Chapter 4. Bivariate Distributions

Math 3215 Summer 2023

Georgia Institute of Technology

Section 1.
Bivariate Distributions of the Discrete Type

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 $Y = u(X)$.

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.

(pmf $f(x) = \mathbb{P}(X = x)$)

Joint distribution

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

Note that

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

Same as before.

Joint distribution

$$f(x,y) = \begin{cases} 0 & x > y \\ 1/36 & x = y \\ 1/18 & x < y \end{cases} \qquad x, y = 1, \dots, 6.$$

Example

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).

$X \backslash Y$	1	2	3	4	5	6	$f_X(x)$
1	$\frac{1}{36}$	$1/18$	$1/18$	--	--	$1/18$	$11/36$
2	0	$1/36$	$1/18$				$9/36 = f_X(2)$
3	0	0	$1/36$				$7/36 = f_X(3)$
4	0	0	0	$1/36$			$5/36$
5	0	0	0	0	$1/36$	$1/18$	$3/36$
6	0	0	0	0	0	$1/36$	$1/36$
$f_Y(y)$	$1/36 = f_Y(1)$	$3/36 = f_Y(2)$	----			$11/36 = f_Y(6)$	

$$\sum_{y=1}^{6} f(1,y) = f_X(1)$$

Marginal distribution

Definition

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

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

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

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

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

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

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

Def X, Y indep. if for any events A, B

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

Marginal distribution

discrete type.

Definition

We say X and Y are independent if

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

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.

X, Y indep. with $f_{X,Y} = f_X \cdot f_Y$

$$\mathbb{E}[X \cdot Y] = \sum_x \sum_y x \cdot y \cdot f_{X,Y}(x,y)$$

$$= \sum_x \sum_y x \cdot y \cdot f_X(x) \cdot f_Y(y)$$

$$= \left(\sum_x x \cdot f_X(x)\right) \cdot \left(\sum_y y f_Y(y)\right)$$

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

But, $\mathbb{E}[X \cdot Y] = \mathbb{E}[X] \cdot \mathbb{E}[Y] \not\Rightarrow$ indep.

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 $y = \underline{\underline{1}}, \underline{\underline{2}}$.

Find the marginal pmfs of X and Y.

Determine whether they are independent.

$$f_X(x) = \sum_{y=1}^{2} \frac{1}{21}(x+y) = \frac{1}{21}\cdot\left((x+1)+(x+2)\right) = \frac{2x+3}{21}.$$

$$f_Y(y) = \sum_{x=1}^{3} \frac{1}{21}(x+y) = \frac{1}{21}\left((1+y)+(2+y)+(3+y)\right) = \frac{3y+6}{21}.$$

$$f_X(x)\cdot f_Y(y) = \frac{1}{(21)^2}\cdot(2x+3)\cdot(3y+6) \overset{?}{=} \frac{1}{21}(x+y)$$

$x=1,\ y=1,$

$$\frac{1}{(21)^2}\cdot 5\cdot 9 \overset{?}{\neq} \frac{1}{21}\cdot 2$$

not equal $\quad X, Y$ dep.

for some x, y

Marginal distribution

Example

Let the joint pmf of X and Y be defined by

$$f(x, y) = \frac{xy^2}{30}$$

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

Find the marginal pmfs of X and Y.

Determine whether they are independent.

$$f_X(x) = \sum_{y=1}^{2} f(x,y) = \frac{1}{30}(x \cdot 1^2 + x \cdot 2^2) = \frac{5x}{30} = \frac{x}{6}.$$

$$f_Y(y) = \sum_{x=1}^{3} \frac{1}{30} x y^2 = \frac{y^2}{30} \cdot (1+2+3) = \frac{y^2}{5}.$$

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

true for all $x = 1, 2, 3$, $y = 1, 2$

$\Rightarrow$ X & 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

ftn joint pmf

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

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

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

$$\| $$

$$\sum_{x_1, x_2} u(x_1, \overset{x_1}{\overset{\shortparallel}{}} x_2) \cdot f(x_1, x_2) = \sum_{x_1} x_1 \underbrace{\sum_{x_2} f(x_1, x_2)}_{= f_{X_1}(x_1)}$$

$$= E[X_1].$$

Recall X, Y discrete RVs

$f_{X,Y}(x,y) = P(X = x, Y = y)$: joint pmf of X, Y.

