

Economet**tricks**: Short guides to econometrics

Trick 03: Review of Distribution Theory

Davud Rostam-Afschar (Uni Mannheim)

Content

1. Joint and marginal bivariate distributions
2. The joint density function
3. The joint cumulative density function
4. The marginal probability density
5. Covariance and correlation
6. The conditional density function
7. Conditional mean aka regression
8. The bivariate normal
9. Useful rules

Bivariate distributions

For observations of two discrete variables $y \in \{1, 2\}$ and $x \in \{1, 2, 3\}$, we can calculate

► the frequencies $n_{x,y}$,



freq. $n_{x,y}$	$y = 1$	$y = 2$	$f(x) = n_x/N$
$x = 1$	1	2	3/10
$x = 2$	1	2	3/10
$x = 3$	0	4	4/10
$f(y) = n_y/N$	2/10	8/10	1

Bivariate distributions

For observations of two discrete variables $y \in \{1, 2\}$ and $x \in \{1, 2, 3\}$, we can calculate

- ▶ the frequencies $n_{x,y}$,
- ▶ conditional distributions $f(y|x)$ and $f(x|y)$,
- ▶
- ▶

freq. $n_{x,y}$	$y = 1$	$y = 2$	$f(x) = n_x/N$	cond. distr. $f(y x)$	$y = 1$	$y = 2$	\sum_y
$x = 1$	1	2	3/10	$f(y x = 1)$	1/3	2/3	1
$x = 2$	1	2	3/10	$f(y x = 2)$	1/3	2/3	1
$x = 3$	0	4	4/10	$f(y x = 3)$	0	1	1
$f(y) = n_y/N$	2/10	8/10	1	$f(y x = 1, x = 2, x = 3)$	1/5	4/5	1

cond. distr.			
$f(x y)$	$f(x y = 1)$	$f(x y = 2)$	$f(x y = 1, y = 2)$
$x = 1$	1/2	1/4	3/10
$x = 2$	1/2	1/4	3/10
$x = 3$	0	1/2	4/10
\sum_x	1	1	1

Bivariate distributions

For observations of two discrete variables $y \in \{1, 2\}$ and $x \in \{1, 2, 3\}$, we can calculate

- ▶ the frequencies $n_{x,y}$,
- ▶ conditional distributions $f(y|x)$ and $f(x|y)$,
- ▶ joint distributions $f(x, y)$, and
- ▶

freq. $n_{x,y}$	$y = 1$	$y = 2$	$f(x) = n_x/N$	cond. distr. $f(y x)$	$y = 1$	$y = 2$	\sum_y
$x = 1$	1	2	3/10	$f(y x = 1)$	1/3	2/3	1
$x = 2$	1	2	3/10	$f(y x = 2)$	1/3	2/3	1
$x = 3$	0	4	4/10	$f(y x = 3)$	0	1	1
$f(y) = n_y/N$	2/10	8/10	1	$f(y x = 1, x = 2, x = 3)$	1/5	4/5	1

cond. distr. $f(x y)$	$f(x y = 1)$	$f(x y = 2)$	$f(x y = 1, y = 2)$	joint distr. $f(x, y)$	$f(x, y = 1)$	$f(x, y = 2)$
$x = 1$	1/2	1/4	3/10	$f(x = 1, y)$	1/10	2/10
$x = 2$	1/2	1/4	3/10	$f(x = 2, y)$	1/10	2/10
$x = 3$	0	1/2	4/10	$f(x = 3, y)$	0	4/10
\sum_x	1	1	1			

Bivariate distributions

For observations of two discrete variables $y \in \{1, 2\}$ and $x \in \{1, 2, 3\}$, we can calculate

- ▶ the frequencies $n_{x,y}$,
- ▶ conditional distributions $f(y|x)$ and $f(x|y)$,
- ▶ joint distributions $f(x, y)$, and
- ▶ marginal distributions $f_y(y)$ and $f_x(x)$.

