Section 3.5: Moments and Moment Generating Functions

Definition 1 *Expected Values of integer powers of X and $X - \mu$ are called* ***moments****. For powers of X these are called* ***moments about zero****. For powers of $X - \mu$ these are called* ***moments about the mean.***

Example 2 *$E(X^4)$ is the fourth moments about 0. $E((X-\mu)^3)$ is called the third moment about the mean.*

Example 3 *The* ***expected value*** *(or* ***mean****) is the first moment about 0.*

$$E(X) = \sum_{x \in X} x * p(x).$$

Example 4 *The* ***variance,*** *$V(X)$, is the second moment about the mean.*

$$\sigma^2 = \sum_{x \in X} (x - E(X))^2 * p(x).$$

Example 5 *Back to the experiment of rolling a pair of dice and summing the faces. The random variable X assigns to each roll its sum. The following table is our previously constructed probability model.*

Sum	2	3	4	5	6	7	8	9	10	11	12
Prob	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

What is the first moment about 0 of the sum of a pair of dice? $E(X) = \sum\limits_{x \in X} x * p(x) = 7$. What is the second moment about the mean of the sum of a pair of dice? $V(X) = \sigma^2 = \sum_{x \in X} (x - E(X))^2 * p(x) = \frac{35}{6}$.

Remark 6 *The third moment about the mean (relative to standard deviation) is a measure of skewness.*

$$\frac{E\left((X-\mu)^3\right)}{\sigma^3} = E\left(\left(\frac{X-\mu}{\sigma}\right)^3\right)$$

For a symmetric distribution, the third moment about the mean is 0.

Example 7 *Compute the third moment about the mean when rolling a pair of dice and summing the faces.*

Sum	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*
$(x - E(X))^3$	*-125*	*-64*	*-27*	*-8*	*-1*	*0*	*1*	*8*	*27*	*64*	*125*
Prob	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

$$\begin{aligned} E\left((X-\mu)^3\right) &= \sum_{x \in X} (x - E(X))^3 * p(x) \\ &= -125 * \frac{1}{36} - 64 * \frac{2}{36} - 27 * \frac{3}{36} - 8 * \frac{4}{36} - 1 * \frac{5}{36} + 0 * \frac{6}{36} \\ &\quad +1 * \frac{5}{36} + 8 * \frac{4}{36} + 27 * \frac{3}{36} + 64 * \frac{2}{36} + 125 * \frac{1}{36} \\ &= 0. \end{aligned}$$

And finally, $\frac{E\left((X-\mu)^3\right)}{\sigma^3} = \frac{0}{\sigma^3} = 0.$

Remark 8 *A negative third moment about the mean indicates left skew. Similarly, A positive third moment about the mean indicates right skew.*

Example 9 *Determine skewness for the pdf defined in the table below.*

X	1	3	6	10
$p(x)$	.4	.3	.2	.1

$E(X) = 1 * .4 + 3 * .3 + 6 * .2 + 10 * .1 = 3.5$
$V(X) = \sum_{x \in X} (x - E(X))^2 * p(x) = (1-3.15)^2 * .4 + (3-3.15)^2 * .3 + (6-3.15)^2 * .2 + (10-3.15)^2 * .1 = 8.1725$. *Thus,* $\sigma = \sqrt{8.1725} = 2.8588$.
So, $E\left((X-\mu)^3\right) = (1-3.5)^3 * .4 + (3-3.5)^3 * .3 + (6-3.5)^3 * .2 + (10-3.5)^3 * .1 = 24.3$.
And finally, $\frac{E\left((X-\mu)^3\right)}{\sigma^3} = \frac{24.3}{2.8588^3} = 1.0401$. *Thus, our pdf exhibits some right skew.*

As you can see from the previous examples, computing moments can involve many steps. Moment generating functions can ease this computational burden. Recall that we've already discussed the expected value of a function, $E(h(x))$. Here our function will be of the form e^{tX}.

Definition 10 *The moment generating function (mgf) of a discrete random variable X is defined to be*

$$\begin{aligned} M_x(t) &= E(e^{tX}) \\ &= \sum_{x \in X} e^{tx} p(x). \end{aligned}$$

We say that $M_x(t)$ exists if it is defined on a symmetric interval $(-t_0, t_0)$. Note that this interval contains 0 in its interior.

Example 11 *Let X be any discrete random variable. What is $M_x(0)$?*

$$\begin{aligned} M_x(0) &= E(e^{0*X}) \\ &= \sum_{x\in X} e^{0*x}p(x) \\ &= \sum_{x\in X} e^{0}p(x) \\ &= \sum_{x\in X} p(x) \\ &= 1. \end{aligned}$$

So $M_x(0)$ is the sum of the probabilities in the discrete random variable.