$f_X(x) = P(X = x) = \sum_y P(X = x, Y = y) = \sum_y f_{X,Y}(x,y)$: marginal pmf of X.

$$f_Y(y) = \mathbb{P}(Y=y) = \sum_x \mathbb{P}(X=x, Y=y)$$
$$= \sum_x f_{X,Y}(x,y)$$

$\underline{\mathbb{E}[u(X,Y)]}$ ex) $u(x,y) = x \cdot y$, $\mathbb{E}[u(X,Y)] = \mathbb{E}[X \cdot Y]$.

$$= \sum_x \sum_y u(x,y) \cdot f_{X,Y}(x,y)$$

$\underline{\text{Ex}}$) $u(x,y) = x+y$

$$\mathbb{E}[u(X,Y)] = \mathbb{E}[X+Y] = \sum_x \sum_y (x+y) \cdot f_{X,Y}(x,y)$$
$$= \sum_x \sum_y x \cdot f_{X,Y}(x,y) + \sum_x \sum_y y \, f_{X,Y}(x,y)$$
$$= \sum_x x \cdot \underbrace{\left(\sum_y f_{X,Y}(x,y) \right)}_{= f_X(x)} + \sum_y y \underbrace{\left(\sum_x f_{X,Y}(x,y) \right)}_{= f_Y(y)}$$
$$= \mathbb{E}[X] + \mathbb{E}[Y].$$

But, $\mathbb{E}[X \cdot Y] \neq \mathbb{E}[X] \cdot \mathbb{E}[Y]$ in general.

Expectations

$\mathbb{E}[X_1+X_2] = \mathbb{E}[X_1] + \mathbb{E}[X_2] = 2\mathbb{E}[X_1] = 2\cdot\left(0\cdot\frac{5}{8}+1\cdot\frac{3}{8}\right) = \frac{3}{4}.$

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, X_2) : the outcome.

$$\begin{cases} f_{X_1,X_2}(0,0) = 3/8 \\ f_{X_1,X_2}(0,1) = f_{X_1,X_2}(1,0) = 2/8 \\ f_{X_1,X_2}(1,1) = 1/8 \end{cases}$$

$X_1 \backslash X_2$	0	1		
0	3/8	2/8	5/8	$= f_{X_1}(0)$
1	2/8	1/8	3/8	$= f_{X_1}(1)$
	$f_{X_2}(0) = 5/8$	$3/8 = f_{X_2}(1)$		

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.

Trinomial distribution

Example

In manufacturing a certain item, it is found that in normal production about 95% of the items are good ones, 4% are "seconds," and 1% are defective.

A company has a program of quality control by statistical methods, and each hour an online inspector observes 20 items selected at random, counting the number X of seconds and the number Y of defectives.

Suppose that the production is normal.

Find the probability that, in this sample of size $n = 20$, at least two seconds or at least two defective items are discovered.

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?

X \ Y	2	3	4	5	6	7	8	$f_X(x)$
1	1/16	1/16	1/16	1/16	0	0	0	1/4
2	0	1/16	1/16	1/16	1/16	0	0	1/4
3	0	0	1/16	1/16	1/16	1/16	0	1/4
4	0	0	0	1/16	1/16	1/16	1/16	1/4
$f_Y(y)$	1/16	2/16	3/16	4/16	3/16	2/16	1/16	

Dependent.

Section 2.
The Correlation Coefficient

Covariance and Correlation coefficient

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

$\mu_Y = \mathbb{E}[Y]$

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

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

rho

$\sigma_X = \sqrt{\text{Var}(X)} = \text{std}(X)$

$\sigma_Y = \sqrt{\text{Var}(Y)}$

Covariance and Correlation coefficient

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

$\Uparrow$ $\Downarrow$

Properties

1. If X and Y are independent, then $\mathrm{Cov}(X, Y) = 0$.
2. $\mathrm{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$.
3. $-1 \le \rho \le 1$.

