Chapter 4. Bivariate Distributions

Math 3215 Spring 2024

Georgia Institute of Technology

Section 1.
Bivariate Distributions of the Discrete Type

2 RVs

Discrete.

Motivation

Suppose that we observe the maximum daily temperature, X, and maximum relative humidity, Y, on summer days at a particular weather station.

We want to determine a relationship between these two variables.

For instance, there may be some pattern between temperature and humidity that can be described by an appropriate curve $Y = u(X)$.

For 1 RV X, $f(x) = \mathbb{P}(X = x)$ ← PMF of X

Joint distribution

Let X and Y be two random variables defined on a discrete sample space.

Let S denote the corresponding two-dimensional space of X and Y, the two random variables of the discrete type.

Definition

The function $f(x, y) = \mathbb{P}(X = x, Y = y)$ is called the joint probability mass function (joint PMF) of X and Y.

"AND"

$$f(x,y) = \mathbb{P}(\{X=x\} \cap \{Y=y\})$$

Joint distribution

$\underbrace{\text{PMF}, \text{ Joint PMF}} = \mathbb{P}(\longrightarrow)$

Note that

- $0 \leq f(x, y) \leq 1$ $\quad = \mathbb{P}(\;)$
- $\sum_{(x,y)\in S} f(x, y) = 1$
- $\mathbb{P}((X, Y) \in A) = \sum_{(x,y)\in A} f(x, y)$

Joint distribution

Example 4 faced

Roll a pair of fair dice.

Let X denote the smaller and Y the larger outcome on the dice.

Find the joint PMF of (X, Y). possible.

$$f(x,y) = \mathbb{P}(X=x, Y=y)$$

$$= \begin{cases} 1/16 & , (x,y) = (1,1) \\ 2/16 & , \quad = (1,2), \cancel{(2,1)} \\ \vdots & , \quad = (1,3) \quad (3,1) \\ \vdots & , \quad = \vdots \end{cases}$$

Marginal distribution

Definition

Let X and Y have the joint probability mass function $f(x, y)$.

The probability mass function of X, which is called the marginal probability mass function of X, is defined by

$$f_X(x) = \sum_y f(x, y) = \mathbb{P}(X = x).$$

$$f_X(x) = \mathbb{P}(X = x) = \sum_y \mathbb{P}(X = x, \underbrace{Y = y}) = \sum_y f(x, y)$$

Marginal

$$f_Y(y) = \sum_x f(x, y)$$

Def We say X, Y are indep. if
any RVs

$$\mathbb{P}(X \in A \cap Y \in B) = \mathbb{P}(X \in A) \cdot \mathbb{P}(Y \in B)$$

and

for all "possible" A, B.

Marginal distribution

Definition $(X, Y$: Discrete$)$

We say X and Y are independent if

$$\text{Joint PMF} = \mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x)\mathbb{P}(Y = y) = \text{Product of Marginal PMFs}$$

for all $(x, y) \in S$.

Equivalently, $f(x, y) = f_X(x) f_Y(y)$ for all x, y.

Otherwise, we say X and Y are dependent.

Marginal distribution

Example

Let the joint PMF of X and Y be defined by

$$f(x, y) = \frac{x+y}{21}$$

for $x = 1, 2, 3$ and $y = 1, 2$.

Find the marginal PMFs of X and Y.

Determine whether they are independent.

$$f_X(x) = P(X = x) = \sum_{y = 1,2} f(x, y)$$

$$= f(x,1) + f(x,2) = \frac{x+1}{21} + \frac{x+2}{21}$$

$$= \frac{1}{21}(2x+3), \quad x = 1, 2, 3$$

$$f_Y(y) = \sum_x f(x, y) = f(1,y) + f(2,y) + f(3,y)$$

$$= \frac{1}{21} \cdot \left((1+y) + (2+y) + (3+y) \right) = \frac{3y+6}{21} = \frac{y+2}{7}, \quad y = 1, 2$$

$x = 1, 2, 3$

$y = 1, 2$

$$f(x, y) \stackrel{?}{=} f_X(x) \cdot f_Y(y)$$

$$\Downarrow$$

$$\frac{1}{21}(x+y) \stackrel{?}{=} \frac{1}{21}(2x+3) \cdot \frac{1}{7} \cdot (y+2)$$

$X = 1, 2, 3$

$y = 1, 2$

No.

$= \frac{1}{21}\cdot\frac{1}{7}(\underset{\sim}{2xy} + \dots)$

$x=1,\ y=1$

$\frac{2}{21} \neq \frac{5}{21}\cdot\frac{3}{7}$

Marginal distribution

Example

Let the joint PMF of X and Y be defined by

$$f(x,y) = \frac{xy^2}{30} = \underbrace{x}_{\text{(function of } x)} \cdot \frac{y^2}{30}$$

a function of y

for $x = 1, 2, 3$ and $y = 1, 2$.

Find the marginal PMFs of X and Y.

Determine whether they are independent.

$f_X(x) = \sum_y f(x,y) = f(x,1) + f(x,2) = \frac{x}{30}\cdot(1^2+2^2) = \frac{x}{6}$

$f_Y(y) = \sum_x f(x,y) = \frac{y^2}{30}\cdot(1+2+3) = \frac{y^2}{5}$

$f(x,y) = \frac{x\cdot y^2}{30} = \frac{x}{6}\cdot\frac{y^2}{5} = f_X(x)\cdot f_Y(y).$

indep.

Expectations

Definition

Let X_1 and X_2 be random variables of the discrete type with the joint PMF $f(x_1, x_2)$ on the space S. If $u(X_1, X_2)$ is a function of these two random variables, then

$$\mathbb{E}[u(X_1, X_2)] = \sum_{(x_1,x_2)\in S} u(x_1, x_2) f(x_1, x_2).$$

In particular, if $u(x_1, x_2) = x_1$, then

$$\mathbb{E}[u(X_1, X_2)] = \mathbb{E}[X_1] = \sum_{(x_1,x_2)\in S} x_1 f(x_1, x_2) = \sum_{x_1} x_1 f_{X_1}(x_1).$$

Examples

$u(X_1, X_2) = X_1 \longrightarrow \mathbb{E}[X_1] = \sum_{x_1, x_2} x_1 \cdot f(x_1, x_2)$

$\quad" \quad = X_2 \longrightarrow \mathbb{E}[X_2] = \sum_{x_1, x_2} x_2 \cdot f(x_1, x_2)$

$\quad" \quad = X_1 + X_2 \longrightarrow \mathbb{E}[X_1 + X_2] = \sum_{x_1, x_2} (x_1 + x_2) f(x_1, x_2)$

$\quad" \quad = X_1 \cdot X_2 \longrightarrow \mathbb{E}[X_1 \cdot X_2] = \sum_{x_1, x_2} x_1 \cdot x_2 f(x_1, x_2)$

$\vdots$

Expectations

Example

There are eight similar chips in a bowl: three marked $(0,0)$, two marked $(1,0)$, two marked $(0,1)$, and one marked $(1,1)$.

A player selects a chip at random.

Let X_1 and X_2 represent those two coordinates.

Find the joint PMF.

Compute $\mathbb{E}[X_1 + X_2]$.

$X_1 \backslash X_2$	0	1	Marginal PMF of X_1
0	3/8	2/8	5/8
1	2/8	1/8	3/8
Marginal PMF f_{X_2}	5/8	3/8	

$$f(x_1, x_2) = \begin{cases} 3/8 & (x_1, x_2) = (0,0) \\ 2/8 & (1,0) \\ 2/8 & (0,1) \\ 1/8 & (1,1) \end{cases}$$

$$\mathbb{E}[X_1 + X_2] = \sum (x_1 + x_2) \cdot f(x_1, x_2)$$

$$= (0+0) \cdot f(0,0) + (1+0) f(1,0) + (0+1) \cdot f(0,1) + (1+1) \cdot f(1,1)$$

$$= 0 \cdot \frac{3}{8} + 1 \cdot \frac{2}{8} + 1 \cdot \frac{2}{8} + 2 \cdot \frac{1}{8}$$

$$= \frac{3}{4}.$$

$$\mathbb{E}[X_1] = \sum_{x_1, x_2} x_1 \cdot f(x_1, x_2)$$

$$= \sum_{x_1} x_1 \cdot f_{X_1}(x_1) = 0 \cdot f_{X_1}(0) + 1 \cdot f_{X_1}(1)$$

$$= \frac{3}{8}$$

$$\mathbb{E}[X_2] = \frac{3}{8} \quad , \quad \mathbb{E}[X_1 + X_2] = \mathbb{E}[X_1] + \mathbb{E}[X_2]$$

Exercise

Roll a pair of four-sided dice, one red and one black.