freq. $n_{x,y}$	$y = 1$	$y = 2$	$f(x) = n_x/N$	cond. distr. $f(y x)$	$y = 1$	$y = 2$	\sum_y
$x = 1$	1	2	3/10	$f(y x = 1)$	1/3	2/3	1
$x = 2$	1	2	3/10	$f(y x = 2)$	1/3	2/3	1
$x = 3$	0	4	4/10	$f(y x = 3)$	0	1	1
$f(y) = n_y/N$	2/10	8/10	1	$f(y x = 1, x = 2, x = 3)$	1/5	4/5	1

cond. distr. $f(x y)$	$f(x y = 1)$	$f(x y = 2)$	$f(x y = 1, y = 2)$	joint distr. $f(x, y)$	$f(x, y = 1)$	$f(x, y = 2)$	marginal pr. $f_x(x)$
$x = 1$	1/2	1/4	3/10	$f(x = 1, y)$	1/10	2/10	3/10
$x = 2$	1/2	1/4	3/10	$f(x = 2, y)$	1/10	2/10	3/10
$x = 3$	0	1/2	4/10	$f(x = 3, y)$	0	4/10	4/10
\sum_x	1	1	1	marginal pr. $f_y(y)$	2/10	8/10	1

The joint density function

Two random variables X and Y have **joint density function**

- ▶ if x and y are discrete

$$f(x, y) = \text{Prob}(a \leq x \leq b, c \leq y \leq d) = \sum_{a \leq x \leq b} \sum_{c \leq y \leq d} f(x, y)$$

- ▶ if x and y are continuous

$$f(x, y) = \text{Prob}(a \leq x \leq b, c \leq y \leq d) = \int_a^b \int_c^d f(x, y) dx dy$$

Example

With $a = 1, b = 2, c = 2, d = 2$ and the following $f(x, y)$

joint distr. $f(x, y)$	$f(x, y = 1)$	$f(x, y = 2)$
$f(x = 1, y)$	1/10	2/10
$f(x = 2, y)$	1/10	2/10
$f(x = 3, y)$	0	4/10

$$\text{Prob}(1 \leq x \leq 2, 2 \leq y \leq 2) = f(y = 2, x = 1) + f(y = 2, x = 2) = 2/5.$$

Bivariate probabilities

For values x and y of two discrete random variable X and Y , the **probability distribution**

$$f(x, y) = \text{Prob}(X = x, Y = y).$$

The axioms of probability require

$$f(x, y) \geq 0,$$

$$\sum_x \sum_y f(x, y) = 1.$$

If X and Y are continuous,

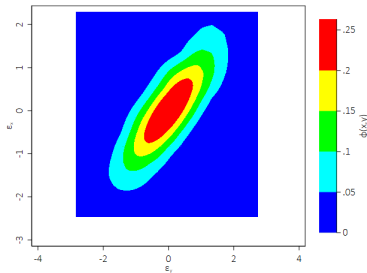
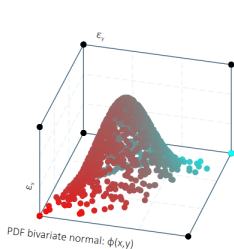
$$\int_x \int_y f(x, y) dx dy = 1.$$

The bivariate normal distribution

The bivariate normal distribution is the joint distribution of two normally distributed variables. The density is

$$f(x, y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} e^{-1/2[(\epsilon_x^2 + \epsilon_y^2 - 2\rho\epsilon_x\epsilon_y)/(1-\rho^2)]}, \quad (1)$$

where $\epsilon_x = \frac{x-\mu_x}{\sigma_x}$, and $\epsilon_y = \frac{y-\mu_y}{\sigma_y}$.



The joint cumulative density function

The probability of a joint event of X and Y have **joint cumulative density function**

- ▶ if x and y are discrete

$$F(x, y) = \text{Prob}(X \leq x, Y \leq y) = \sum_{X \leq x} \sum_{Y \leq y} f(x, y)$$

- ▶ if x and y are continuous

$$F(x, y) = \text{Prob}(X \leq x, Y \leq y) = \int_{-\infty}^x \int_{-\infty}^y f(t, s) ds dt$$

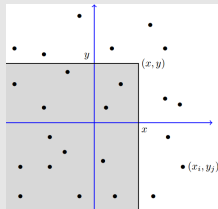
Example

With $x = 2, y = 2$ and the following $f(x, y)$

$f(x, y)$	$f(x, y = 1)$	$f(x, y = 2)$
$f(x = 1, y)$	1/10	2/10
$f(x = 2, y)$	1/10	2/10
$f(x = 3, y)$	0	4/10

$$\text{Prob}(X \leq 2, Y \leq 2) = f(x = 1, y = 1) +$$

$$f(x = 2, y = 1) + f(x = 1, y = 2) + f(x = 2, y = 2) = 3/5.$$



Bivariate probabilities

For values x and y of two discrete random variable X and Y , the **cumulative probability distribution**

