# Generating and Characteristic Functions

February 4, 2022



## Generating and characteristic functions





## Probability generating function

**Definition**

*Let $X$ be a nonnegative integer-valued random variable. The probability generating function of $X$ is defined to be*

$$G_X(s) \equiv \mathbb{E}(s^X) = \sum_{k \geqslant 0} s^k \, \mathbb{P}(X = k)$$

The random variable $s^X$ is a function of the random variable $X$. For each valid value of $s$, where $s \in \mathbb{R}$, we compute the expectation of $s^X$. In this way, we get a one-variable function of $s$.



## Probability generating function

If $X$ takes a finite number of values, then $G_X(s)$ is just a polynomial of the indeterminate $s$:

$$\begin{aligned} G_X(s) &= \sum_{k=0}^{n} s^k \, \mathbb{P}(X = k) \\ &= \mathbb{P}(X = 0) + \mathbb{P}(X = 1)\, s + \cdots + \mathbb{P}(X = n)\, s^n \end{aligned}$$

## Probability generating function

Otherwise, if $X$ takes a countable number of values, then $G_X(s)$ is given by a power series:

$$G_X(s) = \sum_{k \geqslant 0} s^k \, \mathbb{P}(X = k)$$
$$= \mathbb{P}(X = 0) + \mathbb{P}(X = 1)\, s + \cdots + \mathbb{P}(X = k)\, s^k + \cdots$$

The series defining $G_X(x)$ converges at least for all $s \in [-1, 1]$, because if $|s| \leqslant 1$, then

$$\sum_{k \geqslant 0} |s|^k \, \mathbb{P}(X = k) \leqslant \sum_{k \geqslant 0} \mathbb{P}(X = k) = 1$$



## Probability generating function

More generally, there exists a radius of convergence $R$, $1 \leqslant R \leqslant \infty$, such that

$$\sum_{k \geqslant 0} s^k \, \mathbb{P}(X = k)$$

converges absolutely if $|s| < R$ and diverges if $|s| > R$.

- The probability generating function $G_X(s)$ is well-defined for all $s \in [-R, R]$.



## Examples

Let $X$ be a Bernoulli random variable, $X \sim \text{Be}(p)$, such that

$$\mathbb{P}(X = 0) = q, \qquad \mathbb{P}(X = 1) = p,$$

where $q = 1 - p$.

Then

$$G_X(s) = \sum_{k \geqslant 0} s^k \, \mathbb{P}(X = k) = q + p\, s, \quad s \in \mathbb{R}$$



## Examples

If $X \sim \text{Bin}(n, p)$, then

$$\mathbb{P}(X = k) = \binom{n}{k} p^k q^{n-k}, \quad k = 0, 1, \ldots, n$$

Hence

$$G_X(s) = \sum_{k \geqslant 0} s^k \, \mathbb{P}(X = k) = \sum_{k=0}^{n} s^k \binom{n}{k} p^k q^{n-k}$$
$$= \sum_{k=0}^{n} \binom{n}{k} (ps)^k q^{n-k} = (q + ps)^n, \quad s \in \mathbb{R}$$

## Examples

Let $X \sim \text{Po}(\lambda)$.

$$\mathbb{P}(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}, \quad k = 0, 1, 2 \ldots$$

Then

$$G_X(s) = \sum_{k \geqslant 0} s^k \, \mathbb{P}(X = k) = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda s)^k}{k!}$$
$$= e^{-\lambda} e^{\lambda s} = e^{\lambda(s-1)}, \quad s \in \mathbb{R}$$



## Examples

Let $X \sim \text{Ge}(p)$.

$$\mathbb{P}(X = k) = q^{k-1} p, \quad k = 1, 2, \ldots, \quad 0 < p < 1$$

Therefore

$$G_X(s) = \sum_{k \geqslant 0} s^k \, \mathbb{P}(X = k) = \sum_{k=1}^{\infty} s^k q^{k-1} p$$
$$= ps \sum_{k=1}^{\infty} (qs)^{k-1} = \frac{ps}{1 - qs}, \quad |s| < \frac{1}{q}$$



## Unicity

If two nonnegative integer-valued random variables have the same generating function, then they follow the same probability law.

