Lecture 17 - Bivariate Expectation and Conditional Distributions

Defn: Bivariate Expectation

If (X, Y) and $g : \mathbb{R}^2 \to \mathbb{R}$

then

$$E[g(X,Y)] = \begin{cases} \sum_x \sum_y g(x,y) f(x,y) \\ \iint_{\mathbb{R}^2} g(x,y) f(x,y)\, dx\, dy \end{cases}$$

recall: univariate case

Law of Unconscious Stat.

$$E[g(X)] = \begin{cases} \sum_x g(x) f(x) \\ \int g(x) f(x)\, dx \end{cases}$$

Ex. Let (X, Y) where

support $\{2, \dots, 12\}$ $\leftarrow$ X = sum of two dice rolls

$\{0, \dots, 5\}$ Y = abs. value of difference between the rolls

roll 1 : 1
roll 2 : 1 $\Rightarrow$ $X = 2$, $Y = |1-1| = 0$ $\qquad f(2,0) = 1/36$

roll 1 : 5
roll 2 : 1 $\Rightarrow$ $X = 6$, $Y = 4$ $\qquad f(6,4) = 2/36$

(5,1), or (1,5)

$f(3,0) = 0$

Sum (2,2)

(1,1)

$y \backslash x$	2	3	4	5	6	7	8	9	10	11	12
0	1/36		1/36		1/36		1/36		1/36		1/36
1		2/36		2/36		2/36		2/36		2/36	
2			2/36		2/36		2/36		2/36		
3				2/36		2/36		2/36			
4					2/36		2/36				
5						2/36					

x: Sum, y: |diff|

table $f(x,y)$

$$E[XY] = \sum_{x=2}^{12} \sum_{y=0}^{5} xy\, f(x,y) = \boxed{\frac{245}{18}}$$

leave for you to fill in

$$= (2)(0)\overbrace{f(2,0)}^{1/36} + (2)(1)\underbrace{f(2,1)}_{0} \cdots$$

Theorem: E linear

If $g_1 : \mathbb{R}^2 \to \mathbb{R}$ and $g_2 : \mathbb{R}^2 \to \mathbb{R}$

then for any $a, b \in \mathbb{R}$ we have then for any $a, b \in \mathbb{R}$ we have

$$\rightarrow \mathbb{E}\left[a g_1(X, Y) + b g_2(X, Y)\right]$$

$$= a\,\mathbb{E}[g_1(X, Y)] + b\,\mathbb{E}[g_2(X, Y)]$$

Pf. As in univariate case this comes from linearity of addition (discrete) and integration (cts)

$$\iint (a g_1(x,y) + b g_2(x,y)) f(x,y)\,dx\,dy$$

$$= a \iint g_1(x,y) f(x,y)\,dx\,dy + b \iint g_2(x,y) f(x,y)\,dx\,dy$$

Defn: Covariance

We define the covariance between X, Y as

defn
$$\boxed{\mathrm{Cov}(X, Y) = \mathbb{E}\left[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])\right]}$$

$$\mu_X = \mathbb{E}[X] \in \mathbb{R}$$
$$\mu_Y = \mathbb{E}[Y] \in \mathbb{R}$$

