STAT 450: Statistical Theory

Moment Generating Functions

Def'n: The moment generating function of a real valued X is

$$M_X(t) = E(e^{tX})$$

defined for those real t for which the expected value is finite.

Def'n: The moment generating function of $X \in R^p$ is

$$M_X(u) = E[e^{u^t X}]$$

defined for those vectors u for which the expected value is finite.

Formal connection to moments:

$$M_X(t) = \sum_{k=0}^{\infty} E[(tX)^k]/k!$$

$$= \sum_{k=0}^{\infty} \mu_k' t^k/k!$$

Sometimes can find power series expansion of M_X and read off the moments of X from the coefficients of $t^k/k!$.

Example: : X has density

$$f(u) = u^{\alpha-1}e^{-u}1(u > 0)/\Gamma(\alpha)$$

MGF is

$$M_X(t) = \int_0^\infty e^{tu}u^{\alpha-1}e^{-u}du/\Gamma(\alpha)$$

Substitute $v = u(1-t)$ to get

$$M_X(t) = (1-t)^{-\alpha}$$

For $\alpha = 1$ get exponential distribution.

Have power series expansion

$$M_X(t) = 1/(1-t) = \sum_0^\infty t^k$$

Write $t^k = k!t^k/k!$. Coeff of $t^k/k!$ is

$$\mathrm{E}(X^k) = k!$$

Theorem: If M is finite for all $t \in [-\epsilon, \epsilon]$ for some $\epsilon > 0$ then

1. Every moment of X is finite.
2. M is C^∞ (in fact M is analytic).
3. $\mu_k' = \frac{d^k}{dt^k} M_X(0)$.

Note: C^∞ means has continuous derivatives of all orders. Analytic means has convergent power series expansion in neighbourhood of each $t \in (-\epsilon, \epsilon)$.

Theorem: Suppose X and Y have mgfs M_X and M_Y which are finite for all $t \in [-\epsilon, \epsilon]$. If $M_X(t) = M_Y(t)$ for all $t \in [-\epsilon, \epsilon]$ then X and Y have the same distribution.

The proof, and many other facts about mgfs, rely on techniques of complex variables.

MGFs and Sums

If $X_1,\ldots,X_p$ are independent and $Y=\sum X_i$ then the moment generating function of Y is the product of those of the individual X_i:

$$E(e^{tY}) = \prod_i E(e^{tX_i})$$

or $M_Y = \prod M_{X_i}$.

Note: also true for multivariate X_i.

Problem: power series expansion of M_Y not nice function of expansions of individual M_{X_i}.

Related fact: first 3 moments (meaning μ, σ^2 and μ_3) of Y are sums of those of the X_i:

$$\begin{aligned} E(Y) &= \sum E(X_i) \\ \mathrm{Var}(Y) &= \sum \mathrm{Var}(X_i) \\ E[(Y-E(Y))^3] &= \sum E[(X_i - E(X_i))^3] \end{aligned}$$

but

$$E[(Y-E(Y))^4] = \sum\{E[(X_i-E(X_i))^4] - 3E^2[(X_i-E(X_i))^2]\} + 3\left\{\sum E[(X_i-E(X_i))^2]\right\}^2 .$$

It is possible, however, to replace the moments by other objects called **cumulants** which do add up properly.

Theorem: If X and Y are two random variables such that

$$M_X(t) = M_Y(t)$$

for all $t \in (-epsilon, \epsilon)$ (for some $\epsilon > 0$) then X and Y have the same distribution.

Example: If $X_1,\ldots,X_p$ are independent and X_i has a $N(\mu_i,\sigma_i^2)$ distribution then

$$\begin{aligned} M_{X_i}(t) &= \int_{-\infty}^{\infty} e^{tx} e^{-\frac{1}{2}(x-\mu_i)^2/\sigma_i^2} dx/(\sqrt{2\pi}\sigma_i) \\ &= \int_{-\infty}^{\infty} e^{t(\sigma_i z+\mu_i)} e^{-z^2/2} dz/\sqrt{2\pi} \\ &= e^{t\mu_i} \int_{-\infty}^{\infty} e^{-(z-t\sigma_i)^2/2 + t^2\sigma_i^2/2} dz/\sqrt{2\pi} \\ &= e^{\sigma_i^2 t^2/2 + t\mu_i} \end{aligned}$$

If $Y = \sum X_i$ then Y has mgf

$$M_Y(t) = \exp(t\sum \mu_i + t^2 \sum \sigma_i^2/2)$$

which is the mgf of a $N(\sum \mu_i, \sum \sigma_i^2)$. This proves Y has this normal distribution.

Example: 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 (see earlier in course) $S_1 = Z_1^2$ 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

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

which is the moment generating function 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

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

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:

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