Moment Generating functions
and proof of the Central Limit Theorem

Recall CLT: if $X_1, X_2, \dots, X_n, \dots$ is a sequence of independent identically distributed random variables, with mean μ and variance σ^2

$$Y_n = \frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}} \xrightarrow{\text{distribution}} \text{standard normal random variable } Z$$

$$F_{Y_n}(a) \longrightarrow \Phi(a) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-t^2/2}\, dt$$

<u>Recall</u>: X and Y are independent random variables with density functions f_X and f_Y

then $X+Y$ has density function $f_X * f_Y$

$$f_X * f_Y(a) = \int_{-\infty}^{\infty} f_X(x) f_Y(a-x)\, dx$$

Nice fact about convolution: Fourier transform of convolution is the product of the Fourier transforms.

$f(x) \rightarrow$ Fourier transform $\hat{f}(\xi) = \int_{-\infty}^{\infty} e^{-2\pi i \xi x} f(x)\,dx$

$\hat{f}(\xi) \rightarrow$ inverse transform $f(x) = \int_{-\infty}^{\infty} e^{2\pi i \xi x} \hat{f}(\xi)\,d\xi$

One version of Fourier transform in probability theory is the moment generating function

Let X be a random variable

The moment generating function of X is

$$M_X(t) = E[e^{tX}] = \begin{cases} \sum_x e^{tx} P\{X=x\} & \text{discrete} \\ \int_{-\infty}^{\infty} e^{tx} f_X(x)\,dx & \text{continuous} \end{cases}$$

I.e., for continuous X w/ density $f_X(x)$

$$M_X(-2\pi i \xi) = \hat{f}_X(\xi)$$

Why the name? Moments of distribution

$$n\text{th moment} = E[X^n] = \begin{cases} \sum x^n p(x) \\ \int x^n f(x)\,dx \end{cases}$$

("higher order variance")

$$M_X(t) = E[e^{tX}] = E\left[\sum_{n=0}^{\infty} \frac{(tX)^n}{n!}\right] = \sum_{n=0}^{\infty} E[X^n] \frac{t^n}{n!}$$

So

$M_X(t)$ is a function whose Taylor coefficients at $(t=0)$ are exactly the moments of X.

If a_n is any sequence of numbers

$$G(t) = \sum_{n=0}^{\infty} a_n \frac{t^n}{n!} = a_0 + a_1 t + a_2 \frac{t^2}{2} + a_3 \frac{t^3}{6} + \cdots$$

is called exponential generating function of the sequence $\{a_n\}_{n=0}^{\infty}$

$a_0 = G(0) \qquad a_1 = G'(0) \qquad a_2 = G''(0)$

in general $\quad a_n = G^{(n)}(0)$

Point is many combinatorial properties of the sequence a_n translate into properties of generating function

eg. linear recursion relation on a_n's $\iff$ differential equation for the generating function.

Fact: if X and Y are independent R.V.s,

Then $M_{X+Y}(t) = M_X(t) \cdot M_Y(t)$

<u>Proof</u> $E[e^{t(X+Y)}] = E[e^{tX} e^{tY}] = E[e^{tX}]\,E[e^{tY}]$

because e^{tX} and e^{tY} are independent

Ex Binomial dist (n,p)

$$\sum_{k=0}^{n} e^{tk}\binom{n}{k}p^k(1-p)^{n-k} = \sum_{k=0}^{n}\binom{n}{k}(e^t p)^k(1-p)^{n-k}$$

binomial theorem

$$= \left(e^t p + (1-p)\right)^n = M(t)$$

Exponential dist

$$M(t) = \int_0^\infty e^{tx}\,\lambda e^{-\lambda x}\,dx = \lambda\int_0^\infty e^{-(\lambda-t)x}\,dx$$

$$= \lambda\left[-\frac{1}{\lambda-t}e^{-(\lambda-t)x}\right]_0^\infty = \frac{\lambda}{\lambda-t}$$

$$E[X] = M'(0) = \left.\frac{\lambda}{(\lambda-t)^2}\right|_{t=0} = \frac{\lambda}{\lambda^2} = \frac{1}{\lambda}$$

$$E[X^2] = M''(0) = \left.\frac{2\lambda}{(\lambda-t)^3}\right|_{t=0} = \frac{2}{\lambda^2}$$

$$Var(X) = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2}$$

Z Standard normal variable $f_Z(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$

