MS Theory I

Lecture - 7

Review
§ 2.2 Expected Values (average value from dist'n)

Definition 2.2.1 The *expected value* or *mean* of a random variable $g(X)$, denoted by $\mathrm{E}\,g(X)$, is

$$\mathrm{E}\,g(X) = \begin{cases} \int_{-\infty}^{\infty} g(x) f_X(x)\,dx & \text{if } X \text{ is continuous} \\ \sum_{x\in\mathcal{X}} g(x) f_X(x) = \sum_{x\in\mathcal{X}} g(x) P(X=x) & \text{if } X \text{ is discrete,} \end{cases}$$

provided that the integral or sum exists. If $\mathrm{E}|g(X)| = \infty$, we say that $\mathrm{E}\,g(X)$ does not exist. (Ross 1988 refers to this as the $\underbrace{\text{“law of the unconscious statistician.”}}_{\text{LOTUS - see appendix. lecture-5}}$ We do not find this amusing.)

$$E[X] = \begin{cases} \int_{-\infty}^{\infty} x\, f_X(x)\,dx & X \text{ is continuous} \\ \sum_{x\in\mathcal{X}} x\, f_X(x)\,dx & X \text{ is discrete} \end{cases} \Bigg\} \text{ weighted average over values of } X$$

- Note this will be a function of dist'n parameters not a function of x.

Example 2.2.2 (Exponential mean)

$X \sim \text{exponential}(\lambda)$ $\qquad f_X(x) = \frac{1}{\lambda} e^{-x/\lambda} I_{[0,\infty)}(x)$, $\lambda \geq 0$

$$E[X] = \int_0^\infty x \frac{e^{-x/\lambda}}{\lambda}\,dx$$

$$= -xe^{-x/\lambda}\Big|_0^\infty + \int_0^\infty e^{-x/\lambda}\,dx$$

$$= 0 + \int_0^\infty e^{-x/\lambda}$$

$$= -\lambda e^{-x/\lambda}\Big|_0^\infty = 0 - [-\lambda] = \lambda$$

Integration by parts:
$\int_a^b u\,dv = uv\big|_a^b - \int_a^b v\,du$

$u = x \qquad dv = e^{-x/\lambda}/\lambda$
$du = dx \qquad v = -e^{-x/\lambda}$

change of variables $\quad u = -x/\lambda$
$\qquad dx = -\lambda du$

Example 2.2.3 (Binomial mean)

$X \sim \text{bin}(n,p) \qquad P(X=x) = \binom{n}{x} p^x (1-p)^{n-x} * I_{[0,1,\dots,n]}(x)$

$$E[X] = \sum_{x=0}^{n} x \binom{n}{x} p^x (1-p)^{n-x} = \sum_{x=1}^{n} x \binom{n}{x} p^x (1-p)^{n-x}$$

$$= \sum_{x=1}^{n} \underbrace{n\binom{n-1}{x-1}}_{\text{appendix}} p^x (1-p)^{n-x} = \sum_{y=0}^{n-1} n \binom{n-1}{y} p^{y+1} (1-p)^{n-(y+1)}$$

$$= np \underbrace{\sum_{y=0}^{n-1} \binom{n-1}{y} p^y (1-p)^{n-1-y}}_{\text{bin}(n-1,p) \text{ must sum to } 1 \text{ (details later...)}} = np$$

change of variables
$y = x-1 \rightarrow x = y+1$
$x=1 \rightarrow y=0$;
$x=n \rightarrow y=n-1$

Review

Example 2.2.4 (Cauchy mean) → Classic Example
E[X] does not exist.

Cauchy(θ, σ)	
pdf	$f(x\|\theta,\sigma) = \frac{1}{\pi\sigma}\frac{1}{1+\left(\frac{x-\theta}{\sigma}\right)^2}$, $-\infty < x < \infty$; $-\infty < \theta < \infty$, $\sigma > 0$
mean and variance	do not exist
mgf	does not exist
notes	Special case of Student's t, when degrees of freedom $= 1$. Also, if X and Y are independent n$(0,1)$, X/Y is Cauchy.

Cauchy $(\sigma = 1)$

$$f_X(x) = \frac{1}{\pi}\frac{1}{1+x^2} * I_{(-\infty,\infty)}(x)$$

$$E[|X|] = \int_{-\infty}^{\infty} \frac{|x|}{\pi}\frac{1}{1+x^2}\,dx = \frac{2}{\pi}\int_0^{\infty}\frac{x}{1+x^2}\,dx = \frac{\cancel{2}}{\pi}\frac{\log(1+x^2)}{\cancel{2}}\Big|_0^{\infty}$$

$$= \lim_{M\to\infty}\frac{\log(1+M^2)}{\pi} = \infty$$

E[X] does not exist for the Cauchy dist'n.

