Chapter 7: Moment Generating Functions

The moment generating function (mgf) of a random variable X is

① $$M_X(t) = E(e^{tX}) = \begin{cases} \sum_x e^{tx}\, p(x) & \text{if } X \text{ is discete with pmf } p(x) \\ \int_{-\infty}^{\infty} e^{tx} f(x)\,dx & \text{if } X \text{ is continuous with pdf } f(x) \end{cases}$$

⇒ t is a real number

usual formula of E(X)

for real t for which this expression exists. within support

② $M_X(0) \; (t=0) = E(e^{0X}) = E(e^0) = E(1) = 1$

expectation of a constant is that constant

$M_X(t)$ doesn't exist for all values of t and we always mention restrictions

Example 7: Find the mgf of random variable X with the pmf

x	1	2	3
$p(x)$	0.25	0.6	0.15

$$M_X(t) = E(e^{tX}) = \sum_{x=1}^{3} e^{tx}\, p(x)$$

$$= e^{t\cdot 1}(0.25) + e^{t\cdot 2}(0.6) + e^{t\cdot 3}(0.15) =$$

$$= \boxed{(0.25)e^{t} + 0.6\, e^{2t} + 0.15\, e^{3t}, \quad -\infty < t < \infty}$$

a f(t)

Main uses of the mgf:

1. Mgf of a random variable uniquely identifies its distribution.

⊛ So, some functions that you get from calculating mgf can be used to traceback what the main distribution is

⊛ Or sometimes, you will get back $\sum e^{tx} p(x)$ in discrete case so you know p(x) from that

3

2. Finding moments using the mgf.

Recall: The kth moment of X is $E(X^k)$, $k = 1, 2, 3, \ldots$

provided that it exists at $t=0$

$$E(X^k) = \left.\frac{d^k M(t)}{dt^k}\right|_{t=0} \text{ (at } t=0\text{)} = M^{(k)}(t)\Big|_{t=0}$$

kth derivative

- $k=1$, $E(X) = \left.\frac{dM(t)}{dt}\right|_{t=0} = M'(t)\Big|_{t=0}$

- $k=2$, $E(X^2) = \left.\frac{d^2M(t)}{dt^2}\right|_{t=0} = M''(t)\Big|_{t=0}$

If it is not k times differentiable then you can't use this idea

In the textbook:

- The mgfs for common distributions are in Table 7.1 (discrete) and Table 7.2 (continuous) on pp.364-365.
- Study proofs of the mfg formulas of Binomial, Poisson, Normal, and Exponential on pp. 361-364

$\boxed{\text{Do it yourself}}$

Discrete

TABLE 7.1

	Probability mass function, $p(x)$	Moment generating function, $M(t)$	Mean	Variance
Binomial with parameters n, p; $0 \le p \le 1$	$\binom{n}{x} p^x(1-p)^{n-x}$ $x = 0, 1, \ldots, n$	$(pe^t + 1 - p)^n$ for $-\infty < t < \infty$	np	$np(1-p)$
Poisson with parameter $\lambda > 0$	$e^{-\lambda}\frac{\lambda^x}{x!}$ $x = 0, 1, 2, \ldots$	$\exp\{\lambda(e^t - 1)\}$ for $-\infty < t < \infty$	λ	λ
Geometric with parameter $0 \le p \le 1$	$p(1-p)^{x-1}$ $x = 1, 2, \ldots$	$\frac{pe^t}{1-(1-p)e^t}$ for $t < -\ln(1-p)$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
Negative binomial with parameters r, p; $0 \le p \le 1$	$\binom{n-1}{r-1} p^r(1-p)^{n-r}$ $n = r, r+1, \ldots$	$\left[\frac{pe^t}{1-(1-p)e^t}\right]^r$ for $t < -\ln(1-p)$	$\frac{r}{p}$	$\frac{r(1-p)}{p^2}$

Example: Prove/find the mgf of $X \sim \text{Pois}(\lambda)$.