Let X equal the outcome of the red die and let Y equal the sum of the two dice.

Find the joint PMF.

Are they independent?

Section 2.
The Correlation Coefficient

Covariance and Correlation coefficient

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

$\sigma_X = \sqrt{Var(X)}$, $\sigma_Y = \sqrt{Var(Y)}$

Definition

The covariance of X and Y is

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

The correlation coefficient of X and Y is

$$\rho = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}.$$

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

$$= \sum_{x,y} (x - \mu_X) \cdot (y - \mu_Y) \cdot \underbrace{f(x,y)}_{\text{joint PMF.}}$$

<u>Note</u>

$$\text{Cov}(X, X) = \mathbb{E}[(X - \mu_X) \cdot (X - \mu_X)]$$

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

$$\mathrm{Cov}(X,Y) = \mathbb{E}[(X-\mu_X)\cdot(Y-\mu_Y)], \quad \mu_X = \mathbb{E}[X], \quad \mu_Y = \mathbb{E}[Y]$$

$$\rho = \frac{\mathrm{Cov}(X,Y)}{\sigma_X \sigma_Y}$$: Correlation coefficient

$$\sigma_X = \sqrt{\mathrm{Var}(X)}, \quad \sigma_Y = \sqrt{\mathrm{Var}(Y)}$$

(i) $X = Y$, $\mathrm{Cov}(X,Y) = \mathbb{E}[(X-\mu_X)^2] = \mathrm{Var}(X) = \sigma_X^2$

$$\rho = \frac{\mathrm{Cov}(X,X)}{\sigma_X \cdot \sigma_X} = 1$$

(ii) $X = -Y$, $\mathrm{Cov}(X,Y) = -\mathrm{Var}(X) = -\sigma_X^2$

$$\rho = -1$$

(iii) $Y = b\cdot X + c$, $\mathrm{Cov}(X,Y) = b\cdot \mathrm{Cov}(X,X) = b\cdot\sigma_X^2$

$$\Downarrow$$

$$\sigma_Y = |b|\cdot\sigma_X \qquad \rho = \frac{\mathrm{Cov}(X,Y)}{\sigma_X\cdot\sigma_Y} = \begin{cases} 1, & b>0 \\ -1, & b<1 \end{cases}$$

(iv) If X, Y indep. $\left(f(x,y) = f_X(x)\cdot f_Y(y) \quad \forall x,y \right)$

$$\mathrm{Cov}(X,Y) = \sum_{x,y} (x-\mu_X)\cdot(y-\mu_Y)\cdot \underbrace{f(x,y)}_{=f_X(x)\cdot f_Y(y)}$$

$$= \left(\sum_x (x-\mu_X) f_X(x)\right)\left(\sum_y (y-\mu_Y) f_Y(y)\right)$$

$$= \underbrace{\mathbb{E}[(X-\mu_X)]}_{=0}\cdot\underbrace{\mathbb{E}[(Y-\mu_Y)]}_{=0}$$

$$= 0$$

$$\rho = 0$$

Covariance and Correlation coefficient

Properties

1. If X and Y are independent, then $\text{Cov}(X, Y) = 0$.
2. $\text{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$.
3. $-1 \leq \rho \leq 1$. & $\rho^2 = 1$ implies $Y = bX + c$

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

$$= \mathbb{E}[XY - \mu_X \cdot Y - \mu_Y \cdot X + \mu_X \mu_Y]$$

$$= \mathbb{E}[XY] - \mu_X \underbrace{\mathbb{E}[Y]}_{= \mu_Y} \cancel{- \mu_Y \underbrace{\mathbb{E}[X]}_{= \mu_X}} + \cancel{\mu_X \mu_Y}$$

$$= \mathbb{E}[XY] - \mu_X \cdot \mu_Y$$

" $\rho^2 \leqslant 1$ " comes from

$$\left(\mathbb{E}[(X-\mu_X)(Y-\mu_Y)]\right)^2 \leqslant \mathbb{E}[(X-\mu_X)^2] \cdot \mathbb{E}[(Y-\mu_Y)^2]$$

Covariance and Correlation coefficient

Example

Let the joint PMF of X and Y be defined by

$$f(x,y) = \frac{x+2y}{18}$$

for $x = 1, 2$ and $y = 1, 2$.

Compute $\text{Cov}(X, Y)$ and ρ.

$$\text{Cov}(X,Y) = \mathbb{E}[X \cdot Y] - \mathbb{E}[X] \cdot \mathbb{E}[Y]$$

$$\mathbb{E}[XY] = \sum_{x,y} x \cdot y \cdot f(x,y)$$

$$= 1 \cdot 1\, f(1,1) + 1 \cdot 2 \cdot f(1,2) + 2 \cdot 1\, f(2,1) + 2 \cdot 2\, f(2,2)$$

$$= 1 \cdot \frac{3}{18} + 2 \cdot \frac{5}{18} + 2 \cdot \frac{4}{18} + 4 \cdot \frac{6}{18}$$

$$= \frac{1}{18} \cdot (3 + 10 + 8 + 24) = \frac{45}{18}$$

$$\mathbb{E}[X] = 1 \cdot f(1,1) + 1\, f(1,2) + 2\, f(2,1) + 2\, f(2,2)$$

$$= 1 \cdot \frac{3}{18} + 1 \cdot \frac{5}{18} + 2 \cdot \frac{4}{18} + 2 \cdot \frac{6}{18}$$

$$= \frac{1}{18} \cdot (3 + 5 + 8 + 12) = \frac{28}{18}$$

$$\mathbb{E}[Y] = 1 \quad f(1,1) + 2 \cdot f(1,2) + 1 \cdot f(2,1) + 2 \cdot f(2,2)$$

$$= 1 \cdot \frac{3}{18} + 2 \cdot \frac{5}{18} + 1 \cdot \frac{4}{18} + 2 \cdot \frac{6}{18}$$

$$= \frac{1}{18} \cdot (3 + 10 + 4 + 12) = \frac{29}{18}$$

$$\mathrm{Cov}(X, Y) = \frac{45}{18} - \frac{28}{18} \cdot \frac{29}{18}$$

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

The Least Squares Regression Line

Suppose we are trying to see if there is a pattern or a certain relation between two random variables X and Y.

One of natural ways is to consider a linear relation between X and Y, that is, to figure out the best possible slope b such that $Y - \mu_Y = b(X - \mu_X)$ has small errors.

We measure the error by $\mathbb{E}[((Y - \mu_Y) - b(X - \mu_X))^2]$.

Find b, c so that $Y \approx bX + c$

Difference between Y, $bX + c$ is as small as possible

At least, we expect that

$$\mu_Y = \mathbb{E}[Y] = \mathbb{E}[bX + c] = b \cdot \mu_X + c$$

$$c = \mu_Y - b\mu_X$$

$$Y - (bX + c) = (Y - \mu_Y) - b(X - \mu_X)$$

Minimize

$Ax = b$

$\|A\hat{x} - b\|$ minimized

$Y \approx \rho \cdot \frac{\sigma_Y}{\sigma_X} \cdot (X - \mu_X)$

$y = \rho \frac{\sigma_Y}{\sigma_X} (x - \mu_X)$

line of best fit

The Least Squares Regression Line

One can see by some calculus that the error is minimized when

$$b = \rho \frac{\sigma_Y}{\sigma_X}$$

and the minimum error is $\sigma_Y^2(1 - \rho^2)$.

The line $Y - \mu_Y = \rho \frac{\sigma_Y}{\sigma_X}(X - \mu_X)$ is called **the line of best fit**, or **the least squares regression line**.

least square line : $y - \mu_Y = \rho \cdot \frac{\sigma_Y}{\sigma_X}(x - \mu_X)$

$\mu_X = \frac{1}{2} \quad \sigma_X = \sqrt{2 \cdot \frac{1}{4} \cdot \frac{3}{4}} = \frac{\sqrt{6}}{4}$

$\mu_Y = 1 \quad \sigma_Y = \sqrt{2 \cdot \frac{1}{2} \cdot \frac{1}{2}} = \frac{1}{\sqrt{2}}$

$$y - 1 = \rho \cdot \frac{1}{\sqrt{2}} \cdot \frac{4}{\sqrt{6}}\left(x - \frac{1}{2}\right)$$

The Least Squares Regression Line

$\mathrm{Cov}(X,Y) = E[XY] - \frac{1}{2} \cdot 1$

$\rho = \frac{\mathrm{Cov}(X,Y)}{\sigma_X \sigma_Y}$

Example

Let X equal the number of ones and Y the number of twos and threes when a pair of fair four-sided dice is rolled.