$$F(x, y) = \text{Prob}(X \leq x, Y \leq y).$$

The axioms of probability require

$$0 \leq F(x, y) \leq 1,$$

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

$$F(-\infty, y) = 0,$$

$$F(x, -\infty) = 0.$$

The marginal probabilities can be found from the joint cdf

$$f_x(x) = P(X \leq x) = \text{Prob}(X \leq x, Y \leq \infty) = F(x, \infty).$$

The marginal probability density

To obtain the marginal distributions $f_x(x)$ and $f_y(y)$ from the joint density $f(x, y)$, it is necessary to sum or integrate out the other variable. For example,

- ▶ if x and y are discrete

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

- ▶ if x and y are continuous

$$f_x(x) = \int_y f(x, s) ds.$$

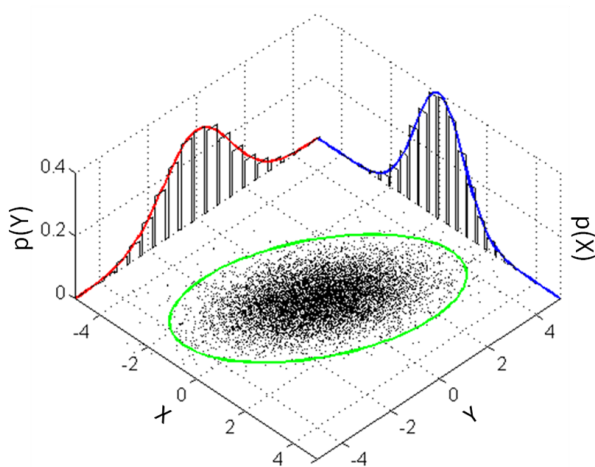
Example

$f(x, y)$	$f(x, y = 1)$	$f(x, y = 2)$	$f_x(x)$
$f(x = 1, y)$	1/10	2/10	3/10
$f(x = 2, y)$	1/10	2/10	3/10
$f(x = 3, y)$	0	4/10	4/10
$f_y(y)$	2/10	8/10	1

$$f_x(x = 1) = f(x = 1, y = 1) + f(x = 1, y = 2) = 3/10.$$

$$f_y(y = 2) = f(x = 1, y = 2) + f(x = 2, y = 2) + f(x = 3, y = 2) = 4/5.$$

The bivariate normal distribution



Why do we care about marginal distributions?

Means, variances, and higher moments of the variables in a joint distribution are defined with respect to the marginal distributions.

► Expectations

If x and y are discrete

$$E[x] = \sum_x x f_x(x) = \sum_x x \left[\sum_y f(x, y) \right] = \sum_x \sum_y x f(x, y).$$

If x and y are continuous

$$E[x] = \int_x x f_x(x) = \int_x \int_y x f(x, y) dy dx.$$

► Variances

$$\text{Var}[x] = \sum_x (x - E[x])^2 f_x(x) = \sum_x \sum_y (x - E[x])^2 f(x, y).$$

Covariance and correlation

For any function $g(x, y)$,

$$E[g(x, y)] = \begin{cases} \sum_x \sum_y g(x, y) f(x, y) & \text{in the discrete case,} \\ \int_x \int_y g(x, y) f(x, y) dy dx & \text{in the continuous case.} \end{cases} \quad (2)$$

The covariance of x and y is a special case:

$$\begin{aligned} \text{Cov}[x, y] &= E[(x - \mu_x)(y - \mu_y)] \\ &= E[xy] - \mu_x \mu_y = \sigma_{xy} \end{aligned}$$

If x and y are independent, then $f(x, y) = f_x(x)f_y(y)$ and

$$\begin{aligned} \sigma_{xy} &= \sum_x \sum_y f_x(x) f_y(y) (x - \mu_x)(y - \mu_y) \\ &= \sum_x (x - \mu_x) f_x(x) \sum_y (y - \mu_y) f_y(y) = E[x - \mu_x] E[y - \mu_y] = 0. \end{aligned}$$

- ▶ correlation $\rho_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$
- ▶ σ_{xy} does not imply independence (except for bivariate normal).

