STAT 801 Lecture 6

Reading for Today's Lecture:
Goals of Today's Lecture:

- Characteristic Functions
- Review of complex arithmetic
- Inversion formulas

Today's notes

Last time defined moments, moment generating functions, cumulants and cumulant generating functions. We established the following relations between moments and cumulants:

$$\begin{aligned}
\kappa_1(X) &= E(X) \\
\kappa_2(X) &= \mathrm{Var}(X) \\
\kappa_3(X) &= E[(X - E(X))^3] \\
\kappa_4(X) &= E[(X - E(X))^4] - 3[\mathrm{Var}(X)]^2
\end{aligned}$$

Example: I am having you derive the moment and cumulant generating function and all the moments of a Gamma rv. Suppose that $Z_1, \ldots, Z_\nu$ are independent $N(0,1)$ rvs. Then we have defined $S_\nu = \sum_1^\nu Z_i^2$ to have a χ^2 distribution. It is easy to check $S_1 = Z_1^2$ has density

$$(u/2)^{-1/2} e^{-u/2}/(2\sqrt{\pi})$$

and then the mgf of S_1 is

$$(1-2t)^{-1/2}$$

It follows that

$$M_{S_\nu}(t) = (1-2t)^{-\nu/2};$$

you will show in homework that this is the mgf of a Gamma$(\nu/2, 2)$ rv. This shows that the χ^2_ν distribution has the Gamma$(\nu/2, 2)$ density which is

$$(u/2)^{(\nu-2)/2} e^{-u/2}/(2\Gamma(\nu/2)).$$

Example: The Cauchy density is

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

and the corresponding moment generating function is

$$M(t) = \int_{-\infty}^{\infty} \frac{e^{tx}}{\pi(1+x^2)} dx$$

which is $+\infty$ except for $t = 0$ where we get 1. This mgf is exactly the mgf of *every* t distribution so it is not much use for distinguishing such distributions. The problem is that these distributions do not have infinitely many finite moments.

This observation has led to the development of a substitute for the mgf which is defined for every distribution, namely, the characteristic function.

Characteristic Functions

Definition: The characteristic function of a real rv X is

$$\phi_X(t) = E(e^{itX})$$

where $i = \sqrt{-1}$ is the imaginary unit.

Aside on complex arithmetic.

The complex numbers are the things you get if you add $i = \sqrt{-1}$ to the real numbers and require that all the usual rules of algebra work. In particular if i and any real numbers a and b are to be complex numbers then so must be $a + bi$. If we multiply a complex number $a + bi$ with a and b real by another such number, say $c + di$ then the usual rules of arithmetic (associative, commutative and distributive laws) require

$$\begin{aligned}(a+bi)(c+di) =& ac + adi + bci + bdi^2 \\ =& ac + bd(-1) + (ad+bc)i \\ =& (ac-bd) + (ad+bc)i\end{aligned}$$

so this is precisely how we define multiplication. Addition is simply (again by following the usual rules)

$$(a+bi) + (c+di) = (a+b) + (c+d)i$$

Notice that the usual rules of arithmetic then don't require any more numbers than things of the form

$$x + yi$$

where x and y are real. We can identify a single such number $x + yi$ with the corresponding point (x, y) in the plane. It often helps to picture the complex numbers as forming a plane.

Now look at transcendental functions. For real x we know $e^x = \sum x^k/k!$ so our insistence on the usual rules working means

$$e^{x+iy} = e^x e^{iy}$$

and we need to know how to compute e^{iy}. Remember in what follows that $i^2 = -1$ so $i^3 = -i$, $i^4 = 1$ $i^5 = i^1 = i$ and so on. Then

$$\begin{aligned} e^{iy} =& \sum_0^\infty \frac{(iy)^k}{k!} \\ =& 1 + iy + (iy)^2/2 + (iy)^3/6 + \cdots \\ =& 1 - y^2/2 + y^4/4! - y^6 + \cdots \\ & + iy - iy^3/3! + iy^5/5! + \cdots \\ =& \cos(y) + i\sin(y) \end{aligned}$$

We can thus write

$$e^{x+iy} = e^x(\cos(y) + i\sin(y))$$

Now every point in the plane can be written in polar co-ordinates as $(r\cos\theta, r\sin\theta)$ and comparing this with our formula for the exponential we see we can write

$$x + iy = \sqrt{x^2 + y^2} e^{i\theta}$$

for an angle $\theta \in [0, 2\pi)$.

