pfsl7(15).tex

Lecture 7. 27.10.2015

2. *Normal* $N(0,1)$. This has MGF

$$\begin{aligned} M(t) &= \frac{1}{\sqrt{2\pi}} \int \exp\{tx - \frac{1}{2}x^2\} dx \\ &= \frac{1}{\sqrt{2\pi}} \int \exp\{-\frac{1}{2}(x-t)^2 + \frac{1}{2}t^2\} dx \quad \text{(completing the square)} \\ &= \exp\{\frac{1}{2}t^2\} . \frac{1}{\sqrt{2\pi}} \int \exp\{-\frac{1}{2}u^2\} du \qquad (u := x - t) \\ &= \exp\{\frac{1}{2}t^2\} \qquad \text{(normal density).} \end{aligned}$$

This holds for all t, real or complex. So here we have $R = \infty$, and M is entire.

CFs

Definition. The *characteristic function (CF)* of a random variable X is

$$\phi(t), \text{ or } \phi_X(t), := E[e^{itX}],$$

where here t is real and $i := \sqrt{-1}$.

Note. 1. The notation here (which is standard) clashes with our use of $\phi(x)$ as the standard normal density. But which is meant will be clear, both from context and from the use of x or t for the argument.

2. The substantial change is from the *real* exponential e^{tX} to the *complex* exponential e^{itX}. Complex numbers are met later than real numbers, and Complex Analysis later than Real analysis. But:

(a) A knowledge of Complex Analysis is essential in Mathematics, so there is no point in trying to avoid it (everything we need, and much more, is on my website under M2P3);

(b) Some parts of Real Analysis – the most relevant to us being the theory of power series (Maclaurin or Taylor series, above) – actually belong to Complex Analysis, as they canot be understood except in a complex setting.

Properties of CFs. The CF has a number of important properties.

1. *Existence.* The CF always exists (the integral defining it always converges). Indeed,

$$|\phi(t)| = |\int e^{itx} dF(x)| \le \int |e^{itx}| dF(x) = \int 1 dF(x) = 1.$$

This property – *existence* of the CF, everywhere – is extremely important, and useful. This advantage vastly outweighs any inconvenience resulting from the use of complex numbers, even if these are initially unfamiliar.
2. *Continuity.* The CF is continuous, indeed uniformly continuous:

$$\begin{aligned} |\phi(t+u) - \phi(t)| &= |\int e^{itx}(e^{itu} - 1)dF(x)| \le |e^{itu} - 1| \int 1dF(x) \\ &= |e^{itu} - 1| \to 0 \qquad (u \to 0). \end{aligned}$$

3. *Uniqueness.* The CF determines the distribution function uniquely (so taking the CF loses no information). This is a general property of Fourier transforms; we quote this.
4. *Inversion formula.* There is an inversion formula (due to Lévy, 1937) giving the distribution function in terms of the CF. We omit this, as the formula is rarely useful.
5. *Continuity theorem* (Lévy, 1937). (i) If F_n, F have CFs ϕ_n, ϕ, and $F_n \to F$ in distribution (see III.1 below), then

$$\phi_n(t) \to \phi(t) \qquad (n \to \infty) \qquad \text{uniformly in } t \text{ on compact sets.}$$

(ii) Conversely, if $\phi_n(t) \to \phi(t)$ pointwise, and the limit function $\phi(t)$ is continuous at $t = 0$, then ϕ is the CF of a distribution function, F say, and $F_n \to F$ in distribution.
6. *Moments.* For a random variable X, the kth *moment* of X is defined by

$$\mu_k := E[X^k].$$

The first moment is the *mean* or *expectation*, $\mu = E[X]$. (We use notation such as μ_X if there are other random variables present. Context will show whether μ denotes a mean or a measure.) If X has k moments (finite), we can expand the exponential e^{itX} in the definition of the CF and get $\sum_{j=0}^{k}(it)^j.E[X^j]/j!$ or $\sum_{j=0}^{k}(it)^j\mu_j/j!$, plus an error term. Analogy with Taylor's Theorem in Real Analysis suggests that this error term should be $o(t^k)$ at $t \to 0$. This is true; we quote it: if X has k moments finite, its CF satisfies

$$\phi^{(j)}(0) = i^j\mu_j \quad (j = 0, 1, \ldots, k), \quad \phi(t) = \sum_{j=0}^{k}(it)^j\mu_j/j! + o(t^k) \quad (t \to 0).$$

7. *Convolutions.* As with the MGF (when it exists!): for X, Y independent,

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

– the CF of an independent sum is the product of the CFs. Thus *addition* of independent random variables (easy) corresponds to *convolution* of distributions (and this involves an integration – awkward, if repeated many times), but to *multiplication* of CFs (or MGFs) – easy.
Examples.
1. $N(0,1)$. By above, the MGF is $e^{\frac{1}{2}t^2}$. Formally replacing t by it here, this suggests that the CF is $e^{-\frac{1}{2}t^2}$:

$$\int_{-\infty}^{\infty} e^{itx}.\frac{e^{-\frac{1}{2}x^2}}{\sqrt{2\pi}}dx = e^{-\frac{1}{2}t^2}.$$

This is in fact true, and the above argument works. This is because the MGF is entire (analytic everywhere), and we can use analytic continuation.

