

# Math 525: Lecture 12

February 13, 2018

## 1 Moment generating functions (a rigorous treatment)

Remember the moment generating function of a random variable  $X$ ? It looked like

$$M(\theta) = \mathbb{E} [e^{\theta X}].$$

A few lectures ago, we worked on the theory of limits of expectations (Monotone Convergence Theorem, Fatou's Lemma, Dominated Convergence Theorem) with the motivation of being able to rigorously treat moment generating functions. Let's do that now.

**Proposition 1.1.** *Suppose  $M(\theta)$  is well-defined in a neighbourhood of the origin. Then,*

$$M^{(k)}(\theta) = \mathbb{E}[X^k e^{\theta X}]$$

*in a neighbourhood of the origin.*

*Proof.* We start with the case of  $k = 1$ . By definition,

$$\begin{aligned} M'(\theta) &= \lim_{h \downarrow 0} \frac{M(\theta + h) - M(\theta)}{h} \\ &= \lim_{h \downarrow 0} \frac{\mathbb{E} [e^{(\theta+h)X}] - \mathbb{E} [e^{\theta X}]}{h} \\ &= \lim_{h \downarrow 0} \mathbb{E} \left[ \frac{e^{(\theta+h)X} - e^{\theta X}}{h} \right]. \end{aligned}$$

We would like to move the limit inside the expectation to get

$$\mathbb{E} \left[ \lim_{h \downarrow 0} \frac{e^{(\theta+h)X} - e^{\theta X}}{h} \right] = \mathbb{E} \left[ \frac{d}{d\theta} e^{\theta X} \right] = \mathbb{E} [X e^{\theta X}].$$

In order to show that this is possible, we will apply the dominated convergence theorem.

Choose  $b > 0$  small enough so that  $M(\pm b)$  is well-defined. Now, pick  $\epsilon$  such that  $0 < \epsilon < b$ . Assume  $\theta$  and  $h$  are small enough so that both  $\theta$  and  $\theta + h$  are in the interval  $(-b + \epsilon, b - \epsilon)$ . For each  $x$ , the mean value theorem tells us that we can find  $\alpha(x) \in (\theta, \theta + h)$  such that

$$\frac{e^{(\theta+h)x} - e^{\theta x}}{h} = x e^{\alpha(x)x}.$$

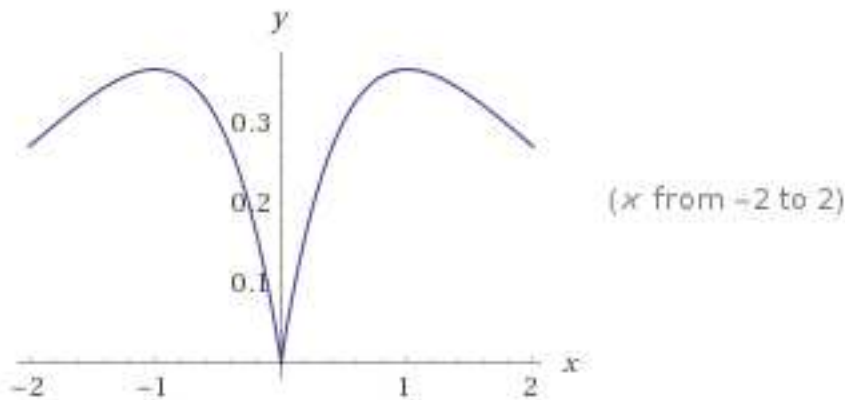


Figure 1:  $x \mapsto |x|e^{-|x|}$

Note that

$$|xe^{\alpha(x)x}| \leq |x|e^{|\alpha(x)x|} = |x|e^{(b-\epsilon)|x|} \leq |x|e^{-\epsilon|x|}e^{b|x|}.$$

The function  $x \mapsto |x|e^{-\epsilon|x|}$  is bounded on the real line (see Figure 1). Therefore, we can find some constant  $C(\epsilon)$  such that

$$|xe^{\alpha(x)x}| \leq C(\epsilon)e^{b|x|} \leq C(\epsilon)(e^{bx} + e^{-bx}).$$

Define the random variable  $Y$  by

$$Y = C(\epsilon)(e^{bX} + e^{-bX}).$$

The inequalities above show that  $Y$  dominates  $|X^{\alpha(X)X}|$ . Moreover,  $Y$  is integrable since

$$\mathbb{E}Y = b(M(b) + M(-b)).$$

The rest follows by the Dominated Convergence Theorem.