Independence: Pdf and cdf from marginal densities

- Two random variables are statistically independent if and only if their joint density is the product of the marginal densities:

$$f(x, y) = f_x(x)f_y(y) \Leftrightarrow x \text{ and } y \text{ are independent.}$$

- If (and only if) x and y are independent, then the marginal cdfs factors the cdf as well:

$$F(x, y) = F_x(x)F_y(y) = \text{Prob}(X \leq x, Y \leq y) = \text{Prob}(X \leq x)\text{Prob}(Y \leq y).$$

Example

$f(x, y)$	$f(x, y = 1)$	$f(x, y = 2)$	$f_x(x)$
$f(x = 1, y)$	1/6	1/6	1/3
$f(x = 2, y)$	1/6	1/6	1/3
$f(x = 3, y)$	1/6	1/6	1/3
$f_y(y)$	1/2	1/2	1

$$f_x(x = 3) \times f_y(y = 2) = 1/3 \times 1/2 = 1/6.$$

$F(x, y)$	$F(x, y = 1)$	$F(x, y = 2)$
$F(x = 1, y)$	1/6	2/6
$F(x = 2, y)$	2/6	4/6
$F(x = 3, y)$	3/6	1

$$\begin{aligned} P(x \leq 2)P(y \leq 2) &= [f(x = 2, y = 1) + f(x = 2, y = 2)] \times \\ &\quad [f(x = 1, y = 2) + f(x = 2, y = 2)] \\ &= [1/6 + 1/6][1/6 + 1/6] = 4/36 = 2/18. \end{aligned}$$

The conditional density function

The **conditional distribution** over y for each value of x (and vice versa) has conditional densities

$$f(y|x) = \frac{f(x,y)}{f_x(x)} \quad f(x|y) = \frac{f(x,y)}{f_y(y)}.$$

The marginal distribution of x averages the probability of x given y over the distribution of all values of y $f_x(x) = E[f(x|y)f(y)]$. If x and y are independent, knowing the value of y does not provide any information about x , so $f_x(x) = f(x|y)$.

Example

cond. distr.				joint distr.			marginal pr.
$f(x y)$	$f(x y=1)$	$f(x y=2)$	$f(x y=1,y=2)$	$f(x,y)$	$f(x,y=1)$	$f(x,y=2)$	$f_x(x)$
$x=1$	1/2	1/4	3/10	$f(x=1,y)$	1/10	2/10	3/10
$x=2$	1/2	1/4	3/10	$f(x=2,y)$	1/10	2/10	3/10
$x=3$	0	1/2	4/10	$f(x=3,y)$	0	4/10	4/10
\sum_x	1	1	1	marginal pr. $f_y(y)$	2/10	8/10	1

$$f(x=3|y=2) = \frac{f(x=3,y=2)}{f_y(y=2)} = 4/10 \times 10/8 = 1/2.$$

$$f_x(x=2) = E_y[f(x=2|y)f(y)] = f(x=2|y=1)f(y=1) + f(x=2|y=2)f(y=2)$$

$$= 1/2 \times 2/10 + 1/4 \times 8/10 = 1/10 + 2/10 = 3/10.$$

Conditional mean aka regression

A random variable may always be written as

$$\begin{aligned}y &= E[y|x] + (y - E[y|x]) \\ &= E[y|x] + \epsilon.\end{aligned}$$

Definition

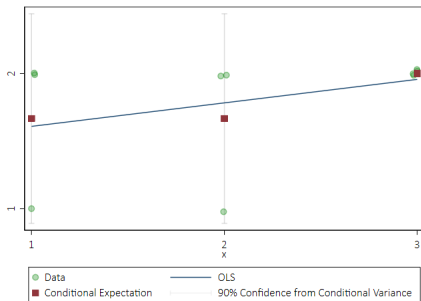
The regression of y on x is obtained from the **conditional mean**

$$E[y|x] = \begin{cases} \sum_y yf(y|x) & \text{if } y \text{ is discrete,} \\ \int_y yf(y|x)dy & \text{if } y \text{ is continuous.} \end{cases} \quad (3)$$

Conditional mean aka regression

Predict y at values of x :

$$\sum_y y f(y|x=1) = 1 \times 2/3 + 2 \times 2/3 = 5/3.$$



Conditional variance

A **conditional variance** is the variance of the conditional distribution:

$$\text{Var}[y|x] = \begin{cases} \sum_y (y - E[y|x])^2 f(y|x) & \text{if } y \text{ is discrete,} \\ \int_y (y - E[y|x])^2 f(y|x) dy, & \text{if } y \text{ is continuous.} \end{cases} \quad (4)$$

The computation can be simplified by using

$$\text{Var}[y|x] = E[y^2|x] - (E[y|x])^2 \geq 0. \quad (5)$$

Decomposition of variance $\text{Var}[y] = E_x[\text{Var}[y|x]] + \text{Var}_x[E[y|x]]$

- ▶ When we condition on x , the variance of y reduces on average.

$$\text{Var}[y] \geq E_x[\text{Var}[y|x]]$$

- ▶ $E_x[\text{Var}[y|x]]$ is the average of variances **within** each x
- ▶ $\text{Var}_x[E[y|x]]$ is variance **between** y averages in each x .

