Contents

1

- Example: The Ratio Model for Geometrical-Optical Illusions I
- Conditional Linear Regression
- The Conditional Regressions
- The Properties of the Residual
- Special Cases
- Parameterization
- Dichotomous Regressors
- Simple and Conditional Regression
- Example: The Ratio-Model for Geometrical-Optical Illusions II

Example: Baldwin Figures

2

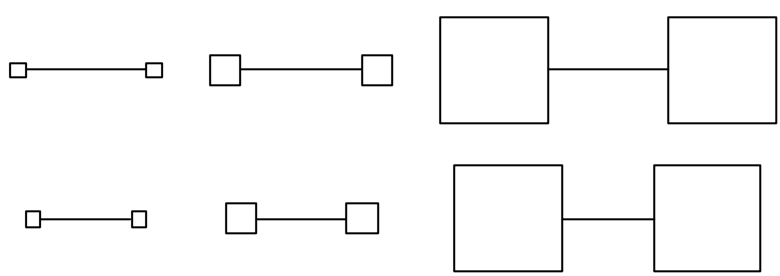

Figure 10.1. Six Baldwin-Figures, that result from the combination of lines of two different lengths and squares of two different sizes.

Example: The Ratio Model for Geometrical-Optical Illusions I

Table 10.1. Context-stimuli combinations and the constant context-stimuli ratios

		Context			
		1	2	4	8
Stimulus	1	1/1	2/1		
	2	1/2	1/1	2/1	
	4		1/2	1/1	2/1
	8			1/2	1/1

Notes. Only those context-stimulus ratios with at least 3 different stimuli are listed. Only within those ratios the hypotheses formulated in the text can be wrong.

Example: The Ratio Model for Geometrical-Optical Illusions I

Empirical studies showed the following model to be the best:

$$E(ln\ Y \mid lnX, Z) = g_0(Z) + g_1(Z) \cdot ln\ X.$$

This means for every context-stimulus ratio z the stochastical power law should hold, i.e.

$$E_{Z=z}(ln\ Y \mid ln\,X) = g_0(z) + g_1(z) \cdot ln\ X$$

where $g_0(z)$ und $g_1(z)$ are real numbers depending on the context-stimulus ratio z.

Conditional Linear Regression: Definition

Definition 10.1. Let X and Y be numerical random variables with finite expected values and variances and Z a random variable, all on a common probability space. Then the regressions $E(Y \mid X, Z)$ are called *conditionally linear in X* with respect to Z and Y is called *conditionally linearly regressively dependent on X with respect to Z*, if there exist two numerical functions $g_0(Z)$ and $g_1(Z)$ of Z such that:

$$E(Y \mid X, Z) = g_0(Z) + g_1(Z) \cdot X.$$

If $g_1(Z) = \gamma_0$, $\gamma_0 \in \mathrm{IR}$, then we call Y *partially linearly regressively dependent on X with respect to Z*.

Conditional Linear Regression: Figure I

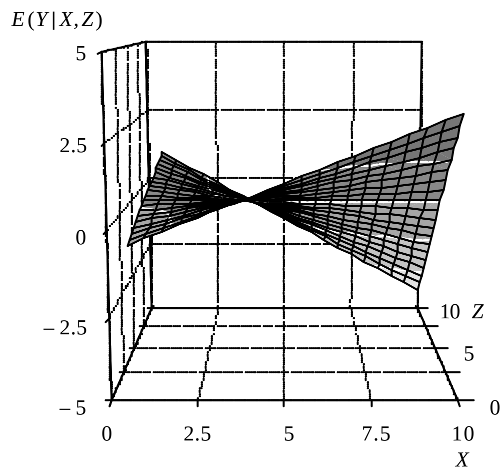

Figure 10.2. Illustration of a conditional linear regression with $g_0(Z) = -0.5 + 0.4 \cdot Z$ and $g_1(Z) = 0.15 - 0.1 \cdot Z$.

The Conditional Regressions

The key to understanding conditional linear regressive dependency is concept of the conditional regression of Y on X given a value z of Z. For a fixed value z of Z we have the equation

$$E_{Z=z}(Y|\ X) = g_0(z) + g_1(z) \cdot X,$$

if $E(Y|\ X) = g_0(Z) + g_1(Z) \cdot X$ is true.