$$= \mathbb{E}[(X - \mu_X)(Y - \mu_Y)]$$

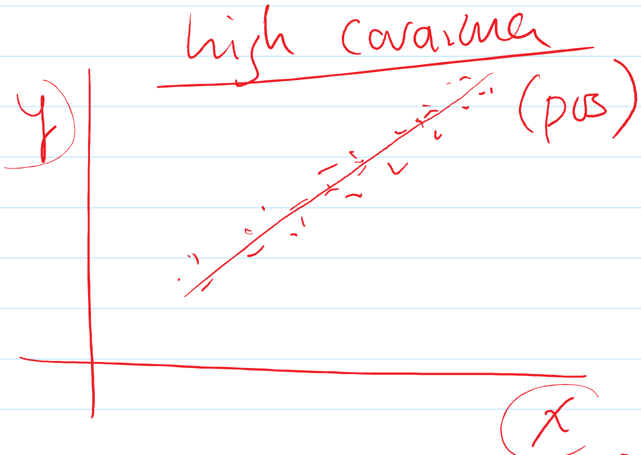
$$= \mathbb{E}[(X-\mu_X)(Y-\mu_Y)]$$

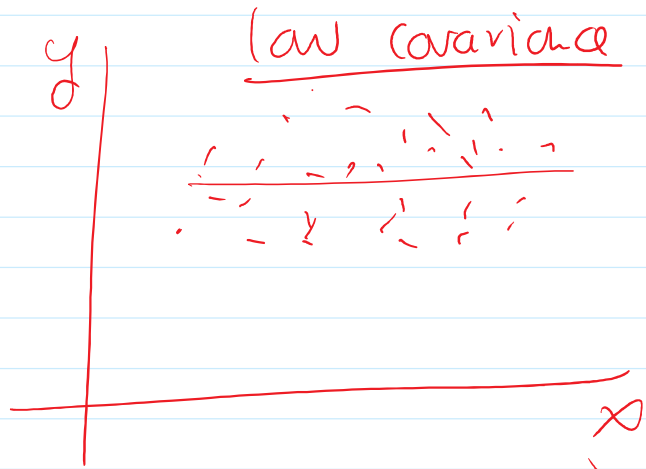
$$\mathbb{E}[g(X,Y)] \text{ where } g(x,y) = (x-\mu_X)(y-\mu_Y)$$

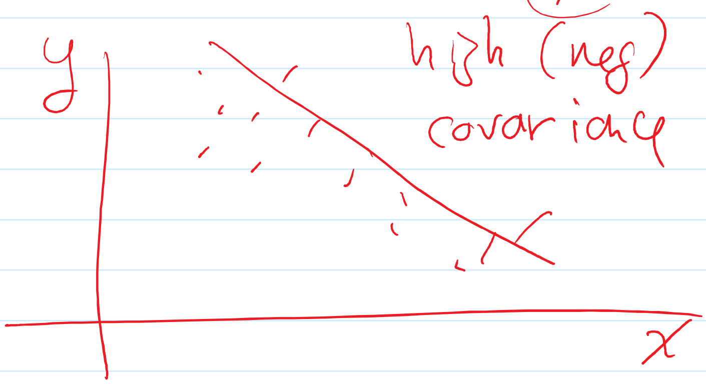
Facts:

recall: $Var(X) = \mathbb{E}[(X-\mathbb{E}[X])^2]$

so $Var(X) = Cov(X,X)$







Defn: Correlation

We re-scale covariance to be in $[-1, 1]$
and define:

and define:

$$\boxed{\mathrm{Cor}(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\,\mathrm{Var}(Y)}}}$$

$$= \frac{\mathrm{Cov}(X,Y)}{sd(X)\,sd(Y)}$$

$$\boxed{\begin{aligned} sd(X) &= \sqrt{\mathrm{Var}(X)} \\ sd(Y) &= \sqrt{\mathrm{Var}(Y)} \end{aligned}}$$

Theorem: If $a, b \in \mathbb{R}$ then

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

different than $\mathbb{E}$: $\mathbb{E}[\underbrace{aX+bY}_{Z}] = a\mathbb{E}[X] + b\mathbb{E}[Y]$

Pf. $Z = aX + bY$

$$\mathrm{Var}(Z) = \mathbb{E}\left[\underbrace{(Z - \mathbb{E}[Z])^2}\right]$$

(1)

$$Z - \mathbb{E}[Z] = aX + bY - a\mathbb{E}[X] + b\mathbb{E}[Y]$$

$$= a(X - \mathbb{E}[X]) + b(Y - \mathbb{E}[Y])$$

(2)

$$(Z - \mathbb{E}[Z])^2 = a^2(X - \mathbb{E}[X])^2 + b^2(Y - \mathbb{E}[Y])^2$$

$$(Z - E[Z])^2 = a^2(X - E[X])^2 + b^2(Y - E[Y])^2 + 2ab(X - E[X])(Y - E[Y])$$

(3) $Var(aX+bY)$

$$Var(Z) = E[(Z - E[Z])^2]$$
$$= a^2 \underbrace{E[(X - E[X])^2]}_{Var(X)} + b^2 \underbrace{E[(Y - E[Y])^2]}_{Var(Y)} + 2ab \underbrace{E[(X - E[X])(Y - E[Y])]}_{Cov(X,Y)}$$
$$= a^2 Var(X) + b^2 Var(Y) + 2ab\, Cov(X,Y).$$

Theorem:

