Section 3.5 Moment generating functions

Thursday, September 8, 2016 1:12 PM

Section 3.5 Moment ...

We have used the well-known fact that $\sum_{x=1}^{m} x^2 = m(m+1)(2m+1)/6$; this can be proven by induction. The variance of X is equal to

$$\begin{aligned}\sigma^2 &= E(X^2) - [E(X)]^2 \\ &= \frac{(m+1)(2m+1)}{6} - \left(\frac{m+1}{2}\right)^2 = \frac{m^2-1}{12}. \ \Box\end{aligned}$$

EXERCISE: Find $\sigma^2 = V(Y)$ in Examples 3.5 and 3.8 (notes).

IMPORTANT RESULT: Let Y be a random variable (not necessarily a discrete random variable). Suppose that a and b are fixed constants. Then

$$V(a + bY) = b^2 V(Y).$$

REMARK: Taking $b = 0$ above, we see that $V(a) = 0$, for any constant a. This makes sense intuitively. The variance is a measure of variability for a random variable; a constant (such as a) does not vary. Also, by taking $a = 0$, we see that $V(bY) = b^2 V(Y)$.

3.5 Moment generating functions

Mean: $E[Y]$

Variance: $E[Y^2] - (E[Y])^2$

TERMINOLOGY: Let Y be a discrete random variable with pmf $p_Y(y)$ and support R. The **moment generating function** (**mgf**) for Y, denoted by $m_Y(t)$, is given by

$$m_Y(t) = E(e^{tY}) = \sum_{y \in R} e^{ty} p_Y(y),$$

provided $E(e^{tY}) < \infty$ for all t in an open neighborhood about 0; i.e., there exists some $h > 0$ such that $E(e^{tY}) < \infty$ for all $t \in (-h, h)$. If $E(e^{tY})$ does not exist in an open neighborhood of 0, we say that the moment generating function does not exist.

$M_Y(0) = E(e^{0Y}) = E(1) = 1$

TERMINOLOGY: We call $\mu'_k \equiv E(Y^k)$ the k**th moment** of the random variable Y:

$E(Y)$	1st moment (mean!)
$E(Y^2)$	2nd moment
$E(Y^3)$	3rd moment
$E(Y^4)$	4th moment
$\vdots$	$\vdots$

REMARK: The moment generating function (mgf) can be used to generate moments. In fact, from the theory of Laplace transforms, it follows that if the mgf exists, it characterizes an infinite set of moments. So, how do we generate moments?

RESULT: Let Y denote a random variable (not necessarily a discrete random variable) with support R and mgf $m_Y(t)$. Then,

$$E(Y^k) = \frac{d^k m_Y(t)}{dt^k}\bigg|_{t=0}.$$

$\frac{d}{dt} e^{ty} = y e^{ty}$

Note that derivatives are taken with respect to t.

Proof. Assume, without loss, that Y is discrete. With $k = 1$, we have

$$\frac{d}{dt} m_Y(t) = \frac{d}{dt}\sum_{y\in R} e^{ty} p_Y(y) = \sum_{y\in R} \frac{d}{dt}\left(e^{ty} p_Y(y)\right) = \sum_{y\in R} \overset{g(y)}{y e^{ty}}\, p_Y(y) = \underbrace{E(Ye^{tY})}_{g(Y)}.$$

$m_Y(t) = 1 + t\mu_1' + \frac{t^2}{2!}\mu_2' + \frac{t^3}{3!}\mu_3' + \cdots$

$\frac{d\,M_Y(t)}{dt} = \mu_1' + \frac{2t}{2!}\mu_2' + \frac{3t^2}{3!}\mu_3' + \cdots$

$\frac{dM_Y(t)}{dt}\Big|_{t=0} = \mu_1'$

$\frac{d^2M_Y(t)}{dt^2}\Big|_{t=0} = \frac{2}{2!}\mu_2' + \cancel{\frac{3!\,t}{3!}\mu_3' + \cdots}$

$= \mu_2'$

Thus,

$$\frac{dm_Y(t)}{dt}\bigg|_{t=0} = E(Ye^{tY})\Big|_{t=0} = E(Y).$$

Continuing to take higher-order derivatives, we can prove that

$$\frac{d^k m_Y(t)}{dt^k}\bigg|_{t=0} = E(Y^k),$$

for any integer $k \geq 1$. See pp 139-140 (WMS) for a slightly different proof. □

ASIDE: In the proof of the last result, we interchanged the derivative and (possibly infinite) sum. This is permitted as long as $m_Y(t) = E(e^{tY})$ exists.

MEANS AND VARIANCES: Suppose that Y is a random variable (not necessarily a discrete random variable) with mgf $m_Y(t)$. We know that

$$E(Y) = \frac{dm_Y(t)}{dt}\bigg|_{t=0}$$

and

$$E(Y^2) = \frac{d^2 m_Y(t)}{dt^2}\bigg|_{t=0}.$$

We can get $V(Y)$ using $V(Y) = E(Y^2) - [E(Y)]^2$.

REMARK: Being able to find means and variances is important in mathematical statistics. **Thus, we can use the mgf as a tool to do this.** This is helpful because sometimes computing

$$E(Y) = \sum_{y \in R} y p_Y(y)$$

directly (or even higher order moments) may be extremely difficult, depending on the form of $p_Y(y)$.

Example 3.10. Suppose that Y is a random variable with pmf

$$p_Y(y) = \begin{cases} \left(\frac{1}{2}\right)^y, & y = 1, 2, 3, ... \\ 0, & \text{otherwise.} \end{cases}$$

Is this pmf valid?