As for the case of  $k > 1$ , we simply apply the argument inductively, noting that

$$\frac{d}{d\theta} [x^{k-1}e^{\theta x}] = x^k e^{\theta x}. \quad \square$$

## 2 Characteristic functions

Moment generating functions, while very useful, have one big flaw: their usefulness depends on whether or not they are well-defined on an interval about the origin. Let's visit an example where the moment generating function fails this condition:

**Example 2.1.** The Cauchy distribution (centred at zero) has probability density function

$$f(x) = \frac{1}{\pi\gamma} \left( \frac{\gamma^2}{(x - x_0)^2 + \gamma^2} \right)$$

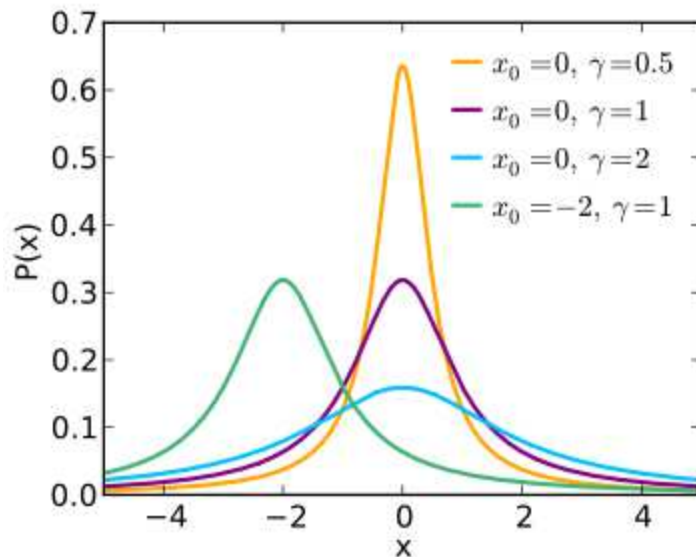


Figure 2: Cauchy probability density

where  $x_0$  specifies the peak of the density and  $\gamma > 0$  is a scaling parameter. Let's take  $x_0 = 0$  and  $\gamma = 1$  for simplicity. Let  $X$  be a random variable with probability density  $f$ . In this case,

$$M(\theta) = \mathbb{E} [e^{\theta X}] = \int_{-\infty}^{\infty} e^{\theta x} f(x) dx = \frac{1}{\pi} \int_{-\infty}^{\infty} e^{\theta x} \frac{1}{x^2 + 1} dx.$$

If  $\theta = 0$ , the above integral becomes  $\frac{1}{\pi} \int_{-\infty}^{\infty} \frac{1}{x^2+1} dx = 1$ . Otherwise, the term  $e^{\theta|x|} \frac{1}{x^2+1}$  grows quickly enough as  $|x| \rightarrow \infty$  for the integral to diverge.

We want a moment generating function that's valid for all distributions. This is our motivation today. First, we need the notion of a complex-valued random variable:

**Definition 2.2.** Let  $U$  and  $V$  be real-valued random variables. Then, we call  $Z = U + iV$  a *complex-valued random variable*.

We define  $\mathbb{E}Z = \mathbb{E}U + i\mathbb{E}V$  whenever  $U$  and  $V$  are integrable. We point out that  $|\mathbb{E}Z| \leq \mathbb{E}|Z|$  (in this case,  $|\cdot|$  is the modulus of a complex number).

**Definition 2.3.** The *characteristic function* of a random variable  $X$  is given by

$$\phi(t) = \mathbb{E} [e^{itX}].$$

Let's prove some properties of the characteristic function.

**Proposition 2.4.** Let  $X$  be a random variable and  $\phi$  be its characteristic function. Then,

1.  $\phi(0) = 1$  and  $|\phi(t)| \leq 1$  for all  $t$  (namely, the characteristic function is defined for all  $t$ ).
2.  $\phi$  is a uniformly continuous function.

