Moment-generating functions and independence

Let $X := (X_1, \ldots, X_n)'$ be a random vector. Its *moment generating function* [written MGF for short] M_X is defined as

$$M_X(\boldsymbol{t}) := \mathrm{E}e^{\boldsymbol{t}'X} \qquad (\boldsymbol{t} \in \mathbf{R}^n).$$

It is the case that $M_X(\boldsymbol{t})$ is a well-defined quantity, but it might be infinite for some, and even all, values of $\boldsymbol{t} \in \mathbf{R}^n$. The following is a hard fact from classical analysis:

Theorem 1 (Uniqueness theorem of MGFs)**.** *Suppose there exists $t > 0$ such that $M_X(\boldsymbol{t}) < \infty$ for all $\boldsymbol{t} \in \mathbf{R}^n$ with $\|\boldsymbol{t}\| \le r$. Then, the distribution of X is determined uniquely by the function M_X. That is, if $\boldsymbol{Y}$ is any random vector whose MGF is the same as M_X, then $\boldsymbol{Y}$ has the same distribution as X.*

We are interested in examples, and primarily those that involve normal distributions in one form or another.

Example 2. If $X \sim \mathrm{N}(\mu\,,\sigma^2)$, then

$$\begin{aligned} M_X(t) = \mathrm{E}e^{tX} &= \int_{-\infty}^{\infty} e^{tx}\, \frac{e^{-(x-\mu)^2/(2\sigma^2)}}{\sqrt{2\pi\sigma^2}}\,dx \\ &= \int_{-\infty}^{\infty} e^{t(\sigma y+\mu)}\, \frac{e^{-y^2/2}}{\sqrt{2\pi}}\,dy \qquad\qquad (y := (x-\mu)/\sigma) \\ &= e^{t\mu}\int_{-\infty}^{\infty} \frac{e^{t\sigma y - y^2/2}}{\sqrt{2\pi}} = e^{t\mu}\int_{-\infty}^{\infty}\frac{e^{-\frac12[y^2-2yt\sigma]}}{\sqrt{2\pi}}\,dy. \end{aligned}$$

We complete the square $[y^2 - 2yt\sigma = (y-t\sigma)^2 - (t\sigma)^2]$ in order to see that

$$M_X(t) = \exp\left(t\mu + \frac{t^2\sigma^2}{2}\right).$$

Therefore, the uniqueness theorem (Theorem 1) tells us that any random variable Y whose MGF is $M_Y(t) = \exp(t\mu + \frac12 t^2\sigma^2)$ is distributed according to $\mathrm{N}(\mu\,,\sigma^2)$. □

Example 3 (MGF of a simple multivariable normal)**.** Suppose $X_i \sim \mathrm{N}(\mu_i\,,\sigma_i^2)$ $(1 \le i \le n)$ are independent. Then, the MGF of $\boldsymbol{X} := (X_1\,,\ldots,X_n)'$ is

$$M_{\boldsymbol{X}}(\boldsymbol{t}) = \mathrm{E}e^{\boldsymbol{t}'\boldsymbol{X}} = \prod_{j=1}^n \mathrm{E}e^{t_jX_j} = \prod_{j=1}^n e^{t_j\mu_j + \frac12 t_j^2\sigma_j^2} = \exp\left(\sum_{j=1}^n t_j\mu_j + \frac12\sum_{j=1}^n t_j^2\sigma_j^2\right)$$

for all $\boldsymbol{t} \in \mathbf{R}^n$. □

Example 4 (MGF of χ_1^2)**.** If X is standard normal, then $Y := X^2 \sim \chi_1^2$. Now

$$\begin{aligned} M_Y(t) = \mathrm{E}e^{tX^2} &= \int_{-\infty}^{\infty} \frac{e^{tx^2 - x^2/2}}{\sqrt{2\pi}}\,dx \\ &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \exp\left(-\frac12(1-2t)x^2\right)\,dx. \end{aligned}$$