The Conditional Regressions: Figure 2

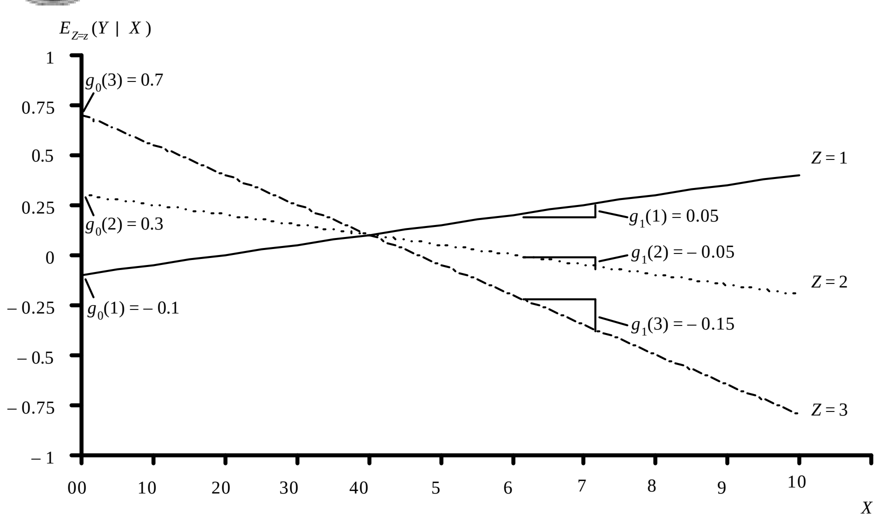

Figure 10.3. Graphs of the conditional linear regressions of Y on X for some values z of Z.

Properties of the Residual

The residual is defined as:

$$e := Y - E(Y \mid X, Z)$$

All of the general properties as discussed earlier hold. E.g.

$$E(e \mid X, Z) = E(e \mid X) = E(e \mid Z) = 0,$$

$$E[e \mid g_0(Z)] = E[e \mid g_1(Z)] = 0$$

$$E(e) = 0,$$

and

$$Cov(e, X) = Cov(e, Z) = Cov[e, g_0(Z)] = Cov[e, g_1(Z)] = 0.$$

Special Cases I

We speak of *conditional regressive independency* of the regressand Y from Z given X, if

$$E(Y \mid X, Z) = E(Y \mid X).$$

This again includes the special case of a simple linear regression where $g_0(Z) = \beta_0$ and $g_1(Z) = \gamma_0$ are constant functions of Z:

$$E(Y \mid X, Z) = \beta_0 + \gamma_0 X = E(Y \mid X).$$

Special Cases II

Multiple linear regression with two regressors is another special case of $E(Y|X) = g_0(Z) + g_1(Z) \cdot X$, in which

$$g_0(Z) = \beta_0 + \beta_1 Z, \quad \beta_0, \beta_1 \in \mathbb{R},$$

$$g_1(Z) = \gamma_0,$$

because this results in the equation:

$$E(Y|X, Z) = \beta_0 + \beta_1 Z + \gamma_0 X.$$

Special Cases III

We still assume that $E(Y|X) = g_0(Z) + g_1(Z) \cdot X$ holds.

If the equations

$$g_0(Z) = \beta_0 + \beta_1 Z, \quad \beta_0, \beta_1 \in \mathbb{R},$$

$$g_1(Z) = \gamma_0 + \gamma_1 Z, \quad \gamma_0, \gamma_1 \in \mathbb{R},$$

hold, this yields

$$\begin{aligned} E(Y|X, Z) &= (\beta_0 + \beta_1 Z) + (\gamma_0 + \gamma_1 Z) \cdot X \\ &= \beta_0 + \beta_1 Z + \gamma_0 X + \gamma_1 X Z. \end{aligned}$$

The function $g_1(Z) = \gamma_0 + \gamma_1 Z$ is called a *linear modification (or moderator) function*.