**Theorem**

*Let $X$ and $Y$ be nonnegative integer-valued random variables such that*

$$G_X(s) = G_Y(s).$$

*Then*

$$\mathbb{P}(X = k) = \mathbb{P}(Y = k) \quad \textit{for all } k \geqslant 0.$$

This result is a special case of the uniqueness theorem for power series.



## Convolution theorem

**Theorem (convolution)**

*Let $X$ and $Y$ be independent, nonnegative, integer-valued random variables, and let $Z = X + Y$. Then*

$$G_Z(s) = G_X(s) \, G_Y(s)$$

Proof: Since $X$ and $Y$ are independent, the random variables $s^X$ and $s^Y$ are also independent. Therefore,

$$G_Z(s) = \mathbb{E}\left(s^Z\right) = \mathbb{E}\left(s^{X+Y}\right)$$
$$= \mathbb{E}\left(s^X s^Y\right) = \mathbb{E}\left(s^X\right) \mathbb{E}\left(s^Y\right) = G_X(s) G_Y(s)$$

## Example

Let $X \sim \text{Bin}(n, p)$ and $Y \sim \text{Bin}(m, p)$ be independent and let

$$Z = X + Y$$

We have

$$G_Z(s) = G_X(s)G_Y(s) = (q + ps)^n(q + ps)^m = (q + ps)^{n+m}$$

Since $G_Z(s)$ is the probability generating function of a $\text{Bin}(n + m, p)$ random variable, we deduce from the unicity theorem that

$$X + Y \sim \text{Bin}(n + m, p)$$



## Convolution theorem

More generally,

**Theorem**

*Let $X_1, X_2, \dots, X_n$ be independent, nonnegative, integer-valued random variables and set $S = X_1 + X_2 + \cdots + X_n$. Then*

$$G_S(s) = G_{X_1}(s)G_{X_2}(s)\cdots G_{X_n}(s).$$



## Convolution theorem

A case of particular importance is:

**Corollary**

*If, in addition, $X_1, X_2, \dots, X_n$ are equidistributed, with common probability generating function $G_X(s)$, then $G_S(s) = (G_X(s))^n$.*

Example: If $X_1, X_2, \dots, X_n$ are independent $\text{Be}(p)$-distributed random variables, then $S = X_1 + \cdots + X_n \sim \text{Bin}(n, p)$ and

$$G_S(s) = (G_X(s))^n = (q + sp)^n$$



## Example: Negative binomial

A coin for which the probability of landing on heads is $p$ is flipped until a total amount of $k$ heads is obtained.

Let $X$ be the number of tosses.

Notice that

$$X = X_1 + X_2 + \cdots + X_k,$$

where $X_i$ is the number of tosses between the $(i-1)$-th and the $i$-th toss showing heads, an so

$$X_i \sim \text{Ge}(p), \quad 1 \leqslant i \leqslant k,$$

and the variables $X_1, \dots, X_n$ are independent.

## Example: Negative binomial

Since the variables $X_1, \dots, X_k$ are independent, we can apply the convolution theorem. Moreover, since $X_1, \dots, X_k$ are identically distributed, we have

$$G_X(s) = G_{X_1}(s)G_{X_2}(s)\cdots G_{X_k}(s)$$
$$= (G_{X_1}(s))^k = \left(\frac{ps}{1-qs}\right)^k, \quad |s| < \frac{1}{q}$$

If we expand $G_X(s)$ as a power series, $G_X(s) = \sum_{k\geqslant 0} a_k s^k$, then

$$a_k = \mathbb{P}(X = k)$$



## Example: Negative binomial

Recall that if $\alpha \in \mathbb{R}$, then the Taylor series expansion about 0 of the function $(1+x)^\alpha$ is

$$\begin{aligned}(1+x)^\alpha &= 1 + \alpha x + \frac{\alpha(\alpha-1)}{2}x^2 + \cdots \\ &\quad + \frac{\alpha(\alpha-1)\dots(\alpha-r+1)}{r!}x^r + \cdots \\ &= \sum_{r\geqslant 0}\binom{\alpha}{r}x^r, \quad x \in (-1,1),\end{aligned}$$

