Report on Characteristic Functions

Fu Tianwen Yao Chaorui Zhao Feng

April 16, 2019

1 Propose of Presentation

In ESTR2002, it is quite common to deal with the sum of independent random variables, or approximate sums of random variables by assuming independence. However, it is pretty difficult and tedious to handle these kind of problems with convolution. With simple knowledge of complex numbers, **Characteristic Functions** can be very helpful by turning convolution into multiplication.
In the presentation, we give a brief introduction about the definition, background and properties related to characteristic functions. A brief proof of *Central Limit Theorem* is also given. However, we did not give any proof for the properties nor the detailed and rigorous proof of CLT because we wanted to keep only the intuition rather than the tedious details of analysis.
To improve our presentation quality, we prepared a handout for our fellow classmates to read during the presentation, and put all the related mathematical statements on the slide. However, since the handout is posted too late, few classmates have them on hand during the presentation; statements on the slide also make the presentation going too fast compared with handwriting them on the whiteboard. Above would be most mistakes in the reflection of our presentation.

2 Major Findings

Below shows the most important findings, which are almost the same as the notes given in Piazza.

2.1 Definition and Relation with Moments

Definition 2.1. Let X be a random variable and denote by F the cumulative distribution function of X (or f the probability density function). The characteristic function $\varphi = \varphi_X$ of X (or of F, in which case we also write φ_F) is defined by [1]

$$\varphi_X(t) := E[e^{itX}] = \int_{-\infty}^{\infty} e^{itx} dF(x) = \int_{-\infty}^{\infty} e^{itx} f(x) dx, t \in \mathbb{R}$$

The i mentioned above is the imaginary unit. If you have no knowledge of complex numbers, all you need to know is simply:

* $i^2 = -1$,
* $e^{i\theta} = \cos\theta + i \sin\theta$. (This formula comes from Taylor Series)

Definition 2.2 (Moment [2])**.** For probability density functions f (or cumulative density function F), the moments are given by

$$\mu'_n = E[X^n] = \int_{-\infty}^{\infty} x^n dF(x) = \int_{-\infty}^{\infty} x^n f(x) dx$$

Theorem 2.1. *Suppose X is a random variable for which the nth moment exists. Suppose $\varphi_X(t)$ is the characteristic function of X, and $\varphi_X(t)$ is n-times differentiable. Then the nth moment*

$$E[X^n] = i^{-n} \left[\frac{d^n}{dt^n} \varphi_X(t) \right]_{t=0}$$

Conclusion. From the theorem one may connect the moments (as in statistics) and the characteristic function (as in probability). It is easy to get the moments from the distribution just directly by definition; however, with characteristic functions, we may retrieve the distribution with all the statistics.

2.2 Properties of Characteristic Function

Theorem 2.2 (Uniqueness Theorem [3]). *Let X be a real random variable with distribution function F and characteristic function φ. Similarly, let Y have distribution function G and characteristic function ψ. If $\varphi(t) = \psi(t)$ for all $t \in \mathbb{R}$ then $F(x) = G(x)$ for all $x \in \mathbb{R}$.*

Remark. From this we may easily conclude the distribution of a random variable if we can prove that its characteristic function is of the same form as a known distribution.

Theorem 2.3 (Inversion Formula [4]). *If $\int_{\mathbb{R}} |\varphi(t)|dt < \infty$ then X has bounded continuous density*

$$f(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-itx} \varphi(t)dt$$

Remark. This gives us a way to calculate the distribution with characteristic functions.

Theorem 2.4. *If X and Y are independent random variables then the characteristic function of their sum is*[1]

$$\varphi_{X+Y}(t) = \varphi_X \cdot \varphi_Y.$$

Remark. This gives us a much simpler way than convolution.

Remark. If X and Y are random variables such that $\varphi_{X+Y} = \varphi_X \cdot \varphi_Y$, then in general we do not conclude X and Y are independent. (See page 13 in [1]) This is called **subindependence**[5].

Corollary 2.4.1. *The product of two characteristic functions is a characteristic function.*

Theorem 2.5. *For any $a, b \in \mathbb{R}$,*

$$\varphi_{aX+b}(t) = e^{ibt} \varphi_X(at).$$

Table 1: Characteristic Functions for Common Distributions[6]

Distribution	PMF/PDF	Characteristic Function
Constant $X \equiv a$	-	$\varphi_X(t) = e^{iat}$.
Binomial $X \sim Binomial(m, p)$	$p_X(n) = \binom{n}{m} p^n (1-p)^{m-n}$	$\varphi_X(t) = (pe^{it} + (1-p))^m$
Poisson $X \sim Poisson(\lambda)$	$p_X(n) = \frac{\lambda^n}{n!} e^{-\lambda}$	$\varphi_X(t) = e^{\lambda(e^{it}-1)}$
Exponential $X \sim Exponential(\lambda)$	$p_X(n) = \lambda e^{-\lambda x}$	$\varphi_X(t) = \frac{\lambda}{\lambda - it}$
Normal $X \sim N(0,1)$	$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$	$\varphi_X(t) = e^{-\frac{t^2}{2}}$
Normal $Y \sim N(\mu, \sigma^2)$	$f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y-\mu)^2}{2\sigma^2}}$	$\varphi_Y(t) = e^{it\mu - \frac{\sigma^2 t^2}{2}}$

2.3 Examples and Applications

Example 2.1. Rain falls on your head at λ drops per second on average. What is the distribution of rain drops on your head in two seconds?