3.  $\phi(t)$  is positive definite. That is, for all  $t_1, \dots, t_n \in \mathbb{R}$  and  $z_1, \dots, z_n \in \mathbb{C}$ ,

$$\sum_{k,j=1}^n \phi(t_k - t_j) z_k \overline{z_j} \geq 0.$$

*Proof.*

1. Note that  $\phi(0) = \mathbb{E}[e^{i0X}] = 1$ . As for the other claim  $\mathbb{E}[|e^{itX}|] \leq 1$ .

2. Note that

$$\begin{aligned} |\phi(t+s) - \phi(t)| &= |\mathbb{E}[e^{i(t+s)X} - e^{itX}]| \leq \mathbb{E}|e^{itX}(e^{isX} - 1)| = \mathbb{E}[|e^{itX}| |e^{isX} - 1|] \\ &= \mathbb{E}|e^{isX} - 1|. \end{aligned}$$

Now, let  $\delta(s) = \mathbb{E}|e^{isX} - 1|$ . Note that  $\delta(s) \leq 2$  and  $\delta(s) \rightarrow 0$  as  $s \rightarrow 0$ . In other words, the right hand side of

$$|\phi(t+s) - \phi(t)| \leq \delta(s)$$

is bounded independently of  $t$ , from which we obtain the desired result.

3. The last claim follows from some algebra:

$$\begin{aligned} \sum_{k,j=1}^n \phi(t_k - t_j) z_k \overline{z_j} &= \mathbb{E} \left[ \sum_{k,j=1}^n e^{i(t_k - t_j)X} z_k \overline{z_j} \right] \\ &= \mathbb{E} \left[ \sum_{k,j} e^{it_k X} z_k e^{-it_j X} \overline{z_j} \right] \\ &= \mathbb{E} \left[ \sum_k \sum_j e^{it_k X} z_k \overline{e^{it_j X} z_j} \right] \\ &= \mathbb{E} \left[ \left( \sum_k e^{it_k X} z_k \right) \overline{\left( \sum_j e^{it_j X} z_j \right)} \right] \\ &\geq 0. \end{aligned}$$

□

Let's now prove that the characteristic function also generates the moments by differentiation. In other words, we can think of the characteristic function as the new and improved "version 2" of the MGF.

**Proposition 2.5.** *Let  $\phi$  be the characteristic function of a random variable  $X$ .*

1. If  $X^k$  is integrable, then  $\phi$  is  $k$  times differentiable. Moreover,

$$\phi(t) = \sum_{j=0}^k \frac{(it)^j}{j!} \mathbb{E}[X^j] + o(t^k) \quad \text{as } t \rightarrow 0. \quad (1)$$

In particular,

$$\phi^{(n)}(0) = i^n \mathbb{E}[X^n] \quad \text{for } n = 1, \dots, k.$$

2. Conversely, if  $\phi^{(2k)}(0)$  exists and is finite, then  $\mathbb{E}[|X|^{2k}] < \infty$ .

*Proof.*

1. Our proof is going to be very similar to the proof for the MGF which we saw earlier. Let's start with  $k = 1$ . Let  $h > 0$  and consider the quantity

$$\begin{aligned} \frac{\phi(t+h) - \phi(t)}{h} &= \frac{\mathbb{E}[e^{i(t+h)X} - e^{itX}]}{h} \\ &= \mathbb{E}\left[e^{itX} \frac{e^{ihX} - 1}{h}\right] \end{aligned} \quad (2)$$

Now, note that

$$\frac{e^{ihX} - 1}{h} = \frac{\cos(hX) - 1}{h} + i \frac{\sin(hX)}{h}.$$

By the mean value theorem, we can find  $\theta$  and  $\theta'$  (depending on  $\omega$ ) between 0 and  $h$  such that

$$\frac{\cos(hX) - \cos(0X)}{h} = \frac{\cos(hX) - 1}{h} = -X \sin(\theta X)$$

and

$$\frac{\sin(hX) - \sin(0X)}{h} = \frac{\sin(hX)}{h} = X \cos(\theta' X).$$

Therefore,

$$\frac{e^{ihX} - 1}{h} = -X \sin(\theta X) + X \cos(\theta' X)$$

and hence

$$\left| \frac{e^{ihX} - 1}{h} \right| \leq |X| |\sin(\theta X) + \cos(\theta' X)| \leq 2|X|.$$

