EE5110: Probability Foundations for Electrical Engineers **July-November 2015**

Lecture 26: Characteristic Functions

Lecturer: Dr. Krishna Jagannathan *Scribe: Aseem Sharma and Ajay M.*

The characteristic function of a random variable X is defined as

$$\begin{aligned} C_X(t) =& \mathbb{E}[e^{itX}] \\ =& \mathbb{E}[\cos(tX)] + i\mathbb{E}[\sin(tX)], \end{aligned}$$

which can also be written as

$$C_X\left(t\right) = \int e^{itx} d\mathbb{P}_X.$$

If X is a continuous random variable with density function $f_X\left(x\right)$, then

$$C_X\left(t\right) = \int e^{itx} f_X\left(x\right) dx.$$

The advantage with the characteristic function is that it always exists, unlike the moment generating function, which can be infinite everywhere except $s = 0$.

Example 1: Let X be an exponential random variable with parameter μ. Find its characteristic function.

Solution: Recall that for an exponential random variable with parameter μ, $f_X(x) = \mu e^{-\mu x}$. Thus, we have

$$\begin{aligned} C_X(t) =& \int\limits_{x=0}^{\infty} \mu e^{-\mu x} e^{itx} dx \\ =& \frac{\mu}{\mu - it}. \end{aligned}$$

We have evaluated the above integral essentially by pretending that $\mu - it$ is a real number. Although this happens to produce the correct answer in this case, the correct method of evaluating a characteristic function is by performing contour integration. Indeed, in the next example, it is not possible to obtain the correct answer by pretending that it is a real number (which is not).

Example 2: Let X be a Cauchy random variable. Find its characteristic function.

Solution: The density function for a Cauchy random variable is

$$f_X(x) = \frac{1}{\pi(1 + x^2)}.$$

Therefore,

$$\begin{aligned} C_X(t) =& \int\limits_{x=-\infty}^{\infty} \frac{e^{itx}}{\pi(1 + x^2)} dx \\ =& e^{-|t|}. \end{aligned}$$

The above expression is not entirely trivial to obtain. Indeed, it requires considering two separate contour integrals for $t > 0$ and $t < 0$, and invoking Cauchy's residue theorem to evaluate the contour integrals. (For details, see *http://www.wpressutexas.net/forum/attachment.php?attac hmentid=408&d=1296667390.*)

However, it is also possible to obtain the characteristic function of the Cauchy random variable by invoking a Fourier transform duality trick from your undergraduate signals and systems course. (Do it!)

Recall also that the moment generating function of a Cauchy random variable does not converge anywhere except at $s = 0$. On the other hand, we find here that the characteristic function for the Cauchy random variable exists everywhere. This is essentially because the integral defining the chracteristic function converges absolutely, and hence uniformly, for all $t \in R$. Characteristic functions are thus particularly useful in handling heavy-tailed random variables, for which the corresponding moment generating functions do not exist.

Let us next discuss some properties of characteristic functions.

26.1 Properties of characteristic functions

26.1.1 Elementary properties

1) If $Y = aX + b$, $C_Y(t) = e^{ibt}C_X(at)$.
2) If X and Y are independent random variables and $Z = X + Y$, then $C_Z(t) = C_X(t)C_Y(t)$.
3) If $M_X(s) < \infty$ for $s \in [-\epsilon, \epsilon]$, then $C_X(t) = M_X(it)$ for all $t \in \mathbb{R}$.

Example 3: Let $X \sim \mathcal{N}(0, 1)$. The moment generating function is

$$M_X(s) = e^{\frac{s^2}{2}}.$$

Then, the characteristic function is

$$C_X(t) = M_X(it) = e^{\frac{-t^2}{2}}.$$

For a non-standard Gaussian, $Y \sim \mathcal{N}(\mu, \sigma^2)$, we can now invoke property 1) and conclude that $C_Y(t) = \exp\left(i\mu t - \frac{\sigma^2 t^2}{2}\right)$.

26.1.2 Defining properties

Theorem 26.1 *A characteristic function $C_X(t)$ satisfies the following properties:*

1) $C_X(0) = 1$ and $|C_X(t)| \leq 1$, $\forall t \in \mathbb{R}$.
2) $C_X(t)$ is uniformly continuous on $\mathbb{R}$, *i.e.*, $\forall t \in \mathbb{R}$, $\exists$ a $\psi(h) \downarrow 0$ as $h \to 0$ such that

$$|C_X(t+h) - C_X(t)| \leq \psi(h).$$

3) $C_X(t)$ is a non-negative definite kernel, *i.e.*, for any n, any real $t_1, t_2, \ldots, t_n$, and any complex $z_1, z_2, \ldots, z_n$, we have

$$\sum_{j,k} z_j C_X(t_j - t_k)\overline{z_k} \geq 0.$$

Proof:

1)

$$|C_X(t)| = |\int e^{itx} d\mathbb{P}_X| \leq \int |e^{itx}| d\mathbb{P}_X = 1.$$

2)

$$\begin{aligned}|\mathbb{E}[e^{i(t+h)X}] - \mathbb{E}[e^{itX}]| &= |\mathbb{E}[e^{itX}(e^{ihX} - 1)]| \\ &\leq \mathbb{E}[|e^{ihX} - 1|].\end{aligned}$$

Let $|e^{ihX} - 1| = y(h)$ and $\mathbb{E}[y(h)] = \psi(h)$. We now need to show that $\psi(h) \downarrow 0$ as $h \downarrow 0$. Note that $y(h) \to 0$ as $h \to 0$. Further,

$$\begin{aligned}y(h) &= |e^{ihX} - 1| \\ &= \sqrt{(\cos(hX) - 1)^2 + (\sin(hX))^2} \\ &= \sqrt{2 - 2\cos(hX)} \\ &= 2\sin\left(\frac{hX}{2}\right) \\ &\leq 2.\end{aligned}$$

Since $y(h)$ is bounded above by 2, applying DCT, we thus have $\psi(h) \to 0$ as $h \to 0$.

