# STA 623 – Fall 2011 – Dr. Charnigo

## Section 2.3: Moments and moment generating functions

*Moments.* Let $X$ be a random variable. For any integer $n \geq 1$, we define the $n^{th}$ moment of $X$ to be $E[X^n]$. For any integer $n \geq 2$, we define the $n^{th}$ central moment of $X$ to be $E[(X-\nu)^n]$, where $\nu := E[X]$ is assumed to exist as a finite number. If the $n^{th}$ moment exists as a finite number, then so do all moments of lower order. Hence, if the $n^{th}$ moment exists as a finite number, then so does the $n^{th}$ central moment.

**Example (moments)**. Suppose $X$ has probability density function

$$f_X(x) := (2\pi)^{-1/2}\sigma^{-1}\exp[-(x-\mu)^2/(2\sigma^2)],$$

where $\mu \in (-\infty, \infty)$ and $\sigma \in (0, \infty)$. Let $g(x) := (x-\mu)/\sigma$ and put $Z := g(X)$. Then for any real $z$ we have

$$P(Z \leq z) = P(X \leq \sigma z + \mu) = \int_{-\infty}^{\sigma z+\mu} (2\pi)^{-1/2}\sigma^{-1}\exp[-(x-\mu)^2/(2\sigma^2)]\ dx.$$

Making the substitutions $y := (x-\mu)/\sigma$ and $dy := dx/\sigma$, we can express the integral as

$$\int_{-\infty}^{z} (2\pi)^{-1/2}\exp[-y^2/2]\ dy.$$

Upon differentiating with respect to $z$, we obtain a probability density function for $Z$ of

$$f_Z(z) := (2\pi)^{-1/2}\exp[-z^2/2].$$

We have

$$E[Z^2] = (2\pi)^{-1/2}\int_{\mathbb{R}} z^2\exp[-z^2/2]\ dz = (2\pi)^{-1/2}\int_{\mathbb{R}} z\exp[-z^2/2]z\ dz.$$

Integrating by parts with $u := z$, $du := dz$, $dv := \exp[-z^2/2]z\ dz$, and $v := -\exp[-z^2/2]$, we obtain

$E[Z^2] =$

Likewise, for any positive integer $p$, we have

$$E[Z^{2p}] = (2\pi)^{-1/2} \int_{\mathbb{R}} z^{2p} \exp[-z^2/2]\, dz = (2\pi)^{-1/2} \int_{\mathbb{R}} z^{2p-1} \exp[-z^2/2] z\, dz.$$

Integrating by parts with $u := z^{2p-1}$, $du := (2p-1)z^{2p-2}\, dz$, $dv := \exp[-z^2/2]z\, dz$, and $v := -\exp[-z^2/2]$, we obtain

$$E[Z^{2p}] = (2\pi)^{-1/2}(2p-1) \int_{\mathbb{R}} z^{2p-2} \exp[-z^2/2]\, dz = (2p-1)E[Z^{2p-2}].$$

Thus, by mathematical induction,

$$E[Z^{2p}] = (2p-1) \times (2p-3) \times \cdots \times 3 \times 1,$$

which some mathematicians denote $(2p-1)!!$, read "double factorial".

Note that $E[Z^{2p-1}]$ must exist as a finite number (because $E[Z^{2p}]$ does) and must equal zero (because $E[Z^{2p-1}]$ is the integral of an odd function over $\mathbb{R}$).

Now we are in a good position to find some moments of $X = \sigma Z + \mu$.

We have

$E[X] =$

$E[X^2] =$

$E[(X-\nu)^2] =$

$E[X^3] = E[(\sigma Z+\mu)^3] = \sigma^3 E[Z^3] + 3\sigma^2\mu E[Z^2] + 3\sigma\mu^2 E[Z] + \mu^3 = 3\sigma^2\mu + \mu^3,$

and

$$E[X^4] = E[(\sigma Z+\mu)^4] = \sigma^4 E[Z^4] + 6\sigma^2\mu^2 E[Z^2] + \mu^4 = 3\sigma^4 + 6\sigma^2\mu^2 + \mu^4.$$

We conclude this set of examples with a cautionary note. Suppose $W$ has probability density function $f_W(w) := \pi^{-1}/(1+w^2)$. Then

$$E[W^2] = \pi^{-1} \int_{\mathbb{R}} w^2/(1+w^2)\, dw \geq \pi^{-1} \int_1^\infty w^2/(1+w^2)\, dw \geq \pi^{-1} \int_1^\infty 1/2\, dw = \infty.$$

On the other hand,

$$E[W^3] = \pi^{-1} \int_{\mathbb{R}} w^3/(1+w^2)\ dw$$

appears to be 0 because the integrand is an odd function. If that is true, then we are in trouble because we have said that the existence of $E[W^3]$ as a finite number should imply the same for $E[W^2]$. How can this apparent contradiction be resolved?

