Math 4501 - Probability and Statistics II

5.4 - The moment generating function technique

Key new fact from last meeting: if $Z \sim N(0,1) \Rightarrow V = Z^2 \sim \chi^2(1)$

consequence: if $X \sim N(\mu, \sigma^2) \rightarrow Z = \frac{X-\mu}{\sigma} \sim N(0,1)$

$\Rightarrow V = Z^2 = \left(\frac{X-\mu}{\sigma}\right)^2 \sim \chi^2(1)$

what is used for

Sec 3.3 Ex 12

Linear combination of independent random variables mgf

Theorem

Let $X_1, X_2, \ldots, X_n$ be independent random variables with respective moment-generating functions $M_{X_i}(t)$ defined on intervals of the form $(-h_i, h_i)$, for some positive constants h_i, $i = 1, 2, \ldots, n$.

The moment-generating function of the linear combination

$$\boxed{Y = \sum_{i=1}^{n} a_i X_i}$$

is given by

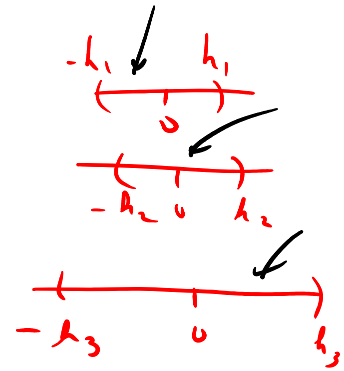
$$\boxed{M_Y(t) = \prod_{i=1}^{n} M_{X_i}(a_i t)},$$

where t satisfies each one of the conditions $-h_i < a_i t < h_i$ for $i = 1, 2, \ldots, n$.

Proof (technique of the proof is important):

Let $X_1, \dots, X_n$ be independent r.v.s with respective mgfs:

$$M_{X_i}(t), \qquad t \in (-h_i, h_i), \quad h_i > 0$$



The m.g.f of $Y = \sum_{i=1}^{n} a_i X_i$ is

$$M_Y(t) \underset{\substack{\text{def} \\ (3501)}}{=} E\left[e^{tY}\right] \underset{\substack{\text{def of} \\ Y}}{=} E\left[e^{t\sum_{i=1}^{n} a_i X_i}\right] = E\left[e^{a_1 t X_1 + a_2 t X_2 + \dots + a_n t X_n}\right]$$

$$\underset{\substack{\text{property of} \\ \text{exponential}}}{=} E\left[e^{a_1 t X_1} \cdot e^{a_2 t X_2} \cdot \dots \cdot e^{a_n t X_n}\right] = \underbrace{E\left[e^{a_1 t X_1}\right]}_{M_{X_1}(a_1 t)} \cdot \underbrace{E\left[e^{a_2 t X_2}\right]}_{M_{X_2}(a_2 t)} \cdot \dots \cdot \underbrace{E\left[e^{a_n t X_n}\right]}_{M_{X_n}(a_n t)}$$

independent r.v.s

$X_1, X_2, \dots, X_n$ independent

$\Rightarrow e^{a_1 t X_1}, e^{a_2 t X_2}, \dots, e^{a_n t X_n}$ are also independent!

CONCLUSION:

$$M_Y(t) = M_{X_1}(a_1 t) \cdot M_{X_2}(a_2 t) \cdots M_{X_n}(a_n t) = \prod_{i=1}^{n} M_{X_i}(a_i t) \longleftarrow$$

$\uparrow$
mgf of Y

where t is such that

$$-h_i < a_i t < h_i \quad \text{for all} \quad i = 1, \dots, n$$

that is

$$-\frac{h_i}{|a_i|} < t < \frac{h_i}{|a_i|} \quad \text{for all} \quad i = 1, \dots, n$$

$M_Y(t)$ is defined for $t \in (-h, h)$ where $h = \min\left\{ \frac{h_i}{|a_i|} \;;\; i = 1, \dots, n \right\}$

Why must the mgf of a r.v. X be defined on an open interval

$$M_X(t) = E\left[e^{tX}\right]$$

if $t=0 \Rightarrow M_X(0) = E\left[e^{0\cdot X}\right] = E[1] = 1.$

$M_X(t)$ is allways defined for $t=0$:

