Math 3501 - Probability and Statistics I

4.3 - Conditional distributions

Motivation

Let X and Y have joint discrete distribution with pmf $f(x, y)$ on the space S, and marginal pmfs $f_X(x)$ and $f_Y(y)$ with spaces S_X and S_Y, respectively.

Define the events

$$A = \{X = x\} \quad \text{and} \quad B = \{Y = y\}\ , \qquad \boxed{(x, y) \in S}$$

and note that

$$A \cap B = \{X = x, Y = y\}\ .$$

Since

$$P(A \cap B) = P(X = x, Y = y) \overset{\text{def}}{=} f(x, y)$$

and

$$P(B) = P(Y = y) = f_Y(y) > 0 \quad (\text{because } y \in S_Y),$$

the conditional probability of A given B is

$$P(X=x \mid Y=y) = \underbrace{P(A \mid B)}_{\text{Chp 1}} = \frac{P(A \cap B)}{P(B)} = \frac{f(x, y)}{f_Y(y)}\ . \leftarrow \text{conditional pmf of } X \text{ given } Y=y$$

Conditional probability mass function

Definition (Conditional probability mass function)

Let X and Y have joint discrete distribution with pmf $f(x, y)$ on the space S, and marginal pmfs $f_X(x)$ and $f_Y(y)$ with spaces S_X and S_Y, respectively.

(a) The *conditional probability mass function of* X, *given that* $Y = y$, is given by

$$g(x \mid y) = \frac{f(x, y)}{f_Y(y)}, \quad \text{provided that } f_Y(y) > 0 .$$

(b) The *conditional probability mass function of* Y, *given that* $X = x$, is given by

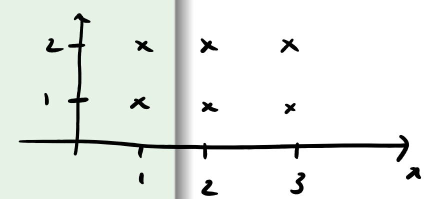
$$h(y \mid x) = \frac{f(x, y)}{f_X(x)}, \quad \text{provided that } f_X(x) > 0 .$$

Example

Let X and Y have the joint pmf

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

S :



Determine:

- the conditional probability mass function of X, given that $Y = y$.
- the conditional probability mass function of Y, given that $X = x$.

$S_X = \{1,2,3\}$

$S_Y = \{1,2\}$

We start by determining the marginal pmfs for X and Y:

$$f_X(x) = \sum_{y \in S_Y} f(x,y) = \sum_{y=1}^{2} \frac{x+y}{21} = \frac{x+1}{21} + \frac{x+2}{21} = \frac{2x+3}{21}, \quad x = 1,2,3$$

$$f_Y(y) = \sum_{x \in S_X} f(x,y) = \sum_{x=1}^{3} \frac{x+y}{21} = \frac{1+y}{21} + \frac{2+y}{21} + \frac{3+y}{21} = \frac{6+3y}{21}, \quad y = 1,2$$

conditional pmf of X given $Y=y$ is $g(x|y) = \dfrac{f(x,y)}{f_Y(y)} = \dfrac{\frac{x+y}{21}}{\frac{6+3y}{21}} = \dfrac{x+y}{6+3y}$ for $x = 1,2,3$

for each $y = 1,2$

we have two functions of x: $g(x|1)$ and $g(x|2)$

$$\left.\begin{array}{l} f(x,y) = P(X=x,\ Y=y) \\ f_Y(y) = P(Y=y) \end{array}\right\} \Rightarrow g(x|y) = \frac{f(x,y)}{f_Y(y)} = \frac{P(X=x \text{ and } Y=y)}{P(Y=y)}$$

$$= P(X=x \mid Y=y)$$

Conditional pmf of Y given $X=x$

$$h(y|x) = \frac{f(x,y)}{f_X(x)} = \frac{\frac{x+y}{21}}{\frac{2x+3}{21}} = \frac{x+y}{2x+3} \quad \text{for } y=1,2 \quad \text{for each } x=1,2,3$$

This represents three functions of y: $h(y|1)$, $h(y|2)$, $h(y|3)$

Alternative approach

$f_Y(y)$	y			
12/21	2	3/21	4/21	5/21
9/21	1	2/21	3/21	4/21
	x	1	2	3
	$f_X(x)$	5/21	7/21	9/21

$$f_X(x) = \begin{cases} 5/21, & \text{if } x = 1 \\ 7/21, & \text{if } x = 2 \\ 9/21, & \text{if } x = 3 \end{cases}$$

$$f_Y(y) = \begin{cases} 9/21 & \text{if } y = 1 \\ 12/21 & \text{if } y = 2 \end{cases}$$

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

	$x=1$	$x=2$	$x=3$	
$y=2$	3/12	4/12	5/12	← $g(x\|2)$
$y=1$	2/9	3/9	4/9	← $g(x\|1)$

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

	$x=1$	$x=2$	$x=3$
$y=2$	3/5	4/7	5/9
$y=1$	2/5	3/7	4/9

Note: the conditional probability mass functions $g(x \mid y)$ and $h(y \mid x)$ satisfy the conditions of a probability mass function.

