ESE 520 Probability and Stochastic Processes

Lecture 11

"Characteristic and moment-generating functions"

Definition 1. For a random variable X its *characteristic function* ϕ_X is defined as

$$\phi_X(t) := E[e^{itX}] = E[\cos tX + i \sin tX], \tag{1}$$

where i is the imaginary unit, i.e., $i^2 = -1$.

Since $|e^{itx}| = 1$, the characteristic function $\phi_X(t)$ exists for all $t \in \mathbb{R}$ for any random variable X.

Definition 2. For a random variable X its *moment-generating function* ψ_X is defined as

$$\psi_X(t) := E[e^{tX}] \tag{2}$$

provided $E[e^{tX}]$ exists (in the textbook, one uses the notation M_X instead of ψ_X).

Given ψ_X exists and is differentiable (the condition $E(X^k) < \infty$ is enough for that), the following fact is true:

$$\frac{d^k}{dt^k}\psi_X(t) = E[X^k e^{tX}], k = 1, 2, ... \tag{3}$$

It follows from (3) that

$$\mu_k = E[X^k] = \psi_X^{(k)}(0), k = 1, 2, ..., \tag{4}$$

where μ_k is the *k-th moment of X*. Formula (4) explains the name "moment-generating function" for ψ.

Similarly, in terms of the characteristic function one has

$$\phi_X^{(k)}(0) = i^k \mu_k, k = 1, 2, ... \tag{5}$$

The formulas (4) and (5) motivate partially the usefulness of functions $\phi_X(t)$ and $\psi_X(t)$ as they allow to calculate the moments of the random variable X. However, the importance of those functions goes far beyond that.

Remark 1. a) If X is a continuous random variable with density $f_X(x)$, then

$$\phi_X(t) = \int_{-\infty}^{\infty} e^{itx} f_X(x)dx, \tag{6}$$

where the integral on right side of (6) represents the *Fourier transform* of $f_X(x)$. Hence, the characteristic function $\phi_X(t)$ of a continuous random variable X is nothing but the Fourier transform of its density function $f_X(x)$.

b) If X is a continuous non-negative random variable with density $f_X(x)$, then

$$\psi_X(t) = \int_{0}^{\infty} e^{tx} f_X(x)dx, \tag{7}$$

where the integral on the right side of (7) represents the *Laplace transform* of $f_X(x)$ "modulo $t \to -t$".

In particular, the importance of characteristic functions is expressed also by the fact that they *uniquely determine the probability distribution* of a random variable as the next Proposition shows.

Proposition 1. a) Let X be a random variable with density $f_X(x)$ and

$$\int_{-\infty}^{\infty} |\phi_X(t)|dt < \infty.$$

Then

$$f_X(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \phi_X(t)dt, \tag{8}$$

where the integral on right side of (8) represents the *inverse Fourier transform* of $\phi_X(t)$.

b) In general, let $a < b$ be points of continuity of the function F_X (cdf of X). Then,

$$F_X(b) - F_X(a) = \lim_{\alpha\to\infty} \frac{1}{2\pi} \int_{-\alpha}^{\alpha} \frac{e^{-ita} - e^{-itb}}{it} \phi_X(t)dt. \tag{9}$$

The following is a straightforward but very important property of characteristic functions.

Proposition 2. Let $X_1, X_2, ..., X_n$ be independent random variables with characteristic functions $\phi_1(t), ..., \phi_n(t)$, respectively. Then the random variable

$$Y := X_1 + ... + X_n \tag{10}$$

has the characteristic function

$$\phi_Y(t) = \prod_{i=1}^{n} \phi_i(t).$$

Indeed: One obtains directly

$$\phi_Y(t) = E[e^{itY}] = E[e^{it(X_1+...+X_n)}] = E[\prod_{i=1}^{n} e^{itX_i}] =$$

- all random variables $e^{itX_i}, i = 1, 2, ..., n$ are independent -

$$\prod_{i=1}^{n} E[e^{itX_i}] = \prod_{i=1}^{n} \phi_i(t).$$

Remark 2. Proposition 2 is also true for the moment-generating functions $\psi_i(t), i = 1, 2, ..., n$.

Example 1. a) Let X be a random variable having a Poisson distribution with parameter λ.