BUT we want to "generate" the moments of X as $M_X^{(n)}(0) = E[X^n]$

To have $M_X(t)$ differentiable at $t=0$, $M_X(t)$ must be defined on some interval containing 0

$$f'(0) = \lim \frac{f(h) - f(0)}{h}$$

Sample mean mgf

$M_Y(t) = \prod_{i=1}^{n} M_{X_i}(a_i t)$

Corollary (Two important special cases of the previous theorem)

If $X_1, X_2, \ldots, X_n$ be a random sample of size n from a distribution with mgf $M(t)$, where $-h < t < h$. ↳ i i d

$M_{X_1}(t) = M_{X_2}(t) = \cdots M_{X_n}(t) = M(t)$

(a) The mgf of the sum $Y = \sum_{i=1}^{n} X_i$ is ← $a_1 = a_2 = \cdots = a_n = 1$

$$M_Y(t) = \prod_{i=1}^{n} M(t) = [M(t)]^n, \qquad -h < t < h\ .$$

(b) The mgf of the sample mean $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ is → $a_1 = a_2 = \cdots = a_n = \frac{1}{n}$

$$\longrightarrow \quad M_{\bar{X}}(t) = \prod_{i=1}^{n} M\left(\frac{t}{n}\right) = \left[M\left(\frac{t}{n}\right)\right]^n, \qquad -h < \frac{t}{n} < h\ .$$

Example

Let $X_1, X_2, \ldots, X_n$ denote the outcomes of n independent Bernoulli trials, each with probability of success p.

Find the mgf of $Y = \sum_{i=1}^{n} X_i$. $\Longrightarrow$ $M_Y(t) = (M(t))^n = (q + pe^t)^n$ $\leftarrow$ m.g.f of Binomial (n, p)

Let $\underbrace{X_1, \ldots, X_n}$ be independent r.v.s with distn. Bernoulli (p), $p \in (0,1)$

$\hookrightarrow$ each takes a value $\begin{cases} 0 \to \text{failure} \\ 1 \to \text{success} \end{cases}$

CONCLUSION

$Y \sim bi(n, p)$

Recall that the m.g.f of each X_i is

$M(t) = q + pe^t$ $\leftarrow$ check formula sheet

without formula sheet

$P(X=0) = 1-p = q$ $\qquad$ $P(X=1) = p$

$M(t) = E[e^{tX}] = f(0) \cdot e^{t \cdot 0} + f(1)\, e^{t \cdot 1} = q \cdot 1 + pe^t = q + pe^t$

For discrete r.v.s we can get the pmf from the mgf as follows

mgf of a r.v. X taking values $x_1, x_2, x_m, \dots$ is

$M(t) = E\left[e^{tX}\right]$ can be written as

$$= p_1 e^{tx_1} + p_2 e^{tx_2} + p_3 e^{tx_3} + \dots$$

$p_1 = P(X=x_1)$ $\quad$ $p_2 = P(X=x_2)$ $\quad$ $p_3 = P(X=x_3)$

Previous example for Bernoulli(p)

$$M(t) = q + pe^t = \underbrace{(1-p)}_{P(X=0)} e^{t.0} + \underbrace{p}_{P(X=1)} e^{t.1}$$

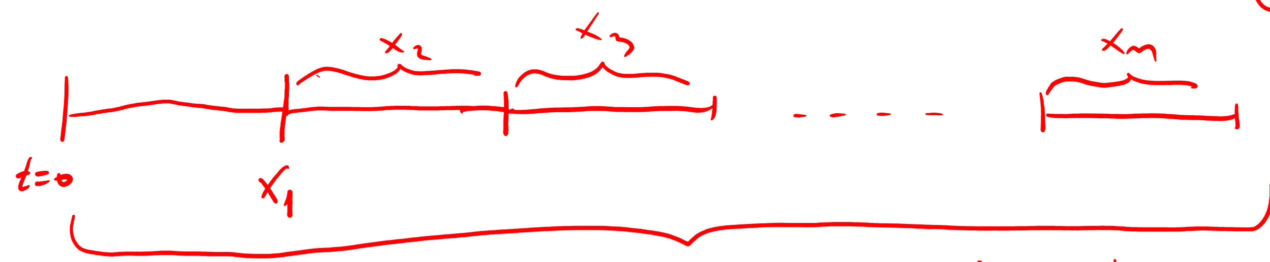
Example

Let $X_1, X_2, \dots, X_n$ be the observations of a random sample of size n from the exponential distribution having mean θ.

