More on Moment Generating Functions

Generating Functions

Moment generating functions are an example of a general class of functions called **generating functions**.

Definition. Given a sequence of numbers (finite or infinite) $a_0, a_1, a_2, \ldots$, we define the **generating function** of the sequence $\{a_j\}$ as

$$G(t) = a_0 + a_1 t + a_2 t^2 + \cdots$$

For certain sequences, we can find a closed form of $G(t)$, and often that closed form will help us understand the sequence at a deeper level.

Examples.

- Let $a_j = 1$ for all $j = 0, 1, 2, \ldots$, then for $|t| < 1$, $G(t) = 1 + t + t^2 + \cdots$ is a geometric series that converges to $G(t) = \frac{1}{1-t}$.
- The generating function of the sequence $\{0, 0, 1, 1, 1, \ldots\}$ is $G(t) = \frac{t^2}{1-t}$.
- For fixed $n \in \mathbb{Z}^+$, the sequence $a_j = \binom{n}{j}$ for $j = 1, \ldots, n$ has the generating function

$$G(t) = \sum_{j=0}^{n} \binom{n}{j} t^j = \sum_{j=0}^{n} \binom{n}{j} t^j 1^{n-j} = (1+t)^n$$

  by the Binomial Theorem.

Probability Generating Function

Definition. Let X be a discrete random variable with support $\mathcal{X} = \{0, 1, 2\ldots,\}$, and define its probability mass function as

$$f_X(k) = P(X = k) = p_k.$$

Then the generating function for the sequence $\{p_0, p_1, p_2, \ldots\}$ is called the **probability generating function** for X:

$$P_X(t) = p_0 + p_1 t + p_2 t^2 + \cdots.$$

Since $\sum_{k=0}^{\infty} p_k = 1$, the series $P_X(t)$ converges absolutely at least for $-1 \leq t \leq 1$. Examining this expression more closely, we see that the probability generating function is equal to the expected value of t^X, $E(t^X)$:

$$P_X(t) = \sum_{k=0}^{\infty} p_k t^k = E(t^X).$$

As with the moment generating function, derivatives of the probability generating function can prove useful. Examine the first derivative:

$$P_X'(t) = \frac{d}{dt} \sum_{k=0}^{\infty} p_k t^k = \sum_{k=0}^{\infty} \frac{d}{dt} p_k t^k = \sum_{k=0}^{\infty} k p_k t^{k-1}.$$

Again, this series converges for at least $-1 < t < 1$, and when $t = 1$, the right side reduces to $\sum_{k=0}^{\infty} k p_k = E(X)$. Thus, the first derivative of the probability generating function evaluated at $t = 1$ is equal to the expected value of X. Similarly, one can show that

$$Var(X) = P_X''(1) + P_X'(1) - \left[P_X'(1)\right]^2.$$

There is an interesting relationship between the probability generating function and the generating function for a discrete distribution's tail probabilities, $q_k = P(X > k)$. Define the generating function for the sequence of tail probabilities as

$$Q_X(t) = q_0 + q_1 t + q_2 t^2 + \cdots .$$

Since $0 < q_k < 1$, $Q_X(t)$ converges for at least $-1 < t < 1$. Then, since $q_{n-1} - q_n = p_n$, for $-1 < s < 1$,

$$Q_X(t) = \frac{1 - P_X(t)}{1 - t}.$$

This result can be used the show the following relations:

- $E(X) = P'(1) = Q(1)$
- $E(X(X-1)) = P''(1) = 2Q'(1)$
- $Var(X) = P''(1) + P'(1) - [P'(1)]^2 = 2Q'(1) + Q(1) - [Q(1)]^2$

Moment Generating Functions

Definition 2.3.6 in Casella and Berger (p. 62) defines the **moment generating function (mgf) of** X as

$$M_X(t) = E(e^{tX}).$$

However, this is only the closed form solution of the moment generating function, and, in fact, this moment generating function is the generating function for not the moments themselves, but for $\mu_k / k!$, where $\mu_k = E(X^k)$ is the kth moment.

Power series definition of moment generating function. If all moments $\mu_r = E(X^r)$ of a random variable X exist, then the **moment generating function** of X is defined as the power series expansion

$$M_X(t) = \sum_{r=0}^{\infty} \frac{\mu_r}{r!} t^r.$$

For an intuitive explanation of how these two definitions are related, take the Taylor series expansion of the exponential function,

$$e^z = \sum_{k=0}^{\infty} \frac{z^k}{k!},$$

plug in tX for z, and then take the expectation to get

$$M_X(t) = E(e^{tX}) = E\left[\sum_{k=0}^{\infty} \frac{(tX)^k}{k!}\right] = \sum_{k=0}^{\infty} \frac{\mu_k}{k!} t^k$$

This power series definition of the moment generating function suggests the following approach to finding the moments of X:

