MICHIGAN STATE UNIVERSITY

AFRE 802
Statistical Methods for Agricultural, Food, & Resource Economists

Multivariate probability distributions – Part 2 of 3
(WMS Ch. 5.3-5.4)

October 10, 2017

Nicole Mason
Michigan State University
Fall 2017

MICHIGAN STATE UNIVERSITY

GAME PLAN

- Return HWs and in-class exercises
- No graded in-class exercise but expect one on Thursday
- Reminder about mid-term (Thurs., Oct. 19)
- Review

Multivariate probability distributions (Part 2 of 3)

1. Finish bivariate probability distributions for continuous RVs
2. Marginal probability distributions
3. Conditional probability distributions
4. Independent random variables

1

MICHIGAN STATE UNIVERSITY

Reminder: Mid-term is next Thurs., Oct. 19

- **Start studying NOW** (if you haven't already)
- Will cover through end of Ch. 5 (multivariate probability distributions)
- One (double-sided) 8.5 x 11 inch cheat sheet
- I will provide tables
- Past mid-terms and answers are on D2L
- Discuss good exam prep and test-taking strategies

2

Review

MICHIGAN STATE UNIVERSITY

Bivariate probability distributions for discrete RVs

Probability distribution :
$$p(y_1, y_2) = P(Y_1 = y_1, Y_2 = y_2)$$

Properties :

1. $0 \le p(y_1, y_2) \le 1$
2. $\sum_{y_1, y_2} p(y_1, y_2) = 1$

EX) $p(y_1, y_2)$ for discrete PDF
(# of customers going to supermarket counter 1 vs. 2)

	y_1		
y_2	0	1	2
0	1/9	2/9	1/9
1	2/9	2/9	0
2	1/9	0	0

3

Review: questions on this?

Bivariate CDFs for discrete RVs

$$F(y_1,y_2) = P(Y_1 \le y_1, Y_2 \le y_2) = \sum_{t_2 \le y_2} \sum_{t_1 \le y_1} p(t_1,t_2)$$

EX) In the supermarket checkout counter example, find:

a. *F(-1, 2)=0 because Y_1 can't be less than zero*

b. *F(1.5, 2)=8/9 because everything but (2,0)*

c. *F(5, 7) =1 because all valid values of Y_1 and Y_2 are less than 5 and 7, respectively*

y_2	y_1: 0	1	2
0	1/9	2/9	1/9
1	2/9	2/9	0
2	1/9	0	0

Review

Both discrete & continuous bivariate CDFs must satisfy similar properties to what we saw in the univariate case:

1. $F(-\infty,-\infty) = 0,\ F(-\infty, y_2) = 0,\ F(y_1,-\infty) = 0$

2. $F(\infty,\infty) = 1$

Bivariate CDF (discrete RVs)

$$F(y_1,y_2) = P(Y_1 \le y_1, Y_2 \le y_2) = \sum_{t_2 \le y_2} \sum_{t_1 \le y_1} p(t_1,t_2)$$

Bivariate CDF (continuous RVs)

$$F(y_1,y_2) = P(Y_1 \le y_1, Y_2 \le y_2) = \int_{-\infty}^{y_2}\int_{-\infty}^{y_1} f(t_1,t_2)\,dt_1\,dt_2$$

where $-\infty < y_1 < \infty, \quad -\infty < y_2 < \infty$

5

Review

Bivariate PDFs (continuous RVs)

$$\boxed{f(y_1, y_2)}$$

Like PDFs for the univariate case (and bivariate probability distributions for discrete RVs), bivariate PDFs must satisfy similar properties:

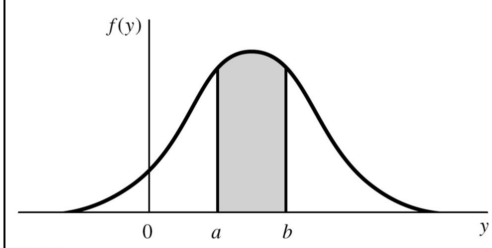
$$1.\ f(y_1, y_2) \geq 0 \quad \text{for all } y_1, y_2$$

$$2.\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(y_1, y_2)\, dy_1\, dy_2 = 1$$