If $a, b, c, d \in \mathbb{R}$ then

$$\boxed{Cov(a + bX, c + dY) = bd\, Cov(X, Y)}$$

recall: $Var(a+bX) = b^2 Var(X)$

$Cov(a+bX, a+bX) = Var(a+bX)$

pf. $Cov(a + bX, Z)$

$$= E[(a + bX - E[a + bX])(Z - E[Z])]$$
$$= E[(\cancel{a} + bX - (\cancel{a} + bE[X])(Z - E[Z])]$$

$$= b\,\mathbb{E}[(X - \mathbb{E}[X])(Z - \mathbb{E}[Z])]$$

$$= b\,\mathrm{Cov}(X, Z) \quad \text{(finish using symmetry)}$$

Theorem: $a, b, c, d \in \mathbb{R}$ then

$$\boxed{\mathrm{Cor}(a + bX, c + dY) = \mathrm{Sign}(b)\,\mathrm{Sign}(d)\,\mathrm{Cor}(X, Y)}$$

$$\mathrm{Sign}(b) = \begin{cases} 1 & \text{if } b > 0 \\ -1 & \text{if } b < 0 \\ 0 & \text{if } b = 0 \end{cases}$$

Alt. $|\mathrm{Cor}(a + bX, c + dY)| = |\mathrm{Cor}(X, Y)|$

E.g. $\mathrm{Cor}(X, -Y) = -\mathrm{Cor}(X, Y)$

Pf. $$\mathrm{Cor}(a + bX, c + dY) = \frac{\mathrm{Cov}(a + bX, c + dY)}{\sqrt{\mathrm{Var}(a + bX)\,\mathrm{Var}(c + dY)}}$$

$$= \frac{bd\,\mathrm{Cov}(X, Y)}{\sqrt{b^2\,\mathrm{Var}(X)\,d^2\,\mathrm{Var}(Y)}}$$

$$= \frac{b}{|b|}\,\frac{d}{|d|}\,\frac{\mathrm{Cov}(X, Y)}{\sqrt{\mathrm{Var}(X)\,\mathrm{Var}(Y)}}$$

Fact:

$$\frac{b}{\;} = \mathrm{Sign}(b)$$

$\frac{b}{|b|} = \text{sign}(b)$ for $b \neq 0$

$= \text{Sign}(b)\,\text{Sign}(d)\,\text{Cor}(X, Y)$

Theorem: $|\underbrace{\text{Cor}(X, Y)}_{\rho}| \le 1$

pf. $\tilde{X} = \dfrac{X - \mu_X}{sd(X)}$ $\quad \text{Var}(\tilde{X}) = 1$

$E[\tilde{X}] = 0$

$\tilde{Y} = \dfrac{Y - \mu_Y}{sd(Y)} \longleftarrow$ linear transformations

$$\underline{|\text{Cor}(X, Y)| = |\text{Cor}(\tilde{X}, \tilde{Y})| = \rho \ge 0}$$

$$\text{Var}(\tilde{X} + \tilde{Y}) = \text{Var}(\tilde{X}) + \text{Var}(\tilde{Y}) + 2\,\text{cov}(\tilde{X}, \tilde{Y})$$

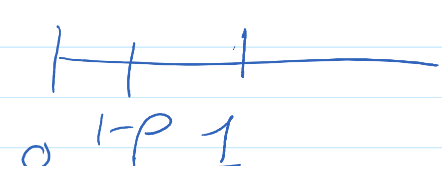
$$\text{Var}(\tilde{X} - \tilde{Y}) = \underbrace{\text{Var}(\tilde{X})}_{1} + \underbrace{\text{Var}(\tilde{Y})}_{1} - 2\,\underbrace{\text{cov}(\tilde{X}, \tilde{Y})}$$

$$\frac{\text{Cov}(\tilde{X}, \tilde{Y})}{\sqrt{1 \cdot 1}} = \text{Cor}(\tilde{X}, \tilde{Y}) = \rho$$

$$0 \le \text{Var}(\tilde{X} \pm \tilde{Y}) = 1 + 1 \pm 2\rho = 2(1 \pm \rho)$$

$$2(1 \pm \rho) \ge 0 \qquad \rho > 0$$

or $\boxed{1 \pm \rho \ge 0} \longrightarrow$



or $|1 \pm \rho| \geq 0$ $0 \quad 1-\rho \quad 1$

$$\boxed{\text{So } |\rho| \leq 1}$$

Theorem: Short-cut ~~formula~~

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

Now: $$\boxed{Cov(X,Y) = E[XY] - E[X]E[Y].}$$

Pf. $Cov(X,Y) = E\left[(X - E[X])(Y - E[Y])\right]$

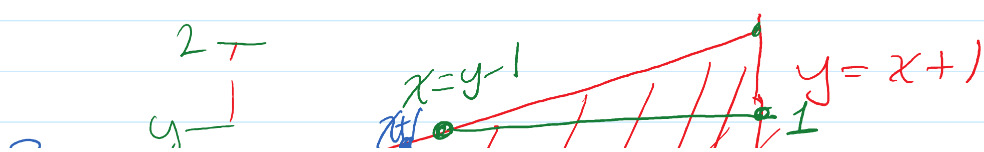
$$= E\left[XY - XE[Y] - YE[X] + E[X]E[Y]\right]$$

$$= \underline{E[XY]} - \cancel{E[X]E[Y]} - \cancel{E[Y]E[X]} + \cancel{E[X]E[Y]}$$