where $\binom{\alpha}{0} \equiv 1$ and, for $r \geqslant 1$,

$$\binom{\alpha}{r} \equiv \frac{\alpha(\alpha-1)\dots(\alpha-r+1)}{r!}$$



## Example: Negative binomial

By identifying $x$ with $-sq$ and $\alpha$ with $-k$ we obtain the series expansion of $G_X(s)$:

$$G_X(s) = (ps)^k(1-qs)^{-k} = (ps)^k\sum_{r=0}^{\infty}\binom{-k}{r}(-qs)^r,$$

where

$$\binom{-k}{r} = \frac{-k(-k-1)\cdots(-k-r+1)}{r!}$$
$$= (-1)^r\binom{k+r-1}{k-1}, \quad r \geqslant 0$$



## Example: Negative binomial

Therefore,

$$G_X(s) = \sum_{r=0}^{\infty}\binom{k+r-1}{k-1}p^k q^r s^{k+r} = \sum_{n=k}^{\infty}\binom{n-1}{k-1}p^k q^{n-k} s^n$$

Hence,

$$\mathbb{P}(X=n) = \begin{cases} 0, & n < k \\ \binom{n-1}{k-1}p^k q^{n-k}, & n = k, k+1, \cdots \end{cases}$$

This is the negative binomial probability law, $X \sim \text{NBin}(k,p)$.

## Some properties

- $G_X(0) = \mathbb{P}(X = 0)$
- $G_X(1) = 1$

Indeed,

$$G_X(1) = \sum_{k \geqslant 0} s^k \mathbb{P}(X = k)\Big|_{s=1} = \sum_{k \geqslant 0} \mathbb{P}(X = k) = 1$$



## Some properties

**Proposition**

*Let $R$ be the radius of covergence of $G_X(s)$. If $R > 1$, then*

$$\mathbb{E}(X) = G_X'(1)$$

Indeed, differentiating the series term by term,

$$G_X'(s) = \frac{d}{ds} \sum_{k \geqslant 0} s^k \mathbb{P}(X = k) = \sum_{k \geqslant 1} k\, s^{k-1} \mathbb{P}(X = k)$$

Hence,

$$G_X'(1) = \sum_{k \geqslant 0} k\, \mathbb{P}(X = k) = \mathbb{E}(X)$$



## Some properties

More generally,

**Proposition**

(a) $\mathbb{E}(X) = G_X'(1) \equiv \lim_{s \to 1^-} G_X'(s)$

(b) $\mathbb{E}(X(X-1)\cdots(X-k+1)) = G_X^{(k)}(1) \equiv \lim_{s \to 1^-} G_X^{(k)}(s)$

If the radius of convergence is $R = 1$, then $\lim_{s \to 1^-} G_X^{(k)}(s)$ could be $\infty$.



## Examples

Let $X \sim \mathrm{Bin}(n, p)$.

$$\mathbb{E}(X) = G_X'(1) = \frac{d}{ds}(q + ps)^n\Big|_{s=1}$$

$$= np\,(q + ps)^{n-1}\big|_{s=1} = np\,(q + p)^{n-1} = np$$

## Examples

Let $X \sim \text{Po}(\lambda)$.

$$\mathbb{E}(X) = G_X'(1) = \frac{d}{ds}\, e^{\lambda(s-1)}\Big|_{s=1} = \lambda\, e^{\lambda(s-1)}\Big|_{s=1} = \lambda$$

Analogously,

$$\mathbb{E}(X(X-1)) = G_X''(1) = \lambda^2\, e^{\lambda(s-1)}\Big|_{s=1} = \lambda^2$$

Hence,

$$E(X^2) = \lambda^2 + \lambda, \qquad \text{Var}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 = \lambda$$



## Examples

If $X \sim \text{Ge}(p)$, then

$$\mathbb{E}(X) = G_X'(1) = \frac{d}{ds}\,\frac{ps}{(1-qs)}\Big|_{s=1} = \frac{p}{(1-qs)^2}\Big|_{s=1} = \frac{1}{p}$$

Analogously,

$$\mathbb{E}(X(X-1)) = G_X''(1) = \frac{2pq}{(1-qs)^3}\Big|_{s=1} = \frac{2q}{p^2}$$