$X_1, \dots, X_n \sim Exp(\theta)$ independent

Find the mgf of $Y = \sum_{i=1}^{n} X_i$ and of $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$.

Recall that the m.g.f of exponential (θ) r.vs is $\boxed{M(t) = (1-\theta t)^{-1}, \quad t < \frac{1}{\theta}}$



$X_1 + X_2 + \dots + X_n$ should be gamma (n, θ)

Previous corollary: $M_Y(t) = [M(t)]^n = [(1-\theta t)^{-1}]^n = \underbrace{(1-\theta t)^{-n}, \quad t < \frac{1}{\theta}}_{\text{m.g.f of gamma } (\alpha = n, \theta)}$

$\Rightarrow Y \sim Gamma(n, \theta)$

For $\bar{X}$ use the 2nd item of the corollary :

$$M_{\bar{X}}(t) = \left[M\left(\frac{t}{n}\right) \right]^{n} = \left[\left(1 - \frac{\theta t}{n} \right)^{-1} \right]^{n} , \quad \frac{t}{n} < \frac{1}{\theta}$$

$$= \underbrace{\left(1 - \frac{\theta}{n} t \right)^{-n} , \quad t < \frac{n}{\theta} = \frac{1}{\theta/n}}_{\text{m.g.f of Gamma}\left(\underset{\alpha}{n},\ \frac{\theta}{n}\right)}$$

CONCLUSION : $\bar{X} \sim \text{Gamma}\left(n, \frac{\theta}{n}\right)$

Similar statements: (try to prove as HW)

① If $X_i \sim \text{Binomial}(m_i, p)$ are independent r.v.s $i = 1, \dots, N$

then $Y = \sum_{i=1}^{N} X_i \sim \text{Binomial}(m_1 + m_2 + \dots + m_N, p)$

② If $X_i \sim \text{Poisson}(\lambda_i)$ are independent r.v.s $i = 1, \dots, N$

then $Y = \sum_{i=1}^{N} X_i \sim \text{Poisson}(\lambda_1 + \lambda_2 + \dots + \lambda_N)$

③ If $X_i \sim \text{Geometric}(p)$ are independent r.v.s $i = 1, \dots, N$

$1 + 1 + 1 + \dots + 1$ N times

then $Y = \sum_{i=1}^{N} X_i \sim \text{Negative Binomial}(N, p)$

④ If $X_i \sim \text{Negative Binomial}(m_i, p)$ independent r.v.s $i = 1, \dots, N$

then $Y = \sum_{i=1}^{N} X_i \sim \text{Negative Binomial}(m_1 + m_2 + \dots + m_N, p)$

Theorem (a 1st explanation for the parameter of the χ^2 distribution)

Let $X_1, X_2, \ldots, X_n$ be independent chi-square random variables with $r_1, r_2, \ldots, r_n$ degrees of freedom, respectively.

Then

$$Y = \sum_{i=1}^{n} X_i = X_1 + X_2 + \cdots + X_n$$

has a $\chi^2(r_1 + r_2 + \cdots + r_n)$ distribution.

