STAT 345/445 Lecture 15

Multiple Random Variables

Bivariate transformations– Section 4.3

Hierarchical Models and Mixture distributions – Section 4.4

Bivariate transformations

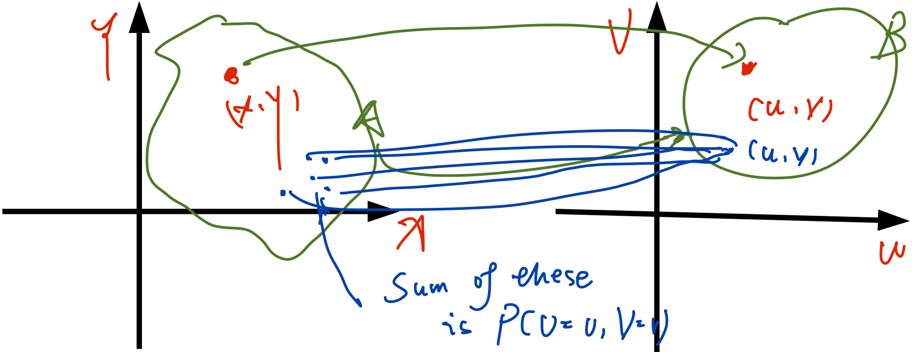
- Suppose (X, Y) is a bivariate random vector with join pdf/pmf $f_{X,Y}(x, y)$
- Have seen already: If X and Y are *independent* then

$$M_{X+Y}(t) = M_X(t)M_Y(t)$$

nice if mgf is recognisable.

- More generally: Let $U = g_1(X, Y)$ and $V = g_2(X, Y)$.
  - What is the joint distribution of (U, V) ?

$P((U, V) \in B) =$



$(X, Y) \longrightarrow (U, V)$

$$(U, V) = (g_1(x, y), g_2(x, y))$$
$$= g(x, y)$$

$$P((U, V) \in B) = P((X, Y) \in A)$$

where $A = \{(x, y) : (g_1(x, y), g_2(x, y)) \in B\}$

Discrete case

- $U = g_1(X, Y)$ and $V = g_2(X, Y)$

$f_{U,V}(u, v) = P(U = u, V = v) = P((X,Y) \in A_{uv}) = \sum_{A_{uv}} f_{XY}(x,y)$

Where $A_{uv} = \{(x,y) : u = g_1(x,y),\ v = g_2(x,y)\}$

I.e. just identify all the (x,y) that correspond to the (u,v) and add up the probabilities.

Example: Two dice

- We throw two dice and observe both $U =$ the sum and $V =$ the absolute difference

  $U = X + Y \quad V = |X - Y|$

- Want the distribution of (U, V) and the marginal distributions of U and V.

- First: Let X be the number of pips observed on die 1 and Y be the number of pips observed on die 2. What is $f(x, y)$?

Two die. Independent

$$f(x,y) = f_X(x) f_Y(y) = P(X=x) P(Y=y) = \frac{1}{6}\frac{1}{6}$$
$$= \frac{1}{36}$$

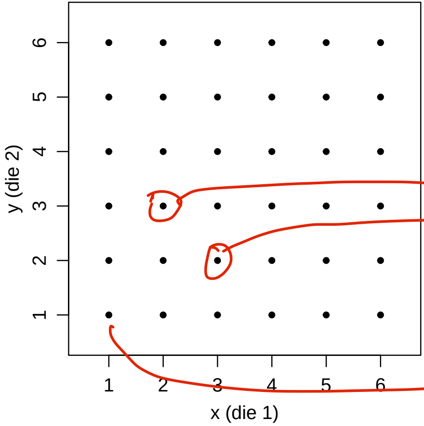
for $x = 1, \cdots 6, \; y = 1, \cdots 6$