Therefore

$$E(X^2) = \frac{2q}{p^2} + \frac{1}{p}, \qquad \text{Var}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 = \frac{q}{p^2}$$



## Examples

$X \sim \text{NBin}(k, p)$.

$$\mathbb{E}(X) = G_X'(1) = \frac{d}{ds}\left(\frac{ps}{1-qs}\right)^k\Bigg|_{s=1}$$

$$= k\left(\frac{ps}{1-qs}\right)^{k-1}\frac{p}{(1-qs)^2}\Bigg|_{s=1} = \frac{k}{p}$$

Notice that this result can also be obtained from $X = \sum_{i=1}^k X_i$, with each $X_i \sim \text{Ge}(p)$, for then

$$\mathbb{E}(X) = \sum_{i=1}^{k}\mathbb{E}(X_i) = \frac{k}{p}$$

## Moment generating function

**Definition**

*The moment generating function of a random variable $X$ is defined by*

$$\Phi_X(t) \equiv \mathbb{E}\left(e^{tX}\right)$$

*for all values of $t$, $t \in \mathbb{R}$, for which this expectation exists.*

- If $X$ takes values in $\{0, 1, 2, \dots\}$, then

$$\Phi_X(t) = \mathbb{E}\left(e^{tX}\right) = G_X(e^t).$$



## Moment generating function

*For more general random variables we have*

$$\Phi_X(t) = \begin{cases} \sum_i e^{tx_i}\,\mathbb{P}(X = x_i), & \text{if } X \text{ is discrete} \\ \int_{-\infty}^{\infty} e^{tx}\, f_X(x)\, dx, & \text{if } X \text{ is continuous} \end{cases},$$

*provided that the sum or the integral converges.*



## Examples

Let $X \sim \text{Bin}(n, p)$.

$$\Phi_X(t) = \sum_{k=0}^{n} e^{tk}\,\mathbb{P}(X = k)$$

$$= \sum_{k=0}^{n} \binom{n}{k} \left(pe^t\right)^k q^{n-k} = \left(q + pe^t\right)^n, \quad t \in \mathbb{R}$$



## Examples

Let $X \sim \text{Po}(\lambda)$.

$$\Phi_X(t) = \sum_{k=0}^{\infty} e^{tk}\,\mathbb{P}(X = k) = \sum_{k=0}^{\infty} e^{tk}\, e^{-\lambda} \frac{\lambda^k}{k!}$$

$$= e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda e^t)^k}{k!} = e^{\lambda(e^t - 1)}, \quad t \in \mathbb{R}$$

- Notice that in the two last examples $\Phi_X(t) = G_X(e^t)$.

## Examples

Let $X \sim \text{Exp}(\mu)$.

$$\Phi_X(t) = \int_{-\infty}^{\infty} e^{tx} f_X(x)\, dx$$
$$= \int_0^{\infty} \mu e^{-(\mu - t)x}\, dx = \frac{\mu}{\mu - t}, \quad t < \mu$$

For a continuos random variable, $\Phi_X(t)$ is related to the Laplace transform of the probability density function $f_X(x)$.



## Examples

Let $Z \sim \text{N}(0,1)$. Then

$$\Phi_Z(t) = \int_{-\infty}^{\infty} e^{tz} f_Z(z)\, dz = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{z^2 - 2tz}{2}}\, dz$$
$$= e^{\frac{t^2}{2}} \underbrace{\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(z-t)^2}{2}}\, dz}_{1} = e^{\frac{t^2}{2}}, \quad t \in \mathbb{R}$$



## Examples

More generally, if $Z \sim \text{N}(0,1)$ and

$$X = \sigma Z + m,$$

then $X \sim \text{N}(m, \sigma^2)$ and

$$\Phi_X(t) = \mathbb{E}\left(e^{tX}\right) = \mathbb{E}\left(e^{t(\sigma Z + m)}\right)$$
$$= e^{tm}\, \mathbb{E}\left(e^{t\sigma Z}\right) = e^{tm}\, \Phi_Z(\sigma t) = e^{\frac{\sigma^2 t^2}{2} + tm}$$



## Unicity

The moment generating function specifies uniquely the probability distribution.