1. Find the moment generating function of X, $M_X(t)$.
2. Expand $M_X(t)$ into a power series in t, i.e., express $M_X(t)$ as

$$M_X(t) = \sum_{k=0}^{\infty} a_k t^k.$$

3. Set $\mu_k = k! a_k$.

Example. It is straightforward to show that the moment generating function of a continuous uniform random variable X over the interval $[a, b]$ is equal to

$$M_X(t) = E(e^{tX}) = \frac{e^{bt} - e^{at}}{(b-a)t}.$$

We can use the Taylor series expansion of the exponential function to write $e^{bt} = \sum_{k=0}^{\infty}(bt)^k/k!$ and $e^{at} = \sum_{k=0}^{\infty}(at)^k/k!$. Substituting these expressions into $M_X(t)$, we get

$$\begin{aligned} M_X(t) &= \frac{1}{(b-a)t}\left[\sum_{k=0}^{\infty}\frac{(bt)^k}{k!} - \sum_{k=0}^{\infty}\frac{(at)^k}{k!}\right] \\ &= \sum_{k=0}^{\infty}\frac{b^k - a^k}{(b-a)k!}t^{k-1} \\ &= 0 + \sum_{k=1}^{\infty}\frac{b^k - a^k}{(b-a)k!}t^{k-1} \\ &= \sum_{j=0}^{\infty}\frac{b^{j+1} - a^{j+1}}{(b-a)(j+1)!}t^j. \end{aligned}$$

Thus, the jth moment of a Uniform distribution over the interval $[a, b]$ is

$$\mu_j = j! \times \frac{b^{j+1} - a^{j+1}}{(b-a)(j+1)!} = \frac{b^{j+1} - a^{j+1}}{(b-a)(j+1)}.$$

Finding moments when the MGF is undefined at $t = 0$

By Theorem 2.3.7 of Casella and Berger (p. 62), if the mgf of X exists for a neighborhood around zero, i.e., for all t in $-h < t < h$ for some $h > 0$, then the kth moment of the distribution of X is equal to the kth derivative of $M_X(t)$ evaluated at $t = 0$. But what if $M_X(t)$ (and its derivatives) are undefined at $t = 0$?

Example (cont). The Uniform distribution over the interval $[a, b]$ is such an example. Its mgf is

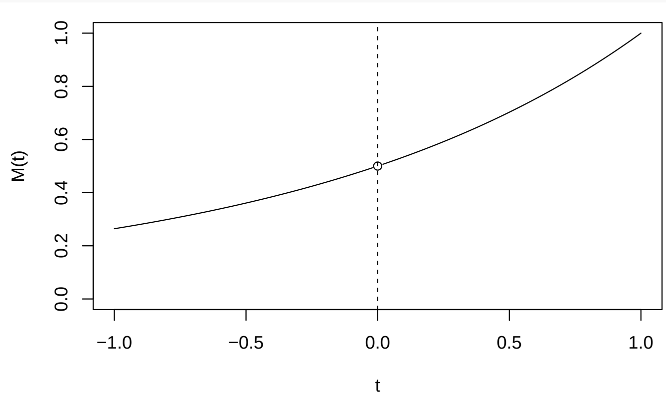
$$M_X(t) = E(e^{tX}) = \begin{cases} \dfrac{e^{bt} - e^{at}}{(b-a)t} & t \neq 0 \\ 1 & t = 0 \end{cases}$$

Theorem 2.3.7 is not applicable in this case. For $t \neq 0$, the first derivative of $M_X(t)$ is

$$M_X'(t) = \frac{-1}{(b-a)t^2}\left[e^{bt} - e^{at}\right] + \frac{1}{(b-a)t}\left[be^{bt} - ae^{at}\right] = \frac{1}{(b-a)t^2}\left[e^{bt}(bt-1) - e^{at}(at-1)\right].$$

For example, if $a = 0$ and $b = 1$, this derivative looks like

```r
curve((1/x^2)*(exp(x)*(x-1)+1), from = -1, to =1,
      xlab = "t", ylab = "M(t)", ylim = c(0,1))
abline(v=0, lty=2)
points(x  = 0, y = 0.5)
```



There is a removable discontinuity in this function at $t = 0$, since

$$\lim_{t\to 0} M_X'(t) = \frac{a+b}{2}.$$

So, just like $M_X(t)$ itself, we define

$$M_X'(t) = \begin{cases} \dfrac{1}{(b-a)t^2}\left[e^{bt}(bt-1) - e^{at}(at-1)\right] & t \neq 0 \\ \dfrac{a+b}{2} & t = 0 \end{cases},$$

and, by the definition of a derivative,

$$E(X) = M_X'(0) = \lim_{t \to 0} \frac{1}{t}\left[M_X'(t) - M_X(0)\right] = \frac{a+b}{2}.$$

Similarly, the kth moment of X for $k \in \mathbb{Z}^+$ is given by

$$E(X^k) = \lim_{t \to 0} M^{(k)}(t).$$

References

Casella, G., and Berger, R. L. (2002). *Statistical Inference*, 2nd ed. Duxbury.

Cohen, D. I. A. (1978). *Basic Techniques of Combinatorial Theory.* John Wiley & Sons, Inc.

Feller, W. (1968). *An Introduction to Probability Theory and Its Applications, Volume 1*, 3rd ed. John Wiley & Sons, Inc.

Žitković, G. (2019). "Lecture 6: Moment-generating functions." https://web.ma.utexas.edu/users/gordanz/notes/mgf_color.pdf