independent r.v.s with

$$\text{E}(X_i) = p, \qquad \text{Var}(X_i) = pq.$$

So

$$S_n \overset{a}{\sim} N(np, npq),$$

i.e., for large n, the binomial c.d.f. is approximated by the c.d.f. of $N(np, npq)$. □

[As a rule of thumb: the approximation is acceptable when n is large and $p \leq \frac{1}{2}$ such that $np > 5$.]

Example 2

As Example 1, but for the χ^2_n distribution.

Solution Let $V_n \sim \chi^2_n$. Then we can write

$$V_n = Z_1^2 + \cdots + Z_n^2,$$

where $Z_1^2, ..., Z_n^2$ are independent r.v.s and

$$Z_i \sim N(0,1), \quad Z_i^2 \sim \chi^2_1; \quad \mathrm{E}(Z_i^2) = 1, \quad \mathrm{Var}(Z_i^2) = 2.$$

So

$$V_n \overset{a}{\sim} N(n, 2n).$$ □

Note: These are not necessarily the 'best' approximations for large n. Thus

(i)

$$\begin{aligned} \mathrm{P}(S_n \le s) &\approx \mathrm{P}\left(Z \le \tfrac{s+\frac{1}{2}-np}{\sqrt{npq}}\right) \quad \text{where } Z \sim N(0,1) \\ &= F_S\left(\tfrac{s+\frac{1}{2}-np}{\sqrt{npq}}\right). \end{aligned}$$

The $\frac{1}{2}$ is a 'continuity correction', to take account of the fact that we are approximating a discrete distribution by a continuous one.

(ii)

$$\sqrt{2V_n} \overset{approx}{\sim} N(\sqrt{2n-1}, 1).$$

6.6 Characteristic function

The MGF does not exist unless all the moments of the distribution are finite. So many distributions (e.g. t,F) do not have MGFs. So another GF is often used.

The *characteristic function* of a continuous r.v. X is

$$C_X(\theta) = \mathrm{E}(e^{i\theta X}) = \int_{-\infty}^{\infty} e^{i\theta x} f(x) dx, \tag{6.27}$$

where θ is real and $i = \sqrt{-1}$. $C_X(\theta)$ *always* exists, and has similar properties to $M_X(\theta)$. The CF uniquely determines the p.d.f.:

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} C_X(\theta) e^{-ix\theta} d\theta \tag{6.28}$$

(cf. Fourier transform). The CF is particularly useful in studying limiting distributions. However, we do not consider the CF further in this module.