Simple Nonlinear Regression

- Example: Stevens' Power Law III
- Linear Quasi-Regression
- Example: Stevens' Power Law IV
- Simple Nonlinear Regression
- Parameterization with Polynomials
- Parameterization as a Means Model
- Testing the Linearity of a Regression
- Logistic Regression

Müller-Lyer-Figure and Baldwin-Figure

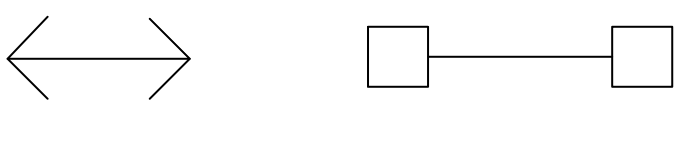

Figure 8.1. Müller-Lyer-Figur (left) and Baldwin-Figure (right). For the purpose of a direct comparison, the lines with the same length but without the context stimulus are presented below each figure.

Stevens' Power Law III

- How can the Stevens' Power Law be tested empirically, i.e. how to test if the dependency really linear regressive?
- How does the perception depend on context stimulus?

Before we can answer these questions we have to introduce further concepts: the linear quasi-regression and the simple nonlinear regression

Linear Quasi-Regression: Definition

Definition 8.1

Let X and Y be numerical random variables of a common probability space, and both with a positive and finite variance. The linear function $Q(Y \mid X) = \alpha_0 + \alpha_1 \cdot X$ of X is called *linear Quasi-Regression*, if for the residual

$$n := Y - (\alpha_0 + \alpha_1 \cdot X), \quad \alpha_0, \alpha_1 \in \mathrm{IR},$$

the following equations hold:

$$E(n) = 0,$$

$$Cov(n, X) = 0.$$

Linear Quasi-Regression: Notes

5

The variable Y is also defined as the sum of a linear function of X and the error variable n. This can be show by transforming the equation to:

$$Y = \alpha_0 + \alpha_1 \cdot X + n.$$

Linear Quasi-Regression: Properties of the Residual

6

$$E(n) = 0$$

$$Cov(n, X) = 0,$$

However

$$E(n \mid X) = 0$$

does not necessarily hold.

Linear Quasi-Regression: Alternative Definition

Definition 8.2.

Under the conditions given in definition 8.1 the *linear quasi regression of Y on X* can be defined as the linear function of *X*, that minimizes the following function of the real-valued a_0 and a_1:

$$LS(a_0, a_1) = E[(Y - (a_0 + a_1 \cdot X))^2]. \qquad \textit{least squares criterion}$$

Those numbers a_0 and a_1, for which the function $LS(a_0, a_1)$ is a minimum, are denoted by a_0 and a_1. The linear quasi regression is then defined by:

$$Q(Y \mid X) := \alpha_0 + \alpha_1 \cdot X.$$

Linear Quasi-Regression: Identification of Coefficients

Both definitions of the linear quasi-regression imply the same equations for the coefficients of the linear quasi-regression. They are identical to the equations for the coefficients of the (real) linear regression:

$$\alpha_0 = E(Y) - \alpha_1 \cdot E(X),$$

$$\alpha_1 = Cov(X, Y)/Var(X).$$

The identification of the coefficient of determination of the linear quasi-regression is also identical:

$$Q^2_{Y|X} := \frac{\alpha_1^2 Var(X)}{Var(Y)} = Kor(X, Y)^2.$$

Only if $E(Y \mid X)$ is really a linear function of *X* the coefficient of determination of the linear quasi-regression is a real coefficient of determination.

Linear Quasi Regression: Illustration

9

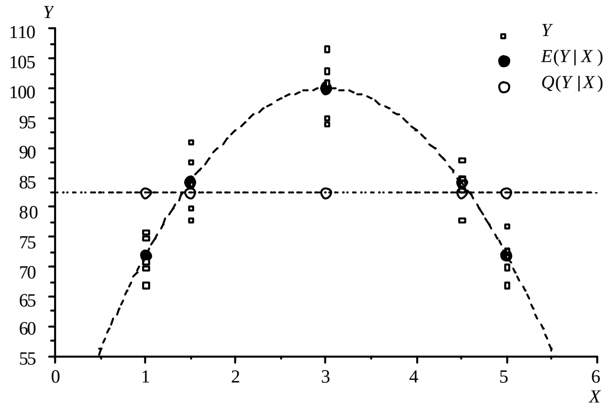

Figure 8.2. Regression and linear Quasi-Regression with parabolic regressive dependency of a variable *Y* from a variable *X*. The (real) regression is the parabola; the conditional expected values are the solid dots. The linear quasi-regression is shown by the straight line parallel to the *X*-axis with the non-solid dots. The squares are the pairs of values of (*X*, *Y*).

