EE5110: Probability Foundations for Electrical Engineers **July-November 2015**

Lecture 25: Moment Generating Function

Lecturer: Dr. Krishna Jagannathan *Scribe: Subrahmanya Swamy P*

In this lecture, we will introduce Moment Generating Function and discuss its properties.

Definition 25.1 *The moment generating function (MGF) associated with a random variable X, is a function, $M_X : \mathbb{R} \to [0, \infty]$ defined by $M_X(s) = \mathbb{E}\left[e^{sX}\right]$.*

The domain or region of convergence (ROC) of M_X is the set $D_X = \{s | M_X(s) < \infty\}$. In general, s can be complex, but since we did not define expectation of complex valued random variables, we will restrict ourselves to real valued s. Note that $s = 0$ is always a point in the ROC for any random variable, since $M_X(0) = 1$.

Cases:

- If X is discrete with pmf $p_X(x)$, then $M_X(s) = \sum_x e^{sx} p_X(x)$.
- If X is continuous with density $f_X(\cdot)$, then $M_X(s) = \int e^{sx} f_X(x)\, dx$.

Example 25.2 *Exponential random variable*

$$f_X(x) = \mu e^{-\mu x}, \qquad x \geq 0,$$

$$M_X(s) = \int_0^\infty e^{sx} \mu e^{-\mu x} dx = \begin{cases} \frac{\mu}{\mu - s}, & \text{if } s < \mu, \\ +\infty, & \text{otherwise.} \end{cases}$$

The Region of Convergence for this example is, $\{s | M_X(s) < \infty\}$, i.e., $s < \mu$.

Example 25.3 *Std. Normal random variable*

$$\begin{aligned} f_X(x) &= \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}, \quad x \in \mathbb{R}, \\ M_X(s) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{sx} e^{\frac{-x^2}{2}} dx, \\ &= e^{\frac{s^2}{2}}, \quad s \in \mathbb{R}. \end{aligned}$$

The Region of Convergence for this example is the entire real line.

Example 25.4 *Cauchy random variable*

$$f_X(x) = \frac{1}{\pi(1 + x^2)}, \quad x \in \mathbb{R},$$

$$M_X(s) = \frac{1}{\pi} \int_{-\infty}^{\infty} e^{sx} \frac{1}{1 + x^2} dx = \begin{cases} 1, & \text{if } s = 0, \\ +\infty, & \text{otherwise.} \end{cases}$$

The Region of Convergence for this example is just the point $s = 0$.

Remark 2: The above examples can be interpreted as follows.

- In Example 25.2, we have the product of two exponentials. Thus, the MGF converges when the product is decreasing.
- In Example 25.3, there is a 'competition' between $e^{-\frac{x^2}{2}}$ and e^{sx}. Since the first term from the Gaussian decreases faster than e^{sx} increases (for any s), the integral always converges.
- In Example 25.4, for $s \neq 0$, an exponential competes with a decreasing polynomial, as a result of which the integral diverges.

It is an interesting question whether or not we can uniquely find the CDF of a random variable, given the moment generating function and its ROC. A quick look at Example 25.4 reveals that if the MGF is finite only at $s = 0$ and infinite elsewhere, it is not possible to recover the CDF uniquely. To see this, one just needs to produce another random variable whose MGF is finite only at $s = 0$. (Do this!) On the other hand, if we can specify the value of the moment generating function even in a tiny interval, we can uniquely determine the density function. This result follows essentially because the MGF, when it exists in an interval, is *analytic*, and hence possesses some nice properties. The proof of the following theorem is rather involved, and uses the properties of an analytic function.

Theorem 25.5 *(Without Proof)*

i) Suppose $M_X(s)$ is finite in the interval $[-\epsilon, \epsilon]$ for some $\epsilon > 0$, then M_X uniquely determines the CDF of X.

ii) If X and Y are two random variables such that, $M_X(s) = M_Y(s) \;\; \forall s \in [-\epsilon, \epsilon], \epsilon > 0$ then X and Y have the same CDF.

25.1 Properties

1. $M_X(0) = 1$.
2. *Moment Generating Property:* We shall state this property in the form of a theorem.

