MOMENT GENERATING FUNCTION AND IT'S APPLICATIONS

ASHWIN RAO

The purpose of this note is to introduce the *Moment Generating Function (MGF)* and demonstrate it's utility in several applications in Applied Mathematics.

1. The Moment Generating Function (MGF)

The Moment Generating Function (MGF) of a random variable x (discrete or continuous) is defined as a function $f_x : \mathbb{R} \to \mathbb{R}^+$ such that:

$$f_x(t) = \mathbb{E}_x[e^{tx}] \text{ for all } t \in \mathbb{R} \tag{1}$$

Let us denote the n^{th}-derivative of f_x as $f_x^{(n)} : \mathbb{R} \to \mathbb{R}$ for all $n \in \mathbb{Z}_{\geq 0}$ ($f_x^{(0)}$ is defined to be simply the MGF f_x).

$$f_x^{(n)}(t) = \mathbb{E}_x[x^n \cdot e^{tx}] \text{ for all } n \in \mathbb{Z}_{\geq 0} \text{ for all } t \in \mathbb{R} \tag{2}$$

$$f_x^{(n)}(0) = \mathbb{E}_x[x^n] \tag{3}$$

$$f_x^{(n)}(1) = \mathbb{E}_x[x^n \cdot e^x] \tag{4}$$

Equation (3) tells us that $f^{(n)}(0)$ gives us the n^{th} moment of x. In particular, $f_x^{(1)}(0) = f'(0)$ gives us the mean and $f_x^{(2)}(0) - (f_x^{(1)}(0))^2 = f''_x(0) - (f'_x(0))^2$ gives us the variance. Note that this holds true for any distribution for x. This is rather convenient since all we need is the functional form for the distribution of x. This would lead us to the expression for the MGF (in terms of t). Then, we take derivatives of this MGF and evaluate those derivatives at 0 to obtain the moments of x.

Equation (4) helps us calculate the often-appearing expectation $\mathbb{E}_x[x^n \cdot e^x]$. In fact, $\mathbb{E}_x[e^x]$ and $\mathbb{E}_x[x \cdot e^x]$ are very common in several areas of Applied Mathematics. Again, note that this holds true for any distribution for x.

2. MGF for the Normal Distribution

Here we assume that the random variables x follows a normal distribution. Let $x \sim \mathcal{N}(\mu, \sigma^2)$.

$$f_{x\sim\mathcal{N}(\mu,\sigma^2)}(t) = \mathbb{E}_{x\sim\mathcal{N}(\mu,\sigma^2)}[e^{tx}] \tag{5}$$
$$= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}} \cdot e^{tx} \cdot dx \tag{6}$$
$$= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{-\frac{(x-(\mu+t\sigma^2))^2}{2\sigma^2}} \cdot e^{\mu t+\frac{\sigma^2 t^2}{2}} \cdot dx \tag{7}$$
$$= e^{\mu t+\frac{\sigma^2 t^2}{2}} \cdot \mathbb{E}_{x\sim\mathcal{N}(\mu+t\sigma^2,\sigma^2)}[1] \tag{8}$$
$$= e^{\mu t+\frac{\sigma^2 t^2}{2}} \tag{9}$$

$$f'_{x\sim\mathcal{N}(\mu,\sigma^2)}(t) = \mathbb{E}_{x\sim\mathcal{N}(\mu,\sigma^2)}[x \cdot e^{tx}] = (\mu+\sigma^2 t) \cdot e^{\mu t+\frac{\sigma^2 t^2}{2}} \tag{10}$$

$$f''_{x\sim\mathcal{N}(\mu,\sigma^2)}(t) = \mathbb{E}_{x\sim\mathcal{N}(\mu,\sigma^2)}[x \cdot e^{tx}] = ((\mu+\sigma^2 t)^2+\sigma^2) \cdot e^{\mu t+\frac{\sigma^2 t^2}{2}} \tag{11}$$

$$f'_{x\sim\mathcal{N}(\mu,\sigma^2)}(0) = \mathbb{E}_{x\sim\mathcal{N}(\mu,\sigma^2)}[x] = \mu \tag{12}$$

$$f''_{x\sim\mathcal{N}(\mu,\sigma^2)}(0) = \mathbb{E}_{x\sim\mathcal{N}(\mu,\sigma^2)}[x] = \mu^2+\sigma^2 \tag{13}$$

$$f'_{x\sim\mathcal{N}(\mu,\sigma^2)}(1) = \mathbb{E}_{x\sim\mathcal{N}(\mu,\sigma^2)}[x \cdot e^{x}] = (\mu+\sigma^2)e^{\mu+\frac{\sigma^2}{2}} \tag{14}$$

$$f''_{x\sim\mathcal{N}(\mu,\sigma^2)}(1) = \mathbb{E}_{x\sim\mathcal{N}(\mu,\sigma^2)}[x \cdot e^{x}] = ((\mu+\sigma^2)^2+\sigma^2)e^{\mu+\frac{\sigma^2}{2}} \tag{15}$$

3. Minimizing the MGF

Now let us consider the problem of minimizing the MGF. The problem is to:

$$\min_{t\in\mathbb{R}} f_x(t) = \min_{t\in\mathbb{R}} \mathbb{E}_x[e^{tx}]$$

This problem of minimizing $\mathbb{E}_x[e^{tx}]$ shows up a lot in various places in Applied Mathematics when dealing with exponential functions (eg: when optimizing the Expectation of a Constant Absolute Risk-Aversion Utility function $U(y) = \frac{-e^{-\gamma y}}{\gamma}$ where γ is the coefficient of risk-aversion and where y is a parameterized function of a random variable x).