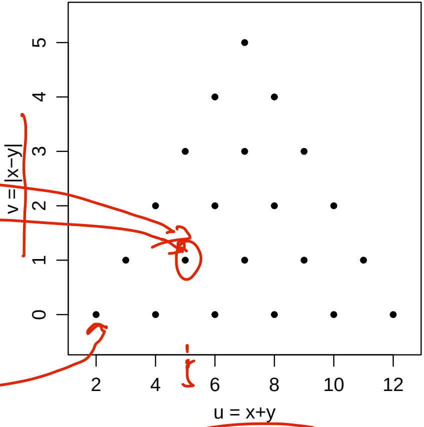
Example: Two dice

$g_1(x,y) = x+y \quad g_2(x,y) = |x-y|$

· : possible outcomes

$$(U, V) = (X + Y, \ |X - Y|)$$





Example: Two dice

- Possible outcomes of U: 2, 3, 4, ..., 12
- Possible outcomes of V: 0, 1, 2, 3, 4, 5

but not all combination e.g. U=3 and V=0

Grouping (X, Y) outcomes with corresponding (U, V) outcomes

U outcomes

2	3	4	5	6	7	8	9	10	11	12
(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)	(2,6)	(3,6)	(4,6)	(5,6)	(6,6)
	(2,1)	(3,1)	(4,1)	(5,1)	(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	
		(2,2)	(2,3)	(2,4)	(2,5)	(3,5)	(4,5)	(5,5)		
			(3,2)	(4,2)	(5,2)	(5,3)	(5,4)			
				(3,3)	(3,4)	(4,4)				
					(4,3)					

V=0 V=2 V=5

V=1 V=3

(x,y) all outcoms each has prob $\frac{1}{36}$

$f(u,v) = \frac{1}{36}$ for V=0 and U=2, 4, 6, 8, 10, 12

$$f(u=3, v=1) = \frac{2}{36}$$

$$f(u,v) = \frac{2}{36}$$ for $v=1$ and $u=3, 5, 7, 9, 11$

$v=2$, $u=4, 6, 8, 10$

$v=3$, $u=$

$v=4$, $u=$

$v=5$, $u=$

Example: Two dice

Determine the pmf of (U, V)

Example: Two dice

pmf of (U, V) in table form:

$f(u,v)$	u											
v	2	3	4	5	6	7	8	9	10	11	12	$f_V(v)$
0	$\frac{1}{36}$	0	$\frac{1}{36}$	0	$\frac{1}{36}$	0	$\frac{1}{36}$	0	$\frac{1}{36}$	0	$\frac{1}{36}$	$\frac{6}{36}$
1	0	$\frac{2}{36}$	0	$\frac{2}{36}$	0	$\frac{2}{36}$	0	$\frac{2}{36}$	0	$\frac{2}{36}$	0	$\frac{10}{36}$
2	0	0	$\frac{2}{36}$	0	$\frac{2}{36}$	0	$\frac{2}{36}$	0	$\frac{2}{36}$	0	0	$8/36$
3	0	0	0	$\frac{2}{36}$	0	$\frac{2}{36}$	0	$\frac{2}{36}$	0	0	0	$6/36$
4	0	0	0	0	$\frac{2}{36}$	0	$\frac{2}{36}$	0	0	0	0	$4/36$
5	0	0	0	0	0	$\frac{2}{36}$	0	0	0	0	0	$2/36$
$f_U(u)$	1	2	3	4	5	6	5	4	3	2	1	$\frac{1}{36}$

Example: Sum of Poisson's

- Let $X \sim \mathrm{Pois}(\lambda_1)$ and $Y \sim \mathrm{Pois}(\lambda_2)$, where X and Y are independent
- Want the distribution of $X + Y$. $(X, Y) \longmapsto (U, V)$
- General strategy to finding the distribution of $U = g(X, Y)$
  - Set $U = g(X, Y)$ and set $V =$ some simple function of X and Y that makes the bi-variate transformation one-to-one. E.g. $V = Y$
  - Find the joint pmf of U and V, $f_{U,V}(u, v)$
  - Then find the marginal pmf of U, $f_U(u) = \sum_v f_{u,v}(u, v)$