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

$$e^{tx} e^{-x^2/2} = e^{tx - x^2/2} = e^{t^2/2} e^{-\frac{t^2}{2} + tx - \frac{x^2}{2}} = e^{t^2/2} e^{-\frac{(x-t)^2}{2}}$$

$$e^{t^2/2} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(x-t)^2}{2}} dx = e^{t^2/2} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-u^2/2} du$$

$$\left(\begin{matrix} u = x - t \\ du = dx \end{matrix}\right)$$

$$= e^{t^2/2}$$

Levy Continuity theorem

If X_n is a sequence of random variables, X some RV.
Then
if $M_{X_n}(t) \to M_X(t)$ for all t

then $X_n \to X$ in distribution
i.e. $F_{X_n}(a) \to F_X(a)$ for all a at which $F_X(a)$ is continuous

Proof CLT if X_i has mean 0 and variance 1

$M(t) = M_{X_i}(t)$ the common MGF of the X_i's (since identially distribution)

Now consider

$$Y_n = \frac{X_1 + \cdots + X_n - n0}{1\sqrt{n}} = \frac{\sum_{i=1}^n X_i}{\sqrt{n}}$$

$$M_{Y_n}(t) = M_{\frac{X_1+\cdots+X_n}{\sqrt{n}}}(t) = M_{\frac{X_1}{\sqrt{n}}}(t) \cdots M_{\frac{X_n}{\sqrt{n}}}(t)$$

$$M_{\frac{X_i}{\sqrt{n}}}(t) = E\left[\exp\left(t\frac{X_i}{\sqrt{n}}\right)\right] = M_{X_i}\left(\frac{t}{\sqrt{n}}\right) = M\left(\frac{t}{\sqrt{n}}\right)$$

$$M_{Y_n}(t) = \left[M\left(\frac{t}{\sqrt{n}}\right)\right]^n$$

Need to show $\lim_{n\to\infty} \left[M\left(\frac{t}{\sqrt{n}}\right)\right]^n = e^{t^2/2}$

$$L(t) := \log M(t)$$

$$\log\left[M\left(\frac{t}{\sqrt{n}}\right)\right]^n = n\,L\left(\frac{t}{\sqrt{n}}\right)$$

$$\log e^{t^2/2} = t^2/2$$

Need to show $\lim_{n\to\infty} n\,L\left(\frac{t}{\sqrt{n}}\right) = \frac{t^2}{2}$

What is the taylor series for $L(t)$

$$L(0) = \log M(0) = \log 1 = 0$$

$$L'(0) = \frac{M'(0)}{M(0)} = \frac{\mu}{1} = \mu = 0$$

$$L''(0) = \frac{M(0)M''(0) - [M'(0)]^2}{[M(0)]^2} = \frac{1(\mu^2+\sigma^2) - \mu^2}{1^2}$$

$$= \sigma^2 = 1$$

$$L(t) = 0 + 0t + 1\frac{t^2}{2} + a_3\frac{t^3}{3!} + a_4\frac{t^4}{4!} + \cdots$$

Now consider what happens when $L(t) \to nL\left(\frac{t}{\sqrt{n}}\right)$

$$t^k \longrightarrow n\left(\frac{t}{\sqrt{n}}\right)^k = n^{1-\frac{k}{2}} t^k$$

ie. k th term is rescaled by $n^{1-\frac{k}{2}}$

$k=2$ $\quad n^0$

$k=3$ $\quad n^{-1/2}$ $\quad$ as $n \to \infty$ all terms but quadratic go away

$k=4$ $\quad n^{-1}$

$$\lim_{n\to\infty} nL\left(\frac{t}{\sqrt{n}}\right) = \frac{t^2}{2}$$