**Theorem**

*Let $X$ and $Y$ be random variables. If there exists $h > 0$ such that*

$$\Phi_X(t) = \Phi_Y(t) \quad \text{for } |t| < h,$$

*then $X$ and $Y$ are identically distributed.*

## Power series expansion

By the following non-rigorous argument one has

$$\Phi'_X(t) = \frac{d}{dt}\mathbb{E}\left(e^{tX}\right) = \mathbb{E}\left(\frac{d}{dt}e^{tX}\right) = \mathbb{E}\left(X\,e^{tX}\right)$$

and therefore

$$\Phi'_X(0) = \mathbb{E}(X)$$

Analogously,

$$\Phi''_X(t) = \frac{d}{dt}\Phi'_X(t) = \frac{d}{dt}\mathbb{E}\left(X\,e^{tX}\right) = \mathbb{E}\left(X^2 e^{tX}\right)$$

and so

$$\Phi''_X(0) = \mathbb{E}(X^2)$$



## Power series expansion

For instance, if $X \sim \text{Exp}(\mu)$,

$$\Phi_X(t) = \frac{\mu}{\mu - t}, \quad t < \mu,$$

then

$$\mathbb{E}(X) = \Phi'_X(0) = \frac{d}{dt}\left(\frac{\mu}{\mu - t}\right)\bigg|_{t=0} = \frac{\mu}{(\mu - t)^2}\bigg|_{t=0} = \frac{1}{\mu}$$

Analogously,

$$\mathbb{E}(X^2) = \Phi''_X(0) = \frac{d}{dt}\left(\frac{\mu}{(\mu - t)^2}\right)\bigg|_{t=0} = \frac{2\mu}{(\mu - t)^3}\bigg|_{t=0} = \frac{2}{\mu^2}$$



## Power series expansion

More generally,

$$\Phi_X(t) = \mathbb{E}\left(e^{tX}\right) = \mathbb{E}\left(1 + t\,X + \frac{(t\,X)^2}{2!} + \cdots + \frac{(t\,X)^k}{k!} + \cdots\right)$$

$$= 1 + \mathbb{E}(X)\;t + \frac{\mathbb{E}(X^2)}{2!}\;t^2 + \cdots + \frac{\mathbb{E}(X^k)}{k!}\;t^k + \cdots$$

This is the Taylor's series expansion of $\Phi_X(t)$,

$$\Phi_X(t) = \sum_{k=0}^{\infty} \frac{\Phi_X^{(k)}(0)}{k!}\;t^k$$

Hence

$$\Phi_X^{(k)}(0) = \mathbb{E}\left(X^k\right)$$



## Power series expansion

**Theorem**

*If $\Phi_X(t)$ converges on some open interval containing the origin $t = 0$, then $X$ has moments of any order,*

$$\mathbb{E}\left(X^k\right) = \Phi_X^{(k)}(0),$$

*and*

$$\Phi_X(t) = \sum_{k=0}^{\infty} \frac{\mathbb{E}(X^k)}{k!}\;t^k$$

## Power series expansion

For instance, let $X \sim \text{Exp}(\mu)$. If $|t| < \mu$, then

$$\Phi_X(t) = \frac{\mu}{\mu - t} = \frac{1}{1-(t/\mu)}$$
$$= 1 + \frac{t}{\mu} + \left(\frac{t}{\mu}\right)^2 + \cdots + \left(\frac{t}{\mu}\right)^n + \cdots$$

Hence,

$$\frac{\mathbb{E}(X^n)}{n!} = \frac{1}{\mu^n}$$

and

$$\mathbb{E}(X^n) = \frac{n!}{\mu^n}$$



## Convolution theorem

The convolution theorem applies also to moment generating functions.

**Theorem**

*Let $X_1, X_2, \ldots, X_n$ be independent random variables and let*

$$S = X_1 + X_2 + \cdots + X_n.$$

*Then,*

$$\Phi_S(t) = \Phi_{X_1}(t)\Phi_{X_2}(t)\cdots\Phi_{X_n}(t)$$



## Examples

Let $X \sim \text{Po}(\lambda_X)$ and $Y \sim \text{Po}(\lambda_Y)$ be independent.

