Lecture 8

MS Theory I

Review

§ 2.3 Moments and Moment Generating Functions

Definition 2.3.1 For each integer n, the nth *moment of* X (or $F_X(x)$), μ'_n, is

$$\mu'_n = \mathrm{E}\, X^n.$$

The nth *central moment of* X, μ_n, is

$$\mu_n = \mathrm{E}(X-\mu)^n,$$

where $\mu = \mu'_1 = \mathrm{E}\, X$.

$X \sim F_X(x)$

Moments	Central Moment
$\mu'_1 = E[X] = \mu$	$\mu_1 = E[X-\mu] = E[X-E[X]] = 0$
$\mu'_2 = E[X^2]$	$\mu_2 = E[(X-\mu)^2] = Var(X)$
$\vdots$	$\vdots$
$E[X] = \text{mean} = \mu$	$E[(X-\mu)^2] = Var(X)$

Definition 2.3.2 The *variance* of a random variable X is its second central moment, $\mathrm{Var}\, X = \mathrm{E}(X - \mathrm{E}\, X)^2$. The positive square root of $\mathrm{Var}\, X$ is the *standard deviation* of X.

Theorem 2.3.4 *If X is a random variable with finite variance, then for any constants a and b,*

$$\mathrm{Var}(aX+b) = a^2\, \mathrm{Var}\, X.$$

Easier Calculation

$$Var(X) = E[(X-E[X])^2] = E[X^2] - (E[X])^2$$

Since: $Var(X) = E[(X-E(X))^2]$

$$= E[X^2 - 2X\,E[X] + (E[X])^2]$$

$$= E[X^2] - 2(E[X])^2 + (E[X])^2$$

Note $E[X]$ is a constant

If b is constant $E[b] = b$

$$Var(X) = E[X^2] - (E[X])^2$$

Moment Generating Functions

- Generate moments
- Characterize a dist'n (used to show convergence)

Definition 2.3.6 Let X be a random variable with cdf F_X. The *moment generating function (mgf) of* X (or F_X), denoted by $M_X(t)$, is

$$M_X(t) = \mathrm{E}\, e^{tX},$$

provided that the expectation exists for t in some neighborhood of 0. That is, there is an $h > 0$ such that, for all t in $-h < t < h$, $\mathrm{E}e^{tX}$ exists. If the expectation does not exist in a neighborhood of 0, we say that the moment generating function does not exist.

More explicitly, we can write the mgf of X as

$$M_X(t) = \int_{-\infty}^{\infty} e^{tx} f_X(x)\,dx \qquad \text{if } X \text{ is continuous,}$$

or

$$M_X(t) = \sum_x e^{tx} P(X = x) \qquad \text{if } X \text{ is discrete.}$$

It is very easy to see how the mgf generates moments. We summarize the result in the following theorem.

Theorem 2.3.7 *If X has mgf $M_X(t)$, then*

$$\mathrm{E}\, X^n = M_X^{(n)}(0),$$

← How to get moments from mgf. (derivatives easier than integrals).

where we define

$$M_X^{(n)}(0) = \left.\frac{d^n}{dt^n} M_X(t)\right|_{t=0}.$$

That is, the nth moment is equal to the nth derivative of $M_X(t)$ evaluated at $t = 0$.

Review

Recall: Gamma distribution (continuous)

Example 2.3.8 (Gamma Mgf).

$$f(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta} * I_{(0,\infty)}(x) \qquad I_{(0,\infty)}(x) = \begin{cases} 1 & 0<x<\infty \\ 0 & \text{else} \end{cases} \qquad \begin{matrix} \alpha>0 \\ \beta>0 \end{matrix}$$

Aside: Since $f(x)$ is a pdf, we can assume that for any $a>0, b>0$

$$\int_0^\infty \frac{1}{\Gamma(a) b^a} x^{a-1} e^{-x/b} dx = 1 \qquad \text{(details later)}$$

$$M_X(t) = E[e^{tX}] = \int_0^\infty \frac{1}{\Gamma(\alpha)\beta^\alpha} e^{tx} x^{\alpha-1} e^{-x/\beta} dx$$

known constants

$$= \frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^\infty x^{\alpha-1} e^{tx} e^{-x/\beta} dx$$

Integrating over x
- t assumed fixe
- $M_X(t)$ will not be a ft'n of x.