Example 12 *Consider flipping an unbalanced coin that lands H 60% of the time. What is the moment generating function for this Bernoulli random variable?*

Coin	H	T
Prob	.6	.4

is the pdf. One can view this representing a success with a 1 and a failure as a 0 for the X values.

X	1	0
$p(x)$	.6	.4

$M_x(t) = E(e^{tX}) = \sum_{x\in X} e^{tx}p(x) = e^{0t}p(0) + e^{1t}p(1) = .4 + .6e^t.$

Remark 13 *Note the that the coefficients of this function are probabilities from the Bernoulli random variable.*

Problem 14 *Show that this will always be true for any Bernoulli random variable.*

Consider any two Bernoulli random variables X and Y. If $M_x(t) = M_y(t)$ then X and Y have the same distribution (and hence $X = Y$) since the probabilities for 0 and 1 are identical. This fact is easy to see for Bernoulli random variables. It also turns out to be true for all discrete distributions. In every $M_x(t)$ the coefficients of our terms are the probabilities in the pdf.

Example 15 *Let* $M_x(t) = .3 + .2e^t + .3e^{2t} + .1e^{3t} + .1e^{4t}$ *be the moment generating function for the random discrete variable* X*. Construct the pdf of* X*.*

X	0	1	2	3	4
$p(x)$	.3	.2	.3	.1	.1

Problem 16 *Is* $.2 + .2e^t + .3e^{2t} + .2e^{3t} + .2e^{4t}$ *a moment generating function for some discrete random variable* X*?*

Example 17 *Back to the experiment of rolling a pair of dice and summing the faces. The random variable* X *assigns to each roll its sum. The following table is our previously constructed probability model.*

Sum	2	3	4	5	6	7	8	9	10	11	12
Prob	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Find the moment generating function
$M_x(t) = \frac{e^{2t}}{36} + \frac{2e^{3t}}{36} + \frac{3e^{4t}}{36} + \frac{4e^{5t}}{36} + \frac{5e^{6t}}{36} + \frac{6e^{7t}}{36} + \frac{5e^{8t}}{36} + \frac{4e^{9t}}{36} + \frac{3e^{10t}}{36} + \frac{2e^{11t}}{36} + \frac{e^{12t}}{36}$.

What else can we do with moment generating functions? How do these functions generate moments? Derivatives!

Definition 18 *Let* $M_x^{(r)}(t)$ *be the* r^{th} *derivative of* $M_x(t)$ *with respect to* t*.* $M_x^{(r)}(0)$ *is the* r^{th} *moment about 0.*

Theorem 19 *Expected value of a distribution is* $M_x^1(0)$*.*

Proof. $M_x(t) = \sum_{x \in X} e^{tx} p(x)$ so $M_x^1(t) = \sum_{x \in X} x e^{tx} p(x)$. Evaluated at $t = 0$, $M_x^1(0) = \sum_{x \in X} x e^{tx} p(x) = \sum_{x \in X} x p(x) = E(X)$. ∎

Example 20 *Recall the following Bernoulli random variable*

X	1	0
$p(x)$	.6	.4

and we know $M_x(t) = .4 + .6e^t$*. What is the expected value of this distribution? Now we know it is also* $M_x^1(0)$*.* $M_x^1(t) = .6e^t$ *and* $M_x^1(0) = .6e^0 = .6$

Example 21 *Use the moment generating function to determine expected value when rolling a pair of dice and summing the faces.*
$M_x^1(t) = \frac{2e^{2t}}{36} + \frac{6e^{3t}}{36} + \frac{12e^{4t}}{36} + \frac{20e^{5t}}{36} + \frac{30e^{6t}}{36} + \frac{42e^{7t}}{36} + \frac{40e^{8t}}{36} + \frac{36e^{9t}}{36} + \frac{30e^{10t}}{36} + \frac{22e^{11t}}{36} + \frac{12e^{12t}}{36}$*. So,* $M_x^1(0) = \frac{2}{36} + \frac{6}{36} + \frac{12}{36} + \frac{20}{36} + \frac{30}{36} + \frac{42}{36} + \frac{40}{36} + \frac{36}{36} + \frac{30}{36} + \frac{22}{36} + \frac{12}{36} = 7$.

Example 22 *What is* $M_x^2(0)$*?*

Since $M_x^1(t) = \sum_{x\in X} xe^{tx}p(x)$ then $M_x^2(t) = \sum_{x\in X} x^2e^{tx}p(x)$. Evaluated at $t=0$, $M_x^2(0) = \sum_{x\in X} x^2e^{tx}p(x) = \sum_{x\in X} x^2p(x) = E(X^2)$.
Recall our shortcut formula for $V(X) = E(X^2) - [E(X)]^2$.

Example 23 *Find the variance and standard deviation of the following Bernoulli random variable*

X	1	0
$p(x)$	.6	.4

.