1. $P_Y(y) \geq 0$ ✓

2. $\sum_{y \in R} P_Y(y) = 1$ ✓

$$\sum_{y=1}^{\infty} \left(\frac{1}{2}\right)^y = \sum_{x=0}^{\infty} \frac{1}{2}\left(\frac{1}{2}\right)^x = \frac{\frac{1}{2}}{1-\frac{1}{2}} = 1$$

$$\left[\sum_{y=0}^{\infty} \left(\frac{1}{2}\right)^y\right] - 1 = \frac{1}{1-\frac{1}{2}} - 1 = 2 - 1 = 1$$

$$\sum_{x=0}^{\infty} ar^x = \frac{a}{1-r} \quad |r|<1$$

Find the mean of Y.

SOLUTION. Using the definition of expected value, the mean of Y is given by

$$E(Y) = \sum_{y \in R} y p_Y(y) = \sum_{y=1}^{\infty} y \left(\frac{1}{2}\right)^y.$$

Finding this infinite sum is not obvious (at least, this sum is not a geometric sum). Another option is to use moment generating functions! The mgf of Y is given by

$$\begin{aligned} m_Y(t) = E(e^{tY}) &= \sum_{y \in R} e^{ty} p_Y(y) \\ &= \sum_{y=1}^{\infty} e^{ty} \left(\frac{1}{2}\right)^y = \sum_{y=1}^{\infty} \left(\frac{e^t}{2}\right)^y = \left[\sum_{y=0}^{\infty} \left(\frac{e^t}{2}\right)^y\right] - 1. \end{aligned}$$

$$\left|\frac{e^t}{2}\right| < 1 \Rightarrow e^t < 2 \Rightarrow t < \ln 2$$

The series $\sum_{y=0}^{\infty}(e^t/2)^y$ is an infinite geometric sum with common ratio $r = e^t/2$. This series converges as long as $e^t/2 < 1$, in which case

$$m_Y(t) = \frac{1}{1 - \frac{e^t}{2}} - 1 = \frac{e^t}{2 - e^t},$$

$$= \left(\frac{2}{2-e^t} - 1\right) = \frac{2}{2-e^t} - \frac{2-e^t}{2-e^t} = \frac{e^t}{2-e^t}$$

for $e^t/2 < 1 \iff t < \ln 2$. Note that $(-h, h)$ with $h = \ln 2$ is an open neighborhood around zero for which $m_Y(t)$ exists. Now,

$$\left(\frac{f(t)}{g(t)}\right)' = \frac{f'(t)g(t) - f(t)g'(t)}{g^2(t)}$$

$$\begin{aligned} E(Y) = \left.\frac{dm_Y(t)}{dt}\right|_{t=0} &= \left.\frac{d}{dt}\left(\frac{e^t}{2 - e^t}\right)\right|_{t=0} \\ &= \left.\frac{e^t(2 - e^t) - e^t(-e^t)}{(2 - e^t)^2}\right|_{t=0} = 2. \ \square \end{aligned}$$

✓

Example 3.11. Let the random variable Y have pmf $p_Y(y)$ given by

$$p_Y(y) = \begin{cases} \frac{1}{6}(3-y), & y = 0, 1, 2 \\ 0, & \text{otherwise.} \end{cases}$$

$$\sum_{y=0}^{2} P_Y(y) = \frac{3}{6} + \frac{2}{6} + \frac{1}{6} = 1$$

Simple calculations show that $E(Y) = 2/3$ and $V(Y) = 5/9$ (verify!). Let's "check" these calculations using the mgf of Y. It is given by

$$(e^{at})' = ae^{at}$$

$$\begin{aligned} m_Y(t) = E(e^{tY}) &= \sum_{y \in R} e^{ty} p_Y(y) \\ &= e^{t(0)}\frac{3}{6} + e^{t(1)}\frac{2}{6} + e^{t(2)}\frac{1}{6} \\ &= \frac{3}{6} + \frac{2}{6}e^t + \frac{1}{6}e^{2t}. \end{aligned}$$

$$\left(\frac{1}{6}e^{2t}\right)' = \frac{1}{6} \times (e^{2t})'$$
$$= \frac{1}{6} \times e^{2t} \times (2t)'$$
$$= \frac{2}{6} \times e^{2t}$$

$$(e^t)' = e^t$$
$$(e^{2t})' = 2e^{2t}$$

Taking derivatives of $m_Y(t)$ with respect to t, we get

$$\begin{aligned} \frac{d}{dt}m_Y(t) &= \frac{2}{6}e^t + \frac{2}{6}e^{2t} \\ \frac{d^2}{dt^2}m_Y(t) &= \frac{2}{6}e^t + \frac{4}{6}e^{2t}. \end{aligned}$$

Thus,

$$\begin{aligned} E(Y) &= \left.\frac{dm_Y(t)}{dt}\right|_{t=0} = \frac{2}{6}e^0 + \frac{2}{6}e^{2(0)} = 4/6 = 2/3 \\ E(Y^2) &= \left.\frac{d^2m_Y(t)}{dt^2}\right|_{t=0} = \frac{2}{6}e^0 + \frac{4}{6}e^{2(0)} = 1 \end{aligned}$$

so that

$$V(Y) = E(Y^2) - [E(Y)]^2 = 1 - (2/3)^2 = 5/9.$$

In this example, it is easier to compute $E(Y)$ and $V(Y)$ directly (using the definition). However, it nice to see that we get the same answer using the mgf approach. □

REMARK: Not only is the mgf a tool for computing moments, but it also helps us to characterize a probability distribution. How? When an mgf exists, it happens to be unique. This means that if two random variables have same mgf, then they have the same probability distribution! This is called the **uniqueness property** of mgfs (it is based on the uniqueness of Laplace transforms). For now, however, it suffices to envision the mgf as a "special expectation" that generates moments. This, in turn, helps us to compute means and variances of random variables.