Let us denote t^* as the value of t that minimizes the MGF. Specifically,

$$t^* = \arg\min_{t\in\mathbb{R}} f_x(t) = \arg\min_{t\in\mathbb{R}} \mathbb{E}_x[e^{tx}]$$

3.1. Minimizing the MGF when x follows a normal distribution.

Here we consider the fairly typical case where x follows a normal distribution. Let $x \sim \mathcal{N}(\mu,\sigma^2)$. Then we have to solve the problem:

$$\min_{t\in\mathbb{R}} f_{x\sim\mathcal{N}(\mu,\sigma^2)}(t) = \min_{t\in\mathbb{R}} \mathbb{E}_{x\sim\mathcal{N}(\mu,\sigma^2)}[e^{tx}] = \min_{t\in\mathbb{R}} e^{\mu t+\frac{\sigma^2 t^2}{2}}$$

From Equation (10) above, we have:

$$f'_{x\sim\mathcal{N}(\mu,\sigma^2)}(t) = (\mu+\sigma^2 t) \cdot e^{\mu t+\frac{\sigma^2 t^2}{2}}$$

Setting this to 0 yields:

$$(\mu+\sigma^2 t^*) \cdot e^{\mu t^*+\frac{\sigma^2 t^{*2}}{2}} = 0$$

which leads to:

$$t^* = \frac{-\mu}{\sigma^2} \tag{16}$$

From Equation (11) above, we have:

$$f''_{x\sim\mathcal{N}(\mu,\sigma^2)}(t) = ((\mu+\sigma^2 t)^2 + \sigma^2)\cdot e^{\mu t + \frac{\sigma^2 t^2}{2}} > 0 \text{ for all } t \in \mathbb{R}$$

which confirms that t^* is a minima.

Substituting $t = t^*$ in $f_{x\sim\mathcal{N}(\mu,\sigma^2)}(t) = e^{\mu t + \frac{\sigma^2 t^2}{2}}$ yields:

$$\min_{t\in\mathbb{R}} f_{x\sim\mathcal{N}(\mu,\sigma^2)}(t) = e^{\mu t^* + \frac{\sigma^2 t^{*2}}{2}} = e^{\frac{-\mu^2}{2\sigma^2}} \tag{17}$$

3.2. **Minimizing the MGF when x is a symmetric binary distribution.** Here we consider the case where x follows a binary distribution: x takes values $\mu+\sigma$ and $\mu-\sigma$ with probability 0.5 each. Let us refer to this distribution as $x \sim \mathcal{B}(\mu+\sigma,\mu-\sigma)$. Note that the mean and variance of x under $\mathcal{B}(\mu+\sigma,\mu-\sigma)$ are μ and σ^2 respectively. So we have to solve the problem:

$$\min_{t\in\mathbb{R}} f_{x\sim\mathcal{B}(\mu+\sigma,\mu-\sigma)}(t) = \min_{t\in\mathbb{R}} \mathbb{E}_{x\sim\mathcal{B}(\mu+\sigma,\mu-\sigma)}[e^{tx}] = \min_{t\in\mathbb{R}} 0.5(e^{(\mu+\sigma)t} + e^{(\mu-\sigma)t})$$

$$f'_{x\sim\mathcal{B}(\mu+\sigma,\mu-\sigma)}(t) = 0.5((\mu+\sigma)\cdot e^{(\mu+\sigma)t} + (\mu-\sigma)\cdot e^{(\mu-\sigma)t})$$

Note that unless $\mu \in$ open interval $(-\sigma,\sigma)$ (i.e., absolute value of mean is less than standard deviation), $f'_{x\sim\mathcal{B}(\mu+\sigma,\mu-\sigma)}(t)$ will not be 0 for any value of t. Therefore, for this minimization to be non-trivial, we will henceforth assume $\mu \in (-\sigma,\sigma)$. With this assumption in place, setting $f'_{x\sim\mathcal{B}(\mu+\sigma,\mu-\sigma)}(t)$ to 0 yields:

$$(\mu+\sigma)\cdot e^{(\mu+\sigma)t^*} + (\mu-\sigma)\cdot e^{(\mu-\sigma)t^*} = 0$$

which leads to:

$$t^* = \frac{1}{2\sigma}\ln\left(\frac{\sigma-\mu}{\mu+\sigma}\right) \tag{18}$$

Note that

$$f''_{x\sim\mathcal{B}(\mu+\sigma,\mu-\sigma)}(t) = 0.5((\mu+\sigma)^2\cdot e^{(\mu+\sigma)t} + (\mu-\sigma)^2\cdot e^{(\mu-\sigma)t}) > 0 \text{ for all } t\in\mathbb{R}$$

which confirms that t^* is a minima.

Substituting $t = t^*$ in $f_{x\sim\mathcal{B}(\mu+\sigma,\mu-\sigma)}(t) = 0.5(e^{(\mu+\sigma)t} + e^{(\mu-\sigma)t})$ yields:

$$\min_{t\in\mathbb{R}} f_{x\sim\mathcal{B}(\mu+\sigma,\mu-\sigma)}(t) = 0.5(e^{(\mu+\sigma)t^*} + e^{(\mu-\sigma)t^*}) = 0.5\left(\left(\frac{\sigma-\mu}{\mu+\sigma}\right)^{\frac{\mu+\sigma}{2\sigma}} + \left(\frac{\sigma-\mu}{\mu+\sigma}\right)^{\frac{\mu-\sigma}{2\sigma}}\right) \tag{19}$$

4. More Applications coming up soon ...