

3.8: Moment-Generating Functions (MGFs) for Discrete Random Variables

The expected value and variance of a random variable are actually special cases of a more general class of numerical characteristics for random variables given by *moments*.

Definition 3.8.1

The r^{th} **moment** of a random variable X is given by

$$E[X^r].$$

The r^{th} **central moment** of a random variable X is given by

$$E[(X - \mu)^r],$$

where $\mu = E[X]$.

Note that the **expected value** of a random variable is given by the **first moment**, i.e., when $r = 1$. Also, the **variance** of a random variable is given the **second central moment**.

As with expected value and variance, the moments of a random variable are used to characterize the distribution of the random variable and to compare the distribution to that of other random variables. Moments can be calculated directly from the definition, but, even for moderate values of r , this approach becomes cumbersome. The next definition and theorem provide an easier way to generate moments.

Definition 3.8.2

The **moment-generating function (mgf)** of a random variable X is given by

$$M_X(t) = E[e^{tX}], \quad \text{for } t \in \mathbb{R}.$$

Theorem 3.8.1

If random variable X has mgf $M_X(t)$, then

$$M_X^{(r)}(0) = \frac{d^r}{dt^r} [M_X(t)]_{t=0} = E[X^r].$$

In other words, the r^{th} derivative of the mgf evaluated at $t = 0$ gives the value of the r^{th} moment.

Theorem 3.8.1 tells us how to derive the mgf of a random variable, since the mgf is given by taking the expected value of a function applied to the random variable:

$$M_X(t) = E[e^{tX}] = \sum_i e^{tx_i} \cdot p(x_i)$$

We can now derive the first moment of the Poisson distribution, i.e., derive the fact we mentioned in Section 3.6, but left as an [exercise](#), that the expected value is given by the parameter λ . We also find the variance.

Example 3.8.1

Let $X \sim \text{Poisson}(\lambda)$. Then, the pmf of X is given by

$$p(x) = \frac{e^{-\lambda} \lambda^x}{x!}, \quad \text{for } x = 0, 1, 2, \dots$$

Before we derive the mgf for X , we recall from calculus the Taylor series expansion of the exponential function e^y :

$$e^y = \sum_{x=0}^{\infty} \frac{y^x}{x!}.$$

Using this fact, we find

$$M_X(t) = E[e^{tX}] = \sum_{x=0}^{\infty} e^{tx} \cdot \frac{e^{-\lambda} \lambda^x}{x!} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{(e^t \lambda)^x}{x!} = e^{-\lambda} e^{e^t \lambda} = e^{\lambda(e^t - 1)}.$$

Now we take the first and second derivatives of $M_X(t)$. Remember we are differentiating with respect to t :

$$\begin{aligned} M'_X(t) &= \frac{d}{dt} \left[e^{\lambda(e^t - 1)} \right] = \lambda e^t e^{\lambda(e^t - 1)} \\ M''_X(t) &= \frac{d}{dt} \left[\lambda e^t e^{\lambda(e^t - 1)} \right] = \lambda e^t e^{\lambda(e^t - 1)} + \lambda^2 e^{2t} e^{\lambda(e^t - 1)} \end{aligned}$$

Next we evaluate the derivatives at $t = 0$ to find the first and second moments of X :

$$\begin{aligned} E[X] &= M'_X(0) = \lambda e^0 e^{\lambda(e^0 - 1)} = \lambda \\ E[X^2] &= M''_X(0) = \lambda e^0 e^{\lambda(e^0 - 1)} + \lambda^2 e^{2 \cdot 0} e^{\lambda(e^0 - 1)} = \lambda + \lambda^2 \end{aligned}$$

Finally, in order to find the variance, we use the alternate formula:

$$\text{Var}(X) = E[X^2] - (E[X])^2 = \lambda + \lambda^2 - \lambda^2 = \lambda.$$

Thus, we have shown that both the mean and variance for the Poisson(λ) distribution is given by the parameter λ .

Note that the mgf of a random variable is a *function* of t . The main application of mgf's is to find the moments of a random variable, as the previous example demonstrated. There are more properties of mgf's that allow us to find moments for functions of random variables.