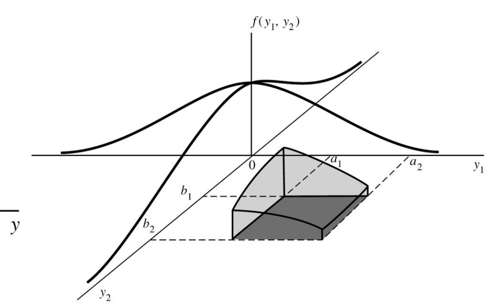
6

Graphical representation of univariate vs. bivariate PDFs

Univariate



Bivariate



Recall that in the univariate case, area under the PDF between a and b = P(a ≤ *Y* ≤ b)

$$\boxed{\begin{aligned} &P(a \leq Y \leq b) \\ &= \int_a^b f(y)\, dy \end{aligned}}$$

Bivariate PDFs are 2-dimensional surfaces, so the volume under the surface corresponds to a probability – e.g., above:

$$\boxed{\begin{aligned} &P(a_1 \leq Y_1 \leq a_2, b_1 \leq Y_2 \leq b_2) \\ &= \int_{b_1}^{b_2}\int_{a_1}^{a_2} f(y_1, y_2)\, dy_1\, dy_2 \end{aligned}}$$

7

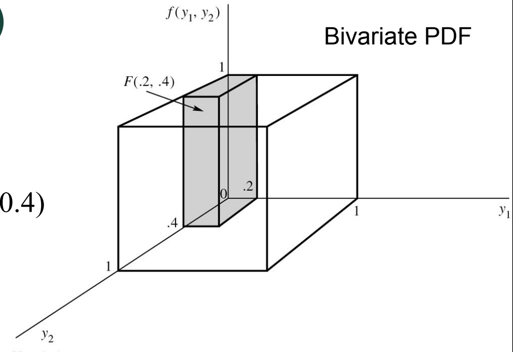
EX1) Finding probabilities from a bivariate PDF (continuous RVs)

$$f(y_1, y_2) = \begin{cases} 1, & 0 \le y_1 \le 1,\ 0 \le y_2 \le 1 \\ 0, & \text{elsewhere} \end{cases}$$

a. Find $F(0.2, 0.4) = P(Y_1 \le 0.2,\ Y_2 \le 0.4)$

b. Find $P(0.1 \le y_1 \le 0.3,\ 0 \le y_2 \le 0.5)$

Bivariate PDF



$$\boxed{\begin{aligned} &P(a_1 \le Y_1 \le a_2, b_1 \le Y_2 \le b_2) \\ &= \int_{b_1}^{b_2} \int_{a_1}^{a_2} f(y_1, y_2)\, dy_1\, dy_2 \end{aligned}}$$

8

EX1) Finding probabilities from a bivariate PDF (continuous RVs)

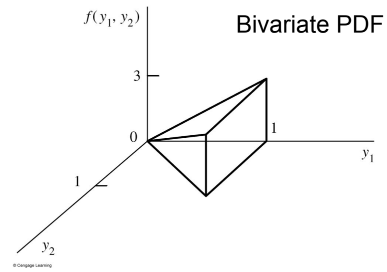
9

EX2) Finding probabilities from a bivariate PDF (continuous RVs)

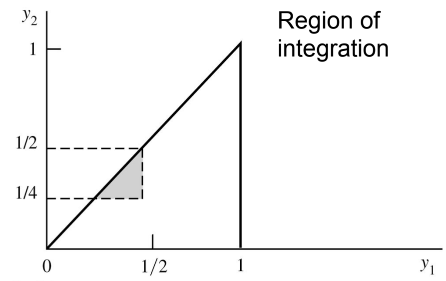
$$f(y_1,y_2)=\begin{cases} 3y_1, & 0\le y_2\le y_1\le 1 \\ 0, & \text{elsewhere}\end{cases}$$

