

## 7 The moment generating function

This is a topic we won't explore very far; suffice to say it is presented here as a curio but in some parts of probability theory it is proves very useful and it often mentioned. Recall the definition of the  $i$ th moment, which for convenience we'll call  $\mu_i$ :

$$\mu_i = \langle X^i \rangle \quad (1)$$

Hence  $\mu$  the mean is  $\mu = \mu_1$ . Now you should also recall the formula for the Taylor expansion of the exponential:

$$e^x = \sum_n \frac{x^n}{n!} \quad (2)$$

Now the **moment generating function**  $m(t)$  is defined as

$$m(t) = \langle e^{tX} \rangle \quad (3)$$

This function 'generates' all the moments because

$$m(t) = \langle e^{tX} \rangle = \sum \frac{t^n}{n!} \langle X^n \rangle = \sum \frac{t^n}{n!} \mu_n \quad (4)$$

so it is equivalent to a sum with all the moment in it. In fact, if we don't worry for a minute about what  $X$  is and just treat it like a constant, we know

$$\frac{d^n}{dt^n} e^{tX} = X^n e^{tX} \quad (5)$$

and so

$$\frac{d^n}{dt^n} m(t) = \langle X^n e^{tX} \rangle \quad (6)$$

so setting  $t = 0$  after differentiating gives

$$\frac{d^n m}{dt^n}(0) = \langle X^n \rangle = \mu_n \quad (7)$$

Of course you might worry about  $X$  not being a constant but, rather, a function; with a bit of effort, which we won't expend here, it can be shown how to set this up so that this doesn't matter.

Why might be do this; well the thing is the moment generating function encodes all the properties of the distribution into one place and this can be useful for proving theorems. We can think of it as an alternative description of the distribution, instead of giving the probabilities  $p(X)$  we give the moments, all wrapped up in a single function.

We just note that it can be calculated as an explicit formula in some cases. For a particularly easy case consider the Bernoulli distribution, that is a random variable with two outcomes, say zero and one, with  $p(1) = p$  and  $p(0) = q = 1 - p$ . Now

$$m(t) = \langle e^{tX} \rangle = pe^t + (1 - p)e^0 = 1 - p + pe^t \quad (8)$$

In fact, though we won't look at any other examples here, it is usually possible to calculate the moment generating function for a distribution.

## Summary

- The **moment generating function** is

$$m(t) = \langle e^{tX} \rangle \quad (9)$$

- It gives the  $n$ th moment

$$\frac{d^n m}{dt^n}(0) = \langle X^n \rangle = \mu_n \quad (10)$$

- For the Bernoulli distribution which is one with probability  $p$  and zero with probability  $q = 1 - p$

$$m(t) = q + pe^t \quad (11)$$