The claim above follows from observing that $0 \leq h(y \mid x)$ and

$$\sum_y h(y \mid x) = \sum_y \frac{f(x,y)}{f_X(x)} = \frac{f_X(x)}{f_X(x)} = 1 \ ,$$

with a similar reasoning applying to $g(x \mid y)$.

Consequence: we can compute conditional probabilities such as

$$P(a < Y < b \mid X = x) = \sum_{\{y : a < y < b\}} h(y \mid x)$$

and conditional expectations such as

$$E[u(Y) \mid X = x] = \sum_y u(y) h(y \mid x)$$

as done with unconditional probabilities and expectations.

EXAMPLE (previous example continued).

For the r.v.s X, Y with joint pmf $f(x,y) = \dfrac{x+y}{21}$, $x = 1,2,3$ and $y = 1,2$

Compute

$$P(X \leq 2 \mid Y=2) = \sum_{x \leq 2} g(x \mid 2) = \sum_{x=1}^{2} g(x \mid 2)$$

$$= \frac{3}{12} + \frac{4}{12} = \frac{7}{12}$$

Conditional mean and conditional variance

The conditional mean of X, given that $Y = y$, is defined as

$$\mu_{X|y} = E[X \mid y] = \sum_x x g(x \mid y) \ ,$$

and the conditional variance of X, given that $Y = y$, is defined as

$$\sigma^2_{X|y} = \mathrm{Var}(X \mid y) = \underbrace{E\left[(X - E[X \mid y])^2 \mid y\right]}_{E\left[(X - \mu_{X|y})^2 \mid y\right]} = \sum_x (x - E[X \mid y])^2 g(x \mid y) \ .$$

The latter may be computed using

$$\sigma^2_{X|y} = E\left[X^2 \mid y\right] - (E[X \mid y])^2 \ .$$

analogue to

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

Similarly, the conditional mean of Y, given that $X = x$, is defined as

$$\mu_{Y|x} = E[Y \mid x] = \sum_y y h(y \mid x) \ ,$$

and the conditional variance of Y, given that $X = x$, is defined as

$$\sigma^2_{Y|x} = \text{Var}(Y \mid x) = E\left[(Y - \underbrace{E[Y \mid x]}_{\mu_{Y|x}})^2 \mid x\right] = \sum_y (y - E[Y \mid x])^2 h(y \mid x) \ .$$

The latter may be computed using

$$\sigma^2_{Y|x} = E\left[Y^2 \mid x\right] - (E[Y \mid x])^2 \ .$$

Example (Previous example continued)

Let X and Y have the joint pmf

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

conditioning on $X=x$

Determine:

- $E[Y \mid x]$ for $x = 1, 2, 3$

$$E[Y \mid X=1] = \sum_{y \in S_y} y \cdot h(y|1) = 1 \cdot h(1|1) + 2 \cdot h(2|1) = 1 \cdot \frac{2}{5} + 2 \cdot \frac{3}{5} = \frac{8}{5}$$

$$E[Y \mid X=2] = \sum_{y \in S_y} y \cdot h(y|2) = 1 \cdot h(1|2) + 2 \cdot h(2|2) = 1 \cdot \frac{3}{7} + 2 \cdot \frac{4}{7} = \frac{11}{7}$$

$$E[Y \mid X=3] = \sum_{y \in S_y} y \cdot h(y|3) = 1 \cdot h(1|3) + 2\, h(2|3) = 1 \cdot \frac{4}{9} + 2 \cdot \frac{5}{9} = \frac{14}{9}$$

X because we don't yet know which values X will take

Note: $E[Y \mid X]$ is the random variable taking the values $E[Y \mid x]$ for $x \in S_X$, with probabilities given by the marginal pmf of X.

- Recall previous example!