Find $P(0\le Y_1\le 0.5,\ Y_2>0.25)$

Bivariate PDF



Region of integration

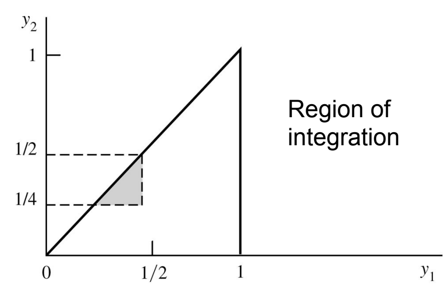


$P(a_1\le Y_1\le a_2, b_1\le Y_2\le b_2)$
$=\int_{b_1}^{b_2}\int_{a_1}^{a_2} f(y_1,y_2)\,dy_1\,dy_2$

EX2) Finding probabilities from a bivariate PDF (continuous RVs)

$$f(y_1,y_2)=\begin{cases} 3y_1, & 0\le y_2\le y_1\le 1 \\ 0, & \text{elsewhere}\end{cases}$$

Find $P(0\le Y_1\le 0.5,\ Y_2>0.25)$



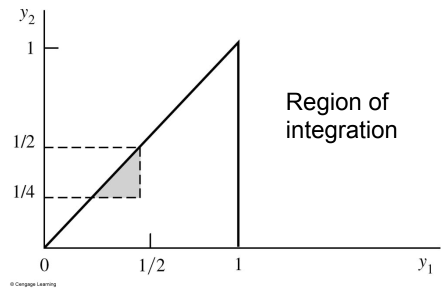
Region of integration

11

EX2) Finding probabilities from a bivariate PDF (continuous RVs)

$$f(y_1, y_2) = \begin{cases} 3y_1, \ 0 \le y_2 \le y_1 \le 1 \\ 0, \ \text{elsewhere} \end{cases}$$

Find $P(0 \le Y_1 \le 0.5,\ Y_2 > 0.25)$



12

Moving beyond the bivariate case

Joint probability distributions for discrete RVs:

$$p(y_1, y_2, \ldots, y_n) = P(Y_1 = y_1, Y_2 = y_2, \ldots, Y_n = y_n)$$

Joint probability density function (PDF) for continuous RVs:

$$f(y_1, y_2, \ldots, y_n)$$

Joint cumulative distribution function (CDF) for discrete & continuous RVs:

$$F(y_1, y_2, \ldots, y_n) = P(Y_1 \le y_1, Y_2 \le y_2, \ldots, Y_n \le y_n)$$

13

Marginal probability distributions

- Recall the supermarket checkout counter example. We derived the **bivariate** probability distribution, $p(y_1, y_2) = P(Y_1=y_1, Y_2=y_2)$:

	y_1			
y_2	0	1	2	$p_2(y_2)$:
0	1/9	2/9	1/9	4/9
1	2/9	2/9	0	4/9
2	1/9	0	0	1/9
$p_1(y_1)$:	4/9	4/9	1/9	

- *Can we get the **univariate** probability distribution for Y_1 (i.e., $p_1(y_1) = P(Y_1=y_1)$ from this bivariate distribution? In other words, what is $P(Y_1=0)$? $P(Y_1=1)$? $P(Y_1=2)$?*
- *How about $p_2(y_2) = P(Y_2=y_2)$?*

14

MICHIGAN STATE UNIVERSITY

Marginal probability distributions

- What we just derived are called "marginal probability distributions"

Marginal probability distributions for discrete RVs :

$$p_1(y_1) = \sum_{\text{all } y_2} p(y_1, y_2) \text{ and } p_2(y_2) = \sum_{\text{all } y_1} p(y_1, y_2)$$

- For continuous RVs:

Marginal probability density functions (PDFs) for continuous RVs :

$$f_1(y_1) = \int_{-\infty}^{\infty} f(y_1, y_2)\, dy_2 \text{ and } f_2(y_2) = \int_{-\infty}^{\infty} f(y_1, y_2)\, dy_1$$

EX) Marginal probability distributions from discrete bivariate probability distribution

From a group of three Republicans, two Democrats, and one independent, a committee of two people is to be randomly selected.

- Y_1: number of Republicans on the committee
- Y_2: number of Democrats on the committee

The bivariate probability distribution is given below. (This is a hypergeometric distribution problem, which we didn't cover.) Use the information in the table below to find the marginal probability distribution of Y_1 and Y_2.