*Variance and standard deviation.* The second central moment $E[(X-\nu)^2]$ is called the variance of $X$. The variance describes how much $X$ fluctuates around its expected value. The standard deviation of $X$ is defined to be the positive square root of the variance. Unlike the variance, the standard deviation is expressed in the same units as $X$. For instance, if $X$ represents systolic blood pressure in mmHg, then the standard deviation of $X$ is expressed in mmHg while the variance is expressed in $(\text{mmHg})^2$.

Three useful results, assuming all expectations and variances referred to exist as finite numbers, are as follows.

1. For any constants $a$ and $b$, $Var[aX+b] = a^2 Var[X]$.
2. A computational formula for the variance is $E[X^2] - (E[X])^2$.
3. If $Var[X] = 0$, then $P(X = E[X]) = P(|X - E[X]| = 0) = 1$.

*A useful result on expectation of indicators, for our next example.* For any random variable $X$ and any set $A \in \mathcal{B}^1$ we have $E[1_{\{X \in A\}}] = P(X \in A)$. To see this, suppose for concreteness that $X$ is continuous with probability density function $f_X(x)$. Then we have

$$E[1_{\{X \in A\}}] = \int_{\mathbb{R}} 1_{\{x \in A\}} f_X(x)\ dx = \int_A f_X(x)\ dx = P(X \in A).$$

**Example (variance and standard deviation)**. Your textbook authors prove the first two useful results above by appealing to linearity of expectation. We can prove the third by contradiction. Indeed, suppose that $E[(X - E[X])^2] = 0$ but that $P(|X - E[X]| > 0) = \epsilon > 0$. Since

$$\{|X - E[X]| > 0\} = \cup_{j=1}^{\infty}\{1/j \leq |X - E[X]| < 1/(j-1)\}$$

we have

$$0 < \epsilon = \sum_{j=1}^{\infty} P(1/j \leq |X - E[X]| < 1/(j-1)).$$

By countable additivity, there must exist $j \in \{1, 2, \ldots\}$ such that

$$0 < \delta = P(1/j \leq |X - E[X]| < 1/(j-1)) \leq P(1/j \leq |X - E[X]|).$$

Then, using monotonicity of expectation (twice) and the useful result on expectation of indicators, we have

$$\begin{aligned} E[(X - E[X])^2] &\geq E[(X - E[X])^2 1_{\{|X - E[X]| \geq 1/j\}}] \\ &\geq E[(1/j)^2 1_{\{|X - E[X]| \geq 1/j\}}] \\ &\geq (1/j)^2 P(|X - E[X]| \geq 1/j) \\ &\geq \\ &> 0. \end{aligned}$$

We have arrived at a contradiction. Therefore, we must conclude that $E[(X - E[X])^2] = 0$ implies $P(|X - E[X]| > 0) = 0$.

*Moment generating function.* The moment generating function of $X$ is defined as $M_X(t) := E[\exp(tX)]$. The moment generating function is potentially useful for three reasons.

1. Suppose there exists $h > 0$ such that $M_X(t) < \infty$ for all $t \in [-h, h]$. Then, for every positive integer $n$, $E[X^n]$ exists as a finite number and is equal to $\frac{d^n}{dt^n} M_X(t)|_{t=0}$.

2. Suppose there exists $h > 0$ such that $M_X(t), M_Y(t) < \infty$ for all $t \in [-h, h]$. If $M_X(t) = M_Y(t)$ for all $t \in [-h, h]$, then $X$ and $Y$ have the same cumulative distribution function: $F_X(u) = F_Y(u)$ for any real $u$.

