

# STAT 830

## Generating Functions

Richard Lockhart

Simon Fraser University

STAT 830 — Fall 2011



# What I think you already have seen

- Definition of Moment Generating Function
- Basics of complex numbers



# What I want you to learn

- Definition of cumulants and cumulant generating function.
- Definition of Characteristic Function
- Elementary features of complex numbers
- How they “characterize” a distribution
- Relation to sums of independent rvs



- **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:

$$\begin{aligned} M_X(t) &= \sum_{k=0}^{\infty} E[(tX)^k] / k! \\ &= \sum_{k=0}^{\infty} \mu'_k t^k / k!. \end{aligned}$$

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



# Moments and MGFs

## 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  $C^\infty$  (in fact  $M$  is analytic).
- 3  $\mu'_k = \frac{d^k}{dt^k} M_X(0)$ .

- Note:  $C^\infty$  means has continuous derivatives of all orders.
- Analytic means has convergent power series expansion in neighbourhood of each  $t \in (-\epsilon, \epsilon)$ .
- The proof, and many other facts about mgfs, rely on techniques of complex variables.



## MGFs and Sums

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

$$M_Y(t) = E(e^{tY}) = \prod_i E(e^{tX_i}) = \prod_i M_{X_i}(t).$$

- 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  $\mu$ ,  $\sigma^2$  and  $\mu_3$ ) of  $Y$  are sums of those of the  $X_i$ :

$$E(Y) = \sum E(X_i)$$

$$\text{Var}(Y) = \sum \text{Var}(X_i)$$

$$E[(Y - E(Y))^3] = \sum E[(X_i - E(X_i))^3]$$



- However:

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

- But related quantities: **cumulants** add up properly.
- Note: log of the mgf of  $Y$  is sum of logs of mgfs of the  $X_i$ .
- **Def'n**: the cumulant generating function of a variable  $X$  by

$$K_X(t) = \log(M_X(t)).$$

- Then

$$K_Y(t) = \sum K_{X_i}(t).$$

- Note: mgfs are all positive so that the cumulant generating functions are defined wherever the mgfs are.



## Relation between cumulants and moments

- So:  $K_Y$  has power series expansion:

$$K_Y(t) = \sum_{r=1}^{\infty} \kappa_r t^r / r!.$$

- **Def'n:** the  $\kappa_r$  are the cumulants of  $Y$ .
- Observe

$$\kappa_r(Y) = \sum \kappa_r(X_i).$$

- Cumulant generating function is

$$\begin{aligned} K(t) &= \log(M(t)) \\ &= \log(1 + [\mu_1 t + \mu_2' t^2 / 2 + \mu_3' t^3 / 3! + \dots]) \end{aligned}$$

- Call quantity in [...]  $x$ ; expand

$$\log(1 + x) = x - x^2/2 + x^3/3 - x^4/4 \dots$$





# Cumulants and moments

- Stick in the power series

$$x = \mu t + \mu'_2 t^2/2 + \mu'_3 t^3/3! + \dots ;$$

- Expand out powers of  $x$ ; collect together like terms.
- For instance,

$$x^2 = \mu^2 t^2 + \mu \mu'_2 t^3 + [2\mu'_3 \mu/3! + (\mu'_2)^2/4] t^4 + \dots$$

$$x^3 = \mu^3 t^3 + 3\mu'_2 \mu^2 t^4/2 + \dots$$

$$x^4 = \mu^4 t^4 + \dots .$$

- Now gather up the terms.
- The power  $t^1$  occurs only in  $x$  with coefficient  $\mu$ .
- The power  $t^2$  occurs in  $x$  and in  $x^2$  and so on.



## Cumulants and moments

- Putting these together gives

$$K(t) = \mu t + [\mu'_2 - \mu^2]t^2/2 + [\mu'_3 - 3\mu\mu'_2 + 2\mu^3]t^3/3! \\ + [\mu'_4 - 4\mu'_3\mu - 3(\mu'_2)^2 + 12\mu'_2\mu^2 - 6\mu^4]t^4/4! \dots$$

- Comparing coefficients to  $t^r/r!$  we see that

$$\kappa_1 = \mu$$

$$\kappa_2 = \mu'_2 - \mu^2 = \sigma^2$$

$$\kappa_3 = \mu'_3 - 3\mu\mu'_2 + 2\mu^3 = E[(X - \mu)^3]$$

$$\kappa_4 = \mu'_4 - 4\mu'_3\mu - 3(\mu'_2)^2 + 12\mu'_2\mu^2 - 6\mu^4 \\ = E[(X - \mu)^4] - 3\sigma^4.$$

- Reference: Kendall and Stuart (or new version called *Kendall's Theory of Advanced Statistics* by Stuart and Ord) for formulas for larger orders  $r$ .