Theorem 3.8.2

Let X be a random variable with mgf $M_X(t)$, and let a, b be constants. If random variable $Y = aX + b$, then the mgf of Y is given by

$$M_Y(t) = e^{bt} M_X(at).$$

Theorem 3.8.3

If X_1, \dots, X_n are independent random variables with mgf's $M_{X_1}(t), \dots, M_{X_n}(t)$, respectively, then the mgf of random variable $Y = X_1 + \dots + X_n$ is given by

$$M_Y(t) = M_{X_1}(t) \cdots M_{X_n}(t).$$

Recall that a binomially distributed random variable can be written as a sum of independent Bernoulli random variables. We use this and Theorem 3.8.3 to derive the mean and variance for a binomial distribution. First, we find the mean and variance of a Bernoulli distribution.

Example 3.8.2

Recall that X has a Bernoulli(p) distribution if it is assigned the value of 1 with probability p and the value of 0 with probability $1 - p$. Thus, the pmf of X is given by

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

In order to find the mean and variance of X , we first derive the mgf:

$$M_X(t) = E[e^{tX}] = e^{t(0)}(1-p) + e^{t(1)}p = 1 - p + e^t p.$$

Now we differentiate $M_X(t)$ with respect to t :

$$\begin{aligned} M'_X(t) &= \frac{d}{dt} [1 - p + e^t p] = e^t p \\ M''_X(t) &= \frac{d}{dt} [e^t p] = e^t p \end{aligned}$$

Next we evaluate the derivatives at $t = 0$ to find the first and second moments:

$$M'_X(0) = M''_X(0) = e^0 p = p.$$

Thus, the expected value of X is $E[X] = p$. Finally, we use the alternate formula for calculating variance:

$$\text{Var}(X) = E[X^2] - (E[X])^2 = p - p^2 = p(1 - p).$$

Example 3.8.3

Let $X \sim \text{binomial}(n, p)$. If X_1, \dots, X_n denote n independent Bernoulli(p) random variables, then we can write

$$X = X_1 + \dots + X_n.$$

In Example 3.8.2, we found the mgf for a Bernoulli(p) random variable. Thus, we have

$$M_{X_i}(t) = 1 - p + e^t p, \quad \text{for } i = 1, \dots, n.$$

Using Theorem 3.8.3, we derive the mgf for X :

$$M_X(t) = M_{X_1}(t) \cdots M_{X_n}(t) = (1 - p + e^t p) \cdots (1 - p + e^t p) = (1 - p + e^t p)^n.$$

Now we can use the mgf of X to find the moments:

$$\begin{aligned} M'_X(t) &= \frac{d}{dt} [(1 - p + e^t p)^n] = n(1 - p + e^t p)^{n-1} e^t p \\ &\Rightarrow M'_X(0) = np \\ M''_X(t) &= \frac{d}{dt} [n(1 - p + e^t p)^{n-1} e^t p] = n(n-1)(1 - p + e^t p)^{n-2} (e^t p)^2 + n(1 - p + e^t p)^{n-1} e^t p \\ &\Rightarrow M''_X(0) = n(n-1)p^2 + np \end{aligned}$$

Thus, the expected value of X is $E[X] = np$, and the variance is

$$\text{Var}(X) = E[X^2] - (E[X])^2 = n(n-1)p^2 + np - (np)^2 = np(1 - p).$$

We end with a final property of mgf's that relates to the comparison of the distribution of random variables.

Theorem 3.8.4

The mgf $M_X(t)$ of random variable X uniquely determines the probability distribution of X . In other words, if random variables X and Y have the same mgf, $M_X(t) = M_Y(t)$, then X and Y have the same probability distribution.

Exercise 3.8.1

Suppose the random variable X has the following mgf:

$$M_X(t) = (0.85 + 0.15e^t)^{33}$$

What is the distribution of X ?

Hint

Use [Theorem 3.8.4](#) and look at [Example 3.8.3](#).

Answer

We found in [Example 3.8.3](#) that the mgf for a binomial distribution is

$$M_X(t) = (1 - p + e^t p)^n,$$

which is the mgf given with $p = 0.15$ and $n = 33$. Thus, $X \sim \text{binomial}(33, 0.15)$

This page titled [3.8: Moment-Generating Functions \(MGFs\) for Discrete Random Variables](#) is shared under a [not declared](#) license and was authored, remixed, and/or curated by [Kristin Kuter](#).