$$M_X(t) = E(e^{tX}) = \sum_{k=0}^{\infty} (e^{tk}) \cdot \frac{\lambda^k e^{-\lambda}}{k!} = \sum_{k=0}^{\infty} \frac{(e^{-\lambda})(e^t \cdot \lambda)^k}{k!}$$

$$= e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda e^t)^k}{k!}$$

$$\circledast \quad \boxed{\sum_{k=0}^{\infty} \frac{a^k}{k!} = e^a \quad \Rightarrow a = \lambda e^t = \text{const}}$$

$$= e^{-\lambda} (e^{\lambda e^t}) = e^{\lambda e^t - \lambda}$$

$$\boxed{= e^{\lambda(e^t - 1)}, \quad -\infty < t < \infty}$$

Since we had to put no restrictions during derivation

TABLE 7.2

	Probability mass function, $f(x)$	Moment generating function, $M(t)$	Mean	Variance
Uniform over (a, b)	$f(x) = \begin{cases} \frac{1}{b-a} & a < x < b \\ 0 & \text{otherwise} \end{cases}$	$\frac{e^{tb} - e^{ta}}{t(b-a)}$ for $t \neq 0$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Exponential with parameter $\lambda > 0$	$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$	$\frac{\lambda}{\lambda - t}$ for $t < \lambda$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Gamma with parameters $(s, \lambda), \lambda > 0$	$f(x) = \begin{cases} \frac{\lambda e^{-\lambda x}(\lambda x)^{s-1}}{\Gamma(s)} & x \geq 0 \\ 0 & x < 0 \end{cases}$	$\left(\frac{\lambda}{\lambda - t}\right)^s$ for $t < \lambda$	$\frac{s}{\lambda}$	$\frac{s}{\lambda^2}$
Normal with parameters (μ, σ^2)	$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2}$ $-\infty < x < \infty$	$\exp\left\{\mu t + \frac{\sigma^2 t^2}{2}\right\}$ for $-\infty < t < \infty$	μ	σ^2

Example: Prove/find the mgf of $X \sim \text{Exp}(\lambda)$.

$$M_X(t) = E(e^{tX}) = \int_{-\infty}^{\infty} e^{tx} f(x)\,dx$$

$$= \int_0^{\infty} e^{tx} \cdot \lambda e^{-\lambda x}\,dx = \lambda \int_0^{\infty} e^{(t-\lambda)x}\,dx$$

$$= \lambda \cdot \frac{e^{(t-\lambda)x}}{(t-\lambda)}\Bigg|_0^{\infty} = \lambda\left(\frac{0-1}{t-\lambda}\right)$$

$t-\lambda<0$ for finite value

$\because$ we want

$$\lambda\left(\cancel{\frac{e^{-\infty}}{--}} - \frac{e^0}{---}\right)$$

0

$$\boxed{= \frac{\lambda}{\lambda - t}, \quad t < \lambda}$$

$\Rightarrow \therefore$ $\boxed{t-\lambda \text{ has to be } < 0}$

Even when $t=\lambda$

$\Rightarrow$ infinite value

4

$f(x)$

$$\frac{1}{\sqrt{2\pi} \times \sqrt{2 \times 968}} \times e$$

$\mu = -18$

$\frac{\sigma^2}{2} = 968$

$$\frac{P(-30 < x < 60)}{P(x > -30)}$$

$$\frac{1}{\sqrt{2\pi} \times \sqrt{2 \times 968}} \times e^{-\frac{(x+18)^2}{2 \times 2 \times 968}} \, dx$$

$$e^{-(x+18)^2} \times (-2)(x+18)$$

-4

$$\frac{1}{\sqrt{2\pi} \times \sqrt{2 \times 968}}$$

$$\Rightarrow \frac{0.569314}{0.60746}$$

⊛ Now, $E(X) = M'(t)\Big|_{t=0} \Rightarrow M'(t) = \left(\frac{\lambda}{\lambda-t}\right)' =$

$$\frac{d}{dt}\left(\frac{\lambda}{\lambda-t}\right) = \lambda\left(\frac{-1}{(\lambda-t)^2}(-1)\right) = \boxed{\frac{\lambda}{(\lambda-t)^2}}$$

so if $t=0 \Rightarrow \frac{1}{\lambda}$

⊛ Now, $E(X^2) = M''(t)\Big|_{t=0}$

$$= \frac{d}{dt}\left(\frac{\lambda}{(\lambda-t)^2}\right) = \lambda\left(-2(\lambda-t)^{-3}\cdot(-1)\right)$$

$$= 2\lambda(\lambda-t)^{-3}$$

Now $t=0 \Rightarrow$

$$\boxed{2\lambda^{-2}}$$

⊛ Variance $= \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$

Additional discussion: If the mgf of random variable X has the form

$$M_X(t) = p_1 e^{v_1 t} + p_2 e^{v_2 t} + \cdots + p_k e^{v_k t} = \sum_{i=1}^{k} p_i e^{v_i t}$$

where

$p_1, p_2, \ldots, p_k$ are positive and $p_1 + p_2 + \cdots + p_k = 1$,

$v_1, v_2, \ldots, v_k$ are some real numbers,

then

X is discrete with pmf

x	v_1	v_2	......	v_k
$P(X=x)$	p_1	p_2	......	p_k

& $\sum p_i = 1$

Example 8: Suppose $M_X(t) = 0.3 + 0.2e^{20t} + 0.5e^{-35t}$. Find $E(X)$, $Var(X)$, and $P(X \geq 0)$.