If $Z = X + Y$, then

$$\Phi_Z(t) = \Phi_X(t)\Phi_Y(t)$$
$$= e^{\lambda_X(e^t-1)}\, e^{\lambda_Y(e^t-1)} = e^{(\lambda_X+\lambda_Y)(e^t-1)}$$

Hence,

$$Z \sim \text{Po}(\lambda_X + \lambda_Y)$$



## Examples

Let $X \sim \text{N}(m_X, \sigma_X^2)$ and $Y \sim \text{N}(m_Y, \sigma_Y^2)$ be independent.

If $Z = X + Y$, then

$$\Phi_Z(t) = \Phi_X(t)\Phi_Y(t)$$
$$= e^{\frac{\sigma_X^2 t^2}{2} + t m_X}\, e^{\frac{\sigma_Y^2 t^2}{2} + t m_Y} = e^{\frac{(\sigma_X^2+\sigma_Y^2)t^2}{2} + t(m_X+m_Y)}$$

Therefore

$$Z \sim \text{N}(m_X + m_Y, \sigma_X^2 + \sigma_Y^2)$$

## Characteristic function

**Definition**

*The characteristic function $M_X(\omega)$ of a random variable $X$ is the complex-valued function of the real argument $\omega$ defined as*

$$M_X(\omega) \equiv \mathbb{E}\left(e^{i\,\omega X}\right) = \mathbb{E}\left(\cos\left(\omega X\right)\right) + i\,\mathbb{E}\left(\sin\left(\omega X\right)\right)$$

- Any random variable has a characteristic function.
- $M_X(\omega)$ is well-defined for all $\omega \in \mathbb{R}$.



## Characteristic function

*If $X$ is a discrete or a continuous random variable, then*

$$M_X(\omega) = \begin{cases} \sum_k e^{i\,\omega x_k}\,\mathbb{P}(X = x_k), & \text{if } X \text{ is discrete} \\ \int_{-\infty}^{\infty} e^{i\,\omega x}\, f_X(x)\, dx, & \text{if } X \text{ is continuous} \end{cases}$$

- If $X$ is a continuous random, $M_X(\omega)$ is the Fourier transform of its probability density $f_X(x)$. (Notice the change of sign in the usual definition of Fourier transform.)
- For discrete random variables, characteristic functions are related to Fourier series.



## Properties

- $|M_X(\omega)| \leqslant M_X(0) = 1$ for all $\omega \in \mathbb{R}$.

  Indeed,

$$|M_X(\omega)| = \left|\mathbb{E}\left(e^{i\,\omega X}\right)\right| \leqslant \mathbb{E}\left(\left|e^{i\,\omega X}\right|\right) = \mathbb{E}(1) = 1$$

  Moreover,

$$M_X(0) = \mathbb{E}\left(e^{i \cdot 0 \cdot X}\right) = \mathbb{E}(1) = 1$$

- $M_X(\omega)$ is uniformly continuous on $\mathbb{R}$.
- If $Y = aX + b$, then $M_Y(\omega) = e^{i\omega b}\, M_X(a\,\omega)$.

## Properties

- $\overline{M_X(\omega)} = M_X(-\omega)$.

$$\overline{M_X(\omega)} = \overline{\mathbb{E}(\cos(\omega X)) + i\,\mathbb{E}(\sin(\omega X))}$$
$$= \mathbb{E}(\cos(\omega X)) - i\,\mathbb{E}(\sin(\omega X))$$
$$= \mathbb{E}(\cos(-\omega X)) + i\,\mathbb{E}(\sin(-\omega X)) = M_X(-\omega)$$

More concisely,

$$\overline{M_X(\omega)} = \overline{\mathbb{E}(e^{i\,\omega X})} = \mathbb{E}\left(\overline{e^{i\,\omega X}}\right)$$
$$= \mathbb{E}\left(e^{-i\,\omega X}\right) = \mathbb{E}\left(e^{i\,(-\omega) X}\right) = M_X(-\omega)$$