We calculate

$$\phi_X(t) = E[e^{itX}] = \sum_{k=0}^{\infty} e^{ikt}\frac{\lambda^k}{k!}e^{-\lambda} = \sum_{k=0}^{\infty} \frac{(\lambda e^{it})^k}{k!}e^{-\lambda} =$$
$$e^{\lambda e^{it}}e^{-\lambda} = e^{\lambda(e^{it}-1)}.$$

Additionally, if we have a sequence of independent random variables $X_1, ..., X_n$ each of which has a Poisson distribution with parameter λ_i, respectively, then Proposition 2 implies that the random variable Y in (10) has again a Poisson distribution with parameter $\lambda = \sum_{i=1}^{n} \lambda_i$ (exercise!).

b) Let X be a random variable with $\mathcal{N}(\mu, \sigma^2)$ distribution. We calculate

$$\phi_X(t) = E[e^{itX}] = \int_{\mathbb{R}} e^{itx}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}dx =$$

$$z = \frac{x-\mu}{\sigma}, dx = \sigma dz$$

$$\int_{\rm I\!R} e^{it(\mu+\sigma z)} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz = e^{it\mu} \int_{\rm I\!R} \sum_{k=0}^{\infty} \frac{(it\sigma z)^k}{k!} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz =$$

$$e^{it\mu} \sum_{k=0}^{\infty} \frac{(it\sigma)^k}{k!} \int_{\rm I\!R} z^k \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz \tilde{=}$$

for k odd the integrals on the right side are equal to zero (why?) so that we have

$$\tilde{=} e^{it\mu} \sum_{k=0}^{\infty} \frac{(it\sigma)^{2k}}{(2k)!} (2k-1)!! =$$

since the even moments of a random variable $\mathcal{N}(0,1)$ were calculated to be equal to $(2k-1)!!$ (earlier in the course).

Noticing that $(2k-1)!! = \frac{(2k)!}{2^k k!}$, we continue

$$= e^{it\mu} \sum_{k=0}^{\infty} \frac{(-t^2\sigma^2)^k}{(2k)!} \frac{(2k)!}{2^k k!} = e^{it\mu} \sum_{k=0}^{\infty} \frac{(-\frac{1}{2}t^2\sigma^2)^k}{k!} =$$

$$e^{it\mu - \frac{1}{2}t^2\sigma^2}.$$

Additionally, if we have a sequence of independent random variables $X_1, ..., X_n$ each of which has a normal distribution $\mathcal{N}(\mu_i, \sigma_i^2)$, respectiviley, then Proposition 2 implies that the random variable Y in (10) has again a normal distribution $\mathcal{N}(\sum_{i=1}^n \mu_i, \sum_{i=1}^n \sigma_i^2)$ (exersice!).

The concept of a characteristic function can then be generalized to a random vector $X = (X_1, ..., X_n)$.

In other words, the function defined as

$$\phi_X(t) := E[e^{it^T X}] = E[e^{i \sum_{k=1}^n t_k X_k}] \tag{11}$$

is said to be the characteristic function of the vector X where $t^T X = \sum_{k=1}^n t_t X_k$ is the dot product of vectors $t = (t_1, ..., t_n)$ and X.

Example 2. Let X be a non-degenerate Gaussian random vector with mean μ and (positive definite) covariance matrix R. Then

$$\phi_X(t) = \exp(it^T\mu - \frac{1}{2}t^T Rt). \tag{12}$$

Indeed:

$$\phi_X(t) = \int_{\mathbb{R}^n} e^{it^T x} \frac{1}{\sqrt{(2\pi)^n \det R}} e^{-\frac{1}{2}(x-\mu)^T R^{-1}(x-\mu)} dx \stackrel{\sim}{=}$$

$$z = R^{-\frac{1}{2}}(x-\mu), dx = \sqrt{\det R} dz$$

$$\stackrel{\sim}{=} \int_{\mathbb{R}^n} e^{it^T(\mu + R^{1/2}z)} \frac{1}{\sqrt{(2\pi)^n}} e^{-\frac{1}{2}z^T z} dz =$$

$$e^{it^T\mu} \int_{\mathbb{R}^n} e^{i[R^{1/2}t]^T z} \frac{1}{\sqrt{(2\pi)^n}} e^{-\frac{1}{2}z^T z} dz.$$