- $E(X) = M'(t)\big|_{t=0} = 4e^{20\cdot(0)} + (-35)(0.5)e^{-35(0)}$
  $= 4 + (-17.5) = \boxed{-13.5}$

- $E(X^2) = M''(t)\big|_{t=0} = \left((0.2)(20^2)(e^{20t}) + (0.5)(-35)^2(e^{-35t})\right)\Big|_{t=0}$
  $= \boxed{692.5}$

- $Var(X) = 692.5 - (-13.5)^2 = \boxed{510.25}$

⑤

$\boxed{\bullet\ M_X(t) = 0.3 + (0.2)e^{20t} + (0.5)(e^{-35t})}$

for $t \in \mathbb{R}$ if not mentioned

$\boxed{\sum_{i=1}^{k} p_i e^{v_i t}}$

$\Downarrow$

form needed

$\sum_x e^{tx} p(x) \Leftarrow = (0.3)e^{0\cdot t} + (0.2)(e^{20t}) + (0.5)(e^{-35t})$

$\Rightarrow$ X has pmf

x	0 ✓	20 ✓	-35 ✗
$P(X=x)$	0.3	0.2	0.5

(All coeff have to be ⊕ and add upto 1)

⊛ Only apply if this holds

$\Rightarrow P(X \geq 0) = 0.3 + 0.2 = \boxed{0.5}$

Example 9: The value of a piece of factory equipment is $100 \cdot 0.5^X$, where X is a random variable with mgf $M_X(t) = \frac{1}{1-2t}$ for $t < \frac{1}{2}$. Find the expected value of this piece of equipment.

Method I

$Y = 100 \cdot (0.5)^X$; $E(Y) = E(100 \cdot (0.5)^X)$

$$\boxed{M_X(t) = \frac{1}{1-2t} = \frac{1/2}{1/2 - t}, \quad t < \frac{1}{2}}$$

$$\boxed{\frac{\lambda}{\lambda - t}, \quad t < \lambda}$$

$$\boxed{X \sim Exp\left(\lambda = \frac{1}{2}\right)}$$

$f_X(x) = \frac{1}{2} e^{\frac{1}{2}x}$, $x \geq 0$ for exp

$$E(100 \cdot (0.5)^X) = \int_0^\infty \left(100 \cdot (0.5)^x\right) \cdot \frac{1}{2} e^{-\frac{1}{2}x} \, dx =$$

$$\boxed{\int a^x dx = \frac{a^x}{\ln a} + c}$$

$$50 \int_0^\infty \left(0.5 e^{-\frac{1}{2}}\right)^x dx = 50 \cdot \frac{\left(0.5 e^{-\frac{1}{2}}\right)^x}{\ln\left(0.5 e^{-\frac{1}{2}}\right)} \Bigg|_0^\infty =$$

$$= \frac{50}{\log(0.5) - \frac{1}{2}} (0 - 1) = \boxed{41.9}$$

Method II

$$E(100 \cdot (0.5)^X) = 100 \, E(0.5^X) = 100 \, \underbrace{E\left(\left(e^{(\ln 0.5)}\right)^X\right)}_{E(e^{tX}) = M_X(t)}$$

6

$$= 100 \cdot M_X(\underbrace{\ln 0.5}_{t}) = (100)\left(\frac{1}{1 - 2\ln(0.5)}\right)$$

$$= \boxed{41.9}$$

given in ques

Doesn't always apply, works only if it is a common known type of distribution (maybe you'll have to rewrite a bit)

Example : X has mgf $M(t) = (0.4 + 0.6e^t)^{10}$

find the mean and variance of X.