Then X and Y have a trinomial distribution.

Find the least squares regression line.

$X \sim \mathrm{Bin}(2, \frac{1}{4}) \qquad Y \sim \mathrm{Bin}(2, \frac{1}{2})$

X: 0, 1, 2 $\qquad Y$: 0, 1, 2

$$f(x,y) = \begin{cases} \left(\frac{1}{4}\right)^2, & x=0, \ y=0 \\ 2 \cdot \left(\frac{1}{4}\right)\left(\frac{1}{4}\right), & x=1, \ y=0 \\ 1 \cdot \left(\frac{1}{4}\right)^2, & x=2, \ y=0 \\ \vdots \end{cases}$$

Trinomial distribution

Consider an experiment with three outcomes, say perfect, seconds, and defective.

Let p_1, p_2, p_3 be the corresponding probabilities.

Repeat the experiment n times and let X, Y be the numbers of perfect and seconds.

We say (X, Y) has **the trinomial distribution**.

Uncorrelated

We say X, Y are uncorrelated if $\rho = 0$.

If X, Y are independent then they are uncorrelated.

However, the converse is not true.

Uncorrelated

Example

Let X and Y have the joint pmf $f(x,y) = \frac{1}{3}$ for $(x,y) = (0,1), (1,0), (2,1)$.

① $\mathrm{Cov}(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\,\mathbb{E}[Y]$

$$E[X \cdot Y] = 0 \cdot 1 \cdot f(0,1) + 1 \cdot 0 \cdot f(1,0) + 2 \cdot 1 \cdot f(2,1) = \frac{2}{3}$$

$$E[X] = 0 \cdot f(0,1) + 1 \cdot f(1,0) + 2 \cdot f(2,1) = 1$$

$$E[Y] = 1 \cdot f(0,1) + 0 \cdot f(1,0) + 1 \cdot f(2,1) = \frac{2}{3}.$$

$$\mathrm{Cov}(X,Y) = \frac{2}{3} - 1 \cdot \frac{2}{3} = 0. \qquad \rho = 0$$

X, Y uncorrelated.

② Indep. ? $\quad f(x,y) \neq f_X(x) \cdot f_Y(y) \quad$ Dependent

$$\frac{1}{3} = f(0,1) \neq f_X(0) \cdot f_Y(1) = \frac{1}{3} \cdot \frac{2}{3}$$

$$f_X(0) = f(0,1) = \frac{1}{3} \qquad f_Y(1) = f(0,1) + f(2,1) = \frac{2}{3}$$

Exercise

The joint pmf of X and Y is $f(x, y) = \frac{1}{6}$, $0 < x + y < 2$, where x and y are nonnegative integers.

Find the covariance and the correlation coefficient.

Section 3.
Conditional Distributions

Conditional distribution

Definition

The conditional probability mass function of X, given that $Y = y$, is defined by

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

$$f_{X|Y}(x|y) = \mathbb{P}(X = x \mid Y = y)$$

$$= \frac{\mathbb{P}(X = x, Y = y)}{\mathbb{P}(Y = y)} = \frac{f(x,y)}{f_Y(y)}$$

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

Conditional distribution

Example

Let the joint pmf of X and Y be defined by

$$f(x, y) = \frac{x + y}{21}$$

for $x = 1, 2, 3$ and $y = 1, 2$. We have shown that

$$f_X(x) = \frac{2x + 3}{21}, \qquad f_Y(y) = \frac{3y + 6}{21}.$$

Find the conditional PMFs.

$$f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)} = \frac{(x+y)/21}{(3y+6)/21} = \frac{x+y}{3y+6}$$

$$f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)} = \frac{(x+y)/21}{(2x+3)/21} = \frac{x+y}{2x+3}$$

In general, $\mathbb{E}[u(Y) \mid X=x] = \sum_y u(y) f_{Y|X}(y|x)$

Conditional distribution

Definition

The conditional expectation of Y given $X = x$ is defined by

$$\mathbb{E}[Y|X = x] = \sum_{y} y f_{Y|X}(y|x).$$

The conditional variance of Y given $X = x$ is defined by

$$\begin{aligned}\mathrm{Var}(Y|X = x) &= \mathbb{E}[(Y - \mathbb{E}[Y|X = x])^2|X = x] \\ &= \mathbb{E}[Y^2|X = x] - (\mathbb{E}[Y|X = x])^2.\end{aligned}$$

Recall

X, Y joint PMF $f(x,y)$

$Y \mid X = x$ Conditional PMF

$$f_{Y|X}(y|x) = P(Y = y \mid X = x) = \frac{f(x,y)}{f_X(x)}$$

$$\mathbb{E}[u(Y) \mid X = x] = \sum_y u(y) \cdot f_{Y|X}(y|x)$$

$$\mathrm{Var}(Y \mid X = x) = \mathbb{E}[(Y - \mathbb{E}[Y|X=x])^2 \mid X = x]$$
$$= \mathbb{E}[Y^2 \mid X = x] - (\mathbb{E}[Y \mid X = x])^2$$

Conditional distribution

Example

Let the joint PMF of X and Y be defined by

$$f(x, y) = \frac{x + y}{21} \qquad f_X(x) = \frac{2x+3}{21}$$

for $x = 1, 2, 3$ and $y = 1, 2.$ $\qquad f_{Y|X}(y|x) = \frac{x+y}{2x+3}$

Find $\mathbb{E}[Y|X = 3]$ and $\mathrm{Var}(Y|X = 3)$.

$$\mathbb{E}[Y \mid X = 3] = \sum_y y \cdot f_{Y|X}(y|3)$$
$$= 1 \cdot f_{Y|X}(1|3) + 2 \cdot f_{Y|X}(2|3)$$
$$= 1 \cdot \frac{1+3}{9} + 2 \cdot \frac{2+3}{9} = \frac{14}{9}.$$

$$\mathbb{E}[Y^2 \mid X = 3] = \sum_y y^2 \cdot f_{Y|X}(y|3)$$
$$= 1^2 \cdot f_{Y|X}(1|3) + 2^2 \cdot f_{Y|X}(2|3)$$
$$= 1^2 \cdot \frac{1+3}{9} + 2^2 \cdot \frac{2+3}{9} = \frac{24}{9}.$$

$$\mathrm{Var}(Y \mid X = 3) = \frac{24}{9} - \left(\frac{14}{9}\right)^2.$$

$h(x) = \mathbb{E}[Y \mid X = x] = \underbrace{\sum_y y \cdot f_{Y|X}(y|x)}_{\text{no } y \text{, still have } x}$ ← a function of x

"Define a new random variable $h(X)$"

$h(X) = \underbrace{\mathbb{E}[Y|X]}$ notation

Contional expectation as a function and a random variable

One can consider $\mathbb{E}[Y|X = x]$ as a function of x.

Say $h(x) = \mathbb{E}[Y|X = x]$

We define a random variable $\mathbb{E}[Y|X] = h(X)$.

