Lesson 13

Moment-Generating and Characteristic Functions: Transform Methods

Famous Transforms such as Laplace

and Fourier Transforms were devised to simplify problems associated with linear differential equations. Similarly, Moment Generating Functions (MGFs), Probability

Generating Functions (PGFs) and Characteristic Functions (CFs) simplify problems about discrete and continuous random variables.

In particular:

- Moments of any order can be calculated by taking derivatives of MGFs rather than integrals involving pdfs, which are harder

- They provide a framework for studying sums of random variables.
- They are very useful for studying convergence of random variables

- They provide an essential framework for studying the spectral theory of random processes
- They are used in studying probability distributions' tails,

rare events, and large deviations.

- They can also be used

when complex r.v.'s are allowed.

Def) (Moment Generating Function)

Assume that X is a r.v. on

some probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

The MGF of X, $M_X : \mathbb{R} \to \mathbb{R}^+$

is defined as

$$M_X(s) = \mathbb{E}\left[e^{sX}\right]$$

Def) (Region of Convergence)

Assume that $M_X(s)$ is the MGF of X. The Region of Convergence of M_X, $ROC(M_X)$ is defined as

$$ROC(M_X) = \{s \mid M_X(s) < \infty\}$$

Note: When X is discrete

$M_X(s) =$

and when X is continuous

$M_X(s) =$

Remark: The above are essentially the same as the definitions of $\mathcal{Z}$ and Laplace transforms of the pmf and pdf, respectively, but with e^{s} instead of z^{-1} and s instead of $-s$.

Note: $S=0$ is always in the $ROC(M_X)$. observe that $E[e^{sX}]\Big|_{s=0} =$

Note: If a discrete random variable has a finite range (i.e. $|\vec{X}(\Omega)| < \infty$), then $ROC(M_X) = \mathbb{R}$. Why?

Example: Assume that $X \sim dU(2, 6)$. Find the MGF of X and its ROC.

Example: Assume that $X \sim Exp(\lambda)$.

Find $M_X(s)$ and its ROC

Example: Assume that $X \sim N(0,1)$. Find $M_X(s)$ and its ROC

Example: Assume that $X \sim \text{Cauchy}(0,1)$.

Find $M_X(s)$ and its ROC

Inversion of MGFs

An interesting question arises:
can we uniquely recover the cdf of a r.v. from its MGF and ROC of MGF?

Generally, the answer is no.
When the MGF is finite at only $s=0$ and infinite elsewhere, we cannot recover the cdf of X from the MGF, as

There are infinitely many r.v.'s that have the same MGF. (Why?)

On the other hand, if the MGF is finite even if only in a tiny region, we can uniquely find the cdf.

In general, ROC(M_X)

is either an isolated point ($s=0$), a bounded interval, or an interval that is bounded from the left, or right. It can also be the whole $\mathbb{R}$.

(It always contains $s=0$).
However, when it is not an isolated point (i.e. $ROC \neq \{0\}$) we have the following Inversion Theorem:

Theorem (Inversion Theorem)

(a) Assume that $M_X(s) < \infty \ \forall s \in [-\varepsilon, \varepsilon]$, $\varepsilon \in \mathbb{R}^+$. Then M_X uniquely determines the cdf of X, F_X.

(b) If X and Y are random variables for which $M_X(s) = M_Y(s) < \infty \ \forall s \in [-\varepsilon, \varepsilon]$, $\varepsilon \in \mathbb{R}^+$, then X and Y have the same cdf.

Proof: It is rather involved

and uses the properties of an analytic function, so it is omitted.

There are formulas for recovering pmf or pdf of a r.v. from its MGF:

$$f_X(x) = \frac{1}{2\pi j}\int_{c-j\infty}^{c+j\infty} e^{-s} M_X(s)\, ds$$

The above formula involves

contour integrals and therefore, we appeal to other methods to find pdfs and cdfs from MGFs. In particular, as we see

later, we use the properties of MGFs and well MGFs to find pdfs. This approach is very similar to the approach we take to invert Laplace and $\mathcal{L}$ transforms

Question: Why $M_X(s) = \mathbb{E}[e^{sX}]$ is called the Moment Generating Function?

$$\left.\frac{d M_X(s)}{ds}\right|_{s=0}$$

$$\left.\frac{d^m M_X(s)}{ds^m}\right|_{s=0} =$$

The m^{th} moment of a r.v. can easily be calculated from its MGF by taking derivatives

Exercise: Find the first four moments of $X \sim \mathcal{N}(0,1)$ from its MGF.

Properties of MGFs

Theorem: X, Y are r.v.s on $(\Omega, \mathcal{F}, \mathbb{P})$:

(a) $Y = aX + b \Rightarrow M_Y(s) = e^{bs} M_X(as)$

(b) If X and Y are independent,

then $M_{X+Y}(s) = M_X(s) M_Y(s)$

(c) If X, Y are independent,

and $\mathbb{P}(Z = X) = p$, $\mathbb{P}(Z = Y) = 1 - p$

then $M_Z(s) = p M_X(s) + (1-p) M_Y(s)$

Proof.