Table 5.2 Joint probability function for Y_1 and Y_2, Example 5.5

	y_2	y_1 = 0	1	2	$p_2(y_2)$: *Total*
	0	0	3/15	3/15	6/15
	1	2/15	6/15	0	8/15
	2	1/15	0	0	1/15
$p_1(y_1)$:	*Total*	3/15	9/15	3/15	1

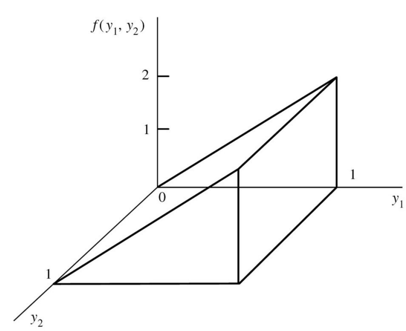
$f_1(y_1) = \int_{-\infty}^{\infty} f(y_1, y_2)\, dy_2$ and $f_2(y_2) = \int_{-\infty}^{\infty} f(y_1, y_2)\, dy_1$

MICHIGAN STATE UNIVERSITY

EX) Marginal PDFs from continuous bivariate PDF

Find the marginal PDFs for Y_1 and Y_2 from the continuous bivariate PDF:

$$f(y_1, y_2) = \begin{cases} 2y_1, & 0 \le y_1 \le 1, 0 \le y_2 \le 1 \\ 0, & \text{elsewhere} \end{cases}$$



Conditional probability distributions for discrete RVs:

What were some formulas we saw for $P(A \cap B)$ in terms of conditional and unconditional probabilities?

$$\boxed{P(A \cap B) = P(A)P(B \mid A) = P(B)P(A \mid B)}$$

So how can write $p(y_1, y_2) = P(Y_1 = y_1, Y_2 = y_2)$, an intersection of 2 events?

$$p(y_1, y_2) = p_1(y_1)p(y_2 \mid y_1) = p_2(y_2)p(y_1 \mid y_2)$$

Solve for the conditional probabilities:

$$p(y_1 \mid y_2) = \frac{p(y_1, y_2)}{p_2(y_2)}$$

and

$$p(y_2 \mid y_1) = \frac{p(y_1, y_2)}{p_1(y_1)}$$

Keep in mind what these expressions mean:

$$P(Y_1 = y_1 \mid Y_2 = y_2) = \frac{P(Y_1 = y_1, Y_2 = y_2)}{P(Y_2 = y_2)}$$

and

$$P(Y_2 = y_2 \mid Y_1 = y_1) = \frac{P(Y_1 = y_1, Y_2 = y_2)}{P(Y_1 = y_1)}$$

EX) Conditional probability distributions for discrete RVs

Continuing with our previous example about a 2-person committee drawn from Democrats, Republicans, and independents (with bivariate probability distribution below), find the conditional probability distribution of Y_1 given that Y_2=1.

Table 5.2 Joint probability function for Y_1 and Y_2, Example 5.5

	y_2	y_1 = 0	y_1 = 1	y_1 = 2	$p_2(y_2)$: Total
	0	0	3/15	3/15	6/15
	1	2/15	6/15	0	8/15
	2	1/15	0	0	1/15
$p_1(y_1)$:	Total	3/15	9/15	3/15	1

$$\boxed{p(y_1 \mid y_2) = \frac{p(y_1, y_2)}{p_2(y_2)}}$$

Conditional probability distributions for continuous RVs:

Recall that for **discrete RVs**, these are:

$$p(y_1 \mid y_2) = \frac{p(y_1, y_2)}{p_2(y_2)} \text{ and } p(y_2 \mid y_1) = \frac{p(y_1, y_2)}{p_1(y_1)}$$

Conditional PDF :

$$f(y_1 \mid y_2) = \frac{f(y_1, y_2)}{f_2(y_2)} \quad \text{and} \quad f(y_2 \mid y_1) = \frac{f(y_1, y_2)}{f_1(y_1)}$$

Conditional CDF :

$F(y_1 \mid y_2)$ means $P(Y_1 \leq y_1 \mid Y_2 = y_2)$

Obtain these values by integrating the conditional PDF over the relevant range.