3)

$$\begin{aligned}\sum_{j,k} z_j C_X(t_j - t_k)\overline{z_k} &= \sum_{j,k} \int z_j e^{i(t_j - t_k)X} \overline{z_k}\, d\mathbb{P}_X \\ &= \sum_{j,k} \int z_j e^{it_jX} (\overline{z_k e^{it_kX}})\, d\mathbb{P}_X \\ &= \mathbb{E}[\sum_{j,k} z_j e^{it_jX} (\overline{z_k e^{it_kX}})] \\ &\geq \mathbb{E}[\sum_j |z_j e^{it_jX}|^2] \\ &\geq 0.\end{aligned}$$

■

The significance of 3) may not be apparent at a first glance. However, these three properties are considered as the defining properties of a characteristic function, because these properties are also *sufficient* for an arbitrary function to be the characteristic function of some random variable. This important result is known as Bochner's theorem, which is beyond our scope.

Theorem 26.2 ***(Bochner's theorem)*** *A function $C(\cdot)$ is a characteristic function of a random variable if and only if it satisfies the properties of theorem 26.1.*

26.2 Inversion Theorems

The following inverse theorems are presented without proof, since the proofs require some sophisticated machinery from harmonic analysis and complex variables. Essentially, they state that the CDF of a random variable can be recovered from the characteristic function.

Theorem 26.3

(i) *Let X be a continuous random variable, having a probability density function $f_X(x)$ and the corresponding characteristic function be*

$$C_X(t) = \int_{-\infty}^{\infty} e^{itx} f_X(x) dx. \tag{26.1}$$

The probability density function, $f_X(x)$ can be obtained from the characteristic function as

$$f_X(x) = \frac{1}{2\pi} \lim_{T\to\infty} \int_{-T}^{T} e^{-itx} C_X(t) dt, \tag{26.2}$$

at every point where $f_X(x)$ is differentiable.

(ii) *The sufficient (but not necessary) condition for the existence of a probability density function is that the characteristic function should be absolutely integrable, i.e.,*

$$\int_{-\infty}^{\infty} |C_X(t)| dt < \infty. \tag{26.3}$$

(iii) *Let $C_X(t)$ be a valid characteristic function of a random variable X with a cumulative distribution function $F_X(x)$. We define,*

$$\hat{F}_X(x) = \frac{1}{2}\left(F_X(x) + \lim_{y\uparrow x} F_X(y)\right) \text{ for some } y, \tag{26.4}$$

then

$$\hat{F}_X(b) - \hat{F}_X(a) = \lim_{T\to\infty} \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{-iat} - e^{-ibt}}{T} C_X(t) dt \quad \forall\ a \text{ and } b. \tag{26.5}$$

In part (iii) above, the function $\hat{F}_X(x)$ coincides with the CDF $F_X(x)$ at all points where the CDF is continuous. At points of discontinuity, it is easy to see that $\hat{F}_X(x)$ takes the value at the mid-point of the right and left limits of the CDF. Equation (26.5) says that the function $\hat{F}_X(x)$ can be recovered from the characteristic function. Finally, since the CDF is right-continuous, we can recover $F_X(x)$ from $\hat{F}_X(x)$.

26.3 Moments from the Characteristic Function

Theorem 26.4

(i) *Let X be a random variable having a characteristic function $C_X(t)$. If $\frac{\mathrm{d}^k C_X(t)}{\mathrm{d}t^k}$ exists at $t = 0$, then* *(a)* $\mathbb{E}[|X^k|] < \infty$ *when* k *is even.*

(b) $\mathbb{E}[|X^k - 1|] < \infty$ *when* k *is odd.*

(ii) If $\mathbb{E}[|X^k|] < \infty$*, then*

$$i^k \mathbb{E}[X^k] = \left. \frac{\mathrm{d}^k C_X(t)}{\mathrm{d}t^k} \right|_{t=0}. \tag{26.6}$$

Further,

$$C_X(t) = \sum_{j=0}^{k} \frac{\mathbb{E}\left[X^j\right]}{j!} (it)^j + \mathcal{O}\left(t^k\right), \tag{26.7}$$

where the error, $\mathcal{O}\left(t^k\right)$ *means that* $\mathcal{O}\left(t^k\right) / \left(t^k\right) \to 0$ *as* $t \to 0$*.*

Note: Since $C_X(t) = \int e^{itx} d\mathbb{P}_X$ converges uniformly, we are justified in 'taking the derivative inside the integral.'

26.4 Exercise:

1. *[Papoulis]* Use characteristic function definition to find the distribution of $Y = aX^2$, if X is Gaussian with zero mean and variance σ^2.
2. *[Papoulis]* Use characteristic function definition to find the distribution of $Y = \sin(X)$, if X is uniformly distributed in $(-\pi/2, \pi/2)$.