↳ $E[Y|X]$ takes the values:

$$E[Y|X] = \begin{cases} E[Y|X=1] = \frac{8}{5} & \text{with probability } f_X(1) = \frac{5}{21} \\ E[Y|X=2] = \frac{11}{7} & \text{with probability } f_X(2) = \frac{7}{21} \\ E[Y|X=3] = \frac{14}{9} & \text{with probability } f_X(3) = \frac{9}{21} \end{cases}$$

Consequence: Because $E[Y \mid X]$ is a random variable, its expectation may exist, in which case it is given by

$$E[E[Y \mid X]] = \sum_x \underbrace{E[Y \mid x]} \underbrace{f_X(x)} \,.$$

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

$$= \frac{33}{21} = E[Y]$$

Law of total probability for expectation

Theorem

Let X and Y be random variables such that $E(Y)$ exists. It holds that

$$\boxed{E[E[Y \mid X]] = E[Y] \ .}$$

Example (Previous example continued)

Let X and Y have the joint pmf

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

Determine:

- $\mathrm{Var}(Y \mid x)$ for $x = 1,2,3$

$$E[Y^2 \mid X=1] = \sum_{y \in S_y} y^2 \cdot h(y \mid 1) = 1^2 \cdot h(1 \mid 1) + 2^2 \cdot h(2 \mid 1) = 1 \cdot \frac{2}{5} + 4 \cdot \frac{3}{5} = \frac{14}{5}$$

$$E[Y^2 \mid X=2] = \sum_{y \in S_y} y^2 \cdot h(y \mid 2) = 1^2 \cdot h(1 \mid 2) + 2^2 \cdot h(2 \mid 2) = 1 \cdot \frac{3}{7} + 4 \cdot \frac{4}{7} = \frac{19}{7}$$

$$E[Y^2 \mid X=3] = \sum_{y \in S_y} y^2 \cdot h(y \mid 3) = 1^2 \cdot h(1 \mid 3) + 2^2 \cdot h(2 \mid 3) = 1 \cdot \frac{4}{9} + 4 \cdot \frac{5}{9} = \frac{24}{9}$$

$$\mathrm{Var}(Y \mid X=1) = E[Y^2 \mid X=1] - (E[Y \mid X=1])^2 = \frac{14}{5} - \left(\frac{8}{5}\right)^2 = \frac{6}{25}$$

$$\mathrm{Var}(Y \mid X=2) = E[Y^2 \mid X=2] - (E[Y \mid X=2])^2 = \frac{19}{7} - \left(\frac{11}{7}\right)^2 = \frac{12}{49}$$

$$\mathrm{Var}(Y \mid X=3) = E[Y^2 \mid X=3] - (E[Y \mid X=3])^2 = \frac{24}{9} - \left(\frac{14}{9}\right)^2 = \frac{20}{81}$$

Note: $\text{Var}(Y \mid X)$ is the random variable taking the values $\text{Var}(Y \mid x)$ for $x \in S_X$, with probabilities given by the marginal pmf of X.

- Recall previous example!

$$\text{Var}(Y|X) = \begin{cases} \text{Var}(Y|X=1) = \frac{6}{25} & \text{with probability } f_X(1) = 5/21 \\ \text{Var}(Y|X=2) = \frac{12}{49} & \text{" " } f_X(2) = 7/21 \\ \text{Var}(Y|X=3) = \frac{20}{81} & \text{" " } f_X(3) = 9/21 \end{cases}$$

Consequence: Because $\text{Var}(Y \mid X)$ is a random variable, its expectation may also exist, in which case it is given by

$$E[\text{Var}(Y \mid X)] = \sum_x \text{Var}(Y \mid x) f_X(x) \ .$$

calculate

$$E[\text{Var}(Y|X)] = \frac{6}{25} \cdot \frac{5}{21} + \frac{12}{49} \cdot \frac{7}{21} + \frac{20}{81} \cdot \frac{9}{21} = \cdots$$

Law of total probability for variances

Theorem

If X and Y are random variables, then

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

provided that all of the expectations and variances exist.

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.

Determine $E[Y]$ and $\mathrm{Var}(Y)$.

We know : $X \sim \text{Poisson}(\underset{\uparrow \; \lambda = 4}{4})$ and $Y \mid X = x \sim \text{Binomial}(\overbrace{x+1}^{n}, p)$

We will use the laws of total probability to determine $E[Y]$ and $\mathrm{Var}(Y)$:

$$E[Y] = E[\underbrace{E[Y|X]}] = E[(X+1)p] = p\,E[X+1] = p(\overbrace{E[X]}^{4}+1) = p(4+1) = 5p$$

Since $Y|X \sim \text{Binomial}(X+1, p)$ then $E[Y|X] = (X+1)\cdot p$

$$\mathrm{Var}(Y) = E[\underbrace{\mathrm{Var}(Y|X)}_{(X+1)p\cdot(1-p)}] + \mathrm{Var}(\underbrace{E[Y|X]}_{(X+1)p}) = E[(X+1)\cdot p(1-p)] + \mathrm{Var}((X+1)\cdot p)$$

$n \cdot p \cdot q \rightarrow (X+1)\,p\cdot(1-p)$

$$= p(1-p)\,E[X+1] + p^2\,\mathrm{Var}(X+1)$$

$$= p(1-p)\cdot\left(E[X]+1\right) + p^2\,\mathrm{Var}(X)$$

$$= p(1-p)\left(4+1\right) + p^2\cdot 4$$

$$= 5p(1-p) + 4p^2 = 5p - p^2$$

if $X \sim \mathrm{Poisson}(\lambda)$

$E[X] = \mathrm{Var}(X) = \lambda$ ← from formula sheet

Math 3501 - Probability and Statistics I

4.4 - Bivariate distributions of the continuous type

extends the concepts introduced in Secs 4.1, 4.2, and 4.3
to the case of continuous-type r.v.s.