$$= \frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^\infty x^{\alpha-1} e^{-x(1/\beta - t)} dx$$

$$= \frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^\infty x^{\alpha-1} e^{-x\cdot\left[\frac{1}{\beta} - \frac{t\beta}{\beta}\right]} dx$$

$$= \frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^\infty x^{\alpha-1} e^{-x/\left[\frac{\beta}{1-\beta t}\right]} dx$$

Define $b = \frac{\beta}{1-\beta t}$

$$= \frac{1}{\Gamma(\alpha)\beta^\alpha} * \frac{b^\alpha}{b^\alpha} * \int_0^\infty x^{\alpha-1} e^{-x/b} dx$$

$$= \left(\frac{b}{\beta}\right)^\alpha \underbrace{\int_0^\infty \frac{1}{\Gamma(\alpha) b^\alpha} x^{\alpha-1} e^{-x/b} dx}_{1 \text{ (since pdf)}}$$

assumes $b>0$
$\Rightarrow (1-\beta t)>0$
$0<\beta t<1$
$\Rightarrow t<\frac{1}{\beta}$

$$= \left(\frac{b}{\beta}\right)^\alpha = \left(\frac{\beta}{1-\beta t} * \frac{1}{\beta}\right)^\alpha = \boxed{\left(\frac{1}{1-\beta t}\right)^\alpha} \quad \Big\{ \text{ for } t<\frac{1}{\beta}$$

$$M_X(t) = \left(\frac{1}{1-\beta t}\right)^\alpha$$

$$E[X] = \frac{d}{dt}\left(\frac{1}{1-\beta t}\right)^\alpha \Bigg|_{t=0} = \frac{d}{dt}(1-\beta t)^{-\alpha} = -\alpha(1-\beta t)^{-\alpha-1}(-\beta)$$

$$= \frac{\alpha\beta}{(1-\beta t)^{\alpha+1}}\Bigg|_{t=0} = \alpha\beta$$

Review

Theorem 2.3.11 *Let $F_X(x)$ and $F_Y(y)$ be two cdfs all of whose moments exist.*

a. *If X and Y have bounded support, then $F_X(u) = F_Y(u)$ for all u if and only if $\mathrm{E}\,X^r = \mathrm{E}\,Y^r$ for all integers $r = 0, 1, 2, \ldots$.*

b. *If the moment generating functions exist and $M_X(t) = M_Y(t)$ for all t in some neighborhood of 0, then $F_X(u) = F_Y(u)$ for all u.*

$M_X(t)$ exists $\forall t$ $-h<t<h$ $h>0$ ← many dist'ns

$$F_X \text{ (or } f_X) \longrightarrow M_X(t) \longrightarrow \mu_r' = E[X^r] \quad r=0,1,2,\ldots$$

bounded support ← fewer dist'ns

a) If X and Y bounded support:

$F_X(u) = F_Y(u)\ \forall u \overset{\text{iff}}{\Longleftrightarrow} E[X^r] = E[Y^r]$ for all $r = 0,1,2,3,\ldots$

b) If mgfs exist $\forall t$ $-h<t<h$ $h>0$

$M_X(t) = M_Y(t) \Longrightarrow F_X(u) = F_Y(u)\ \forall u.$

Finally

Theorem 2.3.12 (Convergence of mgfs) *Suppose $\{X_i,\ i = 1, 2, \ldots\}$ is a sequence of random variables, each with mgf $M_{X_i}(t)$. Furthermore, suppose that*

$$\lim_{i\to\infty} M_{X_i}(t) = M_X(t), \qquad \text{for all } t \text{ in a neighborhood of } 0,$$

and $M_X(t)$ is an mgf. Then there is a unique cdf F_X whose moments are determined by $M_X(t)$ and, for all x where $F_X(x)$ is continuous, we have

$$\lim_{i\to\infty} F_{X_i}(x) = F_X(x).$$

That is, convergence, for $|t| < h$, of mgfs to an mgf implies convergence of cdfs.

example sequence of R.V.'s

$\bar{X}_1 = X_1/1$
$\bar{X}_2 = (X_1+X_2)/2$
$\bar{X}_3 = (X_1+X_2+X_3)/3$
$\vdots$
$\bar{X}_n = \sum_{i=1}^{n} X_i/n$

each $\bar{X}_i$ has own mgf $M_{X_i}(t)$

and cdf $F_{X_i}(t)$

Summary of Thm 2.3.12

$$\lim_{i\to\infty} M_{X_i}(t) = M_X(t) \quad \forall |t|<h \quad \overset{\text{implies}}{\Longrightarrow} \quad \lim_{i\to\infty} F_{X_i}(x) = F_X(x)$$

mgfs converge $\Rightarrow$ cdfs converge

Example: For n large and np small show

Binomial mgf $\rightarrow$ Poisson mgf

Aside: Some Tools needed for this problem

Lemma 2.3.14 *Let $a_1, a_2, \ldots$ be a sequence of numbers converging to a, that is, $\lim_{n\to\infty} a_n = a$. Then*

$$\lim_{n\to\infty}\left(1+\frac{a_n}{n}\right)^n = e^a.$$

We'll use this lemma a few times this year.