$E(X^2) = M_x^2(0) = .6$ *which is* $M_x^2(t) = .6e^t$ *evaluated at* $t=0$. $[E(X)]^2 = .6^2 = 0.36$. *Thus,* $V(X) = .6 - .36 = 0.24$ *and* $\sigma = \sqrt{.24} = 0.489\,90$

Example 24 *Use moment generating functions to find* $V(X)$ *when when rolling a pair of dice and summing the faces.*
$M_x^2(t) = \frac{4e^{2t}}{36} + \frac{18e^{3t}}{36} + \frac{48e^{4t}}{36} + \frac{100e^{5t}}{36} + \frac{180e^{6t}}{36} + \frac{294e^{7t}}{36} + \frac{320e^{8t}}{36} + \frac{324e^{9t}}{36} + \frac{300e^{10t}}{36} + \frac{242e^{11t}}{36} + \frac{144e^{12t}}{36}$, $M_x^2(0) = \frac{4}{36} + \frac{18}{36} + \frac{48}{36} + \frac{100}{36} + \frac{180}{36} + \frac{294}{36} + \frac{320}{36} + \frac{324}{36} + \frac{300}{36} + \frac{242}{36} + \frac{144}{36} = 54.833$.
Thus, $V(X) = 54.833 - 7^2 = 5.833$ *and* $\sigma = \sqrt{5.833} = 2.415\,2$.

Example 25 *Returning to our unbalanced coin that lands H 60% of the time. Let* X *be the number of flips until a tail appears. Find* $M_x(t)$ *and use it to find* $E(X)$ *and* $V(X)$.
While it is still the case that

$flip$	H	T
$p(x)$	.6	.4

,

X	1	0
$p(x)$	.6	.4

is no longer the pdf. Worse yet, X *is no longer finite. It is theoretically possible for* X *to assume any value in* $\mathbb{Z}^+$.

X	1	2	3	$\cdots$	k
$p(x)$	.4	$.6 * .4 = 0.24$	$.6^2 * .4 = 0.144$	$\cdots$	$.6^{k-1} * .4$

So, $M_x(t) = \sum_{x\in X} e^{tx}p(x) = \sum_{x=1}^{\infty} e^{tx} * .6^{x-1} * .4$ *which is almost a geometric series. We need to perform some manipulations and use a limit to get where we need to be. See Section 3.1-3.3 notes for a refresher on the geometric series.*

Bear in mind that our variable is t.

$$\begin{aligned}
M_x(t) &= \sum_{x=1}^{\infty} e^{tx} * .6^{x-1} * .4 \\
&= .4 \sum_{x=1}^{\infty} e^{tx} * .6^{x-1} \\
&= .4 \sum_{x=1}^{\infty} e^{t} e^{t(x-1)} * .6^{x-1} \\
&= .4e^{t} \sum_{x=1}^{\infty} e^{t(x-1)} * .6^{x-1} \\
&= .4e^{t} \sum_{x=1}^{\infty} (.6 * e^{t})^{x-1} \\
&= .4e^{t} \lim_{n\to\infty} \sum_{x=1}^{n} (.6 * e^{t})^{x-1} \\
&= .4e^{t} \lim_{n\to\infty} \frac{1-(.6*e^{t})^{n}}{1-.6*e^{t}} \\
&= .4e^{t} * \frac{1}{1-.6*e^{t}} \\
&= \frac{.4e^{t}}{1-.6*e^{t}}.
\end{aligned}$$

Remark 26 *Note that any moment generating function of the form* $M_x(t) = \sum_{x\in X} e^{tx}p(x) = \sum_{x=1}^{\infty} e^{tx} * .a^{x-1} * .b = \frac{be^{t}}{1-ae^{t}}$*. Since* $a = 1 - b$ *you may see this written as* $\frac{be^{t}}{1-(1-b)e^{t}}$*. In general, for* $M_x(t) = \frac{be^{t}}{1-ae^{t}}$*,*

$$\begin{aligned}
M_x^1(t) &= \frac{be^{t}(1-ae^{t}) - be(-ae^{t})}{(1-ae^{t})^2} \\
&= \frac{be^{t}\left(1-ae^{t}+ae^{t}\right)}{(1-ae^{t})^2} \\
&= \frac{be^{t}}{(1-ae^{t})^2}.
\end{aligned}$$

and

$$\begin{aligned}
M_x^2(t) &= \frac{be^t\left(1-ae^t\right)^2 - be^t(2)\left(1-ae^t\right)\left(-ae^t\right)}{\left(\left(1-ae^t\right)^2\right)^2} \\
&= \frac{be^t[\left(1-ae^t\right)^2 - 2\left(1-ae^t\right)\left(-ae^t\right)]}{\left(1-ae^t\right)^4} \\
&= \frac{be^t[1-2ae^t+a^2e^{2t}-2(-ae^t+a^2e^{2t})]}{\left(1-ae^t\right)^4} \\
&= \frac{be^t[1-2ae^t+a^2e^{2t}+2ae^t-2a^2e^{2t}]}{\left(1-ae^t\right)^4} \\
&= \frac{be^t[1+a^2e^{2t}-2a^2e^{2t}]}{\left(1-ae^t\right)^4} \\
&= \frac{be^t[1-a^2e^{2t}]}{\left(1-ae^t\right)^4} \\
&= \frac{be^t(1-ae^t)(1+ae^t)}{\left(1-ae^t\right)^4} \\
&= \frac{be^t(1+ae^t)}{\left(1-ae^t\right)^3}.
\end{aligned}$$