Example: Let $X \sim \mathcal{N}(0,1)$

and $Y = \sigma X + \mu$. Then $Y \sim \mathcal{N}(\mu, \sigma^2)$.

Find $M_Y(s)$.

Example: Let X_1 and X_2 be independent r.v.'s and $X_1 \sim N(\mu_1, \sigma_1^2)$ and $X_2 \sim N(\mu_2, \sigma_2^2)$. Find $M_{X_1+X_2}(s)$.

Example: Let $X \sim Geo_1(p)$. Find $M_X(s)$.

Generating Functions and Random Sums of Random Variables

Assume that $X_1, X_2, \ldots$ is a sequence of i.i.d r.v.'s

with mean μ and variance σ^2, and $X_i \sim X$. Also assume that N is a r.v. that takes positive integer values and is independent from X_i's.

Assume that

$$Y = \sum_{i=1}^{N} X_i$$

We wish to show that

$$M_Y(s) = M_N(\log(M_X(s)))$$

$$= G_N(M_X(s))$$

and

$$G_Y(z) = G_N(G_X(z))$$

Exercise: Using the previous result

$$M_Y(s) = M_N(\log(M_X(s))),$$

show Wald's Equality

$$\mathbb{E}[Y] = \mathbb{E}[N]\,\mathbb{E}[X].$$

Example: Assume that $X_i \sim \mathrm{Exp}(\lambda)$ and $N \sim \mathrm{Geo}_1(p)$, $p \in (0,1)$. Find the pdf of $Y = \sum_{i=1}^{N} X_i$

Exercise: Let $N \sim Pois(\lambda)$ and $X_i \sim Ber(p)$. What is the pmf of $Y = \sum_{i=1}^{N} X_i$.

The Laplace/Z Transforms and Characteristic Functions

So far, we implicitly assumed that s and z are real numbers in MGFs and PGFs.

However we can use complex s and z to obtain Laplace and z transforms of pdfs and pmfs

If $g(x) = u(x) + jv(x)$, we define

$$\mathbb{E}[g(X)] = \mathbb{E}[u(X)] + j\,\mathbb{E}[v(X)]$$

If X has density f_X:

$$\mathbb{E}[g(X)] = \int_{-\infty}^{+\infty} u(x) f_X(x)\,dx + j\int_{-\infty}^{+\infty} v(x) f_X(x)\,dx$$

$$= \int_{-\infty}^{+\infty} g(x) f_X(x)\,dx$$

Then, if $s = \sigma + j\omega$

$$M_X(s) = \mathbb{E}[e^{sX}] = \int_{-\infty}^{+\infty} e^{sx} f_X(x)\,dx$$

The same procedure can be used to define PGFs for complex X's.

Characteristic Functions

We noted that for some s, $M_X(s)$ may be infinite. A pathologic case is when $M_X(s) = \infty \ \forall s \neq 0$ (e.g. when

$X \sim Cauchy(m, \lambda)$), the MGF doesn't provide enough information to determine F_X.

To cope with this problem,

we consider imaginary values of s, i.e. $s = j\omega$, $\omega \in \mathbb{R}$

Def: $\varphi_X(\omega) = \mathbb{E}[e^{j\omega X}]$

is called the characteristic function of X, a r.v. on $(\Omega, \mathcal{F}, \mathbb{P})$

If X is continuous:

$$\varphi_X(\omega) = \int_{-\infty}^{+\infty} e^{j\omega x} f_X(x)\,dx$$

Note that $\varphi_X(\omega)$ is the Fourier transform of f_X but with change of variable $\omega = -v$

Also,

$$\varphi_X(\omega) = E[e^{j\omega X}]$$

$$= E[\cos \omega X] + j\,E[\sin \omega X]$$

Lemma) $\varphi_X(\omega)$ is well-defined and finite for all $\omega \in \mathbb{R}$, i.e. $|\varphi_X(\omega)| \leq 1$.

Theorem: If X and Y are r.v.'s on $(\Omega, \mathcal{F}, \mathbb{P})$,

(a) $Y = aX + b \Rightarrow \Phi_Y(\omega) = e^{j\omega b} \varphi_X(a\omega)$

(b) X, Y independent $\Rightarrow$

$$\varphi_{X+Y}(\omega) = \varphi_X(\omega)\, \varphi_Y(\omega)$$

(c) If X, Y are independent and

$$I = \begin{cases} X & \text{with prob } p \\ Y & \text{with prob } 1-p \end{cases}$$

$$\varphi_I(\omega) = p\, \varphi_X(\omega) + (1-p)\, \varphi_Y(\omega)$$

Proof: Exercise.

Inversion Theorem For CFs

Theorem: Assume that X, Y are r.v.'s and $\forall \omega$ $\varphi_X(\omega) = \varphi_Y(\omega)$

Then $F_X(x) = F_Y(x)$ $\forall x \in \mathbb{R}$.

