03/09

2.6. Moment generating functions

Def (2.3.6.) The MGF of a r.v. X is

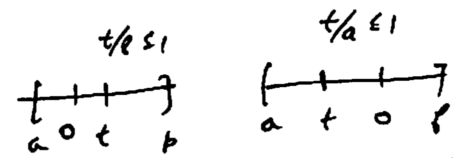
$$M(t) := \mathbb{E}[e^{tX}], \quad t \in \mathbb{R}$$

Prop. (Properties):

(i) $M(0) = 1$ — obvious

(ii) If MGF is finite on the endpoints of an interval $[a,b]$, then it is finite on all points of $[a,b)$.

(ii) Let $t \in (a,b)$. Then either $\frac{t}{a} \le 1$ or $\frac{t}{b} \le 1$ (e.g.



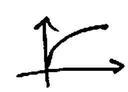
)

Assume $t/b \le 1$. By Jensen's inequality,

$$M(t) = \mathbb{E}[e^{tX}] = \mathbb{E}\left[e^{bX \cdot \frac{t}{b}}\right]$$

$$\le \left[\underbrace{\mathbb{E}\, e^{bX}}_{<\infty}\right]^{t/b}$$

(the function $f(x) = x^{t/b}$ is concave)



$$= M(b)^{t/b} < \infty.$$

Thm 2.3.8 (More properties) Suppose $M(\theta) < \infty$ in some neighborhood of 0.

Then
$$M'(0) = \mathbb{E}[X]$$
$$M''(0) = \mathbb{E}[X^2],$$
$$\vdots$$
$$M^{(n)} = \mathbb{E}[X^n]$$
"generates moments of X.

$$M'(t) = \frac{d}{dt}\mathbb{E}[e^{tX}] = \mathbb{E}\left[\frac{d}{dt}e^{tX}\right] \quad \text{(assuming we can interchange)}$$

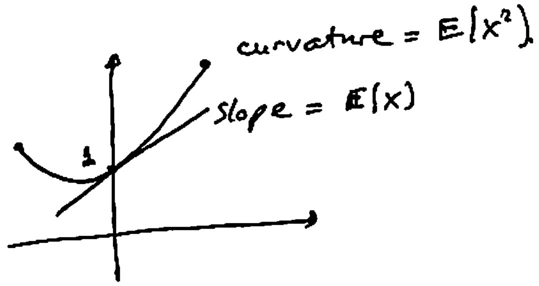
$$= \mathbb{E}[Xe^{tX}].$$

$$\Rightarrow M'(0) = \mathbb{E}[X].$$

Similarly, $M''(t) = \mathbb{E}[X^2 e^{tX}] \Rightarrow M''(0) = \mathbb{E}[X^2]$, etc.

$\boxed{\text{Lemma: MGF is a convex function.}}$

$$M''(t) = \mathbb{E}\big[\underbrace{X^2}_{\geq 0}\underbrace{e^{tX}}_{\geq 0}\big] \geq 0.$$



The local behavior of MGF near 0 is important.

(MGF is determined by moments)

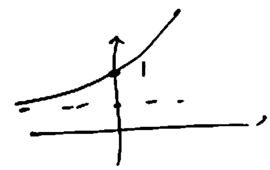
Thm (23.8) ~~Suppose $M(\theta)<\infty$~~

$$M(t)=\sum_{n\ge 0}\frac{t^n}{n!}\,\mathbb{E}[X^n]$$

if the series converges absolutely.

$$M(\theta)=\mathbb{E}[e^{tX}]=\mathbb{E}\left[\sum_{n\ge 0}\frac{t^n}{n!}X^n\right]$$
$$=\sum_{n\ge 0}\frac{t^n}{n!}\,\mathbb{E}[X^n] \qquad \text{(assume we can interchange)}$$

