StatLect

Index > Fundamentals of probability

Moment generating function

by Marco Taboga, PhD

The distribution of a random variable is often characterized in terms of its moment generating function (mgf), a real function whose derivatives at zero are equal to the moments of the random variable. Moment generating functions have great practical relevance not only because they can be used to easily derive moments, but also because a probability distribution is uniquely determined by its mgf, a fact that, coupled with the analytical tractability of mgfs, makes them a handy tool to solve several problems, such as deriving the distribution of a sum of two or more random variables.

It must be mentioned that not all random variables possess a moment generating function. However, all random variables possess a characteristic function, another transform that enjoys properties similar to those enjoyed by the mgf.

Definition

The following is a formal definition.

> **Definition** Let X be a random variable. If the expected value $\mathrm{E}[\exp(tX)]$ exists and is finite for all real numbers t belonging to a closed interval $[-h, h] \subseteq \mathbb{R}$, with $h > 0$, then we say that X possesses a moment generating function and the function
>
> $$M_X(t) = \mathrm{E}[\exp(tX)]$$
>
> is called the **moment generating function** of X.

The next example shows how the mgf of an exponential random variable is calculated.

Example Let X be a continuous random variable with support

$$R_X = [0, \infty)$$

and probability density function

$$f_X(x) = \begin{cases} \lambda \exp(-\lambda x) & \textit{if } x \in R_X \\ 0 & \textit{if } x \notin R_X \end{cases}$$

where λ is a strictly positive number. The expected value $\mathrm{E}[\exp(tX)]$ can be computed as follows:

$$\begin{aligned} \mathrm{E}[\exp(tX)] &= \int_{-\infty}^{\infty} \exp(tx) f_X(x) dx \\ &= \int_0^{\infty} \exp(tx) \lambda \exp(-\lambda x) dx \\ &= \lambda \int_0^{\infty} \exp((t-\lambda)x) dx \qquad \textit{(which is finite only if } t < \lambda) \\ &= \lambda \left[\frac{1}{t-\lambda} \exp((t-\lambda)x) \right]_0^{\infty} \\ &= \lambda \left[0 - \frac{1}{t-\lambda} \right] \\ &= \frac{\lambda}{\lambda - t} \end{aligned}$$

Furthermore, the above expected value exists and is finite for any $t \in [-h, h]$, provided $0 < h < \lambda$. As a consequence, X possesses a mgf:

$$M_X(t) = \frac{\lambda}{\lambda - t}$$

Deriving moments with the mgf

The moment generating function takes its name by the fact that it can be used to derive the moments of X, as stated in the following proposition.

Proposition If a random variable X possesses a mgf $M_X(t)$, then the n-th moment of X, denoted by $\mu_X(n)$, exists and is finite for any $n \in \mathbb{N}$. Furthermore,

$$\mu_X(n) = \mathrm{E}[X^n] = \left. \frac{d^n M_X(t)}{dt^n} \right|_{t=0}$$

where $\left.\frac{d^n M_X(t)}{dt^n}\right|_{t=0}$ is the n-th derivative of $M_X(t)$ with respect to t, evaluated at the point $t = 0$.

Proof

Proving the above proposition is quite complicated, because a lot of analytical details must be taken care of (see e.g. Pfeiffer - 2012). The intuition, however, is straightforward. Since the expected value is a linear operator and differentiation is a linear operation, under appropriate conditions we can differentiate through the expected value:

$$\frac{d^n M_X(t)}{dt^n} = \frac{d^n}{dt^n} \mathrm{E}[\exp(tX)] = \mathrm{E}\left[\frac{d^n}{dt^n} \exp(tX)\right] = \mathrm{E}[X^n \exp(tX)]$$

Making the substitution $t = 0$, we obtain

$$\left.\frac{d^n M_X(t)}{dt^n}\right|_{t=0} = \mathrm{E}[X^n \exp(0 \cdot X)] = \mathrm{E}[X^n] = \mu_X(n)$$

The next example shows how this proposition can be applied.

Example In the previous example we have demonstrated that the mgf of an exponential random variable is

$$M_X(t) = \frac{\lambda}{\lambda - t}$$

The expected value of X can be computed by taking the first derivative of the mgf:

$$\frac{dM_X(t)}{dt} = \frac{\lambda}{(\lambda - t)^2}$$

and evaluating it at $t = 0$:

$$\mathrm{E}[X] = \left.\frac{dM_X(t)}{dt}\right|_{t=0} = \frac{\lambda}{(\lambda - 0)^2} = \frac{1}{\lambda}$$

The second moment of X can be computed by taking the second derivative of the mgf:

$$\frac{d^2 M_X(t)}{dt^2} = \frac{2\lambda}{(\lambda - t)^3}$$

and evaluating it at $t = 0$:

$$\mathrm{E}\left[X^2\right] = \left.\frac{d^2 M_X(t)}{dt^2}\right|_{t=0} = \frac{2\lambda}{(\lambda-0)^3} = \frac{2}{\lambda^2}$$

And so on for higher moments.