Contional expectation as a function and a random variable

$$\begin{cases} f_X(x) = \frac{2x+3}{21} \\ f_Y(y) = \frac{y+2}{7} \\ f_{Y|X}(y|x) = \frac{x+y}{2x+3} \end{cases}$$

Example

Let the joint pmf of X and Y be defined by

$$f(x,y) = \frac{x+y}{21}$$

for $x = 1, 2, 3$ and $y = 1, 2$. One can see that $\mathbb{E}[Y|X=1] = \frac{8}{5}$, $\mathbb{E}[Y|X=2] = \frac{11}{7}$, $\mathbb{E}[Y|X=3] = \frac{14}{9}$

Find the PMF of $\mathbb{E}[Y|X]$ and $\mathbb{E}[\mathbb{E}[Y|X]]$.

$$Z = \mathbb{E}[Y|X] = h(X), \quad h(x) = \mathbb{E}[Y|X=x]$$

$$= \begin{cases} \frac{8}{5}, & x = 1 \\ \frac{11}{7}, & x = 2 \\ \frac{14}{9}, & x = 3 \end{cases}$$

$$f_Z(z) = \mathbb{P}(Z = z)$$

$$= \mathbb{P}(h(X) = z)$$

$$= \begin{cases} \frac{5}{21}, & z = \frac{8}{5} \\ \frac{7}{21}, & z = \frac{11}{7} \\ \frac{9}{21}, & z = \frac{14}{9} \end{cases}$$

$$\mathbb{P}(h(X) = \tfrac{8}{5}) = \mathbb{P}(X=1) = f_X(1) = \frac{5}{21}$$

$$f_X(2)$$

$$\mathbb{E}[Z] = \mathbb{E}[\mathbb{E}[Y|X]] = \frac{\cancel{5}}{21}\cdot\frac{8}{\cancel{5}} + \frac{\cancel{7}}{21}\cdot\frac{11}{\cancel{7}} + \frac{\cancel{9}}{21}\cdot\frac{14}{\cancel{9}}$$

$$= \frac{33}{21} = \frac{11}{7}$$

$$\mathbb{E}[Y] = \sum_y y \cdot f_Y(y) = 1 \cdot \frac{(1+2)}{7} + 2 \cdot \frac{(2+2)}{7} = \frac{11}{7}$$

Contional expectation as a function and a random variable

Theorem

1. $\mathbb{E}[\mathbb{E}[Y|X]] = \mathbb{E}[Y]$
2. $\mathrm{Var}(Y) = \mathbb{E}[\mathrm{Var}(Y|X)] + \mathrm{Var}(\mathbb{E}[Y|X])$

$$\mathbb{E}[\mathbb{E}[Y|X]] = \sum_x \mathbb{E}[Y|X=x] \cdot P(X=x)$$

$$= \sum_x \left(\sum_y y \, f_{Y|X}(y|x)\right) \cdot f_X(x)$$

$$= \sum_{x,y} y \cdot \frac{f(x,y)}{\cancel{f_X(x)}} \cdot \cancel{f_X(x)}$$

$$= \sum_{x,y} y \cdot f(x,y) = \mathbb{E}[Y].$$

Note $\quad \mathbb{E}[u(X)\,\mathbb{E}[Y|X]] = \mathbb{E}[u(X)Y]$

Y

Contional expectation as a function and a random variable

$X \sim \text{Pois}(4)$

$Y | X = x \sim \text{Bin}(x, p)$

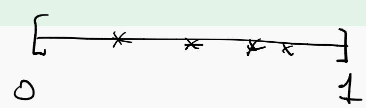
Example

$\lambda = 4$

Let X have a Poisson distribution with mean 4, and let Y be a random variable whose conditional distribution, given that $X = x$, is binomial with sample size $n = x$ and probability of success p.

Find $\mathbb{E}[Y]$ and $\text{Var}(Y)$.

$x = 4$



X = # of customers $\sim$ Poisson.

Y = # of male customers

$$\mathbb{E}[Y] = \mathbb{E}[\mathbb{E}[Y | X]]$$

$$= \mathbb{E}[X \cdot p] = p \cdot \mathbb{E}[X] = 4p$$

$$\text{Var}(Y) = \text{Var}(\mathbb{E}[Y | X]) + \mathbb{E}[\text{Var}(Y | X)]$$

$$= \text{Var}(X \cdot p) + \mathbb{E}[X \cdot p(1-p)]$$

$$= p^2 \cdot \text{Var}(X) + p(1-p)\,\mathbb{E}[X] = 4p^2 + p(1-p) \cdot 4 = 4p$$

$$\mathbb{E}[Y|X] = a + bX$$

$$\mu_Y = \mathbb{E}[Y] = \mathbb{E}[\mathbb{E}[Y|X]] = a + b\,\mathbb{E}[X] = a + b\mu_X$$

$$\mathbb{E}[XY] = \mathbb{E}[X \cdot \mathbb{E}[Y|X]] = \mathbb{E}[aX + bX^2] = a\,\mathbb{E}[X] + b\,\mathbb{E}[X^2]$$

Linear case

a function of x

Suppose $\mathbb{E}[Y|X=x]$ is linear in x, that is, $\mathbb{E}[Y|X=x] = a + bx$.

Then we have $\mu_Y = a + b\mu_X$ and $\mathbb{E}[XY] = a\mu_X + b\mathbb{E}[X^2]$.

Solving for a,, we have

$$a = \mu_Y - \rho\frac{\sigma_Y}{\sigma_X}\mu_X, \qquad b = \rho\frac{\sigma_Y}{\sigma_X}.$$

Thus,

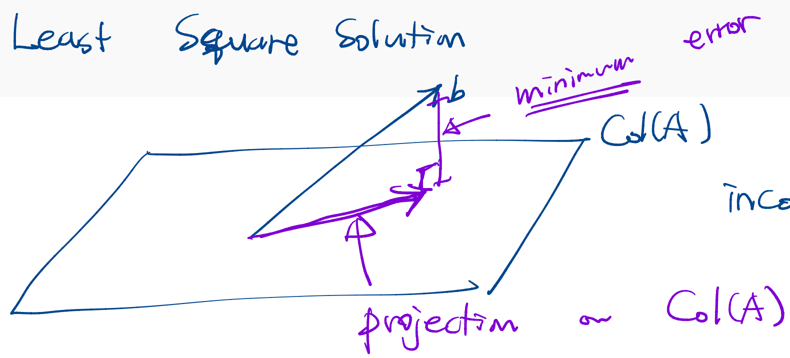
$$\mathbb{E}[Y|X=x] = \mu_Y + \rho\frac{\sigma_Y}{\sigma_X}(x - \mu_X).$$

Linear Regression.

Idea

Least Square Solution

$$Ax = b$$



inconsistent = no solution.

Conditional Expectation is a "projection".

Linear case

$$\binom{n}{x, y} = \frac{n!}{x!\, y!\, (n-x-y)!} = \binom{n}{x} \cdot \binom{n-x}{y}$$

Example

Let X and Y have the trinomial distribution with parameters n, p_X, p_Y, that is, the joint pmf is given by

$$f(x, y) = \binom{n}{x, y} p_X^x p_Y^y (1 - p_X - p_Y)^{n-x-y}.$$

Find $\mathbb{E}[Y|X = x]$.

$(p_X + p_Y + p_Z = 1)$

Experiment w/ three outcomes A, B, C
p_X p_Y p_Z

Repeat n times

X = # of A happens $\quad X \sim Bin(n, p_X)$

Y = # of B happens $\quad Y \sim Bin(n, p_Y)$

$Y | X = x \sim$??

Example $\quad n = 6, \quad x = 2 \qquad Y | X = 2 \sim Bin\left(4, \dfrac{p_Y}{p_Y + p_Z}\right)$

4 Experiments. $\quad P(B \text{ happens} \mid A \text{ does not happen}) = \dfrac{p_Y}{p_Y + p_Z}$

$$Y \mid X = x \sim Bin\left(n - x, \frac{p_Y}{p_Y + p_Z}\right)$$

$$\Rightarrow \mathbb{E}[Y \mid X = x] = (n - x) \cdot \frac{p_Y}{p_Y + p_Z}$$

$p_Y + p_Z = 1 - p_X$

$$\mathbb{E}[XY] = \mathbb{E}[X \cdot \mathbb{E}[Y \mid X]]$$

$$= \mathbb{E}\left[X \cdot (n - X) \cdot \frac{p_Y}{1 - p_X}\right]$$

$$= \left(\frac{p_Y}{1 - p_X}\right) \cdot \mathbb{E}[X(n - X)] = \ldots$$

$Cov(X, Y) = ?$

Trinomial distribution

Consider an experiment with three outcomes, say perfect, seconds, and defective.

Let p_1, p_2, p_3 be the corresponding probabilities.

Repeat the experiment n times and let X, Y be the numbers of perfect and seconds.

We say (X, Y) has **the trinomial distribution**.

Exercise

A miner is trapped in a mine containing 3 doors.

The first door leads to a tunnel that will take him to safety after 3 hours of travel.

The second door leads to a tunnel that will return him to the mine after 5 hours of travel.