Properties of Expectation

Theorem 2.2.5 *Let X be a random variable and let $a, b,$ and c be constants. Then for any functions $g_1(x)$ and $g_2(x)$ whose expectations exist,*

Proof in C&B pg. 57

a. $E(ag_1(X) + bg_2(X) + c) = aE\,g_1(X) + bE\,g_2(X) + c$.

b. *If $g_1(x) \geq 0$ for all x, then* $E\,g_1(X) \geq 0$.

c. *If $g_1(x) \geq g_2(x)$ for all x, then* $E\,g_1(X) \geq E\,g_2(X)$.

d. *If $a \leq g_1(x) \leq b$ for all x, then $a \leq E\,g_1(X) \leq b$.*

$$E[aX+b] = aE[X] + b \;; \quad a \,\&\, b \text{ constants}$$

$$E[X - E[X]] = 0$$

example $X \sim \text{bin}(n,p)$

$$E[X] = np$$

$$E[X - np] = E[X] - np = 0$$

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
	$\mu_1' = E[X] = \mu$	$\mu_1 = E[X-\mu] = E[X-E[X]] = 0$
	$\mu_2' = E[X^2]$	$\mu_2 = E[(X-\mu)^2] = \mathrm{Var}(X)$
	$\vdots$	$\vdots$
	$E[X] = \text{mean} = \mu$	$E[(X-\mu)^2] = \mathrm{Var}(X)$

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.

- Variance measure of spread
- Large Variance $\Rightarrow$ X more variable
- $\mathrm{Var}(X) = 0$ $\quad E[(X-E[X])^2] = 0$
  $X = E[X]$ with prob. 1 (no variation in X)
- Small $\mathrm{Var}(X) \Rightarrow X$ less variable and close to $E[X] = \text{mean}$
- If X has units: age, height, ... $\quad \mathrm{Var}(X)$: Units^2
- Standard Deviation $= \sqrt{\mathrm{Var}(X)}$ (same units as X)

Example 2.3.3 (Exponential variance)

$X \sim \text{exponential}(\lambda)$

$$E[X] = \lambda$$

$$f_X(x) = \frac{e^{-x/\lambda}}{\lambda} * I_{(0,\infty)}(x)$$

$$\mathrm{Var}(X) = E[(X-\lambda)^2] = \int_0^\infty (x-\lambda)^2 \frac{1}{\lambda} e^{-x/\lambda}\, dx$$

$$= \int_0^\infty (x^2 - 2x\lambda + \lambda^2) \frac{1}{\lambda} e^{-x/\lambda}\, dx$$

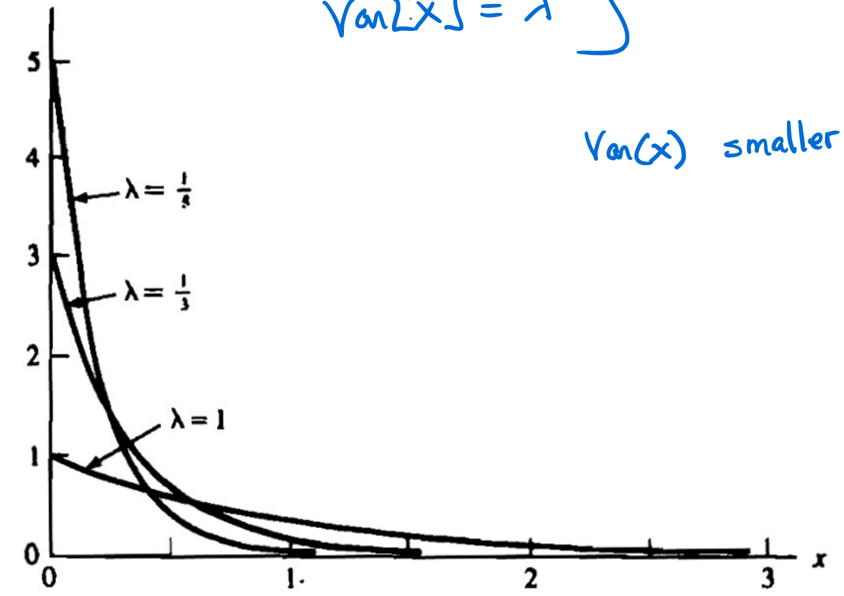
$$= \underbrace{\int_0^\infty \frac{x^2}{\lambda} e^{-x/\lambda} dx}_{2\lambda^2} \underbrace{- 2\lambda \int_0^\infty \frac{x}{\lambda} e^{-x/\lambda} dx}_{-2\lambda E[X] = -2\lambda^2} + \underbrace{\int_0^\infty \lambda e^{-x/\lambda} dx}_{\lambda^2}$$