Solution. Let X, Y be two independent $Poisson(\lambda)$ random variables. Let $Z = X + Y$. Notice that characteristic functions for X (and respectively Y) is $\varphi_X(t) = e^{\lambda(e^{it}-1)}$. Therefore we have $\varphi_Z(t) = \varphi_{X+Y}(t) = (\varphi_X(t))^2 = e^{2\lambda(e^{it}-1)}$. By uniqueness of characteristic functions we know that $Z \sim Poisson(2\lambda)$.

Remark. By similar ideas, one can show that the sum of two independent poisson random variables has a possion distribution with an expectation of the sum of both expectations.

Example 2.2. $X_1 \sim N(\mu_1, \sigma_1^2), X_2 \sim N(\mu_2, \sigma_2^2)$. X_1, X_2 are independent. Distribution of $Y = X_1 + X_2$?

Solution. Similarly to last example

$$\varphi_Y(t) = \varphi_{X_1+X_2}(t) = e^{it\mu_1 - \frac{\sigma_1^2 t^2}{2}} \cdot e^{it\mu_2 - \frac{\sigma_2^2 t^2}{2}}$$

Therefore we have

$$\varphi_Y(t) = e^{it(\mu_1+\mu_2) - \frac{(\sigma_1^2+\sigma_2^2)t^2}{2}}$$

which implies that $Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$

[1] Properties 2.4 and 2.5 on come from Bisgaard and Zoltan's book [1].

2.4 Central Limit Theorem

First we prove a useful lemma[2]:

Lemma 2.1. *For any random variable X with $E[X] = 0, Var[X] = 1$, we have $\varphi_X(t) = 1 - \frac{1}{2}t^2 + o(t^2)$.*

Proof. Directly from the definition and Taylor Series we have

$$\varphi_X(t) = E[e^{itX}] = E[1 + itX + \frac{1}{2}i^2t^2X^2 + o((tX)^2)]$$

By linearity of expectation,

$$\varphi_X(t) = E[1] + itE[X] - \frac{1}{2}t^2E[X^2] + o(t^2)$$

Also

$$E[X^2] = Var[X] - (E[X])^2 = 1$$

Therefore

$$\varphi_X(t) = 1 - \frac{1}{2}t^2 + o(t^2)$$

□

Theorem 2.6 (Central Limit Theorem[7])**.** *If X_i are independent identically distributed random variables with $E[X_i] = \mu, Var[X_i] = \sigma^2$, then $S_n^* = \frac{1}{\sigma\sqrt{n}}\sum_{i=1}^{n}(X_i - \mu)$ converges weakly to $N(0, 1)$.*

Remark. Note that theorem 2.6 is equivalent to the one given in class.

Proof. By Lemma 2.1 we denote the characteristic function of $\frac{X_i-\mu}{\sigma}$ by $\varphi(t)$, then $\varphi(t) = 1 - \frac{1}{2}t^2 + o(t^2)$. Therefore the characteristic function of S_n^* is

$$\varphi^n(t/\sqrt{n}) = [1 - \frac{t^2}{2n} + o(t^2/n)]^n$$

Take $n \to \infty$ and we get the characteristic function of $\lim_{n\to\infty} S_n^*$

$$\Phi(t) = \lim_{n\to\infty} \varphi^n(t/\sqrt{n}) = \lim_{n\to\infty} [1 - \frac{t^2}{2n}]^n = e^{-\frac{t^2}{2}}$$

Therefore $\lim_{n\to\infty} S_n^*$ converges to $N(0, 1)$. □

References

[1] T. M. Bisgaard and S. Zoltán, *Characteristic functions and moment sequences : positive definiteness in probability.* Huntington, N.Y.: Nova Science Publishers, 2000.

[2] Wikipedia contributors, "Moment (mathematics) — Wikipedia, the free encyclopedia." `https://en.wikipedia.org/w/index.php?title=Moment_(mathematics)&oldid=886051824`, 2019. [Online; accessed 8-March-2019].

[3] M. J. Wichura, " Characteristic functions." `https://galton.uchicago.edu/~wichura/Stat304/Handouts/L12.cf2.pdf`. [Online; accessed 4-April-2019].

[4] R. Durrett, "Probability: Theory and examples." `https://services.math.duke.edu/~rtd/PTE/PTE5_011119.pdf`, 2019. [Online; accessed 5-April-2019].

[5] Wikipedia contributors, "Subindependence — Wikipedia, the free encyclopedia." `https://en.wikipedia.org/w/index.php?title=Subindependence&oldid=874173978`, 2018. [Online; accessed 7-April-2019].

[2]The proofs for lemma 2.1 and theorem 2.6 are not rigorous and serves only for intuitively demonstrating the usage of Characteristic Functions. See the references[7, 8] for detailed proofs.

[6] Aktuella kurssidor vid Matematiska institutionen Stockholms universitet , “Lecture 10: Characteristic Functions.” `https://kurser.math.su.se/pluginfile.php/5149/mod_resource/content/1/lecture-10b.pdf`. [Online; accessed 4-April-2019].

[7] D. L. McLeish, “Characteristic functions and the central limit theorem.” `http://sas.uwaterloo.ca/~dlmcleis/s901/chapt6.pdf`. [Online; accessed 5-April-2019].

[8] R. Su, “Lecture 7.” `http://www.math.nus.edu.sg/~matsr/ProbI/Lecture7.pdf`. [Online; accessed 5-April-2019].