3. Suppose there exists $h > 0$ such that $M_X(t), M_{X_1}(t), M_{X_2}(t), \ldots < \infty$ for all $t \in [-h, h]$. If $M_{X_i}(t) \stackrel{i\to\infty}{\to} M_X(t)$ for all $t \in [-h, h]$, then the cumulative distribution functions of $X_1, X_2, \ldots$ converge to the cumulative distribution function of $X$ at all points where the latter is continuous: $F_{X_i}(u) \stackrel{i\to\infty}{\to} F_X(u)$ for any real $u$ at which $F_X(u)$ is continuous.

**Example (moment generating function)**. To see why the first result above holds, suppose for concreteness that $X$ is continuous with probability density function $f_X(x)$. For $t \in (-h, h)$ we have

$$\begin{aligned}
\frac{d^n}{dt^n} M_X(t) &= \frac{d^n}{dt^n} \int_{\mathbb{R}} \exp[tx] f_X(x)\ dx \\
&= \int_{\mathbb{R}} \frac{\partial^n}{\partial t^n} \exp[tx] f_X(x)\ dx \\
&= \int_{\mathbb{R}} x^n \exp[tx] f_X(x)\ dx.
\end{aligned}$$

(The second equality above, interchange of differentiation and integration, will be justified next week.) Hence,

$$\frac{d^n}{dt^n} M_X(t)|_{t=0} = \int_{\mathbb{R}} x^n f_X(x)\ dx = E[X^n].$$

Suppose $X$ has probability density function

$$f_X(x) := (2\pi)^{-1/2} \sigma^{-1} \exp[-(x-\mu)^2/(2\sigma^2)],$$

where $\mu \in (-\infty, \infty)$ and $\sigma \in (0, \infty)$. Put $Z := (X - \mu)/\sigma$. Then $Z$ has probability density function

$$f_Z(z) := (2\pi)^{-1/2} \exp[-z^2/2],$$

from which we can calculate the moment generating function of $Z$:

$$\begin{aligned}
M_Z(t) &= (2\pi)^{-1/2} \int_{\mathbb{R}} \exp[tz] \exp[-z^2/2] \; dz \\
&= (2\pi)^{-1/2} \int_{\mathbb{R}} \exp[(-z^2 + 2tz)/2] \; dz \\
&= (2\pi)^{-1/2} \int_{\mathbb{R}} \exp[(-z^2 + 2tz - t^2)/2] \exp[t^2/2] \; dz \\
&= (2\pi)^{-1/2} \exp[t^2/2] \int_{\mathbb{R}} \exp[-(z-t)^2/2] \; dz \\
&=
\end{aligned}$$

We can then calculate the moment generating function of $X = \sigma Z + \mu$:

$$\begin{aligned}
M_X(t) &= E[\exp(tX)] = E[\exp(t\sigma Z + t\mu)] = \exp(t\mu) E[\exp(t\sigma Z)] \\
&= \exp(t\mu) M_Z(t\sigma) =
\end{aligned}$$

To see an application of the third result above, suppose $X_i$ has probability density function

$$\frac{1}{\Gamma[\alpha_i]\beta_i^{\alpha_i}} x^{\alpha_i - 1} \exp[-x/\beta_i] 1_{\{x>0\}}$$

for $i \in \{1, 2, \ldots\}$, where $\alpha_i := i$ and $\beta_i := 1/i$. One can show that

$$M_{X_i}(t) = \left[\frac{1}{1 - \beta_i t}\right]^{\alpha_i} = \left[\frac{1}{1 - t/i}\right]^i$$

for $t \in (-i, i)$. Taking this for granted, we have

$$M_{X_i}(t) = \left[1 + \frac{t/i}{1 - t/i}\right]^i \stackrel{i \to \infty}{\to} \exp[t].$$

On the other hand, $\exp[t]$ is clearly $M_X(t)$, where $X$ is a random variable that equals 1 with probability 1. As such, we conclude that for any $u \neq 1$

$$P(X_i \leq u) \stackrel{i \to \infty}{\to} P(X \leq u).$$

In particular,

$$\begin{aligned}
P(X_i \leq u) &\stackrel{i \to \infty}{\to} 0 \;\; \text{for} \;\; u < 1 \;\; \text{and} \\
P(X_i \leq u) &\stackrel{i \to \infty}{\to} 1 \;\; \text{for} \;\; u > 1.
\end{aligned}$$