I will remind you when we need it

proof of CLT.

Poisson(λ)

pmf	$P(X=x\|\lambda) = \frac{e^{-\lambda}\lambda^x}{x!};\quad x=0,1,\ldots;\quad 0\le\lambda<\infty$
mean and variance	$\mathrm{E}X=\lambda, \quad \mathrm{Var}\,X=\lambda$
mgf	$\boxed{M_X(t)=e^{\lambda(e^t-1)}}$

← Easier to calculate than binomial(n,p).

The approximation: $P(X=x) \approx P(Y=x)$

for $X \sim \text{bin}(n,p)$, $Y \sim \text{Pois}(\lambda)$, $\lambda = np$ — n 'large', np small

One rule: $n \geq 20$, $p \leq .05$, $np < 10$

Poisson (λ) pmf: $P(Y=y) = \frac{e^{-\lambda}\lambda}{y!} * I_{[0,1,2,\dots)}(y)$

Recall $\underset{\text{Binomial}}{M_X(t)} = [pe^t + (1-p)]^n$ and $\underset{\text{Poisson}}{M_Y(t)} = e^{\lambda(e^t-1)}$

If $\lambda = np$ then $p = \lambda/n$

$$M_X(t) = [pe^t + (1-p)]^n = [1 + \tfrac{1}{n}(e^t-1)(np)]^n$$

if $p = \frac{\lambda}{n}$ $\quad = [1 + \frac{1}{n}(e^t-1)\lambda]^n$

[Now use Lemma 2.3.14

define:
$a = a_n = (e^t-1)\lambda$

Then $\lim_{n\to\infty} M_X(t) = \boxed{e^{\lambda(e^t-1)}} = M_Y(t)$

$\therefore \lim_{n\to\infty} F_X(u) = F_Y(u) \quad u = \{0,1,2,\dots\}$

Note: we must recognize $e^{\lambda(e^t-1)}$ is Mgf of Poisson for this to be useful Thm.

Finally:

Theorem 2.3.15 *For any constants a and b, the mgf of the random variable $aX+b$ is given by*

$$M_{aX+b}(t) = e^{bt}M_X(at).$$

Proof:

$$\begin{aligned}
M_{(aX+b)}(t) &= E[e^{(aX+b)t}] && \text{a, b constants}\\
&= E[e^{(aX)t}e^{bt}] && \text{properties of exponents}\\
&= e^{bt}E[e^{(aX)t}] && \text{assuming } t \text{ fixed: } e^{bt} \text{ constant}\\
&= e^{bt}M_X(at).\\
&= e^{bt}\int_{-\infty}^{\infty} e^{(at)x} f_X(x)\,dx && \text{continuous dist'n}
\end{aligned}$$

or

$$e^{bt}\sum_x e^{(at)x} P(X=x) \qquad \text{discrete dist'n}$$

§2.4 Differentiating under integral sign

- We will frequently need to interchange intergration & differentiation.

- Need to know when legitimate to interchange $\int$ & $\frac{d}{d\theta}$

- Assume interested in calculating:

$$\frac{d}{d\theta}\int_{a(\theta)}^{b(\theta)} f(x,\theta)\,dx \qquad -\infty < a(\theta),\ b(\theta) < \infty$$

- Use application of Fundamental Thm of Calculus and chain rule.

Theorem 2.4.1 (Leibnitz's Rule) *If $f(x,\theta)$, $a(\theta)$, and $b(\theta)$ are differentiable with respect to θ, then*

$$\frac{d}{d\theta}\int_{a(\theta)}^{b(\theta)} f(x,\theta)\,dx = f(b(\theta),\theta)\frac{d}{d\theta}b(\theta) - f(a(\theta),\theta)\frac{d}{d\theta}a(\theta) + \int_{a(\theta)}^{b(\theta)} \frac{\partial}{\partial\theta} f(x,\theta)\,dx.$$