Now let $u = R^{1/2}t$ so that $[R^{1/2}t]^T z = \sum_{i=1}^n u_i z_i$ and we obtain

$$\phi_X(t) = e^{it^T\mu} \int_{\mathbb{R}^n} \prod_{i=1}^n (\frac{1}{\sqrt{(2\pi)}} e^{iu_i z_i} e^{-\frac{1}{2}z_i^2}) dz =$$

$$e^{it^T\mu} \prod_{i=1}^n (\int_{\mathbb{R}} \frac{1}{\sqrt{(2\pi)}} e^{iu_i z_i} e^{-\frac{1}{2}z_i^2} dz_i).$$

The integral in parentheses is the characteristic function of 1-dimensional normal distribution $\mathcal{N}(0,1)$ calculated earlier so that

$$\phi_X(t) = e^{it^T\mu} \prod_{i=1}^n e^{-u_i^2/2} = e^{it^T\mu} e^{-\frac{1}{2}u^T u} = e^{it^T\mu} e^{-\frac{1}{2}t^T Rt}.$$

The notion of characteristic function allows to give a definition of an *arbitrary Gaussian vector* $X = (X_1, ..., X_n)$, *not necessarily a non-degenerate one.*

Definition 3. A random vector $X = (X_1, ..., X_n)$ is said to be a *(jointly) Gaussian vector* with mean μ and covariance matrix R if its characteristic function $\phi_X(t)$ is of the form (12).

Note: R is no longer assumed to be positive definite. If it is, then we know that it will be the characteristic function of a non-degenerate Gaussian vector.

Example 3. Let $R = 0$ be a zero matrix. Then $\phi_X(t) = e^{it^T\mu}$ is the characteristic function of the vector $X = \mu$=const which is also a Gaussian vector by definition.

Example 4. Let X be a random variable with Pareto distribution, that is its pdf is of the form

$$f(x) = \begin{cases} \frac{2}{x^3}, & x \geq 1 \\ 0, & x < 1. \end{cases}$$

Calculate the moment-generating function of X (if it exists).

Solution: For any $t \leq 0$, $\psi_X(t) = E[e^{tX}] \leq 1$ so that $\psi_X(t)$ exists.

For any $t > 0$, we have that

$$\psi_X(t) = E[e^{tX}] = E[\sum_{n=0}^{\infty} \frac{(tX)^n}{n!} \geq E[\frac{(tX)^3}{3!}$$

since $tX > 0$.

However,

$$E[\frac{(tX)^3}{3!} = \frac{t^3}{3!} E[X^3] = \frac{2t^3}{3!} \int_1^{\infty} \frac{x^3}{x^3} dx = \infty$$

implying that $\psi_X(t)$ does not exist for all $t > 0$.

Example 5. a) Let X be a random variable having an exponential distribution with parameter $\lambda > 0$. In other words, its pdf is given by

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0 \\ 0, & x \leq 0. \end{cases}$$

We find

$$\psi_X(t) = E[e^{tX}] = \int_0^{\infty} \lambda e^{tx} e^{-\lambda x} dx =$$

$$\lambda \int_0^{\infty} e^{(t-\lambda)x} dx = \frac{\lambda e^{(t-\lambda)x}}{t-\lambda} |_0^{\infty} = [t < \lambda] = \frac{\lambda}{\lambda - t}.$$

b) Let X and Y two independent exponential random variables with parameter λ. Define $Z := X - Y$. Find the $\psi_Z(t)$.

First, using independence, we obtain that

$$\psi_Z(t) = E[e^{tZ}] = E[e^{t(X-Y)}] = E[e^{tX}e^{-tY}] = \psi_X(t)\psi_Y(-t).$$

By part a),

$$\psi_Z(t) = \psi_X(t)\psi_Y(-t) = \frac{\lambda}{\lambda - t}\frac{\lambda}{\lambda - (-t)} = \frac{\lambda^2}{\lambda^2 - t^2}$$

which would be the moment-generating function of so-called *Laplace probability distribution* with parameter λ defined for all t such that $|t| < \lambda$.