proof)

$$\begin{aligned}
\mathrm{Cov}(X, Y) &= \mathbb{E}\big[(X - \mathbb{E}[X]) \cdot (Y - \mathbb{E}[Y])\big] \\
&= \mathbb{E}\big[XY - \mathbb{E}[X] \cdot Y - \mathbb{E}[Y] \cdot X + \mathbb{E}[X] \cdot \mathbb{E}[Y]\big] \\
&= \mathbb{E}[XY] - \mathbb{E}\big[\mathbb{E}[X] \cdot Y\big] - \mathbb{E}\big[\mathbb{E}[Y] \cdot X\big] + \mathbb{E}[X] \cdot \mathbb{E}[Y] \\
&= \boxed{\mathbb{E}[XY] - \mathbb{E}[X] \cdot \mathbb{E}[Y]} - \cancel{\mathbb{E}[Y] \cdot \mathbb{E}[X]} + \cancel{\mathbb{E}[X] \cdot \mathbb{E}[Y]}
\end{aligned}$$

Covariance and Correlation coefficient

$\overline{X} = X - \mu_X = X - \mathbb{E}[X]$ $\Rightarrow \operatorname{Cov}(X,Y)$

$\overline{Y} = Y - \mu_Y = X - \mathbb{E}[Y]$ $= \mathbb{E}[\overline{X} \cdot \overline{Y}]$

Properties

1. If X and Y are independent, then $\operatorname{Cov}(X, Y) = 0$.
2. $\operatorname{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$.
3. $-1 \leq \rho \leq 1$.

$0 \leq \mathbb{E}\left[\underbrace{(\overline{Y} - t\,\overline{X})^2}_{\geq 0}\right] = \mathbb{E}\left[\overline{Y}^2 - 2t \cdot \overline{X} \cdot \overline{Y} + t^2\,\overline{X}^2\right]$

for all t

$= \underbrace{\mathbb{E}[\overline{Y}^2]}_{\substack{\parallel \\ \operatorname{Var}(Y)}} - 2t\,\underbrace{\mathbb{E}[\overline{X}\,\overline{Y}]}_{\substack{\parallel \\ \operatorname{Cov}(X,Y)}} + t^2 \cdot \underbrace{\mathbb{E}[\overline{X}^2]}_{\substack{\parallel \\ \operatorname{Var}(X)}}$

minimize in t

$a = \sigma_X^2$

$b = \operatorname{Cov}(X,Y)$

$c = \sigma_Y^2$

$\underbrace{a t^2 - 2bt + c}_{\text{minimum?}} = \underline{c - \frac{b^2}{a}} = \sigma_Y^2 - \frac{\operatorname{Cov}(X,Y)^2}{\sigma_X^2} \geq 0$

$2at - 2b = 0$ $\quad t = \frac{b}{a} = \frac{\operatorname{Cov}(X,Y)}{\sigma_X^2}$

$$\sigma_Y^2 - \frac{\mathrm{Cov}(X,Y)^2}{\sigma_X^2} \geq 0$$

$$1 \geq \left(\frac{\mathrm{Cov}(X,Y)}{\sigma_X \cdot \sigma_Y}\right)^2$$

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

Covariance and Correlation coefficient

Properties

1. If X and Y are independent, then $\mathrm{Cov}(X, Y) = 0$.
2. $\mathrm{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$.
3. $-1 \leq \rho \leq 1$.

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 $\mathrm{Cov}(X,Y)$ and ρ.