Characterization of a distribution via the moment generating function

The most important property of the mgf is the following.

Proposition (Equality of distributions) Let X and Y be two random variables. Denote by $F_X(x)$ and $F_Y(y)$ their distribution functions and by $M_X(t)$ and $M_Y(t)$ their mgfs. X and Y have the same distribution (i.e., $F_X(x) = F_Y(x)$ for any x) if and only if they have the same mgfs (i.e., $M_X(t) = M_Y(t)$ for any t).

Proof

For a fully general proof of this proposition see e.g. Feller (2008). We just give an informal proof for the special case in which X and Y are discrete random variables taking only finitely many values. The "only if" part is trivial. If X and Y have the same distribution, then

$$M_X(t) = \mathrm{E}[\exp(tX)] = \mathrm{E}[\exp(tY)] = M_Y(t)$$

The "if" part is proved as follows. Denote by R_X and R_Y the supports of X and Y and by $p_X(x)$ and $p_Y(y)$ their probability mass functions. Denote by A the union of the two supports:

$$A = R_X \cup R_Y$$

and by $a_1, \ldots, a_n$ the elements of A. The mgf of X can be written as

$$\begin{aligned} M_X(t) &= \mathrm{E}[\exp(tX)] \\ &= \sum_{x \in R_X} \exp(tx) p_X(x) \quad \text{(by the definition of expected value)} \\ &= \sum_{i=1}^{n} \exp(ta_i) p_X(a_i) \quad \text{(because } p_X(a_i) = 0 \text{ if } a_i \notin R_X\text{)} \end{aligned}$$

By the same token, the mgf of Y can be written as:

$$M_Y(t) = \sum_{i=1}^{n} \exp(ta_i) p_Y(a_i)$$

> If X and Y have the same mgf, then for any t belonging to a closed neighborhood of zero
>
> $$M_X(t) = M_Y(t)$$
>
> and
>
> $$\sum_{i=1}^{n} \exp(ta_i) p_X(a_i) = \sum_{i=1}^{n} \exp(ta_i) p_Y(a_i)$$
>
> Rearranging terms, we obtain
>
> $$\sum_{i=1}^{n} \exp(ta_i)[p_X(a_i) - p_Y(a_i)] = 0$$
>
> This can be true for any t belonging to a closed neighborhood of zero only if
>
> $$p_X(a_i) - p_Y(a_i) = 0$$
>
> for every i. It follows that that the probability mass functions of X and Y are equal. As a consequence, also their distribution functions are equal.

It must be stressed that this proposition is extremely important and relevant from a practical viewpoint: in many cases where we need to prove that two distributions are equal, it is much easier to prove equality of the moment generating functions than to prove equality of the distribution functions. Also note that equality of the distribution functions can be replaced in the proposition above by equality of the probability mass functions (if X and Y are discrete random variables) or by equality of the probability density functions (if X and Y are continuous random variables).

More details

The following sections contain more details about the mgf.

Moment generating function of a linear transformation

Let X be a random variable possessing a mgf $M_X(t)$. Define

$$Y = a + bX$$

where $a, b \in \mathbb{R}$ are two constants and $b \neq 0$. Then, the random variable Y possesses a mgf $M_Y(t)$ and

$$M_Y(t) = \exp(at) M_X(bt)$$

Proof

By the very definition of mgf, we have

$$\begin{aligned} M_Y(t) &= \mathrm{E}[\exp(tY)] \\ &= \mathrm{E}[\exp(at + btX)] \\ &= \mathrm{E}[\exp(at)\exp(btX)] \\ &= \exp(at)\mathrm{E}[\exp(btX)] \\ &= \exp(at) M_X(bt) \end{aligned}$$

Obviously, if $M_X(t)$ is defined on a closed interval $[-h, h]$, then $M_Y(t)$ is defined on the interval $\left[-\frac{h}{b}, \frac{h}{b}\right]$.

