IB Paper 7: Probability
How do Moment Generating Functions Work?

James Brind
jb753@cam.ac.uk

Lent Term 2018

Abstract

This note shows how the definition of moment generating functions leads to short cuts to calculating mean, variance and higher moments of a distribution. Expanding the generating function as a power series, is demonstrated that if it is differentiated n times, the leading order term includes a component in $\mathbb{E}[X^n]$ which can be used to calculate the nth moment.

Continuous generating functions

Start with the continuous case because the algebra is tidier. We have a continuous random variable X. Define a function f of a 'transform variable' s which takes the form,

$$f(s) = \exp(-sX).$$

The exponential can also be written as a power series (see Maths Data Book),

$$f(s) = 1 + (-sX) + \frac{(-sx)^2}{2!} + \frac{(-sx)^3}{3!} + \mathscr{O}(s^4).$$

Now define the continuous generating function $g(s)$ to be the expectation of f, across all possible values of X with s fixed,

$$g(s) = \mathbb{E}[f(s)] = 1 - s\mathbb{E}[X] + \frac{s^2}{2}\mathbb{E}[X^2] - \frac{s^3}{6}\mathbb{E}[X^3] + \mathscr{O}(s^4).$$

Calculate the derivatives of g, and evaluate at $s = 0$ to keep only the leading order term,

$$g'(s) = -\mathbb{E}[X] + s\mathbb{E}[X^2] - \frac{s^2}{2}\mathbb{E}[X^3] + \mathscr{O}(s^3) \quad \Rightarrow \quad g'(0) = -\mathbb{E}[X],$$
$$g''(s) = \mathbb{E}[X^2] - s\mathbb{E}[X^3] + \mathscr{O}(s^2), \quad \Rightarrow \quad g''(0) = \mathbb{E}[X^2].$$

So the mean and variance are,

$$\boxed{\mathbb{E}[X] = -g'(0), \quad \text{and} \quad \mathbb{V}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = g''(0) - g'(0)^2}$$

This procedure works for any nth moment, because if the generating function is differentiated n times, the leading order term includes a component in $\mathbb{E}[X^n]$.

Discrete generating functions

We have a discrete random variable X. Define a function f of a 'transform variable' z which for a particular value of X takes the form,

$$f(z) = z^X.$$

Differentiating with respect to z, with X fixed,

$$\begin{aligned} f'(z) &= X z^{X-1}, \\ f''(z) &= X(X-1)z^{X-2}. \end{aligned}$$

Write $f(z)$ as a Taylor series about $z = 1$. Using the Maths Data Book notation, $x = 1$ and $h = z - 1$. We have,

$$f(z) = f(1) + (z-1)f'(1) + \frac{(z-1)^2}{2!} f''(1) + \mathcal{O}((z-1)^3).$$

Substituting for the derivatives of f evaluated at $z = 1$,

$$f(z) = 1 + (z-1)X + \frac{(z-1)^2}{2} X(X-1) + \mathcal{O}((z-1)^3).$$

Now define the discrete generating function $g(z)$ to be the expectation of f, across all values of X with z fixed,

$$g(z) = \mathbb{E}[f(z)] = 1 + (z-1)\mathbb{E}[X] + \frac{(z-1)^2}{2}\mathbb{E}[X(X-1)] + \mathcal{O}((z-1)^3).$$

Calculate derivatives of g and then evaluate at $z = 1$ to keep the leading order term,

$$\begin{aligned} g'(z) &= \mathbb{E}[X] + (z-1)\mathbb{E}[X(X-1)] + \mathcal{O}((z-1)^2) \quad \Rightarrow \quad g'(1) = \mathbb{E}[X], \\ g''(z) &= \mathbb{E}[X(X-1)] + \mathcal{O}((z-1)) \quad \Rightarrow \quad g''(1) = \mathbb{E}[X(X-1)] = \mathbb{E}[X^2 - X] = \mathbb{E}[X^2] - \mathbb{E}[X]. \end{aligned}$$

So the mean and variance are, in these terms,

$$\boxed{\mathbb{E}[X] = g'(1),}$$

$$\mathbb{V}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \big(\mathbb{E}[X^2] - \mathbb{E}[X]\big) + \mathbb{E}[X] - \mathbb{E}[X]^2,$$

$$\boxed{\mathbb{V}[X] = g''(1) + g'(1) - g'(1)^2.}$$

This procedure works for any nth moment, because if the generating function is differentiated n times, the leading order term includes a component in $\mathbb{E}[X^n]$.