see appendix (i) $\qquad\qquad$ see appendix (ii) Lecture - 6

$X \sim \text{exponential}(\lambda)$ $\quad f_X(x) = \frac{e^{-x/\lambda}}{\lambda} \cdot I_{[0,\infty)}(x)$; $\lambda > 0$

$E[X] = \lambda$
$Var[X] = \lambda^2$ } Function of parameter

$f_X(x)$



Var(x) smaller, smaller values λ.

Figure 2.3.1. *Exponential densities for* $\lambda = 1, \frac{1}{3}, \frac{1}{5}$

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

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

Proof: $Var(aX+b) = E[((aX+b) - E[(aX+b)])^2]$

$= E[((aX+\cancel{b}) - (aE[X] + \cancel{b}))^2]$

$= E[(a(X) - aE[X])^2]$

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

$= a^2 Var(X)$ //

Easier Calculation

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

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

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

Note $E[X]$ is a constant

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

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

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

Example 2.3.5 (Binomial Variance)

$X \sim \text{bin}(n,p)$ $\quad f_X(x) = P(X=x) = \binom{n}{x} p^x (1-p)^{n-x} \cdot I_{[0,1,2,\dots,n]}(x)$

Know $E[X] = np$ (page 3)

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

Calculate $E[X^2] = \sum_{x=0}^{n} x^2 \binom{n}{x} p^x (1-p)^{n-x}$

aside

$$x^2 \binom{n}{x} = x^{\cancel{2}} \frac{n(n-1)!}{\cancel{x}(x-1)!(n-x)!} = xn\binom{n-1}{x-1}$$

$$E[X^2] = n \sum_{x=1}^{n} x \binom{n-1}{x-1} p^x (1-p)^{n-x}$$

change of variables

$y = x-1 \quad x=1 \quad y=0$

$x = y+1 \quad x=n \quad y=n-1$

$$= n \sum_{y=0}^{n-1} (y+1) \binom{n-1}{y} p^{y+1} (1-p)^{n-1-y}$$

$$= np \underbrace{\sum_{y=0}^{n-1} y \binom{n-1}{y} p^y (1-p)^{n-1-y}}_{E[y],\ y \sim \text{bin}(n-1,p)} + np \underbrace{\sum_{y=0}^{n-1} \binom{n-1}{y} p^y (1-p)^{n-1-y}}_{=1 \text{ since pdf sum bin}(n-1,p)}$$

$$= np((n-1)p) + np$$

$$= n(n-1)p^2 + np$$

$$Var(X) = n(n-1)p^2 + np - (np)^2$$

$$= \cancel{n^2p^2} - np^2 + np - \cancel{n^2p^2}$$

$$= np(1-p).$$

Can continue to higher order moments

- sometime interested in 3rd or 4th order moments
- rarely higher order moments of interest.

Moment Generating Functions

- Genate Moments; many cases easier than directly.

→ Help Characterize Dist'ns

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*

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

where we define

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

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

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

Proof

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

$$E[X^n] = \frac{d^n}{dt^n} M_X(t)\Big|_{t=0}$$

assumes we can interchange $\int$ + $\frac{d}{dt}$. (details soon).

$$d/dt\, M_X(t) = \frac{d}{dt}\int_{-\infty}^{\infty} e^{tx} f_X(x)\,dx = \int_{-\infty}^{\infty}\left(\frac{d}{dt} e^{tx}\right) f_X(x)\,dx = \int_{-\infty}^{\infty}(x e^{tx}) f_X(x)\,dx$$

$$= E[X e^{tX}]$$

$$d/dt\, M_X(t)\Big|_{t=0} = E[X e^{tX}]\Big|_{t=0} = E[X]$$

Similarly $\frac{d^n}{dt^n} M_X(t)\Big|_{t=0} = E[X^n e^{tX}]\Big|_{t=0} = E[X^n]$ //