$\Rightarrow$ $E(X) = M'(t)\Big|_{t=0}$; $E(X^2) = M''(t)\Big|_{t=0}$

$\Rightarrow$ $Var(X) = E(X^2) - (E(X))^2$ $\Rightarrow$ solve

$\Rightarrow$ $X \sim Binom(n=10, p=0.6)$ $\Rightarrow$ from the table above

$\boxed{((1-p) + pe^t)^n}$

$\Rightarrow$ $E(X) = np = 6$

$Var(X) = np(1-p) = 6 \times 0.4 = 2.4$

Make sure to mention they are independent

MGF of Sums of Independent Random Variables

7

Proposition: If $X_1, X_2, \dots, X_n$ are independent, then for any functions $g_1(X_1), g_2(X_2), \dots, g_n(X_n)$

$$E(g_1(X_1) \cdot g_2(X_2) \cdot \dots \cdot g_n(X_n)) = E(g_1(X_1)) \cdot E(g_2(X_2)) \cdot \dots \cdot E(g_n(X_n))$$

$g_i(X) = e^{tX_i}$ and i from 1 to n

$$E\left(e^{t(X_1+X_2+\dots X_n)}\right) = E(e^{tX_1}) \cdot E(e^{tX_2}) \dots\dots E(e^{tX_n})$$

$$= M_{X_1}(t) \cdot M_{X_2}(t) \dots M_{X_n}(t)$$

Corollary: If $X_1, X_2, \dots, X_n$ are independent, then

8

$$\boxed{M_{X_1+X_2+\dots+X_n}(t) = M_{X_1}(t) \cdot M_{X_2}(t) \cdot \dots \cdot M_{X_n}(t)}$$

$E\left(e^{t(X_1+X_2+\dots X_n)}\right)$

and moreover,

This whole thing becomes t

$$\boxed{M_{a_1X_1+a_2X_2+\dots+a_nX_n}(t) = M_{X_1}(a_1t) \cdot M_{X_2}(a_2t) \cdot \dots \cdot M_{X_n}(a_nt)}$$

Recall: If $X_1, X_2 \dots X_n$ are independent & $X_i \sim N(\mu_i, \sigma_i^2)$ $(i = 1, 2, \dots n)$

then $\sum_{i=1}^{n} a_i X_i \sim N\left(\sum_{i=1}^{n} a_i \mu_i, \sum_{i=1}^{n} a_i^2 \sigma_i^2\right)$

Proof: $M_{a_1X_1+\dots a_nX_n}(t) = M_{X_1}(a_1t) \dots\dots M_{X_n}(a_nt)$

since $X_1, X_2 \dots X_n$ are independent

$$= \left(e^{\mu_1(a_1t) + \frac{\sigma_1^2}{2}(a_1t)^2}\right) \cdot \dots\dots (\dots) \cdot \left(e^{\mu_n(a_nt) + \frac{\sigma_n^2}{2}(a_nt)^2}\right)$$

for normal

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

$M(t) = e^{\mu t + \frac{\sigma^2}{2}t^2}$

$$= e^{(\mu_1(a_1t) + \dots \mu_n(a_nt)) \left(\frac{\sigma_1^2}{2}(a_1t)^2 + \dots \frac{\sigma_n^2}{2}(a_nt)^2\right)}$$

$$= e^{t\underbrace{(a_1\mu_1 + \dots a_n\mu_n)}_{\mu} + \frac{t^2}{2}\underbrace{(a_1^2\sigma_1^2 + \dots\dots a_n^2\sigma_n^2)}_{\sigma^2}}$$

$$= e^{\mu t + \frac{\sigma^2 t^2}{2}} = \text{mgf of } N\left(\sum_{i=1}^{n} a_i\mu_i, \sum_{i=1}^{n} a_i^2\sigma_i^2\right)$$

Variance

mean

$\Rightarrow a_1X_1 + \dots a_nX_n \sim N(\Sigma \dots, \Sigma \dots)$

Example 10: Let X_1, X_2, X_3 be a random sample from a distribution with the pmf

$$p(x) = \begin{cases} 1/3 & x = 0 \\ 2/3 & x = 1 \\ 0 & \text{otherwise} \end{cases}$$ $\Rightarrow$ ∴ Not taking this value

Find the mgf of $Y = X_1 + X_2 + X_3$ and determine the distribution of Y.

X_1, X_2, X_3 are iid (independent & identically distributed)

x	0	1
$P(x)$	1/3	2/3

$\Rightarrow X_i \sim \text{Bernoulli}\left(\frac{2}{3}\right)$

$\sim \text{Binom}\left(n=1, p=\frac{2}{3}\right)$

$$\boxed{M_Y(t) = M_{X_1+X_2+X_3}(t) = M_{X_1}(t)\, M_{X_2}(t)\, M_{X_3}(t)}$$

b/c X_1, X_2, X_3 are independent

same for all X_1, X_2, X_3 ∴ X

$$M_X(t) = E(e^{tX}) = \sum_x e^{tx} p(x)$$

$$= e^{t\cdot 0}\left(\frac{1}{3}\right) + e^{t\cdot 1}\left(\frac{2}{3}\right) = \boxed{\frac{1}{3} + \frac{2}{3}e^t}$$

$$\boxed{\left((1-p) + pe^t\right)^n}$$

Binomial form

$$M_Y(t) = \left(\frac{1}{3} + \frac{2}{3}e^t\right)^3$$

$$\boxed{\Rightarrow Y \sim \text{Binom}\left(n=3, p=\frac{2}{3}\right)}$$

(from table)

