STAT 801: Mathematical Statistics

Inversion of Generating Functions

Previous theorem is non-constructive characterization. Can get from ϕ_X to F_X or f_X by **inversion**. See homework for 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

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

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

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

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

MGF and characteristic function related formally:

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

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.

SO: there is an inversion formula for M_X using a complex *contour integral*:

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.

Do algebra with such integrals via usual theorems of calculus. 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. Methods of complex integration permit us to replace C by any other path which starts and ends at the same place. Sometimes can choose path to make it easy to do the integral approximately; this is what saddlepoint approximations are. Inversion formula is called the inverse Laplace transform; the mgf is also called the Laplace transform of the distribution or cdf or 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:

$$\begin{aligned} F_T(x) - F_T(0) &= \int_0^x f_T(y)dy \\ &= \int_0^x \frac{1}{2\pi} \int_{-\infty}^{\infty} \times \prod (1 - 2it\lambda_j)^{-1/2} e^{-ity} dt dy \end{aligned}$$

Multiply

$$\phi(t) = \left[\frac{1}{\prod(1-2it\lambda_j)}\right]^{1/2}$$

top and bottom by the complex conjugate of the denominator:

$$\phi(t) = \left[\frac{\prod(1+2it\lambda_j)}{\prod(1+4t^2\lambda_j^2)}\right]^{1/2}$$

The complex number $1+2it\lambda_j$ is $r_j e^{i\theta_j}$ where $r_j = \sqrt{1+4t^4\lambda_j^2}$ and $\tan(\theta_j) = 2t\lambda_j$ This allows us to rewrite

$$\phi(t) = \left[\frac{\prod r_j e^{i\sum\theta_j}}{\prod r_j^2}\right]^{1/2}$$

or

$$\phi(t) = \frac{e^{i\sum\tan^{-1}(2t\lambda_j)/2}}{\prod(1+4t^2\lambda_j^2)^{1/4}}$$

Assemble this to give

$$F_T(x) - F_T(0) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{e^{i\theta(t)}}{\rho(t)}\int_0^x e^{-iyt}dydt$$

where

$$\theta(t) = \sum\tan^{-1}(2t\lambda_j)/2$$

and $\rho(t) = \prod(1+4t^2\lambda_j^2)^{1/4}$. But

$$\int_0^x e^{-iyt}dy = \frac{e^{-ixt}-1}{-it}$$

We can now collect up the real part of the resulting integral to derive the formula given by Imhof. I don't produce the details here.

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

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

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

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

so that

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

It now follows that

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

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

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

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

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

3) Saddlepoint approximation from MGF inversion formula

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

(limits of integration indicate contour integral running up imaginary axis.) Replace contour (using complex variables) with line $Re(z) = c$. ($Re(Z)$ denotes the real part of z, that is, x when $z = x + iy$ with x and y real.) Must choose c so that $M(c) < \infty$. Rewrite inversion formula using cumulant generating function $K(t) = \log(M(t))$:

$$2\pi i f(x) = \int_{c-i\infty}^{c+i\infty} \exp(K(z) - zx)dz\,.$$

Along the contour in question we have $z = c + iy$ so we can think of the integral as being

$$i\int_{-\infty}^{\infty} \exp(K(c+iy) - (c+iy)x)dy$$

Now do a Taylor expansion of the exponent:

$$K(c+iy) - (c+iy)x = K(c) - cx + iy(K'(c) - x) - y^2K''(c)/2 + \cdots$$

Ignore the higher order terms and select a c so that the first derivative

$$K'(c) - x$$

vanishes. Such a c is a saddlepoint. We get the formula

$$2\pi f(x) \approx \exp(K(c) - cx) \times \int_{-\infty}^{\infty} \exp(-y^2K''(c)/2)dy.$$

The integral is just a normal density calculation and gives $\sqrt{2\pi/K''(c)}$. The saddlepoint approximation is

$$f(x) = \frac{\exp(K(c) - cx)}{\sqrt{2\pi K''(c)}}.$$

Essentially the same idea lies at the heart of the proof of Sterling's approximation to the factorial function:

$$n! = \int_0^\infty \exp(n\log(x) - x)dx$$

The exponent is maximized when $x = n$. For n large we approximate $f(x) = n\log(x) - x$ by

$$f(x) \approx f(x_0) + (x - x_0)f'(x_0) + (x - x_0)^2 f''(x_0)/2$$

and choose $x_0 = n$ to make $f'(x_0) = 0$. Then

$$n! \approx \int_0^\infty \exp[n\log(n) - n - (x-n)^2/(2n)]dx$$

Substitute $y = (x - n)/\sqrt{n}$ to get the approximation

$$n! \approx n^{1/2} n^n e^{-n} \int_{-\infty}^\infty e^{-y^2/2} dy$$

or

$$n! \approx \sqrt{2\pi} n^{n+1/2} e^{-n}$$

This tactic is called Laplace's method. Note that I am being very sloppy about the limits of integration; to do the thing properly you have to prove that the integral over x not near n is negligible.