20

MICHIGAN STATE UNIVERSITY

EX) Conditional probability distributions for continuous RVs

Find the conditional probability, $P(Y_1 \leq 0.5 \mid Y_2 = 1.5)$, for the continuous bivariate PDF:

$$f(y_1, y_2) = \begin{cases} 0.5, & 0 \leq y_1 \leq y_2 \leq 2 \\ 0, & \text{elsewhere} \end{cases}$$

Conditional PDF :

$$f(y_1 \mid y_2) = \frac{f(y_1, y_2)}{f_2(y_2)}$$

21

Independent random variables

*What is $P(A \cap B)$ if events A and B are **independent**?*

$$\boxed{P(A \cap B) = P(A)P(B)}$$

Can use similar approach to see if two RVs, Y_1 and Y_2, are independent:

CDF : $F(y_1, y_2) = F_1(y_1)F_2(y_2)$

Discrete probability distribution : $p(y_1, y_2) = p_1(y_1)p_2(y_2)$

Continuous probability distribution : $f(y_1, y_2) = f_1(y_1)f_2(y_2)$

22

MICHIGAN STATE UNIVERSITY

EX) Checking for independence – discrete RVs

Continuing with our previous example about a 2-person committee drawn from Democrats, Republicans, and independents (with bivariate probability distribution below), is Y_1 independent of Y_2? Can check any point, but let's try (0,0).

Table 5.2 Joint probability function for Y_1 and Y_2, Example 5.5

	y_2	y_1: 0	1	2	$p_2(y_2)$: *Total*
	0	0	3/15	3/15	6/15
	1	2/15	6/15	0	8/15
	2	1/15	0	0	1/15
$p_1(y_1)$:	*Total*	3/15	9/15	3/15	1

Discrete probability distribution if independent : $p(y_1, y_2) = p_1(y_1)p_2(y_2)$

23

MICHIGAN STATE UNIVERSITY

EX) Checking for independence – continuous RVs

Suppose Y_1 and Y_2 have the continuous bivariate PDF below. Are these two RVs independent?

$$f(y_1,y_2)=\begin{cases} 6y_1y_2^2,\ 0\le y_1\le 1,\ 0\le y_2\le 1 \\ 0,\ \text{elsewhere}\end{cases}$$

Continuous probability distribution if indepedent : $f(y_1,y_2)=f_1(y_1)f_2(y_2)$

24

Another way to check for independence

Y_1 and Y_2 are independent RVs if:

$f(y_1,y_2)=g(y_1)h(y_2)$

where $g(y_1)$ is a nonnegative function of y_1 alone and $h(y_2)$ is a nonnegative function of y_2 alone ***AND only if $f(y_1,y_2)>0$ for $a\le y_1\le b$ and $c\le y_2\le d$ for constants a,b,c,d***.

EX) Suppose Y_1 and Y_2 have the continuous bivariate PDF below. Are these two RVs independent?

$$f(y_1,y_2)=\begin{cases} 2y_1,\ 0\le y_1\le 1,\ 0\le y_2\le 1 \\ 0,\ \text{elsewhere}\end{cases}$$

25

MICHIGAN STATE UNIVERSITY

Cannot always use this method – e.g.:

$$f(y_1, y_2) = \begin{cases} 8y_1y_2, & 0 \le y_1 \le 1,\ 0 \le y_2 \le y_1 \\ 0, & \text{elsewhere} \end{cases}$$

We <u>cannot</u> use the alternative method here because the range over which y_2 has positive probability is a function of y_1.

If we found the marginal distributions for y_1 and y_2, we would see that:

$$f_1(y_1) = 4y_1^3$$

and

$$f_2(y_2) = 4y_2(1 - y_2^2)$$

So $f(y_1, y_2) \neq f_1(y_1) f_2(y_2)$, thus Y_1 and Y_2 not independent

26

MICHIGAN STATE UNIVERSITY

Homework:

- WMS Ch. 5 (part 2 of 3)
  - Marginal & conditional probability distributions: 5.19, 5.20, 5.26
  - Independent random variables: 5.45, 5.46, 5.52
- Reminder: Ch. 5 HW will <u>not</u> be collected

Next class:

- Multivariate probability distributions, cont'd (Part 3 of 3)
  - Expected values, variances, and covariances

Reading for next class:

- WMS Ch. 5 (sections 5.5-5.8, 5.11, 5.12)

27