Proof (one more example of m.g.f technique)

How to specify $\chi^2(r)$ from its m.g.f

Recall that a $\chi^2(r)$ distrib. is a gamma distr. with $\alpha = \frac{r}{2}$, $r \in \mathbb{N}$, and $\theta = 2$

and so, the mgf of $\chi^2(r)$ is $M(t) = (1-2t)^{-\frac{r}{2}}$, $t < \frac{1}{2}$

To find the distn. of $Y = \sum_{i=1}^{m} X_i$, we compute the m.g.f of Y:

$$\alpha = \frac{\sum_{i=1}^{m} x_i}{2}$$

<u>Option 1:</u> Use the theorem proved earlier to get

$$M_Y(t) = \prod_{i=1}^{m} \underset{\substack{\uparrow \\ a_1 = a_2 = \cdots = a_m = 1}}{M_{X_i}(t)} = \prod_{i=1}^{m} \underbrace{(1-2t)^{-x_i/2}}_{\substack{\uparrow \\ \text{m.g.f of } \chi^2(x_i) \text{ as in} \\ \text{the previous slide}}} = (1-2t)^{-\left(\sum_{i=1}^{m} x_i\right)/2}$$

m.g.f of $\chi^2\left(\sum_{i=1}^{m} x_i\right)$

$$\Rightarrow Y \sim \chi^2(x_1 + x_2 + \cdots + x_m)$$

worth remembering

Option 2: If we do not remember the theorem statement!!!

Since $X_i \sim \chi^2(r_i)$, then $M_{X_i}(t) = (1-2t)^{-r_i/2}$, $t < \frac{1}{2}$

then for $Y = \sum_{i=1}^{m} X_i$ we have

$$M_Y(t) \underset{\text{def}}{=} E\left[e^{tY}\right] = E\left[e^{t\sum_{i=1}^{m} X_i}\right] \overset{\text{expand sum + distribute}}{=} E\left[e^{tX_1 + tX_2 + \cdots + tX_m}\right]$$

$$\underset{\text{exponentn}}{=} E\left[e^{tX_1} \cdot e^{tX_2} \cdot \ \cdots \ \cdot e^{tX_m}\right] \underset{\text{independence}}{=} \underbrace{E\left[e^{tX_1}\right]}_{M_{X_1}(t)} \cdot \underbrace{E\left[e^{tX_2}\right]}_{M_{X_2}(t)} \cdots \underbrace{E\left[e^{tX_m}\right]}_{M_{X_m}(t)}$$

$$= M_{X_1}(t) \cdot M_{X_2}(t) \cdots M_{X_m}(t) = (1-2t)^{-r_1/2} \cdot (1-2t)^{-r_2/2} \cdots (1-2t)^{-r_m/2}$$

$$= (1-2t)^{-(r_1+r_2+\cdots+r_m)/2} \Rightarrow Y \sim \chi^2(r_1+r_2+\cdots+r_m)$$

Corollary (IMPORTANT!)

Let $Z_1, Z_2, \ldots, Z_n$ be independent random variables with $N(0,1)$ distribution.

Then

$$W = \sum_{i=1}^{n} (Z_i)^2 = Z_1^2 + Z_2^2 + \cdots + Z_n^2$$

has a $\chi^2(n)$ distribution.

Proof: Last class: if $Z_i \sim N(0,1) \Rightarrow Z_i^2 \sim \chi^2(1)$

$Z_1, Z_2, \ldots, Z_n$ independent $\Rightarrow Z_1^2, Z_2^2, \ldots, Z_n^2$ independent

$\Rightarrow$ previous theorem $\quad W = \sum_{i=1}^{n} \underbrace{Z_i^2}_{\chi^2(1)} \sim \chi^2(n)$

Corollary (IMPORTANT!)

Let $X_1, X_2, \dots, X_n$ be independent random variables with $N(\mu_i, \sigma_i^2)$ distributions, $i = 1, 2, \dots, n$.

Then

$$W = \sum_{i=1}^{n} \left(\frac{X_i - \mu_i}{\sigma_i} \right)^2$$

has a $\chi^2(n)$ distribution.

Proof : $X_1, X_2, \dots, X_n \sim N(\mu_i, \sigma_i^2)$ independent r.v.s $\Rightarrow$

$\Rightarrow Z_i = \frac{X_i - \mu_i}{\sigma_i} \sim N(0,1)$ independent r.v.s

$\Rightarrow Z_i^2 \sim \chi^2(1)$ independent r.v.s

$\Rightarrow W = \sum_{i=1}^{n} Z_i^2 = \sum_{i=1}^{n} \left(\frac{X_i - \mu_i}{\sigma_i} \right)^2 \sim \chi^2(n)$