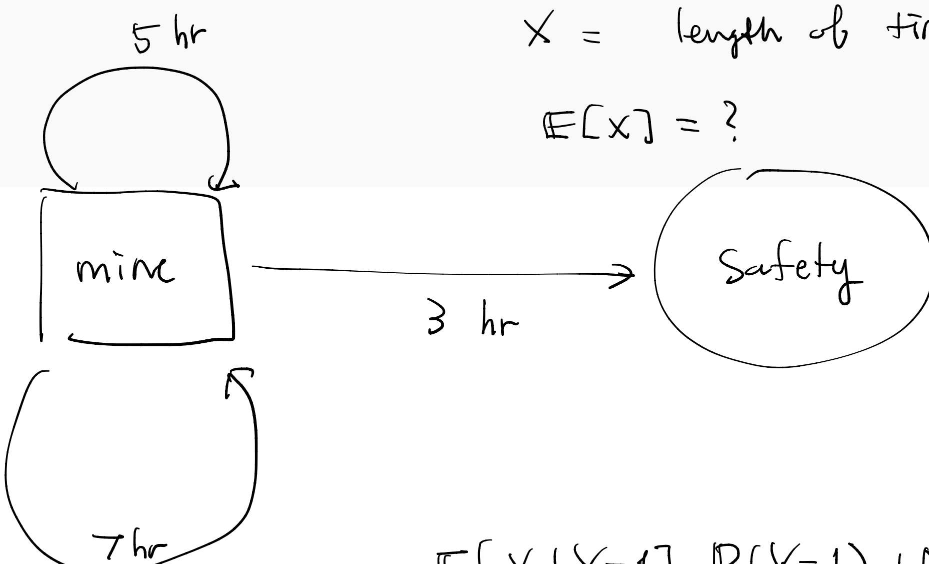
The third door leads to a tunnel that will return him to the mine after 7 hours.

If we assume that the miner is at all times equally likely to choose any one of the doors, what is the expected length of time until he reaches safety?

$Y = \begin{cases} 1 \\ 2 \\ 3 \end{cases}$

X = length of time until safety

$\mathbb{E}[X] = ?$



$$\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X \mid Y]] = \mathbb{E}[X \mid Y=1] \cdot \mathbb{P}(Y=1) + \mathbb{E}[X \mid Y=2]\,\mathbb{P}(Y=2) + \mathbb{E}[X \mid Y=3]\,\mathbb{P}(Y=3)$$

$$= \frac{1}{3}\left(\mathbb{E}[X \mid Y=1] + \mathbb{E}[X \mid Y=2] + \mathbb{E}[X \mid Y=3]\right)$$

$$= \frac{1}{3}\left(3 + (5 + \mathbb{E}[X]) + (7 + \mathbb{E}[X])\right)$$

$$3\,\mathbb{E}[X] = 15 + 2\,\mathbb{E}[X] \qquad \therefore\ \mathbb{E}[X] = 15.$$

$$\mathbb{E}[Y \mid X = x] = \sum_y y \cdot \underbrace{f_{Y|X}(y|x)}_{= \frac{f(x,y)}{f_X(x)}} = h(x)$$

$$\mathbb{E}[Y \mid X] = h(X)$$

$$\mathbb{E}[Y] = \mathbb{E}[\,\mathbb{E}[Y \mid X]]$$

$$\mathbb{E}[XY] = \mathbb{E}[\,X \cdot \mathbb{E}[Y \mid X]]$$

$$\mathrm{Var}(Y) \left(= \mathbb{E}[(Y-\mu_Y)^2] = \mathbb{E}[\,\mathbb{E}[(Y-\mu_Y)^2 \mid X]]\right)$$

$$= \mathbb{E}[\mathrm{Var}(Y \mid X)] + \mathrm{Var}(\mathbb{E}[Y \mid X])$$

If X, Y indep. $\quad \mathbb{E}[Y \mid X] = \mathbb{E}[Y]$

$$\left(\begin{aligned} f(x,y) &= f_X(x) \cdot f_Y(y) \\ f_{Y|X}(y|x) &= f_Y(y) \end{aligned}\right)$$

$$\mathbb{E}[X \mid X = x] = x \qquad \mathbb{E}[X \mid X] = X$$

$$\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$$

$$\mathrm{Var}(X+Y) = \mathbb{E}[(X+Y)^2] - (\mathbb{E}[X+Y])^2$$

$$= \mathbb{E}[X^2 + 2XY + Y^2] - \left((\mathbb{E}[X])^2 + 2\,\mathbb{E}[X]\,\mathbb{E}[Y] + (\mathbb{E}[Y])^2\right)$$

$$= \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\,\mathrm{Cov}(X,Y)$$

$$\mathrm{Cov}(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\,\mathbb{E}[Y]$$

If X, Y indep, $\quad \mathrm{Cov}(X,Y) = 0$

$$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$$

Section 4.
Bivariate Distributions of the Continuous Type

Recall X is Conti. RV if it has a PDF.

Joint PDF

Definition joint PDF

An integrable function $f(x, y)$ is **the joint probability density function** of two random variables X, Y if

- $f(x, y) \geq 0$
- $\iint f(x, y)\, dxdy = 1$
- $\mathbb{P}((X, Y) \in A) = \iint_A f(x, y)\, dxdy$

The marginal density functions for X, Y are

$$f_X(x) = \int f(x, y)\, dy, \qquad f_Y(y) = \int f(x, y)\, dx.$$

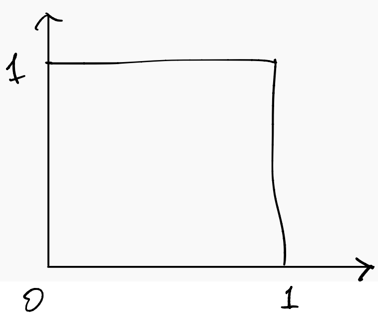
Joint PDF

Example

Let X and Y have the joint PDF

$$f(x,y) = \frac{4}{3}(1 - xy)$$

for $0 < x, y < 1$. Find f_X, f_Y, and $\mathbb{P}(Y \le \frac{X}{2})$.



inequalities

$0 < x, y < 1$ defines a region (support) where $f(x,y) > 0$

$$\begin{cases} x > 0 \\ y > 0 \\ x < 1 \\ y < 1 \end{cases}$$

$\Downarrow$

$x = 0$, $y = 0$, $x = 1$, $y = 1$ give the bdry

$$f_X(x) = \int f(x,y)\, dy$$

$$= \int_0^1 \frac{4}{3}(1 - xy)\, dy = \frac{4}{3}\left[y - \frac{x}{2} y^2 \right]_0^1 = \underline{\frac{4}{3}\left(1 - \frac{x}{2}\right)}.$$

$0 < x < 1$

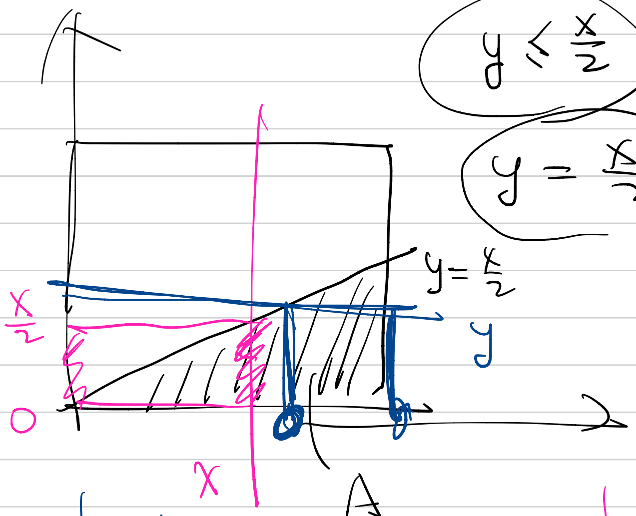
$f_Y(y) = \frac{4}{3}\left(1 - \frac{y}{2}\right)$ for $0 < y < 1$

$$P\left(Y \le \frac{X}{2}\right) = \int_0^1 \int_0^{\frac{x}{2}} f(x,y) \, dy \, dx$$

$$= P\left((X,Y) \in A\right) = \iint_A f(x,y) \, dx \, dy$$

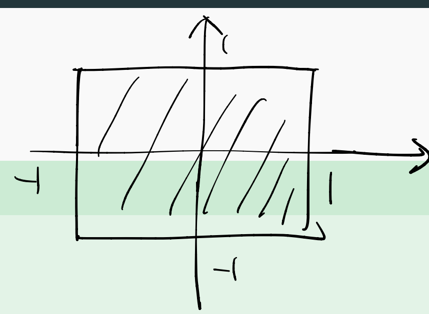
$y \le \frac{x}{2}$

$y = \frac{x}{2}$



$$\int_0^{\frac{1}{2}} \int_{2y}^{1} \qquad dx \, dy = \int_0^1 \int_0^{\frac{x}{2}} \frac{4}{3}(1 - xy) \, dy \, dx$$

Joint PDF



Example

Let X and Y have the joint PDF

$$f(x, y) = \frac{3}{2}x^2(1 - |y|)$$

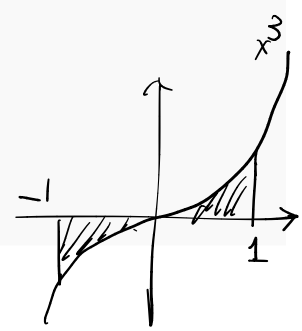
for $-1 < x, y < 1$.