Conditional expectations and variances

- ▶ $E[y|x = 1] = 1.67$, $E[y|x = 2] = 1.67$, and $E[y|x = 3] = 2$
- ▶ $V[y|x = 1] = 0.22$, $V[y|x = 2] = 0.22$, and $V[y|x = 3] = 0$

Example

$f(y x)$	$y = 1$	$y = 2$	
$f(y x = 1)$	1/3	2/3	1
$f(y x = 2)$	1/3	2/3	1
$f(y x = 3)$	0	1	1

$$E[y|x = 1] = 1/3 \times 1 + 2/3 \times 2 = 5/3$$

$$E[y|x = 2] = 1/3 \times 1 + 2/3 \times 2 = 5/3$$

$$E[y|x = 3] = 0 \times 1 + 1 \times 2 = 2$$

$f(x, y)$	$f(x, y = 1)$	$f(x, y = 2)$	$f_x(x)$
$f(x = 1, y)$	1/10	2/10	3/10
$f(x = 2, y)$	1/10	2/10	3/10
$f(x = 3, y)$	0	4/10	4/10
$f_y(y)$	2/10	8/10	1

$$V[y|x = 1] = 1^2 \times 1/3 + 2^2 \times 2/3 - (5/3)^2 = 2/9$$

$$V[y|x = 2] = 1^2 \times 1/3 + 2^2 \times 2/3 - (5/3)^2 = 2/9$$

$$V[y|x = 3] = 1^2 \times 0 + 2^2 \times 1 - 2^2 = 0$$

alternatively (requiring more differences)

$$V[y|x = 1] = (1 - 5/3)^2 \times 1/3 + (2 - 5/3)^2 \times 2/3 = 2/9$$

Conditional expectations and variances

Average of variances **within** each x , $E[V[y|x]]$ is less or equal total variance $E[y]$.

Example

- Use the conditional mean to calculate $E[y]$:

$$\begin{aligned}E[y] &= E_x[E[y|x]] = E[y|x=1]f(x=1) + E[y|x=2]f(x=2) + E[y|x=3]f(x=3) \\&= 5/3 \times 3/10 + 5/3 \times 3/10 + 2 \times 4/10 = 9/5.\end{aligned}$$

$$E[y] = \sum_y f_y(y) = 1 \times 2/10 + 2 \times 8/10 = 9/5.$$

- Variation in y , $V[y|x=1] = 0.22$, $V[y|x=2] = 0.22$, and $V[y|x=3] = 0$ due to variation in x , is on average $E[V[y|x]] = 3/10 \times 2/9 + 3/10 \times 2/9 + 4/10 \times 0 = 2/15$.
- For each conditional mean $E[y|x=1] = 5/3$, $E[y|x=2] = 5/3$, and $E[y|x=3] = 2$, y varies with $V[E[y|x]] = E[(E[y|x])^2] - (E[y|x])^2 = 3/10 \times (5/3)^2 + 3/10 \times (5/3)^2 + 4/10 \times (2)^2 - (9/5)^2 = 2/75$.
- $E[V[y|x]] + V[E[y|x]] = V[y] = 2/75 + 2/15 = 4/25$.

With degree of freedom correction ($n - 1$) (as reported in software):

$$E[V[y|x]] + V[E[y|x]] = V[y] = 2/75/(10 - 1) \times 10 + 2/15/(10 - 1) \times 10 = 8/45.$$

Properties of the bivariate normal

Recall bivariate normal distribution is the joint distribution of two normally distributed variables. The density is

$$f(x, y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} e^{-1/2[(\epsilon_x^2 + \epsilon_y^2 - 2\rho\epsilon_x\epsilon_y)/(1-\rho^2)]}, \quad (6)$$

where $\epsilon_x = \frac{x-\mu_x}{\sigma_x}$, and $\epsilon_y = \frac{y-\mu_y}{\sigma_y}$.

