The Moment Generating Function

Guy Lebanon

January 30, 2006

Definition 1. *The k-moment of a RV X is $\mathsf{E}(X^k)$.*

Definition 2. *The moment generating function (mgf) of the RV X is the function*

$$m : \mathbb{R} \to \mathbb{R} \quad m(t) = E(\exp(tX)).$$

The reason for the name mgf stems from the fact that its derivatives, evaluated at 0 produce the moments. For example, using the Taylor series expansion of e^z around 0 we have

$$m'(0) = \frac{d}{dt}\mathsf{E}\left(1 + tX + \frac{t^2X^2}{2!} + \cdots\right)\Big|_{t=0} = \mathsf{E}\left(\frac{d}{dt}\left(1 + tX + \frac{t^2X^2}{2!} + \cdots\right)\right)\Big|_{t=0} = \mathsf{E}(X).$$

Similarly, the second derivative produces the second moment

$$m''(0) = \frac{d^2}{dt^2}\mathsf{E}\left(1 + tX + \frac{t^2X^2}{2!} + \cdots\right)\Big|_{t=0} = \mathsf{E}\left(\frac{d^2}{dt^2}\left(1 + tX + \frac{t^2X^2}{2!} + \cdots\right)\right)\Big|_{t=0} = \mathsf{E}(X^2).$$

and so on for the higher order derivatives.

In addition to producing the moments of X, the mgf is useful in identifying the distribution of X. The following theorem asserts that two RVs with the same mgf have the same distribution. We will see numerous applications of this useful role of the mgf.

Theorem 1. *Let X, Y be two RVs with mgfs m, n. If $m(t) = n(t)$ for all $t \in (-\epsilon, +\epsilon)$ (for some $\epsilon > 0$) then the two RVs have equal distributions (and the same cdf, pdf and pmf).*

Proposition 1. *The mgf of a Normal RV $X \sim N(\mu, \sigma^2)$ is $m(t) = \exp(\mu t + t^2\sigma^2/2)$*

Proof.

$$\begin{aligned} m(t) &= \frac{1}{\sqrt{2\pi\sigma^2}}\int_{\mathbb{R}} e^{tr}e^{-(r-\mu)^2/(2\sigma^2)}dr = \frac{1}{\sqrt{2\pi\sigma^2}}\int_{\mathbb{R}} e^{t(u+\mu)}e^{-u^2/(2\sigma^2)}du = \frac{1}{\sqrt{2\pi\sigma^2}}e^{t\mu}\int_{\mathbb{R}} e^{tu}e^{-u^2/(2\sigma^2)}du \\ &= \frac{1}{\sqrt{2\pi\sigma^2}}e^{t\mu}\int_{\mathbb{R}} e^{-(u^2 - tu2\sigma^2)/(2\sigma^2)}du \end{aligned}$$

where we used the change of variables $u = r - \mu$. To solve the above integral we complete the square by multiplying and dividing by $e^{t^2\sigma^2/2}$

$$\begin{aligned} m(t) &= \frac{1}{\sqrt{2\pi\sigma^2}}e^{t\mu}e^{t^2\sigma^2/2}\int_{\mathbb{R}} e^{-(u^2 - tu2\sigma^2 + t^2\sigma^4)/(2\sigma^2)}du = \frac{1}{\sqrt{2\pi\sigma^2}}e^{t\mu}e^{t^2\sigma^2/2}\int_{\mathbb{R}} e^{-(u-t\sigma^2)^2)/(2\sigma^2)}du \\ &= e^{t\mu + t^2\sigma^2/2}. \end{aligned}$$

□

Theorem 2. *If $X \sim N(\mu, \sigma^2)$ then $(X - \mu)/\sigma \sim N(0, 1)$.*

Proof. By the above proposition mgf of a $N(0,1)$ RV is $e^{t^2/2}$. On the other hand, to find the mgf of $(X-\mu)/\sigma$ we can repeat the above computation (not quite so, but similarly)

$$\begin{aligned}
m(t) = \mathsf{E}(\exp(t(X-\mu)/\sigma)) &= \frac{1}{\sqrt{2\pi\sigma^2}}\int_{\mathbb{R}} e^{t(r-\mu)/\sigma}e^{-(r-\mu)^2/(2\sigma^2)}dr = \frac{1}{\sqrt{2\pi\sigma^2}}\int_{\mathbb{R}} e^{tu/\sigma}e^{-u^2/(2\sigma^2)}du \\
&= \frac{1}{\sqrt{2\pi\sigma^2}}\int_{\mathbb{R}} e^{-(u^2-2tu\sigma)/(2\sigma^2)}du = \frac{1}{\sqrt{2\pi\sigma^2}}e^{t^2/2}\int_{\mathbb{R}} e^{-(u^2-2tu\sigma+t^2\sigma^2)/(2\sigma^2)}du \\
&= \frac{1}{\sqrt{2\pi\sigma^2}}e^{t^2/2}\int_{\mathbb{R}} e^{-(u-t\sigma)^2/(2\sigma^2)}du = e^{t^2/2}.
\end{aligned}$$

Since the two mgfs agree, by the theorem above, the distributions of $(X-\mu)/\sigma$ and $Z \sim N(0,1)$ are identical. □

Another consequence of the theorem above is that if all the moments of a RV X exists, they characterize completely the mgf (since the moments are derivatives of the mgf that compose its Taylor series) and by the above theorem, the moments also completely characterize the distribution, as well as the cdf, pdf, and pmf.

Theorem 3. *Let $X_1,\ldots,X_n$ be independent RVs with mgfs $m_1,\ldots,m_n$. The mgf of $\sum_{i=1}^n X_i$ is $\prod_{i=1}^n m_i(t)$.*

Proof. The mgf $m(t)$ of the sum $\sum_{i=1}^n X_i$ is by definition

$$m(t) = \mathsf{E}\left(\exp\left(t\sum_i X_i\right)\right) = \mathsf{E}\left(\prod_i \exp(tX_i)\right) = \prod_i \mathsf{E}(\exp(tX_i)) = \prod_i m_i(t).$$

□

Theorem 4. *Let $X_i \sim N(\mu_i,\sigma_i)$ be independent RVs, for $i=1,\ldots,n$ and $a_1,\ldots,a_n \in \mathbb{R}$. Then $\sum_{i=1}^n a_iX_i \sim N(\sum a_i\mu_i, \sum a_i^2\sigma_i^2)$.*

Proof. Recall that the mgf of $N(\mu,\sigma^2)$ is $e^{\mu t+\sigma^2t^2/2}$. The mgf of a_iX_i is $\mathsf{E}(\exp(ta_iX_i))$ which is the mgf of X_i at ta_i which equals $e^{\mu ta_i+\sigma_i^2t^2a_i^2/2}$. This is also the mgf of a $N(a_i\mu, a_i^2\sigma_i^2)$ and by the mgf uniqueness theorem, this means that $a_iX_i \sim N(a_i\mu, a_i^2\sigma_i^2)$. Finally, since the mgf of the sum of RVs is the product of the mgfs

$$\prod_i e^{\mu ta_i+\sigma^2t^2a_i^2/2} = e^{\sum_i \mu ta_i+\sigma_i^2t^2a_i^2/2} = e^{t(\sum_i \mu a_i)+t^2(\sum \sigma_i^2a_i^2)/2}$$

which is the mgf of $N(\sum a_i\mu_i, \sum a_i^2\sigma_i^2)$ and again by the mgf uniqueness theorem the theorem is proved. □