Find $\mathbb{E}[X]$ and $\mathbb{E}[Y]$.

$$\mathbb{E}[X] = \int_{-1}^{1}\int_{-1}^{1} x \cdot f(x,y)\, dx\, dy$$

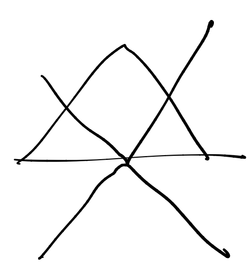
$$= \int_{-1}^{1}\int_{-1}^{1} \frac{3}{2} x^3 (1-|y|)\, dx\, dy$$

$$= \frac{3}{2}\int_{-1}^{1} (1-|y|) \Big(\underbrace{\int_{-1}^{1} x^3\, dx}_{=0}\Big)\, dy = 0.$$

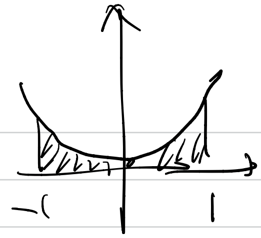


$$\int_{-1}^{1} (1-|y|)\, dy = 1$$

$1 - |y|$

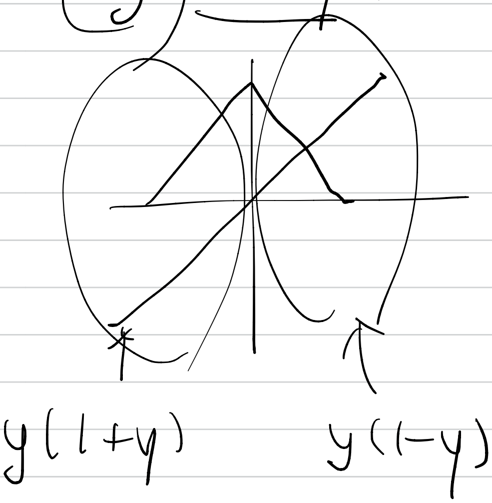


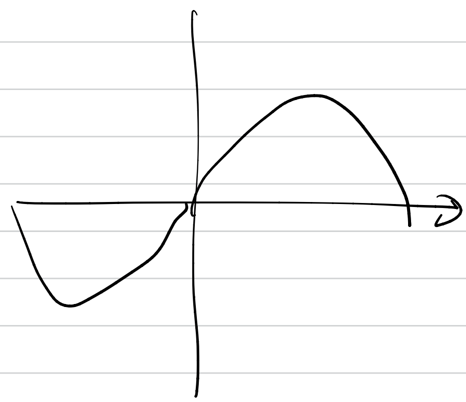
$$E[Y] = \int_{-1}^{1}\int_{-1}^{1} y \; f(x,y)\, dx\, dy$$



$$= \frac{3}{2}\underbrace{\left(\int_{-1}^{1} y(1-|y|)\,dy\right)}\underbrace{\left(\int_{-1}^{1} x^2\,dx\right)}_{= 2\int_0^1 x^2 dx \;=\; \frac{2}{3}} = 0$$

$y(1-|y|)$





$$y(1-|y|) = \begin{cases} y(1-y) & \text{if } y \geq 0 \\ y(1+y) & \text{if } y < 0 \end{cases}$$

$$\int_{-1}^{1} y(1-|y|)dy = \int_0^1 y(1-y)\,dy + \underbrace{\int_{-1}^{0} y(1+y)\,dy}_{= -\int_0^1 x(1-x)\,dx} = 0$$

$y = -x$

f is even if $f(-x) = f(x)$

$$\int_{-a}^{a} f(x)\,dx = 2\int_0^a f(x)\,dx$$

f is odd if $f(-x) = -f(x)$

$$\int_{-a}^{a} f(x)\,dx = 0$$

Independent random variables

Definition

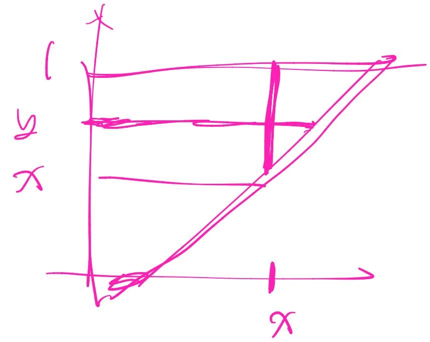
Two random variables X, Y with joint pdf are independent if and only if $f(x, y) = f_X(x) f_Y(y)$.

$$f_X(\boxed{x}) = \int f(x,y)\, dy$$

$$= \int_x^1 2 \, dy$$

$$= 2(1-x)$$

$$f_Y(y) = 2y = \int_0^y 2\, dy = 2y$$



Independent random variables

$X > 0$ $\quad x = 0$

$y > x$ $\quad y = x$

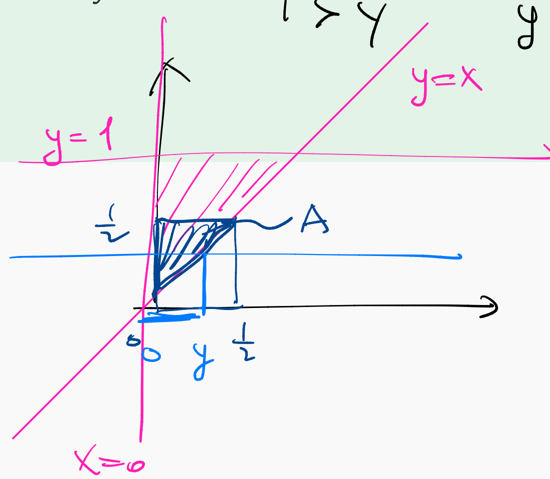
$1 > y$ $\quad y = 1$

Example

Let X and Y have the joint pdf $f(x, y) = 2$ for $0 < x < y < 1$.

Compute $\mathbb{P}(0 < X, Y < \frac{1}{2})$.

Are they independent? No.



$$\mathbb{P}(0 < X, Y < \tfrac{1}{2})$$

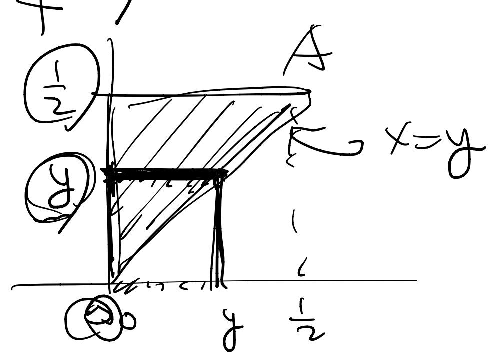
$$= \iint_A 2 \, dx\, dy$$

Method 1

$$= 2 \int_0^{\frac{1}{2}} \int_0^{y} 1 \, dx\, dy$$

Method 2

$$= 2 \cdot (\text{Area of } A) = \frac{1}{4}$$



X, Y have joint PDF

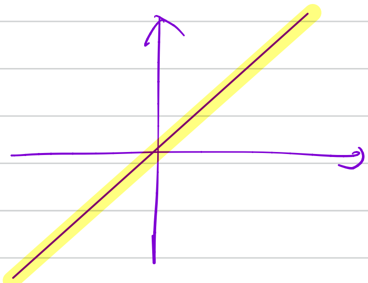
$$\begin{cases} f(x,y) \geq 0 \\ \int_{\mathbb{R}}\int_{\mathbb{R}} f(x,y)\,dx\,dy = 1 \\ \mathbb{P}((X,Y) \in A) = \iint_A f(x,y)\,dx\,dy \end{cases}$$

X, with PDF f_X (conti.)

$$\mathbb{P}(X = 5) = 0 = \lim_{\varepsilon \downarrow 0} \int_{5-\varepsilon}^{5+\varepsilon} f_X(x)\,dx$$

X, Y have joint PDF

$$\Rightarrow \quad \mathbb{P}(X = Y) = 0 = \iint_A f(x,y)\,dx\,dy$$



If X conti. $Y = X$ conti.