We will need from time to time a couple of other definitions:

Definition: The **modulus** of the complex number $x + iy$ is

$$|x + iy| = \sqrt{x^2 + y^2}$$

Definition: The **complex conjugate** of $x + iy$ is $\overline{x + iy} = x - iy$.

Notes on calculus with complex variables. Essentially the usual rules apply so, for example,

$$\frac{d}{dt} e^{it} = ie^{it}$$

We will (mostly) be doing only integrals over the real line; the theory of integrals along paths in the complex plane is a very important part of mathematics, however.

End of Aside

Since

$$e^{itX} = \cos(tX) + i\sin(tX)$$

we find that

$$\phi_X(t) = E(\cos(tX)) + iE(\sin(tX)) .$$

Since the trigonometric functions are bounded by 1 the expected values must be finite for all t and this is precisely the reason for using characteristic rather than moment generating functions in probability theory courses.

Theorem 1 *For any two real rvs X and Y the following are equivalent:*

1. *X and Y have the same distribution, that is, for any (Borel) set A we have*

$$P(X \in A) = P(Y \in A)$$

2. *$F_X(t) = F_Y(t)$ for all t.*

3. *$\phi_X = E(e^{itX}) = E(e^{itY}) = \phi_Y(t)$ for all real t.*

Moreover, all of these are implied if there is a positive ϵ such that for all $|t| \le \epsilon$

$$M_X(t) = M_Y(t) < \infty .$$

Inversion

The previous theorem is a non-constructive characterization. It does not show us how to get from ϕ_X to F_X or f_X. For cdfs or densities with reasonable properties, however, there are effective ways to compute F or f from ϕ. In homework I am asking you to prove the following basic **inversion** formula: If X is a random variable taking only integer values then for each integer k

$$\begin{aligned} P(X = k) &= \frac{1}{2\pi}\int_0^{2\pi} \phi_X(t)e^{-itk}dt \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \phi_X(t)e^{-itk}dt . \end{aligned}$$

The proof proceeds from the formula

$$\phi_X(t) = \sum_k e^{ikt} P(X = k)\,.$$

Now suppose that X has a continuous bounded density f. Define

$$X_n = [nX]/n$$

where $[a]$ denotes the integer part (rounding down to the next smallest integer). We have

$$\begin{aligned} P(k/n \le X < (k+1)/n) &= P([nX] = k) \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \phi_{[nX]}(t)e^{-itk}dt\,. \end{aligned}$$

Make the substitution $t = u/n$, and get

$$nP(k/n \le X < (k+1)/n) = \frac{1}{2\pi}\int_{-n\pi}^{n\pi} \phi_{[nX]}(u/n)e^{iuk/n}du$$

Now, as $n \to \infty$ we have

$$\phi_{[nX]}(u/n) = E(e^{iu[nX]/n}) \to E(e^{iuX})$$

(by the dominated convergence theorem – the dominating random variable is just the constant 1). The range of integration converges to the whole real line and if $k/n \to x$ we see that the left hand side converges to the density $f(x)$ while the right hand side converges to

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} \phi_X(u)e^{-iux}du$$

which gives the inversion formula

$$f_X(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \phi_X(u)e^{-iux}du$$

Many other such formulas are available to compute things like $F(b) - F(a)$ and so on.

All such formulas are sometimes referred to as Fourier inversion formulas; the characteristic function itself is sometimes called the Fourier transform of the distribution or cdf or density of X.