## Example, $N(0,1)$

- **Example:**  $X_1, \dots, X_p$  independent,  $X_i \sim N(\mu_i, \sigma_i^2)$ :

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

- So cumulant generating function is:

$$K_{X_i}(t) = \log(M_{X_i}(t)) = \sigma_i^2 t^2/2 + \mu_i t.$$

- Cumulants are  $\kappa_1 = \mu_i$ ,  $\kappa_2 = \sigma_i^2$  and every other cumulant is 0.
- Cumulant generating function for  $Y = \sum X_i$  is

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

which is the cumulant generating function of  $N(\sum \mu_i, \sum \sigma_i^2)$ .



## Chi-squared distributions

- **Example:** Homework: derive moment and cumulant generating function and moments of a Gamma rv.
- Now suppose  $Z_1, \dots, Z_\nu$  independent  $N(0, 1)$  rvs.
- By definition:  $S_\nu = \sum_1^\nu Z_i^2$  has  $\chi_\nu^2$  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

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

which is (homework) moment generating function of a Gamma( $\nu/2, 2$ ) rv.

- SO:  $\chi_\nu^2$  dstbn has Gamma( $\nu/2, 2$ ) density:

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



# Cauchy Distribution

- **Example:** The Cauchy density is

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

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.

- Every  $t$  distribution has exactly same mgf.
- So: can't use mgf to distinguish such distributions.
- Problem: these distributions do not have infinitely many finite moments.
- So: develop substitute for mgf which is defined for every distribution namely, the characteristic function.



## Aside on complex arithmetic

- Complex numbers: add  $i = \sqrt{-1}$  to the real numbers.
- Require all the usual rules of algebra to work.
- So: if  $i$  and any real numbers  $a$  and  $b$  are to be complex numbers then so must be  $a + bi$ .
- Multiplication: 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.



## Complex aside, slide 2

- Addition: follow usual rules to get

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

- Additive inverses:  $-(a + bi) = -a + (-b)i$ .
- Multiplicative inverses:

$$\begin{aligned}\frac{1}{a + bi} &= \frac{1}{a + bi} \frac{a - bi}{a - bi} \\ &= \frac{a - bi}{a^2 - abi + abi - b^2i^2} = \frac{a - bi}{a^2 + b^2}.\end{aligned}$$

- Division:

$$\frac{a + bi}{c + di} = \frac{(a + bi)(c - di)}{(c + di)(c - di)} = \frac{ac - bd + (bc + ad)i}{c^2 + d^2}.$$

- Notice: usual rules of arithmetic don't require any more numbers than

$$x + yi$$

where  $x$  and  $y$  are real.



## Complex Aside Slide 3

- **Transcendental functions:** For real  $x$  have  $e^x = \sum x^k/k!$  so

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

- How to compute  $e^{iy}$ ?
- Remember  $i^2 = -1$  so  $i^3 = -i$ ,  $i^4 = 1$   $i^5 = i^1 = i$  and so on. Then

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

- We can thus write

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





## Complex Aside Slide 4, Argand diagrams

- Identify  $x + yi$  with the corresponding point  $(x, y)$  in the plane.
- Picture the complex numbers as forming a plane.
- 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

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

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

- Multiplication revisited:  $x + iy = re^{i\theta}$ ,  $x' + iy' = r'e^{i\theta'}$ .

$$(x + iy)(x' + iy') = re^{i\theta} r'e^{i\theta'} = rr'e^{i(\theta+\theta')}.$$



## Complex Aside Slide 4, Argand diagrams

- We will need from time to time a couple of other definitions:
- **Definition:** The **modulus** of  $x + iy$  is

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

- **Definition:** The **complex conjugate** of  $x + iy$  is  $\overline{x + iy} = x - iy$ .
- Some identities:  $z = x + iy = re^{i\theta}$  and  $z' = x' + iy' = r'e^{i\theta'}$ .
- Then

$$z\bar{z} = x^2 + y^2 = r^2 = |z|^2$$

$$\frac{z'}{z} = \frac{z'\bar{z}}{|z|^2} = rr'e^{i(\theta' - \theta)}$$

$$\overline{re^{i\theta}} = re^{-i\theta}.$$



# Notes on calculus with complex variables

- Essentially 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.
- FACT: (not used explicitly in course). If  $f : \mathbb{C} \mapsto \mathbb{C}$  is differentiable then  $f$  is analytic (has power series expansion).

**End of Aside**



# Characteristic Functions

- **Def'n:** The characteristic function of a real rv  $X$  is

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

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

- 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  $t$ .
- This is precisely the reason for using characteristic rather than moment generating functions in probability theory courses.



## Theorem

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

Moreover, all these are implied if there is  $\epsilon > 0$  such that for all  $|t| \leq \epsilon$

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