① $f_X(x) = f(x,1) + f(x,2) = \frac{x+2}{18} + \frac{x+4}{18} = \frac{2x+6}{18}$

$= \frac{x+3}{9}$

$f_Y(y) = f(1,y) + f(2,y) = \frac{1+2y}{18} + \frac{2+2y}{18}$

$= \frac{3+4y}{18}$

② $E[X] = 1\cdot\frac{(1+3)}{9} + 2\cdot\frac{(2+3)}{9} = \frac{14}{9}$

$E[Y] = 1\cdot\frac{(3+4)}{18} + 2\cdot\frac{(3+8)}{18} = \frac{29}{18}$

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

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

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

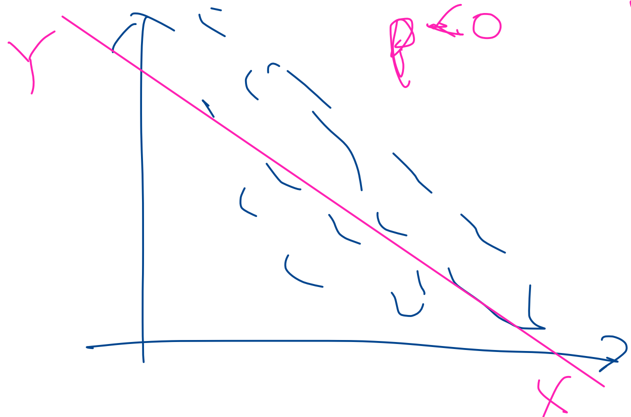
$$= \frac{1}{18}(3 + 10 + 8 + 24) = \frac{45}{18} = \frac{5}{2}$$

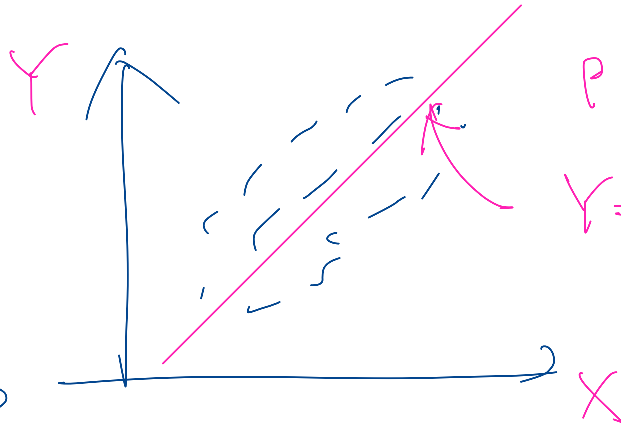
$$Cov(X, Y) = E[X \cdot Y] - E[X] \cdot E[Y]$$

$$= 5/2 - (14/9) \cdot (29/8).$$

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

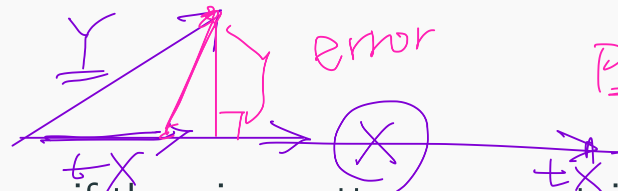
(skip...)





$$Y=\mu_Y+\rho\frac{\sigma_Y}{\sigma_X}(X-\mu_X)$$

The Least Squares Regression Line



projection = conditional expectation

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 $\underbrace{Y-\mu_Y}_{\bar{Y}} = b\underbrace{(X-\mu_X)}_{\bar{X}}$ has small errors.

$$\bar{Y} = b\bar{X}$$

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

$$\min_b \mathbb{E}[(\bar{Y}-b\bar{X})^2]$$

$$b=\frac{Cov(X,Y)}{Var(X)}=\rho\cdot\frac{\sigma_Y}{\sigma_X}$$

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

The Least Squares Regression Line

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

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

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.

The Least Squares Regression Line

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.

Uncorrelated

$E[X \cdot Y] = E[X] \cdot E[Y]$

$Cov(X, Y) = 0$

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

If X, Y are independent then they are uncorrelated.

However, the converse is not true.

X, Y positively correlated if $\rho > 0$

" negatively " if $\rho < 0$

Uncorrelated

Example

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

$$E[X] = 0 \cdot \frac{1}{3} + 1 \cdot \frac{1}{3} + 2 \cdot \frac{1}{3} = 1$$

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

$$E[X \cdot Y] = 2 \cdot 1 \cdot \frac{1}{3} = \frac{2}{3}$$

$$Cov(X,Y) = E[XY] - E[X] \cdot E[Y]$$

$$= \frac{2}{3} - 1 \cdot \frac{2}{3} = 0.$$

X, Y uncorrelated.

$P(X=0) = \frac{1}{3}$ $\quad$ $P(Y=1) = \frac{2}{3}$

$P(X=0, Y=1) = \frac{1}{3}$ $\quad$ not indep.