Sum of 2 independent Poisson

$X \sim \text{Poisson}(\lambda_1)$ $Y \sim \text{Poisson}(\lambda_2)$

Want distribution of $U = X + Y$

Side note: Saw in Lecture 14

that $X+Y \sim \text{Poisson}(\lambda_1 + \lambda_2)$ using mgf

Let's define $V = Y$

$$(X, Y) \longrightarrow (X+Y, Y)$$

First determine the join distribution of (X, Y)

$f(x,y) = f_X(x) f_Y(y)$ $\quad x = 0, 1, 2, \cdots$ and

$y = 0, 1, 2, \cdots$

Since X, Y are indep

$$= \frac{e^{-\lambda_1} \lambda_1^{x}}{x!} \frac{e^{-\lambda_2} \lambda_2^{y}}{y!}$$

$u = \overbrace{x+y}^{g_1(x,y)}$

$A_{uv} = \{(x,y) ; u = g_1(x,y), v = g_2(x,y)\}$ $\quad v = y = g_2(x,y)$

$= \{(\overbrace{u-v}^{x}, \overbrace{v}^{y})\}$ Just one value

Solve for x and y

$y = v$ and $x = u - y = u - v$

$$f_{u,v}(u,v) = f_{x,y}(u-v, v) = \frac{e^{-\lambda_1}\lambda_1^{u-v}}{(u-v)!} \frac{e^{-\lambda_2}\lambda_2^{v}}{v!}$$

Support? $\longrightarrow$ For what values of u and v?

$u = 0, 1, 2 \ldots$

$v = 0, 1, 2 \ldots$

exp: $u = 0 \quad v = 3$

Note that $y \le x+y$, so $V \le U$

e.g. $U = 0, 1, 2 \ldots$

$V = 0, 1, 2, \ldots U$

To find $f_U(u) \longrightarrow f_U(u) = \sum_{\text{all } v} f_{U,V}(U, V)$

$$= \sum_{v=0}^{u} \frac{e^{-(\lambda_1+\lambda_2)} \lambda_1^{u-v} \lambda_2^{v}}{(u-v)!\, v!}$$

$$= \frac{1}{u!} e^{-(\lambda_1+\lambda_2)} \sum_{v=0}^{u} \frac{1 \cdot u!}{(u-v)!\, v!} \lambda_1^{u-v} \lambda_2^{v}$$

$$= \frac{e^{-(\lambda_1+\lambda_2)}}{u!} \sum_{v=0}^{u} \binom{u}{v} \lambda_1^{u-v} \lambda_2^{v}$$

$$= \frac{e^{-(\lambda_1+\lambda_2)}}{u!} (\lambda_1+\lambda_2)^{u}, \text{ for } u = 0, 1, 2, 3 \ldots$$

pmf of Poisson$(\lambda_1+\lambda_2)$

Binomial Theorem

$$\sum_{x=0}^{n} \binom{n}{x} a^{n-x} b^{x} = (a+b)^{n}$$

Continuous case

$(x,y) \to (u,v)$ one-to-one transformation

- $U = g_1(X, Y)$ and $V = g_2(X, Y)$
- Let

$$\mathcal{A} = \{(x, y) : f_{X,Y}(x, y) > 0\}$$

Support of (x,y)

$$\mathcal{B} = \{(u, v) : u = g_1(x, y), v = g_2(x, y) \text{ for some } (x, y) \in \mathcal{A}\}$$

Support of (u,v)

- If $g_1(x, y), g_2(x, y))$ is a one-to-one function from $\mathcal{A}$ to $\mathcal{B}$, then

$$A_{uv} = \{(x, y) : u = g_1(x, y), v = g_2(x, y)\}$$

contains only one point i.e. only one (x,y) for which $u=g_1(x,y)$ and $v=g_2(x,y)$

- Then we can find inverse functions

$$x = h_1(u, v) \qquad \text{and} \qquad y = h_2(u, v)$$

Continuous case

determinante

- Jacobian:

$$J = \left| \begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{bmatrix} \right| = \frac{\partial x}{\partial u}\frac{\partial y}{\partial v} - \frac{\partial y}{\partial u}\frac{\partial x}{\partial v}$$

where

$$\frac{\partial x}{\partial u} = \frac{\partial}{\partial u} h_1(u,v), \quad \frac{\partial x}{\partial v} = \frac{\partial}{\partial v} h_1(u,v),$$
$$\frac{\partial y}{\partial u} = \frac{\partial}{\partial u} h_2(u,v), \quad \frac{\partial y}{\partial v} = \frac{\partial}{\partial v} h_2(u,v),$$

