

# STATS310A - Lecture 17

Persi Diaconis  
Scribed by Michael Howes

11/18/21

## Contents

<b>1 Characteristic functions</b>	<b>1</b>
1.1 Definitions . . . . .	1
1.2 Properties . . . . .	1
1.3 Motivation . . . . .	2

## 1 Characteristic functions

Today we will be discussing the use of Fourier analysis in probability. In this area, the key objects of study are characteristic functions. The plan is to first define characteristic functions and state some of their properties. We will then use characteristic functions in some interesting examples and then we will prove the properties we used.

### 1.1 Definitions

**Definition 1.** Let  $\mu$  be a probability measure on  $\mathbb{R}$  the *characteristic function* of  $\mu$  is the function  $\phi_\mu : \mathbb{R} \rightarrow \mathbb{C}$  given by

$$\phi_\mu(t) = \int_{\mathbb{R}} e^{itx} \mu(dx) = \int_{\mathbb{R}} \cos(tx) \mu(dx) + i \int_{\mathbb{R}} \sin(tx) \mu(dx).$$

**Definition 2.** If  $X$  is a random variable on  $\mathbb{R}$  with distribution  $\mu$ , the *characteristic function* of  $X$  is defined to be  $\phi_X = \phi_\mu$ .

Note that if  $X \sim \mu$ , then

$$\phi_X(t) = \mathbb{E}[e^{itX}] = \mathbb{E}[\cos(tX)] + i\mathbb{E}[\sin(tX)].$$

We will regularly interchange  $\phi_\mu$  and  $\phi_X$ .

### 1.2 Properties

Some properties of characteristic functions are

- (1) The integral  $\phi_X(t)$  is well defined for all  $t$  and  $X$ .
- (2) The function  $|\phi_X|$  is bounded by 1 and  $\phi_X(0) = 1$ .
- (3) The function  $\phi_X$  is continuous.

*Proof.* For all  $t, h \in \mathbb{R}$  we have

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

By the dominated convergence theorem,

$$\mathbb{E} [|e^{ihX} - 1|] \xrightarrow{h \rightarrow 0} 0.$$

Thus  $\phi_X$  is in fact uniformly continuous. □

(4) If  $X$  and  $Y$  are independent random variables, then

$$\phi_{X+Y}(t) = \phi_X(t)\phi_Y(t).$$

Equivalently, if  $\mu$  and  $\nu$  are two probability measures on  $\mathbb{R}$ , then

$$\phi_{\mu * \nu}(t) = \phi_\mu(t)\phi_\nu(t),$$

where  $\mu * \nu$  is the convolution of  $\mu$  and  $\nu$ .

(5) If the  $k^{th}$  moment  $\mathbb{E}[X^k]$  exists, then  $\phi$  is  $k$  times differentiable at 0 and

$$\mathbb{E}[X^k] = i^k \phi_X^{(k)}(0).$$

The converse is not necessarily true. The differentiability of  $\phi_X$  does not imply the existence of moments.

(6) If  $\phi_\mu(t) = \phi_\nu(t)$  for all  $t$ , then  $\mu = \nu$ . This is called the uniqueness theorem.

(7) Let  $\mu_n$  be a sequence of measures, then  $\mu_n \Rightarrow \mu$  if and only  $\phi_{\mu_n}(t) \rightarrow \phi_\mu(t)$  for all  $t$ . This is called the continuity theorem.

### 1.3 Motivation

Regarding characteristic functions, one might ask “what’s it about?”. One answer is that we defined  $\mu_n \Rightarrow \mu$  by  $\int f d\mu_n \rightarrow \int f d\mu$  for all continuous bounded  $f$ . The continuity theorem says we only need to consider the very special functions  $f_t(x) = e^{itx}$ . The usefulness of this can be seen in the following example.

**Example 1.** Let’s return to our card guessing ESP example. Suppose we have card  $\{1, 2, \dots, n\}$  which are shuffled and turned over one at a time. We have seen that with complete feedback there is probability  $\frac{1}{n-i+1}$  of correctly guessing the  $i^{th}$  card to be turned over. We previously considered examples where guessing card  $i$  correctly results in a constant reward. We will now consider what happens if rewards are proportionate to difficulty.

With this in mind, define independent random variables  $(\tilde{X}_i)_{i=1}^\infty$  by

$$\tilde{X}_i = \begin{cases} i & \text{with probability } \frac{1}{i}, \\ 0 & \text{with probability } 1 - \frac{1}{i}. \end{cases}$$

Then  $\mathbb{E}[\tilde{X}_i] = 1$  and

$$\text{Var}(\tilde{X}_i) = i^2 \left( \frac{1}{i} \left( 1 - \frac{1}{i} \right) \right) = i - 1.$$

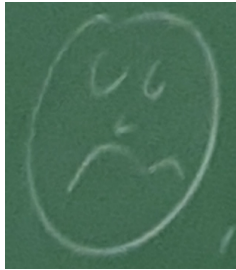
Next define  $X_i = \tilde{X}_i - 1 = \tilde{X}_i - \mathbb{E}[\tilde{X}_i]$  so that  $X_i$  has mean 0. We can try to apply the central limit theorem to  $X_i$ . To do this we need to check Lindeberg's condition. With notation as in the statement of the central limit theorem we have

$$S_n = \sum_{i=1}^n X_i, \quad \sigma_i = \sqrt{V(X_i)} = \sqrt{i-1} \quad \text{and} \quad s_n^2 = \sum_{i=1}^n \sigma_i^2 = \binom{n}{2}.$$

For  $\varepsilon = \frac{\sqrt{2}}{2}$  we have

$$\begin{aligned} \frac{1}{s_n^2} \sum_{i=1}^n \int_{\{|X_n| > \varepsilon s_n\}} |X_i|^2 d\mathbb{P} &= \frac{2}{n(n-1)} \sum_{i=1}^n \int_{\{|X_n| > \varepsilon \sqrt{n(n-1)/2}\}} |X_i|^2 d\mathbb{P} \\ &= \frac{2}{n(n-1)} \sum_{i=1}^n \int_{\{|X_n| > \frac{1}{2} \sqrt{n(n-1)}\}} |X_i|^2 d\mathbb{P} \\ &\geq \frac{2}{n^2} \sum_{i=\lceil n/2 \rceil}^n \int_{\{|X_n| > \frac{1}{2} n\}} |X_i|^2 d\mathbb{P} \\ &= \frac{2}{n^2} \sum_{i=\lceil n/2 \rceil}^n \frac{1}{i} (i-1)^2 \\ &\not\rightarrow 0. \end{aligned}$$

This Lindeberg's condition does not hold for  $\varepsilon = \frac{\sqrt{2}}{2}$  and thus the central limit theorem does not hold:



But we can still use characteristic functions! Note that

$$\begin{aligned} \mathbb{E}[e^{itX_j}] &= e^{it(j-1)} \frac{1}{j} + e^{-it} \left(1 - \frac{1}{j}\right) \\ &= \frac{e^{-it}}{j} (e^{itj} + j - 1) \end{aligned}$$