(X, Y) do not have joint PDF.

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

in terms of x

in terms of y

Conditional densities and Conditional Expectation

Definition

The conditional density of Y given $X = x$ is defined by

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

As in the discrete case, the conditional expectation and the conditional variance are defined by

$$\mathbb{E}[Y|X = x] = \int y f_{Y|X}(y|x)\, dy,$$

$$\mathrm{Var}(Y|X = x) = \mathbb{E}[(Y - \mathbb{E}[Y|X = x])^2|X = x].$$

$\mathrm{Var}(Y) = \mathbb{E}[\mathrm{Var}(Y|X)] + \mathrm{Var}(\mathbb{E}[Y|X])$

$\mathbb{E}[Y] = \mathbb{E}[\mathbb{E}[Y|X]]$

Conditional densities and Conditional Expectation

$f(x,y) > 0$ defined

"support" or "domain"

Example

Let X and Y have the joint PDF $f(x,y) = 2$ for $0 < x < y < 1$.

Then, $f_X(x) = 2(1-x)$ for $0 < x < 1$ and $f_Y(y) = 2y$ for $0 < y < 1$.

Find $\mathbb{E}[X|Y=y]$ and $\mathbb{E}[Y|X=x]$. $= \frac{1+x}{2}$

$\{0 < x < y < 1\}$ defines the region where $f(x,y) = 2$

otherwise $f(x,y) = 0$.

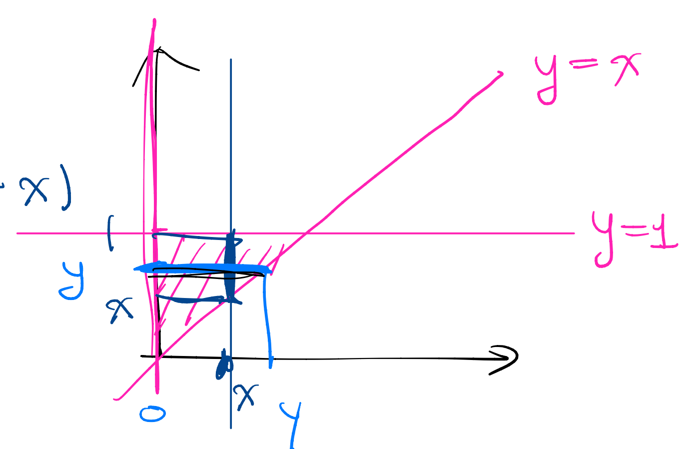
Three ineq.

$$\begin{cases} x > 0 \\ y > x \\ y < 1 \end{cases} \Rightarrow \begin{cases} x = 0 \leftarrow y\text{-axis} \\ y = x \\ y = 1 \end{cases}$$

define the bdry

$$f_X(\underset{\text{fixed}}{x}) = \int_x^1 f(x,y)\,dy = 2(1-x)$$

$$f_Y(\underset{\text{fixed}}{y}) = \int_0^y f(x,y)\,dx = 2y$$



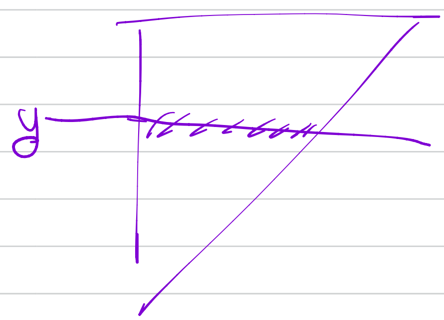
$$\mathbb{E}[X \mid \underset{\substack{\uparrow \\ \text{fixed}}}{Y=y}] = \int_0^y x \cdot \underbrace{f_{X|Y}(x|y)}_{= \frac{f(x,y)}{f_Y(y)} = \frac{\cancel{2}}{\cancel{2}y}}^{=\frac{1}{y}} dx$$

$$= \int_0^y \frac{x}{y}\, dx$$

$$= \frac{1}{y} \cdot \frac{1}{2} y^2 = \frac{y}{2} \quad \text{for } 0 < y < 1$$

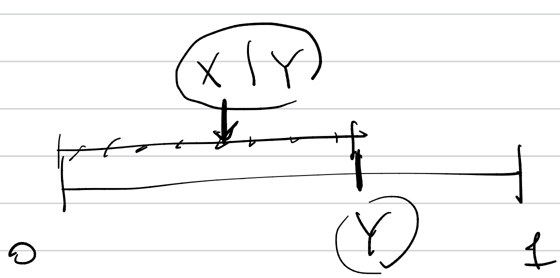
$$f_{X|Y}(x|\overset{\substack{\text{fixed} \\ \downarrow}}{y}) = \frac{\overset{=f(x,y)}{2}}{\underset{\substack{\shortparallel \\ f_Y(y)}}{2y}} = \frac{1}{y}, \qquad 0 < X < y$$

<u>$X \mid Y=y$</u> distribution?

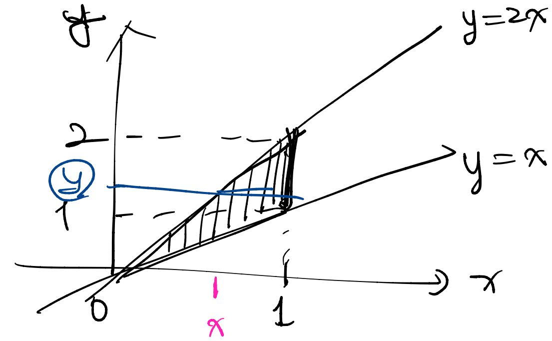


$X \mid Y=y \sim \text{Unif}(0, y)$

$\mathbb{E}[X \mid Y=y] = \frac{y}{2}$



$Y \sim \text{Unif}(0,1)$



$$f(x,y) = \overset{\frac{1}{x}}{f_{Y|X}(y|x)} \cdot \overset{1}{f_X(x)} = \frac{1}{x}$$

$$f_Y(y) = \int f(x,y)\, dx = \int \frac{1}{x}\, dx$$

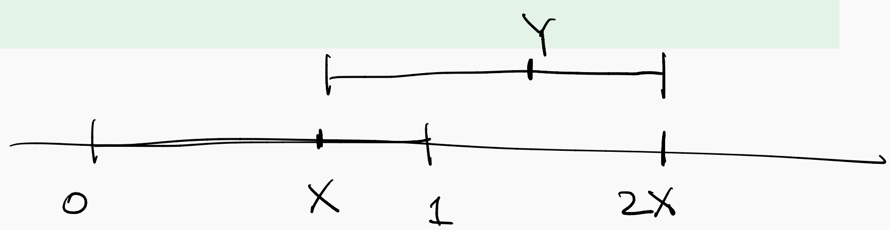
Conditional densities and Conditional Expectation

$$f_Y(y) =$$

Example

Let X be $U(0,1)$, and let the conditional distribution of Y, given $X = x$ be $U(x, 2x)$.

Find $\mathbb{E}[Y]$ and $\text{Var}(Y)$.



$$Y | X = x \sim U(x, 2x)$$

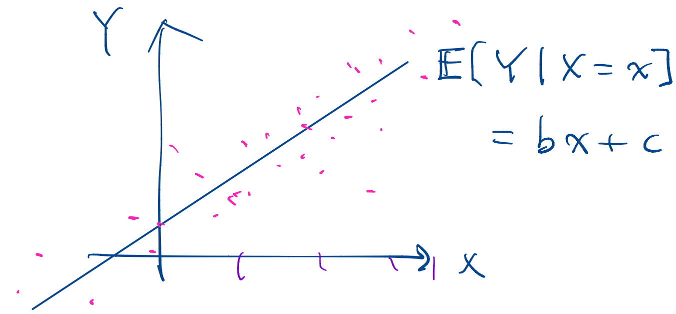
$$\mathbb{E}[Y] = \mathbb{E}[\mathbb{E}[Y | X]]$$
$$= \mathbb{E}\left[\frac{1}{2}(X + 2X)\right] = \frac{3}{2}\mathbb{E}[X] = \frac{3}{4}.$$

Exercise

Let $f(x, y) = 2e^{-x-y}$, $0 < x \leq y < 0$, be the joint pdf of X and Y.