Then

One-to-one bivariate transformations

$$f_{U,V}(u,v) = f_{X,Y}(\underbrace{h_1(u,v)}_{x}, \underbrace{h_2(u,v)}_{y})\,|J|$$

← absolute value

for $(u,v) \in B$

Example

- Let X and Y be independent and identically distributed (iid) random variables with $X \sim \mathrm{Expo}(\beta)$ and $Y \sim \mathrm{Expo}(\beta)$
- Set

$$U = \frac{X}{X+Y} \qquad \text{and} \qquad V = X + Y$$

Find the joint distribution of (U, V) and the marginal distributions of U and V

Sum of exponentials

$X \sim \mathrm{Expo}(\beta)$, $Y \sim \mathrm{Expo}(\beta)$, indep $\mathbb{I}_{(0,\infty)}(x)\,\mathbb{I}_{(0,\infty)}(y)$

$\Rightarrow f_{X,Y}(x,y) = f_X(x) f_Y(y) = \frac{1}{\beta} e^{-x/\beta} \frac{1}{\beta} e^{-y/\beta}$ for $x>0$, $y>0$

$u = \frac{x}{x+y}$ and $v = x+y$

Solving for x and y: e.g. $u = \frac{x}{v} \Rightarrow x = uv$ put in for x
$= h_1(u,v)$

and $y = v - x = v - uv = v(1-u) = h_2(u,v)$
put in for y

Jacobian: $\frac{\partial x}{\partial u} = \frac{\partial}{\partial u} uv = v$, $\frac{\partial x}{\partial v} = \frac{\partial}{\partial v} uv = u$

$\frac{\partial y}{\partial u} = \frac{\partial}{\partial u} v(1-u) = -v$, $\frac{\partial y}{\partial v} = \frac{\partial}{\partial v} v(1-u) = 1-u$

$$\Rightarrow J = \left| \begin{bmatrix} v & u \\ -v & 1-u \end{bmatrix} \right| = v(1-u) - (-v u) = v - vu + vu$$

$$= v$$

$x > 0$ and $y > 0$

$$\Rightarrow v = x + y \in (0, \infty)$$

$$u = \frac{x}{x+y} \in (0, 1)$$

support

$$f_{U,V}(u,v) = f_{X,Y}(h_1(u,v), h_2(u,v))\,|J|$$

$$= f_{X,Y}\big(uv, v(1-u)\big)\,|v|$$

$$= \frac{1}{\beta^2} e^{-uv/\beta} e^{-v(1-u)/\beta} \quad * \text{ for } 0 \le u \le 1 \text{ and } v \ge 0$$

$*$ To figure out support

$$I_{(0,\infty)}(uv)\, \underbrace{I_{(0,\infty)}(v(1-u))}$$

$v(1-u) > 0$

$v > vu$

$u < 1$

$$f_{U,V}(u,v) = |v| \frac{1}{\beta^2} e^{-v/\beta} I_{(0,\infty)}(v) I_{(0,1)}(u)$$

Notice: ① U, V indep. (!!)

Since $f(u,v) = g(u) \cdot h(v), \ \forall u, v$

where $g(u) = I_{(0,1)}(u)$ ← uniform (0,1) pdf

and $h(v) = \frac{1}{\beta^2} v e^{-v/\beta} I_{(0,\infty)}(v)$ ← Gamma(2, β)

② $U = \frac{X}{X+Y} \sim \text{Uniform}(0,1)$, recall the pdf of Gamma(α, β)

$$f(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta}$$

note: $\Gamma(2) = 1! = 1$

③ $V = X+Y \sim \text{Gamma}(2, \beta)$

Example

- Let X and Y be independent and identically distributed (iid) random variables with $X \sim \mathrm{N}(0,1)$ and $Y \sim \mathrm{N}(0,1)$
- Find the distribution of $\dfrac{X}{Y}$ $U = \frac{X}{Y}$, want $f_U(u)$, X, Y iid $N(0,1)$