Inversion of CFs

For a r.v. X with pdf $f_X(x)$

$$f_X(x) = \frac{1}{2\pi} \lim_{T \to \infty} \int_{-T}^{T} e^{-j\omega x} \Phi_X(\omega)\, d\omega$$

$\forall x$ at which $f_X(x)$ is differentiable.

Lemma: If $\mathbb{E}[|X|^k] < \infty$,

then $\left.\frac{d^k \Phi_X(\omega)}{d\omega}\right|_{\omega=0} = j^k \mathbb{E}[X^k]$

Proof: Exercise

Exercise: Assume $X \sim N(0,1)$ and

$Y = \sigma X + \mu$, which is known

to have a $N(\mu, \sigma^2)$ distribution

a) Find the MGF and CF of

Y.

(b) Let $X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2)$ and $X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$ be independent r.v.'s. Find $M_{X_1+X_2}(s)$, $\varphi_{X_1+X_2}(\omega)$, and the distribution of X_1+X_2.

Generating Functions for Random Vectors

Recall that joint pdfs/cdfs/pmfs capture the interaction and convey information about

dependence of multiple r.v.'s, something that cannot be obtained from marginal distributions.

The same is true for

generating functions.

To capture the dependence between multiple random variables, one has to use multivariate generating functions.

Def) Let $X_1, \dots, X_n$ be r.v.'s on $(\Omega, \mathcal{F}, \mathbb{P})$. Assume that $s_1, s_2, \dots, s_n$ are real variables

The multivariate MGF of

$X_1, \ldots, X_n$ is a function of $s_1, s_2, \ldots, s_n$ and is defined as:

$$M_{X_1, X_2, \ldots, X_n}(s_1, s_2, \ldots, s_n) = \mathbb{E}\left[e^{s_1 X_1 + s_2 X_2 + \cdots + s_n X_n}\right]$$

Using the vector notation

$$\underline{s} = \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{bmatrix}, \quad \underline{X} = \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix}$$

$$M_{\underline{X}}(\underline{s}) = \mathbb{E}\left[e^{\underline{s}^T \underline{X}}\right]$$

Inversion Property

Inversion Property of single r.v.'s is inherited by multivariable r.v.'s.

Theorem: If $\underline{X}$ and $\underline{Y}$ are two random vectors and

$M_{\underline{X}}(s) = M_{\underline{Y}}(s) \quad \forall s \in C,$

where C is a neighborhood of $s = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}_{n\times 1}$, then $F_{\underline{X}}(x) = F_{\underline{Y}}(x)$

$\forall x \in \mathbb{R}^n$

Lemma: If

X, Y are independent r.v.'s

$$M_{X,Y}(s_1, s_2) = M_X(s_1)\, M_Y(s_2)$$

Proof:

Example: Assume that X, Y are two r.v.'s and their joint transform is

$$M_{X,Y}(s_1, s_2) = \mathbb{E}[e^{s_1 X + s_2 Y}]$$

and assume that $Z = s_1 X + s_2 Y$.

Write $M_{X,Y}$ in terms of M_Z

Joint Characteristic Functions

Def) Assume that $\underline{X} = \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix}$.

$$\varphi_{\underline{X}}(\omega) = \mathbb{E}\left[e^{j\omega^T \underline{X}}\right] = \mathbb{E}\left[e^{j(\omega_1 X_1 + \cdots + \omega_n X_n)}\right]$$

where $\omega = \begin{bmatrix} \omega_1 \\ \vdots \\ \omega_n \end{bmatrix}$.

When $\underline{X}$ has the joint pdf $f_{\underline{X}}$,

$\varphi_{\underline{X}}(\omega)$ is just the n-dimensional Fourier Transform of $f_{\underline{X}}$

$$\varphi_{\underline{X}}(\omega) = \int_{\mathbb{R}^n} e^{j\omega^T \underline{x}} f_{\underline{X}} \, d\underline{x}$$

and the joint density

can be recovered using the multivariate Fourier transform:

$$f_{\underline{X}}(x) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{-j\omega^T x} \varphi_{\underline{X}}(\omega) d\underline{\omega}$$

Example: Show that if $\underline{X} = \begin{bmatrix} x_1 \\ \vdots \\ X_n \end{bmatrix}$, the joint moment

$$E[X_k X_l] = -\frac{\partial^2}{\partial\omega_k \partial\omega_l} E\left[e^{j\omega^T \underline{X}}\right]\Big|_{\underline{\omega}=0}$$

$$= -\frac{\partial^2}{\partial\omega_k \partial\omega_l} \varphi_{\underline{X}}(\underline{\omega})\Big|_{\underline{\omega}=0}$$

Theorem: If $\underline{X} = [X_1, \ldots, X_n]^T$

X_i's are independent, if and only if $\varphi_{\underline{X}}(\omega) = \prod_{i=1}^{n} \varphi_{X_i}(\omega_i)$.

Proof: Use Fourier Transforms and inverse Fourier Transforms.