Ex. (X,Y) has pdf of

$$\boxed{f(x,y) = 1} \quad \text{for } 0 < x < 1 \text{ and } x < y < x+1$$

Calculate



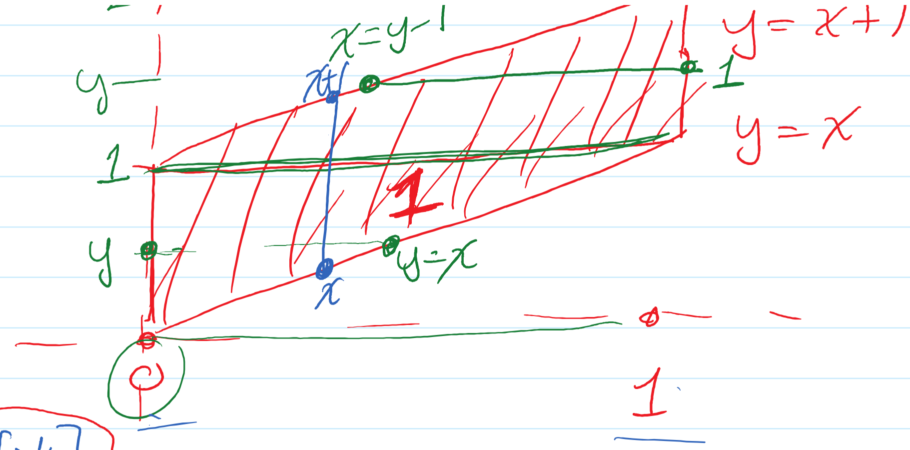
Calculate

$Cor(X, Y) = ?$

$$= \frac{Cov(X,Y)}{\sqrt{Var(X)\,Var(Y)}}$$

$$= \frac{E[XY] - E[X]E[Y]}{\sqrt{Var(X)\,Var(Y)}}$$

get these from marginal dists



Marginal of X:

$$f(x) = \int f(x,y)\,dy = \int_x^{x+1} 1\,dy = y\Big|_x^{x+1} = (x+1)-x = 1$$

for $0 < x < 1$

i.e. $X \sim U(0,1)$

$$\boxed{E[X] = \tfrac{1}{2} \quad Var(X) = \tfrac{1}{12}}$$

Marginal of Y

$$f(y) = \int f(x,y)\,dx = \begin{cases} \int_0^y 1\,dx = y & 0 < y < 1 \\ \int_{y-1}^{1} 1\,dx = 2-y & 1 \le y < 2 \end{cases}$$

$= (y \quad 0 < y < 1$

$$= \begin{cases} y & 0 < y < 1 \\ 2-y & 1 \le y < 2 \end{cases}$$

Exercise: $\mathbb{E}[Y] = 1 \quad Var(X) = 1/6$

Last part: $\mathbb{E}[XY]$

$$\mathbb{E}[XY] = \int_0^1 \int_x^{x+1} xy\, f(x,y)\, dy\, dx = \text{Calc III} = 7/12$$

exercise

All together:

$$Cor(X,Y) = \frac{7/12 - (1/2)(1)}{\sqrt{1/12 \cdot 1/6}} = \text{some number} > 0$$

Conditional Distributions

Recall: $\boxed{\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}}$

If X, Y discrete let $A = \{X = x\}$, $B = \{Y = y\}$

$$\longrightarrow \mathbb{P}(X = x | Y = y) = \frac{\mathbb{P}(X = x, Y = y)}{\mathbb{P}(Y = y)}$$

$$\rightarrow \mathbb{P}(X=x \mid Y=y) = \frac{\mathbb{P}(X=x, Y=y)}{\mathbb{P}(Y=y)}$$

Defn: Conditional PMF

The conditional pmf of X given $Y=y$ is defined as

$$\boxed{f_{X|Y=y}(x) = f(x|y) = \frac{f(x,y)}{f_Y(y)}}$$

Consider $Z = $ "$X \mid Y=y$" as a univariate random variable w/ pmf $f(x|y)$.