Strategy: Set $V = Y$, find $f_{U,V}(u,v)$ and then

$f_U(u) = \int_{-\infty}^{\infty} f_{U,V}(u,v)\,dv$, Support $(x,y) \in \mathbb{R}^2 \Rightarrow (u,v) \in \mathbb{R}^2$

Inverse function: $y = v = h_2(u,v)$

and $x = uy = uv = h_1(u,v)$

$$\frac{\partial x}{\partial u} = \frac{\partial}{\partial u} uv = v \qquad \frac{\partial y}{\partial u} = \frac{\partial}{\partial u} v = 0$$

$$\frac{\partial x}{\partial v} = \frac{\partial}{\partial v} uv = u \qquad \frac{\partial y}{\partial v} = \frac{\partial}{\partial v} v = 1$$

$$\Rightarrow J = \left| \begin{bmatrix} v & u \\ 0 & 1 \end{bmatrix} \right| = v - 0 = v$$

Joint pdf for x, y:

$$f_{x,y}(x,y) = f_x(x) f_y(y) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \frac{1}{\sqrt{2\pi}} e^{-y^2/2}$$

$$= \frac{1}{2\pi} e^{-\frac{1}{2}(x^2+y^2)}, \quad x, y \in R$$

$$\Rightarrow f_{u,v}(u,v) = \frac{1}{2\pi} e^{-\frac{1}{2}(u^2v^2+v^2)} \cdot |v|, \quad u,v \in \mathbb{R}$$

$$\Rightarrow f_u(u) = \int_{-\infty}^{\infty} \frac{1}{2\pi} |v| e^{-v^2\left(\frac{u^2+1}{2}\right)} dv$$

(Symmetric around 0)
$k(v) = k(-v)$

$$= \underbrace{\int_{-\infty}^{0} \frac{1}{2\pi} (-v) e^{-v^2\left(\frac{u^2+1}{2}\right)} dv}_{= \int_{0}^{\infty} \frac{1}{2\pi} v e^{-v^2\left(\frac{u^2+1}{2}\right)} dv} + \int_{0}^{\infty} \frac{1}{2\pi} v e^{-v^2\left(\frac{u^2+1}{2}\right)} dv$$

$$= 2\int_{0}^{\infty} \frac{1}{2\pi} v e^{-v^2\left(\frac{u^2+1}{2}\right)} dv$$

$$\left(\frac{d}{dv} e^{-v^2} = -2v e^{-v^2} \right)$$

$$= \frac{1}{\pi}\left[-\frac{e^{-y^2\left(\frac{u^2+1}{2}\right)}}{2\cdot\frac{u^2+1}{2}} \right]_0^{\infty}$$

$$= \frac{1}{(u^2+1)\pi}\left(-\lim_{y\to\infty} e^{-y^2\left(\frac{u^2+1}{2}\right)} + e^{-0} \right)$$

$$= \frac{1}{\pi(u^2+1)}, \text{ for } u \in \mathbb{R}$$

pdf of Cauchy

$$u = \frac{x}{y} \sim \text{Cauchy}(0,1)$$

$$\frac{d}{dy}$$

$$\int 2y\cdot c\cdot e^{-y^2 c}\,dy = -e^{-y^2 c}$$

$$\Downarrow$$

$$\int y e^{-y^2 c}\,dy = -\frac{e^{-y^2\cdot c}}{2c}$$

Hierarchical Models and Mixture distributions

- We can construct a joint pdf/pmf using a
  - *conditional* pdf/pmf and a
  - *marginal* pdf/pmf
- From the definition of a conditional pmf/pdf follows:

$$\begin{aligned} & f(x,y) = f(x \mid y)\, f_Y(y) \\ \text{and} \quad & f(x,y) = f(y \mid x)\, f_X(x) \end{aligned}$$

A powerful modeling tool - only have to think about one thing at a time.