## Examples

Let $X \sim \text{Bin}(n, p)$. Then,

$$M_X(\omega) = \left(q + p\,e^{i\omega}\right)^n$$

If $X \sim \text{Po}(\lambda)$, then

$$M_X(\omega) = e^{\lambda\left(e^{i\omega}-1\right)}$$

If $X \sim \text{N}(m, \sigma^2)$, then

$$M_X(\omega) = e^{i\omega m - \frac{1}{2}\sigma^2\omega^2}$$



## Characteristic function and moments

**Theorem**

*If $\mathbb{E}(X^n) < \infty$ for some $n \geqslant 1$, then*

$$M_X(\omega) = \sum_{k=0}^{n} \frac{\mathbb{E}(X^k)}{k!}(i\,\omega)^k + o\left(|\omega|^n\right) \quad \text{as} \quad \omega \to 0.$$

*So,*

$$\mathbb{E}(X^k) = \frac{M_X^{(k)}(0)}{i^k} \quad \text{for} \quad k = 1, 2, \ldots, n.$$

In particular, if $\mathbb{E}(X) = 0$ and $\text{Var}(X) = \sigma^2$, then

$$M_X(\omega) = 1 - \frac{1}{2}\sigma^2\omega^2 + o\left(\omega^2\right) \quad \text{as} \quad \omega \to 0.$$



## Characteristic function and moments

Indeed, in the case that $X$ has moments of any order we have

$$M_X(\omega) = \mathbb{E}\left(e^{i\omega X}\right)$$
$$= \mathbb{E}\left(\sum_{k=0}^{\infty} \frac{(i\,\omega X)^k}{k!}\right) = \sum_{k=0}^{\infty} \frac{i^k\,\mathbb{E}(X^k)}{k!}\,\omega^k$$

This is the Taylor series expansion for $M_X(\omega)$. Therefore $i^k\mathbb{E}(X^k) = M_X^{(k)}(0)$ and hence

$$\mathbb{E}(X^k) = \frac{M_X^{(k)}(0)}{i^k}$$

## Convolution theorem

**Theorem**

*Let $X_1, X_2, \dots, X_n$ be independent random variables and let*

$$S = X_1 + X_2 + \cdots + X_n.$$

*Then*

$$M_S(\omega) = M_{X_1}(\omega) M_{X_2}(\omega) \cdots M_{X_n}(\omega)$$



## Convolution theorem

Proof:

$$\begin{aligned} M_S(\omega) &= \mathbb{E}\left(e^{i\omega S}\right) \\ &= \mathbb{E}\left(e^{i\omega(X_1+X_2+\cdots+X_n)}\right) \\ &= \mathbb{E}\left(e^{i\omega X_1} e^{i\omega X_2} \cdots e^{i\omega X_n}\right) \\ &= \mathbb{E}\left(e^{i\omega X_1}\right) \mathbb{E}\left(e^{i\omega X_2}\right) \cdots \mathbb{E}\left(e^{i\omega X_n}\right) \\ &= M_{X_1}(\omega) M_{X_2}(\omega) \cdots M_{X_n}(\omega) \end{aligned}$$



## Convolution theorem

**(Remark)**

*We have essentially the convolution theorem for Fourier transforms.*

*If $X$ and $Y$ are continuous and independent random variables and $Z = X + Y$, then $f_Z = f_X * f_Y$, that implies*

$$\mathcal{F}(f_Z) = \mathcal{F}(f_X) \cdot \mathcal{F}(f_Y),$$

*that is to say,*

$$M_Z(\omega) = M_X(\omega) M_Y(\omega)$$



## Inversion

**Theorem (Inversion of the Fourier transform)**

*Let $X$ be a continuous random variable with density $f_X(x)$ and characteristic function $M_X(\omega)$. Then*

$$f_X(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega x} M_X(\omega)\, d\omega$$

*at every point $x$ at which $f_X(x)$ is differentiable.*

- To obtain $f_X(x)$ from $M_X(\omega)$ usually requires contour integration in the complex plane.

## Inversion

In the discrete case, $M_X(\omega)$ is related to Fourier series.

