Lecture note on moment generating functions

Ernie Croot

October 23, 2008

1 Introduction

Given a random variable X, let $f(x)$ be its pdf. The quantity (in the continuous case – the discrete case is defined analogously)

$$\mathbb{E}(X^k) \ = \ \int_{-\infty}^{\infty} x^k f(x) dx$$

is called the kth moment of X.

The "moment generating function" gives us a nice way of collecting together all the moments of a random varaible X into a single power series (i.e. Maclaurin series) in the variable t. It is defined to be

$$M_X(t) \ := \ \mathbb{E}(e^{Xt}) \ = \ \mathbb{E}\left(\sum_{k=0}^{\infty} \frac{X^k t}{k!} \right).$$

Thinking here of the t as a constant, at least from the perspective of taking expectations, we use the linearity of expectation to conclude that

$$\mathbb{E}(e^{Xt}) \ = \ \sum_{k=0}^{\infty} \frac{\mathbb{E}(X^k) t^k}{k!}$$

So, the coefficients of the powers of t give us moments divided by $k!$.

It is fairly easy to see that if we take the kth derivative of the moment generating function, and set $t = 0$, the result will be the kth moment. In symbols, this is

$$\left(\frac{d}{dt} \right)^k M_X(t) \Big|_{t=0} \ = \ k\text{th moment of } X.$$

Caution!: It may be that the moment generating function does not exist, because some of the moments may be infinite (or may not have a definite value, due to integrability issues). Also, even if the moments are all finite and have definite values, the generating function may not converge for any value of t other than 0. All that said, these convergence issues and infinities rarely come up in the sort of problems we will consider in this course (and rarely come up in real-world problems).

Take home message. I expect you to know everything from this section.

2 Some examples

Example 1. We saw in class that if X is a Bernoulli random variable with parameter p, then

$$M_X(t) \ = \ \mathbb{E}(e^{Xt}) \ = \ e^{0\cdot t}(1-p) + e^{1\cdot t}p \ = \ pe^t + 1 - p.$$

Also, because $X^k = X$, it is clear that the kth moment of X, $k \geq 1$, is the same as the first moment, which is just p; indeed, taking the kth derivative of $M_X(t)$ and setting $t = 0$ we find that

$$k\text{th moment} \ = \ pe^t\Big|_{t=0} \ = \ p.$$

Example 2. Let us compute the moment generating function for a normal random variable having variance σ^2 and mean $\mu = 0$. Note that the pdf for such a random variable is just

$$f(x) \ = \ \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2\sigma^2}.$$

So, we have that

$$\begin{aligned} M_X(t) \ = \ \mathbb{E}(e^{Xt}) \ &= \ \int_{-\infty}^{\infty} e^{xt} \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2\sigma^2} dx \\ &= \ \int_{-\infty}^{\infty} \frac{e^{-(1/2\sigma^2)(x^2 - 2\sigma^2 tx)}}{\sqrt{2\pi}\sigma} dx. \end{aligned}$$

Now we complete $x^2 - 2\sigma^2 tx$ to a square by adding to it (and then subtracting) $\sigma^4 t^2$, and so

$$\begin{aligned} M_X(t) &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} e^{-(1/2\sigma^2)((x-\sigma^2 t)^2 - \sigma^4 t^2)} dx \\ &= e^{\sigma^2 t^2/2} \left(\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\sigma^2 t)^2/2\sigma^2} dx \right). \end{aligned}$$

Now we recognize that the last integral here equals 1, since it is the integral of the pdf for the normal random variable $N(\sigma^2 t, \sigma^2)$ over the full interval $(-\infty, \infty)$. So,

$$M_X(t) = e^{\sigma^2 t^2/2}.$$

Just to make sure you understand how moment generating functions work, try the following two example problems.

Problem 1. Compute the moment generating function for the random variable X having uniform distribution on the interval $[0, 1]$.

Problem 2. Based on your answer in problem 1, compute the fourth moment of X – i.e. $\mathbb{E}(X^4)$.

Take home message. I expect you to know how to compute the moment generating function of some basic random variables, like those with Bernoulli and uniform distribution. I **do not** expect you to know how to derive the MGF for normal random variables for the purposes solving a problem on an exam. Though, I may give you the MGF of some random variable on an exam, and then ask you to compute moments of that r.v.

3 From moment generating functions to distributions

A key and profoundly useful fact about moment generating functions is the following.

Key Fact. Suppose that X and Y are two random variables having moment generating functions $M_X(t)$ and $M_Y(t)$ that exist for all t in some interval

$[-\delta, \delta]$. Then, if

$$M_X(t) \;=\; M_Y(t), \text{ for all } t \in [-\delta, \delta],$$

we must have that X and Y have the same cumulative distribution; that is,

$$P(X \le a) \;=\; P(Y \le a), \text{ for all } a \in \mathbb{R}.$$

This beautiful fact can be exploited to pin down the distribution of the sum of various types of independent random variables, due to the fact that the exponential in the definition of MGFs allows us to transform sums of random variables into products of expectations. More specifically, we have the following immensely useful "reproductive property" of the moment generating function.

Claim. Suppose that $X_1, ..., X_k$ are independent random variables. Then,

$$\begin{aligned} M_{X_1+\cdots+X_k}(t) &= \mathbb{E}(e^{(X_1+\cdots+X_k)t}) \\ &= \mathbb{E}(e^{X_1 t} e^{X_2 t} \cdots e^{X_k t}) \\ &= \mathbb{E}(e^{X_1 t}) \cdots \mathbb{E}(e^{X_k t}) \\ &= M_{X_1}(t) \cdots M_{X_k}(t). \end{aligned}$$

We note that we used the fact that the X_i's are independent when we rewrote the expectation of a product of exponentials as a product of expectations. This is just the familiar fact that if $Z_1, ..., Z_k$ are independent, then

$$\mathbb{E}(Z_1 \cdots Z_k) \;=\; \mathbb{E}(Z_1) \cdots \mathbb{E}(Z_k).$$

Take home message. Know everything from this section.

4 An application to sums of independent normal random variables

Using the moment generating function we computed for $N(0, \sigma^2)$ in the previous section, we now use it to show the following.

Claim. Suppose that $X_1, ..., X_k$ are independent normal random variables with means $\mu_1, ..., \mu_k$ and variances $\sigma_1^2, ..., \sigma_k^2$, respectively. Then,

$$Y = X_1 + \cdots + X_k$$

is normal with mean $\mu_1 + \cdots + \mu_k$ and variance $\sigma_1^2 + \cdots + \sigma_k^2$.

Basically, to show that this is the case, we just compute the moment generating function for Y, and show that it is the same as that of

$$N(\mu_1 + \cdots + \mu_k, \sigma_1^2 + \cdots + \sigma_k^2).$$

We will do this only for the case $\mu_1 = \cdots = \mu_k = 0$ (the more general case is really no more difficult!).

Combining together various facts from previous sections, we find that

$$\begin{aligned} M_Y(t) \ = \ M_{X_1}(t) \cdots M_{X_k}(t) \ &= \ e^{\sigma_1^2 t^2/2} \cdots e^{\sigma_k^2 t^2/2} \\ &= \ e^{(\sigma_1^2 + \cdots + \sigma_k^2)t^2/2}, \end{aligned}$$

which is clearly the moment generating function for

$$N(0,\ \sigma_1^2 + \cdots + \sigma_k^2),$$

so we are done (note that the moment generating function exists for all t, not merely for $t \in [-\delta, \delta]$).

Problem 3. See if you can figure out how to handle the more general case of where the X_i's have means μ_i that may or may not equal 0.

Take home message. Know everything from this section.