An alternative argument proves the above by using Cauchy's Theorem (see e.g. [BF] p.21, or M2P3 L1 and 26-27).
1a. $N(\mu, \sigma^2)$. This has CF $\exp\{i\mu t - \frac{1}{2}\sigma^2 t^2\}$. For, if $X \sim N(\mu, \sigma^2)$, $(X - \mu)/\sigma \sim N(0, 1)$, so

$$E[e^{it(X-\mu)/\sigma}] = e^{-\frac{1}{2}t^2}: \qquad E[e^{it(X-\mu)}] = e^{-\frac{1}{2}\sigma^2 t^2}; \qquad E[e^{itX}] = e^{i\mu t - \frac{1}{2}\sigma^2 t^2}.$$

$$\phi'(t) = (i\mu - \sigma^2 t)\phi(t), \qquad \phi'(0) = i\mu = iE[X]: \quad E[X] = \mu.$$

So $var(X) = E[(X - \mu)^2]$; $X - \mu$ has CF $\phi_0(t) = e^{-\frac{1}{2}\sigma^2 t^2}$. As

$$\phi_0'(t) = -\sigma^2 t e^{-\frac{1}{2}\sigma^2 t^2}, \quad \phi_0''(t) = -\sigma^2 e^{-\frac{1}{2}\sigma^2 t^2} - \sigma^2 t e^{-\frac{1}{2}\sigma^2 t^2},$$

$$\phi_0''(0) = -\sigma^2, \quad = -E[(X - \mu)^2] = -var(X): \qquad var(X) = \sigma^2.$$

2. $\chi^2(n)$. The MGF is $1/(1-2t)^{\frac{1}{2}n}$ in $t < \frac{1}{2}$, and this extends to the half-plane $Re\ t < \frac{1}{2}$ in the complex t-plane. By analytic continuation, this is enough to give the CF as $1/(1 - 2it)^{\frac{1}{2}n}$ for all real t. Observe how the singularity in the MGF is avoided in the CF!
3. *Cauchy distribution.* Here

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

The mean does not exist! (The integral $\int xf(x)dx$ diverges logarithmically, so does not exist as a Lebesgue integral. So the mean does not exist – despite the symmetry, which means that the median and mode both exist and are 0.) One can show (Problems) that the CF is

$$\phi(t) = \int_{-\infty}^{\infty} \frac{e^{ixt}}{\pi(1+x^2)} dx = e^{-|t|}.$$

This is not differentiable at 0 – reflecting the non-existence of the mean, by Property 6 above on moments.

4. *Symmetric exponential distribution*, $SE(x)$. Here

$$f(x) = \frac{1}{2} e^{-|x|}.$$

This has CF (Problems)

$$\phi(t) = \int_{-\infty}^{\infty} e^{ixt} . \frac{1}{2} e^{-|x|} dx = 1/(1+t^2).$$

Note. 1. The CF is the *Fourier transform* of the density, when this exists (and the *Fourier-Stieltjes transform* of the distribution function, in general.

The MGF is related to the *Laplace transform* (or Laplace-Stieltjes transform), where one has e^{-xt} in place of e^{xt}. These *integral transforms* are useful, and widely used; you may have met them already.

2. Notice the similarity between the last two examples! Apart from a factor of 2π, it looks as if we are 'Fourier transforming twice and getting back to where we started'. This is indeed the case, and is an instance of a general result, the *Fourier Integral Theorem.*

Higher dimensions

In n dimensions, the arguments x, t of the density and MGF or CF are both n-vectors – column-vectors, say. We replace e^{ixt}, etc. by $e^{it^T x}$, using the superscript T for 'transpose'. Thus t^T is a row-vector, and $t^T x$ is a scalar. The theory goes through as before. We will return to this later in connection with the very important *multivariate normal* (multinormal) distribution.

PGFs

We often meet random variables X which take only non-negative integer values. The distribution is then specified by

$$p_n := P(X = n), \qquad n = 0, 1, \ldots; \qquad \sum_{n=0}^{\infty} p_n = 1.$$