**Theorem**

*If $X$ is an integer-valued random variable, then*

$$\mathbb{P}(X = k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-ik\omega} M_X(\omega)\, d\omega$$



## Inversion and unicity

**Theorem**

*Let $X$ have probability distribution function $F_X(x)$ and characteristic function $M_X(\omega)$. Let*

$$\overline{F}_X(x) = \frac{F_X(x) + F_X(x^-)}{2}$$

*Then*

$$\overline{F}_X(b) - \overline{F}_X(a) = \lim_{T\to\infty} \frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-ia\omega} - e^{-ib\omega}}{i\omega} M_X(\omega)\, d\omega.$$

$M_X$ specifies uniquely the probability law of $X$. Two random variables have the same characteristic function if and only if they have the same distribution function.



## Joint characteristic functions

**Definition**

*The joint characteristic function of the random variables $X_1$, $X_2$, ..., $X_n$ is defined to be*

$$M_X(\omega_1, \omega_2, \ldots, \omega_n) \equiv \mathbb{E}\left(e^{i(\omega_1 X_1 + \omega_2 X_2 + \cdots + \omega_n X_n)}\right)$$

Using vectorial notation one can write $\omega = (\omega_1, \omega_2, \cdots, \omega_n)^t$, $X = (X_1, X_2, \cdots, X_n)^t$ and

$$M_X(\omega^t) = \mathbb{E}\left(e^{i\omega^t X}\right)$$



## Joint moments

The joint characteristic function allows as to calculate joint moments. For instance, given $X$, $Y$,

$$M_{XY}(\omega_1, \omega_2) = \mathbb{E}\left(e^{i(\omega_1 X + \omega_2 Y)}\right)$$

Therefore,

$$\frac{\partial M_{XY}(\omega_1, \omega_2)}{\partial \omega_1} = i\, \mathbb{E}\left(X\, e^{i(\omega_1 X + \omega_2 Y)}\right)$$

$$\frac{\partial^2 M_{XY}(\omega_1, \omega_2)}{\partial \omega_1 \partial \omega_2} = i^2\, \mathbb{E}\left(XY\, e^{i(\omega_1 X + \omega_2 Y)}\right)$$

## Joint moments

Hence,

$$\left.\frac{\partial^2 M_{XY}(\omega_1,\omega_2)}{\partial\omega_1\partial\omega_2}\right|_{(\omega_1=0,\omega_2=0)} = i^2\,\mathbb{E}(XY) = -\mathbb{E}(XY)$$

More generally,

$$\mathbb{E}\left(X^k Y^l\right) = \frac{1}{i^{k+l}} \left.\frac{\partial^{k+l} M_{XY}(\omega_1,\omega_2)}{\partial^k\omega_1\,\partial^l\omega_2}\right|_{(\omega_1,\omega_2)=(0,0)}$$



## Marginal characteristic functions

Marginal characteristic functions are easily derived from the joint characteristic function.

For instance, given $X$, $Y$:

$$M_X(\omega) = \mathbb{E}\left(e^{i\omega X}\right)$$
$$= \mathbb{E}\left(e^{i(\omega_1 X+\omega_2 Y)}\right)\Big|_{(\omega_1=\omega,\omega_2=0)} = M_{XY}(\omega,0)$$

Analogously,

$$M_Y(\omega) = M_{XY}(0,\omega)$$



## Independent random variables

**Theorem**

*The random variables $X_1, X_2, \ldots, X_n$ are independent if and only if*

$$M_X(\omega_1,\omega_2,\ldots,\omega_n) = M_{X_1}(\omega_1)M_{X_2}(\omega_2)\cdots M_{X_n}(\omega_n)$$

The necessity of the condition is easily proved. Indeed, if the variables are independent, then

$$M_X(\omega_1,\omega_2,\ldots,\omega_n)$$
$$= \mathbb{E}\left(e^{i(\omega_1X_1+\omega_2X_2+\cdots+\omega_nX_n)}\right) = \mathbb{E}\left(e^{i\omega_1X_1}\,e^{i\omega_2X_2}\cdots e^{i\omega_nX_n}\right)$$
$$= \mathbb{E}\left(e^{i\omega_1X_1}\right)\,\mathbb{E}\left(e^{i\omega_2X_2}\right)\cdots\mathbb{E}\left(e^{i\omega_nX_n}\right)$$
$$= M_{X_1}(\omega_1)M_{X_2}(\omega_2)\cdots M_{X_n}(\omega_n)$$