If $t \ge 1/2$, then the preceding is infinite. Otherwise, a change of variables $[y = \sqrt{1-2t}\,x]$ tells us that it is equal to

$$\int_{-\infty}^{\infty} \frac{e^{-y^2/2}}{\sqrt{2\pi}}\,\frac{dy}{\sqrt{1-2t}} = \frac{1}{\sqrt{1-2t}}.$$

In other words, $M_{X^2}(t) = \infty$ if $t \ge 1/2$ and $M_{X^2}(t) = (1-2t)^{-1/2}$ if $t < 1/2$. □

Example 5 (MGF of χ^2_n). Let $X_1, \ldots, X_n \sim \mathrm{N}(0\,,1)$ be independent, and consider the χ^2_n random variable $Y := \sum_{i=1}^n X_i^2$. Its MGF is

$$M_Y(t) = \prod_{j=1}^n M_{X_j^2}(t) = \begin{cases} (1-2t)^{-n/2} & \text{if } t < 1/2, \\ \infty & \text{if } t \geq 1/2. \end{cases}$$

According to Theorem 1, this is a formula for the MGF of the χ^2_n distribution, and identifies that distribution uniquely. □

Theorem 6 (Independence theorem of MGFs). *Let $\boldsymbol{X}$ be a random n-vector with a MGF that is finite in an open neighborhood of the origin $\mathbf{0} \in \mathbf{R}^n$. Suppose there exists $r = 1, \ldots, n$ such that*

$$M_{\boldsymbol{X}}(\boldsymbol{t}) = M_{\boldsymbol{X}}(t_1\,, \ldots, t_r\,, 0\,, \ldots, 0) \cdot M_{\boldsymbol{X}}(0\,, \ldots, 0\,, t_{r+1}\,, \ldots, t_n)$$

for all $\boldsymbol{t} \in \mathbf{R}^n$. Then, $(X_1\,, \ldots, X_r)$ and $(X_{r+1}\,, \ldots, X_n)$ are independent.

Proof. Let $\tilde{\boldsymbol{X}}$ denote an independent copy of $\boldsymbol{X}$. Define a new random vector $\boldsymbol{Y}$ as follows:

$$\boldsymbol{Y} := \begin{pmatrix} X_1 \\ \vdots \\ X_r \\ \tilde{X}_{r+1} \\ \vdots \\ \tilde{X}_n \end{pmatrix}.$$

Then,

$$M_{\boldsymbol{Y}}(\boldsymbol{t}) = M_{\boldsymbol{X}}(t_1\,, \ldots, t_r\,, 0\,, \ldots, 0) \cdot M_{\boldsymbol{X}}(0\,, \ldots, 0\,, t_{r+1}\,, \ldots, t_n).$$

According to the condition of this theorem, $\boldsymbol{X}$ and $\boldsymbol{Y}$ have the same MGF's, and therefore they have the same distribution (Theorem 1). That is, for all sets $A_1, \ldots, A_n$,

$$\mathrm{P}\left\{X_1 \in A_1\,, \ldots, X_n \in A_n\right\} = \mathrm{P}\left\{Y_1 \in A_1\,, \ldots, Y_n \in A_n\right\},$$

which is, by construction equal to

$$\mathrm{P}\left\{X_1 \in A_1\,, \ldots, X_r \in A_r\right\} \cdot \mathrm{P}\left\{\tilde{X}_{r+1} \in A_{r+1}\,, \ldots, \tilde{X}_n \in A_n\right\}.$$

Since $\tilde{\boldsymbol{X}}$ has the same distribution as $\boldsymbol{X}$, this proves that

$$\mathrm{P}\left\{X_1 \in A_1\,, \ldots, X_n \in A_n\right\} = \mathrm{P}\left\{X_1 \in A_1\,, \ldots, X_r\right\} \cdot \mathrm{P}\left\{X_{r+1} \in A_{r+1}\,, \ldots, X_n \in A_n\right\},$$

which has the desired result. □