Exercise

(0, 0) (1, 0) (0, 1)

(2, 0) (0, 2) (1, 1)

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

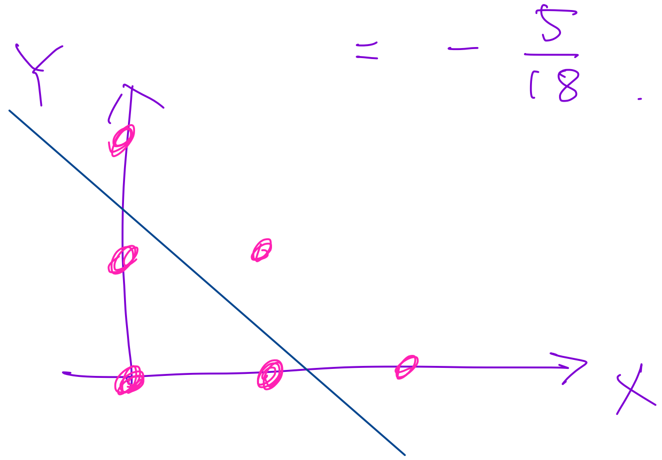
Find the covariance and the correlation coefficient. (σ?.)

$$E[X] = 0 \cdot \frac{3}{6} + 1 \cdot \frac{2}{6} + 2 \cdot \frac{1}{6} = \frac{2}{3} = E[Y]$$

$$E[XY] = 1 \cdot 1 \cdot \frac{1}{6} = \frac{1}{6}.$$

$$\operatorname{Cov}(X, Y) = \frac{1}{6} - \left(\frac{2}{3}\right)^2 = \frac{1}{6} - \frac{4}{9} = \frac{3-8}{18}$$

$$= -\frac{5}{18}.$$



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

$$E[X^2] = 0 \cdot \frac{3}{6} + 1^2 \cdot \frac{2}{6} + 2^2 \cdot \frac{1}{6}$$

$$= 1$$

$\operatorname{Var}(Y)$ ''

$$\operatorname{Var}(X) = 1 - \left(\frac{2}{3}\right)^2 = \frac{5}{9}$$

Section 3.
Conditional Distributions

Conditional distribution

Definition

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

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

$$\| $$

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

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

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.

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

$$= \frac{(x+y)/\cancel{21}}{(3y+6)/\cancel{21}} = \frac{x+y}{3(y+2)}.$$

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

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

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

Conditional distribution

$f_X(x) = \dfrac{2x+3}{21}$

$f_{Y|X}(y|x) = \dfrac{x+y}{2x+3}$

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

$$\mathbb{E}[Y \mid X=3] = \sum_y y \cdot \underline{f_{Y|X}(y \mid 3)}$$

$$= \sum_y y \cdot \frac{(3+y)}{9} = 1 \cdot \frac{4}{9} + 2 \cdot \frac{5}{9}$$

$$= \frac{14}{9}.$$

$$\mathbb{E}[Y^2 \mid X=3] = 1^2 \cdot \frac{4}{9} + 2^2 \cdot \frac{5}{9} = \frac{24}{9}$$

$$\mathrm{Var}(Y \mid X=3) = \mathbb{E}\left[\left(Y - \mathbb{E}[Y \mid X=3]\right)^2 \mid X=3\right]$$

$$= \mathbb{E}[Y^2 \mid X=3] - \left(\mathbb{E}[Y \mid X=3]\right)^2$$

$$= \frac{24}{9} - \left(\frac{14}{9}\right)^2 = \frac{1}{81}(216 - 196)$$

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

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

Contional expectation as a function and a random variable

Theorem

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

Contional expectation as a function and a random variable

Example

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 $n = x + 1$ and probability of success p.

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

Linear case

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 case

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

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?

Section 4.
Bivariate Distributions of the Continuous Type

Joint PDF

Definition

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 $\mathbb{P}(Y \leq \frac{X}{2})$.

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 $\mathbb{E}[Y]$.

Independent random variables

Definition

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

Independent random variables

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?

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

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

Conditional densities and Conditional Expectation

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

Conditional densities and Conditional Expectation

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 $\text{Var}(Y)$.

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

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

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

Bivariate normal distribution

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

Definition

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

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

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

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

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 $\mathrm{Var}(Y|X = x)$.