Notice that if $a(\theta)$ and $b(\theta)$ are constant, we have a special case of Leibnitz's Rule:

partial derivative

$$\underbrace{\frac{d}{d\theta}\int_a^b f(x,\theta)\,dx}_{\text{ft'n of }\theta} = \int_a^b \underbrace{\frac{\partial}{\partial\theta} f(x,\theta)\,dx}_{\text{ft'n }x,\theta}.$$

Example Beta $(\alpha,1)$

Beta(α,β)	
pdf	$f(x\mid\alpha,\beta) = \frac{1}{B(\alpha,\beta)}x^{\alpha-1}(1-x)^{\beta-1}, \quad 0\le x\le 1, \quad \alpha>0, \quad \beta>0$
mean and variance	$\mathrm{E}X = \frac{\alpha}{\alpha+\beta}, \quad \mathrm{Var}\,X = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$
mgf	$M_X(t) = 1 + \sum_{k=1}^{\infty}\left(\prod_{r=0}^{k-1}\frac{\alpha+r}{\alpha+\beta+r}\right)\frac{t^k}{k!}$
notes	The constant in the beta pdf can be defined in terms of gamma functions, $B(\alpha,\beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$. Equation (3.2.18) gives a general expression for the moments.

$$f(x\mid\alpha,\beta{=}1) = \frac{\Gamma(\alpha+1)}{\Gamma(\alpha)\Gamma(1)}x^{\alpha-1}(1-x)^{\beta-1} * I_{[0,1]}(x)$$

$$= \frac{\alpha\Gamma(\alpha)}{\Gamma(\alpha)0!}x^{\alpha-1} * I_{[0,1)}(x) = \alpha x^{\alpha-1} I_{[0,1]}(x)$$

$$f(x|\alpha,\beta=1) = \alpha x^{\alpha-1} I_{[0,1]}(x)$$

want $\dfrac{d}{d\alpha}\displaystyle\int_0^1 g(x,\alpha)\,\alpha x^{\alpha-1}\,dx$

$g(x,\alpha)$ often something like $\log f(x|\alpha)$... (more later).

here $a(\theta)=1$ and $b(\theta)=0$ (both constant)!

so $$\frac{d}{d\alpha}\int_0^1 g(x,\alpha)\,\alpha x^{\alpha-1}dx = \int_0^1 \frac{\partial}{\partial\alpha}\left[g(x,\alpha)\,\alpha x^{\alpha-1}\right]dx$$

Example $U(0,\theta)$

Caution:
Sample space is ft'n of parameters that define dist'n!

Uniform(a,b)

pdf	$f(x\|a,b)=\frac{1}{b-a},\quad a\le x\le b$
mean and variance	$\mathrm{E}X=\frac{b+a}{2},\quad \mathrm{Var}\,X=\frac{(b-a)^2}{12}$
mgf	$M_X(t)=\frac{e^{bt}-e^{at}}{(b-a)t}$
notes	If $a=0$ and $b=1$, this is a special case of the beta ($\alpha=\beta=1$).

$$f(x) = \frac{1}{\theta} I_{[0,\theta]}(x) \qquad a(\theta)=0,\quad b(\theta)=\theta$$

$$\frac{d}{d\theta}\int_0^\theta g(x,\theta)\frac{1}{\theta}\,dx \neq \int_0^\theta\left[\frac{\partial}{\partial\theta} g(x,\theta)\frac{1}{\theta}\right]dx$$

$$\frac{d}{d\theta}\int_0^\theta g(x,\theta)\frac{1}{\theta}\,dx = \frac{g(\theta,\theta)}{\theta}\frac{d}{d\theta}\theta - \frac{g(0,\theta)}{\theta}\frac{d}{d\theta}(0) + \int_0^\theta \frac{\partial}{\partial\theta}\frac{g(x,\theta)}{\theta}\,dx$$

More theorems in C&B to cover when $0<x<\infty$ or $-\infty<x<\infty$
$a(\theta)$"="$-\infty$ or $b(\theta)$"="∞

Can we assume:

$$\frac{d}{d\theta}\int_{-\infty}^{\infty} f(x,\theta)\,dx = \int_{-\infty}^{\infty} \frac{\partial}{\partial\theta} f(x,\theta)\,dx \ ?$$

The property illustrated ~~for the exponential distribution~~ above holds for a large class of densities, which will be dealt with in Section 3.4.

'exponential families'
"obviously" the sample space/support won't be a ft'n of θ.

Similarly for 'exponential families' we can assume that we can interchange differentiation and summation.
$\frac{d}{d\theta}$ and $\sum_{i=1}^{\infty}$ or $\sum_{i=1}^{n}$.

More in chapter-3

3 Common Families of Distributions