With $M_x(t) = \dfrac{.4e^t}{1-.6*e^t}$, we know $M_x^1(t) = \dfrac{.4e^t}{\left(1-.6e^t\right)^2}$ and thus $M_x^1(0) = \dfrac{.4e^0}{\left(1-.6*e^0\right)^2} = 2.5$. We expect to flip our unbalanced coin 2.5 times before seeing a tail. With $M_x^2(t) = \dfrac{.4e^t(1+.6e^t)}{\left(1-.6e^t\right)^3}$ and now $M_x^2(0) = \dfrac{.4e^0(1+.6*e^0)}{\left(1-.6*e^0\right)^3} = 10.0$. Thus, $V(X) = M_x^2(0) - (M_x^1(0))^2 = 10 - (2.5)^2 = 3.75$ which yields a standard deviation of $\sigma = \sqrt{3.75} = 1.9365$.

We've done this for general a and b. Frequently this geometric series sets $a = 1-b$ as probabilities. Consider p to be the probability of success. What is the expected number of attempts before the first success?

$$\begin{aligned}
M_x^1(0) &= \frac{pe^0}{\left(1-(1-p)e^0\right)^2} \\
&= \frac{p}{\left(1-(1-p)\right)^2} \\
&= \frac{p}{p^2} \\
&= \frac{1}{p}
\end{aligned}$$

What about standard deviation?

$$\begin{aligned} M_x^2(0) &= \frac{pe^0(1+(1-p)e^0)}{(1-(1-p)e^0)^3} \\ &= \frac{p(1+(1-p))}{(1-(1-p))^3} \\ &= \frac{p(1+(1-p))}{p^3} \\ &= \frac{(1+(1-p))}{p^2} \end{aligned}$$

so

$$\begin{aligned} V(X) &= \frac{(1+(1-p))}{p^2} - \left(\frac{1}{p}\right)^2 \\ &= \frac{(1+(1-p))-1}{p^2} \\ &= \frac{1-p}{p^2} \end{aligned}$$

and $\sigma = \frac{\sqrt{1-p}}{p}$.

We've done the hard, general form work so now we can plug and chug using the formulae. All we have to do is recognize that the moment generating function is a geometric series.

Example 27 *Roll a pair of fair dice until you get a sum of 12. Find the mean and standard deviation for the number of rolls until one obtains the sum of 12.*

sum	12	$\neq 12$
$p(x)$	$\frac{1}{36}$	$\frac{35}{36}$

X	1	2	3	$\cdots$	k
$p(x)$	$\frac{1}{36}$	$\frac{35}{36} * \frac{1}{36} = \frac{35}{1296}$	$\left(\frac{35}{36}\right)^2 * \frac{1}{36} = \frac{1225}{46\,656}$	$\cdots$	$\left(\frac{35}{36}\right)^{k-1} * \frac{1}{36}$

$M_x(t) = \sum_{x \in X} e^{tx} p(x) = \sum_{x=1}^{\infty} e^{tx} * \left(\frac{35}{36}\right)^{x-1} * \frac{1}{36}$.

$M_x^1(t) = \frac{\frac{1}{36}e^t}{\left(1-\frac{35}{36}e^t\right)^2}$ *and thus* $M_x^1(0) = \frac{\frac{1}{36}e^0}{\left(1-\frac{35}{36}e^0\right)^2} = 36.0.$

$M_x^2(t) = \frac{\frac{1}{36}e^t(1+\frac{35}{36}e^t)}{\left(1-\frac{35}{36}e^t\right)^3}$ *and thus* $M_x^2(0) = \frac{\frac{1}{36}e^0(1+\frac{35}{36}e^0)}{\left(1-\frac{35}{36}e^0\right)^3} = 2556.0$ *so* $V(X) = 2556 - 36^2 = 1260$ *and* $\sigma = \sqrt{1260} = 35.496$

We can also note that the probability of success is $p = \frac{1}{36}$ *so* $\mu = \frac{1}{\frac{1}{36}} = 36$ *and* $\sigma = \frac{\sqrt{1-\frac{1}{36}}}{\frac{1}{36}} = 35.496$.