Simple Nonlinear Regression: Parameterization as a Polynomial Regression Model

10

The example in Figure 8.2 showed a regression, which was not a linear function but a *quadratic function*

$$E(Y \mid X) = \alpha_0 + \alpha_1 \cdot X + \alpha_2 \cdot X^2,$$

of *X*. In general there can cubic or even polynomials of higher order

$$E(Y \mid X) = \alpha_0 + \alpha_1 \cdot X + \alpha_2 \cdot X^2 + \ldots + \alpha_{n-1} \cdot X^{n-1}.$$

In this case however, the regressor *X* has to have at least *n* different values, because you can always put a polynomial of the order (*n*–1) through *n* points.

Simple Nonlinear Regression: Parameterization as a Means Regression Model I

If the regressor X of a regression $E(Y|X)$ has only n different values $x_1, ..., x_n$ the regression $E(Y|X)$ can be parameterizised as a means model. First, we need n indicator variables:

$$I_i = \begin{cases} 1, \text{if } X = x_i \\ 0, \text{otherwise} \end{cases}, \quad i = 1, ..., n.$$

The variables I_i (so-called dummy variables) indicate with the value 1, if the regressor X takes the value x_i. Note that these indicator variables I_i are functions of X, and that all $I_1, \ldots, I_n$ contain the same information as the regressor X. That is: $E(Y|X) = E(Y|I_1, ..., I_n)$. Hence we can use both notations.

Simple Nonlinear Regression: Parameterization as a Means Regression Model II

The following parameterization represents a saturated model:

$$E(Y \mid X) = \mu_1 \cdot I_1 + ... + \mu_n \cdot I_n .$$

The usage of the symbols μ_i makes sense, because the μ_i represent the conditional expected values, i.e.

$$\mu_i = E(Y|X = x_i), \quad \text{for } i = 1, ..., n.$$

This can be easily be shown by

$$\begin{aligned}
E(Y|X = x_1) &= \mu_1 \cdot 1 + \mu_2 \cdot 0 + ... + \mu_n \cdot 0 &&= \mu_1, \\
E(Y|X = x_2) &= \mu_1 \cdot 0 + \mu_2 \cdot 1 + \mu_3 \cdot 0 + ... + \mu_n \cdot 0 &&= \mu_2, \\
&\vdots \\
E(Y|X = x_n) &= \mu_1 \cdot 0 + \mu_2 \cdot 0 + ... + \mu_n \cdot 1 &&= \mu_n.
\end{aligned}$$

Testing the Linearity of a Regression I

To test the linearity of a regression $E(Y \mid X)$ you first calculate a linear quasi-regression and its coefficient of determination $Q^2_{Y|X} = \alpha_1^2 \; Var(X) \,/\, Var(Y)$ with a program for simple linear regression. Then you calculate the regression $E(Y \mid X)$ and the coefficient of determination $R^2_{Y|X}$ via a saturated parameterization. The test of the linearity of the regression is done by comparing $Q^2_{Y|X}$ with $R^2_{Y|X}$. If the regression $E(Y \mid X)$ is linear, then:

$$H_0\!: \;\; R^2_{Y|X} - Q^2_{Y|X} = 0$$

Testing the Linearity of a Regression II

The test of linearity can be accomplished by the following test statistic:

$$F = \frac{(\hat{R}^2_{Y|X_1,\ldots,X_n} - \hat{Q}^2_{Y|X})/(n-m)}{(1-\hat{R}^2_{Y|X_1,\ldots,X_n})/(N-n)}$$

where n is the number of parameters of the saturated parameterization, m the number of parameters of the restricted, the quasi linear parameterization and N the number of the sample size.

Logistic Regression I

If Y is a dichotomous regressand with values 0 and 1, then the regression $E(Y \mid X)$ is also the conditional probability function $P(Y = 1 \mid X)$. If X is not a dichotomous, but a continuous variable, then the regression $E(Y \mid X)$ can not be parameterized as a linear regression, because a straight line is inconsistent with the range [0, 1] of a (conditional) probability.
In this case the logistic linear parameterization is often used:

$$P(Y = 1 \mid X) = \frac{exp(\gamma_0 + \gamma_1 X)}{1 + exp(\gamma_0 + \gamma_1 X)}$$

Logistic Regression III

In the linear case, the *logit* may be written:

$$ln \frac{P(Y = 1 \mid X)}{P(Y = 0 \mid X)} = \gamma_0 + \gamma_1 X$$

Logistic Regression II

If X is continuous and Y is dichotomous the regression might not be a logistic linear one. The more general case would be a polynomial parameterization:

$$P(Y=1|X) = \frac{exp(\gamma_0 + \gamma_1 X + \gamma_2 X^2 + ... + \gamma_{n-1} X^{n-1})}{1 + exp(\gamma_0 + \gamma_1 X + \gamma_2 X^2 + ... + \gamma_{n-1} X^{n-1})}.$$

Logistic Regression III

The equation

$$P(Y=1 \mid X) = \frac{exp(\lambda_1 I_1 + \lambda_2 I_2 + ... + \lambda_n I_n)}{1 + exp(\lambda_1 I_1 + \lambda_2 I_2 + ... + \lambda_n I_n)}$$

is the saturated parameterisation if the regressor X has only n different values $x_1, ..., x_n$.