Examples (a) $X\sim \mathrm{Ber}(p)\Rightarrow$



$$M(t)=e^{t\cdot 1}p+e^{t\cdot 0}(1-p)=pe^{t}+1-p.$$
$$M'(t)=pe^{t}\Rightarrow \mathbb{E}[X]=M'(0)=pe^{0}=p.$$

(b) $X\sim \mathrm{Poisson}(\lambda)$.

$$M(\theta)=\sum_{k=0}^{\infty}\underbrace{\frac{\lambda^k}{k!}e^{-\lambda}}_{=p(k)}\,e^{tk}=e^{-\lambda}\sum_{k=0}^{\infty}\frac{(\lambda e^{t})^k}{k!}$$
$$=\boxed{e^{\lambda(e^{t}-1)}}$$

$$\Rightarrow \mathbb{E}[X]=M'(0)=\lambda e^{t}\cdot e^{\lambda(e^{t}-1)}\Big|_{t=0}=\boxed{\lambda}$$
$$\mathbb{E}[X^2]=M''(0)=\lambda+\lambda^2 \Rightarrow \mathrm{Var}(X)=\lambda.$$

(c) $X\sim N(0,1)$

$$M(\theta)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{tx}e^{-x^2/2}\,dx=\cdots=\boxed{e^{t^2/2}}.$$

~~$\Rightarrow M'(0$~~ $\Rightarrow \mathbb{E}[X]=M'(0)=te^{t^2/2}\Big|_{t=0}=0.$

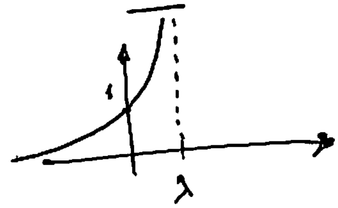
$$\mathbb{E}[X^2]=M''(0)=(t^2+1)e^{t^2/2}\Big|_{t=0}=1. \quad ✓$$

(d) $X \sim N(\mu, \sigma^2) \Rightarrow$

$$M(t) = \exp\left(\mu t + \frac{1}{2}\sigma^2 t^2\right). \qquad (\text{check!})$$

(e) Compute MGF of $\text{Exp}(\lambda)$, (Only exists on some interval!)

show $M(t) = \frac{\lambda}{\lambda - t}$, for $t < \lambda$.



(f) MGF of Cauchy $= \infty \quad \forall t \neq 0$

Prop If $X_1, \dots, X_n$ are indep. r.v's then

$$M_{X_1 + \dots + X_n}(t) = M_{X_1}(t) \cdots M_{X_n}(t), \quad \forall t \in \mathbb{R}.$$

$$\ulcorner \quad \mathbb{E}\left[e^{t(X_1 + \dots + X_n)}\right] = \mathbb{E}\left[e^{tX_1} \cdots e^{tX_n}\right] \underset{\text{indep}}{=} \mathbb{E}\left[e^{tX_1}\right] \cdots \mathbb{E}\left[e^{tX_n}\right] \quad \lrcorner.$$

Ex (g) $X \sim \text{Gamma}(n, \lambda) \Rightarrow M(t) = \left(\frac{\lambda}{\lambda - t}\right)^n$

since $X = X_1 + \dots + X_n$, $X_i \sim \text{Exp}(\lambda)$ indep.

Another proof of:

Thm Let $X_i \sim N(\mu_i, \sigma_i^2)$ be independent. Then

$$\sum_1^n X_i \sim N(\Sigma \mu_i, \Sigma \sigma_i^2).$$

$$\begin{aligned}
M_{\Sigma X_i}(t) &= \prod_i M_{X_i}(t) \quad \text{(by independence)} \\
&= \prod_i \exp\left(\mu_i t + \tfrac{1}{2}\sigma_i^2 t^2\right) \quad \text{(p. 80)} \\
&= \exp\left[\underbrace{(\Sigma \mu_i)}_{\mu} t + \tfrac{1}{2}\underbrace{(\Sigma \sigma_i^2)}_{\sigma} t^2\right] \\
&= M_X(t) \quad \text{for } X \sim N(\mu, \sigma^2).
\end{aligned}$$

~~Thm Let X_i~~

Thm (Uniqueness) MGF determines the distribution of X uniquely.