   Theorem 25.6 *Supposing $M_X(s) < \infty$ for $s \in [-\epsilon, \epsilon]$, $\epsilon > 0$ then,*

   $$\frac{d}{ds} M_X(s)\Big|_{s=0} = \mathbb{E}[X]. \tag{25.1}$$

   More generally,

   $$\frac{d^m}{ds^m} M_X(s)\Big|_{s=0} = \mathbb{E}[X^m] \; ; \;\; m \geq 1.$$

   Proof: (25.1) can be proved in the following steps.

   $$\frac{d}{ds} M_X(s) = \frac{d}{ds}\mathbb{E}[e^{sX}] \overset{(a)}{=} \mathbb{E}[\frac{d}{ds} e^{sX}] = \mathbb{E}[X e^{sX}],$$

   where, (a) is obtained by the interchange of the derivative and the expectation. This follows from the use of basic definition of the derivative, and then invoking the DCT; see Lemma 25.7 (d). ■

Lemma 25.7 *Suppose that X is a non-negative random variable and $M_X(s) < \infty$, $\forall s \in (-\infty, a]$, where a is a positive number, then*

(a) $\mathbb{E}[X^k] < \infty$, *for every* k.

(b) $\mathbb{E}[X^k e^{sX}] < \infty$, *for every* $s < a$.

(c) $\frac{e^{hX}-1}{h} \leq Xe^{hX}$.

(d) $\mathbb{E}[X] = \mathbb{E}[\lim_{h\downarrow 0} \frac{e^{hX}-1}{h}] = \lim_{h\downarrow 0} \frac{\mathbb{E}[e^{hX}]-1}{h}$.

Proof: Given that X is a non-negative random variable with a Moment Generating Function such that $M_X(s) < \infty$, $\forall s \in (-\infty, a]$, for some positive a.

(a) For a positive number a, $x^k \leq e^{ax}$, $\forall k \in \mathbb{Z}^+ \cup \{0\}$. Therefore, $\mathbb{E}[X^k] = \int x^k \mathrm{d}\mathbb{P}_X \leq \int e^{ax}\mathrm{d}\mathbb{P}_X$. However, $\int e^{ax}\mathrm{d}\mathbb{P}_X = M_X(a) < \infty$. Therefore, $\mathbb{E}[X^k] < \infty$.

(b) For s $< a$, $\exists \epsilon > 0$ such that $M_X(s+\epsilon) < \infty \Rightarrow \int e^{sx}e^{\epsilon x}\mathrm{d}\mathbb{P}_X < \infty$. But since $\epsilon > 0$, as $x \to \infty$, $x^k \leq e^{\epsilon x}$. Therefore, $\mathbb{E}[X^k e^{sX}] = \int x^k e^{sx}\mathrm{d}\mathbb{P}_X \leq \int e^{sx}e^{\epsilon x}\mathrm{d}\mathbb{P}_X < \infty \Rightarrow \mathbb{E}[X^k e^{sX}] < \infty$.

(c) To prove that $\frac{e^{hX}-1}{h} \leq Xe^{hX}$.
Let $hX = Y$. Therefore, re-arranging the terms, we need to prove that $e^Y - Ye^Y \leq 1$. Or equivalently, it is enough to prove that, $g(Y) = e^Y(Y-1) \geq -1$.
$g(Y)$ has a minima at $Y = 0$, and the minimum value, *i.e.,* $g(0) =$ -1.
$\Rightarrow g(Y) \geq -1$,
$\Rightarrow e^Y(Y-1) \geq -1$.
Hence proved.

(d) Define $X_h = \frac{e^{hX}-1}{h}$.
$\lim_{h\downarrow 0} X_h = X$ i.e. $X_h \to X$ point-wise. Since $\mathbb{E}[X^k e^{sX}] < \infty$ is true, when $s = h$ and $k = 1$, we get $\mathbb{E}[Xe^{hX}] < \infty$. Since X_h is dominated by Xe^{hX}, $\mathbb{E}[Xe^{hX}] < \infty$ and $\lim_{h\downarrow 0} X_h = X$, applying DCT we get $\mathbb{E}[X] = \mathbb{E}[\lim_{h\downarrow 0} X_h] = \mathbb{E}[\lim_{h\downarrow 0} \frac{e^{hX}-1}{h}] = \lim_{h\downarrow 0} \mathbb{E}\left[\frac{e^{hX}-1}{h}\right] = \lim_{h\downarrow 0} \frac{\mathbb{E}[e^{hX}]-1}{h}$. Therefore, $\mathbb{E}[X] = \mathbb{E}[\lim_{h\downarrow 0} \frac{e^{hX}-1}{h}] = \lim_{h\downarrow 0} \frac{\mathbb{E}[e^{hX}]-1}{h}$.
Hence proved.

■

3. If $Y = aX + b$, $a, b \in \mathbb{R}$, then $M_Y(s) = e^{sb}M_X(as)$. For example, $X \sim \mathcal{N}(0,1)$, $Y = \sigma X + \mu$ $\Rightarrow Y \sim \mathcal{N}(\mu, \sigma^2) \Rightarrow M_Y(s) = e^{\mu s}e^{\sigma^2 \frac{s^2}{2}}$, $s \in \mathbb{R}$.

