Characteristic Functions

Fu Tianwen Yao Chaorui Zhao Feng

April 7, 2019

1 Sum of Two Independent Random Variables

Let's consider the following problems:

Problem 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?

Problem 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 $Y = X_1 + X_2$?

With what we have already learned, it is really easy to calculate the statistics(expectation, variance, ...) of the sums. But what is the distribution? Are they the same as the parts or some strange distribution? One may suggest calculating them by convolution. But is there some simpler idea? That leads to the topic we're going to talk about today – characteristic functions.

2 Characteristic Functions

2.1 Definition

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)

2.2 Properties

Theorem 2.1 (Uniqueness Theorem [2])**.** *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.

Even if we cannot find the distribution in the lookup table, we may also retrieve it from characteristic functions by hand:

Theorem 2.2 (Inversion Formula [3])**.** *If $\int_{\mathbb{R}} |\varphi(t)| dt < \infty$ then X has bounded continuous density*

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

Theorem 2.3. *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**[4].

Corollary 2.3.1. *The product of two characteristic functions is a characteristic function.*

Theorem 2.4. *For any $a, b \in \mathbb{R}$,*

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

2.3 Common Distributions and Their Characteristic Functions

First we try to derive the characteristic function for $X \sim Exponential(\lambda)$. Then

$$\begin{aligned}
\varphi_X(t) &:= E[e^{itX}] \\
&= \int_{-\infty}^{\infty} e^{itx} f(x)dx \\
&= \int_{0}^{\infty} e^{itx} \cdot \lambda e^{-\lambda x} dx \qquad \text{By distribution of x} \\
&= \frac{\lambda}{it - \lambda} e^{(it-\lambda)x}\bigg|_0^{\infty} \\
&= \frac{\lambda}{\lambda - it} \qquad \text{By Squeeze Theorem}
\end{aligned}$$

where the last two steps can be justified by

$$\lim_{x\to\infty} |e^{(it-\lambda)x}| = \lim_{x\to\infty} |e^{-\lambda x}| = 0 \text{ for positive } \lambda$$

Characteristic functions of other distributions can be seen in table 1.

Table 1: Characteristic Functions for Common Distributions[5]

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 $Y \sim N(\mu, \sigma^2)$	$f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y-\mu)^2}{2\sigma^2}}$	$\varphi_Y(t) = e^{it\mu - \frac{\sigma^2 t^2}{2}}$

2.4 Bridge between Probability and Statistics

So what ”characteristics” are these weird ”characteristic functions” talking about? To understand this, first we need some knowledge about **moments**.

Definition 2.2 (Moment [6])**.** For probability density functions f (or cumulative density function F), the moments are given by

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

[1] Properties 2.3 and 2.4 on come from Bisgaard and Zoltan's book [1].

Remark. Not to be confused with mean μ.

Relation with statistics. Observe that the first moment is simply the expectation. The second moment is related to the variance: $Var[X] = E[X^2] - (E[X])^2$. What about the third moment? A related concept is **skewness**[7]:

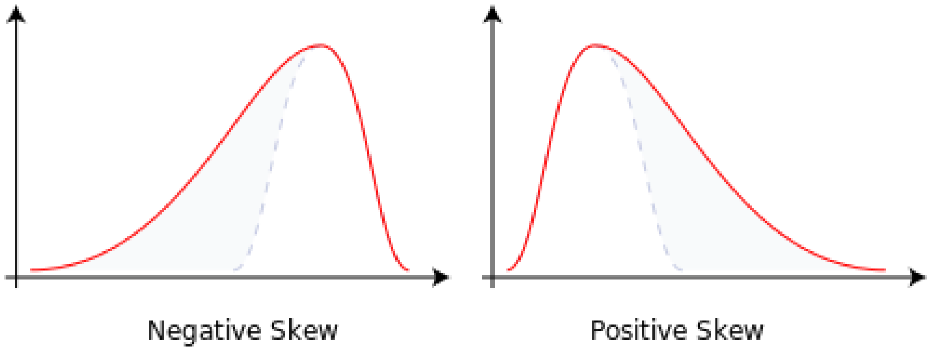

Figure 1: Two distributions with the same expectation and variance, but different skewness

Now let's consider the values of derivatives of $\varphi_X(t)$ at 0:

$$\begin{aligned}
\varphi_X(t) &= \int_{-\infty}^{\infty} e^{itx} f(x)dx; & \varphi_X(0) &= 1 \\
\varphi_X'(t) &= i\int_{-\infty}^{\infty} xe^{itx} f(x)dx; & \varphi_X'(0) &= iE[X] \\
\varphi_X''(t) &= i^2\int_{-\infty}^{\infty} x^2e^{itx} f(x)dx; & \varphi_X''(0) &= i^2E[X^2] \\
&\dots
\end{aligned}$$

By simple induction one can get the following:

Theorem 2.5. *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*

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

Therefore with infinite statistics we may reconstruct the distribution by characteristic functions.

3 Solutions to the Problems

With the powerful characteristic functions, the sums of independent random variables are just a piece of cake.

Solution to Problem 1. Our intuition tells us that it should be $Poisson(2\lambda)$.

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.

Solution to Problem 2. Similarly to previous problem

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

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

Remark. This shows that the sum of two independent normal random variables is another normal random variable, the mean and standard deviation of which is the same as calculated statistically from the two components.

4 Other Examples

Example 4.1. Show that $Poisson(\lambda)$ is the same as $\lim_{n\to\infty} Binomial(n, \frac{\lambda}{n})$.

Solution. Take the limit of the characteristic function of the binomial distribution.

$$\begin{aligned}\lim_{n\to\infty}\varphi_B(t) &= \lim_{n\to\infty}(\frac{\lambda}{n}e^{it} + (1-\frac{\lambda}{n}))^n \\ &= \lim_{n\to\infty}(1+(e^{it}-1)\frac{\lambda}{n})^n \\ &= e^{\lambda(e^{it}-1)}\end{aligned}$$

Remark. This is not a standard way of defining Poisson distribution; however, this gives us a nice intuition of Poisson distribution and can be seen as a proof for the formula of the characteristic function of it.

5 Central Limit Theorem

First we prove a useful lemma[2]:

Lemma 5.1. *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 5.1 (Central Limit Theorem[8])**.** *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 5.1 is equivalent to the one given in class.

Proof. By Lemma 5.1 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 $N(0,1)$. □

[2]The proofs for lemma 5.1 and theorem 5.1 are not rigorous and serves only for intuitively demonstrating the usage of Characteristic Functions. See the references[8, 9] for detailed proofs.

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

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

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

[5] 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].

[6] 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].

[7] Wikipedia contributors, “Skewness — Wikipedia, the free encyclopedia.” `https://en.wikipedia.org/w/index.php?title=Skewness&oldid=886052950`, 2019. [Online; accessed 8-March-2019].

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

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