Find $f_X(x)$ and $f_Y(y)$. Are X and Y independent?

Section 5.
The Bivariate Normal Distribution



Motivation

Let X be a random variable.

We construct a random variable Y in the following way:

The conditional distribution of Y given $X = x$ satisfies

1. it is normal for each x
2. $\mathbb{E}[Y|X = x]$ is linear in x
3. $\mathrm{Var}(Y|X = x)$ is constant in x

mean variance

$$Y \mid X = x \sim N\left(\rho \frac{\sigma_Y}{\sigma_X}(x - \mu_X) + \mu_Y,\ \sigma_Y^2(1-\rho^2)\right)$$

linear regression line

$$\mathbb{E}[Y \mid X = x] = bx + c \overset{\uparrow}{=} \rho \cdot \frac{\sigma_Y}{\sigma_X}(x - \mu_X) + \mu_Y$$

+ previous class

$$\sigma_Y^2 = \mathrm{Var}(Y) = \mathrm{Var}(\mathbb{E}[Y \mid X]) + \mathbb{E}[\underbrace{\mathrm{Var}(Y \mid X)}_{\text{Const.}}]$$

$$\Rightarrow \quad \mathrm{Var}(Y \mid X = x) = \sigma_Y^2 (1 - \rho^2)$$

size of error.

$$f_{Y|X}(y|x) = \frac{1}{\sqrt{2\pi} \cdot \sigma_Y \sqrt{1-\rho^2}} \exp\left(-\frac{1}{2\sigma_Y^2(1-\rho^2)}(x - \qquad)^2\right)$$

$\mu_X = \mathbb{E}[X]$, $\mu_Y = \mathbb{E}[Y]$, $\sigma_X^2 = Var(X)$, $\sigma_Y^2 = Var(Y)$

ρ = correlation coefficient $= \frac{Cov(X,Y)}{\sqrt{Var(X) \cdot Var(Y)}} = \frac{\mathbb{E}[XY] - \mu_X \mu_Y}{\sigma_X \cdot \sigma_Y}$

Motivation

Then, $Y|X = x$ is normal with mean $\mu_Y + \rho\frac{\sigma_Y}{\sigma_X}(x - \mu_X)$ and variance $\sigma_Y^2(1-\rho^2)$.

The conditional density is

$$f_{Y|X}(y|x) = \frac{1}{\sigma_Y\sqrt{2\pi}\sqrt{1-\rho^2}} \exp\left(-\frac{(y - (\mu_Y + \rho\frac{\sigma_Y}{\sigma_X}(x-\mu_X)))^2}{2\sigma_Y^2(1-\rho^2)}\right)$$

+

$X \sim N(\mu_X, \sigma_X^2)$

$$f(x,y) = f_{Y|X}(y|x) \cdot \underline{f_X(x)}$$

$$\| $$

$$\frac{1}{\sqrt{2\pi} \cdot \sigma_X} e^{-\frac{1}{2\sigma_X^2}(x-\mu_X)^2}$$

Bivariate normal distribution

If X itself has normal distribution, (X, Y) is called **a bivariate normal random variables**.

$E[Y \mid X=x]$

$= bx + c$ $\longrightarrow$ $Y \mid X = x \sim N(\quad , \quad)$

$Var(Y \mid X=x)$ $\qquad X \sim N(\mu_x, \sigma_x^2)$

$=$ constant

Bivariate normal distribution

$$\mu = \begin{pmatrix} E[X] \\ E[Y] \end{pmatrix}$$

Definition

We say (X, Y) has a bivariate normal distribution with mean vector $\begin{pmatrix} \mu_X \\ \mu_Y \end{pmatrix}$ and covariance matrix $\begin{pmatrix} \sigma_X^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_X\sigma_Y & \sigma_Y^2 \end{pmatrix} = \Sigma$ if its joint pdf is given by

$$f(x, y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp\left(\underbrace{-\frac{1}{2(1-\rho^2)} \left(\frac{\bar{x}^2}{\sigma_X^2} - 2\frac{\rho\bar{x}\bar{y}}{\sigma_X\sigma_Y} + \frac{\bar{y}^2}{\sigma_Y^2} \right)}_{-\frac{1}{2}(X-\mu_X \quad Y-\mu_Y)\,\Sigma^{-1}\begin{pmatrix} X-\mu_X \\ Y-\mu_Y \end{pmatrix}} \right)$$

where $\bar{x} = x - \mu_X$ and $\bar{y} = y - \mu_Y$.

$$\begin{pmatrix} \underbrace{Cov(X, X)}_{Var(X)} & Cov(X, Y) \\ Cov(Y, X) & \underbrace{Cov(Y, Y)}_{Var(Y)} \end{pmatrix}$$

(X, Y) bivariate Normal

$$\begin{cases} Y \mid X = x & \sim N\left(\rho \frac{\sigma_Y}{\sigma_X}(x-\mu_X) + \mu_Y, \ \sigma_Y^2(1-\rho^2)\right) \\ X & \sim N(\mu_X, \sigma_X^2) \\ Y & \sim N(\mu_Y, \sigma_Y^2) \end{cases}$$

Bivariate normal distribution

Example

Let us assume that in a certain population of college students, the respective grade point averages, say X and Y, in high school and the first year of college have a bivariate normal distribution with parameters $\mu_X = 2.9$, $\mu_Y = 2.4$, $\sigma_X = 0.4$, $\sigma_Y = 0.5$, and $\rho = 0.6$.

Find $\mathbb{P}(2.1 < Y < 3.3 | X = 3.2)$.

$\rho \cdot \frac{\sigma_Y}{\sigma_X}(3.2 - \mu_X) + \mu_Y = m$

$Y \mid X = 3.2 \sim N(m, s^2) \sim W$

$\sigma_Y^2(1-\rho^2) = s^2$

$P(2.1 < Y < 3.3 \mid X = 3.2)$

$= \mathbb{P}(2.1 < W < 3.3)$

$= \mathbb{P}\left(\frac{2.1 - m}{s} < \frac{W - m}{s} < \frac{3.3 - m}{s}\right)$ $\quad \frac{W-m}{s} \sim N(0,1)$

$= \Phi\left(\frac{3.3-m}{s}\right) - \Phi\left(\frac{2.1-m}{s}\right)$

Use the table.

Recall

(X, Y) Bivariate Normal

① $Y \mid X = x \sim N\left(\rho \frac{\sigma_Y}{\sigma_X}(x - \mu_X) + \mu_Y ,\ \sigma_Y^2 (1-\rho^2) \right)$

② $X \sim N(\mu_X, \sigma_X^2)$

$$f(x,y) = \frac{1}{2\pi \sigma_X \sigma_Y \sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)} \left(\frac{\bar{X}^2}{\sigma_X^2} - 2\rho \frac{\bar{X}}{\sigma_X} \cdot \frac{\bar{Y}}{\sigma_Y} + \frac{\bar{Y}^2}{\sigma_Y^2} \right) \right)$$

$\bar{X} = x - \mu_X \qquad \bar{Y} = y - \mu_Y$.

- X, Y uncorrelated if $\begin{cases} \rho = 0 = \dfrac{Cov(X,Y)}{\sqrt{Var(X)\, Var(Y)}} \\ Cov(X,Y) = 0 \end{cases}$

(positively correlated if $\rho > 0$
negatively " if $\rho < 0$)

X, Y Indep $\Rightarrow$ X, Y uncorrelated
$\not\Leftarrow$ in general

- (X, Y) Bivariate Normal & $\rho = 0$

$\Rightarrow$ (X, Y) Indep .

Bivariate normal distribution

Theorem

If X and Y have a bivariate normal distribution with correlation coefficient ρ, then X and Y are independent if and only if $\rho = 0$.

Exercise

For a female freshman in a health fitness program, let X equal her percentage of body fat at the beginning of the program and Y equal the change in her percentage of body fat measured at the end of the program.

Assume that X and Y have a bivariate normal distribution with

$\mu_X = 24.5$, $\mu_Y = -0.2$, $\sigma_X = 4.8$, $\sigma_Y = 3$, and $\rho = -0.32$.

Find $\mathbb{P}(1.3 < Y < 5.8)$, $\mathbb{E}[Y|X = x]$, and $\mathrm{Var}(Y|X = x)$.