4. If X and Y are independent and $Z = X + Y$, then $M_Z(s) = M_X(s)M_Y(s)$.
Proof: $\mathbb{E}[e^{sZ}] = \mathbb{E}[e^{sX+sY}] = \mathbb{E}[e^{sX}e^{sY}]$=$\mathbb{E}[e^{sX}]\mathbb{E}[e^{sY}]$. ■

Consider the following examples:

(a) $X_1 \sim N(\mu_1, \sigma_1^2)$; $X_2 \sim N(\mu_2, \sigma_2^2)$; and X_1, X_2 are independent. $Z = X_1 + X_2$;

$$
\begin{aligned}
M_{X_1}(s) &= e^{\left(\mu_1 s + \frac{\sigma_1^2 s^2}{2}\right)}, \\
M_{X_2}(s) &= e^{\left(\mu_2 s + \frac{\sigma_2^2 s^2}{2}\right)}, \\
M_Z(s) &= M_{X_1}(s)M_{X_2}(s), \\
&= e^{\left((\mu_1+\mu_2)s + \frac{(\sigma_1^2+\sigma_2^2)s^2}{2}\right)}.
\end{aligned}
$$

$$\Rightarrow Z \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2).$$

(b) $X_1 \sim \exp(\mu); X_2 \sim \exp(\lambda)$, $\lambda \neq \mu$ and X_1, X_2 are independent. $Z = X1 + X2$;

$$\begin{aligned}
M_{X_1}(s) &= \frac{\mu}{\mu - s}, \\
M_{X_2}(s) &= \frac{\lambda}{\lambda - s}, \\
M_Z(s) &= M_{X_1}(s) M_{X_2}(s), \\
&= \frac{\mu\lambda}{(\mu - s)(\lambda - s)}, \quad \text{ROC is } \; s < \min(\lambda, \mu) \\
\Rightarrow f_Z(x) &= \frac{\mu}{\mu - \lambda}\lambda e^{-\lambda x} - \frac{\lambda}{\mu - \lambda}\mu e^{-\mu x}, \\
&= \left(\frac{\mu\lambda}{\mu - \lambda}\right)\left(e^{-\lambda x} - e^{-\mu x}\right), \quad x \geq 0.
\end{aligned}$$

5. $Z = \sum_{i=1}^{N} X_i$, X_i are i.i.d and N is independent of X_i.

$$\begin{aligned}
M_Z(s) = \mathbb{E}[e^{sZ}] &= \mathbb{E}\left[\mathbb{E}\left[e^{sZ}|N\right]\right], \\
&= \mathbb{E}\left[(M_X(s))^N\right],
\end{aligned}$$

If we write in terms of the PGF and MGF of N, then,

$$\begin{aligned}
M_Z(s) &= G_N(M_X(s)), \\
&= M_N(\log M_X(s)).
\end{aligned}$$

For example, $X_i \sim \exp(\mu); N \sim \text{Geom}(p)$ and $Z = \sum_{i=1}^{N} X_i$. Then the distribution of Z is computed as follows:

$$\begin{aligned}
M_X(s) &= \frac{\mu}{\mu - s}, \quad s < \mu, \\
G_N(\xi) &= \frac{p\xi}{1 - (1-p)\xi}, \quad |\xi| < \frac{1}{1-p}, \\
M_Z(s) &= G_N(M_X(s)), \\
&= \frac{p\left(\frac{\mu}{\mu - s}\right)}{1 - (1-p)\left(\frac{\mu}{\mu - s}\right)}, \\
&= \frac{\mu p}{\mu p - s}, \quad s < \mu p, \\
\Rightarrow Z &\sim \exp(\mu p).
\end{aligned}$$

25.2 Exercise

1. (a) *[Dimitri P.Bertsekas]* Find the MGF associated with an integer-valued random variable X that is uniformly distributed in the range $\{a, a+1, ..., b\}$.

(b) *[Dimitri P.Bertsekas]* Find the MGF associated with a continuous random variable X that is uniformly distributed in the range $[a, b]$.

2. *[Dimitri P.Bertsekas]* A non-negative interger-valued random variable X has one of the following MGF:

   (a) $M(s) = e^{2(e^{e^s - 1} - 1)}$.

   (b) $M(s) = e^{2(e^{e^s} - 1)}$.

   (a) Explain why one of the 2 cannot possibly be a MGF.

   (b) Use the true MGF to find $\mathbb{P}(X = 0)$.

3. Find the variance of a random variable X whose moment generating function is given by

$$M_X(s) = e^{3e^s - 3}$$