Moment generating function of a sum of mutually independent random variables

Let X_1, ..., X_n be n mutually independent random variables. Let Z be their sum:

$$Z = \sum_{i=1}^{n} X_i$$

Then, the mgf of Z is the product of the mgfs of X_1, ..., X_n:

$$M_Z(t) = \prod_{i=1}^{n} M_{X_i}(t)$$

Proof

This is easily proved by using the definition of mgf and the properties of mutually independent variables:

$$\begin{aligned} M_Z(t) &= \mathrm{E}[\exp(tZ)] \\ &= \mathrm{E}\left[\exp\left(t\sum_{i=1}^{n} X_i\right)\right] \\ &= \mathrm{E}\left[\exp\left(\sum_{i=1}^{n} tX_i\right)\right] \\ &= \mathrm{E}\left[\prod_{i=1}^{n} \exp(tX_i)\right] \\ &= \prod_{i=1}^{n} \mathrm{E}[\exp(tX_i)] && \text{(by mutual independence)} \\ &= \prod_{i=1}^{n} M_{X_i}(t) && \text{(by the definition of mgf)} \end{aligned}$$

Solved exercises

Some solved exercises on moment generating functions can be found below.

Exercise 1

Let X be a discrete random variable having a Bernoulli distribution. Its support R_X is

$$R_X = \{0, 1\}$$

and its probability mass function $p_X(x)$ is

$$p_X(x) = \begin{cases} p & \text{if } x = 1 \\ 1-p & \text{if } x = 0 \\ 0 & \text{if } x \notin R_X \end{cases}$$

where $p \in (0, 1)$ is a constant. Derive the moment generating function of X, if it exists.

Solution

By the definition of moment generating function, we have

$$\begin{aligned} M_X(t) &= \mathrm{E}[\exp(tX)] \\ &= \sum_{x \in R_X} \exp(tx) p_X(x) \\ &= \exp(t \cdot 1) \cdot p_X(1) + \exp(t \cdot 0) \cdot p_X(0) \\ &= \exp(t) \cdot p + 1 \cdot (1-p) \\ &= 1 - p + p\exp(t) \end{aligned}$$

Obviously, the moment generating function exists and it is well-defined because the above expected value exists for any $t \in \mathbb{R}$.

Exercise 2

Let X be a random variable with moment generating function

$$M_X(t) = \frac{1}{2}(1 + \exp(t))$$

Derive the variance of X.

Solution

We can use the following formula for computing the variance:

$$\text{Var}[X] = \text{E}[X^2] - \text{E}[X]^2$$

The expected value of X is computed by taking the first derivative of the moment generating function:

$$\frac{dM_X(t)}{dt} = \frac{1}{2}\exp(t)$$

and evaluating it at $t = 0$:

$$\text{E}[X] = \left.\frac{dM_X(t)}{dt}\right|_{t=0} = \frac{1}{2}\exp(0) = \frac{1}{2}$$

The second moment of X is computed by taking the second derivative of the moment generating function:

$$\frac{d^2M_X(t)}{dt^2} = \frac{1}{2}\exp(t)$$

and evaluating it at $t = 0$:

$$\text{E}[X^2] = \left.\frac{d^2M_X(t)}{dt^2}\right|_{t=0} = \frac{1}{2}\exp(0) = \frac{1}{2}$$

Therefore,

$$\begin{aligned}\text{Var}[X] &= \text{E}[X^2] - \text{E}[X]^2 \\ &= \frac{1}{2} - \left(\frac{1}{2}\right)^2 \\ &= \frac{1}{2} - \frac{1}{4} \\ &= \frac{1}{4}\end{aligned}$$

Exercise 3

A random variable X is said to have a Chi-square distribution with n degrees of freedom if its moment generating function is defined for any $t < \frac{1}{2}$ and it is equal to

$$M_X(t) = (1-2t)^{-n/2}$$

Define

$$Y = X_1 + X_2$$

where X_1 and X_2 are two independent random variables having Chi-square distributions with n_1 and n_2 degrees of freedom respectively. Prove that Y has a Chi-square distribution with $n_1 + n_2$ degrees of freedom.

Solution

The moment generating functions of X_1 and X_2 are

$$M_{X_1}(t) = (1-2t)^{-n_1/2}$$
$$M_{X_2}(t) = (1-2t)^{-n_2/2}$$

The moment generating function of a sum of independent random variables is just the product of their moment generating functions:

$$\begin{aligned} M_Y(t) &= (1-2t)^{-n_1/2}(1-2t)^{-n_2/2} \\ &= (1-2t)^{-(n_1+n_2)/2} \end{aligned}$$

Therefore, $M_Y(t)$ is the moment generating function of a Chi-square random variable with $n_1 + n_2$ degrees of freedom. As a consequence, Y has a Chi-square distribution with $n_1 + n_2$ degrees of freedom.

References

Feller, W. (2008) An introduction to probability theory and its applications, Volume 2, Wiley.

Pfeiffer, P. E. (1978) Concepts of probability theory, Dover Publications.