Parameterization: General

In general the number of values of Z is of no concern, no matter if Z is one-dimensional or multidimensional. However, to analyse a conditional linear regression with PC programs for multiple linear regression, the functions $g_0(Z)$ and $g_1(Z)$ have to be linear functions of the type:

$$g_0(Z) = \beta_0 + \beta_1 Z_1 + \beta_2 Z_2 + ... + \beta_{k-1} Z_{k-1}$$

$$g_1(Z) = \gamma_0 + \gamma_1 Z_1 + \gamma_2 Z_2 + ... + \gamma_{k-1} Z_{k-1}$$

where each of the variables $Z_1, Z_2, ..., Z_{k-1}$ is a function of Z.

Parameterization as Polynomials

If Z has only k different values the function $g_0(Z)$ and the modification function $g_1(Z)$ can always be presented as a polynomial of the order (k–1):

$$g_0(Z) = \beta_0 + \beta_1 Z + \beta_2 Z^2 + ... + \beta_{k-1} Z^{k-1}$$

$$g_1(Z) = \gamma_0 + \gamma_1 Z + \gamma_2 Z^2 + ... + \gamma_{k-1} Z^{k-1} .$$

The parameters β_i and γ_i, $i = 1, ..., k - 1$, are real numbers. This is called a *polynomial parameterization* of the conditional linear regression.

Parameterization with Indicator Variables

If Z has only k different values the function:

$$g_0(Z) = \beta_0 + \beta_1 I_1 + ... + \beta_j I_j + ... + \beta_{k-1} I_{k-1}$$

as well as the function

$$g_1(Z) = \gamma_0 + \gamma_1 I_1 + ... + \gamma_j I_j + ... + \gamma_{k-1} I_{k-1}.$$

can be represented as a weighted sum of indicator variables I_j, where each indicator variable I_j indicates with the value 1 that the j-th value of Z occurs, and 0 otherwise.

Dichotomous Regressors I

If the regressor X is dichotomous, then the equation of the contidional regression is always true. E.g. if X has only the values 0 and 1, then the conditional regression coefficients $g_1(z)$ are to be interpreted as the mean differences between the groups represented by X:

$$g_1(z) = E_{Z=z}(Y \mid X = 1) - E_{Z=z}(Y \mid X = 0)$$

or, which is equivalent:

$$g_1(z) = E(Y \mid X = 1, Z = z) - E(Y \mid X = 0, Z = z).$$

Dichotomous Regressors II

If the regressors X and Z are both dichotomous, then the following equation is always true:

$$E(Y \mid X, Z) = (\beta_0 + \beta_1 Z) + (\gamma_0 + \gamma_1 Z) \cdot X$$
$$= \beta_0 + \beta_1 Z + \gamma_0 X + \gamma_1 Z \cdot X.$$

Hence, this is a saturated parameterization. Filling in the values of X and Z results in the following system of linear equations:

$$E(Y \mid X = 1, Z = 1) = \beta_0 + \beta_1 + \gamma_0 + \gamma_1$$
$$E(Y \mid X = 1, Z = 0) = \beta_0 + \gamma_0$$
$$E(Y \mid X = 0, Z = 1) = \beta_0 + \beta_1$$
$$E(Y \mid X = 0, Z = 0) = \beta_0.$$

Dichotomous Regressors III

Solving this system of linear equation leads to:

$$\gamma_0 = E(Y|X = 1, Z = 0) - E(Y|X = 0, Z = 0)$$
$$\beta_1 = E(Y|X = 0, Z = 1) - E(Y|X = 0, Z = 0)$$
$$\gamma_1 = [E(Y|X = 1, Z = 1) - E(Y|X = 0, Z = 1)] - [E(Y|X = 1, Z = 0) - E(Y|X = 0, Z = 0)]$$

Hence, the conditional regression coefficients can be interpreted as follows:

$$g_0(0) = E(Y \mid X = 0, Z = 0) = \beta_0$$
$$g_1(0) = E(Y \mid X = 1, Z = 0) - E(Y|X = 0, Z = 0) = \gamma_0$$
$$g_0(1) = E(Y \mid X = 0, Z = 1) = \beta_0 + \beta_1$$
$$g_1(1) = E(Y \mid X = 1, Z = 1) - E(Y|X = 0, Z = 1) = \gamma_0 + \gamma_1.$$