Inversion of the Moment Generating Function

The moment generating function and the characteristic function are related formally by

$$M_X(it) = \phi_X(t)$$

When M_X exists this relationship is not merely formal; the methods of complex variables mean there is a "nice" (analytic) function which is $E(e^{zX})$ for any complex $z = x + iy$ for which $M_X(x)$ is finite. All this means that there is an inversion formula for M_X. This formula requires a complex *contour integral*. In general if z_1 and z_2 are two points in the complex plane and C a path between these two points we can define the path integral

$$\int_C f(z)dz$$

by the methods of line integration. When it comes to doing algebra with such integrals the usual theorems of calculus still work. The Fourier inversion formula was

$$2\pi f(x) = \int_{-\infty}^{\infty} \phi(t)e^{-itx}dt$$

so replacing ϕ by M we get

$$2\pi f(x) = \int_{-\infty}^{\infty} M(it)e^{-itx}dt$$

If we just substitute $z = it$ then we find

$$2\pi i f(x) = \int_C M(z)e^{-zx}dz$$

where the path C is the imaginary axis. This formula becomes of use by the methods of complex integration which permit us to replace the path C by any other path which starts and ends at the same place. It is possible, in some cases, to choose this path to make it easy to do the integral approximately; this is what saddlepoint approximations are. This inversion formula is called the inverse Laplace transform; the mgf is also called the Laplace transform of the distribution or of the cdf or of the density.

Applications of Inversion

1): Numerical calculations

Example: Many statistics have a distribution which is approximately that of

$$T = \sum \lambda_j Z_j^2$$

where the Z_j are iid $N(0,1)$. In this case

$$\begin{aligned} E(e^{itT}) &= \prod E(e^{it\lambda_j Z_j^2}) \\ &= \prod (1-2it\lambda_j)^{-1/2} . \end{aligned}$$

Imhof (*Biometrika*, 1961) gives a simplification of the Fourier inversion formula for

$$F_T(x) - F_T(0)$$

which can be evaluated numerically.

2): The central limit theorem (in some versions) can be deduced from the Fourier inversion formula: if $X_1, \ldots , X_n$ are iid with mean 0 and variance 1 and $T = n^{1/2}\bar{X}$ then with ϕ denoting the characteristic function of a single X we have

$$\begin{aligned} E(e^{itT} &= E(e^{in^{-1/2}t\sum X_j}) \\ &= \left[\phi(n^{-1/2}t)\right]^n \\ &\approx \left[\phi(0) + n^{-1/2}t\phi'(0) + n^{-1}t^2\phi''(0)/2 + o(n^{-1}\right]^n \end{aligned}$$

But now $\phi(0) = 1$ and

$$\phi'(t) = \frac{d}{dt}E(e^{itX_1}) = iE(X_1 e^{itX_1})$$

So $\phi'(0) = E(X_1) = 0$. Similarly

$$\phi''(t) = i^2 E(X_1^2 e^{itX_1})$$

so that

$$\phi''(0) = -E(X_1^2) = -1$$

It now follows that

$$\begin{aligned} E(e^{itT}) &\approx [1 - t^2/(2n) + o(1/n)]^n \\ &\to e^{-t^2/2} \end{aligned}$$

With care we can then apply the Fourier inversion formula and get

$$\begin{aligned}
f_T(x) &= \frac{1}{2\pi i}\int_{-\infty}^{\infty} e^{-itx}[\phi(tn^{-1/2})]^n dt \\
&\to \frac{1}{2\pi i}\int_{-\infty}^{\infty} e^{-itx}e^{-t^2/2}dt \\
&= \frac{1}{\sqrt{2\pi}}\phi_Z(-x)
\end{aligned}$$

where ϕ_Z is the characteristic function of a standard normal variable Z. Doing the integral we find

$$\phi_Z(x) = \phi_Z(-x) = e^{-x^2/2}$$

so that

$$f_T(x) \to \frac{1}{\sqrt{2\pi}}e^{-x^2/2}$$

which is a standard normal random variable.

This proof of the central limit theorem is not terribly general since it requires T to have a bounded continuous density. The central limit theorem itself is a statement about cdfs not densities and is

$$P(T \le t) \to P(Z \le t)$$