Example: Binomial - Poisson mixture

- Say
  - $Y =$ number of seeds spread by a plant in a plot
  - $X =$ number of seeds that survive to become seedlings
- A reasonable model would be

$$X \mid Y \sim \mathrm{Binomial}(Y, p)$$
$$Y \sim \mathrm{Pois}(\lambda)$$

The n parameter is now a random variable.

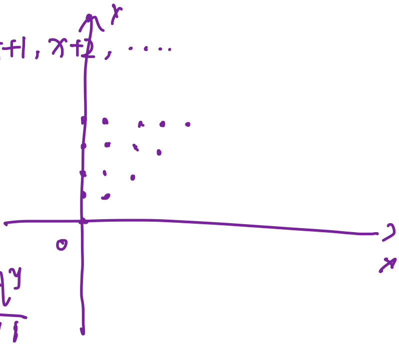
- Can show that the marginal distribution of X is $\mathrm{Pois}(\lambda p)$

Binomial, Poisson mixture.

$$f_{X,Y}(x,y) = f(x|y) \cdot f(y)$$
$$= \binom{y}{x} p^x (1-p)^{y-x} \cdot e^{-\lambda} \frac{\lambda^y}{y!}$$

for $x = 0, 1, 2, \dots, y$ and $y = 0, 1, 2, 3 \dots$

or $x = 0, 1, 2, \dots$ and $y = x, x+1, x+2, \dots$



2) $f_X(x) = \sum_y f(x,y)$

$$= \sum_{y=x}^{\infty} \binom{y}{x} p^x (1-p)^{y-x} e^{-\lambda} \frac{\lambda^y}{y!}$$

Suming over y, so x is a constant.

$$= e^{-\lambda} p^x \frac{1}{x!} \sum_{y=x}^{\infty} \frac{y!}{(y-x)!} \frac{\lambda^y}{y!} (1-p)^{y-x} \quad (*)$$

Substitution $t = y - x$, $\Rightarrow y = t + x$

and when $y = x, x+1, \ldots,$ we get, $t = 0, 1, 2, \ldots$

$$(*) = \frac{e^{-\lambda} p^{x}}{x!} \sum_{t=0}^{\infty} \frac{\lambda^{t+x}}{t!} (1-p)^{t}$$

$$= \frac{e^{-\lambda} p^{x} \lambda^{x}}{x!} e^{\lambda(1-p)} \underbrace{\sum_{t=0}^{\infty} \frac{(\lambda(1-p))^{t}}{t!} e^{-\lambda(1-p)}}_{\text{pmf of Poisson } (\lambda(1-p))}$$

$$= \frac{(p\lambda)^{x} e^{-\lambda+\lambda-\lambda p}}{x!}$$

$$= \frac{e^{-\lambda p} (\lambda p)^{x}}{x!} \quad \text{for } x = 0, 1, 2, \ldots$$

= pmf of Poisson(λp)

Poisson - Gamma mixture = Negative binomial

- If

$$\begin{aligned} Y \mid X &\sim \mathrm{Pois}(X) \\ X &\sim \mathrm{Gamma}(\alpha, \beta) \end{aligned}$$

then the marginal distribution of Y is Negative binomial

$\Rightarrow$ HW

Finite Mixture distributions

- Let $f_i(x)$, $i = 1, \ldots, n$ be pdfs/pmfs and let $p_i > 0$ and $p_1 + p_2 + \cdots + p_n = 1$. Then

Y ~ Discrete with outcomes $1, 2, \cdots, n$, and $f(x|y) = f_y(x)$

$$f(x) = \underbrace{p_1 f_1(x)}_{\geq 0} + \underbrace{p_2 f_2(x)}_{\geq 0} + \cdots + \underbrace{p_n f_n(x)}_{\geq 0}$$