Recall: Gamma distribution (continuous)

Gamma(α, β)		
pdf	$f(x\mid\alpha,\beta) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1}e^{-x/\beta}$,	$0 \le x < \infty$, $\quad \alpha, \beta > 0$
mean and variance	$\mathrm{E}X = \alpha\beta$, $\quad \mathrm{Var}\, X = \alpha\beta^2$	
mgf	$M_X(t) = \left(\frac{1}{1-\beta t}\right)^\alpha$, $\quad t < \frac{1}{\beta}$	
notes	Some special cases are exponential ($\alpha = 1$) and chi squared ($\alpha = p/2$, $\beta = 2$). If $\alpha = \frac{3}{2}$, $Y = \sqrt{X/\beta}$ is *Maxwell*. $Y = 1/X$ has the *inverted gamma distribution*. Can also be related to the Poisson (Example 3.2.1).	

sample space (for $0 \le x < \infty$)

parameter space (for $\alpha, \beta > 0$)

$\Gamma(\alpha+1) = \alpha\Gamma(\alpha) \quad \alpha > 0$

$\Gamma(n) = (n-1)! \quad$ if $n = 1, 2, 3 \ldots$

$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t}\, dt$

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

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

Gamma mgf cont.

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

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

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

Example 2.3.9 (Binomial Mgf):

Binomial(n, p)

sample space; parameter space (n known)

pmf	$P(X = x\|n,p) = \binom{n}{x}p^x(1-p)^{n-x};$	$x = 0,1,2,\ldots,n; \quad 0 \le p \le 1$
mean and variance	$\mathrm{E}X = np, \quad \mathrm{Var}\,X = np(1-p)$	p= probability success 1-p = probability failure n bernoulli trials (often n known) $X \sim \text{bin}(n,p)$
mgf	$\boxed{M_X(t) = [pe^t + (1-p)]^n}$	
notes	Related to Binomial Theorem (Theorem 3.2.2). The *multinomial* distribution (Definition 4.6.2) is a multivariate version of the binomial distribution.	

Theorem 3.2.2 (Binomial Theorem) *For any real numbers x and y and integer $n \ge 0$,*

$$(x+y)^n = \sum_{i=0}^{n}\binom{n}{i}x^i y^{n-i}. \tag{3.2.4}$$

$$M_X(t) = \sum_{x=0}^{n} e^{tx}\binom{n}{x}p^x(1-p)^{n-x} = \sum_{x=0}^{n}\binom{n}{x}(pe^t)^x(1-p)^{n-x}$$

by binomial Thm: $= (pe^t + (1-p))^n$ //

Moment Generating Function
- generates moments
- Mg function also can characterize a dist'n.

Hmmm... Mgf $\longrightarrow$ infinite set of moments $\xrightarrow{???}$ dist'n

Moment Generating Function
- generates moments
- Mg function also can characterize a dist'n.

Hmmm... Mgf $\longrightarrow$ infinite set of moments $\xrightarrow{???}$ dist'n

Example (Nonunique Moments) lognormal dist'n.

$$f_1(x) = \frac{1}{\sqrt{2\pi}\, x} e^{-(\log(x))^2/2} I_{[0,\infty)}(x)$$

$$I_{[0,\infty)}(x) = \begin{cases} 1 & 0 \le x < \infty \\ 0 & \text{else} \end{cases}$$