Dichotomous Regressors: Figure

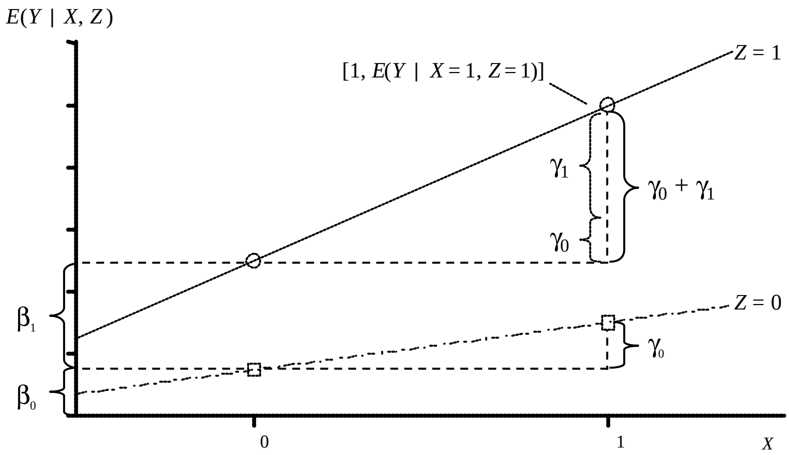

Figure 10.4. Conditional linear regression with dichotomous regressors, both with values 0 and 1.

Simple and Conditional Regression

Suppose Y is *conditionally linearly regressively dependent* on X given Z. Can we then conclude that Y is *linearly* regressively dependent on X? The answer is: generally not.

But if

$$E[g_0(Z)|X] = E[g_0(Z)] \quad \text{and} \quad E[g_1(Z)|X] = E[g_1(Z)]$$

then

$$E(Y|X) = \alpha_0 + \alpha_1 X$$

where

$$\alpha_0 = E[g_0(Z)] \quad \text{and} \quad \alpha_1 = E[g_1(Z)].$$

Example: The Ratio-Model for Geometrical-Optical Illusions II

For the three context-stimulus ratios we define the following indicator variables:

$$I_{1/2} := \begin{cases} 1, & \text{if the context-stimulus ratio equals } 1/2 \\ 0, & \text{otherwise,} \end{cases}$$

$$I_{1/1} := \begin{cases} 1, & \text{if the context-stimulus ratio is } 1/1 \\ 0, & \text{otherwise,} \end{cases}$$

Both functions

$$g_0(Z) := \beta_0 + \beta_1 I_{1/2} + \beta_2 I_{1/1} \quad \text{and} \quad g_1(Z) := \gamma_0 + \gamma_1 I_{1/2} + \gamma_2 I_{1/1}$$

have then different values for each of the three context-stimulus ratios.

Example: The Ratio-Model for Geometrical-Optical Illusions II

This leads to the following equation:

$$\begin{aligned} E(ln\ Y \mid ln\ X, Z) &= \beta_0 + \beta_1 I_{1/2} + \beta_2 I_{1/1} + (\gamma_0 + \gamma_1 I_{1/2} + \gamma_2 I_{1/1})\ ln\ X\,, \\ &= \beta_0 + \beta_1 I_{1/2} + \beta_2 I_{1/1} + \gamma_0\ ln\ X + \gamma_1 I_{1/2}\ ln\ X + \gamma_2 I_{1/1}\ ln\ X. \end{aligned}$$

Another interesting aspect is the following

$$Std_{Z=z}(Y \mid X) = exp[g_0(z)] + X^{g_1(z)} \cdot Std(d \mid Z = z)$$

Coefficient of Determination

The identification formula for the coefficient of determination is as follows:

$$R^2_{Y|X,Z} = \frac{Var[g_0(Z)] + Var[g_1(Z) \cdot X] + 2\,Cov[g_0(Z),\, g_1(Z) \cdot X]}{Var(Y)}$$

A statistic for the strength of the conditional linear regressive dependency of the regressand *Y* of the regressand *X* or variance proportion that is accounted by *X* additionally to *Z*:

$$R^2_{Y|X,Z} - R^2_{Y|Z}$$