That is, ~~assume~~ let X, Y be r.v's such that

$$M_X(t) = M_Y(t) \quad \forall t \in \mathbb{R}.$$

Then X, Y have same distribution.

(For discrete r.v's taking finitely many values):

X takes values x_i with prob. p_i, $i = 1, \dots, n$

Y : y_j, q_j, $j = 1, \dots, m$.

Assumption $\Rightarrow$

$$\sum_{i=1}^{n} p_i e^{t x_i} = \sum_{j=1}^{m} q_j e^{t y_j} \quad \forall t \in \mathbb{R} \qquad (*)$$

WLOG $x_n = \max_{i \le n} x_i$, $y_m = \max_{i \le m} y_j$.

As $t \to \infty$, $\displaystyle\sum_{i=1}^{n} p_i e^{t x_i} = (1 + \underbrace{r(t)}_{\downarrow 0}) p_n e^{t x_n}$ $\quad \left(\text{set } r(t) = \sum_{i=1}^{n} \frac{p_i}{p_n} \underbrace{e^{t(x_i - x_n)}}_{\downarrow 0}\right)$

$$\sum_{j=1}^{m} q_j e^{t y_j} = (1 + \underbrace{s(t)}_{\downarrow 0}) q_m e^{t y_m}$$

$$(*) \Rightarrow \quad \frac{p_n}{q_m} e^{t(x_n - y_m)} = 1 \quad \forall t \in \mathbb{R}.$$

$\Rightarrow x_n = y_m$, $p_n = q_m$.

Iterate $(x_{n-1}, y_{m-1}) \Rightarrow$ QED.

~~**Def** (Characteristic functions) Let X be a r.v. The characteristic function of X is~~

~~$$\phi(t) := \mathbb{E}[e^{itX}] = M(it), \quad t \in \mathbb{R}$$~~

~~Remark: $|e^{itX}| = 1 \ \forall t \Rightarrow$ characteristic function is finite $\forall t \in \mathbb{R}$~~ ~~as oppos~~ ~~In contrast to MGF.~~

Characteristic Functions

Def Let X be a r.v. The char function of X is
$$\phi(t) := \mathbb{E}[e^{itX}] = \underset{\substack{\uparrow \\ \text{MGF}}}{M(it)}, \quad t \in \mathbb{R}$$

Advantage over MGF: char. function is defined $\underline{\forall t \in \mathbb{R}}$, since $|e^{itx}| = 1$.

Ex (a) $X \sim N(0,1) \Rightarrow \phi(t) = e^{-t^2/2}$

(b) $X = \pm 1$ with prob $\frac{1}{2}$ each $\Rightarrow \phi(t) = \cos(t)$ (Check!)

(c) $X \sim$ cauchy, i.e. pdf $f(x) = \frac{1}{\pi(1+x^2)}$, $x \in \mathbb{R}$.

$\Rightarrow \phi(t) = e^{-|t|}$ ~~(can be checked using Fourier inv~~

(recall MGF of X is $= \infty$).

Char. functions share many properties with MGF, including:
$$\phi_{\Sigma X_i}(t) = \phi_{X_1}(t) \cdots \phi_{X_n}(t)$$
if X_i are independent.

~~Rem Remark If X has dens~~

Remark Char. function also determines the distr. of X uniquely.

Remark. Let X have density $f(x)$. Then
$$\phi(x) = \int_{-\infty}^{\infty} e^{itx} f(x)\,dx = \hat{f}(-t), \quad \text{\underline{Fourier transform} of } f.$$

Fourier inversion formula:
$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \phi(t)\,dt$$
allows to compute density from char. function.