Distribution of Y

Joint Moment Generating Functions

9

The joint moment generating function of random variables $X_1, X_2, \dots, X_n$ is

$$\boxed{M_{X_1, X_2, \dots, X_n}(t_1, t_2, \dots, t_n) = E(e^{t_1X_1+t_2X_2+\cdots+t_nX_n})}$$

for real $t_1, t_2, \dots, t_n$ for which this expression exists.

Marginal mgf of X_1

$$M_{X_1}(t_1) = E(e^{t_1X_1}) = E(e^{t_1X_1+0X_2+\cdots+0X_n})$$
$$= M_{X_1,X_2\dots X_n}(t_1, 0, \dots, 0)$$

connecting joint with individual

Marginal mgf of X_i

$$\boxed{M_{X_i}(t_i) = M_{X_1,\dots X_i,\dots X_n}(0,\dots 0, t_i, 0\dots 0)}$$

Proposition

RVs $X_1, \dots, X_n$ are independent $\Leftrightarrow$

$$M_{X_1,\dots X_n}(t_1, \dots t_n) = M_{X_1}(t_1)\cdots M_{X_n}(t_n)$$

for all $t_1, \dots t_n$

Something similar to what we have studied

Example: Let X and Y be independent RVs with mgf $M(t) = e^{\frac{t^2}{2}}$

i) Let $U = X+Y$ and $V = X-Y$

$\boxed{\text{Joint} = \text{Marg}\cdot\text{Marg}?}$

Find the joint mgf of U & V. Are U & V independent?

$$M_{U,V}(t_1,t_2) = E(e^{t_1 U + t_2 V}) = E\left(e^{t_1(X+Y) + t_2(X-Y)}\right) =$$

$$= E\left(e^{(t_1+t_2)X + (t_1-t_2)Y}\right) = M_{X,Y}(t_1+t_2,\ t_1-t_2)$$

$$\boxed{= M_X(t_1+t_2)\, M_Y(t_1-t_2)} \quad (\because X, Y \text{ are independent})$$

$$= e^{(t_1+t_2)^2/2} \cdot e^{(t_1-t_2)^2/2} = \boxed{e^{t_1^2+t_2^2}}$$

from question

$$-\infty < t_1 < \infty, \quad -\infty < t_2 < \infty$$

Checking for independence

ii) $\ni$ $\boxed{M_{U,V}(t_1,t_2) = e^{t_1^2+t_2^2}}$

$M_U(t_1) = M_{(U,V)}(t_1, 0) = e^{t_1^2}$ $(-\infty < t_1 < \infty)$

$M_V(t_2) = M_{(U,V)}(0, t_2) = e^{t_2^2}$ $(-\infty < t_2 < \infty)$

Marginal mgfs

$$M_{U,V}(t_1,t_2) = e^{t_1^2+t_2^2} = M_U(t_1)\cdot M_V(t_2) = e^{t_1^2}\cdot e^{t_2^2}$$

$\boxed{-\infty < t_1, t_2 < \infty}$

Since these are equal by chance $\Rightarrow$ U & V are independent

⊛ X, Y $M(t) = e^{t^2/2}$

∵ same distribution so taking together

$= e^{0t + \frac{1^2}{2}t^2}$ mgf of $N(\mu, \sigma^2)$

but μ here $= 0$, & $\sigma = 1$

mgf of

$\Rightarrow X, Y \sim N(0,1)$

⊛ U, V $M_U(t)$ or $M_V(t)$ $= e^{t^2} = e^{0 \cdot t + \frac{2}{2}t^2}$

$\Rightarrow U, V \sim N(0,2)$

⊛ Relation b/w joint mgf & joint pmf

⊛ $M_{X_1, \dots X_n}(t_1, \dots t_n) = E\left(e^{t_1 X_1 + \dots + t_n X_n}\right)$

$X_1 \dots X_n$ have joint pmf $p(x_1, \dots x_n)$

$$= \sum_{x_1} \dots \sum_{x_n} e^{t_1 x_1 + \dots + t_n x_n} \cdot p(x_1, \dots x_n)$$

⊛ Just like mgf identifies distribution joint mgf will identify joint distribution as long as they exist.