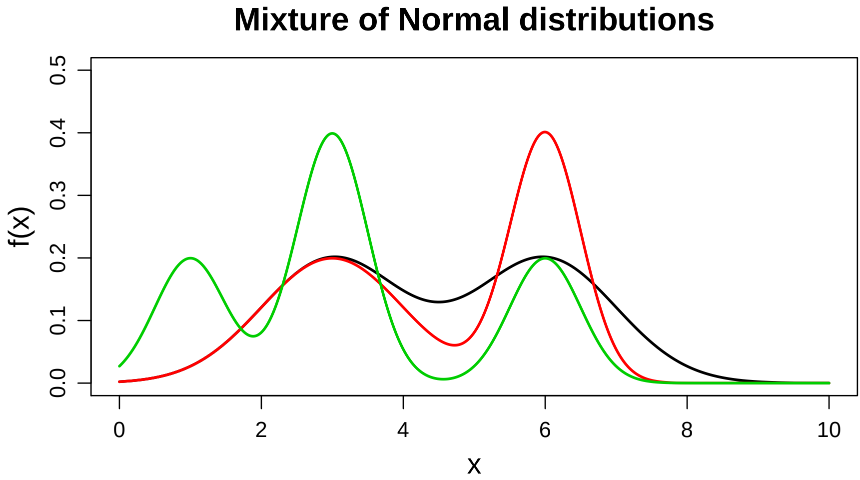
is a pdf/pmf

$$\int f(x) = \int p_1 f_1(x) + \cdots + p_n f_n(x)\, dx$$

$$= p_1 \int f_1(x) dx + \cdots + p_n \int f_n(x) dx = p_1 + p_2 \cdots + p_n = 1$$

$\Rightarrow f(x)$ is a pdf

Mixture of Normals



Iterative expections

$\Rightarrow E(g(X)) = E(E(g(X)|Y))$

- Can obtain marginal expectation and variance from a hierarchical model without first finding the marginal pdf or pmf

Theorem: Iterative expectation

$$E(X) = E(\underbrace{E(X \mid Y)}_{\int x f(x|y)\,dx}) \longrightarrow E\left(\int x f(x|y)\,dx\right)$$

If the expectations exist

marginal mean

$= \int x f(x)\,dx$

proof...

pdf. pmf cdf

- "E" is used for expectation for any distribution. Similar to "$\int$" or "F"
- Usually clear from context what distribution the E refers to.
- If in doubt, it can be useful to put the name of the random variable as a subscript:

$$E_X(X) = E_Y(E_{X|Y}(X \mid Y))$$

So $E(X)$ is $\int x f(x)\,dx$

$E(X|Y=y)$ is $\int x f(x|y)\,dx$

$E(Y)$ is $\int y f(y)\,dy$

w.r.t $f_Y(y)$

$$E(X) = E(\underbrace{E(X|Y)}_{\text{w.r.t } f(x|y)})$$

A note on conditional expectations

- Note the difference between e.g. $E(Y|X=x) = \frac{2}{3}x$

random variable $\longrightarrow$ $E(X \mid Y)$ and $E(Y|X) = \frac{2}{3}X$

$E(X \mid Y = y)$ or $E(X \mid y)$

  - $E(X \mid Y)$ is a function of Y (remember example 2 in lecture 14)
  - $E(X \mid Y)$ is a *random variable*
  - $E(X \mid Y = y)$ is a scalar (whose value depends on y)

- Remember: Any function of a random variable is also a random variable!

In Lecture 19, example 2 we found

$\mathbb{E}(Y \mid X = x) = \frac{2}{3}x$ $\leftarrow$ a number

(depends on the value of x)

$\mathbb{E}(Y \mid X) = \frac{2}{3}X$ $\leftarrow$ a function of a random variable

$\Rightarrow$ $\mathbb{E}(Y \mid X)$ is a random variable.

Similarly, $Var(Y|X)$ is also a random variable.