The covariance is $\sigma_{xy} = \rho_{xy}\sigma_x\sigma_y$, where

- ▶ $-1 < \rho_{xy} < 1$ is the correlation between x and y
- ▶ $\mu_x, \sigma_x, \mu_y, \sigma_y$ are means and standard deviations of the marginal distributions of x or y

Properties of the bivariate normal

If x and y are bivariate normally distributed

$$(x, y) \sim N_2[\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho_{xy}]$$

- ▶ the marginal distributions are normal

$$f_x(x) = N[\mu_x, \sigma_x^2]$$

$$f_y(y) = N[\mu_y, \sigma_y^2]$$

- ▶ the conditional distributions are normal

$$f(y|x) = N[\alpha + \beta x, \sigma_y^2(1 - \rho^2)]$$

$$\alpha = \mu_y - \beta\mu_x; \beta = \frac{\sigma_{xy}}{\sigma_x^2}$$

- ▶ $f(x, y) = f_x(x)f_y(y)$ if $\rho_{xy} = 0$: x and y are independent if and only if they are uncorrelated

Useful rules

- ▶ $\rho_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$
- ▶ $E[ax + by + c] = aE[x] + bE[y] + c$
- ▶ $Var[ax + by + c] = a^2 Var[x] + b^2 Var[y] + 2abCov[x, y] = Var[ax + by]$
- ▶ $Cov[ax + by, cx + dy] = acVar[x] + bdVar[y] + (ad + bc)Cov[x, y]$
- ▶ If X and Y are uncorrelated, then
$$Var[x + y] = Var[x - y] = Var[x] + Var[y].$$

Useful rules

- ▶ Linearity

$$E[ax + by|z] = aE[x|z] + bE[y|z].$$

- ▶ Adam's Law / Law of Iterated Expectation

$$E[y] = E_x[E[y|x]]$$

- ▶ Adam's general Law / Law of Iterated Expectation

$$E[y|g_2(g_1(x))] = E[E[y|g_1(x)]|g_2(g_1(x))]$$

- ▶ Independence

If x and y are independent, then

$$E[y] = E[y|x],$$

$$E[g_1(x)g_2(y)] = E[g_1(x)]E[g_2(y)].$$

Useful rules

- ▶ Taking out what is known

$$E[g_1(x)g_2(y)|x] = g_1(x)E[g_2(y)|x].$$

- ▶ Projection of y by $E[y|x]$, such that orthogonal to $h(x)$

$$E[(y - E[y|x])h(x)] = 0.$$

- ▶ Keeping just what is needed (y predictable from x needed, not residual)

$$E[xy] = E[xE[y|x]].$$

- ▶ Eve's Law (EVVE) / Law of Total Variance

$$\text{Var}[y] = E_x[\text{Var}[y|x]] + \text{Var}_x[E[y|x]]$$

- ▶ ECCE law / Law of Total Covariance

$$\text{Cov}[x, y] = E_z[\text{Cov}[y, x|z]] + \text{Cov}_z[E[x|z], E[y|z]]$$

Useful rules

- ▶ $Cov[x, y] = Cov_x[x, E[y|x]] = \int_x (x - E[x]) E[y|x] f_x(x) dx.$
- ▶ If $E[y|x] = \alpha + \beta x$, then $\alpha = E[y] - \beta E[x]$ and $\beta = \frac{Cov[x,y]}{Var[x]}$
- ▶ Regression variance $Var_x[E[y|x]]$, because $E[y|x]$ varies with x
- ▶ Residual variance $E_x[Var[y|x]] = Var[y] - Var_x[E[y|x]]$, because y varies around the conditional mean
- ▶ Decomposition of variance $Var[y] = Var_x[E[y|x]] + E_x[Var[y|x]]$
- ▶ Coefficient of determination = $\frac{\text{regression variance}}{\text{total variance}}$
- ▶ If $E[y|x] = \alpha + \beta x$ and if $Var[y|x]$ is a constant, then

$$Var[y|x] = Var[y] (1 - Corr^2[y, x]) = \sigma_y^2 (1 - \sigma_{xy}^2)$$

References I

GREENE, W. H. (2011): *Econometric Analysis*. Prentice Hall, 5 edn.

PISHRO-NIK, H. (2014): *Introduction to Probability, Statistics, and Random Processes*. Kappa Research LLC.