We have proved that the argument in the expectation (2) is dominated by an integrable random variable (namely,  $2|X|$ ). Therefore, we can take limits to get

$$\begin{aligned} \lim_{h \downarrow 0} \frac{\phi(t+h) - \phi(t)}{h} &= \mathbb{E} \left[ e^{itX} \lim_{h \downarrow 0} \frac{e^{ihX} - 1}{h} \right] \\ &= \mathbb{E} \left[ e^{itX} \lim_{h \downarrow 0} \frac{1 + (ihX) + (ihX)^2 + \dots - 1}{h} \right] \\ &= \mathbb{E} \left[ iX e^{itX} \lim_{h \downarrow 0} \frac{(ihX)^2 + \dots}{h} \right] \\ &= i \mathbb{E} [X e^{itX}]. \end{aligned}$$

Now, what about the case of  $k > 1$ ? This is handled by induction. Suppose that

$$\phi^{(k-1)}(t) = i^{k-1} \mathbb{E} [X^{k-1} e^{itX}]$$

for some  $k$ . Consider

$$\begin{aligned} \frac{\phi^{(k-1)}(t+h) - \phi^{(k-1)}(h)}{h} &= \frac{i^{k-1} \mathbb{E} [X^{k-1} e^{i(t+h)X}] - i^{k-1} \mathbb{E} [X^{k-1} e^{ihX}]}{h} \\ &= i^{k-1} \mathbb{E} \left[ \frac{X^{k-1} e^{itX} (e^{ihX} - 1)}{h} \right]. \end{aligned}$$

We can follow the same arguments to show that the argument in the expectation is bounded by  $2|X|^k$  and take limits to get

$$\lim_{h \downarrow 0} \frac{\phi^{(k-1)}(t+h) - \phi^{(k-1)}(h)}{h} = i^k \mathbb{E} [X^k e^{itX}].$$

This establishes

$$\phi^{(k)}(t) = i^k \mathbb{E} [X^k e^{itX}].$$

Now, applying Taylor's theorem to expand  $e^{itX}$ , we obtain the form (1).

2. We'll just handle the case of  $k = 1$ . Recall that for a function  $f$  that is twice differentiable at  $x$ ,

$$f''(x) = \lim_{h \downarrow 0} \frac{f(x-h) - 2f(x) + f(x+h)}{h^2}.$$

Now, suppose  $\phi''(0) = A$  for some finite number  $A$ . Then,

$$\begin{aligned} A &= \lim_{h \downarrow 0} \frac{\phi(-h) - 2\phi(0) + \phi(h)}{h^2} \\ &= \lim_{h \downarrow 0} \frac{\phi(-h) - 2 + \phi(h)}{h^2} \\ &= \lim_{h \downarrow 0} \mathbb{E} \left[ \frac{e^{-ihX} + e^{ihX} - 2}{h^2} \right] \\ &= 2 \lim_{h \downarrow 0} \mathbb{E} \left[ \frac{\cos(hX) - 1}{h^2} \right]. \end{aligned}$$

Now  $1 - \cos t \geq 0$ . Moreover, for  $|t| \leq 2$ , it turns out that  $\cos t \leq 1 - t^2/4$  (check this). Therefore, if  $|hx| \leq 2$ ,

$$\frac{1 - \cos(hx)}{h^2} \geq \frac{(hx)^2}{4h^2} = \frac{x^2}{4}$$

and hence

$$(1 - \cos(hX)) \geq \frac{X^2}{4} I_{\{|X| \leq 2/h\}}.$$

Therefore,

$$|A| = 2 \lim_{h \downarrow 0} \mathbb{E} \left[ \frac{1 - \cos(hX)}{h^2} \right] \geq \frac{1}{2} \lim_{h \downarrow 0} \mathbb{E} [X^2 I_{\{|X| \leq 2/h\}}].$$

Now, as  $h \downarrow 0$ ,  $X^2 I_{\{|X| \leq 2/h\}} \rightarrow X^2$  pointwise. Therefore, by the Monotone Convergence Theorem,

$$2|A| \geq \mathbb{E}[X^2],$$

as desired.

□