Stefan Grosskinsky, Emma Southall 10.10.2019

Stochastic Modelling and Random Processes

Hand-out 2

Characteristic function, Gaussians, LLN, CLT

Let X be a real-valued random variable with PDF f_X. The **characteristic function** $\phi_X(t)$ is defined as the Fourier transform of the PDF, i.e.

$$\phi_X(t) = \mathbb{E}\big(e^{itX}\big) = \int_{-\infty}^{\infty} e^{itx} f_X(x)\, dx \quad \text{for all } t \in \mathbb{R}\ .$$

As the name suggests, ϕ_X uniquely determines (characterizes) the distribution of X and the usual inversion formula for Fourier transforms holds,

$$f_X(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}\phi_X(t)\, dt \quad \text{for all } x \in \mathbb{R}\ .$$

The characteristic function plays a similar role as the probability generating function for discrete random variables. Moments can be recovered via

$$\tfrac{\partial^k}{\partial t^k}\phi_X(t) = (i)^k\mathbb{E}(X^k e^{itX}) \quad \Rightarrow \quad \mathbb{E}(X^k) = (i)^{-k}\tfrac{\partial^k}{\partial t^k}\phi_X(t)\big|_{t=0}\ . \tag{1}$$

Also, if we add independent random variables X and Y, their characteristic functions multiply,

$$\phi_{X+Y}(t) = \mathbb{E}\big(e^{it(X+Y)}\big) = \phi_X(t)\,\phi_Y(t)\ . \tag{2}$$

Furthermore, for a sequence $X_1, X_2, \ldots$ of real-valued random variables we have

$$X_n \to X \quad \text{in distribution, i.e.} \quad f_{X_n}(x) \to f_X(x)\ \forall x \in \mathbb{R} \quad \Leftrightarrow \quad \phi_{X_n}(t) \to \phi_X(t)\ \forall t \in \mathbb{R}\ . \tag{3}$$

A real-valued random variable $X \sim N(\mu, \sigma^2)$ has **normal** or **Gaussian** distribution with mean $\mu \in \mathbb{R}$ and variance $\sigma^2 \geq 0$ if its PDF is of the form

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\Big(-\frac{(x-\mu)^2}{2\sigma^2}\Big)\ .$$

Properties.

- The characteristic function of $X \sim N(\mu, \sigma^2)$ is given by

$$\phi_X(t) = \frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty}\exp\Big(-\frac{(x-\mu)^2}{2\sigma^2} + itx\Big)\, dx = \exp\Big(i\mu t - \frac{1}{2}\sigma^2 t^2\Big)\ .$$

  To see this (try it!), you have to complete the squares in the exponent to get

$$-\frac{1}{2\sigma^2}\big(x - (it\sigma^2 + \mu)\big)^2 - \frac{1}{2}t^2\sigma^2 + it\mu\ ,$$

  and then use that the integral over x after re-centering is still normalized.

- This implies that linear combinations of independent Gaussians $X_1,\ X_2$ are Gaussian, i.e.

$$X_i \sim N(\mu_i, \sigma_i^2),\ a, b \in \mathbb{R} \quad \Rightarrow \quad aX_1 + bX_2 \sim N\big(a\mu_1 + b\mu_2, a^2\sigma_1^2 + b^2\sigma_2^2\big)\ .$$

  This holds even for correlated X_i, where the variance is $a^2\sigma_1^2 + b^2\sigma_2^2 + 2ab\mathrm{Cov}(X_1, X_2)$ with covariance $\mathrm{Cov}(X_1, X_2) = \mathbb{E}\big((X_1 - \mu_1)(X_2 - \mu_2)\big)$ and the mean remains unchanged.

Let $X_1, X_2, \ldots$ be a sequence of iidrv's with mean μ and variance σ^2 and set $S_n = X_1 + \ldots + X_n$. The following two important limit theorems are a direct consequence of the above.

Weak law of large numbers (LLN)

$$S_n/n \to \mu \quad \textit{in distribution as } n \to \infty\ .$$

There exists also a strong form of the LLN with almost sure convergence which is harder to prove.

Central limit theorem (CLT)

$$\frac{S_n - \mu n}{\sigma\sqrt{n}} \to N(0,1) \quad \textit{in distribution as } n \to \infty\ .$$

The LLN and CLT imply that for $n \to \infty$, $\quad S_n \simeq \mu n + \sigma\sqrt{n}\,\xi \quad$ with $\quad \xi \sim N(0,1)$.

Proof. With $\phi(t) = \mathbb{E}\big(e^{itX_i}\big)$ we have from (2)

$$\phi_n(t) := \mathbb{E}\big(e^{itS_n/n}\big) = \big(\phi(t/n)\big)^n\ .$$

(1) implies the following Taylor expansion of ϕ around 0:

$$\phi(t/n) = 1 + i\mu\frac{t}{n} - \frac{\sigma^2}{2}\frac{t^2}{n^2} + o(t^2/n^2)\ ,$$

of which we only have to use the first order to see that

$$\phi_n(t) = \Big(1 + i\mu\frac{t}{n} + o(t/n)\Big)^n \to e^{it\mu} \quad \text{as } n \to \infty\ .$$

By (3) and uniqueness of characteristic functions this implies the LLN.
To show the CLT, set $\quad Y_i = \dfrac{X_i - \mu}{\sigma} \quad$ and write $\quad \tilde{S}_n = \displaystyle\sum_{i=1}^n Y_i = \frac{S_n - \mu n}{\sigma}$.
Then, since $\mathbb{E}(Y_i) = 0$, the corresponding Taylor expansion (now to second order) leads to

$$\phi_n(t) := \mathbb{E}\big(e^{it\tilde{S}_n/\sqrt{n}}\big) = \Big(1 - \frac{t^2}{2n} + o(t^2/n)\Big)^n \to e^{-t^2/2} \quad \text{as } n \to \infty\ ,$$

which implies the CLT. $\square$

Multivariate case.
All of the above can be generalized to multivariate random variables in $\mathbb{R}^d$. $\mathbf{X} = (X_1, \ldots, X_d)$ is a **multivariate Gaussian** if it has PDF

$$f_X(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^d \det\Sigma}} \exp\Big(-\frac{1}{2}\langle \mathbf{x} - \boldsymbol{\mu} | \Sigma^{-1} | \mathbf{x} - \boldsymbol{\mu}\rangle \Big) \quad \text{with} \quad \mathbf{x} = (x_1, \ldots, x_d)\ ,$$

where $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_d)$ is the vector of means $\mu_i = \mathbb{E}(X_i)$ and $\Sigma = (\sigma_{ij} : i,j = 1, \ldots, d)$ is the **covariance matrix** with entries

$$\sigma_{ij} = \mathrm{Cov}(X_i, X_j) = \mathbb{E}\big((X_i - \mu_i)(X_j - \mu_j)\big)\ .$$

Σ is symmetric and invertible (unless in degenerate cases with vanishing variance). The characteristic function of $\mathbf{X}$ is given by

$$\phi_X(\mathbf{t}) = \mathbb{E}\big(e^{i\langle \mathbf{t}|\mathbf{X}\rangle}\big) = \exp\Big(i\langle \mathbf{t}|\boldsymbol{\mu}\rangle - \frac{1}{2}\langle \mathbf{t}|\Sigma|\mathbf{t}\rangle \Big)\ , \quad \mathbf{t} \in \mathbb{R}^d\ .$$

(In the above notation $\langle \mathbf{x}|$ is a row and $|\mathbf{x}\rangle$ a column vector, and scalar products are written as $\langle \mathbf{x}|\mathbf{x}\rangle$.)