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

Multivariate probability distributions – Part 1 of 3

(WMS Ch. 5.1-5.2)

October 5, 2017

Nicole Mason
Michigan State University

Fall 2017

GAME PLAN

Collect Ch. 4 HW

Review

Graded in-class exercise

Multivariate probability distributions (Part 1 of 3)

1. Bivariate probability distributions (discrete & continuous)
2. Multivariate probability distributions (discrete & continuous)

1

Review: standard normal distribution

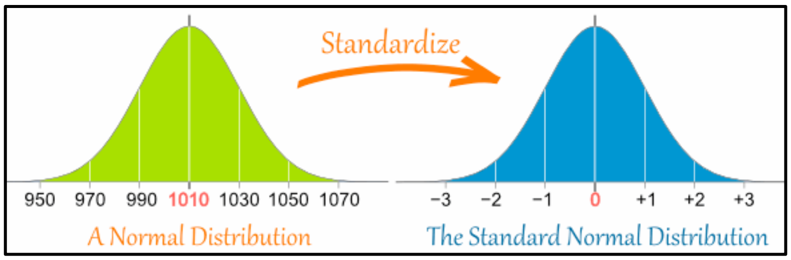
MICHIGAN STATE UNIVERSITY

Can convert any normal RV to standard normal

- Suppose $Y \sim N(\mu, \sigma^2)$, then

$$\boxed{Z = \frac{Y-\mu}{\sigma}, \quad Z \sim N(0,1) \; = \text{standard normal}}$$

- Once converted to standard normal → use Table 4 (P(Z>z))



2

Review: gamma distribution

MICHIGAN STATE UNIVERSITY

- **Non-negative** and **skewed to the right** (graphs on next slide)

$$f(y) = \begin{cases} \dfrac{y^{\alpha-1} e^{-y/\beta}}{\beta^{\alpha}\Gamma(\alpha)}, & 0 \le y \le \infty \\ 0, & \text{elsewhere} \end{cases}$$

$$\text{where} \quad \Gamma(\alpha) = \int_0^{\infty} y^{\alpha-1} e^{-y}\, dy$$

$$\text{and} \quad \alpha > 0,\ \beta > 0$$

$\Gamma(\alpha)$: "gamma function"

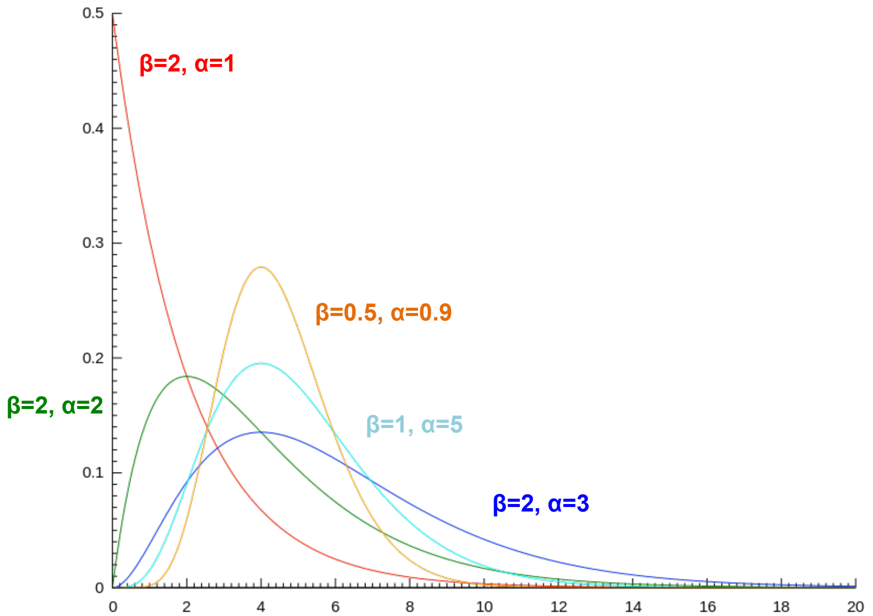
$\Gamma(1)=1$

$\Gamma(n)=(n-1)!$ if n is an integer

$$\boxed{E(Y) = \alpha\beta} \qquad \boxed{V(Y) = \alpha\beta^2}$$

3

Examples of gamma PDFs



4

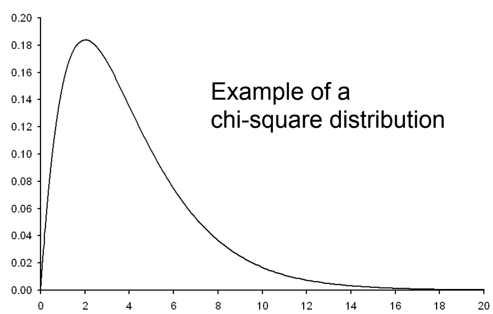
Review: gamma special case #1 – chi square

MICHIGAN STATE UNIVERSITY

1. **Chi-square** (χ^2) distribution: gamma w/ $\alpha = \frac{\nu}{2}$, $\beta = 2$
where ν is referred to as the number of degrees of freedom

$$E(Y) = \alpha\beta = \frac{\nu}{2} \cdot 2 = \nu$$

$$V(Y) = \alpha\beta^2 = \frac{\nu}{2} \cdot 2^2 = 2\nu$$



5

Review: gamma special case #2 – exponential

MICHIGAN STATE UNIVERSITY

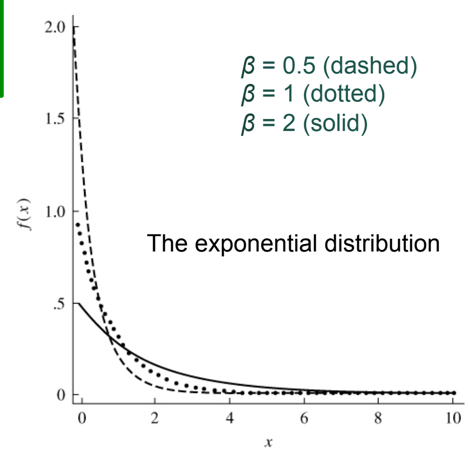
2. **Exponential** distribution: gamma w/ $\alpha = 1$, $\beta > 0$

$$f(y) = \begin{cases} \dfrac{e^{-y/\beta}}{\beta}, & 0 \le y \le \infty \\ 0, & \text{elsewhere} \end{cases}$$

$$F(y) = \begin{cases} 0, & y < 0 \\ 1 - e^{-y/\beta}, & y \ge 0 \end{cases}$$

$$E(Y) = \alpha\beta = 1\beta = \beta$$

$$V(Y) = \alpha\beta^2 = 1\beta^2 = \beta^2$$



Distribution	Probability Function	Mean	Variance
Uniform	$f(y) = \dfrac{1}{\theta_2 - \theta_1};\ \theta_1 \le y \le \theta_2$	$\dfrac{\theta_1 + \theta_2}{2}$	$\dfrac{(\theta_2 - \theta_1)^2}{12}$
$F(y) = 0$ for $y < \theta_1$; $\dfrac{y-\theta_1}{\theta_2-\theta_1}$ for $\theta_1 \le y \le \theta_2$; 1 for $y > \theta_2$			
Normal	$f(y) = \dfrac{1}{\sigma\sqrt{2\pi}}\exp\left[-\left(\dfrac{1}{2\sigma^2}\right)(y-\mu)^2\right]$ $-\infty < y < +\infty$	μ	σ^2
	$Z = \dfrac{Y-\mu}{\sigma}$, $\quad Z \sim N(0,1)$ = standard normal		
Exponential	$f(y) = \dfrac{1}{\beta}e^{-y/\beta};\quad \beta > 0$ $0 < y < \infty$	β	β^2
$F(y) = \begin{cases} 0, & y < 0 \\ 1 - e^{-y/\beta}, & y \ge 0 \end{cases}$			
Gamma	$f(y) = \left[\dfrac{1}{\Gamma(\alpha)\beta^\alpha}\right] y^{\alpha-1}e^{-y/\beta};$ $0 < y < \infty$ where $\Gamma(\alpha) = \int_0^\infty y^{\alpha-1}e^{-y}\,dy$ $\Gamma(n) = (n-1)!$ if n is an integer	$\alpha\beta$	$\alpha\beta^2$
Chi-square	$f(y) = \dfrac{(y)^{(v/2)-1}e^{-y/2}}{2^{v/2}\Gamma(v/2)};$ $y > 0$	v	$2v$

Graded in-class exercise

Multivariate probability distributions

- We are often interested in the intersection of 2 (or more) events
  - EX) Blackjack
    - A: drawing an ace
    - B: drawing a face card
  - EX) Egg producer
    - A: A chicken produces *n* eggs
    - B: *y* of those eggs are bad
- Need to understand the joint probability distributions (bivariate = 2, multivariate ≥ 2)
- We'll discuss these for discrete & continuous RVs

MICHIGAN STATE UNIVERSITY

Bivariate probability distribution for discrete RV

Example

- Roll a pair of dice. *How many ordered pairs of #s?*
  - mn=6*6=36
- Let (y_1, y_2) represent the (# of 1st die, # on 2nd die)
  - EX) Pair of 1s: (1, 1)
  - EX) 2 on 1st die, 3 on 2nd die: (2, 3)
- Consider all pairs (y_1, y_2). *What is the probability of each (ordered) pair?*
  - 1/36

→Bivariate probability distribution for this example:

$$p(y_1, y_2) = P(Y_1 = y_1, Y_2 = y_2) = 1/36,$$

where $y_1 = 1, 2, 3, 4, 5, 6, \quad y_2 = 1, 2, 3, 4, 5, 6$

10

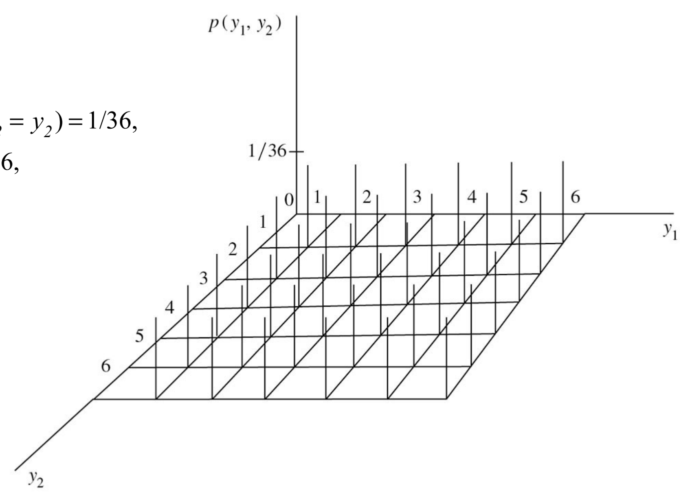
MICHIGAN STATE UNIVERSITY

Graphical representation of rolling a pair of die bivariate probability distribution

$$p(y_1, y_2) = P(Y_1 = y_1, Y_2 = y_2) = 1/36,$$

where $y_1 = 1, 2, 3, 4, 5, 6,$

$y_2 = 1, 2, 3, 4, 5, 6$



MICHIGAN STATE UNIVERSITY

Bivariate (or joint) probability distribution for discrete RVs (in general)

$$\boxed{\begin{array}{l} p(y_1, y_2) = P(Y_1 = y_1, Y_2 = y_2) \\ \text{where } -\infty < y_1 < \infty, \quad -\infty < y_2 < \infty \end{array}}$$

The usual rules for probabilities apply, but now they apply to the joint probability, $p(y_1, y_2)$. *What is the interpretation?*

$$1.\ 0 \le p(y_1, y_2) \le 1$$

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

Process for finding the bivariate probability distribution for discrete RVs: (1) find all (ordered) pairs of values; (2) assign a probability to each pair.

12

MICHIGAN STATE UNIVERSITY

EX) In the toss 2 dice experiment, what is $P(2 \le Y_1 \le 3,\ 1 \le Y_2 \le 2)$?

Rolls that satisfy this condition:

(2, 1)

(2, 2)

(3, 1)

(3, 2)

Probability of each is 1/36

Mutually exclusive so can add up the probabilities → (1/36)+(1/36)+(1/36)+(1/36)=4/36=1/9

13

Bivariate probability distribution (discrete RVs)

MICHIGAN STATE UNIVERSITY

EX) A supermarket has <u>3 checkout counters</u>. Two customers arrive at the counters at different times when the counters are serving no other customers. (Both can choose the same counter.)

- Let Y_1 denote the number of customers (out of the two customers) who choose counter 1.
- Let Y_2 denote the number of customers who choose counter 2.

Find the joint probability function of Y_1 and Y_2.

------

1. Define the sample space. Let {*i, j*} denote that the 1st customer chooses counter *i*, 2nd chooses counter *j*.
   E.g., {1,3} =1st customer chooses counter 1, 2nd customer chooses counter 3. *How many pairs of counter choices are there?*
   - 9
2. Map these into (y_1, y_2) pairs (in a table is easiest). Assign probabilities to each (y_1, y_2) pair.

14

Find the joint probability function of Y_1 and Y_2.

------

1. Define the sample space. Let {*i, j*} denote that the 1st customer chooses counter *i*, 2nd chooses counter *j*.
   E.g., {1,3} =1st customer chooses counter 1, 2nd customer chooses counter 3.
2. Map these into (y_1, y_2) pairs (in a table is easiest). Assign probabilities to each (y_1, y_2) pair.

Sample space
<u>(counters chosen)</u>

	y_1 (# of customers choosing counter 1)		
y_2 (# of customers choosing counter 2)	0	1	2
0			
1			
2			

15

Bivariate CDF (discrete RVs)

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

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

EX) In the toss 2 dice experiment, what is F(2, 3) = P(Y_1 ≤ 2, Y_2 ≤ 3)?
Recall that there are 36 possible ordered pairs, each w/ probability 1/36.

= *p(1,1) + p(1,2) + p(1,3) + p(2,1) + p(2,2) + p(2,3)* = 6*(1/36) = 1/6

16

Bivariate cumulative distribution function (CDF) (discrete RVs)

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

EX) In the supermarket checkout counter example, find:

a. F(-1, 2)

b. F(1.5, 2)

c. F(5, 7)

	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

Bivariate CDF (discrete RVs)

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

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 (continuous RVs)

$$\boxed{\begin{array}{l} F(y_1,y_2) = P(Y_1 \le y_1, Y_2 \le y_2) = \int_{-\infty}^{y_2}\int_{-\infty}^{y_1} f(t_1,t_2)\,dt_1\,dt_2 \\ \text{where } -\infty < y_1 < \infty, \quad -\infty < y_2 < \infty \end{array}}$$

MICHIGAN STATE UNIVERSITY

Bivariate PDFs (continuous RVs)

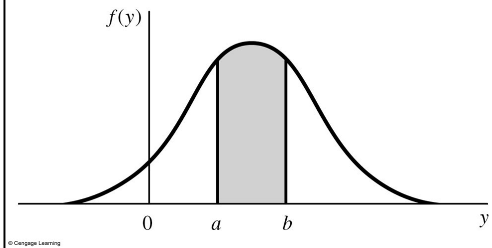
$$\boxed{f(y_1,y_2)}$$

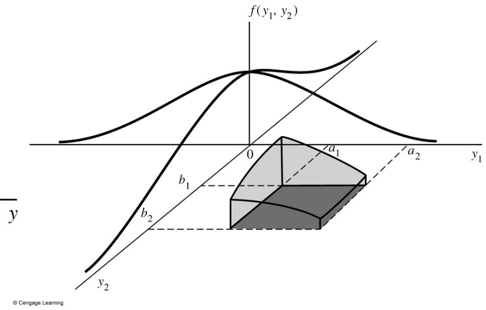
Like PDFs for the univariate case (and bivariate probability distributions for discrete RVs), bivariate PDFs must satisfy similar properties:

1. $0 \le f(y_1,y_2) \le 1 \quad \text{for all } y_1, y_2$
2. $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(y_1,y_2)\,dy_1\,dy_2 = 1$

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)

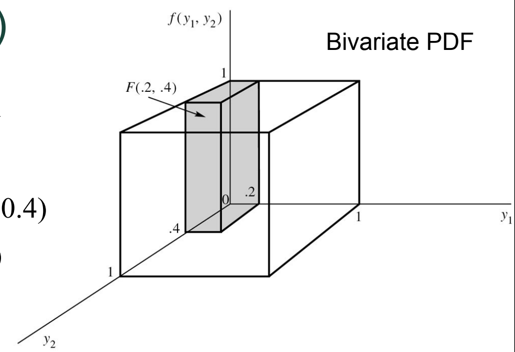
$$\boxed{P(a \le Y \le b) = \int_a^b f(y)\,dy}$$

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

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

20

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

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

21

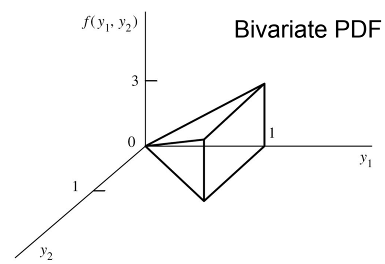
EX1) Finding probabilities from a bivariate PDF (continuous RVs)

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

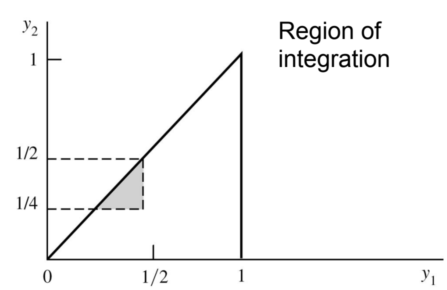
$$f(y_1, y_2) = \begin{cases} 3y_1, \ 0 \le y_2 \le y_1 \le 1 \\ 0, \ \text{elsewhere} \end{cases}$$

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

Bivariate PDF

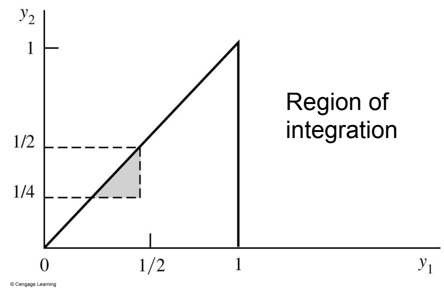


Region of integration



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

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

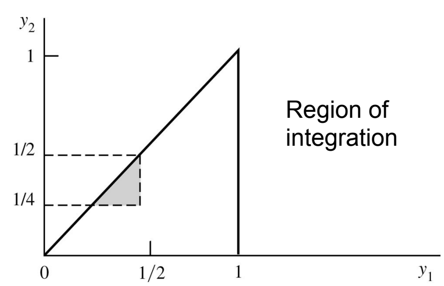


$$f(y_1, y_2) = \begin{cases} 3y_1, & 0 \le y_2 \le y_1 \le 1 \\ 0, & \text{elsewhere} \end{cases}$$

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

24

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

25

MICHIGAN STATE UNIVERSITY

Moving beyond the bivariate case

Joint probability distributions for discrete RVs :

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

Joint probability density function (PDF) for continuous RVs :

$$f(y_1, y_2, ..., y_n)$$

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

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

26

MICHIGAN STATE UNIVERSITY

Homework:

- WMS Ch. 5 (part 1 of 3)
  - Bivariate & multivariate probability distributions: 5.1, 5.2, 5.4, 5.7, 5.8
- Ch. 5 will not be collected (b/c would be due the class before the midterm)

Next class:

- Multivariate probability distributions, cont'd (Part 2 of 3)
  - Marginal & conditional probability distributions
  - Independent RVs

Reading for next class:

- WMS Ch. 5 (sections 5.3-5.4)

27