Joint probability density function

Definition (Joint probability density function)

Let X and Y be two continuous random variables, and let S denote the corresponding two-dimensional space of X and Y (also referred to as the support of X and Y).

The *joint probability density function* (abbreviated joint pdf) of X and Y is an integrable real-valued function $f(x, y)$ with the following properties:

(a) $f(x, y) \geq 0$ for all $(x, y) \in \mathbb{R}^2$ with $f(x, y) = 0$ when $(x, y) \notin S$.

f(x,y) is only positive on the support S of X and Y

(b) $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)\, \mathrm{d}x\mathrm{d}y = 1.$ $\iff$ $\iint_S f(x,y)\,dx\,dy = 1$

(c) $P[(X, Y) \in A] = \iint_A f(x, y)\, \mathrm{d}x\mathrm{d}y$ where $A \subset \mathbb{R}^2$.

Example

Determine the constant c that makes the function

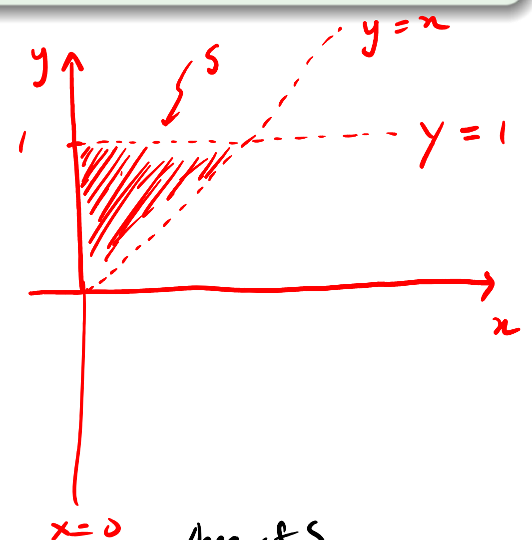
$$\boxed{f(x,y) = c\ , \quad 0 < x < y < 1}$$

into a joint pdf of two continuous random variables X and Y.

First identify the support of X and Y

$$S = \{ (x,y) \in \mathbb{R}^2 : \quad 0 < x < y < 1 \}$$

It is helpful to sketch S



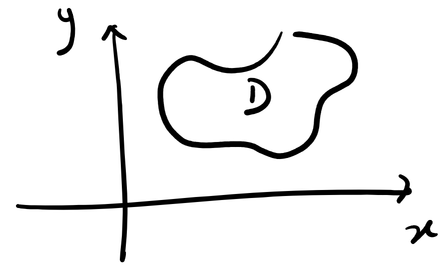
We must have

i) $c > 0$

ii) $$\iint_S f(x,y)\,dx\,dy = 1 \iff \iint_S c\,dx\,dy = 1 \iff c \overbrace{\iint_S 1 . dx\,dy}^{\text{Area of } S} = 1 \iff c . \frac{1}{2} = 1$$

$$\iff \boxed{c = 2}$$

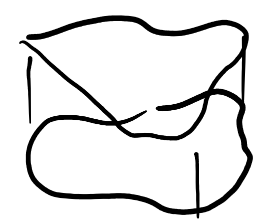
Recall for a region D on the plane



$$\text{Area}(D) = \iint_D 1 \, dx\, dy$$

which also happens to be equal to the volume of the cylinder of height 1 and base D

Volume = Area(S)



$$\text{Volume} = \text{Area}(S) \times \underbrace{\text{height}}_{=1} = \text{Area}(S)$$

Example (Continued)

Let X and Y have the joint pdf

$c = 2$

$$f(x, y) = 2\,, \quad 0 < x < y < 1\,.$$

Determine $P(X \leq 1/2)$.

$$P\left(X \leqslant 1/2\right) =$$

$$= \int_0^{1/2} \int_x^1 f(x,y)\, dy\, dx$$

$$= \int_0^{1/2} \int_x^1 2\, dy\, dx = \int_0^{1/2} \Big[2y\Big]_{y=x}^{y=1} dx$$

$$= \int_0^{1/2} 2 - 2x\, dx = \Big[2x - x^2\Big]_{x=0}^{x=1/2} = 1 - \frac{1}{4} = \frac{3}{4}$$