$$f_2(x) = f_1(x)[1 + \sin(2\pi \log(x)) * I_{[0,\infty)}(x)$$

ICBST (it can be shown that) C&B example 2.3.10

- $E[X_1^r] = E[X_2^r]$
- 2 distinct dist'ns with same moments

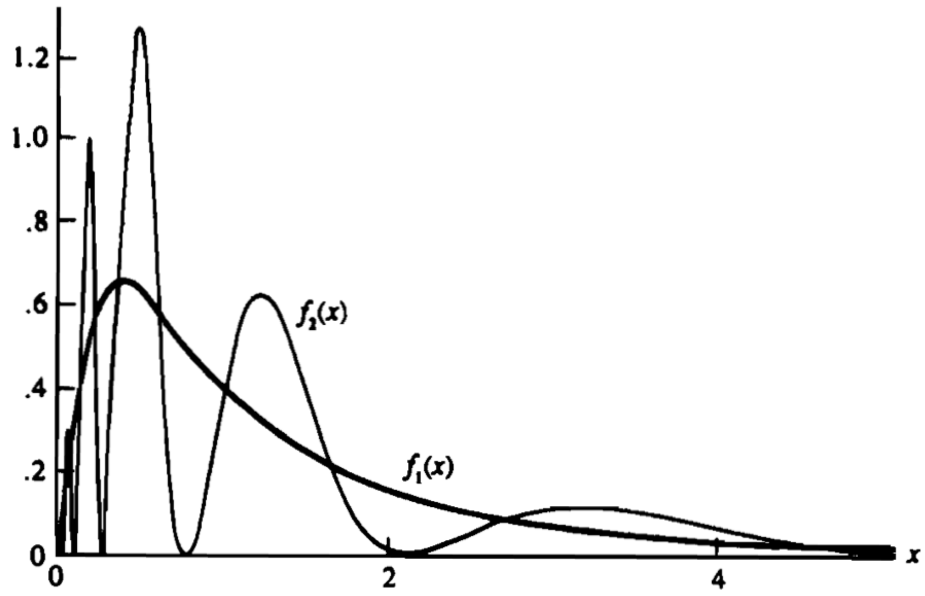


Figure 2.3.2. *Two pdfs with the same moments:* $f_1(x) = \frac{1}{\sqrt{2\pi}x} e^{-(\log x)^2/2}$ *and* $f_2(x) = f_1(x)[1 + \sin(2\pi \log x)]$

Mgf $\xrightarrow{\text{characterizes}}$ infinite set of moments $\overset{\text{may not characterize}}{\not\longrightarrow}$ dist'n

$$\text{Mgf} \xrightarrow{\text{characterizes}} \text{infinite set of moments} \overset{\text{may not characterize}}{\not\longrightarrow} \text{dist'n}$$

BUT

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.*

(a)

$$\text{Mgf} \xrightarrow{\text{characterizes}} \text{infinite set of moments} \xrightarrow[\text{if bounded support.}]{\text{characteizes}} \text{dist'n}$$

Bounded Support: (Sam's basic def'n):
Support: Set of x where $f_X(x) \neq 0$.
Bounded: Support of X is $[a,b]$; a, b finite

few dist'ns

(b) $\left.\begin{array}{l} F_X(x) \\ F_Y(y) \end{array}\right\rangle$ 2 cdfs; moments exist.

and $M_X(t) = M_Y(t)$ for all t in some neighborhood of 0.

then: $F_X(u) = F_Y(u) \quad \forall u.$

Note: Here mgf (not moments) defining cdfs.

many dist'ns

$$\Rightarrow M_X(t) = M_Y(t) \longrightarrow F_X(u) = F_Y(u) \quad \forall u.$$

$\left.\begin{array}{l} \{-h < t < h\} \\ h > 0 \end{array}\right\}$ for all t in neighborhood of 0.

Summary of thm 2.3.11

characterizes / characteizes

Mgf $\longrightarrow$ infinite set of moments $\dashrightarrow$ dist'n

if bounded support.

$\{-h<t<h\}$, $h>0$ — for all t in neighborhood of 0.

Character izes

Another very useful MGF Theorem: (Proof beyond C&B scope...)

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

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

mgfs converge $\Rightarrow$ cdfs converge

Example (Binomial mgf → Poisson mgf) (n large, np small)

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 $P(X=x|\lambda) = \frac{e^{-\lambda}\lambda^x}{x!}; \quad x = 0,1,\ldots; \quad 0 \le \lambda < \infty$ ← Easier to calculate than binomial (n,p).

mean and variance $\mathrm{E}X = \lambda, \quad \mathrm{Var}\,X = \lambda$

$\boxed{\textit{mgf} \quad M_X(t) = e^{\lambda(e^t-1)}}$

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

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

One rule: $n \ge 20$, $p \le .05$, $np < 10$

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

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 $p = \lambda/n$

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

if $p = \frac{\lambda}{n}$ $\quad = \left[1 + \tfrac{1}{n}(e^t-1)\lambda\right]^n$ $\quad$ Now use Lemma 2.3.14

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

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

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

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:

$$M_{(aX+b)}(t) = E\left[e^{(aX+b)t}\right]$$ a, b constants

$$= E\left[e^{(aX)t} e^{bt}\right]$$ properties of exponents

$$= e^{bt} E\left[e^{(aX)t}\right]$$ assuming t fixed: e^{bt} constant

$$= e^{bt} M_X(at).$$

$$= e^{bt} \int_{-\infty}^{\infty} e^{(at)x} f_X(x)\, dx$$ continuous dist'n

or

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