Ex. Joint PMF of X and Y

values of $f(x,y)$

marginal of Y	y \ x	0	1	2
7/18	30	0	3/18	4/18
6/18	20	2/18	4/18	0
5/18	10	2/18	3/18	0
	marg. X	4/18	10/18	4/18

$4/18 \qquad 1/18 \qquad 1/18 \qquad$ marg. X

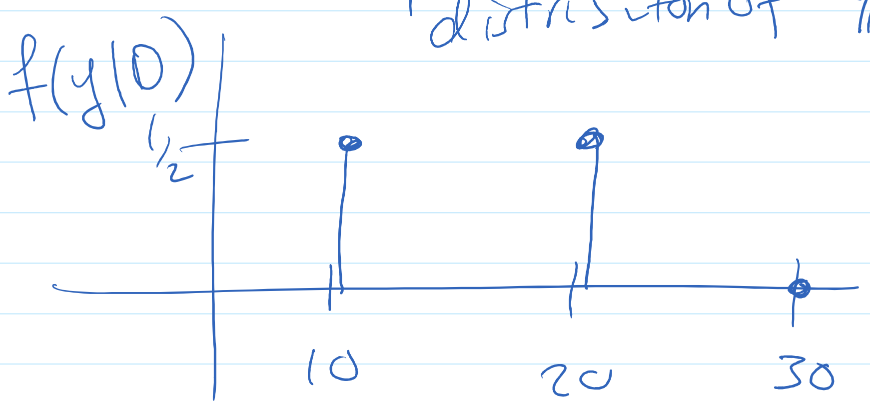
Dist of $Y \mid X = x$

$$f(y \mid X=0) = \frac{f(0,y)}{f_X(0)}$$

$f(10 \mid X=0)$	$f(20 \mid X=0)$	$f(30 \mid X=0)$
$\frac{f(0,10)}{f_X(0)} = \frac{2/18}{4/18}$	$\frac{f(0,20)}{f_X(0)} = \frac{2/18}{4/18} = 1/2$	$\frac{f(0,30)}{f_X(0)} = 0$

$= 1/2$

distribution of $Y \mid X = 0$



Cts Case?

$$P(X=x \mid Y=y) = \frac{P(X=x, Y=y)}{P(Y=y)}$$

problem! $P(Y=y) = 0$ in cts case.

Defn: Conditional Dist (cts case)

Defn: Conditional Dist (cts case)

If X, Y continuous then the conditional pdf of X given $Y = y$ is

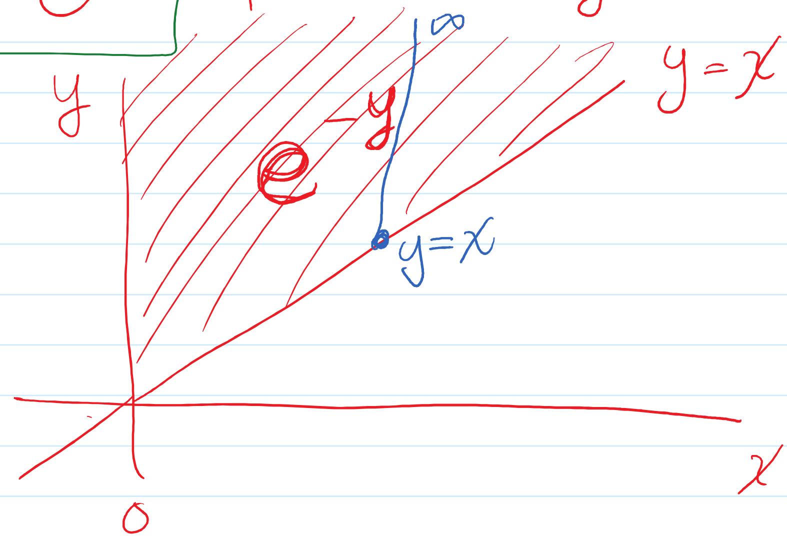
$$f(x|y) = \frac{f(x,y)}{f_Y(y)}$$

$f(x,y)$ ← joint pdf

$f_Y(y)$ ← marginal pdf

Ex. X, Y has joint pdf of

$$\boxed{f(x,y) = e^{-y}} \quad \text{for } 0 < x < y$$



Want: $f(y|x)$

i.e. pdf of $Y | X = x$

Formula:

$$f(y|x) = \frac{f(x,y)}{f_X(x)}$$

Marginal of X: $f_X(x) = \int f(x,y)\,dy = \int_x^\infty e^{-y}\,dy = e^{-x}$

So $f(y|x) = \dfrac{e^{-y}}{e^{-x}} = e^{x-y}$ for $0 < x < y$

looks like an exponential dist.

Called: shifted exponential: