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

Wrap-up of multivariate probability distributions

October 17, 2017

Nicole Mason
Michigan State University
Fall 2017

GAME PLAN

- Hand back graded in-class exercises
- Midterm details
  - Covers material through today (except conditional exp. & var.)
  - Please bring blank paper with the last 4 digits of your PID written on each sheet, cheat sheet, pencil, calculator (preferably NOT your phone)
  - Format similar to last year's midterm but slightly shorter
- Review material from last class and complete discussion of multivariate probability distributions
  - Additional covariance example (discrete RVs) – questions?
  - Rules for expected values, variances, and covariances of linear functions of RVs
  - Conditional expectations & variances
- Q & A / review based on your questions

1

Review: Independent random variables

Recall that two events A and B are **independent** if:

$P(A \cap B) = P(A)P(B)$

Two RVs, Y_1 and Y_2, are **independent** if:

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

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

or

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

2

Review: The expected value of a function of RVs

For the ***bivariate*** case,
Discrete RVs :
$$E\left[g(Y_1, Y_2)\right] = \sum_{\text{all } y_2} \sum_{\text{all } y_1} g(y_1, y_2)p(y_1, y_2)$$
Continuous RVs :
$$E\left[g(Y_1, Y_2)\right] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(y_1, y_2) f(y_1, y_2)\, dy_1\, dy_2$$

Rules

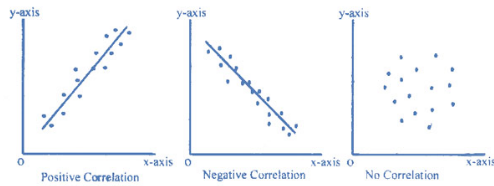
1. $E\left[cg(Y_1, Y_2)\right] = cE\left[g(Y_1, Y_2)\right]$ for any constant c
2. $E\left[g_1(Y_1, Y_2) + g_2(Y_1, Y_2) + \ldots + g_k(Y_1, Y_2)\right]$
$= E\left[g_1(Y_1, Y_2)\right] + E\left[g_2(Y_1, Y_2)\right] + \ldots + E\left[g_k(Y_1, Y_2)\right]$
3. $E\left(c_1Y_1 + c_2Y_2 + \ldots + c_kY_k\right) = c_1E\left(Y_1\right) + c_2E\left(Y_2\right) + \ldots + c_kE\left(Y_k\right)$
4. If Y_1 and Y_2 are independent then, $E(Y_1Y_2) = E(Y_1)E(Y_2)$
and, more generally, $E\left[g(Y_1)h(Y_2)\right] = E\left[g(Y_1)\right]E\left[h(Y_2)\right]$

3

Review: Covariance & correlation of 2 RVs

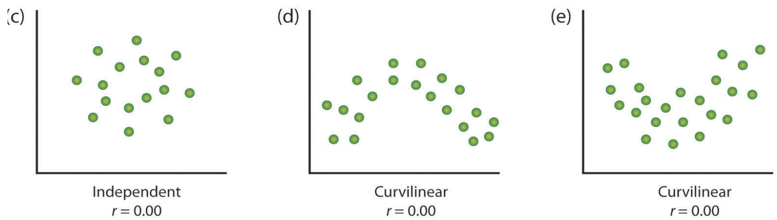
$$Cov(Y_1,Y_2) = E\left[(Y_1-\mu_1)(Y_2-\mu_2)\right] = E(Y_1Y_2) - E(Y_1)E(Y_2) = E(Y_1Y_2) - \mu_1\mu_2$$

$$Corr(Y_1,Y_2) \equiv \rho = \frac{Cov(Y_1,Y_2)}{\sigma_1\sigma_2}$$



Independence implies zero cov/corr

BUT **zero cov/corr does NOT imply independence**. *Why?*



4

Any questions on this example?

MICHIGAN STATE UNIVERSITY

Calculating the covariance – example #2

Show that Y_1 and Y_2 are dependent but have zero covariance.

	y_1		
y_2	-1	0	$+1$
-1	1/16	3/16	1/16
0	3/16	0	3/16
$+1$	1/16	3/16	1/16

$$E\left[g(Y_1,Y_2)\right] = \sum_{\text{all } y_2}\sum_{\text{all } y_1} g(y_1,y_2)p(y_1,y_2)$$

$$Cov(Y_1,Y_2) = E(Y_1Y_2) - E(Y_1)E(Y_2)$$

Rules for the expected value, variance, and covariance of linear functions of RVs:

The bivariate case (see WMS pp. 271-273 for proof & multivariate case)

Random variables Y_1 and Y_2, and constants a_1, a_2, b_1 and b_2 :

$$\boxed{1.\ E(a_1Y_1 + b_1 + a_2Y_2 + b_2) = a_1E(Y_1) + b_1 + a_2E(Y_2) + b_2}$$

EX) $E(3Y_1 - 2 - 8Y_2 + 5)$

$$\boxed{2.\ V(a_1Y_1 + b_1 + a_2Y_2 + b_2) = a_1^2V(Y_1) + a_2^2V(Y_2) + 2a_1a_2Cov(Y_1, Y_2)}$$

EX) $V(Y_1 + Y_2)$

EX) $V(Y_1 - Y_2)$

EX) $V(3Y_1 - 2 - 8Y_2 + 5)$

$$\boxed{3.\ Cov(a_1Y_1 + b_1, a_2Y_2 + b_2) = a_1a_2Cov(Y_1, Y_2)}$$

EX) $Cov(Y_1, -Y_2)$

EX) $Cov(3Y_1 - 2, -8Y_2 + 5)$

6

MICHIGAN STATE UNIVERSITY

Conditional expectations

Motivation

- Covariance and correlation measure the linear relationship (linear dependence) between two RVs and treat them symmetrically
- In applied economics, we often want to explain one RV (*Y*) in terms of another RV (*X*)
- Call *Y* the “explained” variable, *X* the “explanatory” variable
- Recall conditional probability distributions and PDFs: *p(y|x)* and *f(y|x)*
- We are often interested in the **conditional expectation** (a.k.a. the **conditional mean**): ***E(Y|X=x)*** or, for shorthand, ***E(Y|X)*** or sometimes ***E(Y|x)***

7

MICHIGAN STATE UNIVERSITY

Conditional expectations & variances - formulas

Conditional expectation of Y given X

Discrete RVs:

$$E(Y \mid X=x) = E(Y \mid X) = \sum_{\text{all } y} y\, p(y \mid x)$$

Continuous RVs:

$$E(Y \mid X=x) = E(Y \mid X) = \int_{-\infty}^{\infty} y\, f(y \mid x)\, dy$$

How would you use the E[g(Y)|X] formula to find the conditional variance, V(Y|X)?

$$V(Y \mid X) = E(Y^2 \mid X) - \left[E(Y \mid X)\right]^2$$

Conditional expectation of $g(Y)$ given X

Discrete RVs:

$$E\left[g(Y) \mid X=x\right] = E\left[g(Y) \mid X\right] = \sum_{\text{all } y} g(y)\, p(y \mid x)$$

Continuous RVs:

$$E\left[g(Y) \mid X=x\right] = E\left[g(Y) \mid X\right] = \int_{-\infty}^{\infty} g(y)\, f(y \mid x)\, dy$$

8

Conditional expectations & variances

Gist: treat the variable you are conditioning on as a constant

EX) Suppose $E(u \mid X) = 0$ and $Y = \beta_0 + \beta_1 X + u$, what is $E(Y \mid X)$?

RULES

1. $E\left[g(Y) \mid Y\right] = g(Y)$ for any function $g(.)$

   EX) $E(Y^2 \mid Y)$

2. $E\left[g(X)Y \mid X\right] = g(X)E(Y \mid X)$

   EX) $E(2X^2 Y \mid X)$

3. If X and Y are independent, then $E(Y \mid X) = E(Y)$ and $V(Y \mid X) = V(Y)$

4. If $E(Y \mid X) = E(Y)$, then $Cov(X,Y) = 0$

5. $E\left[E(Y \mid X)\right] = E(Y)$ "the law of iterated expectations"

See next slide for example

MICHIGAN STATE UNIVERSITY

$$\boxed{5.\ E\left[E(Y \mid X)\right] = E(Y) \quad \text{"the law of iterated expectations"}}$$

EX) If $E(WAGE \mid EDUC) = 4 + 0.6\,EDUC$ and $E(EDUC) = 11.5$, find $E(WAGE)$.

10

MICHIGAN STATE UNIVERSITY

Homework:

- WMS Ch. 5 (part 3 of 3) – NOT collected
  - Expected value of a function of RVs & special theorems: 5.72, 5.74
  - Covariance: 5.89, 5.91, 5.92 (Hint: $E(Y_1)=0.25$ and $E(Y_2)=0.5$)
  - Expected values, variances, covariances, and correlations of linear functions of RVs: 5.102, 5.103 (consult Theorem 5.12), 5.110
  - Conditional expectations: none but review & internalize the rules
  - Include the various rules on your cheat sheet

After midterm:

- Sampling distributions, estimation, hypothesis testing, and intro to linear regression/OLS
- Before our next class after the midterm, please watch https://www.youtube.com/watch?v=BwE2a18Th4c&feature=youtu.be - covers the sampling distribution of the sample mean, the Law of Large Numbers and the Central Limit Theorem

11