Therom. $E(X) = E(E(X|Y))$

Proof: $\underbrace{E(E_{X|Y}(X|Y))}_{=g(Y)} = \int_{-\infty}^{\infty} E_{X|Y}(X|Y) f_Y(y)\,dy$

i.e. just a function of Y

$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X|Y}(x|y)\,dx\, f_Y(y)\,dy$

$= (*)$

$$(X) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x f_{X|Y}(x|y)\, f_Y(y)\, dx\, dy$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x f(x,y)\, dy\, dx$$

switch order.

$$= \int_{-\infty}^{\infty} x \underbrace{\int_{-\infty}^{\infty} f(x,y)\, dy}_{= f_X(x)}\, dx = \int_{-\infty}^{\infty} x f_X(x)\, dx$$

$$= E(X)$$

Iterative variances

Theorem: Iterative variances

$$V(X) = E(V(X \mid Y)) + V(E(X \mid Y)) \geq 0$$

If the expectations exist

proof... First recall from Lecture 14 that often we have $V(X|Y=y) < V(X)$

From Theorem: $V(X) \leq E(V(X|Y))$, Since a variance is never negative.

- Remember: $V(X \mid Y)$ and $E(X \mid Y)$ are random variables
- With clarifying subscripts:

$$V_X(X) = E_Y(V_{X|Y}(X \mid Y)) + V_Y(E_{X|Y}(X \mid Y))$$

So, conditional var is on average smaller than the marginal variance.

Proof: recall $\mathrm{Var}(x) = E(x^2) - E(x)^2$

$$\Rightarrow V(g(x)) = E(g(x)^2) - E(g(x))^2$$

$$E(V(x|y)) + V(E(x|y)) = E[E(x^2|y) - E(x|y)^2]$$

$$+ E(E(x|y)^2) - \underbrace{E(E(x|y))}_{= E(x)}{}^2 =$$

$= E(x)^2$

$$= \underbrace{E(E(X^2|Y))}_{\shortparallel \atop E(X^2)} - E(E(X|Y)^2) + E(E(X|Y)^2) - E(X)^2$$

$$= E(X^2) - E(X)^2 = V(X)$$

Example

- Suppose

$$E(Y)=\lambda \qquad V(Y)=\lambda$$

$$X \mid Y \sim \text{Binomial}(Y, p) \implies E(X|Y)=Yp, \quad V(X|Y)=Yp(1-p)$$
$$\Longleftarrow \quad Y \sim \text{Pois}(\lambda)$$

What is the (marginal) expected value and variance of X?

$$E(X)=E(E(X|Y))=E(Yp)=pE(Y)=p\lambda$$

$$V(X)=E(V(X|Y))+V(E(X|Y))$$

$$=E(Yp(1-p))+V(Yp)=p(1-p)E(Y)+p^2V(Y)$$

$$=p(1-p)\lambda+p^2\lambda=p\lambda-p^2\lambda+p^2\lambda=p\lambda, \quad \text{fits. since}$$

$X \sim \text{Pois}(p\lambda)$

Example

$E(X|Y) = nY$

$\Rightarrow V(X|Y) = nY(1-Y)$

- Suppose

$E(Y) = \frac{\alpha}{\alpha+\beta}$

$$X \mid Y \sim \text{Binomial}(n, Y)$$

$$Y \sim \text{Beta}(\alpha, \beta)$$

$\Leftarrow$ $Var(Y) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$

What is the (marginal) expected value and variance of X?

$E(X) = E(E(X|Y)) = E(nY) = nE(Y) = \frac{n\alpha}{\alpha+\beta}$

$Var(X) = E(V(X|Y)) + V(E(X|Y))$

$= E(nY(1-Y)) + V(nY)$

$$= nE(Y(1-Y)) + n^2 V(Y)$$

$$= nE(Y) - nE(Y^2) + n^2 V(Y) \quad \cdots (\ast)$$

$$E(Y^2) = Var(Y) + E(Y)^2$$

$$(\ast) = nE(Y) - n\left(V(Y) + E(Y)^2\right) + n^2 V(Y)$$

$$= \frac{n\alpha}{\alpha+\beta} - n\left(\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} + \frac{\alpha^2}{(\alpha+\beta)^2}\right) + n^2\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

$$= \frac{n\alpha\beta(n+\alpha+\beta)}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

We get $X \sim \ldots$ 不建议

we can get $Y|x \ldots$